Reliability of Tanoak Volume Equations When Applied To Different Areas

Norman H. Pillsbury, *Natural Resources Management Department, California Polytechnic State University, San Luis Obispo, California 93407;* **Philip M. McDonald,** *Pacific Southwest Forest and Range Experiment Station, USDA Forest Service, Redding, CA 96001; and* **Victor Simon,** *Statistics Department, California Polytechnic State University, San Luis Obispo, California 93407.*

ABSTRACT. *Tree volume equations for tanoak* (Lithocarpus densiflorus) *were developed for seven stands throughout its natural range and compared by a volume prediction and a parameter difference method. The objective was to test if volume estimates from a species growing in a local, relatively uniform habitat could be applied more widely. Results indicated that they could not. Diameter slope coefficients and intercepts differed at the 5% probability level. Tree height slope coefficients did not differ, but when all slope coefficients (dbh and height) were combined, a significant difference was present. Overall, volume equations from different populations were significantly different, and the general rule of matching standard volume tables to specific stands is reinforced. West. J. Appl. For. 10 (2):72–78.*

Foresters are often reminded to apply a tree volume table to the geographic range where the data were collected. The farther the population of trees from the sample collection area, the less accurate the equation.

In practice, convenience and economics often play a role in the decision to use or not use a volume table in a different area. What if equations do not exist for the area where tree measurements are needed? A substitute, often the only equation available, is used.

How reliable is the next best equation? Would a substitute equation show a significant difference between the predicted volume and the cut volume? In this article we will examine these questions based on data for tanoak *(Lithocarpus densiflorus)* in California and southern Oregon. Tanoak (Figure 1) was selected for study for two reasons: (1) we noticed that tanoak trees (not the shrub form) grew on good sites with deep soils and abundant precipitation. They tended to form even-aged stands, and were codominant in crown position with trees that were uniform in height. This uniformity suggested the possibility that, at least over most of its range, tree height and volume would be similar, and a volume table for tanoak would have broad application (McDonald and Tappeiner 1988), and (2) volume data had previously been measured at seven different locations throughout its range and were available for analysis.

Study Area and Methods

Tanoak is found from southwestern Oregon to central California (Burns and Honkala 1990). It occurs on more than 2.5 million ac in California. The estimated volume of live, sound tanoak trees 5.0 in. in dbh and larger is 1.9 billion ft^3 (Bolsinger 1988). In general, tanoak is a reliable indicator of productive sites in southwest Oregon (Atzet et al. 1982).

The location for the data sets used in this study is shown in Figure 2. These locations are representative of the range of tanoak. Each data set was assumed to be representative of the diameter, height, and volume characteristics of trees at that site. This was known to be true for two of the studies in which we were involved and appears to be the case for the rest of the studies.

Before data sets could be compared, they were converted to the same units and the same utilization standard. Differences in utilization standards were corrected as summarized in Table 1, so the analysis would be based on diameter at breast height, total height (0" top), and total volume inside bark (ib) to a 0" top excluding stump calculated by Smalian's cubic volume formula, Volume (cubic units) = $\Sigma[(h/2)(A_b + A_u)]$, where h is the segment length, A is the basal area at the small (b) and large (u) end of the log for cut trees.

Figure 1. A typical stand of tanoak *(Lithocarpus densiforus)* located at Challenge Experimental Forest, California.

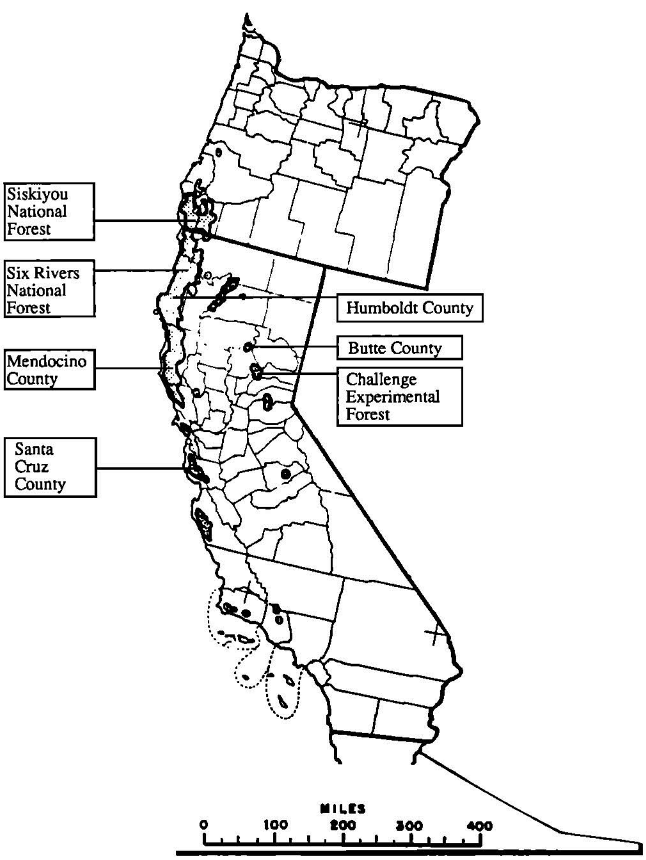

Figure 2. The range of tanoak *(Lithocarpus densiforus)* and the location of the sites used in this study.

Variables measured at each location included diameter, height, and cut volume. A tree volume equation was calculated for each location.

Two methods of analysis were used.

(1) *Volume prediction method:* First, each volume equation was applied to the data from the other locations. This allowed us to take a "practical" look at the differences between predicted and cut volumes and judge the importance these differences would have on a field survey.

(2) *Parameter difference method:* Secondly, we tested the volume equation coefficients for each of the seven independent samples of tanoak, to determine if the coefficients were statistically different. The hypothesis tested was that no significant difference existed among tree volume equations at the 5% confidence level.

Volume Equation Development

The model used to develop the volume equation is nonlinear of the form

$$y = b_0 x_1^{b1} x_2^{b2} e \qquad (1)$$

which is intrinsically linear, meaning that it can be expressed as a multiple linear regression model by using a suitable transformation (Montgomery and Peck 1982, Cunia 1964). Applying a Log_{10} transformation to both sides of the equation gives

Table 1 Summary of data sets available for study

Population location	Variables measured	Adjustments in data set[1]
1. Butte County	Total height Volume inside bark to 4" top Sample size = 42	None Volume corrected to 0" top
2. Challenge Experimental Forest	Total height Total volume (ib) to 0" top Sample size = 143	None None
3. Humboldt County	Total height Volume inside bark to 4" top Sample size = 120	None Volume corrected to 0" top
4. Mendocino County	Total height Volume inside bark to 4" top Sample size = 38	None Volume corrected to 0" top
5. Santa Cruz County	Total height Total volume (ib) to 0" top Sample size = 20	None None
6. Siskiyou National Forest	Height to 4" top Volume inside bark to 4" top Sample size = 43	Height corrected to 0" top Volume corrected to 0" top
7. Six Rivers National Forest	Total height Total volume (ib) to 0" top Sample size = 30	None None

1 Volumes were adjusted by applying a regression equation developed from the Challenge data set where both total volume to a 0" top and volume to a 4" top were known. The cut volume data for four locations were adjusted from a 4" top to a 0" top. Note that the volume which occurs above the 4" diameter point is very small (less than 1 ft^3). A similar process was used to adjust merchantable tree height to total tree height for the Siskiyou data set.

$$\text{Log } y = \text{Log } b_0 + b_1 \text{ Log } x_1 + b_2 \text{ Log } x_2 + \text{Log } e \quad (2)$$

or

$$y' = b_0' + b_1 x_1' + b_2 x_2' + e' \quad (3)$$

where

y' = Log of total volume
x_1' = Log of diameter at breast height
x_2' = Log of total height
b_i = Regression coefficients

and e' = Log of e (error term) and is assumed to be normally and independently distributed with mean = 0 and variance = s^2.

The equation may be visualized as a flat plane sloping upward in three-dimensional space (Figure 3 in nonlogarthmic form). b_0' is the relative "elevation" of the plane at the point where the log of diameter at breast height and that of total height equal one.

Transformed equations were computed using a least squares regression program. Estimates of volume equation coefficients b_0 (antilog b_0'), b_1, and b_2, for the original model were then obtained for each sample.

Data point outliers were detected by use of an approximate t-test (Lund 1975). We tested the null hypothesis of each sample containing no outliers, versus the alternative hypothesis of the observation with the largest studentized residual being an outlier. If the residual was larger than the appropriate critical value (significance level of 0.05), the corresponding observation was considered to be an outlier and was discarded from the sample. The equation was computed

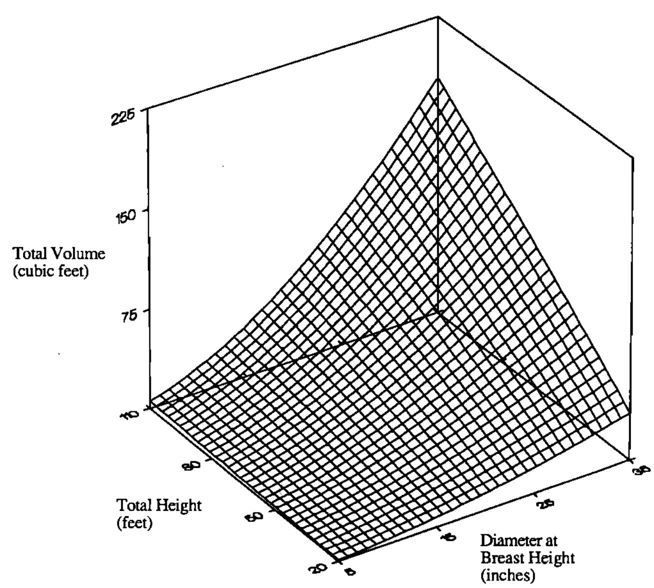

Figure 3. Response surface for tree volume equation. This figure shows the relationship for tanoak *(Lithocarpus densiflorus)* at Challenge Experimental Forest California. Vol^3 = 0.0034006 *(dbh)*$^{1.79980}$ *(Total Height)*$^{0.95124}$.

again without the outlier and studentized residuals, and again compared with the critical value, until all residuals passed the test of no outliers. Based on this analysis a total of four data points were identified as outliers and dropped from the database (436 points retained). The final volume equation coefficients are shown in Table 2.

Comparison of Volume Equations from Different Areas

How reliable would these volume equations be if used in a different area? One method to determine the reliability of each equation was to estimate the total volume of the sample trees at each of the other six locations and compare it to the measured sample volume from that location. Earlier this analysis was referred to as the volume prediction method. The results shown in Table 3 are intended to be compared in a relative manner and not as absolute values.

Results of the Volume Prediction Analysis

We expected that the volume equation developed for each location would be the "best" predictor for that sample; however, this was not true in all locations, most notably the Humboldt sample where both the Butte and Siskiyou prediction equations did better. The variation among the predictions range from 0.5 to 39%. Although equations that predict volume within 10 to 15% or less are desirable, the actual acceptance level is subjective and depends on many factors. Based on experience we selected 12% as the acceptable cutoff point for this study. Of the 42 possible predictions in this analysis, 22 fell within 12% of the sample volume. In general terms, the tree volume models were able to predict the sample volume at the other 6 locations about one-half of the time.

Some bias is introduced because sample sizes are not equal and the volume distribution is not the same among locations. This also assumes that equal variance exists which is unlikely. This is always a risk when using the "next best" equation, but it is unlikely that such bias could ever be avoided.

Also, it should be noted that sample trees measured at each location were not in plots and not randomly selected; rather the trees were subjectively picked to represent all dbh and volume classes expected to occur where the equation would be applied. A random sample of trees from each location would provide a better comparison although substantial differences will also occur from stand-to-stand if the diameter distribution is different. However, the data sets that we used provide a reasonable approximation of how well equations developed for other areas would perform.

Analysis of Volume Equation Coefficients

A second method of analysis is the parameter difference method. This approach is designed to test the volume equation coefficients (shown in Table 2 as b_0, b_1, b_2) to determine if they are statistically different among sample locations. Possible differences among location equations were determined by fitting a "full" linear regression model (combined data from all locations) and a "reduced" linear regression model (designed to test specific coefficients in location equations).

To test for equality of coefficients among equations, qualitative variables (variables with no natural scale or measurement) were used to account for the effect of different locations on the response variable, which was logarithm of total volume. Each qualitative variable was assigned a set of levels corresponding to different locations by using an indicator variable that could only take the values 0 and 1.

The transformed volume equations (Equations 4–10) can be visualized as regression planes in three-dimensional space (Figure 3). The analysis tests (1) if the planes are the same, that is, do b_0, b_1, b_2 have the same values for all equations; (2) if the planes have the same slopes (b_1, and b_2 are equal), meaning they are parallel; and (3) if the planes have a common intercept b_0, or the same relative "elevation."

From the volume equations (Table 2) there was no reason to expect that either intercepts or slopes would be equal among equations. Therefore, a set of values was assigned to each of the coefficients b_0, b_1, and b_2 as follows:

Table 2. Volume equation coefficients for tanoak data sets from seven locations in California and southern Oregon.

	Volume equation coefficients						
Location	Intercept (b_0)	dbh (b_1)	Total Ht (b_2)	N	$\%R^2$	SE	Equation no.
1 Butte County	0.0094282	1.78921	0.75325	42	94.9	1.14	(4)
2 Challenge Experimental Forest	0.0034006	1.79980	0.95124	143	99.0	1.15	(5)
3 Humboldt County	0.0027677	1.91609	0.93641	120	99.6	1.07	(6)
4 Mendocino County	0.0020745	1.91820	0.99333	38	99.7	1.04	(7)
5 Santa Cruz County	0.0047481	1.93389	0.81907	20	98.9	1.10	(8)
6 Siskiyou National Forest	0.0011316	1.81368	1.21613	43	99.3	1.09	(9)
7 Six Rivers National Forest	0.0024962	1.71686	1.14367	30	99.0	1.19	(10)

KEY: Total tree volume (V) inside bark excluding stump in cubic feet is predicted by these coefficients of the form: $V = b_0 D^{b1} H^{b2}$ where D is diameter at breast height in inches, and H is total height in feet. N is the sample size, $\%R^2$ is the proportion of variation in total volume explained by the model, SE is the standard error of the estimate in cubic feet.

Table 3 Comparison of tanoak volume equations to sample volumes from different locations in California.

Butte

Rank	Location	Sample Volume	Percent Deviation
BUTTE VOLUME:		**762**	0.0
1	Butte	749	- 1.7
2	***Santa Cruz***	724	- 5.0
3	***Six Rivers***	809	6.2
4	Humboldt	650	- 14.7
5	Siskiyou	636	- 16.5
6	Challenge	624	- 18.1
7	Mendocino	619	- 18.8

Challenge

Rank	Location	Sample Volume	Percent Deviation
CHALLENGE VOLUME:		**3091**	0.0
1	Challenge	3139	1.6
2	***Mendocino***	3240	4.8
3	***Humboldt***	3330	7.7
4	Butte	3501	13.3
5	Siskiyou	3554	15.0
6	Santa Cruz	3558	15.1
7	Six Rivers	4298	39.0

Humboldt

Rank	Location	Sample Volume	Percent Deviation
HUMBOLDT VOLUME:		**4440**	0.0
1	***Butte***	4378	- 1.4
2	***Siskiyou***	4576	3.1
3	Humboldt	4299	- 3.2
4	***Santa Cruz***	4580	3.2
5	***Mendocino***	4196	- 5.5
6	***Challenge***	3944	- 11.2
7	Six Rivers	5424	22.2

Santa Cruz

Rank	Location	Sample Volume	Percent Deviation
SANTA CRUZ VOLUME:		**563**	0.0
1	Santa Cruz	565	0.4
2	***Butte***	566	0.5
3	***Siskiyou***	584	3.7
4	***Humboldt***	533	- 5.3
5	***Mendocino***	519	- 7.8
6	***Challenge***	512	- 9.1
7	Six Rivers	715	27.0

Siskiyou

Rank	Location	Sample Volume	Percent Deviation
SISKIYOU VOLUME:		**3321**	0.0
1	***Santa Cruz***	3291	- 0.9
2	Siskiyou	3269	- 1.6
3	***Humboldt***	3098	- 6.7
4	***Mendocino***	3034	- 8.6
5	***Butte***	3029	- 8.8
6	Six Rivers	3782	13.9
7	Challenge	2794	- 15.9

Six Rivers

Rank	Location	Sample Volume	Percent Deviation
SIX RIVERS VOLUME:		**1723**	0.0
1	Six Rivers	1645	- 4.5
2	Santa Cruz	1467	- 14.9
3	Siskiyou	1402	- 18.6
4	Butte	1374	- 20.3
5	Humboldt	1364	- 20.8
6	Mendocino	1327	- 23.0
7	Challenge	1237	- 28.2

Mendocino

Rank	Location	Sample Volume	Percent Deviation
MENDOCINO VOLUME:		**738**	0.0
1	Mendocino	732	- 0.8
2	***Challenge***	729	- 1.2
3	***Humboldt***	757	2.6
4	***Siskiyou***	799	8.3
5	***Santa Cruz***	817	10.7
6	Butte	832	12.7
7	Six Rivers	994	34.7

Explanation: Location (Siskiyou), cut volume (3321 cu. ft.), and % variance from sample volume ------>

Rank 1-7 show the predicted volume for this location (Siskiyou) based on equations 5-11 (Table 2).

Location names in bold italics predicted the sample volume within 12 percent.

Siskiyou

Rank	Location	Sample Volume	Percent Deviation
SISKIYOU VOLUME:		**3321**	0.0
1	***Santa Cruz***	3291	- 0.9
2	Siskiyou	3269	- 1.6
3	***Humboldt***	3098	- 6.7
4	***Mendocino***	3034	- 8.6
5	***Butte***	3029	- 8.8
6	Six Rivers	3782	13.9
7	Challenge	2794	- 15.9

Percent Deviation is the percent deviation from the total volume of the sample data.

y' = response variable: the logarithm of total volume
x_1' = logarithm of diameter at breast height
x_2' = logarithm of total height

The full model tested was

$$\begin{aligned} y' = {} & b_0' + b_1x_1' + b_2x_2' + [\, b_3x_3 + b_4x_4 + b_5x_5 + b_6x_6 + b_7x_7 \\ & + b_8x_8] + [\, b_9x_1'x_3 + b_{10}x_1'x_4 + b_{11}x_1'x_5 + b_{12}x_1'x_6 + b_{13}x_1'x_7 \\ & + b_{14}x_1'x_8] + [\, b_{15}x_2'x_3 + b_{16}x_2'x_4 + b_{17}x_2'x_5 + b_{18}x_2'x_6 \\ & + b_{19}x_2'x_7 + b_{20}x_2'x_8] + e \end{aligned} \quad [11]$$

where x_3 through x_8 are indicator variables as follows:

x_3 = 1, if observation is from Humboldt location; 0, if not
x_4 = 1, if Butte location; 0, if not
x_5 = 1, if Mendocino location; 0, if not
x_6 = 1, if Santa Cruz location; 0, if not
x_7 = 1, if Siskiyou location; 0, if not
x_8 = 1, if Six Rivers location; 0, if not

and e (error term) is assumed to be normally and independently distributed with mean 0 and variance s^2.

The reduced models are based on each location. For the Challenge location the variables x_3 through $x_8 = 0$. The reduced model becomes $y' = b_0' + b_1x_1' + b_2x_2' + e$. For the Humboldt data the reduced model is

$$\begin{aligned} y' &= b_0' + b_1x_1' + b_2x_2' + b_3(1) + b_9x_1'(1) + b_{15}x_2'(1) + e \\ &= (b_0' + b_3) + (b_1 + b_9)x_1' + (b_2 + b_{15})x_2' + e \end{aligned}$$

since $x_3 = 1$.

b_3 represents the change in intercept (b_0') associated with changing from the Challenge to the Humboldt model. b_9 is the difference in dbh slope (b_1) between Humboldt and Challenge. Likewise, b_{15} is the change in height slope (b_2). All other location models (Table 4) are similarly defined using the appropriate indicator variables.

Note that the independent variables in the full model [Equation (11)] are grouped by brackets corresponding to

Table 4. Summary of location models (reduced models) used to test for equality of coefficients among locations

Location	Model	Equation no.
Challenge	$y' = b_0' + b_1x_1' + b_2x_2' + e$	[12]
Humboldt	$y' = (b_0' + b_3) + (b_1 + b_9)x_1' + (b_2 + b_{15})x_2' + e$	[13]
Butte	$y' = (b_0' + b_4) + (b_1 + b_{10})x_1' + (b_2 + b_{16})x_2' + e$	[14]
Mendocino	$y' = (b_0' + b_5) + (b_1 + b_{11})x_1' + (b_2 + b_{17})x_2' + e$	[15]
Santa Cruz	$y' = (b_0' + b_6) + (b_1 + b_{12})x_1' + (b_2 + b_{18})x_2' + e$	[16]
Siskiyou	$y' = (b_0' + b_7) + (b_1 + b_{13})x_1' + (b_2 + b_{19})x_2' + e$	[17]
Six Rivers	$y' = (b_0' + b_8) + (b_1 + b_{14})x_1' + (b_2 + b_{20})x_2' + e$	[18]

specific differences among equations. For example, in the first group of brackets are variables related to changes in intercept; in the second group, changes in dbh slopes; in the third, changes in height slopes. The purpose is:

1 To test if all location equations are equal. This would mean variables in all 3 bracketed groups are insignificant, or b_3 through $b_{20} = 0$.
2 Provided equations are not equal, the next test is to determine if they have equal slopes. In this case, variables in the last 2 groups would be insignificant, b_9 through $b_{20} = 0$.
3 Provided slopes are not equal, the next tests determine if dbh slopes are equal, or height slopes are equal. Variables in either group 2 (b_9 through $b_{14} = 0$), or group 3 (b_{15} through $b_{20} = 0$) would be insignificant.
4 Provided equations are not equal, it was next tested to determine if they have equal intercepts. Variables in group 1 would be insignificant in this case, b_3 through $b_8 = 0$.

This procedure reduces possible differences among equations to certain groups of independent variables.

All computations were performed in single-precision on the CDC Cyber 170. The BMDP 1982 least squares regression program supplied the output from the design matrix (Dixon 1981). The design matrix contains 436 observations in the full model format.

The F-statistic was computed for each subset of coefficients tested as follows (Neter and Wasserman 1974):

$$F_0 = \frac{(SSE_{RM} - SSE_{FM}) \div (d.f._{RM} - d.f._{FM})}{SSE_{FM} \div d.f._{FM}} \qquad (19)$$

where:

SSE = sum of squared errors (e)
FM = full model, Equation (11)
RM = reduced model, Equations (12) through (18)
$d.f.$ = degrees of freedom

A significantly larger SSE from the reduced model would tend to make F_0 large. F_0 was compared with a critical value in a table of the F-distribution. Significance was at the 5% level. Hypotheses tested and results (Table 5) follows.

Table 5. Statistical results of test to determine the equality of coefficients among locations for tanoak tree volume equations.

Hypothesis	Test statistic	Critical value (a =0.05)	Conclusion
1 H_o All regression equations are equal H_o: $b_3=b_4= \ldots =b_{20}=0$ versus H_A: At least 1 b#0	F_o=15.683	1.161	Reject H_o; Equations not equal
2 H_o All slope coefficients are equal H_o: $b_9=b_{10}= \ldots =b_{20}=0$ versus H_A: At least 1 b#0	F_o=4.094	1.770	Reject H_o; Slopes not equal
3 H_o All dbh slope coefficients are equal H_o: $b_9=b_{10}= \ldots =b_{14}=0$ versus H_A: At least 1 b#0	F_o=2.891	2.100	Reject H_o; dbh slopes not equal
4 H_o All height slope coefficients are equal H_o: $b_{15}=b_{16}= \ldots =b_{20}=0$ versus H_A: At least 1 b#0	F_o=1.662	2.100	Fail to reject H_o; Height slopes equal
5 H_o All intercept coefficients are equal H_o: $b_3=b_4= \ldots =b_8=0$ versus H_A: At least 1 b#0	F_o=4.612	2.100	Reject H_o; Intercepts not equal

Results of Parameter Difference Analysis

The results of this test shows F_0 is larger than the critical value and led to the rejection of the null hypothesis H_0 at the 5% level (item 1, Table 5), implying that there is a difference among equations. Volume equations from different populations (locations) are statistically different. Detected differences involve the dbh slope and intercept coefficients (items 3 and 5). Note that the height slope coefficient was not rejected (item 4) implying greater similarity among height slope coefficients. However, when all slope coefficients (dbh and height) were combined (item 2) the null hypothesis was rejected.

Supplement information provided by the BMDP program indicated that the model which we used appears to be unbiased and adequate for this test. The coefficient of multiple determination R^2, measuring the proportion of variation in response explained by the full model, is 99.3%. A plot of the residuals versus predicted response showed no obvious model defects. The normal probability plot indicated a possible light-tailed distribution of residuals, but no strong departure from the normality assumption.

Discussion and Conclusions

We have tested tanoak volume equations from different areas using two methods of analysis: (1) volume prediction (how well they predict volumes from different areas), and (2) parameter difference (were their coefficients statistically different among the sampled regions).

The volume equations used both diameter at breast height and total height as predictors of cubic volume. The equations were termed "acceptable" for use in different areas if they predicted the sample volume within 12% of the cut volume. In general, the seven tanoak volume equations were not reliable as predictors of sample volumes outside their areas. Furthermore, no one equation consistently performed well when applied to the other six locations. We also tested the equation coefficients in several ways and found, at the 5% probability level, that the hypothesis of equation similarity was rejected.

It is our conclusion that "conventional wisdom" should be heeded; that large and often unacceptable errors can result when applying a tree volume equation to trees in different areas. The evidence as it relates to tanoak is particularly convincing.

Unfortunately volume tables continue to be applied to areas outside of the geographic range where the data were collected. Our analysis shows that the volume error could be as high as 40% for the worst case scenario (Table 3). Regardless of how one proceeds, the next best equation is not likely to be reliable. If possible, we recommend that volume equations be developed for the area of use to avoid serious errors If this is not possible, a rough calculation of the percent error is needed. To make this calculation, cut at least 20 trees representing the range of diameters in the stand and carefully measure them for volume. Develop a ratio of the "cut volume" to "predicted volume." This ratio will provide an approximation of the applicability of the volume equation being considered. Also, by multiplying the inventory volume (predicted) by this ratio, the adjusted inventory volume would better represent the stand.

Literature Cited

Atzet, T., D. Wheeler, B. Smith, G. Riegal, and J. Franklin. 1982. Tanoak series of the Siskiyou Region of Southwestern Oregon. FIR Rep. 6(3):6–7

Bolsinger, C.L. 1988. The hardwoods of California's timberlands, woodlands, and savannas. USDA For. Serv., Resour. Bull. PNW-RB-148. 148 p.

Burns, R., and B. Honkala, tech. coords. 1990. Silvics of North America 2 Hardwoods. Agric. Handb. 654. USDA For. Serv., Wash., DC. Vol 2 877 p.

Cunia, T. 1964. Weighted least squares method and construction of volume tables. For. Sci. 10(2):181–191.

Dixon, W.J. 1981. BMDP statistical software, 1981. University of California Regents, Berkeley. P. 23, 237–250.

Lund, R.E. 1975. Tables for an approximate test for outliers in linear models Techometrics 17(4):473–476.

McDonald, P.M., and J.C. Tappeiner II. 1988. Silviculture, ecology, and management of tanoak in northern California. P. 64–70 *in* Proc. symp. on Multiple-use management of California's hardwood resources, Plumb, T.R., and N.H. Pillsbury (eds.). USDA For. Serv. Gen. Tech. Rep. PSW-100.

Montgomery, D.C., and E.A. Peck. 1982. Introduction to linear regression analysis. Wiley, New York. P. 79, 81, 110, 146, 21–218.

Neter, J., and W. Wasserman. 1974. Applied linear statistical models. Richard Irwin, Inc., Homewood, Ill. 842 p.