Regional Locale and Its Influence on the Prediction of Loblolly Pine Diameter Distributions

Matthew B. Russell, Harold E. Burkhart, Ralph L. Amateis, and Stephen P. Prisley

Data gathered from intensively and nonintensively managed loblolly pine (Pinus taeda L.) plantations were used to model the diameter distributions of stands across the southeastern United States. Weibull scale and shape parameters were predicted using stand density, site index, and stand age as covariates. Including geographic locale (latitude and longitude) of the stand improved the diameter distribution prediction from 1.9 to 16.9% when two common goodness-of-fit-statistics were applied to the models. Cumulative distribution function regression methods performed up to 13% better than a moment-based parameter recovery approach for estimating the parameters of the diameter distribution. The resultant models indicate that for a given set of stand conditions, plantations at northern latitudes exhibit a distribution shifted toward larger diameter classes; however, nonintensively managed plantations at eastern locales exhibit a reverse trend: diameter distributions were predicted to shift toward smaller diameter classes, with a larger mean diameter predicted to occur at western locales. These results highlight the importance of quantifying differences in management practices and the gain from incorporating regional locale information in predicting loblolly pine growth and yield throughout its natural range.

Keywords: latitude, longitude, cumulative distribution function regression, Weibull distribution, Pinus taeda

Loblolly pine (Pinus taeda L.) plantations have been established across different physiographic regions throughout the southeastern United States. With a natural range extending from southern New Jersey south to central Florida and west to eastern Texas (Baker and Langdon 1990), plantations are subject to a diverse set of edaphic and climatic conditions. Regionwide data sets are available to study two generations of plantations: those exemplifying more intensive management strategies (with fertilization, competition control, and planting genetically improved stock) and those representative of less intensive methods (no intermediate fertilization or competition control, planting nonimproved stock). The results of intensive management in loblolly pine plantations have been noted, e.g., growth responses from fertilization were greater in larger diameter trees (Carlson et al. 2008). In terms of geographic variability, Hasenauer et al. (1994) found that at a given site quality, the maximum amount of basal area that a stand can carry varies across the range of loblolly pine. Although climate variables, such as annual precipitation, have been shown to be effective predictors of basal area growth (Amateis and Burkhart 2008), geographic variables, such as latitude and longitude, are advantageous to use as predictors in loblolly pine growth equations because they are easy to obtain and have been shown to account for growth and yield differences across the southeastern United States (Amateis et al. 2006, Russell et al. 2010). Stand-level basal area prediction models developed by Amateis et al. (2006) show that plantations with northerly latitudes and westerly longitudes display greater basal area.

Diameter distribution models offer insight into stand structure and are commonly estimated using the Weibull probability density function. Diameter distribution methods are advantageous in that they provide insight into forest stand structure that is valuable in guiding management decisions. By using diameter distributions, stand attributes can be disaggregated into size classes, individual tree volume equations can subsequently be used to estimate total stand yield, and tree mortality can be assigned within distribution models (e.g., Cao 1997). Cao (2004) evaluated several methods for predicting the Weibull density parameters and concluded that a cumulative distribution regression approach outperformed that of a traditional moment-based parameter recovery method. Despite expansive ranges in loblolly pine plantation ownerships and the need to quantify regionwide trends in productivity, diameter distribution models are generally developed for use in a specific region (e.g., Baldwin and Feduccia 1987, Knowe 1992, Lee and Coble 2006).

The primary objectives of this analysis were to (1) evaluate the influence of geographic locale in predicting the diameter distribution of loblolly pine stands, and (2) quantify the effects of intensive management practices on loblolly pine diameter distributions.

Data

Data from intensively managed plantations (IMPs) were acquired from 172 permanent plot locations across the natural range of loblolly pine (Amateis et al. 2006, Russell et al. 2010). These plots are representative of intensive silvicultural practices conducted in...
loblolly pine plantations at the time of their establishment. Common management strategies of the sampled stands include site preparation, planting genetically improved stock, and applying fertilization and competition control as needed. Twenty-seven different half-sibling families, some single families and others mixed, were planted among the locations. Each location consisted of one nonthinned control plot and two plots with thinning, with the heavily thinned plot also receiving a pruning treatment. Plots were established in 1996–1999 in stands ranging in age from 3 to 8 years and are being remeasured every 2 years.

Data from nonintensively managed plantations (NIMPs) were obtained from a regionwide thinning study established at 186 permanent locations in stands planted with nonimproved stock on cutover, site-prepared lands (Burkhart et al. 1985). These plots did not receive midrotation fertilizer applications or competition control treatments. Each location consisted of a control treatment, a light thin treatment, and heavy thin treatment. Plots were established in 1980–1982 in stands ranging in age from 8 to 25 years and were remeasured every 3 years for 21 years, totaling seven remeasurements.

Only nonthinned, control plots were used from both data sets. Given the differences in stand development of the two data sets, and to compare diameter distribution models with data from a similar range of ages, the last remeasurement was used from IMPs and the initial measurement was used from NIMPs. This resulted in 142 IMPs and 152 NIMPs, respectively. The resultant 142 IMPs spanned approximately 7° of latitude and 18° of longitude, whereas NIMPs spanned approximately 8° of latitude and 20° of longitude (Figure 1). Summary statistics for data used from both data sets are found in Table 1.

As the two data sets differed in terms of the period under investigation (i.e., approximately 1970–1982 for NIMPs and 1994–2006 for IMPs), one might suspect that climate conditions influencing growth might differ during the two periods. After selecting a subset of representative locations throughout the natural range of the species and obtaining climate data from the PRISM Climate Group (2011), analyses did not indicate that the climate variables mean maximum temperature, mean minimum temperature, or mean annual precipitation differed appreciably between the two study periods.

### Methods and Results

#### Model Development

Two methods, parameter recovery and cumulative distribution function, were used and compared for estimating the parameters of a Weibull function.

#### Parameter Recovery

The Weibull function has long been recognized as a suitable probability density function for estimating the diameter distribution of even-aged, single-species forest stands (Bailey and Dell 1973). In forest growth and yield studies, the cumulative distribution function for the Weibull commonly takes the following form:

$$ F(x) = 1 - \exp\left[-\left(\frac{dbh_i - a}{b}\right)^c\right] $$  

where \(a\), \(b\), and \(c\) are the location, scale, and shape parameters, respectively; and \(dbh_i\) is dbh (cm) of tree \(i\) in plot \(j\). The location parameter \(a\) was estimated from \(\hat{D}_0\), the predicted minimum dbh of the stand, by subtracting one-half the width of the diameter class from the predicted minimum diameter. Hence,

$$ a = \hat{D}_0 - 1 $$

because 2-cm diameter classes were used. Minimum dbh was estimated and predicted separately for the IMP and NIMP data sets:

$$ \hat{D}_0 = \exp[a_1 + a_2\bar{D}] $$

where \(\bar{D}\) is observed arithmetic mean dbh (cm). The \(b\) and \(c\) parameters of Equation 1 were recovered on estimating the predicted arithmetic (\(\hat{D}\)) and quadratic (\(\hat{D}_q\)) mean dbh using moment-based parameter recovery (PR) techniques (Cao et al. 1982). The natural logarithm of stand density (\(N\); trees ha\(^{-1}\)) and the average height of dominant and codominant trees (\(H_d\); m) were two significant covariates in predicting \(\hat{D}_q\):

$$ \hat{D}_q = \exp[b_1 + b_2\ln(N) + b_3\ln(H_d)] $$

where \(b_1\) and \(b_2\) are coefficients to be estimated.

Geopositional variables were evaluated to determine their effectiveness in improving the accuracy of Equation 4. Latitude and longitude proved significant when scaled and centered around the approximate mean values in the data (34° latitude, −84° longitude) to avoid collinearity with the intercept term when used in regression (Amateis et al. 2006). Elevation, aspect, and physiographic region were not significant predictors. A final equation which included geographic locale took the following form:

$$ \hat{D}_q = \exp[b_1 + b_2\ln(N) + b_3\ln(H_d) + b_4LAT + b_5LONG] $$

where \(LAT\) = latitude (decimal degrees) −34 and \(LONG\) = longitude (decimal degrees, expressed as a negative value west of the prime meridian) +84. \(LAT\) and \(LONG\) were insignificant covariates.
in predicting \( \hat{D}_0 \) (Equation 3) for IMPs and NIMPs. To ensure that \( \hat{D}_q > \hat{D} \), the equation

\[ \hat{D} = \hat{D}_q - \exp(c_0) \]  

was used to predict arithmetic mean dbh, where \( A^{-1} \) is the inverse of stand age (years), and \( c_1 \) and \( c_2 \) are parameters to be estimated. Seemingly unrelated regression was used to estimate coefficients for Equations 3–6 (Borders 1989). A full-model versus reduced-model \( F \)-test concluded that for IMPs, the \( \hat{D}_q \) and \( \hat{D} \) separate regression coefficients were needed for both IMPs and NIMPs (\( \alpha = 0.05 \); Table 2).

Cumulative Distribution Function

The objective of the cumulative distribution function (CDF) method for parameter estimation is to minimize the error sums of squares with respect to Equation 1. The \( \hat{b} \) and \( \hat{c} \) parameters were iteratively searched to minimize the following function:

\[ \sum_{i} \sum_{j} \left( \hat{F}_{ij} - \hat{F}_{ij} \right)^2 / n_i \]

where \( \hat{F}_{ij} \) is the observed cumulative probability of tree \( j \) in plot \( i \); \( \hat{F}_{ij} \) is the right-hand side of Equation 1; \( n_i \) is the number of trees in plot \( i \); and \( n \) is the number of plots. The Weibull location parameter \( a \) was estimated using Equations 2 and 3, and the equations that predicted the scale (\( \hat{b} \)) and shape (\( \hat{c} \)) parameters best were Equations 7 and 8, respectively:

\[ \hat{b} = \exp\left[ b_1 + b_2RS + b_3\ln(N) + b_4A^{-1} \right] \]  

\[ \hat{c} = \exp\left[ c_1 + c_2RS + c_3SI + c_4A^{-1} \right] \]  

where \( RS = (10000/N)^{0.5}/H_d \) is the relative spacing of trees within the plot. For NIMPs, the SI variable was not significant for predicting the shape parameter \( \hat{c} \), which was likely due to the negative correlation observed between \( RS \) and \( SI \) (\( P < 0.001 \)); \( \text{LAT} \) and \( \text{LONG} \) were additional significant covariates in predicting \( \hat{b} \) and \( \hat{c} \):

\[ \hat{b} = \exp\left[ b_1 + b_2RS + b_3\ln(N) + b_4A^{-1} + b_5\text{LAT} + b_6\text{LONG} \right] \]  

\[ \hat{c} = \exp\left[ c_1 + c_2RS + c_3SI + c_4A^{-1} + c_5\text{LAT} + c_6\text{LONG} \right] \]

For IMPs, the \( \text{LONG} \) variable was not significant in predicting \( \hat{c} \), which may have been due to the positive correlation between \( \text{LONG} \) and \( A \) (\( P = 0.0068 \)) and negative correlation between \( \text{LONG} \) and \( RS \) (\( P = 0.0056 \)) in the plots examined. Parameter estimates for Equations 7–10 are found in Table 3.

Model Performance

Two goodness-of-fit statistics were computed for the PR and CDF methods: the Anderson-Darling (AD) statistic (Anderson and Darling 1954) for the \( \hat{a} \)th plot:

\[ AD_i = -n_i - \frac{1}{n_i} \sum_{j=1}^{n_i} (2j - 1)[\ln(u_j) + \ln(1 - u_{i-j+1})] \]
where \( u_j = f(d_{bh,j}) \) = the CDF evaluated at \( dbh_j \) (where \( dbh \) values are sorted in ascending order \( \{dbh_1 \leq dbh_2 \leq \ldots \leq dbh_n\} \) for each plot), and the Reynolds et al. (1988) error index (EI):

\[
EI_i = \sum_{j=1}^{k_i} |n_{i,k} - \hat{n}_{i,k}|
\]

where \( n_{i,k} \) and \( \hat{n}_{i,k} \) are the observed and predicted number of trees per hectare in the \( k \)th diameter class, respectively. Both measures compared observed and predicted diameter attributes: AD uses the predicted diameter evaluated from the cumulative distribution function for each tree within a plot and compares it with the total number of observed trees in the plot, whereas EI measures the sum of absolute differences between observed and predicted number of trees within each diameter class for the \( k \)th plot. AD and EI are computed for each plot and then averaged to compare error estimates for each fitting method (Cao 2004).

Mean squared error (MSE) values for the PR method for \( \hat{D}_0 \) and \( \hat{D}_j \) (Equations 3, 4, and 6—the baseline equations) were 4.83, 1.20, and 1.36 for IMPs and 2.65, 1.07, and 1.15 for NIMPs, respectively. For these same variables, MSEs were 4.84, 1.01, and 0.96 for IMPs and 2.65, 0.84, and 0.89 for NIMPs, respectively, when geographic locale was included (Equations 3, 5, and 6).

Including geographic locale improved the prediction of the scale (\( \hat{b} \)) and shape (\( \hat{c} \)) parameters for both data sets (Table 4). For the CDF method, AD improved by 1.9 and 16.9% for IMPs and NIMPs, respectively. Similarly, EI values improved by 1.8 and 6.0% for IMPs and NIMPs, respectively, when geographic locale was included in the model.

CDF fitting methods outperformed PR methods for both data sets (Table 4). AD improvement ranged from 5.8 to 10.6% depending on the data set and whether or not geographic locale was included in the model. Similarly, EI improvement ranged from 4.7 to 13.0% when the CDF method was used over the PR method.

Using a representative set of stand conditions (1,400 trees ha\(^{-1}\), relative spacing = 0.2; site index = 21 m; age = 15 years), Equations 9 and 10 were used to generate diameter distributions at selected geographic coordinates found throughout the natural range of loblolly pine (Figure 2). Results showed that for a given set of stand conditions, IMPs and NIMPs displayed a modal value of the Weibull density shifted to a larger diameter moving from south to north across the range of loblolly pine. Trends were similar moving from east to west locales for NIMPs, and an opposite trend was observed for IMPs.

**Table 3.** Estimates (standard errors in parentheses) for scale (\( \hat{b} \)) and shape (\( \hat{c} \)) parameters (Equations 7–10) using cumulative distribution function fitting techniques for intensively (IMP) and nonintensively (NIMP) managed loblolly pine plantations with and without geographic locale variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Parameter</th>
<th>IMP</th>
<th>NIMP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>No geographic locale</td>
<td>With geographic locale</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7.3059 (0.08)</td>
<td>7.0395 (0.08)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2.7303 (0.06)</td>
<td>-2.7839 (0.07)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.5740 (0.01)</td>
<td>-0.5400 (0.01)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.9711 (0.19)</td>
<td>1.5557 (0.19)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.0115 (0.001)</td>
<td>0.5740 (0.01)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.0033 (0.0004)</td>
<td>0.0214 (0.01)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.7680 (0.19)</td>
<td>2.0851 (0.18)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.3202 (0.48)</td>
<td>-3.1588 (0.47)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.0214 (0.01)</td>
<td>-0.0383 (0.007)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.9771 (1.25)</td>
<td>7.6803 (1.29)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.0212 (0.006)</td>
<td>-0.0103 (0.002)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Parameter</th>
<th>No geographic locale</th>
<th>With geographic locale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>6.4413 (0.04)</td>
<td>6.9574 (0.04)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.8854 (0.03)</td>
<td>-2.2786 (0.03)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.4710 (0.01)</td>
<td>-0.5365 (0.005)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.0460 (0.11)</td>
<td>-0.3520 (0.10)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.0298 (0.0007)</td>
<td>0.0087 (0.0003)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.8926 (0.03)</td>
<td>0.3636 (0.07)</td>
</tr>
</tbody>
</table>

**Table 4.** Means* (and standard deviations) of goodness-of-fit statistics for intensively (IMP) and nonintensively (NIMP) managed loblolly pine plantations with and without geographic locale using moment-based parameter recovery (PR) and cumulative distribution function (CDF) fitting techniques.

<table>
<thead>
<tr>
<th>Data Method</th>
<th>Anderson-Darling</th>
<th>Error index</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No geographic locale</td>
<td>PR</td>
<td>8.96 (12.35)</td>
</tr>
<tr>
<td>With geographic locale</td>
<td>CDF</td>
<td>8.01 (7.24)</td>
</tr>
<tr>
<td>NIMP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No geographic locale</td>
<td>PR</td>
<td>3.64 (3.43)</td>
</tr>
<tr>
<td>With geographic locale</td>
<td>CDF</td>
<td>3.43 (3.42)</td>
</tr>
</tbody>
</table>

* The smaller the value, the better the fit.

Discussion

Results of this analysis showed that geographic locale information can be used directly in the prediction of diameter distributions of loblolly pine plantations. Latitude and longitude, in addition to traditional measures of stand density, site index, and stand age, offered improved predictions for quantifying loblolly pine diameter distributions for intensively and nonintensively managed plantations.

In agreement with Cao (2004), we also found that a simultaneous estimation of parameters via the cumulative distribution function regression technique outperformed a moment-based parameter recovery method. Differences between IMPs and NIMPs were shown; these differences were likely due to the impacts of management practices on diameter growth potential, such as the planting of improved genetic stock and fertilization. For IMPs and NIMPs, models using a common set of stand conditions predicted that plantations at northern latitudes would exhibit greater mean stand diameter, as indicated by a shifting of the distribution toward larger diameter classes (Figure 2). These results agree with those of Amateis et al. (2006), who showed that for a given set of stand conditions, stand-level basal area increased at northern locales throughout the natural range of loblolly pine. We postulate that these trends may be due to the ability for stands at northern latitudes to continue stem basal area growth later into the growing season after height growth has ceased. Contrasting trends were predicted across the range of...
longitude for differing management schemes: NIMPs displayed a shifted distribution toward larger diameter classes as one moved toward western locales, whereas IMPs displayed distributions slightly shifted toward larger diameter classes as one moved toward eastern locales. This shift is reflected in the Weibull scale ($\hat{b}$) parameter, as parameter estimates in the equations predict significant estimates with contrasting signs (i.e., a positive value for the LONG variable for IMPs and a negative value for NIMPs), which leads to differing results in the scale parameter across the range of longitude (Figure 2). We suspect that this contrast may be due to the dynamics between stand age and average stand diameter between IMPs and NIMPs. Because of the planting of improved genotypes and practices of fertilization and competition control, IMPs generally displayed a greater average stand diameter yet were slightly younger than their NIMP counterparts (Table 1). Although we used a host of stand variables to characterize the magnitude of the scale parameter, in other applications mean stand diameter has solely been used to characterize the scale parameter (Cao et al. 1982, Nord-Larsen and Cao 2006). Many studies have reported that an increase in stand age results in an increase in the magnitude of the scale parameter (Liu et al. 2004, Jiang and Brooks 2009), which we similarly observed for NIMPs but not for IMPs.

Stand-level (Amateis et al. 2006) and individual tree-level (Russell et al. 2010) loblolly pine height equations generally predicted opposite trends compared with diameter attributes: that is, loblolly pine heights were predicted to be greater at south and east locales in the region for a common set of stand conditions. The significance of site index and regional variables in Equation 10 indicates that traditional site productivity measures in combination with variables that capture the multitude of climatic and abiotic factors (i.e., latitude and longitude) can be used to quantify loblolly pine diameter distributions. Furthermore, regional trends were dependent on the level of management. Regional variation on the potential volume of loblolly pine has been reported (Hasenauer et al. 1994), and the degree to which these various stand- and tree-level attributes affect other growth and yield equation components, such as tree taper and biomass, should be evaluated. To more specifically address yield-climate interactions, tools such as the PRISM climate mapping system (PRISM Climate Group 2011) can be used to obtain climatic information. Such climate variables representing past and present climate conditions are generally used within empirical growth equations (e.g., Amateis and Burkhart 2008); however, growth models that incorporate future possible climatic scenarios are becoming available (Crookston et al. 2010). Parameters presented in Table 3

Figure 2. Predicted diameter distributions (Equations 9 and 10) for loblolly pine stands at east ($34^\circ, -77^\circ$), center ($34^\circ, -84^\circ$), and west ($34^\circ, -95^\circ$) locales for nonintensively (a) and intensively (b) managed plantations, and at south ($32^\circ, -84^\circ$), center ($34^\circ, -84^\circ$), and north ($36^\circ, -84^\circ$) locales for nonintensively (c) and intensively (d) managed plantations. (Stand assumes age of 15 years with 1,400 trees ha$^{-1}$, relative spacing of 0.2, site index 21 m.)

Stand levels (Amateis et al. 2006) and individual tree levels (Russell et al. 2010) loblolly pine height equations generally predicted opposite trends compared with diameter attributes: that is, loblolly pine heights were predicted to be greater at south and east locales in the region for a common set of stand conditions. The significance of site index and regional variables in Equation 10 indicates that traditional site productivity measures in combination with variables that capture the multitude of climatic and abiotic factors (i.e., latitude and longitude) can be used to quantify loblolly pine diameter distributions. Furthermore, regional trends were dependent on the level of management. Regional variation on the potential volume of loblolly pine has been reported (Hasenauer et al. 1994), and the degree to which these various stand- and tree-level attributes affect other growth and yield equation components, such as tree taper and biomass, should be evaluated. To more specifically address yield-climate interactions, tools such as the PRISM climate mapping system (PRISM Climate Group 2011) can be used to obtain climatic information. Such climate variables representing past and present climate conditions are generally used within empirical growth equations (e.g., Amateis and Burkhart 2008); however, growth models that incorporate future possible climatic scenarios are becoming available (Crookston et al. 2010). Parameters presented in Table 3
allow a user to predict the characteristics of diameter distributions using latitude and longitude as surrogates for more complex climatic factors for intensively or nonintensively managed loblolly pine plantations.

**Literature Cited**


