Stem Profile Equations for Several Commercially Important Timber Species in Wisconsin

By

Tim Hart, B.S.

Southern Illinois University, 2006

A thesis

Submitted in partial fulfillment of the

Requirements of the degree

MASTER OF SCIENCE

IN

NATURAL RESOURCES (FORESTRY)

College of Natural Resources

UNIVERSITY OF WISCONSIN

Stevens Point, WI

May 2009

APPROVED BY THE GRADUATE COMMITTEE OF:

Committee Chairman
Dr. Paul F. Doruska-Associate Professor-Forestry

Dr. Michael Demchik-Associate Professor-Forestry

Dr. Nathan Wetzel-Associate Professor-Mathematical Studies

Abstract

Current estimates of standing tree merchantable volume within the Lake States frequently rely upon composite volume tables and/or approximation formulas. The composite volume tables used, and the approximation formulas that are explicitly derived from them, were created over 50 years ago. Within that time, standards of merchantability have changed, along with stand characteristics based on forest management activities. Therefore, the need exists to develop more accurate estimations of volume to benefit the timber industry. In this project, equations were developed to predict outside bark stump diameter (at 0.5 ft.) and outside bark diameters along the entire tree bole for several commercially important hardwood species in Wisconsin. This was accomplished by taking paired height/diameter measures on standing trees using Laser Technology's Criterion RD 1000 electronic dendrometer. Profile equations were developed for ash (*Fraxinus* spp.), American basswood (*Tilia americana* L.), sugar maple (*Acer saccharum* Marsh.), and bigtooth aspen (*Populus grandidentata* Michx.). The equation for sugar maple had an approximate R^2 of 0.883 and a mean absolute error (MAE) of 0.8522 in. The equation for ash had an approximate R^2 of 0.934 and a MAE of 0.9154 in. For aspen, the equation had an adjusted R^2 of 0.921 and a MAE of 0.7878. The equation for basswood had an approximate R^2 of 0.942 and a MAE of 0.7850. Volumes were predicted from the selected equations for each sample tree. Smalian's formula was applied to consecutive section of each tree as measured with the Criterion RD 1000 and summed. Predicted volumes obtained from this study were compared to volumes from commonly used composite tables. Not surprisingly, predicted volumes from the two methods differed greatly when applied to the project dataset.

Acknowledgments

The author would like to thank the following individuals and groups for their help in completion of this thesis:

My committee members, Dr. Michael Demchik and Dr. Nathan Wetzel for their input throughout the project.

Dr. Paul Doruska, my committee chair, for spending a considerable amount of time helping me understand the project as a whole, teaching aspects of the statistical analysis involved, and being there for all my random questions. Without his help, this project would have never reached completion.

The forestry community within Wisconsin for finding study locations for use, providing meetings, maps, and their own time to help me out.

My late father, Gene Hart, for teaching me to excel in everything I do, which eventually led me to pursue a graduate degree.

My mother, Pamela Hart, and sister, Megan Hart, for being there throughout the duration of my academic career, making sure I never quit and got everything done on time.

And lastly, all my friends throughout the various places I have been, for giving me a break every once in a while from writing my thesis or analyzing data. Without random phone calls, impromptu games of softball, or a quick beer, I would have gone mad staring at my computer screen.

This material is based upon work supported by the Cooperative State Research, Education, and Extension Service, U.S. Department of Agriculture, under McIntire-Stennis Award No. WIS01235.

Table of Contents

Abstract iii

Acknowledgments iv

Table of Contents v

List of Tables vii

List of Figures viii

Introduction ..... - 1 -

Literature Review ..... - 4 -

General Background ..... - 4 -

Taper ..... - 4 -

Stem Form ..... - 5 -

Developing Taper Equations ..... - 7 -

Early Taper Equations ..... - 7 -

Segmented Polynomial Taper Equations ..... - 10 -

Variable Exponent Form Taper Equations ..... - 14 -

Compatible Volume Equations ..... - 18 -

Nonlinear Seemingly Unrelated Regression ..... - 19 -

Nonlinear Mixed Effect Modeling ..... - 20 -

Summary ..... - 22 -

Objectives ..... - 23 -

Methods ..... - 24 -

Data Collection ..... - 24 -

Taper Equation Development ..... - 27 -

Analysis ..... - 27 -

Outside Bark Stump Equation ..... - 28 -

Outside Bark Profile Equation ..... - 29 -

Regression Diagnostics ..... - 30 -

Volume Determination ..... - 36 -

Results ..... - 39 -

Data Summarization ..... - 39 -

Outside Bark Stump Equation ..... - 39 -
Outside Bark Profile Equation ..... - 45 -
Volume Determination ..... - 57 -
Final Model Selection ..... - 57 -
Discussion ..... - 71 -
Volume Comparisons ..... - 71 -
Fitting other Models ..... - 73 -
Data Issues ..... - 74 -
Project Strengths and Weaknesses ..... - 76 -
Summary and Conclusion ..... - 78 -
Literature Cited ..... - 82 -

List of Tables

Table		Page
1	Mean and standard deviation (in parentheses) of tree attributes by species	40
2	Distribution of sugar maple in the dataset by DBH and total height class	41
3	Distribution of ash in the dataset by DBH and total height class	42
4	Distribution of aspen in the dataset by DBH and total height class	43
5	Distribution of basswood in the dataset by DBH and total height class	44
6	Fit statistics for stump equation models (20, 21) by species and equation	46
7	Regression coefficients for equation (20) by species	47
8	Fit statistics for base model equation (4) by species, equation, and covariance structure	51
9	Fit statistics for base model equation (14) by species, equation, and covariance structure	52
10	Fit statistics for base model equation (16) by species, equation, and covariance structure	53
11	Regression coefficients for equation (4) by species using first-order autoregressive covariance structure	54
12	Regression coefficients for base equation model (14) by species using first-order autoregressive covariance structure	56
13	Regression coefficients for base equation model (16) by species using first-order autoregressive covariance structure	58
14	Fit statistics for volume determination	59

List of Figures

Figure		Page
1	County map of Wisconsin illustrating study site locations	25
2	Observed versus estimated values of outside bark diameters from equation (34) for a sugar maple tree on site 1 with a DBH of 11.45 in. and a total height of 66 ft.	61
3	Observed versus estimated values of outside bark diameters from equation (34) for a sugar maple tree on site 4 with a DBH of 12.2 in. and a total height of 76 ft.	62
4	Observed versus estimated values of outside bark diameters from equations (35) and (20) for an ash tree on site 3 with a DBH of 9.8 in. and a total height of 81 ft.	63
5	Observed versus estimated values of outside bark diameters from equations (35) and (20) for an ash tree on site 3 with a DBH of 12 in. and a total height of 85 ft.	64
6	Observed versus estimated values of outside bark diameters from equations (4) and (20) for an aspen tree on site 2 with a DBH of 9.6 in. and a total height of 85 ft.	66
7	Observed versus estimated values of outside bark diameters from equations (4) and (20) for an aspen tree on site 1 with a DBH of 10.75 in. and a total height of 70 ft.	67
8	Observed versus estimated values of outside bark diameters from equations (39) and (20) for a basswood tree on site 6 with a DBH of 13 in. and a total height of 75 ft.	68
9	Observed versus estimated values of outside bark diameters from equations (39) and (20) for a basswood tree on site 3 with a DBH of 12.95 in. and a total height of 76 ft.	69

Introduction

One use of the timber inventory process is to predict the amount and value of timber within a forested area. Accurate and precise estimation of standing tree merchantable volume is an important and vital calculation derived from the timber inventory process. Current estimations of standing tree merchantable volume within the Lake States frequently rely on composite volume tables and/or approximation formulas.

The composite volume tables and corresponding taper tables used in the Lake States were developed by Gevorkiantz and Olsen (1955). The volumes calculated from some of these tables apply to all species. Ideally, a large sample set of data, taken across species and region, should have compensating errors that balance the inherent variation among species and within the region. Due to forest parcelization and different management goals in small stands, the compensating factors will likely decrease, rendering the composite tables less accurate. Also, the tables were developed during the 1950's when standards of merchantability were much different than those of the present. Fifty years of forest management may have changed stem form from that which was available when the tables were created (Ek and others 1986). For their provided standards, the Lake States composite tables give a relatively accurate account of volume within a forest.

Approximation formulas for stand volume, similar to the composite tables, are species independent. Ek and Burk (1986) claim that since derivation of the formulas are based upon the Lake State's composite volume tables, compensating error factors into its application and can potentially lead to a decrease in accuracy of volume estimates for small stands. Currently, approximation formulas used in the Lake States only apply to point sampling inventories with a 10-factor prism (Ek and Burk 1986). With the increasing need of forest inventories upon smaller tracts of timber, fixed radius plots may be a more appropriate technique to use. The approximation formulas do not apply to fixed radius plot inventories.

Over the last 30 years, other regions of the country have begun using taper equations (also known as profile equations), instead of relying upon composite tables and approximation formulas (Fang and others 2000). Tree form is known to vary among species and geographical regions (Avery and Burkhart 2002). As a result, taper equations can be more accurate than other volume estimations because they are species specific, region specific, and volume can easily be derived from them. Volume to any merchantable limit can be estimated because stem diameter is traced along the entire bole. Wright's (1923) goal of developing general taper tables for species per region to build local volume tables to meet the industry's needs is closer than ever to being realized.

Taper equations have been derived from the Lake States composite tables, but they are not species specific (Burk and Ek 1999). Publicly available species specific Lake States taper equations are generally lacking, with only some published data on red pine (*Pinus resinosa* Ait.) compiled by Byrne and Reed (1986). Furthermore, there are no publicly available region specific taper equations for use on timber species in Wisconsin within the National Volume Estimator Library (NVEL) (National Volume Estimator Library 2008).

The Great Lakes Stem Profile Modeling Project, established by the Ontario Ministry of Natural Resources and the Ontario Forest Research Institute, is beginning to create taper equations for the Lake States. Most work is being conducted in Michigan and Ontario; however, due to the region specific nature of taper equations, equations for trees in Wisconsin should be developed.

Taper equations will be developed for four important timber species and/or species groups of Wisconsin. Ash (*Fraxinus* spp.) will be examined in preparation of the impact of the emerald ash borer (*Agrilus planipennis* Fairmaire). American basswood (*Tilia americana* L.), sugar maple (*Acer saccharum* Marsh.) and bigtooth aspen (*Populus grandidentata* Michx.), have different stem forms and are commercially important. Note that American basswood is hereafter referred to as basswood and bigtooth aspen is hereafter referred to as aspen, if not specifically identified as such. All species selected for use in this study encompass a large area of land within Wisconsin. The sugar maple/beech/yellow birch forest type covers about 2.3 million acres of forested land, while the hard maple/basswood and aspen forest types cover 1.4 million acres and 2.8 million acres, respectively (Miles 2007). Four sets of developed equations should serve as an adequate starting point for taper equation development in Wisconsin until further studies of a similar nature can occur.

Literature Review

General Background

The composite volume tables presented in USDA Technical Bulletin 1104 have been widely used since their development by Gevorkiantz and Olsen (1955). The application of the composite tables is based upon standards of utilization common during the 1950's (Ek and others 1986). The tables were derived from a large sampling of trees throughout the Lake States and based upon the idea that form and taper were closely related to stand conditions. Taper and form variation among Lake State species was a less important idea, because compensating error, which accounts for individual variation, was factored in the tables' development (Burk and Ek 1999). Utilization of trees to lower stump heights and smaller upper stem diameters has increased the need for more accurate volume estimations (Ek and others 1986).

Approximation formulas are derived explicitly from composite volume tables. Compensating error limitations imposed upon the Gevorkiantz and Olsen (1955) tables adhere firmly to any estimated value originated from an approximation formula. Though the computed values are highly correlated, the need for accurate assessments of timber quantity is the driving force behind present timber inventories (Ek and Burk 1986).

Taper

Tree taper refers to the rate of decrease in diameter with increasing height up the stem (Newnham 1992). Taper can be affected by age, diameter, height, and locality (Wright 1923). Taper equations are developed from paired height/diameter measurements taken at various points along the entire tree bole (Westfall 2004). There are two important reasons as to why study in this area remains a high priority. First, there is no single theory that adequately explains variation in stem form for all trees, hence, no universally accepted taper equation. Second, and a more critical factor, is the ability of taper equations to estimate both total and merchantable tree stem volumes from easily measured variables (Newnham 1988). Taper equations are useful to predict volume to any merchantability limit and stem diameter at any height (Byrne and Reed 1986). Single trees are merchandized into multiple products, thus, knowing volumes of and diameters at specific heights is important (Jordan and others 2005). With a greater proportion of each tree being harvested, as well as new manufacturing facilities being built, accurate estimations of volume are paramount to new merchantability standards within the industry (Newnham 1988).

Stem Form

There is a long history of the scientific community attempting to model tree form so that the entire stem profile, accounting for individual tree variability, could be described more accurately (Behre 1923). Grosenbaugh (1966) stated that a tree stem may assume an infinite number of shapes, making the development of a simple, accurate equation difficult (Newnham 1988). Though the bole of a tree does not approximate a singular geometric solid of revolution, it can be segmented into several, making possible the accurate calculation of volume content. The lower bole is neiloid in shape, the middle portion is often in the shape of a paraboloid, and the upper portion is conical (Husch and others 1972).

Stem form is directly related to the same conditions that determine the size and distribution of the crown along the tree bole. According to Larson (1963) stand density (Härdtl 1938, Stoate 1942), crown class (Vorreiter 1954), site quality (Schmeid 1918, Burger 1931), inheritance (Metzger 1896, 1908), and thinning/pruning (Flury 1903, Bickerstaff 1946) affect stem form along the entire tree.

The lower bole section is the most variable segment among individual trees, due to many complex factors. Though difficult to determine, butt swell should not be an eliminated factor when quantifying stem form (Behre 1923). Larson (1963) also states that crown development (Gevorkiantz and Hosley 1929), inhibited root growth (Hartig 1892), longer persistence in cambial activity (Mer 1892, Knight 1961), structural stability (Metzger 1893, Laitakari 1929), and shallowness of soil (Davis and Richards 1934) may all be contributing factors to the inconsistency of butt swell between trees. Treated as an anomaly in the past, butt swell is now being utilized for wood products (Newnham 1988).

Taper is commonly used when describing stem form. According to Wright (1923) the taper of a forest tree was described by Jonson (1910, 1911) to be the absolute form quotient, which is the ratio of diameter at breast height (DBH 4.5 ft. above ground) to diameter at half the height between breast height and the top of a tree. A more commonly used form quotient in the United States today is the Girard Form Class, which is the ratio of stem diameter inside bark at the top of the first full log (17.3 ft.) to DBH outside bark (Avery and Burkhart 2002). The measurement derived is a percentage (when the ratio is multiplied by 100) and can more easily determine accurate accounts of volume during timber inventories by accounting for taper. Volume tables by form class can be regularly found in published booklets (Mesavage and Girard 1956).

Developing Taper Equations

Early Taper Equations

According to Kozak (1988) early taper equations derived by Höjer (1903), Jonson (1910, 1911), and Behre (1923) concentrated on the merchantable portion of the stem and were relatively simple mathematical formulas. The approach used to describe diameter changes along the tree bole was a single function of different forms. Höjer's 1903 equation, developed from measurements of Norway Spruce (*Picea abies* [L.] Karst.), to express the diameter of any tree at any point along its stem took the following form:

$$\frac{d}{D} = C\ log\frac{c+1}{c} \qquad (1)$$

where: *d*=diameter at distance "*l*" from the tip,

l=distance from top of tree,

D=diameter at breast height (4.5 feet), and

C,*c*=constants.

While working with this formula, Jonson (1910, 1911) used it as a basis for developing taper tables and tables of cubic volume for trees of all sizes because of its close conformity to the actual taper of forest trees (Behre 1923). However, Behre explains that Jonson found that with Scotch Pine (*Pinus sylvestris* L.), a "biological constant" would be needed in Höjer's equation to better fit the tops of trees which led him to a new formula, the absolute form quotient formula:

$$q = \frac{d}{D} \qquad (2)$$

where: *q*=absolute form quotient,

d=the diameter at the middle of the stem above breast height, and all other terms as previously defined.

The absolute form quotient gives a value for the taper class of the tree, most often referred to as a percentage, which will give the diameter at any point along a tree bole expressed as a ratio of the breast height diameter (Behre 1923). The advantage of this expression of the variation in stem form is that classification of form class is made independent of height, being that the two form determining diameters are always in relation to one another (Claughton-Wallin and McVicker 1920). This led Jonson to believe that all conifers and hardwoods taper with the same law and adhere to the same set of general taper curves, which is in accordance with the mechanical theory of tree from (Behre 1924).

Still without good fits in the upper stems of trees, Behre (1923) proceeded to present a new equation for stem curve which would be more consistent in its conformity to nature. The equation has the form of the ordinary hyperbola when expressed in terms similar to Höjer's equation:

$$\frac{d}{D} = \frac{l}{a+bl} \tag{3}$$

where: *a*,*b*=constants,

a+*b*=1,

and all other terms as previously defined.

Bruce (1972) attempted to transform Behre's hyperbolic equation, but still found poor fits at different points along the tree stem. Hilt (1980) claims that this taper system does have the advantage of being simpler to apply by practicing foresters than newer, more complex taper systems (Wiant and Charlton 1984). Höjer's, Jonson's, and Behre's formulas were all found to be relatively accurate, except for within the region of butt swell (Husch and others 1972).

Before Höjer, Jonson, and Behre, several Europeans attempted to express tree form by a single equation of a known solid of revolution, but were unsuccessful due to the tree conforming to several different solids at different points along the bole (Behre 1923). Gray (1956) demonstrated that the main section of the tree bole could be a quadratic paraboloid, but that alone gave poor fits at the base (Newnham 1992).

Other early notions of taper favored the thought that any two trees, regardless of species, having the same merchantable length, DBH, and diameter inside bark (DIB) at merchantable top, will have the same diameter at any paired points throughout their length (Lemieux 1936). Wright (1923) was one of the first to report a need to develop separate taper tables for different species and for different size classes within the same species.

An early and widely used polynomial taper model was developed by Kozak, Munro, and Smith in British Columbia for several tree species (Kozak and others 1969). The model is a simple parabolic function of the form:

$$d_i^2/D_i^2 = \beta_0 + \beta_1(h_i/H_i) + \beta_2(h_i^2/H_i^2) + \varepsilon_i \tag{4}$$

where: H_i=total tree height, and

b_0,b_1,b_2=regression coefficients to be estimated.

This model form has been found to fit well for over 85% of a tree bole, but has performed poorly within the top and butt swell sections of a tree (Avery and Burkhart 2002). Kozak's model can easily be integrated over the length of the tree bole to produce total or merchantable volume estimates.

Two variable form approaches are currently being used today to describe tree taper. The first approach is a taper equation with a step function so that the bole is separated into segments by inflection points, while the second is a single continuous function for the entire bole (Fang and others 2000). Also, fully compatible volume equations are being used to develop taper equations through iteration. Recently, the use of nonlinear seemingly unrelated regression and mixed effects modeling has also aided the development of taper equations by increasing predictive precision.

Segmented Polynomial Taper Equations

Tree form is highly variable between species, as well as individuals. Complex models may be needed in order to correctly describe tree form (Max and Burkhart 1976). Early polynomial systems of taper tried to develop equations based upon a single polynomial of various powers for red alder (*Alnus rubra* Bong.) in the Pacific Northwest and for several species groups in British Columbia (Bruce and others 1968, Kozak and others 1969).

To establish a more complete model for tree form, a tree can be partitioned into separate models for each geometric solid it includes. The geometric solids should then be fused together to finalize the model (Newnham 1992). The segmented model must be continuous at inflection points to work properly (Fang and others 2000). As one single segmented polynomial model, the whole may easily be analyzed by regression techniques (Max and Burkhart 1976).

When using loblolly pine (*Pinus taeda* L.) and slash pine (*Pinus elliottii* Englem.) from the coastal plains of eastern U.S., the segmented polynomial taper model appears to describe the stem profile quite accurately (Fang and others 2000). Compared with simpler taper-volume estimation systems, the segmented system has greater ability to predict volumes to a top diameter. Also, the segmented system is more compatible, meaning that the taper equation developed can easily define an associated volume equation. Volume calculated by integration of the model is usually equal to that of computed volume equation solutions. According to Byrne and Reed (1986), when developing equations for red pine from the upper Midwest and loblolly pine from the coastal plains of eastern U.S., systems based on a segmented taper equation outperformed all other simpler systems, especially when estimating volume to a top diameter.

Early models suggested the use of two mathematical functions; one describing the upper bole and the other describing the lower bole. The whole bole system of dual equations is set under the restrictions that diameter is equal to zero at the top of the bole, diameter equals that of diameter of inside bark at breast height, and that the two equations uniformly join at the inflection point. This system, when evaluating 32 species groups in British Columbia, has a large bias when accounting for butt swell, but is nearly perfect when predicting diameters inside bark along the tree bole (Demaershalk and Kozak 1977).

Max and Burkhart (1976) found that within a segmented polynomial taper equation for loblolly pine in the coastal plains of eastern U.S., a complex quadratic-linear (first degree polynomial)-quadratic model (each equation corresponding to a different tree bole segment) was sufficient to describe stem taper in plantations, while a quadratic-quadratic-quadratic model was good for use in naturally occurring stands. The difference between complex models being used for plantation and natural stands is due to the relatively shorter height of plantation trees (Max and Burkhart 1976). In a study of

Appalachian hardwoods, comparing several taper models, the Max and Burkhart model was suggested as a possible model to be used because of its consistency and goodness of fit within the lower boles, though it is not a simple model to use (Martin 1981). In 1985, Burkhart and Walton tried to increase the precision of the model by including crown ratio (length of live crown divided by total tree length) as a tree variable to the already used variables of DBH and height. The lower join points were found to be around 10-15%, while the upper join points were found at 75-85%. For broad ranges of crown ratio, it was found that inclusion of an extra variable was unwarranted, while for extremes in stand density, the variable's inclusion could be justified (Burkhart and Walton 1985).

Segmented polynomial models can be written as a series of grafted submodels. Gallant and Fuller (1973) provide a simple model written as:

$$y_i = f(x_i) + e_i \tag{5}$$

where: y_i=independent variable to be estimated,

x_i=dependent variable,

$$\begin{aligned} f(x) &= f_1(x,\beta_1), \quad a \le x \le \alpha_1 \\ &= f_2(x,\beta_2), \quad \alpha_1 < x \le \alpha_2 \\ &\;\;. \\ &\;\;. \\ &\;\;. \\ &= f_r(x,\beta_r), \quad \alpha_{r-1} < x \le b. \end{aligned}$$

α_i=join points, and

β_i=parameter to be estimated.

The submodels are grafted together at the join points by imposing restrictions on the model. Restrictions are imposed so that *f* is continuous and has continuous first or higher order derivatives (Max and Burkhart 1976).

Using the Gallant and Fuller method (1973), Max and Burkhart (1976) developed the following model (6) to account for three segmented sections of a tree conforming to different geometric solids of revolution. Each submodel herein is quadratic:

$$y_i = \beta_1(x_1 - 1) + \beta_2(x_i^2 - 1) + \beta_3(\alpha_1 - x_1)^2 I_1 + \beta_4(\alpha_2 - x_1)^2 I_2 + \varepsilon_i \qquad (6)$$

where: $y=d^2/D^2$,

d=diameter inside bark at any given height h,

h=height above the ground,

D=diameter at breast height outside bark,

H=total tree height from ground to tip,

$x=h/H$,

$I_1=1, \quad \alpha_1-x_i \geq 0$

$=0, \quad \alpha_1-x_i<0$

$I_2=1, \quad \alpha_2-x_i \geq 0$

$=0, \quad \alpha_1-x_i<0$, and

$\beta_1, \beta_2, \beta_3, \beta_4, \alpha_1, \alpha_2$=parameters to be estimated.

For a different type of segmented system, Liu (1980) used a complex model of cubic spline functions, applied with Reinsch's (1976) algorithm (for the smoothing of isolated errors in measured data), to describe stem taper of yellow poplar (*Liriodendron tulipifera* L.) in eastern Kentucky. A set of cubic polynomial segments with smooth joints, given by a set of coordinates corresponding to relative radii and positional heights, approximated intervals along the tree bole. When evaluating 51 points along the tree bole, the cubic spline function estimated stem taper quite accurately. A weakness of this system is the complicated mathematical procedures to determine taper.

Variable Exponent Form Taper Equations

As a tree's height position along the bole increases, stem form changes from base to leader. The geometric solids of revolution assumed to be compatible with stem form do not abruptly change at the join points, but gradually assume the shape of a curve. Exponents used within other taper equation systems are subject to the constrained values of conventional geometric solids (1 for a cone, 2 for a quadratic paraboloid, 3 for a cubic paraboloid, and 2/3 for a neiloid) (Newnham 1988). The variable exponent form taper equation employs one continuous function describing the entire tree bole from ground to tip with a changing exponent (Kozak 1988).

The use of variable exponent form taper systems has both advantages, and disadvantages. An advantage is that only one equation is necessary to determine the taper curve. Regression is easily applied, since fewer parameters are used (Newnham 1988). Kozak (1988) found that when using a variable exponent in a continuous function, the shape of the stem, from ground to tip, is less biased than other systems when using several different species groups in British Columbia. According to Martin (1981), in a study of Appalachian hardwoods, the Kozak 1988 variable exponent model was not as precise as the Max and Burkhart 1976 segmented model, but was just as accurate and simpler to use. Muhairwe and others (1994) attempted to improve the precision of Kozak's 1988 model by adding tree, stand, and site characteristics such as crown class, site class, breast height age, quadratic mean diameter, and crown ratio into the exponent of the equation for Douglas fir (*Pseudotsuga menziesii* [Mirb.] Franco), western redcedar (*Thuja plicata* Donn ex D. Don), quaking aspen (*Populus tremuloides* Michx.), and lodgepole pine (*Pinus contorta* Douglas ex Louden). Except for crown ratio and quadratic mean diameter for lodgepole pine, all other variables gave only marginal improvements indicating an unjustifiable cost in measuring for them. The marginal increase in precision may be due to the fact that many of these variables are highly correlated with *D/H* measures, which are already accounted for in the exponent (Muhairwe and others 1994).

Kozak's (1988) variable exponent taper equation model is derived through several equations. The function providing the general shape describing the change of diameter from ground to top is as follows:

$$Y = X^c \tag{7}$$

assuming that: $Y=d_i/DI$,

$$X = (1 - \sqrt{h_i/H})/(1 - \sqrt{p})$$

where: d_i=diameter inside bark at h_i,

h_i=height from ground, $0 \leq h_i \leq H$,

H=total height of the tree,

$p=(HI/\mathrm{H})*100$,

HI=height of the inflection point from ground, and

DI=diameter inside bark at the inflection point.

Demaerschalk and Kozak (1977) indicated that the inflection point, where the relationship between d_i/D and h_i/H changes from neiloid to paraboloid, could be expressed as a percentage (*p*) of the total height of the tree. They found the inflection point to range between 20 to 25% of total height from ground for all commercial tree species of British Columbia. Perez and others (1990) found that when using Kozak's 1988 variable exponent model on oocarpa pine (*Pinus oocarpa* Shiede ex Schltdl.) in central Honduras, different locations of the inflection point did not change the predictive ability of the model.

By expressing the exponent, C, as a function of h_i/H, as in:

$$C = 1.0/(h_i/H + k) \tag{8}$$

where: k=parameter to be estimated, and

all other terms as previously defined,

and by varying k, a family of curves describing the shape of tree boles can be found. Obtaining a good fit between d_i/D and h_i/H would, however, needs a more complicated form of the exponent. Expressing the exponent as a multiple curvilinear regression yields:

$$C = b_0 + b_1 Z + b_2 Z^2 + b_3/Z + b_4 \ln(Z + 0.001) + b_5 \sqrt{Z} + b_6 e^Z + b_7(D/H) \tag{9}$$

where: $Z=h_i/H$,

D=diameter outside bark at breast height,

$b_0,b_1,b_2,b_3,b_4,b_5,b_6,b_7$=parameters to be estimated, and

all other terms as previously defined.

An analysis of equation (9) for 33 species groups in British Columbia found the best subset of variables being:

$$C = b_2 Z^2 + b_4 \ln(Z + 0.001) + b_5 \sqrt{Z} + b_6 e^Z + b_7(D/H) \tag{10}$$

By substituting equation (10) into equation (7) and renumbering the coefficients, we get:

$$d_i/DI = X^{b_1 Z^2 + b_2 \ln(Z+0.001) + b_3\sqrt{Z} + b_4 e^Z + b_5(D/H)} \tag{11}$$

From equation (11), diameter inside bark (DI) is not known but can be estimated from diameter outside bark at breast height (D). From the data in Kozak’s 1988 study, it was found that:

$$DI = a_0 D^{a_1} a_2^D \tag{12}$$

where: a_0, a_1, a_2=parameters to be estimated, and

all other terms as previously defined,

was the best at calculating *DI*. By substituting equation (12) into equation (11) and rearranging the terms, we get:

$$d_i = a_0 D^{a_1} a_2^D X^{b_1 Z^2 + b_2 \ln(Z+0.001) + b_3\sqrt{Z} + b_4 e^Z + b_5 (D/H)} \tag{13}$$

where: all terms are as previously defined.

Using logarithmic transformations, equation (13) can be linearized:

$$\begin{aligned} ln(\, d_i) = ln(a_0) + a_1\, ln(D) + ln(a_2)\, D + b_1\, ln(X)\, Z^2 \\ + b_2\, ln(X)\, ln(Z + 0.001) + b_3\, ln(X)\sqrt{Z} \\ + b_4\, ln(X)\, e^Z + b_5\, ln(X)\,(D/H) + \varepsilon_i \end{aligned} \tag{14}$$

where: $X = (1 - \sqrt{h_i/H})/(1 - \sqrt{p})$,

all other terms as previously defined.

The properties of this model are that d_i=0 when h_i/H=1.0 and d_i=*DI* when h_i/H=*p*.

There are two theoretical problems associated with Kozak's 1988 variable exponent model. Firstly, the model contains several polynomial terms and transformations of the same regressor variable, leading to several independent variables potentially having high multicollinearity. Secondly, each tree measured has multiple observations, violating the independent error term assumption through autocorrelation (Kozak 1997). Kozak addresses these problems with a new model of the following form:

$$d_i = a_0 D^{a_1} H^{a_2} X_i^{b_1 X_i^{\frac{1}{10}} + b_2 Z_i^4 + b_3 \arcsin\,(Q_i) + b_4 [1/e^{D/H}] + b_5 DBH^{X_i}} \tag{15}$$

where: $Q_i = (1 - \sqrt{h_i/H})$,

p=1.3/H, and

all other terms as previously defined.

Using logarithmic transformations, equation (15) can be linearized:

$$\ln(d_i) = \ln(a_0) + a_1 \ln(D) + a_2 \ln(H) + b_1 \ln(X)\, X^{1/10} + b_2 \ln(X)\, Z^4 +$$

$$b_3 \ln(X) \arcsin(Q) + b_4 \ln(X) \left(1/e^{D/H}\right) + b_5 \ln(X)\, DBH^X + \varepsilon_i \qquad (16)$$

where: all terms are as previously defined.

When studying the effects of multicollinearity and autocorrelation, Kozak's new model was shown to predict values with a higher rate of precision than that of his previously published variable exponent equation (Kozak 1997).

The variable exponent form taper equation is still, however, sensitive to errors when estimating diameter near ground level. Since terms are chosen empirically, rather than geometrically, it is not possible to transpose the equation to determine height for a given diameter (Newnham 1988). The most serious problem concerning the variable exponent taper equation is that merchantable volume for a given height/diameter pair can only be obtained by iteration (Kozak 1988). A separate volume equation (Smalian, Huber, Newton, etc.) needs to be used between partitioned height/diameter pair sections, and when summed together, estimation of the volume to a merchantable top can be expressed.

Compatible Volume Equations

Variable top merchantable volume equations can be used to estimate volume to a variety of merchantable tops. A merchantable volume equation implicitly defines an associated taper function (Clutter 1980). The taper equation derivation is based on the fact that volume estimates acquired by integration of taper equations should be equal to that of an existing volume equation. The total volume, which is estimated from the summation of the sections whose volume is obtained using a taper equation, is identical to the volume yielded by a volume equation (Demaerschalk 1973). Volume equations and their associated taper equations are then fully compatible, giving identical outcomes for total volume calculations (Demaerschalk 1972). For practical application of merchantable volume equations, diameter at different heights along the bole or heights at given diameters should be known (Clutter 1980). A problem with this system is the fact that an equation that is best for taper, may not be inherently best for volume (Demaerschalk 1973). One must decide where precision in estimation needs to lie. A compatible system of volume and taper is important to the timber industry where existing volume equations will continue to be used (Demaerschalk 1972).

Nonlinear Seemingly Unrelated Regression

Nonlinear equations, when combined into a set, having a contemporaneous cross-equation error correlation, are known as a nonlinear seemingly unrelated regression system. Though taper, total volume, and merchantable volume equations appear to be unrelated, the equations are related through the correlation in the errors. Each equation within the set can use its own independent variables, have differently weighted functions, and share parameters, ensuring a sense of conformity. When error and coefficient restrictions are specified, parameter estimations and predictions of the equations become dependable (Jordan and others 2005). The parameter estimations are asymptotically more efficient than the usually used least squares method (Zellner 1962).

Jordan and others (2005) used the following structural equations (17) for a system of nonlinear models. When used concurrently, the application minimizes the error associated with diameter estimation.

$$y_1=f_1(X_1,\beta_1)+\varepsilon_1$$

$$y_2=f_2(X_2,\beta_2)+\varepsilon_2$$

$$\vdots$$

$$y_k=f_k(X_k,\beta_k)+\varepsilon_k \tag{17}$$

where: yi=is a vector containing the dependent variable from the *i*th equation,

X_i=is a matrix containing the independent variables from the *i*th equation,

$B_{i=}$is the parameter vector for the *i*th equation, and

ε_i=is the random error vector for the *i*th equation.

Nonlinear Mixed Effect Modeling

Recent studies included mixed effect modeling when developing taper equations. Mixed effect modeling allows for taper equations to include both fixed and random parameters. Using repeated and random features, correlations between dependent variables and random effects among subjects can be approached concurrently (Zhang 2002). When applied, mixed effect modeling increases flexibility and efficiency in estimation for regional conditions. New information about individuals can be incorporated into the response variable, allowing for an adjusted answer. The time and cost necessary to employ widespread sampling can be reduced when taper equations are based on the adjustment of parameters. When applied to Max and Burkhart's segmented polynomial system (1976), mixed effect modeling increased the fit of their taper equation, especially in the lower bole section (Trincado and Burkhart 2006). When covariance structures are heterogeneous and/or correlations exist among stochastic error terms, mixed effect modeling yields an unbiased and consistent estimation of parameters (Zhang 2002).

Trincado and Burkhart (2006) used a nonlinear mixed effects model written in the form of a two stage model. Model (6) accounts for the variation within the tree, while model (7) describes the between tree variations. Within tree variation model is as follows:

$$R_i(\beta_i, \xi) = \sigma^2 G_i^{1/2}(\beta_i, \theta) \Gamma_i(\rho) G_i^{1/2}(\beta_i, \theta) \quad (18)$$

where: $R_i(\beta_i, \xi)$=is a variance-covariance structure,

$G_i^{1/2}(\beta_i, \theta)$=is an ($n_i$ x n_i) diagonal matrix characterizing interindividual variance,

$\Gamma_i(\rho)$= is an (n_i x n_i) matrix that describes the correlation pattern within the measurements of the ith individual,

ξ=represents a vector of unknown parameters$[\sigma, \theta', \rho']^{\mathrm{T}}$, and

B_i=parameter to be estimated.

This assumes that the within tree variance is heterogeneous and that residuals are uncorrelated.

Between tree variation is modeled as follows:

$$\beta_i = \mathrm{A}_i \beta + \mathrm{B}_i \mathrm{b}_i \qquad \mathrm{b}_i \sim \mathrm{N}(0, \mathrm{D}) \quad (19)$$

where: A_i=is a design matrix of size (r x p) for fixed effects,

B_i=is a design matrix of size (r x q) for the random effects,

β=a vector of fixed population parameters of size (p x 1), and

B_i=a vector of random effects of size (q x 1) associated with the ith individual and assumed to be multivariate normally distributed with $E(b_i)=0$ and variance-covariance structure D.

Summary

In summary, many forms of taper equations have been used to predict outside bark diameters along an entire tree bole. Some of these equations were used in this study, while others were not. The reasons as to why certain model forms were not chosen for investigation will be addressed in the discussion chapter of this thesis.

Objectives

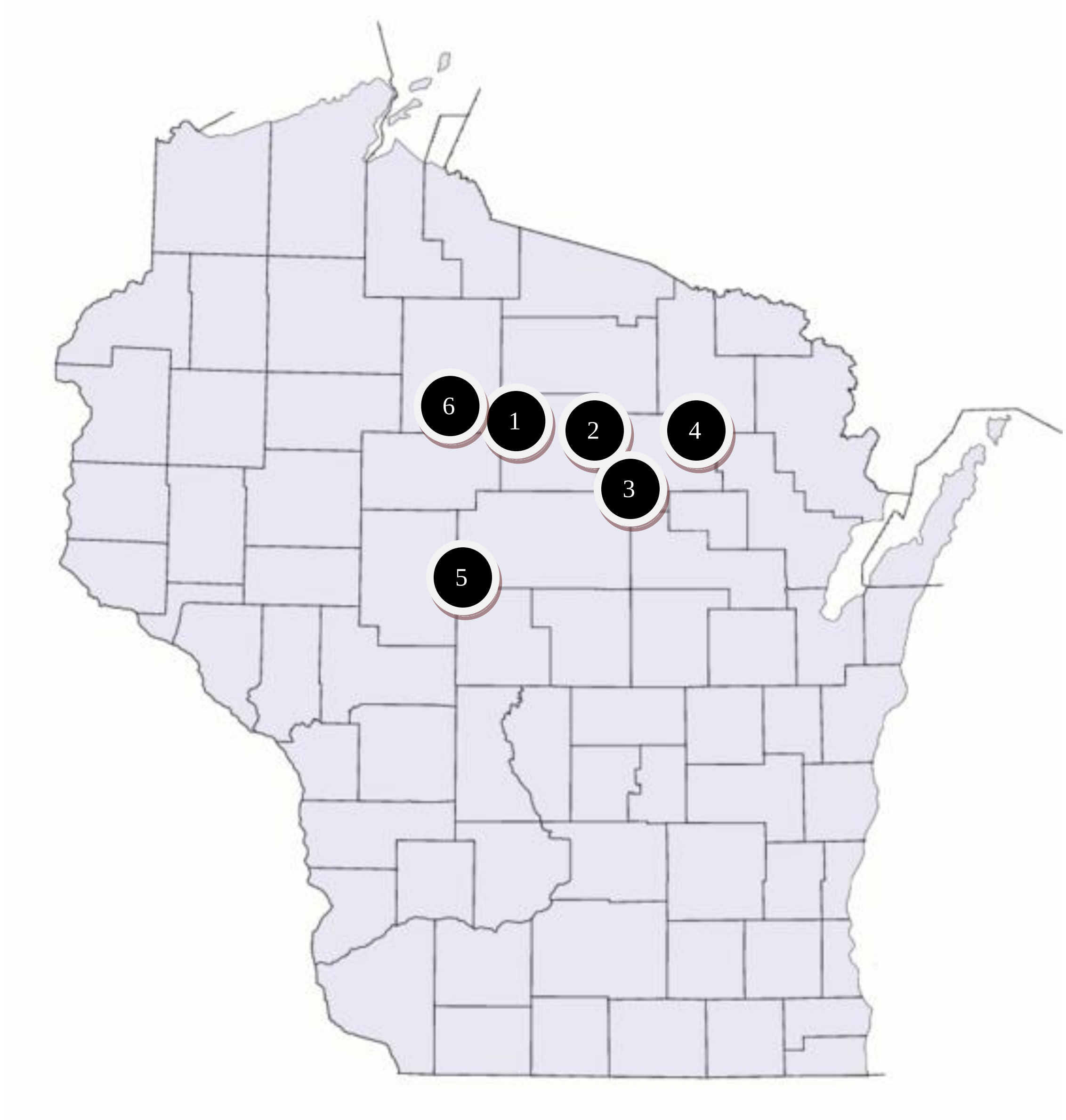
The objectives of this study are to develop species specific outside-bark profile equations, from standing trees, for several commercially important timber species in Wisconsin. Equations will be developed to estimate outside bark diameter at the stump and outside bark diameter for any height at or above 4.5 feet above ground. The specific species being used are sugar maple (*Acer saccharum* Marsh.), American basswood (*Tilia americana* L.), and bigtooth aspen (*Populus grandidentata* Michx.). Additionally, ash (*Fraxinus* spp.) will be examined.

Methods

Data Collection

Throughout 2008, six northern hardwood stands were visited in north central Wisconsin for data collection. Six study sites were included in this study; two sites at Treehaven, owned and operated by the University of Wisconsin-Stevens Point College of Natural Resources, near Tomahawk, WI, one site provided by, and located in, Lincoln County, one site provided by Kretz Lumber Company Inc. near Crandon, WI, one site provided by Plum Creek Timber Company Inc. near Spirit, WI, and the final site at the The Highgrounds Veterans Memorial Park, near Nellsville, WI (Figure 1). It was within the goals of the research project to obtain a sample size of roughly 100 trees per species. Since this project is a part of a larger project involved with measuring felled trees, sites were somewhat limited to areas where data for all parts of the project could be collected. The inclusion of certain merchantable sized trees was based upon harvesting activities. Therefore, this is an observational study, as opposed to an experimental one. A variety of site conditions and a range of tree diameters and heights were sought. This broad range of characteristics should increase the applicability of developed taper equations.

The tools used during the extent of the field work for this project include calipers, vertex hypsometers, Spiegel Relaskops, and Laser Technology Inc. Criterion RD 1000 dendrometers (hereafter referred to as the Criterion RD 1000). Calipers were used for the lower bole diameter measurements of standing trees. The Haglof Vertex III Hypsometer was used to assist with distance measurements when using the Criterion RD 1000. The Spiegel Relaskop was used to estimate the total heights of trees, as well as height to a



Key
1. West end of Treehaven Field Station, off of County Road H
2. North end of Treehaven Field Station, off of Pickerel Road
3. Lincoln County, WI forest, near County Road A
4. Kretz Lumber Company Inc. site, off Rolling Stone Road, near Mole Lake, WI
5. The Highgrounds Veterans Memorial Park, Nellsville, WI
6. Plum Creek Timber Company Inc. site, off of Highway 86, near Spirit, WI

Figure 1 County map of Wisconsin illustrating study site locations.

10 in. and 3 in. top. The Criterion RD 1000 is a forestry tool specifically created for timber inventories. Measurements obtained with the Criterion RD 1000 were upper stem outside bark diameters and heights, both of which are slope corrected. When measuring standing trees, two Criterion RD 1000s were used to provide paired upper stem diameter/height measures perpendicular to one another, which were then averaged. The measuring device provides heights, based on distance away and the angle of tilt, and diameters, based on a scope that is able to be toggled to meet the outer edges of the tree bole. The units were placed between 20 and 40 feet away, depending on the height of the tree and visibility to the upper diameters.

Several attributes were recorded for each tree included within the study:

1. Location.
2. Species.
3. Diameter at breast height, using calipers and the Criterion RD 1000.
4. Outside-bark diameter every foot along the bole from 0.5 feet aboveground to 4.5 feet above ground, using calipers and the Criterion RD 1000.
5. Outside-bark diameter at least every 3 feet above DBH above ground and at any point along the bole where a noticeable change in diameter occurs, measured with the Criterion RD 1000.
6. Height to a 10 inch top, 3 inch top, and total height, using the Spiegel Relaskop and the Criterion RD 1000.

All measured diameters, regardless of the tool used to measure, were taken twice at locations perpendicular to one another.

Taper Equation Development

Taper equations will be developed for each species such that estimation of volume to a merchantable height can be derived from a tree's DBH measure and an estimate of merchantable height. The two attributes recorded, commonly taken during a timber inventory, will easily become applicable to the developed equations. Equations will be developed for:

1. Outside bark diameter at the stump, and
2. Outside bark diameter for any height along the tree bole.

Both equations will be used when determining volume for each sample tree. Volume will be determined from using only equations developed for outside bark diameter for any height along the tree bole. Also, volume will be predicted from the summation of equations developed for outside bark diameter at the stump, predicting bole volume from 0.5 feet up to DBH, and the equations developed for outside bark diameter at any height along the tree bole, predicting bole volume from DBH to the top of the tree. Both sets of volume predictions will be compared to check for accuracy, and whether or not the use of developed outside bark diameter at stump equations are warranted. Previously published model forms of taper equations will be used for prediction of outside bark diameter for any height along the tree bole, which will then be compared to assist in development of equations for four Wisconsin timber species.

Analysis

After data collection was completed, data were analyzed using Microsoft Excel® and SAS® computer programs. Several techniques were employed to parameterize the profile equations. While nonlinear seemingly unrelated regression, nonlinear mixed effect modeling, polynomial regression, and segmented regression have been used in prior development of taper equations, two variable form exponent approaches and one polynomial model were used in this study. The model forms not selected for use in this study will be discussed in the discussion section of this thesis. Through the use of the SAS® statistical software package, the method with the least bias and smallest error was suggested for use as the best taper equation. Accuracy of developed profile equations was assessed by regression diagnostics where appropriate.

Outside Bark Stump Equation

An equation was developed to estimate stump diameter (at 0.5 feet) for four tree species in Wisconsin to aid forestry personnel and landowners in volume estimation of a tree's region of buttswell. Two models were developed for investigation of stump diameter, which were then fit to the dataset for each tree species to check for significant parameters, normal distribution, and to apply regression diagnostics. The two model forms developed to examine stump diameter include:

$$Stump\ DOB_i = \beta_0 + \beta_1 D_i + \varepsilon_i \tag{20}$$

$$Stump\ DOB_i = \beta_0 + \beta_1 (D_i^2 H_i) + \varepsilon_i \tag{21}$$

where: *Stump* DOB_i=outside bark diameter (inches) at 0.5 feet,

ε_i=error term, and

all other terms as previously defined.

Least squares regression was used to fit the developed models and slopes were deemed significantly different from zero if the p-value of the resultant significance tests were

<0.05. The most appropriate model was selected based on significant slopes, the highest adjusted R^2 value, and lowest MAE. Iteration, through the use of Smalian's formula on the lower bole section of the tree, from 0.5 feet up to the DBH, was used to determine volume.

Outside Bark Profile Equation

An equation was developed from previously published model forms to estimate diameter at given heights along the length of a tree bole for four tree species in Wisconsin as to aid forest managers, industry personnel, and landowners in volume estimation to differing merchantable tops. As mentioned in the literature review, there are many different types of model forms to choose from to estimate diameter at different heights along a tree bole. Three models chosen for investigation for each tree species were fit to the dataset to check for significant parameters, to find a pattern of normal distribution, and to apply regression diagnostics. The three model forms chosen to examine include:

$$d^2/D^2 = b_0 + b_1(h/H) + b_2(h^2/H^2) \tag{4}$$

$$\begin{aligned} ln(\,d_i) = {} & ln(a_0) + a_1\,ln(D) + ln(a_2)\,D + b_1\,ln(X)\,Z^2 \\ & + b_2\,ln(X)\,ln(Z + 0.001) + b_3\,ln(X)\sqrt{Z} \\ & + b_4\,ln(X)\,e^Z + b_5\,ln(X)\,(D/H) \end{aligned} \tag{14}$$

$$\begin{aligned} ln(\,d_i) = {} & ln(a_0) + a_1\,ln(D) + a_2\,ln(H) + b_1\,ln(X)\,X^{\frac{1}{10}} \\ & + b_2\,ln(X)\,Z^4 + b_3\,ln(X)\,arcsin(Q) \\ & + b_4\,ln(X)\,(1/e^{D/H}) + b_5\,ln(X)\,DBH^X \end{aligned} \tag{16}$$

Terms within the models not significantly different from zero were excluded from the model and further investigation commenced without said terms. Least squares regression was used to fit the chosen models and parameter estimates were deemed insignificantly different from zero if the p-value of the resultant significance tests were ≥0.05. Since multiple observations were taken upon the same experimental unit, high degrees of serial autocorrelation could be expected. Mixed linear models of equations (4, 14, and 16) were used to vary the covariance structure of the matrices (Quinn and Keough 2002). First-order autoregressive and compound symmetry covariance structures, along with the standard identity matrix covariance structure, were applied to find the Akaike Information Criterion (AIC) and -2 log likelihoods. Models with the lowest AIC and -2 log likelihood have a lesser degree of serial autocorrelation and weigh more heavily in final model selection. Once appropriate models were identified, iteration, through the use of Smalian's formula on individual tree segments, was used to assess accuracy in volume determination.

Regression Diagnostics

After the several models were chosen for use in development of the taper equations, many diagnostics were observed for each model to determine the equation with the most accurate estimation of taper. An evaluation of selected models included fit statistics calculated from ordinary residuals y_i-$\hat{y}_i$, where y_i and $\hat{y}_i$ are the observed and predicted values of the dependent variable, respectively. Regression analysis will lead to a model that most highly predicts the dependent variable. Diagnostics specifically used for this research include Shapiro-Wilk's test of normality of errors, approximate R^2, mean absolute error, PRESS residual and the PRESS statistic, DFFITS, variance inflation factor, and the Durbin-Watson test for positive autocorrelation (Myers 1990).

To test for normal distribution of the error term, the Shapiro-Wilk test was used in the model selection process (Quinn and Keough 2002). This test corresponds to the following hypothesis when the residual error terms are checked:

H_0: Errors are normal

H_1: Errors are not normal

α=0.05

The Shapiro-Wilk test examines the residual error for data that are not normally distributed in the case of this study. Therefore, the null hypothesis should not be rejected for data if normality is desired. Models that suppose the alternative hypothesis are discarded or transformed to meet the assumption. When the Shapiro-Wilk test was not available in SAS® procedures, the Kolmogorov-Smirnov test will be used to check for normality.

The coefficient of determination (R^2), often used to measure the fit of the regression line, represents the variability of the Y variable that can be predicted by the independent variables (Myers 1990). The R^2 is a value that lies between 0 and 1, and is usually expressed as a percentage (0-100%) of the variation in the response data that is explained by the model. A high degree of predictive ability of the independent variable is achieved the closer R^2 approaches 1 (or 100%). In this study, some form of diameter at a given height above ground was the dependent variable while many other observed measurements were the independent variables. When simple linear regression was used in developing a taper equation, the following model was used to calculate R^2:

$$R^2 = \frac{SS_{Reg}}{SS_{Total}} * 100\% \tag{22}$$

where: SS_{Reg}=Sum of squares regression-sum of squared differences between the estimated regression line and the mean of the Y variable, and

SS_{Total}=Sum of squares total-sum of squared differences between each of the observations and the mean of the Y variable.

For equations with transformed data, each R^2 value was calculated by hand, to get transformed data back into original units, by taking the difference between SS_{Total} and SS_{Error}, predicting the SS_{Reg}. Because of this, all coefficients of determination for equations 14 and 16 are approximate R^2 values.

For equations developed using multiple linear regression, adjusted R^2 was used instead of R^2. With additional independent variables (more parameters) added to a model, R^2 will remain constant or increase regardless of the improved/decreased fit of the model. To take account of the predictive ability of a model with additional parameters and compare multiple linear regression models, adjusted R^2 should be used. The adjusted R^2 can be calculated using the following equation:

$$Adjusted\ R^2 = 1 - \left(\frac{n-1}{n-p}\right)\left(\frac{SSE}{SST}\right) \tag{23}$$

where: Adjusted R^2=percent variation of Y explained by multiple X's,

n=number of observations,

p=number of parameters,

SSE=Sum of squared error-sum of squared differences between the actual and estimated dependent variable values, and

all other terms as previously defined.

As additional parameters are added to a model, the value of the adjusted R^2 will decrease unless the additional parameters add to the usefulness of the model by decreasing the SSE. This measure takes into account the true effect of any new independent variables introduced to a model.

Another statistic used for comparing models is mean absolute error (MAE), which is determined from the regression residuals. MAE is calculated from the following equation:

$$MAE = \frac{\sum|y_i - \hat{y}_i|}{n} \tag{24}$$

where: y_i=actual value of observation i,

$\hat{y}_i$=estimated value of observation i,

n=number of observations, and

i=1,2,3,..n.

The calculated mean of each residual's absolute value leads to a calculation of average deviation from the regression. Absolute value is used so that errors on the positive and negative side do not cancel each other out when averaging them together. When several models are compared in this manner, the equation with the lowest MAE is generally chosen because this test shows that actual values of observations are closer, on average, to the estimated regression line.

Predicted sum of squares (PRESS) was used as a form of validation and to determine the influence of individual observations when comparing models (Myers 1990). In the dataset, each observation is individually withheld, using the remaining n-1 observations to estimate coefficients of each regression model. The deleted response is estimated n times, one for each observation, so as to yield a prediction error or PRESS residual. A PRESS residual is calculated by:

$$PRESS_i = Y_i - \hat{Y}_{i,-i} \tag{25}$$

where: $PRESS_i$=PRESS residual for observation i,

Y_i=actual value of observation i,

$\hat{Y}_{i,-i}$=estimated value of observation i from a regression equation that was constructed without using that observation, and

i=1, 2, 3…n.

The value given is a measure of how much influence each observation has on the calculated regression. If the absolute value of the PRESS residual is significantly larger than the ordinary residual for a given observation, then that observation may be highly influential in the construction of the regression.

A PRESS statistic can be calculated from all (n) PRESS residuals using the following equation:

$$PRESS = \sum_{i=1}^{n}(Y_i - \hat{Y}_{i,-i}) \qquad (26)$$

where: All terms are as previously defined.

A large PRESS statistic may occur if one or a few PRESS residual values are large. When comparing models, the one with the lowest PRESS statistic will be chosen.

To determine the influence each observation i had upon the predicted value of $\hat{y}_i$, the diagnostic DFFITS was used (Myers 1990). The value of DFFITS for each observation i is the number of estimated standard errors the fitted value of $\hat{Y}_i$ changes without said i^{th} observation. This diagnostic can check for outliers within the dataset. DIFFITS can be calculated by the following equation:

$$DFFITS_i = \frac{\hat{y}_i - \hat{y}_{i,-i}}{s_{-i}\sqrt{h_{ii}}} \qquad (27)$$

where: $DFFITS_i$=Difference in fits for observation i,

s_{-i}=Residual standard deviation calculated without observation i,

h_{ii}=i^{th} diagonal of the hat matrix, and

all other terms as previously defined.

An absolute value of 2 for a DFFITS calculation should be closely examined, as it may possibly be an outlier within the dataset (Myers 1990).

Variance inflation factor (VIF) was the diagnostic used to determine the level of collinearity among each independent variable. Collinearity exists when independent variables are correlated, leading to regression coefficients that are highly dependent on the dataset generating them and poor predictive ability of the model (Myers 1990). Therefore, slight changes in the dataset can drastically change the estimated regression coefficients. VIF can be calculated by:

$$VIF = \frac{1}{1-R_i^2} \tag{28}$$

where: VIF=variance inflation factor for the i^{th} regressor, and

R_i^2=the coefficient of multiple determination of the regression produced

by regressing the dependent variable against the other regressor variables.

The value calculated is the increase over the ideal case where regressor variables are orthogonal. A VIF value greater than ten would suggest high collinearity and should be examined closely (Quinn and Keough 2002). This diagnostic was weighed heavily when choosing an appropriate taper model.

With repeated measures upon each sample tree contained within the study, autocorrelation might exist, impacting the value of parameter estimates. The Durbin-Watson test for positive autocorrelation was used to detect autocorrelation for each proposed model.

Lack of fit tests were planned to check the adequacy of the models (Myers 1990). This test would have been possible if there were several instances of repeated measurements within a species at each height above the ground with matching independent variables of total height and DBH. This test could not be done to assess observational and experimental errors, as such instances did not occur.

Partitioning a dataset into two series of data, a fitting sample and a validation sample, known as data splitting, was a common practice used during regression analysis (Myers 1990). The fitting sample data is used with models to estimate regression coefficients, which are then used to estimate samples within the validation dataset. The problem with such a procedure revolves around the fact that with a splitting of data, fewer samples are used during model development and construction. Kozak and Kozak (2003) concluded that the data splitting procedure does not provide any additional information about model building compared with the respective statistics derived explicitly from the entire dataset. Myers (1990) states that the PRESS statistic can be used in lieu of data splitting, where it is similar to the procedure but uses the entire dataset. With this information, data splitting was not used for validation purposes during this project.

Volume Determination

To determine the adequacy of volume determination for our proposed models, Smalian's formula was used to compute cubic volume (feet) from observed and estimated variables. Avery and Burkhart (2002) present the formula as:

$$Cubic\ foot\ volume = \frac{(B+b)}{2} L$$

(29)

where: B=cross sectional area (sq. ft.) at large end of log (0.005454*d^2),

b=cross sectional area (sq. ft.) at small end of log (0.005454*d^2), and

L=log length (ft.).

Though not as accurate as Huber's or Newton's formulas on more traditional applications to 16 foot logs, Smalian's formula is easy to apply using the data previously collected and will be fairly accurate giving the short stem segments used herein. Prior measurements (Y) were taken at standard increments (X) along the tree bole, yielding cross sectional areas at large and small ends of various bole sections. These sections ranged in size from 0.5 ft. to 4 ft. in length, thereby increasing the accuracy of Smalian's formula to that of Huber's or Newton's formula. By summing all sections for a single sample tree, an estimated volume is achieved.

$$cubic\ foot\ volume = \sum_{i=1}^{n} (\frac{B+b}{2} L) \tag{30}$$

where: n=number of segments in the tree stem, and

all other terms are as previously defined.

This method is done for every sample tree with observed (y_i) diameters from data collection and estimated ($\hat{y}_i$) diameters from taper models. Lower bole volumes (from stump to DBH) will be estimated with the outside bark stump equation and will be added to volume estimates (DBH to finite top) from selected taper models. These will be compared to volume estimates using only the taper models.

To compare model accuracy, mean absolute error of estimated total volume to observed total volume, for all trees within a species, will be used. The model exhibiting the lowest MAE will be weighed heavily in choosing a sufficient model.

Results

Data Summarization

The target sample size was to obtain 100 individual trees per species (400 total) during 2008. However, due to logistical difficulties, only 312 trees could be measured and analyzed in this study. All 312 trees were used in testing and fitting of regression equations. Table 1 summarizes the tree level attributes by species. The attributes include sample size, average DBH (in.), average total tree height (ft), average height to a ten inch diameter top (ft), average height to a three inch diameter top (ft), and average Girard form class. Standard deviations of each attribute are included in the table. The distribution of trees per species in the regression dataset based on 1" DBH class and 10-ft total height class are shown in Tables 2-5.

Outside Bark Stump Equation

Two model forms were developed to examined to obtain the outside bark stump diameter of four tree species in Wisconsin. Residual error terms of the regression fits were checked for non-normal distribution using the Shapiro-Wilk test for normality. Both models were found to reject the null hypothesis, meaning that residual error terms were not normally distributed. However, the residual error terms did assume a bell shaped curve that was highly peaked in the middle. This will be further addressed in the discussion section of this thesis.

For each developed model examined, all non-intercept parameter estimates needed to be significantly different from zero at the α=0.05 level. For both models and

Table 1 Mean and standard deviation (in parentheses) of tree attributes by species

Species	n	DBH	THT	HT10	HT3	FC
Sugar maple	80	9.5(3.02)	66.6(12.2)	15.6(17.0)	52.4(10.8)	86(8)
Ash	84	11.7(3.83)	80.4(13.9)	24.3(19.6)	60.4(15.4)	85(7)
Aspen	82	10.4(2.75)	75.3 (9.3)	17.5(15.3)	59.1 (9.2)	89(4)
Basswood	66	12.9(3.19)	80.7(11.7)	30.2(17.7)	68.5(11.2)	88(4)

n=Sample size
DBH=Diameter (in.) at breast height
THT=Total tree height (ft.)
HT10=Height (ft.) of tree to a 10 inch diameter top
HT3=Height (ft.) of tree to a 3 inch diameter top
FC= Girard form class (%)

Table 2 Distribution of sugar maple in the dataset by DBH and total height class

Total height by 10-ft class	1 inch DBH class 6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Total
40		3														3
50	6	2	4	1												13
60		5	5	4	4											18
70		4	8	2	4	6	1	2		1	1					29
80		1	2		1		4		1				1			10
90				1			1			3		1			1	7
Total	6	15	19	8	9	6	6	2	1	4	1	1	1	0	1	80

Table 3 Distribution of ash in the dataset by DBH and total height class

Total height	1 inch DBH class																	
by 10-ft class	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	Total
50	1	1																2
60	2	1	3	1	3	1												11
70		2	2	5	1	3	2			1								16
80				1	4	4	5	5	3	1	2	1	1					27
90					1	1		1	2	2	3	1	1		1			13
100								2	1		3	2	2	1	2			13
110														1			1	2
Total	3	4	5	7	9	9	7	8	6	4	8	4	4	2	3	0	1	84

Table 4 Distribution of aspen in the dataset by DBH and total height class

Total height	1 inch DBH class												
by 10-ft class	6	7	8	9	10	11	12	13	14	15	16	17	Total
50				1									1
60	1	3	4	3				1					12
70	2	5	7	6	2	1	2	2	1				28
80				4	6	7	5	6	2	1	1	1	33
90				1	1	1				1	2		6
100									1	1			2
Total	3	8	11	15	9	9	7	9	4	3	3	1	82

Table 5 Distribution of basswood in the dataset by DBH and total height class

Total height by 10-ft class	1 inch DBH class														
	6	7	8	9	10	11	12	13	14	15	16	17	18	19	Total
60	2		2	1			1	1							7
70			2	3		3	1	4		1					14
80				1	3	4	1	5	3	1		4	1	1	24
90					2	1	1	2	2		2	1	1	1	13
100								1	1	2	3			1	8
Total	2	0	4	5	5	8	4	13	6	4	5	5	2	3	66

all species, parameter estimates were significant. Regression diagnostics including R^2 and MAE for all species for both models are included in Table 6. Higher R^2 and lower MAE values for all species were seen for equation (20) compared to equation (21). Therefore, equation (21) was discarded and equation (20) was selected for use as the outside bark stump equation. Regression coefficients for each species using equation (20) are included in Table 7.

Outside Bark Profile Equation

Three previously published model forms were examined to obtain profile equations for four tree species in Wisconsin. Residual error terms of the regression fits were checked for non-normal distribution using either the Shapiro-Wilk test for normality or the Kolmogorov-Smirnov test for normality. All models were found to reject the null hypothesis, meaning that residual error terms of the regression fits were not normally distributed. However, the residual error terms did assume a bell shaped curve, being symmetrical and non-skewed.

For each model form examined, all non-intercept parameter estimates needed to be significantly different from zero at the $\alpha=0.05$ level. For the fit of equation (4), all parameter estimates for all species were significantly different from zero. Except for basswood, the fit of equation (14) for all other species had one term that was not significantly different from zero. Since insignificant parameters should not be included, a new version of equation (14) was developed without said insignificant terms:

Table 6 Fit statistics for stump equation models (20, 21) by species and equation

Species	R^2	MAE (in.)
Sugar maple		
Equation (20)	0.7841	1.1586
Equation (21)	0.7679	1.1704
Ash		
Equation (20)	0.9480	0.8960
Equation (21)	0.8669	1.4726
Aspen		
Equation (20)	0.9440	0.7224
Equation (21)	0.9005	0.9757
Basswood		
Equation (20)	0.8646	1.5270
Equation (21)	0.7940	1.7779

Table 7 Regression coefficients for equation (20) by species

Species	Parameter	Estimate	Standard Error	Pr>|t|
Sugar maple	b_0	-1.2617	0.8968	0.1645
	b_1	1.4248	0.0950	<0.0001
Ash	b_0	-1.0370	0.5223	0.0522
	b_1	1.4659	0.0455	<0.0001
Aspen	b_0	-1.6779	0.0162	0.0004
	b_1	1.4282	0.0428	<0.0001
Basswood	b_0	-1.2513	0.0366	0.2720
	b_1	1.5070	0.0861	<0.0001

$$ln(\,d_i) = ln(a_0) + a_1\,ln(D) + b_1\,ln(X)\,Z^2$$

$$+b_2\,ln(X)\,ln(Z + 0.001) + b_3\,ln(X)\,\sqrt{Z}$$

$$+b_4\,ln(X)\,e^Z + b_5\,ln(X)\,(D/H) \tag{31}$$

where: all terms as previously defined.

All terms for equation (31), for all species other than basswood, were found to be significantly different from zero. For equation (16), all species had at least one parameter that was not significantly different from zero. Several new versions of equation 16 were developed for each species to contain only significant parameters. For sugar maple:

$$ln(\,d_i) = ln(a_0) + a_1\,ln(D) + a_2\,ln(H) + b_1\,ln(X)\,X^{\frac{1}{10}}$$

$$+b_2\,ln(X)\,Z^4 + b_3\,ln(X)\,arcsin(Q) + b_5\,ln(X)\,DBH^X \tag{32}$$

$$ln(\,d_i) = ln(a_0) + a_1\,ln(D) + a_2\,ln(H) + b_1\,ln(X)\,X^{\frac{1}{10}}$$

$$+b_2\,ln(X)\,Z^4 + b_3\,ln(X)\,arcsin(Q) + b_4\,ln(X)\,(1/e^{D/H}) \tag{33}$$

$$ln(\,d_i) = ln(a_0) + a_1\,ln(D) + a_2\,ln(H) + b_1\,ln(X)\,X^{\frac{1}{10}} + b_2\,ln(X)\,Z^4 \tag{34}$$

where: all terms are as previously defined.

For ash:

$$ln(\,d_i) = ln(a_0) + a_1\,ln(D) + b_1\,ln(X)\,X^{\frac{1}{10}} + b_2\,ln(X)\,Z^4$$

$$+b_3\,ln(X)\,arcsin(Q) + b_4\,ln(X)\,(1/e^{D/H}) \tag{35}$$

$$ln(\,d_i) = ln(a_0) + a_1\,ln(D) + b_1\,ln(X)\,X^{\frac{1}{10}} + b_2\,ln(X)\,Z^4$$

$$+b_4\,ln(X)\,(1/e^{D/H}) \tag{36}$$

where: all terms as previously defined.

For aspen:

$$ln(\,d_i) = ln(a_0) + a_1\,ln(D) + a_2\,ln(H) + b_1\,ln(X)\,X^{\frac{1}{10}}$$

$$+b_2\,ln(X)\,Z^4 + b_4\,ln(X)\left(1/e^{D/H}\right) \tag{37}$$

$$ln(\,d_i) = ln(a_0) + a_1\,ln(D) + b_1\,ln(X)\,X^{\frac{1}{10}}$$

$$+b_2\,ln(X)\,Z^4 + b_5\,ln(X)\,DBH^X \tag{38}$$

where: all terms as previously defined.

For basswood:

$$ln(\,d_i) = ln(a_0) + a_1\,ln(D) + a_2\,ln(H) + b_1\,ln(X)\,X^{\frac{1}{10}}$$

$$+b_2\,ln(X)\,Z^4 + b_4\,ln(X)\,(1/e^{D/H}) + b_5\,ln(X)\,DBH^X \tag{39}$$

$$ln(\,d_i) = ln(a_0) + a_1\,ln(D) + a_2\,ln(H) + b_1\,ln(X)\,X^{\frac{1}{10}}$$

$$+b_2\,ln(X)\,Z^4 + b_4\,ln(X)\left(1/e^{D/H}\right) \tag{37}$$

where: all terms as previously defined.

The parameter estimates for all models [equations (4, 14, 31-39)], for their respective species, were examined to see if they were significantly different from zero while simultaneously being evaluated for serial autocorrelation. Fits for all equations produced a low Durbin-Watson test statistic (largest value was 0.977), and when interpolated to standard Critical Values of the Durbin-Watson Test Statistic table (Myers 1990), inference could be made to reject the null hypothesis and accept that each equation has positive autocorrelation. To address the autocorrelation, covariance structures of each equation were manipulated to produce the lowest AIC value, hence a lower level of autocorrelation. First-order autoregressive and compound symmetry covariance structures were chosen over other structures because of the use of dependent data in the dataset. Both structures were applied to each equation for each respective species. Tables 8-10 contain fit statistics (when applicable) comparing approximate/adjusted R^2, MAE, AIC, PRESS, maximum VIF, and Durbin-Watson test statistics for each equation, for each species, with original (identity matrix) and selected covariance structures (only when all parameter estimates are significantly different from zero) to assist in evaluating the best equation for each base model form [equations (4, 14, 16)].

For equation (4), fits of the equation for all species under the identity matrix covariance structure produced a high degree of positive autocorrelation and should therefore not be selected. All species under the first-order autoregressive covariance structure yielded the lowest AIC values when compared with the compound symmetry covariance structure. The compound symmetry covariance structure did produce lower MAE and higher approximate R^2 values [both back transformed into original units of measure (in.)] than the first-order autoregressive covariance structure for all species, but due to high positive autocorrelation and lowest AIC value, the first-order autoregressive covariance structure was selected as the best form of equation (4). As seen in Table 8, multicollinearity is not a problem with this equation because all variance inflation factors values remained near or under ten. Parameter estimates and standard errors for fits of equation (4) under the first-order autoregressive covariance structure for all species are included in Table 11.

As with equation (4), fits of equations (14) and (31) for all species under the identity matrix covariance structure produced a high degree of positive autocorrelation and should therefore not be selected for use. For all species, fits of equations using a

Table 8 Fit statistics for base model equation (4) by species, equation, and covariance structure

Species	Adjusted R^2	MAE (in.)	AIC	PRESS	Max VIF	DW
Sugar maple						
Equation (4)						
IM	0.7582	0.9694	-242.2	96.004	10	0.865
AR(1)	0.8620	1.1604	-1999.2			
CS	0.8780	0.9637	-332.0			
Ash						
Equation (4)						
IM	0.8503	1.0034	-1370.2	70.24	10	0.887
AR(1)	0.8970	1.2283	-3569.1			
CS	0.9190	1.0026	-1406.1			
Aspen						
Equation (4)						
IM	0.8975	0.6840	-2633.6	34.08	10	0.688
AR (1)	0.9210	0.7878	-4859.2			
CS	0.9320	0.6770	-2865.1			
Basswood						
Equation (4)						
IM	0.8138	1.0909	-805.2	70.06	11	0.969
AR(1)	0.8500	1.5092	-2710.3			
CS	0.8960	1.0875	-823.8			

IM=identity matrix covariance structure
AR (1)=first-order autoregressive covariance structure
CS=compound symmetry covariance structure
Max VIF=maximum VIF associated with any single parameter for stated equation
DW=Durbin-Watson test statistic

Table 9 Fit statistics for base model equation (14) by species, equation, and covariance structure

Species	Approximate R^2	MAE (in.)	AIC	PRESS	Max VIF	DW
Sugar maple						
Equation (31)						
IM	0.914	0.7257	-1018.9	63.42	96120	0.721
AR(1)	0.905	0.7850	-2418.5			
CS	0.916	0.7178	1383.8			
Ash						
Equation (31)						
IM	0.954	0.7007	-2533.5	41.31	81417	0.558
AR(1)	0.947	0.7617	-4649.8			
CS	0.954	0.7039	-2912.8			
Aspen						
Equation (31)						
IM	0.955	0.5116	-2529.8	35.51	86905	0.439
AR (1)	0.947	0.5845	-4763.5			
CS	0.955	0.5162	-2946.8			
Basswood						
Equation (14)						
IM	0.953	0.6779	-1639.2	44.49	75301	0.908
AR(1)	0.948	0.7268	-2379.6			
CS	0.953	0.6752	-1829.6			

Table 10 Fit statistics for base model equation (16) by species, equation, and covariance structure

Species	Approximate R^2	MAE (in.)	AIC	PRESS	Max VIF	DW
Sugar maple						
Equation (32)						
IM	0.902	0.7859	-940.3	65.95	125	0.803
CS	0.905	0.7801	-1120.9			
Equation (33)						
IM	0.905	0.7795	-943.3	66.32	513	0.813
Equation (34)						
IM	0.900	0.7712	-893.8	67.74	3	0.814
AR (1)	0.883	0.8522	-2052.0			
CS	0.901	0.7665	-1081.1			
Ash						
Equation (35)						
IM	0.951	0.7691	-2423.3	43.59	740	0.553
AR(1)	0.934	0.9154	-4439.3			
Equation (36)						
IM	0.950	0.7673	-2425.5	43.61	635	0.557
AR (1)	0.945	0.8231	-4351.2			
CS	0.950	0.7690	-2839.3			
Aspen						
Equation (37)						
IM	0.943	0.6023	-2137.3	43.26	529	0.525
CS	0.944	0.5968	-2422.6			
Equation (38)						
IM	0.939	0.6171	-2081.1	44.18	10	0.538
AR (1)	0.891	0.9077	-4165.6			
CS	0.934	0.6166	-2361.0			
Basswood						
Equation (37)						
IM	0.929	0.7898	-1376.7	51.54	297	0.869
AR(1)	0.930	0.8209	-2274.7			
CS	0.930	0.7829	-1527.4			
Equation (39)						
IM	0.943	0.7464	-1429.7	49.91	368	0.837
AR (1)	0.942	0.7850	-2320.8			
CS	0.944	0.7381	-1584.4			

Note: For equations (32), (35), and (37) (basswood), the identity matrix covariance structure fit statistics were included to show PRESS, Max VIF, and DW though not all parameter estimates were significantly different from zero.

Table 11 Regression coefficients for equation (4) by species using first-order autoregressive covariance structure

Species	Parameter	Estimate	Standard Error	Pr>|t|
Sugar maple	b_0	1.5559	0.0326	<0.0001
	b_1	-2.7574	0.1133	<0.0001
	b_2	1.1782	0.0895	<0.0001
Ash	b_0	1.5467	0.0259	<0.0001
	b_1	-3.0965	0.0889	<0.0001
	b_2	1.5777	0.0660	<0.0001
Aspen	b_0	1.3660	0.0162	<0.0001
	b_1	-2.3467	0.0589	<0.0001
	b_2	1.0084	0.0457	<0.0001
Basswood	b_0	1.5805	0.0366	<0.0001
	b_1	-2.7456	0.1185	<0.0001
	b_2	1.2159	0.0918	<0.0001

compound symmetry covariance structure produce lower MAE and higher approximate R^2 values than when using a first-order autoregressive covariance structure. First-order autoregressive covariance structures did produce the lowest AIC values for all species, and because of the autocorrelation issue, were selected as the best form of equation (14). DFFITS values for these equations, using the identity matrix covariance structure, did not exceed two for any species. As seen in Table 9, variance inflation factor values are extremely high. Multicollinearity is a major concern regarding the use of this base equation model. Parameter estimates and standard errors for fits of equations (14 and 31) (respectively for species) using a first-order autoregressive covariance structure are included in Table 12.

Once again for fits of equations (32-39), all species using an identity matrix covariance structure produced a high degree of positive autocorrelation and should not be selected for use. For all species the use of first-order autoregressive covariance structures produced the lowest AIC values, dealing with the issue of autocorrelation better than a compound symmetry covariance structure. Even though compound symmetry covariance structures produced lower MAE and higher R^2 values for all equations, they were not selected because of the importance of autocorrelation. Since equation (34) produced the lowest AIC value and had VIF less than 10, it was selected as the best equation to use for sugar maple for base model equation (16). Even though VIF values were high, equation (35) had the lowest AIC values for ash and was selected as the best equation to use for base model equation (16). Aspen, using equation (38) had the lowest AIC values and acceptable VIF values, therefore being the selected as the best equation for base model equation (16). For Basswood, equation (39) was chosen as the best fit of base model

Table 12 Regression coefficients for base equation model (14) by species using first-order autoregressive covariance structure

Species	Parameter	Estimate	Standard Error	Pr>|t|
Sugar maple				
Equation (31)	a_0	0.2808	0.1025	0.0076
	a_1	0.8241	0.0451	<0.0001
	b_1	0.9954	0.2753	0.0003
	b_2	-0.3632	0.0540	<0.0001
	b_3	3.3627	0.5690	<0.0001
	b_4	-1.4604	0.3066	<0.0001
	b_5	0.5948	0.0574	<0.0001
Ash				
Equation (31)	a_0	0.1152	0.0657	0.0834
	a_1	0.8875	0.0265	<0.0001
	b_1	1.4025	0.1767	<0.0001
	b_2	-0.5429	0.0345	<0.0001
	b_3	5.7288	0.3759	<0.0001
	b_4	-2.4830	0.2004	<0.0001
	b_5	0.7056	0.0384	<0.0001
Aspen				
Equation (31)	a_0	0.3178	0.0736	<0.0001
	a_1	0.8141	0.0313	<0.0001
	b_1	1.0111	0.1609	<0.0001
	b_2	-0.3962	0.0314	<0.0001
	b_3	4.3206	0.3359	<0.0001
	b_4	-1.8116	0.1801	<0.0001
	b_5	0.5391	0.0367	<0.0001
Basswood				
Equation (14)	a_0	-0.2709	0.2415	0.2662
	a_1	1.2088	0.1621	<0.0001
	a_2	-0.0297	0.0134	0.0300
	b_1	2.0700	0.2437	<0.0001
	b_2	-0.5707	0.0488	<0.0001
	b_3	5.3179	0.4998	<0.0001
	b_4	-2.5619	0.2711	<0.0001
	b_5	0.4568	0.0565	<0.0001

equation (16) because of having the lowest AIC value. All selected equations, using the identity matrix covariance structure, did not have and DIFFITS values over two. Regression coefficients and standard error estimates for fits of equations (34), (35), (38), and (39) using a first-order autoregressive covariance structure are included in Table 13.

Volume Determination

Smalian's formula was used to iterate the volumes for all sample trees included in this study. Volume was determined for each species using the acceptable forms of taper equation models selected in the previous section, as well as the acceptable taper equation models with the lower bole section volume determined using the accepted outside diameter bark stump equation. MAE and percent error are included for each species using each form of volume determination in Table 14.

For sugar maple, equation (34) and (20) produced the lowest MAE (2.5940 cu. ft.) while equation (31) produced the lowest percent error (14.16%). Equations (31) and (20) together had the lowest MAE (3.3490 cu. ft.) while equation (31) had the lowest percent error (11.30%) out of all equations for use with ash. For aspen, equations (31) and (20) together produced the lowest MAE (2.0257 cu. ft.) and percent error (9.00%). Equations (39) and (20) had the lowest MAE (2.8773 cu. ft.) while equation (14) had the lowest percent error (8.21%) for basswood.

Final Model Selection

Based on regression diagnostics and fits when determining volume, an equation, or set of equations from previously determined acceptable equations, were selected for

Table 13 Regression coefficients for base equation model (16) using first-order autoregressive covariance structure

Species	Parameter	Estimate	Standard Error	Pr>|t|
Sugar maple				
Equation (34)	a_0	1.6465	0.3032	<0.0001
	a_1	0.9242	0.0563	<0.0001
	a_2	-0.3263	0.0888	0.0002
	b_1	0.4529	0.0193	<0.0001
	b_2	0.2948	0.0052	<0.0001
Ash				
Equation (35)	a_0	0.3520	0.0700	<0.0001
	a_1	0.9364	0.0284	<0.0001
	b_1	1.2350	0.0386	<0.0001
	b_2	0.6350	0.0273	<0.0001
	b_3	0.5068	0.0682	<0.0001
	b_4	-0.7045	0.0451	<0.0001
Aspen				
Equation (38)	a_0	0.6843	0.1020	<0.0001
	a_1	0.7602	0.0436	<0.0001
	b_1	0.4321	0.0186	<0.0001
	b_2	0.2381	0.0057	<0.0001
	b_5	0.0699	0.0083	<0.0001
Basswood				
Equation (39)	a_0	1.0547	0.2437	<0.0001
	a_1	0.9694	0.0393	<0.0001
	a_2	-0.1795	0.0659	0.0084
	b_1	0.6968	0.0371	<0.0001
	b_2	0.4532	0.0283	<0.0001
	b_4	-0.3282	0.0515	<0.0001
	b_5	0.0684	0.0093	<0.0001

Table 14 Fit statistics for volume determination

Species	MAE (cu. ft.)	Percent error
Sugar maple		
Equation (4)	3.3502	16.02%
Equation (31)	2.8046	14.16%
Equation (34)	2.6596	14.92%
Equation (4, 20)	3.4315	16.49%
Equation (31, 20)	2.7622	14.30%
Equation (34, 20)	2.5940	15.27%
Ash		
Equation (4)	6.2301	30.00%
Equation (31)	3.6434	11.30%
Equation (35)	4.3017	12.04%
Equation (4, 20)	6.7599	23.01%
Equation (31, 20)	3.3490	11.55%
Equation (35, 20)	3.9856	11.71%
Aspen		
Equation (4)	3.3877	13.08%
Equation (31)	2.1792	9.37%
Equation (38)	4.0617	14.64%
Equation (4, 20)	3.6146	13.94%
Equation (31, 20)	2.0257	9.00%
Equation (38, 20)	3.7098	13.47%
Basswood		
Equation (4)	10.4494	25.92%
Equation (14)	3.2743	8.21%
Equation (39)	3.2279	9.02%
Equation (4, 20)	11.2058	27.86%
Equation (14, 20)	3.0579	8.06%
Equation (39, 20)	2.8773	8.49%

each species to predict outside bark diameters and volume within a standing tree. It is important to note that equations (14) and (31) were discarded from final selection due to their high VIF values even though regression diagnostic and fits for these equations showed very promising results. The high degree of multicollinearity attached to these equations makes them reflect well upon the original dataset. If a new dataset were to be introduced, these equations would not perform as well as this study indicates.

For sugar maple, equation (34) was selected as the final model to use. Since equation (31) was discarded, equation (34) contained the next lowest MAE and AIC values compared to equation (4). The R^2 value was also larger for equation (34) than for equation (4). Also, for volume determination, equation (34) performed well by itself, without the outside bark stump equation. MAE was low for this equation and percent error was very close to the lowest value for all equations. Two examples, comparing observed diameters to predicted diameters of sugar maple trees, using equation (34) are included in Figures 2 and 3. Figure 2 shows a tree with a DBH of 11.45 in. and a total height of 66 ft., which was one of the more accurately predicted trees, having a MAE of 0.40 cu. ft. when comparing observed to predicted volume. Figure 3 shows a tree with a DBH of 12.2 in. and a total height of 76 ft., which was one of the worst predicted trees in terms of diameter estimation, having a MAE of 5.55 cu. ft. when predicting volume.

Equation (35) performed the best for ash and was selected as the final model equation to use. With equation (31) being discarded, equation (35) had the lowest AIC and MAE values and highest R^2 value when compared to equation (4). The use of equation (20) for the stump diameter did aid equation (35) in determining volume, and therefore, was used. Figures 4 and 5 show observed diameters compared to predicted

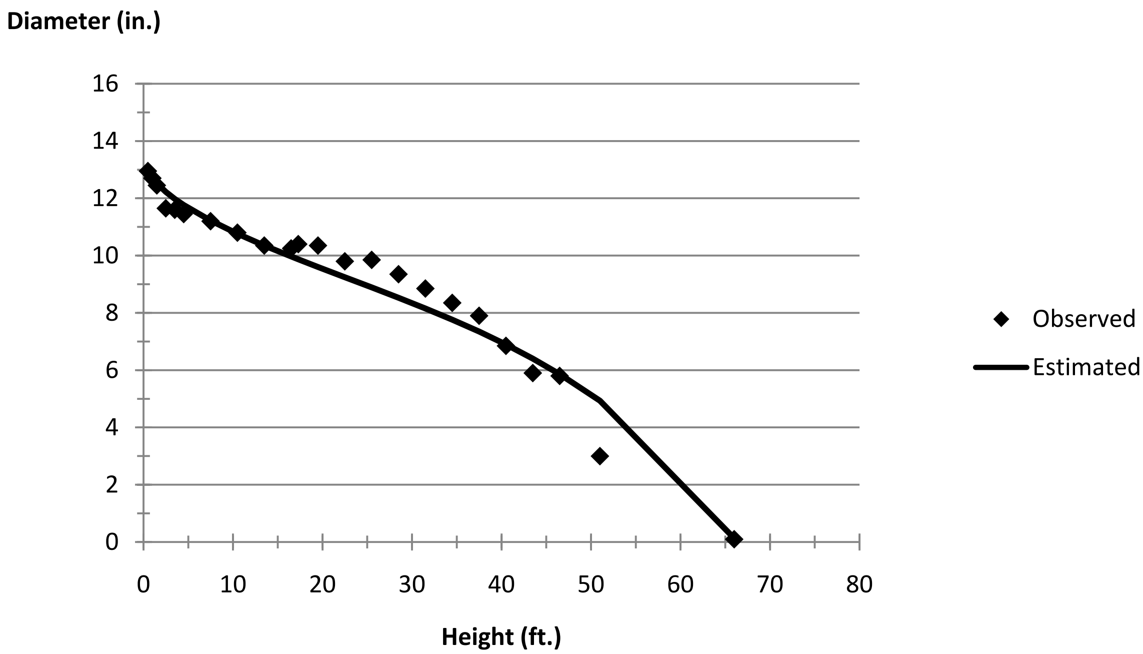


Figure 2 Observed versus estimated values of outside bark diameters from equation (34) for a sugar maple tree on site 1 with a DBH of 11.45 in. and a total height of 66 ft.

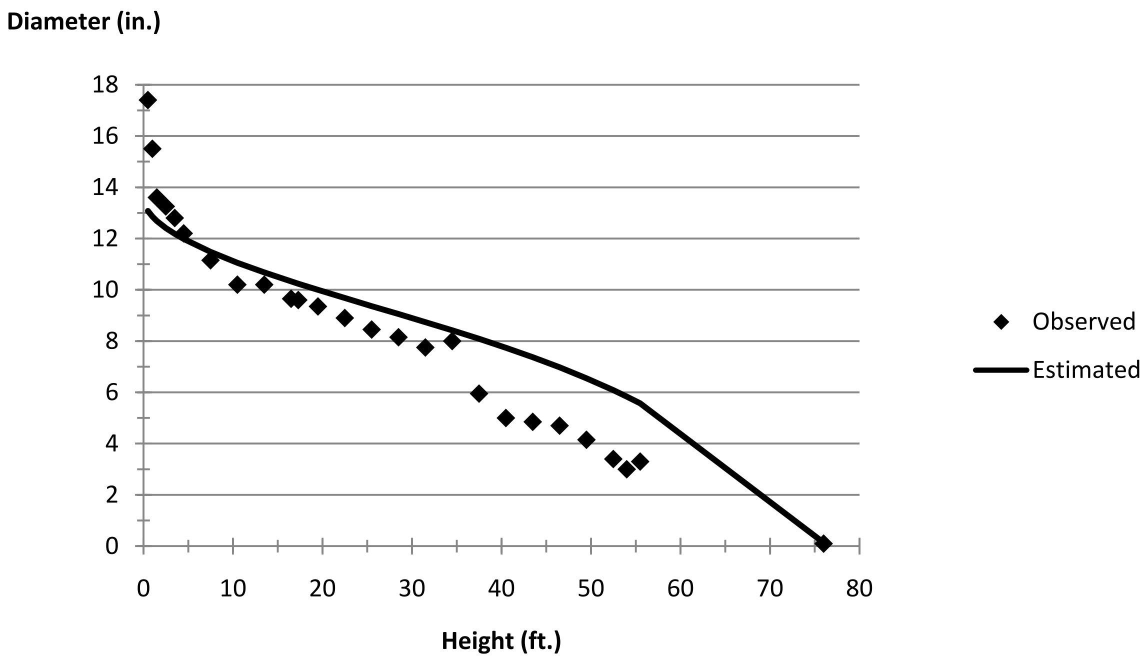


Figure 3 Observed versus estimated values of outside bark diameters from equation (34) for a sugar maple tree on site 4 with a DBH of 12.2 in. and a total height of 76 ft.

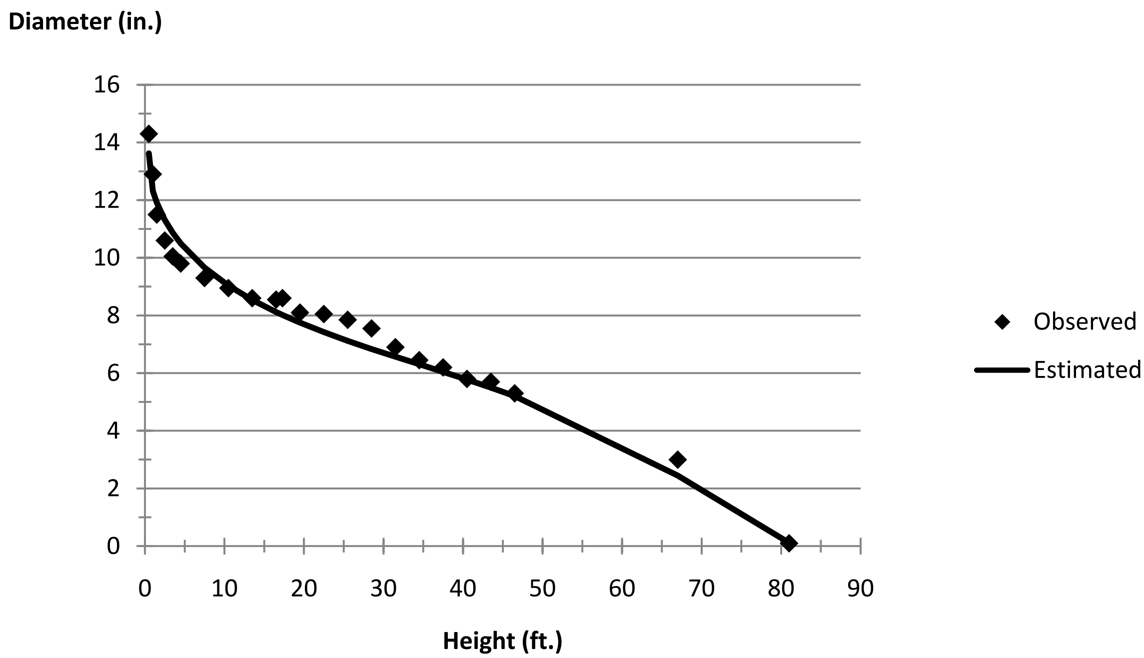


Figure 4 Observed versus estimated values of outside bark diameters from equations (35) and (20) for an ash tree on site 3 with a DBH of 9.8 in. and a total height of 81 ft.

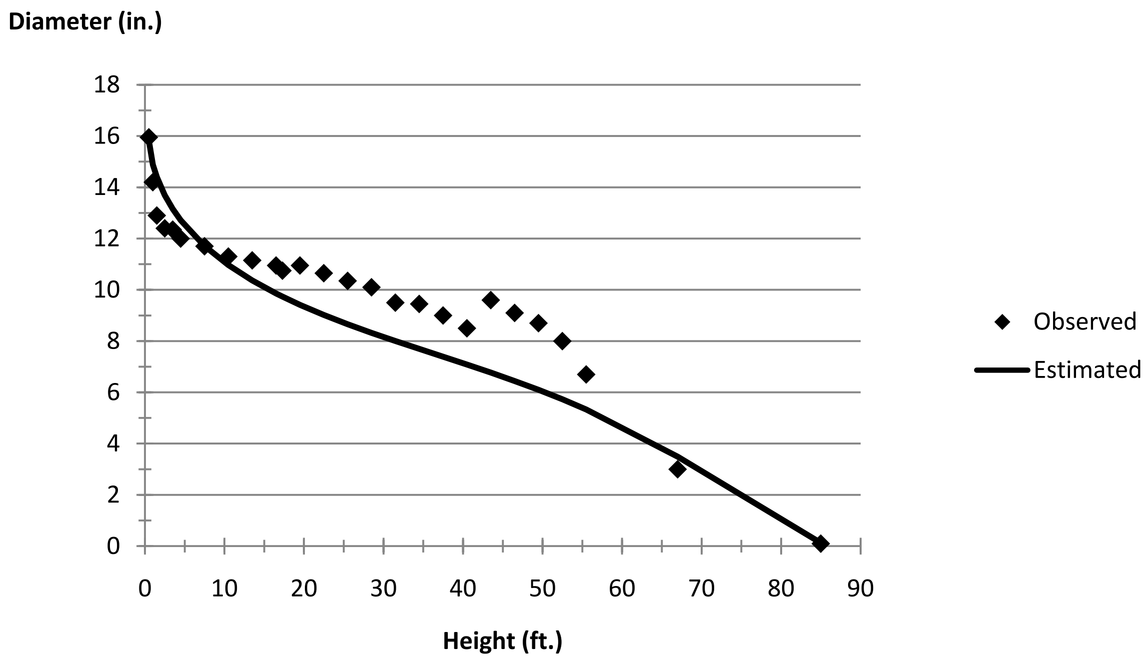


Figure 5 Observed versus estimated values of outside bark diameters from equations (35) and (20) for an ash tree on site 3 with a DBH of 12 in. and a total height of 85 ft.

diameters for two different ash trees, using equation (35) and (20). An ash tree with a 9.8 in. DBH and a total height of 81 ft. (Figure 4) was predicted quite accurately having a MAE of 0.49 cu. ft. when comparing observed volume to predicted volume. Another ash tree with a DBH of 12 in. and a total height of 85 ft. (Figure 5) was one of the worst predicted trees in terms of volume determination, having a MAE of 6.52 cu. ft.

For aspen, equation (4) performed the best out of all equations. Overall, the AIC value was the lowest for this equation, and besides for equation (31), equation (4) had the next lowest MAE value and next highest R^2 value. When determining volume, the addition of equation (20) to equation (4) did aid in closer estimates to observed volumes. Therefore, the outside bark stump equation should be used when determining volume for aspen. Two examples, comparing observed diameters to predicted diameters of aspen trees, using equations (4) and (20) are included in Figures 6 and 7. Figure 6 shows a tree with a 9.6 in. DBH and a total height of 85 ft., which had accurately predicted diameters and a MAE of 0.47 cu. ft when predicting volume. Figure 7 shows a tree with a DBH of 10.75 in. and a total height of 70 ft., which was not predicted as accurately, having a MAE of 9.67 cu. ft. when determining volume.

Equation (39) performed the best overall for basswood. AIC was the highest for this equation, but it was not far from the lowest value obtained by equation (4). Besides for equation (14), equation (39) produced the next lowest MAE value and next highest R^2 value. Together, equations (39) and (20) had the lowest MAE value when determining volume out of any equation. Equation (20) did help increase the accuracy of volume determination and should be used. Figures 8 and 9 show observed diameters compared to predicted diameters for two different basswood trees, using equations (39) and (20). A

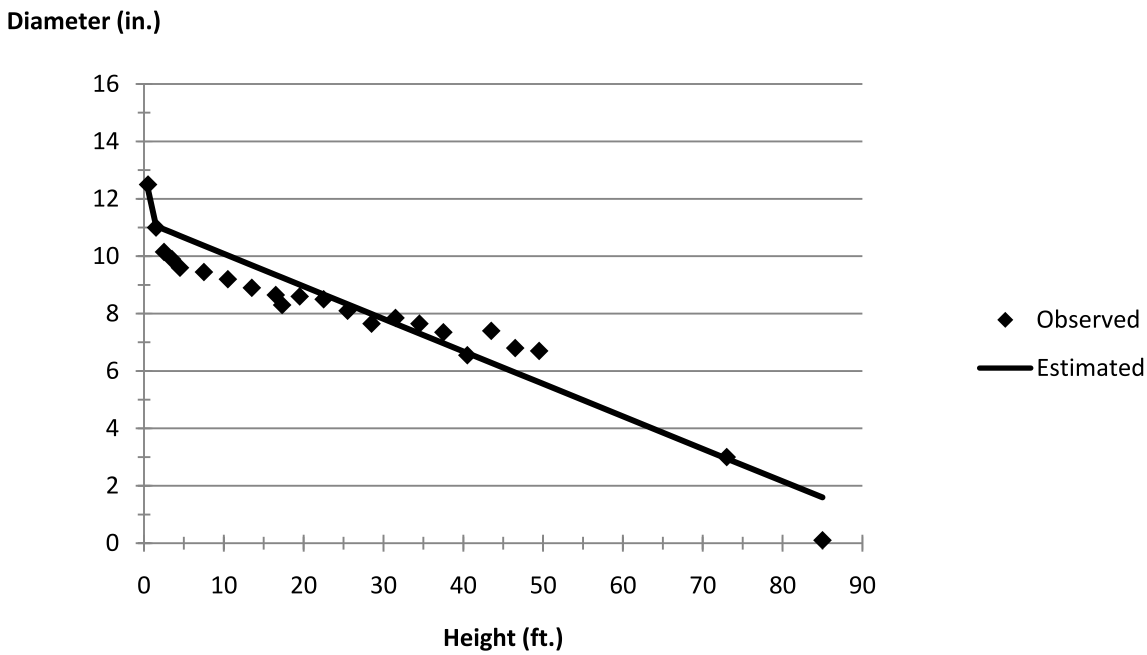


Figure 6 Observed versus estimated values of outside bark diameters from equations (4) and (20) for an aspen tree on site 2 with a DBH of 9.6 in. and a total height of 85 ft.

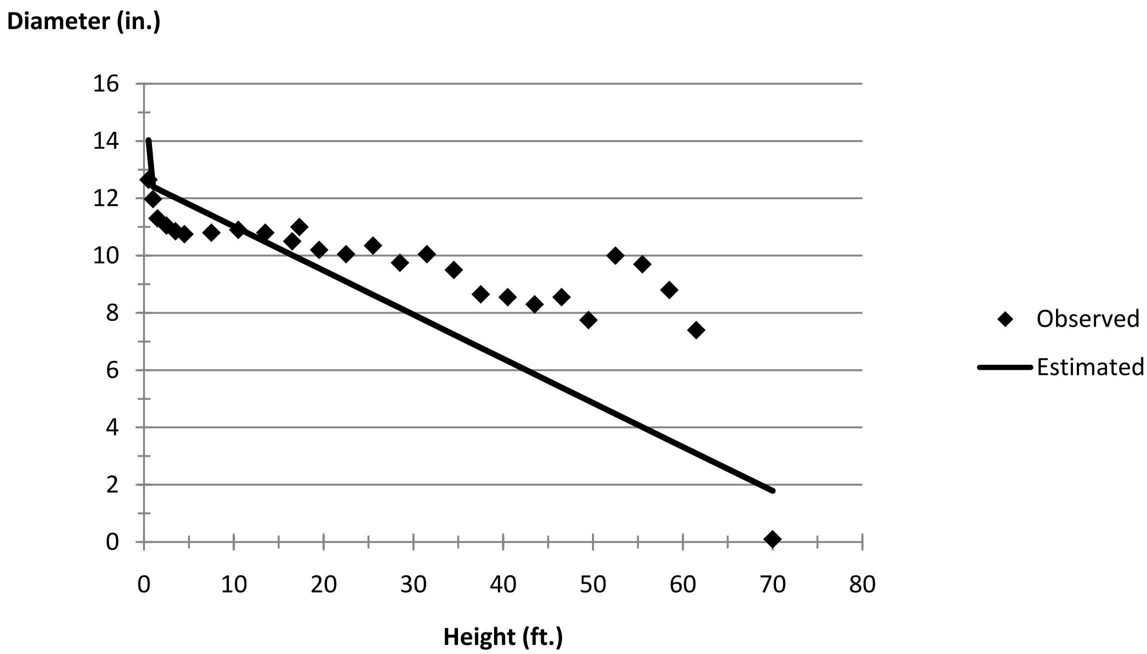


Figure 7 Observed versus estimated values of outside bark diameters from equations (4) and (20) for an aspen tree on site 1 with a DBH of 10.75 in. and a total height of 70 ft.

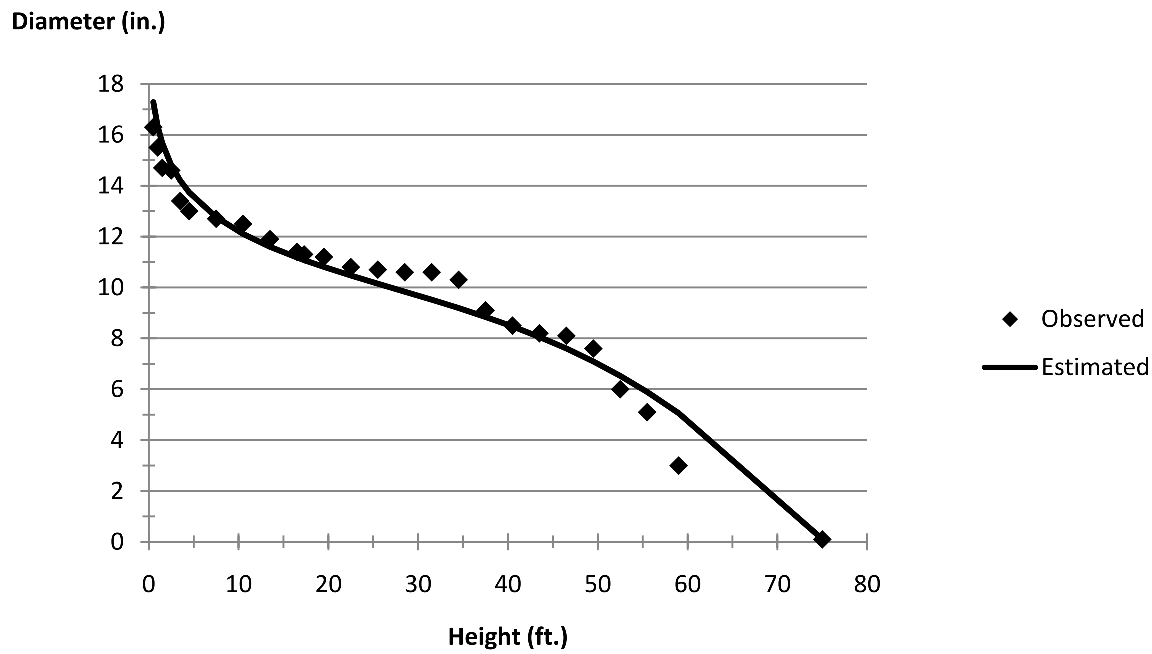


Figure 8 Observed versus estimated values of outside bark diameters from equations (39) and (20) for a basswood tree on site 6 with a DBH of 13 in. and a total height of 75 ft.

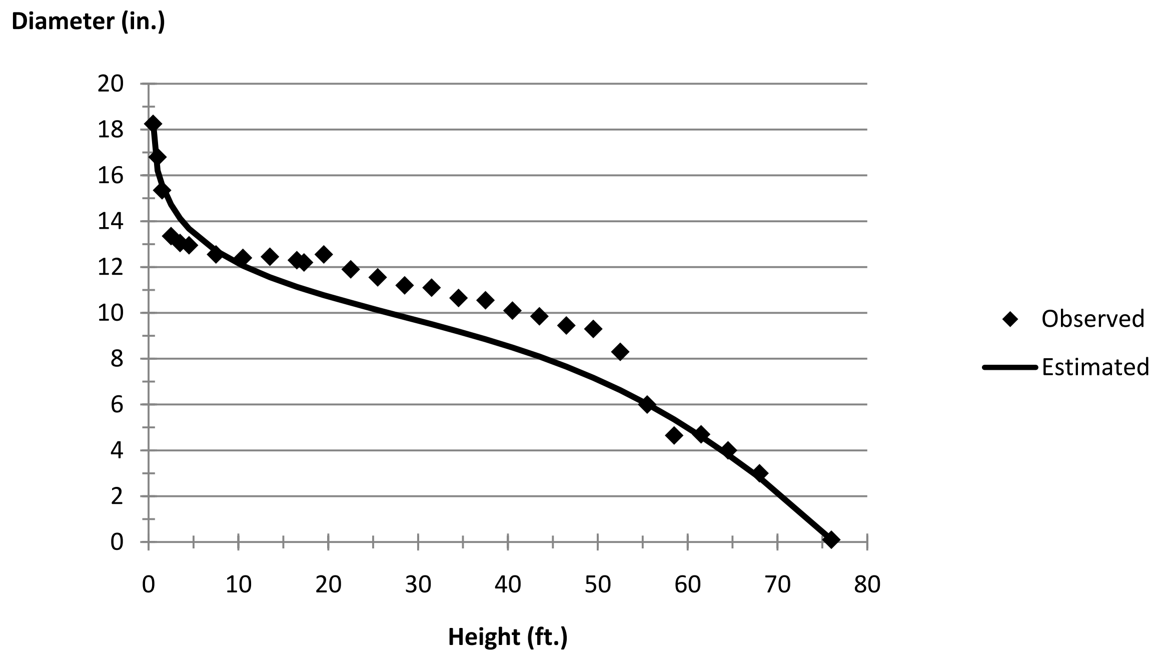


Figure 9 Observed versus estimated values of outside bark diameters from equations (39) and (20) for a basswood tree on site 3 with a DBH of 12.95 in. and a total height of 76 ft.

basswood tree with a DBH of 13 in. and a total height of 75 ft. (Figure 8) had accurately predicted diameters and a MAE of 0.01 cu. ft. when predicting volume. A basswood tree with a 12.95 in. DBH and a total height of 76 ft. (Figure 9) was not predicted as accurately and had a MAE of 5.70 cu. ft. when predicting volume.

Discussion

Several issues prompted by the nature of this study deserve additional discussion. This section explicitly addresses those points of discussion.

Volume Comparisons

To validate the research conducted in this study, and to check for accuracy when determining stem cubic foot volumes, two-sided paired t-test (α=0.05) were used to compare paired volume estimates. Volumes estimated from application of the selected profile equations (by species) were compared to volumes obtained through applying Smalian's formula to the Criterion RD 1000 data directly (observed volumes). The same comparison was made for volumes estimated with Gevorkiantz and Olsen's (1955) Table 3: Composite table: gross peeled volume in cubic feet, entire stem, by total height. All comparisons were made within three tree size class categories: pulpwood (D < 11.6 in.), small sawtimber (11.6 in. $\leq$ D $\leq$ 15.5 in.) and large sawtimber (D > 15.5 in.). It is important to note that the volumes predicted from the selected equations should more closely resemble the observed volumes because both values came from the same dataset. Also, Gevorkiantz and Olsen's Table 3 is gross peeled volume, meaning that bark is not included in the overall volume. Because of this, predictions of volume should be underestimated when compared to observed volumes, as all volumes obtained with the dataset included in this study are from outside bark measures.

In the sugar maple analysis, equation (34) significantly overpredicted stem volumes (p-value 0.005) with an average difference of 0.6 cu.ft. (5% error) whereas the Gevorkiantz and Olsen estimates significantly underestimated volumes (p-value <

0.0001) by an average difference of 2 cu.ft (15% error) for the pulpwood size class. Both methods significantly underpredicted stem volumes in the small sawtimber category (p-values of 0.0121 and 0.0335, respectively), each resulting in about 13% error. In the large sawtimber category, neither method resulted in significant differences in volumes estimates (p-values of 0.1484 and 0.3737, respectively).

When analyzing ash, neither the combined use of equations (35) and (20) or the Geovorkiantz and Olsen estimates resulted in significant differences when estimating volume in the pulpwood size class (p-values of 0.7682 and 0.0550, respectively). Both methods significantly underpredicted stem volumes in the small sawtimber class (p-values of <0.0001 and 0.0308, respectively), with equations (35) and (20) together having an average difference of 3.8 cu. ft. (10% error), while Gevorkiantz and Olsen estimates had an average difference of 2.2 cu. ft. (6% error). Only predicted volumes from equations (35) and (20) were found to be significantly different when compared to observed volumes of large sawtimber (p-values of 0.0006 and 0.4414, respectively). These estimates of stem volume were underestimated with about 13% error and an average difference of 9.2 cu. ft.

For basswood, equations (39) and (20) together were not found to be significantly different in the pulpwood, small sawtimber, or large sawtimber size classes (p-values of 0.1517, 0.6504, and 0.0524 respectively). Gevorkiantz and Olsen estimates of stem volume led to significantly underestimated volumes for all size classes (p-values of <0.0001, <0.0001, and 0.0067, respectively). These estimates had an average difference of 3.8 cu. ft. (19% error), 4.5 cu. ft. (12% error), and 6.1 cu. ft (9% error), respectively.

In the aspen analysis, equations (4) and (20) significantly overpredicted stem volumes (p-value 0.0007) with an average difference of 1.1 cu.ft. (7% error) whereas the Gevorkiantz and Olsen estimates significantly underestimated volumes (p-value < 0.0001) by an average difference of 2.4 cu.ft (15% error) for the pulpwood size class. The same significant trend was noticed for the small sawtimber size class where equations (4) and (20) overpredicted volumes (p-value of <0.0001) with about 18% error (6.2 cu. ft.), while Gevorkiantz and Olsen estimates were underestimated (p-value of 0.0013) with about 9% error (3.2 cu. ft.). In the large sawtimber size class, only equations (4) and (20) together were found to be significantly different (p-values of 0.0165 and 0.8939, respectively) with an overestimate of volume by 13 cu. ft. (25% error).

Based on these comparisons, the equations developed herein outperformed the Gevorkiantz and Olsen (1955) Table 3 in some instances. In other cases they did not, indicating the need for refinement of the profile equations. Nonetheless, an advantage of using the profile equations developed herein is that they can be used to estimate diameter at any point along the bole, which in turn can be used to estimate volume to any top diameter. This allows flexibility with respect to changing merchantability limits/standards.

Fitting other Models

Fits of segmented polynomial taper models to the dataset were not extensively used in this study. As stated by Martin (1981) for his work on Appalachian hardwoods, variable exponent models were not as precise as segmented polynomial models, but were just as accurate and simpler to use. A segmented polynomial model was attempted with the regression dataset, but led to poor results. Insignificant parameters were detected when using both two and three inflection points along the bole of a tree. With insignificant parameters in a segmented system, the model becomes invalid since each parameter is tied to a segmented section of the tree. Based on insignificant fits to these data, segmented polynomial taper models were dismissed from further consideration.

Compatible volume equations and nonlinear seemingly unrelated regression were not used in this study because of the lack of compatibility. It was anticipated within this study to include compatible taper and volume equations, but because a compatible volume equation does not inherently imply the best fit equation for taper, compatibility was excluded. This study aimed at finding the best taper equation possible from the chosen models.

Nonlinear fixed effect modeling was not used in this study because of the exclusion of random parameters. The benefits of nonlinear fixed effect modeling are that flexibility and efficiency in estimation for regional conditions are increased when random parameters are included to deal with the variability between sites. This study encompassed a narrow range, insufficient in size to carry a large amount of variability. There was no warranting cause to include this type of modeling.

Data Issues

For both the outside bark stump equations and the outside bark profile equations, residual error terms of the regression fits were found to be not normally distributed. This means that the residual error terms of dependent variables do not assume a bell shaped curve across the levels of dependent variables. Not having normal distribution can result in problems with homogeneity of variance and linearity (Quinn and Keough 2002). Since distribution was not positively skewed and several equations were already in a transformed state, further transformations of dependent variables would not alleviate the problem. Data show that residual error terms for all equations in this study did assume a bell shaped curve; however it was highly peaked in the middle, leading to a non normal distribution. A bell shaped curve, though peaked, is still marginally acceptable; skewed data would be problematic.

Autocorrelation posed a problem for all outside bark profile equations, as seen by low Durbin-Watson test statistics. Since repeated measures were taken on each sample tree, all dependent variables are not truly dependent from each other. Each paired height/diameter measure taken or predicted was highly influenced by the previous measure. In this case, errors are correlated and the errors within the covariance matrix are no longer ideal (Myers 1990). Myers (1990) goes on to state that autocorrelation can result in an estimate of variance that is underestimated, which can inflate *t*-statistics on coefficients. This would increase the type I error, making parameter estimates appear to be significantly different from zero when they truly are not. Changing the covariance structure to a first-order autoregressive covariance structure, AIC values were drastically lowered. This structure works because constant variance is assumed across the diagonal of the matrix and dependent covariates are increasingly less correlated with increasing distance. Just as diameter estimates at DBH are less influential to diameter estimates at the top of a tree compared to a diameter estimate ten feet below the top, this covariance structure works well when dealing with autocorrelation.

Multicollinearity was also a problem which arose from regression analysis.

Multicollinearity appears when dependent variables are highly correlated, having linear dependencies (Myers 1990). When this is the situation, coefficient estimates are tied to the dataset in which they came from. Small changes to dataset can drastically affect the coefficient estimate. This would decrease the applicability of any model produced with a high degree of multicollinearity. For this reason, equations (14) and (31) were discarded. VIF values for those equations were extremely higher than those for other equations used.

Project Strengths and Weaknesses

This study provides Wisconsin's forestry community with more accurate and applicable taper equations for estimating outside bark diameter for four species of trees within the region. These equations are needed in Wisconsin because of the limiting factors associated with current taper/volume estimation methods used during forest inventories. The results of this project should help to increase the amount of taper equation scientific literature available for application in Wisconsin.

Throughout the course of this study, several possible weaknesses of the project occurred. First, several research assistants were used to collect data from standing trees using the Criterion RD 1000. In total, six people were used, none of which were previously familiar with the new forestry tool. Field trials with the equipment were conducted prior to engaging in data collection, but user error may have attributed to some inconsistent data, coincidently impacting analysis.

Data for the project were collected throughout the entire year of 2008. Some data were collected in March and April, before leaf-out, while the rest was collected in early to late summer. Because of the nature of using the Criterion RD 1000, measurements of upper bole diameters were sometimes difficult to take with dense canopy vegetation. Errors associated with upper stem diameter estimation when leaves were present may have impacted analysis. It is the recommendation of the author to acquire standing tree measurements after leaf senescence in the fall and before leaf-out in the spring, though problems with seasonal variation and access may be a hindrance.

Other tree and stand attributes may have been helpful in acquiring more accurate taper equations for the species used in this study. Live crown ratio, stand density, soil conditions, etc. could have been measured to reduce variability and pinpoint differences between sample trees. However, few studies of taper on hardwood trees have included these extra stand and tree attributes.

Stands included in this study were in a fairly narrow geographical range. Any conclusions deduced from this project should only be applied to datasets of similar origins to the dataset used for this study. The region specific nature of taper equations infers that these conclusions may be of little use outside the region of Wisconsin.

Lastly, this study is a small part of a larger project. A part of the larger project includes comparing modeled equations from this study to equations modeled from felled trees of the same species. Through this study, it is known that taper equations can be developed from standing trees using the Criterion RD 1000, but it is not yet known how accurate they are to felled tree measures taken with calipers. When completed, that work should shed further light on these results.

Summary and Conclusion

A large component of Wisconsin's forests is comprised of the four tree species examined in this study. With current trends in the Lake States, higher degrees of utilization are more common, requiring more precise and accurate estimates of merchantable volume. This focus towards greater utilization increases the need for equations that can accurately predict stem profiles (and inherently produce merchantable volumes) better than composite volume tables from the 1950's and approximation formulas. Once these equations have been developed and/or tested in the field, Wisconsin's forestry community can use them during forest inventories including examined species. In this study, three previously published model forms to estimate tree taper were applied to a dataset including aspen, sugar maple, ash, and basswood collected in north central Wisconsin to check for predictive ability of paired height/diameters and volume estimation.

Six study sites were included in this study; two sites at Treehaven, owned and operated by the University of Wisconsin-Stevens Point College of Natural Resources, near Tomahawk, WI, one site provided by, and located in, Lincoln County, one site provided by Kretz Lumber Company Inc. near Crandon, WI, one site provided by Plum Creek Timber Company Inc. near Spirit, WI, and the final site at the The Highgrounds Veterans Memorial Park, near Nellsville, WI. Data were collected for all species included in the study at all sites when applicable. Analysis was applied to 312 total trees that were sampled throughout 2008.

Paired height/diameter measurements were taken on all sample trees included within the study, which were then used as the composition of the regression dataset.

Paired height/diameter measurements were taken at 0.5, 1.0, 1.5, 2.5, 3.5, 4.5 feet in the lower bole section and every three feet beyond that until the top of the tree was reached. Tree level attributes including DBH, total height, height to a ten inch diameter top, height to a three inch diameter top, and Girard form class were also collected and individually averaged for each tree species.

An equation was developed to estimate outside bark stump diameter at 0.5 feet for each species, and a second set of equations were developed to predict outside bark diameters along the entire tree bole for each species. All equations were developed using the entire dataset of sample trees from the six sites used in this study. Sample trees for each species encompassed a large degree of variability in an attempt to increase the applicability of developed equations.

Two model forms for the outside bark stump equation were compared using regression diagnostics to find the best predictor of diameter at 0.5 feet for each species. Independent variables of D and D^2H were used in model building. The recommended equation to use had R^2 values of 0.7841, 0.9480, 0.9440, 0.8646 and MAE values of 1.1586, 0.8960, 0.7224, 0.15270 inches for sugar maple, ash, aspen and basswood, respectively.

Three base model forms for outside bark profile equations were compared using regression diagnostics to find the most accurate predictor of paired diameter/height estimates. Since parameter estimates for some equations were not significantly different from zero, several versions of the original model forms were developed to test for regression diagnostics. With repeated measures being taken on each sample tree, positive autocorrelation became evident in all equations used. To address this situation, the covariance error structure for each equation was changed to first-order autoregressive covariance structure, and AIC values for each equation became dramatically lower. Multicollinearity also became a problem in model evaluation because of linear dependencies shown by high VIF values. Certain equations were discarded based solely on their high degree of multicollinearity. Acceptable forms of equations for each species were then further scrutinized after predicted volumes were estimated and compared to observed volumes.

Volumes, in cubic feet, of sample trees were determined through iteration using Smalian's formula. Each sectional increment from paired height/diameter measure provided a cross sectional area at large and small ends of various bole sections. These were then summated to estimate total cubic foot volume in each standing tree. This method was used to estimate volume for all predicted values of acceptable profile equations. Also, volumes for trees were estimated by adding the selected outside bark stump equation, for lower bole volume estimates, to the volumes of upper bole estimates from the acceptable taper equations. All forms of volume determination were compared using MAE and percent error.

Final models were selected for each species based on regression diagnostics from fitting models to the dataset and accuracy when determining volume. For sugar maple, the selected equation had an approximate R^2 value of 0.883, a diameter estimation MAE value of 0.8522 inches, an AIC value of -2052, a volume estimation MAE of 2.6596 cubic feet, and a percent error of 14.92%. For ash, the final selected equation had an approximate R^2 value of 0.934, a diameter estimation MAE value of 0.9154 inches, an AIC value of -4439.3, a volume estimation MAE of 4.3017 cubic feet, and a percent error of 12.04%. For aspen, the selected final equation had as approximate R^2 value of 0.897, a diameter estimation MAE value of 1.2283 inches, an AIC value of -3569.1, a volume estimation MAE of 2.3254 cubic feet, and a percent error of 9.91%. For basswood, the final selected equation had an approximate R^2 value of 0.942, a diameter estimation MAE value of 0.785 inches, an AIC value of -2320.8, a volume estimation MAE of 3.2279 cubic feet, and a percent error of 9.02%.

The equations developed for each species in this thesis should help forestry professionals and landowners in determining diameters at merchantable limits and volumes of standing trees. It is hoped that this work will be used in Wisconsin when applicable.

Literature Cited

Avery and Burkhart. 2002. Forest Measurements 5th ed. New York: McGraw Hill. 456 p.

Behre, C.E. 1923. Preliminary notes on the study of tree form. J. of For. 21:507-511.

Behre, C.E. 1924. Is taper based on form quotient independent of species and size? J. of For. 22:282-290.

Bruce, D. 1972. Some transformations of the Behre equation of tree form. For. Sci. 18(2):164-166.

Bickerstaff, A. 1946. Knot-free red pine by debudding. Can., Dept. Mines and resources, For. Serv. Silv. Res. Note 76. 10pp.

Bruce, D.R., R.O. Curtis, and C. Vancoevering. 1968. Development of a system of taper and volume tables for red alder. For. Sci. 14:339-350.

Burger, H. 1931. Einfluss der Herkunft des Samens auf die Eigenshaften forstlicher Holzgewächse. III. Die Föhre. Mitt. Schweiz. Centralanstalt forstl. Versuchsw. 16:153-230.

Burk, T.E. and A.R. Ek. 1999. Approximation, evaluation, and extension of the Lake States composite taper and volume tables. North. J. Appl. For. 16(2):65-71.

Burkhart, H.E. and S.B. Walton. 1985. Incorporating crown ratio into taper equations for loblolly pine trees. For. Sci. 31(2):478-484.

Byrne, J.C., and D.D. Reed. 1986. Complex compatible taper and volume estimation systems for red and loblolly pine. For. Sci. 32(2):423-443.

Claughton-Wallin, H. and F. McVicker. 1920. The Jonson “Absolute Form Quotient” as an expression of taper. J. of For. 18:346-357.

Clutter, J.L. 1980. Development of taper functions from variable top merchantable volume equations. For. Sci. 26:117-120.

Davis, T.A.W. and G. W. Richards. 1934. The vegetation of Morabilti Creek. Brit. Guiana J. Ecol. 22:126-132.

Demaerschalk, J.P. 1972. Converting volume equations to compatible taper equations. For. Sci. 18:241-245.

Demaerschalk, J.P. 1973. Integrated systems for the estimation of tree taper and volume. Can. J. For. Res. 3(90):90-94.

Demaerschalk, J.P., and A. Kozak. 1977. The whole bole system: a conditioned dual-equation system for precise prediction of tree profiles. Can. J. For. Res. 7:488-497.

Ek, A.R. and T.E. Burk. 1986. Field notes: why number of sticks plus number of trees divided by two may equal cords per acre. North. J. Appl. For. 3:133-134.

Ek, A.R., T.D. Droessler, and M.T. Cheky. 1986. Taper equations for the Lake States composite volume tables and their application. Dept. of For. Resource., Univ. Minnesota, St. Paul, MN. Staff Pap. Ser. Rep. No. 57. 48p.

Fang, Z., B.E. Borders, and R.L. Bailey. 2000. Compatible volume-taper models for loblolly and slash pine based on a system with segmented-stem form factors. For. Sci. 46(1):1-12.

Flury, P. 1903. Einfluss verschiedener Dutchforstungsgrade auf Zuwachs und Form der Fichte und Buche. Mitt. Schweiz. Centralanstalt forstl. Versuchsw. 7:1-246.

Gallant, A.R and W.A. Fuller. 1973. Fitting segmented polynomial models whose join points have to be estimated. J. of Am. Stat. Assoc. 68 pp. 144-147.

Gevorkiantz, S.R. and N. W. Hosley. 1929. Form and development of white pine stands in relation to growing space. Harvard Univ. For. Bull. 13. 83 pp.

Gevorkiantz, S.R., and L.P. Olsen. 1955. Composite volume tables for timber and their application in the Lake States. Tech. Bull. No. 1104. U.S. Department of Agriculture, Forest Service, Lake States Forest Experiment Station. 51 p.

Gray, H.R. 1956. The form and taper of forest tree stems. Imp For Inst Pap 32, 79 p.

Grosenbaugh, L.R. 1966. Tree form: Definition, interpolation, extrapolation. For. Chron. 42:444-457.

Härdtl, H. 1938. Die formbildende Wirkung der Eigenlast von Stamm und Zweigen bei cinem Nadelbaum. Forstwiss. Centr. 60:343-352; 424-429.

Hartig, R. 1892. Ueber den Entwicklungsgang der Fichte im geschlossenen Bestande nach Höhe, Form und Inhalt. Forstl.-naturw. Z. 1:169-185.

Hilt, D.E. 1980. Taper-based system for estimating stem volumes of upland oaks. USDA For. Serv. Resp. Pap. NE-458. 11 p.

Höjer, A.G. Tallens och granens tillväxt. Bihang till Fr. Lovėn. Om våra barrskogar. Stockholm, 1903.

Husch, B.M., C.I. Miller, and T.W. Beers. 1972. Forest mensuration. New York: Ronald Press Co. 410 p.

Jonson, T. 1910. Taxatoriska undersökningar om skogsträdens form. I Granens stamform. Skogsvårdsföreningens Tidskrift. Häfte 11, pp. 285-328.

Jonson, T. 1911. Taxatoriska undersökningar om skogsträdens form. II. Tallens stamform. Skogsvårdsföreningens Tidskrift. Häfte 9-10, pp. 285-329.

Jordan, L., K. Berenhaut, R. Souter, and R.F. Daniels. 2005. Parsimonious and completely compatible taper, total, and merchantable volume models. For. Sci. 51(6):578-584.

Knight, R.C. 1961. Taper and secondary thickening in stems and roots of the apple. East Malling Res. Sta., Ann. Rept. 1960 65-71.

Kozak A., and R. Kozak. 2003. Does cross-validation provide additional information in the evaluation of regression models? Can. J. For. Res. 33:976-987.

Kozak, A. 1988. A variable exponent taper equation. Can. J. For. Res. 18:1363-1368.

Kozak, A. 1997. Effects of multicollinearity and autocorrelation on the variable exponent taper functions. Can. J. For. Res. 27:619-629.

Kozak, A., D.D. Munro, and J.H.G. Smith. 1969. Taper functions and their application in forest inventory. For. Chron. 45:1-6.

Laitakari, E. 1929. Über die Fähigkeit der Bäume sich gegen Sturmgefahr zu schützen. Acta Forest. Fennica 34(34):1-29.

Larson, P.R. 1963. Stem form development of forest trees. For. Sci. Monogr. 5.

Lemieux, F.J. 1936. Log rules, taper tables, and volume tables for use in the south. J. of For. 34(11):970-974.

Liu, C.J. 1980. Log volume estimation with spline approximation. For. Sci. 26:361-369.

Martin, A.J. 1981. Taper and volume equations for selected Appalachian hardwood species. For. Serv. Res. Paper NE-490. 22 pages.

Max. T.A., and H.E. Burkhart. 1976. Segmented polynomial regression applied to taper equations. For. Sci. 22(3)283-289.

Mer, M.E. 1892. Sur les causes de variation de la densitė des bois. Bull. Soc. Bot. France 39:95-105.

Mesavage, C., and J.W. Girard. 1956. Tables for estimating boardfoot volume of timber. USDA For. Serv. 94 p. Washington, D.C.

Metzger, K. 1893. Der Wind als massgebender Faktor für das Wachsthum der Bäume. Mündener forstl. 3:35-86.

Metzger, K. 1896. Form und Wachstum der Waldbäume im Kichte der Darwinschen Lehre. Allgem. Forst- u. Jagdztg. 72:224-233.

Metzger, K. 1908. Über das Konstruktionsprinzip des sekundären Holzkörpers. Naturw. Z. Forst- u. Landwirtsch. 6:249-273.

Miles, Patrick D. Dec-10-2007. Forest inventory mapmaker web-application version 2.1. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station. [Available only on internet: www.ncrs2.fs.fed.us/4801/fiadb/index.htm]

Muhairwe, C.K., V. M. LeMay, and A. Kozak. 1994. Effects of adding tree, stand, and site variables to Kozak's variable-exponent taper equation. Can. J. For. Res. 24:252-259.

Myers, R.H. 1990. Classical and modern regression with applications. Massachusetts: PWS-KENT. 488 p.

National Volume Estimator Library. 2008. Fort Collins, Co: U.S. Department of Agriculture, Forest Service, Forest Management Service Center. Available at http://www.fs.fed.us/fmsc/measure/volume/nvel/index.php (verified 4-25-2008)

Newnham, R.M. 1988. A variable-form taper function. Canadian Forest Service, Pettawa Natl. For. Inst. Rep., PI-X-83.

Newnham, R.M. 1992. Variable form taper functions for 4 Alberta tree species. Can. J. For. Res. 24:252-259.

Perez, D.N., H.E. Burkhart, and C.T. Stiff. 1990. A variable from taper function for pinus oocarpa shiede in central Hondura. For. Sci. 36(1):186-191.

Quinn, G.P. and M.J. Keough. Experimental Design and Data Analysis for Biologists. New York: Cambridge University Press, 2002.

Reinsch, C.H. 1967. Smoothing by spline functions. Numerische Mathematik 10(3):177-

183.

Schmeid, H. 1918. Vergleichbare Schaftformdarstullungen. Centr. ges. Forstw. 44: 193-202.

Stoate, T.N. 1942. Taper in *Pinus pinaster*. Australian For. 6:59-66.

Trincado, G., and H.E. Burkhart. 2006. A generalized approach for modeling and localizing stem profile curves. For. Sci. 52(6):670-682.

Vorreiter, L. 1954. Stammform und Holzeigenschaften der Bayernwald-Fichte. Holz Roh- u. Werkstoff 12:47-54.

Wiant, Jr., H.V. and P.M. Charlton. 1984. Estimating volumes of upland hardwoods with the Behre hyperboloid. J. of For. 82(3):173-174.

Westfall, James A. 2004. Modifying Taper-Derived Merchantable Height Estimates to Account for Tree Characteristics In: Proceedings of the sixth annual forest inventory and analysis symposium; 2004 September 21-24; Denver, CO. Gen. Tech. Rep. WO-70. Washington, DC: U.S. Department of Agriculture Forest Service. 126p.

Wright, W.G. 1923. Investigation of taper as a factor in measurement of standing timber. J. of For. 21:569-581.

Zellner, A. 1962. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. J. Am. Stat. Assoc. 57:348-368.

Zhang, Y., B.E. Borders, and R.L. Bailey. 2002. Derivation, fitting, and implication of a compatible stem taper volume weight system for intensively managed, fast growing loblolly pine. For. Sci. 48(3)595-607.