

# An evaluation of alternative dispersal functions for trees

DAVID F. GREENE\*§, CHARLES D. CANHAM†, K. DAVID COATES‡ and PHILIP T. LEPAGE§

\*Department of Geography, Concordia University, 1455 de Maisonnueve Boulevard, Montreal, Quebec, H3G 1M8, Canada, †Institute of Ecosystem Studies, PO Box AB, Millbrook, New York, USA, ‡British Columbia Forest Service, Research Section, Bag 6000, Smithers, British Columbia, V0J 2N0, Canada, and §Groupe de Recherche en Écologie Forestière Interuniversitaire, Université du Québec à Montreal, Montreal, Quebec, Canada

## Summary

- 1** We compared three commonly used empirical seed/seedling dispersal functions for trees (lognormal, 2Dt, and two-parameter Weibull) by analysis of published studies where the location of the source is known, as well as by inverse modelling within an old growth hardwood forest in southern Quebec. Almost all the species were wind-dispersed.
- 2** For the discrete source studies, the lognormal was clearly superior, while for the inverse modelling the performance of the three dispersal functions was somewhat more even. We speculate that collisions with boles spuriously enhanced the likelihood of the 2Dt and the Weibull with inverse modelling, as both these functions assume that the greatest seed/seedling density will occur at the base of the maternal parent bole.
- 3** We conclude that the lognormal function is to be preferred because, as well as providing a framework for mechanistic interpretation, it tends to provide a closer approximation to observed dispersal curves.
- 4** We also argue that mean distances travelled by seed crops are far more extensive than indicated by previous studies that used the Weibull function.

*Key-words:* anemochory, inverse modelling, recruitment, seed dispersal, tree regeneration

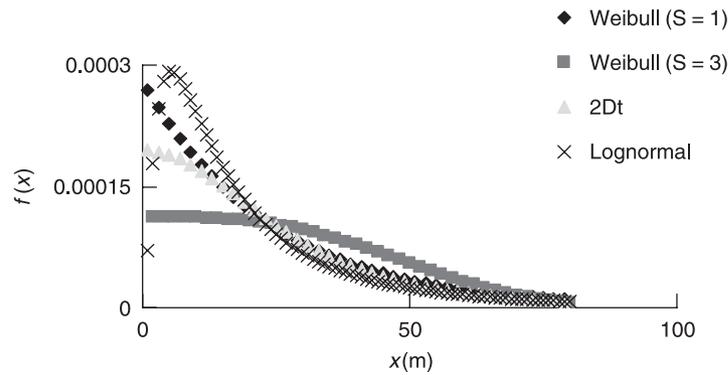
*Journal of Ecology* (2004) **92**, 758–766

## Introduction

Information about seed dispersal and recruitment is crucial for understanding the genetic structure of plant populations, plant invasions and, in some cases, species coexistence (reviewed in Nathan & Muller-Landau 2000). Further, the recruitment subroutine is an essential part of stand dynamics simulators being developed by foresters to predict stand density and volume (LePage *et al.* 2000). Nonetheless, empirical delineation of seed and seedling dispersal curves within forests has been a difficult task because the individual dispersal curves of conspecific trees usually overlap. There are a number of methods available for determining individual dispersal curves (Greene & Calogeropoulos 2002), but by far the most economical is the inverse modelling approach pioneered by Ribbens *et al.* (1994). Under this approach, maximum likelihood methods are used to estimate the terms of the dispersal function, given the spatial distribution and sizes of potential parent trees around each sample location.

The inverse modelling approach has now been used in a number of different studies, but with disagreement among practitioners over the most appropriate functional form of the dispersal curve. Ribbens *et al.* (1994) used a two-parameter Weibull function (sometimes referred to as the exponential family; Clark *et al.* 1999). Clark *et al.* (1999) proposed a composite dispersal function (the '2Dt' function) that was exponential in shape, but with a normally distributed variable for the scale parameter. They argued that this function was a better descriptor of dispersal curves than the two-parameter Weibull used by Ribbens *et al.* (1994). Stoyan & Wagner (2001) claimed the lognormal was superior to the Weibull. Meanwhile, other authors (e.g. LePage *et al.* 2000 for the Weibull, Tanaka *et al.* 1998 for the lognormal) have simply adopted one or another of these functions, intuiting, perhaps, that they will perform about equally well. Nonetheless, the choice of the function is critical; as noted by Nathan & Muller-Landau (2000), some functions have far tails that are too thin to permit meta-population persistence (let alone a migrational velocity sufficient to explain the Holocene record).

There is general agreement on the basic expression for the dispersal curve:



**Fig. 1** The three dispersal functions ( $f(x)$ ) from Table 1 with the median distance travelled set at 40 m for all three. For the 2Dt and lognormal we use values typical of the results for discrete sources (Table 2):  $S = 1$  and  $L = 1630$  for the 2Dt, and  $L = 40$  (the median distance) and  $S = 1$  for the lognormal. Two curves are depicted for the Weibull to show the difference in near-tail convexity as a function of the  $S$  value:  $S = 1$  or 3 and  $L = 0.042$  or  $0.0000057$ , respectively.

$$Q_{Dx} = a_1 B^b f_{s_g(x,y)} s_{s(x,y)} f_{s_h(x,y)} f_1(x) \quad \text{eqn 1}$$

where  $Q_{Dx}$  is the deposition density (seedlings  $\text{m}^{-2}$ ) at distance  $x$  from a source producing  $Q$  seeds,  $B$  is the basal area of a tree (or some other allometrically related size measure),  $f_1(x)$  is the seed dispersal kernel,  $f_{s_g(x,y)}$  is the density-dependent mortality in Cartesian space  $(x,y)$  after abscission but prior to germination, and  $a_1$  and  $b$  are empirical coefficients. The coefficient  $a_1$  translates tree size (basal area) into seed production ( $Q$ ) as  $Q = a_1 B^b$ . Analogously, we can imagine  $Q$  asexual buds along the roots of species such as *Fagus* or *Populus* that reproduce via root sprouts. The function  $s_g(x,y)$  is problematic as it represents the density response of granivores to the density of all the conspecific seeds as well as allospecific seeds, and the preferences and densities of the granivore species themselves. In equation 1, the surviving seed densities are subsequently reduced further by two types of losses: density-independent mortality at germination, and shortly thereafter, conditioned by seedbed-related mortality ( $s_{s(x,y)}$ ) and density-dependent herbivory ( $f_{s_h(x,y)}$ ) on the germinants and seedlings. There is no longer any doubt regarding the importance of density-dependent mortality (e.g. Harms *et al.* 2000) at both the seed and seedling stage, and thus the shape and scale of the realized recruitment kernel ( $f(x)$ ) may well look quite different from the original seed dispersal kernel ( $f_1(x)$ ). To date no study of inverse modelling has yet tried to parameterize this complete density-dependent function. Instead, modellers have used simplified versions:

$$F_{Dx} = aB^b f(x) \quad \text{eqn 2a}$$

(e.g. Ribbens *et al.* 1994), where  $f(x)$  is now the realized kernel with the original seed dispersal curve (*sensu* Nathan & Muller-Landau 2000) modified by predation, or, in a slightly more complex form:

$$F_{Dx} = aB^b f(x) s_{s(x,y)} \quad \text{eqn 2b}$$

**Table 1** The three  $f(x)$  with  $x$  the distance from the source tree. Recall that this is for *density* of seeds or seedlings vs. distance from a source plant and thus a summation of any of them from 0 to infinity must be multiplied by  $2\pi x$  to yield unity. Note that  $N$  below, a function of  $L$  and  $S$ , is merely a normalizer for the modified Weibull because it is not based on a true probability distribution.  $S$  is the shape parameter and  $L$  is the scale parameter. Familiarly, for the lognormal  $L$  is the median distance travelled while  $S$  is the standard deviation of the logarithms of the distance travelled

	$f(x)$
Weibull	$(1/N) \exp(-Lx^S)$
2Dt	$S/(\pi L [1 + (x^2/L)]^{S+1})$
Lognormal	$[1/((2\pi)^{1.5} S x^2)] \exp(-(\ln(x/L))^2/(2S^2))$

(LePage *et al.* 2000), with the seedbed-generated survivorship term now explicit.

It is this latter form (equation 2b) that we will use in the subsequent analyses.

#### THE ALTERNATIVE DISPERSAL KERNELS ( $F(X)$ )

We examine three dispersal kernels for tree seeds or seedlings. Each is a two-parameter distribution: the modified Weibull (as used by Ribbens *et al.* 1994), the lognormal (introduced by Greene & Johnson 1989), and the 2Dt (as proposed by Clark *et al.* 1999) (Fig. 1 and Table 1). In each case we will deal with *density* per distance (i.e. the kernel divided by  $2\pi x$ ). We restrict ourselves here to closed-form expressions, and thus ignore mechanistic individual-trajectory models such as those of Nathan *et al.* (2001) or Tackenberg (2003) for wind or Murray (1988) for bird defaecation of fruits. Further, we ignore 'mixed-model' empirical formulations that require more than two parameters as advocated by Bullock & Clarke (2000) among others.

The Weibull can take on very different shapes depending on the value of the parameter  $S$ . When this shape parameter is 3 (as with Ribbens *et al.* 1994 or LePage

*et al.* 2000) or higher, densities of recruits are relatively invariant with distance at first but then decline quite sharply (Fig. 1). The far tail is very thin. With decreasing  $S$ , the initial 'plateau' in density becomes less marked and the tail more extensive:  $S = 2$  for example leads to the right half of a Gaussian distribution and  $S = 1$  is the familiar negative exponential. Note that the modified Weibull, unlike the 2Dt or lognormal, requires a normalizer (Table 1) as it is not a true probability distribution (i.e. does not sum to 1) when the function is multiplied by  $2\pi x$  to create recruits per annulus rather than recruits per area per annulus.

The shape of the 2Dt is affected less dramatically by changes in the shape parameter,  $S$ . With increasing  $S$ , the density of recruits in the near tail becomes greater while the far tail density becomes both thicker and more extensive. For the lognormal, increases in  $S$  thicken the far tail while pushing the modal density back towards the source tree. Meanwhile, the scale parameter, denoted as  $L$  in Table 1, merely affects such distance measures as the median or mean distance travelled.

Clark *et al.* (1999) argued that only the 2Dt had the right shape for both the near and far tails of the  $f(x)$ . Reasonably, they argued that post-Holocene migration velocities required an extensive far tail containing an unspecified but presumably large fraction of the  $Q$  seeds produced. They pointed out that the modified Weibull can never place a sufficiently great proportion of  $Q$  in the far tail unless the shape parameter  $S$  (Table 1) is very small ( $< 1.0$ ). But in that case, the near tail would be concave near the source. They then claimed that near-tail concavity is never observed for point sources. Both Ribbens *et al.* (1994) and LePage *et al.* (2000) used a much larger value ( $S = 3$ ) for the Weibull, forcing the curve to maintain convexity near the source tree, but of course producing a much foreshortened far tail. Thus, Clark *et al.* (1999) argued, the 2Dt is dramatically better than the Weibull because it can simultaneously capture the shape of the curve both 'near and far'. Similarly, they argued that the lognormal, which places the modal deposition away from the source, cannot be a useful expression because the near tail must be convex. For empirical justification of this assertion of the near tail convexity, Clark *et al.* (1999) cite a pair of modelling studies that lack empirical data and two examples of inverse modelling. Realistically, inverse modelling cannot be used to prove the near-bole convexity because neither of the cited studies tested alternate forms that were not convex near the bole.

In a preliminary analysis, we found that the three dispersal terms can be quite similar over a limited range. For example, constraining them so that the median dispersal distance is 20 m and the 95th percentile occurs at 100 m, the three functions predict seed or seedling densities within 1.5-fold of each other from about 6 m to 200 m from a tree. Thus, at the scale at which ecologists have sampled (and this includes every data set discussed here), the three functions will tend to differ primarily in the very near tail, that is, close to the maternal parent.

(According to Greene & Calogeropoulos (2002) they will also differ substantially in the far tail.)

Our purpose is to compare the three candidate dispersal terms in two ways. First, we will look at the handful of point source studies (a single tree well-isolated from other conspecifics but nonetheless deep within a forest) that exist in the literature and ask which of the three alternative dispersal functions ( $f(x)$ ) is a better expression for empirical data. Secondly, we will repeat Clark's comparison of the 2Dt and Weibull, but include the lognormal now, as we examine inverse modelling results from an eastern hardwood forest.

## Methods

We examined seeds and seedlings of 10 species of trees at St Hilaire, an old growth hardwood-dominated forest in southern Quebec near the United States border. It suffered moderate damage from the 1961 and 1984 ice storms (Melanson & Lechowicz 1987), and heavy damage from the 1998 event (Proulx & Greene 2001).

In early April 1997, we placed a 270-m transect perpendicular to the contour lines in steeply dissected terrain so as to cross the bands of major abundance of each of the major species. That is, unlike all other earlier work with inverse modelling, we sought to ensure that longer dispersal distances (although not, it should be stressed, the extreme far tail crucial to migration) for each species would be prominent. The transect began on the far side of a ridge (dominated by *Acer saccharum* and *Ostrya virginiana*), went over the ridge (*O. virginiana* and *Quercus rubra*, with some *Pinus resinosa*), descended a steep south-west-facing slope (*Q. rubra* and, somewhat unusually for this slope position, a large number of *Tsuga canadensis*), reached the footslope (*Fagus grandifolia*, *A. saccharum*, some *Betula populifolia*), entered a 70-m wide bog (*Betula alleghaniensis* and *Acer rubrum*), and then began up a very steep east-facing slope (*A. saccharum*). In the bog we only sampled plots that were clearly above the high-water line from the previous year (this level was not exceeded during 1997). All the species except *F. grandifolia* and *Q. rubra* are wind dispersed, and almost all had a good to very good seed crop the previous autumn.

## ANALYSIS OF THE DISPERSAL OF THE 1996 SEED CROP

We sifted through the litter in late April (just after snowmelt) and early May to tally the seeds (about 4 person-hours per  $m^2$ ) that resulted from the mast year of 1996. We counted only filled seeds (inspected visually after cutting), but did not test for germinability. We used 4- $m^2$  plots spaced every 10 m along the 270 m transect for a total of 28 plots. Source trees greater than 5 cm d.b.h. were mapped (species, diameter, Cartesian coordinates) within 50 m on either side of the transect and 50 m beyond the two ends of the transect, for a total area of 3.7 ha.

We also used abscised ovulate cones of *T. canadensis* as proxy diaspores. First, we wished to examine dispersal of the cones themselves, because Ribbens *et al.* (1994) had claimed that most hemlock seeds stayed inside the cones, which in turn had a mean dispersal distance, they speculated, of only a few metres. These cones can also serve as examples of wind dispersal of fruits that are not well designed for anemochory, and yet are not subject to re-entrainment by granivores or frugivores. Thus, *Fagus* nuts or *Prunus* fruits that escaped predation while still attached to branches might be expected to travel in a similar manner (at least for the primary dispersal event) as these cones.

#### ANALYSIS OF THE 1997 GERMINANTS DERIVED FROM THE 1996 SEED CROP

In August we tallied the germinants along this same transect (1 m to the side of the seed plots) in 3.14-m<sup>2</sup> plots. Following LePage *et al.* (2000), the percentage cover for each seedbed type was recorded. The seedbed types were (i) rock and fresh wood, (ii) rotted wood, (iii) exposed mineral soil or humus, (iv) thin leaf litter (< 3 cm), (v) thick leaf litter (≥ 3 cm), and (vi) deep moss (> 3 cm). The two leaf litter categories comprised over 90% of the transect area.

#### ANALYSIS OF THE 2001 GERMINANTS

The January 1998 ice storm reduced the basal area of these stands at St Hilaire by about 35% (Hooper *et al.* 2001). Consequently there were few seeds or germinants in the next 2 years. Nonetheless, many species enjoyed good crops in southern Quebec in the autumn of 2000 and a few species provided us with enough germinants the following summer to permit further analysis. Censusing was done in 43, 3.14-m<sup>2</sup> plots 1 m to the west of the original 1997 seed plots in August of 2001 and seedbed coverage was again estimated.

#### INVERSE MODELLING

The 10 species analysed were those for which we obtained at least 50 censused seeds (or germinants). While we include seedbed effects for the germinant analyses of these species, these are, of course, ignored for the seeds. We follow LePage *et al.* (2000) with two exceptions. First, we examined two other dispersal functions (2Dt, lognormal) in addition to the two-parameter Weibull. Secondly, rather than set the exponent (*b*) on basal area in equation (2a) to a constant 1.0 as in previous studies, we let this parameter vary so that we could estimate (rather than assume) the scaling of seed production to tree size. Maximum likelihood estimates of model parameters were derived using simulated annealing (a global optimization algorithm) (Goffe *et al.* 1994). Seed and seedling count data were assumed to be Poisson distributed (as in Ribbens *et al.* 1994 and LePage *et al.* 2000). Clark *et al.* (1999) tested

use of the negative binomial distribution to account for clumping and hyperdispersion in the count data, but found that the Poisson was an appropriate distribution for wind-dispersed species with moderately large mean dispersal distances (as is the case for virtually all of the analyses reported here).

#### PUBLISHED CURVES FOR DISCRETE SOURCES

We found eight examples of published point source dispersal curves (all but one for seedlings rather than seeds) to which we added data for *Tilia americana* (germinants near an isolated tree about 1 km from our transect) at St Hilaire. The citations and species are listed in Table 2; all the species are wind-dispersed. The criterion for selection was that there was a single tree, well isolated from conspecifics, but deep inside a forest. Data were subjected to the same inverse modelling software as described above.

In addition, we used a detailed study of post-fire recruitment of *Pinus strobus* (Wood 1932) where there were only about 10 trees, well isolated from one another. What makes this study different from the others is that the wind speeds following fire will be much larger than in an intact forest in full leaf, but perhaps only somewhat faster than in a winter deciduous forest (Greene & Johnson 1996).

Finally, we used an example of dispersal by *Picea engelmannii* from discrete area sources into adjacent clearcuts (data from Squillace 1954; Crossley 1955; Ronco 1970; Greene & Johnson 1996). The source trees were simply arrayed as a uniformly spaced orchard for the inverse modelling estimation of the  $f(x)$  parameters.

## Results

#### LIKELIHOOD COMPARISONS

The lognormal was generally the best of the three models as a descriptor of the empirical point source curves, having the highest likelihood in six of the nine point source analyses (Table 2). In five of these analyses, however, the magnitude of the differences in log likelihood between the best and worst model was very small (< 2), indicating that all three of the models were roughly equally likely as descriptions of the data. The lognormal function had a much higher likelihood than the worst model in the other three analyses in which it performed best, while the 2Dt model had the highest likelihood in the one other analysis where likelihoods differed widely among the three models (Table 2). The general superiority of the lognormal in these analyses occurred despite the fact that some of these studies were not ideal for examining near-bole deposition. For example, the studies on *Tachigalia 2/Lonchocarpus/Platypodium* accumulated results in 5-m intervals and we cannot be sure if the mode was next to the bole or a few metres away. Finally, for the two examples in Table 2 where we treat a patch source (*Pinus strobes*, Wood 1932)

**Table 2** Log likelihood for the discrete source studies. The sources are: Rudis *et al.* (1978) (*Pinus strobus* no. 1 and *Pinus strobus* no. 2); Wood (1932) (*Pinus strobus* stand); Kitajima & Augspurger (1989) (*Tachigalia versicolor* no. 1 and no. 2); present study (*Tilia americana*); Ford *et al.* (1983) (*Betula uber*); Stoyan & Wagner (2001) (*Fraxinus excelsior*); Augspurger & Hogan (1983) (*Lonchocarpus pentaphyllus*); Augspurger (1983) (*Platypodium elegans*); and various sources (given in text) for *Picea engelmannii*. As in Table 1, the shape parameter is denoted S while the location parameter is L. For L for the Weibull, the values given have been multiplied by 10 000.  $\bar{x}$  is the mean distance travelled by seed or seedling

Species	Model	L	S	Likelihood	$\bar{x}$
<i>Pinus</i> no. 1	Lognormal	10.4	0.71	-12.88	13.4
	Weibull	86	1.86	-14.41	12.1
	2Dt	686	3.93	-14.29	13.0
<i>Pinus</i> no. 2	Lognormal	95	1.20	-24.07	195
	Weibull	0.70	2.96	-22.14	18.9
	2Dt	2583	2.99	-23.34	30.0
<i>Pinus</i> stand	Lognormal	39	1.10	-158.23	72
	Weibull	0.90	1.79	-160.76	174
	2Dt	279 141	7.85	-160.70	176
<i>Tachigalia</i> no. 1	Lognormal	34	0.85	-91.70	48.3
	Weibull	7.04	2.02	-72.13	31.9
	2Dt	15 916	12.7	-70.77	32.3
<i>Tachigalia</i> no. 2	Lognormal	34	0.89	-56.23	50.1
	Weibull	83.8	1.42	-68.97	34.8
	2Dt	1824	1.60	-60.94	40.3
<i>Tilia</i>	Lognormal	50	0.95	-30.76	78.6
	Weibull	12.2	1.80	-34.76	40.1
	2Dt	6665	3.53	-33.94	43.3
<i>Betula uber</i>	Lognormal	331	1.56	-15.43	> 1000
	Weibull	2719	0.53	-15.31	169
	2Dt	556	0.00065	-16.0	> 10 000
<i>Fraxinus</i>	Lognormal	46.7	0.84	-12.10	66.5
	Weibull	25.7	1.57	-14.59	47.6
	2Dt	3235	1.24	-14.30	68.7
<i>Lonchocarpus</i>	Lognormal	16.9	0.95	-19.82	26.5
	Weibull	1683	0.90	-20.04	18.1
	2Dt	94	0.37	-19.94	> 1000
<i>Platypodium</i>	Lognormal	55.2	0.93	-23.30	85
	Weibull	3.25	2.05	-27.71	43.7
	2Dt	10 934	4.5	-27.52	47.8
<i>Picea</i> area source	Lognormal	43.3	1.14	-115.86	83.0
	Weibull	20 843	0.31	-124.10	49.7
	2Dt	1667	0.97	-119.46	66.7

or area source (the *Picea* studies of dispersal from forest edges), the lognormal had a significantly higher likelihood than either the Weibull or 2Dt.

For the eight St Hilaire studies of seeds, and for the cones of *Tsuga*, differences among the three models in log likelihood were relatively minor, with the Weibull having the highest likelihood in four cases, the lognormal in three, and the 2Dt in two cases (Table 3). The differences among the three models in log likelihood for the analyses of the distribution of germinants were minor (< 2) for four of the eight cases. Where there were larger differences in likelihood, the lognormal was the best model in three of four cases, and the 2Dt model was superior in the remaining case (Table 3).

## Discussion

### PROBLEMS WITH ESTIMATION

Each of the three models at times produced unlikely results. For example, the 2Dt produced absurdly exten-

sive tails ( $\bar{x} > 10$  km) for 5 of 17 St Hilaire studies and one (*B. uber*) of the discrete studies. This occurred when the shape parameter, S, became very small (typically S was much less than 0.01). This problem may have arisen just as frequently in the original data set of Clark *et al.* (1999), although we cannot be sure as the authors arbitrarily constrained the parameter to values  $S > 0.5$ .

In a few cases the lognormal was capable of producing an S value so small relative to the mode that the recruitment curve was, essentially, a torus around the source tree. There are two examples of this toroidal shape in Table 3 (*Tsuga* germinants and *Quercus*). It is possible that these examples represent severe predation on seeds and seedlings near parent boles. However, none of the point source evaluations led to a torus.

When the Weibull produced a very large S it must be regarded as an unlikely result as it precludes long-distance dispersal (e.g. *Pinus* no. 2, see Table 2). In general, we suggest that the high fixed value of S (set to 3) in Ribbens *et al.* (1994) significantly underestimates

**Table 3** Comparison of the log likelihood for the three dispersal functions ( $f(x)$ ): the 2Dt, the lognormal, and the Weibull, using St Hilaire studies of the 1996 seed crop (seeds), 1997 germinants arising from that crop, and 2001 germinants. Ovulate cones of *Tsuga canadensis* are included. Parameter symbols are as in Table 1. L is the scale parameter; S is the shape parameter; b is the exponent on basal area (equation 1). The L values for the Weibull have been multiplied by 10 000. The highest likelihood models are indicated in bold.  $\bar{x}$  is the estimated mean distance

Species	$f(x)$	L	S	Likelihood	$\bar{x}$	b
<i>Acer pennsylvanicum</i> (seeds)	2Dt	132	0.0063	<b>-100.10</b>	10 000	5.04
	Lognormal	84.0	1.44	-101.95	237	4.0
	Weibull	11 426	0.36	-103.07	120	4.0
<i>Acer rubrum</i> (seeds)	2Dt	1.65	0.003	-34.63	> 10 000	0.99
	Lognormal	11.1	1.72	<b>-34.62</b>	48.9	0.82
	Weibull	17 512	0.378	-35.87	27.0	1.16
<i>Acer saccharum</i> (seeds)	2Dt	3439	2.96	-85.92	34.8	2.66
	Lognormal	5.4	1.09	-87.92	27.9	0.0003
	Weibull	139	1.26	<b>-83.98</b>	41.7	2.93
<i>Betula alleghaniensis</i> (seeds)	2Dt	358	1.0	-54.45	29.5	1.32
	Lognormal	20.8	0.983	-54.32	33.8	1.78
	Weibull	2304	0.76	<b>-54.0</b>	26.9	1.82
<i>Betula populifolia</i> (seeds)	2Dt	1999	1.47	-57.55	45.5	2.90
	Lognormal	34.6	0.72	-57.7	44.9	2.29
	Weibull	0.97	2.42	<b>-56.2</b>	36.2	2.20
<i>Fagus grandifolia</i> (seeds)	2Dt	1280	2.38	-80.77	24.7	0
	Lognormal	22.1	0.58	<b>-80.62</b>	26.1	0
	Weibull	272	1.26	-83.37	24.1	0
<i>Ostrya virginiana</i> (seeds)	2Dt	120.5	0.91	-41.03	19.8	2.36
	Lognormal	11.5	0.81	<b>-39.76</b>	15.9	2.58
	Weibull	2323	0.84	-42.45	16.9	2.47
<i>Tsuga canadensis</i> (seeds)	2Dt	136 366	25.47	<b>-313.22</b>	65.8	1.05
	Lognormal	70.2	0.80	-320.03	96.7	1.26
	Weibull	44.2	1.34	-315.72	70.94	0.968
<i>Tsuga canadensis</i> (cones)	2Dt	2499	4.99	-389.65	21.5	1.36
	Lognormal	15.4	0.69	-392.20	19.6	1.60
	Weibull	64.12	62.65	<b>-384.44</b>	24.0	1.10
<i>Acer pennsylvanicum</i> (1997 germinants)	2Dt	13 234	2.41	-26.51	78.7	13.8
	Lognormal	63.7	0.58	<b>-25.53</b>	75.4	15.6
	Weibull	103	1.20	-27.09	66	10.1
<i>Acer pennsylvanicum</i> (2001 germinants)	2Dt	187	0.0001	-66.74	> 10 000	16.8
	Lognormal	18.6	1.94	<b>-64.02</b>	122	19.7
	Weibull	0.73	0.21	-65.09	108	22.4
<i>Acer saccharum</i> (1997 germinants)	2Dt	8797	0.29	-80.76	> 10 000	5.08
	Lognormal	46.7	0.31	<b>-77.27</b>	49	3.17
	Weibull	0.87	1.26	-81.09	124	4.64
<i>Fagus grandifolia</i> (2001 germinants)	2Dt	123	0.01	<b>-67.02</b>	> 10 000	1.89
	Lognormal	13.4	1.91	-67.40	82.8	1.02
	Weibull	0.53	0.63	-67.95	59	0.60
<i>Ostrya virginiana</i> (2001 germinants)	2Dt	191	2.34	-20.41	10	2.75
	Lognormal	6.8	0.21	<b>-20.40</b>	7	2.32
	Weibull	16 897	0.39	-20.71	23	3.53
<i>Pinus resinosa</i> (2001 germinants)	2Dt	0.32	1.17	-33.93	14	0.73
	Lognormal	9.4	0.81	<b>-33.45</b>	13	0.70
	Weibull	243	1.47	-35.02	14	1.15
<i>Quercus rubra</i> (1997 germinants)	2Dt	16 011	2.98	-79.86	74.9	0.03
	Lognormal	96.3	0.018	<b>-72.94</b>	96.3	3.97
	Weibull	80.1	1.22	-79.56	73.9	0.15
<i>Tsuga canadensis</i> (1997 germinants)	2Dt	3276	27.3	-116.37	9.8	0.0037
	Lognormal	9.4	0.29	<b>-109.49</b>	9.8	0.059
	Weibull	1.66	2.81	-147.09	16.7	0.50

the true dispersal capacity of the tree species concerned, for instance the very small  $\bar{x}$  of 4 m for *T. canadensis*. Ribbens *et al.* (1994) suggested that this result was a consequence of seeds tending to remain in the ovulate cones. However, when we analysed our data using their methods (20-m search radius, Weibull with S fixed at 3.0, etc.), intact hemlock cones at St Hilaire had an  $\bar{x}$  of

25 m, while the seeds from that same crop had an estimated  $\bar{x}$  of 85 m. Further, an examination of 75 *Tsuga* ovulate cones on the snow in mid-April 1997 indicated that only about 5% of the viable seeds were still inside the cones.

If we are to include recruitment subroutines in stand dynamics simulations, then, despite the problems with

all three dispersal terms, we must choose a function and parameterize it. Our results suggest that the lognormal is to be preferred to other two-parameter functions for wind dispersed species and, perhaps, for passively dispersed animals such as mites (Jung & Croft 2001). There is considerable empirical support for the lognormal function and it also has the merit of serving as a platform for mechanistic interpretation of seed dispersal (Greene & Calogeropoulos 2002), while dispersal functions such as the Weibull and  $2Dt$  are opaque to such interpretations (Nathan *et al.* 2002a). Clark *et al.* (1999, p. 1477) argued that the parameters required for mechanistic interpretation of dispersal within forests are 'difficult to obtain', although default turbulence parameters for averaged conditions (as opposed to any one short event) are readily available for forest conditions (e.g. Greene & Johnson 1996; Nathan *et al.* 2002a; Calogeropoulos *et al.* 2003). Further, while such averages undoubtedly lead to problems in prediction of dispersal by a particular tree in a particular place and time, the alternative would be to extrapolate empirical data from a limited field study to other places and times.

Greene & Johnson (1989) have argued that dispersal by wind can be well-approximated by a lognormal function (*long-term* horizontal wind speeds are assumed to be lognormally distributed, as with Nathan *et al.* 2001 or Greene & Johnson 1989), and that, *contra* Clark *et al.* (1999), a single tree within a forest will behave acceptably as a point source (Nathan *et al.* 2002b). Random walk models of dispersal (e.g. Tackenberg 2003), the individual trajectory approach of Nathan *et al.* (2002b), or the closed-form advection-diffusion function of Okubo & Levin (1989) lead to dispersal curves that can be expressed as a lognormal. The lognormal has two parameters (Table 1). For tree seeds, the location parameter  $L$  is the median distance travelled by a seed crop from a single tree (a simple function of the terminal velocity, median horizontal and vertical wind speeds, and the release height). The shape parameter  $S$  is the standard deviation of the logarithms of the distances travelled. This variation, largely a function of vertical and horizontal winds (cf. Greene & Johnson 1992; Nathan *et al.* 2002b; Tackenberg 2003), is expected to place the mode within a few metres of the source tree. Meanwhile, the typical ranges of tree height, intraforest wind speed and terminal velocity for wind-dispersed tree species lead to the conclusion that the median distance,  $L$ , should be 15–45 m from a tree (Greene *et al.* 1999). Most of the  $S$  values in Tables 2 and 3 are around 1.0, and one might recommend  $S = 1$  as a default value for wind-dispersed tree species in closed forests.

Clark *et al.* (1999) argued that the meteorological approach outlined by Nathan *et al.* (2002a) and others was incorrect for the near tail because the individual tree canopy was not truly a point source (as required by the simple formulation above). We suggest that this criticism is relatively unimportant, as the great majority of tree seeds are produced in the upper fifth of the forest

(references in Greene & Johnson 1996) and crowns within a closed forest are quite narrow (Canham *et al.* 1994, 1999). Narrow crowns with seeds concentrated near the top will produce a population of dispersal trajectories that only trivially pushes the modal density back towards the bole. Indeed, we argue just the opposite of Clark *et al.* (1999): an open-grown tree (or, for that matter, an open-grown shrub, such as used by Bullock & Clarke 2000) is more likely to have a wide variation in release height and lateral extent and so is more likely to depart from the point source idealization.

Another criticism of Clark *et al.* (1999), at least for wind dispersed trees, is that the lognormal function demands a 'skip distance', i.e. an empty interval between the bole and the first censused seeds or seedlings. They argue that this 'skip distance is not expected in real stands' (Clark *et al.* 1999, p. 1477). But typical  $S$  values in the lognormal are sufficiently large (around 1.0) that we would expect (with, say, a median distance,  $L$  in Table 1, of 30 m and  $Q = 10\,000$ ) about one seed  $m^{-2}$  as close as 0.5 m to the bole, and thus there would not be any dramatic 'skip distance'. Further, although the lognormal does not permit any seeds to fall at  $x = 0$  m, one might regard this, perhaps pedantically, as a touch of added realism: the Weibull and  $2Dt$  unrealistically place the mode exactly at  $x = 0$ , which is, by definition, the centre of the bole at ground level.

In the case of dispersal of seeds by animals, the simplest mechanistic models (e.g. Murray 1988) are based on the time ( $t$ ) of the seed travel, and are fundamentally of the form:

$$f(x) = dQ/dx = dt/dx dQ/dt$$

where  $dt/dx$  is the net velocity of the animal away from the maternal plant and  $dQ/dt$  is the passage time through or on the animal (Murray 1988). Given the *multiplicative* process explicit in the above equation, we expect a right-skewed curve (e.g. the simulation results of Murray (1988) for birds dispersing tree fruits), and thus a mode shifted away from the source. Such a curve can be approximated by a lognormal (Greene & Calogeropoulos 2002) if movement is not limited by small territory size.

One last useful feature of the lognormal is that by permitting the modal density to occur away from the bole it is the only one of the three functions entertained here that could accommodate large Janzen-Connell effects. In the tropics this is certainly an important characteristic given the strong density-dependence observed there (Harms *et al.* 2000).

#### A SPECULATION CONCERNING COLLISIONS

The  $2Dt$  and Weibull perform relatively better with the inverse modelling within diffuse sources rather than with discrete sources. One wonders why. We can imagine a second process that would *seemingly* place the mode near the source tree but it would only be an

artifact of the inverse modelling approach. As shown by Greene (1990), an autorotating winged seed colliding with a branch will, at worse, suffer a slight vertical acceleration before the terminal velocity and ambient horizontal speeds are re-achieved. After collision with a bole, however, a propagule such as an *Acer* samara can again reach autorotation, but it will, if the wind direction does not suddenly change, simply encounter the bole again at a lower height. This general argument ought to apply to the more massive plumed or winged seeds (but not to low-inertia propagules such as pollen or spores or the least massive seeds). This trapping effect of woody plants has been demonstrated recently by Bullock & Moy (2004), with the crowns of shrubs acting as 'traps' for wind-dispersed seeds within a matrix of short grasses. We can estimate the probability of collision with a bole in a simple quantitative way as follows. Let us idealize flight trajectories as rectilinear paths starting at the abscission height ( $z_a$ ) and descending into the bole space only at some lower height ( $z_b$ ; thus the bole space extends from  $z = 0$  to  $z_b$ ). Above the bole space we have ramified branches that cannot permanently stop a seed's further travel, while within it we have only untapered vertical boles of diameter  $d$ . The constant probability of collision with a bole per metre of travel is  $dN_D$  (where  $N_D$  is the density of canopy trees). For a seed that would have travelled to the distance  $x$  in the absence of collisions, the descent into the bole space begins at the earlier distance  $b$  (and  $b = x(1 - (z_b/z_a))$ ). Thus for a seed the probability of colliding ( $p$ ) during dispersal rises with  $x$  as:

$$P = 1 - (1 - (dN_D))^{g x} \quad \text{eqn 4}$$

where  $g = xz_b/z_a$ . Further, the fraction of the  $Q$  crop involved in collisions ( $Q_c/Q$ ) is:

$$Q_c/Q = \int_b^x f(x)p \quad \text{eqn 5}$$

Let us assume a forest with  $N_d = 0.04 \text{ m}^{-2}$  (400 canopy trees  $\text{ha}^{-1}$ ) and  $d = 0.3 \text{ m}$  (thus about  $28 \text{ m}^2$  of basal area  $\text{ha}^{-1}$ ). While much depends on species and long-term disturbance history, a reasonable range for  $z_b$  (the height of the vertical bole space) is 5–10 m. According to equation 5 this means that 13–23% of seeds will be deposited at the bases of canopy trees via collisions. Some of these trees will, of course, be conspecifics (otherwise one need not have bothered with inverse modelling), and  $f(x)$  such as the Weibull or 2Dt, which demand a modal deposit at the base of the tree, will misconstrue the seeds as representing poor dispersal. Thus, when we use inverse modelling these relatively poor  $f(x)$  will begin to rival the lognormal as a predictor, and they will seriously underestimate the mean distance travelled. Interestingly, after running a simulation with the lognormal ( $L = 30 \text{ m}$  and  $S = 1$ ) and allowing for collisions, the shape of individual dispersal curves is still reasonably lognormal. This is because the great

majority of the collisions occur near  $x$  rather than much earlier due to the fact that much of the initial trajectory occurs above the trunk space.

Note that something analogous could occur for endozoochory with bats or birds roosting in a tree of the same species as that from which they have recently eaten fruits (Schupp *et al.* 2002). The deposited seeds will certainly be interpreted by the 2Dt or Weibull as very poorly dispersed diaspores derived from the individual tree used for roosting.

In summary, the lognormal was the highest likelihood model in 18 of the 28 analyses of both point source and inverse modelling data sets, but the differences in likelihood among the three models were often very small.

## Acknowledgements

Field assistance was provided by Charlotte Lamontagne, Emma Arnold, Annick Malletto, Stephane Auger and Melanie McCavour. The work was funded by NSERC and the Sustainable Forest Management Network. We thank the two referees and the editor for their very useful suggestions.

## References

- Augsburger, C.K. (1983) Seedling escape from fungal pathogens. *Journal of Ecology*, **71**, 759–772.
- Augsburger, C.K. & Hogan, K.P. (1983) Wind dispersal of fruits with variable seed number in a tropical tree (*Lonchocarpus pentaphyllus*, Leguminosae). *American Journal of Botany*, **70**, 1031–1037.
- Bullock, J.M. & Clarke, R.T. (2000) Long distance seed dispersal by wind: measuring and modelling the tail of the curve. *Oecologia*, **124**, 506–521.
- Bullock, J.M. & Moy, I.L. (2004) Plants as seed traps: interspecific interference with dispersal. *Acta Oecologia*, **25**, 35–41.
- Calogeropoulos, C., Greene, D.F., Messier, C. & Brais, S. (2003) Refining tree recruitment models. *Canadian Journal of Forest Research*, **33**, 41–46.
- Canham, C.D., Coates, K.D., Bartemucci, P. & Quaglia, S. (1999) Measurement and modeling of spatially-explicit variation in light transmission through interior cedar-hemlock forests of British Columbia. *Canadian Journal of Forest Research*, **29**, 1775–1783.
- Canham, C.D., Finzi, A.C., Pacala, S.W. & Burbank, D.H. (1994) Causes and consequences of resource heterogeneity in forests: interspecific variation in light transmission by canopy trees. *Canadian Journal of Forest Research*, **24**, 337–349.
- Clark, J.S., Silman, M., Kern, R., Macklin, E. & HilleRis-Lambers, J.H. (1999) Seed dispersal near and far: patterns across temperate and tropical forests. *Ecology*, **80**, 1475–1494.
- Crossley, D.I. (1955) *The Production and Dispersal of Lodgepole Pine Seed*. Technical Note Number 25. Forest Research Division, Department of Northern Affairs and Natural Resources, Forestry Branch, Ottawa.
- Ford, R.H., Sharik, T.L. & Feret, P.P. (1983) Seed dispersal of the endangered Virginia round-leaf birch (*Betula uber*). *Forest Ecology and Management*, **6**, 115–128.
- Goffe, W.L., Ferrier, G.D. & Rogers, J. (1994) Global optimization of statistical functions with simulated annealing. *Journal of Econometrics*, **60**, 65–99.

- Greene, D.F. (1990) The aerodynamics and dispersal of winged and plumed seeds. PhD dissertation, Department of Biology, University of Calgary, Calgary, Alberta, Canada.
- Greene, D.F. & Calogeropoulos, C. (2002) Dispersal of seeds by animals and wind. *Dispersal Ecology* (eds J. Bullock, R. Kenward & R. Hails.), pp. 3–23. Blackwell Press, Oxford.
- Greene, D.F. & Johnson, E.A. (1989) A micrometeorological model of the dispersal of winged or plumed seeds from a point source. *Ecology*, **70**, 339–347.
- Greene, D.F. & Johnson, E.A. (1992) Can the variation in samara mass and terminal velocity on an individual plant affect the distribution of dispersal distances? *American Naturalist*, **139**, 825–838.
- Greene, D.F. & Johnson, E.A. (1996) Wind dispersal of seeds from a forest into a clearing. *Ecology*, **77**, 595–609.
- Greene, D.F., Zasada, J.C., Sirois, L., Kneeshaw, D., Morin, H., Charron, I. et al. (1999) A review of the regeneration dynamics of boreal forest tree species. *Canadian Journal of Forest Research*, **29**, 824–839.
- Harms, K., Wright, S.J., Calderon, O., Hernandez, A. & Herre, E.A. (2000) Pervasive density-dependent recruitment enhances seedling diversity in a tropical forest. *Nature*, **404**, 493–495.
- Hooper, M.C., Arai, K. & Lechowicz, M.J. (2001) Impact of a major ice storm on an old-growth hardwood forest. *Canadian Journal of Botany*, **79**, 70–75.
- Jung, C. & Croft, B.A. (2001) Aerial dispersal of phytoseiid mites (Acari: Phytoseiidae): estimating falling speed and dispersal distances for adult females. *Oikos*, **94**, 182–190.
- Kitajima, K. & Augspurger, C.K. (1989) Seed and seedling ecology of a monocarpic tropical tree, *Tachigalia versicolor*. *Ecology*, **70**, 1102–1114.
- LePage, P.T., Canham, C.D., Coates, K.D. & Bartemucci, P. (2000) Seed abundance versus substrate limitation of seedling recruitment in northern temperate forests of British Columbia. *Canadian Journal of Forest Research*, **30**, 415–427.
- Melanson, S. & Lechowicz, M.J. (1987) Differences in the damage caused by glaze ice on codominant *Acer saccharum* and *Fagus grandifolia*. *Canadian Journal of Botany*, **65**, 1157–1159.
- Murray, K.G. (1988) Avian seed dispersal of three neotropical gap-dependent plants. *Ecological Monographs*, **58**, 271–298.
- Nathan, R., Horn, H.S., Chave, J. & Levin, S.A. (2002a) Mechanistic models for tree seed dispersal by wind in dense forest and open landscapes. *Seed Dispersal and Frugivory: Ecology, Evolution, and Conservation* (eds D.J. Levey, W.R. Silva & M. Galetti.), pp. 69–82. CAB International Press, Wallingford.
- Nathan, R., Katul, G., Horn, H.S. & Levin, S.A. (2002b) Mechanisms of long-distance dispersal of seeds by wind. *Nature*, **418**, 409–413.
- Nathan, R. & Muller-Landau, H. (2000) Spatial patterns of seed dispersal, their determinants and consequences for recruitment. *Trends in Evolution and Ecology*, **15**, 278–285.
- Nathan, R., Safriel, U.N. & Noy-Meir, I. (2001) Field validation and sensitivity analysis of a mechanistic model for tree seed dispersal by wind. *Ecology*, **82**, 374–388.
- Okubo, A. & Levin, S.A. (1989) A theoretical framework for data analysis of wind dispersal of seeds and pollen. *Ecology*, **70**, 329–338.
- Proulx, O. & Greene, D.F. (2001) The effect of the magnitude of ice accretion and tree size on tree damage during ice storms. *Canadian Journal of Forest Research*, **31**, 1758–1767.
- Ribbens, E.J., Silander, J.A. & Pacala, S.W. (1994) Seedling recruitment in forests: calibrating models to predict patterns of tree seedling dispersion. *Ecology*, **75**, 1794–1806.
- Ronco, F. (1970) *Engelmann Spruce Seed Dispersal and Seedling Establishment in Clearcut Forest Openings in Colorado – a Progress Report*. Research Note Number RM-168. Rocky Mountain Forest and Range Experiment Station, Fort Collins, Colorado.
- Rudis, V.A., Ek, A.R. & Balsiger, J.W. (1978) Within-stand seedling dispersal for isolated *Pinus strobus* within hardwood stands. *Canadian Journal of Forest Research*, **8**, 10–13.
- Schupp, E.W., Milleron, T. & Russo, S.E. (2002) Dissemination limitation and the origin and maintenance of species-rich tropical forests. *Seed Dispersal and Frugivory: Ecology, Evolution, and Conservation* (eds D.J. Levey, W.R. Silva & M. Galetti.), pp. 69–82. CAB International Press, Wallingford.
- Squillace, A.E. (1954) *Engelmann Spruce Seed Dispersal into a Clear-Cut Area*. Research Note Number INT-11. Intermountain Forest and Range Experiment Station, Ogden, Utah.
- Stoyan, D. & Wagner, S. (2001) Estimating the fruit dispersion of anemochorous trees. *Ecological Modelling*, **145**, 35–47.
- Tackenberg, O. (2003) Modelling long distance dispersal of plant diaspores by wind. *Ecological Monographs*, **73**, 173–189.
- Tanaka, H., Shibata, M. & Nakashizuka, T. (1998) A mechanistic approach for evaluating the role of wind dispersal in tree population dynamics. *Journal of Sustainable Forestry*, **6**, 155–174.
- Wood, O.M. (1932) An example of white pine reproduction on burned lands in northeastern Pennsylvania. *Journal of Forestry*, **30**, 838–845.

Received 26 January 2004

revision accepted 15 June 2004

Handling Editor: Mark Williamson