Forest models defined by field measurements: I. The design of a northeastern forest simulator[1]

STEPHEN W. PACALA[2]
Department of Ecology and Evolutionary Biology, The University of Connecticut, Storrs, CT 06268, U.S.A.
CHARLES D. CANHAM
Institute of Ecosystem Studies, Millbrook, NY 12545, U.S.A.
AND
J.A. SILANDER, JR.
Department of Ecology and Evolutionary Biology, The University of Connecticut, Storrs, CT 06268, U.S.A.

PACALA, S.W., CANHAM, C.D., and SILANDER, J.A., JR. 1993. Forest models defined by field measurements: I. The design of a northeastern forest simulator. Can. J. For. Res. **23**: 1980–1988.

We introduce a new spatially explicit model of forest dynamics. The model is constructed from submodels that predict an individual tree's growth, survival, dispersal, and recruitment, and submodels that predict the local availability of resources. Competition is entirely mechanistic; plants interfere with one another only by depleting resources. We also describe maximum likelihood methods for estimating each of the submodels from data collected in the field. Over the past two years, we collected the necessary data for the dominant tree species in the Great Mountain Forest (Norfolk, Conn.). We report estimates of submodels for each species, and show that the calibrated population dynamic model predicts the structure and dynamics of natural forests. Finally, we contrast our model with the JABOWA–FORET family of forest models.

PACALA, S.W., CANHAM, C.D., et SILANDER, J.A., JR. 1993. Forest models defined by field measurements: I. The design of a northeastern forest simulator. Can. J. For. Res. **23** : 1980–1988.

Nous présentons un nouveau modèle spatialement explicite de dynamique forestière. Le modèle est construit à partir de sous-modèles qui prédisent la croissance des arbres individuellement, la survie, la dispersion, le recrutement et la disponibilité locale des ressources. La compétition est entièrement mécaniste, les plantes compétitionnant entre elles uniquement par l'épuisement des ressources. Nous décrivons aussi les méthodes pour estimer avec le maximum de probabilité chacun des sous-modèles à partir des données collectées sur le terrain. Au cours des deux dernières années, nous avons ramassé les données nécessaires pour les espèces d'arbre dominant dans la "Great Mountain Forest" située à Norfolk au Connecticut. Nous présentons des estimations des sous-modèles pour chaque espèce et nous montrons que le modèle calibré de dynamique de population prédit la structure et la dynamique des forêts naturelles. Finallement, nous comparons notre modèle à la famille de modèles forestiers JABOWA–FORET.

[Traduit par la rédaction]

Introduction

Models of the dynamics of forests are perhaps the most widely studied class of models in the ecological literature. The vast majority of forest models are derived from JABOWA, which was developed by Botkin et al. (1972) for northern hardwood forests in New Hampshire. This model was later modified and extended, primarily by H. Shugart and his colleagues, to forests ranging from boreal regions to the tropics. Shugart's earliest model was called FORET (Shugart and West 1977), and so we refer to the class of models derived from JABOW as "JABOWA–FORET" models.

A JABOWA–FORET model consists of one or more spatial cells, each occupying 0.01–0.1 ha. Each cell contains many trees and the spatial positions of individuals within the cell are unspecified. The model is constructed from submodels that predict the growth, mortality, and recruitment of each tree, and the resources (light, water, and nutrients) available to each tree. The submodels are designed to be consistent with available physiological information and to be calibrated primarily using information already published in the forestry literature. This feature may be viewed as a principal strength of JABOWA–FORET models and has contributed greatly to the rapid extension of the approach to forests across the globe.

JABOWA–FORET models have been subjected to an extensive series of tests and are capable of reproducing critical features of natural forests. For example, the models are able to reproduce the observed altitudinal zonation in New Hampshire (Botkin et al. 1972), the response of southern Appalachian forests to chestnut blight (Shugart and West 1977), the changes in species composition and population structure during succession (reviewed in Shugart 1984), and the response of forests to post-Pleistocene climate change (Davis and Botkin 1985).

Nevertheless, the information necessary to estimate several critical features of the models directly is simply not in the published literature. These features are currently inferred from the available information and assigned, or are omitted altogether. Generally, the range of dynamic behavior that a model is capable of exhibiting grows with the complexity of the model. The blessing of complex models is that they can hope to capture nature's complexity. The curse of complex models is that they can sometimes predict observed phenomena even if they are fundamentally wrong. For example, the "Game of Life" is a cellular automaton (model in which spatial cells change state due to interactions with neighbouring cells) used primarily by computer hobbyists. Although this model contains many fewer parameters than JABOWA–FORET, it is capable of universal computation. That is, the Game of Life is capable of exhibiting any phenomenon that can be simulated on a computer, including every phenomenon that has ever been predicted by JABOWA–FORET.

[1]This paper was presented at the International Conference on Forest Vegetation Management: Ecology, Practice and Policy held April 27 – May 1, 1992, Auburn University, Auburn, Alabama, U.S.A., and has undergone the Journal's usual peer review.

[2]Author to whom all correspondence should be addressed at the following address: Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544, U.S.A.

Printed in Canada / Imprimé au Canada

No one would suggest that the Game of Life is an appropriate model of forest dynamics because the rules governing state transitions do not describe (at least in any obvious way) processes known to be important in forests. A Game of Life that could exhibit forest dynamics under some set of conditions would provide neither explanation nor accurate prediction of the response to novel conditions (such as the response of forests to global climate change). The Game of Life is a good mimic, but a poor teacher and prophet.

For these reasons, we initiated a project two years ago to design a forest model from components that could be estimated directly from data collected in the field, and to collect the necessary information for the dominant tree species in a forest in northwestern Connecticut: American beech (*Fagus grandifolia* Ehrh.) (FAGR), eastern hemlock (*Tsuga canadensis* (L.) Carr.) (TSCA), sugar maple (*Acer saccharum* Marsh.) (ACSA), red maple (*Acer rubrum* L.) (ACRU), yellow birch (*Betula lutea* Michx.f.) (BELU), white ash (*Fraxinus americana* L.) (FRAM), white pine (*Pinus strobus* L.) (PIST), black cherry (*Prunus serotina* Ehrh.) (PRSE), red oak (*Quercus rubra* L.) (QURU), and white oak (*Quercus alba* L.) (QUAL). Although we attempted to start from scratch when formulating this model, the result is properly viewed as an intellectual descendent of JABOWA–FORET. Like JABOWA–FORET, our model is mechanistic. Plant performance is a function of light, water, and nutrient availability and competition occurs only through resource depletion. Also like JABOWA–FORET, our model makes population dynamic forecasts by predicting the fate of each individual using growth, mortality, recruitment, and resource submodels. The principal distinction between our model and JABOWA–FORET is that our submodels were designed *simultaneously* with maximum likelihood estimators necessary to estimate them from simple field measurements. Thus, the submodels represent a compromise between what one would like to know and what can be known with reasonable effort. To date we have produced a calibrated model for forests of seven of the 10 species (the above list minus PRSE, QURU, and QUAL).

The purpose of this paper is to report the design of our forest model and to contrast it with the design of JABOWA–FORET. Because our model is intended as a vehicle for extrapolating from tractable field measurements to community dynamics, an understanding of the model's design requires an understanding of the associated empirical and statistical methods. We thus provide an overview of the methods and estimated submodels described in detail elsewhere (Ribbens et al. 1993; Canham et al. 1994; S.W. Pacala et al.[3]; R. Kobe et al. [4]). We also review some initial tests and analyses of the model and discuss these in relation to JABOWA–FORET models.

JABOWA–FORET model

Growth submodels

In a JABOWA–FORET model, individual trees have sigmoid growth under ideal conditions, with a species-specific asymptotic size. An individual's actual growth rate is determined by multiplying its ideal growth rate by a series of species-specific "growth modifying functions" (GMF) that describe the effects of resource availability and climate on growth. Here, we focus solely on the light GMF, because we have completed the calibration of our model only for the case in which light is the sole resource causing within-stand variation in performance.

Foresters have published extensive tables that place tree species into categories of shade tolerance (i.e., Baker 1949). Physiological ecologists have shown that photosynthetic rates of shade-intolerant species typically increase slowly with increasing light to a high asymptote while rates for shade-tolerant species increase rapidly to a low asymptote (Bazzaz 1979). In JABOWA–FORET models, each tree species is assigned to one of two or three shade-tolerance classes, and each class is assigned a different light GMF. Like photosynthetic responses, the GMF for shade intolerants increases slowly to a high asymptote and the GMF for shade tolerants increases rapidly to a low asymptote (reviewed in Shugart 1984). It is thus *assumed* that published shade-tolerance classification can be used to predict the functional response of photosynthesis to light, and that the functional response of growth to light is congruent with the photosynthetic response. Moreover, it is assumed that all species in the same shade-tolerance class share the same growth response.

Mortality submodels

Mortality may be caused by either by density-independent factors such as windstorm or fire or by density-dependent factors such as shading or nutrient competition. Many JABOWA–FORET models have detailed components, based on field data, that govern mortality caused by density-independent factors (see Shugart 1984). Resource-dependent mortality, in contrast, is typically governed in a simple way. The probability of mortality increases to an admittedly arbitrary value of $0.368 \cdot \text{year}^{-1}$ if an individual's radial growth rate falls to less than $1\ \text{mm} \cdot \text{year}^{-1}$ (Shugart 1984) or less than 10% of its growth rate under ideal conditions (Solomon 1986). The same growth dependent mortality function is used for all species in a model.

Recruitment submodels

In most JABOWA–FORET models, juveniles are not produced by the trees in the model. Rather, in each iteration, recruits are drawn from a fixed list of species. These models contain neither dispersal nor the important population dynamic feedback between abundance and recruitment. The selection of recruits may be entirely random, or may be modified by factors such as the level of herbivory or the availability of microsites that favor germination and establishment (Shugart 1984). Several of the most recent versions (e.g., Smith and Urban 1988; Urban 1990) now include both "closed" recruitment and dispersal. In these, some recruits are drawn from a fixed list, while the others are produced by the trees in the model and disperse among spatial cells.

Resource submodels

Light availability is governed by Beer's Law, with each individual shading all shorter individuals in the same spatial cell (Shugart 1984). Usually, sunlight is assumed to come from directly overhead.

SORTIE model

We call our model SORTIE, to acknowledge its descendence from FORET (sortie and foray are synonyms) and its reliance

[3]S.W. Pacala, C.D. Canham, J.A. Silander, Jr., and R.K. Kobe. Sapling growth as a function of resources in a north temperate forest. Can. J. For. Res. In review.

[4]R. Kobe, S.W. Pacala, J.A. Silander, Jr., and C.D. Canham. Juvenile tree survivorship as a component of shade tolerance. Ecology. In review.

on high-level sorting algorithms. In SORTIE, each tree occupies a unique spatial position (there are no spatial cells), the light and growth submodels account for seasonal and diurnal movement of the sun, and all recruits are produced by the trees in the model. Below we describe the estimation of the submodels. Unless otherwise specified, all data comes from the 2400 ha Great Mountain Forest in Norfolk, Conn. This site is located at 400–500 m elevation in the northwest corner of Connecticut.

Growth submodels

To estimate a growth submodel for a species, we simply regress radial growth (taken from rings) against diameter and measurements of local light, water, and nutrient availability. Whole-season light availability is obtained by fish-eye photography (see Canham 1988*a*), water by time-domain reflectometry (TDR), and nutrient availability (standing pools and mineralization rates) by standard buried bag techniques. Again, we focus in this paper solely on light. Although we have collected and analyzed all of the nutrient and water data, we have not completed the estimation of all components of the model relating to nutrients and water.

Our measure of local light availability is Canham's (1988*a*) "general light index" (GLI). This index integrates knowledge of the seasonal and diurnal movement of the sun, the mix of diffuse and direct-beam radiation, and the local spatial distribution of canopy openness (taken from a digitized canopy photograph) into a single index of whole season light availability (in units of percentage of full sun). We have confirmed with arrays of quantum sensors that this index is correlated, with slope not significantly different from one, to total photosynthetically active radiation in closed and "gappy" canopies dominated by a suite of different species (Canham et al. 1994).

To date, we have estimated growth submodels for individuals ranging in height from a few cm (seedlings) to 5 m. For each of 50–60 individuals per species, we took a fish-eye photograph directly above the individual (at the top and center of the crown) in midsummer and harvested it in late fall. We then regressed the width of the most recent annual ring against GLI. We tried a series of regression models (S.W. Pacala et al., to be published, see footnote 3), and the best was

$$[1]\quad \text{Ring Width} = \text{Radius}\left(\frac{P_1\ \text{GLI}}{\frac{P_1}{P_2} + \text{GLI}}\right) + \alpha$$

where P_1 is the asymptotic relative growth rate (ring width/radius), P_2 is the slope of the relative growth rate at zero light, and α is a normally distributed random variable with zero mean and variance: $\sigma^2 = C$ [Predicted Ring Width]^D, where C and D are estimated constants. Regressions in which radius was raised to an arbitrary (estimated) power or in which ring width reached zero at nonzero GLI fit the data nonsignificantly better for all species (likelihood ratio tests, S.W. Pacala et al., to be published, see footnote 3).

The functions for eight species in Fig. 1 (parameter values in Table 1) show striking interspecific variation that does not conform to the assumptions of JABOWA–FORET models. In particular, the most shade tolerant species (TSCA and FAGR) have low-light radial growth rates that are lower than those of some more intolerant species (i.e., BELU). Also, the conifers have higher high-light radial growth rates that the angiosperms.

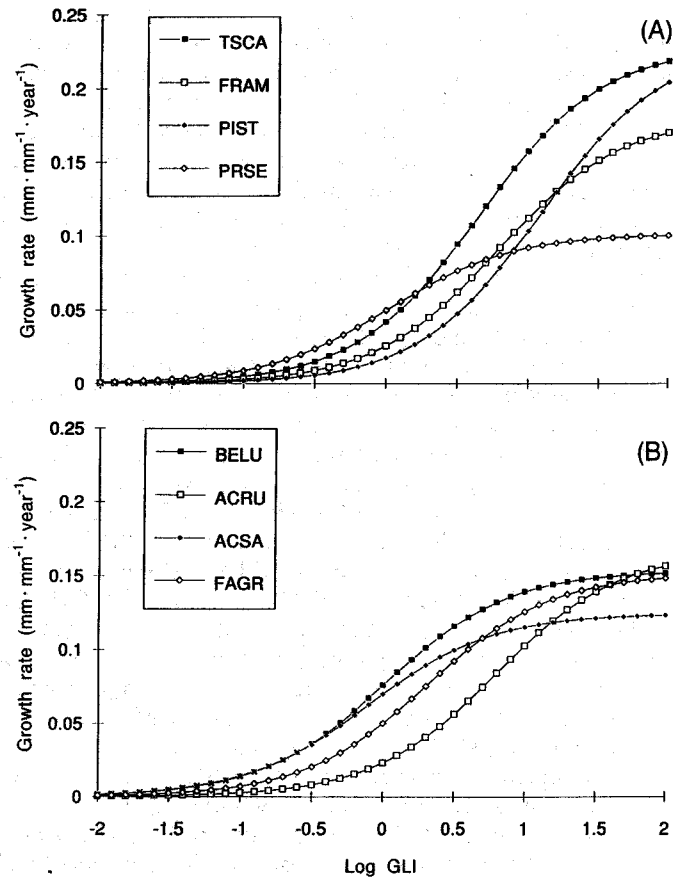

FIG. 1. Growth rate (ring width/radius per year (mm · mm^{-1} · year^{-1})) as a function of light level ($\log_{10}$(GLI)) for eight species. The curves were produced by eq. 1, using the parameter values in Table 1.

Although the data indicate that ring widths of saplings are proportional to radius at any given light level, this relationship cannot extend to canopy trees. At present, we assume the constant area increment law in the model (see Phipps 1967): the cross sectional area of a ring cannot exceed 63 cm^2 (a ring width of 2 mm for a tree 100 cm in diameter). Radial growth is given by the *smaller* of the ring widths predicated by eq. 1 and the constant area increment law.

To translate radial growth rates into height growth, we regressed height against diameter using samples of 50–100 individuals of each species (S.W. Pacala et al., to be published, see footnote 3). These ranged in height from a few cm to canopy height. Again, we tried a series of regression equations, and the best was

$$[2]\quad \text{Height (m)} = A\left[1 - \exp\left(-\frac{S}{A}\text{Diameter (cm)}\right)\right] + \alpha$$

where A is the asymptotic height, S is the slope of the function at zero diameter and, as before, α is a normal random variable with zero mean and variance that increases as a power law with the predicted mean. The parameter estimates in Table 1, to some extent, explain the counterintuitive interspecific variation among the growth functions in Fig. 1. For example, TSCA will not overtop the shade intolerants at high light because of its small value of S. However, the overall pattern is still at odds with the assumptions of JABOWA–FORET; TSCA

TABLE 1. Parameter estimates and 95% confidence limits

	FAGR	TSCA	ACSA	BELU	ACRU	FRAM	PIST
P_1	0.152	0.229	0.125	0.154	0.167	0.181	0.230
	(0.183, 0.132)	(0.262, 0.201)	(0.149, 0.105)	(0.181, 0.125)	(0.222, 0.127)	(0.252, 0.126)	(0.296, 0.188)
P_2	0.075	0.051	0.159	0.150	0.027	0.030	0.019
	(0.115, 0.052)	(0.071, 0.036)	(0.227, 0.112)	(0.227, 0.105)	(0.040, 0.018)	(0.039, 0.023)	(0.026, 0.015)
A	34.6	29.6	24.8	23.2	25.7	32.5	38.4
	(42.7, 29.5)	(39.0, 23.9)	(27.5, 22.2)	25.1, 21.1)	(27.9, 23.3)	(35.2, 29.9)	(44.5, 33.4)
S	1.06	0.73	1.87	1.89	1.89	1.69	1.00
	(1.12, 1.02)	(0.77, 0.69)	(1.98, 1.78)	(2.01, 1.76)	(2.01, 1.79)	(1.79, 1.58)	(1.06, 0.93)
U	4.27	5.74	8.84	3.13	7.66	15.80	4.70
	(5.22, 3.59)	(7.08, 4.64)	(11.16, 7.03)	(3.68, 2.67)	(8.99, 6.54)	(20.84, 11.36)	(5.92, 3.75)
V	0.001	0.250	1.58	0.638	1.06	1.90	0.44
	(0.068, 0.000)	(0.349, 0.154)	(1.80, 1.38)	(0.850, 0.460)	(1.19, 0.94)	(2.16, 1.69)	(0.59, 0.29)
h ($\times 10^{-4}$)	19.57	59.91	7.44	0.00	3.63	0.32	1.03
	(34.70, 8.01)	69.34, 50.75)	(11.49, 4.75)	(0.01, 0.00)	(4.77, 2.65)	(0.52, 0.03)	(1.91, 0.37)
E_i	2.75	2.75	0.92	0.92	0.92	0.92	0.92
	(2.77, 2.40)	(2.77, 2.40)	(0.93, 0.91)	(0.93, 0.91)	(0.93, 0.91)	(0.93, 0.91)	(0.93, 0.91)
W	0.152	0.100	0.107	0.109	0.108	0.095	0.087
	(0.166, 0.138)	(0.111, 0.089)	(0.118, 0.097)	(0.122, 0.096)	(0.115, 0.101)	(0.104, 0.085)	(0.097, 0.077)
D	0.664	0.846	0.580	0.539	0.488	0.319	0.413
	(0.726, 0.602)	(0.891, 0.801)	(0.623, 0.537)	(0.611, 0.467)	(0.546, 0.430)	(0.380, 0.258)	(0.457, 0.369)

NOTE: Upper and lower confidence limits are given in parentheses. See the text for an explanation of symbols in the first column with the exception of W (canopy radius (m) / DBH (cm)) and D (canopy thickness (m) / tree height (m)).

and FAGR have slower height growth in deep shade than several of the more intolerant species (i.e., BELU).

Mortality submodels

As in JABOWA–FORET, resource-dependent mortality is modeled indirectly, through an effect of growth rate on the probability of mortality. The difficulty in estimating mortality submodels is that mortality rates are low, and so an enormous sample would be required to regress mortality rate against growth rate using standard techniques (binomial regression). However, together with R. Kobe, we have developed an alternative that requires growth rates from a sample of only 35–50 standing dead individuals and 35–50 live individuals, and a count of live and dead individuals along transects (R. Kobe et al., to be published, see footnote 4). Let the distribution of growth rates in a site be $X(g)$, and the mortality function be $M(g)$, where g is growth rate. Then, the distribution of growth rates for standing dead individuals is

$$[3]\quad Y_D(g) = \frac{X(g)M(g)}{\int_{-\infty}^{\infty} M(g)X(g)dg}$$

and the distribution for live individuals is

$$[4]\quad Y_L(g) = \frac{[1 - M(g)]X(g)}{\int_{-\infty}^{\infty} [1 - M(g)]X(g)dg}$$

We assume a flexible distribution for $X(g)$ (usually a γ-density) and a functional form for $M(g)$ and then derive the conditional distributions $Y_D(g)$ and $Y_L(g)$ using eqs. 3 and 4. We then derive a maximum-likelihood estimator from these conditional distributions and the binomial distribution (for the transect counts). This estimator provides estimates and confidence limits for the parameters of the mortality submodel, $M(g)$, and the parameters of $X(g)$.

The reason that the method works is that the sample of dead individuals will contain more low growth rates that the sample of live individuals, if the probability of mortality increases as growth rate decreases. The estimator finds the parameter values of $M(g)$ that optimally reshape $X(g)$ into $Y_L(g)$ and $Y_D(g)$.

When collecting standing dead individuals, R. Kobe et al. (to be published, see footnote 4) restricted the sample to individuals that had died within 2 or 3 years, using a series of characters. The function $M(g)$ gives the probability of mortality over this length of time.

R. Kobe et al. (to be published, see footnote 4) tried a series of different functional forms for $M(g)$, and the best was

$$[5]\quad M(g) = \exp(-U[\text{Average Ring Width (mm)}])^V)$$

where U and V are constants and average ring width is the arithmetic average over the previous 5 years. Better fits were obtained from a 5-year average than from any other interval between 1 and 10 years. R. Kobe et al. (to be published, see footnote 4) also considered two-parameter functions in which mortality approached a value less than one at $g = 0$ and functions that included W/R^z as an index of growth, where z is an estimated constant, W is average ring width, and R is stem radius. The former provided poorer fits and the latter yielded nonsignificantly better fits and estimates of z close to zero for each species. Because ring width increases with radius at any given light level (eq. 1), eq. 5 indicates that light-dependent mortality decreases with size.

In contrast to the assumption in JABOWA–FORET models, large interspecific variation among estimated mortality functions is evident in Fig. 2 (R. Kobe et al. (to be published, see footnote 4), parameter values in Table 1). These functions must be evaluated in the context of the growth functions. For example, the fact that the shade-tolerant species, ACSA, has relatively high mortality at low rates of growth does not mean that this species has high mortality at low light. This is because ACSA has a high rate of growth at low light. The

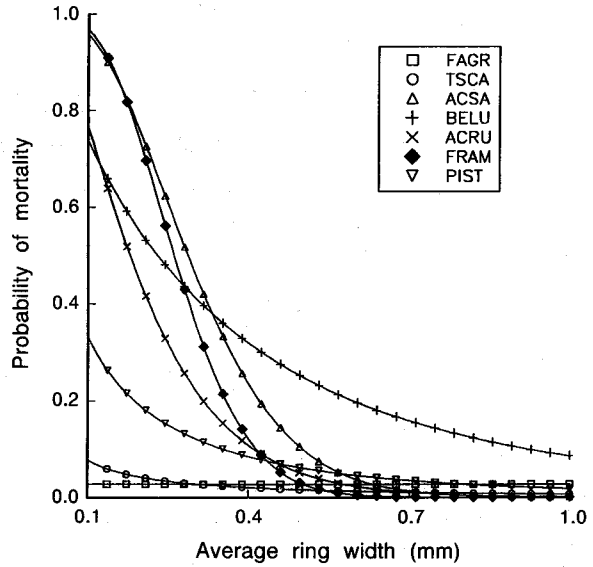

FIG. 2. Probability of mortality as a function of growth rate (average ring width (mm) over the previous 5 years) for seven species. The curves were produced by eq. 5, using the parameter values in Table 1.

order of mortality rates for 2 cm diameter saplings at 1% light (GLI) is FAGR < TSCA < ACSA < BELU < PIST < ACRU < FRAM, which agrees well with conventional wisdom (R. Kobe et al., to be published, see footnote 4).

R. Kobe et al. (to be published, see footnote 4) also produced replicate estimates for four species (FAGR, ACSA, BELU, and FRAM) from a site in Michigan, as well as two species (FRAM and ACSA) from a Connecticut site with a very different soil (limestone parent material instead of granite). The estimates for these sites agree remarkably well with the functions in Fig. 2, suggesting that trees have a highly deterministic starvation response. The one exception is that ACSA's function indicates greater shade tolerance on calcareous soil. This result is perhaps responsible for the dominance of ACSA on calcareous soils in the region (R. Kobe et al., to be published, see footnote 4).

Like the mortality submodels of JABOWA–FORET, our mortality submodels also include purely random disturbance. Each individual has a constant probability of dying from density-independent factors, in addition to its growth-dependent probability of mortality. We also have developed subroutines that create catastrophic disturbances (windthrow or fire) with user-specified frequencies, severities, and sizes. We have not, to date, developed species-specific functions for density-independent disturbance (e.g., owing to differential susceptibility to fire or windstorm), although it would be possible to do so given appropriate data.

Recruitment submodels

Our recruitment submodels predict the "seedling shadow" of an individual tree: the density of seedlings produced by the tree as a function of the size of the tree and the distance from the tree. We, together with E. Ribbens, have developed a maximum likelihood method that permits one to estimate a recruitment submodel for each species from a map of adults in a stand and a census of seedlings in square-meter quadrats placed within the stand (Ribbens et al. 1993).

Suppose that the mean density of seedlings (number per square meter) produced by a tree falls off with distance as

$$[6]\quad \text{Seedling Density} = F(\text{DBH})^2\exp(-h\ \text{distance}^3)$$

where F and h are constants. Diameter at breast height (DBH) is raised to the second power and distance to the third, because these numbers provided better fits than any other integers from 1 to 4 (Ribbens et al. 1993). Now, let d_{ij} be the distance between the ith tree and the jth seedling quadrate. The *total* density of seedlings expected in the jth quadrat (N_j) is then

$$[7]\quad N_j = \sum_{i=0}^{q} FD_i^2\exp(-hd_{ij}^3)$$

where D_i is the DBH of the ith conspecific tree and q is the number of conspecific trees in the stand.

Ribbens et al. (1993) developed a maximum likelihood estimator in which the mean of a Poisson random variable for the jth quadrat was N_j, and used this method to estimate values and confidence limits of h and F for each species in a range of sites (parameter values in Table 1). The method works by finding values of F and h that bring the seedling shadows of the mapped trees into optimal agreement with the spatial distribution of conspecific seedlings.

Although the distance decay of the seedling shadow is strongly conserved for each species across replicate samples, the total number of seedlings produced does vary substantially among sites and years, presumably because of masting and stochastic variation in pre-establishment mortality (Ribbens et al. 1993). The estimated submodels show that mean dispersal distance decreases with the degree of shade tolerance. The mean dispersal distances (m) are: TSCA, 4.1; FAGR, 5.9; PRSE, 8.0; ACSA, 8.2; QUAL, 8.6; QURU, 8.7; ACRU, 10.4; PIST, 15.8; FRAM, 23.3; BELU, >50.0 (Ribbens et al. 1993).

Resource submodels

To be useful in a population dynamic model, a resource submodel must predict the same measure of local resource availability that is used to predict growth. We selected GLI as an index of local light availability in the growth submodel, rather than some simpler measure, because numerous studies have shown that the growth of saplings in nature depends critically on nonvertical shading and illumination. In Connecticut, a canopy gap located 25 m north of a sapling has no effect, but a gap located 25 m south may have a dramatic effect (see Canham 1988*a*). Moreover, subtle changes in the timing, number, and placement of gaps about a location apparently alter the competitive ranking of shade-tolerant species during regeneration (Canham 1988*b*). Thus, the problem in designing a light submodel is to predict the GLI for a location from the sizes, locations and species identities of trees in the vicinity.

We, together with A. Finzi and D. Burbank, developed this submodel in five steps (see Canham et al. 1994). First, we developed regression equations for each species that give canopy diameter and depth from DBH (parameter values in Table 1). These equations supplement the DBH–height equations described earlier. Second, we determined the locations and diameters of all trees in six circular stands (50 m diameter). These stands were chosen to provide a wide range of local species composition. We used the regression equations to construct a three-dimensional map of the canopy in each stand

(assuming cylindrical crowns). Third, we took a series of fish-eye photographs at several heights throughout each stand. After digitizing the photographs, we divided each into a series of small regions with different zenith and azimuth angles and determined the percentage of sky visible in each region (canopy openness of the region). Fourth, consider the linear path corresponding to the mean zenith and azimuth angle of a region. We used the canopy maps to determine the number of crowns of each species intercepted by the path, and the length of path occupied by each species. This provided a data set with several thousand lines, each containing a percentage of sky visible along a path, and the numbers and path lengths of each tree species along the path. We then developed a maximum-likelihood estimator, based on the β-distribution, to provide estimates and confidence limits for species-specific light extinction coefficients. We selected estimates based on the number of crowns intercepted, rather than the path length intercepted, because the former provided better fits than the latter. Thus, a crown of species i reduces canopy openness by the factor $\exp(-E)$, where E is the light extinction coefficient for the species (estimates in Table 1). Finally, to validate the method, we mapped a single larger stand (2 ha) with a heterogeneous species composition, and *measured* GLIs at a series of locations. We then *calculated* GLIs using the regression equations and estimated light extinction coefficients, and showed that these predict the actual values (Canham et al. 1994).

Note that there are only two different light extinction coefficients in Table 1. The most shade tolerant species (FAGR and TSCA) have a higher coefficient than the remaining species. Although the method can yield separate estimates for each species, pooling species into the two classes in Table 1 provides a nonsignificantly poorer fit (likelihood ratio test, Canham et al. 1994). Also, although the extinction coefficients in Table 1 may appear small, it must be remembered that these are averages for trees in *closed* canopy and that they predict measured values of GLI.

The shade cast by an individual tree is determined both by the tree's crown geometry and its light extinction coefficient. In order of decreasing shade cast: TSCA > FAGR > ACSA > BELU > ACRU > PIST > FRAM (Canham et al. 1994). This order confirms Horn's (1971) result that the shade cast by an individual increases with the degree of shade tolerance.

The population dynamic model

The model is written in C and contains the diameter, species identity, and x- and y-coordinates of every individual. At the beginning of each iteration, SORTIE uses the resource submodels to determine the resources available to each individual. For example, to determine the amount of light available to an individual, the model computes a GLI from a fish-eye photograph taken above the individual. The procedure for calculating GLIs relies on efficient algorithms to identify trees that will shade a location, and to compute the set of zenith and azimuth angles shaded by each tree. SORTIE then uses the local resource availabilities together with the growth submodels to grow each plant, and uses these growth rates together with the mortality submodels to determine each plant's probability of survival. After pseudorandom coin tosses determine which plants are killed, SORTIE uses the recruitment submodels to determine the number and spatial positions of all recruits produced by every tree. This completes one interaction. Each iteration has a duration of 5 years (forecasts the state of the forest 5 years into the future). By repeatedly iterating the model, we can forecast long-term changes in the abundance, age and size structure, and spatial distribution of all species.

Both speed and memory place limits on the spatiotemporal scales that SORTIE is able to simulate. Because the model tracks the state of every tree, the memory limit is determined primarily by the number of trees that can be accommodated. The record for an individual tree occupies approximately 20 bytes. Assuming approximately 6000 individuals greater than 5 years of age per hectare, each megabyte of memory can store the records for between 8 and 9 ha. In SORTIE, over 90% of the execution time is required to calculate GLIs. JABOWA–FORET models are typically much faster because they assume that light comes from a single direction (usually vertical). A PC operating at 0.2 megaflops executes a simulated year of SORTIE at a rate of approximately 1 $h \cdot km^{-2}$. An IBM RS 6000 operating at approximately 20 megaflops requires roughly 0.01 $h \cdot km^{-2} \cdot year^{-1}$.

A note on the recruitment submodels in the current version

The current version of the model does not use estimates of F described in the recruitment submodels section for two reasons. First, we do not yet have a sufficiently long time series of values to characterize the high temporal variability in reproduction. Three to 5 years typically separate mast years, and we currently have only 2 years of data. Second, because the time step in the model is 5 years, juveniles that enter the plot are actually 5 years of age. We cannot reliably use our method of estimating mortality submodels for trees less than 5 years of age. Because small individuals may disappear rapidly after dying (or be consumed), the sample of small standing-dead individuals in a location is likely to be highly biased. Thus, we currently have no reliable way of converting estimates of F into numbers of 5-year-old juveniles. What is needed is a growth-dependent mortality function for ages 1–5. Fortunately, this interval is short enough to be covered by an experiment and the necessary experiment is in progress. We expect to be able to produce mortality submodels for 1–5 year olds within 3 years.

The problem of the unknown F-values is perhaps not overly severe because preliminary sensitivity analyses indicate that the model is less sensitive to changes in F than to changes in the other parameters. We currently set values for F in two ways. First, the published literature contains information on seed production and survival through 5 years of age of some of the species. We choose values that are consistent with this information. Second, we can adjust the values of F to yield average densities of 1–5 year olds as reported in the literature. Fortunately, these two methods are generally in close agreement. In the preliminary test described below, F-values were chosen so that the relative fecundities of trees of identical DBH were 1 (FAGR) : 5 (BELU and PIST) : 3 (all other species). The lower fecundity of FAGR reflects the predominantly clonal reproduction (root sprouting) of this species in Great Mountain Forest. The higher fecundities of BELU and PIST reflect the high rate of seed production of these species (Houle and Payette 1990; Wendel and Smith 1990).

A preliminary test of the model

Two difficulties in testing a forest simulation model are that successional return to old-growth forest takes a long time, and that species composition during succession depends on

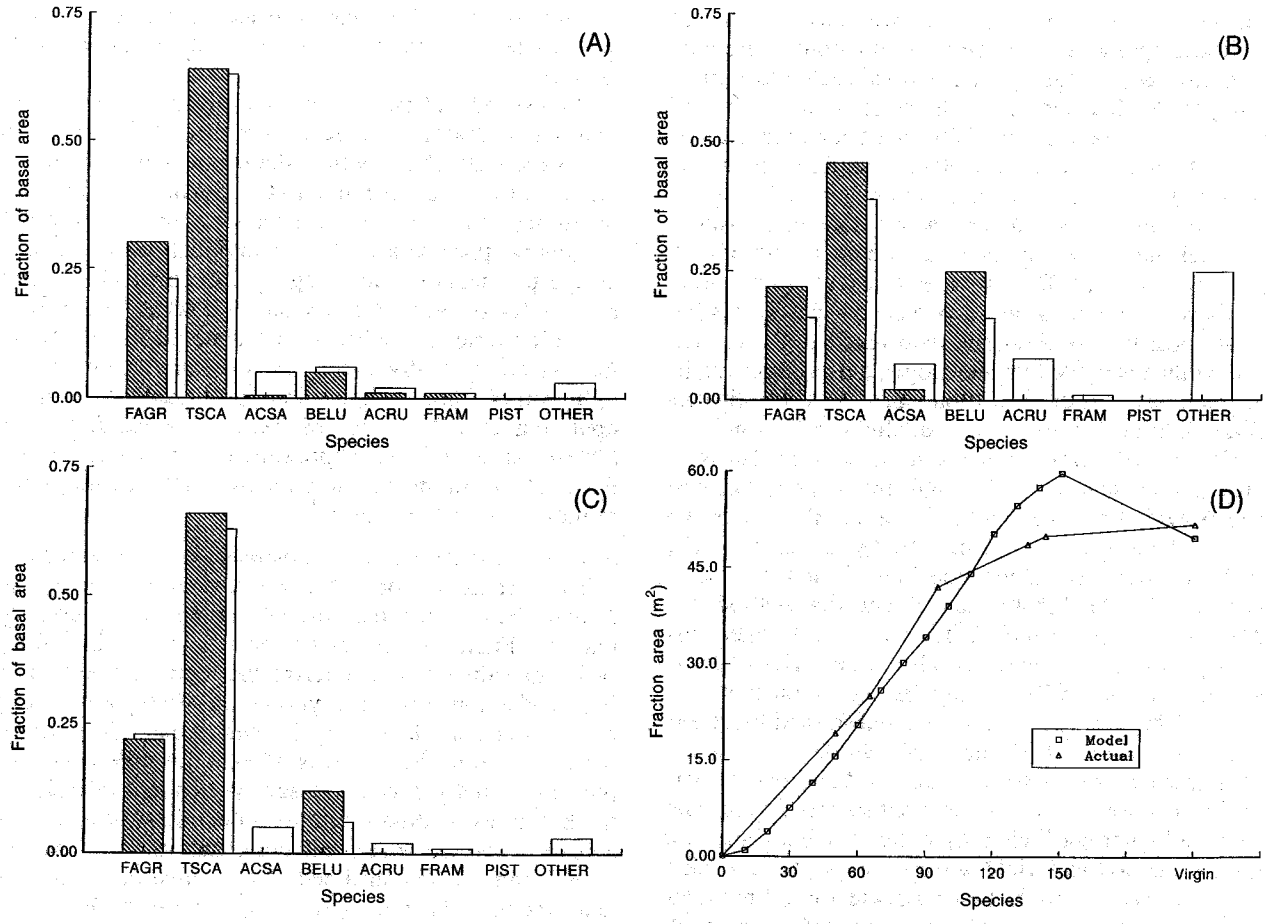

FIG. 3. Predicted versus actual succession following clear-cutting of hemlock–beech old growth. (A) Initial composition. (B) Composition 50 years after clear-cutting. (C) Composition at 1000 years. (D) The predicted and actual accumulation of total basal area (ha^{-1}) following clear-cutting. For Figs. 3A–3C open bars give the species composition as reported in the literature; shaded bars give the species composition predicted by the model.

the initial composition of a disturbed site and its surroundings. Because of the lack of long-term historical data on succession, one is left with three alternatives for testing the model: (*i*) an historical reconstruction of a forest (based on old stumps, tip-up mounds, and the pollen record), (*ii*) data from a successional sere, or (*iii*) data from a short chronosequence. Here, we focus on the second two alternatives.

We selected the test data sets as follows. We reviewed studies that report the composition of old-growth forests in locations similar to our Connecticut sites (sites in southern New England, New York, and Pennsylvania reported in Hough and Forbes 1943; Nichols 1913; Potzger 1946). These studies confirm that there are basically two types of old-growth: hemlock–beech dominated old-growth and white pine dominated old-growth, and that the species composition of each type is conserved across the region. We computed the arithmetic average of the species compositions reported for each type. The average composition for hemlock–beech old-growth is given by the open bars in Figs. 3*a* and 3*c*, and the average composition for white pine old-growth is given by the open bars in Fig. 4. In Figs. 3 and 4, the "other" category contains primarily QURU, QUAL, and PRSE. Because we have not yet completed estimating all submodels for these species, we normalized the species composition of the FAGR–TSCA–ACSA–BELU–ACRU–FRAM–PIST community (open bars) to sum to one.

We then selected the two studies reporting species composition both initially (before disturbance) and during succession. Hough and Forbes (1943) report species compositions of stands 40–50 years after clear-cutting of hemlock–beech old-growth in Pennsylvania. Similarly, Kelty (1984, 1986) reports species compositions of stands in the Harvard Forest (central Massachusetts) 40–50 years after they were leveled by the 1938 hurricane. Although the Harvard Forest sites were not technically old-growth, they had been abandoned in the early 19th century and had recovered a species composition identical to the hemlock–beech old-growth by the time of the hurricane (Kelty 1984, 1986). The compositions of the two successional stands were very similar, and we averaged them to produce the open bars in Fig. 3*b*.

To augment these studies, we also selected a series that report compositions and basal areas of successional stands located in or near the Great Mountain Forest (where the model was calibrated). Because none of these reports the initial species composition, we report only the total basal area ($m^2 \cdot ha^{-1}$) at each successional stage (Fig. 3*d*). Our hope is that total basal area is less sensitive to initial condition than is species composition. The basal areas in Fig. 3*d* come from the following sources: 50 years (average of stands in northwestern Connecticut and the Harvard forest; Kelty 1984; Stephens and Waggoner 1980; G.R. Stephens, unpublished), 65 and 95 years (average of stands in the Great Mountain Forest; Stephens and Hill 1971;

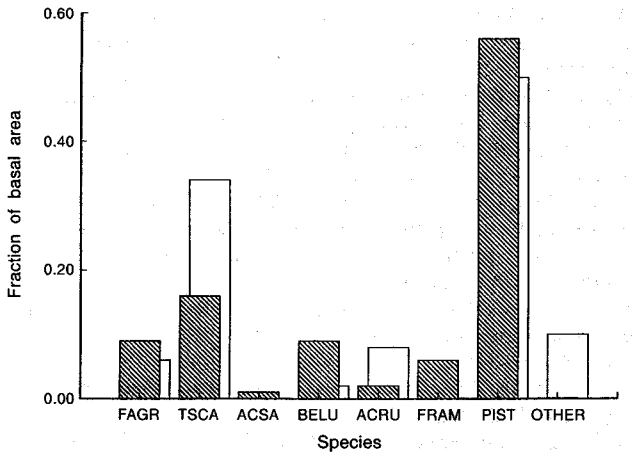

FIG. 4. Predicted (shaded bars) versus actual (open bars) species composition with periodic large-scale disturbance. There were four 0.25-ha catastrophic disturbances in the 1-ha plot during the 1000-year run. The initial composition was the same as in Fig. 3. See text for additional information.

G.R. Stephens, unpublished), 135 years (average of stands in Catlin Woods in northwestern Connecticut; Kelty 1984), and 143 years (average of stands in Catlin Woods; G.R. Stephens, unpublished). Finally, the "virgin" basal area is the average for hemlock-beech old-growth reported in Hough and Forbes (1943) and Potzger (1946).

We tested the model's ability to predict these data with two kinds of simulations. First, we simulated a 4-ha area with an initial species composition equal to the actual average composition of hemlock–beech old-growth (open bars in Fig. 3*a*, but minus the other category). The composition in the central hectare of the modeled stand is given by the shaded bars in Fig. 3*a*. We then clear-cut this central hectare, killing all canopy trees and 0, 25, 50, 75, or 100% of saplings. Because the different rates of sapling mortality during clear-cutting produced qualitatively similar results (with marginally more rapid return to old-growth composition for the lower rates), we report only the results for a rate of 75%. At 50 years, the central clear-cut contained the composition shown by the shaded bars in Fig. 3*b*, and at 1000 years by the shaded bars in Fig. 3*c*. The composition was essentially unchanged for 500 to 1000 years. Note the close correspondence between the predicted (shaded) and actual (open) compositions. The loss of FRAM, ACRU, and ACSA is, perhaps not overly serious because the small simulated plot contained only one or two trees of these species initially. Also, the only disturbances in the simulation were single-tree gaps created by the random deaths of canopy trees. In contrast natural stands inevitably have experienced some multitree disturbances during their development.

The temporal sequence of total basal areas in the central hectare of the simulation is shown by the circles in Fig. 3*d*. The point opposite "virgin" corresponds to year 1000, but total basal area was essentially unchanged from years 300 to 1000. Again, there is obviously close correspondence between the predicted and actual pattern. The discrepancy between predicted and actual values of peak basal area may reflect an artificial lack of lateral inhibition among canopy trees in the model (recall that fish-eye photographs are taken in the model only directly above the center of each tree's crown).

We replicated the above simulation three times with different random number seeds. The results of the replicates were virtually identical to the results in Fig. 3.

The second group of simulations were designed to determine if the model can predict the composition of white pine dominated old-growth. We simulated 1-ha stands with the same initial composition as in Fig. 3 (hemlock–beech old-growth), but included four 0.25-ha catastrophic disturbances that occurred at random times and locations during the 1000-year runs. As before, the predictions of replicate runs were similar, and so we include the results of only one run (Fig. 4). Note that the model correctly predicts the dominance of PIST and the subdominance of TSCA and FAGR.

Discussion

This study shows that it is possible to calibrate a mechanistic forest-simulation model with currently available levels of funding. The result predicts coexistence without relying on an unchanging list of species from which recruits are drawn. It predicts the observed composition of old-growth communities, with and without periodic catastrophic disturbance. Finally, it predicts observed successional change.

The community-level predictions of the model must be caused by interspecific variation among the estimated submodels. Several interspecific trade offs are evident in Table 1 and Figs. 1 and 2. A traditional ordering of the species by the degree of shade tolerance is: (FAGR, TSCA) > ACSA > (BELU, ACRU) > (FRAM, PIST). As shade tolerance defined by this order increases: (*i*) mean dispersal distance generally decreases (Ribbens et al. 1993), (*ii*) survival at low light generally increases (R. Kobe et al., to be published, see footnote 4; S.W. Pacala et al., to be published, see footnote 3), (*iii*) shade cast by an individual tree generally increases (Canham et al. 1994), (*iv*) the rate of height growth at high light generally decreases (S.W. Pacala et al., to be published, see footnote 3), and (*v*) rates of height growth at low light, or rates of radial growth at any light generally neither monotonically increase nor decrease (S.W. Pacala et al., to be published, see footnote 3).

The challenge now is to determine how these relationships among the "strategies" of species determine the predicted community-level patterns. We are approaching this problem both through formal sensitivity analysis and by studying the modeled community the way that an empirical scientist studies a forest. We make measurements, formulate hypotheses (often in the form of simple analytically tractable models), and test the hypotheses with simulated experiments. Though this effort is still in its early stages, it is already clear that no source of interspecific variation has little effect on the predicted composition and structure of the forest. Whenever we give the species a common growth, mortality, or recruitment submodel, predicted community-level patterns bear little resemblance to actual patterns (Ribbens et al. 1993; Canham et al. 1994; R. Kobe et al., to be published, see footnote 4; S.W. Pacala et al., to be published, see footnote 3). The implication is that theoretical and empirical studies should focus on ensembles of trade offs such as those in *i–v* above, to determine how interactions among trade offs shape community dynamics and structure.

We again emphasize that SORTIE has much in common with JABOWA–FORET. Once could view SORTIE as a version of JABOWA–FORET that includes explicit spatial positions of trees, a more complex model of light availability, purely closed

recruitment (all recruits produced by trees in the model), and directly calibrated submodels of growth, mortality, and dispersal. Both models are apparently able to predict the dynamics and structure of real forests. JABOWA–FORET can undoubtedly do so under a broader range of conditions than SORTIE because it has been under development for a longer period of time and because it includes features such as nutrient cycling, the effects of climate, and herbivory.

Still, some of the trade offs evident in *i–v* above either contradict assumptions of JABOWA–FORET or are not included in JABOWA–FORET. One example is that BELU actually grows faster than FAGR and TSCA under low-light levels and yet has lower survivorship at low light than either species. We suspect that this property is an essential reason for BELU's position as the third most abundant species in hemlock–beech old-growth. A second example is that the most shade tolerant species in SORTIE have the shortest dispersal. We have written a graphical display for SORTIE that includes a map showing the location, size, and species identity of each tree. Because of short dispersal, the spatial distribution of TSCA and FAGR produced by the model exhibits the pronounced within-species clumping and between-species spatial segregation so evident in natural stands. Theoretical work (Pacala 1986) shows that this phenomenon promotes coexistence. By reducing the number of between-species contacts, short dispersal reduces the mean level of interspecific competition. We suspect that this mechanism is an essential reason for the coexistence of TSCA and FAGR in old-growth forest.

Clearly, it is not enough for models to be merely predictive. If it were, the Game of Life would suffice as a general model. For over 20 years, JABOWA–FORET models have generated significant insights (i.e., Bormann and Likens 1979*a*, 1979*b*). What is more, these models are of proven applied value. But only by tightening the coupling between forest models and empirical studies, will we arrive at an explanatory and predictive theory.

Acknowledgements

We thank the Childs family for their generous hospitality and for the use of the facilities at Great Mountain Forest. We thank the National Science Foundation (BSR-8918616), the Department of Energy (DE-FG02-90ER60933), and NASA for funding the project. We are especially grateful to our project manager Diane Burbank, and to graduate students J. Hill, R. Kobe, E. Ribbens, and A. Finzi, for their hard work, insight, and friendship. Finally, we thank the numerous undergraduate assistants who worked long hours in the field and laboratory.

Baker, F.S. 1949. A revised tolerance table. J. For. **47**: 179–181.

Bazazz, F.A. 1979. The physiological ecology of plant succession. Annu. Rev. Ecol. Syst. **10**: 351–372.

Bormann, F.H., and Likens, G.E. 1979*a*. Pattern and process in a forested ecosystem. Springer-Verlag, New York.

Bormann, F.H., and Likens, G.E. 1979*b*. Catastrophic disturbance and the steady state in northern hardwood forests. Am. Sci. **67**: 660–669.

Botkin, D.B., Janak, J.F., and Wallis, J.R. 1972. Rationale, limitations, and assumptions of a northeastern forest simulator. IBM J. Res. Dev. **16**: 106–116.

Canham, C.D. 1988*a*. An index for understory light levels in and around canopy gaps. Ecology, **69**: 1634–1638.

Canham, C.D. 1988*b*. Growth and architecture of shade-tolerant trees: response to canopy gaps. Ecology, **69**: 786–795.

Canham, C.D., Finzi, A.C., Pacala, S.W., and Burbank, D.H. 1994. Causes and consequences of resource heterogeneity in forests: interspecific variation in light transmission by canopy trees. Can. J. For. Res. **24**. In press.

Davis, M.B., and Botkin, D.B. 1985. Sensitivity of cool-temperature forests and their fossil pollen to rapid temperature change. Quat. Res. **23**: 327–340.

Horn, H.S. 1971. The adaptive geometry of trees. Princeton University Press, Princeton, N.J.

Hough, A.F., and Forbes, R.D. 1943. The ecology and silvics of forests in the high plateaus of Pennsylvania. Ecol. Monogr. **13**: 301–320.

Houle, G., and Payette, S. 1990. Seed dynamics of *Betula alleghaniensis* in a deciduous forest of northeastern North America. J. Ecol. **78**: 677–690.

Kelty, M.J. 1984. The development and productivity of hemlock–hardwood forest in southern New England. Ph.D. dissertation, Yale University, New Havan, Conn.

Kelty, M.J. 1986. Development pattens in two hemlock–hardwood stands in southern New England. Can. J. For. Res. **10**: 885–891.

Nichols, G.E. 1913. The vegetation of Connecticut. II. Virgin forest. Torreya, **13**: 1991–215.

Pacala, S.W. 1986. Neighborhood models of plant population dynamics. II. Multi-species models of annuals. Theor. Pop. Biol. **29**: 262–292.

Phipps, R.L. 1967. Annual growth of suppressed chestnut oak and red maple, a basis for hydrological inference. U.S. Geol. Surv. Prof. Pap. 485-C.

Potzger, J.E. 1946. Phytosociology of the primeval forest in central-northern Wisconsin and upper Michigan, and a brief postglacial history of the lake forest formation. Ecol. Monogr. **16**: 211–250.

Ribbens, E., Silander, J.A., Jr., and Pacala, S.W. 1993. Recruitment in forests: calibrating models to predict patterns of tree seedling dispersal. Ecology. In press.

Shugart, H.H. 1984. A theory of forest dynamics. Springer-Verlag, New York.

Shugart, H.H., and West, D.C. 1977. Development of an Appalachian deciduous forest model and its application to assessment of the impact of the chestnut blight. J. Environ. Manage. **5**: 161–179.

Smith, T.M., and Urban, D.L. 1988. Scale and resolution of forest structural pattern. Vegetatio, **74**: 143–150.

Solomon, A.M. 1986. Transient response of forests to CO_2-induced climate change: simulation modeling experiments in eastern North America. Oecologia, **68**: 567–579.

Stephens, G.R., and Hill, D.E. 1971. Damage, drought, defoliation, and death in unmanaged Connecticut forests. Agric. Exp. Stn. New Haven, Conn. Bull. 718.

Stephens, G.R., and Waggoner, P.E. 1980. A half century of natural transitions in mixed hardwood forests. Agric. Exp. Stn. New Haven, Conn. Bull. 783.

Urban, D.L. 1990. A versatile model to simulate forest pattern: a user's guide to ZELIG version 1.0. Environmental Sciences Department, University of Virginia, Charlottesville.

Wendel, G.W., and Smith, H.C. 1990. *In* Silvics of forest trees of the United States. *Compiled by* R.M. Burns and B.H. Hankala. U.S. Dep. Agric. Agric. Handb. 654. Washington D.C. pp. 476–488.