Filling the gap: A compositional gap regeneration model for managed northern hardwood forests

James D.A. Millington [a,*], Michael B. Walters [b,c], Megan S. Matonis [d], Jianguo Liu [b]

[a] Department of Geography, King's College London, Strand, London WC2R 2LS, UK
[b] Center for Systems Integration and Sustainability, Department of Fisheries and Wildlife, Michigan State University, USA
[c] Department of Forestry, Michigan State University, USA
[d] Department of Forest and Rangeland Stewardship, Colorado State University, USA

ARTICLE INFO

ABSTRACT

Regeneration of trees in canopy gaps created by timber harvest is vital for the sustainability of many managed forests. In northern hardwood forests of the Great Lakes region of North America, regeneration density and composition are highly variable because of multiple drivers that include browsing by herbivores, seed availability, and physical characteristics of forest gaps and stands. The long-term consequences of variability in regeneration for economic productivity and wildlife habitat are uncertain. To better understand and evaluate drivers and long-term consequences of regeneration variability, simulation models that combine statistical models of regeneration with established forest growth and yield models are useful. We present the structure, parameterization, testing and use of a stochastic, regression-based compositional forest gap regeneration model developed with the express purpose of being integrated with the US Forest Service forest growth and yield model 'Forest Vegetation Simulator' (FVS) to form an integrated simulation model. The innovative structure of our regeneration model represents only those trees regenerating in gaps with the best chance of subsequently growing into the canopy (i.e., the tallest). Using a multi-model inference (MMI) approach and field data collected from the Upper Peninsula of Michigan we find that 'habitat type' (a proxy for soil moisture and nutrients), deer density, canopy openness and basal area of mature ironwood (*Ostrya virginiana*) in the vicinity of a gap drive regeneration abundance and composition. The best model from our MMI approach indicates that where deer densities are high, ironwood appears to gain a competitive advantage over sugar maple (*Acer saccharum*) and that habitat type is an important predictor of overall regeneration success. Using sensitivity analyses we show that this regeneration model is sufficiently robust for use with FVS to simulate forest dynamics over long time periods (i.e., 200 years).

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

In forests managed by uneven-aged selection silviculture, juvenile shade-tolerant trees are expected to regenerate naturally beneath forest-canopy gaps created by periodic harvesting of single to small groups of trees (Arbogast, 1957; Nyland, 1998). However, in uneven-aged managed northern hardwood forests in the Great Lakes region of North America regeneration abundance and composition is highly variable (e.g., Matonis et al., 2011). Environmental drivers of this variation include browsing by herbivores (Horsley et al., 2003; Long et al., 2007), seed availability (Ribbens et al., 1994; Garrett and Graber, 1995), physical characteristics of forest gaps and stands (Kern, 2011), competition between seedlings and other vegetation (Randall, 2007; Matonis et al., 2011), and interactions between these factors (Powers and Nagel, 2009; Matonis et al., 2011). Variation in regeneration of dominant canopy species can lead to substantial changes in forest stand composition and structure over time (Seagle and Liang, 2001; Pedersen and Wallis, 2004), thus altering habitat for wildlife (Millington et al., 2011), and posing economic concerns for the timber industry (Donovan, 2005; Racevskis and Lupi, 2006). Predicting regeneration abundance and composition due to environmental drivers is therefore important both for understanding long-term forest dynamics and to inform forest management actions.

To study tree population dynamics, many forest gap models have been developed (e.g., Botkin et al., 1972; Shugart, 1984; Urban, 1990; Bugmann et al., 1996; Larocque et al., 2011). Gap models represent the establishment, growth and senescence of individual trees in and around forest gaps through time. These models are primarily used to investigate ecological questions such as how species interactions influence succession. In contrast, forest growth and yield models are designed and used primarily for

---

* Corresponding author. Tel.: +44 (0) 207 848 2604.
E-mail address: james.millington@kcl.ac.uk (J.D.A. Millington).

planning and managing commercial forests (Liu and Aston, 1995; Weiskittel et al., 2011). Examples include the US Forest Service 'Forest Vegetation Simulator' (FVS, Crookston and Dixon, 2005) and the ORGANON model (Hann, 2011). The focus of forest growth and yield models on overstory trees (for merchantable timber) is at the expense of the representation of other processes such as understory regeneration. Model users often must provide forest growth and yield models with data on new trees regenerating following a timber harvest. For example, beyond the Rocky Mountains (for which a regeneration module has been developed, Ferguson and Carlson, 1993), managers using FVS must specify the species, density, and size of expected new trees at each simulated cycle. Such an approach assumes managers have good knowledge of previous regeneration abundance and composition and that it will not change through time (or that planting will drive regeneration). However, observed spatial (e.g., Matonis et al., 2011) and temporal variation (e.g., Sage et al., 2003) in the multiple environmental drivers of natural regeneration suggest that future regeneration is unlikely to be stable. To overcome these shortcomings, hybrid approaches which combine statistical models of regeneration with established forest growth and yield models are useful for understanding and predicting variation in tree regeneration and its long-term impacts on forest stand conditions (Weiskittel et al., 2011).

Predicting regeneration abundance and composition has long challenged forest scientists and managers (Liu and Aston, 1999), with a variety of approaches used to statistically model regeneration. These approaches include probabilistic (Monserud and Ek, 1977), multi-level regression (Miina and Saksa, 2006), artificial neural networks (Hasenauer and Kindermann, 2002), nearest neighbour imputation (Hassani et al., 2004), and expert-based (Vickers et al., 2011) models. Given the inherent variability in forest regeneration, methods that incorporate stochasticity alongside deterministic functions are particularly important (Miina and Heinonen, 2008). Furthermore, if the regeneration model is to be coupled with a growth and yield model for continued simulation of tree growth, harvest, and mortality, regeneration model structure must consider the requirements of the growth and yield model and the specific questions it will be used to address.

To better understand and predict the consequences of environmental drivers of regeneration in managed uneven-aged northern hardwood forests over the long term, and to contribute to identifying management practices for meeting regeneration targets, we have developed a stochastic, regression-based compositional forest gap regeneration model. Coupling this regeneration model with FVS (Ontario variant, Lacerte et al., 2006; Sharma et al., 2008) and other submodels (e.g., representing deer populations) results in an integrated forest simulation model for evaluating impacts of scenarios of timber and wildlife management on tree regeneration and forest dynamics over long time periods (Millington et al., 2013). Here, we present the parameterization and evaluation of our stochastic gap regeneration model, which takes an innovative approach to represent only trees with the best chance of growing into the canopy and which is specifically designed to provide input to FVS. The model simulates the initial height of the tallest saplings 10 years following gap creation (potentially either advanced regeneration or gap colonizers), and grows them until they are at least 7 m in height when they are passed to FVS for continued simulation. Our approach does not aim to produce a thorough mechanistic model of regeneration dynamics, but rather is one that is sufficiently mechanistically based to allow us to reliably predict regeneration for trees most likely to recruit to canopy positions from readily collectable field data. We also investigate and illustrate regeneration responses (species composition and abundance) to environmental conditions implied by the model by examining scenarios based on observed environmental conditions.

2. Methods

2.1. Data collection and preparation

The data used to develop the forest gap regeneration model presented here were collected from 166 gaps in 28 northern hardwood stands in the central Upper Peninsula (UP) of Michigan, USA (see Appendix A for a summary of stand attributes and Millington et al., 2010 and Matonis et al., 2011 for wider study area details). The overstory canopies of these stands are dominated by commercially valuable sugar maple (*Acer saccharum* Marsh.), but include other species such as American basswood (*Tilia americana* L.), black cherry (*Prunus serotina* Ehrh.), paper birch (*Betula papyrifera* Marshall), red maple (*Acer rubrum* L.), white ash (*Fraxinus americana* L.), and yellow birch (*Betula alleghaniensis* Britton). Ironwood (*Ostrya virginiana* (Mill.) K. Koch) is also prevalent in these stands (Matonis et al., 2011), but compared with other common tree species, the relatively small maximum size of ironwood relegates it to subcanopy status and low economic value. Sugar maple and ironwood are the dominant species regenerating in northern hardwood stands in our central UP study area, representing approximately 75% of trees 2–7 m tall across harvest gaps we surveyed.

Our approach to data collection and modelling is influenced by the intention to provide regeneration data to the FVS model, as part of an integrated simulation model. This integrated simulation model will run at a temporal resolution of 10 years, both because it can take this long for seedlings to recruit to heights beyond the influence of deer (e.g., Matonis, 2009) and because this represents a reasonable timber harvest return interval for our study area. Consequently, we selected stands for data collection that had been harvested between 8 and 12 years prior to surveying (we used a range of gap ages centred around 10 years to maximize the size of our sample). In each stand, six harvest gaps were sampled unless fewer existed or very high sapling densities made data collection prohibitively time-intensive (which was the case in only five of the 28 stands, see Matonis et al., 2011).

At each stand, 'habitat type' – a categorical proxy for soil moisture and nutrient regimes (Burger and Kotar, 2003) routinely used by forest managers in our study area – was determined using diagnostic assemblages of understory vegetation. Habitat type classifications for surveyed stands were: *AOCa* (mesic soil moisture regime; rich to very rich soil nutrient regime), *ATD-Hp* (mesic; medium to rich), *ATD* (mesic; medium to rich), *ATM* (dry-mesic to mesic; medium), or *TMC* (mesic to wet-mesic; medium). Winter deer density was estimated for each stand using spring-time faecal pellet surveys to account for the influence of browse on regeneration (Cote et al., 2004). Pellet group counts from ten transects (50 m × 4 m oriented in a 'bow tie' shape around each stand centre; see Millington et al., 2010) were averaged and converted to an estimate of deer density (deer km^{-2}, after Hill, 2001). To account for seed production potential, estimates of the basal area (m^2 ha^{-1}) of mature (potentially seed bearing) sugar maple and ironwood basal area (trees with DBH > 20 cm and DBH > 5 cm respectively) were estimated for trees in a fixed radius plot (20 m radius from centre of gap). We used a radius of 20 m as the majority of seeds for most northern hardwood species have been found to fall within 20 m of their tree of origin (Ribbens et al., 1994).

In each gap, heights (to the nearest 0.25 m) and species of all saplings (trees 2–7 m tall) in a 154 m^2 circular plot (7 m radius centred in the gap) were recorded. Age at 0.05 m and 1.4 m height was determined for one sugar maple and one ironwood sapling in each of three height strata (1–2 m, 2–4 m and > 4 m) to estimate height

growth rate and to assess whether the sapling was present prior to gap creation ('advanced regeneration') or not ('gap colonizer'). For the sampled trees 2–7 m tall, 95% of sugar maple and 92% of ironwood saplings were 'advanced regeneration'. Extended gap area (EGA, m^2) was estimated (i.e., the area between the boles of gap-edge trees; Runkle, 1981), and converted to canopy gap area (CGA, m^2) using the relationship (after Ménard et al., 2002):

$$CGA = 0.0762 \times EGA^{1.309} \tag{1}$$

As this equation predicts canopy gap area at the time of gap formation, we adjusted the estimate of CGA to account for the lateral growth of overstory branches over a 10-year period (as we were measuring regeneration approximately 10 years after harvest). We assumed that gaps were shaped as an ellipse with a length:width ratio of 1.25 (based on our empirical measurements) and then reduced the length and width of the canopy gap area by 8.97 cm per year (the average rate of lateral branch extension for hardwood species; Webster and Lorimer, 2005) for 10 years. Canopy openness (%), a proxy for light availability (Kobe and Hogarth, 2007), was estimated from hemispherical photographs taken in each gap (Canham et al., 1990 and see Matonis et al., 2011).

2.2. Regeneration model structure

The regeneration model takes a phenomenological approach to simulate heights of gap saplings 10 years following timber harvest and grows them until they are 7 m tall (see Millington et al., 2011 for details), at which point they are passed to FVS for continued growth and harvest. Initial regeneration sapling heights (i.e., 10 years after harvest) are estimated as a function of canopy openness (Eq. (2) in Millington et al., 2011) and subsequent growth rate is modelled as a function of tree height, time since the gap was created, and whether the sapling was advanced regeneration or gap colonizer (Eq. (3) in Millington et al., 2011). All saplings modelled as having initial height ≥ 3 m are assumed to be advanced regeneration (as 100% of empirically sampled saplings with height ≥ 3 m were advanced regeneration) and the probability that a modelled sapling with height < 3 m is advanced regeneration is estimated as a function of sapling height and time since the gap was created (Eq. (4) in Millington et al., 2011).

In our regeneration model we assume that each gap contains space for a total of n, 7 m tall trees. We simulate n or less trees because we are only concerned with the first trees that grow to 7 m to completely fill the gap at that height, thereby out-competing any other trees in the gap (i.e., the n 'gap winners'; see Millington et al., 2011). We estimate n for each gap using:

$$n = \frac{CGA}{CA_7} \tag{2}$$

where CA_7 is the canopy area of a tree 7 m tall (equal to 13.79 m^2 based on our empirical data for sugar maple; see Millington et al., 2011) and n is rounded down to the nearest integer.

For each of the n spaces in a gap ('gap spaces'), we estimate the probability that the gap space is in one of four states 10 years after harvest. The four gap-space states we consider are: (i) occupied by a > 2 m tall sugar maple tree (SM); (ii) occupied by a > 2 m tall ironwood tree (IW); (iii) occupied by a > 2 m tall tree of another species (OT); or (iv) not occupied by a tree > 2 m tall (i.e., empty, ET). We use these categories as sugar maple and ironwood are the dominant regenerating species in our study area and because the absence of regeneration (i.e., empty spaces) is concerning to forest managers. We include 'other species' for completeness. The other species category is heavily dominated by red maple, white ash, and black cherry (representing approximately 88% of non-sugar maple and non-ironwood saplings 2–7 m tall). We consider only

trees with height > 2 m, as these trees have successfully regenerated beyond the influence of deer browse and competing non-tree vegetation. Considering shorter trees is not necessarily more useful for identifying and modelling 'gap winners', as taller trees are more likely to outcompete shorter trees and the probabilities of shorter trees negotiating strong negative stressors such as deer browse are smaller and less discernible (i.e., fewer trees negotiate these stressors and identifying which individuals will succeed is more difficult). Furthermore, our parameters for height growth of regenerating trees (Appendix B in Millington et al., 2011), while significantly different from zero, are not significantly different between species. This helps to justify the assumption that species composition does not change in the 2–7 m height interval, as our approach implies.

The gap-level probability for each of the four gap-space states (i.e., composition probabilities) is estimated by a regression model for composition data (Aitchison, 1982, 1986). Our raw composition data are a vector for each of our empirical gaps specifying the proportion of all saplings with height > 2 m that were sugar maple, ironwood, or other species (i.e., SM, IW, and OT). If the total number of trees with height > 2 m is denoted by t, the proportion of empty spaces (ET) equals zero if $t > n$, otherwise ET = $(n - t)/n$. These raw composition data provide information on the ratios of the components (i.e., gap-space states). The use of standard statistical methods with raw composition data can lead to spurious correlation effects, in part due to the absence of an interpretable covariance structure (Aitchison, 1986). However, transforming composition data, for example by taking logarithms of ratios (log-ratios), enables a mapping of the data onto the whole of real space and the use of standard unconstrained multivariate analyses (Aitchison and Egozcue, 2005). We transformed our composition data with a centred log-ratio transform using the 'aComp' scale in the 'compositions' package (van den Boogaart and Tolosana-Delgado, 2008) in R (R Development Core Team, 2009). These transformed data were then ready for use in a standard multivariate regression model. A centred log-ratio transform is appropriate in our case as our composition data are proportions (not amounts) and the difference between components is relative (not absolute). The 'aComp' transformation uses the centred log-ratio scalar product (Aitchison, 2001) and worked examples of the transformation computation can be found in Tolosana-Delgado et al. (2005).

Our regression model then has the general form:

$$clr(y) \sim \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_i x_i \tag{3}$$

where $clr(y)$ is the centred log-ratio transformed vector of gap-space state probabilities and β_i is the parameter estimate of environmental predictor variable x_i. This regression model can be used to stochastically predict the regeneration composition in gaps 10 years after harvest, given values of environmental predictors for each gap. For each gap, proportions of gap-spaces in the four states predicted by the regression model are interpreted as the probability that each of the n gap-spaces is in the corresponding state. The state of each gap-space is then predicted by comparing a uniform random value from the interval [0,1] to the cumulative gap-space state probabilities (gap-space is assigned the state for which the random value is closest to, but smaller than, the corresponding cumulative probability; e.g., see Table 1).

We take a multi-model inference approach to select the regression model with the lowest Bayesian Information Criterion (BIC, Schwarz, 1978) value and the highest subsequently derived model weighting ($BICwt$, see Millington and Perry, 2011). The model with these criteria is deemed the 'best' model. Our model selection evaluated all possible models for five variables (i.e., $2^5 = 32$). The five predictor variables used in the model are based on our previous work (Matonis et al., 2011, and see rationales provided in Section 2.1) and can be readily derived from output of the integrated simulation model to which this regeneration model will contribute. The

Table 1
Stochastic prediction of gap-space states. At the gap-level, a regression model predicts the proportions of each gap-space state given environmental predictors. These proportions are interpreted as the cumulative probability of gap-spaces being in each state. In this hypothetical example, there are five gap-spaces (i.e., $n = 5$, see Eq. (2)) and so five uniform random values are found (one for each gap-space). Each gap-space is assigned the state for which the random value is closest to, but smaller than, the corresponding cumulative probability.

Gap-level				
Gap-space state	SM	IW	ET	OT
Predicted proportion	0.30	0.30	0.30	0.10
Cumulative probability	0.30	0.60	0.90	1.00
Gap-space level				
Gap-space identifier	1	2	3	4	5
Uniform random value	0.62	0.65	0.93	0.26	0.42
Predicted state	ET	ET	OT	SM	IW

five variables are forest habitat type (*HabitatType*, with *AOCa* as the reference category), deer density (*DeerDensity*), canopy openness (*CO*), basal area of mature sugar maple trees (*AcerBA*) and ironwood trees (*OstryaBA*). Matonis et al. (2011) also found that cover of competing, non-tree, ground vegetation was important for determining tree regeneration. However, we currently do not represent the presence of non-tree vegetation in our simulation model because data on understory composition were not collected and are not available in stand inventory data used to parameterize the integrated simulation model.

Using the variables identified as belonging to the 'best' model, we fit a Bayesian linear regression model and for all variables we use Bayesian Model Averaging (BMA) to find the averaged model (Hoeting et al., 1999; Millington and Perry, 2011). For our Bayesian regression model we used uninformative priors for all parameter estimates ($\sim$Normal[$\mu = 0$, $\sigma^2 = 10,000$]). Model estimation was performed using WinBUGS v.1.4.3 (Spiegelhalter et al., 2003) through the 'R2WinBUGS' package (Sturtz et al., 2005) using three parallel chains with randomly selected starting values for 10,000 iterations (burn-in of 5000 and a thinning rate of 1). All models showed evidence of convergence (i.e., Rhat < 1.01).

2.3. Model testing and analysis

To measure models' predictive performance we simulate compositions of our 166 observed gaps using the 'best', 'averaged', and 'null' models and compare predicted compositions to observed gap compositions. The null model assumes that probabilities for each state in all gaps are equal to the mean observed gap proportions, which are 27% in SM, 21% in IW, 37% in ET, and 15% in OT. As our approach is probabilistic we run 1000 simulations for each model, determining gap-space states by comparing a uniform random value from the interval [0,1] to cumulative modelled probabilities (as described above and see Table 1). For each simulation, we calculate the mean Kendall rank correlation coefficient (tau) between observed and simulated ranks of gap-space state probabilities across all gaps. In each gap, the state with highest observed proportion or simulated probability is ranked as first, and lowest is ranked as fourth. Kendall's tau is calculated using these rankings for each gap and the mean for all gaps is then found. We also calculate the proportion of gaps in a simulation for which *all* gap-space states are ranked correctly (i.e., ranking of probabilities in the gap perfectly matches the ranking of observed state proportions) and the proportion of gaps in which each individual gap-space state is correctly ranked. For each of these model performance measures we calculate mean and maximum values for all simulations. We use paired *t*-tests to evaluate significant differences in performance measures between the null and other models. To ensure we are not over-fitting our 'best' model and to evaluate its performance beyond data used to calibrate it, we use *k*-fold cross validation (Hastie et al., 2001) with $k = 10$, re-fitting the model in each of 1000 runs.

To investigate and illustrate regeneration responses (species composition and abundance) to environmental conditions, we use the model to examine scenarios with different values for driving variables in all gaps and stands. For these scenarios we use parameters of the 'best' model. We examine the influence of each environmental variable in turn, setting gap values at 5th, 25th, 75th and 95th percentiles of empirical values (Table 2). For example, to examine the influence of deer density on regeneration composition and abundance, we run the model assuming that all gaps have deer densities (deer km^{-2}) of 4.4 (5th), 10.9 (25th), 21.0 (75th) and 50.4 (95th percentile). As above, we simulate each scenario 1000 times, generating a new set of uniform random values in the interval [0,1] to compare to estimated cumulative probabilities for each gap state. The R script used to fit models and run all tests and analyses described above can be found in the online supplementary material (Appendix B).

3. Results

3.1. Model selection and parameter estimates

Our model selection analysis shows that the model which minimizes *BIC* (the 'best' model) contains all variables except basal area of mature sugar maple (*AcerBA*, Table 3). This 'best' model has a probability of 0.478 that it is the 'true' model, given that truth is in the set of models considered (Link and Barker, 2006; Millington and Perry, 2011). Habitat type and basal area of ironwood are present in all of the eight models with lowest *BIC* values, and deer density, canopy openness and basal area of mature sugar maple are present in half of those eight. Signs for means of posterior parameter distributions for variables in the 'best' model and the averaged model from BMA (Table 4) are consistent with our previous findings for saplings 1–2 m tall (Matonis et al., 2011), with the exception of canopy openness (described below). For example, in the 'best' model deer density has a negative relationship with sugar maple proportion and a positive relationship with ironwood proportion, with parameter estimate credible intervals (CI) not encompassing zero at the 90% and 95% confidence levels respectively (Table 4). Furthermore, basal area of ironwood is positively related to ironwood proportion and negatively related to other species' proportion (95% CI) and negatively related to empty proportion (90% CI) but has no relationship with sugar maple

Table 2
Variable values used in scenario analyses. Values are for the 5th, 25th, 75th and 95th percentiles of the available data for each variable used as boundary conditions to the model. Variables are defined in Section 2.2 of the text.

Variable	5th	25th	Empirical mean	75th	95th
DeerDensity (km^{-2})	4.40	10.90	18.36	21.00	50.40
CO (%)	4.45	7.94	12.52	16.33	24.63
OstryaBA (m^2 ha^{-1})	0.00	0.03	0.31	0.59	1.28

Table 3

Model selection results for composition models. An '×' indicates the inclusion of the corresponding variable in the model. Also shown for each model are the Bayesian Information Criterion (*BIC*) value, difference between model *BIC* and minimum *BIC* (Δ*BIC*) and the model weight (*BICwt*). The *BICwt* can be interpreted as the probability that the model is true, given that truth is in the set of models considered. Variables are defined in Section 2.2 of the text.

Model	HabitatType	DeerDensity	CO	OstryaBA	AcerBA	BIC	ΔBIC	BICwt
1	×	×	×	×		497.5	0.0	0.478
2	×		×	×		498.7	1.2	0.265
3	×	×	×	×	×	499.5	2.0	0.177
4	×		×	×	×	501.3	3.8	0.073
5	×	×		×		507.6	10.1	0.003
6	×			×		508.3	10.8	0.002
7	×	×		×	×	509.3	11.8	0.001
8	×			×	×	510.6	13.1	0.001

proportion. Habitat type *TMC* is positively related to sugar maple and other species proportions (95% and 90% CI respectively) and negatively related to ironwood and empty proportions (95% and 90% CI respectively). Habitat type *ATD* is positively related to ironwood proportion and negatively related to empty proportion (95% CI) and habitat type *ATM* is positively related to ironwood and sugar maple proportions (95% and 90% CI respectively) and negatively related to empty proportion (95% CI). Counter-intuitively, and in contrast to previous findings for 1–2 m saplings (Matonis et al., 2011), canopy openness (*CO*) is negatively related to sugar maple proportion and positively related to empty proportion (95% CI).

3.2. Predictive performance

According to *t*-tests, all performance measures for both 'best' and averaged models are statistically better than the null model (significant at $p < 0.001$). Our performance measures also show however, that the 'best' model outperforms the averaged model (Table 5). The 'best' model has a mean Kendall tau (maxima in parentheses) across all simulations of 0.228 (0.352), produces gap space state probabilities which rank all four states perfectly in 16.1% (24.7%) of gaps, and correctly ranks the sugar maple, ironwood, other tree species and empty space state probabilities in 40.6% (56.6%), 45.1% (54.2%), 50.0% (58.4%) and 45.3% (56.6%) of

Table 4

Averaged and 'best' composition models. Mean and standard deviation of posterior parameter distributions (mean and SD) are presented for each variable for each gap-space state for both models. Also presented for the averaged model (from Bayesian Model Averaging) is the posterior probability that the variable is present in the true model (*p*). Variables for which the 95% credible interval of a parameter estimate does not encompass 0.0 are shown in bold, and for the 90% credible interval in *italics*. Variables are defined in Section 2.2 of the text.

	Variable	Averaged			Best	
		p	Mean	SD	Mean	SD
Acer saccharum	Intercept	**1.000**	**0.325**	**0.113**	**0.467**	**0.108**
	HabitatTypeATD	0.114	0.015	0.050	0.081	0.093
	HabitatTypeATD-Hp	0.162	−0.020	0.053	−0.067	0.080
	HabitatTypeATM	0.355	0.085	0.133	*0.231*	*0.122*
	HabitatTypeTMC	0.236	0.055	0.118	**0.314**	**0.129**
	DeerDensity	0.170	−0.001	0.002	*−0.005*	*0.003*
	CO	0.202	−0.002	0.004	**−0.012**	**0.005**
	OstryaBA	0.034	−0.003	0.021	0.002	0.088
	AcerBA	0.156	−0.002	0.004		
Ostrya virginiana	Intercept	1.000	0.066	0.121	0.057	0.094
	HabitatTypeATD	0.527	0.100	0.112	**0.174**	**0.079**
	HabitatTypeATD-Hp	0.292	−0.039	0.071	−0.089	0.068
	HabitatTypeATM	0.464	0.111	0.141	**0.239**	**0.110**
	HabitatTypeTMC	0.439	−0.107	0.142	**−0.244**	**0.112**
	DeerDensity	0.317	0.001	0.002	**0.005**	**0.002**
	CO	0.224	−0.002	0.004	−0.005	0.004
	OstryaBA	**1.000**	**0.369**	**0.081**	**0.400**	**0.079**
	AcerBA	0.132	0.001	0.003		
Empty	Intercept	**1.000**	**0.321**	**0.123**	**0.304**	**0.117**
	HabitatTypeATD	**0.987**	**−0.317**	**0.102**	**−0.341**	**0.102**
	HabitatTypeATD-Hp	0.101	0.011	0.043	0.100	0.086
	HabitatTypeATM	**0.983**	**−0.408**	**0.138**	**−0.437**	**0.136**
	HabitatTypeTMC	0.406	−0.115	0.163	*−0.271*	*0.141*
	DeerDensity	0.059	0.000	0.001	0.000	0.003
	CO	0.801	0.012	0.008	**0.016**	**0.006**
	OstryaBA	0.221	−0.036	0.081	*−0.181*	*0.099*
	AcerBA	0.042	0.000	0.002		
Other species	Intercept	**1.000**	**0.216**	**0.037**	**0.177**	**0.086**
	HabitatTypeATD	0.092	0.006	0.027	0.086	0.070
	HabitatTypeATD-Hp	0.040	0.001	0.011	0.053	0.058
	HabitatTypeATM	0.053	−0.004	0.026	−0.030	0.095
	HabitatTypeTMC	0.240	0.040	0.083	*0.196*	*0.103*
	DeerDensity	0.038	0.000	0.000	−0.001	0.002
	CO	0.041	0.000	0.001	0.001	0.004
	OstryaBA	**1.000**	**−0.223**	**0.061**	**−0.216**	**0.069**
	AcerBA	0.039	0.000	0.001		

Table 5

Performance measures for composition models. 'Best' is the best model selected using multi-model inference, 'Averaged' is the Bayesian averaged model, 'Null' is the model that assumes the empirical mean of regeneration proportions, and 'Best-CV' is the performance of the best model for k-fold cross validation with $k = 10$. Values are means across 1000 replicates (values in brackets are maxima). See text (Section 2.3) for descriptions of model performance measures ($Rank_X$ are proportion of gaps in which each gap-space state was ranked correctly, where SM = sugar maple, IW = ironwood, ET = empty and OT = other).

Model	Tau	Perfect ranking	$Rank_{SM}$	$Rank_{IW}$	$Rank_{ET}$	$Rank_{OT}$
Best	0.228 (0.352)	0.161 (0.247)	0.406 (0.566)	0.451 (0.542)	0.500 (0.584)	0.453 (0.566)
Averaged	0.177 (0.297)	0.139 (0.253)	0.394 (0.554)	0.437 (0.560)	0.476 (0.578)	0.420 (0.536)
Null	0.040 (0.183)	0.109 (0.181)	0.357 (0.488)	0.391 (0.512)	0.406 (0.512)	0.408 (0.506)
Best-CV	0.181 (0.516)	0.143 (0.247)	0.394 (0.727)	0.436 (0.773)	0.478 (0.773)	0.436 (0.733)

gaps respectively. Cross validation results indicate that the 'best' model also out-performs the null model when predicting gaps not used for model fitting (Table 5, 'Best-CV').

3.3. Scenario analysis

Scenario analysis results show that greatest differences in simulated regeneration abundance and composition are due to the habitat type variable (*HabitatType*, Fig. 1a). For example, we observe 5.4% gap spaces empty for *ATM* compared with 53.6% for *ATD-Hp*. The deer density variable (*DeerDensity*) does little to influence simulated regeneration abundance (i.e., ET), but does drive shifts in composition between sugar maple and ironwood. Our model predicts that as deer density increases, SM decreases and IW increases (Fig. 1b). When deer density is very high (~50 deer km^{-2}) in all our gaps, the percentage of gap-spaces in the SM state drops to 11.1%, compared with 25.8% using empirical mean deer densities for each gap and 32.4% for very low densities (~4 deer km^{-2}) in all gaps (Fig. 1b). The canopy openness variable (*CO*) has a greater influence on regeneration abundance than composition, with increases in *CO* increasing ET and decreasing both SM and IW (Fig. 1c). The variable for basal area of mature ironwood (*OstryaBA*) has little influence on SM, but increases IW at the expense of ET and OT as it increases (Fig. 1d). Although there is little difference in proportions
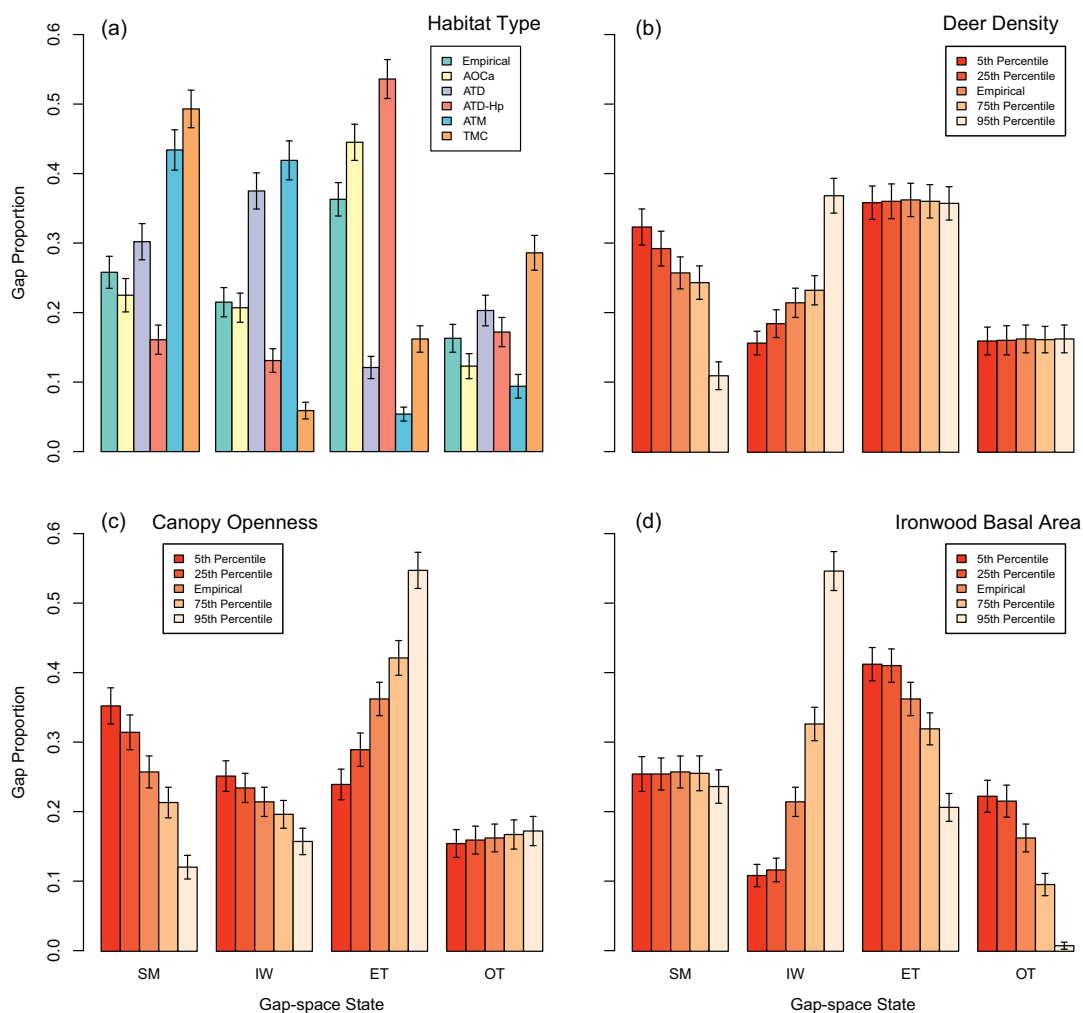

Fig. 1. Scenario analysis results. (a) Habitat type, (b) deer density, (c) canopy openness, (d) ironwood basal area. Percentile values used for environmental variables are presented in Table 1. Results are mean gap proportions of each gap-space state for all gaps averaged over 1000 simulation replicates. Error bars provide 95% confidence (i.e., 1.96 × standard deviation of mean gap proportions from 1000 simulations).

of gaps in the IW state between 5th and 25th percentiles (10.9% and 11.7%, respectively) as the difference between these percentile values is low (Table 2), differences for the 75th and 95th percentiles are much larger (32.7% and 54.7%, respectively) as ironwood basal area increases (to greater than 1 m^2 ha^{-1}, Table 2).

4. Discussion

4.1. Predictors of northern hardwood regeneration

Our model results show that variation in forest habitat type (a proxy for soil moisture and nutrients), deer density, canopy openness and basal area of seed-bearing ironwood trees are important predictors of regeneration composition and abundance in harvest gaps of northern hardwood forest stands. This is demonstrated by the selection of these variables via multi-model inference (Table 3), the credible intervals of parameters in the best model identified (Table 4), and the results of our scenario analyses (Fig. 1). Variables for habitat type (HabitatType) and ironwood basal area (OstryaBA) are consistently selected for inclusion in the model according to the Bayesian Information Criterion (BIC), and variables for canopy openness (CO) and deer density (DeerDensity) are present in four of the best eight models (Table 3). Given that deer density is only selected once in the best two models and that $\Delta BIC < 2$ for these two models, it could be argued that this variable should not be included in the best model. We argue that its inclusion is justified here for several reasons. First, deer are included in the model with greatest probability (0.48) of being the most appropriate model (Table 3); second, deer are a key factor of management interest in northern hardwood (and other) forests and it is therefore useful to include a representative variable in the model to explore impacts of management alternatives; and third, our scenario analyses (discussed below) indicate that deer density is important as a predictor of regeneration composition (although not abundance). The limited predictive power of AcerBA, which was not included in our best model, may be due to limited variation in sugar maple basal area between stands in our empirical data (as this factor is controlled by timber management). However, large variability in regeneration at other tree-height classes across central Upper Michigan despite this limited variation in AcerBA (e.g., Matonis et al., 2011 found that densities of sugar maple 1–2 m tall saplings in forest gaps ranged from 0 to 260 per 100 m^2) implies that other factors are influencing regeneration densities.

Our model shows that sugar maple and ironwood respond oppositely to variation in environmental predictors. Ironwood appears to gain a competitive advantage over sugar maple at high deer density. Furthermore, the proportion of ironwood regeneration in a gap is greater on habitat types where sugar maple comprises a smaller proportion. This might be driven by variation in deer densities across habitat types (see Matonis et al., 2011) or differences between habitat types in the suitability for ironwood vs. sugar maple. However, with the exception of the TMC habitat type, relationships between habitat type and sugar maple and ironwood are consistent but opposite to those for empty gap-spaces (Table 4). This pattern, with the fact that sugar maple and ironwood compose the vast majority of regenerating trees, highlights the importance of habitat type as a predictor of overall regeneration abundance.

The only counter-intuitive finding from our analysis is the negative impact that canopy openness has on the abundance of sugar maple and ironwood saplings. It would likely be expected that greater canopy openness would result in greater proportions of sugar maple and ironwood and lower proportions of empty gap spaces, as greater canopy openness implies greater light availability for the growth of regenerating trees. One reason for our counter-intuitive finding may be that tree density declines with increasing gap size due to increasing seed limitation (i.e., increasing distance to seed sources), increasing competing vegetation, and/or greater mortality rates in the hotter/drier conditions of larger gaps. Alternatively, declining sapling density with increasing gap size may be an artefact of how we have analyzed our data to structure the model. The vast majority of 2–7 m tall trees we sampled (e.g., 95% of sugar maple) were 'advanced regeneration' (i.e., present in the forest stand prior to creation of the gap by harvest). Assuming advanced regeneration grew larger and faster in larger gaps (due to greater light availability and hence greater canopy openness) and that self-thinning occurred amongst advanced regeneration in those gaps, then empty spaces (i.e., ET) would be the only gap-space state that would increase in proportion. That is, competitive exclusion (i.e., mortality) of smaller trees no longer present in the gap when we collected our data results in greater numbers of empty gap spaces (i.e., higher proportion of ET). Although this self-thinning effect in larger gaps may produce counter-intuitive results at the gap-space level (given how we have structured our regeneration model), our previous results indicate that at the stand level and through time the number of regenerating trees reaching maturity in the canopy produces appropriate stand basal areas (Millington et al., 2011).

4.2. Prediction of regeneration at the gap level

The predictive performance of our best model is better than that of a null model using all measures we considered (Table 5). However, in absolute terms our model performs seemingly poorly. For example, it is able to rank gap-space states perfectly on average only 16% of the time. The difficulties of predicting forest regeneration – in part due to difficulties associated with measuring processes that influence germination, growth, damage and mortality of trees – have been previously acknowledged (e.g., Miina and Heinonen, 2008). We suggest two reasons to explain our model performance, the first related to the system we are studying and the second to the scale and level of aggregation at which we are considering it. First, as highlighted above, regeneration densities in northern hardwood stand harvest gaps across the central Upper Peninsula of Michigan are highly variable, despite limited variation in seed availability (i.e., AcerBA controlled by timber management). Although our model accounts for several other factors likely influencing regeneration densities, it does not account for potential spatial variation in regeneration within gaps due to factors such as light availability or soil moisture (although with our gap-centred plots we believe effects would be limited). Furthermore, our model does not account for competing, non-tree, ground vegetation which has been found to be important for determining tree regeneration (see Section 2.2). These missing predictors require data which is usually unavailable to forest managers and cannot be provided as variables by the growth and yield model to which we couple this regeneration model. Second, our model testing is at the gap-level. Regeneration components of many models representing forest dynamics are tested using only stand-level aggregate measures (e.g., total stand basal area, total stand stem density) simulated through time (e.g., Seagle and Liang, 2001; Larocque et al., 2011; Holm et al., 2012). We are unable to make similar comparison with our model, both because we do not have historical data with which to make such a comparison and because such stand-level aggregate measures in our study system are strongly driven by forest management (i.e., timber harvest prescriptions). Until such assessments are made, use of our model outside the limits of our empirical data may not be possible (unless re-parameterized for the area of interest).

However, as we highlighted at the outset, the development of our regeneration model was pursued with the intention of incorporating it into an integrated forest tree regeneration, growth and harvest model to examine impacts of alternative timber and
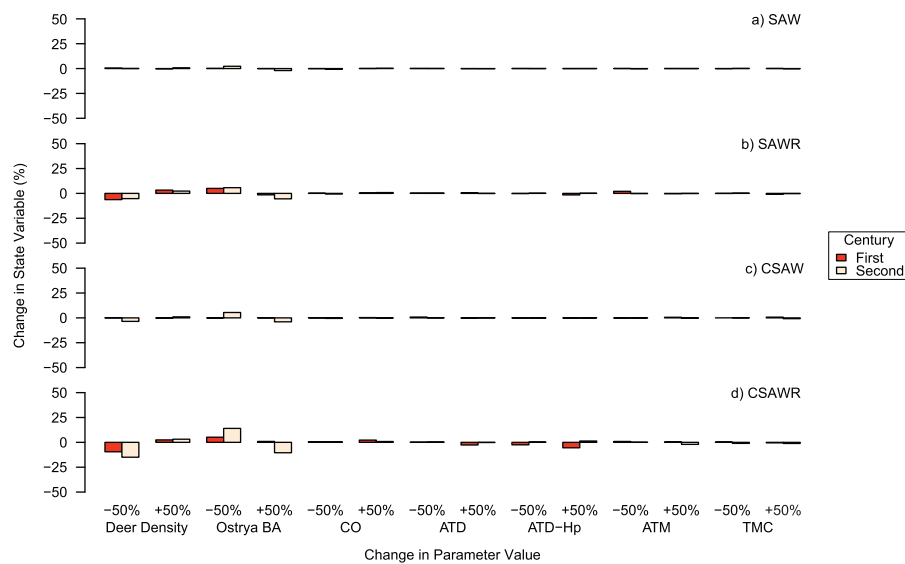
Fig. 2. Sensitivity analysis results from using the regeneration model in an integrated forest regeneration, growth and harvest simulation model. (a) mean stand merchantable standing timber, (b) mean stand merchantable standing timber from regenerated trees, (c) mean stand merchantable harvested timber, and (d) mean stand merchantable harvested timber from regenerated trees. The effect of varying each regeneration model parameter estimate ±50% in turn on state variables is shown for the first and second simulated centuries.

wildlife (deer) management actions over long time periods. A full description of how our regeneration model is used in that integrated simulation model can be found in Millington et al. (2013). When our regeneration model is used in the integrated model, we find that stand-level aggregated measures such as standing merchantable timber vary through time at expected levels (Figure 1 of Millington et al., 2013). To demonstrate how regeneration influences the stand-level aggregate measures via the integrated model, we present the results of a sensitivity analysis which varies each parameter value in the regeneration model (i.e., those in Table 4) in turn by +50% and −50%. The integrated model was run for 200 years from initial empirical conditions (see Millington et al., 2013) with operationally realistic harvest prescriptions rules as described by Millington et al. (2011, Prescription B in Table 1). The response of the integrated model state variables (Fig. 2) show that for stand-level aggregates, measures of standing and cut merchantable timber from all and regeneration-only trees are most sensitive to *OstryaBA* and *DeerDensity*. However, all variables are insensitive to changes in parameter estimates (i.e., for 50% increase/decrease in parameter estimates, stand-level aggregate measures change <50%). Stand-level measures of timber that account only for simulated regeneration (standing and cut) are most sensitive, but change only 5–15%. This insensitivity in parameter estimates suggests that the regeneration model is sufficiently robust for use in the integrated model for simulation of forest management dynamics over long time periods.

Future development and improvement of the regeneration model should focus in part on improving parameter estimates to reduce the range of credible intervals. This improvement could come through better quantification of the factors influencing regeneration. For example, our estimates of deer density used to produce parameter estimates in the regeneration sub-model come from deer pellet counts (Millington et al., 2010)

which themselves provide only a proxy for deer density. Considering other potential environmental factors (such as snow fall which can control soil moisture availability during spring-time, Henne et al., 2007) should also help to improve the regeneration sub-model.

5. Conclusions

We have presented the structure, parameterization, testing and use of a stochastic, regression-based compositional forest gap regeneration model developed with the express purpose of being integrated with a forest growth and yield model (FVS). The innovative structure of the model represents only those trees regenerating in gaps created by timber harvest with the best chance of subsequently growing into the canopy. A multi-model inference approach identified that habitat type, deer density, canopy openness and basal area of mature ironwood in the vicinity of a gap drive regeneration abundance and composition. The best model of those examined shows that where deer densities are high, ironwood appears to gain a competitive advantage over sugar maple and that habitat type is an important predictor of overall regeneration success. In its current state the regeneration model is sufficiently robust for use with FVS to simulate forest dynamics over long time periods (as explored in Millington et al., 2013).

Appendix A.

Description of northern hardwood stands sampled in this study ($n = 28$) in the central Upper Peninsula of Michigan, USA.

Stand ID	UTM Easting UTM Northing	Habitat type[a]	Deer density[b]	Canopy openness[c]	Ironwood basal area[d]	Sugar maple basal area[e]	Relative density[f,g]	Year of harvest
803	461,407 5,088,766	ATD-Hp	12.4	11.2 (11.2)	0.25	22.34	SM 97% BP 2% YB 1%	2000
806	449,656 5,100,197	AOCa	10.9	12.6 (6.7)	0.03	21.16	SM 97% BA 3% IW < 1%	1999
808	464,248 5,082,430	ATD-Hp	13.9	14.1 (14.0)	0.18	5.09	BA 54% SM 22% WA 20%	2000
814	443,446 5,101,948	AOCa	7.8	21.4 (14.1)	0.10	22.18	SM 79% BA 20% YB 1%	1999
819	452,458 5,100,012	AOCa	27.0	16.9 (9.0)	0.00	25.73	SM 100%	1999
824	433,048 5,170,449	TMC	3.2	19.3 (17.3)	0.62	19.67	SM 78% YB 13% RM 4%	1999
826	475,013 5,139,239	ATM	12.2	14.6 (19.3)	0.01	8.68	SM 97% RM 3%	1999
829	475,972 5,130,553	ATD	5.3	8.1 (5.0)	0.00	16.85	SM 85% BC 6% RM 6%	1998
837	405,133 5,131,185	AOCa	21.5	14.5 (14.8)	0.04	18.06	SM 69% RM 17% AE 8%	1998
842	478,993 5,126,810	ATD	15.3	6.2 (9.5)	0.00	24.43	SM 78% RM 20% HM 1%	1997
851	437,221 5,111,274	ATD	15.4	12.3 (15.3)	0.20	13.24	BA 65% SM 24% QA 9%	2000
1042	401,125 5,117,269	TMC	61.6	13.5 (14.6)	0.00	10.78	SM 36% RM 36% HM 22%	1996
1043	401,029 5,117,607	AOCa	16.6	9.1 (8.5)	0.03	21.14	SM 94% RM 3% NC 2%?	1999
1081	413,026 5,135,891	ATM	8.0	7.6 (6.3)	0.20	18.00	SM 77% RM 18% BA 5%	1996
1133	436,239 5,090,816	AOCa	17.2	13.5 (11.4)	0.69	11.63	SM 78% BA 12% YB 8%	1999
1272'	446,459 5,085,661	ATD-Hp	27.6	9.0 (12.3)	0.52	16.03	SM 59% BA 33% WA 8%	1998
4024	452,336 5,074,499	ATD-Hp	15.5	18.3 (15.5)	0.82	16.08	SM 79% HM 11% BA 9%	2000
4042	457,134 5,078,785	AOCa	4.4	10.1 (13.7)	0.20	17.61	SM 84% BA 11% NC 4%	1999
4102	446,405 5,099,004	AOCa	50.4	7.7 (13.5)	1.31	17.47	SM 97% BA 2% WA 1%	1996
4332	454,413 5,126,758	ATM	14.0	24.6 (10.4)	0.00	24.79	SM 94% BC 4% YB 2%	1998
4342	458,611 5,124,129	ATD	7.5	18.0 (10.8)	0.01	25.95	SM 98% RM 1% YB 1%	1998
4422	458,436 5,073,974	ATD-Hp	9.7	11.7 (20.3)	0.12	11.21	SM 84% HM 8% AB 6%	2000
4423	457,874 5,073,803	ATD-Hp	21.0	9.7 (5.1)	0.03	20.25	SM 95% WA 4% HM < 1%	1997
4429	458,262 5,074,127	AOCa	13.7	15.3 (23.0)	0.68	13.20	SM 84% HM 5% AB 4%	1999

Stand ID	UTM Easting UTM Northing	Habitat type[a]	Deer density[b]	Canopy openness[c]	Ironwood basal area[d]	Sugar maple basal area[e]	Relative density[f,g]	Year of harvest
4524	434,902 5,104,419	ATD-Hp	18.1	9.4 (10.95)	0.88	17.75	SM 83% BA 15% RM 1%	1998
5032	437,620 5,082,189	ATD-Hp	12.8	8.9 (16.09)	1.28	20.90	SM 88% BA 11% IW 1%	1997
5104	460,065 5,115,757	ATD	20.1	10.4 (7.7)	0.07	26.42	SM 97% RM 3% IW < 1%	1998
5292	407,479 5,121,325	AOCa	21.3	11.1 (12.5)	0.14	20.58	SM 95% RM 4% BA 1%	1997

[a] Habitat type is a proxy for soil moisture and nutrient regimes (see Burger and Kotar, 2003): AOCa (mesic soil moisture regime; rich to very rich soil nutrient regime), ATD-Hp (mesic; medium to rich), ATD (mesic; medium to rich), ATM (dry-mesic to mesic; medium), and TMC (mesic to wet-mesic; medium).

[b] Winter deer density (km^{-2}) was estimated for using spring-time faecal pellet surveys (see Matonis et al., 2011).

[c] Canopy openness (%) was estimated from hemispherical photographs taken in each gap. Mean gap value is shown with range in brackets.

[d] Basal area (m^2 ha^{-1}) of mature (> 10 cm diameter at breast height, DBH, breast height = 1.4 m) ironwood (Ostrya virginiana) trees.

[e] Basal area (m^2 ha^{-1}) of mature (> 20 cm DBH) sugar maple (Acer saccharum) trees.

[f] Relative abundance is provided for the three most abundant overstory tree (DBH > 20 cm) species, calculated using basal area.

[g] AB = American beech (Fagus grandifolia Ehrh.), AE = American elm (Ulmus americana L.), BA = American basswood (Tilia americana L.), BC = black cherry (Prunus serotina Ehrh.), BP = balsam poplar (Populus balsamifera L.), HM = eastern hemlock (Tsuga canadensis (L.) Carrière), IW = ironwood (Ostrya virginiana (Mill.) K. Koch), NC = northern white-cedar (Thuja occidentalis L.), QA = quaking aspen (Populus tremuloides Michx.), RM = red maple (Acer rubrum L.), SM = sugar maple (Acer saccharum Marsh.), WA = white ash (Fraxinus americana L.), YB = yellow birch (Betula alleghaniensis Britton).

Appendix B. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ecolmodel.2012.12.033.

References

Aitchison, J., 1982. The statistical analysis of compositional data. Journal of the Royal Statistical Society: Series B 44 (2), 139–177.

Aitchison, J., 1986. The Statistical Analysis of Compositional Data. Monographs on Statistics and Applied Probability. Chapman & Hall Ltd., London, UK.

Aitchison, J., 2001. Simplicial inference. In: Viana, M.A.G., Richards, D.S.P. (Eds.), Algebraic Methods in Statistics and Probability. Contemporary Mathematics Series, vol. 287. American Mathematical Society, Providence, RI, pp. 1–22.

Aitchison, J., Egozcue, J.J., 2005. Compositional data analysis: where are we and where should we be heading? Mathematical Geology 35 (7), 829–850.

Arbogast Jr., C., 1957. Marking Guides for Northern Hardwoods Under the Selection System. USDA Forest Service Station Paper LS-SP-56, St. Paul, MN, USA.

Botkin, D.B., Janak, J.F., Wallis, J.R., 1972. Some ecological consequences of a computer model of forest growth. Journal of Ecology 60 (3), 849–872.

Bugmann, H.K.M., Yan, X., Sykes, M.T., Martin, P., Lindner, M., Desanker, P.V., Cumming, S.G., 1996. A comparison of forest gap models: model structure and behaviour. Climatic Change 34, 289–313.

Burger, T.L., Kotar, J., 2003. A Guide to Forest Communities and Habitat Types of Michigan. The Department of Forest Ecology and Management, University of Wisconsin-Madison, WI, USA.

Canham, C.D., Denslow, J.S., Platt, W.J., Runkle, J.R., Spies, T.A., White, P.S., 1990. Light regimes beneath closed canopies and tree-fall gaps in temperate and tropical forests. Canadian Journal of Forest Research 20, 620–631.

Cote, S.D., Rooney, T.P., Tremblay, J.P., Dussault, C., Waller, D.M., 2004. Ecological impacts of deer overabundance. Annual Review of Ecology, Evolution and Systematics 35, 113–147.

Crookston, N.L., Dixon, G.E., 2005. The forest vegetation simulator: a review of its structure, content, and applications. Computers and Electronics in Agriculture 49, 60–80.

Donovan, G., 2005. Chronic regeneration failure in northern hardwood stands: a liability to certified forest landowners. In: Cook, B. (Ed.), Proceedings of the Forests & Wildlife-Striving for Balance Michigan Society of American Foresters Conference. St. Ignace, MI, pp. 125–130.

Ferguson, D.E., Carlson, C.E., 1993. Predicting regeneration establishment with the Prognosis Model. Res. Paper INT-467. U.S. Department of Agriculture, Forest Service, Intermountain Research Station, Ogden, UT.

Garrett, P.W., Graber, R.E., 1995. Sugar Maple Seed Production in Northern New Hampshire. Northeastern Forest Experiment Station Research Paper NE-697.

Hann, D.W., 2011. ORGANON User's Manual Edition 9.0. Oregon State University, Department of Forest Engineering, Resources, and Management, Corvallis, OR.

Hasenauer, H., Kindermann, G., 2002. Methods for assessing regeneration establishment and height growth in uneven-aged mixed species stands. Forestry 75, 385–394.

Hassani, B.T., LeMay, V., Marshall, P.L., Temesgen, H., Zumrawi, A.-A., 2004. Regeneration imputation models for complex stands of southeastern British Columbia. Forestry Chronicle 80, 271–278.

Hastie, T., Tibshirani, R., Friedman, J.H., 2001. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer, New York, NY.

Henne, P.D., Hu, F.S., Cleland, D.T., 2007. Lake-effect snow as the dominant control of mesic-forest distribution in Michigan, USA. Journal of Ecology 95, 517–529.

Hill, H.R., 2001. Deer Pellet Group Surveys. Michigan Department of Natural Resources Wildlife Division Report No. 3349, Lansing, MI, USA.

Hoeting, J.A., Madigan, D., Raftery, A.E., Volinsky, C.T., 1999. Bayesian model averaging: a tutorial. Statistical Science 14, 382–401.

Holm, J.A., Shugart, H.H., Van Bloem, S.J., Larocque, G.R., 2012. Gap model development, validation, and application to succession of secondary subtropical dry forests of Puerto Rico. Ecological Modelling 233, 70–82.

Horsley, S.B., Stout, S.L., deCalesta, D.S., 2003. White-tailed deer impact on the vegetation dynamics of a northern hardwood forest. Ecological Applications 13, 98–118.

Kern, C., 2011 The role of harvest gaps in the plant diversity of a northern hardwood forest of northern Wisconsin, USA. PhD Thesis. University of Minnesota.

Kobe, R.K., Hogarth, L., 2007. Evaluation of irradiance metrics with respect to predicting sapling growth. Canadian Journal of Forest Research 37, 1203–1213.

Lacerte, V., Larocque, G.R., Woods, M., Parton, W.J., Penner, M., 2006. Calibration of the forest vegetation simulator (FVS) model for the main forest species of Ontario, Canada. Ecological Modelling 199, 336–349.

Larocque, G.R., Archambault, L., Delisle, C., 2011. Development of the gap model ZELIG-CFS to predict the dynamics of North American mixed forest types with complex structures. Ecological Modelling 222, 2570–2583.

Link, W.A., Barker, R.J., 2006. Model weights and the foundations of multimodel inference. Ecology 87, 2626–2635.

Liu, J., Aston, P.S., 1995. Individual-based simulation models for forest succession and management. Forest Ecology and Management 73, 157–175.

Liu, J., Aston, P.S., 1999. Simulating effects of landscape context and timber harvest on tree species diversity. Ecological Applications 9, 186–201.

Long, Z.T., Pendergast, T.H., Carson, W.P., 2007. The impact of deer on relationships between tree growth and mortality in an old-growth beech-maple forest. Forest Ecology and Management 252, 230–238.

Matonis, M.S., 2009. Multiple gap-, stand-, and landscape-scale factors affect regeneration in managed northern hardwood forests. M.S. Thesis. Michigan State University, East Lansing, MI, USA.

Matonis, M.S., Walters, M.B., Millington, J.D.A., 2011. Gap-, stand-, and landscape-scale factors contribute to poor sugar maple regeneration after timber harvest. Forest Ecology and Management 262 (2), 286–298.

Ménard, A., Dubé, P., Bouchard, A., Canham, C.D., Marceau, D.J., 2002. Evaluating the potential of the SORTIE forest succession model for spatio-temporal analysis of small-scale disturbances. Ecological Modelling 153, 81–96.

Miina, J., Saksa, T., 2006. Predicting regeneration establishment in Norway spruce plantations using a multivariate multilevel model. New Forests 32, 265–283.

Miina, J., Heinonen, J., 2008. Stochastic simulation of forest regeneration establishment using a multilevel multivariate model. Forest Science 54, 206–219.

Millington, J.D.A., Walters, M.B., Matonis, M.S., Laurent, E.J., Hall, K.R., Liu, J., 2011. Combined long-term effects of variable tree regeneration and timber management on forest songbirds and timber production. Forest Ecology and Management 262, 718–729.

Millington, J.D.A., Walters, M.B., Matonis, M.S., Liu, J., 2013. Modelling for forest management synergies and trade-offs: northern hardwood tree regeneration, timber and deer. Ecological Modelling 248, 103–112.

Millington, J.D.A., Walters, M.B., Matonis, M.S., Liu, J., 2010. Effects of local and regional landscape characteristics on wildlife distribution across managed forests. Forest Ecology and Management 259, 1102–1110.

Millington, J.D.A., Perry, G.L.W., 2011. Multi-model inference in biogeography. Geography Compass 5 (7), 448–463.

Monserud, R.A., Ek, A.R., 1977. Prediction of understory tree height growth in northern hardwood stands. Forest Science 23, 391–400.

Nyland, R.D., 1998. Selection system in northern hardwoods. Journal of Forestry 96, 18–21.

Pedersen, B.S., Wallis, A.M., 2004. Effects of white-tailed deer herbivory on forest gap dynamics in a wildlife preserve, Pennsylvania, USA. Natural Areas Journal 24, 82–94.

Powers, M.D., Nagel, L.M., 2009. Pennsylvania sedge cover, forest management and deer density influence tree regeneration dynamics in a northern hardwood forest. Forestry 82 (3), 241–254.

R Development Core Team (2009). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/

Racevskis, L.A., Lupi, F., 2006. Comparing urban and rural perceptions of and familiarity with the management of forest ecosystems. Society and Natural Resources 19, 479–495.

Randall, J.A., 2007. Deer and sedge effects on tree seedling dynamics in northern temperate forests. Dissertation. Department of Forestry, Michigan State University, East Lansing, Michigan, USA.

Ribbens, E., Silander, J.A., Pacala, S.W., 1994. Seedling recruitment in forests: calibrating models to predict patterns of tree seedling dispersion. Ecology 75, 1794–1806.

Runkle, J.R., 1981. Gap regeneration in some old-growth forests of the eastern United States. Ecology 62, 1041–1051.

Sage, R.W., Porter, W.F., Underwood, H.B., 2003. Windows of opportunity: white-tailed deer and the dynamics of northern hardwood forests of the northeastern US. Journal for Nature Conservation 10, 213–220.

Schwarz, G., 1978. Estimating the dimension of a model. The Annals of Statistics 6, 461–464.

Seagle, S.W., Liang, S., 2001. Application of a forest gap model for prediction of browsing effects on riparian forest succession. Ecological Modelling 144, 213–229.

Sharma, M., Parton, J., Woods, M., Newton, P., Penner, M., Wang, J., Stinson, A., Bell, F.W., 2008. Ontario's forest growth and yield modelling program: advances resulting from the Forestry Research Partnership. Forestry Chronicle 84, 694–703.

Shugart, H.H., 1984. A Theory of Forest Dynamics. Springer-Verlag, New York.

Spiegelhalter, D.J., Thomas, A., Best, N.G., Gilks, W.R., Lunn, D., 2003. BUGS: Bayesian Inference Using Gibbs Sampling. MRC Biostatistics Unit, Cambridge, England http://www.mrc-bsu.cam.ac.uk/bugs (accessed 29.11.11).

Sturtz, S., Ligges, U., Gelman, A., 2005. R2WinBUGS: a package for running WinBUGS from R. Journal of Statistical Software 12, 1–16.

Tolosana-Delgado, R., Otero, N., Pawlowsky-Glahn, V., 2005. Some basic concepts of compositional geometry. Mathematical Geology 37 (7), 673–680.

Urban, D.L., 1990. A Versatile Model to Simulate Forest Pattern: A User's Guide to ZELIG Version 1.0. University of Virginia, Charlottesville, Virginia.

van den Boogaart, K.G., Tolosana-Delgado, R., 2008. compositions: a unified R package to analyze compositional data. Computers & Geosciences 34 (4), 320–338.

Vickers, L.A., Fox, T.R., Loftis, D.L., Boucugnani, D.A., 2011. Predicting forest regeneration in the Central Appalachians using the REGEN expert system. Journal of Sustainable Forestry 30 (8), 790–822.

Webster, C.R., Lorimer, C.G., 2005. Minimum opening sizes for canopy recruitment of midtolerant tree species: a retrospective approach. Ecological Applications 15 (4), 1245–1262.

Weiskittel, A.R., Hann, D.W., Kershaw, J.A., Vanclay, J.K., 2011. Forest Growth and Yield Modeling. Wiley-Blackwell, Chichester.