Comparing Parameter Estimation Techniques for Diameter Distributions of Loblolly Pine in a GxE Study

B.C. Smith, B.P. Bullock, and S.E. McKeand[1]

Diameter distributions play an important role in stand modeling. Frequency by size class can be estimated from a distribution function with estimated parameters. A number of different distribution functions have been utilized to model diameter distributions, including the Beta, Lognormal, Johnson's S_b, and Weibull. The Weibull function has been widely used due to its flexibility in modeling reverse-J, skewed, and unimodal shapes and because integration is not required to estimate frequencies. For these reasons the Weibull function was chosen for this study. Two of the main methods which have been used to estimate the parameters of the Weibull function are parameter prediction, in which each parameter is directly predicted from stand-level variables with a regression equation, and parameter recovery, where selected percentiles of the distribution are predicted and used to equate parameters to moments of the distribution. The purpose of this study was to determine the parameter estimation technique most appropriate for young plantation-grown loblolly pine, and to present parameter estimation equations incorporating necessary genotype and environmental information.

MATERIALS AND METHODS

The SETRES-2 study site is located in the Sandhills of NC in Scotland County on very well drained, infertile soil. Established in November 1993, the study is a split-split-plot design with two silvicultural treatments, fertilized and unfertilized, and two provenances (Atlantic Coastal Plain and Lost Pines Texas) with five open-pollinated loblolly pine families from each. The study is divided into nine blocks, with two main treatment plots each containing two provenance sub-plots, with five family sub-sub-plots nested within each provenance plot. 100 trees were planted in each family plot with a rectangular spacing of 1.5 m by 2.1 m, for a total of 18,000 trees. Several measurements, including height and diameter, were taken for every tree at ages 0, 1, 2, 3, 4, 5, 6, 8, and 10. Diameter measurements were taken for the interior 64 trees of each family plot at age 11. Due to the size of the trees relative to the buffer size between treatment plots inducing potential edge effects, only measurements from the interior 64 trees were used for this study.

Theoretical distributions were fitted to the empirical diameter distributions from ages five to 11 using the two parameter Weibull distribution, with the probability density function (p.d.f.) given in Equation (1). The fit method utilized was maximum likelihood estimation (MLE).

$$f(x) = \frac{c}{b}\left(\frac{x}{b}\right)^{c-1} \exp\left[-\left(\frac{x}{b}\right)^{c}\right] \quad \text{for } x \geq 0, b > 0, c > 0 \tag{1}$$

where b = scale parameter

c = shape parameter

x = stem diameter

[1] Graduate Student, Assistant Professor, and Professor, respectively, Department of Forestry and Environmental Resources, NC State University, Campus Box 8002, Raleigh, NC 27695-8002.

Parameter prediction and recovery models were constructed to estimate the parameters produced by the MLE distribution. Incorporation of various stand level variables as well as family and provenance indicator variables in the models were tested. Regression equations were fit simultaneously using seemingly unrelated regression due to correlated errors.

As Shiver (1988) pointed out, because there is no known underlying distribution function with specific parameters, what are important in modeling diameter distributions are not the individual predicted parameters but rather how well the combinations of predicted parameters reproduce the empirical distributions. Rather than checking the ability of each model to estimate individual parameters, it is desirable to test the fit of a model's distribution from the estimated parameters. In light of this, the goodness-of-fit of the MLE and the predicted distributions to the empirical distributions were evaluated with an error index proposed by Reynolds et al. (1987). The index is a weighted measure of how well a distribution predicts the number of trees in individual diameter classes. The error index, e.i., is given by:

$$e.i. = N\sum_{i=1}^{j} w\left(x_i\right)\left|\hat{F}\left(x_i\right) - F^*\left(x_i\right)\right|$$

where $\left(x_i\right)$ = the centerpoint of the ith d-class

$\hat{F}\left(x_i\right)$ = the proportion of distribution predicted in the ith d-class (2)

$F^*\left(x_i\right)$ = the proportion of distribution observed in the ith d-class

$w\left(x_i\right)$ = weight of x_i

N = trees per hectare

When $w\left(x_i\right)$ is set to basal area (m^2) of diameter x_i, the error index units are basal area per hectare predicted in the incorrect diameter class. If the multiplier N is dropped from the error index equation, the resulting index yields the percent of BA/ha which was predicted in the incorrect diameter class. For this study, two centimeter diameter classes were used.

RESULTS AND DISCUSSION

All of the final models tested performed similarly. The inclusion of a provenance indicator variable was not found to be significant in any of the models. The addition of certain family indicator variables slightly improved the fit of the parameter prediction model, but this would be of limited usefulness as the study families are random typical families. What is interesting to note is that the rank of the mean shape and scale parameters remains fairly stable over time. The percent of the basal area misclassified by the model decreased as age increased for all the models, indicating the accuracy of the fitted distributions are improving over time. The mean e.i. over all treatments, provenances, and ages was lower for the parameter prediction techniques than the parameter recovery techniques.

The recommended parameter prediction model is given in Equation (3), and the recommended parameter recovery model is given in Equation (4). The mean e.i. through age 10 of Equation (3) was 1.291 m^2/ha, with a minimum of 0.107 and maximum of 4.251 m^2/ha, compared to a mean e.i. for Equation (4) of 1.389 m^2/ha, minimum of 0.103 and maximum of 6.137 m^2/ha. For comparison, the mean e.i. of the MLE distributions was 1.075 m^2/ha, with a minimum of 0.061 and maximum of 4.446 m^2/ha. A direct comparison of the methods was not possible through age

11 due to the lack of dominant height data, but of the parameter recovery models Equation (4) still performed best at age 11.

$$\hat{b} = -0.01393 + 119.02733\left(\overline{\text{BA}}\right)^{\frac{1}{2}}$$
$$\hat{c} = 7.08377 + 4.90044\left(\text{Fertilizer}\right) - 7.97094\left(\text{RS}\right) - 0.34313\left(\text{Fertilizer}\right)\left(\text{Age}\right) \tag{3}$$

where,

$\hat{b}$ = the estimated Weibull scale parameter

$\hat{c}$ = the estimated Weibull shape parameter

$\overline{\text{BA}}$ = mean basal area, m^2

Fertilizer = fertilization indicator variable, 1 if fertilized, else 0

RS = relative spacing, $\dfrac{\sqrt{10{,}000/\left(\text{trees per ha}\right)}}{\text{mean dominant height, m}}$

$$\ln\left(\hat{D}_{25}\right) = 5.89988 + 0.68360\ \ln\left(\overline{\text{BA}}\right) - 0.20221\ \ln\left(\text{Age}\right)$$
$$\ln\left(\hat{D}_{95}\right) = 3.87163 + 0.35889\ \ln\left(\overline{\text{BA}}\right) + 0.19035\ \ln\left(\text{Age}\right)$$
$$\hat{b} = \sqrt{\frac{\left(\overline{D_q}\right)^2}{\Gamma\left(1+2/\hat{c}\right)}} \qquad \hat{c} = \frac{2.28823}{\ln\left(\hat{D}_{95}\right) - \ln\left(\hat{D}_{25}\right)} \tag{4}$$

where,

$\hat{b}$ = the estimated Weibull scale parameter ; $\hat{c}$ = the estimated Weibull shape parameter

$\overline{\text{BA}}$ = mean basal area in m^2; Age = age of stand, years

$\overline{D_q}$ = quadratic mean diameter in cm; $\Gamma(\cdot)$ = the gamma function

D_i = the i^{th} percentile of the diameter distribution in m, for i = 25, 95

Due to the use of different stand level variables as inputs in the models, Equation (3) is recommended for parameter estimation if dominant height data are available. Otherwise, Equation (4) should be used.

REFERENCES

Reynolds, M.R. Jr., T.E. Burk and W.C. Huang. 1988. Goodness-of-fit tests and model selection procedures for diameter distribution models. For. Sci. 34(2):373-399.

Shiver, B.D. 1988. Sample sizes and estimation methods for the Weibull distribution for unthinned slash pine plantation diameter distributions. For. Sci. 34(3):809-814.