韓國林學會誌 79 (3) : 260-268. 1990.
Jour. Korean For. Soc. 79(3) : 260-268. 1990.

The Use of Ridge Regression for Yield Prediction Models with Multicollinearity Problems[1]

Man Yong Shin[2]

收穫豫測 Model 의 Multicollinearity 問題點 解決을 爲한 Ridge Regression의 利用

申 萬 鏞[2]

ABSTRACT

Two types of ridge regression estimators were compared with the ordinary least squares (OLS) estimator in order to select the "best" estimator when multicollinearity existed. The ridge estimators were Mallows's (1973) C_p-like statistic, and Allen's (1974) PRESS-like statistic. The evaluation was conducted based on the predictive ability of a yield model developed by Matney *et al.* (1988). A total of 522 plots from the data of the Southwide Loblolly Pine Seed Source study was used in this study

All of ridge estimators were better in predictive ability than the OLS estimator. The ridge estimator obtained by using Mallows's statistic performed the best. Thus, ridge estimators can be recommended as an alternative estimator when multicollinearity exists among independent variables.

Key words : Yield prediction models, multicollinearity, ridge regression.

要 約

收穫 豫測 model이 multicollinearity 問題點을 가질때 보다 정확한 推定式을 얻기 위하여 두 종류의 ridge estimator와 最小 自乘法(OLS)의 추정치를 比較하였다. 본 硏究에서 使用된 ridge estmator는 Mallows's(1973)Cp-like statistic과 Allens's(1974) PRESS-like statistic 이었다. 위의 세가지 estimator 豫測 能力 評價는 Matney 等(1988)에 의하여 開發된 收穫 model을 利用하여 比較하였다. 사용되어진 資料는 美國 南部 테에다 소나무 試驗林의 總522個 plot을 利用하였다.

두 個의 ridge estimator가 最小 自乘法에 의한 추정치 보다 收穫 豫測 能力이 優秀하였으며, 특히 Mallows's statistic에 의한 ridge estimator가 가장 優秀하였다. 따라서 ridge estimator는 收穫 豫測 model의 獨立 變數 間에 multicollinearity 問題點이 있을때 最小 自乘法에 의한 추정치를 代置할 수 있는 estimator로서 推薦할 수 있었다.

INTRODUCTION

Foresters are often required to make estimates of wood volume yield. Yield estimation accomplishes a key role in supporting management plans and determining the amount of cutting in the forest. Therefore, accurate yield prediction is essential to effective forest management planning.

Multiple linear regression techniques have been

[1] 接受 1990年 4月 16日 Received on April 16, 1990
[2] 慶熙大學校 產業大學 College of Industry, Kyung-Hee University

employed in the development of yield prediction models since Mackinney and Chaiken(1939) first applied them to loblolly pine stands. Model parameters usually have been estimated using the ordinary least squares(OLS) method, that produces estimates with lower variance than other linear unbiased estimators. However, the OLS estimators can have large variance when multicollinearity exists among variables in the data.

Yield prediction models require stand variables such as age, density, and site index as independent variables. Since the yield models are developed by multiple linear regression techniques, the presence of multicollinearity should be considered in the estimation of parameters for the prediction models. If high correlation exists between some of the independent variables, then the regression model is said to contain multicollinearity between these variables.

Problems can arise depending on the degree of multicollinearity that the regression model exhibits (Marquardt 1970 ; Kmenta 1971). When high multicollinearity is involved in a regression model, there are some adverse dffects on parameter estimates such as imprecise estimates and incorrect signs of regression coefficients.

To avoid most of the pitfalls of the OLS method in the presence of multicollinearity, biased estimation techniques such as ridge regression, principal components regression, and Stein-rule estimators have been used. Since the 1970's, much research has been conducted on obtaining biased estimators with better overall performance than OLS when multicollinearity is present(McDonald and Galarneau 1975 ; Gunst and Mason 1977 ; Dempster *st al*. 1977 ; Bare and Hann 1981).

The concerns of multicollinearity have been recently addressed in forestry. Mitchell and Hann (1979) discussed ridge regression methodology for dealing with multicollinearity and also presented an algorithm for obtaining the coefficients in ridge regression. Bare and Hann(1981) concluded, in the development of a basal area growth model for ponderosa pine, that the use of ridge regression produced precise and stable estimates of model parameters.

In this study, ridge regression were evaluated to select the "best"estimator in predictive ability of yield models.

PAST WORK

Least squares estimator is an unbiased estimator of the regression parameters and has the smallest variance of all unbiased linear functions. However, the least squares estimator can be extremely unstable when there exists multicollinearity in the data. To obtain appropriate estimators under conditions of multicollinearity, therefore, considerable attention has been focused on biased estimation of the parameters of a linear regression model.

A number of alternatives to OLS may be preferable although they produce biased estimates. The objection to bias may not be strong depending upon the intended use of the regression models(Hocking 1976). The important issue would appear to be whether or not the resulting estimators perform better than the OLS estimation method.

Ridge Regression

Ridge regression sacrifices unbiasedness to obtain parameter estimates that have a smaller mean squared error(MSE). The ridge estimator proposed by Hoerl and Kennard (1970) is

$$b_{RR} = (X'X + kI)^{-1}X'y \qquad (1)$$

where

b_{RR} = the ridge estimator,

X = standardized matrix of independent variables,

X' = transpose of X,

y = standardized dependent variable vector,

I = identity matrix, and

k = ridge parameter.

Since the 1970's, there has been much interest in ridge regression. The concept of ridge regression has been examined by many researchers(Marquardt 1970 ; Mayer and Willke 1973 ; McDonald and Schwing 1973 ; McDonald and Galarneau 1975). Much of the discussion centered around the choice of the constant k. It is recognized that the OLS estimator is unlikely to be a satisfactory estimator when the design matrix (X'X) is badly conditioned due to multicollinearity. Ridge regression can be used to remedy this problem. The important step in ridge regression is to choose a value for k such that the ridge estimator has smaller mean squared error than the OLS estimator. To improve the coefficents of the models, numerous methods such as ridge trace and variance inflation factor have been proposed for determining the value of k.

Some other criteria have been proposed to select k when the prediction capability of the model is more important than the precision of coefficients of the models. Research on this topic has been sketchy so far. Myers (1986) summarized general criteria to select the value of k for prediction performance of regression models. The criteria are Mallows's (1973) C_p-like statistic, and Allens's (1974) PRESS-like statistic.

C_p was proposed by Mallows (1973) as a criterion for selecting a regression model. C_p is a measure of total squared error. Mallows's criterion in a ridge regression context, C_k, has been used by some researchers to select k. Erikson (1983) used ridge regression to directly estimate lagged effects in marketing and discussed the C_k statistic as one of the prediction criteria for ridge regression. Li (1986) discussed the asymptotic optimality of C_k in the setting of ridge regression.

Allen (1974) proposed PRESS (predicted residual sum of squares) as a cross validation technique for the selection of a suitable regression model. When prediction capability is an important criterion for a choice of k, a PRESS-like statistic can be used in ridge regression. This statistic is very similar to the PRESS statistic in OLS. The method consists of dropping one observation at a time, estimating the model, and predicting its left-out observation. The sum of squares of the predicted residuals is computed for each choice of k. Delaney and Chatterjee (1986), using Monte Carlo simulation technique, evaluated several methods of choosing ridge parameter k including the PRESS-like statistic. Erickson (1983) also reviewed the PRESS-like statistic and compared it with other prediction criteria.

Bare and Hann (1981) introduced ridge regression to the field of forestry, using it to select independent variables during the development of a basal area growth model for ponderosa pine. They concluded that the use of ridge regression produced a meaningful predictive model with interpretable coefficients. However, no study so far has been done to improve the predictive capability of yield models based on data with multicollinearity problems.

MATERIALS AND METHODS

Data

Date for this study came from the Southwide Loblolly Pine Seed Source Study, which was established in 1952-1953 to determine the genetic variation associated with geographic variation for loblolly pine (Wells and Wakeley 1966). A total of 522 plots was available from this data set. Each plot contained 121 trees on a 2m×2m spacing. The inner 49 trees on each plot were measured at 1, 3, 5, 10, 15, 20, and 25 years after planting, although the last three measurements at some locations were made at age 16, 22, and 27 instead.

Height of the 49 measurement trees on each plot was noted when planted, and survival was recorded the first May and June thereafter. Diameter at breast height was recorded starting at the tenth growing season.

Total cubic-meter volume outside bark per hectare was computed using Burkhart *et al*.'s (1972) individual tree volume equation. Also, the mean height of the tallest 50 percent of surviving trees at each age was used as average height of the dominants and codominants for each plot. This approach was employed by Golden *et al*. (1981) on the same data set because crown class data were not available.

Since only data after the tenth growing season are generally available for the development of growth and yield models, data collected before age 10 were not used to estimate parameters of yield prediction models. Furthermore, remeasurements from these permanent plots formed time series data. The autocorrelation among the error terms of the time series data was detected ($p>0.1$) by Durbin-Watson test (Neter *et al*.1985). To remove the effect of autocorrelation problems on yield prediction models,

Table 1. Data summary of stand variables for the fit and test data sets.

Variable[a]	Number of obs.	Minimum	Maximum	Mean
		Fit data set		
Age (years)	261	10	27	18
H_d	261	6.0	23.7	14.4
N	261	60	2927	1334
V	261	7.8	189.9	72.7
		Test data set		
Age (years)	261	10	27	18
H_d	261	4.8	23.8	14.9
N	261	121	2865	1186
V	261	6.9	189.0	73.4

[a] Notations :
H_d = Average height of the dominant and codominants in meter.
N = Number of trees per hectare.
V = Total volume per hectare in cubic-meter outside bark.

only one age class from each plot was randomly selected. This process was adopted to simulate the temporary plot data similar to those used for developing yield models.

Yield prediction data for this study were divdied randomly into a fit data set and a test data set. Regression coefficients of the model were estimated from the fit data set. The test data set was used to validate the ability of the yield models to accurately predict volume yield for an independent data set. The fit data set consisted of 261 plots randomly selected from a total of 522 plots available. The remaining 261 plots were withheld to form the test data set. This half-and-half data splitting method is popular when the collection of new data is neither practical nor possible for model validation (Snee 1977). The fit and test data sets were found to be similar in stand attributes (Table 1.)

Procedure

The process of data standardization was employed before fitting the model. Standardization is merely a transformation on variables that eliminates all units of measurements and forces the standardized variables to have the same mean and the same amount of variability.

Model form for yield prediction

The model form developed by Matney *et al*. (1988) for yield prediction was used for this study :

$$\ln(V) = b_0 + b_1(1/A) + b_2 \ln(H_d)/A + b_3 \ln(N)/A + b_4 \ln(H_d) \quad (2)$$

where

V = total cubic-meter volume outside bark per hectare,

A = total stand age in years,

H_d = average height of the dominants and codominants in meter,

N = number of surviving trees per hectare, and

ln(X) = natural logarithm of x.

Multicollinearity diagnostics

Multicollinearity means that the model has redundant information because of linear dependuncy among independent variables. In this study, four diagnostics (simple correlations among independent variables, variance inflation factors (VIFs), system of eigenvalues of X'X, and variance decomposition proportions) were used to detect the strength of the linear dependencies and how much the variance of each regression coefficient is inflated.

Correlation is a measure of the intensity of association. In multiple regression, however, the simple correlations do not always underscore the extent of the multicollinearity problem because multicollinearity often involves associations among multiple independent variables. Even though the simple correlations do not indicate the extent of multicollinearity, they may provide guideline values to see which one-on-one associations exist (Myers 1986). The values of simple correlations among independent variables are presented in Tabpe 2. As a general rule if the correlation coefficient between the values of two independent variables is greater than 0.8 or 0.9, then multicol-

Teble 2. Simple correlations among independent variables used in the yield prediction model.

Variable	1/A	$\ln(H_d)/A$	$\ln(N)/A$	$\ln(H_d)$
1/A	1.0000	0.9863	0.9646	-0.9085
$\ln(H_d)/A$		1.0000	0.9472	-0.8385
$\ln(N)/A$			1.0000	-0.8858
$\ln(H_d)$				1.0000

linearity is a problem (Judge *et al*. 1988). In this study, the absolute values of correlation coefficient among independent variables ranged from 0.8385 to 09863, signifying a degree of multicollinearity.

The VIFs represent the inflation that each regression coefficient experiences above the ideal level if the correaltion matrices were an identity matrix. They provide more a productive approach for detection than do simple correlation. They indicate which coefficients are adversely affected and to what extent. The VIF is given by

$$VIF = 1/(1-R^2_i) \quad (3)$$

where

R^2_i = coefficient of determination when X_i is regressed on the remaining inedpendent variables.

It is generally known that if VIF exceeds 10 there should be at least some concern with multicollinearity (Myers 1986). As shown in Table 3, the VIFs of variables 1/A and $\ln(H_d)/A$ were 222.1 and 120.8, respectively, indicating that a multicollinearity problem should be suspected.

Eigenvalues of the correlation matrix can also be used to detect the multicollinearity problem. A near-zero eigenvalue indicates a strong linear dependency. Multicollinearity can be measured in the condition number of correlation matrix which is given by

$$\phi_i = \sqrt{\frac{\lambda max}{\lambda_i}} \quad (4)$$

where

ϕ_i = the condition number of the ith eigenvalue,

λmax = the largest eigenvalue of the correlation matrix, and

λ_i = the ith eigenvalus of the correlation matrix.

A large condition number is evidence that the regression coefficients are unstable. When the condition number exceeds 30, multicollinearity should be suspected (Belsley *et al*. 1980). Table 4 shows that the smallest eigenvalue in this study had a condition number of 36.38, signifying a multicollinearity problem.

Table 3. Variance inflation factor analysis for the fit data set.

Variable	Variance inflation factor
1/A	222.1
$\ln(H_d)/A$	120.8
$\ln(N)/A$	14.5
$\ln(H_d)$	18.8

Table 4. Condition numbers and variance proportions for the fit data set as multicollinearity diagnostics.

Eigen-value	Condition number	Variance proportion			
		1/A	$\ln(H_d)/A$	$\ln(N)/A$	$\ln(H_d)$
3.767380	1.0000	0.0003	0.0006	0.0046	0.0033
0.177395	4.6084	0.0006	0.0115	0.0090	0.2114
0.052379	8.4809	0.0075	0.0270	0.9475	0.0214
0.002846	36.3821	0.9916	0.9609	0.0389	0.7640

It should be emphasized that a serious multicollinearity does not deposit its effect on only one regression coefficient. The variance decomposition proportions should be analyzed to determine what proportion of the variance of each coefficient is attributed to each dependency. The variance decomposition proportions are computed as follows :

$$p_{ji} = (v^2_{ij}/\lambda_j)/c_{ii} \quad (5)$$

where

p_{ji} = variance decomposition proportion of b_i,

v_{ij} = the ith element in the eigenvector associated with the jth eigenvalue,

λ_j = the jth eigenvalue, and

$c_{ii} = Var(b_i)/\sigma^2$.

According to the analysis of variance proportions in this study (Table 4), the precision of estimating regression coefficients for 1/A and $\ln(H_d)/A$ was damaged by the linear dipendency with high variance proportions for the smallest eigenvalue. It seems that the variable $\ln(H_d)$ does not have a lot of variation. Thus, based on the analysis of variance proportions, the variables 1/A and $\ln(H_d)/A$ basically seem to be the same.

From the above diagnostics, some multicollinearities were detected in the data. As a result, an alternative estimation method to OLS should be recommended for the yield prediction model.

Ridge Regression

The performance of the ridge regression estimator depends on how well the ridge parameter k is determined. Obviously, in yield prediction models with multicollinearity, the prediction capability should be improved by using an appropriate value for k. In this study, two criteria for choosing k were Mallows's

(1973) C_p-like statistic, and Allens's (1974) PRESS-like statistic. Mallows's criterion in a ridge regression context is

$$C_k = SSE_k/\hat{\sigma}^2 - n + 2 + 2tr(H_k) \qquad (6)$$

where

SSE_k = the sum of squared error using ridge regression,

$\hat{\sigma}^2$ = the mean squared error from OLS estimation,

n = number of observation,

H_k = hat matrix in ridge regression, which is computed by $X(X'X+kI)^{-1}X'$, and

$tr(H_k)$ = trace of the hat matrix for ridge regression.

The PRESS-like statistic, a modification of Allens's PRESS, used in this study is given by

$$PR(Ridge) = (1/n) \sum_{i=1}^{n} [e^2_{i,k}/(1-h_{ii,k})^2] \qquad (7)$$

where

$e_{i,k}$ = the ith residual for specific value of k, and

$h_{ii,k}$ = the ith diagonal elements of hat matrix.

Most ridge regression is applied to the standardized form of the model. The ridge estimator (1) in standardized form is given by

$$b^*_{RR} = (R_{xx} + kI)^{-1} r_{xy} \qquad (8)$$

where

R_{xx} = the correlation matrix of indendent variables, and

r_{xy} = the vector of simple correlation of the independent variables and the dependent variable.

For different values of k from 0 to 1, the two criteria C_k statistic, and PR(Ridge) were computed using the standardized from of the data. A value of k which minimized the statistic was chosen for each criterion. The parameters of the yield prediction model were then estimated from the equation(8), resulting in two yield equations.

Evaluation criteria

Parameter estimates of the yield prediction model were obtained from the fit data set using each of the ridge estimation methods. In addition, the OLS technique was employed to estimate the parameters of the model. Thus, three final equations were evaluated to determine which method provided the "best"results in terms of prediction performance of the model under the multicollinearity situation.

To evaluate the estimation methods, candidate estimators were compared based on the following three evaluation criteria.

1. *Mean difference*, which is a measure of bias of a model.

$$\overline{Diff} = (1/n) \sum_{i=1}^{n} Diff_i$$

where

$Diff_i = y_i - \hat{y}_i$ = difference between the ith observed and predicted volume per hectare, and

n = the number of observations.

2. *Mean absolute difference*, which is a measure of precision of a model.

$$|\overline{Diff}| = (1/n) \sum_{i=1}^{n} |Diff_i|$$

3. Mean squared difference, which is similar to the mean absolute difference, but is more sensitive to outliers.

$$\overline{Diff^2} = (1/n) \sum_{i=1}^{n} (Diff_i)^2$$

These statistics were computed separately for the test data and the pooled data (both fit and test data sets). The test data represented an independent data set, whereas the pooled data were regarded as the representative of the population.

The evaluation criteria were computed based on volume per hectare rather than the logarithm of volume which was the dependent variable in the yield model. This was because volume per hectare was really the variable of interest.

The final three equations were ranked relative to one another based on each criterion, with rank 1 corresponding to the smallest value. Then the overall rank was calculated as the sum of the ranks over three criteria. The "best"system of yield prediction equation was the one with the smallest overall rank.

RESULTS AND DISCUSSION

The minimum values for C_k, and the PRESS-like statistic were obtaned when k was 0.00013, and 0.00065, respectively. These k values were conservative (close to zero). Hocking (1976) reported that for his data, C_k statistic was more conservative in producing a smaller k value than the ridge trace and VIF criteria. In this study, the PRESS-like criterion produced the least conservative (largest k) biased estimation of the coefficients.

Three sets of coefficients of the yield prediction model (2) were obtained from the fit data set (Table 5). The three estimation methods were OLS, two ridge estimators based on different criteria of choosing k. The results of evaluation on the test data set and the pooled data set are presented in Table 6.

The ranks based on the three criteria are presented in Table 7. The overall ranks were similar for both the test data set and pooled data set, indicating that each estimator performed consistently for an independent data as well as the population.

For both validation data sets, ridge estimators performed better than OLS estimators. Especially, the C_k criterion produced the best improvement in prediction capability of the model. The PRESS-like statistic also provided some improvements of prediction over the OLS and ranked second, in both validation data sets (Table 6). However, the ridge estimators gained 1.2 to 3.8 cubic meter per hectare in mean difference and mean absolute difference for both validation data sets. This amount of improvement by ridge estimators over OLS estimators may not be meaningful in the applications for small areas. These results were similar to those obtained by Delaney and Chatterjee (1986), who compared ridge estimators to OLS estimator through Monte Carlo simulations. They concluded that, for the predictive ability, the OLS estimator performed as well as the ridge estimator from PRESS-like statistic. Judge *et al*. (1988) also discussed a near-exact multicollinearity situation in which the ill effects of small eigenvalues were cancelled out, resulting in good predictions from the OLS estimator.

However, yield prediction models are usually applied to large areas. Thus, this study showed that the use of ridge estimators provides the meaningful improvement of precision for the yield prediction when data have a multicollinearity problem.

Table 5. Parameter estmates of the yield prediction model from three different estimation methods.

Estimators a/	Parameter estimates				
	b_0	b_1	b_2	b_3	b_4
OLS	0.2038	-61.7273	3.0816	8.7655	1.8884
C_k	0.0998	-59.9155	2.6386	8.7320	1.9162
PR (Ridge)	-0.2487	-53.9804	1.2856	8.6080	2.0048

a/ Notation :
OLS = Ordinary least squares estimator,
C_k = Ridge estimator based on Mallow's (1973) statistic (k = 0.00013), and
PR (Ridge) = Ridge estimator based on Appens's (1974) PRESS-like statistic (k = 0.00065).

CONCLUSION

This study was conducted to select the "best" estimation method of linear regression yield models with multicollinearity. Attention has been focused on biased estimation techniques for dealing with multicollinearity. Two ridge estimators were compared to select the best estimator in predictive ability of yield models with the OLS estimator.

Table 6. Evaluation statistics from three estimtion methods for the test data set and the pooled data set.

Estimator	*Test data set*			*Pooled data set*		
	$\overline{\text{Diff}}$ a/	$\lvert\overline{\text{Diff}}\rvert$ b/	$\overline{\text{Diff}^2}$ c/	$\overline{\text{Diff}}$	$\lvert\overline{\text{Diff}}\rvert$	$\overline{\text{Diff}^2}$
OLS	8.22	27.3	763.5	10.61	29.1	837.4
C_k	7.09	24.1	711.2	9.01	25.3	783.9
PR (Ridge)	7.16	25.2	739.3	9.66	26.8	790.7

a/ $\overline{\text{Diff}} = (1/n) \sum_{i=1}^{n} \text{Diff}_i$, where $\text{Diff}_i = y_i - \hat{y}_i$ = difference between the ith observed and predicted volume per hectare.

b/ $\lvert\overline{\text{Diff}}\rvert = (1/n) \sum_{i=1}^{n} \lvert \text{Diff}_i \rvert$.

c/ $\overline{\text{Diff}^2} = (1/n) \sum_{i=1}^{n} (\text{Diff}_i)^2$.

Table 7. Ranks[a] of evaluation statistics from three estimation methods for the test data set and the pooled data set.

Estimator	Test data set: $\overline{Diff}$	Test data set: $\overline{\|Diff\|}$	Test data set: $\overline{Diff^2}$	Test data set: Total	Poolled data set: $\overline{Diff}$	Poolled data set: $\overline{\|Diff\|}$	Poolled data set: $\overline{Diff^2}$	Poolled data set: Total	Rank sum	Overall rank
OLS	3	3	3	9	3	3	3	9	18	3
C_k	1	1	1	3	1	1	1	3	6	1
PR(Ridge)	2	2	2	6	2	2	2	6	12	2

a: Number to represent relative performances of three estimation methods (1 being the best and 3 being the worst). The overall ranks were determined by the sum of the ranks over three evaluation statistics.

Based on three evaluation statistics, ridge estimators were better than the OLS in their performances. However, care should be focused on the method of choosing ridge parameter k. In this case, the choice of k in ridge regression should be restricted to prediction-oriented selection criteria such as C_k, and PRESS-like statistics.

Ridge estimator with k based on the C_k statistic was the "best" in the predictive ability. This ridge estimator thus can be recommended as an alternative to the OLS estimator when there exists multicollinearity in the data for yield prediction models.

LITERATURE CITED

1. Allen, D.M. 1974. The relationship between variable selection, data augmentation, and a method of prediction. Technometrics 16 : 125-127.
2. Bare, B.B. and D.W. Hann. 1981. Applications of ridge regression in forestry. Forest Sci. 27 : 338-348.
3. Belsley, D.A., E.Kuh, and R.E.Welsch. 1980. Regression Diagnostics. John Wiley & Sons, New York. 292 p.
4. Burkhark, H.E., R.C.Parker., M.R.Strub, and R.G.Oderwald. 1972. Yield of old-field loblolly pine plantations. Va. Poly. Inst. and State Univ. Pub. FWS-3-72, 51 p.
5. Delaney, N.J. and S. Chatterjee. 1986. Use of the bootstrap and cross-validation in ridge regression. J. Busi. & Econ. Assoc. 72 : 77-93.
6. Dempster, A.P., M. Schatzoff and N. Wermuth. 1977. A simulation study of alternatives to ordinary least squares. J. Am. Stat. Assoc. 72 : 77-93.
7. Erickson, G.M. 1983. Using ridge regression to directly estimate lagged effects in marketing. J. Am. Stat. Assoc. 76 : 766-773.
8. Golden, M.S., H. Meldahl, S.A.Knowe, and D. B. Boyer. 1981. Predicting site index for old-field loblolly pine plantations. South. J.Appl. For. 5(3) : 109-114.
9. Gunst, R.F. and R.L. Mason. 1977. Biased estimation in regression : An evaluation using mean squared error. J. Am. Stat. Assoc. 72 : 616-628.
10. Hocking, R.R. 1976. The analysis and selection of variables in linear regression. Biometrics 32 : 1-50.
11. Hoerl, A.E. and R.W. Kennard. 1970. Ridge regression : Biased estimation for non-orthogonal problems. Technometrics 12 : 55-67.
12. Judge, G.G., R.C. Hill, W.E. Griffiths, H. Lutkepohl, and T. Lee. 1988. Introduction to the Theory and Practice of Econometrics. John Wiley & Sons, New York. 1024 P.
13. Kmenta, J. 1971. Elements of Econometrics. MacMillan, New York. 665 P.
14. Li, K.C. 1986. Asymptotic optimality of C_L, and generalized cross-validation in ridge regression with application to spline smoothing. The Annals of Stat. 14 : 1101-1112.
15. Mackinney, A.L and L.E. Chaiken. 1939. Volume, yield, and growth of loblolly pine in the Mid-Atlantic Coastal Region. U.S. For. Serv. Tech. Note No. 33, 30 p.
16. Mallows, C.L. 1973. Some comments on C_p. Technometrics 15 : 661-675.
17. Marquardt, D.W. 1970. Generalized inverses,

ridge regression, biased linear estimation, and nonlinear estimation. Technometrics 12 : 591-612.

18. Matney, T.G., A.D. Sullivan, J.R. Ledbetter, and R.M. Farrar. 1988. Stand-level cubic-foot volume ratio equations for planted loblolly pine on site-prepared land in the mid-South Gulf Coastal plain. South. J. Appl. For. 12 : 7-11.

19. Mayer, L.S. and T.A. Willke. 1973. On biased estimation in linear models. Technometrics 15 : 497-508.

20. McDonald, G.C. and D.I. Galarneau. 1975. A monte carlo evaluation of some ridge-type estimators. J. Am.Stat.Assoc. 78 : 407-416.

21. McDonald, G.C. and R.C. Sehwing. 1973. Instabilities of regression estimates relating air pollution to mortality. Technometrics 12 : 591-612.

22. Mitchell, B.R. and D.W. Hann. 1979. A computer program for applying ridge regression techniques to multiple linear regression. USDA For. Serv. Gen. Tech. Rep. INT-51, 25 p.

23. Myers, R.H. 1986. Classical and Modern Regression with Applications. Duxbury Press, 359 p.

24. Neter, J., W. Wasserman, M.H. Kutner. 1985. Applied Linear Statistical Models. Homewood, Illinois. Irwin. 1127p.

25. Snee, R.D. 1977. Validation of regression models : Methods and examples. Technometrics 19 : 415-428.

26. Wells, O.O. and P.C. Wakeley. 1966. Geographic variation in survival, growth, and fusiform-rust infection of planted loblolly pine. Forest Sci. 11 : 1-40.