

## MODELING SEED DISPERSAL DISTANCES: IMPLICATIONS FOR TRANSGENIC *PINUS TAEDA*

CLAIRE G. WILLIAMS,<sup>1,4</sup> SHANNON L. LADEAU,<sup>1,5</sup> RAM OREN,<sup>2</sup> AND GABRIEL G. KATUL<sup>2,3</sup>

<sup>1</sup>Duke University, Department of Biology, Durham, North Carolina 27708 USA

<sup>2</sup>Duke University, Nicholas School of the Environment and Earth Sciences, Durham, North Carolina 27708 USA

<sup>3</sup>Duke University, Department of Civil and Environmental Engineering, Durham, North Carolina 27708 USA

**Abstract.** Predicting forest-tree seed dispersal across a landscape is useful for estimating gene flow from genetically engineered (GE) or transgenic trees. The question of biocontainment has yet to be resolved, although field-trial permits for transgenic forest trees are on the rise. Most current field trials in the United States occur in the Southeast where *Pinus taeda* L., an indigenous species, is the major timber commodity. Seed dispersal distances were simulated using a model where the major determinants were: (1) forest canopy height at seed release, (2) terminal velocity of the seeds, (3) absolute seed release, and (4) turbulent-flow statistics, all of which were measured or determined within a *P. taeda* plantation established from seeds collected from wild forest-tree stands at the Duke Forest near Durham, North Carolina, USA. In plantations aged 16 and 25 years our model results showed that most of the seeds fell within local-neighborhood dispersal distances, with estimates ranging from 0.05 to 0.14 km from the source. A fraction of seeds was uplifted above the forest canopy and moved via the long-distance dispersal (LDD) process as far as 11.9–33.7 km. Out of  $10^5$  seeds produced per hectare per year, roughly 440 seeds were predicted to be uplifted by vertical eddies above the forest canopy and transported via LDD. Of these, 70 seeds/ha traveled distances in excess of 1 km from the source, a distance too great to serve as a biocontainment zone. The probability of LDD occurrence of transgenic conifer seeds at distances exceeding 1 km approached 100%.

**Key words:** Duke Forest (North Carolina, USA); gene flow; genetically modified conifers; gymnosperms; long-distance dispersal; pine plantation; *Pinus taeda*; seed dispersal; transgenic trees.

### INTRODUCTION

Risk analysis of gene flow from genetically modified or transgenic forest trees requires reliable estimates of seed dispersal distances. One class of forest trees, pines, disperses seeds on the order of kilometers in old-growth forests (Nathan et al. 2002) but dispersal distances have not yet been predicted for young pine plantations, which reproduce at least a decade before timber harvest.

Our case study is based on *Pinus taeda*, a major commodity forest species in the United States indigenous to the southeastern United States (Baker and Langdon 1990). Dispersal distances are relevant because molecular domestication is rapidly changing the genetic composition of private forests often surrounded

by wild *P. taeda* forests. *Pinus taeda* can now be clonally produced from genetically identical somatic embryos as part of highly intensive plantation silvicultural systems and, more recently, genetically transformed *P. taeda* is being field tested although not yet planted on a commercial scale (Mann and Plummer 2002) and transgenic field tests are destroyed at reproductive onset.

### Predicting seed dispersal

Most seeds fall near their source, a dispersal process defined as local-neighborhood diffusion (LND) (Hengeveld 1989). A small fraction is lifted above the canopy by vertical updrafts then moved far from source via long-distance dispersal (LDD) (Fig. 1; Nathan et al. 2002). While no formal definition exists, LDD is often quantified using one of two quantities: (1) the distance traveled by the 99.9th percentile of seeds ( $D_{max}$ ) and (2) the distance traveled by the 99th percentile of the seeds ( $D_{99}$ ). LND events constitute an advancing wave of seedling regeneration around the

Manuscript received 17 December 2004; revised 26 April 2005; accepted 11 May 2005. Corresponding Editor: C. R. Linder.

<sup>4</sup> E-mail: claire.williams@duke.edu

<sup>5</sup> Present address: Smithsonian Environmental Research Center, P.O. Box 28, Edgewater, Maryland 21037 USA.

TABLE 1. A summary of seed dispersal models, listed by increasing complexity, within type.

Model type	Comments/description
<b>Mechanistic</b>	
Ballistic models	Only consider transport by mean wind speed. Reproduce the dispersal kernel mode reasonably well, but not long-distance dispersal (LDD).
Analytical solutions for simplified advection-diffusion equations	The vertical inhomogeneity introduced by the canopy on the flow statistics is not considered. The resulting kernels are heavy tailed but unbounded in variance.
Coupled Eulerian–Lagrangian closure models, CELC	Vertical inhomogeneity introduced by the canopy elements, atmospheric stability, and organized turbulent excursions are all resolved. The effect of vertical velocity skewness is not addressed. Computationally demanding.
Large-eddy simulations, LES	All the dynamics of large-scale eddies transporting seeds are explicitly considered at short time scales. Computationally prohibitive for simulating more than a few hours of real time.
<b>Empirical</b>	
Power-law and exponential models	Monotonically decaying kernels with limited number of fitting parameters. Kernels do not distinguish between dispersal patterns near and far from the source. Kernel parameters are typically fitted to dispersal data collected near the seed source, making inferences about LDD unrealistic.
Mixed models	Able to account for the fact that the dispersal-kernel behavior near the seed source is different from the dispersal-kernel behavior far from the source. The number of parameters needed ranges from 4 to 6, but a broad range of dispersal distances must be collected for reliably determining them.
The 2Dt (bivariate Student's <i>t</i> distribution) model	This is the most popular empirical and heavy-tailed kernel, parsimonious, but not skewed. Some justification for its usage, as it may be connected with Tsallis's thermostat entropy.

periphery of the source but LDD events are the source of remote satellite colonies (Moody and Mack 1988, Nichols and Hewitt 1994, LeCorre et al. 1997, Clark et al. 1998).

Several mechanistic and empirical models have been proposed for modeling dispersal and these include attempts to separate LND and LDD dispersal processes (summarized in Table 1). Given the absence of dispersal data for genetically modified seeds, empirical models may not be viable. However, mechanistic models can be adapted to predict seed dispersal if seed traits, release height, and the flow statistics are known. We chose the coupled Eulerian-Lagrangian closure (CELC) model because it accounts for the turbulent velocity excursions inside the forest canopy, particularly in the vertical direction, at fine scales (~10 seconds) and because it has been verified across a wide range of ecosystems (Nathan et al. 2002, Soons et al. 2004). In this approach, first and second moments of the velocity statistics inside the forest canopy are first computed using higher-order Eulerian moment-closure principles (Katul and Albertson 1998). Next, a three-dimensional Lagrangian trajectory model generates random velocity fluctuations at fine time scales with statistical moments identical to those computed by the Eulerian approach. The resulting synthetic flow field is used to disperse seeds in three dimensions within and above the forest canopy at one or more transport levels above the earth's surface. Generating the flow field us-

ing large-eddy simulations (LES) (Table 1) might be more desirable than a CELC model. The drawback is that LES adds a steep computational cost, prohibitive for this application, requiring more than 10 seconds of computer time to simulate 1 second of turbulence. The CELC model is thus deemed optimum in terms of predictive skills and computational demands (Table 1).

In this study, seed dispersal distances were quantified using the CELC model. Variables for predicting seed dispersal and spread are shown in a schematic diagram (Fig. 1). Seed dispersal distances were predicted for plantation conditions using plantation height at seed release, terminal velocity for *P. taeda* seeds, wind sta-

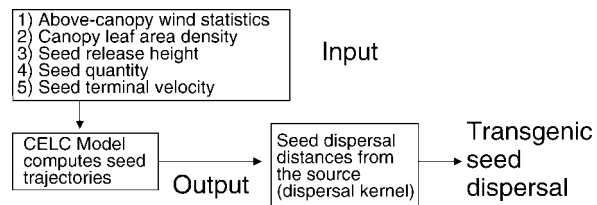


FIG. 1. Schematic diagram of the CELC (coupled Eulerian–Lagrangian closure) model input/output used for predicting seed dispersal kernels. Actual measurements for leaf-area density (and index), seed-release height and terminal velocity, and wind statistics above the canopy were for a *Pinus taeda* plantation situated at the Blackwood division of the Duke Forest near Durham, North Carolina, USA.

tistics above the plantation's canopy, leaf-area density, and drag properties of the canopy.

#### METHODS AND MATERIALS

##### *Measurements from a wild Pinus taeda plantation*

Seed dispersal distances were computed from reproductive data measured in a *Pinus taeda* plantation established from wild forest-tree seed. The plantation is located in the free-air CO<sub>2</sub> enrichment (FACE) pine plantation in the Duke Forest, near Durham, North Carolina, USA (35°97' N, 79°09' W). Trees in this well-studied plantation have shallow root systems that limit water and nutrient availability (Oren et al. 1998, Oren et al. 2001)

*Site attributes.*—Planting density is 1730 stems/ha and the genetic composition is local wild forest-tree seed source (LaDeau and Clark 2001). The mean canopy height ( $h$ ) at 16 years or reproductive onset is 14.6 m and projected to increase to 21.3 m by age 25 (harvest age). Leaf-area index (LAI), defined as foliage area per unit ground area, is 4.6 m<sup>2</sup>/m<sup>2</sup> for the 16-yr-old stand during the dispersal season (Schäfer et al. 2003). At 21 years the measured LAI is 5.5 m<sup>2</sup>/m<sup>2</sup> and represents the expected maximum LAI for this stand. Hence, for the purposes of model calculations and noting the limited nutrient availability, we assumed LAI equals 5.5 m<sup>2</sup>/m<sup>2</sup> for the 25-yr-old stand.

Based on fecundity data from 1998–2000, we estimate an annual seed production of  $\sim 10^5$  seeds/ha for age 16 and used this value for all simulations. This was an approximation based on the following observed release of quantities of sound seed at the Duke Forest plantation: 6321 seeds in 1998, 85 289 seeds in 1999, and 145 493 seeds in 2000.

*Seed attributes.*—Seed terminal velocity for *P. taeda* seeds is  $0.7 \pm 0.23$  m/s (mean  $\pm$  SD) and was determined from a nearby stand (Nathan et al. 2002). The mean seed-release height in the model calculations was set to  $0.8 h$  (at 16 years) and  $0.9 h$  (at 25 years) to reflect age-related changes in crown seed distribution.

*Wind statistics.*—As part of a long-term carbon-flux monitoring initiative, the three components of the wind velocity are collected (since 1998) above the canopy using a CSAT3 triaxial sonic anemometer (Campbell Scientific, Logan, Utah, USA) at 10 Hz and averaged every 30 min. To extrapolate these wind statistics to the 25-yr-old case, we fitted a Weibull function to the 30-min measured friction velocity ( $u_*$ ) histogram (pdf [probability density function]) for 1998–2003 data, given by

$$\text{pdf}(u_*) = \frac{cu_*^{c-1}}{b^c} \exp\left[-\left(\frac{u_*}{b}\right)^c\right]$$

and assumed that the parameters of the resulting Weibull distribution ( $b = 2.11$  and  $c = 0.456$ ) remain stationary.

##### *A predictive CELC model for seed dispersal*

To model seed trajectory, we used a Lagrangian approach originally applied to seed dispersal by Nathan et al. (2002) and Soons et al. (2004) in which the seed coordinate at position  $x_i$  is described as

$$x_i(t + \Delta t) = x_i(t) + \int_t^{t+\Delta t} (u_i - \delta_{i,3} V_t) dt$$

$$i = 1, 2, 3. \quad (1)$$

Here  $u_i$  is the seed velocity along direction  $x_i$  with  $x_1$ ,  $x_2$ , and  $x_3$  being the longitudinal, lateral, and vertical velocities, respectively,  $V_t$  is the seed terminal velocity,  $\Delta t$  is the discrete simulation time interval, and  $t$  is time. The model, with instantaneous velocity in the horizontal and vertical directions, has random variables possessing many statistical attributes of canopy and surface-layer turbulence. These attributes are aimed at reproducing the well-known spatial and temporal coherence of canopy eddies. We use a Lagrangian stochastic-dispersion approach to generate the turbulent velocity excursions based on the assumption that the evolution of the position and velocity of a fluid element are described by a Markov process. Basic concepts, details, and many references to the use of this approach in atmospheric sciences have been reviewed in Rodean (1996). Using the Markov process assumption, the evolution of  $u_i$ , needed to compute fluid element trajectories  $x_i$ , can be expressed by

$$du'_i = \alpha_i(x_i, u_i, t)dt + \beta_{ij}(x_i, u_i, t)d\Omega_j \quad (2)$$

where  $u'_i$  is, as before, the instantaneous turbulent velocity along direction  $x_i$ ,  $\alpha_i$  are the drift coefficients,  $\beta_{ij}$  are the random acceleration coefficients, and  $d\Omega$  is a Gaussian random variable with zero mean and variance  $dt$ .

For stationary and inhomogeneous near-Gaussian turbulent flows,  $\alpha_i$  and  $\beta_{ij}$  are estimated from Thomson's (1987) "simplest solution" using the procedure outlined in Rodean (1996) to preserve the so-called "well-mixed" condition. The well-mixed condition states that if the concentration of a species is initially uniform in a turbulent flow, it will remain uniform if there are no sources or sinks, consistent with the second law of thermodynamics. Thomson's (1987) "simplest solution" in three dimensions results in the following:

$$du'_1 = \left[ -\frac{C_{\delta} \langle \varepsilon \rangle}{2} (\lambda_{11} u'_1 + \lambda_{13} u'_3) + \frac{\partial \langle \bar{u}_1 \rangle}{\partial x_3} u'_3 + \frac{1}{2} \frac{\partial \langle \bar{u}'_1 u'_3 \rangle}{\partial x_3} \right] dt$$

$$+ \left[ \frac{\partial \langle \bar{u}'_1 u'_1 \rangle}{\partial x_3} (\lambda_{11} u'_1 + \lambda_{13} u'_3) + \frac{\partial \langle \bar{u}'_1 u'_3 \rangle}{\partial x_3} u'_1 \right. \\ \left. + \lambda_{33} u'_3 \right] \frac{u'_3}{2} dt + \sqrt{C_{\delta} \langle \varepsilon \rangle} d\Omega$$

$$du'_2 = \left[ -\left( \frac{C_{\delta} \langle \varepsilon \rangle}{2} + \frac{1}{2} \frac{\partial \langle \bar{u}'_2 u'_2 \rangle}{\partial x_3} \right) (\lambda_{22} u'_2) \right] dt + \sqrt{C_{\delta} \langle \varepsilon \rangle} d\Omega$$

$$\begin{aligned}
du'_3 = & \left[ -\frac{C_o \langle \varepsilon \rangle}{2} (\lambda_{13} u'_1 + \lambda_{33} u'_3) + \frac{1}{2} \frac{\partial \langle \overline{u'_3 u'_3} \rangle}{\partial x_3} \right] dt \\
& + \left[ \frac{\partial \langle \overline{u'_1 u'_3} \rangle}{\partial x_3} (\lambda_{11} u'_1 + \lambda_{13} u'_3) \right. \\
& \left. + \frac{\partial \langle \overline{u'_3 u'_3} \rangle}{\partial x_3} (\lambda_{13} u'_1 + \lambda_{33} u'_3) \right] \frac{u'_3}{2} dt + \sqrt{C_o \langle \varepsilon \rangle} d\Omega
\end{aligned} \tag{3}$$

where  $C_o$  ( $\approx 5.0$ ) is a similarity constant (related to the Kolmogorov constant), the overbar indicates time averaging,  $\langle \cdot \rangle$  indicates spatial averaging, and  $\lambda_{11}$ ,  $\lambda_{13}$ ,  $\lambda_{22}$ , and  $\lambda_{33}$ , are given by the following equations:

$$\lambda_{11} = \frac{1}{\langle \overline{u'_1 u'_1} \rangle - \frac{\langle \overline{u'_1 u'_3} \rangle^2}{\langle \overline{u'_3 u'_3} \rangle}}$$

$$\lambda_{22} = \langle \overline{u'_2 u'_2} \rangle^{-1}$$

$$\lambda_{33} = \frac{1}{\langle \overline{u'_3 u'_3} \rangle - \frac{\langle \overline{u'_1 u'_3} \rangle^2}{\langle \overline{u'_1 u'_1} \rangle}}$$

$$\lambda_{13} = \frac{1}{\langle \overline{u'_1 u'_3} \rangle - \frac{\langle \overline{u'_1 u'_1} \rangle \langle \overline{u'_3 u'_3} \rangle}{\langle \overline{u'_1 u'_1} \rangle}}$$

where  $\langle \overline{u_1} \rangle$  is the Eulerian mean longitudinal velocity (defined such that  $\langle \overline{u_2} \rangle = 0$ );  $\langle \overline{u'_1 u'_1} \rangle (= \sigma_u^2)$ ,  $\langle \overline{u'_2 u'_2} \rangle (= \sigma_v^2)$ , and  $\langle \overline{u'_3 u'_3} \rangle (= \sigma_w^2)$  are the Eulerian variances of the three velocity components (subscript  $v$  represents the lateral velocity and subscript  $w$  represents the vertical velocity);  $\langle \overline{u'_1 u'_3} \rangle (= \langle \overline{w' u'} \rangle)$  is the Reynolds (mean momentum) stress; and  $\langle \varepsilon \rangle$  is the mean turbulent kinetic-energy dissipation rate (note that both index and meteorological notations are used interchangeably). These flow statistics can be computed from a one-dimensional Eulerian second-order closure model developed for canopy flows and described next. We estimate  $dt$  from the Lagrangian integral time scale  $T_L$  given by

$$T_L = \frac{\frac{1}{2}(\sigma_u^2 + \sigma_v^2 + \sigma_w^2)}{\langle \varepsilon \rangle}.$$

For our simulations, we set  $dt = 0.05T_L$ , which is sufficiently fine to ensure numerical convergence. In this model, the existence of a coherent time scale  $T_L$  ensures that turbulent excursions are autocorrelated. That is, a fluid element carrying a seed experiencing an updraft is likely to experience an updraft in the next time step as compared to a downdraft. A recent study by Pope (2000) demonstrated that such a model for the Lagrangian acceleration is in close agreement with direct numerical simulation (DNS) calculations in rough-wall boundary-layer flows, and this further added confidence about the realism of our approach to this particular problem.

The closure model for computing  $\langle \overline{u} \rangle$ ,  $\langle \overline{u'w'} \rangle$ ,  $\langle \overline{u'^2} \rangle$ ,  $\langle \overline{v'^2} \rangle$ ,  $\langle \overline{w'^2} \rangle$ , simplifies to the following set of ordinary differential equations (ODEs) upon temporal and spatial averaging of the conservation-of-momentum equations and following standard closure approximations of canopy flows (Katul and Albertson 1998):

mean momentum:

$$0 = -\frac{d\langle \overline{u'w'} \rangle}{dz} - C_d a(z) \langle \overline{u} \rangle^2$$

tangential stress budget:

$$\begin{aligned}
0 = & -\langle \overline{w'^2} \rangle \frac{d\langle \overline{u} \rangle}{dz} + 2 \frac{d}{dz} \left( q \lambda_1 \frac{d\langle \overline{u'w'} \rangle}{dz} \right) \\
& - \frac{q \langle \overline{u'w'} \rangle}{3 \lambda_2} + C q^2 \frac{d\langle \overline{u} \rangle}{dz}
\end{aligned}$$

longitudinal velocity variance:

$$\begin{aligned}
0 = & -2 \langle \overline{u'w'} \rangle \frac{d\langle \overline{u} \rangle}{dz} + \frac{d}{dz} \left( q \lambda_1 \frac{d\langle \overline{u'^2} \rangle}{dz} \right) + 2 C_d a(z) \langle \overline{u} \rangle^3 \\
& - \frac{q}{3 \lambda_2} \left( \langle \overline{u'^2} \rangle - \frac{q^2}{3} \right) - \frac{2}{3} \frac{q^3}{\lambda_3}
\end{aligned}$$

lateral velocity variance:

$$0 = \frac{d}{dz} \left( q \lambda_1 \frac{d\langle \overline{v'^2} \rangle}{dz} \right) - \frac{q}{3 \lambda_2} \left( \langle \overline{v'^2} \rangle - \frac{q^2}{3} \right) - \frac{2}{3} \frac{q^3}{\lambda_3}$$

vertical velocity variance:

$$0 = \frac{d}{dz} \left( 3 q \lambda_1 \frac{d\langle \overline{w'^2} \rangle}{dz} \right) - \frac{q}{3 \lambda_2} \left( \langle \overline{w'^2} \rangle - \frac{q^2}{3} \right) - \frac{2}{3} \frac{q^3}{\lambda_3}$$

where  $q = \sqrt{\langle \overline{u'_1 u'_1} \rangle}$  is a characteristic velocity scale related to the turbulent kinetic energy;  $C_d$  is a drag coefficient ( $\sim 0.2$ );  $a(z)$  is a leaf-area density (recall that  $\text{LAI} = \int_0^h a(z) dz$ );  $\lambda_j = a_j L_{ws}$ , with  $L_{ws}$  a characteristic length scale specified using the formulation in Katul and Albertson (1998) and not permitted to increase at a rate larger than  $k = 0.4$ , the von Karmann constant; and  $a_1$ ,  $a_2$ ,  $a_3$ , and  $C$  are determined so that the flow conditions well above the canopy reproduce well-established surface-layer similarity relations. With estimates of the four constants ( $a_1$ ,  $a_2$ ,  $a_3$ ,  $C$ ), as described in Table 2 and the Appendix, the above five ODEs can be solved iteratively for the five flow variables  $\langle \overline{u} \rangle$ ,  $\langle \overline{u'w'} \rangle$ ,  $\langle \overline{u'^2} \rangle$ ,  $\langle \overline{v'^2} \rangle$ ,  $\langle \overline{w'^2} \rangle$ , which are used to drive the Lagrangian model described earlier. The Appendix describes the computational grid, numerical scheme, and boundary conditions.

#### Model implementation

In the CELC model calculations,  $10^5$  seeds were released for each stand age. For each seed, a  $u_*$  was drawn from the following:

$$u_* = b \left[ \log \left( \frac{1}{1 - \mu} \right) \right]^{1/c}$$

where  $\mu \in [0, 1]$  is a uniformly distributed random number. If the modeled velocity at release height did

not exceed 0.5 m/s, another  $u_*$  was drawn. (This velocity is estimated as the minimum velocity necessary to overcome frictional forces and commence rotating the cups of a standard cup anemometer. Hence, dislodging seeds would require at minimum this threshold velocity.)

The trajectory of each seed was computed using Eq. 1 with the velocity fluctuations being modeled from the numerical solution of Eq. 3. For each  $u_*$  the  $\lambda$  parameters in Eq. 3 were computed from the second-order closure model. Once the forest floor intercepted a seed, its distance from the release point was determined. The dispersal kernel was then computed as the histogram of all these distances (i.e.,  $10^5$  distances). An example of the resulting kernel is shown in Fig. 2 for illustration. We draw attention to four features about this kernel.

1) More than six decades of distances and five decades of probability densities were simulated by CELC; hence, “extreme” value statistics were well resolved.

2) Because of a preset threshold velocity (0.5 m/s) on seed dislodging, none of the seeds will fall exactly at the source location.

3) Most of the seeds will fall between  $10^1$  and  $10^2$  m (mode is  $\sim 25$  m).

4) Few seeds moved via meso-transport dispersal—or dispersal distances farther than the height of the atmospheric boundary layer ( $\sim 10^3$  m during daytime conditions).

We summarized the kernel statistics for each stand age in Table 3 and then discussed these summaries. Note that while the  $u_*$  distributions and terminal velocity were identical for the two stand ages, both seed-release height and leaf-area densities differed. The effect of seed-release height was accounted for in the initial conditions of Eq. 1. The effect of leaf-area density differences was accounted for in the term  $C_d a$  of the Eulerian model, which then affects the  $\lambda$  parameters in Eq. 3, and ultimately the trajectories in Eq. 1.

## RESULTS

Dispersal distance ( $D_{99}$ ) from source widened more than four-fold as seed-release height increased from age 16 to 25 years (Table 3). Most released seeds were subject to local neighborhood dispersion (LND) and these distances ranged from 0.05 to 0.14 km (Table 3). The fraction of seeds caught in updrafts above the canopy traveled at a maximum distance ( $D_{\max}$ ) of 11.9 km in the 16-yr-old stand and at a distance of 33.7 km in the older 25-yr-old stand (Table 3).

Using these coupled Eulerian–Lagrangian closure (CELC) simulation results for a stand at age 16 years and assuming release of  $10^5$  seeds/ha annually, we predicted roughly 440 seeds ( $10^5 \times 0.0044$ ) will be uplifted and transported via long-distance dispersal (LDD) (Table 3). About 70 seeds ( $10^5 \times 0.0007$ ) will reach distances in excess of 1 km from source (Table 3), a distance too great to serve as a biocontainment

TABLE 2. Dimensionless closure constants used in the Eulerian model calculations.

Constant	Value
$a_1$	0.30
$a_2$	1.58
$a_3$	20.8
$\alpha$	0.07
$C$	0.12
$C_d$	0.2

*Notes:* The dimensionless constants  $a_1$ ,  $a_2$ , and  $a_3$  relate the characteristic length scales describing the flux-transport term, the return-to-isotropy term, and the viscous-dissipation term in the second-order closure model to the basic mixing length scale, respectively. They are computed assuming that the dimensionless standard deviations of the longitudinal ( $A_u$ ), lateral ( $A_v$ ), and vertical ( $A_w$ ) velocity components, respectively, are  $A_u = 2.2$ ,  $A_v = 2.0$ , and  $A_w = 1.4$  above the canopy (as discussed in Katul et al. [2006]). Here, these velocity standard deviations are normalized by the friction velocity ( $u_*$ ) above the canopy. The dimensionless constant  $\alpha$  relates the mixing-length scale of the model to the canopy leaf-area density and drag coefficient,  $C_d$ . The constant  $C$  is a similarity constant related to the pressure–velocity interactions and can be computed from  $a_1$ ,  $a_2$ ,  $a_3$ ,  $A_u$ ,  $A_v$ , and  $A_w$ .

zone. Of these, roughly four seeds will survive as first-year seedlings, assuming germination rates of 6% (DeSteven 1991). Each year, a 10-ha field trial at a 16-yr-old stand would disperse 40 seeds that would colonize at distances in excess of 1 km from source.

Results were similar for a stand at age 25 years where nearly 420 ( $10^5 \times 0.0042$ ) seeds were uplifted and transported via LDD (Table 3). About 25% of the escaped seeds reached distances in excess of 1 km (e.g., Fig. 2), and of these roughly six seeds per year germinated. For a 10-ha transgenic planting at age 25 years, this would translate into 60 seedlings colonizing at least 1 km from source each year.

## DISCUSSION

We used computer simulation based on measurements from a plantation established by wild forest-tree seeds because experiments using transgenic trees is prohibited under U.S. regulatory statutes (*available online*).<sup>6</sup> Long-distance dispersal distances of *Pinus taeda* seeds ranging from 11.9 to 33.7 km were consistent with an anecdotal report of pine seed dispersal at distances of 8–25 km in treeless areas of the Southern Hemisphere (Richardson et al. 1994). Long-distance seed-dispersal estimates are greater by an order of magnitude when compared to average seed-dispersal distances of 60 to 90 m for *P. taeda* (Baker and Langdon 1990). The explanation is that use of average dispersal distance is less informative about the dispersal processes due to its bimodal nature.

However, the mode of the predicted CELC kernel in Fig. 2 was 20–50 m, quite close to the average dispersal distances reported elsewhere (Baker and Langdon

<sup>6</sup> <http://www.aphis.usda.gov/brs>

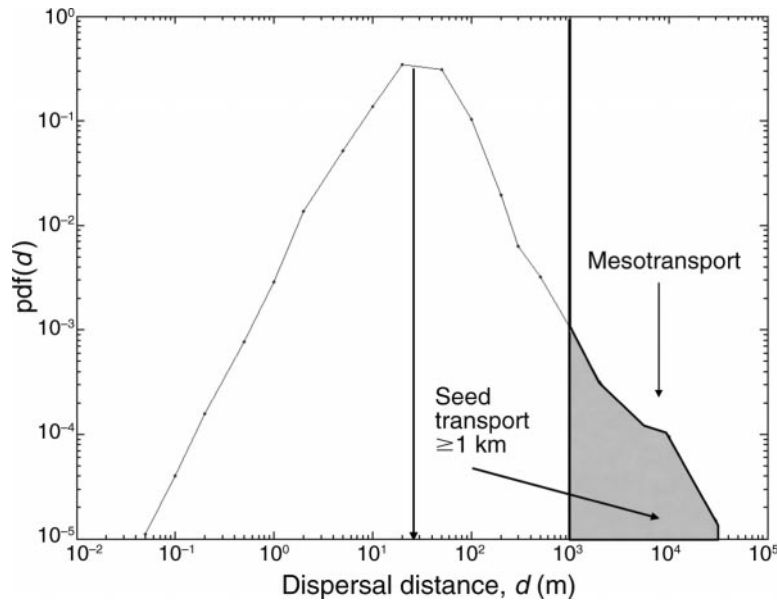


FIG. 2. The variation of the modeled seed dispersal kernel (pdf, probability density function) expressed as a function of dispersal distance ( $d$ ) from the release source for the 25-yr-old stand is shown on a log-log scale, given the wide range in values of distance and probability. The pdf should be interpreted as the fraction of released seeds (dimensionless) dispersing at a distance  $d$  from the source. Note that the area under the entire density function is unity if all the released seeds fall on the ground. The percentage of seeds moving at meso-transport distances of 1 km or greater are shown by the shaded area. Note that the log-log presentation is not "area preserving."

1990). In any event, long-distance dispersal of pine seeds, like long-distance pollen dispersal (Katul et al. 2006) can move on the order of kilometers.

Our predicted results, when extrapolated to transgenic seed dispersal distances, are conservative for three reasons:

1) Model parameters were based on measurements from a moderately low-fertility site (Oren et al. 2001) so improved conditions on more fertile sites or on fertilized sites will accelerate seed production and rate of height growth with a predictable enhancement of absolute seed production and seed escapes via long-distance dispersal (LDD).

2) LDD predictions reported here excluded sonic anemometer data from occasional hurricane-force winds

occurring during autumn seed dispersal because equipment power outages typically occur during hurricanes. As a result, LDD seed movement may be underestimated by use of the fitted Weibull function for the 30-min sonic anemometer velocity measurements.

3) Seed production in *P. taeda* plantations is underestimated for elevated atmospheric CO<sub>2</sub> conditions. Elevated CO<sub>2</sub> triples seed production (LaDeau and Clark 2001) and range expansion is predicted to double for *P. taeda* (Iverson and Prasad 2002). Elevated CO<sub>2</sub> predictions over the next 50 years directly challenge our model's assumption of constant fecundity over time.

What are the implications for gene-flow dynamics for genetically modified or transgenic forests? Our

TABLE 3. Input and output variables for CELC model (Fig. 1) for *Pinus taeda* seed dispersal distances.

Variable	16-yr-old stand	25-yr-old stand
Input		
Seed-release height (m)	12.3	19.0
Leaf-area index, LAI	4.1	5.5
Output		
Local-neighborhood diffusion, LND (km)	0.05	0.14
$D_{99}$ (km)†	0.79	3.59
$D_{max}$ (km)‡	11.90	33.70
Fraction of uplifted seeds	0.0044	0.0042
Fraction of seeds dispersed > 1 km	0.0007	0.0011

† The distance traveled by 99% of the seeds.

‡ The distance traveled by 99.9% of the seeds.

modeling results confirm that gene flow for transgenic conifers is more complex when compared to transgenic food crops (Williams 2005). Each year, more transgenic conifer seeds and pollen will travel farther from the same plantation for at least three reasons: (1) long-distance dispersal occurs at micro-, meso-, and possibly macro-transport distances, (2) seed production starts at least a decade prior to timber harvest and annually recurs for the same plant, and (3) forest canopy heights increase with age, which in turn increases the dispersal distances.

First, *Pinus taeda* disperses unwanted seeds as well as pollen on the order of kilometers. Establishing of remote satellite colonies via LDD seed movement has high probability of success because dispersed *P. taeda* seeds retain viability for a few months prior to germination (Little and Somes 1959, Baker and Langdon 1990). Gene flow from transgenic conifer seeds deserves closer scrutiny in conjunction with pollen movement.

Second, uplifted seed movement depends on realistic seed production within the Duke Forest plantation. Our estimates are conservative because producing  $10^5$  seeds per year is low for a species characterized by heavy annual seed production at mature ages (Cain and Shelton 2001). Annual production of 74 000 sound seeds per hectare is poor for a mature *P. taeda* forest, and 198 000 sound seeds per hectare is a good crop (Baker and Langdon 1990). Consider too that numbers of LDD seeds will increase when sound seed yields exceed 1.6 million per hectare per year for *P. taeda* old-growth forests (Cain and Shelton 2001).

Third, the annual height increment of the forest canopy with age increases dispersal distances. Even though we held released seed quantity constant between plantation ages 16 and 25 years, dispersal distance still increased with age. This was a net outcome of two age-related changes in forests: (1) added annual height for seed release increased the distance of LDD seed transport but (2) increased LAI (leaf-area index) attenuated wind speed and other flow statistics within the canopy, decreasing the distance. Seed-release height proved to be the more important of these two factors, resulting in increased LDD distances with age.

Only a few escaped transgenic seedlings can cause colonization and spread if the transgene is favored by selection (Meagher et al. 2003, Williams and Davis 2005). Our results here merely point to the inadequacy of biocontainment zones for *P. taeda*, contrary to the recommendations for transgenic eucalypt field trials (Linacre and Ades 2004).

These seed dispersal distance predictions can be used for ecological risk analyses associated with the commercial release of transgenic *P. taeda* (Williams and Davis 2005). Focusing on long-distance dispersal (LDD) distances is more realistic for this type of risk analysis than using averages for seed dispersal distances because seed dispersal and satellite colonization

of transgenic conifers moving farther from source poses problems for adjacent landowners who manage for indigenous forests or practice less intensive silviculture. Risk analysis should focus on LDD distances, taking into account that there are no public conservation banks or in situ repositories for commodity forest species in the United States to serve as insurance against harm to indigenous conifer forests. Our results suggest that escape of transgenic *P. taeda* seeds is a certainty under the planting conditions represented in this study.

#### ACKNOWLEDGMENTS

Support from the USDA-Forest Service Southern Research Station Grant Number SRS 04-DG-1133010 and National Science Foundation Grant Number 0454650 to C. Williams is acknowledged. G. Katul, C. Williams, and R. Oren acknowledge support from the Center on Global Change at Duke University. Measurements and seed dispersal modeling were supported by the Office of Science (BER), U.S. Department of Energy through the Terrestrial Carbon Processes Program (TCP) Grant Number DE-FG02-00ER63015 and Grant Number DE-FG02-95ER62083, and through BER's Southeast Regional Center (SERC) of the National Institute for Global Environmental Change (NIGEC) under through the Southeast Regional Center (SERC) of the National Institute for Global Environmental Change (NIGEC) under Cooperative Agreement Number DE-FC02-03ER63613 to R. Oren and G. Katul. C. G. Williams and G. G. Katul made equal contributions to this paper.

#### LITERATURE CITED

- Baker, J. B., and O. G. Langdon. 1990. *Pinus taeda* L., loblolly pine. Pages 497–512 in R. M. Burns and B. H. Honkata, technical coordinators. *Silvics of North America: 1. Conifers*. Agriculture Handbook 654. USDA Forest Service, Washington, D.C., USA.
- Cain, M. D., and M. G. Shelton. 2001. Twenty years of natural loblolly and shortleaf pine seed production on the Crossett Experimental Forest in southeastern Arkansas. *Southern Journal of Applied Forestry* 25:40–45.
- Clark, J., C. Fastie, G. Hurtt, S. T. Jackson, C. Johnson, G. King, M. Lewis, J. Lynch, S. Pacala, I. C. Prentice, E. W. Schupp, T. Webb, III, and P. Wyckoff. 1998. Reid's paradox of rapid plant migration. *BioScience* 48:13–24.
- DeSteven, D. 1991. Experiments on mechanisms of free establishment in old-field succession: seedling survival and growth. *Ecology* 72:1076–1088.
- Hengeveld, R. 1989. *Dynamics of biological invasions*. Chapman and Hall, London, UK.
- Iverson, L. R., and A. M. Prasad. 2002. Potential redistribution of tree species habitat under five climate change scenarios in the eastern United States. *Forest Ecology and Management* 155:205–222.
- Katul, G. G., and J. D. Albertson. 1998. An investigation of higher-order closure models for a forested canopy. *Boundary Layer Meteorology* 89:47–74.
- Katul, G. G., C. G. Williams, M. Siqueira, D. Poggi, A. Porporato, H. McCarthy, and R. Oren. 2006. Spatial modeling of transgenic conifer pollen. Chapter 10 in C. G. Williams, editor. *Landscapes, genomics and transgenic conifers*. Springer-Verlag, New York, New York, *in press*.
- La Deau, S. L., and J. S. Clark. 2001. Rising CO<sub>2</sub> levels and the fecundity of forest trees. *Science* 292:95–98.
- LeCorre, V., N. Machon, R. J. Petit, and A. Kremer. 1997. Colonization with long-distance seed dispersal and genetic structure of maternally inherited genes in forest trees: a simulation study. *Genetical Research* 69:117–125.

- Linacre, N. A., and P. K. Ades. 2004. Estimating isolation distances for genetically modified trees in plantation forestry. *Ecological Modelling* **179**:247–257.
- Little, S., and H. A. Somes. 1959. Viability of loblolly pine seed stored on the forest floor. *Journal of Forestry* **57**:848–849.
- Mann, C. C., and M. L. Plummer. 2002. Biotechnology—forest biotechnology edges out of the lab. *Science* **295**:1626–1629.
- Meagher, T. R., F. C. Belanger, and P. R. Day. 2003. Using empirical data to model transgene dispersal. *Philosophical Transactions of the Royal Society London, Series B* **358**:1157–1162.
- Moody, M. E., and R. N. Mack. 1988. Controlling the spread of plant invasions: the importance of nascent foci. *Journal of Applied Ecology* **25**:1009–1021.
- Nathan, R., G. G. Katul, H. S. Horn, S. M. Thomas, R. Oren, R. Avissar, S. W. Pacala, and S. A. Levin. 2002. Mechanisms of long-distance dispersal of seeds by wind. *Nature* **418**:409–413.
- Nichols, R. A., and G. M. Hewitt. 1994. The genetic consequences of long-distance dispersal during colonization. *Heredity* **72**:312–317.
- Oren, R., D. Ellsworth, K. Johnsen, N. Phillips, B. Ewers, C. Maier, K. Schäfer, H. McCarthy, G. Hendrey, S. McNulty, and G. G. Katul. 2001. Soil fertility limits carbon sequestration by forest ecosystems in a CO<sub>2</sub>-enriched atmosphere. *Nature* **411**:469–472.
- Oren, R., B. E. Ewers, P. Todd, N. Phillips, and G. G. Katul. 1998. Water balance delineates the soil layer in which moisture affects canopy conductance. *Ecological Applications* **8**:990–1002.
- Pope, S. B. 2000. *Turbulent flows*. Cambridge University Press, Cambridge, UK.
- Richardson, D. M., P. A. Williams, and R. J. Hobbs. 1994. Pine invasions in the Southern Hemisphere: determinants of spread and invadability. *Journal of Biogeography* **21**:511–527.
- Rodean, H. 1996. *Stochastic Lagrangian models of turbulent diffusion*. American Meteorological Society, Boston, Massachusetts, USA.
- Schäfer, K. V. R., R. Oren, D. S. Ellsworth, C. T. Lai, J. D. Herrick, A. C. Finzi, D. D. Richter, and G. G. Katul. 2003. Exposure to an enriched CO<sub>2</sub> atmosphere alters carbon assimilation and allocation in a pine forest ecosystem. *Global Change Biology* **9**:1378–1400.
- Soons, M. B., G. Heil, R. Nathan, and G. G. Katul. 2004. Determinants of long-term seed dispersal by wind in grasslands. *Ecology* **85**:3069–3079.
- Thomson, D. J. 1987. Criteria for the selection of stochastic models of particle trajectories in turbulent flows. *Journal of Fluid Mechanics* **180**:529–556.
- Williams, C. G. 2005. Framing the issues on transgenic pine forests. *Nature Biotechnology* **23**:1–3.
- Williams, C. G., and B. H. Davis. 2005. Rate of transgene spread via long-distance seed dispersal in *Pinus taeda*. *Forest Ecology and Management* **217**:95–102.

#### APPENDIX

Description of the coupled Eulerian–Lagrangian closure (CELC) model computations (*Ecological Archives* A016-008-A1).