USE OF THE WEIBULL FUNCTION TO PREDICT FUTURE DIAMETER DISTRIBUTIONS FROM CURRENT PLOT DATA

Quang V. Cao

ABSTRACT

The Weibull function has been widely used to characterize diameter distributions in forest stands. The future diameter distribution of a forest stand can be predicted by use of a Weibull probability density function from current inventory data for that stand. The parameter recovery approach has been used to "recover" the Weibull parameters from diameter moments or percentiles. The Moment method involves arithmetic or quadratic mean diameter, and diameter variance, whereas the Percentile method includes diameter percentiles. The Hybrid method is a combination of both methods, requiring both diameter moments and percentiles. Results based on data from loblolly pine plantations showed that the three methods involving the predicted quadratic mean diameter performed better than the rest, and that the two methods involving the predicted 31st and 63rd percentiles performed the poorest.

INTRODUCTION

The Weibull function was introduced by Bailey and Dell (1973) to model diameter distributions in forest stands. It has since become popular because it is flexible enough to fit shapes commonly found in both uneven-aged and even-aged stands, and also because the calculation of proportions of trees in diameter classes is straightforward. The parameter recovery approach (Hyink and Moser 1983) has been found to perform better than the parameter prediction approach, in which the Weibull parameters are predicted directly. In the parameter recovery approach, the Weibull parameters are "recovered" from diameter moments (arithmetic and quadratic diameters, and diameter variance), diameter percentiles (e.g. 25th, 50th, 31st, 63rd, or 95th), or a combination of both.

The objective of this study was to evaluate ten parameter recovery methods to predict the parameters of Weibull functions that modeled diameter distributions of a future stand. The Weibull parameters were recovered from future stand attributes, which were predicted from current stand attributes by use of regression.

DATA

Data were from the Southwide Seed Source Study, which involved 15 loblolly pine (*Pinus taeda* L.) seed sources planted at 13 locations across 10 southern states (Wells and Wakeley 1966). Seedlings were planted at a 6 ft x 6 ft spacing. Each plot of size 0.04 acre consisted of 49 trees measured four times at ages 10, 15 or 16, 20 or 22, and 25 or 27. A subset (100 plots) of the original data was randomly selected as the fit data set, to be used for fitting the models. Furthermore, only one growing period was randomly chosen from each plot. The fit data set therefore contained growth periods from age 10 to age 15 (33 plots), from age 15 to age 20 (33 plots), and from age 20 to age 25 (34 plots). Another 100 plots were randomly selected from the remaining original data in the same manner to form a validation data set. Table 1 shows summary statistics for stand attributes at the end of each growth period for the fit and validation data sets.

METHODS

The Weibull probability density function (pdf), used in this study to characterize diameter distribution, has the following form:

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

where a, b, and c are the location, scale, and shape parameters, respectively, and x is tree diameter at breast height.

PARAMETER RECOVERY METHODS

The Weibull location parameter (a) must be smaller than the predicted minimum diameter in the stand ($\hat{D}_0$). We set $a = 0.5\,\hat{D}_0$ since Frazier (1981) found that this gave best results in terms of goodness-of-fit. The other Weibull parameters,

Quang V. Cao, Professor, School of Renewable Natural Resources, Louisiana State University Agricultural Center, Baton Rouge, LA 70803

b and c, were recovered from the moments of the diameter distribution (Moment method), the diameter percentiles (Percentile method), or a combination of both (Hybrid method). The following parameter recovery methods were evaluated:

Moment methods

- Method 1 ($\hat{\bar{D}}$ and $\hat{D}\mathrm{var}$)
- Method 2 ($\hat{D}q$ and $\hat{D}\mathrm{var}$)

Percentile methods

- Method 3 ($\hat{D}_{31}$ and $\hat{D}_{63}$)
- Method 4 ($\hat{D}_{50}$ and $\hat{D}_{95}$)
- Method 5 ($\hat{D}_{25}$, $\hat{D}_{50}$, and $\hat{D}_{95}$)
- Method 6 ($\hat{D}_{31}$, $\hat{D}_{50}$, and $\hat{D}_{63}$)

Hybrid methods

- Method 7 ($\hat{\bar{D}}$ and $\hat{D}_{95}$)
- Method 8 ($\hat{D}q$ and $\hat{D}_{95}$)
- Method 9 ($\hat{D}q$, $\hat{D}_{25}$, and $\hat{D}_{95}$)
- Method 10 ($\hat{D}q$, $\hat{D}_{25}$, $\hat{D}_{50}$, and $\hat{D}_{95}$)

The symbols $\hat{\bar{D}}$, $\hat{D}q$, $\hat{D}\mathrm{var}$, $\hat{D}_{25}$, $\hat{D}_{31}$, $\hat{D}_{50}$, $\hat{D}_{63}$, and $\hat{D}_{95}$ denote predicted values of average diameter, quadratic mean diameter, diameter variance, and the 25th, 31st, 50th, 63rd, and 95th diameter percentiles, respectively. In method 10 (Bailey et al. 1989), the a parameter was computed from

$$a = \frac{\hat{D}_0 n^{1/3} - \hat{D}_{50}}{n^{1/3} - 1}$$

where n is number of trees in the plot. Systems of equations for the ten methods are shown in Table 2.

EVALUATION

The error index (Reynolds et al. 1988), used to evaluate how well each method performed for the validation data, is defined as:

$$EI = \frac{1}{m} \sum_i \sum_k \left| n_{ik} - \hat{n}_{ik} \right|$$

where n_{ik} and $\hat{n}_k$ are, respectively, the observed and predicted number of trees/ha in diameter class k for the i^{th} plot, and m is the number of plots. The smaller the error index, the better the distribution fits the data.

RESULTS AND DISCUSSION

The future stand survival was predicted from:

$$\hat{N} = N_0 \left(\frac{A}{A_0}\right)^{1.30125-1} \exp\left[-0.01958\left(A^{1.30125} - A_0^{1.30125}\right)\right]$$

; $R^2 = 0.5775$, where $\hat{N}$ is the predicted future stand survival, N_0 is the current number of trees/ha, A and A_0 are future and current stand age, respectively.

The diameter moments and percentiles were simultaneously predicted by use of seemingly unrelated regression (SUR). The equations used were of the following general form:

$$\hat{y} = y_0\{1 + \exp[b_1 + b_2 RS_0 + b_3 \ln(Hd_0) + b_4 \ln(N_0)]\}$$

where $\hat{y}$ is the predicted future moment/percentile, y_0 is the current moment/percentile, $RS_0 = [(10000/N_0)^{0.5}]/Hd_0$ is the current relative spacing, and Hd_0 is the current dominant height. The parameter estimates obtained from the fit data are presented in Table 3.

Table 4 shows the error index computed for each method from the validation data. The methods formed three groups based on their error index results, with some overlaps in between. The first group (methods 2, 9, and 10) produced the best results by scoring the lowest error indices. This group consisted of three methods, all of which involved $\hat{D}q$. These results showed that $\hat{D}q$ seemed to be a better central measure of the Weibull distribution than either $\hat{\bar{D}}$ or $\hat{D}_{50}$. An exception was method 8 ($\hat{D}q$ and $\hat{D}_{95}$), which was ranked 8th among ten methods.

The lowest-ranked group (methods 3 and 6) produced the highest values of error index, i.e. poorest fit. Both of these methods involved $\hat{D}_{31}$ and $\hat{D}_{63}$, suggesting that these two percentiles should not be used in recovering the Weibull parameters.

CONCLUSIONS

The analyses shown in this study revealed that the predicted quadratic mean diameter played an important role in recovering parameters of the Weibull that characterized the future diameter distribution of loblolly pine plantations. On the other hand, both methods involving the predicted 31st and 63rd percentiles performed the poorest among all methods. Method 2 ($\hat{D}q$ and $\hat{D}\mathrm{var}$) produced the lowest error index and should be considered as a parameter recovery method for other data sets.

LITERATURE CITED

Bailey, R.L., Dell, T.R. 1973. Quantifying diameter distributions with the Weibull function. Forest Science 19:97-104.

Bailey, R.L., Burgan, T.M., Jokela, E.J. 1989. Fertilized midrotation-aged slash pine plantations--stand structure and yield prediction models. Southern Journal of Applied Forestry 13:76-80.

Frazier, J.R. 1981. Compatible whole-stand and diameter distribution models for loblolly pine plantations. Ph.D. Diss., Va. Polytech. Inst. & State Univ., Blackburg, VA. 125 p.

Hyink, D.M., Moser, J.W. 1983. A generalized framework for projecting forest yield and stand structure using diameter distributions. Forest Science 29:85-95.

Reynolds, M.R., Burk, T.E., Huang, W.C. 1988. Goodness-of-fit tests and model selection procedures for diameter distribution models. Forest Science 34: 373-399.

Wells, O.O., Wakeley, P.C. 1966. Geographic variation in survival, growth, and fusiform-rust infection of planted loblolly pine. Forest Science Monograph 11. 40 p.

Table 1—Means (and standard deviations) of stand attributes at the end of each growth period for the fit and validation data

Age	Number of plots	Dominant height (m)	Number of trees per ha	Basal area (m^2/ha)
		Fit Data		
15	34	12.7 (1.8)	1843 (605)	28.9 (7.4)
20	33	16.4 (1.4)	1291 (250)	34.3 (8.4)
25	33	18.5 (2.0)	1346 (263)	40.1 (8.1)
		Validation Data		
15	34	13.1 (1.7)	1742 (579)	30.0 (6.3)
20	33	16.3 (2.2)	1305 (254)	34.0 (6.9)
25	33	18.9 (2.0)	1350 (255)	41.1 (6.2)

Table 2—Summary of ten parameter recovery methods

Method	Equation for a	Equation for b and c [1/]
	Moment methods	
Method 1 ($\hat{\bar{D}}$ and $\hat{D}\,\mathrm{var}$)	$a = 0.5\,\hat{D}_0$	$b = (\hat{\bar{D}} - a)/G_1$ c is obtained from $b^2(G_2 - G_1^2) - \hat{D}\,\mathrm{var} = 0$
Method 2 ($\hat{D}q$ and $\hat{D}\,\mathrm{var}$)	$a = 0.5\,\hat{D}_0$	$b = \frac{-aG_1}{G_2} + \sqrt{\left(\frac{a}{G_2}\right)^2 (G_1^2 - G_2) + \frac{\hat{D}q^2}{G_2}}$ c is obtained from $b^2(G_2 - G_1^2) - \hat{D}\,\mathrm{var} = 0$
	Percentile methods	
Method 3 ($\hat{D}_{31}$ and $\hat{D}_{63}$)	$a = 0.5\,\hat{D}_0$	$c = \frac{\ln\left(\frac{\ln(1-0.63)}{\ln(1-0.31)}\right)}{\ln(\hat{D}_{63} - a) - \ln(\hat{D}_{31} - a)}$ $b = \frac{\hat{D}_{63} - a}{[-\ln(1-0.63)]^{1/c}}$
Method 4 ($\hat{D}_{50}$ and $\hat{D}_{95}$)	$a = 0.5\,\hat{D}_0$	$c = \frac{\ln\left(\frac{\ln(1-0.95)}{\ln(1-0.50)}\right)}{\ln(\hat{D}_{95} - a) - \ln(\hat{D}_{50} - a)}$ $b = \frac{\hat{D}_{50} - a}{[-\ln(1-0.50)]^{1/c}}$
Method 5 ($\hat{D}_{25}$, $\hat{D}_{50}$, and $\hat{D}_{95}$)	$a = 0.5\,\hat{D}_0$	$c = \frac{\ln\left(\frac{\ln(1-0.95)}{\ln(1-0.25)}\right)}{\ln(\hat{D}_{95} - a) - \ln(\hat{D}_{25} - a)}$ $b = \frac{\hat{D}_{50} - a}{[-\ln(1-0.50)]^{1/c}}$
Method 6 ($\hat{D}_{31}$, $\hat{D}_{50}$, and $\hat{D}_{63}$)	$a = 0.5\,\hat{D}_0$	$c = \frac{\ln\left(\frac{\ln(1-0.63)}{\ln(1-0.31)}\right)}{\ln(\hat{D}_{63} - a) - \ln(\hat{D}_{31} - a)}$ $b = \frac{\hat{D}_{50} - a}{[-\ln(1-0.50)]^{1/c}}$

Table 2—(Continued) Summary of ten parameter recovery methods

Method	Equation for *a*	Equation for *b* and *c*
	Hybrid methods	
Method 7 ($\hat{\bar{D}}$ and $\hat{D}_{95}$)	$a = 0.5\,\hat{D}_0$	$b = \dfrac{\hat{D}_{95} - a}{[-\ln(1-0.95)]^{1/c}}$ c is obtained from $a + bG_1 - \hat{\bar{D}} = 0$
Method 8 ($\hat{D}q$ and $\hat{D}_{95}$)	$a = 0.5\,\hat{D}_0$	$b = \dfrac{\hat{D}_{95} - a}{[-\ln(1-0.95)]^{1/c}}$ c is obtained from $b^2G_2 + 2abG_1 + a^2 - \hat{D}q^2 = 0$
Method 9 ($\hat{D}q$, $\hat{D}_{25}$, and $\hat{D}_{95}$)	$a = 0.5\,\hat{D}_0$	$c = \dfrac{\ln\left(\dfrac{\ln(1-0.95)}{\ln(1-0.25)}\right)}{\ln(\hat{D}_{95} - a) - \ln(\hat{D}_{25} - a)}$ $b = -aG_1/G_2 + [(a/G_2)^2(G_1^2 - G_2) + \hat{D}q^2/G_2]^{0.5}$
Method 10 ($\hat{D}q$, $\hat{D}_{25}$, $\hat{D}_{50}$, and $\hat{D}_{95}$)	$a = \dfrac{\hat{D}_0 n^{1/3} - \hat{D}_{50}}{n^{1/3} - 1}$	$c = \dfrac{\ln\left(\dfrac{\ln(1-0.95)}{\ln(1-0.25)}\right)}{\ln(\hat{D}_{95} - a) - \ln(\hat{D}_{25} - a)}$ $b = -aG_1/G_2 + [(a/G_2)^2(G_1^2 - G_2) + \hat{D}q^2/G_2]^{0.5}$

[1/] $G_k = \Gamma(1 + k/c)$, where $\Gamma(\cdot)$ is the gamma function.

Table 3—Parameter estimates for predicting future diameter moments and percentiles

Variable	b_1	b_2	b_3	b_4	R^2
$\hat{D}_0$	-2.29785	4.43916			0.3583
$\hat{\bar{D}}$	3.13645		-1.29982	-0.21411	0.9242
$\hat{D}\mathrm{var}$	2.50950		-1.32765		0.6975
$\hat{D}q$	3.14000		-1.30057	-0.21404	0.9309
$\hat{D}_{25}$	1.08057		-1.16174		0.8554
$\hat{D}_{31}$	1.54103		-1.35951		0.8740
$\hat{D}_{50}$	4.85622		-1.46683	-0.39758	0.9041
$\hat{D}_{63}$	1.30993		-1.19434		0.8862
$\hat{D}_{95}$	3.65865		-1.15230	-0.31641	0.9179

Equation: $\hat{y} = y_0\{1 + \exp[b_1 + b_2 RS_0 + b_3 \ln(Hd_0) + b_4 \ln(N_0)]\}$

Table 4—Error index for each method from the validation data

Rank	Method	Moments/percentiles	EI [1/]
1	2	$\hat{D}q$ and $\hat{D}\mathrm{var}$	663 a
2	9	$\hat{D}q$, $\hat{D}_{25}$, and $\hat{D}_{95}$	670 a
3	10	$\hat{D}q$, $\hat{D}_{25}$, $\hat{D}_{50}$, and $\hat{D}_{95}$	674 a
4	4	$\hat{D}_{50}$, and $\hat{D}_{95}$	720 b
5	1	$\hat{\bar{D}}$ and $\hat{D}\mathrm{var}$	726 bc
6	5	$\hat{D}_{25}$, $\hat{D}_{50}$, and $\hat{D}_{95}$	727 bc
7	7	$\hat{\bar{D}}$ and $\hat{D}_{95}$	731 bcd
8	8	$\hat{D}q$ and $\hat{D}_{95}$	743 bcd
9	6	$\hat{D}_{31}$, $\hat{D}_{50}$, and $\hat{D}_{63}$	755 cd
10	3	$\hat{D}_{31}$ and $\hat{D}_{63}$	760 d

[1/] Values with the same letter are not different at the 5% level.