Forest Science, Vol. 39, No. 4, pp. 670–679

Modeling Multiplicative Error Variance: An Example Predicting Tree Diameter from Stump Dimensions in Baldcypress

BERNARD R. PARRESOL

ABSTRACT. In the context of forest modeling, it is often reasonable to assume a multiplicative heteroscedastic error structure to the data. Under such circumstances ordinary least squares no longer provides minimum variance estimates of the model parameters. Through study of the error structure, a suitable error variance model can be specified and its parameters estimated. This error model is used to construct a covariance matrix which in turn is used to form an estimated generalized least squares estimator of the forest model parameters. The theory is illustrated with data on baldcypress (*Taxodium distichum* [L.] Rich.). A multiple linear regression equation is developed for predicting diameter at 3 m from solid-wood stump diameter (i.e., diameter inside the fluting) and stump height. By modeling the error structure, standard errors on three of the four coefficients from the tree diameter-stump dimensions regression were reduced by 13 to 50%. The effect on prediction confidence intervals is graphically illustrated. FOR. SCI. 39(4):670–679.

ADDITIONAL KEY WORDS. Heteroscedasticity, consistency, estimated generalized least squares, prediction confidence intervals.

FOREST MODELERS ARE OFTEN FACED WITH heteroscedasticity in their data. As is well known, under ordinary least squares (OLS) parameter estimates are no longer minimum variance, though they are still unbiased (Draper and Smith 1981, Neter et al. 1985). Further, prediction confidence intervals are no longer reliable. The solution to the problem is to weight each observation by the inverse of its variance. This then achieves homogeneity of variance. But what if the variance of each observation is unknown? The problem then becomes one of estimating the proper weight for each observation. In instances where there are replicated observations across the independent variable(s) the Minimum Norm Quadratic Unbiased Estimation (MINQUE) theory of Rao (1970, 1971a, 1971b) can be applied to estimate each σ_i^2. In many allometric relationships, errors are inherently lognormally distributed; hence a logarithmic transformation properly weights the observations and corrects for heteroscedasticity as well as nonnormality (Baskerville 1972, Beauchamp and Olson 1973, Flewelling and Pienar 1981).

It is often the case that the error variance (or disturbance) is functionally related to the predictor variable in a simple linear regression. Under multiple regression, it is possible for the disturbance to be functionally related with two or more predictor variables. Harvey (1976) and Judge et al. (1988) have shown that if the

Reprinted from *Forest Science*, Vol. 39, No. 4, November 1993.

error variance is a function of a small number of unknown parameters, and these parameters can be consistently estimated, then estimated generalized least squares (EGLS) estimation will provide asymptotically efficient estimates of the model parameters. In this paper I use a method suggested by Harvey (1976) for obtaining consistent estimates of the parameters of a variance model. In particular I derive a functional form for the error variance from a tree diameter-stump dimensions regression for baldcypress, fit the variance model as suggested by Harvey, then use EGLS to determine the coefficients for the baldcypress regression model.

THE MULTIPLICATIVE ERROR MODEL

Consider the general linear statistical model

$$y = X\beta + e$$

where X is a $(T \times K)$ observable nonstochastic matrix, β is a $(K \times 1)$ vector of parameters to be estimated, y is a $(T \times 1)$ observable random vector, and e is a $(T \times 1)$ unobservable random vector with properties

$$E[e] = 0 \text{ and } E[ee'] = \Phi = \sigma^2\Psi$$

and Ψ is a $(T \times T)$ diagonal matrix. Heteroscedasticity exits when the diagonal elements of Ψ are not all identical. In the general heteroscedastic specification $\Phi = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_T^2)$, where T unknown variances must be estimated with only T observations. Reasonable variance estimates cannot be obtained under such circumstances unless some further assumption is made that reduces the number of unknown parameters on which the variances depend. The assumption that often arises is that the error variance is related to one or more of the explanatory variables. Specifically, we assume that each σ_t^2 is an exponential function of P explanatory variables and hence

$$E[e_t^2] = \sigma_t^2 = \exp[z_t'\alpha] \; t = 1, 2, \ldots, T \tag{1}$$

where $z_t' = (z_{t1}z_{t2} \ldots z_{tP})$ is a $(1 \times P)$ vector containing the tth observation on P nonstochastic explanatory variables and $\alpha = (\alpha_1\alpha_2 \ldots \alpha_P)'$ is a $(P \times 1)$ vector of unknown coefficients. The first element in z_t is taken as unity ($z_{t1} \equiv 1$), and the other z's could be identical to, or functions of, the x's.

This specification reduces the problem of estimating T σ_t^2's to that of estimating the P dimensional vector α. The modeler now must choose not only x_t, the variables which explain changes in y_t, but also z_t, the variables that explain changes in the variance of y_t. The relevant z's may be obvious, and experience or past work may suggest the proper variables. Often graphical analyses of the data and/or residuals will reveal the appropriate z's.

Function (1) can be written as

$$\sigma_t^2 = \exp[\alpha_1] \cdot \exp[\alpha_2 z_{t2}] \ldots \exp[\alpha_P z_{tP}] \tag{2}$$

which shows that the components of the variance are related in a multiplicative fashion, hence the term *multiplicative heteroscedasticity*. A useful convention is the parameterization of the scale factor σ^2 as $\exp(\alpha_1)$, or $\alpha_1 = \ln \sigma^2$. This means the expression in (2) can be written as

$$\sigma_t^2 = \sigma^2 \exp[\mathbf{z}_t^{*\prime} \boldsymbol{\alpha}^*] \tag{3}$$

where $\mathbf{z}_t^{*\prime} = (z_{t2}, \ldots, z_{tP})$ and $\boldsymbol{\alpha}^* = (\alpha_2, \ldots, \alpha_P)'$. This will prove useful shortly. The covariance matrix can be written

$$\boldsymbol{\Phi} = \begin{bmatrix} \exp(\mathbf{z}_1'\boldsymbol{\alpha}) & & & \\ & \exp(\mathbf{z}_2'\boldsymbol{\alpha}) & & \\ & & \ddots & \\ & & & \exp(\mathbf{z}_T'\boldsymbol{\alpha}) \end{bmatrix}$$

$$= \sigma^2 \boldsymbol{\Psi} = \sigma^2 \begin{bmatrix} \exp(\mathbf{z}_1^{*\prime}\boldsymbol{\alpha}^*) & & & \\ & \exp(\mathbf{z}_2^{*\prime}\boldsymbol{\alpha}^*) & & \\ & & \ddots & \\ & & & \exp(\mathbf{z}_T^{*\prime}\boldsymbol{\alpha}^*) \end{bmatrix} \tag{4}$$

Estimating $\boldsymbol{\alpha}$

We first take logarithms of Equation (1) to obtain

$$\ln \sigma_t^2 = \mathbf{z}_t'\boldsymbol{\alpha} \tag{5}$$

Since the σ_t^2 are not known, we use instead the squares of the OLS residuals. These residuals are likely to reflect the size of σ_t^2, that is, large when σ_t^2 is large and small when σ_t^2 is small. Adding $\ln \hat{e}_t^2$ to both sides of Equation (5) yields

$$\ln \hat{e}_t^2 + \ln \sigma_t^2 = \mathbf{z}_t'\boldsymbol{\alpha} + \ln \hat{e}_t^2$$

or

$$\ln \hat{e}_t^2 = \mathbf{z}_t'\boldsymbol{\alpha} + v_t \tag{6}$$

where $v_t = \ln \hat{e}_t^2 - \ln \sigma_t^2 = \ln(\hat{e}_t^2/\sigma_t^2)$. In matrix notation Model (6) can be written as $\mathbf{q} = \mathbf{Z}\boldsymbol{\alpha} + \mathbf{v}$ where the vector $\mathbf{q} = (\ln \hat{e}_1^2 \ln \hat{e}_2^2 \ldots \ln \hat{e}_T^2)'$. One way to estimate $\boldsymbol{\alpha}$ is to apply OLS to model (6) which yields $\hat{\boldsymbol{\alpha}} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{q}$. What are the properties of $\hat{\boldsymbol{\alpha}}$. Since $\hat{\boldsymbol{\alpha}} = \boldsymbol{\alpha} + (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{v}$, the properties of $\hat{\boldsymbol{\alpha}}$ depend on those of $\mathbf{v}$. The finite sample properties of the elements in $\mathbf{v}$ are complicated. Using asymptotic properties Harvey (1976) showed that if the e_t's are normally distributed then

$$E[v_t] = -1.2704$$

$$var(v_t) = E[(v_t - E[v_t])^2] = 4.9348$$

$$cov(v_t, v_\tau) = 0 \; for \; t \neq \tau$$

Thus, the elements in $\mathbf{v}$ satisfy the OLS assumptions of homoscedasticity and no autocorrelation, but the assumption of 0 mean is violated. Fortunately, the consequences of having a disturbance with nonzero mean are not great. The consequence is that the intercept α_1 will not be consistently estimated, but the remaining elements in $\hat{\boldsymbol{\alpha}}$ will be consistent or unbiased.

ESTIMATED GENERALIZED LEAST SQUARES

Substituting $\hat{\boldsymbol{\alpha}}$ for $\boldsymbol{\alpha}$ in expression (4) we obtain the estimated covariance matrix $\hat{\boldsymbol{\Phi}} = \hat{\sigma}^2\hat{\boldsymbol{\Psi}}$. The EGLS estimator is formed as

$$\hat{\boldsymbol{\beta}} = (X'\hat{\boldsymbol{\Phi}}^{-1}X)^{-1}X'\hat{\boldsymbol{\Phi}}^{-1}y = (X'\hat{\boldsymbol{\Psi}}^{-1}X)^{-1}X'\hat{\boldsymbol{\Psi}}^{-1}y \tag{7}$$

Fortunately, $\hat{\boldsymbol{\beta}}$ does not depend on $\hat{\alpha}_1$, which can be factored out as a proportionality constant. Since $\hat{\boldsymbol{\alpha}}^* = (\hat{\alpha}_2, \hat{\alpha}_3, \ldots, \hat{\alpha}_P)'$, the estimator $\hat{\boldsymbol{\beta}}$ only depends on the consistently estimated elements of $\hat{\boldsymbol{\alpha}}$. The covariance matrix of the parameter estimates is consistently estimated by $\hat{\sigma}^2(X'\hat{\boldsymbol{\Psi}}^{-1}X)^{-1}$, where $\hat{\sigma}^2 = (y - X\hat{\boldsymbol{\beta}})'\hat{\boldsymbol{\Psi}}^{-1}(y - X\hat{\boldsymbol{\beta}})/(T - K)$. The usual hypothesis tests and interval estimates can be based on this result. For prediction intervals on some future value y_0 the sampling error is estimated by $\hat{\psi}_0\hat{\sigma}^2(1 + \mathbf{x}_0'(X'\hat{\boldsymbol{\Psi}}^{-1}X)^{-1}\mathbf{x}_0\hat{\psi}_0^{-1})$, where $\hat{\psi}_0$ is the scalar $\exp(\mathbf{z}_0^{*\prime}\hat{\boldsymbol{\alpha}}^*)$.

TESTING FOR MULTIPLICATIVE HETEROSCEDASTICITY

We have examined the model $y_t = \mathbf{x}_t'\boldsymbol{\beta} + e_t$, where $E[e_t^2] = \sigma^2 \cdot \exp(\mathbf{z}_t^{*\prime}\boldsymbol{\alpha}^*)$. Treating this model as an alternative to one with homoscedastic errors is equivalent to testing $H_0{:}\boldsymbol{\alpha}^* = \mathbf{0}$ against $H_1{:}\boldsymbol{\alpha}^* \neq \mathbf{0}$. Let $\boldsymbol{R}$ be the matrix $(Z'Z)^{-1}$ with its first row and first column removed. If the e_t's are normally distributed then $\hat{\boldsymbol{\alpha}}^* \sim N[\boldsymbol{\alpha}^*, 4.9348\boldsymbol{R}]$ and the following statistic (Judge et al. 1988), based on the distribution of quadratic forms in normal variables, tests the above null hypothesis

$$\frac{\hat{\boldsymbol{\alpha}}^{*\prime}\boldsymbol{R}^{-1}\hat{\boldsymbol{\alpha}}^*}{4.9348} \sim \chi^2_{(P-1)} \tag{8}$$

Note that the numerator is the regression (or explained) sum of squares obtained when estimating $\boldsymbol{\alpha}$ and this test is asymptotically equivalent to the F test for testing that all coefficients, except the intercept, are 0.

BALDCYPRESS MENSURATION AND DATA COLLECTION

Recent publications (Hotvedt et al. 1985, Parresol et al. 1987, Parresol and Hotvedt 1990) advocate the use of diameter measured at 3 m above the ground (D_3) as the reference diameter for second-growth baldcypress (*Taxodium distichum* [L.] Rich.) because D_3 is superior to the use of "normal diameter" (diameter measured 50 cm above butt-swell) for estimating tree volume and stem profile. Previous studies on the relationship of stump dimensions to tree diameter for baldcypress (e.g., McClure 1968) have used normal diameter as the reference diameter. The importance of the stump dimensions-reference diameter relationship for trespass cases, tree removal studies, etc., indicate the need for defining this relationship for the new reference diameter for baldcypress, namely D_3.

Data were collected on 157 downed trees from 26 sites (25 sites with 6 trees and 1 site with 7) located throughout the South Delta region of Louisiana. The area in and around the Atchafalaya Basin was sampled more intensively due to its proportionally higher concentration of baldcypress. Solid-wood stump diameter (D_s) was measured by inscribing with an expandable hoop the largest possible circle or ellipse (excluding bark) inside the flutes (Figure 1). If the shape was elliptical then the length of the long and short axes were geometrically averaged. D_s was taken at stump heights (H_s) of 0.30, 0.91, and 1.52 m on each tree. D_3 was also measured on each tree. All diameters were measured to the nearest 0.25 cm. Stump diameter ranged from 12.7 to 66.5 cm and D_3 ranged from 10.4 to 57.4 cm.

MODEL DEVELOPMENT

TREE DIAMETER-STUMP DIMENSIONS

After examining scatter plots of D_3 over D_s by stump height, the following simple linear model was hypothesized for the relationship at each stump height:

$$D_3 = \gamma_0 + \gamma_1 D_s + e \tag{9}$$

This model was fitted to the 157 data points at each H_s using OLS regression. The coefficients and variances are given in Table 1. To ascertain the relationship of D_3 to D_s and H_s, the intercept and slope coefficients from Model (9) were examined to see how they changed with height. The intercept and slope values from Table 1 are plotted over H_s in Figure 2. There is a sharp decline in the intercept value as stump heights increase from 0.3 to 0.91 m then a gradual decline for stumps of 0.91 to 1.52 m, finally the intercept becomes negative. A hyperbola models this trend well, so in Model (9) γ_0 is replaced by $\beta_0 + \beta_1/H_s$. The slope values have

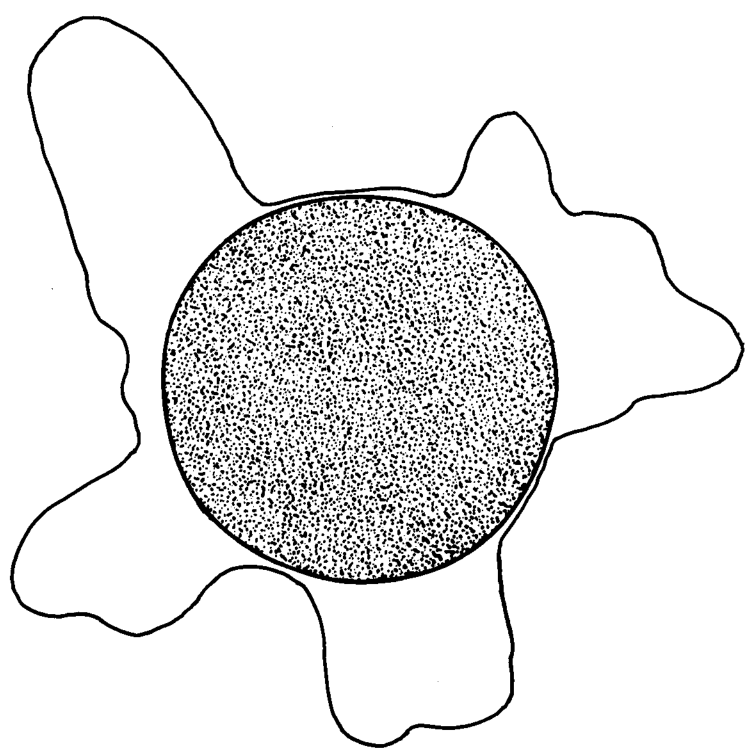

FIGURE 1. Top view of cypress stump showing inscribed circle containing solid-wood diameter.

TABLE 1.

Coefficients and variance for regression of tree diameter on stump diameter by stump height.

H_s	$\hat{\gamma}_0$	$\hat{\gamma}_1$	$\hat{\sigma}^2$
0.30 m	2.994	0.788	52.224
0.91 m	0.013	0.835	26.124
1.52 m	−0.138	0.872	12.530

a near-perfect straight line relationship with H_s, so γ_1 is replaced with $\beta_2 + \beta_3 H_s$. The following model results:

$$D_3 = \beta_0 + \beta_1 \frac{1}{H_s} + (\beta_2 + \beta_3 H_s) D_s + e$$

or

$$D_3 = \beta_0 + \beta_1 \frac{1}{H_s} + \beta_2 D_s + \beta_3 D_s H_s + e \tag{10}$$

Error Variance

An examination of the variances by stump height in Table 1 indicates a strong decreasing heteroscedastic trend with increasing stump height. A plot of these variances (Figure 2c) suggests a negative exponential trend with H_s, that is, $\sigma_t^2 = \sigma^2 \exp[-kH_{s(t)}]$. Scatter plots of the residuals over D_s for each of the three regressions revealed a similar fan pattern of increasing error variance. This type of heteroscedasticity is common and is usually modeled as a power function (Draper and Smith 1981, Neter et al. 1985), that is, $\sigma_t^2 = \sigma^2 D_{s(t)}^k$. Combining these two heteroscedastic trends into one multiplicative error model gives

$$\hat{e}^2 = \exp[\alpha_1 + \alpha_2 \ln D_s - \alpha_3 H_s] \cdot \exp[v]$$

or

$$\ln \hat{e}^2 = \alpha_1 + \alpha_2 \ln D_s - \alpha_3 H_s + v \tag{11}$$

RESULTS

Model (10) was fitted with OLS and the resultant OLS residuals were used as the dependent variable in the error variance model. The results of applying OLS to (11) are given in Table 2. Using the χ^2 statistic in (8) we have 490.543/4.9348 = 99.4 (Prob < 0.001), and we can conclude that there is significant heteroscedasticity in the OLS residuals. The same conclusion is obtained from the F test in Table 2. The vector $\hat{\alpha}^*$ is significant and is used to construct $\hat{\Psi}$. Model (10) was refitted using the EGLS estimator in (7). The parameter estimates and their corresponding standard errors from OLS and EGLS are listed in Table 3. The intercept term ($\hat{\beta}_0$) was affected most by applying the variance model. This is not surprising due to the large weight applied to smaller diameter trees. Its standard

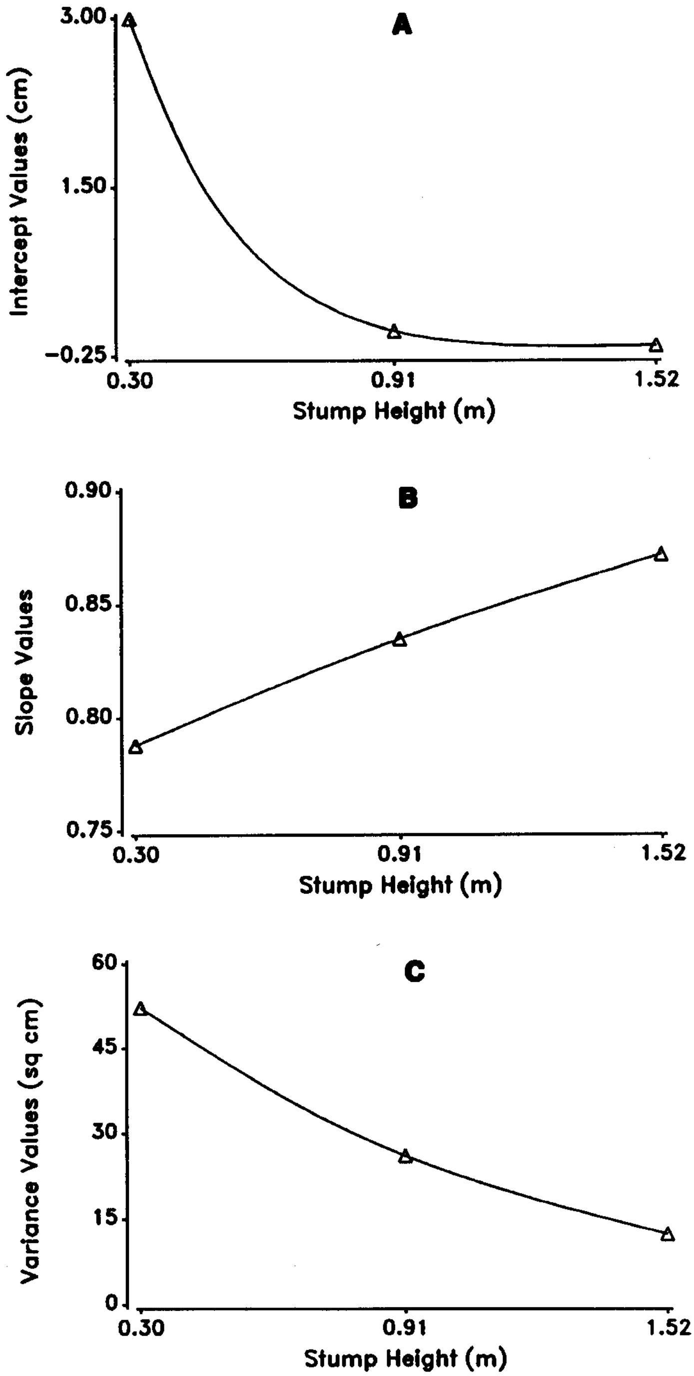

FIGURE 2. Plots of (A) intercepts, (B) slopes, and (C) variances from the regressions of D_3 on D_s by H_s.

error was halved and consequently its importance in Model (10) was established. Under OLS this parameter would erroneously be removed from the model. Reductions in standard errors of 14% and 13% occurred for the coefficients of the $1/H_s$ and D_sH_s terms in Model (10). These reductions, though not as dramatic as

TABLE 2.

Error variance regression model.

		Analysis of variance			
Source	df	Sum of squares	Mean square	F value	Prob > F
Model	2	490.543	245.271	58.96	<0.001
Error	468	1947.006	4.160		
C Total	470	2437.549			
		Parameter estimates			
Parameter	Estimate	Std err	t value	df	Prob > |t|
α_1	−3.014	0.940	−3.208	1	0.001
α_2	1.794	0.260	6.895	1	<0.001
α_3	1.564	0.189	−8.284	1	<0.001

for the intercept, still reflect considerable gains in efficiency and precision by modeling the error structure and using EGLS.

Predicted values and their confidence intervals for Model (10) under OLS and the multiplicative heteroscedastic specification fitted with EGLS are plotted in Figure 3. As is readily seen, predicted values are nearly identical whether one does or does not correct for heteroscedasticity. However, inferences are vastly different on the predicted values. Under OLS the width of the intervals are very uniform across the stump diameter-stump height combinations. This is counter-intuitive based on the scatterplots of the data and the positive test for heteroscedasticity. By modeling the error structure, confidence intervals displayed in Figure 3 under EGLS reflect the changing variances with stump diameter and stump height.

DISCUSSION

Heterogeneity is a common phenomenon inherent in many types of biological populations (Pielou 1977, Seber 1986). Thus, if we assume that $E[\boldsymbol{ee}'] = \sigma^2\boldsymbol{I}$ when in fact $E[\boldsymbol{ee}'] = \sigma^2\boldsymbol{\Psi}$ (a diagonal matrix), we will obtain a biased estimator of the covariance matrix for $\hat{\boldsymbol{\beta}}$, which, through hypothesis tests and interval estimation, could result in misleading inferences about $\boldsymbol{\beta}$. The OLS and EGLS results concerning the intercept parameter in Model (10) illustrates this point. By

TABLE 3.

Parameter estimates and their standard errors for baldcypress tree diameter-stump dimensions regression using OLS and EGLS.

	OLS				EGLS				
Parameter	Estimate	Std err	t-value	Prob > |t|	Estimate	Std err	t-value	Prob > |t|	Reduction in std err
β_0	−1.242	1.142	−1.088	0.277	−1.169	0.574	−2.036	0.042	50%
β_1	1.321	0.477	2.768	0.006	0.816	0.412	1.981	0.048	14%
β_2	0.762	0.035	21.801	<0.001	0.799	0.035	22.603	<0.001	0%
β_3	0.076	0.030	2.579	0.010	0.057	0.026	2.238	0.026	13%

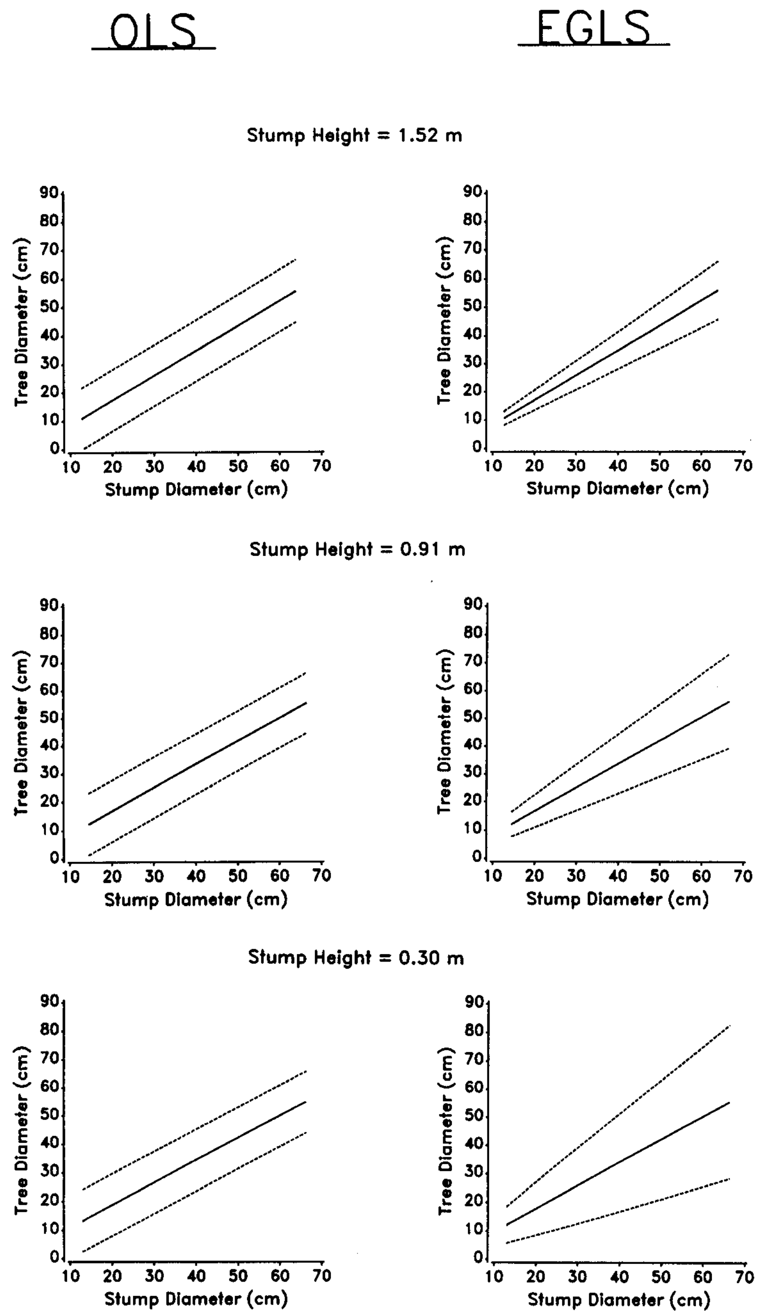

FIGURE 3. Comparison of predicted values (solid lines) and their confidence intervals (dashed lines) across D_s by H_s for Model (10) under OLS and EGLS.

modeling the error variance and using estimated generalized least squares an asymptotically efficient estimate of $\boldsymbol{\beta}$ can be obtained. Also, prediction confidence intervals will reflect the nonconstant nature of the error relation.

In applied work the functional form of multivariate data and the corresponding error variance can often be ascertained by disaggregating the data, building submodels, and graphing the results. This is the approach I used with the baldcypress data. The graphs of the intercept, slope, and error variance values from the simple linear functions aided in determining an appropriate model for baldcypress tree diameter from stump dimensions.

LITERATURE CITED

BASKERVILLE, G.L. 1972. Use of logarithmic regression in the estimation of plant biomass. Can. J. For. 2:49–53.

BEAUCHAMP, J.J., and J.S. OLSON. 1973. Corrections for bias in regression estimates after logarithmic transformation. Ecology 54:1403–1407.

DRAPER, N.R., and H. SMITH. 1981. Applied regression analysis. Ed. 2. Wiley, New York. 709 p.

FLEWELLING, J.W., and L.V. PIENAAR. 1981. Multiplicative regression with lognormal errors. For. Sci. 27:281–289.

HARVEY, A.C. 1976. Estimating regression models with multiplicative heteroscedasticity. Econometrica 44:461–465.

HOTVEDT, J.E., Q.V. CAO, and B.R. PARRESOL. 1985. Tree-volume and stem-profile functions for baldcypress. South. J. Appl. For. 9:227–232.

JUDGE, G.G., R.C. HILL, W.E. GRIFFITHS, H. LÜTKEPOHL, and T. LEE. 1988. Introduction to the theory and practice of econometrics. Ed. 2. Wiley, New York. 1024 p.

MCCLURE, J.P. 1968. Predicting tree d.b.h. from stump measurements in the southeast. USDA For. Serv. Res. Note SE-99. 4 p.

NETER, J., W. WASSERMAN, and M.H. KUTNER. 1985. Applied linear statistical models. Ed. 2. Richard D. Irwin, Homewood, IL. 1127 p.

PARRESOL, B.R., J.E. HOTVEDT, and Q.V. CAO. 1987. A volume and taper prediction system for baldcypress. Can. J. For. Res. 17:250–259.

PARRESOL, B.R., and J.E. HOTVEDT. 1990. Diameter measurement in baldcypress. For. Ecol. Manage. 33/34:509–515.

PIELOU, E.C. 1977. Mathematical ecology. Wiley, New York. 385 p.

RAO, C.R. 1970. Estimation of heteroscedastic variances in linear models. J. Am. Stat. Assoc. 65:161–172.

RAO, C.R. 1971a. Estimation of variance and covariance components—MINQUE theory. J. Multiv. Anal. 1:257–275.

RAO, C.R. 1971b. Minimum variance quadratic unbiased estimation of variance components. J. Multiv. Anal. 1:445–456.

SEBER, G.A.F. 1986. A review of estimating animal abundance. Biometrics 42:267–292.

Manuscript received July 17, 1992

AUTHOR AND ACKNOWLEDGMENTS

Bernard R. Parresol is a mathematical statistician, USDA Forest Service, Institute for Quantitative Studies, Southern Forest Experiment Station, 701 Loyola Ave., New Orleans, LA 70113. The data used in this manuscript was collected by the author while a graduate student at Louisiana State University. This research was supported by USDA Forest Service forest management research funds. The author wishes to thank the anonymous reviewers for their helpful comments and suggestions.