Predicting oak regeneration success at the stem exclusion stage of stand development in upland hardwood forests

Lake E. Graboski[a,b], Kim C. Steiner[a,*], Marc E. McDill[a], James C. Finley[a]

[a] *Department of Ecosystem Science and Management, The Pennsylvania State University, University Park, PA 16802, United States*
[b] *Now with Mountain Research LLC, Altoona, PA 16601, United States*

ABSTRACT

Oaks (*Quercus* spp.) are becoming less abundant in most of the Central Hardwood Region of the eastern United States, and this is creating shifts in forest composition that will likely have important economic and ecological consequences. In large measure, these changes originate with deficiencies in the oak regeneration cohort preceding stand-replacement disturbances such as timber harvest. To sustain the oak resource, managers need better information on the connection between pre-harvest and early stand conditions and regeneration outcomes. In this study, we used direct observations of oak seedling dominance in the stem exclusion stage of stand development (mean age = 17.4 years) to model the probability of successful regeneration during stand initiation as a function of stand conditions before and after harvest (ages −1, 1, 4, and 7 years). For pre-harvest conditions, the most predictive model was based solely on the aggregate height of advance regeneration oak seedlings > 15 cm in sample plots. As expected, post-harvest models were more predictive, and increasingly more predictive with the passage of time, and they were optimized by contrasting the height of the plot-dominant oak seedling with the heights of competing tree species. The predictive power of post-harvest models increased most between ages 1 and 4 years and only slightly between ages 4 and 7, indicating that age 4 is an optimal time to evaluate opportunities to favor oak regeneration with early silvicultural interventions. Of the two most common competitors, black birch (*Betula lenta*) had the more inhibitory effect on the success of oak regeneration when it was present. However, red maple (*Acer rubrum*) was the more important competitor because of its very high frequency of occurrence in plots occupied by oak seedlings.

1. Introduction

The abundance of oak (*Quercus* spp.) has significantly declined in recent decades throughout most of eastern North America but especially in the Central Hardwood Region (CHR), where oaks have been historically a dominant and nearly ubiquitous forest component (Clark, 1993; McWilliams et al., 2002; Knott et al., 2019). Many factors are likely involved because the declines are not occurring equally everywhere and in all species (Fei et al., 2011), but it is believed that the regional trend is largely a result of changes in disturbance regimes that favor the regeneration of other species (Abrams, 1992; Nowacki and Abrams, 2008). This trend is more pronounced on mesic sites, where species such as red maple (*Acer rubrum* L.), black birch (*Betula lenta* L.), yellow-poplar (*Liriodendron tulipifera* L.), and black cherry (*Prunus serotina* Ehrh.) often replace oaks following harvest (McGee and Hooper, 1970; Beck and Hooper, 1986; Gould et al., 2005; Fei and Steiner, 2009; Swaim et al., 2016). Despite decades of research focused on improving

oak regeneration under contemporary conditions (Clark, 1993), significant reductions in oak importance values continue to occur between successive forest inventories (Fei et al., 2011; Knott et al., 2019). These trends are significant because, compared with most other genera within the region, oaks are disproportionately valuable both culturally and economically, and they are ecologically "keystone" in the sense that they have a disproportional effect on other species including birds, mammals, insects, and other plants (Fralish, 2004; Tallamy and Shropshire, 2009). Given the importance of oaks and their declining abundance, promoting better oak regeneration is one of the most significant silvicultural challenges in eastern North America.

Although stump sprouts may routinely contribute as much as 75% of future oak stocking in regenerated stands (McIntyre, 1936; Gould et al., 2003), seedlings are a necessary complement to stump sprouts where healthy and successful oak regeneration is desired. Early research demonstrated that successful oak seedling regeneration is usually dependent upon the density and size of oak advance

regeneration, i.e., seedlings present when a harvest or other stand-replacing disturbance occurs (Trimble and Hart, 1961; Sander and Clark, 1971). Further research improved our understanding of these relationships and led to the development of oak seedling regeneration models for specific regions including the Missouri Ozarks (Sander et al., 1984) and the Appalachian Mountains of North Carolina (Loftis, 1990a) and Pennsylvania (Gould et al., 2006; Steiner et al., 2008).

To our knowledge, all models of oak regeneration in eastern North America are based on pre-harvest measurements of oak seedlings coupled with subsequent measurements at ≤8 years after harvest and assumed relationships between these measurements and success as dominant or co-dominant stems at approximately age 20 (Sander et al., 1984; Loftis, 1990a; Gould et al., 2006). Interest is drawn to success at age 20 because by this age most well-stocked upland oak stands in this region are in the stem exclusion (SE) stage of stand development, i.e., the culmination of stand initiation. It is a time when intense competition prevents the establishment of new stems, less competitive trees begin to die from insufficient sunlight, and trees that remain form the basis for the future forest. Previous research has shown that oaks in competitive crown positions at SE are expected to continue to dominate their immediate competition and are much more likely to survive than trees in lower crown classes (Ward and Stephens, 1994; Johnson et al., 2009).

In contrast to SE, the regenerative process of stand initiation is one of rapid changes in species composition and stand structure (Oliver and Larson, 1996). The developmental uncertainties of the full period of stand initiation in oak stands (Beck and Hooper, 1986; Gould et al., 2005; Swaim et al., 2016), ending with stem exclusion, have yet to be modeled with empirical data. Here we present a study of 23 stands and their replacements in the central Appalachians (Fig. 1), from mature, oak-dominated forests with advance oak regeneration, through various

ages after harvest, and culminating with the stem exclusion stage of their replacement stands. The results have relevance to pre-harvest expectations of oak regeneration success and to decisions about silvicultural interventions early in stand development for the purpose of benefitting oak regeneration.

2. Methods

2.1. Stand conditions

Stands were 12–28 ha in size and located on state-owned forest land managed by the Pennsylvania Bureau of Forestry. Prior to harvest, oaks dominated the overstory of every study stand and contributed a mean percentage of 68% to the total basal area, with a range from 9.7 to 22.7 m^2 ha^{-1} for oak trees ≥5 cm in DBH (diameter at 1.37 m above the ground). Chestnut oak (*Quercus montana* Willd.), northern red oak (*Q. rubra* L.), and white oak (*Q. alba* L.) accounted for 42%, 22%, and 20% of oak basal area respectively, but black oak (*Q. velutina* Lam.) and scarlet oak (*Q. coccinea* Muenchh.) were also present in some stands. Red maple was the most common non-oak species, with an average of 15% of total basal area. The study encompassed much of central Pennsylvania with most stands (20) in the Ridge and Valley physiographic region, where oak-dominated stands are the norm, and three located in the Appalachian Plateau region, where oak-dominated stands are less common (Fig. 1). Across this area, annual precipitation ranges from 999 to 1042 mm, average summer temperatures from 20.1 to 21.5 °C, and average winter temperatures from −1.2 to −2.5 °C (Pennsylvania State Climatologist, 2019). All stands were situated on slopes or ridgetops and typically had stony, well- to moderately well-drained soils derived from shale or sandstone parent material (Soil Survey Staff, 2019). Stand site indices (base age 50 years) varied from 14.6 to 22.9 m. Although we lack a detailed history of each stand, they all arose from regeneration following the end of widespread logging around the turn of the last century (Abrams and Nowacki, 1992).

2.2. Data collection

Measurements began one year before harvest by recording the species and diameter at breast height (DBH) of all overstory trees (DBH > 5 cm) within 200-m^2 permanent "superplots" in a gridded arrangement of 12–30 per stand. Site quality (site index, SI) was determined using the average height and age of 3–4 mature trees in each stand and standardized to northern red oak site index at a base age of 50 years (Carmean et al., 1989). To increase the resolution of site quality determinations, we used Arcmap10 GIS to overlay these plots with an integrated moisture index (IMI) (Iverson et al., 1997; Peters et al., 2013) that functions as an estimate of long-term soil moisture availability.

Trees with DBH < 5 cm (advance regeneration) and other vegetation were measured in four 4-m^2 permanent plots nested within each superplot. These smaller plots are the focus of the regeneration models described below. Each tree seedling was tallied by species and height class (≤5 cm, 6–15 cm, 16–30 cm, 31–61 cm, 62–91 cm, 92–122 cm, and 123–152 cm) and assigned a height at the midpoint of its class. Seedlings taller than 1.52 m were measured to the nearest 0.3 m. The percentage of ground shaded by mountain-laurel (*Kalmia latifolia* L.) (%C Laurel) and by rhizomatous ferns (%C Fern) was visually estimated to the nearest 5% for each plot in 2 strata (ground to 1.52 m and 1.53–6.1 m). These species were the most abundant non-tree vegetation and are known to impede the growth of seedlings in the study area based on previous research (Fei et al., 2010; Campbell et al., 2015). The entire study comprised 518 superplots containing a total of 2044 4-m^2 regeneration plots that were measured on all dates.

Between the years 1996 and 2003 all stands were harvested via a single removal, leaving behind an average 4.6 m^{2}ha^{-1} (range 1.2–10.1 m^2) of residual basal area. Stands were revisited 1, 4, and 7
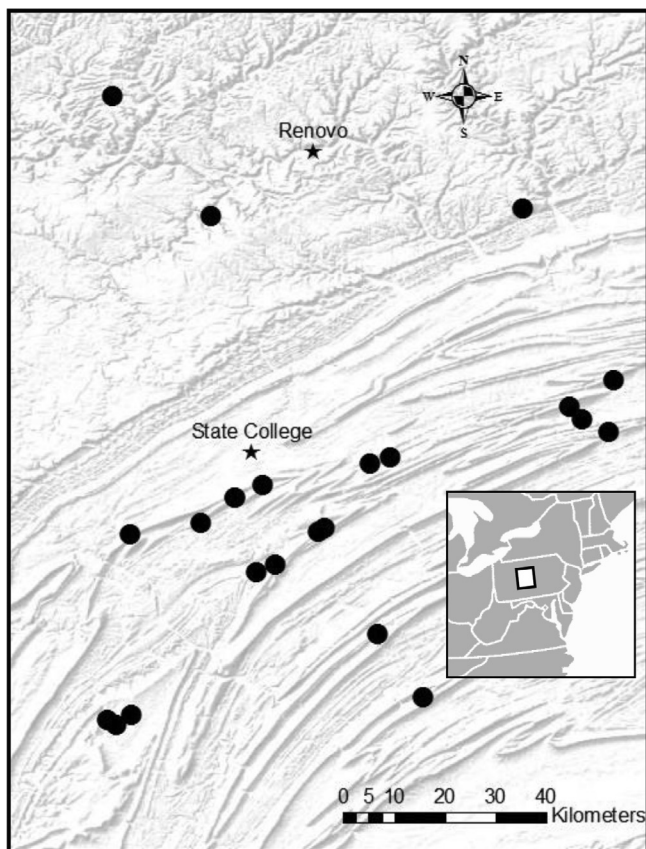

Fig. 1. Study stand locations (black dots): 3 in the Appalachian Plateau physiographic region (near Renovo, Pennsylvania) and 20 in the Ridge and Valley (near State College, Pennsylvania).

growing seasons after harvest to remeasure tree regeneration and percentage cover of competing vegetation within each regeneration plot. In year 7, some size measurements (about 14% of all seedlings) were shifted from height to DBH values. In order to obtain height values for some analyses described below, a height-DBH relationship was established using linear regression based on a data set of 840 observations of seedlings on which both the height and DBH continued to be measured.

Oak regeneration success was measured between the ages of 15 and 20 years after harvest (average 17.4 years), at which time all stands were in the stem exclusion stage of stand development, with regeneration visually estimated to form a closed canopy over at least 75% of the plots in each instance. A combination of crown class designation (Helms, 1998) and relative height was used to determine whether plots were successfully stocked with one or more oak stems of seedling origin. A plot was regarded as successfully stocked if it contained one or more "dominant" or "codominant" oaks of seedling origin, or an "intermediate" class oak seedling whose height was at least 85% as great as the tallest tree within a 2.07-m radius centered on the plot. The last-named criterion amounted to an average height deficit of ≤1.7 m and was based upon stem dissections in similar stands that showed oaks capable of overcoming moderate height deficits at age 15 to become dominant or codominant trees at age 30 (Zenner et al., 2012). Studies of oak success in the Ozark region of the United States have employed an 80% relative height criterion for projecting regeneration success of upland oaks from 5- to 10-year data (Sander et al., 1984; Spetich et al., 2002).

An average of only 63% of growing space was available for regeneration because residual trees from the previous stand occupied 37% of the growing space at SE, based upon calculations using the maximum tree area equation of Gingrich (1967). We used two methods to calculate the expected percentage contribution of oak-stocked plots at SE to full stocking of available growing space. In Method 1 we determined the minimum tree area occupied by all successful oak seedlings at SE based upon their mean DBH and then divided this area by the total available growing space to obtain the percentage contribution of oak seedlings. In Method 2 we estimated the actual percentage of seedling-origin oak stocking in each stand by applying Gingrich's (1967) minimum tree area equation to all stems on 4-m^2 plots > 1.37 m tall, divided this result by the percentage of growing space available for regeneration, and regressed those values against percentage of plots stocked by oak seedlings.

Tree regeneration was aggregated into four species groups: all oak species (Oak), red maple (Maple), black birch (Birch), and all other species (Other). Red maple and black birch were of interest because of their frequency of occurrence (90% and 39% of plots at age 7, respectively) and because of their frequent and abundant presence in regenerating oak stands within this region (Nowacki and Abrams, 1992; Gould et al., 2005; Fei and Steiner, 2007). The remaining species were grouped together as Other for modeling purposes because each was found in ≤16% of all plots and was absent from many stands (Table 2). Dominance in the regeneration cohort was measured either by aggregating the heights of all seedling-origin regeneration in the plot, by species, as a composite measure of size and abundance (AgHt) (Fei et al., 2006) or simply by the height of the single tallest stem of the species or species group within the plot (H-Dom).

2.3. Data analysis

Our objective was to develop models predictive of oak seedling regeneration at SE, using presence or absence of successful oak stocking in 4-m^2 regeneration plots as the response variable and candidate predictor variables as summarized in Table 1. In order to document changes in important predictors and predictive power over time, four separate models were fit, each using only measurements collected at a specified time (pre-harvest, year 1, year 4, or year 7) to predict the response. The probability of success was modeled with generalized linear regression using either a logit link function (Equation (1)) or a negative exponential link function (Eq. (2)), depending on the quality of fit.

Eq. (1): Model utilizing logit link function:

$$\pi = \frac{e^{(\alpha+\beta X+\gamma Z+\varepsilon)}}{(1 + e^{(\alpha+\beta X+\gamma Z+\varepsilon)})} \tag{1}$$

Eq. (2): Model utilizing negative exponential link function:

$$\pi = 1 - e^{-(\alpha+\beta X++\gamma Z)+\varepsilon} \tag{2}$$

where π = probability of oak seedling success, α = study-wide intercept (fixed), β = slope coefficient(s) of the predictor variable(s) (X), γ = stand (Z) intercepts (random), and ε = residual error.

A combination of methods was used to account for the possibility that the model intercept and fixed-effect slopes were affected by a lack of independence between plot observations clustered within superplots and stands. Model structure accommodated clustering of plots within stands by including a random intercept for each stand in the fitted model. However, an insufficiency of data points prevented the estimation of random slope parameters for stands and of random intercept and slope parameters for superplots. To compensate for superplot random effects, a single, liberal estimate of the superplot intra-class correlation coefficient (ICC) for regeneration plot data was used to adjust the significance tests and confidence intervals of parameter estimates reported in the final models (Hedges, 2007). Calculations showed that 95% of ICC values for predictor variables were < 0.06 (i.e., plots were nearly independent for these variables), so we conservatively used a value of 0.10 for adjustment. By construction, superplot variables had an ICC of 1.0, reducing their sample size by a factor of 4. Models were fit using Gauss-Hermite quadrature with 5 quadrature points to maximize the model likelihood (Bates et al., 2015). Due to convergence warnings when pre-harvest models were fit with the negative exponential link function, predictions from these models were checked with a complementary Bayesian model (Su and Yajima, 2015). Trace plots showed that Bayesian estimates converged within 10,000 simulations and their parameter estimates were within 2% of those obtained by models fit with Gauss-Hermite quadrature.

Model development began with the selection of a single predictor (H-DomOak or AgHt-Oak) that best represented the well-established relationship between advance regeneration and oak seedling success (Sander, 1971; Johnson et al., 2002; Gould et al., 2006), or more generally between oak regeneration at time 2 as a function of oak regeneration at time 1. Both variables were tested for fit with a range of exponents above and below 1.0. AgHt-Oak was a good predictor at all ages and clearly the best predictor, using the negative exponential function, for pre-harvest data. For post-harvest data (years 1, 4, and 7), however, H-DomOak using the logistic link function was as predictive of success as AgHt-Oak, and this variable was chosen for those models because of its simplicity of measurement.

From this foundation, model development progressed by adding additional predictor variables in a forward step-wise process, selecting the one at each step that reduced the Akaike's Information Criterion (AIC) the most. The new variable was retained if it met each of two criteria: (1) the effect was significant (P < 0.15) by the Wald test and (2) the predictive power of the new model was a significant (P < 0.15) improvement. The process continued until no new parameters satisfied both criteria. The Wald test was performed following parametric bootstrapping with 1000 iterations to reduce bias in the estimate of standard error of parameters. Model predictive power was determined based upon the area (AUC) under its receiver operator characteristic (ROC) curve. The AUC measures the power of regressions that predict a continuous probability for binary data where the conventional R-squared statistic is not applicable (Faraway, 2016; Harrell, 2015). The greater the area under the ROC curve, the better the model (Faraway, 2016; Harrell, 2015; Xavier et al., 2011), with AUC = 1.0 indicating a perfect model. Delong's test was used to determine whether the AUC of

Table 1
Description of candidate independent variables tested for modeling oak seedling success and the direction of their expected effect (in parentheses) on the probability of success.

Level	Variable	Definition
Plot (4 m^2)	AgHt-Oak	Aggregate height of oak seedlings ≥15 cm (m/m^2) (+)
	AgHt-Maple	Aggregate height of all red maple seedlings (m/m^2) (−)
	AgHt-Birch	Aggregate height of all black birch seedlings (m/m^2) (−)
	AgHt-Other	Aggregate height of all seedlings of other species (m/m^2) (−)
	H-DomOak	Height of the tallest oak seedling (m) (+)
	H-DomMaple	Height of the tallest red maple seedling (m) (−)
	H-DomBirch	Height of the tallest black birch seedling (m) (−)
	H-DomOther	Height of the tallest seedling of other species (m) (−)
	%C fern	Percentage cover of rhizomatous fern (−)
	%C laurel	Percentage cover of mountain-laurel (−)
Superplot (200 m^2)	#Trees-Maple	Density of red maple trees > 5 cm DBH before harvest (number/ha) (−)
	IMI	Integrated Moisture Index (−)
Stand	SI	Northern red oak site index (height (m) at age 50) (−)

Table 2
Frequency of occurrence of tree species in regeneration plots (percentage of plots) and the mean height of the tallest tree of each species within plots (m).

Species group	Pre-Harvest	Year 1	Year 4	Year 7
	frequency/mean height			
Oak	78/0.42	70/0.49	73/1.25	73/2.33
Red maple	92/0.38	87/0.47	92/1.13	90/2.01
Black birch	14/0.95	24/0.66	38/1.75	39/3.51
Other	60/0.46	65/0.50	73/1.37	71/2.61
Other includes:				
Blackgum	10/0.23	11/0.59	11/1.62	10/3.46
Serviceberry	16/0.31	10/0.34	7/0.88	6/1.47
Sassafras	6/0.35	5/0.41	5/1.23	5/3.90
Black cherry	5/0.20	2/0.53	5/1.17	6/3.52
Miscellaneous	13/0.63	37/0.56	45/1.54	44/3.99

the larger model was significantly greater than the AUC of the smaller model (Delong et al., 1988). The percentage deviance explained by each variable was calculated by removing the variable of interest from the full model and determining the difference in model deviance (−2 times the model's log-likelihood) between the reduced model and the full model, then dividing this difference by the deviance of the null model (intercept only) and multiplying by 100%. While not the same as a partial R^2, the percentage deviance explained provides a metric of relative importance among predictors.

All data were analyzed using R statistical software version 3.5.0 (R Core Team, 2018). The "pROC" package in R was used to calculate areas under the ROC curve (Xavier et al., 2011). The "glmer" function in "lme4" package (Bates et al., 2015) was used to fit main models using Gauss-Hermite quadrature. The "R2JAGS" package (Su and Yajima, 2015) was used to fit complementary Bayesian models. Kendall's rank correlations (τ) and Pearson correlations (r) among predictors were calculated using the "Corrplot" package (Wei and Simko, 2017).

3. Results

Of the 2044 plots measured, 228 were successfully stocked with a seedling origin oak at SE, representing a study-wide average success probability of 11%. Most (90%) of these plots contained a seedling-origin oak with a codominant or dominant crown class and the remainder were plots with an "intermediate" oak at least 85% the height of its nearest competitor. Chestnut oak accounted for 40% of all successfully stocked plots, followed by scarlet oak (20%), white oak (16%), and black oak (14%).

Northern red oak seedlings accounted for the smallest percentage (9%) of successes even though northern red oak was the second most abundant oak overstory species in the antecedent stands. Northern red

oak was the only oak species whose advance regeneration was positively but weakly correlated with IMI (τ = 0.1), indicating that it was associated with more mesic conditions than other oak seedlings. H-DomOak in year 7 was lower for northern red oak compared to other oak species (1.60 vs 2.40 m), and plots with a dominant northern red oak in year 7 more often contained a competing black birch (a mesophyte and usually the tallest of competing species) than plots dominated by a different oak species (41% vs 33%).

Of the competing non-oak regeneration, red maple seedlings were the most frequently encountered and were present both before and after harvest in never less than 87% of all plots in any year (Table 2). Black birch was less common than red maple, especially prior to harvest, but black birch appeared in 39% of plots by year 7, and the tallest black birch seedlings in plots were taller than the tallest red maples, on average. Black birch was also taller on average than Other species (primarily blackgum [*Nyssa sylvatica* Marshall], sassafras [*Sassafras albidum* (Nutt.) Nees], serviceberry [*Amelanchier* spp.], and black cherry), although the plot-tallest of some infrequent species in the other category were taller than black birch.

The pre-harvest foundation model, based on AgHt-Oak, was not improved by the addition of non-oak variables, so this became the final model (Table 3). In this model, the average plot – one that contained advance oak regeneration totaling 0.31 m/m^2 of aggregate height – had a 14% probability of becoming stocked with oak at SE (Fig. 2). By comparison, a plot with 1.0 m/m^2 AgHt-Oak had a 24% probability of successful oak recruitment at SE. These are study-wide average predictions and do not account for random stand-to-stand differences, which amounted to as much as 20% of the variation in successfully stocked plots (Fig. 2).

As explained above, H-DomOak was the foundational predictive variable in models for years 1, 4, and 7 after harvest, and the predictive power of these models was improved by the addition of variables that represented competing tree regeneration (H-DomBirch, H-DomMaple, and H-DomOther). Each of these additional variables had significant negative associations with oak success (Table 3) and, with one exception, they satisfied both criteria for inclusion in every post-harvest model. The exception was H-DomMaple in year 1, which failed to significantly increase AUC (P = 0.223). The variable was nonetheless retained for consistency. As with the pre-harvest model, there were large differences among stands that could not be attributed to any measured variable but instead were modeled by random-effect stand intercepts (Fig. 3).

The probability of oak seedling success was substantially reduced by the presence of a black birch within the plot in year 1 (Table 3, Fig. 3). For example, the probability of successful stocking was 20% for a plot with a 0.5 m oak seedling (the average height of the tallest oak in a plot in that year, Table 2) and no non-oak competitors, but the probability

Table 3

Fixed effect estimates, standard errors, and adjusted probability values displayed on the linear predictor scale. Percentage of null deviance explained is the increase in model deviance (2*-(ln(likelihood))) realized when the effect is removed from the full model, expressed as a percentage of the deviance of the null model. The higher the value, the more important the variable.

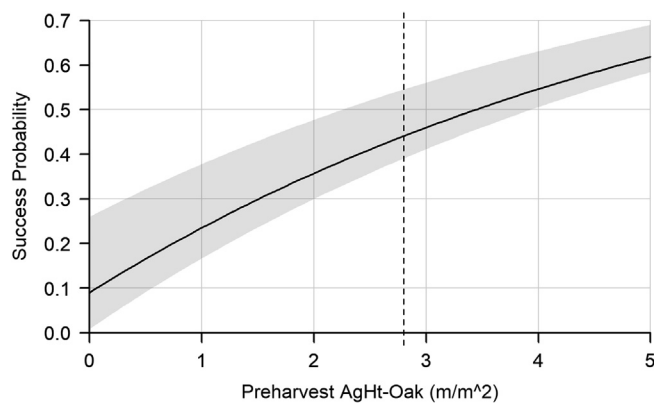| Model	Parameter	Estimate	Adjusted standard error	Adjusted P	% of null deviance explained
Pre- Harvest	Intercept	0.094			
	AgHt-Oak (m/m^2)	0.174	0.032	< 0.0001	5.19
Year 1	Intercept	−3.598			
	√(H-DomOak) (m)	3.082	0.348	< 0.0001	9.19
	√(H-DomBirch) (m)	−2.169	0.654	0.0009	1.38
	H-DomMaple (m)	−0.447	0.233	0.0550	0.38
	H-DomOther (m)	−0.735	0.291	0.0117	0.66
Year 4	Intercept	−3.512			
	H-DomOak (m)	1.543	0.130	< 0.0001	18.18
	H-DomBirch (m)	−0.552	0.151	0.0002	1.48
	H-DomMaple (m)	−0.439	0.117	0.0002	1.43
	H-DomOther (m)	−0.286	0.107	0.0074	0.70
Year 7	Intercept	−4.277			
	H-DomOak (m)	1.022	0.081	< 0.0001	23.19
	H-DomBirch (m)	−0.219	0.071	0.0021	0.98
	H-DomMaple (m)	−0.241	0.065	0.0002	1.36
	H-DomOther (m)	−0.114	0.057	0.0460	0.38

Fig. 2. Probability that a plot will be stocked with oak at SE as a function of pre-harvest aggregate height of oak seedlings. The shaded area encompasses the variability among stands expressed by the maximum and minimum stand intercepts. Most plots (99%) had less than 2.8 m/m^2 aggregate height oak seedlings as marked by the vertical dashed line.

Fig. 3. Probability that a plot will be stocked with oak at SE as a function of the height of the tallest oak seedling in the plot in years 1, 4, and 7, in the absence of competing tree species (solid line) and in the presence of non-oak competitors (dashed lines) with study-wide mean heights according to species (Table 2). Shaded areas encompass the variability among stands expressed by the maximum and minimum stand intercepts.

was reduced to 3% in the presence of a 0.7 m black birch (the average height of the tallest black birch in a plot, Table 2). Black birch continued to have significant negative effects on oak success in years 4 and 7. Red maple and other tree competitors also significantly reduced oak regeneration success, but in most cases the unit-height effects were smaller for these species than for black birch (Table 3), and in all years these species were shorter on average than black birch (Table 2). However, red maple was arguably a more important competitor than black birch because it was a much more frequent competitor of oak seedlings across all stands and in all years (Table 2).

Adding other parameters to the models failed to significantly improve prediction. We found no significant relationship between SI and oak regeneration success, perhaps due to sample size and a narrow range of values (~4 m) for most stands. Other variables (%C Fern, %C Laurel, IMI, #Trees-Maple) had significant correlations with tree regeneration variables already in the model (Kendall's Tau, P < 0.05) but did not contribute significantly to the models. For example, #Trees-Maple was negatively correlated with AgHt-Oak and appears to have limited the accumulation of oak advance regeneration (Fig. 4), but it had no relationship with oak success in the pre-harvest model where AgHt-Oak was the predictor. Similarly, IMI was positively correlated with H-DomBirch, and negatively correlated with H-DomOak and H-
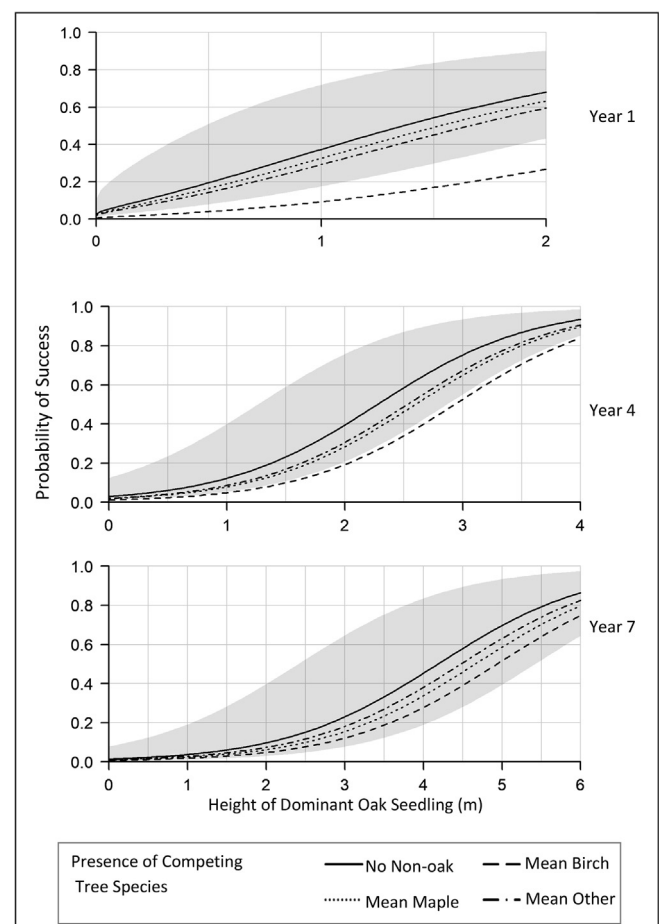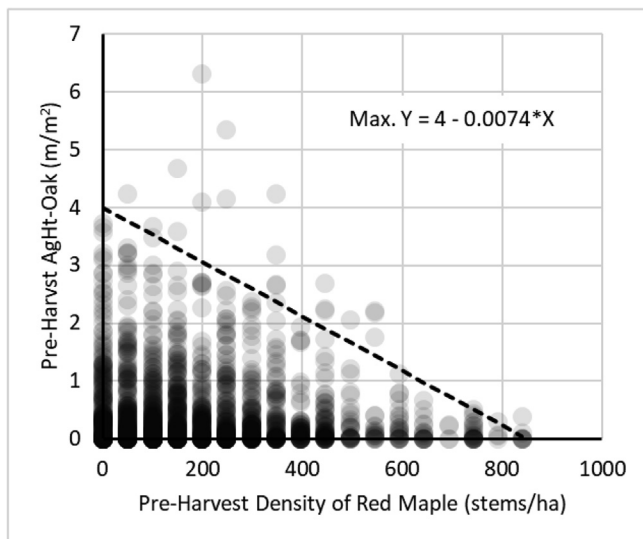
Fig. 4. Relationship between advance oak regeneration (AgHt-Oak) and density of red maple trees in the pre-harvest overstory (DBH $\geq$ 5 cm). Virtually all (99%) observations (dots) fall below the hand-fitted line along what appears to be an approximate limit on pre-harvest AgHt-Oak caused by increasing numbers of red maple trees in the forest canopy.
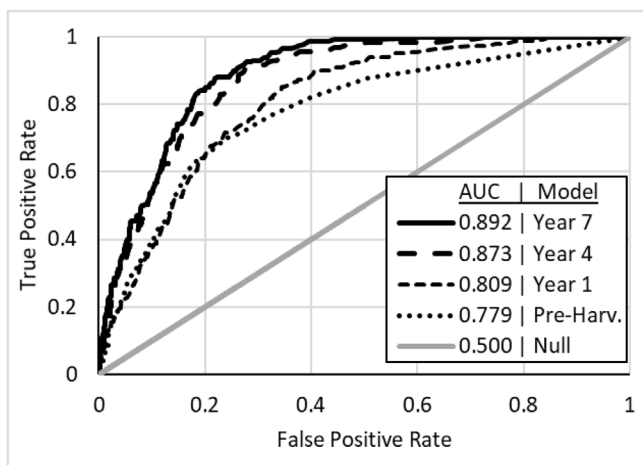

Fig. 5. Receiver operator characteristic curves of models developed to predict success from pre-harvest, year 1, year 4, and year 7 measurements. Increased area under the curve (AUC) indicates increased power. True Positive Rate = sensitivity and False Positive Rate = 1 - specificity of observations classified as successes or failures as the probability threshold for success was increased over the range of possible values from 0 to 1.

DomMaple, but it did not contribute independently to predicting oak success.

As one would expect, the year 7 model was more predictive of oak stocking at SE than those fit with earlier data (Fig. 5). Between the pre-harvest and year 1 models there was a small increase in AUC from 0.78 to 0.81 (P = 0.03), indicating a modest increase in the ability to discriminate between plot success or failure at SE. However, there was a large increase in predictive power between year 1 and year 4 models (AUC increase from 0.81 to 0.87, P < 0.0001), followed by a smaller but still significant (P = 0.003) increase between year 4 and year 7 models. As time progressed from year 1 to year 7 most of the improvement in the model was due to an increase in the percentage deviance explained by H-DomOak rather than by the heights of competitors. Despite their significance, non-oak predictors each accounted for a small amount of the deviance in the data relative to H-DomOak (Table 3).
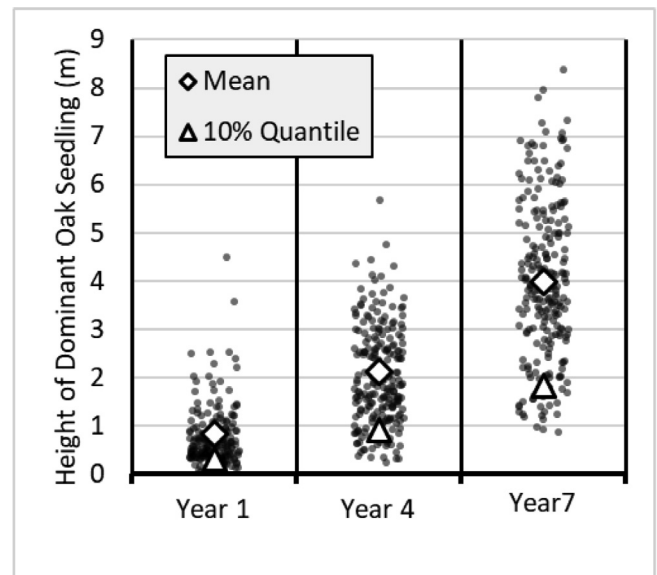

Fig. 6. Heights of dominant oak seedlings in years 1, 4, and 7 in plots that went on to be successfully stocked with oak at SE. A small amount of random noise (lateral spread) was added to observations to better reveal data density.

A wide range of dominant oak seedling heights in years 1, 4, and 7 was associated with plot success at SE (Fig. 6). The variability was lowest in year 1, when most seedlings were < 1 m tall, but by year 7 their average height had increased to 4 m and some seedlings were as tall as 8 m. Some seedlings < 0.15 m in year 1 were successful at SE, even though seedlings this small did not add explanatory power to the pre-harvest model. By year 7, no oak seedling shorter than ~1.0 m led to a successfully stocked plot at SE.

Each percentage of plots with at least one successful oak contributed 1.7–2.6% to total regeneration stocking, depending on the method of calculation. The mean DBH of plot-dominant oaks in successfully stocked plots at SE was 9.7 cm, corresponding to an estimated growing space requirement of 4.26 m^2 (Gingrich, 1967). On average, 1486 such trees per hectare of stand area not stocked with residual trees were required to fill the space available for regeneration (i.e., not stocked with residual trees). Assuming no more than one successful oak in each plot, a 1% increase in the proportion of oak-stocked plots equated to a 1.7% increase in the contribution of seedling-origin oaks to total regeneration stocking, or 19% total for a study-wide average. This is likely an underestimate because it does not account for instances of two or more successful oak seedings occupying a 4-m^2 plot. When stocking was calculated from regeneration > 1.37 m tall, each percentage increase in plot success equated to a 2.6% increase in the oak seedling proportion of regeneration stocking, equating to a study-wide average of 29% at SE. This method accounts for instances of two or more successful oak seedlings in a plot, but it overestimates stocking to the extent that it includes contributions from intermediate and suppressed trees that will soon die.

4. Discussion

These models show that successful, seedling-origin oak regeneration at the stem exclusion stage of stand development can be predicted before harvest from the plot-aggregate height of advance regeneration oak seedlings or, similarly, at stand ages 1, 4, and 7 years after harvest, using the height of the plot-tallest oak seedling together with variables representing non-oak tree seedling competitors. The models are good fits to measured outcomes over the range of probabilities, with moderate positive and negative deviations at the high end of the scales, where class sizes were small (Fig. S1). As expected, models were less

and less predictive at ever earlier ages of development, and the pre-harvest model explained only 5.2% of model deviance in plot-level outcomes. However, when plot-level predictions were binned and averaged – as they were in Fig. S1, and as they would be in making stand-level decisions based on samples of 25 or more plots – the composite predictions became accurate. The principal source of unpredictability was random (i.e., unexplained) stand effects, as illustrated in Figs. 2 and 3.

Our results corroborate previous findings (Gould et al., 2006; Fei et al., 2006) that aggregate height is a useful index of oak seedling density (a composite of number and size) in predicting regeneration success. Given the possibility of non-linear allometric relationships between different size-related variables contributing to competitive advantage (e.g., height vs. leaf area), we expected that models could be improved by calculating aggregate height using some exponent other than 1.0. This was not the case, however, perhaps because of the small sizes involved or distortions in normal allometric relationships imposed by the shaded understory conditions. Similarly, because advance regeneration oak seedlings often die back and re-sprout in the understory, basal stem diameter is sometimes regarded as a more accurate than height as a representation of seedling potential (Sander, 1971, 1972; Miller et al., 2017). To test this, we conducted a post-hoc analysis of the effect of substituting basal diameter in place of height in the calculation of aggregate size, but the substitution did not improve model performance. Earlier research has shown that the percentage of advance oak seedlings > 5 years old is low in stands like those in this study (Steiner et al., 1993), indicating that few seedlings live long enough for the relationship between height and diameter to become distorted through repeated dieback. However, advance regeneration of oak species can become considerably older and presumably larger (Tryon and Powell, 1984), and aggregate height may not be the best predictor of regeneration success under those circumstances.

Competing tree regeneration was considerably less predictive of future oak stocking than were the size and abundance of oak seedlings, but the heights of non-oak competitors contributed importantly to model predictive power in post-harvest models. Red maple was found on 9 of every 10 plots from pre-harvest through age 7 and was the most common non-oak species encountered (Table 2). Red maple is probably the most important single competitor to oaks in general throughout the eastern U.S. (Fei and Steiner, 2007), and its abundance and negative relationship with oak seedling success in the current study underscores red maple's importance as a competitor in regenerating oak stands. Although less abundant than red maple, black birch was found on 38% of all plots by year 4, and where the species was present it was usually considerably taller than the tallest red maple seedling (Table 2). H-DomBirch was negatively correlated with H-DomOak in all years ($\tau = -0.1$, P < 0.001), and although black birch was far less common in plots than red maple it was size-for-size a stronger competitor against oak until year 7 based on model coefficients (Table 3). Where the species is native, black birch has long been considered a strong competitor in stands regenerating after timber harvest in eastern portions of the CHR (McGee and Hooper, 1970; Beck and Hooper, 1986; Johnson et al., 2009). However, black birch often does not sustain its dominance over oak after about age 20 (Oliver, 1978; Steiner et al., 2018), which means that our models may underestimate the competitive position of oaks in plots that also contain black birch at SE.

#Trees-Maple was included among predictor variables as a surrogate for the density of nearby red maple sprouts based on the observation that > 90% of red maple trees in the study region sprout after cutting (Fei and Steiner, 2009). (Stump sprouts of other non-oak species were infrequent and not considered as potential model components.) However, while measurements of red maple seedlings improved the predictive power of models, #Trees-Maple did not. The fact that this estimate of red maple stump sprout density was spatially imprecise with respect to regeneration plots may have contributed to its non-significance. However, because red maple stump sprouts necessarily

originated from pre-harvest overstory trees, it is likely that any direct effect of competition from red maple stump sprouts was overshadowed by the effect of red maple's pre-harvest suppression of oak advance regeneration. This relationship (Fig. 4) illustrates the role of red maple as a suppressor of advance oak regeneration in addition to its role as a competitor in regenerating stands. Recalling that red maple averaged only 15% of pre-harvest stand basal area and was a minor component to oak in every stand, we attribute its influence on advance regeneration to the common presence of hundreds of small red maple stems per ha. in suppressed and intermediate canopy positions (Fei and Steiner, 2009). The presence of low shade from small woody stems is known to impede the growth and survival of oak seedlings beneath a forest canopy (Loftis, 1990b; Lorimer, 1994).

The largest increase in model predictive power occurred between years 1 and 4 and was followed by a much more modest increase between years 4 and 7. This agrees with other research within the Central Hardwood Region (Zenner et al., 2012; Swaim et al., 2016; Vickers et al., 2019) in suggesting that useful forecasts of future stratification of trees into crown classes can be made as early as year 4. Early silvicultural intervention in the form of crop tree release or cleaning (Nyland, 1996) is sometimes advocated to favor oak and other desired species in developing hardwood stands. To our knowledge, there are no studies of cleaning in 4-year-old stands, but the method has significantly increased survival and diameter growth of codominant and intermediate oaks when applied as early as age 7 (Allen and Marquis, 1970; Ward, 2013). We observed wide variation in the heights of plot-dominant oak seedlings that led to successful stocking at SE (Fig. 6), but most were > 1 m tall by year 4 and > 2 m tall by year 7. These values are reasonable benchmarks for minimum competitive height requirements for oak seedlings in those years, but they do not assure success (Fig. 3). Our results indicate that success probabilities of oak seedlings even in the 2- to 3-m height class at year 4 could be improved by 50–100% if taller stems of black birch and red maple were removed within a small (ca. 1 m) radius.

The antecedent stands in this study had an average of 56% oak stocking at the time of harvest, and it is too early to say whether their replacements will reach that level. Our estimates of stocking at SE did not include contributions from stump sprouts, and previous research indicates that during the next 20 years the oak component in some stands will gain relative basal area at the expense of red maple and black birch (Steiner et al., 2018). If these stands do show a decline in oak importance compared to their predecessors, it will be attributable primarily to northern red oak. This species made up 22% of all oak basal area in the previous overstory, but it was underrepresented in advance regeneration (14% of AgHt-Oak) and accounted for only 9% of plots that ultimately regenerated to oak. These numbers indicate a comparative failure of northern red oak to establish a strong cohort of advance regeneration and to dominate in post-harvest competition. This species' tendency toward greater abundance on relatively mesic sites (elevated IMI) may have exposed it to heavier non-oak competition both before and after harvest (Adams and Anderson, 1980; Johnson et al., 2002), and the results tend to support that conclusion.

Our findings provide new information about the relationship between advance oak seedling regeneration and oak stocking at the stem exclusion stage of stand development in 15- to 20-year-old stands. A model based solely on plot aggregate height of advance regeneration was optimal, presumably because other factors known to affect the success of oak seedlings (e.g., competition and site quality) were prefigured in the occurrence and size of advance regeneration. Results show that even advance regeneration seedlings that are small at the time of harvest can contribute importantly to successful oak regeneration, in agreement with the findings of some previous studies (Ross et al., 1986; Gould et al., 2006) but in contrast with others (Sander, 1971, 1972; Sander et al., 1984). These differences emphasize that oak regeneration models are perhaps best suited to the specific regions in which the data are collected. After harvest, the structure of predictive

models changed when, at ages 1, 4, and 7, the size of the tallest oak seedling became increasingly determinate of stocking success, and the sizes of competing species (especially black birch and red maple) became important as inhibiting factors. The power of these models to predict the ultimate success of oak seedlings based upon conditions before and 1–7 years after stand disturbance can contribute to better-informed decisions about harvest scheduling and early interventions to favor oak regeneration. In our view, the best opportunity for future model improvements lies in understanding the causes of stand-to-stand differences in regeneration success that remained unexplained in this study.

CRediT authorship contribution statement

Lake E. Graboski: Conceptualization, Methodology, Formal analysis, Investigation, Writing - original draft. **Kim C. Steiner:** Conceptualization, Methodology, Writing - review & editing, Supervision, Funding acquisition. **Marc E. McDill:** Methodology, Writing - review & editing, Data curation. **James C. Finley:** Methodology, Writing - review & editing, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foreco.2020.118093.

References

Abrams, M.D., 1992. Fire and the development of oak forests in eastern North America. Bioscience 42, 346–353.

Abrams, M.D., Nowacki, G.J., 1992. Historical variation in fire, oak recruitment, and post-logging accelerated success in central Pennsylvania. Bull. Torrey Bot. Club 119, 19–28.

Adams, D.E., Anderson, R.C., 1980. Species response to a moisture gradient in central Illinois forest. Am. J. Bot. 67, 381–392.

Allen, R.H. Jr., Marquis, D.A., 1970. Effect of thinning on height and diameter growth of oak & yellow-poplar saplings. USDA Forest Service Res. Pap. NE-173, p. 11.

Bates, D., Maechler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67, 1–48.

Beck, D.E., Hooper, R.M., 1986. Development of a southern Appalachian hardwood stand after clearcutting. South. J. Appl. For. 10, 168–172.

Campbell, A.J., Steiner, K.C., Finley, J.J., Leites, L., 2015. Limitations on regeneration potential after even-aged harvests in mixed-oak stands. For. Sci. 61, 874–881.

Carmean, W.H., Hahn, J.T., Jacobs, R.D., 1989. Site index curves for tree species in the eastern United States. USDA Forest Service Gen. Tech. Rep. NC-128, p. 142.

Clark, F.B., 1993. An historical perspective of oak regeneration. In: Loftis, David, McGee, Charles E. (Eds.). Oak regeneration: serious problems, practical recommendations. USDA Forest Service Gen. Tech. Rep. SE-84, p. 319.

DeLong, E.R., DeLong, D.M., Clarke-Pearson, D.L., 1988. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44, 837–845.

Faraway, J.J., 2016. Extending the Linear Model With R, second ed. Taylor and Francis Group LLC Ch2 Binary Response.

Fei, S., Gould, P.J., Steiner, K.C., Finley, J.C., 2006. Aggregate height- a composite measure of stand density for tree seedling populations. For. Ecol. Manage. 223, 336–341.

Fei, S., Steiner, K.C., 2007. Evidence for increasing red maple abundance in the eastern United States. For. Sci. 53, 473–477.

Fei, S., Steiner, K.C., 2009. Rapid capture of growing space by red maple. Can. J. For. Res. 39, 1444–1452.

Fei, S., Gould, P., Kaeser, M., Steiner, K.C., 2010. Distribution and dynamics of the invasive native hay-scented fern. Weed Sci. 58, 408–412.

Fei, S., Kong, N., Steiner, K.C., Moser, W.K., Steiner, E.B., 2011. Change in oak abundance in the eastern United States from 1980 to 2008. For. Ecol. Manage. 262, 1370–1377.

Fralish, J.S., 2004. The keystone role of oak and hickory in the central hardwood forest. In: Spetich, M.A. (Ed.), Upland oak ecology symposium: history, current conditions, and sustainability. USDA Forest Service Gen. Tech. Rep. SRS-73, p. 331.

Gingrich, S.F., 1967. Measuring and evaluating stocking and stand density in upland hardwood forests in the central states. Forest Science 13, 38–53.

Gould, P.J., Steiner, K.C., Finley J.C., McDill, M.E., 2003. Regenerating mixed-oak stands in Pennsylvania: a quarter-century retrospective. In: Van Sambeek, J.W., Dawson, J. O., Ponder, F., Jr., Loewenstein, E.F., Fralish, J.S., (Eds.), Proceedings, 13th Central Hardwood Forest conference. USDA Forest Service Gen. Tech. Rep. NC-234, pp. 254–258.

Gould, P.J., Steiner, K.C., Finley, J.C., McDill, M.E., 2005. Developmental pathways following the harvest of oak-dominated stands. For. Sci. 51, 76–90.

Gould, P.J., Steiner, K.C., McDill, M.E., Finley, J.C., 2006. Modeling seedling-origin oak regeneration in the central Appalachians. Canadian J. For. Resour. 36, 833–844.

Harrell, F.E. Jr., 2015. Regression modeling strategies, second ed. Springer Series in Statistics. Springer International, Switzerland (Ch 10 Binary Logistic Regression).

Hedges, L.V., 2007. Correcting a significance test for clustering. J. Educ. Behav. Stat. 32, 151–179.

Helms, J.A., 1998. The Dictionary of Forestry. Pub. Bethesda, MD Society of American Foresters, p. 210.

Iverson, L.R., Dale, M.E., Scott, C.T., Prasad, A., 1997. A GIS-derived integrated moisture index to predict forest composition and productivity of Ohio forests (U.S.A.). Landscape Ecol. 12, 331–348.

Johnson, P.S., Shifley, S., Rodgers, R., 2009. The Ecology and Silviculture of Oaks, second ed. CAB Publishing, CAB International, Wallingford, Oxon, UK, pp. 580.

Knott, J.A., Desprez, J.M., Oswalt, C.M., Fei, S., 2019. Shifts in forest composition in the eastern United States. For. Ecol. Manage. 433, 176–183.

Loftis, D.L., 1990a. Predicting post-harvest performance of advance red oak reproduction in the southern Appalachians. For. Sci. 36, 908–916.

Loftis, D.L., 1990b. A shelterwood method for regenerating red oak in the southern Appalachians. For. Sci. 36, 917–929.

Lorimer, C.G., Chapman, J.W., Lambert, W.D., 1994. Tall understory vegetation as a factor in the poor development of oak seedlings beneath mature stands. J. Ecol. 82, 227–237.

McGee, C.E., Hooper, R.M., 1970. Regeneration after clearcutting in the southern Appalachians. USDA Forest Service Research Paper SE- 70, p. 12.

McIntyre, A.C., 1936. Sprout groups and their relation to the oak forests of Pennsylvania. J. Forest. 34, 1054–1058.

McWilliams, W.H., O'Brian, R.A., Reese, G.C., Waddell, K.L., 2002. Distribution and abundance of oaks in North America. In: McShea, W.J., Healy, W.M. (Eds.), Oak forest ecosystems: ecology and management for wildlife. Johns Hopkins University Press, Baltimore, MD, pp. 13–32.

Miller, G.W., Brose, P., Gottschalk, K.W., 2017. Advanced oak seedling development as influenced by shelterwood treatments, competition control, deer fencing, and prescribed fire. J. Forest. 115, 179–189.

Nowacki, G.J., Abrams, M.D., 1992. Community, edaphic, and historical analysis of mixed-oak forests in the Ridge and Valley province of central Pennsylvania. Can. J. For. Res. 22, 790–800.

Nowacki, G.J., Abrams, M.D., 2008. The demise of fire and "mesophication" of forests in the eastern United States. Bioscience 58, 123–138.

Nyland, R.D., 1996. Silviculture Concepts and Applications. McGraw-Hill, New York, pp. 633.

Oliver, C.D., 1978. The development of northern red oak in mixed species stands in central New England. Yale University School of Forestry and Environmental Studies Bulletin 91, p. 63.

Oliver, C.D., Larson, B.C., 1996. Forest Stand Dynamics. McGraw-Hill, New York, pp. 520.

Pennsylvania State Climatologist. 2019. Regional climate trends for Pennsylvania from 1980 to 2018. Regions 7 and 5. Accessed online on 11/24/2019 at < http://climate.met.psu.edu > .

Peters, M.P., Iverson, L.R., Matthews, S.N., Prasad, A.M., 2013. Wildfire hazard mapping: exploring site conditions in eastern US wildland–urban interfaces. Int. J. Wildland Fire 22, 567–578.

R Core Team, 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available online at < http://www.r-project.org > . Accessed 10/07/2019.

Ross, M.S., Sharik, T.L., Smith, D.W., 1986. Oak regeneration after clear felling in southwest Virginia. For. Sci. 32, 157–169.

Sander, I.L., 1971. Height growth of new oak sprouts depends on size of advance reproduction. J. Forest. 69, 809–811.

Sander, I.L., 1972. Size of oak advance reproduction: key to growth following harvest cutting. USDA Forest Service Research Paper NC-79, p. 11.

Sander, I.L., Clark, F.B., 1971. Reproduction of upland hardwood forests in the central United States. USDA. Forest Service. Agricultural Handbook 405, p. 25.

Sander, I.L., Johnson, P.S., Rogers, R., 1984. Evaluating oak advance reproduction in the Missouri Ozarks. USDA Forest Service Research paper NC-251, p. 19.

Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Web Soil Survey. Available online at < https://websoilsurvey.sc.egov. usda.gov/ > . Accessed 10/07/2019.

Spetich, M.A., Dey, D.C., Johnson, P.S., Graney, D.L., 2002. Competitive capacity of Quercus rubra L. planted in Arkansas' Boston Mountains. Forest Science 48, 504–517.

Steiner, K.C., Abrams, M.D., Bowersox, T.W., 1993. Advance reproduction and other stand characteristics in Pennsylvania and French stands of northern red oak. Proceedings of the 9th Central Hardwood Forest Conference, USDA Forest Service Gen. Tec. Rep. NC-161, pp. 473–483.

Steiner, K.C., Finley, J.C., Gould, P.J., Fei, S., McDill, M.E., 2008. Oak regeneration guidelines for the central Appalachians. North. J. Appl. For. 25, 5–11.

Steiner, K.C., Stein, B.S., Finley, J.C., 2018. A test of the delayed oak dominance hypothesis at mid-rotation in developing upland stands. For. Ecol. Manage. 408, 1–8.

Swaim, T.J., Dey, D.C., Saunders, M.R., Weigel, D.R., Thornton, C.D., Kabrick, J.M., Jenkins, M.A., 2016. Predicting the height growth of Quercus species (Quercus) reproduction over a 23-year period following clearcutting. For. Ecol. Manage. 364, 101–112.

Tallamy, D.W., Shropshire, K.J., 2009. Ranking lepidopteran use of native versus introduced plants. Conserv. Biol. 23, 941–947.

Su, Y.S., Yajima, M., 2015. R2jags: a package for running jags (Just Another Gibbs Sampler). R. package version 0.5-7. Available online at https://CRAN.R-project.org/package=R2jags. Accessed 10/07/2019.

Trimble, G.R., Hart, G. 1961. An appraisal of early reproduction after cutting in Northern Appalachian Hardwood Stands. USDA Forest Service Northeastern Forest Experiment Station. Paper NE-162, p. 25.

Tryon, E.H., Powell, D.S., 1984. Root ages of advance hardwood reproduction. For. Ecol. Manage. 8, 293–298.

Vickers, L.A., Larsen, D.R., Knapp, B.O., Kabrick, J.M., Dey, D.C., 2019. Height development milestones for canopy recruitment after overstory removal in the Missouri Ozarks. For. Ecol. Manage. 445, 122–133.

Ward, J.S., 2013. Precommercial crop tree release increases upper canopy persistence and diameter growth of oak saplings. North. J. Appl. For. 30, 156–163.

Ward, J.S., Stephens, G.R., 1994. Crown class transition rates of maturing northern red oak (Quercus rubra L.). For. Sci. 40, 221–237.

Wei, T., Simko, V., 2017. Corrplot: visualization of a correlation matrix. R package version0.84. Available online at < github.com/taiyun/corrplot > . Accessed 10/07/2019.

Xavier, R., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., Müller, M., 2011. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinf. 12, 77.

Zenner, E.K., Heggenstaller, D.J., Brose, P.H., Peck, J.E., Steiner, K.C., 2012. Reconstructing the competitive dynamics of mixed-Quercus neighborhoods. Canadian J. For. Resour. 42, 1714–1723.