PREDICTING TREE VOLUMES WITHOUT TAPER CURVES

by

Evelyn Lacaya

M.Sc., University of Calgary, 1983

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in the Department

of

Mathematics and Statistics

SIMON FRASER UNIVERSITY

December 1987

APPROVAL

Name: Evelyn Lacaya

Degree: Master of Science

Title of thesis: Predicting Tree Volumes Without Taper Curves

Examining Committee:

Chairman: Dr. G. Bojadziev

Dr. R. Routledge
Senior Supervisor

Dr. T. Swartz

Dr. L. Weldon

Dr. C. Villegas
External Examiner

Date Approved: December 8, 1987

ABSTRACT

Estimates of wood volume in standing trees are commonly obtained by first constructing a taper equation that predicts the shape of the bole, and by integrating this equation to predict volume. In this study, a method for estimating wood volume based on a statistical model that directly relates volume to such quantities as diameter at breast height and total height is considered. This method does not require the use of taper curves. Both the total tree volume and volume to any merchantable standard of utilization expressed at top diameter limits are estimated using the model. Logarithmic transformation of the proposed volume equation leads to a linear regression model relating the logarithms of the variables. The proposed method is a feasible alternative to those based on estimating taper equations for situations in which the very large data sets required for estimating the equation cannot be collected.

ACKNOWLEDGMENTS

My sincere thanks to Dr. Rick Routledge for his ideas, guidance, great patience and the time he spent with me in the preparation of this work. I would also like to thank him for his excellent supervision during my studies.

My thanks to the British Columbia Ministry of Forests for providing the data for this project.

I would also like to thank Simon Fraser University for the financial support that I have received.

Finally, my special thanks to my husband Rick, for his encouragement and support.

DEDICATION

To my family

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

In forestry, considerable importance has been given to the development of estimation schemes to predict volume for each individual tree and for the whole stand. The main reason for this is economic. Forest industries and other organizations often need periodic inventories to determine the quantity of wood available for utilization. The traditional measure of such quantity is volume.

It has been a practice for foresters to rely on prepared volume tables (Husch, *et. al.*, 1972). These tables predict the volume of standing trees of given species, diameters and heights that would be obtained if trees where felled, bucked and scaled as logs. Aside from providing cubic volume for the entire main stem, the tables also sometimes give either the volume from stump to a given merchantable diameter or the volume of cut lumber obtainable from a tree.

There are difficulties that may arise when one depends upon using these prepared volume tables. Husch, *et al.* (1972) give the following summary of these problems: 1) Any individual tree may not have the same height, diameter, and form characteristics as those used in the construction of the volume table employed. 2) Volume tables are constructed for specific units of measurement and utilization standards, and it is often impossible or unwieldy to try to change to these specifications.

3) Any errors in the volume tables are incorporated into the estimate and in most cases no assessment of their magnitude is included. In addition to these difficulties, one must consider that the volume table is based on established estimation schemes which are in turn based on a specific model. Thus, the reliability of such tabulations depends upon the model selected as the basis for the construction.

In modern practice, equations are generally used to predict tree volume rather than obtaining the values from tables (Avery and Burkhart, 1983). Numerous volume equations, varying in complexity have been proposed and used. Foresters are constantly searching for models that are accurate but quick and easy to use. These models must be simple and flexible enough to calculate all parameters based on only a few tree characteristics which are important and easy to measure such as diameter outside bark at breast height and total height (Demaerschalk, 1972).

The traditional approach has been to make an educated guess of tree volume based on directly measurable characteristics such as total height, HT and diameter at breast height, DBH. (DBH is the diameter measurement at 1.3 meters from the ground). The educated guess is obtained by fitting a regression equation to sample observations relating volume to height and diameter at breast height.

The more recent studies concentrate on equations that describe the shape of the tree or taper curve. A taper equation predicts the diameter at any point along the stem as a function of height at that point, total height and diameter at breast height. An estimate of total volume of wood in the merchantable portion of the bole can then be extracted by integrating the taper equation. The main reasoning behind the use of taper curves is that if tree profile can be accurately described, the volume for any merchantability limit or segment can be computed by integration. The use of taper curves also allows the determination of the height of a given diameter, the diameter at a given height, and the volume between two diameter limits or two heights.

There are disadvantages in using taper curves. One is that tree profiles are highly irregular; hence taper curves are modelled by complex functions containing several undetermined parameters. The model given by Demaerschalk and Kozak (1977) currently used by the British Columbia Ministry of Forests contains eight parameters some of which are very complicated to estimate. If the parameters are not estimated accurately, then neither is volume. Also, when a complex model is used, it becomes more difficult to solve the associated equation for height and volume for a given merchantable diameter. Another disadvantage of using taper curves is that whenever a taper curve is estimated, the parameters define an 'average' tree profile for the data. Trees coming from different species or from single species in a single stand may have different forms. Also in a stand of trees, dominant trees tend to be taller and larger in diameter and take a more conical form than the suppressed trees which tend to be more paraboloidal (Reed and Byrne, 1985). This introduces bias for all trees having a form different from the average form. Several attempts have been made to eliminate bias. Different equations have been proposed to model different parts of the tree rather than using one equation for the whole tree bole. The dual - equation model proposed by Demaerschalk ad Kozak (1977) has minimized bias but has not completely eliminated it.

Foresters have argued that the simpler regression method in the past is flawed. Max and Burkhart (1976) suggested that a more complex model is needed to adequately describe the tree bole. Bruce, Curtis and Vancouvering (1968) state, "in the past, the principal difficulty in developing complex taper curves was the computational labor involved. This handicap has been largely overcome by the electronic computer." However, a recent study by Martin (1984) has shown that the simple formulas based on regular geometric solids usually outperform formulas based on more complex taper curves in predicting total and merchantable volumes.

The objective of this project is to look at quick and reliable methods for estimating tree volumes that do not use taper curves. Geometric considerations will motivate the construction of the prediction equations for: 1) total tree volume, and 2) merchantable volume for a given upper diameter limit.

Total volume is defined as the volume inside bark from the stump to the tree top (British Columbia Forestry Handbook, 1983). This quantity is used in determining the total stumpage fee that an individual or company has to pay the government when logging public forests (British Columbia Forestry Handbook, 1983). In British Columbia, a forest license grants the licensee the right to harvest a specific volume each year. As of 1983, the provincial government charged the licensee an annual rental of $0.25 per cubic meter of allowable cut of timber (British Columbia Forestry Handbook, 1983).

Merchantable volume is the amount of wood between ground level and the terminal position of the last usable portion of the tree. Buyers are usually interested in the merchantable volume and the amount of cut lumber that can be extracted from standing trees.

In this study, the volume equation model is formulated by considering the geometric nature of the tree. By taking logarithms, the volume equation results in a linear regression model of the variables (throughout the project, logarithms mean the natural logarithm, i.e. the base used is e). This regression model is fitted to the data. Back transformation of this fitted model then gives the estimated volume equation formula.

Survey of Literature

Tree volume equations have been in the literature for many years. Only a survey of popular estimation equations will be included.

The following symbols will be used throughout the discussion of volume equations.

V = total volume

DBH = diameter at breast height

HT = total tree height.

Schumacher and Hall (1933) proposed the logarithmic volume equation

$$\log V = b_0 + b_1 \log(DBH) + b_2 \log(HT).$$

This equation is appropriate for predicting total volume. To predict merchantable volume , the equation is conditioned to pass through an appropriate point by translating the axes. In 1976, the British Columbia Ministry of Forest constructed volume tables for all commercial species in the province based on this equation (British Columbia Handbook of Forestry, 1983).

Cunia (1964) and Moser and Beers (1969) advocated the use of weighted least squares to solve the volume equation

$$V = \lambda DBH^{\alpha} HT^{\beta}$$

where γ, α, β are the unknown parameters. They each used a different weighting scheme to stabilize the variance.

Honer (1965) proposed the transformed variable function

$$V = \frac{DBH^2}{(a + b/HT)}$$

to express the volume, diameter and height relationship. A linear relationship between DBH^2/V and $1/HT$ was observed, permitting an ordinary least square solution to the equation $DBH^2/V = a + b/HT$.

Newnham (1967) tested what he referred to as a modified combined-variable formula for calculating the total volume of trees on data of 11 species. This formula is,

$$V = a_1 + a_2 DBH^{\alpha} HT^{\beta}$$

where the a_1 and a_2 are regression coefficients and parameters α and β are obtained by another regression using Schumacher's and Hall's formula:

$$Log(V) = \gamma + \alpha Log(DBH) + \beta Log(HT).$$

He compared the results with that from the combined formula $V = a_1 + a_2 DBH^2 HT$ proposed by Spurr (1952). The modified formula gave a small improvement in accuracy. This may be attributed to the fact that there is more freedom in the use of the two extra parameters α and β.

The search for accurate taper models began late in the 19th century and since then many models have been developed. (James and Kozak, 1984).

Bruce, Curtis and Vancouvering (1968) used a polynomial regression equation for red alder, expressing the ratio of squared upper stem diameter inside bark(DIB) to squared diameter outside bark as a function of DBH, HT, and the 3/2, 3rd, 32nd, and 40th power of relative height. The equation can be integrated to find appropriate volumes.

Kozak, Munro and Smith (1969) proposed the parabolic equation for upper stem diameter d, at height h,as

$$(d/DBH)^2 = b_0 + b_1 h/HT + b_2 (h/HT)^2.$$

The least square solution was derived by imposing the condition that the sum of the coefficients b_0, b_1, and b_2 be zero. This was needed to ensure that when h equals HT, the estimated diameter is equal to zero (since $b_0 = -b_1 - b_2$). They concluded that no practical improvement can be gained in volume estimation for any measurement of form in addition to DBH and HT measurements.

Ormerod (1971) estimated the upper bole diameters of tree species that have a prolonged, undivided stem (i.e., that have an excurrent form), by the equation

$$d = DBH(HT - h/HT - k)^{\rho} , \quad \rho > 0.$$

where d is the estimated diameter, h is the height of estimation, and ρ is the unknown parameter. Due to the changes in form along the length, the equation may not provide an adequate description of the bole. The equation was modified by Ormerod (1973) as to be what he called a step function,

$$d_i = (D_i - C_i)(H_i - h/H_i - k)^{\rho_i} + C_i \quad , \rho_i > 0$$

where: H_i is the height to top of section i, C_i is the section diameter intercept, D_i is the section measured diameter at height k_i, and p_i is the fitted exponent on the closed interval $[h=H_{i-1}, h=H_i]$. The "step" function resulted in a better description of the bole.

Max and Burkhart (1976) tested the use of segmented polynomial regression and found three submodels to be best. One submodel was used for the upper bole portion which was assumed to be a cone, another for the the middle part assumed to be paraboloid a and a third for the lower portion assumed to be a neiloid frustum (a solid more concave than the cone frustum). The submodels were restricted so as to produce identical diameters at join points and diameters that decreased monotonically from the base of the stem to the tip. The first-order derivatives were constrained to match at the join points.

Demaerschalk and Kozak (1977) developed taper equations consisting of two mathematical functions, one describing the upper bole and another, the lower bole. The two functions were joined at an inflection point and were continuous at that point. The whole system was based on height and breast-height diameter measurements. Bias was observed at heights near the ground level.

Demaerschalk and Kozak's model was developed from a set of inside bark measurements. To obtain such data, trees must be felled, bucked, and measured - a costly undertaking. James and

Kozak (1984) developed a system to estimate volumes from standing trees following Demaerschalk and Kozak's taper system. Aside from diameter inside bark at breast height and height measurements, their scheme requires only the outside bark diameters. The relationship of bark thickness to diameter at various points on the stem can be predicted from a reference bark thickness at breast height. All the outside bark diameters in the data set were then reduced by double bark thickness at breast height expressed as a percentage of DBH to obtain the corresponding inside bark diameters at other parts of the tree. Using Demaerschalk and Kozak's equation, volumes were then estimated.

Reed and Byrne (1985) developed a volume estimation system, giving total as well as merchantable volume to a height or diameter limit. The scheme was based on Ormerod's model. The parameter ρ however, was made to vary depending on the tree's total height and diameter at breast height. With low HT and DBH ratio (about 30:1, when expressed in common units), the tree form was considered to be a cone and ρ=1. A tree with HT and DBH ratio about 90:1 was considered a parabola with ρ= 1/2. Most of the trees however, fell in between these two extremes. For these cases, ρ was calculated using the formula

$$\rho = 1 - \frac{(HT/DBH) - 30}{120}$$

where HT and DBH were expressed in common units. The merchantable volume was estimated from the product of the total volume and a ratio R determined from the equation

$$R = 1 - (1 - h/HT)^{2\rho+1}$$

or

$$R = 1 - (1 - 1.3/HT)^{2\rho+1} \cdot (d/DBH)^{2\rho+1/\rho}$$

where h and d are the merchantable height and diameter respectively. The overall performance of the volume estimation system presented to evaluate the prediction ability varied by species.

Therein and Camire (1986) proposed a model based on the algebraic equation

$$Y = b_0 + b_1X + b_2X^2 + b_3X^3 + b_4X^4 + b_5X^5 + b_6X^6.$$

where

Y= d/DR	Dr= reference diameter
	d=diameter to evaluate
X= h/Hr	Hr= reference height
	h = length from d to top
b_i	i =0,1,...,6 parameters to evaluate.

Six measured diameters and corresponding heights were needed to obtain six equations to solve six unknown parameters for each tree. After the parameters were calculated, the volume was then obtained by integration. Since seperate equations had to be fitted for different trees , an increase in precision to the estimate was noted.

Martin (1984), compared true log volume to the estimates derived using taper equations and several other volume equations. His purpose was to see how closely these predicted volumes estimate the true volume. He used the water displacement technique in determining true volumes of logs. Each log was immersed in a specially built xylometer filled with water. The log volume is equal to that of displaced water. The results were compared in precision and accuracy with the estimated volume using fourteen different equations. When estimates were based on actual measurements (heights, lengths, and diameters) Huber, and Newton's followed by Smalian's formula did the best jobs in predicting volumes; these are all based on simple geometric models, the volume formulas of which are given on Table II.1 of the next chapter. Ormerod's taper based volume function ranked next in ability to predict volume accurately. Some of the equations showed widely different levels of performance. The results indicated that volume equations not based on taper curves gave the best overall performance in predicting total and merchantable volumes.

CHAPTER II

DATA

The data for this project were made available by the British Columbia Ministry of Forests (BCMF). Coastal and Interior Douglas Fir *(Pseudotsuga Menziesii,* [Mirb] Franco) and White Spruce *(Picea glauca,* [Moench] Voss) are the tree species being analyzed.

The data were gathered at various locations in the province and from a variety of ages and sites (good, medium, and poor). Measurements on individual trees included among others, total height (HT), diameter at breast height (DBH) and bark thickness. Depending upon its overall height, each tree was divided into eight to twelve sections. For each section, measurements were taken for lengths, upper and lower inside bark diameters and bark thickness. All measurements were in meters.

Diameter at breast height (DBH) is the diameter measurement outside bark at 1.3 meters above the point of germination, which is usually on the ground level. In the BCMF data set, DBH was measured using a diameter tape which was graduated to read the diameter of the trees to the nearest 0.1 cm when the tape was wrapped around the trunk. For trees of abnormal swell at the 1.3 meter mark, an average diameter above and below the swelling was used.

Total tree height HT is the vertical distance from the point of germination to the tree tip. Smaller trees were measured directly using a rod or a pole. For taller trees, a hypsometer was used. There are several forms and variations of this instrument, but the construction is based on the geometric principle of similar triangles.

Two measurements of bark thickness were made at breast height diametrically opposite each other. These were then averaged. Bark thickness was determined using a gauge that was pushed through the bark. Diameter inside bark (DIB) was obtained by subtracting two times the bark thickness from DBH.

In the calculation of volume, mensurationists often consider the whole bole as a composite of geometric solids. A typical example is shown on the Figure II.1.

The shape of the tip approximates a cone or a paraboloid, the central sections resemble frusta of paraboloids, and in some cases frusta of cones, and the butt resembles a frustum of a neiloid. The stump from the ground level to .3 m in height is considered to be a cylinder in the calculation of volume (British Columbia Forestry Handbook, 1983). Formulas that are often useful in calculating the volume of these solids are given in Table II.1.

Figure II.1:

Geometric Forms Assumed by Portions of Tree Stem

(Source: Forest Mensurations, Husch et.al., 1972)

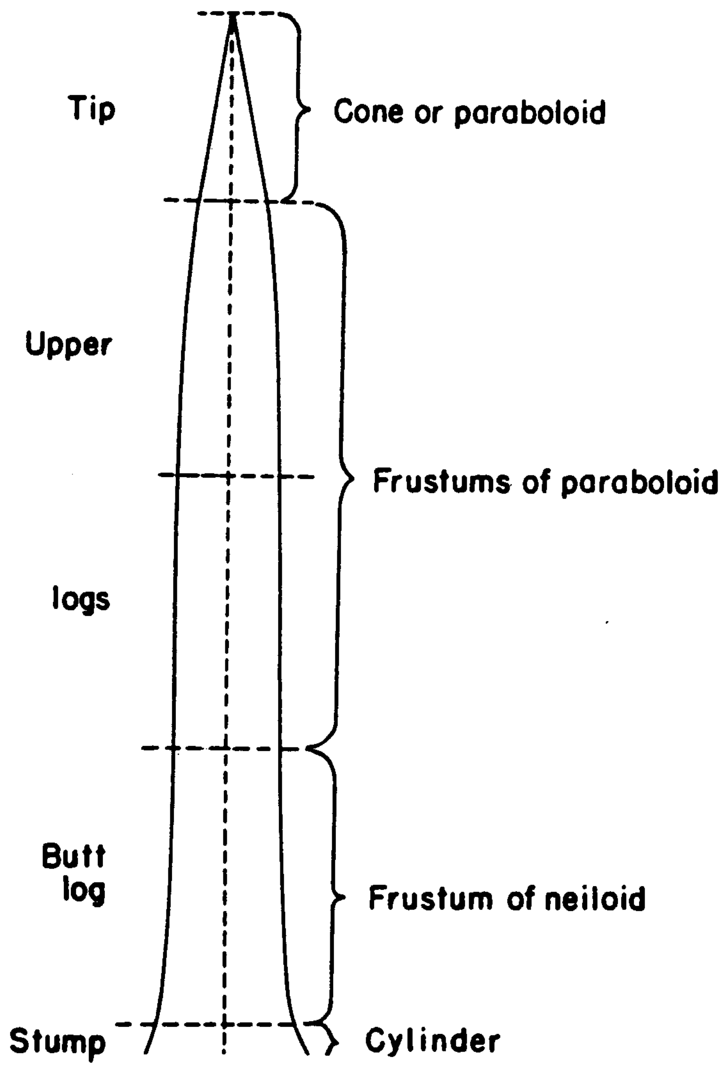


Table II.1

Equations to Compute Cubic Volume of Important Solids

Geometrical Solid	Equation for Volume, V, in Cubic Units
Cylinder	$V = A_b h$
Paraboloid	$V = \frac{1}{2}(A_b h)$
Cone	$V = \frac{1}{3}(A_b h)$
Paracone	$V = \frac{2}{5}(A_b h)$
Neiloid	$V = \frac{1}{4}(A_b h)$
Paraboloid Frustum	$V = \frac{h}{2}(A_b + A_u)$ (Smalian's Formula) $V = h(A_m)$ (Huber's Formula)
Cone Frustum	$V = \frac{h}{4}[A_b + (A_b A_u)^{1/2} + A_b A_u]$
Neiloid, Cone or Paraboloid Frustum	$V = \frac{h}{6}(A_b + 4A_m + A_u)$ (Newton's Formula)

h = height
A_b = cross-sectional area at base
A_m = cross-sectional area of middle
A_u = cross-sectional area of top

(Source: Husch, et.al, 1972)

In the calculation of volume, BCMF assumed the topmost section to be a paracone (a solid that is between a cone and a parabola) in the calculation of volume. The volume of the lowermost section was calculated assuming a cylinder from the ground to the lowest height observation. The sections in between were assumed to be frusta of paraboloids and were computed using Smalian's formula. Section volumes were automatically summed up to give the observed total cubic volume inside bark. Table II.2 on page 19, gives a summary of the dimensions and volumes of trees used in the analysis.

The trees used in the analysis are very common species in British Columbia. Coastal Douglas Fir trees grow along the southern and eastern side of Vancouver Island, on the Gulf Islands and on the adjacent coastal mainland occupying about 8,500 square km of forest land. The interior type, occupying about 48,000 square km is found in the south central third of British Columbia (British Columbia Forestry Handbook, 1983). The interior tree is characterized by a cylindrical, long and branch-free trunk with a short, flat crown. By contrast, the coastal form has a short-tapering trunk and a long, limby crown which can grow as high as 45 to 92 meters while the interior form is much stockier and seldom exceeds 43 meters (Hosie, 1979).

White Spruce can be found almost everywhere in Canada and can grow in a variety of soils and climates. The tree has a pronounced, uniform, conical crown with branches that spread or droop slightly covering the trunk. In dense stands, where there is little light, it gradually sheds its lower branches.

Douglas Fir is the one of the best known timber producing trees in the world market. It is used for a variety of purposes

- for lumber, veneer, poles, pulp and many other uses. White Spruce is one of the most important sources of pulpwood and lumber in Canada.

Although botanists refer to both Interior and Coastal Douglas Fir as the same species, throughout this project, for convenience, we refer to them as different tree species.

BCMF designated two age groupings, immature and mature for all trees in the province (British Columbia Forestry Handbook, 1983). The immature are up to age 135, while the mature types are over 135. The data in this project are a collection of both mature and immature type trees.

Table II.2: Summary of Data

	Number of Trees	Diameter (in meters) Range	Average	Height (in meters) Range	Average	Volume (in cubic meters) Range	Average
Coastal Douglas Fir	581	.054 - 2.164	.5928	6.28 - 76.72	33.46	.0070 - 77.313	5.47
Interior Douglas Fir	1504	.021 - 1.357	.3521	2.62 - 44.10	20.83	.0005 - 44.10	1.13
White Spruce	4789	.041 - 1.107	.3035	3.02 - 54.47	23.73	.0027 - 15.23	1.14

CHAPTER III

THE MODEL

As discussed in Chapter 1, Martin's study showed a simpler estimation scheme giving a better result than the more complex taper-based equations. This was the motivation for developing a simple model that relates volume to field measurements of height and diameter.

The model is an application of the multiple regression technique that predicts a function of the volume V in terms of functions of diameter at breast height (DBH) and height (HT) measurements. The development of the model is guided by considering the geometric nature of tree boles. A thorough discussion of this aspect will be given in the latter part of the chapter.

First, we show an example that illustrates a problem that arises when multiple regression is applied without considering the geometry of the problem.

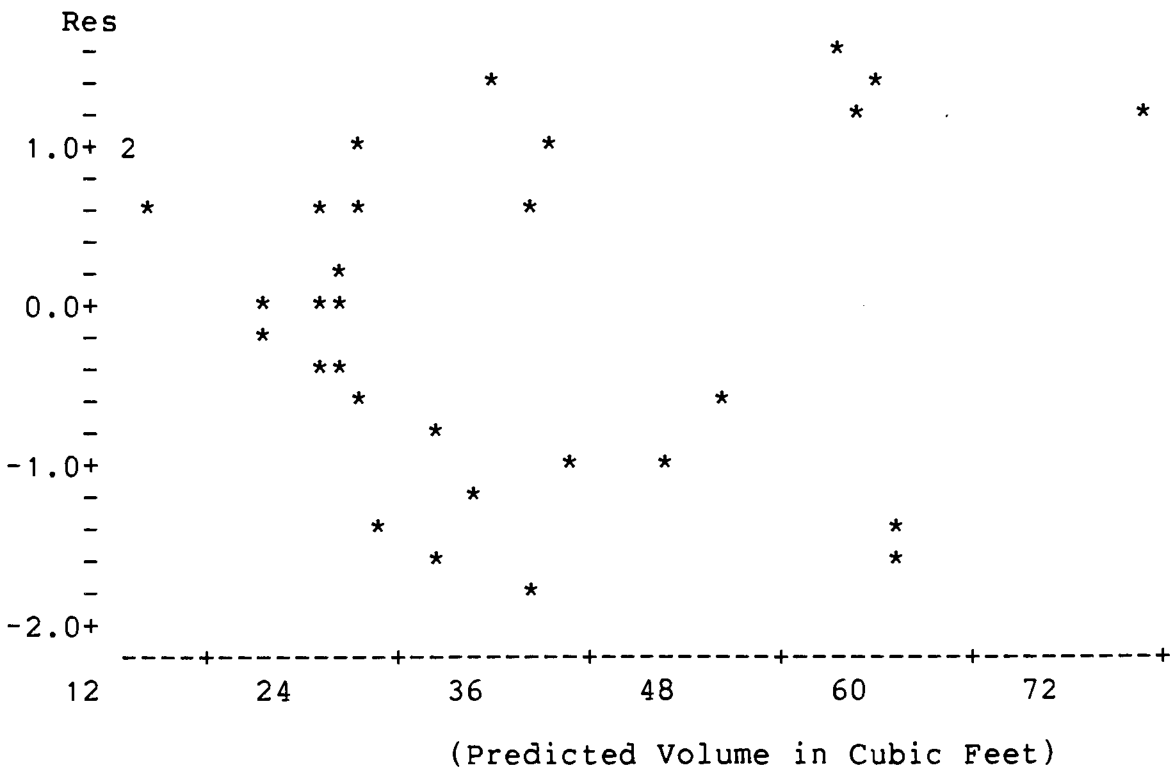
In the handbook for the general statistical package MINITAB (Ryan, et al., 1976), a modelling procedure for tree volume is presented. The suggested approach is to try predicting V by using expressions that are increasingly complicated polynomials in DBH and HT until the prediction errors look like independent draws from a normal population with mean zero and a constant standard deviation. For their data set on 31 black cherry trees

(*Prunus serotina*, Ehrh.), the procedure yields the prediction equation:

$$V_i = a + b\, DBH^2{}_i + c\, HT_i + e_i ,$$

where a,b,c are the estimated regression coefficients and the e_i- terms are the prediction errors. To assess the validity of the regression model, the prediction errors are plotted against the predicted values in Figure III.1. The plot shows error terms apparently satisfying the usual assumption of the regression model.

Figure III.1: Residual Plot for Cherry Trees



When the same model is applied to another species of trees, a different plot is evident. Fitting the same model to 25 black spruce trees *(Picea mariana, [Mill] B.S.P.)* resulted in a bowl shaped trend with increasing variance as shown in Figure III.2.

Figure III.2: Residual Plot for Black Spruce Trees

```
    Res
     -                                                      *
  3.2+
     -
     -
     -
     -                                             *
  1.6+
     -
     - *
     -   2*
     -   ***  *    *                          *
  0.0+       * *          *
     -           **           *
     -               *  *
     -                *             *                      *
     -
 -1.6+
     -                                  *         *
     -
       +---------+---------+---------+---------+---------+---------+---
0.0      4.0       8.0      12.0      16.0      20.0
                      (Predicted Volume in Cubic Feet)
```

The two residual plots clearly indicate the flaw of the modelling scheme. One regression model may work well for one species of trees but not for another.

III.1 Predicting Total Volume

Considering the geometric nature of the volume prediction problem will help overcome the difficulties presented previously. Foresters have long recognized that if all trees were to have the same shape then the volume would be related to DBH and HT through the equation, (Schumacher and Hall, 1933)

$$V_{tot} = f\ DBH^2 HT, \qquad (3.1)$$

where V_{tot} is the total tree volume, and the constant of proportionality, f, is called a form factor. Husch, et al. (1972) defined the form factor as the ratio of tree volume to the volume of some geometrical solid such as a cylinder, a cone, or a cone frustum, that has the same diameter and height as the tree. The form factor is used as a multiplier of the volume of a standard geometric solid to obtain the tree volume. If the geometric solid is chosen to be a square prism, then

$$f = (V_{tot})/(DBH^2 HT).$$

If the correlation between the form factor with height and diameter is zero , then the volume increases directly as height and the square of DBH. But this does not seem to be the case as the form factor depends weakly on DBH and HT. Figures III.3 and III.4 show the plot of the form factor $V_{tot}/(DBH^2 HT)$ vs. HT and DBH respectively for the Coastal Douglas Fir, while Figures III.5 and III.6 are the correponding plots in logarithmic scale.

FIGURE III.3

Plot of Form Factor vs. DBH

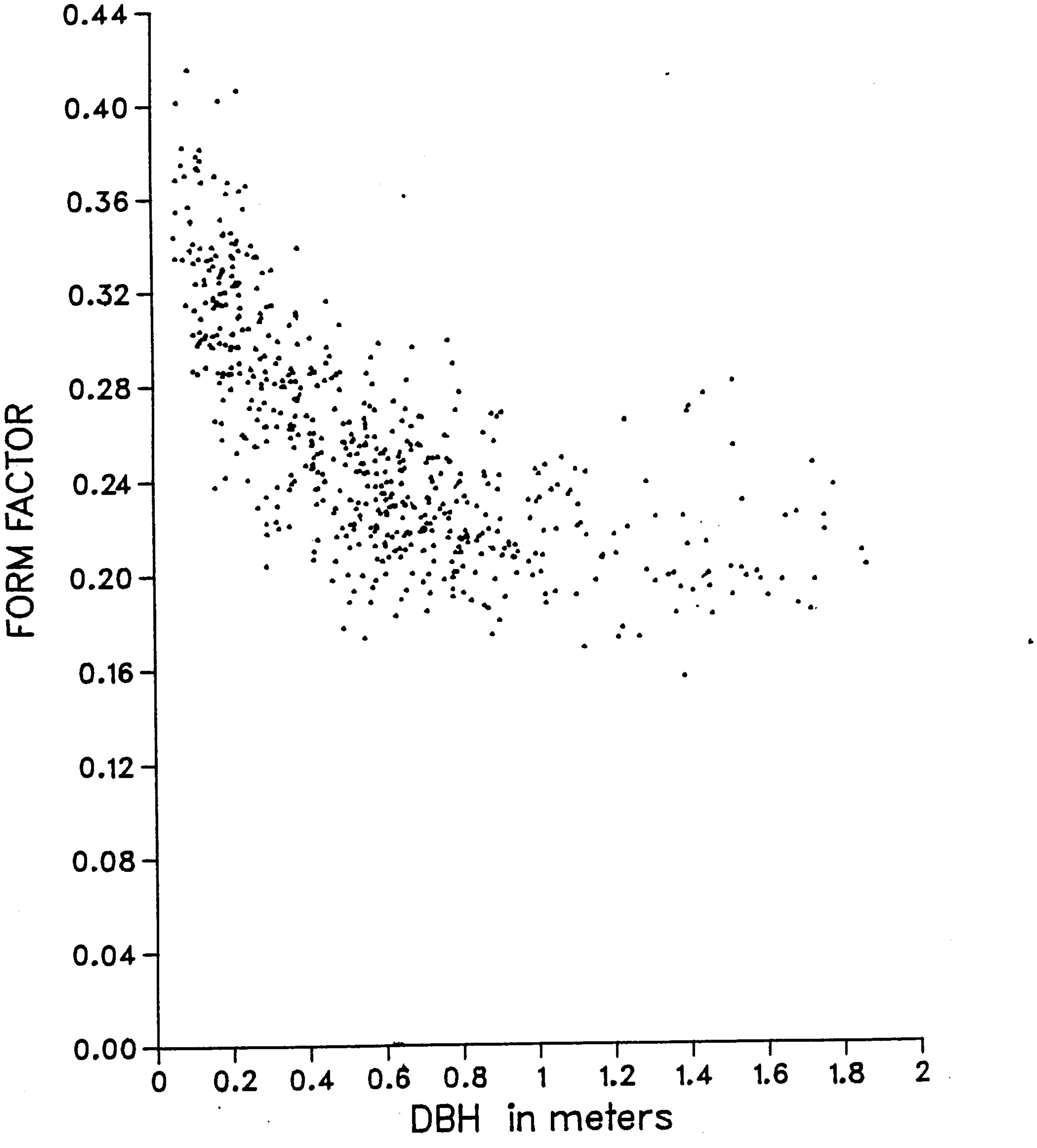


FIGURE III.4

Plot of Form Factor vs. HEIGHT

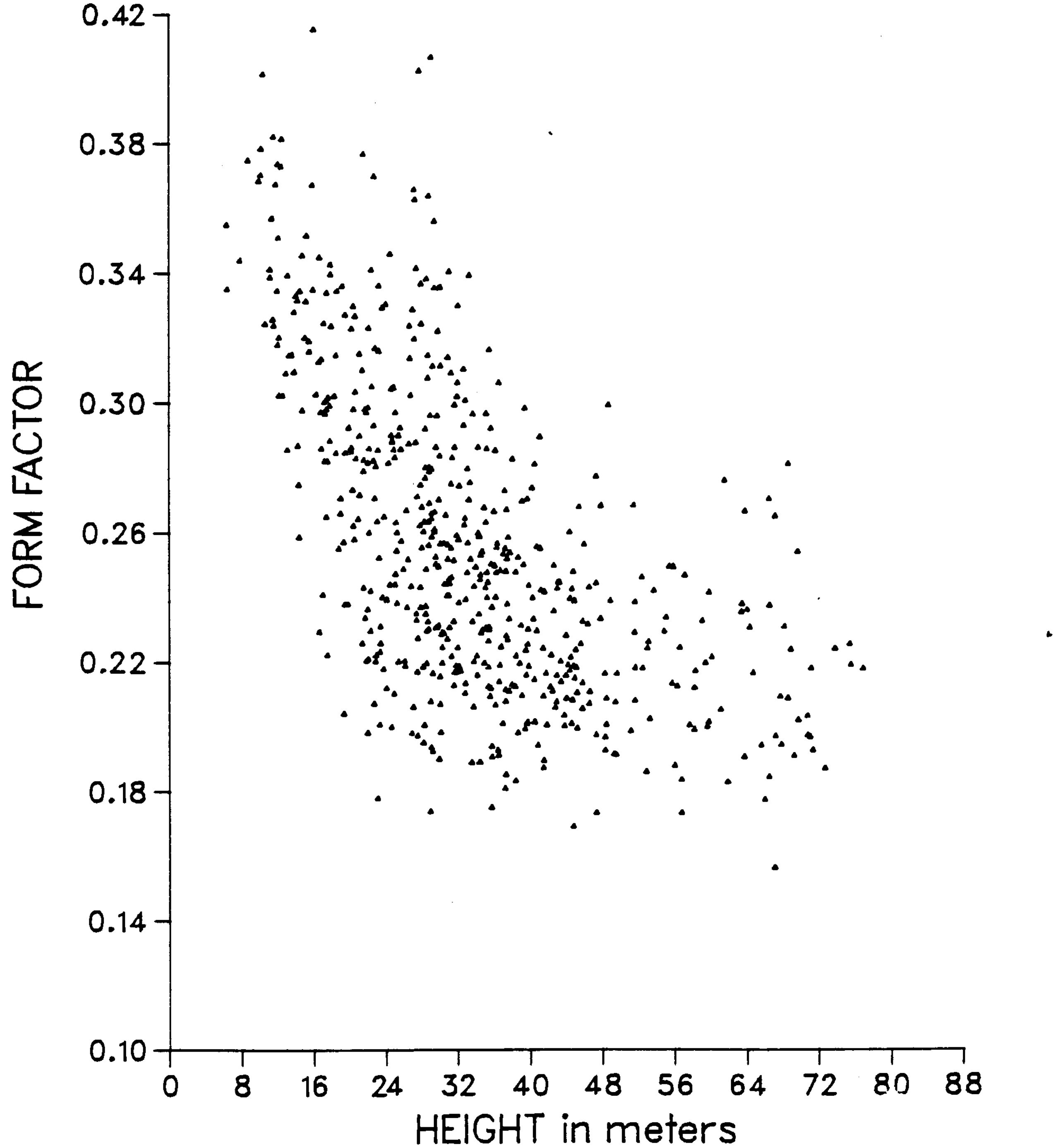


Figure III.5
Plot of Log(Form factor) vs. Log(DBH)

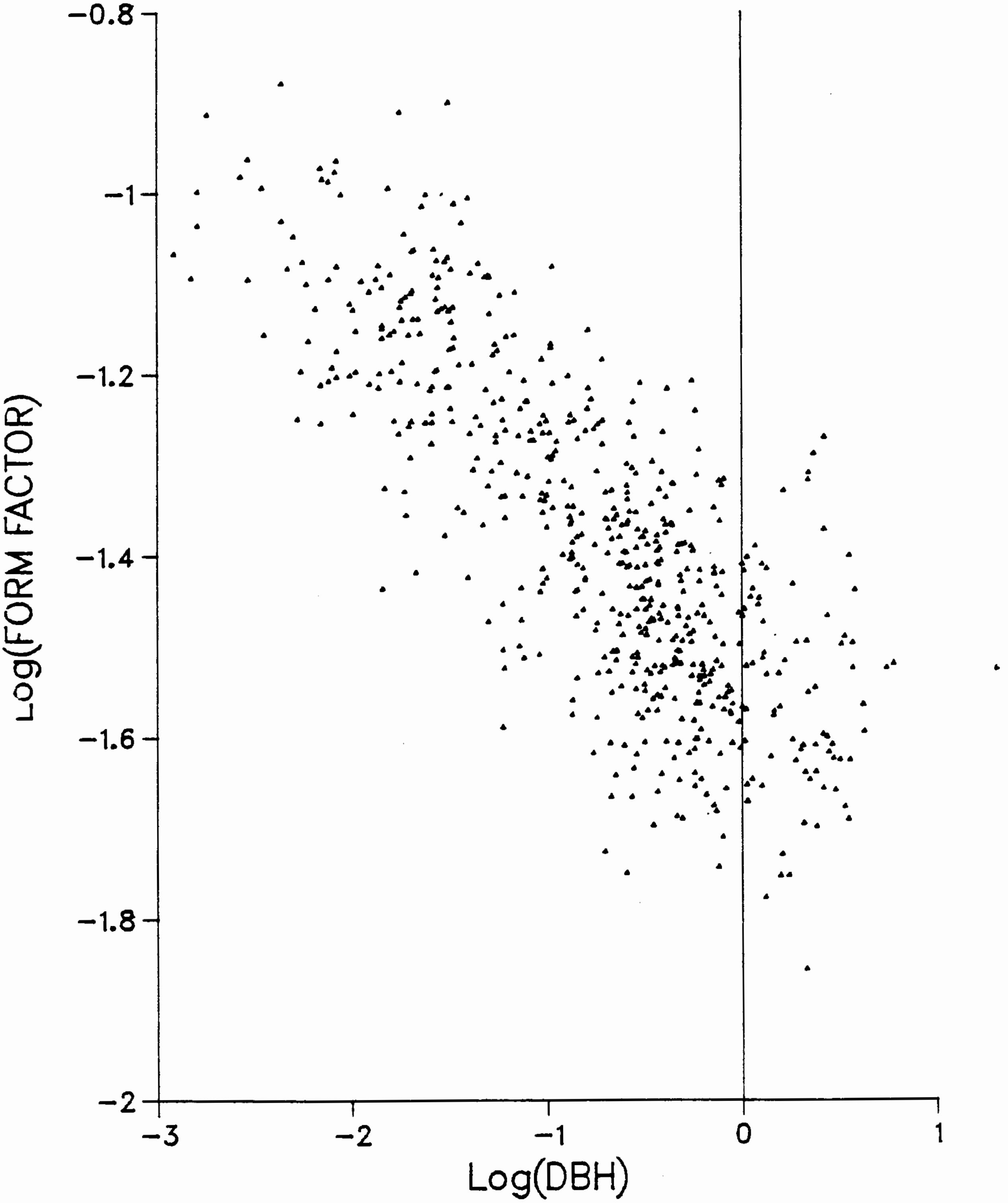
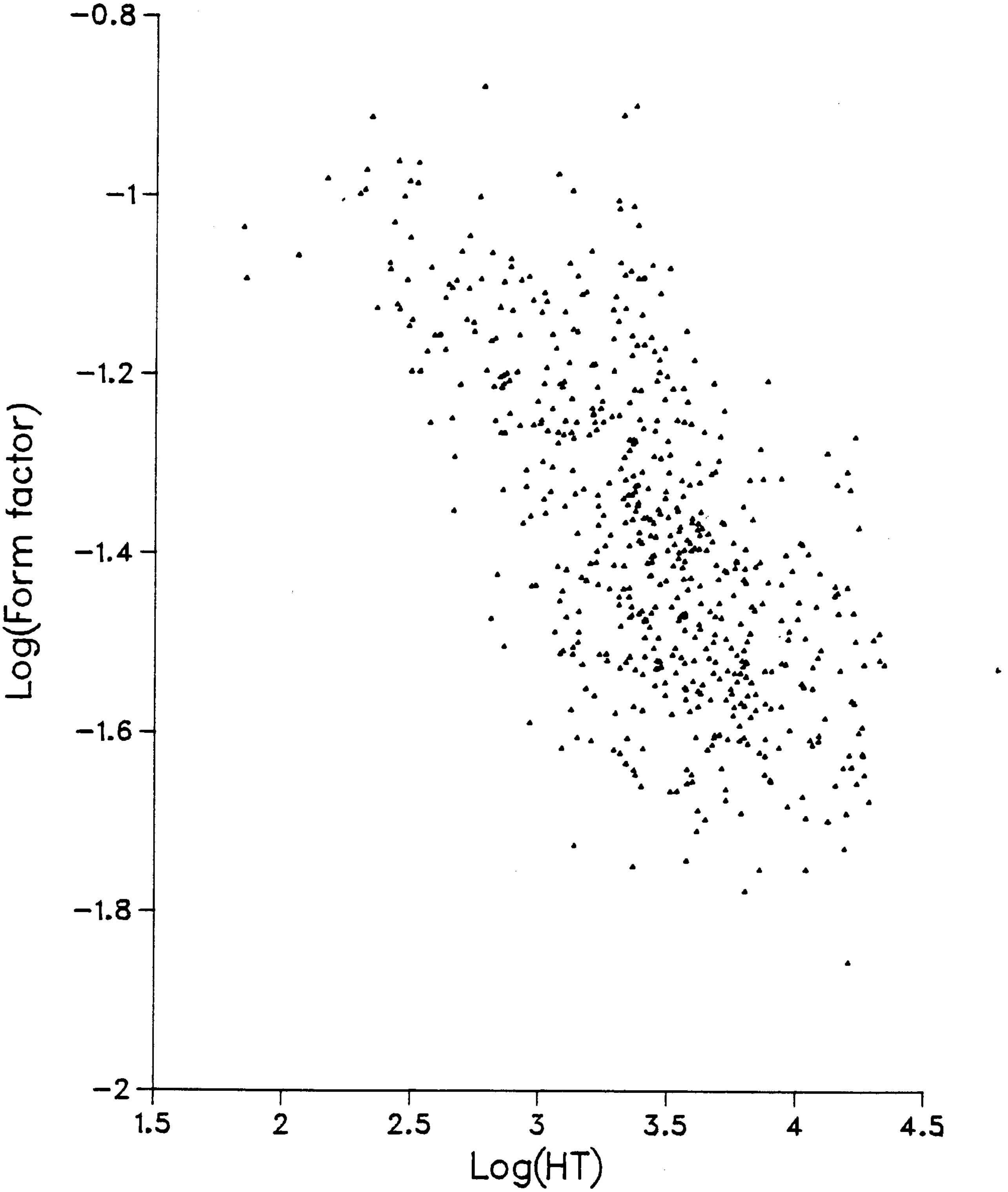


Figure III.6
Plot of Log(Form factor) vs. Log(HT)



We shall use equation (3.1) as the basis for the proposed model but shall permit the form factor to be a function of DBH and HT. At this point, it is difficult to know the exact functional dependence of f on DBH and HT. We could however, start with the observation that the form factor is influenced by the overall size of the tree. Large trees are shaped more like a cone while small trees are shaped more like a paraboloid. This was indicated by Reed and Byrne (1985) when they noted that trees having HT and DBH ratio of 30:1 are formed like a cone while those having a higher ratio of 90:1 are formed like a paraboloid. The dependence of the form factor to tree size as characterized by HT and DBH is also evident in Figures III.1 and III.2. It is then reasonable to assume that the functional form of f be given by

$$f = \lambda \ DBH^{\gamma} HT^{\theta} \tag{3.2}$$

Substituting equation 3.2 to equation 3.1 gives the following,

$$\begin{aligned} V_{tot} &= f \ DBH^2 HT \\ &= \lambda \ DBH^{\gamma} HT^{\theta} (DBH^2 HT) \\ &= \lambda \ DBH^{\gamma+2} HT^{\theta+1} \end{aligned}$$

or,

$$V_{tot} = \lambda_1 \ DBH^{\alpha_1} HT^{\beta_1} \tag{3.3}$$

where

$$\lambda_1 = \lambda$$

$$\alpha_1 = \gamma + 2$$

$$\beta_1 = \theta + 1$$

Methods to fit equation 3.3 to data include non-linear regression on the raw results or linear regression on the logarithmically transformed data. The main drawback in the use of the least-squares method in solving the non-linear equation (3.3) is the non-homogeneity of variance. Cunia (1964) had shown that the variance is usually a function of the quantity DBH^2xHT (see Figure III.2). Thus, the deviation from the true regression function of the volume of large trees has a disproportionate effect on the estimation of the parameters by the method of least squares. Cunia suggested that the best set of weights for equation (3.3) is $1/z$, where $z = (DBH^2HT)$ in keeping with the Gauss-Markov theorem and the dependence of the variance on DBH^2HT. Furthermore, the least-squares regression coefficient of variable V_{tot}/z is equivalent to the weighted least squares regression coefficient of V_{tot}.

The logarithmic transformation stabilizes such a variance (Draper and Smith, 1981). However, fitting a logarithmic form introduces bias in the predicted volume (Husch, *et al.*, 1972). Bias or systematic error results from the fact that the transformed regression equation passes through the arithmetic mean of the independent and dependent variables which are the geometric means of the original variables (Avery and Burkhart, 1983). The prediction bias is introduced since the arithmetic means are always greater than the geometric means.

On the other hand, the simplicity of linear regression on transformed data makes it a more appealing method than the weighted least squares method. The logarithmic transformation simultaneously linearizes the model and stabilizes the variance. Despite the bias induced by the logarithmic transformation, the resulting prediction errors have been found to be smaller than the errors which resulted in using taper-based volume equations. This was shown by Parresol, et. al. (1985) in comparison of several volume equations most of them taper-based, for predicting total volume of Bald Cypress trees (*Taxodium distichum* (L.) Rich.).

Taking the logarithmic form of equation (3.3) gives

$$\log(V_{tot}) = \gamma_1 + \alpha_1 \log(DBH) + \beta_1 \log(HT) \qquad (3.4)$$

where $\gamma_1 = \log(\lambda_1)$. Equation 3.4 is the model used for fitting the data. The coefficients γ_1, α_1, and β_1 are estimated by the method of least squares. Total volume is then estimated by

$$V_{tot} = \lambda_1 DBH^{\alpha_1} HT^{\beta_1}.$$

The logarithmic form of the volume equation is the same as the one proposed by Schumacher and Hall (1933) as discussed in Chapter I. However, in the estimation of merchantable volume, there is a difference between the prediction scheme proposed in this project and that proposed by Schumacher and Hall (1933).

III.2. Predicting Merchantable Volume

Fitting a separate regression model directly to predict merchantable as opposed to total volume in terms of HT and DBH would produce inconsistent estimates for some trees. (The inconsistency arises when the predicted merchantable volume is larger than the predicted total volume). Foresters often insist on obtaining consistent estimates for merchantable volume. Also in this approach, we need to fit new regressions every time the merchantability standard changes.

To estimate the volume of the merchantable portion, we only need to estimate the volume of the unusable portion near the top of the bole. This is a relatively small volume compared to the total tree volume. An illustration in Figure III.7 shows the plot of the percentage of the volume above the given diameter d to the total volume for each of the 581 Coastal Douglas Fir trees. If the merchantable diameter limit is set at 20 cm, it can be inferred from the same plot that the maximum top volume is about 10% of the total volume. Thus an error of of about 20% in the estimate of the top volume would induce a maximum error of 2% in the estimate of the total merchantable volume.

What is needed then is a rough estimate of the shape of the top part of the tree. This could be approximated by a simple geometric solid such as a cone, parabola, or a paracone. If D_m is the specified merchantable diameter limit, H_m is the length from the top of the tree to the diameter limit, then the general form of the top volume V_{top} of the solids is

$$V_{top} = k\ D_m{}^2 H_m.$$

In the equation for top volume, the independent variable is H_m since D_m is set at a fixed value. Furthermore, H_m depends on HT and DBH. Figures III.8 and III.9 show the plot of H_m vs. HT and DBH respectively, while Figures III.10 and III.11 are the corresponding plots in logarithmic scale. If the functional dependence of H_m is assumed to be

$$H_m = c\ DBH^{\alpha_2}\ HT^{\beta_2},$$

where c is the constant of proportionality, then we can then model V_{top} as

$$V_{top} = \lambda_2 DBH^{\alpha_2}\ HT^{\beta_2} \qquad , (3.6)$$

where $\lambda_2 = k \times c$. Transforming equation 3.6 to its logarithmic form , we obtain

$$\log(V_{top}) = \gamma_2 + \alpha_2 \log(DBH) + \beta_2 \log(HT) \qquad (3.7)$$

where $\gamma_2 = \log(\lambda_2)$.

The parameters of the regression equation 3.7 are also obtained by the method of least squares.

The merchantable volume V_m is estimated by taking the difference between the predicted total and the unusable top volumes.

Figure III.7
Percentage of Top Volume To Total Volume Above Given Diameter for Coastal Douglas Fir

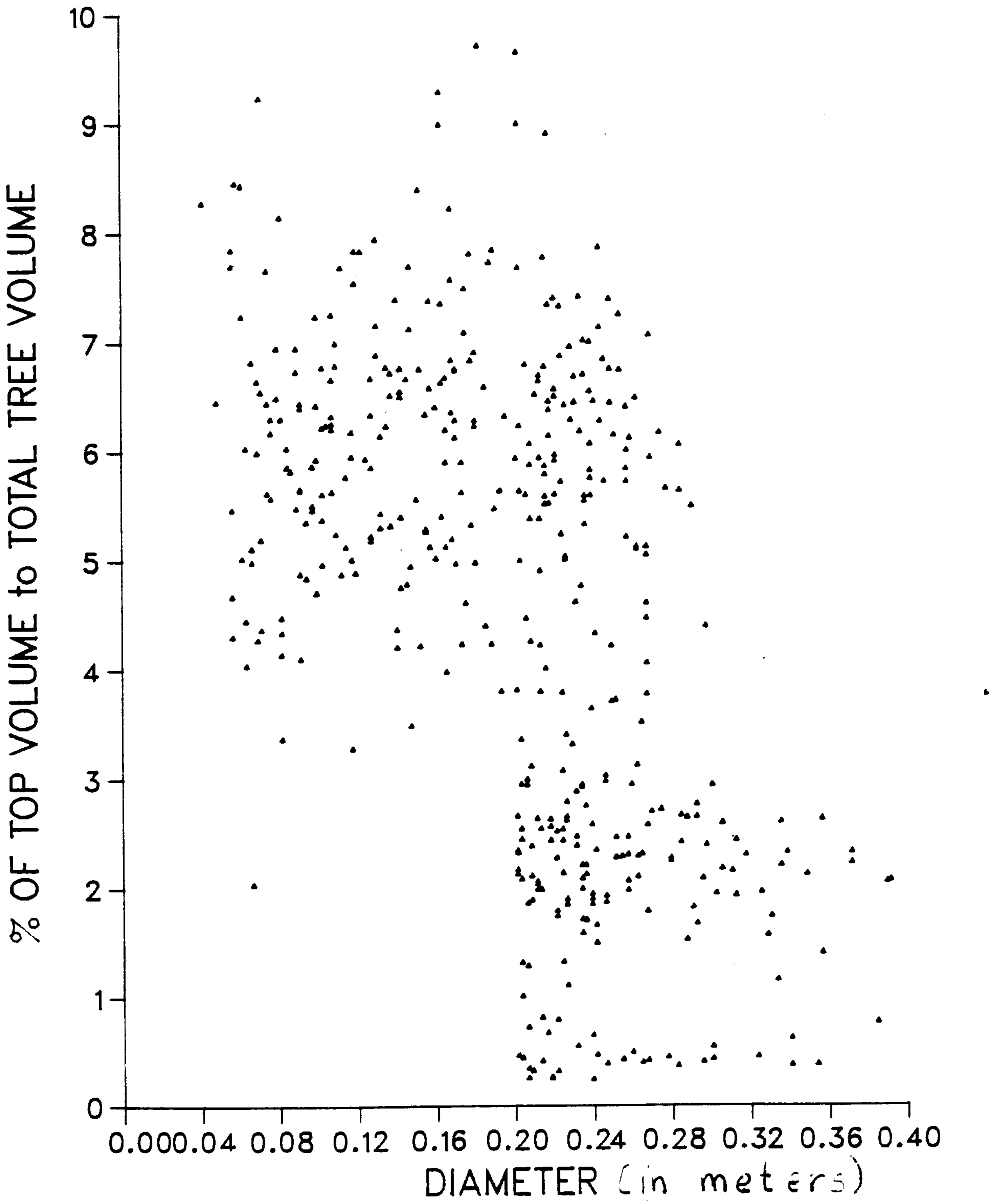


Figure III.8
Plot of hm vs. DBH

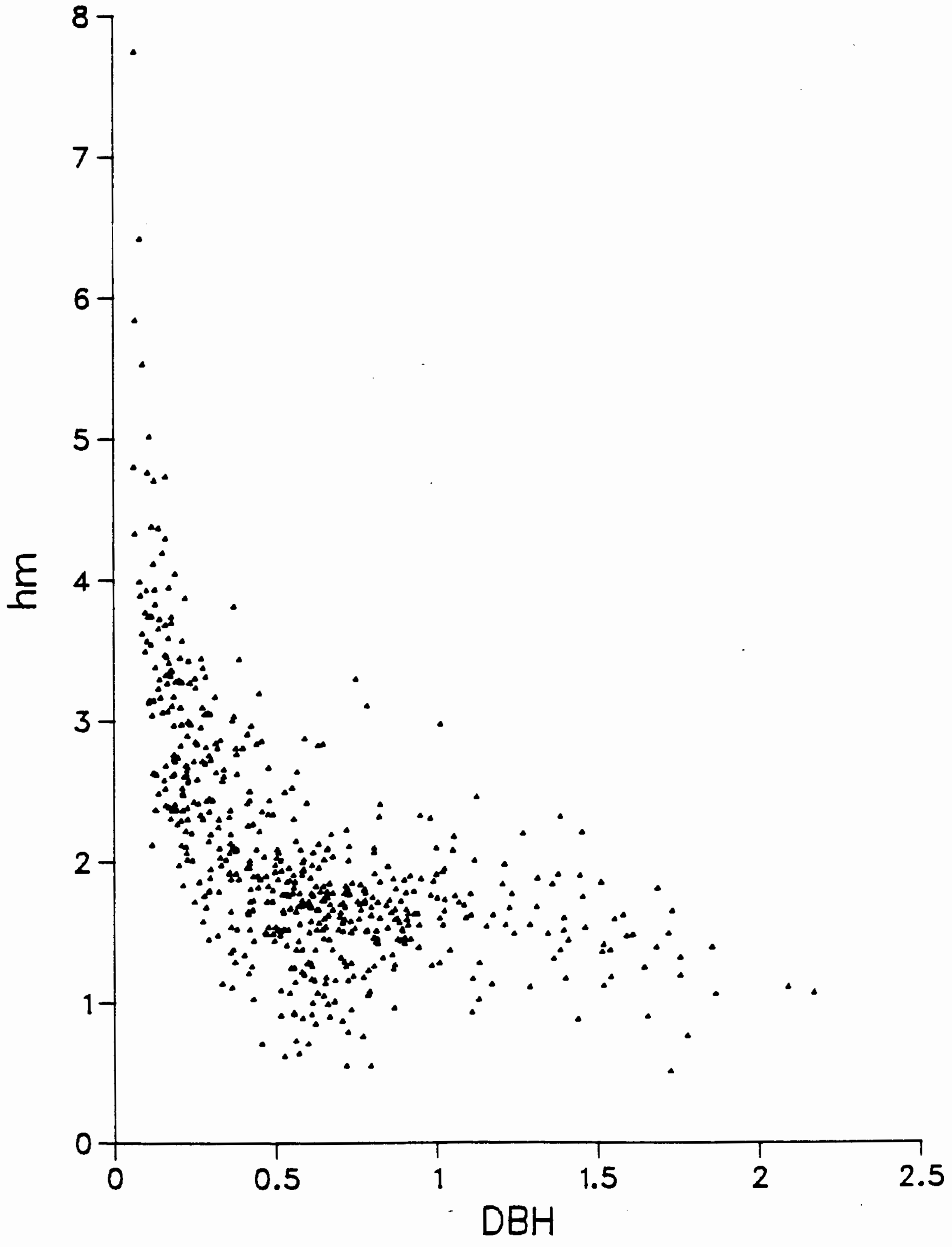


Figure III.9
Plot of hm vs. HT

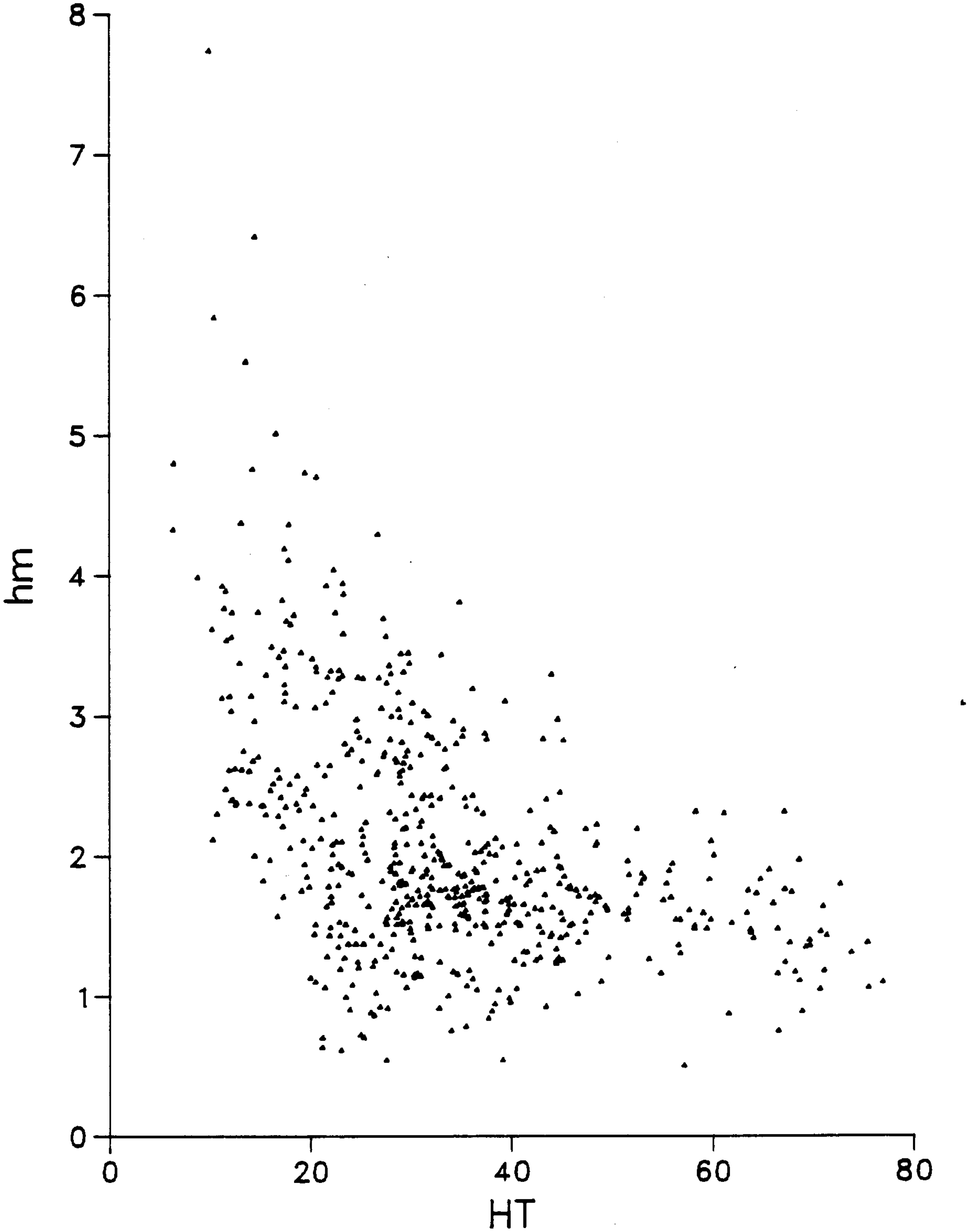


Figure III.10
Plot of Log(hm) vs. Log(DBH)

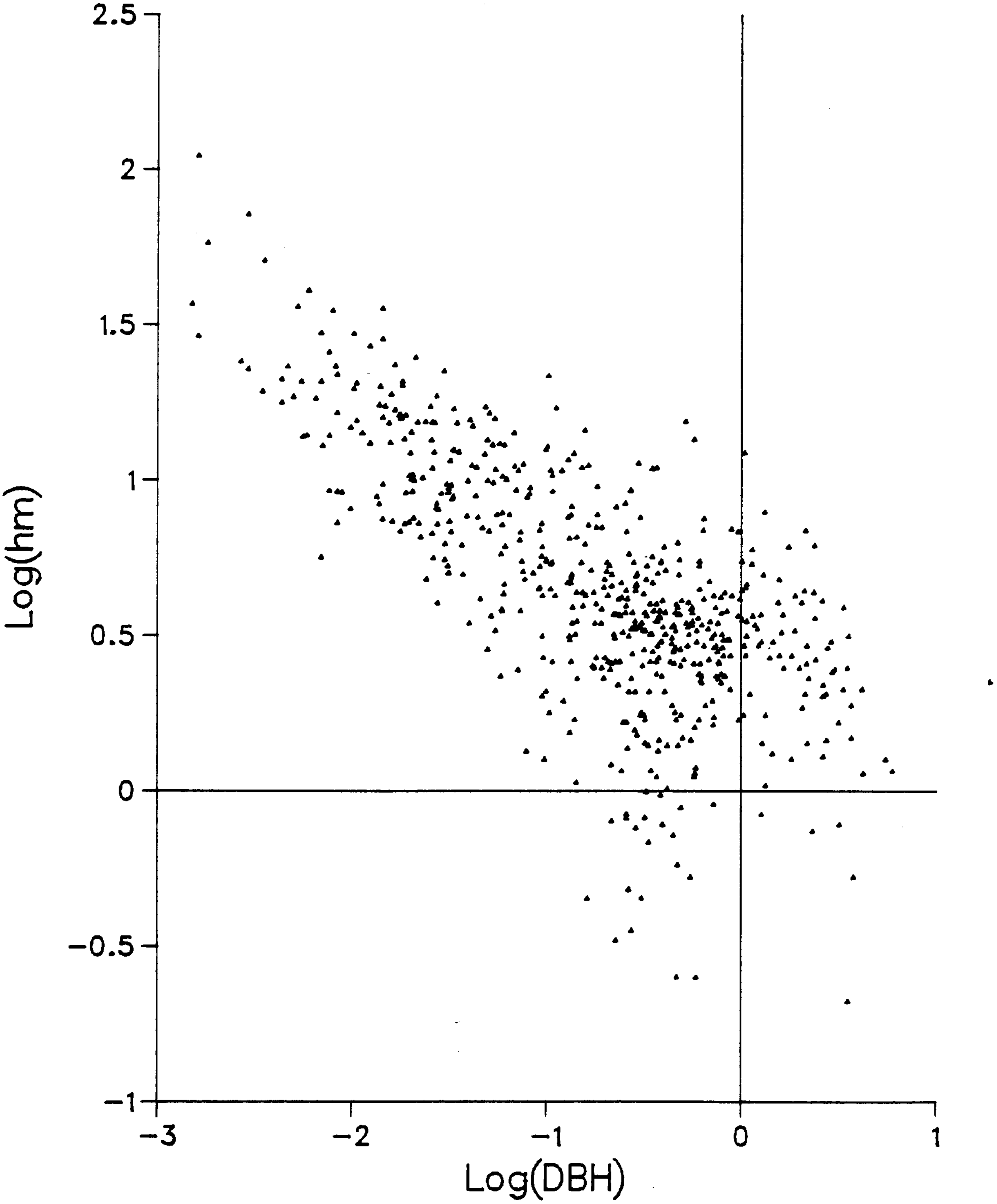
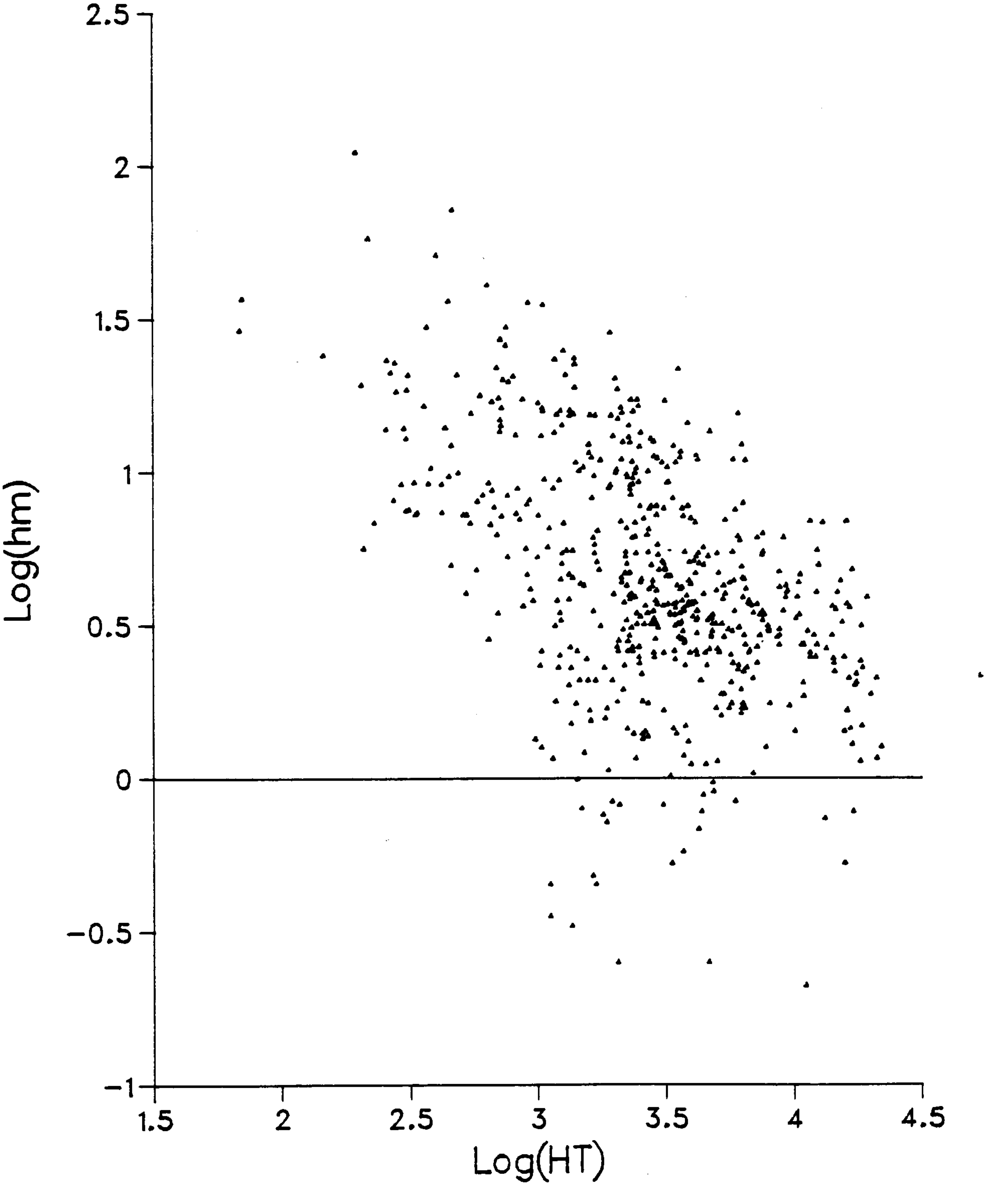


Figure III.11
Plot of Log(hm) vs. Log(HT)



CHAPTER IV

FITTING THE MODEL EQUATIONS

As discussed in the previous chapter, the total tree volume is estimated using

$$V_{tot} = \lambda_1 DBH^{\alpha_1} HT^{\beta_1} e^{\epsilon}$$

where the parameters are derived by fitting the regression model

$$\log(V_{tot}) = \gamma_1 + \alpha_1 \log(DBH) + \beta_1 \log(HT) + \epsilon.$$

Similarly, the parameters for the unusable top volume equation are derived by fitting the regression model,

$$\log(V_{top}) = \gamma_2 + \alpha_2 \log(DBH) + \beta_2 \log(HT) + \epsilon.$$

The parameters in the regression model are estimated by the method of least squares. This method chooses the best fitting plane to be that plane which minimizes the sum of the squares of the distances between the observed and those predicted by the fitted model (Kleinbaum,<u>et. al.</u>, 1978).

Appropriate checks are to be made to assess the adequacy of the regression model. To do this, we shall construct residual plots and consider the following measures:

i) Coefficient of multiple determination R^2 defined by

$$R^2 = 1 - \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 \Big/ \sum_{i=1}^{n} (Y_i - \overline{Y})^2$$

and

ii) Root mean square error $\sqrt{MSE}$ defined by the equation

$$\sqrt{MSE} = [\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 / (n-3)]^{1/2}$$

where

$$Y_i = \log(V_{tot})_i$$

$$\hat{Y}_i = \log(\hat{V}_{tot})_i$$

$$\bar{Y} = (\sum_{i=1}^{n} Y_i)/n$$

R^2 indicates the proportionate reduction of the variance in $\log(V_{tot})$ due to using log(DBH) and log(HT) to predict $\log(V_{tot})$. When the assumptions of the regression model are met, the model gives a good fit to the data if R^2 is close to 1 and a poor fit when R^2 is close to 0. When R^2 is 1, all the data points lie on the fitted plane. In the absence of any linear component to the relationship between $\log(V_{tot})$ with log(DBH) and log(HT), R^2 is 0.

The standardized residuals will be plotted against the predicted values. The standardized ith residual is e_i/s where e_i represents the difference between the actual and the predicted value and s is the square root of the mean square error. When the assumptions of the regression model hold, the plot of the standardized residuals is like one that is based on a random sample from a N(0,1) distribution.

Although using studentized residual plot is better in analysing residual plots when the variance of each residual is not constant, this method was not used in this project. In this method, the ith residual is scaled by its standard deviation, $\{(1 - r_{ii})s^2\}^{1/2}$ where r_{ii} is the ith diagonal of the matrix $R = X(X'X)^{-1}X'$, and the X matrix contains the columns of 1's, log(DBH) and log(HT). In a large majority of data set however, the e_i plot and the e_i/s plot tend to reflect the same general features that are found in the more correct studentized residual plots (Draper and Smith, 1981).

Root mean square error $\sqrt{MSE}$ measures the accuracy of the estimates. If the $\sqrt{MSE}$ is 0, the model fits the data perfectly and every observed point lies on the response plane.

Since our purpose is to predict both the total and merchantable volumes, trees having breast height diameter smaller than the specified diameter limit were excluded from the data set. The log volume estimates derived from the remaining data will not apply to trees with DBH similar to those rejected. Although we were truncating our data for predicting total volume, in a way we were simulating an actual forest setting of a more or less homogeneous stand of mature trees, in which the trees are usually of similar sizes and have large DBH.

After the total and top volume were estimated, the merchantable volume was obtained by taking the difference of these two quantities.

Samples of 60 trees were selected randomly from each of the three data sets and the logarithmic volume equation was fitted to these smaller samples. The sampling that was performed was like a double sample since the data sets were samples from a much larger population of trees. This was done to simulate the real situation where a volume formula (or table) is constructed using a possibly small sample from the stand in question or a related one. The DBH and HT are measured for a usually large sample of the stand whose volume is to be estimated.

When the volume equations were constructed using the smaller samples, the transformed equations,

$$\hat{V}_{tot} = \lambda_1 DBH^{\alpha_1} HT^{\beta_1}, \quad \text{and}$$

$$\hat{V}_{top} = \lambda_2 DBH^{\alpha_2} HT^{\beta_1},$$

were used to estimate the total volume and top volume for each tree in the data set for each species. The predicted total and merchantable volumes of all trees belonging to the same species were summed up. The deviations of these sums from the observed totals were calculated.

An assumption was made that the sample trees from a particular species belong to the same stand of trees. In this project, the total stand volume V is referred to as the sum of total volumes of all trees in the stand, while the merchantable stand volume V_m is the sum of all merchantable volumes.

The process of sample selection, constructing volume equations based on the sample and estimating total stand volume as well as merchantable stand volume was simulated 1000 times. The congruential uniform random number generator GGUBS in the IMSLS package was used in selecting the samples.

Bias, which is the mean of the deviations of the predicted from the observed was computed. The root mean square error $\sqrt{MSE}$ and the standard deviation S.D. of the deviations were also calculated. The formulas for these statistics are given by the following:

$$Bias = (\sum_{i=1}^{n} d_i)/n$$

$$\sqrt{MSE} = \{\sum_{i=1}^{n} (d_i)^2/n\}^{1/2}$$

$$SD = \{\sum_{i=1}^{n} (d_i - \bar{D})^2 /n\}^{1/2}$$

where $d_i = V - \hat{V}_i$

$\bar{D}$ = bias

V = observed stand volume

$\hat{V}_i$ = estimated stand volume

n = 1000 (number of simulations)

When dealing with the deviations from the predicted merchantable stand volumes, V_m is substituted for V in the formulas for Bias, $\sqrt{MSE}$, and S.D.

The 1000 simulations were repeated when the volume equations were constructed based on a larger sample of 100 trees. The purpose for this was to determine whether an increase in sample size would have a considerable decrease in the amount of bias, $\sqrt{MSE}$ and S.D.

CHAPTER V

RESULTS

The merchantable diameter limit was specified at 20 cm. Preliminary screening of the data sets indicated that a substantial number of trees had DBH smaller than 20 cm. Table V.1 gives a summary of the original number of trees in the data available and the number remaining after discarding trees with small DBH.

Table V.1: Summary of the Number of Trees with DBH Greater than 20 cm

Tree Species	Number of trees in the data set	Number of trees with DBH > 20 cm
Coastal D.F.	581	463
Interior D.F.	1504	1000
White Spruce	4789	3120

The parameter estimates and the values of $\sqrt{MSE}$ and R^2 that resulted after fitting the log-total volume equation to trees with DBH greater than 20 cm in each species are in Table V.3. The corresponding standardized residual plot for the Coastal Douglas Fir is in Figure V.1, for the Interior Douglas Fir in

Figure V.2 and White Spruce in Figure V.3. As shown in Table V.3, the R^2's are close to 100% and root mean square errors close to 0 for the three species. The standardized residual plots do not indicate any systematic trends. This means that the regression model for log-total volume is appropriate to the logarithmically transformed data. Table V.4 gives the standard deviations of the parameter estimates.

The parameter estimates and the fit statistics for the log-top volume regression model are in Table V.5, while the standard deviation of the estimates are in Table V.6. The regression model does not seem to give as good a fit; R^2 ranges from 72.9% to 81.0% for the three species. The residual plots in Figures V.4 to V.6 reveal some mild heteroscedasticity. The residuals corresponding to small predicted log volumes are more widely spread.

Back transformation of the fitted model, i.e. taking exponentials on both sides of the fitted equations, gives the estimated volume equations. These were then employed to estimate the total and merchantable volumes for each tree. The stand volumes were obtained by adding the estimated tree volumes. The observed and estimated total and merchantable stand volumes are summarized in Table V.2.

The fitted equations underestimated both the total and merchantable stand volumes. The presence of the bias can be attributed to using linear regression on the transformed data since transformation introduces bias in the predicted volumes (Husch, et al., 1972). In our problem, the value of DBH and HT are to be measured for many trees, the volume of each is to be estimated, and these estimates are to be summed. Hence, a systematic error present in the estimate for each tree will be present in every single term in the total. However, despite the presence of the systematic errors, the overall bias for the predicted stand volumes are relatively small compared to the actual stand volume. Following the results on Table V.2, the bias for total stand volume was 2.0% of the observed stand volume for Coastal Douglas Fir, .94% for Interior Douglas Fir and .19% for White Spruce. The biases for the merchantable stand volumes were 2.09% for Coastal Douglas Fir, .91% for Interior Douglas Fir and .14% for White Spruce.

The previous results were obtained by constructing the volume equations based on all the sample trees on each species. A smaller sample of 60 trees were selected randomly to construct the volume equations and were applied to estimate total and merchantable volumes for the whole stand. This was done 1000 times. Each time, the deviations of the estimated total and merchantable stand volumes were computed. The histograms of the deviations of the estimated total stand and merchantable volumes for Coastal Douglas Fir trees are given in Figure V.7, for Interior Douglas Fir trees in Figure V.9, and for White Spruce in Figure V.11. Bias, root mean square error ($\sqrt{MSE}$), and standard deviation (SD) of all the deviations are in Table V.7

for the stand volume estimates of the three species. The overall bias, i.e. the mean of all the stand deviations that resulted after the 1000 simulations was negative. This result again demonstrated the presence of the systematic error and that the volume equations underestimated the stand volume.

The simulations were repeated with the sample size used to fit the regression model increased from 60 to 100. The histogram of the deviations for the total and merchantable stand volume are in Figure V.8 for Coastal Douglas Fir, Figure V.10 for Interior Douglas Fir and Figure V.12 for White Spruce. The corresponding bias, $\sqrt{MSE}$ and SD are in Table V.8. An increase in sample size did not give any significant reduction in the amount of bias. However, there was a significant reduction in the chance error as given by a smaller $\sqrt{MSE}$.

Figures V.7 to V.12 indicate that the distribution of the deviations of the estimated total and merchantable stand volumes are approximately symmetric. The mean or the overall bias and SD of each distribution are in Tables V.7 and V.8. A normal probability plot of the 1000 deviations was also done and revealed a linear pattern. The deviations approximately follow a normal distribution and the results indicated different means and variance for each species. The deviations of the merchantable volume estimates are more or less of the same magnitude and followed the same symmetric patterns as the total stand volume. Although there was concern about the presence of a mild heteroscedasticity after fitting the log-top volume in

Figures V.4 to V.6, the overall bias in the estimate of the merchantable stand volume was smaller than the bias of the total stand volume (see Tables V.7 to V.10). This result is to be expected since the only difference between the top and the merchantable volumes is the unusable top volume. This is a small proportion of the overall volume.

Table V.2: Summary of Observed and Predicted Total and Merchantable Stand Volumes (in cubic meters)

	Coastal Douglas Fir	Interior Douglas Fir	White Spruce
Total Stand Volume	3144.8	1592.9	5163.5
Predicted Stand Volume	3081.9	1577.9	5153.6
% Error	2.0%	0.94%	0.19%
Merchantable Stand Volume	3101.6	1485.5	4735.7
Predicted Merchantable Volume	3036.8	1471.9	4728.8
% Error	2.1%	0.91%	0.15%

Table V.3: Regression Coefficients, R^2 and $\sqrt{MSE}$ of Logarithmic Equation for Total Volume:

$$\log(V_{tot}) = \gamma_1 + \alpha_1 \log(DBH) + \beta_1 \log(HT) + \log(\epsilon)$$

Tree Type	γ_1	α_1	β_1	R^2	$\sqrt{MSE}$
Coastal D. F.	-2.3903	1.6853	1.2254	99.2%	0.1079
Interior D.F.	-1.9529	1.7880	1.1334	98.4%	0.1125
White Spruce	-1.8147	1.7642	1.1269	99.0%	0.0837

Table V.4: Standard Deviation of Estimates

Tree Type	$SD(\gamma_1)$	$SD(\alpha_1)$	$SD(\beta_1)$
Coastal D.F.	0.1139	0.0193	0.0296
Interior D.F.	0.0678	0.0193	0.0296
White Spruce	0.0437	0.0079	0.0110

Table V.5: Regression Coefficients, R^2 and $\sqrt{MSE}$ of Logarithmic Equation for Top Volume

$$\log(V_{top}) = \gamma_2 + \alpha_2 \log(DBH) + \beta_2 \log(HT) + \log(\epsilon)$$

Tree Type	γ_2	α_2	β_2	R^2	$\sqrt{MSE}$
Coastal D. F.	-6.4909	-1.1768	0.9719	72.9%	0.2161
Interior D.F.	-5.8701	-1.1267	0.8183	78.8%	0.1517
White Spruce	-5.6285	-1.0802	0.7524	81.0%	0.1145

Table V.6: Standard Deviation of Parameter Estimates

Tree Type	SD(γ_2)	SD(α_2)	SD(β_2)
Coastal D.F.	0.2280	0.0387	0.0593
Interior D.F.	0.0914	0.0189	0.0249
White Spruce	0.0599	0.0109	0.0151

Table V.7: Fit Statistics for Total Stand Volume Estimates Based on Equations Constructed from 60 Sample Trees Based on 1000 Simulations (expressed as percentages of observed stand volume)

Tree Type	Bias	SD	$\sqrt{MSE}$
Coastal D.F.	-1.91%	2.18%	2.90%
Interior D.F.	-0.79%	1.95%	2.10%
White Spruce	-0.23%	1.52%	1.55%

Table V.8: Fit Statistics for Merchantable Stand Volume Estimates Based on Equations Constructed from 60 Sample Trees Based on 1000 Simulations (expressed as percentages of observed merchantable stand volume)

Tree Type	Bias	SD	$\sqrt{MSE}$
Coastal D.F.	-1.90%	2.22%	2.92%
Interior D.F.	-0.76%	2.15%	2.28%
White Spruce	-0.19%	1.73%	1.75%

Table V.9: Fit Statistics for Total Stand Volume Estimates Based on Equations Constructed from 100 Sample Trees Based on 1000 Simulations (expressed as percentages of observed stand volume)

Tree Type	Bias	SD	$\sqrt{MSE}$
Coastal D.F.	-1.92%	1.61%	2.50%
Interior D.F.	-0.79%	1.52%	1.78%
White Spruce	-0.18%	1.16%	1.17%

Table V.10: Fit Statistics for Merchantable Stand Volume Estimates Based on Equations Constructed from 100 Sample Trees Based on 1000 Simulations (expressed as percentages of observed merchantable stand volume)

Tree Type	Bias	SD	$\sqrt{MSE}$
Coastal D.F.	-1.91%	1.64%	2.51%
Interior D.F.	-0.90%	1.67%	1.90%
White Spruce	-0.13%	1.32%	1.11%

Figure V.1
Plot of Residuals vs. Fitted Log Total Volume for Coastal Douglas Fir Trees

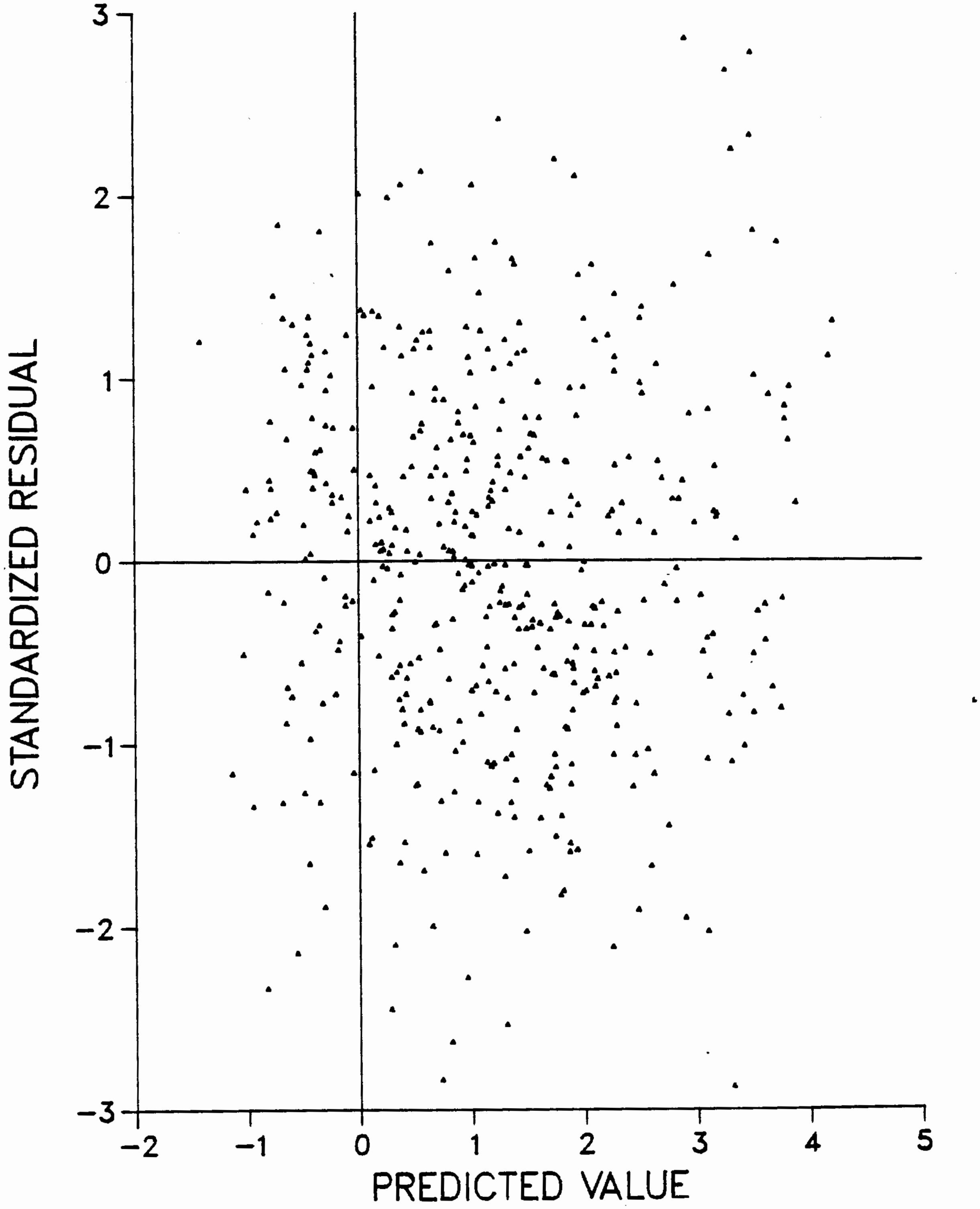


Figure V.2
Plot of Residuals vs. Fitted Log Total Volume
for Interior Douglas Fir Trees

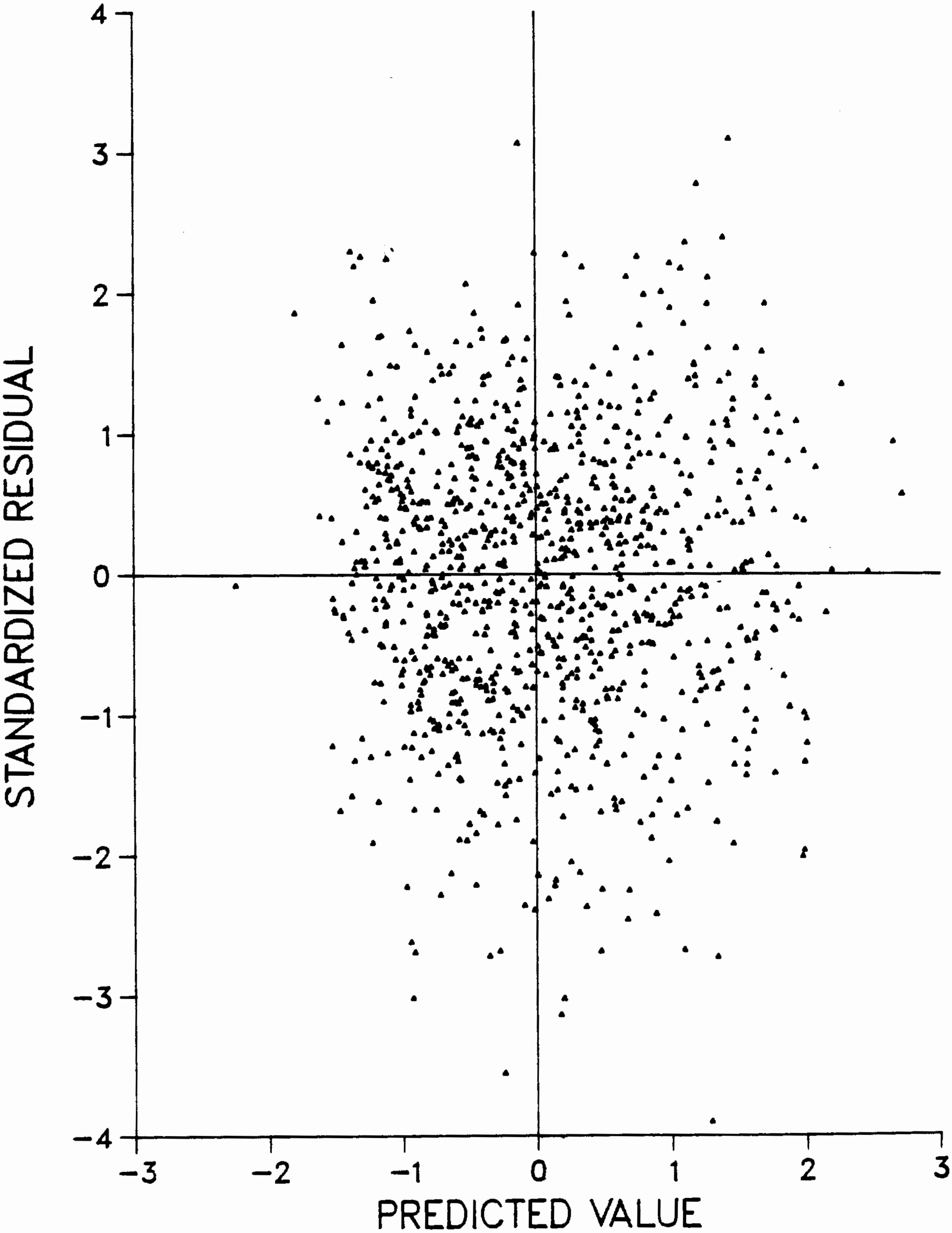


Figure V.3
Plot of Residuals vs. Fitted Log Total Volume
White Spruce

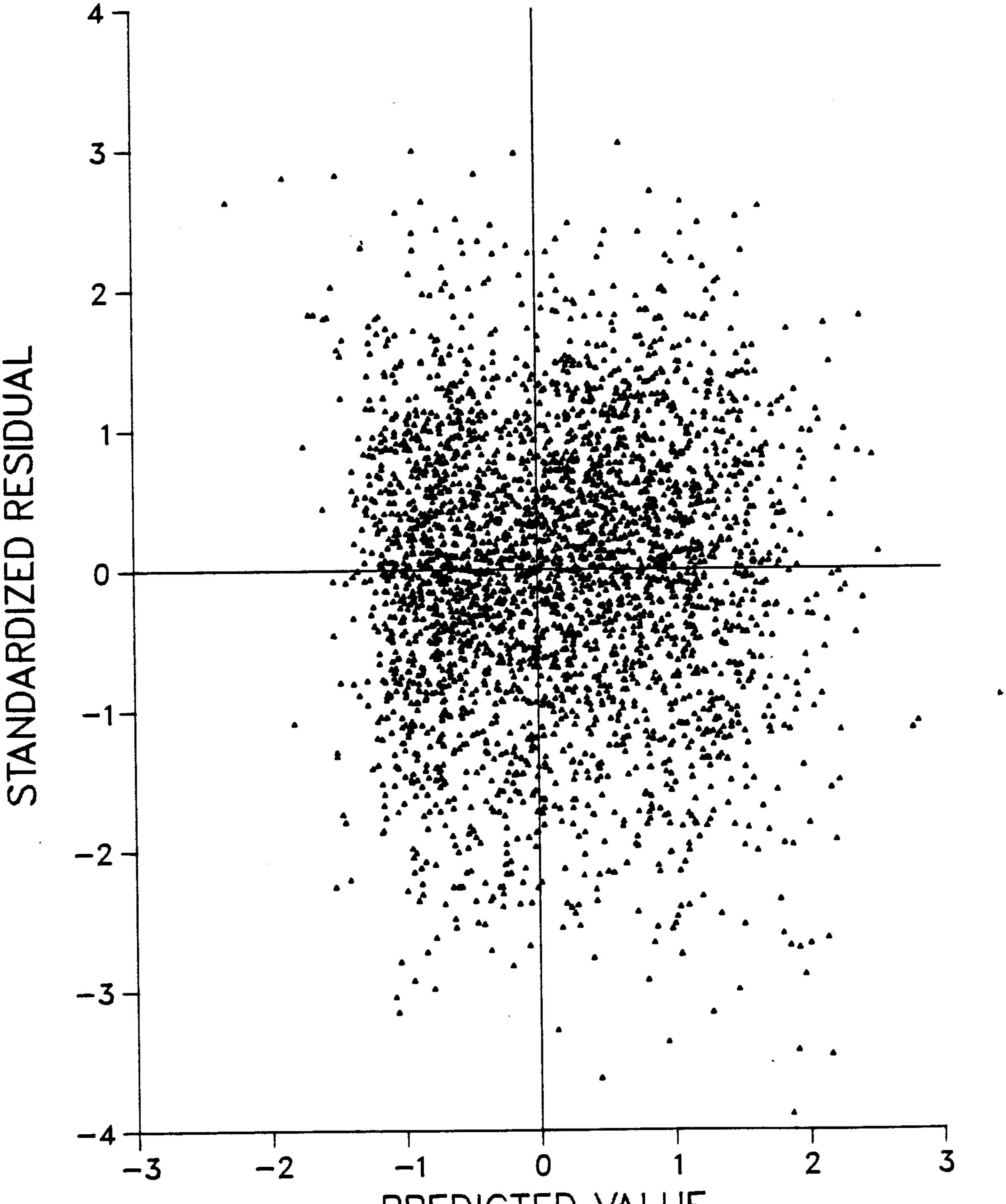


Figure V.4
Plot of Residuals vs. Fitted Log Top Volume for Coastal Douglas Fir Trees

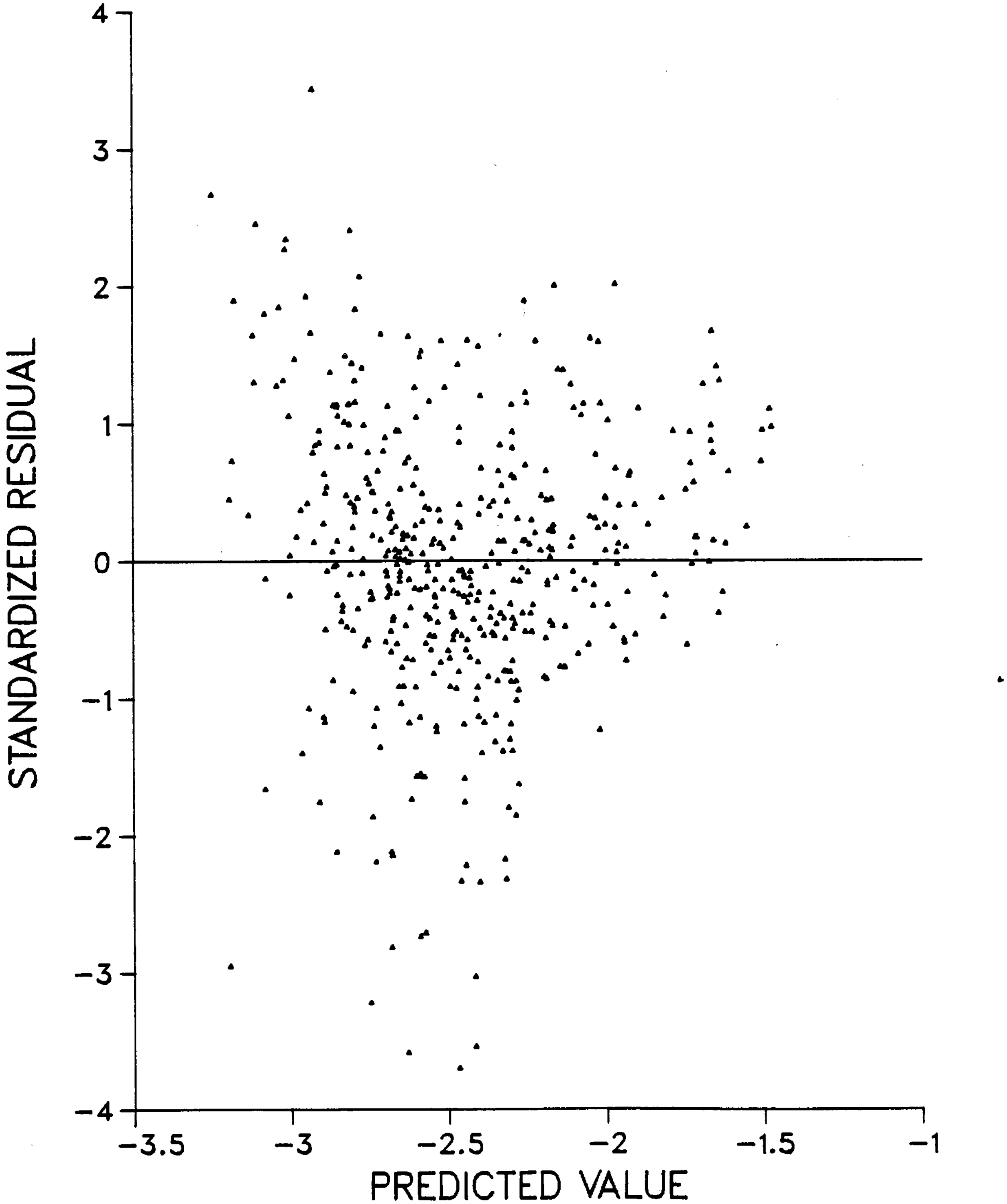


Figure V.5
Plot of Residuals vs. Fitted Log Top Volume for Interior Douglas Fir Trees

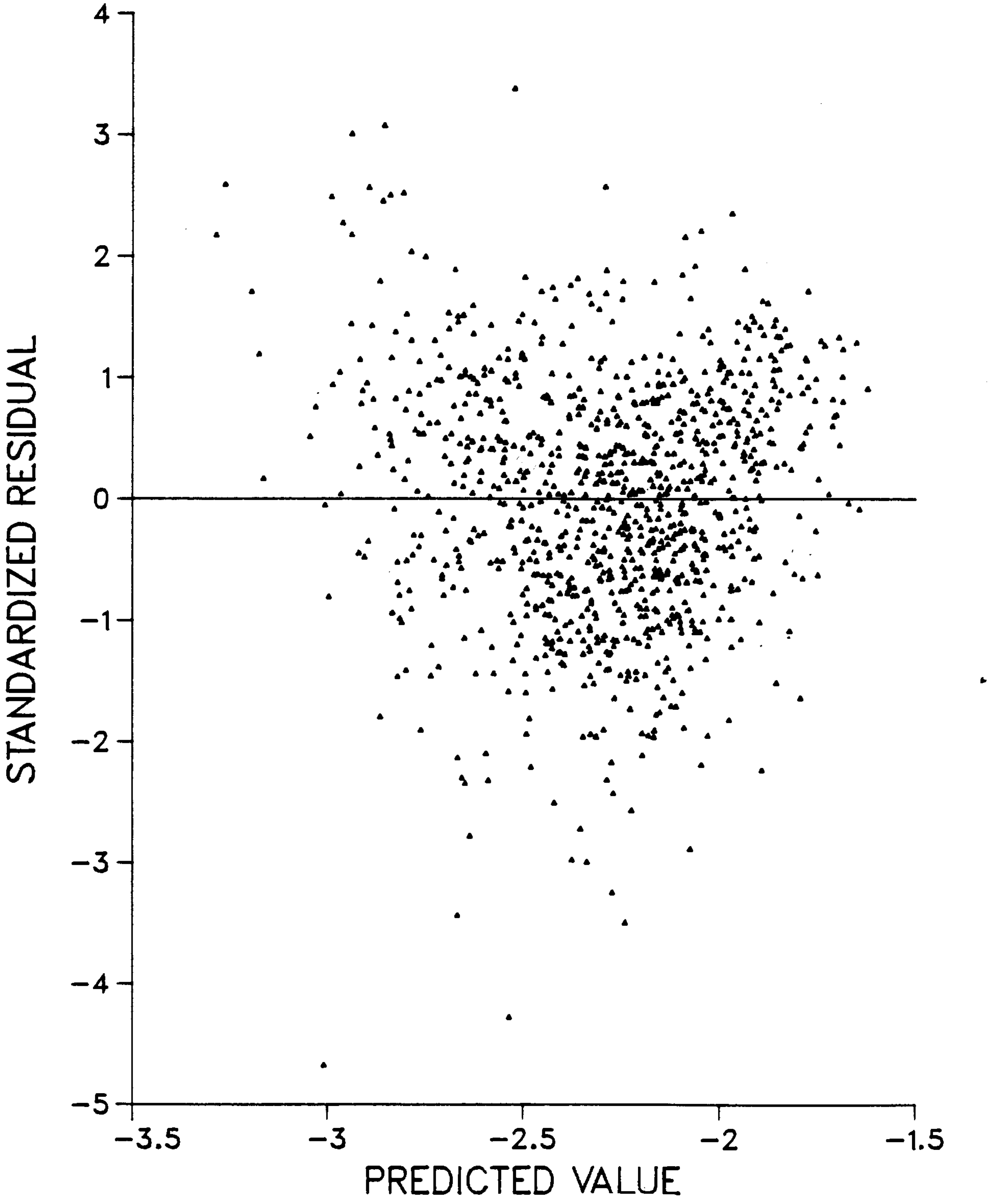


Figure V.6
Plot of Residuals vs. Fitted Log Top Volume
White Spruce

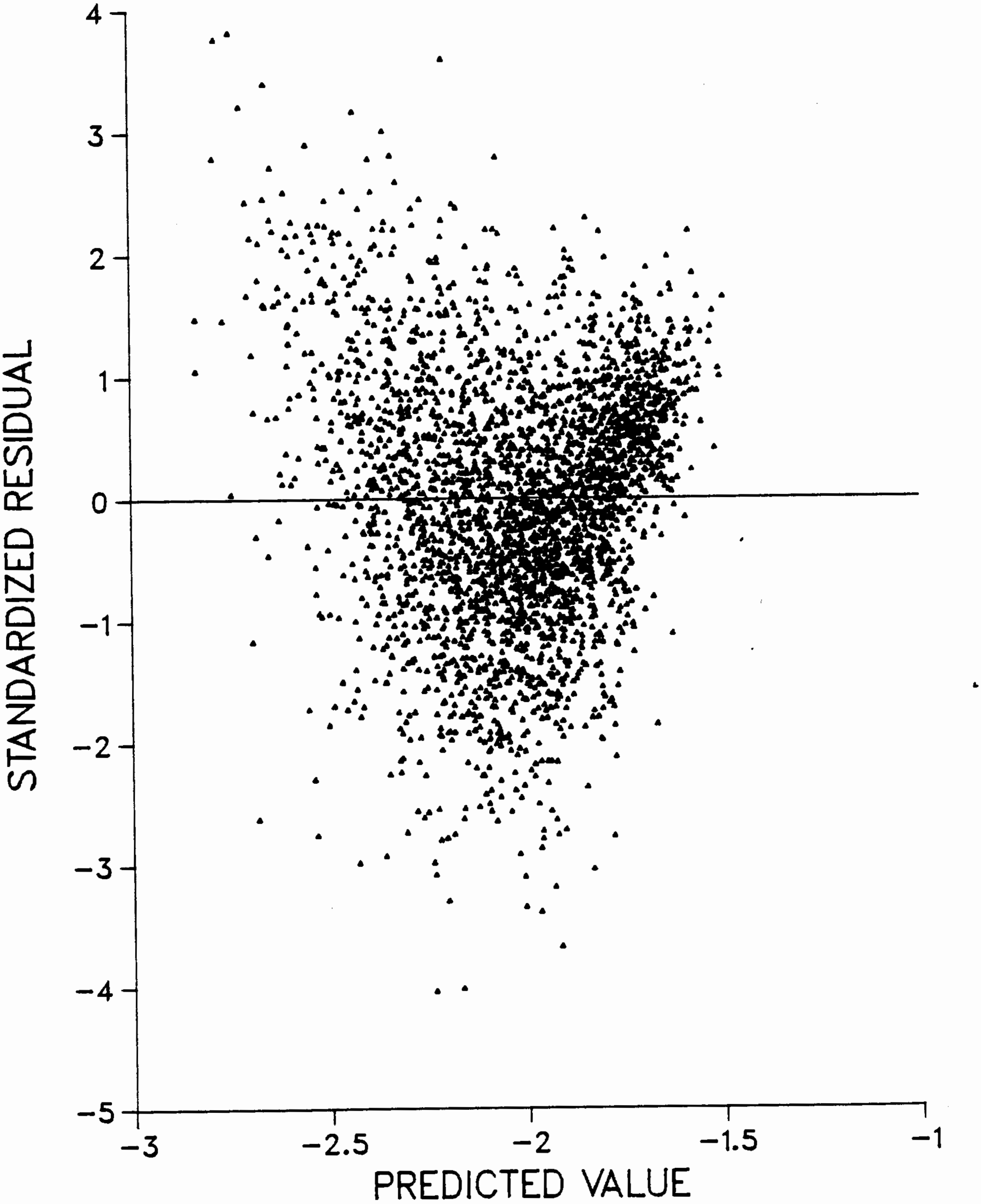


Figure V.7
Histogram of Deviations of the Estimated Total and Merchantable Stand Volumes Based on Equations Constructed from 60 Coastal Douglas Fir Trees

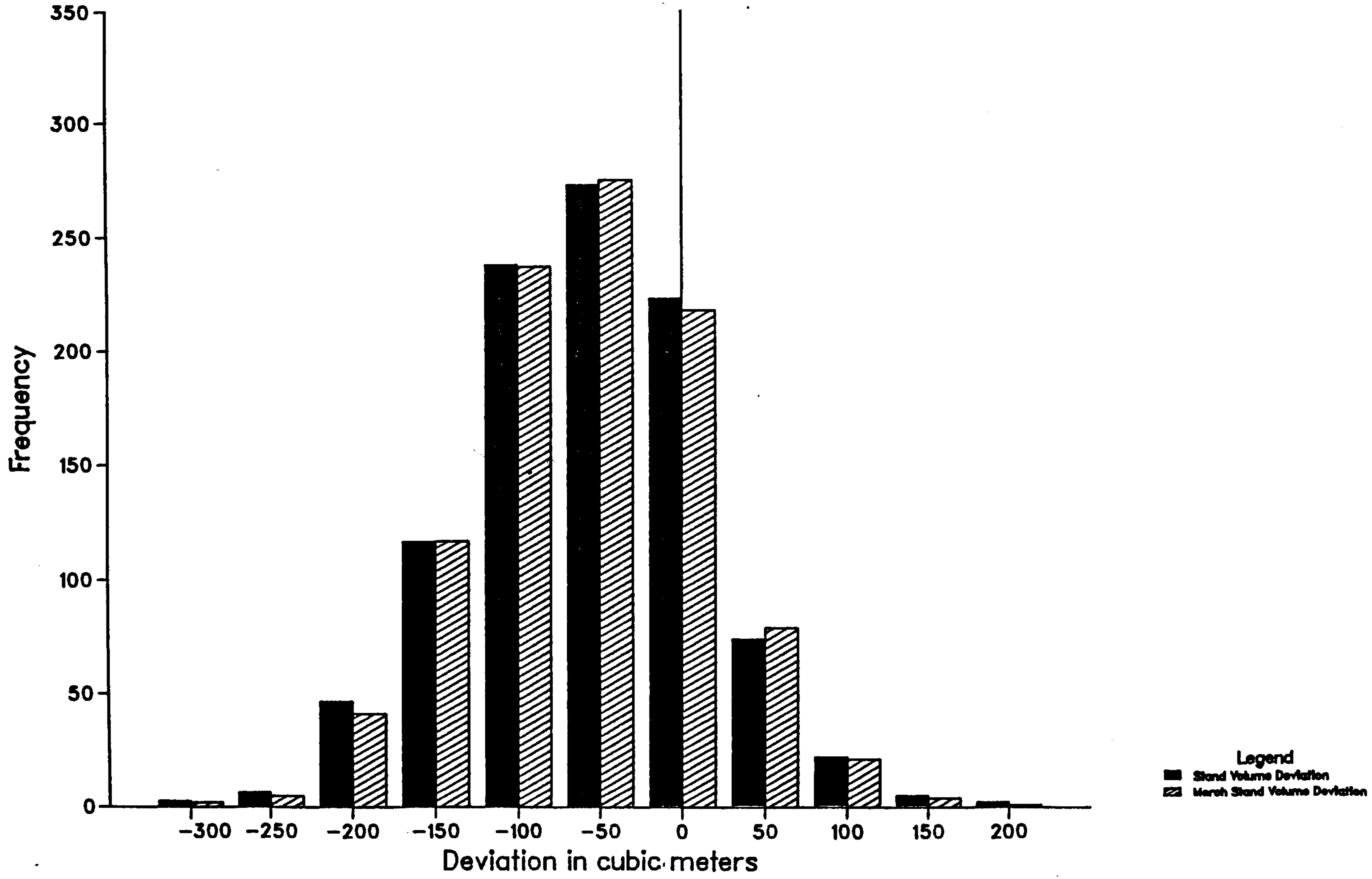


Figure V.8
Histogram of Deviations of the Estimated Total and Merchantable Stand Volumes Based on Equations Constructed from 100 Coastal Douglas Fir Trees

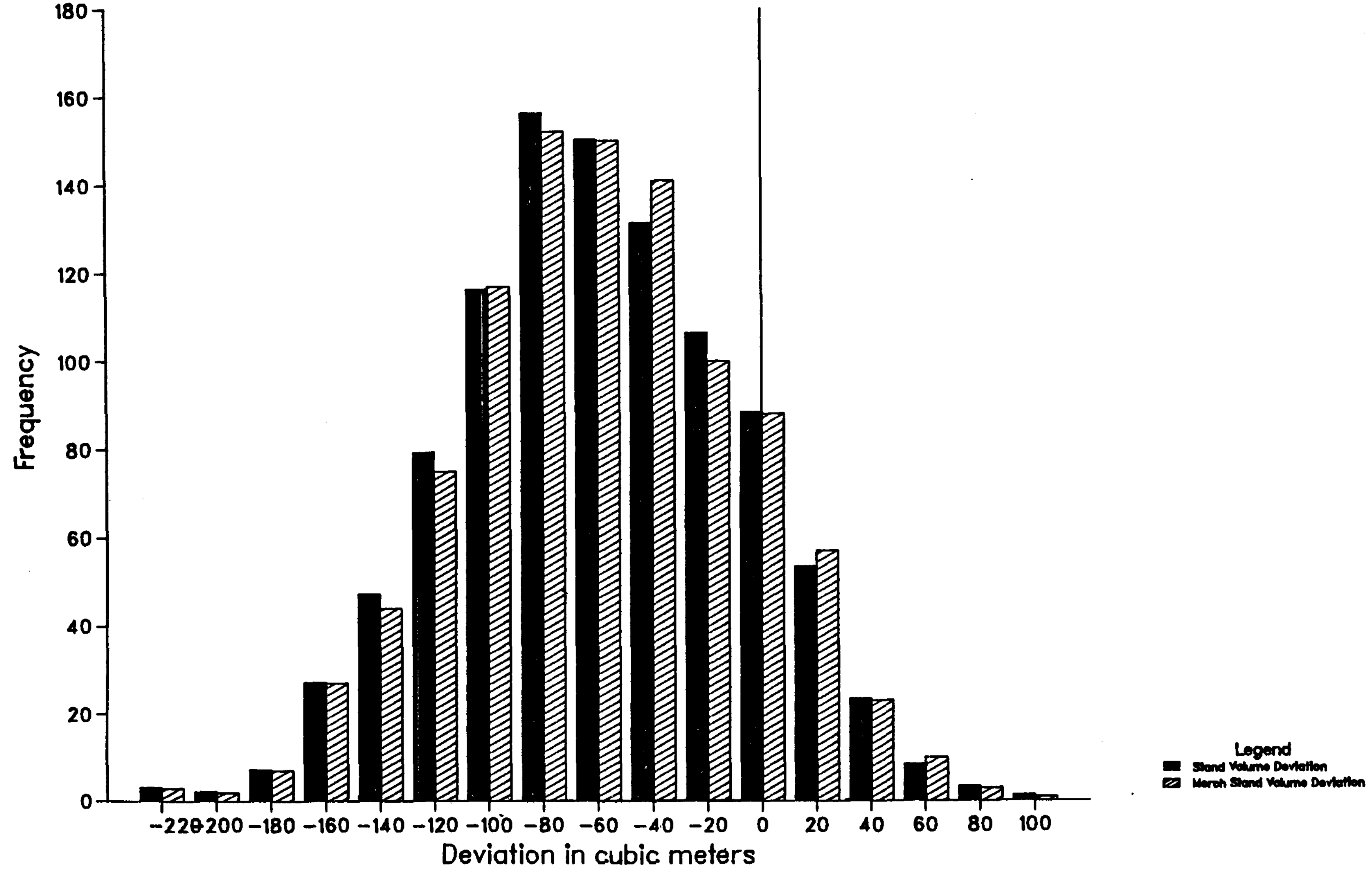


Figure V.9
Histogram of Deviations of the Estimated Total and Merchantable Stand Volumes Based on Equations Constructed from 60 Interior Douglas Fir Trees

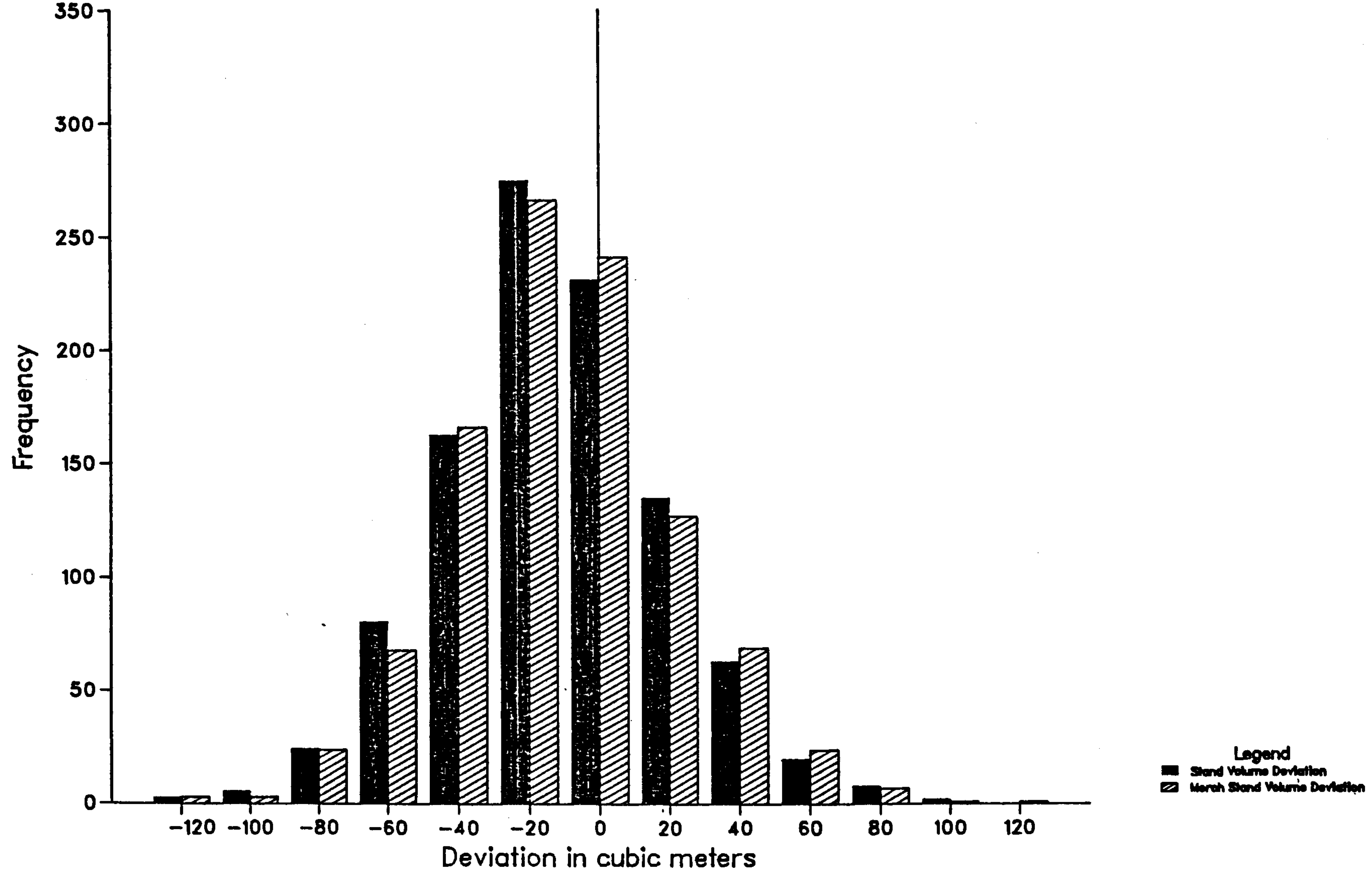


Figure V.10
Histogram of Deviations of the Estimated Total and Merchantable Stand Volumes Based on Equations Constructed from 100 Interior Douglas Fir Trees

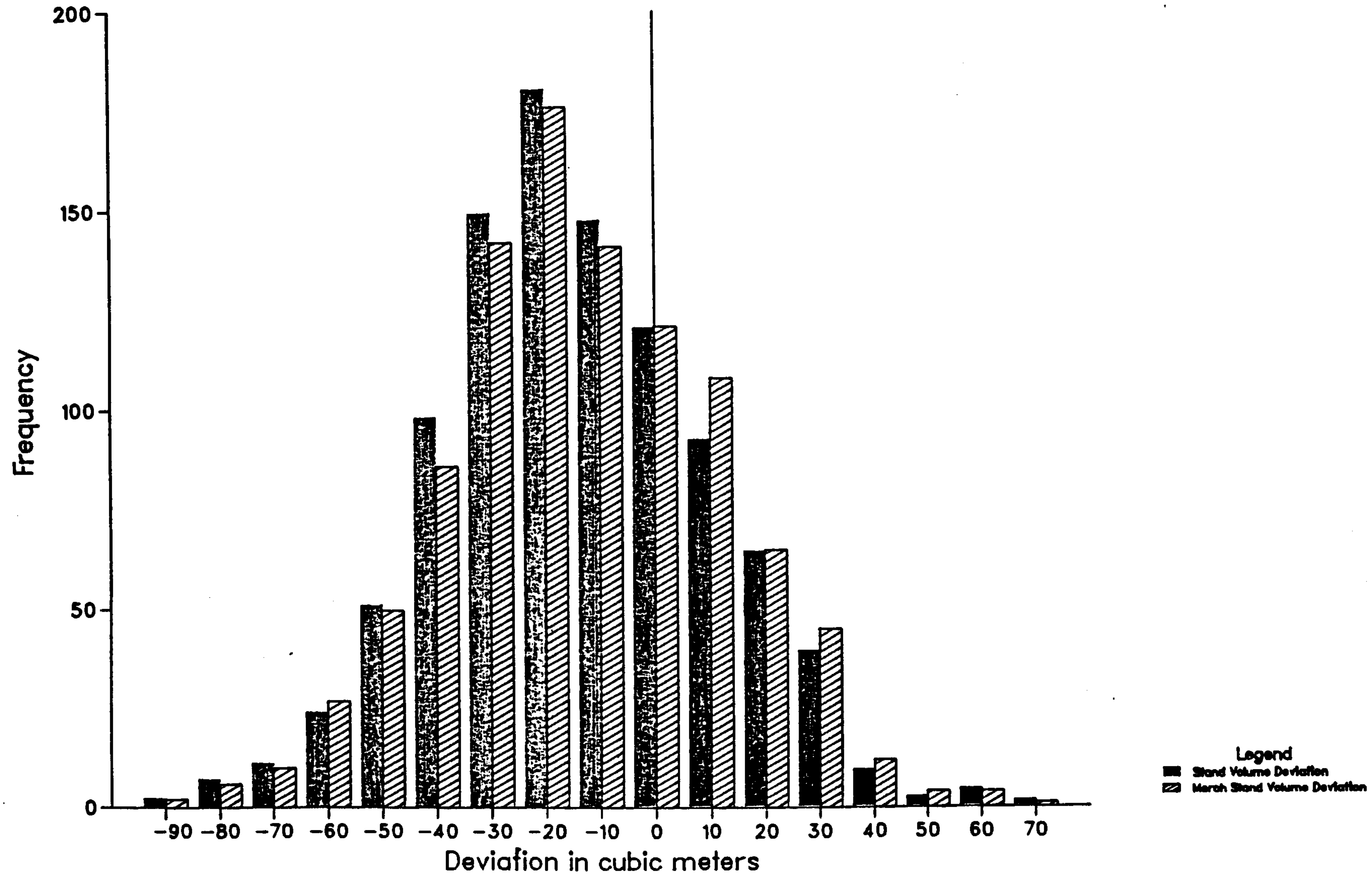


Figure V.11
Histogram of Deviations of the Estimated Total and Merchantable Stand Volumes Based on Equations Constructed from 60 White Spruce Trees

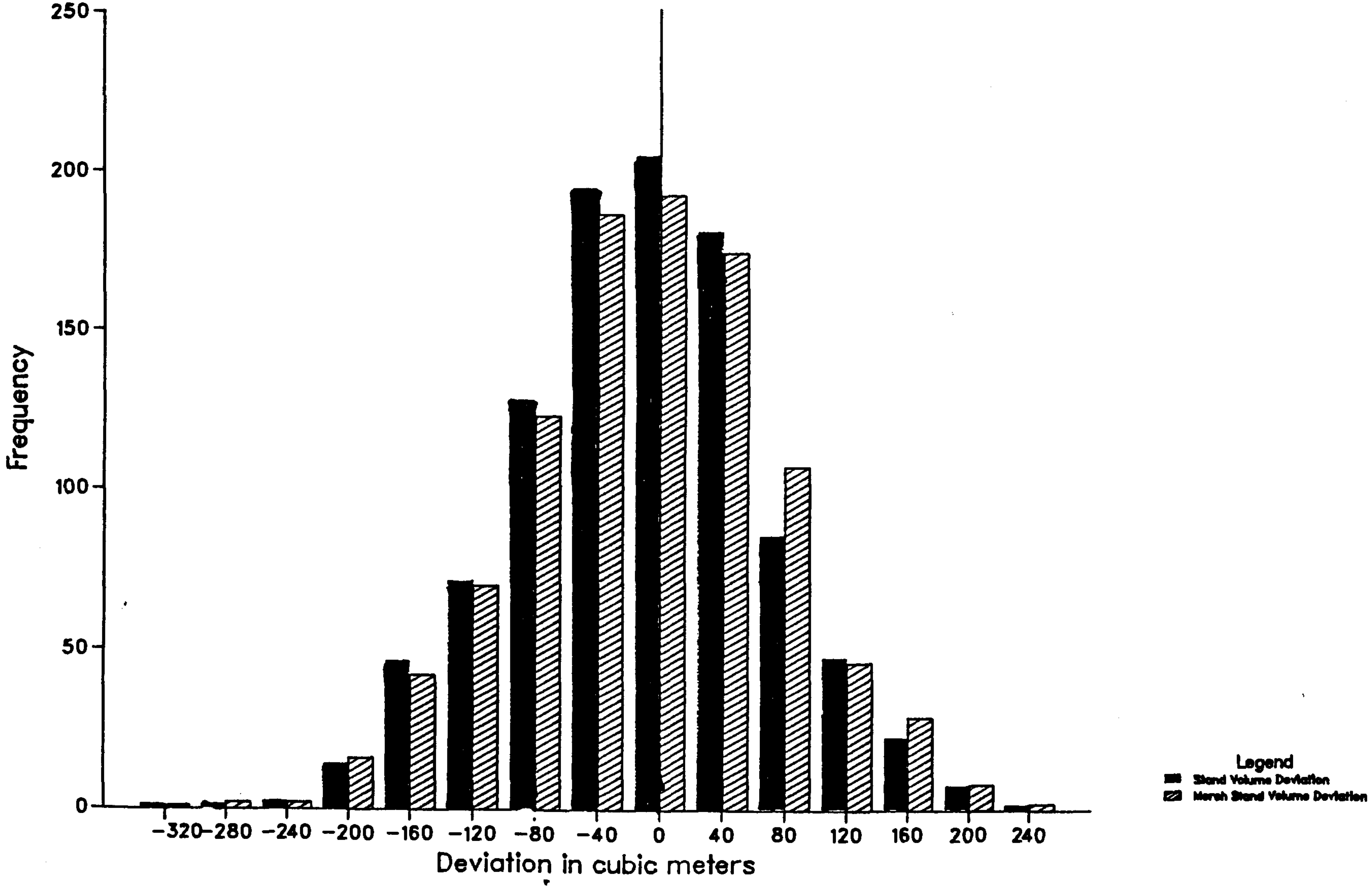
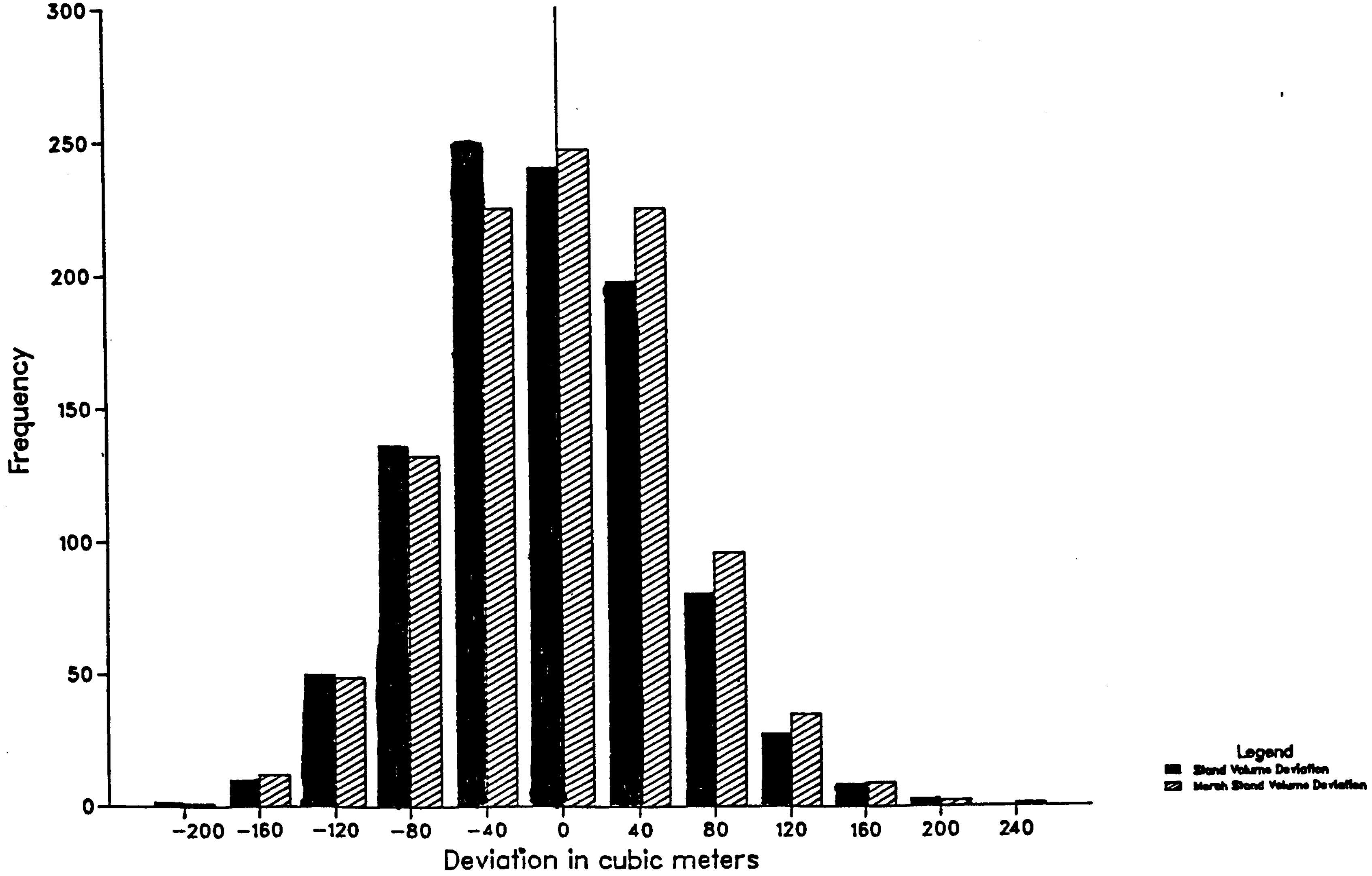


Figure V.12
Histogram of Deviations of the Estimated Total and Merchantable Stand Volumes Based on Equations Constructed from 100 White Spruce Trees



CHAPTER VI

FURTHER ANALYSES

It is also worthwhile to examine the effects of several factors in developing volume equations. A comparison of the three species with regard to their separate regression of log of total volume on log(DBH) and log(HT) was made. To do this, all the data from the three species were combined and indicator variables were used to identify them. The regression model for total volume was then fitted to the combined data. (BMDP program P1R with species as the grouping variable was used.) The analysis showed the estimated equations corresponding to the three species are significantly different from the individual regressions with an F-ratio of 297.93 with 6 and 4574 degrees of freedom (p-value is 0.0000). Since the coefficients of the fitted regression equation are dependent on the form factor, the difference in the volume equations could be attributed to the difference in form factors for the three species.

As mentioned in Chapter II, trees in the data set came from different site qualities (poor, medium and good). Within each species, a comparison was made of the regression equations corresponding to site quality. By using indicator variables to identify site index, a single regression equation was fitted to the data for each species. The results indicated that within each species, the fitted regression equations are significantly different for different site indices. For the Coastal Douglas

Fir, p-value was 0.00620, for Interior Douglas Fir the p-value was .00647, and .00001 for White Spruce.

Another comparison was made of the regression equations corresponding to immature and mature age categories for each species. Results indicated that the estimated equations were significantly different with the p-values, 0.0000, 0.0170, and 0.0000 for the Coastal, Interior Douglas Fir, and White Spruce trees respectively.

The resulting volume estimates using the proposed scheme are biased. As mentioned in Chapter III, the bias is due to the logarithmic transformation. Since $\log(V_{tot})$ is $N(\mu,\sigma^2)$, V_{tot} is lognormal with expectation $E(V_{tot}) = \exp(\mu+\sigma^2/2)$. The bias can be reduced by multiplying the predicted total volume by $\exp(\hat{\sigma}^2/2)$, where $\hat{\sigma}^2$ is the estimated variance.

An attempt was made to develop an estimation scheme that predicts the amount of cut lumber that can be extracted from standing trees (the common unit for this volume is board-foot). The method tested was to model the volume of wood wasted due to shrinkage, saw-kerf, slabs and edgings. Once this wasted volume is estimated and subtracted from the merchantable volume, the volume of cut lumber is obtained. Following the International Log Rule, shrinkage is commonly taken to be a constant fraction of the total volume. The shape of the bole is irrelevant to its calculation. The loss due to saw-kerf is a constant fraction of the overall volume. Losses due to slabs and edgings are difficult to calculate. While saw-kerf and shrinkage deductions are related to total tree volume, slabs and edgings allowances are closely related to log diameter. The International Log Rule handles the slabs and edgings deductions by viewing them as a collar having an average thickness of 1.8 cm and is assumed to be the same for all logs. The total volume of such collar is proportional to $(DBH)H_m$, where H_m is the length of the merchantable portion of the bole. H_m depends upon the fixed diameter limit d_m, HT and DBH. Hence, the volume of this collar is a function of HT and DBH.

The model equation proposed was similar to that used in predicting total volume and the unusable top volume, i.e. an equation relating the volume lost due to slabs and edgings, saw kerf and shrinkage to diameter at breast height and total height. The volume equation was transformed to its logarithmic form and was fitted to the sample trees. The residual plots for the three species revealed curvilinear trends. This indicated that the regression model was not appropriate. However, when the log volume was regressed directly on log(DBH) and log(Hm) where Hm is the *merchantable* height, the residual plots behaved as if they were random draws from an N(0,1) distribution. Also, R^2 was close to 100% and root mean square was close to 0. This shows that the model works if instead of using HT in the volume equation, the merchantable height Hm is substituted. The disadvantage of this method is that every time we have to estimate the merchantable height for the given diameter limit before the regression model can be applied. To obtain an estimate of the merchantable height corresponding to a given diameter, we need to know the taper curve which in turn requires a lot of data to develop. A simple estimation scheme similar to those developed for total and merchantable volume which only required DBH and HT may not be applicable here.

CHAPTER VII

CONCLUSIONS

Logarithmic transformation of the proposed volume equation resulted in an appropriate regression model that predicts log volume in terms of log(DBH) and log(HT). Although there was a concern about bias in the volume estimates as a result of doing the transformation, the study revealed that this bias is often minimal when the constructed equation is applied to all trees in the stand.

The regression model can be fitted to a smaller data set and a further increase in sample size only gave a slight reduction in the chance component of the error but no reduction in bias. When sampling cost is prohibitive, an estimation scheme that needs a large of amount of data may not be feasible to apply. This gives an advantage of the proposed method over the use of taper curves. In the latter, to get more precise results, an enormous amount of data is needed to develop the taper equation.

The trees providing data for each species in this study were of different ages and were collected from different areas in the province. These data do not really represent samples from a homogeneous stand of trees. In spite of this limitation, the proposed scheme for estimating stand volume remains feasible. As mentioned earlier, site quality and age are important factors in considering the volume equation. The estimate may further be improved when the volume equation is constructed and applied to a homogeneous stand of trees. Fitting the model to a homogeneous stand of trees will result in a better R^2. The equation constructed from a homogeneous stand however, will not be reliable when extrapolated to another stand.

The results of this study further support Martin's findings that simple equations with only a few parameters involved give good volume estimates. The model used in this project is simple to apply and contains few parameters.

The top volume prediction equations may be further improved by adding to the equation other terms in DBH, HT, and diameter and height measurements at the unusable top portion or by using a weighted least squares method. However, since we only need a rough estimate of the top volume, the use of a complicated model is not worth pursuing at this stage.

The presence of bias in the volume estimates should not deter one from using the proposed model. Most taper-based volume equations in literature are biased too (see Chapter I). As a consequence of the study, it is worth exploring in the future a method that would compensate for the bias induced using the estimation schemes presented in this project.

REFERENCES CONSULTED

Avery, T.E. and H.E. Burkhart. Forest Measurements. New York: McGraw-Hill, 1983.

BMDP Statistical Software. Ed. W.J. Dixon. Berkeley: University of California Press, 1981.

British Columbia Forestry Handbook. Ed. Susan B. Watts. Vancouver: University of British Columbia Press, 1983.

Bruce, D., R. Curtis and C. Vancoevering. "Development of a System of Taper and Volume Tables for Red Alder." Forest Science, 14 (1968), 339-350.

Cunia, T. "Weighted Least-Squares Method and Construction of Volume Tables." Forest Science, 10 (1964), 180-191.

Demaerschalk, J.P. and A. Kozak. "The Whole-Bole System: A Conditioned Dual-Equation System for Precise Prediction of Tree Profiles." Canadian Journal of Forestry Research, 7 (1977), 488-497.

Draper, N.R. and H. Smith. Applied Regression Analysis, 2nd ed. New York: John Wiley and Sons Inc., 1981.

Honer, T.G. "A New Cubic Foot Volume Function." Forestry Chronicle, 41 (1965), 476-493.

Hosie, R.C. Native Trees in Canada. Ontario: Fitzhenry and Whiteside Ltd., 1979.

Husch, B., C.I. Miller and T.W. Beers. Forest Mensuration. New York: Ronald Press Company, 1972.

James, C.A. and A. Kozak. "Fitting Taper Equation from Standing Trees." Forestry Chronicle, 60 (1984), 157-161.

Kleinbaum, David and Kupper, Lawrence. Applied Regression and Other Multivariate Methods. North Scituate, Massachusets: Duxbury Press, 1978.

Kozak, A., D.D. Munro and J.H.G. Smith. "Taper Functions and Their Application to Forest Inventory." Forest Science, 45 (1969), 278-283.

Martin, A. J. "Testing Volume Equation Accuracy with Water Displacement Techniques." Forest Science, 30 (1984), 41-50.

Max, T.A. and H.E. Burkhart. "Segmented Polynomial Regression Applied to Taper Equations." Forest Science, 22 (1976), 283-289.

Moser, J.W. and T.W. Beers. "Parameter Estimation in Nonlinear Volume Equations." Journal of Forestry, 67 (1969), 878-879.

Newnham, R.M. "A Modification to the Combined-Variable Formulas for Computing Tree Volumes." Journal of Forestry, 65 (1967), 719-720.

Ormerod, D.W. "A Simple Bole Model." Forestry Chronicle, 49 (1973), 136-138.

Parresol, B.R., J.E. Hotvedt and Q.V. Cao. "A Volume and Taper Prediction System for Bald Cypress." Canadian Journal of Forestry Research, 17 (1983), 250-259.

Reed, D.D. and J.C. Byrne. "A Simple, Variable Form Volume Estimation System." Forestry Chronicle, 61 (985), 87-90.

Ryan, T.A., B.L. Joiner and B.F. Ryan. Minitab Student Handbook. Duxbury Press, 1976.

Schumacher, F.X. and F.S. Hall. "Arithmetic Expression of Timber-Tree Volume." Journal of Agricultural Research, 47 (1933), 719-734.

Spurr, S.H. Forest Inventory. New York: The Ronald Press Company, 1952.

Therien, G. and C. Camire. "Use of Taper Curve Solved by an Algebraic Method in Silviculture Research." Forestry Chronicle, 62 (1986), 529-532.