



A model of hardwood tree colonization among forest fragments: predicting migration across human-dominated landscapes

Nina Hewitt, Guy R. Larocque, David Greene & Martin Kellman

To cite this article: Nina Hewitt, Guy R. Larocque, David Greene & Martin Kellman (2019) A model of hardwood tree colonization among forest fragments: predicting migration across human-dominated landscapes, *Écoscience*, 26:1, 35-51, DOI: [10.1080/11956860.2018.1515596](https://doi.org/10.1080/11956860.2018.1515596)

To link to this article: <https://doi.org/10.1080/11956860.2018.1515596>



View supplementary material [↗](#)



Published online: 14 Sep 2018.



Submit your article to this journal [↗](#)



Article views: 96



View related articles [↗](#)



View Crossmark data [↗](#)

ARTICLE



A model of hardwood tree colonization among forest fragments: predicting migration across human-dominated landscapes

Nina Hewitt^a, Guy R. Larocque^b, David Greene^c and Martin Kellman^d

^aGeography Department, University of British Columbia, Vancouver, BC, Canada; ^bCanadian Forest Service Natural Resources Canada, Laurentian Forestry Centre, Quebec City, QC, Canada; ^cDepartment of Forestry and Wildland Resources, Humboldt State University, Arcata, CA, USA; ^dDepartment of Geography, York University, Toronto, Canada

ABSTRACT

We developed a model of hardwood tree colonization in forest patches. We began with a basic model of species' recruitment density calculated as a function of seed production and juvenile survivorship. Survivorship probability was expressed as a function of seed size, using seed-sowing data for a wide variety of species. To account for dispersal, we used an exponential distance-decay function based on empirical colonization data for species classified by dispersal mechanism and seed mass. The basic model reasonably predicted observed recruitment densities at or near forest edges, except for small seeded, wind-dispersed species with strong establishment constraints, for which it over-predicts. Our dispersal term yielded predictions that were not statistically different from observed colonization. However, species with large seeds and unspecialized dispersal mechanism appear to have distinct thresholds beyond which *no* dispersal occurs. Further research should better account for establishment constraints among small-seeded species, dispersal constraints among large-seeded species and unspecialized dispersers, and improve the dispersal functions to better reflect vectors such as birds. Nevertheless, the present model is adequate for the prediction of colonization probabilities in fragmented forests, requiring only an estimate of the abundance of source trees of a species and the mean inter-fragment distances.

RÉSUMÉ

Nous avons développé un modèle de colonisation d'arbres feuillus dans des parcelles forestières. Nous avons commencé avec un modèle de base de densité de recrutement d'espèces calculé en fonction de la production de graines et du taux de survie juvénile. La probabilité du taux de survie a été exprimée en fonction de la dimension des graines, en utilisant des données d'ensemencement pour une grande variété d'espèces. Afin de tenir compte de la dispersion, nous avons utilisé une fonction exponentielle de distance-décroissance basée sur des données empiriques de colonisation pour des espèces classées selon le mécanisme de dispersion et la masse des graines. Le modèle de base a raisonnablement prédit les densités de recrutement observées en bordure forestière ou à proximité, excepté pour les espèces à petites graines dispersées par le vent avec de fortes contraintes d'établissement, pour lesquelles le modèle a surestimé. Notre terme de dispersion a produit des prédictions qui n'étaient pas statistiquement différentes de la colonisation observée. Cependant, les espèces à grosses graines et des mécanismes de dispersion non spécialisés semblent avoir des seuils au-delà desquels aucune dispersion ne se produit. Des recherches additionnelles devraient mieux tenir compte des contraintes d'établissement parmi les espèces à petites graines, des contraintes de dispersion parmi les espèces à grosses graines et des mécanismes de dispersion non spécialisés, et améliorer les fonctions de dispersion pour mieux refléter les vecteurs comme les oiseaux. Néanmoins, le présent modèle est adéquat pour la prédiction de probabilités de colonisation dans des forêts fragmentées, nécessitant seulement une estimation de l'abondance d'arbres sources d'une espèce et des distance moyennes entre les fragments.

ARTICLE HISTORY

Received 25 March 2018
Accepted 31 July 2018

KEYWORDS

Eastern forests; tree colonization; seed dispersal; forest fragmentation; tree migration; seed size



MOTS-CLÉS


forêts de l'est; colonisation arborescente; dispersion des graines; fragmentation forestière; migration arborescente; taille des graines

Introduction

To remain competitive, plant species must extend their ranges poleward at velocities commensurate with the rate of climate warming (Loarie et al. 2009; Sandel et al.

2011). However, not only is the present rate of warming faster than in the early Holocene (Malcolm et al. 2002; Chen et al. 2011), but human-induced ecosystem fragmentation increases the challenge since migration

CONTACT Nina Hewitt  nina.hewitt@ubc.ca  Geography Department, University of British Columbia, Vancouver Campus, 240A-1984 West Mall V6T 1Z2, Vancouver, BC, Canada

 Supplemental data can be accessed [here](#).

2018 Université Laval

will have to occur via inter-fragment dispersal rather than across continuous habitats (Pearson and Dawson 2005). Thus, migration during the next few centuries will need to occur at rates greater than those inferred from palynological studies of the Holocene (e.g. Davis 1981) and under altered biogeographical contexts.

While predictive models will be essential to estimate species' migration potentials (Nathan et al. 2003; Trakhtenbrot et al. 2005; Levey et al. 2008; Morin et al. 2008), these models presently share a number of limitations. First, they invariably lack validation using empirical or experimental data on inter-patch seed dispersal (Neilson et al. 2005; Pearson and Dawson 2005; Jones and Muller-Landau 2008; Petit et al. 2008). In particular, there is little knowledge of the behaviour of seed-depositing animal vectors moving among fragments (Anand and Langille 2010; Nathan et al. 2008; Cousens et al. 2010). There is evidence that birds disperse seed distances of several hundreds of metres in temperate forests (van der Pijl 1982; Johnson and Adkisson 1985; Dzwonko and Loster 1992), while non-volant dispersers such as rodents achieve dispersal distances over only few tens of metres (Sork 1984; Abbott and Quink 1970), but how volant and non-volant animal dispersers affect species' colonization abilities needs quantification. A second problem is that almost all available models address only the dispersal phase and ignore the subsequent establishment phase of colonization (e.g. Malcolm et al. 2002; Dullinger et al. 2004; Pearson and Dawson 2005). This is important because, for wind-dispersed tree species at least, there will be an interdependence, mediated by seed mass, among dispersal and juvenile survivorship. As shown by Greene and Johnson (1992, 1993), seed mass controls the terminal velocity of wind-dispersed diaspores, and thus is strongly negatively related to dispersal capacity. For animal dispersal, there is evidence that heavy-seeded species may achieve only limited distances of transport (Ranney and Johnson 1977; Darley-Hill and Johnson 1981; Hewitt and Kellman 2002b), though Thomson et al. (2011) found no relationship between dispersal and seed size. At the same time, for all vectors, seed mass is proportional to juvenile survivorship (Greene and Johnson 1998, 2000; Moles and Westoby 2006), particularly in shady forest environments (Hewitt 1998; Bruun and Ten Brink 2008).

Any complete model of recruitment by seed necessarily has a minimum of three components. These are source strength (a function of mean plant seed production and source plant density), dispersal (which dilutes the source strength in space), and juvenile survivorship. To fill the knowledge gaps identified above, in this

paper, we develop and test a full model of inter-fragment tree colonization in a southern Ontario hardwood forest for a large number of animal- and wind-dispersed tree species, varying dramatically in seed mass. To test and calibrate the model, we employ one of the few available, comprehensive empirical and experimental datasets on inter-fragment medium- to long-distance temperate tree dispersal and colonization (Hewitt and Kellman 2002a, 2002b, 2004).

Materials and methods

Empirical and experimental data

Data were obtained from empirical and experimental studies of tree species composition, seed dispersal and colonization in a fragmented eastern deciduous forest system near Lake Erie, Southern Ontario (42° 40' N; 80° 20' W). Forest fragments existed in a matrix of rural, agricultural land, that varied in size between 1 and 49 ha separated by mean distances of 215 m to the four nearest neighbouring woodlots. The sites were colonized by about 40 hardwood species, which represented much of the regional tree species diversity.

We measured hardwood colonization using established seedlings growing in pine (*Pinus strobus* and *P. resinosa*) plantations, the 'seed traps' in our study system (Hewitt and Kellman 2002a; and see Appendix). The seed sources were the reproductively mature conspecifics in the surrounding landscape. Observed recruitment density in our model (F_D) was the mean species' seedling density recorded in 16, 100 × 16 m plots positioned lengthwise along the edges of 12 pine plantations. Recruits were smaller than 1.5 cm in diameter at 1 cm above root collar level and ring counts of up to 5 stems of each of 21 hardwood species indicated that these ranged in age from 4 to 11 years with a median age of 7.9 years (Hewitt and Kellman 2002a).

Conspecific source tree density (N_D) was measured within 25 metre-wide distance zones to a distance of 150 m from the edges of the plots in plantations (Appendix). Earlier analyses indicated that N_D was significantly related to seedling densities in plantations (F_D) (Hewitt and Kellman 2002a), supporting the assumption that N_D generally reflected seed supply to plantation 'seed traps'. Due to the large number of trees being enumerated, and because tree canopies were often out of visible range, reproductive maturity was indicated by stem diameter based on information in the literature (Burns and Honkala 1990). Minimum sizes for tree enumeration within source areas were greater than or equal to 10 cm in diameter at breast

height (dbh) for canopy species and greater than or equal to 5 cm in dbh for tolerant species characteristically limited to the understory. Some individuals may begin to set seed when they are as small as 5 cm dbh, if open-grown. Smaller individuals of all species were checked for signs of flowering and fruiting and this confirmed that few were reproducing, and any that did appear to be were included in the stem counts (Hewitt 1999). Source trees were typically contained in woodlots, either isolated from or connected to plantations, but also within fencerows, riparian patches, or, very occasionally, as isolated individuals in a non-forested surrounding landscape. Reproductive individuals were occasionally present in older plantations that were invaded by hardwoods for longer periods and were included within the 0–25 m distance zone.

Basal area (B_m) was measured on extensive point-quarter samples in the study area that encompassed fragments within which tree densities for N_D had been enumerated (Hewitt 1999; Richart and Hewitt 2008). Mean seed mass (m) was determined using the USDA compendium of Schopmeyer (1974), which lists air-dried masses (dried at 20°C for 3 months). Seed mass includes the embryo and endosperm plus seed coat. For *Acer*, *Fraxinus* and *Liriodendron*, the mass estimates include persistent woody wing material, and these values have been reduced by 15% (Greene and Johnson 1994).

Seedling survivorship (S) was determined from experimental sowing results on 10 hardwood species in pine plantations (Hewitt and Kellman 2002a) and encompassed 2 years, from sowing to final census in August of the second growing season. Results were expressed as proportion of viable seed planted. Seed predation by large vertebrates, especially large rodents (squirrels, chipmunks) was controlled with 1.3-cm diameter wire mesh. A separate seed-sowing experiment on juvenile survivorship at forest fragment edges for *Fagus grandifolia*, *Carya cordiformis* and *Juglans nigra* indicated comparable survivorship for *F. grandifolia* and *J. nigra*, but *C. cordiformis* experienced significantly greater survival at woodlot edges. For the first two species, results from plantations and woodlots were pooled. Granivory by squirrels and other larger vertebrates on nut-tree species was measured in a separate experiment on three species spanning a range of seed sizes and genera: *J. nigra*, *C. cordiformis*, *F. grandifolia* (Hewitt and Kellman 2004) and these results were used to estimate losses for all nut tree species according to taxonomic and seed mass similarity (see Table 1). Additional details of the study system, field and lab methodology may be found in Hewitt and Kellman (2002a, 2002b, 2004).

The basic model

Our basic sexual recruitment model (equation 1) is a combination of three terms: Q_D , the density of seeds within the area source population; $f(x)$, seed production; x , the dispersal term which is the distance to an area source; and S , the juvenile survivorship rate (see Appendix). Thus, F_{Dx} , the density of recruits at distance x perpendicular from the edge of the area source, is given as:

$$F_{Dx} = Q_D f(x) S \quad (1)$$

Seed production. Greene and Johnson (1994) showed that the individual mean annual seed production (Q) for North American trees was a positive function of basal area (B) raised to the power 0.92, and an inverse function of seed mass, m , raised to the exponent -0.58 .

$$Q = 3067 m^{-0.58} B^{0.92} \quad (2a)$$

Muller-Landau et al. (2008) recently reported an exponent on seed mass for tropical tree seed production (-0.60) very close to the -0.58 we use here. Converting [2a] to the seed density of an area source (Q_D), we use the mean value of B to the power 0.92 and multiply by tree density (N_D). Thus,

$$Q_D = 3067 t m^{-0.58} N_D B_m^{0.92} \quad (2b)$$

Where t is the span of years over which the seed rain has been accumulated. We calculated B from species-specific diameter measures in our point quarter sample data (see *Empirical and Experimental Data*, above). For canopy species, only stems greater than 10 cm in dbh were used to compute basal area (as with the stem counts around plantations used to calculate N_D , above), to isolate reproductively mature stem sizes. Only stems greater than 5-cm dbh were used for four understory species (*Amelanchier arborea*, *Carpinus caroliniana*, *Ostrya virginiana*, *Cornus florida*) that reach reproductive maturity at smaller sizes. The number of stems captured in the point quarter sample varied among the species, and several species had few or no stems ($n < 5$) (Table 1). Species for which no stems were available in the sample (5 species, Table 1) were assigned the average B_m of congeneric species' means, except for *Liriodendron tulipifera*, for which congeneric species were lacking, and for which the average of all canopy species' means was used. In addition to the above, we computed a single mean B_m value for all canopy and understory species, respectively (B_m , Canopy spp. = average of all canopy species' means; B_m , Understory spp. = average of all understory species' means) and tested the use of this measure in the model, given the limited sample sizes for many species.

Table 1. Species attributes used in the seedling survivorship and colonization density calculations.

Species	Shade tolerance	Disperser	Seed mass (g)	ND (stems/m ²)	Proportion of seed producers	BD (m ² /m ²) = ND*Bm _{0.92}	n for Bm calculations	FD – Observed (stems/m ²)	Granivory (proportion)
<i>Acer negundo</i>	2	W	0.03264	0.000004518	0.47	0.000000242	1	0.00578750	-
<i>Acer rubrum</i>	2	W	0.016745	0.000960843	0.47	0.000074948	624	0.35792500	-
<i>Acer saccharinum</i>	2	W	0.2159	0.000220382	0.47	0.000073624	9	0.00081250	-
<i>Acer saccharum</i>	1	W	0.0629	0.000601908		0.000059859	55	0.04144583	-
<i>Amelanchier arborea</i>	1	B	0.00567	0.000022590		0.000000168	14	0.01219583	-
<i>Betula alleghaniensis</i>	3	W	0.001	0.000359438		0.000021811	26	0.01875000	-
<i>Carpinus caroliniana</i>	1	U	0.015	0.000088353		0.000000684	27	0.00045833	-
<i>Carya cordiformis</i>	4	R	2.91	0.000072791			0	0.01925417	0.292
<i>Carya ovata</i>	3	R	4.536	0.000058735		0.000001589	3	0.01335833	0.292
<i>Cornus florida</i>	1	B	0.1008	0.000239960		0.000002351	225	0.02785000	-
<i>Fagus grandifolia</i>	1	B	0.284	0.000360944		0.000021152	17	0.00920833	0.129
<i>Fraxinus amer/pennsyl</i>	3	W	0.03825	0.000629518	0.237	0.000035167	29	0.43875000	-
<i>Juglans cinerea</i>	4	R	15.12	0.000008534			0	0.00013750	0.878
<i>Juglans nigra</i>	4	R	11.34	0.000076807		0.000001461	1	0.00020833	0.878
<i>Liriodendron tulipifera</i>	4	W	0.02754	0.000038153			0	0.01837500	-
<i>Nyssa sylvatica</i>	2	B	0.58	0.000041165	0.378	0.000001452	5	0.00012500	-
<i>Ostrya virginiana</i>	1	U	0.015	0.000059237		0.000000534	22	0	-
<i>Populus deltoides</i>	5	W	0.00097	0.000034137	0.5		0	0.00187500	-
<i>Populus grandidentata</i>	5	W	0.00015	0.000152108	0.5	0.000013227	63	0.00034583	-
<i>Populus tremuloides</i>	5	W	0.00017	0.000049699		0.000003102	20	0	-
<i>Prunus serotina</i>	3	B	0.0945	0.000206325		0.000011783	83	0.10823333	-
<i>Quercus alba</i>	3	R	3.024	0.000512048		0.000049610	301	0.01250000	0.292
<i>Quercus bicolor</i>	3	R	3.77	0.000029116		0.000002400	10	0	0.292
<i>Quercus macrocarpa</i>	3	R	6.048	0.000014558		0.000000582	2	0.00086250	0.292
<i>Quercus rubra</i>	3	R	2.59	0.000405622		0.000055370	183	0.05370417	0.292
<i>Quercus velutina</i>	3	B	3.09	0.000324799		0.000043915	34	0.02025000	0.292
<i>Sassafras albidum</i>	4	B	0.091	0.000112952	0.5	0.000004889	16	0.01722917	-
<i>Tilia americana</i>	2	U	0.082	0.000058735		0.000007801	4	0.00025000	-
<i>Ulmus spp.</i>	3	W	0.0067	0.000099398		0.000003737	23	0.00135833	-

Note: Proportion of seed producing stems are those likely to be seed producers among dioecious and polygamo-dioecious species. Values are 1 (monoecious species) unless otherwise indicated. Data from Clark and Burns and Honkala (1990). Granivory is proportion of seed consumed by vertebrates among nut tree species' seed (Hewitt 1999). Shade-tolerance class was derived from Baker (1949). Where more than one tolerance class listed, Burns and Honkala (1990) was used to identify the class most applicable to the seedling stage. Classes are 1 very tolerant, 2 tolerant, intermediate, 4 intolerant, 5 very intolerant. Dispersers were: W wind, B bird, R rodent, U unspecialized, and represent the primary dispersal mechanism indicated for the species (Darley-Hill and Johnson 1981; Johnson and Adkisson 1985; Burns and Honkala 1990). Species that were both bird and rodent dispersed were classified as bird dispersed to indicate the potential for long-distance transport that was important in explaining inter-fragment dispersal migration.

To account for the lower proportional individual seed production in non-monoecious species, we adjusted N_D according to the fraction of females (for dioecious species), or potentially reproducing individuals in the population (for polygamo-dioecious species) reported in the literature. We employed the values of Clark et al. (2004) for *Nyssa sylvatica* and *Fraxinus spp.* and for dioecious species within the genus *Acer* (*A. negundo*, *A. rubrum* and *A. saccharinum*). For *Sassafras albidum* and *Populus spp.*, we set the proportion at 0.5, as reported to be approximately correct for these species (Burns and Honkala 1990) (Table 1).

Survivorship. Greene and Johnson (1998) used the forestry literature on experimental seeding to estimate juvenile survivorship as a function of seed mass for a wide range of North American species. Although granivory rates fluctuated widely among studies, they used

an average value of 0.43 for survival through the granivory stage. Their empirical result for undisturbed forest substrates was:

$$S = 0.43(1 - \exp(-0.33m^{0.76})) \quad (3)$$

Where 0.43 is an empirical constant representing seed loss to granivores. While the relationship between seed mass and survivorship is undoubtedly more complex than the expression in equation (3) (Moles and Westoby 2006), we will attempt to show here that seed mass functions provide a reasonable approximation for eastern North American hardwood species at least. To test the accuracy of the seed mass coefficient (0.33) and exponent (0.76) with respect to the study species, we determined the relationship between seed mass and survivorship for 10 study species for which we had obtained experimental survivorship data within

the study system. We removed the 0.43 granivory term when examining survivorship and seed mass since this had been controlled in our experiment using wire mesh (see *Statistical Analysis of the Basic Model*, below). Granivory was later added into the basic model to test its relationship to empirically observed colonization, although we applied our own term developed from experimental information in the study system. We can now recast equation (1) as:

$$F_{Dx=0} = 3067 m^{-0.58} N_D B_m^{0.92} G (1 - \exp(-0.33 m^{0.76})) t \quad (4)$$

Where $x = 0$ is an average distance to a woodlot edge of 0 m, and t time in years over which dispersal and colonization were measured, i.e., number of annual cohorts. Since the established seedlings used to measure colonization were uneven-aged recruits that had accumulated over a mean of 7.9 years (Hewitt and Kellman 2002a) and thus emphasized younger cohorts due to age-related mortality, we adjusted t when comparing to the field data. Recall that our survivorship equation reflects survival from seed to the seedling stage at the end of a second growing season when the majority of mortality would be expected to occur. However, particularly for establishment in shade (i.e., beneath conifers in the plantations), recruits face a thinning dynamic. Hett (1971), for sugar maple showed roughly a halving of numbers from the second to the 12th year in understory conditions. We had recruits varying in age from 2 to a median of 8 years (6 years beyond our juvenile S estimate). We therefore divided the cohort number by $\frac{1}{2}$ on the basis of Hett (1971), in effect dividing seed supply by half (alternatively, we could have halved the expected survivorship). Thus $t = 3.95$ ($7.9/2$), which we have rounded to $t = 4$.

G is vertebrate seed predation for large-seeded nut-tree species, as indicated in Table 1. Effects of invertebrate or small animal predation that apply to smaller-seeded species were encompassed in our experimental measures used to calibrate expected survivorship (S) since these animals would be able to circumvent wire mesh in our seed sowing experiment.

The empirical data used to assess the basic model represented pooled colonization densities in plantations for which the nearest woodlot edge (area source) was located to an average of 47.8 m away from the plot edge (Range: 0–210; Standard Deviation: 53.9 m) (Hewitt and Kellman 2002a). In evaluating the basic model against field data, we therefore adjusted for distance effects broadly by multiplying expected F_{Dx} by 0.205. This figure was based on Greene and Johnson's (1996) area source model for seeds

dispersing from a tree canopy roughly the same distance (43 m) into an adjacent area. The adjustment assumes that source individuals were distributed randomly away from the woodlot edges. The term reflects an average terminal velocity of 0.8 m/s for anemochorous species (Greene and Johnson 1996) and was also used to adjust for dispersal of zoochores. In the subsequent analysis of dispersal (below), we test for and incorporate dispersal relationships in the model.

Statistical analysis of the basic model

We modelled the relationship between seed mass and survivorship using equation (3) (Greene and Johnson 1998), a non-linear (log-log) relationship because we hypothesized that the rate of increase in survival probability would diminish with increasing seed mass. The granivory factor (0.43) was removed from the predicted survivorship equation because large granivores had been excluded in the seed sowing experiment with wire mesh. However, seed losses to invertebrates, fungi and small granivores capable of circumventing mesh (e.g., mice) would be encompassed in experimental results and the resulting model, though these losses would tend to be relatively minor in the post-dispersal stage at longer distances due to low conspecific seed densities (Greene and Johnson 2000; Hewitt and Kellman 2004). We revised model fit, first using a 3-parameter fit, and then assuming a constant asymptote with two fitted parameters (seed mass and experimentally observed survival). For dioecious species, N_D was calibrated according to proportion of seed producing adults in the average population (Table 1).

We used linear regression analysis to test for a log-log relationship among observed and expected colonization density (F_D). Three species (*Ostrya virginiana*, *Populus tremuloides* and *Quercus bicolor*) were not observed as seedlings in plantations and were excluded from analyses (Table 1). While zero values for these species may, in fact, represent valid observations indicating poor colonization ability, we could not rule out the possibility that the plantation environment was responsible for their absence. Thus, analyses were made for 27 of the 30 species enumerated in the source areas around plantations.

Dispersal

In order to evaluate effects of distance to seed sources on F_D and how these varied among species, we broke down and examined the empirical colonization data in relation to distance to nearest conspecific. Data on species' observed seedling densities in each of the 16 plantation plots surveyed were expressed according to

the distance zone (0–25 m, 25–50 m, 50–75 m, etc. to 150 m) containing the nearest reproductively mature conspecific(s). Zones were identified by their midpoint. If a species in a plantation had no conspecific presence within 150 m, we assumed that seed came from just beyond the furthest zone and set the distance to 175 m. While this may underestimate dispersal distance in some cases, it allowed for a conservative estimate of dispersal probabilities and how these compared among species over distances of greater than the 150 m to conspecifics that we measured. We computed a metric for colonization designed to control for variation in conspecific tree density among species, such that colonization was the ratio of seedling densities in plantations (stems/m²) to the mean density of reproducing conspecifics in the landscape (stems/m²). Thus, we expressed species' seedling density within plantations as a proportion of the species' pooled adult tree densities, adjusted for proportion that would be seed producing in the landscape (N_D , Table 1). This represents a measure of the average number of seedling colonists per reproductively mature conspecific in the landscape, at particular distances from the nearest potential conspecific source. Recall that seedling colonists have accrued over an average of $t = 7.9$ year period. We pooled species' F_D values for plantations having the nearest conspecific in the same distance zone prior to computing the dependent variable.

Due to the nature of the empirical data, there were often few or no observations of colonization for particular distance to nearest conspecifics within species. Thus, it was necessary to group similar species and determine group-wide distance-dispersal relationships. We investigated the role of dispersal mechanism (wind, bird, rodent, unspecialized) and seed size (measured by mass) (small, medium, large) on species-specific colonization frequency. These parameters were indicated in the literature to be key controls on long-distance dispersal (LDD) and colonization, and were found by Hewitt and Kellman (2002b) to be significantly related to probability of colonization (simple presence/absence data) of species in the study system. Hewitt and Kellman (2002b) determined that species fell into the following, combined, dispersal mechanism and seed mass (DM-SM) categories: wind-dispersed, small seeded (wind-small); wind-medium; bird-medium or simply 'bird'; rodent-medium; rodent-large. Seed mass categories were: small < 0.01 g; medium = 0.01–5.0 g; large > 5.0 g. We adopted these groupings, and further investigated relationships between colonization density and distance by DM-SM classes, and whether these varied among DM-SM classes, using Analysis of Covariance (PROC GLM in SAS).

The principal dispersal mechanism of each species was determined from the literature (Table 1 and references therein). Most bird-dispersed species (e.g., *Fagus grandifolia*, some *Quercus* species) are also mammal (rodent) dispersed. However, since birds are well known to achieve longer-distance transport than rodents, and medium to long-distance dispersal was the issue of concern in a model of inter-fragment colonization, these species were classified as bird-dispersed and distinguished from strictly rodent-dispersed species. Three species (*Ostrya virginiana*, *Tilia americana* and *Carpinus caroliniana*) lack obvious or effective adaptations for either wind or animal dispersal, although they may be dispersed by either mechanism, and were classified as species of 'unspecialized' dispersal.

Comparison of predicted to observed colonization with distance from seed sources

We used an exponential model to estimate seedling shadow curves for each DM-SM class. This was achieved using the LINEST function in Microsoft Excel (Microsoft ©) which calculates the least squares values of colonization frequency in relation to distance to nearest conspecific using the exponential equation, $y = c e^{(-b x)}$, where 'c' is the initial value of y (when $x = 0$) and b is the rate of change in y per unit x. A small constant was added to zero values (0.01 stems/ha; 0.001 to F_{Dx}/N_D) prior to exponential modelling.

We focused on dispersal distances of 50 m or greater in order to emphasize tail differences. We used the exponential relationship to extrapolate to distances of 300 m in order to encompass the 'tails' of the dispersal curves where dispersal differences among species would be most pronounced (Portnoy and Willson 1993). The 'tail' is defined as the set of distance categories beyond the last clear mode of the curve (Portnoy and Willson 1993), where the modes for all DM-SM classes were below ca. 50–75 m.

To incorporate dispersal effects into the colonization model we first calculated F_D values using the basic model for each species at a source ($F_{Dx=0}$) (Table S1). These were effectively values of 'c', the intercept, in our exponential decay equation $y = c e^{(-b x)}$. We then adjusted values according to our exponential decay functions. That is, we computed F_D values for each distance to seed source (x) using the exponential decay term ('b') obtained for each DM-SM class.

The predicted colonization values for each species, scaled according to distance to nearest seed source were compared to observed colonization densities using Repeated Measures Analysis of Covariance (Proc GLM in SAS) with predicted F_D and observed

F_D as repeat measures dependent variables, DM-SM as a class variable, and distance to nearest conspecific as a covariate. All statistical analyses in the study were conducted in SAS version 9.1 (©SAS Institute 2003, SAS Inst. Cary, NC, USA) except for the examinations of relationships among seed mass and observed survivorship, which employed SYSTAT (© Systat Software Inc. 2009, San Jose, CA).

Results

Seed mass and seedling survivorship

Regression analysis indicated a significant relationship between seed mass and seedling survivorship among the 10 species involved ($r^2 = 0.40$ $F = 5.436$; $p = 0.048$; log values). A non-linear model of the relationship between seed mass and observed vs. predicted survival indicated that the model parameters in equation (3) (Greene and Johnson 1998) produced a reasonable fit. However, the fit was improved by a non-linear mean function in which the predicted values assumed a constant asymptote with two parameters fitted (seed mass and survival) (Figure 1). The revised parameters produce the survivorship equation:

$$S = 0.65(1 - \exp(-0.481m^{0.3589})) \quad (5)$$

A regression between observed and predicted survivorship in the new model was very strong ($r^2 = 0.75$; $F = 24.057$; $p = 0.0011$).

We recast the model for predicted colonization as:

$$F_{Dx=0} = 3067 m^{-0.58} N_D B_m^{0.92} G^{0.65} (1 - \exp(-0.481m^{0.3589}))t \quad (6)$$

Where $t = 4$; G is granivory by large rodents for nut tree species only (see Table 1). Seed losses to smaller predators were encompassed in the 0.65 multiplier in the S portion of the equation.

The basic model

The regression of log Observed versus Expected F_{Dx} ($x = 43m$) on 27 species was highly significant ($r^2 = 0.28$; $F = 9.95$; $p = 0.0042$). The relationship improved somewhat when we employed a single, species' mean basal area measure for (i) canopy species and (ii) understory species, respectively ($r^2 = 0.30$; $F = 11.01$; $p = 0.0028$) (Figure 2).

The model explained a significant proportion of the variation in colonization density. While it under-predicted slightly (mean difference, observed-expected values: 0.020 stems/ m^2 ; $SD = 0.0997$), the difference between observed and expected was not significant (paired T-test: $T = 1.054$; $p = 0.3015$, two-tail). Residuals were somewhat large (Studentized residuals > 1.25 ; = residual/SE) for five species: *Fraxinus americana/pennsylvanicum* and *A. rubrum* (observed F_D greater than expected; implies colonization is under-predicted) and for *Nyssa sylvatica*, *P. grandidentata*, *T. americana* (observed F_D lower than expected; colonization overpredicted).

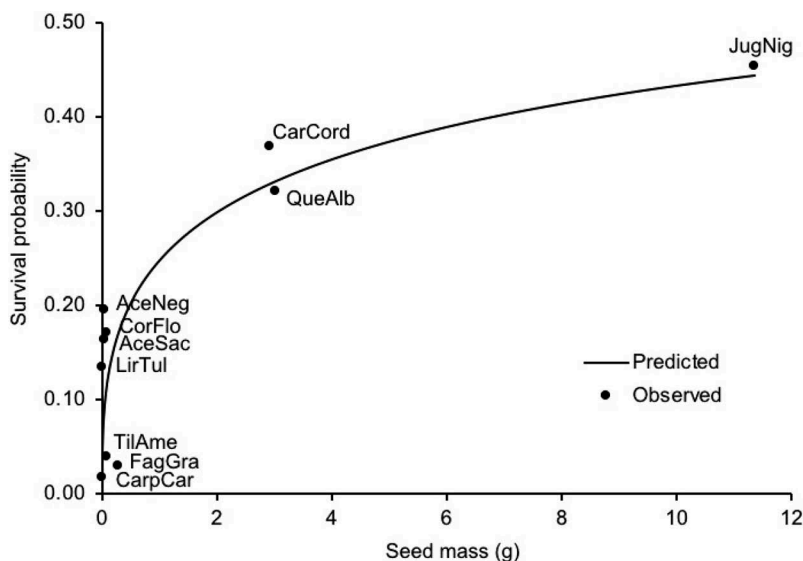


Figure 1. Relationship between species mean seed mass and observed survivorship to 2 years in a seedling establishment experiment. Species are labeled with the first three letters of the genus and species (see Table 1 for full names).

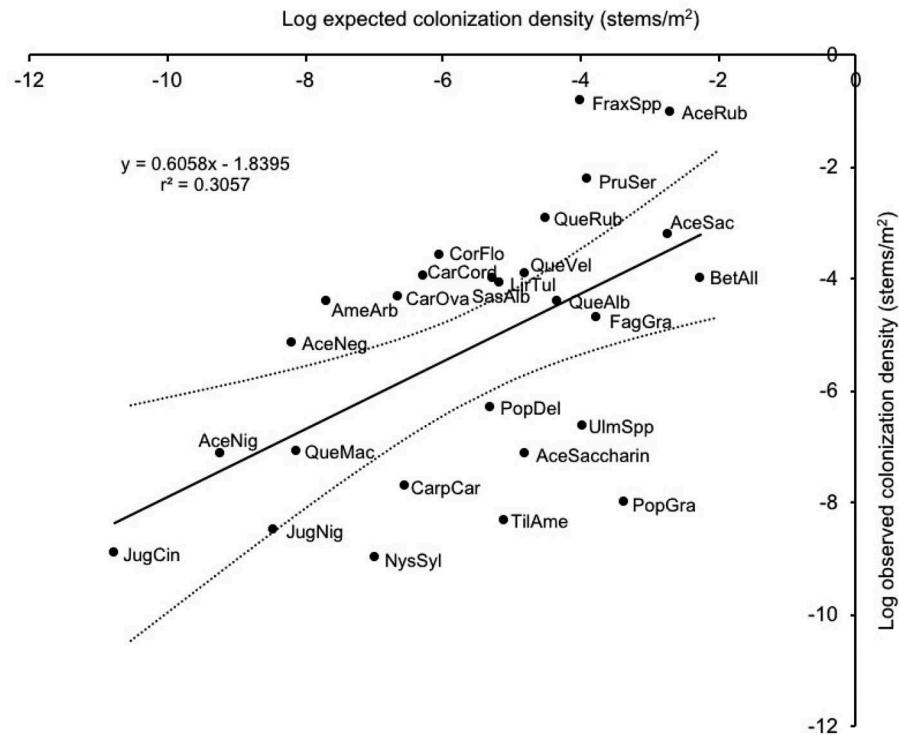


Figure 2. Relationship between Log observed vs. Log expected colonization at an average distance of 43 m from a source edge ($F_{Dx=43m}$) on 27 species whose seedling densities in plantations and adult densities in the surrounding landscape had been enumerated. The 95% confidence interval is indicated with a dashed line. The difference between observed and expected colonization density was not significant (paired T-test: $T = 1.054$; $P = 0.3015$, two-tail), indicating that the slope of the line was not significantly different from 1. Species are labeled with the first three letters of the genus and species (see Table 1 for full names).

Colonization patterns in relation to distance from seed sources

The Analysis of Covariance (PROC GLM in SAS) of Log colonization (ratio of F_D to N_D) versus DM-SM class and distance to seed source (covariate) was significant ($df = 11$; $F = 2.77$; $p = 0.0034$). As might be expected, there was a strong significant negative relationship between Log colonization probability and distance ($df = 1$; $F = 7.57$; $p = 0.007$). Colonization probability differed significantly among the DM-SM classes ($df = 5$; $F = 3.30$; $p = 0.0082$), but there was no interaction between distance and DM-SM class ($df = 5$; $F = 1.68$; $p = 0.1459$). Wind-dispersed (medium seed), bird-dispersed, and rodent-dispersed (medium seed) species had significantly greater mean colonization probabilities than species that were wind dispersed (small seed), rodent dispersed (large seed) and those lacking specialized dispersal mechanisms (Figure 3).

These differences in colonization ability were evident in the exponential decay curves fitted to the empirical colonization probabilities of different DM-SM classes (Figure 4). Values of 'b' in our empirically based exponential decay functions for each DM-SM

class were: wind-small = 0.0444; wind-medium = 0.0416; bird = 0.0324; rodent-medium = 0.0487; rodent-large = 0.0457; unspecialized = 0.0368. These parameter values average 0.0416, which is remarkably similar to that computed from empirical data by Greene and Calogeropoulos (2002). They obtained a value of 0.0415 in their dispersal term for seeds per m^2 at distance x , ' $= \exp(-gx)/n$ ', where 'g' was equivalent to our 'b', and n was a normalizer in functions that do not sum to 1.

The graphs of observed vs. predicted colonization against distance for species grouped into DM-SM classes suggested that our exponential decay terms ('b') effectively adjusted predicted colonization densities in relation to distance (x) to the nearest seed source (Figure 5). Repeated Measures Analysis of Covariance (PROC GLM) indicated a significant difference between species' predicted and observed F_D (log values) at varying distances from the nearest seed source (Table 2). There was also a significant interaction between the 'dummy' variable for vs. Predicted vs. Observed values (PredVsObs) and both (1) distance and (2) DM-SM class, indicating that there was a change in the relationship between predicted vs. observed over distance within DM-SM classes, and

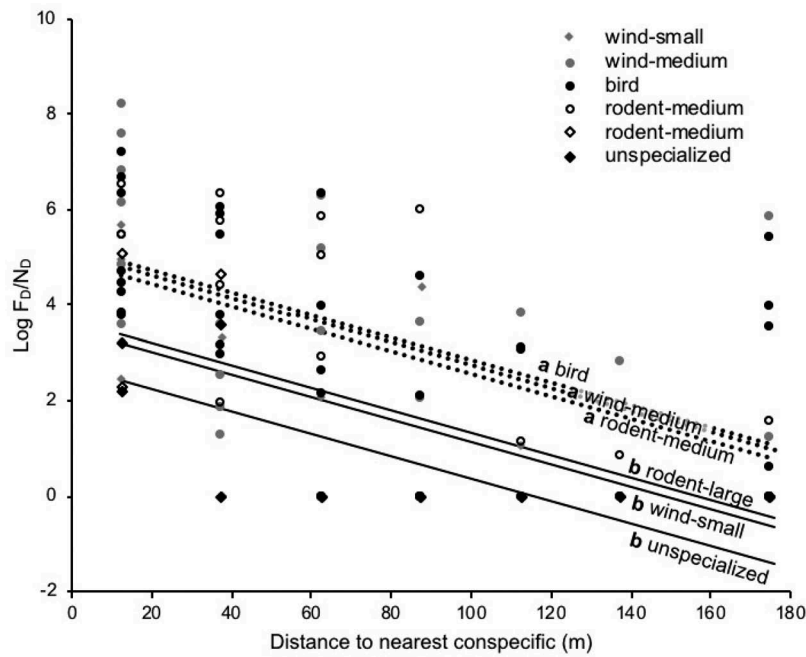


Figure 3. Relationship between $\text{Log } F_D/N_D$ and the distance to nearest conspecific seed source for species within six classes of dispersal mechanism-seed mass class (DM-SM, Appendix 1). 1 was added to F_D/N_D prior to logging to adjust for zero values ($= \text{Log } (F_D/N_D + 1)$). N_D values were first adjusted according to the estimated proportion of seed producers among dioecious and polygamodioecious species. DM-SM classes with the same letter are not significantly different ($P > t$ is < 0.002). Slope in the ANCOVA model was constant (-0.02354) for all species, given the lack of interaction between distance and DM-SM class.

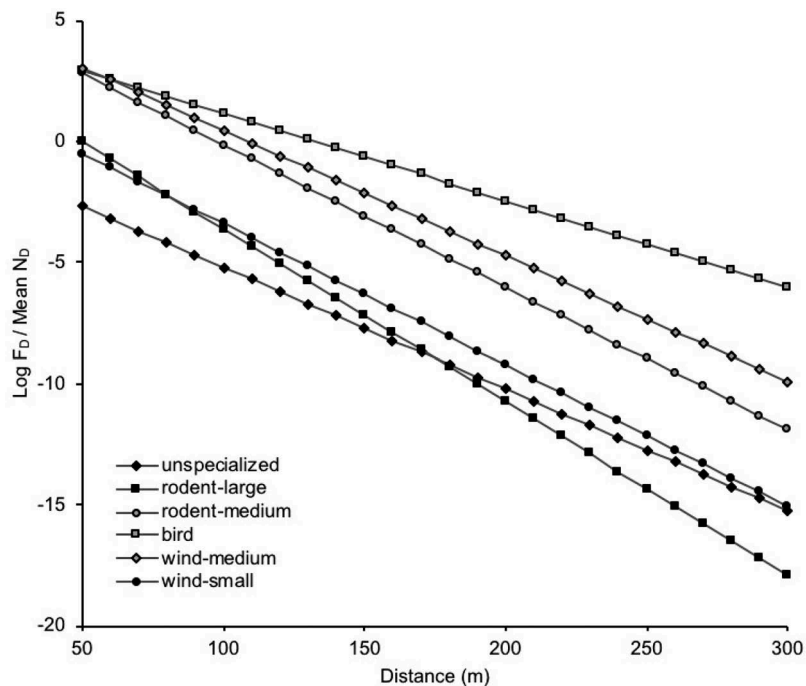


Figure 4. Log values for the modelled exponential relationships between observed colonization density and distance to nearest conspecific. Distances from 50 to 300 m are shown. Modelled relationships are for species grouped according to dispersal mechanism and seed mass (DM-SM) class (see labels). Predictions are modelled on the ratio of measured colonization density to mean adult conspecific densities in the landscape and represent number of colonists per mature conspecific at that distance from a source edge, averaged across DM-SM class.

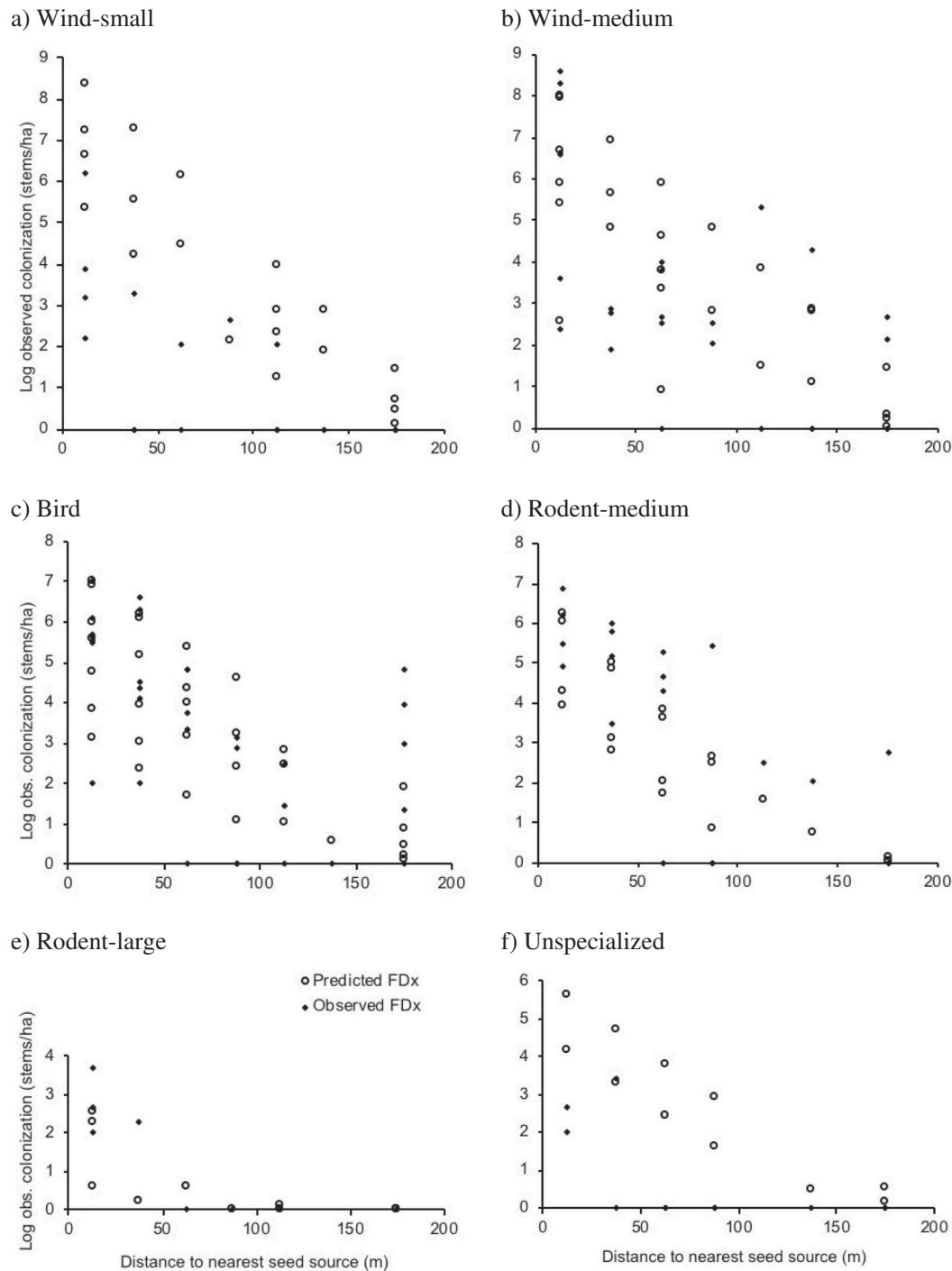


Figure 5. Predicted and observed Colonization (F_D) (log values + 1) vs. distance to nearest conspecific seed source. Predicted (expected $F_{Dx=0}$) values were first calculated for each species using the basic model, as above (multiplied by $t = 4$ to account for the average number of annual cohorts measured in the observed data and age-dependent mortality among these), then they were adjusted for distance effects by multiplying expected $F_{Dx=0}$ by empirically derived rates of exponential decline for each DM-SM class (Figure 4).

that predicted vs. observed values varied among DM-SM class. There was no significant interaction among all three variables (Table 2).

To determine significant differences among predicted and observed F_D and distance, two separate

Repeated Measures ANOVA (Proc GLM), with Tukey tests for differences among means, were performed on log values of predicted and observed F_D at two distance ranges: a) 0–50 m (near) and b) 100–175 m (far) from the nearest potential seed sources. Results indicated a

Table 2. Results of the Repeated Measure Analysis of Covariance for differences between Log Predicted F_{Dx} and Log Observed F_{Dx} , and interactions among these ('PredVsObs', the 'dummy' variable for the former comparison) to DM-SM grouping and Distance to nearest adult conspecific.

Source	df	F-value	Prob > F
Predicted vs. Observed F_{Dx}	1	14.88	0.0002
PredVsObs * DM-SM	5	5.81	<0.0001
PredVsObs * Distance	1	4.99	0.0277
PredVsObs * Distance * DM-SM	5	1.06	0.3852

significant effect of predicted vs. observed on colonization density only at the 'near' distances ($F = 7.47$; $df = 1$; error = 38; $p = 0.0095$), no significant effect 'far' ($F = 0.04$; $df = 1$; error = 35; $p = 0.8355$), with a significant interaction between predicted vs. observed and DM-SM for both models. For 'near' colonization, predicted values differed significantly from observed only within the wind-small (small-seeded wind

dispersed) category and only at the near (0–50 m) range (Figure 6). For this group, predicted values were higher than observed ($p \leq 0.05$), and the difference among the raw mean values was more than an order of magnitude (1226.81 stems/ha vs 86.15 stems/ha), although the variances were high (Table 3).

Although not significant, differences among predicted and observed values for some of the other DM-SM groupings are notable. At the near distance range (0–50 m), among unspecialized species, the model predicted much higher colonization densities (119.31 vs 12.5 stems/ha; Table 3). At the far distance range (100–175 m), no significant differences were found between mean predicted and observed values, and mean values of both predicted and observed were low. However, wind-small remained overpredicted by the model (though not significantly), whereas wind-medium, bird and rodent-medium had more than 3

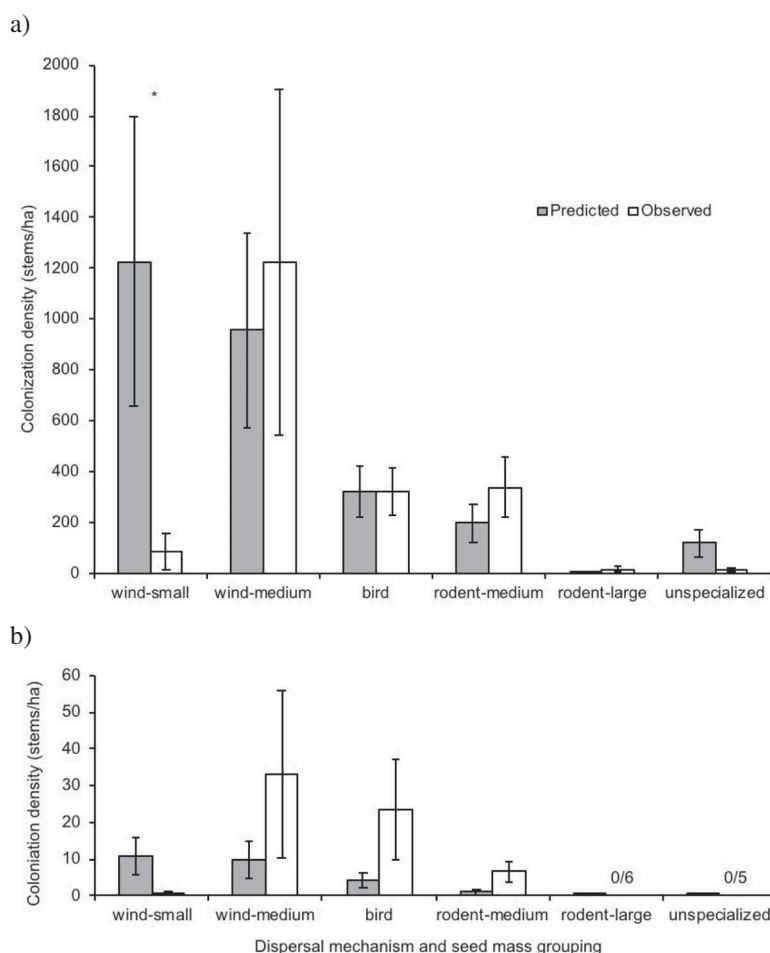


Figure 6. Mean log values of predicted F_{Dx} (grey bars) and observed F_{Dx} (white bars) at a) distances of 0–50 m and, b) > 100 m of the nearest potential seed source. Standard error bars are shown. The seed sources were adult conspecifics in the surrounding landscape, enumerated within 25 m wide distance zones. Asterisks indicate pairs for which observed values differed from expected. For rodent-large and unspecialized, the ratio indicates that 0 plantations had colonists out of a possible 5 or 6 plantations (the denominator) with seed sources > 100 m away.

Table 3. Means and Standard errors (SE) for predicted and observed colonization (stems/ha) by DM SM class ‘near’ (0–50 m; top chart) and ‘far’ (100–175 m or greater; bottom chart) from the nearest conspecific adult. Means within the same distance range having the same letter were not significantly different, according to statistical the analyses on Log values (Figure 5). Values are stems/ha. N = number of species with available data at each of the 2–3 distance zones involved (near, 0–50m involves 2 distance zones; far, > 100 m involves 3 distance zones), though there may have been several plantations for which values were pooled for each N.

DM-SM Class	Predicted		Observed		N
	Mean \pm SE		Mean \pm SE		
Near conspecific adults (within 0–50 m)					
Wind-small	1226.81 \pm 569.51	a	86.15 \pm 69.09	bc	7
Wind-medium	957.89 \pm 383.23	ab	1224.55 \pm 679.65	ab	9
Bird	321.24 \pm 102.40	ab	320.96 \pm 90.76	ab	13
Rodent-medium	196.87 \pm 74.94	ab	336.56 \pm 117.89	ab	7
Rodent-large	5.47 \pm 2.92	bc	17.17 \pm 7.74	bc	4
Unspecialized	11931 \pm 55.03	abc	12.5 \pm 6.44	bc	4
Far from conspecific adults (> 100–175 m away)					
Wind-small	10.80 \pm 4.98	a	0.67 \pm 0.67	ab	10
Wind-medium	9.80 \pm 5.06	ab	33.02 \pm 22.79	ab	9
Bird	4.15 \pm 1.87	ab	23.37 \pm 13.75	ab	9
Rodent-medium	1.07 \pm 0.76	ab	6.53 \pm 2.98	ab	5
Rodent-large	0.03 \pm 0.02	b	0	B	6
Unspecialized	0.51 \pm 0.17	ab	0	B	5

times the observed colonization relative to predicted (stems per ha; Table 3). Rodent-large and unspecialized species, though close to their predicted values, had no observed colonization at these distances (> 100 m) from conspecific seed sources. Further, observed values for these two groups were significantly lower than observed values for the other four. Since plantation plots were pooled with respect to distance to nearest source, observations for these species represents multiple ‘failed trials’. For example, among rodent-large species: *J. cinerea*, seedlings were absent from all 13 of the 150 m² plantation plots at distances of > 100 m from conspecifics; *J. nigra* was absent from all 10 plots; *Q. macrocarpa*: absent from all 15 plots; unspecialized: *C. caroliniana* and *T. americana*: absent from all of 6 plots. In fact, no rodent-large or unspecialized species were found in plots further than 50 m from a conspecific (Table S2).

Discussion

The basic model

We devised a general model for tree species colonization that functions well for fragmented eastern deciduous forests. The basic model, equation (6), which employed easily obtained parameters (e.g., seed mass, tree abundance), explained a significant proportion of aggregate colonization frequency of species in the study system at or near to a source edge (Figure 2). By comparing model predictions to empirical colonization occurring near or adjacent to seed sources, thereby controlling for dispersal effects, we were able to isolate the roles of seed production and seedling establishment. The strong correspondence between predicted

and observed values lent support to these parameters (Figure 6(a)). While the model thus performs sufficiently well to be of use in management or as part of a larger model of migration, it is worthwhile to examine the seeming deficiencies.

One problem was the under-prediction for two of the most common taxa in the system (e.g., *F. americana/pennsylvanicum*, *A. rubrum*)(Figure 2). This is likely due to the presence of source trees within 25 m of all sink areas, rather than the typical mean distance of 48 m for other taxa. But this is really a study-specific problem rather than an inherent difficulty with the model.

Our revised survivorship equation (5) was adjusted to reflect experimental relationships between seed mass and survivorship among a subset of 10 species in the study system, relationships that proved to be highly significant. We refined the term for seed loss to granivores relative to equation (3), Greene and Johnson (1998), using the results of seed sowing experiments in the study system that excluded large rodent seed predators with wire mesh. The revised equation contains an adjustment for seed losses to invertebrate, fungal, or other small seed predators (the multiplier ‘0.65’ in equation (5), to replace ‘0.43’, equation (3)) that was applied to all species. We further included an experimentally determined term, G, for large rodent (squirrel, chipmunk) predation that was applied only to nut-tree species that are targets (Table 1). This nuanced, species-specific approach to seed predation resulted in modest but significant improvements to model predictions relative to those achieved with the earlier version of S, equation (3) ($r^2 = 0.306$, Figure 2, compared to $r^2 = 0.240$; Hewitt, unpubl. data).

These advantages notwithstanding, our survivorship equation had a smaller seed mass exponent (0.36) than Greene and Johnson (1998; 0.76, equation (3)), which led to an expectation for juvenile survivorship several times higher than that of Greene and Johnson (1998) for the smallest-seeded, wind-dispersed species (*Populus*, *Betula*), and thus a correspondingly large over-prediction of the density of colonists for these species. However, the seed mass exponent improved the predictions for other, larger-seeded taxa in our study. We suggest that a more flexible mathematical expression of the relationship between seed size and juvenile survivorship is required. The discrepancy relates to the precise establishment requirements of small-seeded species in our study. They require open, unshaded sites with bare soil or a presence of rotting wood up off the forest floor – microsites lacking in our plantation environments. Conversely, one can use a separate seed mass exponent for species with very small seeds (say, seed mass < 10 mg) when clement seedbeds are rare.

A major omission in our model is shade tolerance. While much of the area we surveyed for recruits was shaded by conifers, light levels were comparable to those of woodlot edges in the study system (Hewitt and Kellman 2002a). Clearly, where there are few canopy gaps, tolerance of shade will be an important determinant. Thus, if the area examined had a large fraction of closed canopy forest with little light at the forest floor, the model would perform less effectively. Nevertheless, experimental seedling establishment of four species suggested that survivorship among plantations, woodlot edges and interior plots did not differ greatly (Hewitt and Kellman 2002a).

There are sources of variation in recruitment abundance that no simple model such as ours can hope to remedy. One example is seed production. Most North American tree species are masting, and exhibit dramatic inter-annual variation in seed supply. While the time period for recruitment used here smooths over much of this year-to-year variation, nonetheless there will still be substantial variation for mean seed production even with time periods of 4 or more years as shown by Greene and Johnson (2004).

An equally variable parameter is juvenile survivorship. Greene and Johnson (1998) provided an example with *Picea glauca*, collating 32 sowing studies on a single seedbed type (mineral soil) with ample light. Astonishingly, for this single substrate type and single species, one standard deviation to either side of the median provided more than an order of magnitude difference in survivorship. Presumably, much of this variation is due to differences in granivore abundance

from one site (or year) to another, as well as to variation in summer precipitation. The situation worsens as we consider seedbeds. Small-seeded species require a lack of litter and a shallow or absent duff layer. In undisturbed forests, this kind of seedbed is invariably limited to rotted wood (Hewitt and Kellman 2004; Greene and Johnson 1998). But certain forest types will have little or no rotted wood substrates; e.g. our conifer plantations. This is a problem for the migration of small-seeded species across fragmented landscapes where so many of the woodlots have experienced large-scale harvesting over several decades.

Another problem is relating seed production to the basal area of trees. While tree size is the single best predictor of intra-specific seed supply in any one year and place, nonetheless, it typically explains only about 30% of the variation (Calogeropoulos et al. 2003; Nanos et al. 2010). Our results suggest that tree abundance (N_D) combined with allometric relationships with seed mass were the critical variables in predicting seed production and that colonization may be modelled using simple canopy or sub-canopy stem size averages combined with tree density measures. Adjusting N_D according to the female fraction for dioecious species, sensu Clark et al. (2004) improved predictions relative to assuming all stems to be seed-producing (Figure 2; $r^2 = 0.306$ vs. $r^2 = 0.298$). And while our results account for a large proportion of variability in the dispersal and colonization process, the limitations we identify imply that refinements of any modelling approach that ascribes mean values to such inherently unpredictable local events are unlikely to raise the proportion of variance explained much higher than what we achieved.

Dispersal and colonization

We employed empirical measures of species' dispersal ability to create distance-dispersal relationships across hardwood taxa in the study system. By expressing colonization frequency (seedlings/ha) as a ratio of the species' mean tree abundance (N_D) around plantations (i.e., with the metric, N_D/F_D), interspecific differences in adult conspecific density and seed load were controlled for, and dispersal and colonization effects isolated. The dispersal patterns summarized here correspond well to the limited empirical information on LDD reported in the literature for eastern forest hardwood species (Table S3). For example, Johnson (1988), using established seedlings in relation to isolated trees or tree-rows, found similar dispersal maxima and relative species rankings among three species

we examined, *T. americana*, *A. saccharum* and *F. pennsylvanica* (Table S3).

The limitations of the field data employed to assess dispersal are summarized in Hewitt and Kellman (2002b), notably the inability to identify the precise location(s) of individuals that supplied seed. In the present study, since adult conspecifics, the seed sources, were surveyed to distances of only 150 m, plantations lacking conspecifics within the surveyed area were assumed to have been supplied by individuals just beyond, ca. 175 m away. Extensive tree inventories in the study system (Hewitt and Kellman 2002a, 2004) suggested that this assumption is reasonable since species occurred regularly throughout the landscape, source areas to the 16 plantation plots often intersected, and conspecifics were thus likely to be present not far outside the surveyed source area. The effect of our having collapsed seed sources, which potentially occurred across a wider area, into a single 175 m zone, would be compensated for in the model by the smoothing effect of our exponential dispersal function. In sum, we provide a conservative estimate of LDD probabilities among species, notwithstanding the challenges of measuring landscape scale dispersal.

Our study highlights how LDD and colonization ability vary according to species' dispersal mechanism and seed mass. There were stark contrasts in dispersal ability among (a) bird and wind-medium species, which repeatedly colonized sites beyond 150 m from source trees, and (b) rodent-large and unspecialized species, which colonized none beyond 50 m. Model predictions (adjusted for dispersal effects) were not significantly different from observed values in each DM-SM class, except for wind-small species in the 'near' category, a finding we attribute to seedling establishment constraints within this group. These results suggest that the basic model equation combined with distance decay function produced reasonably accurate dispersal predictions across the distances measured. However, the significant interaction between distance and dispersal mechanism in the global Repeated Measures ANOVA, and comparison of the mean values, suggest that the model is not consistent across dispersal distances. Specifically, there were some taxa that were somewhat over-predicted at the source (i.e., by the basic model) but under-predicted at longer distances. These discrepancies between model predictions and observations for longer distances, though not statistically significant, have theoretical and practical implications, prompting an examination of the dispersal function.

The importance of birds as an LDD mechanism generally, and specifically in fragmented landscapes,

has empirical support in the limited literature available (Table S3). This is attributed to birds' ability to travel long distances and direct seed to suitable habitat (Johnson and Adkisson 1985; Darley-Hill and Johnson 1981), which is in contrast to wind, which scatters seed somewhat randomly. This creates a problem in choosing a dispersal function, as wind dispersal from an area source (i.e., a collection of conspecific point sources) is generally best described by a negative exponential at the scale of 0 to 200 m while for animal dispersal a function such as the 2dt of Clark et al. (1999) may be more appropriate. We suggest therefore that subsequent modelling of inter-fragment colonization should investigate separate vectors for animal vs. wind-dispersed species (Muller-Landau et al. 2008).

The dispersal function for large-seeded rodent-dispersed species and species of unspecialized dispersal mechanism, species with lowest chances of colonizing distant sites, requires refinement. While mean predicted values for longer distances were within a few stems/ha of observed values, the comparison is misleading given that observed values were all zero. As noted above, colonization was not found beyond 50 m of a conspecific (Table S2). While distance effects in the model were calculated as a gradual exponential decrease in colonization, this indicates a 'threshold' distance beyond which dispersal and colonization for these species is improbable. The dispersal parameter must reflect relevant thresholds. Based on our empirical data and the available literature (Table S3), and pending further research, we suggest that colonization for rodent-large and unspecialized species be modelled as '0' beyond threshold distances of 75 m, an ample dispersal maximum given the evidence. That is, for these particular species, if $F_{Dx} \leq 75$ m, the dispersal parameters are as stated above; if $F_{Dx} > 75$ m, the dispersal term = 0. More fundamentally, dispersal of unspecialized or large seed rodent-dispersed species among fragment in mid-latitude forests may be unimportant.

In conclusion, the present model provides a relatively simple method for examining colonization probabilities for a large proportion of eastern deciduous forest species. It is one of the few existing models to have been calibrated using independent experimental data on seedling establishment, as well as empirical dispersal and colonization data across a wide range of dispersal distances (0–175 m) and dispersal mechanisms, wind and non-wind (and see e.g., Clark et al. 1999, Table S3, covering distances of up to 60 m). The model reflects dispersal across patchy or fragmented landscapes that are increasingly common (and see Nathan et al. 2005; Schurr et al. 2008, for dispersal by

wind in heterogeneous landscapes). In light of the growing pressures on tree species not only to maintain metapopulations but to migrate across landscapes, the model is timely.

Acknowledgments

We thank the Long Point Regional Conservation Authority for permission to work in the Backus Woods Forest Preserve and the many private landowners who granted us access to their land, some of whom provided valuable insights into the landscape history. We thank the undergraduate and graduate students who assisted with field work that contributed to this model, including M. Richart, S. Richart, P. Moti and J. Biddulph. E. Johnson provided helpful feedback on the modelling approach. We are grateful to two anonymous reviewers for helpful comments on the manuscript.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The research was supported, in part, by an internal Summer Research Grant and a Research Assistantship from De Paul University, College of Liberal Arts and Sciences (to NH).

References

- Abbott HG, Quink TF. 1970. Ecology of eastern white pine seed caches made by small forest mammals. *Ecology*. 51:271–278.
- Anand M, Langille A. 2010. A model-based method for estimating effective dispersal distance in tropical plant populations. *Theor Pop Biol*. 77:219–222.
- Baker FS. 1949. A revised tolerance table. *J for*. 47:179–181.
- Bruun HH, Ten Brink DJ. 2008. Recruitment advantage of large seeds is greater in shaded habitats. *Écoscience*. 15:498–507.
- Burns RM, Honkala BH. 1990. *Silvics of North America*, vol. 2: hardwoods. Agricultural handbook # 654. Washington (DC): US Department of Agriculture.
- Calogeropoulos C, Greene DF, Messier C, Brais S. 2003. Refining tree recruitment models. *Can J For Res*. 33:41–46.
- Chen I-C, Hill JK, Ohlemueller R, Roy DB, Thomas CD. 2011. Rapid range shifts of species associated with high levels of climate warming. *Science*. 333:1024–1026.
- Clark JS, LaDeau S, Ibanez I. 2004. Fecundity of trees and the colonization competition hypothesis. *Ecol Monographs*. 74:415–442.
- Clark JS, Silman M, Kern R, Macklin E, HilleRisLambers J. 1999. Seed dispersal near and far: patterns across temperate and tropical forests. *Ecology*. 80:1475–1494.
- Cousens RD, Hill J, French K, Bishop ID. 2010. Towards better prediction of seed dispersal for animals. *Funct Ecol*. 24:1163–1170.
- Darley-Hill S, Johnson WT. 1981. Acorn dispersal by Blue Jay (*Cyanocitta cristata*). *Oecologia* 50:231–233.
- Davis MB. 1981. Quaternary history and the stability of forest communities. In: West DC, Shugart HH, Botkin DB, editors. *Forest succession: concepts and application*. New York: Springer; p. 132–153.
- Dullinger S, Dirnböck T, Grabherr G. 2004. Modelling climate change-driven treeline shifts: relative effects of temperature increase, dispersal and invisibility. *J Ecol*. 92:241–252.
- Dzwonko Z, Loster S. 1992. Species richness and seed dispersal to secondary woods in southern Poland. *J Biogeogr*. 19:195–204.
- Greene DF, Calogeropoulos C. 2002. Dispersal of seeds by animals and wind. In: Bullock J, Kenward R, Hails R, editors. *Dispersal ecology*. Oxford (UK): Blackwell; p. 3–23.
- Greene DF, Johnson EA. 1992. Can the variation in samara mass and terminal velocity on an individual plant affect the distribution of dispersal distances? *Am Nat*. 139:825–838.
- Greene DF, Johnson EA. 1993. Seed mass and dispersal capacity in wind-dispersed diaspores. *Oikos* 67:69–74.
- Greene DF, Johnson EA. 1994. Estimating the mean annual seed production of trees. *Ecology*. 75:642–647.
- Greene DF, Johnson EA. 1996. Wind dispersal of seeds from a forest into a clearing. *Ecology*. 77:595–609.
- Greene DF, Johnson EA. 1998. Seed mass and early survivorship of tree species in upland clearings and shelterwoods. *Can J For Res*. 28:1307–1316.
- Greene DF, Johnson EA. 2000. Tree recruitment from burn edges. *Can J For Res*. 30:1264–1274.
- Greene DF, Johnson EA. 2004. Modelling the temporal variation in North American tree seed production. *Can J For Res*. 34: 65–75.
- Hett JM. 1971. A dynamic analysis of age in sugar maple seedlings. *Ecology*. 52:1071–1074.
- Hewitt N. 1998. Seed size and shade tolerance: a comparative analysis of North American temperate trees. *Oecologia* 114:432–440.
- Hewitt N. 1999. Plant dispersal and colonization in fragmented forest systems [dissertation]. Toronto: York University.
- Hewitt N, Kellman M. 2002a. Tree seed dispersal among forest fragments: I. Conifer plantations as seed traps. *J Biogeogr*. 29:337–350.
- Hewitt N, Kellman M. 2002b. Tree seed dispersal among forest fragments: II. Dispersal abilities and biogeographic controls. *J Biogeogr*. 29:351–363.
- Hewitt N, Kellman M. 2004. Factors influencing tree colonization in fragmented forests: an experimental study of introduced seeds and seedlings. *For Ecol Manag*. 191:39–59.
- Johnson WC. 1988. Estimating dispersibility of *Acer*, *Fraxinus*, and *Tilia* in fragmented landscapes from patterns of seedling establishment. *Landsc Ecol*. 1:175–187.
- Johnson WC, Adkisson CS. 1985. Dispersal of beech nuts by Blue jays in fragmented landscapes. *Am Midl Nat*. 113:319–324.
- Jones AF, Muller-Landau HC. 2008. Measuring long-distance seed dispersal in complex natural environments: an evaluation and integration of classical and genetic methods. *J Ecol* 96:642–652.
- Levey DJ, Tewksbury JJ, Bolker BM. 2008. Modelling long-distance dispersal in heterogeneous landscapes. *J Ecol*. 96:599–608.

- Loarie SR, Duffy PB, Hamilton H, Asner GP, Field CB, Ackerly DD. 2009. The velocity of climate change. *Nature* 462:1052–1055.
- Malcolm JR, Markham A, Neilson RP, Garaci M. 2002. Estimated migration rates under scenarios of global climate change. *J Biogeogr.* 29:835–849.
- Moles AT, Westoby M. 2006. Seed size and plant strategy across the whole life cycle. *Oikos*. 113:91–105.
- Morin X, Viner D, Chuine I. 2008. Tree species range shifts at a continental scale: new predictive insights from a process-based model. *J Ecol.* 96:784–794.
- Muller-Landau HC, Wright SJ, Caldrón O, Condit R, Hubbell SP. 2008. Interspecific variation in primary seed dispersal in a tropical forest. *J Ecol.* 96:653–667.
- Nanos N, Larson K, Milleróna M, Sjöstedt-de Luna S. 2010. Inverse modelling for effective dispersal: do we need tree size to estimate fecundity? *Ecol Model.* 221:2415–2424.
- Nathan R, Perry G, Cronin JT, Strand AE, Cain ML. 2003. Methods for estimating long-distance dispersal. *Oikos*. 103:261–273.
- Nathan R, Sapir N, Trakhtenbrot A, Katul GG, Bohrer G, Otte M, Avissar R, Soons MB, Horn HS, Wikelskiand M, et al. 2005. Long-distance biological transport processes through the air: can nature's complexity be unfolded *in silico*? *Divers Dist.* 11:131–137.
- Nathan R, Schurr FM, Spiegel O, Steinitz O, Trakhtenbrot A, Tsoar A. 2008. Mechanisms of long-distance dispersal. *Trends Ecol Evol.* 23:638–647.
- Neilson OP, Pitelka LF, Solomon AM, Nathan R, Guy F, Midgley GF, Fragoso JMV, Lischke H, Thompson K. 2005. Forecasting regional to global plant migration response to global change. *Bioscience*. 55:749–759.
- Pearson RG, Dawson TE. 2005. Long-distance plant dispersal and habitat fragmentation: identifying conservation targets for spatial landscape planning under climate change. *Biol Conserv.* 123:389–401.
- Petit RJ, Hu FS, Dick CW. 2008. Forests of the past: A window to future changes. *Science*. 320:1450–1452.
- Portnoy S, Willson MF. 1993. Seed dispersal curves: behaviour in the tail of the distribution. *Evol Ecol.* 7:25–44.
- Ranney JW, Johnson WC. 1977. Propagule dispersal among forest islands in south-eastern South Dakota. *Prairie Nat.* 9:17–24.
- Richart M, Hewitt N. 2008. Forest remnants in the long point region, Southern Ontario: tree species diversity and size structure. *Landsc Urban Plan.* 86:25–37.
- Sandel B, Arge L, Dalsgaard B, Davies RG, Gason KJ, Sutherland WJ, Svenning JC. 2011. The influence of late quaternary climate-change velocity on species endemism. *Science*. 334:660–664.
- SAS Institute. 2003. Version 7, fifth edition. Cary (N Carolina (US)) SAS Institute.
- Schopmeyer CS, editor. 1974. Seeds of woody plants in the United States. Washington (DC): US Department of Agriculture, Government Printing Office.
- Schurr FM, Steinitz O, Nathan R. 2008. Plant fecundity and seed dispersal in spatially heterogeneous landscape environments: models, mechanisms and estimation. *J Ecol.* 96:628–641.
- Sork VL. 1984. Examination of seed dispersal and survival in red oak, *quercus rubra* (Fagaceae), using metal-tagged acorns. *Ecology*. 65:1020–1022.
- SYSTAT Software Inc. 2009. Version 13. San Jose (CA).
- Thomson FJ, Moles AT, Auld TD, Kingsford RT. 2011. Seed dispersal distance is more strongly correlated with plant height than with seed mass. *J Ecol.* 99:1299–1307.
- Trakhtenbrot A, Nathan R, Perry G, Richardson DM. 2005. The importance of long-distance dispersal in biodiversity conservation. *Divers Distrib.* 11:173–181.
- Van der Pijl L. 1982. Principles of dispersal in higher plants. New York: Springer.

Appendix

List of terms, abbreviations and variables used in the model, in the measurement of colonization frequency in

the study system, and in species' classification into distinct groups with respect to colonization ability. Further information about each of these is given throughout the text.

Term or Abbreviation	Definition
Terms, variables and parameters relating to the model and its analysis	
Basic model	The model that predicts colonization density at or near a forest edge as a function of seed production and survivorship (equation (1), recast as equation (4)). Seed production is modelled as an inverse function of seed mass (m), and a direct function of tree density and tree basal area; Survivorship (S) is modelled as a direct function of seed mass based on experimentally determined relationships (see "S" below). The basic model lacks a specific function for dispersal, which is added via the dispersal term.
Dispersal term	An empirically derived exponential function to describe the relationship between colonization density and distance to nearest conspecific seed source. This replaces $f(x)$ in equation (1) and is added to the equation (4) for each species, depending on their dispersal mechanism and seed size classes (see DM-SM groups, below). Thus, there are 6 different versions of this function, one for species in each of the DM-SM groups (e.g., wind-small, wind-medium, etc.).
F_{Dx}	Colonization density (stems/ m^2) at distance x , perpendicular to an area source.
N_D	Conspecific tree density (stems/ m^2) in the landscape or area source that supplies seed.
B, B_m	Basal area (B) and mean basal area (B_m) of a species. B_m was computed from diameter breast height (dbh) measures from a point quarter sample of trees in the study area (see Table 1 for number of stems sampled per species).
m	Species' seed mass (g).
S	The survivorship equation in the basic model. This predicts proportion of individuals surviving from the post-dispersal seed stage to the end of the second growing season and is based on a relationship between species' seed mass and experimental survivorship for 10 species in the study area.
G	The granivory term in the basic model to account for seed losses to rodent predators (squirrels and chipmunks) that target large seed, nut tree species. Additionally, the survivorship equation has a multiplier (0.65) that accounts for seed losses to smaller seed predators including invertebrates and fungi.
t	Number of annual cohorts represented by colonizing seedlings.
PredVsObs	The "dummy" variable created to model the interaction between Log Predicted F_D and Log Observed F_D in ANCOVA analysis, Table 2.
Terms relating to measuring dispersal and colonization and species' classification	
LDD	Long distance dispersal of seed. We classify LDD as dispersal > 100–150 m from a source, because colonization frequencies at these distances occurred well within the "tails" (Portnoy and Willson 1993) of the measured distance-density distributions of seedling colonists (Figures 3 and 5).
Area source	An area of forest supplying seed to adjacent areas. Represents a collection of reproducing conspecifics.
Plantation	Conifer plantations in the study area that served as "seed traps" for hardwood species via seedling recruits enumerated within large (15 m x 100 m) plots ($n = 16$) positioned along the length of plantation edges (Hewitt and Kellman 2002b).
Distance zones	Concentric zones of specified distance from plantations plot edges (0–25 m; 25–50 m; etc.) in which mature conspecific hardwood trees, the potential seed sources to plantations, were enumerated.
Seedling/recruit	The established hardwood seedlings in plantations used to represent dispersal and colonization events. Seedlings ranged in age from 1 to a median age of 8 years across the species, and thus represented up to 8 average annual cohorts of colonists.
Conspecific adult	The potential seed sources for hardwood species' seedlings recruited in plantations. These individuals were surveyed within the concentric distance zones around plantations.
DM-SM group	Dispersal mechanism-seed size group into which species were classified on the basis of dispersal mechanisms and seed size class (Table 1). Specific categories are indicated below.
Wind-small	Wind-dispersed species with small (< 0.01 g) seed size
Wind-medium	Wind-dispersed species with medium (0.01–5.0 g) seed size
Bird	Bird-dispersed species, which all had medium seed size
Rodent-medium	Rodent-dispersed species with medium (0.01–5.0 g) seed size
Rodent-large	Rodent-dispersed species with medium (> 5.0 g) seed size
Unspecialized	Species lacking specialized mechanisms for effective long-distance dispersal. Seed mass for these species ranged between 0.015–0.082 g (Table 1).