Developing volume and taper equations

for

***Styrax tonkinensis* in Laos**

A thesis
submitted in partial fulfilment
of the requirements for the degree of

Master of Forestry Science

By

Khamsene Ounekham

University of Canterbury
New Zealand

2009

Contents

Abstract

A volume equation for predicting individual tree volume, and a taper function for describing a stem profile were developed for a little known species, *Styrax tonkinensis* (Siam benzoin) in northern Laos. The species has high potential commercial value and can make an important contribution to the local economy. It can provide two different types of products, a non-wood product (benzoin resin) and timber. In Laos, the most important product is currently resin, and the use of timber for commercial purposes is rare. One reason is that information about the timber is not available. In Vietnam, on the other hand, the species is an import pulpwood species.

Data used in this study came from 73 trees. Trees were purposely selected to ensure coverage of a full range of tree sizes. Measurement was undertaken only on over-bark diameters due to some constraints, limitations and problems during the field data collection. However, due to the importance of under-bark volume for this species, a small available dataset was used to build a bark model as an interim guide to the errors associated with using over-bark models for estimating under-bark volumes. From this bark model, errors in estimating under-bark volumes of trees with diameters at breast height between 10cm and 17 cm were approximately 18%.

Nineteen individual volume models, and 7 individual taper functions were compared for bias and precision. Collective names for the volume equations tested include single-entry, double-entry, logarithmic, combined variables. Most volume models had similar bias but a few were clearly biased. The models with similar bias were further evaluated by four common statistics including bias, standard error of estimates, standard deviation of residuals and mean absolute deviation. The results showed that a five parameter model was ranked first, and was the most precise model. However, the magnitudes of difference in prediction errors between this model and other models, particularly the three parameter model were not significant. For practical purposes, the simpler model was preferred.

Seven taper functions tested here belong to three different groups including single taper equations, compatible taper equations and segmented taper equations. Evaluation of taper equations used the same residual analysis procedures and criteria as those applied with volume equations. Graphical residual analysis showed that most taper models had similar precision with their errors in diameter predictions being similar in range. However, some models showed obvious bias. The most highly ranked taper model was a compatible taper model of polynomial form. It was the least biased model. The second ranked model was a single, simple model. This latter model is relatively simple to apply, but it is not compatible with the volume model, yielding slightly different estimates of volume if it is integrated and rotated around the longitudinal axis of a tree. However, if the sole purpose is to describe tree taper, it is the best model to use.

List of Tables

List of Figures

1. Introduction

1.1 Background

Tree volume and taper equations are useful and important for forestry, and they are lacking for commercial species in developing nations. They are simple methods and tools that can be used to obtain individual tree volume and the volumes of entire stands. Such information is vital for forest management.

Volume equations have been used to estimate tree and stand volume, and have played a crucial role in forest inventories and management for more than a hundred years. Studies of tree volume began in the early nineteenth century. Around 1804 Heinrich Cotta was the first forester to introduce the concept of a volume table (Clark 1902). However, an extensive study to collect data for constructing the first volume table was carried out many years later. This early study was mainly of Norway spruce.

Taper functions were introduced and used much later. The advantage of taper functions over volume equations is that taper functions can describe changes in diameter up a tree stem, and therefore provide estimates of dimensions of logs that might be cut from stems. Volumes of any specific log length can be obtained by integrating a taper function. In many cases, this is more convenient than volume ratio equations that are usually limited to a particular height such as commercial height.

Even though both volume equations and taper functions have been studied for many years, they continue to attract forest research. One reason is that there is no single theory in volume and taper equations that can be used satisfactorily for all tree species (Clutter *et al.* 1983; Muhairwe 1999), and no single taper model is best for all purposes (McClure and Czaplewski 1986; Coa *et al.* 1980). Another reason is that both volume and taper equations are required to be increasingly accurate and flexible in their predictions. Forest measurement needs to be improved because market requirements for timber have become more specific in recent years.

Volumes of current growing stock and future growth potential are both vital information for forest management. The former can be obtained through forest inventories and the latter can be estimated or projected from a current inventory by growth and yield models (Methol 2001). Individual tree volumes are primary data for estimating stand volume per hectare (or stand volume for a fixed area), and are directly linked to forest inventory. Effective tools for estimating individual tree volume are volume equations and taper functions. They are acceptably accurate, easy and cheap methods (Philip 1994). While most volume equations developed can limitedly provide total and/or merchantable stem volume, wider ranges of information can be obtained from taper functions. According to Methol (2001) taper functions can be used to estimate the following tree variables:

- diameter (either under- or over-bark) at any point of the stem;
- height at which a given diameter occurs a long the stem;
- total volume (either under- or over-bark);
- merchantable volume (either under- or over-bark) to any merchantable height or minimum upper-stem diameter and from any stump height; and
- individual log volumes.

Volume and taper functions have been scarcely studied in Laos. Elsewhere, a considerable amount of work on volume and taper has been done (Clutter *et al.* 1983; Kozak *et al*. 1969; Max and Burkhart 1976; Newnham 1988). No volume or taper equation for any species in Laos has been published.

Tree species that have been the focus for past studies, particularly for taper equations, are softwoods (Max and Burkhart 1976; Cao *et al*. 1980; Fang *et al*. 2000). Relatively small numbers of taper equations have been developed for hardwoods. Examples of such equations are the study of Appalachian hardwoods conducted in locations in the United State of America (Jiang *et al*. 2005).

Building volume and taper equations warrants study, particularly when the methodology is applied to species which have not been previously studied in this way.

Styrax tonkinensis is a subject species for the study described here. The species is an angiosperm, and belongs to Styracaceae family. More details about it are left for the literature review section (section 2 of this paper). Scientific knowledge about the species is very limited, particularly in Laos, even though *S. tonkinensis* has important social value and potentially high commercial value. The main economic value of the species currently arises from its gum. Timber products are not used commercially in Laos at the moment. However, it has high potential value for industrial uses. The the timber is good for pulp and has been used for pulping elsewhere. Vietnam has commercially planted this species for many years (Pinyopusarerk 1994; Williamson 1989).

1.2. Study objectives

The overall objective of this study was to find reliable indirect methods for tree volume estimation for *S. tonkinensis,* a species that has a potential commercial value, and is important to minority groups in the mountainous areas in northern Laos. The specific objectives of study were to:

- develop taper and volume models for *S. tonkinensis* that can explicitly state the relationship between tree volume and dbh, and also tree volume and dbh and height, by fitting regression equations to sample tree data;
- compare the performances of different forms of volume equations in predicting volumes of *S. tonkinensis;*
- compare the performances of different forms of taper functions in describing tree profiles of *S. tonkinensis*; and
- select the most suitable compatible volume and taper equations from amongst those tested.

1.3 Notations

The following notation will be used hereafter. Other definitions specific to a particular equation will be listed with the equation:

D	is diameter at breast height (1.3 m above ground) over bark (cm)
H	is total tree height (m)
d	is stem diameter over bark at height ‘h’ (cm)
h	is height up the stem from ground (m)
h_D	is tree height at breast height equivalent to 1.3 m for this study (other study may differ height i.e. 1.37 m)
V	is estimated total stem volume over bark from stump (m^3). Stump height for this study is 0.15m.
F	is tree form factor
ln	is natural logarithms
$\beta_{i\ (i=1,2,3....)}$	is coefficient
$Bias$	is mean residual
SEE	is the standard error of estimate
SDR	is the standard deviation of residual
MAD	is the mean absolute deviation
G	is tree basal area
K	is a constant number $\dfrac{\pi}{40{,}000}$

2 Literature review

2.1 Tree profile and log volumes

Modelling individual stem volume is a critical foundation of forest mensuration, growth and yield estimation, and forest valuation. Individual stem volume equations are used to translate forest inventory measurements of height and diameter at breast height into wood volume and to generate predicted volumes from modelled estimates of future heights and diameters at breast height. In addition, taper equations, models of profiles of stems, are critical for estimating the sizes and shapes of logs that might be obtained from a tree, and these estimates are required to compute log value. It is essential that volume and taper models are as unbiased and accurate as possible.

Trees have various sizes, forms and shapes. A single stem also consists of different geometric segments. Some portions of a stem may be cylindrical, while others may be conoid or other geometric solids. Some parts of a stem may be also affected by irregular forms such as butt swell. Stem volume, therefore, is difficult to estimate accurately. However, foresters usually treat tree stems as common geometric solids in with volumes that can calculated using the ordinary existing volume formulae. Three common solids of revolution applied to the tree stems are neiloids, conoids and paraboloids (Avery and Burkhart 1994). Volumes of these solids of revolution can be calculated using formulae as follows:

$$\text{paraboloid} = \frac{A}{2}L$$

$$\text{conoid} = \frac{A}{3}L$$

$$\text{neiloid} = \frac{A}{4}L,$$

where A is cross-sectional area, and L is log length.

The lower bole portion is generally assumed to be a neiloid frustum, the middle portion a paraboloid frustum, and the upper portion a cone (Hush *et al.* 1972).

According to Avery and Burkhart (1994), the principal problem encountered when computing log volumes is that of accurately determining the elusive average cross-sectional area, because volumes for all solids of revolution are computed from the product of their average cross section and length. Two commonly used formulae (Huber's formula and Smalian's formula) define average cross-sectional area in different ways. Huber's formula treats cross-sectional area at the midpoint as the average, and thus:

$$\mathrm{V}_{\mathrm{log}} = A_{1/2}L$$

where $\mathrm{V}_{\mathrm{log}}$ is log volume, and L is as previously defined.

Smalian's formula, on the other hand, uses cross-sectional areas of both ends of the log. Therefore, the average cross section is the mean of two end cross-sectional areas, and thus:

$$\mathrm{V}_{\mathrm{log}} = \frac{A_1 + A_2}{2}L$$

where A_1 and A_2 are cross-sectional areas at large end and at small end, respectively, and other notations are as previously defined.

One other log volume formula occasionally used for log volume estimation is Newton's formula. The average cross-sectional area used in this formula is slightly more complicated than those used in the previous two formulae. This formula is more difficult to apply in practice. It requires three measurements, one at both ends plus the midpoint of the log and thus:

$$\mathrm{V}_{\mathrm{log}} = \frac{A_1 + A_{1/2} + A_2}{6}L$$

where all notations are as previously defined.

All three formulae can provide identical results if logs are perfectly cylindrical (Avery and Burkhart 1994). In addition, when logs are short it is also found that the volumes estimated by three formulae are essentially equivalent.

Of the three formulae, Smanlian's formula is the easiest to apply, but it can be shown to be inaccurate with some irregular shapes of logs. Avery and Burkhart pointed out that the formula introduced errors when it was used to estimate volumes of butt logs having flared ends. Some problems are shared by both Smalian's formula and Huber's formula. If the log is not a frustum of a quadratic paraboloid and not a cylinder, then the use of either Smanlian's formula or Huber's formula will introduce errors (Philip 1994). However, given some problems encountered by the other two formulae, Smalian's formula is usually preferred by researchers. The use of midpoint cross sectional area by Huber's formula and Newton's formula can cause a problem in practice. The midpoints of logs in piles or ricks are often inaccessible and cannot be measured (Avery and Burkhart 1994). Even though Newton's formula is more accurate than the other two methods, its use is limited and it is the least favoured formula among the three.

2.2 Volume tables and equations

A volume table is a tabulation that can be used to obtain the estimated volumes of single trees of given dimensions (Avery and Burkhart 1994). In modern practice, equations are generally used to predict tree volumes rather than hardcopy tables. An early volume equation was introduced during 1930's by Schamacher *et al.* (Laar and Akca 2007). The form of equation is:

$$V = \beta_0 D^{\beta 1} H^{\beta 2}$$

where V is stem volume, D is diameter at breast height, H is total tree height, and $\beta_{0\text{-}2}$ are estimated parameters.

This equation is a non-linear volume equation that can be linearized by a logarithmic transformation of the dependent and independent variables. The resultant equation is:

$$\ln(V) = \beta_0 + \beta_1\ln(D) + \beta_2\ln(H)$$

Nowadays there are many different forms of volume equations of both linear and non-linear form. One of the most common forms is Spurr's volume equation for a linear combined variable model (Bi and Hamilton 1998). This equation has a form:

$$V = \beta_0 + \beta D^2 H$$

One merely substitutes the combined variable of 'diameter squared times height' for the quantity X in the basic equation for a straight-line relationship. Solution of the equation is by simple linear regression techniques. Regression methods are favoured over other traditional methods like tabular and graphic methods that have become obsolete for several reasons (Laar and Akca 2007). One is that it eliminates the necessity to read off the estimated volume from a graph or to interpolate in a table. More importantly, the parameters of the equation can be stored in the memory of a computer and retrieved for volume calculation anytime.

In most cases estimates and profiles of under-bark volume are required, but there are cases where over-bark volume is essential, such as when trees are used for fuel.

2.3 Classification of volume tables

Even though nowadays stem volumes are commonly calculated with volume equations, the term 'volume table' has persisted in forestry usage as a generic term meaning tabulations or equations that show the contents of standing trees (Laasasenaho, *et al*. 2005, and Avery and Burkhart 1994). According to (Laar and

Akca 2007) volume tables (equations) can be classified based on the number of entries to the table and predictor variables of the volume function:

- Single-entry volume table (one-way table)
- Multiple-entry volume table (two-way table and three-way table)

A single-entry volume table was first developed towards the end of 19th century for all-aged forests in France and adapted for management of mixed uneven-aged forests of Switzerland (Laar and Akca 2007). The term ”local table” is sometimes used to refer to this kind of volume table. Normally diameter at breast height (D) or basal area (G) is required for constructing a single-entry volume table. The relationship between tree volume and D for many species has been well documented. Generally, Tree volumes have a curvilinear relationship with D but are approximately linearly related to D squared (Avery and Burkhart 1994). Therefore, the volume-basal area line is actually a simple, linear relationship of volume on basal area and the equation can be expressed:

$$V = \beta_0 + \beta_1(D^2) \quad \text{or} \quad V = \beta_0 + \beta_1(G^2)$$

Most single-entry volumes are simple and easy to apply, but their uses are limited to local conditions. Tree dimensions that are more difficult to measure such as height or form are usually not required. Thus, single-entry tables are particularly useful for quick forest inventories and are low cost in use (Philip 1994). Elimination of height and form determinations also tends to assure greater uniformity in volume estimates, particularly when two or more field parties are cruising within the same project area (Avery and Burkhart 1994).

Because trees of a given diameter class, particularly those from different stands, can vary in their heights and forms, use of single-entry volume tables can introduce bias. Thus most volume tables of this type have to be restricted to a small range of diameters in a specific stand at a specific age (Philip 1994). It is usually necessary to construct single-entry tables for each broad site class encountered when soils and topography are notably varied.

Multiple-entry volume tables include double-entry and triple-entry volume tables. The former is sometimes referred to as ‘standard volume tables’. The standard volume tables use both ‘D’ and ‘H’ as table entries, while triple-entry volume tables have a third variable of tree’s dimensions such as form, diameter at a particular height or taper. Some papers reported that the addition of a third predictor variable reduced the amount of unexplained variation and improved the accuracy of volume estimates, while other studies found that the addition of a third predictor variable did not significantly improve the quality of prediction (Laar and Akca 2007). It is clear, however, that triple-entry volume tables are not as widely available as double-entry volume tables.

Double-entry volume tables are probably the most common form of volume table (Philip 1994). A large number of volume equations, which are linear in their parameters, have been proposed with D, D^2, H, H^2 and interaction terms as independent variables, and individual tree volume as the dependent variable. The volume equations in Table 2.1 are taken from a list of candidate volume equations obtained from the available literature that will be tested with the data set created for this study. A full list of them is in the method section. All but equations 8, and 9 are linear in their parameters.

Table 2.1: Examples of double-entry volume equations

Equation no.	models	References
5	$V = b_1D^2H$	Clutter *et al*. (1983)
6	$V = b_0 + b_1D^2H$	
7	$V = b_0 + b_1D^2 + b_2H + b_3D^2H$	
8	$V = b_1D^{b2}H^{b3}$	
9	$V = b_0 + b_1D^{b2}H^{b3}$	
10	$V = D^2/(b_0 + b_1H^{-1})$	
11	$V = b_0 + b_1D^2H + b_2D^3H$	Bi and Hamilton (1998)
12	$V = b_0 + b_1D^2H + b_2D^3H + b_3D$	
13a	$V = b_0 + b_1D^2H + b_2D^2H^2$	
13b	$V = b_1D^2H + b_2D^2H^2$	
14	$V = b_0 + b_1D^2H + b_2D^2H^2 + b_3H$	
15	$V = b_0 + b_1D^2H + b_2D^3H + b_3D^2H^2$	
16	$V = b_0 + b_1D^2H + b_2D^3H + b_3D^2H^2 + b_4D$	
17	$V = b_0 + b_1D^2H + b_2D^3H + b_3D^2H^2 + b_4H$	
18	$V = b_0 + b_1D^2H + b_2D^3H + b_3D^2H^2 + b_4D + b_5H$	

The equations included in Table 1.1 consist of all three categories of double-entry volume equations defined by Philip (1994):

- simple combined variable model
- multiple regressions with powers of D and H
- logarithmic forms

These three categories of volume equations have been used frequently. Each category has some advantages and limitations. One of the advantages of a logarithmic form is that it can directly handle heterogeneity of the variance, while ordinary least squares techniques cannot.

2.4 Linear and non-linear volume equations

Performances of non-linear or logarithmic volume equations are not affected by non-homogeneity of the variance. Using this form of volume equation can surmount the problem of non-homogeneity of the variance that exists in most tree volume data. Because volume data usually include different sizes of trees from very small ones to very large ones, deviations from the regression function of the volumes of the large trees have a disproportionate effect on the estimation of the least squares regression coefficients from a sample (Cunia 1964). In other words, one of the common assumptions underlying least squares methods of regression, that all data points contribute equally to the estimation of the regression coefficients, is not satisfied. Even though heterogeneity of variance does not necessarily introduce bias, it may increase model statistics such as the standard errors of regression coefficients and inply that estimates for small trees are less precise than they actually are. In addition, when the residual error increases with the size of prediction, estimates for small trees may be biased because measurements of small trees would have less influence on estimated coefficients than those of large trees. This may be one reason that unweighted least squares techniques are fully efficient only in the absence of heteroscedasticity, a term denoting a correlation between average error maginitude and the magnitude of the predicted value of a model (Furnival 1961).

2.5 Weighted least squares

Cunia (1964) argued that using the logarithms of tree volume equations was not the best option to surmount the problem of heterogeneity of variance. One important drawback of this method is that by taking logarithms the estimation of the arithmetic mean is automatically replaced by the estimation of the geometric mean. Because the first one is always larger than the second, the results are definitely biased. An alternative common technique to combat the non-homogeneity of the variance in tree volume construction is to use weighted least squares. A common weight factor for volume that uses both variable predictors D and H (interaction term of D^2H) is $\left(\frac{1}{D^2H}\right)^2$. That is because the variance of stem volume tends to increase in proportion to D^2H as reported by Furnival (1961), Wright (1964), and Meng and Tsai (1986). For single-entry volume tables, on the other hand, Meng and Tsai (1986) suggested that the weight factor of $\frac{1}{D^2}$ is more appropriate than $\frac{1}{D^4}$ which was used by some authors.

2.6 Taper

Taper can be defined as the rate of narrowing in diameter along the tree stem of a given form (Gray 1956). It can be expressed as a function of height above ground level, total tree height, and diameter at breast height (Clutter *et al.* 1983). Taper equations are very useful as they can provide information about diameter at any height, and height at any diameter based only on commonly taken tree measurements (Byrne and Reed 1986). They can be also used to predict individual log or sectional volumes to any height along stem. Therefore, merchantable heights and volumes can be estimated directly using a taper equation. Furthermore, taper equations can be used to derive volume equations by integration when the equation is rotated around the longitudinal axis of a tree (Bruce *et al.* 1968), Byrne and Reed (1986). They can be also compatible to volume equations that are used to estimate single tree volumes. A compatible taper equation developed from volume-taper equation system assures that estimated volume obtained by integrating the taper equation is equal to the estimate obtained with a volume function.

2.7 Classification of taper equations

During the past century, form and taper have been studied widely all over the world. Many different forms of taper equations have been developed for various species, particularly for softwoods. A simple form of taper depicts the entire stem profile with a single equation. A complex taper model usually consists of sub-models that attempts to describe different portions of the tree profile with different sub-models, and uses complicated variable predictors.

According to Methol (2001) taper equations can be grouped into four categories; namely, single functions, segmented polynomial models, within-tree variable from (or variable exponent equation), and between-tree variable form functions. Single functions include the form of polynomial equations that represent the whole bole with one single continuous function. The taper models by Bruce *et al.* (1968), Ormerod (1973), Hilt (1980), and Gordon *et al.* (1995) are examples of this kind. Segmented polynomial taper models consist of sequence of grafted sub-models describing different segments. A taper model by Max and Burkhart (1976) is an example of this kind. The other two types of taper equations, within-tree variable form (or variable exponent equation), and between-tree variable form functions have not widely used. This study will not examine any equation of these latter two types.

2.8 Single functions

An equation presented by Kozak *et al*. (1969) is an example of the relatively simple parabolic function with three estimated parameters. This equation is one of the candidates model tested with data set in this study (model 19).

(19) $$d^2 = D^2\left(\beta_1\left(\frac{h}{H}-1\right)+\beta_2\left(\frac{h^2}{H^2}-1\right)\right)$$

This taper equation was developed based on the basic relationship of a parabolic function. It was conditioned by $\left(\frac{h}{H}-1\right)$ and $\left(\frac{h^2}{H^2}-1\right)$ as predictor variables so that when h equals H, estimated diameter is exactly zero. The related model without imposing such a condition has the form:

$$d^2 = D^2\left(\beta_0+\beta_1\frac{h}{H}+\beta_2\frac{h^2}{H^2}\right)$$

By imposing the condition, bias generated by the model can be reduced because unexplained variation at the top of a tree is restricted.

Other candidate taper equations of single form will be examined in the studies described here, including models 20 by Sharma and Oderwald (2001), 21 by Ormerod (1973), and 22 (polynomial series model):

(20) $$d^2 = D^2\left(\frac{h}{h_D}\right)^{2-\beta}\left(\frac{H-h}{H-h_D}\right)$$

$$d = D\left(\frac{H-h}{H-h_D}\right)^{\beta}, \beta > 0 \qquad (21)$$

$$d = D\left(\beta_0 + \beta_1\left(\frac{h}{H}\right) + \beta_2\left(\frac{h}{H}\right)^2 + \ldots \beta_n\left(\frac{h}{H}\right)^n\right) \qquad (22)$$

Equations 20, and 21 are so conditioned that when h = H, d= 0, and when h = h_D (h_D is breast height equal to 1.3 m), d = D. For equation 21, when β takes on a value of one, the resulting tree profile is conic and when β is one-half the resulting tree profile is parabolic (Reed and Byrne 1985). When β is greater than one but less than one-half i.e. three-fourths, a tree shape is also between a cone and a parabola 'paracone'.

Another common form of single taper functions is a polynomial series. Equation 22 represents the general form of this kind. Even though very high degree polynomials have been used in some studies the most common ones are around fifth-degree polynomial (Figueired-Filho *et al.* 1996). Volume equations with additional complex powers i.e. an equation by Bruce *et al.* (1968) can improve the standard error of estimate (Kozak *et al.* 1969). However, Kozak *et al.* (1969) argued that, for practical purposes it appeared that the real advantage of complex taper equations was little.

Relatively simple taper equations can effectively describe the general taper of trees, however, they often fail to describe the entire stem profile well (Max and Burkhart 1976; Newnhan 1992). Some equations are better for describing the profile along the mid stem portion of the tree, but they are inadequate for describing the area near the butt and at the very top sections of the tree (Jiang *et al.* 2005). Martin (1981) indicated that although no single equation form was best at predicting diameter, height and volume the Max and Burkhart segmented polynomial was best overall. This model was also ranked best amongst six models evaluated by Cao *et al.* (1980).

Taper equations are compatible with tree volume equations if they are derived from each other. The coefficients of the derived volume equations can be written in terms of the taper equation coefficients (Byrne and Reed 1986). Compatible volume and taper equations can also written by derived the expression of taper from an existing total volume equation (Goulding and Murray 1976), and from an existing volume ratio equation (Clutter *et al.* 1983; Reed and Green 1984). The theory of compatible taper and volume systems was first introduced in the early 1970s (Demaerschalk 1972). Before his introduction of the concept it was common that tree volume and taper equations were both in use for a given population, and that volumes obtained from the tree volume equation were not equal to volumes obtained by integration of the taper equation (Methol 2001). The study will examine a compatible taper and volume system that has the form of polynomials (equation 23). The process used for deriving compatible taper equations is described in the methods section.

(23) $$d^2 = \frac{V_{ml}(\beta_1 z_1 + \beta_2 z_2 + \beta_3 z_3 + \beta_4 z_4 + \beta_5 z_5)}{kH}$$

where $z_n = (n+1)\left(\frac{H-h}{h}\right)^n$

k is a constant number $\frac{\pi}{40{,}000}$

V_{ml} is the volume of a tree estimated by the volume model

other notations are as previously defined

2.9 Segmented taper models

Segmented taper models were first introduced by Max and Burkhart (1976). This type of taper model may provide a better description of the stem profile than that provided by single taper models, especially in the high-volume butt region (Cao *et al.* 1980). As previously described, it is generally assumed that a tree stem can be divided into three geometric shapes (form the top to bottom: a cone shape, a frustum of a paraboliod and a frustum of a neiloid). Segmented models describe these shapes by fitting each one with a different equation and then mathematically joining them to produce an overall segmented function (Diéguez-Aranda *et al*. 2006). Equation 24 was the classic segmented model proposed by Max and Burkhart (1976). It is one of the popular taper models of this kind.

(24) $$d^2 = D^2\left(\beta_1\left(\frac{h}{H}-1\right)+\beta_2\left(\frac{h^2}{H^2}-1\right)+\beta_3\left(a_1-\frac{h}{H}\right)^2 I_1+\beta_4\left(a_2-\frac{h}{H}\right)^2 I_2\right)$$

$I_1 = 1$ if $\frac{h}{H} \le a_1$ and $I_1 = 0$ otherwise

$I_2 = 1$ if $\frac{h}{H} \le a_2$ and $I_2 = 0$ otherwise

While taper and volume equations are very common in developed countries, they are less common in developing countries. The study described here involved developing a volume and taper equations for a native species that has high potential value in timber for commercial uses in Laos, *Styrax tonkinensis*.

2.10 *Styrax tonkinensis*

2.10.1 General information

Styrax tonkinensis belongs to the genus *Styrax* and the family Styracaceae. It has a narrow range of natural distribution, naturally growing only some parts of Southeast Asia. The countries that reportedly have natural growth forests of this species include Vietnam, Thailand and Laos. In Laos, the species grows naturally in four provinces in the mountainous northern areas of the country; Luang Prabang, Phong Saly, Houaphan and Oudomxay. Relatively little is know about the species. Closely related species that are in the same genus, produce similar gum, and are better known include *S. paralleloneurum*, *S. benzoin* and *S. hypoglauca* in Indonesia, Malaysia and China.

Some studies classified *S. tonkinensis* as an important non-timber forest product (NTFP) species, because of current use of its gum. The gum extracted from *S. tonkinensis* is known as Lao or Siam benzoin. Laos is currently the only country that exports significant products of Lao benzoin. The exact amounts of gum exported each year cannot be accurately estimated, but it is believed to be many hundreds of tonnes.

S. tonkinensis is an angiosperm with a soft wood that is used only by local people for household purposes such as buildings and firewood. So far, there have been no reports of trade in timber products in Laos, and therefore, it has no recognized commercial value. In Vietnam, however, timber has been used for pulp and commercial plantations have been established (Pinyopusarerk 1994). It is one of major plantation species in Vietnam. Difficult access to *S. tonkinensis* forests precludes any type of log transportation, and this may be a main limitation on use of timber. The species grows naturally in very remote mountainous terrains where there are neither roads nor other means of transportation.

2.10.2 Previous research

In Laos, studies of *S. tonkinensis* are rare, and only undertaken in recent years. The most comprehensive study project so far was an experimental trial at two sites in Luang Prabang Province. The results of this trial were documented by Kashio and Johnson (2001). This paper also presented the results from other research conducted 2 years prior to the trial, on the relationship between tree size and benzoin production, and the methods of harvesting gum. According to the authors of this paper, much research and study on *S. tonkinensis* is required, particularly in topics that could promote successful plantation establishment.

Other studies published include the study by Pinyopusarerk (1994), about the distribution, ecology and silvicultural management; Takeda and Shinya (2004) about traditional trapping of the benzoin gum by local people in Luang Prabang Province, and the regeneration of the species after fallow cultivation; and Satoshi (2004) about the importance of *S. tonkinensis* for mountainous people in Laos.

In Vietnam, however, there have been more studies of the species over a long period. The research and study papers cited in Kashio and Johnson (2001) include Lam Cong Dinh (1964), Le Quang Dang (1966), Hoang Chuong (1974), Doan Van Nhung *et al* (1978), Nguyen Ba Chat (1979) and Anon (1983). Vietnam is the only country known to have large areas of plantations of *S. tonkinensis*. Over 20,000 ha have been planted.

The following descriptions of characteristics of *S. tonkinensis* are drawn or summarized from the report by Kashio and Johnson (2001).

2.10.3 Species characteristics and ecology

S. tonkinensis has not been studied widely. Information about it is scare or inaccessible. As mentioned above, most studies about the species were carried out in Vietnam. Access to those reports is very difficult. However, a few papers are available i.e. Kashio and Johnson (2001); Williamson (1989); Jøker (2000). According to these papers some important information about the species can be summarised as follows:

> *"S. tonkinensis* is a deciduous tree up to 25 m tall and 30 cm in diameter with a clear bole for about two-thirds of total tree height. The bark is generally gray, smooth, and 6-9 mm thick when young, but becomes brown and rough with longitudinal fissures with age. *S. tonkinensis* is a fast growing species that can reach sexual maturity at 4-5 years of age. Under management, trees can grow very fast. Thai Van Trung (1975) reported that a mean height of 18-25 m and DBH of 20-24 cm could be obtained at 10 years, and height increments of 3 m during the first three years have been observed on good sites.", from Kashio and Johnson (2001).

S. tonkinensis is a pioneer species. Within its range it regenerates well in gaps. In Laos, the species is naturally found at 800-1,600 m elevation in the northern part of the country. Clearing forest for temporary agriculture such as shifting cultivation can promote regeneration of the species. It can sometimes occur in almost pure stands over many hectares. Burning is also believed to promote seed germination. *S. tonkinensis* trees produce a lot of seeds throughout their lives. Each year, a mature tree can produce up to 40 kg of fruit, and 2-3 kg of fruit contains 8,000-9,000 seeds.

Even though *S. tonkinensis* can grow well naturally through gaps created by forest disturbance, i.e. shifting cultivation plots, tree growth can be significantly improved by silvicultural practices. Important techniques and procedures applied to the species were described by Kashio and Johnson (2001). Seeds are best collected from trees when they have matured but before they fall to the ground. Storage techniques are required to keep seeds at an appropriate moisture content which is 30 %. Planting can be done directly from seeds or transplanted from seedlings previously grown in a nursery. High stock planting was recommended as *S. tonkinensis* is a pioneer species. Two or three thinnings may be required before the plantation reaches the minimum rotation age of approximate 10 years. *S. tonkinensis* is regarded as a 'fast-growing species' has a high productive capacity and no major health problems. However, according to Williamson (1989) one negative feature of the species is a high wood to pulp ratio (about 8.6 m^3 per tonne). Another undesirable feature is that the species is susceptible to defoliation from insects at a young age.

The uses of *S. tonkinensis* were summarised by Jøker (2000):

> "Because the wood is light and soft with a density of 410-450 kg/m3 (at 15% moisture), it is not thought to be suitable for construction. In Vietnam it is an important pulpwood species and yield and quality of the pulp is comparable with many commercial pulpwood species. An important non-wood product is the benzoin resin that is tapped from the trunk. Although the market has decreased, it is still an important contribution to the local economy for people in the highlands of Laos."

Estimates of current uses of species in Vietnam are difficult to access from outside Vietnam. Currently *S. tonkinensis* timber is not used commercially in Laos. However, a study of production from one major paper mill in Vietnam by Williamson (1989) found that the mill used 112,000 tonnes of *S. tonkinensis* annually. Paper production from that mill accounted for 30% of Vietnam's total production at that time. Given that the species has been a major part of the reafforestation programme in Vietnam, it is expected that the current use should be much higher than 20 years ago.

Because the wood of *S. tonkinensis* has high potential commercial value in Laos, and the non-wood product from this species (benzoin resin) is an important contribution to the local economy for people in the highlands of Laos, it is worth promoting plantations of the species for commercial purposes.

The species was chosen for this study of volume and taper equations primarily because of its high potential commercial value and because it has been neglected by researchers in Laos. Details and specific objectives of the study were already expressed in the objective section (section 1.2).

3 Study area and methodology

3.1 Study area

3.1.1 Geographical

Data used in this study were collected from five locations. All five locations are in Nam Bak district, Luang Prabang province (Figure 3.1). Nam Bak town is located in a mountainous area, about 130 km north of the provincial capital Luang Prabang. The study area is part of the Nam Ou (Ou River) watershed. The Nam Ou is the largest tributary of Mekong River. The Mekong River is one of the Asia's longest river running through many countries including China, Laos, Thailand, Cambodia and Vietnam. The study sites range in altitude from 400 – 900 meters above sea level.

3.1.2 Climate, soil and vegetation

The climate in the study area is a wet and dry monsoon tropical climate. The mean annual rainfall is about 1,400 millimeters. The annual weather pattern is characterized by two main seasons. The rainy season is between the months of May and October and account for 90% of annual precipitation. The variation in annual rainfall from year to year is wide ranging from as low as 1,000 millimeters up to above 2,000 millimeters.

The mean temperature is about 25°C. the coolest period is between the months of December and January, while the hottest period is between May and June.

Soil data from the study area is insufficient. Limited study has been carried out. It was presumed that the prevalent soil texture was clay-loam with a higher portion of loam in the upper soil layers.

Most of the forests in Nam Bak district are secondary forests in different stages of succession. They are influenced by the practice of shifting cultivation. Abandoned cultivation plots are usually occupied by pioneer species in the first few years. Then, vegetation can develop towards a forest cover if the plots are left for a longer period (Kashio and Dennis 2001).

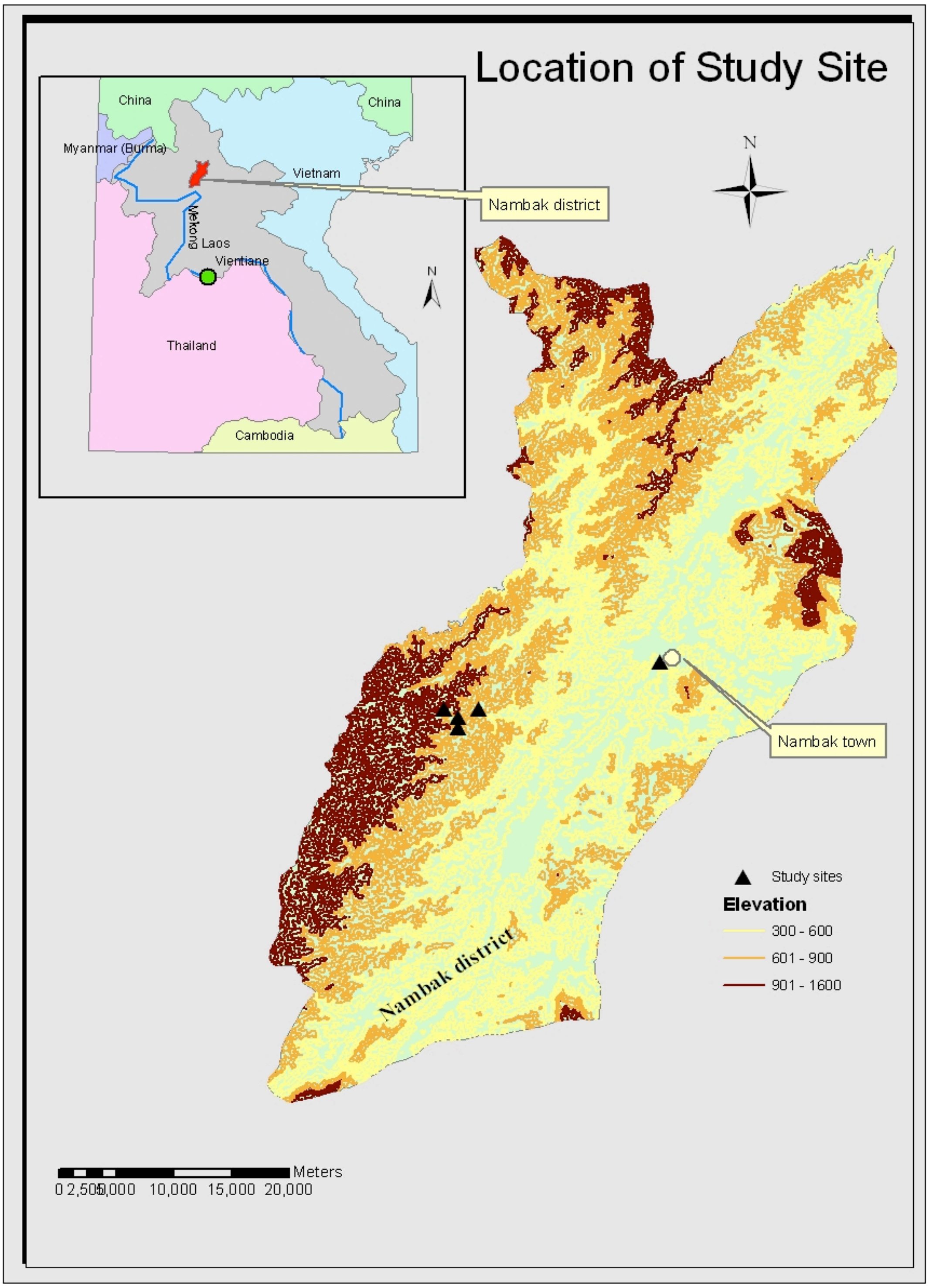


Figure 3.1: Map showing the location of study sites

3.2 Data collection

Sample data were collected from 73 trees in five different locations. Three sites were natural forests and two were plantations. All sites were located in relatively close proximity within the same district (see map 1). Thirty-five trees were from two plantations (one site had 10 and the other had 25). Thirty-eight were from three stands of natural forests (two sites had 12 trees each and the third one had 14 trees). The trees were subjectively selected to ensure a representative distribution by diameter and height classes.

Before felling the trees, diameters at breast height (1.3 m above ground level), at 0.75m and at 0.15m were measured using a diameter tape to the nearest 0.1 cm on each tree. After felling, total length was measured to the nearest 0.01 m, and then diameters along the upper tree boles were measured to the nearest 0.1cm at 3m intervals starting from a height of 3m above ground level[1].

Individual tree volumes were the sum of volume sections. The volume sections were calculated by using Smalian's formula presuming the top end section of each tree was a cone (see section 2.1). Table 3.1 shows a summary of the data. Plots of height against diameter at breast height, and relative height against relative diameter are shown in Figure 3.2 and Figure 3.3, respectively.

Table 3.1: A summary of the data

		Height				**DBH**			
sites	**no. of trees**	**mean**	**max**	**min**	**SD**	**mean**	**max**	**min**	**SD**
Plantation 1	10	12.05	16.72	8.43	2.15	14.29	25.10	10.00	4.35
Plantation 2	25	12.51	15.32	12.52	2.28	15.10	18.50	11.20	2.48
Natural 1	12	13.05	16.92	10.31	1.94	14.98	20.00	10.50	2.96
Natural 2	14	16.59	18.06	14.59	1.17	24.61	31.00	19.50	3.26
Natural 3	12	16.34	18.19	13.40	1.31	21.39	28.90	15.20	3.56
Total	**73**	**13.85**	**18.19**	**8.43**	**2.69**	**17.66**	**31.00**	**10.00**	**5.480**

[1] Diameters were measured over-bark owing to a decision taken under difficult circumstances. Please see section 3.4 for an explanation

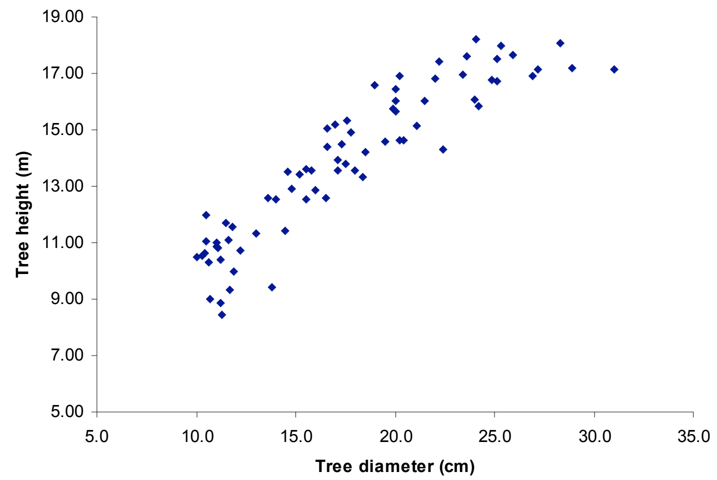


Figure 3.2: Height plotted against diameter at breast height for all 73 sample trees

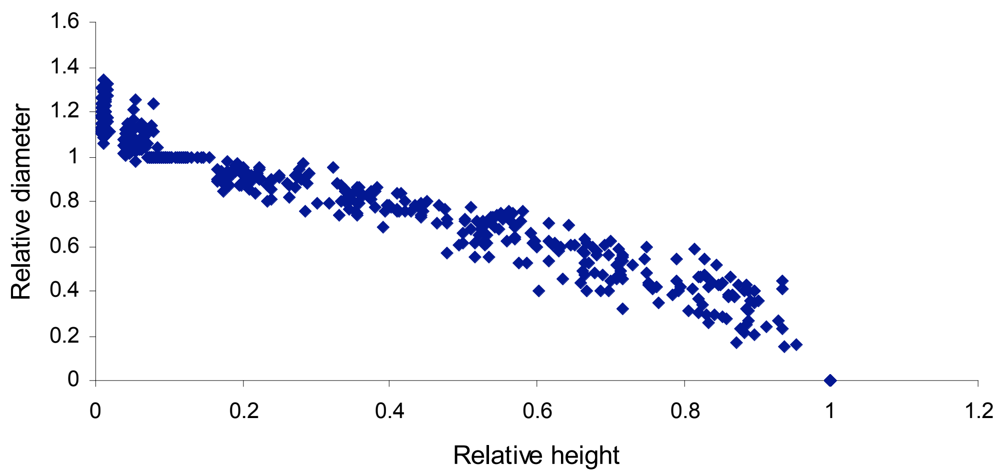


Figure 3.3: Relative diameter (the quotient between an upper-stem diameter and D) plotted against relative height (the quotient between above ground level to an upper-stem diameter and total tree height) for 571 sections.

3.3 Models

3.3.1 Volume models

Stem volume is a function of a tree's height (*H*), diameter at breast height (*D*) and tree shape or form factor (*F*). The general formula for the volume of a tree is

$$V = \frac{\pi}{4} D^2 HF ,$$

and a common expression for tree volume equations is

$$V = f(D, H, F) .$$

However, *F* is rarely used in tree volume model construction. Even though form (*F*) is required in some formulae for the volume of a tree, it is not a truly independent variable; like volume, form factor is usually estimated from other measurements of a tree's dimensions and form factor can be neither measured nor calculated without first measuring the volume (Philip 1994). Form quotients are ratios of diameters at specified heights to tree diameter at breast height, and cylindrical form factor is defined as the ratio of total volume to the volume of a cylinder with diameter equal to tree diameter at breast height and height equal to the tree height (Clutter *et al.* 1983).

Most volume models are constructed from one variable, *D* only, or two variables, *D* and *H*. The two parameter volume equations are the most widely used. Such equations have functional form

$$V = f(D)$$

$$V = f(D, H) .$$

Volume equations with more than two independent variables are not widely built and used. The third and fourth independent variables can be crown height, site index, ratios of D_i to *D* or other tree dimensions.

This study included tests on various volume equations of one parameter and two parameters limited to the functional forms $V = f(D)$ and $V = f(D,H)$. Most equations were directly obtained from the available literature. Some models were modified specifically for the study. All volume models are included in Table 3.2.

Table 3.2: A list of candidate volume equations

no.	Models	References	Type
1	$\ln(V) = b_0 + b_1 D$	Laar and Akca 2007	single-entry
2	$\ln(V) = b_0 + b_1 \ln(D)$		
3	$V = b_0 + b_1 D^2$		
4	$V = b_0 + b_1 D + b_2 D^2$		
5	$V = b_1D^2H$	Clutter *et al*. (1983)	double-entry
6	$V = b_0 + b_1D^2H$		
7	$V = b_0 + b_1D^2 + b_2H + b_3D^2H$		
8	$V = b_1D^{b2}H^{b3}$		
9	$V = b_0 + b_1D^{b2}H^{b3}$		
10	$V = D^2/(b_0 + b_1H^{-1})$		
11	$V = b_0 + b_1D^2H + b_2D^3H$	Bi and Hamilton (1998)	
12	$V = b_0 + b_1D^2H + b_2D^3H + b_3D$		
13a	$V = b_0 + b_1D^2H + b_2D^2H^2$		
13b	$V = b_1D^2H + b_2D^2H^2$		
14	$V = b_0 + b_1D^2H + b_2D^2H^2 + b_3H$		
15	$V = b_0 + b_1D^2H + b_2D^3H + b_3D^2H^2$		
16	$V = b_0 + b_1D^2H + b_2D^3H + b_3D^2H^2 + b_4D$		
17	$V = b_0 + b_1D^2H + b_2D^3H + b_3D^2H^2 + b_4H$		
18	$V = b_0 + b_1D^2H + b_2D^3H + b_3D^2H^2 + b_4D +$		

3.3.1.1 Single-entry volume models

Single-entry volume models are constructed by relating tree volume to one of a tree's dimensions. Diameter at breast height or basal area (*G*) is the most commonly used dimension. The relationship between *D* and *G* is a direct mathematical expression. The formula is

$$G = \frac{\pi}{4} D^2 .$$

Four functional forms of single-entry volume equations were tested with the data set. Two were of logarithmic form (Table 3.2).

3.3.1.2 Double-entry volume models

The second portion of the study focused on the tests on double-entry volume models. Double-entry volume equations are probably the most common form of volume model (Philip 1994). Many forms and equations of this type were tested with the data set. Amongst the candidate models tested were six equation forms commonly used for estimation of stem volumes given by Clutter *et al*. (1983) and the various modification forms of generalized combined variables listed in Bi and Hamilton (1998). They were listed in Table 3.2.

3.3.2 Taper models

A tree's taper can be expressed as a function of diameter at breast height (D), total height (*H*) and upper stem height (*h*). The common expression for functional form of a taper equations is

$$d = f(D, H, h).$$

Taper equations can be represented by a single simple quadratic model describing a whole bole length of a tree, or by a system of many complex models describing different sections of trees separately. A taper function is also related to a volume function. A taper equation, in which a volume equation is presented in the process of constructing the equation, is known as a compatible taper equation, and has the property that it will produce the same volume as the volume equation if it is integrated and rotated around an axis representing height of the tree.

Taper equations can be classified by different methods. One classification method classifies them into two groups: (i) **compatible taper models** and (ii) **non-compatible models**. Another method of classifying taper equations was described by Methol (2001). This method groups them into four categories according to the taper model components, characteristics of independent variables and the process of model construction. The four categories were: (a) **single functions**; (b) **segmented taper models**; (c) variable-**exponent taper equations**; and (d) **between-tree variable form models**.

Both compatible and non-compatible taper equations were examined in this study. They were within the first two groups, single functions and segmented taper models. Compatible taper models generated in this study were polynomial equations that were within the group of single functions. One model of segmented polynomial equation consisting of three sub-models was tested with the data. All taper equations tested in this study included in Table 3.3.

Table 3.3: A list of candidate taper equations

no.	Models	References	Type
19	$d^2 = D^2\left(\beta_1\left(\frac{h}{H}-1\right)+\beta_2\left(\frac{h^2}{H^2}-1\right)\right)$	Kozak *et al.* (1969)	***(1a)***
20	$d^2 = D^2\left(\frac{h}{h_D}\right)^{2-\beta_1}\left(\frac{H-h}{H-h_D}\right)$	Sharma and Oderwald (2001)	
21	$d = D\left(\frac{H-h}{H-h_D}\right)^{\beta_1}$, $\beta_1 > 0$	Ormerod (1973)	
22	$d = D\left(\beta_0+\beta_1\left(\frac{h}{H}\right)+\beta_2\left(\frac{h}{H}\right)^2+\ldots\beta_n\left(\frac{h}{H}\right)^n\right)$		***(1b)***
23	$d^2 = \frac{V_{ml}(\beta_1 z_1+\beta_2 z_2+\beta_3 z_3+\beta_4 z_4+\beta_5 z_5)}{kH}$, where V_{ml} is volume estimated by volume model ***Note***: In this study, two functional forms of V_{ml} were tested: (i) **23a** with ***volume model 4*** and; (ii) **23b** with ***volume model 13a***.	Goulding and Murray (1976), and Gordon (1983).	***(2)***
24	$d^2 = D^2\left(\beta_1\left(\frac{h}{H}-1\right)+\beta_2\left(\frac{h^2}{H^2}-1\right)+\beta_3\left(a_1-\frac{h}{H}\right)^2 I_1+\beta_4\left(a_2-\frac{h}{H}\right)^2 I_2\right)$ *where* $I_1 = 1$ if $\frac{h}{H} \le a_1$ and $I_1 = 0$ otherwise $I_2 = 1$ if $\frac{h}{H} \le a_2$ and $I_2 = 0$ otherwise	Max and Burkhart (1976)	***(3)***

<u>Type</u>*:** ***(1a) Single simple functions; ***(1b)*** Single polynomial functions; ***(2)*** Compatible taper model; and ***(3)*** Segmented polynomial taper equation.

3.3.2.1 Single functions

Single taper equations tested with the data set consisted of three single simple models (equations 19, 20, and 21), and one polynomial series model (equation 22).

3.3.2.2 Compatible taper models

Compatible polynomial form of taper was tested with the data set (model 23). The theory and development of compatible polynomial taper equations were described by Goulding and Murray (1976), and Gordon (1983).

Model 23 was tested by using two different volume models, in order to compare the influences of volume models on such compatible systems. Two volume models were one-entry and double-entry volume equations. The two volume equations were:

$$V_{(volume\ model\ 4)} = 0.0062478622 - 0.0048205045\mathrm{d} + 0.0008340514\mathrm{d}^2$$
$$V_{(volume\ model\ 13a)} = -\ 0.0011383579 + 0.000052703\mathrm{d}^2\mathrm{h} - 0.0000008312\mathrm{d}^2\mathrm{h}^2$$

Note: **Model 23** that was tested with $V_{(volume\ model\ 4)}$ and $V_{(volume\ model\ 13a)}$ is referred to **model 23a** and **model 23$_b$**, respectively. Estimating parameters for the two compatible volume systems followed the same procedure. There were five parameters to be estimated for each of two compatible taper models. The Polynomial series of notations $z_n = (n+1)\left(\frac{H-h}{h}\right)^n$ were used.

3.3.2.3 Segmented polynomial taper equation

The segmented polynomial taper with three sub-models was examined in this study (model 24). The terminology and methodology of segmented polynomial regression were described by Max and Burkhart (1976). All three sub-models are quadratic. Two join-point parameters were represented by I_1 and I_2.

3.3.3 Model fitting

All statistical analyses for both volume models and taper equations were carried out with the SAS statistical software package. The primary analyses of all models used the least squares method. Further analyses of the compatible taper equations used restricted least squares, while further analyses of some volume models used weighted least squares. Two SAS options of the linear regression and nonlinear regression were chosen to assist the requirement of the above analyses. The subsequent SAS programmes within the two regression options employed were the GLM procedure, the REG procedure and the NLIN procedure.

3.3.4 Evaluating models

Statistical interferences based on t and F distributions were not sufficient to evaluate the volume models, and were invalid for the taper models because of the nature and characteristics of the data. Since taper data were obtained by taking diameters at several positions along the same tree, they were not independent. The quality of a volume model cannot be fully illustrated by examining t and F. The volume model that has apparently good precision, with strongly significant t and F values, can be biased.

Residual analysis was necessary for evaluating both volume models and taper models in this study. Fit statistics of residual analyses were used along side with the graphical methods of residual plotting. Fit statistics used were among the most common ones. They included bias (mean residual), the standard error of estimate (SEE), the standard deviation of the residuals (SDR) and the mean absolute deviation (MAD). These statistics are defined as:

$$Bias = \frac{\sum_{i=1}^{n}\left(y_i - \hat{y_i}\right)}{n} \qquad SEE = \sqrt{\frac{\sum_{i=1}^{n}\left(y_i - \hat{y_i}\right)^2}{n-p}}$$

$$SDR = \sqrt{\frac{\sum_{i=1}^{n}(y_i - y_i^{\wedge})^2 - \frac{\left(\sum_{i=1}^{n}(y_i - y_i^{\wedge})\right)^2}{n}}{n-1}} \qquad MAD = \frac{\sum_{i=1}^{n}\left|y_i - y_i^{\wedge}\right|}{n}$$

where y_i is observation

$y_i^{\wedge}$ is prediction

n is number of observations

p is the numbers of estimated parameters

Graphs of residuals were plotted against predicted values and independent variables, and these were examined for evidence of bias. In addition, frequency distributions of residuals were examined for departures from normality.

3.3.5 Estimating parameters for volume models

Parameters for volume models were estimated by using PROC GLM and PROC NLIN. Estimating parameters for equations 8, and 9 that are non-linear used the NLIN procedure. On the other hand, the GLM procedure was used for the other linear volume equations.

According to Cunia (1964) in fitting the tree volume data by using the GLM procedure cannot effectively estimate the variance due to heteroscedasticity. One way of correcting for non-homogeneity of the variance is to estimate the model parameters using weighted least squares.

This study applied the weighted least squares method using weighting factors (ω) of: $\omega = \frac{1}{D^2H}$, and $\omega = \frac{1}{D^2}$. While the first (ω) was used for those volume equations of two parameters, the latter (ω) was used for those volume equations of one parameter.

In contrast, for the NLIN procedure there are fewer problems with heteroscedasticity. Cunia (1964) suggested that using logarithmic forms of volume equations is another way to overcome the problems of non-homogeneity of the variance. Two volume equations of this study were tested with the NLIN procedure.

The t-test was used to determine the significance of estimated parameters. The estimated parameters that were not significant had the absolute t-value very small around zero. The insignificant estimated parameters were removed from the models.

3.3.6 Estimating parameters for taper models

The processes and procedures of estimating the parameters for four taper equations of the single functions were similar to those of the volume equations. The taper equations 19 and 22 were fitted using the GLM procedure and the REG procedure, respectively. The NLIN procedure, on the other hand, was used to fit the equations 20 and 21.

Compatible taper equations were fitted using the REG procedure. This procedure allows a restriction command to be added into the analysis. The compatible taper and volume system used in this study adapted a variable '*z*' (previously defined), and a variable '*y*'. The variable '*y*' was defined as:

$$y = \frac{kd^2 H}{V_{ml}}$$

The initial parameters were estimated for the variable '*z*' by establishing the relationships between '*y*' and '*z*', and based on the theory of the least squares method. The restriction was applied to the estimated parameters to make the model compatible.

The NLIN procedure was used for the segmented taper equation, equation 24. Three steps were set in the model procedure with the conditions of $I_{1,\text{and } 2}$. The first step was to fit the model to the lower bole portions of trees by conditioning both $I_{1,\text{and } 2}$ equal zero. Both parameters a_1 and a_2 ($0 > a_1 < a_2$) were not useful. The following step allowed the procedure to fit the middle bole portions by conditioning I_1 equal to one and I_2 equal to zero. Only parameters a_1 alongside with parameters $b_{(1-4)}$ were used in this step. The last step allowed both parameters a_1 and a_2 to fit the equation to the upper bole portions.

3.4 Problems and constraints

Usually under-bark volume and taper receive more attention from researchers than over-bark volume and taper. For example taper equations constructed by Bruce *et al.* (1968); Ormerod (1973); Goulding and Murray (1976); McClure and Czaplewski (1986); Newnhan (1992); and Kozak (1997) were all under-bark models. While some taper equations have been constructed for both under-bark and over-bark tree tapers, a few were constructed for over-bark tree tapers only (Methol 2001). The later included the taper functions by Diéguez-Aranda *et al.* (2006), and Reed and Byrne (1985).

Initially this study was designed to model both under- and over bark stem volume, as well as both under- and over-bark tree taper. However, there were some constraints, limitations and problems during field data collection that forced a change in the original proposal. I had to drop the measurement of bark thickness in order to speed up field work during a brief return to Laos during my studies, in order to get enough tree samples. One factor was that the study was delayed by an unplanned road incident. Another factor was that the study sites were very remote and located in rough terrain with undesirable conditions for field work. I was assured that a bark gauge was available at a field station that was my last point of departure for the remote field sites, when in fact a bark gauge was not available at the field station. I began stripping bark to measure it with a ruler, but this consumed more time and resources than were available. Therefore, to collect enough or sufficient the numbers of sample trees would have been impossible if under-bark measurement was undertaken. Dropping measurements of bark thickness was a better option than reducing the numbers of sample trees because the numbers of sample trees proposed were already minimal for effective volume and taper modelling for a new species.

Measurement of bark was carried out on few trees. At each measuring point along the stem two bark measurements were taken by peeling it off, and measuring it using a ruler with millimeter marks. Double bark thickness was recorded as a sum of the two measurements. There were only 20 data points. They were measured from 4 trees that the diameters of breast height of around 14 cm. A summary of the data and the results of an analysis of bark thickness are in the Appendix 1. Data were too few to be used for building a reliable bark model for the taper and volume study described here. Under bark volumes of some sample trees were estimated using a bark model built from this small data set (Appendix 1), particularly those with 'D' smaller than 14 cm.

Bark measured during this study had a slightly curvilinear relationship with diameter. Non-linear forms of equations were prime candidates, and it was found that the best form of equation was: $Y = \beta_0 + \beta_1 d^{\beta_2}$, where *Y* is double bark thickness and *d* is over-bark diameter at a range of points up a stem. The parameters for this model were estimated using the NLIN procedure, and residual analysis followed the same procedures as those used for non-linear volume equations. Courbet & Houllier (2002) implied a model of double bark thickness of *Cedrus atlantica* by modelling the ratio of over-bark to under-bark diameter as a function of distance from tree apex:

$$\frac{D_o}{D_i} = c_1 + \frac{c_2}{x^{c_3}}$$

where D_o = over-bark diameter, D_i = under-bark diameter, X= distance from stem apex, and c_1, c_2 and c_3 were coefficients. Their model implies an almost linear increase in bark thickness with diameter when it is applied to estimates of diameter from a fitted taper equation.

4 Results

4.1 Estimated parameters

Estimated parameters of fitted models are presented in Table 4.1. The first 19 models are volume models and the seven remaining models at the bottom of the table are taper models. Two equations of each type had one estimated parameter. Two taper models had six estimated parameters and that was the highest number of parameters in a single equation tested in this study. One volume model also had six estimated parameters.

Many volume models contained estimated parameters that were not useful. Such parameters had very small t-values and were not significant at the 5% level ($p=0.05$). Insignificant parameters are marked with asterisks in Table 4.1. The models that contained insignificant estimated parameters were not considered further. This will be discussed further in the sections on stem volume.

Insignificant parameters were excluded from models during the parameter estimating process of taper models. Therefore, they are not included in Table 4.1. Parameters for taper models were estimated using a slightly different technique to that used for volume models. Fewer taper models than volume models were tested, and the forms of each taper model differed markedly from each other. Form of polynomial taper (equation 22) was one of them. The best form of polynomial taper equation 22 for data set was with the maximum power (n) equal to five. More details on the taper results will be presented in the taper equation sections.

Table 4.1: Estimated parameters for both volume models and taper equations

Model	Estimated parameters								Type
	β_0	β_1	β_2	β_3	β_4	β_5	a_1	a_2	
1	–4.307128116	0.138775542							Volume
2	–8.777960846	2.451721028							Volume
3	–0.0320955366	0.0006970493							Volume
4	0.0062478622	–0.0048205045	0.0008340514						Volume
5		0.0000394999							Volume
6	0.0076987481	0.0000380232							Volume
7	–0.0155762999	0.000157317*	0.0014547596*	0.0000283177					Volume
8		0.000103	2.0207	0.6289					Volume
9	–0.0142	0.000192	1.9350	0.5170					Volume
10	388.37623	19340.5766							Volume
11	0.0035692069	0.0000425543	–0.0000001729*						Volume
12	–0.0068662056	0.0000383096	–0.0000000887*	0.001306919*					Volume
13a	–0.0011383579	0.0000527030	–0.0000008312						Volume
13b		0.0000511539	–0.000000746						Volume
14	–0.006234753	0.0000518431	–0.0000008075	0.0005518114*					Volume
15	–0.0014589324	0.0000526977	–0.0000000399	–0.0000007717*					Volume
16	0.0673786113	0.0001070346	–0.0000003951	–0.000002439	–0.0099818225				Volume
17	–0.0080354411	0.0000514999	0.0000000323*	–0.000000846	0.0007748548*				Volume
18	0.1435781767	0.0001464542	–0.0000009870	–0.0000030718	–0.0160811635	–0.0040220457*			Volume
19		–2.015015689	0.726953554						Taper
20		2.1074							Taper
21		0.6124							Taper
22	1.21557	–3.44075	15.94318	–39.63581	43.57723	–17.65705			Taper
23a		1.66332	–6.34498	16.78882	–18.03822	6.93107			Taper
23b		1.66643	–6.29986	16.59056	–17.77828	6.82116			Taper
24		– 1.2665	0.1842	7.0547	70.063		0.186	0.0517	Taper

4.2 Volume equations

Examination of the estimated parameters revealed that some volume equations could be eliminated from the list of candidates, and they were not examined further. These equations contained estimated parameters that had very small t-values that were not significant at the 5% level (p=0.05). Those insignificant parameters were distributed within 7 models including models 7, 11, 12, 14, 15, 17 and 18.

When fitting candidate equations to data, insignificant variables are usually removed from the models, and new tests on fitting are carried out again until there are no more insignificant parameters in the models. Because seven models that contained the insignificant parameters found in this study were related to other candidate models, removing one or more variables of these models transformed them to other candidate models. For example, two independent variables (D^2 and H or β_1 and β_2) in model 7 were insignificant, and excluding these variables transformed the model 7 into the candidate model 6. Removing the insignificant variable D^3H from model 11 also transformed it into the candidate model 6. Meanwhile, removing variable H from model 14 transformed it into model 13a which was the final candidate volume model. Model 13a will be examined in details in the later sections.

4.2.1 Independent variable D^2H

Table 4.2 presents the variables and t-values of twelve combined variable equations. Significant parameters are shown in bold font. All models contained the independent variable D^2H. The models were different forms of combined variable volume equations. Variable D^2H was highly significantly related to stem volume. It was significant in all twelve volume models. Bi and Hamilton (1998) explain that this variable represents the volume of a cylinder of diameter D and height H. Stem volume is directly related to the cylindrical volume by the coefficient of this variable that varies with stem form, that is, the solid shape of the stem.

Table 4.2: The significant parameters at 5% level and t-values

Model	D^2H		D^3H		D^2H^2		D^2		D		H	
	t-value	Pr > | t |	t-value	Pr > | t |	t-value	Pr > | t |	t-value	Pr > | t |	t-value	Pr > | t |	t-value	Pr > | t |
5	**99.75**	<0.0001										
6	**71.2**	<0.0001										
7	**5.13**	0.0001					1.66	0.1021			1.12	0.2653
11	**12.55**	<0.0001	−1.35	0.1803								
12	**4.65**	<0.0001	−0.45	0.6526					0.57	0.5728		
13a	**8.09**	<0.0001			**−2.26**	0.0269						
13b	**19.65**	<0.0001			**−4.52**		<0.0001					
14	**7.58**	<0.0001			**−2.16**	0.0342					0.44	0.6615
15	**8.04**	<0.0001	−0.27	0.7852	−1.80	0.0766						
16	**4.12**	0.0001	−1.82	0.0736	**−2.78**	0.071			**−2.16**	0.0344		
17	**6.99**	<0.0001	0.13	0.8959	−1.77	0.0809					0.37	0.7156
18	**3.93**	0.0002	**−2.15**	0.0349	**−3.16**	0.0024			**−2.60**	0.0116	−1.46	0.1479

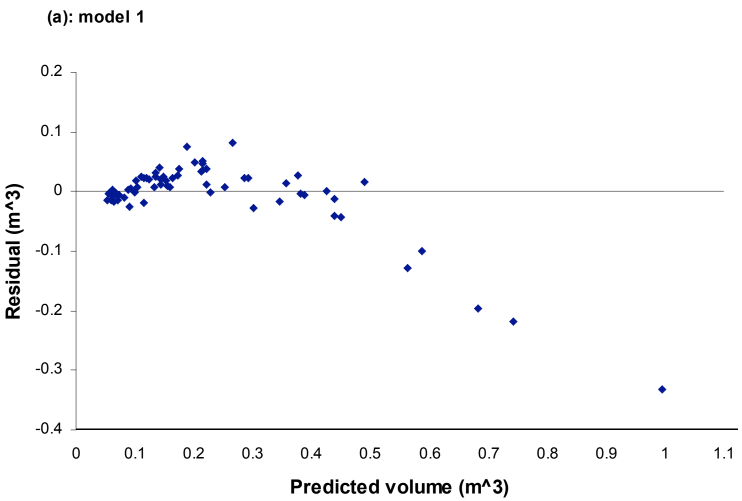
4.2.2 Graphs of residuals

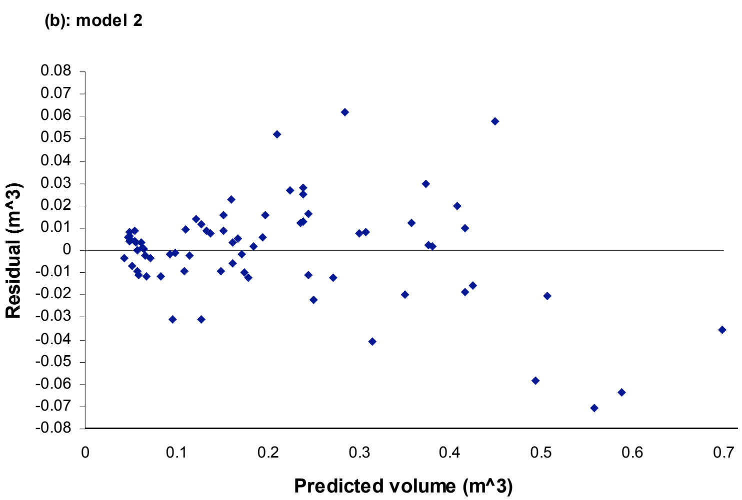
Seven volume models were eliminated as they contained insignificant parameters. The remaining models were evaluated by analyzing errors in their volume predictions. For each model, graphs of residuals were plotted, and common statistics of fit Bias (mean residual), SEE (standard error of estimate), SDR (standard deviation of the residual) and MAD (mean absolute deviation) were calculated. Figure $4.1_{\textbf{(a-l)}}$ presents plots of residuals. Fit statistics are presented in Table 4.3. Figure 4.1a shows the residual plot of volume model 1, and Figure 4.1b shows that of volume model 2. Figure 4.1c shows the residual plot of volume model 3 and so on. It was clear from these residual plots that model 1 was a biased model. It was also the least precise model. Model 2 was also biased, but less obviously so than model 1. Fit statistics of these two models also indicated that they were biased. The models were excluded from ranking, and were not included in Table 4.3.

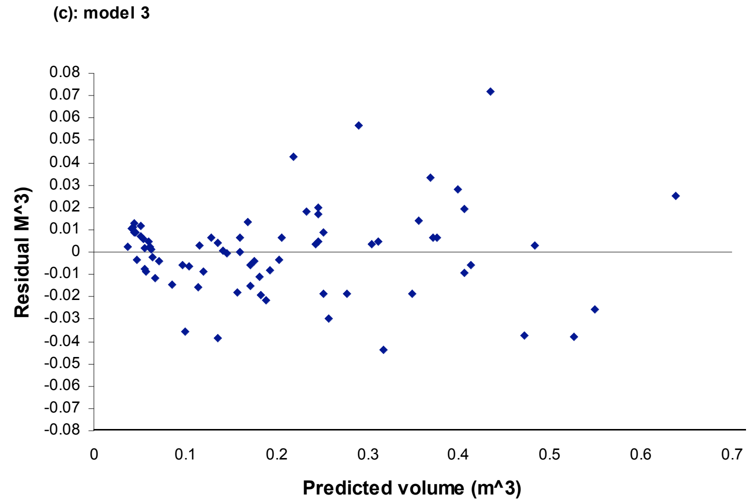
Even though analyzing a graph of residuals can be a useful tool to evaluate volume model performances, it is ineffective in some circumstances. This study found that the graphs of residuals from ten models (models 3, 4, 5, 6, 8, 9, 10, 13a, 13b, and 16) were not significant different, and they could not be used to distinguish their precision and accuracy in tree volume predictions. Evaluating the performances of these models required the other statistics tools, and this will be discussed in the next section. Meanwhile, the observation on insignificant differences in the graphs of residuals from the ten models can be described as follows:

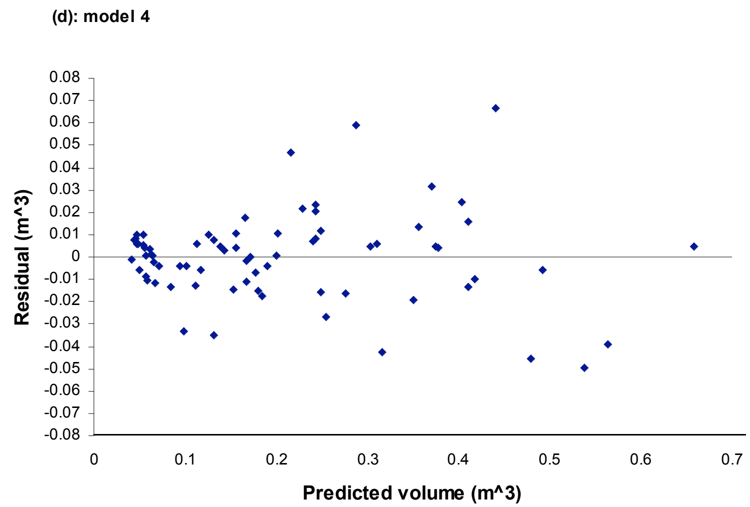
Models 9, 13, and 16 were amongst the best candidate volume models. They were slightly less biased than other models. On the other hand, models 5, and 6 were among the most biased models. However, graphs of residuals from the best group and from most biased group were not obviously different. They performed similarly for predicting the volumes of trees. It could be observed that the models performed particularly well when predicting the volumes of small trees but moderately in predicting the volumes of big trees, as more biased were observed at the upper ends than lower ends in those graphs of residuals. Moreover, biased patterns in their residual plots were similar, and errors in their volume predictions were within the similar range. Also, they all overestimated slightly the volumes of big trees.

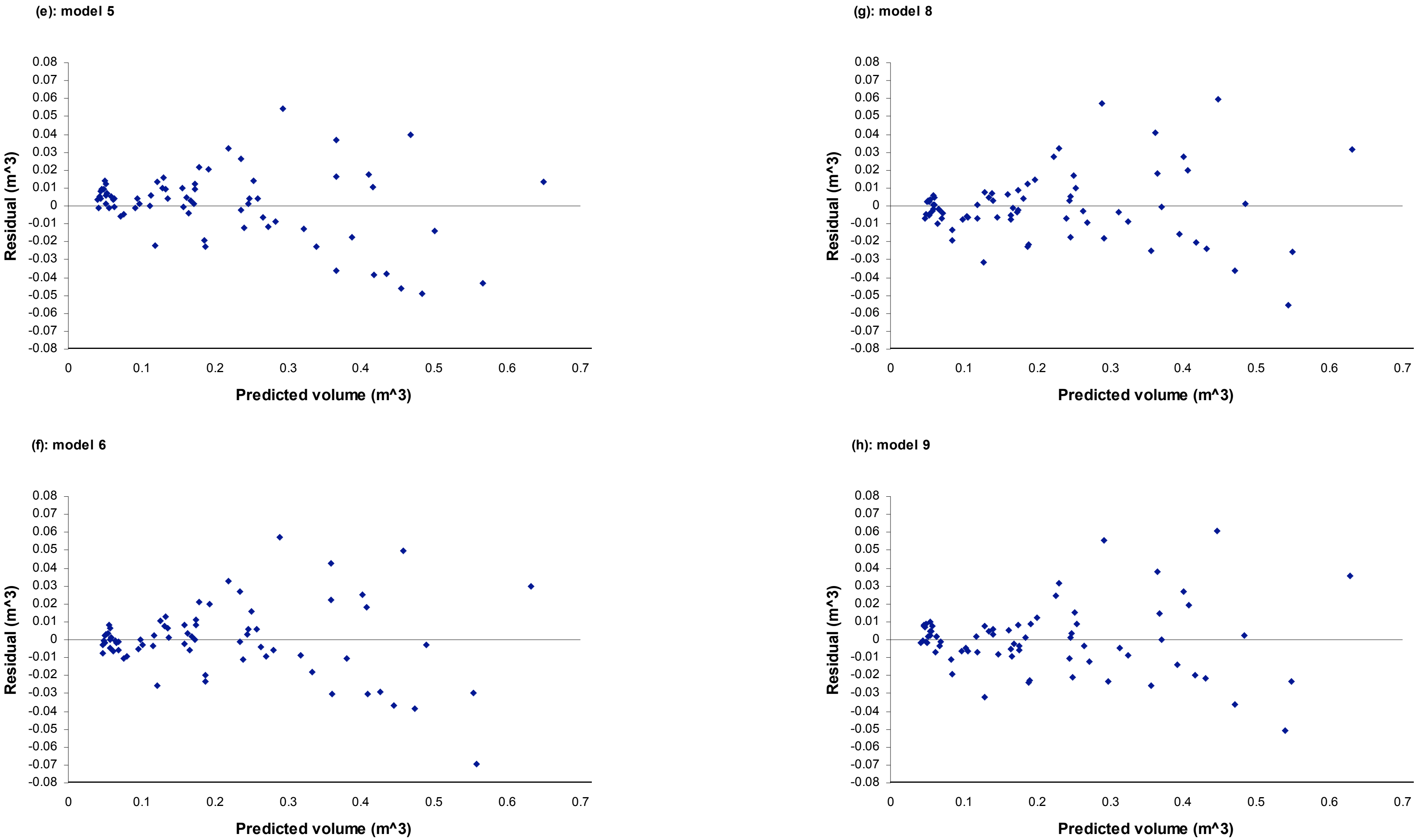
Figure 4.1: Plots of residuals from volume equations



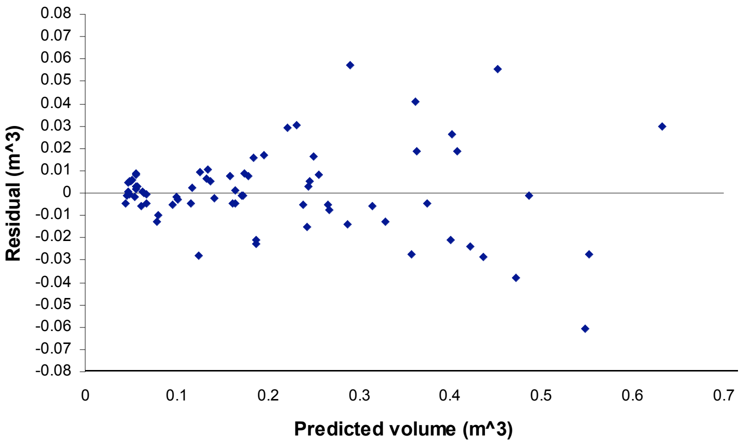




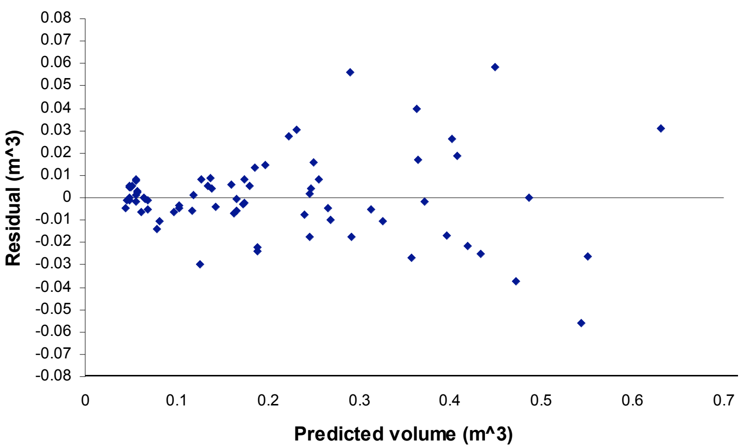



(e): model 5
(g): model 8
(f): model 6
(h): model 9
Residual (m^3)
Predicted volume (m^3)
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
-0.01
-0.02
-0.03
-0.04
-0.05
-0.06
-0.07
-0.08
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7

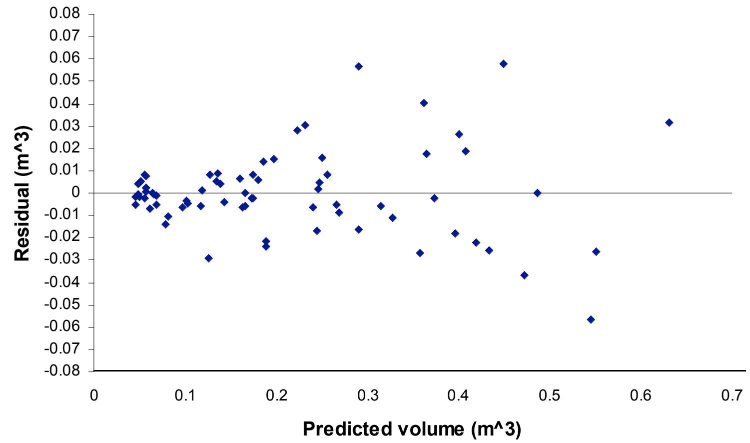
(i): model 10



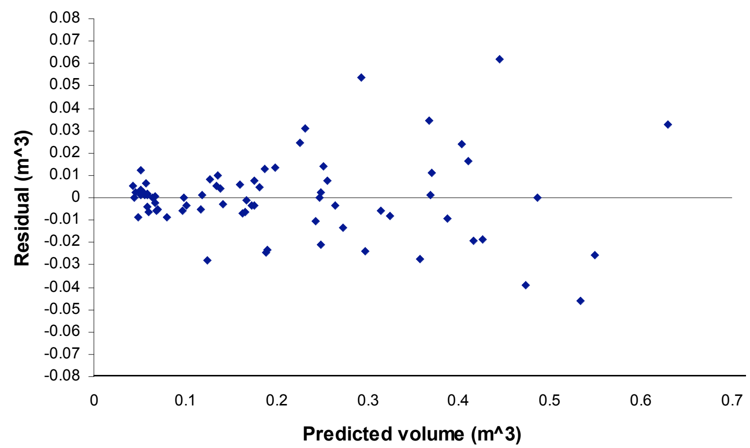
(j): model 13a



(k): model 13b



(l): model 16



4.2.3 Fit Statistics of residuals

After examining the graphs of residuals, it was clear that fitted equations 1, and 2 were biased. It was not necessary to calculate their statistics of fit (bias, SEE, SDR and MAD), and compared them with the statistics of the other equations. Therefore, only ten equations were left for examining the fit statistics of residuals. Bias, SEE, SDR and MAD calculated for these equations are shown in Table 4.3. The table also shows the results of model ranking.

Various statistical results were used to rank the volume models. Models were firstly ranked four times by bias, SEE, SDR and then MAD. Ranking values were then summed and divided by four to give an average value for overall model ranking. The best model by the individual statistics was assigned a rank of one, while the poorest model was assigned a rank of ten. Therefore, if the best model was ranked first by all four statistics, the model would have an average value of one for overall model ranking. The model with smallest average value was ranked first by overall model ranking. This scheme was used as a preliminary screening technique only, and then the highest ranked models were evaluated carefully to explore the consequences of their relative bias and imprecision.

The best model by overall model ranking was model 16. This model was ranked seventh by bias, but first by other three statistics. Models 13a, and 13b were ranked second and third by overall model ranking, respectively.

Some models generated similar values in every statistics calculated that made it hard to distinguish them. Most models generated very small values in bias statistics. The other three statistics were also very similar among some of models.

Bias

Models 8, and 10 were the two of the most biased models. These two models were significantly biased when compared with model 4 that was ranked at number one by this statistic. The value of bias for model 4 was 1.7×10^{-17}, while the values of bias for model 8, and 10 were 0.00059 and 0.0051, respectively. Model 9, on the other hand, was ranked at eight with a bias value of $2.1 \times 7 \times 10^{-9}$.

The bias statistic was not useful for determining differences among the seven other models (model 3, 4, 5, 6, 13a, 13b, and 16). These models generated very small bias. Model 16 ranked at number seven was very little different from model 4 ranked at number one in value of bias. The former model had a bias value of 2.2×10^{-16}, while the later model had a value of 1.7×10^{-17}.

SEE, SDR and MAD

By examining SEE, SDR and MAD, it was clear that model 5 was the most imprecise model for predicting the volume of trees. The model was ranked at the bottom by all three statistics. Other imprecise models included models 8, and 3. Model 3 is a single-entry volume model using 'D' as an independent variable, while model 8 is a logarithmic form.

The top ranking models by these statistics were models 16, 13a, 13b and 9. They were ranked first, second, third and fourth respectively. Models 13a, and 13b were ranked second and third by MAD, respectively. In contrast, model 13b was ranked second, while model 13a was ranked third by SDR. Model 9 was ranked third, fourth and fifth by SDR, SEE and MAD, respectively.

Three models ranked at the middle by overall statistics were model 6, 10, and 4. Model 6 was ranked fifth, while model 10, and 4 were ranked sixth and seventh, respectively. Model 4 is a single-entry model. This model was the least biased model ranking first by bias.

While the differences in these statistics between the poorest model (model 5) and best models (model 16) were clear, these differences were not substantial among the 4 highest ranked models (models 16, 13a, 13b and 9). Values of SEE and SDR for all four models were around 0.018, while their values of MAD were around 0.012.

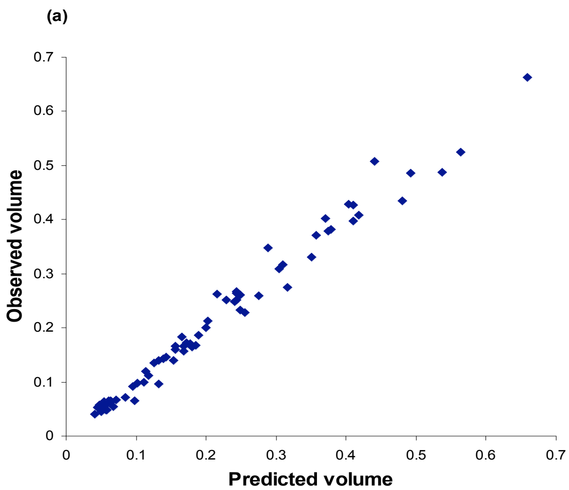
Table 4.3: The statistics of Bias, SEE, SDR and MAD, and model ranking for volume equations

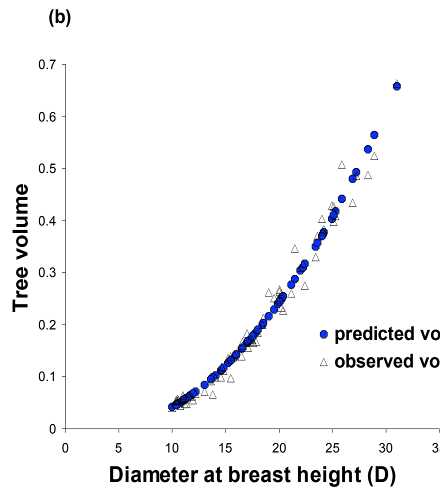
Models	Bias	Rank	SEE	Rank	SDR	Rank	MAD	Rank	Overall rank
3	-3.23×10^{-17}	2	0.020046	8	0.019906	8	0.014174	9	7 (6.75)
4	1.73×10^{-17}	1	0.020343	9	0.020059	9	0.013954	8	7 (6.75)
5	1.101×10^{-16}	4	0.021112	10	0.021112	10	0.014402	10	10 (8.50)
6	8.54×10^{-17}	3	0.019772	7	0.019634	7	0.013092	7	5 (6.00)
8	-0.0005943	10	0.018819	6	0.018546	5	0.012664	6	9 (7.00)
9	2.10×10^{-9}	8	0.018731	4	0.018337	3	0.012618	5	4 (5.00)
10	0.0005119	9	0.018810	5	0.018672	6	0.012587	4	6 (6.25)
13a	1.33×10^{-16}	5	0.018585	3	0.018325	2	0.012375	2	2 (3.00)
13b	1.53×10^{-16}	6	0.018534	2	0.018405	4	0.012420	3	3 (3.75)
16	2.23×10^{-16}	7	0.018130	1	0.017620	1	0.01171	1	1 (2.50)

Note: In overall rank column, the number in parenthesis is the average value of four statistics. The model with smallest value ranked first (the number outside parenthesis).

4.2.4 Equations 4, and 13a

Equation 4 was the best model of single-entry volume equations tested. It was ranked seventh by overall ranking. Figure $4.2_{(a\text{-}c)}$ illustrates the predicted volumes of all trees in data set by this equation against the observed volumes. The figure also shows that the volumes had a curvilinear relationship with diameter at breast height (D) but were seemingly linearly related to D squared (D^2). Both D and D^2 were included in the equation.





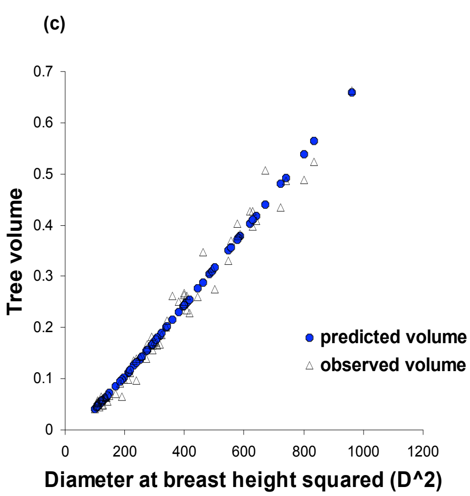
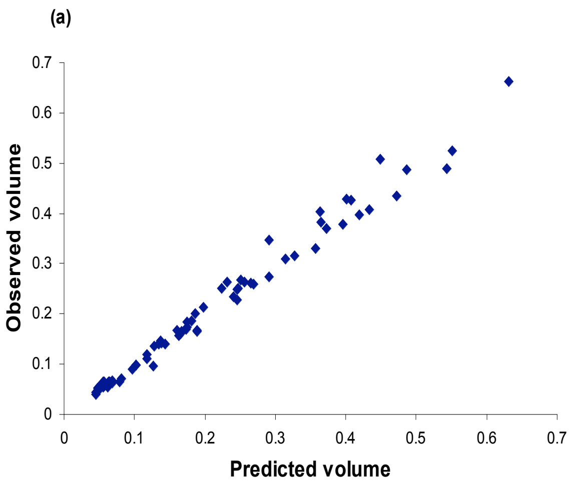
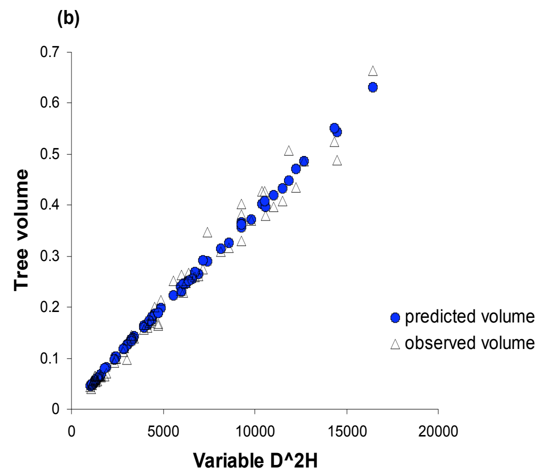


Figure 4.2: Observed volume, and predicted volume by equation 4. (a) predicted volume against observed volume; (b) predicted volume and observed volume against D and (c) predicted volume and observed volume against D squared (D^2).

Meanwhile, equation 13a was the highest ranked double-entry volume equation among those tested. It was second to model 16 by overall ranking. However, the two equations were not substantially different in predicting the volumes of trees. The complexity of model 16 that has many estimated parameters (5 in total) makes it less desirable. Figure $4.3_{(a\text{-}c)}$ presents the performance of model 13a, and also illustrates a relationship between independent variables of this equation (D^2H and D^2H^2) and the volumes of trees within data set. Both variables were plotted against both observed and predicted volumes.





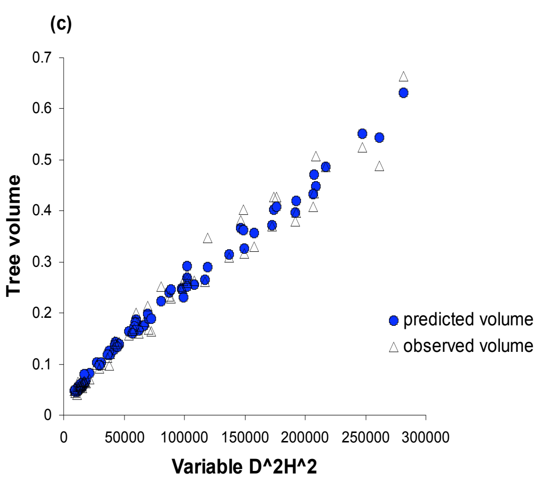


Figure 4.3: Observed volume, and predicted volume by equation 13a. (a) predicted volume against observed volume; (b) predicted volume and observed volume against D^2H and (c) predicted volume and observed volume against D^2H^2.

4.3 Taper equations

4.3.1 Intercept parameter (β_0)

Estimated parameters for the seven taper equations tested are shown in the Table 4.1. All parameters were significant at the 5% level. Most equations do not contain an intercept term (β_0). However, when fitting the equations to data, some candidate models were tested with intercept terms included. The results showed that intercepts (β_0) were either insignificant or deteriorated the performances of the models. The intercept parameter was significant in taper equation 19 but, by including it, the model did not give zero values for diameters when h equalled H. This increased the overall bias in its predictions. Meanwhile, β_0 was insignificant in equation 21. On the other hand, the intercept term was required for model 22. The equation could not function properly without this parameter.

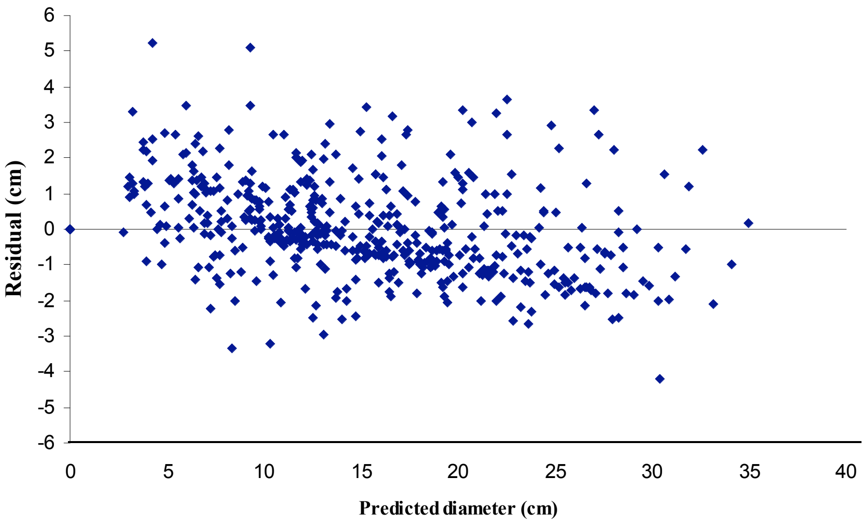
4.3.2 Residual analysis

Graphs of residuals were plotted for each candidate taper model, and they are presented in Figure 4.4. It was initially concluded that precisions in their predictions were similar. Errors in the predictions of all models were distributed between ± 5 cm. Only a few data points from some models were outside this range. However, some models were more biased than others. Compatible model (model 23) and one of three simple function models (model 20) by Sharma and Oderwald (2001) were the least biased and most precise models. On the other hand, two simple models including model 21 by Ormerod (1973), and model 19 by Kozak *et al*. (1969) were obviously biased. Model 22, a polynomial model with the maximum power of five (n=5) was moderately biased. The other taper equation with moderate bias was a segmented model by Max and Burkhart (1976). Models 20, 23a, and 23b performed similarly, and were the most precise and least biased. There were no substantial differences in the residual distributions of the three later models.

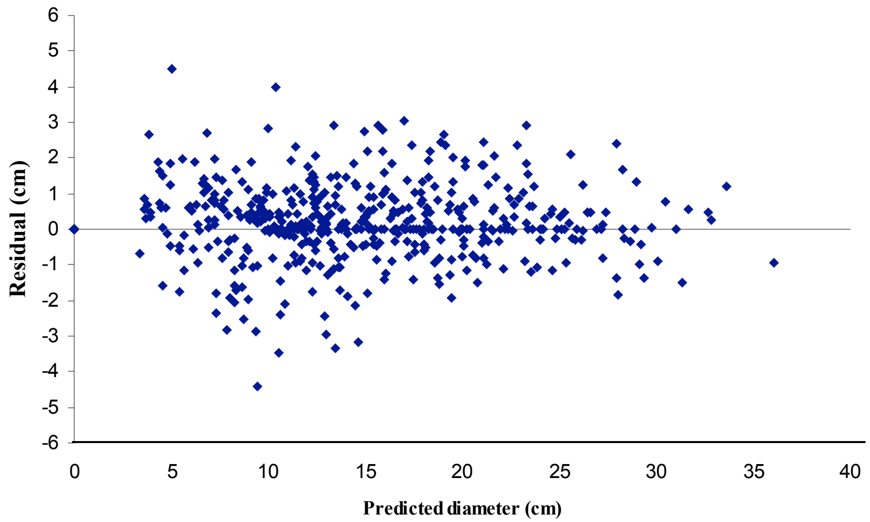
The most obviously biased models were eliminated from further evaluation. Further evaluation of these models used the results of fit statistics.

Figure 4.4: Plots of residuals from taper equations

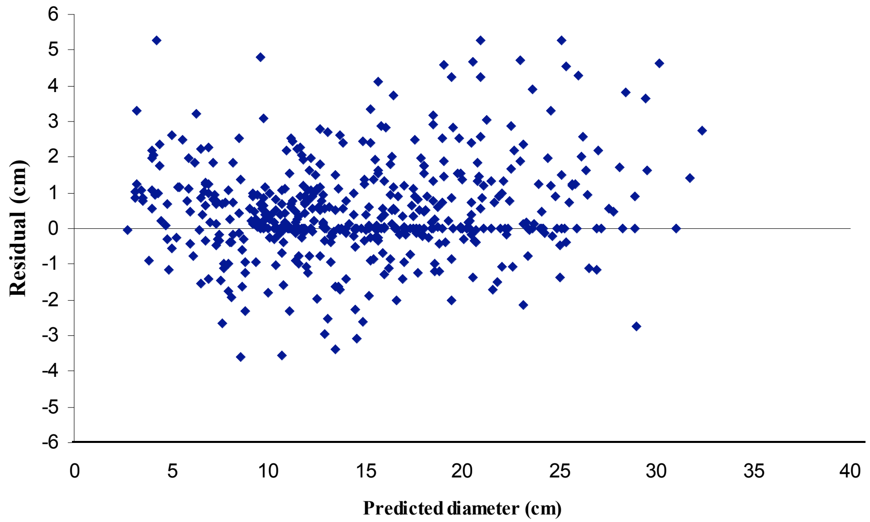
(a): model 19

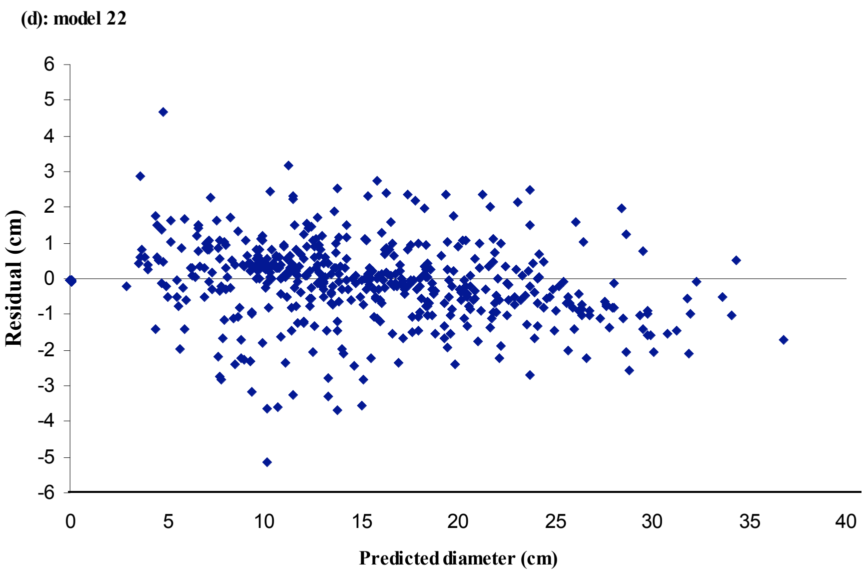


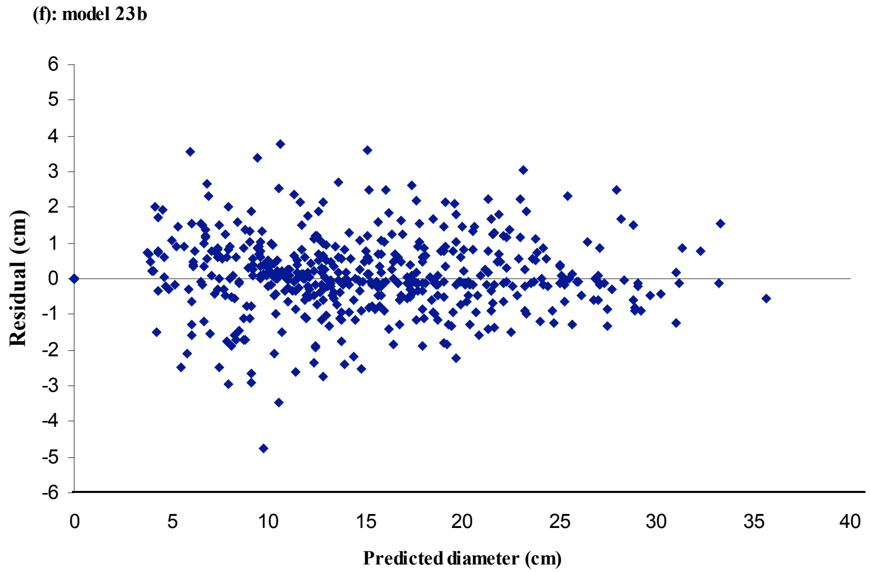
(b): model 20

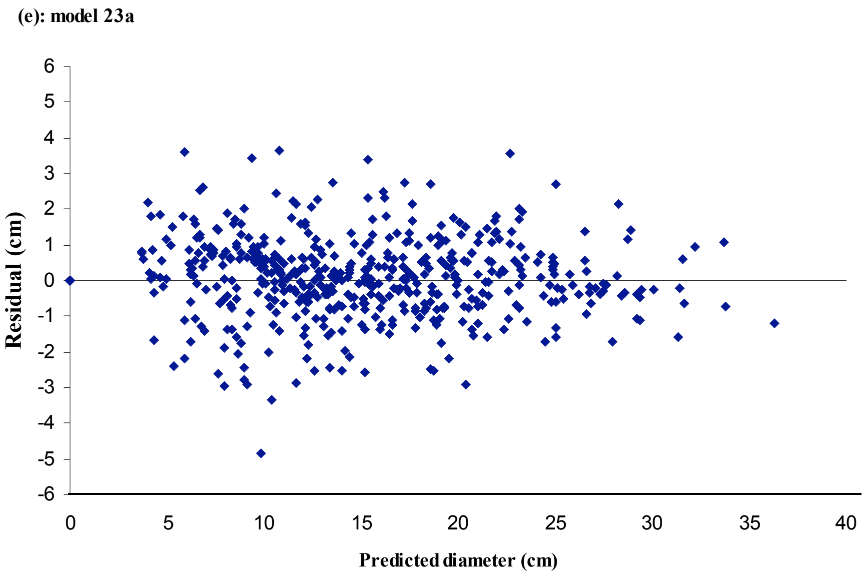


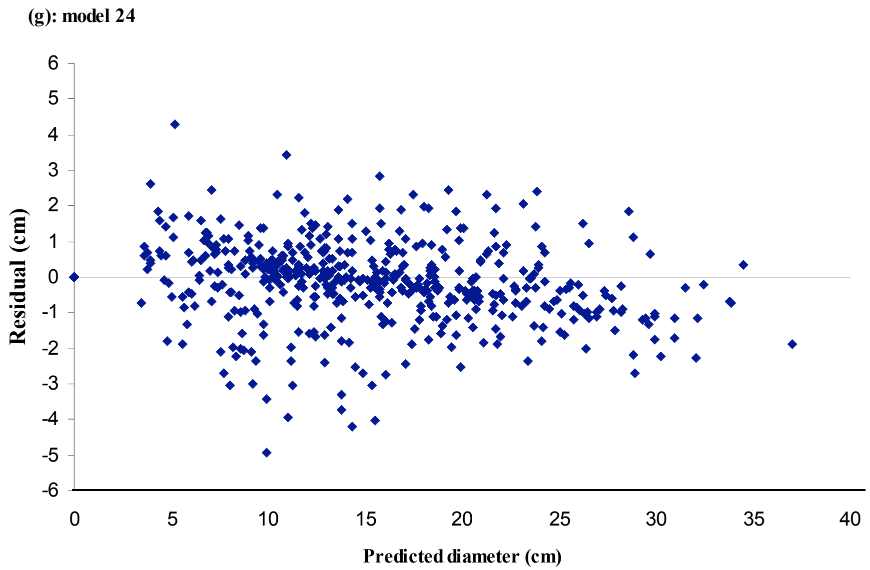
(c): model 21




(d): model 22
Residual (cm)
Predicted diameter (cm)


(f): model 23b
Residual (cm)
Predicted diameter (cm)


(e): model 23a
Residual (cm)
Predicted diameter (cm)


(g): model 24
Residual (cm)
Predicted diameter (cm)

4.3.3 Fit statistics

Bias, SEE, SDR and MAD calculated for equations 20, 23a, 23b, and 24 are shown in Table 4.3. The table also shows model ranks. The other three models (model 19, 21, and 22) were excluded; models 19, and 21 were obviously biased (see Figure 4.4) and model 22 generated errors in predicting diameter at particular portions of stems. The diameters at the tops of trees (h = H) were not zero when predicted by this latter model. Therefore, only 4 equations were left for examining the fit statistics of residuals.

Table 4.4: The statistics of Bias, SEE, SDR and MAD, and model ranking for taper equations

Models	Bias	Ranking	SDR	Ranking	SEE	Ranking	MAD	Ranking	Overall ranking
20	0.19826	4	1.09604	2	1.11386	2	0.79287	2	2.5
23a	0.08520	1	1.14527	3	1.15309	3	0.86969	4	2.75
23b	0.08685	2	1.08287	1	1.09075	1	0.78699	1	1.25
24	-0.15391	3	1.16294	4	1.17905	4	0.86362	3	3.5

The results of fit statistics showed that model 23b was the most highly ranked model. It was ranked first by three out of four statistics. The bias statistic ranked this model at number 2. The other compatible taper equation (model 23a) was ranked first by the bias statistic. Model 20 was the second ranked model, ranking second by SDR, SEE and MAD. However, this model was the most biased model. The value of bias for this model was 0.19826 cm, while the value of bias for model 23a that ranked at the top was 0.0852 cm. It was interesting to observe that bias and MAD ranked some models in opposite ways. Model 23a was ranked first by bias, but last (at number 4) by MAD. Meanwhile, model 20 was ranked fourth by bias, but second by MAD. Both bias and MAD are related to the average value of residuals. MAD, however, is not affected by the minus and plus residuals cancelling each other out, as pointed out by Kozak and Smith (1993). The other two statistics (SDR and SEE), on the other hand, ranked all models in the same way. Both ranked model 24 last at number 4 and model 23b first. The second and third models ranked by SDR were also ranked at number 2 and 3 by SEE, respectively.

After examining the graphs of residuals and fit statistics, it was clear that model 24 was not as precise as the other three models, and it was more biased. However, this model was further examined for estimating diameter at different portions of the stems. The statistics calculated for this test are presented in Table 4.4.

It was clear that model 20 was less biased in predicting lower and upper portions than middle portions of stems. The bias values for the middle portion of relative height between 30% and 60% ranked 0.41 cm - 0.64 cm, while the bias values for the lower and upper portions was around 0.1 cm or less. Models 23b, and 23a were more constant in their predictions. The values of bias for different parts of stems were not substantially different. Two high values of bias were at the lower part and upper stem portions. The highest value of bias for model 23b was 0.48 occurring at the portion of relative height 70% - 80%, but the two portions located on either side of this portion had very small values of bias. The second highest value of bias for this was 0.31 occurring at the lower part of stems (portion of 10% - 20%). Values of bias for the remaining portions by this model were similar and low. Bias from model 23a had the similar pattern to model 23b.

MAD statistics for the lower portion were smaller than the upper portions for all four models. The values of MAD for lower portions (lower than 50%) of all models were less than 1.00 cm, while the most values for upper portions were greater than 1.00 cm. SDR and SEE statistics were similar in their interpretation of residuals. Where the SDR value was high SEE was high. Meanwhile, if SDR was low SEE was also low. Overall all four models had the lower values of SDR and SEE for the lower portions of stem than the upper portions.

4.3.4 Equations 20 and 23b

Equation 20 was ranked top for non-compatible taper equations. It was the most precise and least biased model among non-compatible models. It performed similarly to compatible taper models 23b. Figure 4.5 illustrates these models predicting diameters graphically for different sizes of trees.

Table 4.5: The statistics of Bias, SEE, SDR and MAD for estimating diameter at different portions of stems

		Model 20				Model 23a				Model 23b				Model 24			
Relative height	**n**	**Bias**	**SDR**	**SEE**	**MAD**	**Bias**	**SDR**	**SEE**	**MAD**	**Bias**	**SDR**	**SEE**	**MAD**	**Bias**	**SDR**	**SEE**	**MAD**
<5%	109	0.178	1.052	1.067	0.857	-0.013	1.109	1.130	0.864	-0.005	1.052	1.072	0.856	-0.197	1.141	1.186	0.970
5 - <10%	82	0.140	0.605	0.621	0.274	-0.093	0.701	0.725	0.515	-0.059	0.586	0.604	0.353	-0.139	0.792	0.831	0.607
10 - <20%	56	0.120	0.534	0.547	0.324	0.338	0.790	0.893	0.667	0.312	0.573	0.679	0.432	-0.084	0.642	0.679	0.428
20 - <30%	39	0.138	0.594	0.610	0.500	0.086	0.743	0.791	0.618	0.101	0.600	0.644	0.490	0.005	0.623	0.668	0.481
30 - <40%	34	0.641	1.019	1.209	0.994	0.156	1.143	1.231	0.938	0.178	0.987	1.070	0.783	-0.066	1.011	1.100	0.802
40 - <50%	25	0.414	0.931	1.022	0.816	-0.263	0.868	0.995	0.635	-0.167	0.888	0.990	0.650	-0.059	1.034	1.164	0.702
50 - <60%	42	0.526	1.206	1.318	1.050	0.394	1.191	1.322	1.023	0.290	1.209	1.310	0.973	-0.005	1.235	1.318	0.954
60 - <70	32	-0.233	1.634	1.651	1.295	-0.006	1.608	1.723	1.302	0.072	1.603	1.719	1.285	-0.663	1.719	2.016	1.284
70 - <80%	28	0.005	1.277	1.277	1.008	0.521	1.305	1.527	1.024	0.478	1.279	1.482	1.017	-0.388	1.314	1.520	1.094
80 - <90%	43	0.090	1.740	1.743	1.449	0.018	1.762	1.852	1.440	-0.033	1.736	1.825	1.407	-0.165	1.818	1.946	1.483
90 - <100%	8	0.620	2.002	2.108	1.387	-0.157	1.982	3.039	1.532	-0.145	1.996	3.059	1.567	0.486	2.008	3.880	1.439

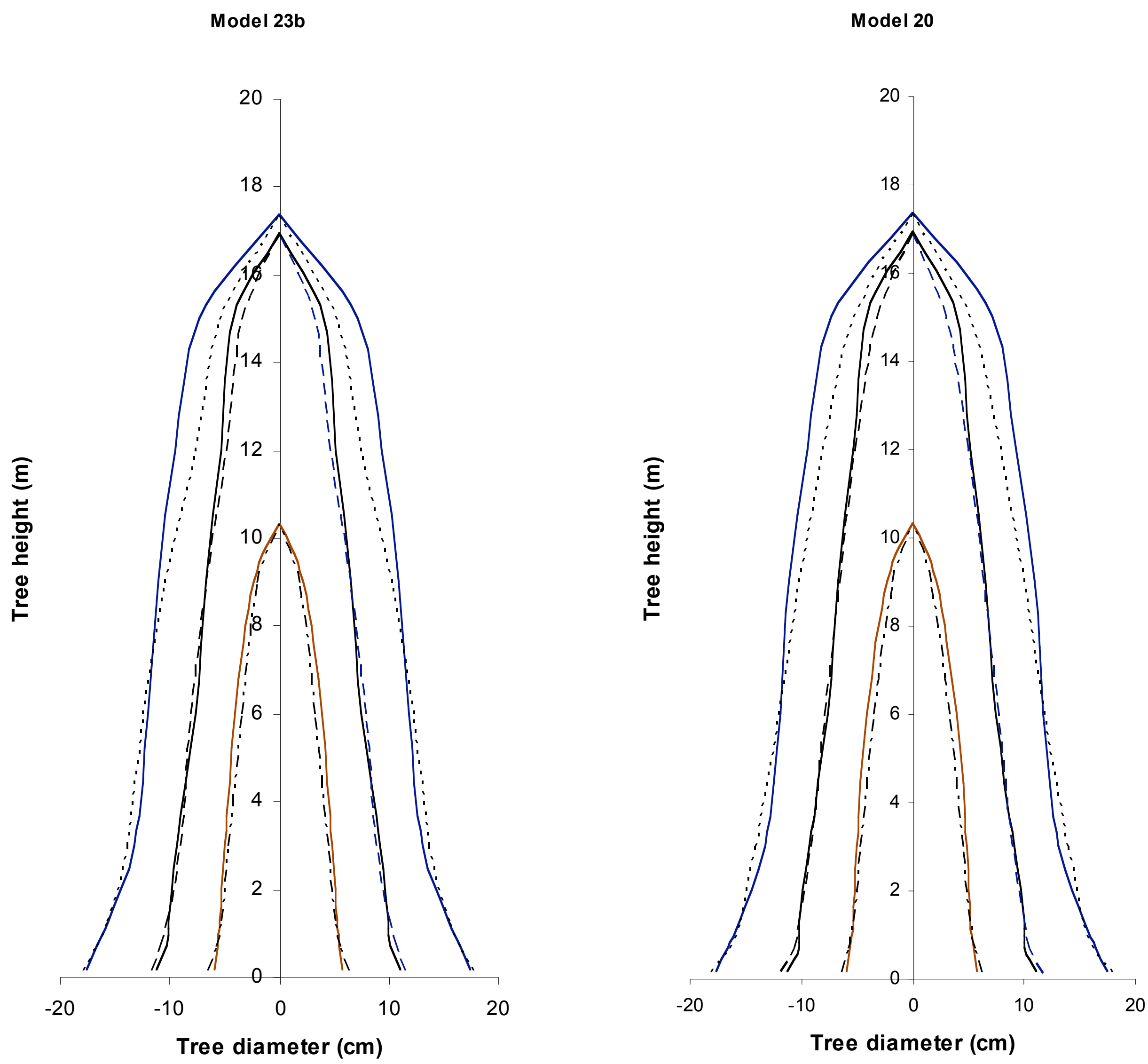


Figure 4.5: Tree profiles generated from observed data and from predicted data (observed data with solid lines and predicted data with broken lines) by equation 23b (left) and 20 (right). Three different sizes of trees were presented, an innermost one with observed D = 10.6 cm and H = 10.31 m, a middle one with observed D = 20.2 cm and H = 16.92 m and an outermost one D = 31.0 cm and H = 17.36 m.

4.3.5 Tree volume generated by taper equation 20

One of the non-compatible taper equations (equation 20) was tested for predicting volume and the results were compared with observed volumes and the volumes predicted by the best volume equation (equation 13a). This taper equation was the best model of non-compatible equations tested. Residual plots of the two models are presented in Figure 4.6. The two models predicted the volumes of trees similarly. Their errors in predictions were almost the same. The patterns of bias in the two residual plots were similar and their precisions were within the same ranges. However, the volumes by these two methods were not exactly equal. These were expected results. Non-compatible taper equations do not normally generate exactly the same volumes of trees as volume equations predict.

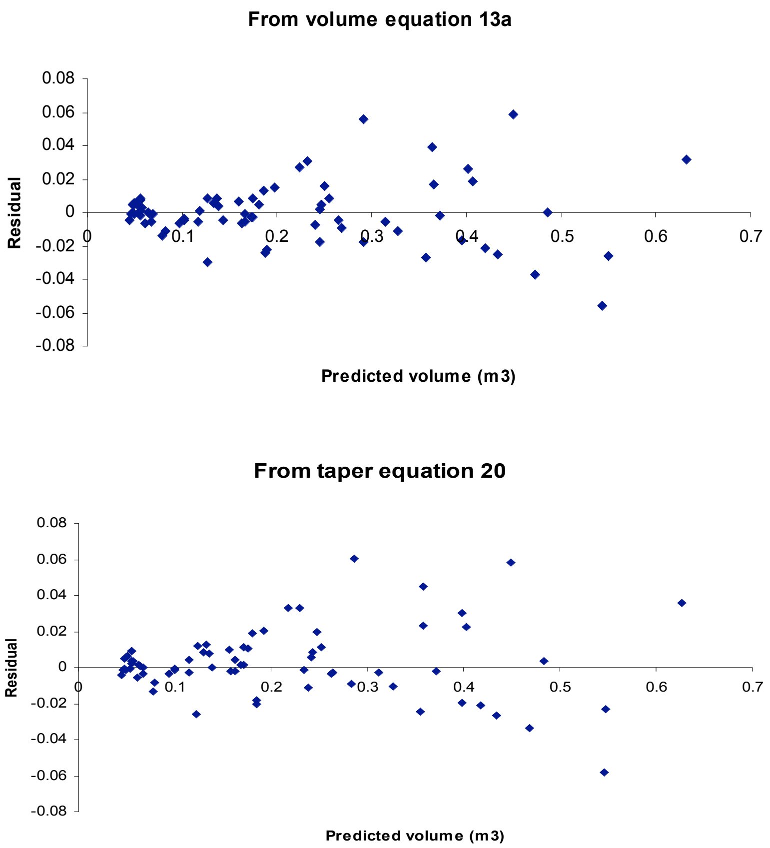


Figure 4.6: Residual plots from two predicted volumes by volume model (above) and taper model (below).

Comparing the two volumes by subtracting one from the other showed that volume equation 13a provided slightly greater values for almost every tree in the data set than taper equation 20 did. The differences between the two volumes were very small. The biggest difference was 0.0081 m^3 and only four data points exceeded 0.006 m^3. Figure 4.7 shows differences between the models. Predicted volume from equation 13a was subtracted by volumes from equation 20. The volumes of a few trees were predicted equally by the two models.

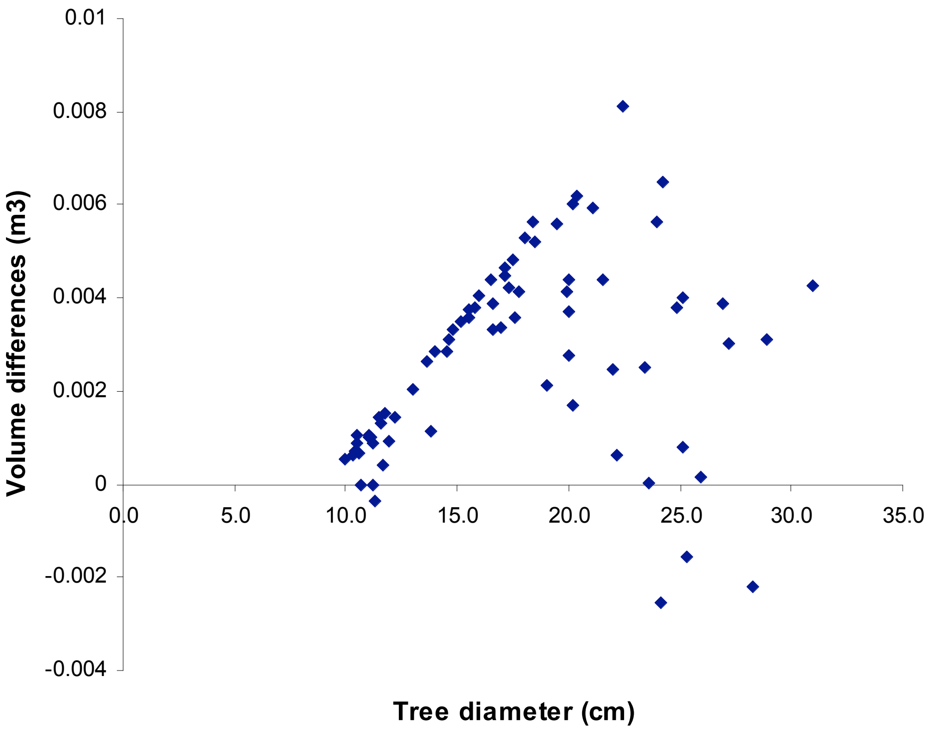


Figure 4.7: Plot of the differences between the predicted volume by volume equation (model 13a), and the volume generated by taper equation (model 20).

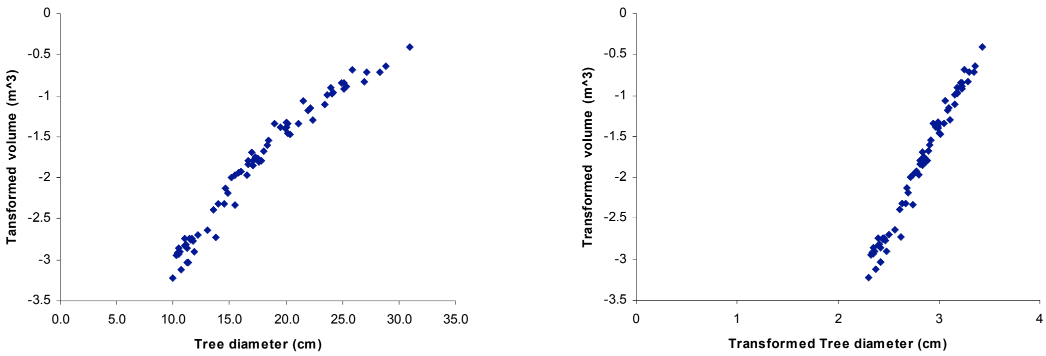
5. Discussion

5.1 Volume equations

5.1.1 Single-entry volume equations

Most of the single-entry models did not perform well with the data set. Models 1, and 2 were the two lowest ranked models amongst all types of volume models tested in this study. Both models are log-linear line volume equations. Apart from common problems of transforming data into logarithms explained by Cunia (1964); and Vanclay and Skovsgaard (1997), the obvious problem with taking logarithms in this study could be explained simply by plotting tree volumes transformed against D or D transformed. In the case of model 1, transforming volume data into logarithms (natural logarithms) did not make the relationship between stem volume and D more linear (Figure 5.1a). Therefore, when the model was fitted by linear regression methods, volume predictions were clearly biased. In the case of model 2, the relationship was improved by transforming both volumes and diameters into natural logarithms (Figure 5.2b). However, when the model was fitted by linear regression, residual analysis showed that it was biased in predicting the volumes of bigger trees (Figure 4.1b).

Figure 5.1: Graphs of stem volume against tree diameter: (a) only volume transformed into natural logarithms; (b) both volume and diameter transformed into natural logarithms.

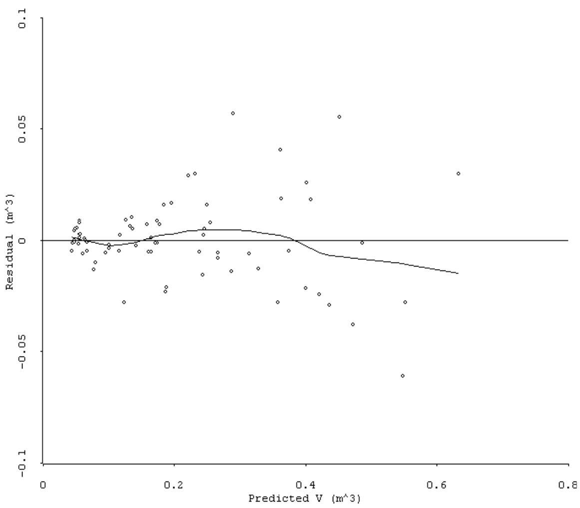


The other two models for single-entry volume equations (models 3, and 4) behaved similarly to each other. However, the performances of these two models were not as precise as some of the double-entry volume models. Most statistics tested, and residual plots did not clearly show differences between the two models (Figure 4.1c, and d; Table 4.3). Bias statistics showed that model 4 was slightly less biased, but the magnitude of the difference in bias statistics was too small to be used to justify the conclusion that model 4 was better than model 3. It implied that the presence of variable D in equation 4 had little effect on the outcome of prediction. However, a t-test showed that this variable was significant at the 5% level. Given that the formula used to calculate the standard error of estimate (SEE) is directly influenced by the numbers of estimated parameters, model 3 could be preferred over model 4 but the results shown in Table 4.3 did not favour model 3 in this study. The SEE of model 3 was not significantly different from that of model 4. This statistic is commonly used for measuring overall predictive value of a model, along with goodness of fit, with low values indicting better fits (Akindele and LeMay 2006). More importantly, these two models were ranked equally at number 7 by overall rank. The performances of the two models are discussed further below where they are compared with other types of volume equations tested in this study.

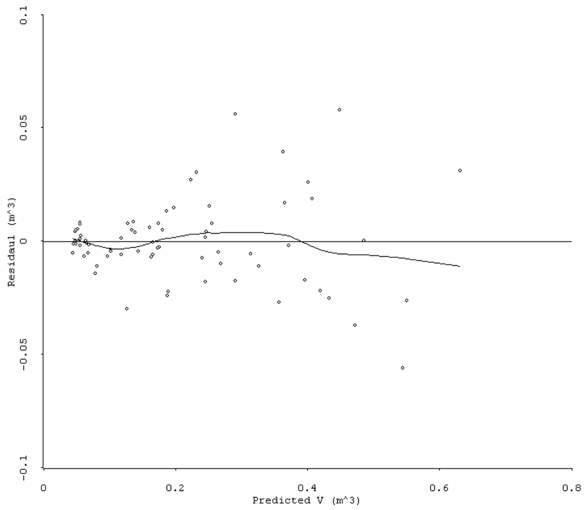
5.1.2 Double-entry volume equations

The combined variable equation of Spurr (1952) and the logarithmic of Schamacher and Hall (1933) are classic volume models and commonly used, often without question, when developing stem volume equations (Bi and Hamilton 1999). Thus statistical analysis to ensure the most appropriate model specification has sometimes been neglected. With this study, while one of these classic models performed considerably well, the other was not very consistent in its prediction. Model 8, a logarithmic form by Schamacher and Hall (1933), performed inconsistently through different classes of sample trees and, thus, it showed more bias than many models tested including two single-entry models (models 3, and 4). Overall it was ranked at number 9 out of 10 final candidate volume equations. Bias statistics ranked it last at number 10 (Table 4.3). Residual plots revealed that, in general, the model was as precise as the other top models; its error was in a similar range to those of the other models (Figure 4.1g). However, this model slightly overestimated the volumes of trees with individual total volumes less than 0.1 m^3. Another classic model (model 6) by Spurr (1952), on the other hand, was moderately precise compared to other models. It was ranked at number 5 overall. This model also slightly overestimated volumes of small trees, but its bias was less obvious than that of model 8. Another interesting volume equation tested in this study is the Honer transformed variable (model 10). It performed well in predicting the volumes of small trees. Even though this model was ranked at number 6, residual plots revealed that it was similar in precision and bias to model 13a (combined variable equation) that was ranked at the top. Figures 5.2a, and 5.2b were residual plots from two models with Loess smoothing lines added.

Figure 5.2: Residual plots with Loess smoothing to show average bias: (a) Honer transformed variable (model 10); (b) combined variable equation (model 13a).



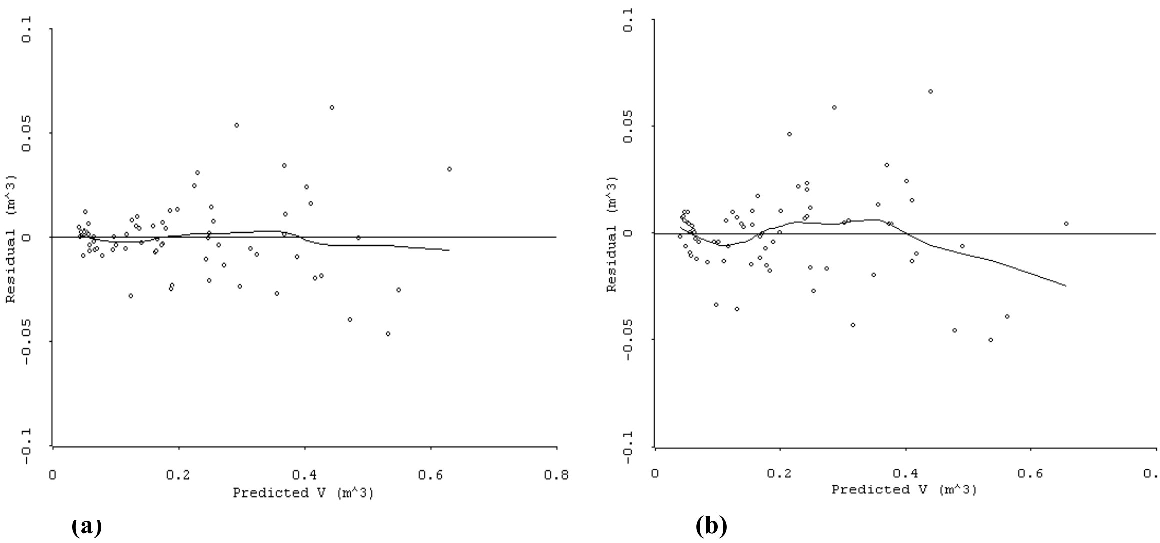
(a)



(b)

Model 16 by Bi and Hamilton (1999) was the most precise and least biased model. It was ranked first by three statistics including SEE, SDR and MAD. Residual plots with Loess smooth lines added also showed that the model was less biased than models 6, and 13a. However, it was quite interesting that the bias statistic ranked this model at number 7 below models 3, and 4. Model 4 was actually more biased than model 16 (Figure 5.3). It was also less precise. One well known problem of the bias statistic used is that it performs ineffectively when many negative and positive residuals cancel each other (Koza and Smith 1993; Muhairwe 1999). In the case of model 4 (Figure 5.3b), many imprecise residuals located at both above and below a horizontal reference line did not equally concentrate along the predicted volume scale. Even though the bias statistic of this model was very low (Table 4.3) with negative and positive residuals cancelling each other out, the model could be still obviously biased in graphical residual plots.

Figure 5.3: Residual plots with Loess smoothing of bias: (a) combined variable volume equation (model 16); (b) single-entry volume equation (model 4).



Model 13a was ranked second to model 16. However, differences between two models were not substantial. Figures 5.2b and 5.3a show that bias patterns of two models were similar, but the Loess smoothing line of model 16 was slightly closer to the zero reference line. For practical purposes, model 16 was more difficult to apply than model 13a as it contained more estimated parameters than model 13a. Consequently, model 13a was considered to be the best model amongst two parameter volume equations tested, and it was selected for further study of compatible volume and taper equation system. Single-entry volume model 4 was also selected for comparison between the two systems of deriving compatible taper equations (see the methods section).

Three main points can be made to sum up the overall findings in this volume section:

- Most models tested performed similarly to each other. Residual analysis showed little differences in magnitudes of bias and precision amongst many models;
- Overall, double-entry volume equations were more precise and less biased than single-entry ones; and
- Overall combined variable volume equations were more consistent in predicting the volumes of different stems. Some of them were among the most precise models, even though many of them were eliminated at the very beginning of model screening because they contained insignificant estimated parameters.

As mentioned in the previous sections, various forms of volume equations have been developed by other researchers. Some of those models were very precise and little biased. For example, Teshome (2005) developed over-bark volume model for *Cupressus lusitanica* in Munessa forest, Ethiopia. His model form is: $V = \beta_0 + \beta_1 \left(\frac{H}{D} \right)^{\beta_2} D^2 H$. It was very precise that error terms were within $\pm 1 \times 10^{-4} m^3$. Compared to his volume model, all the models developed in this study were very imprecise in his study. Another study of over bark volume model was carried by RonDeux and Pauwels (2000). They developed over bark volume model for small trees of larch (Larix sp.) in the southern part of Belgium. The volume equation form that was the best fit to their data was constant form factor $V = b_1 D^2 H$. Unfortunately, they used different residual analysis criteria to evaluate their model (R^2 and CRV, which is the residual standard deviation divided by the mean volume), and it is not compatible with the ones used in the study reported here. Meanwhile, Fowler (1997) developed volume models for red pine species in Michigan, USA using a logarithmic volume equation ($V = \beta_0 D^{\beta_1} H^{\beta_2}$). One model evaluation method used by Fowler

(1997) was not applied in this study. That method was to use an independent data set to validate the model. Even though it is the most effective model evaluation criteria, many studies can not employ it because they do not have enough data.

To conclude the discussion for the volume equation study, it is important to highlight that even though over bark model volume equations have been developed for some species, and are useful for particular cases, they can not provide estimates of timber volumes. For this we need under bark volume models. Some species have a high ratio of bark volume over total volume over bark. For such species bark models are of particular importance. It was very unfortunate that the implementation of field data collection for this study did not go according to plan (see the methodology section). Very few bark data could be collected. The bark thickness model built in this study was built with data from too few trees to be reliable, therefore, it can be regarded as a guide for academic purposes only (see Appendix 1). To model bark thickness effectively, more data are required.

The bark model in Appendix 1 was applied to estimate under bark volumes of small trees with the largest diameter (i.e. at height 0.15 m) were less than 17.5 cm. The range of diameters was within the diameter range of four trees that were used to build the model. Twenty-seven sample trees were classified into the above diameter range, and their under bark volumes were calculated using the bark model. The results are presented in Table A1.2 in Appendix 1. The differences between over- and under bark volumes were around 18 %. The maximum was 18.8 %, while the minimum was 17.6%. Users of the models developed here might be tempted to simply apply a correction factor of 0.82 to estimates of volume to calculate under-bark volume, but it should be noted that bark measurements were available for only four trees, and such a practice may lead to severe bias. Clearly developing a bark model is a crucial future step for the studies reported here.

5.2 Taper equations

5.2.1 Single taper functions

Most taper functions tested in this study were precise in their predictions of over bark diameters of *S. tonkinensis*. Some models, however, showed more bias than the others. Overall, simple single taper functions tested in this study performed poorly. Three single taper functions (models 19, 21, and 22) were eliminated during the first screening test by examining residual plots because they were obvious biased. Only one model of this kind (model 20) was relatively unbiased, and it was ranked at the top among the best taper models of all types. However, these results were not surprising. Single taper functions usually fail to describe entire tree profiles adequately (Max and Burkhart 1976; Jiang *et al.* 2005).

Model 19 by Kozak *et al.* (1969) assumed a constant influence of section height on taper regardless of tree height. It underestimated small diameters but overestimated the bigger diameters (Figure 5.4a). Given that the sizes of diameter at breast height 'D' of samples were between 10 cm and 25 cm, many smaller diameters were likely to be the diameters near the tops of the trees while the bigger diameters were likely to be the diameters near the butts of the trees. The model, therefore, was inadequate for describing two portions at the both ends of the tree bole. However, when residuals were plotted against stem heights, it was found that most biases occurred at the lowest stem portions near the butts. The model was relatively unbiased in predicting the diameters at the top portions of the stems. Thus, the bias at the lower end in the residual plot (Figure 5.4a) could come from the diameters of the lower-middle portions of small trees.

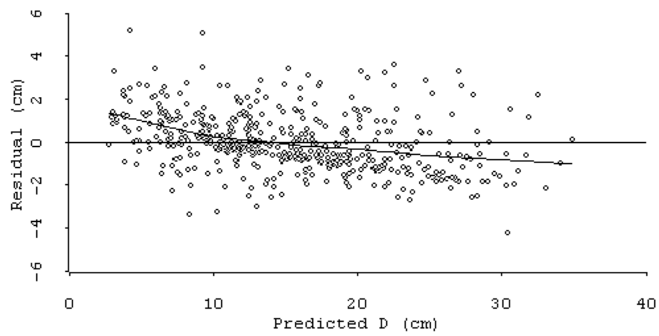


Figure 5.4a: Residual plot with smooth line showing bias pattern (model 19).

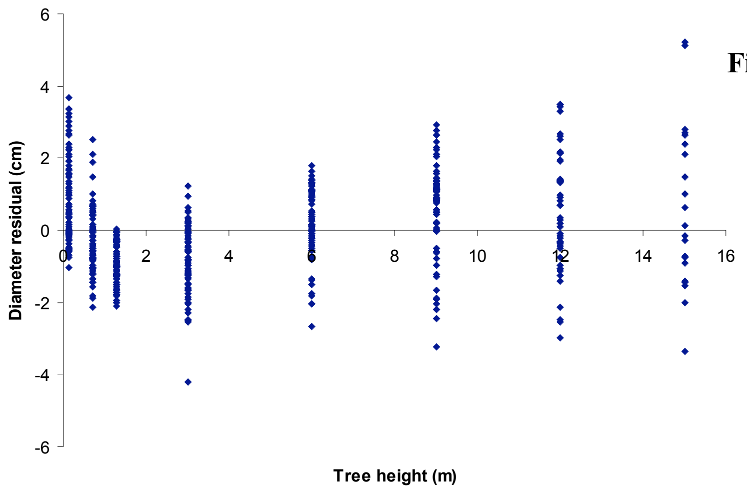


Figure 5.4b: Plot of residuals against stem height

Model 21 by Ormerod (1973), on the other hand, was biased because it underestimated the diameters of all stem sizes. A majority of residuals were located above a horizontal reference line. The results were contrary to those from a study by Reed and Byrne (1985) in which the same taper function was applied to jack pine. In their study, it was presumed that the model overestimated as bias (average residual) was negative. Values of estimated parameter (β_1) in their study, and in this study were different. A value of β_1 less than 0.5 was estimated in their study, while this study estimated it around 0.61. Taper of *S. tonkinensis* may not be related to those of the jack pine, but when β_1 takes on the value less than 0.5 this implies tree stems are cylindrical. This may have resulted in overestimates the diameters of jack pine. When β_1 takes on 0.61 (between a conic and parabolic shape) this may have caused the diameters of *S. tonkinensis* to be underestimated.

Model 22 was similarly precise and biased when compared with model 19, but the bias was less obvious. It slightly underestimated small diameters, but slightly overestimated bigger diameters. Another problem of this polynomial taper function is that it incorrectly predicted the diameters at the top by estimating ‘d’ with some non-zero values at this point.

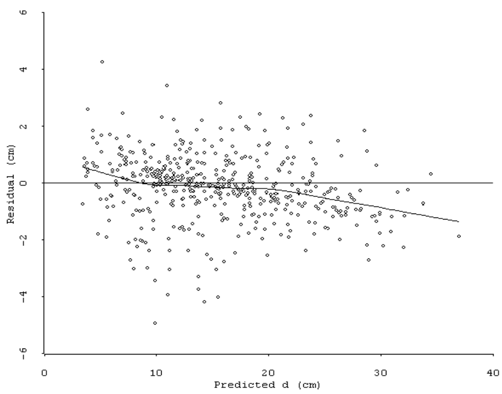
5.2.2 Segmented taper functions

Model 24, a classic segmented taper equation by Max and Burkhart (1976) is of particular interest. It has been tested with many species by a number of people including Methol (2001), Muhairwe (1999), Diéguez-Aranda *et al.* (2006), Jiang *et al.* (2005), Sharma and Oderwald (2001); and Cao *et al.* (1980). The results were mixed. Some found it ranked at the top, while others found it less precise and more biased than other forms of taper functions. Moreover, some studies investigated the model with more than one species and found that it performed differently with each species (Muhairwe 1999). Nevertheless, this model is considered to have many advantages and it is probably the most used segmented taper model (Methol 2001). It is a relatively simple model amongst the segmented models, and is often able describe a tree profile adequately. In comparison with single taper functions, however, estimating the equation coefficients for this model is a complicated computing process. Thus, it is worth describing the procedure for estimating equation coefficients (see Appendix 2).

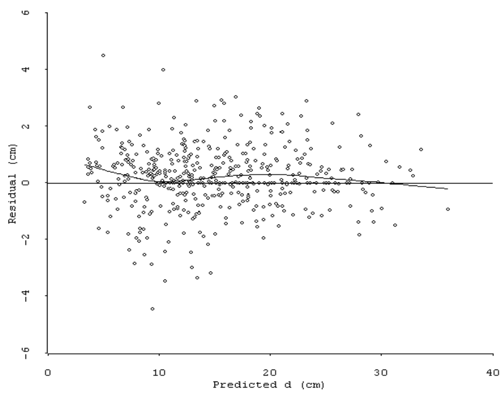
In this study, the model by Max and Burkhart (1976) was ranked below the two models that showed the least bias and most precision in prediction. They include compatible taper models by (Gordon 1983), and model 20 by Sharma and Oderwald (2001). This model was not consistent in predicting all sizes of stem diameters. Graphical residual analysis by Loess smoothing line showed that the bias was very obvious at the upper end along the predicted volume scale line which the model overestimated the diameters (Figure 5.5a). Meanwhile, at the other end, it slightly underestimated the small diameters of the stems.

5.2.3 Model 20

The performance of model 20 was the least biased amongst non-compatible taper equations (Figure 5.5b). However, this model can be developed alongside with volume model to make it a compatible taper equation. The theory and development of such a system were proposed by Sharma and Oderwald (2001). Compatible volume and taper equation systems investigated in this study followed the theory proposed by Gordon (1983). Nevertheless, the most interesting point about this model is that the only estimated parameter (β_1) is very sensitive. Two is a critical value for β_1; if it is equal to 2, the function takes a purely parabolic form to depict a tree shape (Sharma and Oderwald 2001); if it is greater than 2, the function predicts the diameter at the butt much larger than diameter at breast height 'D'; and if it is smaller than 2, the function predicts the diameter at the butt smaller than diameter at breast height 'D' (Figure 5.6). In the last case, one would not expect it to accurately depict a tree's shape.



(a)



(b)

Figure 5.5: Plots of residuals against predicted volumes; (a) for model 24 and (b) for model 20

S. tonkinensis' taper took a shape that meant the β_1 was greater than 2. The value estimated for β_1 in this study was 2.1074 (Table 4.1). This finding is compatible to the result presented by Sharma and Oderwald (2001) in which the value of β_1 was

2.1852 and 2.0056 for over- and under bark taper equations, respectively. Their taper functions were developed for loblolly pine species in the Coastal Plain of North Carolina, USA, the standard errors of estimate for the over bark taper model was 1.38 cm. If only this statistic was used as a criterion for evaluation model quality, the taper equation developed in this study here was more precise than their over bark model. SEE calculated for the model in this study was 1.11 which is lower than their value.

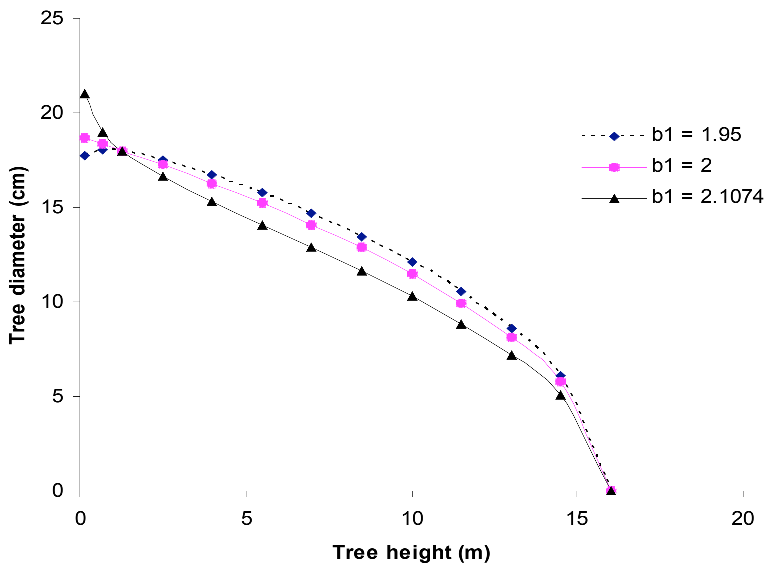
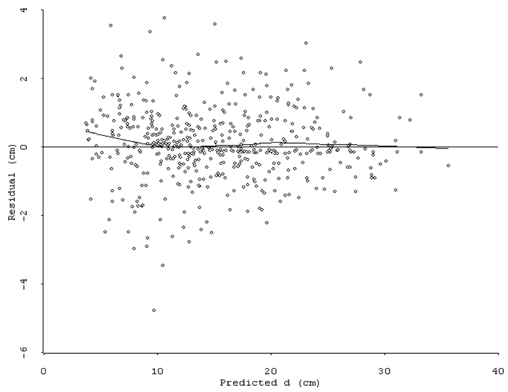


Figure 5.6: Tree profiles generated from model 20 using artificial data to test a function shape with three different values of b_1 (b_1 = 2.1074 was value estimated from the data of the sample trees). Adopted from Sharma and Oderwald (2001)

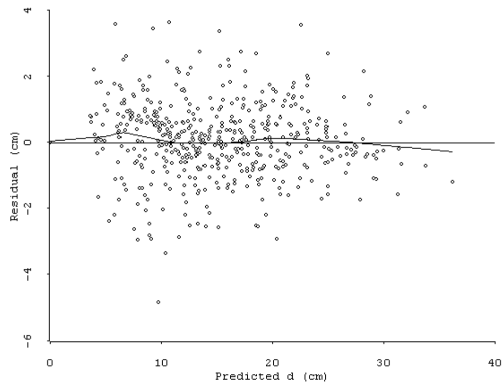
This model not only showed the most precise and least biased in prediction the diameters of all sizes, but also was considerably more precise in prediction of tree volume when integrated to generate stem volume. It is worth describing the procedure of volume generation in detail (see Appendix 3).

5.2.4 Compatible taper functions

Ideally, a volume estimation system should be compatible, i.e., the volume computed by integration of the taper equation from the ground to the top of the tree should be equal to that calculated by a total volume equation (Demaerschalk 1972). However, some studies found that this type of taper equation has been ranked behind many non-compatible taper functions using common precision and bias evaluation criteria (i.e. Cao *et al.* 1980). In addition, Cao *et al.* (1980) pointed out that if the sole purpose was to describe tree taper, the use of compatible taper was not a good option. This study, on the other hand found that a compatible volume and taper equation system was the least biased taper model, even though its precision was not substantially different from the other types of taper models. The residual plot with Loess smoothing lines added showed no obvious bias (Figure 5.7). Using different volume models to derive a compatible taper equation did not significantly affect the model quality, even though bias patterns were not identical. Figure 5.7 compares the residual plots of two compatible taper models that used different volume equations.



(a)



(b)

Figure 5.7: Plots of residuals against predicted volumes; (a) for model 23b and (b) for model 23a

Four main points can be made to sum up the overall findings in this taper section:

- The precisions of most taper models were similar. All models had most residuals distributed between –0.4cm and +0.4 cm (only few data points outside this range);
- Most models tested performed best in prediction of the medium sizes of tree diameters but worst in prediction of the small sizes;
- Compatible taper models were the least biased model. They were relatively consistent in prediction of all diameter sizes; and
- If the sole purpose is to describe tree taper, the best model to use is model 20. It is simple to apply, and was ranked amongst the best taper models tested in this study.

6. Conclusions

6.1 Volume models

Amongst nineteen individual tree volume models tested during the study, two models, each with three estimated parameters, were judged to be best overall when using bias and precision as criteria to predict over-bark volume of *S. tonkinensis*. The first model (model 4) was a single entry model with diameter at breast height as an independent variable, while the second one (model 13a) was a double-entry model that used both diameter at breast height and height as independent variables:

(4) $$V = 0.0062478622 - 0.0048205045D + 0.0008340514D^2$$

(13a) $$V = - 0.0011383579 + 0.000052703D^2H - 0.0000008312D^2H^2$$

Precision and bias of these two models were similar. Errors in their volume predictions were within a similar range (+0.06 m^3/tree and – 0.06m^3/tree), except that one data point from model 4 was 0.066. Models were equivalently effective at predicting the volumes of small trees. Based on the analysis of residuals, three out of four statistics showed that model 13a was a better model for the entire dataset. Those statistics included bias (mean residual), the standard error of estimate (SEE), the standard deviation of the residuals (SDR) and mean absolute deviation (MAD). Model 4 was slightly less biased than model 13a. The differences in each statistic between the two models, however, were not significant.

Other forms and equations tested were inferior to the two models above. They were either too complicated, significantly biased, or less precise.

6.2 Taper equations

Amongst three different forms of taper models tested, a compatible taper model (model 23b) was the most precise and least biased model for describing the over-bark stem profile of *S. tonkinensis*. Residual analysis showed that the compatible taper model was less biased, more accurate and more precise than the other two models tested. The equation for the model is:

$$d^2 = \frac{V(1.66604z_1 - 6.29782z_2 + 16.58655z_3 - 17.77485z_4 + 6.82008z_5)}{kH} \tag{23}$$

where $z_n = (n+1)\left(\frac{H-h}{h}\right)^n$

k = 0.00007854

other notations are as previously defined.

The model had residuals all distributed between –0.4cm and +0.4 cm, except one data point which was –0.48 cm. This was the narrowest range of the errors generated by the three models tested. Graphs of residuals against predicted values showed the compatible taper model was the least biased among the models tested.

The other two types of models (a simple polynomial model and a segmented polynomial model) were more obviously biased. The simple model was the least precise and the most biased model among those tested.

6.3 Compatible taper systems

While two compatible volume systems (model 23a, and 23b) were similar in bias, the compatible volume system using volume equation 13a (model 23b) was more precise than the system using model 4 (model 23a). The system using volume equation 4 also had residuals all but one distributed between –0.4cm and +0.4 cm. Residual plotting, however, showed that this system was less precise as the residuals were less compacted. Fit statistics of residual analyses also indicated that distributions of residuals from the system using model 13a were more precise and less biased than the other system.

References

Akindele, S.O. and V.M. LeMay. 2006. Development of tree volume equations for common timber species in tropical rain forest area of Nigeria. Forest Ecology and Management. 226: 41-48.

Avery, T.E. and H.E. Burkhart. 1994. Forest measurement. 4th ed. McGraw Hill Book Co., New York. 408pp.

Bi HuiQuan, and F. Hamilton. 1998. Stem volume equations for native tree species in southern New South Wales and Victoria. Australian Forestry 61(4): 275-286.

Bruce, D. R.O. Curtis and C. Vancoevering. 1968. Development of a system of taper and volume tables for red alder. Forest Science 14(3): 339-350.

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

Cao, Q.V. and H.E. Burkhart and T.A. Max. 1980. Evaluation of two methods for cubic-volume prediction of loblolly pine to any merchantable limit. Forest Science 26(1): 71-80.

Clark, J.F. 1902. Volume Tables and the bases on which they may be built. Forestry Quarterly 1: 6-11.

Clutter, J.L., J.C. Fortson, L.V. Pienaar, G.H. Brister and R.L. Bailey. 1983. Timber management: A quantitative approach. John Wiley & Sons, USA. 333 pp.

Courbet, F. and F. Houllier. 2002. Modelling the profile and internal structure of tree stem: Application to *Cedrus atlantica* (Manetti). Ann. For. Sci. 59: 63-80.

Cunia, T. 1964. Weighted least squares method and construction of volume tables. Forest Science 10: 180-191.

Diéguez-Aranda, U., F. Castedo-Dorado, J. G. Álvarez-González & A. Rojo. 2006. Comaptible taper function for Scots pine plantations in northwestern Spain. Canadian Journal of Forest Research 36: 1190-1205.

Demaerschalk, J.P. 1972. Converting Volume Equations to Compatible Taper Equations Forest Science 18(3):241-245.

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

Figueiredo-Filho, A., B.E. Borders and K.L. Hitch. 1996. Taper equations for *Pinus taeda* plantations in Southern Brazil. Forest Ecology and Management 83: 39-46.

Fowler, G.W. 1997. Individual tree volume equations for Red Pine in Michigan. North. J. Appl. For. 14(2): 53-58.

Furnival, G.M. 1961. An index for comparing equations used in constructing volume tables. Forest Science 7(4): 337-341.

Gordon, A. 1983. Comparison of compatible polynomial taper equations. New Zealand Journal of Forestry Science. 13 (2): 146-155.

Gordon, A.D., C. Lundgren and E. Hay. 1995. Development of a composite taper equation to predict over-and under-bark diameter and volume of *Eucalyptus saligna* in New Zealand. New Zealand Journal of Forestry Science. 25 (3): 318-327.

Goulding, C.J. and J.C Murray. 1976. Polynomial taper equations that are compatible with tree volume equations. New Zealand Journal of Forestry Science. 5 (3): 313-322.

Gray, H.R. 1956. The form and taper of forest-tree stems. Imp. For. Inst. Pa. No. 32, pp 1–79

Hilt, D.E. 1980. Taper-based system for estimating stem volumes for upland oak. USDA Forest Service. Research Paper. NE-458. 12pp.

Husch, B., C.I. Miller and T.W. Beers. 1972. Forest mensuration. 2nd Edition. Ronald Press, New York. 410 pp.

Jiang, L., J.R. Brooks and J. Wang. 2005. Compatible taper and volume equations for yellow-poplar in West Virginia. Forest Ecology and Management 213: 399-409.

Jøker, D. 2000. Seed leaflet No. 44. Danida Forest Seed Centre.

Kashio, M. and V.J. Dennis. 2001. Monograpgh on benzoin: Balsamic resin from *Styrax* species. Food and Agriculture Organization of the United Nations Regional Office for Asia and the Pacific, Bankok, Thailand.

Kozak, A. 1997. Effects of multicollinearity and autocorrelation on the variable-exponent taper functions. Canadian Journal of Forest Research 27: 619-629.

Kozak, A. and J.H.G. Smith. 1993. Standards for evaluating taper estimation systems. The Forestry Chronicle 69(4): 438-444.

Kozak, A., D.D. Munro and J.H.G. Smith. 1969. Taper functions and their application in forest inventory. The Forestry Chronicle 45(4): 278-283.

Laar, A.V. and A. Akça. 2007. Forest mensuration. Springer. Dordrecht, The Netherlands.

Laasasenaho, J., T. Melkas and S. Alde′n. 2004. Modelling bark thickness of *Picea abies* with taper curves. Forest Ecology and Management 206: 35-47.

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

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

McClure, J.P. and R.L. Czaplewski. 1986. Compatible taper equation for loblolly pine. Canadian Journal of Forest Research 16: 1272-1277.

Meng, C.H. and W.Y. Tsai. 1986. Selections of weights for a weighted regression of tree volume. Canadian Journal of Forest Research 16: 671-673.

Methol, R.J. 2001. Comparisons of approaches to modelling tree taper, stand structure and stand dynamics in forest plantations. Ph.D.thesis. University of Canterbury.298 pp.

Muhairwe, C.K.1999. Taper equations for Eucalyptus pilularis and Eucalyptus grandis for the north coast in New South Wales, Australia. Forest Ecology and Management 113: 251-269.

Newnham, R.M. 1988. A variable-form taper function. Petawawa National Forestry Institute, Forestry Canada, Information Report, PI-X-83.

Newnhan, R. M. 1992. Variable-form taper functions for four Alberta tree species. Canadian Journal of Forest Research 16: 109-114.

Ormerod, D.W.1973. A simple bole model. The Forestry Chronicle 49(3): 136-138.

Philip, M.S. 1994. Measuring tree and forests. 2nd ed. CAB International, Wallingford, UK. 310pp.

Pinyopusarerk, K. 1994. Styrax tonkinensis: Taxonomy, ecology, silviculture and uses. Australian Centre for International Agricultural Research Technical Reprot 31, Canbera.

Reed, D.D. and E.J. Green. 1984. Compatible stem taper and volume ratio equations. Forest Science. 30: 977-990.

Reed, D.D. and J.C. Byrne 1985. A simple, variable form volume estimation system. The Forestry Chronicle 61(2): 87-90.

Rondeux, J. and D Pauwels. 2000. Volume tables for small trees of larch (Larix sp.) in the southern part of Belgium. Forestry 73 (1): 91-93.

Satoshi, Y. 2004. Forest, ethnicity and settlement in the mountainous area of northern Laos. Southeast Asian Studies. 42(2): 132-156.

Schumacher, F.X. and F.S. Hall. 1933. Logarithmic expression of timber-tree volume. Journal of Agriculture Research 47(9): 719-734.

Sharma, M. and R.G. Oderwald. 2001. Dimensionally compatible volume and taper equations. Canadian Journal of Forest Research 31: 797-803.

Spurr, S.H. 1952. Forest Inventory. Ronald Press, New York. 476pp.

Takeda, Shinya. 2004. Management of non-timbre forest products in Laos: Fallow forest growing Styrax tonkinensis for benzoin production. *In* Furukawa, H., M. Nishibuchi, Y. Kono and Y. Kaida (Editors). 2004. Ecological destruction, health, and development. Kyoto University Press, Kyoto.

Teshome, T. 2005. Analysis of individual tree volume equations for *Cupressus lusitanica* Munessa forest, Ethiopia. South African Forestry Journal 203: 72-31.

Vanclay, J. and J.P. Skovsgaard. 1997. Evaluating forest growth models. Ecological Modelling 98: 1-12.

Williamson, M. 1989. Bai Bang Pulp and Paper Mill project, Vietnam. New Zealand Forestry 33(4): 19-20.

Appendix 1: Bark model

Table A1.1: a summary of the bark data

		Bark thickness statistics			
Stem height 'd' (m)	***n***	**Mean**	**Max (cm)**	**Min (cm)**	**SD**
0.15	4	1.27	1.4	1.2	0.096
0.70	4	1.25	1.3	1.2	0.058
1.30	4	1.15	1.3	1.0	0.129
3	3	0.93	1.0	0.9	0.058
6	3	0.67	0.8	0.6	0.115
9	3	0.57	0.6	0.5	0.058
12	1	0.5	0.5	0.5	-
15					

Note: All trees were less than 15 m in height, and their diameters at breast height and the largest diameter (i.e. at 1.5 m height) were between 14 cm and 17.4 cm,.

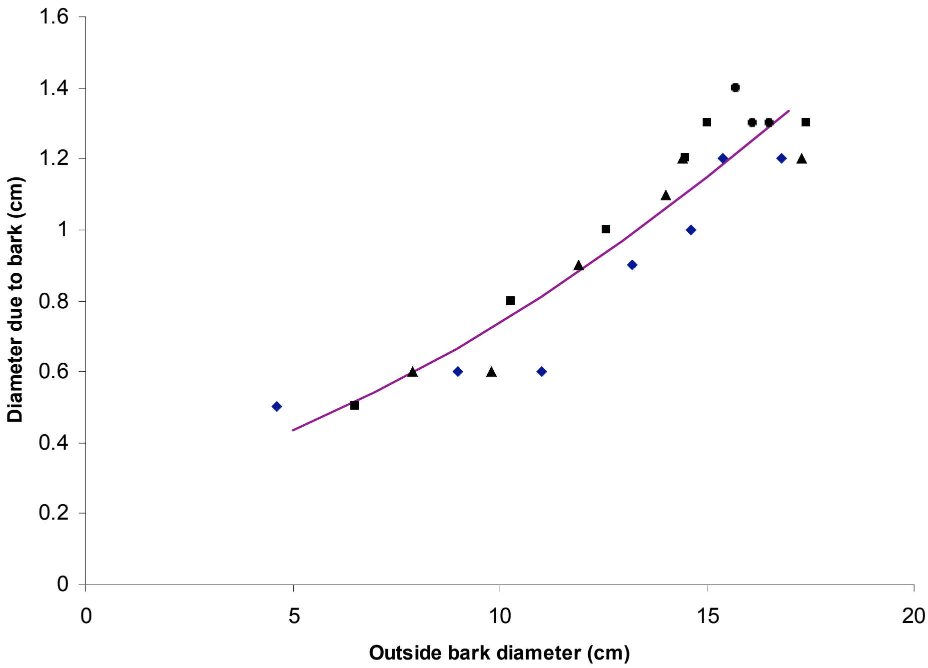


Figure A1.1: Diameter due to bark (sum of the two bark measurements) versus over-bark diameter for four trees. The four trees are represented by different symbols. The line shows a non-linear model between the two variables.

Note: bark model was $Y = 0.2814 + 0.0115d^{1.5944}$

where Y, d are 2 x bark thickness and stem diameter of the stem respectively.

As a guide to the errors associated with using over-bark models, the bark thickness model developed from the data shown in Figure A1.1 were applied to trees in the sample that had diameters at 0.15 m less than 17.5 cm (Table A1.2).

Table A1.2: Over-bark and under-bark volumes of small trees with the largest diameter (i.e. at height 0.15 m) were less than 17.5 cm.

No.	Tree height(m)	DBH (cm)	Over-bark volume (m^3)	Under-bark volume (m^3)	% of difference
1	11.54	11.8	0.062907982	0.053500204	17.6
2	9.31	11.7	0.064539862	0.054762586	17.9
3	10.5	10	0.04012935	0.034037713	17.9
4	10.87	11	0.06406858	0.054295496	18.0
5	8.43	11.3	0.047904764	0.040597022	18.0
6	10.31	10.6	0.054851933	0.046464369	18.1
7	11.34	13	0.071032231	0.060169775	18.1
8	11.04	10.5	0.057347186	0.048576317	18.1
9	12.54	14	0.098325641	0.083252027	18.1
10	10.98	11	0.059347956	0.050244092	18.1
11	10.82	11.1	0.05978593	0.05061093	18.1
12	11.96	11.5	0.064747655	0.054810591	18.1
13	12.59	13.6	0.09086216	0.076902001	18.2
14	10.55	10.3	0.052431591	0.044357097	18.2
15	11.09	11.6	0.063972373	0.054113373	18.2
16	10.39	11.2	0.057302104	0.04847012	18.2
17	11.96	10.5	0.053382158	0.045153017	18.2
18	10.62	10.4	0.054431203	0.045986897	18.4
19	13.5	14.6	0.119552602	0.100981038	18.4
20	9.02	10.7	0.044286018	0.037403354	18.4
21	12.84	16	0.145561749	0.122920746	18.4
22	8.85	11.2	0.047994032	0.040508266	18.5
23	11.4	14.5	0.098948584	0.083496379	18.5
24	10.73	12.2	0.067459525	0.056895726	18.6
25	10	11.9	0.05503014	0.046394513	18.6
26	12.6	16.5	0.139353982	0.117396593	18.7
27	9.43	13.8	0.065214131	0.054902858	18.8

Note: percentages of differences were calculated by using equation A1.1.

$$p = \frac{V_{over} - V_{under}}{V_{under}} \times 100 \qquad \text{(A1.1)}$$

where p, V_{over}, and V_{under} are the percentage, volume over bark and volume under bark, respectively.

A plot of error versus diameter at breast height (DBH) is shown in Figure A1.2. The error averaged 18 %. While this figure might be used as an interim guide to the likely magnitude of errors associated with estimating under-bark volumes for small trees using the models developed during the study reported here, it should be noted that bark measurements came from only 4 trees. Any under-bark volume estimates derived by combining the over-bark volume model with this adjustment factor should therefore be regarded as potentially biased. A wider ranging study of bark thickness of *S. tonkinensis* is urgently required.

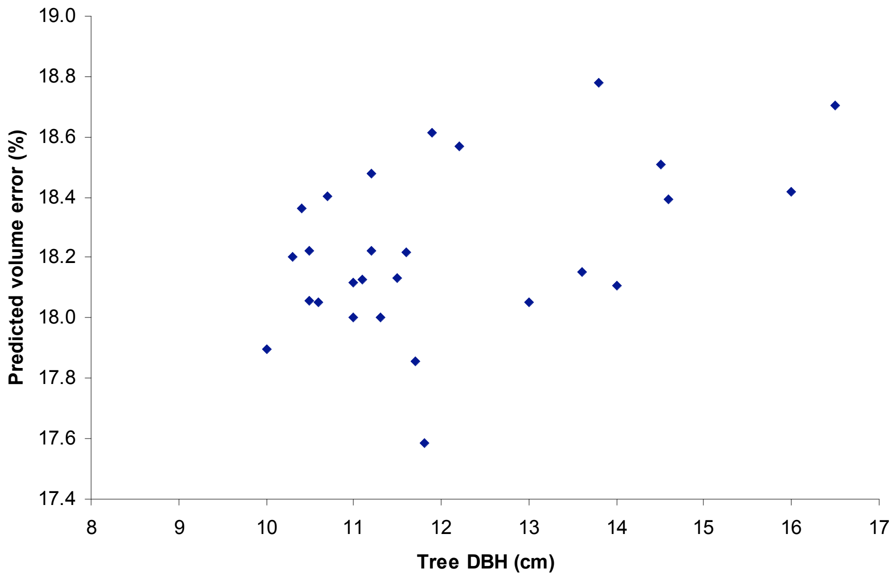


Figure A1.2: Plot of predicted volume errors against tree DBH

Courbet & Houllier (2002) implied a model of double bark thickness of *Cedrus atlantica* by modelling the ratio of over-bark to under-bark diameter as a function of distance from tree apex:

$$\frac{D_o}{D_i} = c_1 + \frac{c_2}{x^{c_3}}$$

where D_o = over-bark diameter, D_i = under-bark diameter, X= distance from stem apex, and c_1, c_2 and c_3 were coefficients.

Their model implies an almost linear increase in bark thickness with diameter when it is applied to estimates of diameter from a fitted taper equation. When an inverted form of this equation was applied to estimate bark thickness among trees where bark was measured during the studies reported here, a plot of predicted versus residual values indicated that although the model was biased, a simple change to one or more of the parameters might enable it to fit data from *S. tonkinensis* (Figure A1.3). It is therefore a candidate for future studies of bark thickness.

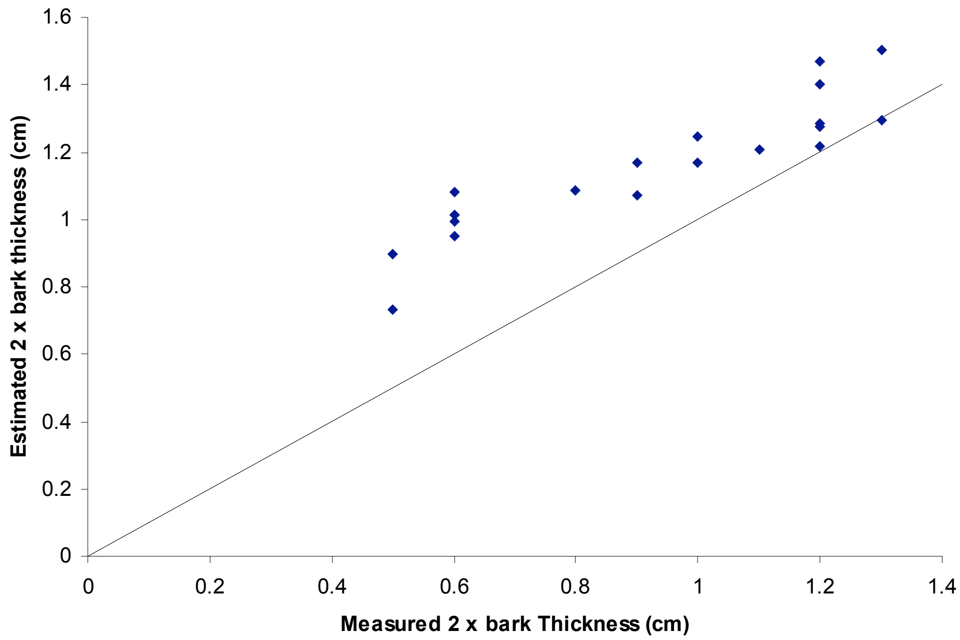


Figure A1.3: Estimated 2 x bark thickness using Courbet and Houlliet's (2002) model versus actual 2 x bark thickness. The line shows a one to one correspondence.

Using a bark thickness model derived from measurements of *S. tonkinensis* along with the over-bark volume and taper models created during the study reported here will enable estimates of under-bark volume, but at a cost when compared to direct models of under-bark volume and taper. The extra cost will arise from combining imprecision of over-bark models with imprecision of the bark thickness model. When taper and volume studies are done in future for this species, measurements of bark thickness on the sampled trees will allow us to directly model under-bark volume and taper.

Appendix 2: **T**he procedure for estimating equation coefficients for segmented taper model (model 24) by Max and Burkhart (1976).

The list below provides the complete steps and the code that used for estimating equation coefficients for Max and Burkhart taper function in this study:

```
libname avolume 'p:\thesis\data analysis\input data' ;
data temp;
set avolume.table_combinedv2

    d = diaob ;
    ht = topheight ;
    h = height ;
    y = (d/dbh)**2 ;
    x = h/ht ;
    x1 = x – 1 ;
    x2 = x*x – 1 ;

proc nlin ;
    parms b1 = 0.002 b2 = - 0.002 b3 = 0.05  b4 =35 a1 = 0.95 a2 = 0.05 ;
    if x ge a1 then do ;
model y = b1*x1 + b2*x2 ;
    der.b1 = x1 ;
    der.b2 = x2 ;
    der.b3 = 0  ;
    der.b4 = 0  ;
    der.a1 = 0 ;
    der.a2 = 0 ;
    end ;
    else do ;
    if a2 le x lt a1 then do ;
model y = b1*x1 + b2*x2 + b3*(a1 - x)**2 ;
    der.b1 = x1 ;
    der.b2 = x2 ;
    der.b3 = (a1 - x)**2  ;
    der.b4 = 0  ;
    der.a1 = 2*b3*(a1 - x)  ;
    der.a2 = 0 ;
    end ;
    else do ;
model y = b1*x1 + b2*x2 + b3*(a1 - x)**2 + b4*(a2-x)**2  ;
    der.b1 = x1 ;
    der.b2 = x2 ;
    der.b3 = (a1 - x)**2  ;
    der.b4 = (a2 - x)**2  ;
    der.a1 = 2*b3*(a1 - x)  ;
    der.a2 = 2*b4*(a2 - x)  ;

    output out = stats r = resid p = pred;
    end;  end ;

proc gplot ;
    plot resid * pred/ vref = 0 ;

run ;
```

Appendix 3: **T**he procedure for generating stem volume from taper function

Taper function (model 20):

$$d^2 = D^2\left(\frac{h}{h_D}\right)^{2-\beta_1}\left(\frac{H-h}{H-h_D}\right),$$ where $b_1 = 2.1074$; and $h_D = 1.3$ m

Thus, $$d^2 = D^2\left(\frac{h}{1.3}\right)^{-0.1074}\left(\frac{H-h}{H-1.3}\right) = D^2\left(\frac{h^{-0.1074}}{0.972215373}\right)\left(\frac{H-h}{H-1.3}\right)$$

Solving the above function, so that:

$$d^2 = D^2\left(\frac{Hh^{-0.1074}}{0.972215373H-1.2638}\right)-\left(\frac{h^{0.8926}}{0.972215373H-1.26388}\right)$$

If the function is rotated around the X-axis and the frustra is generated.

Because

$$V = k\int_0^H d^2\,dh$$

Replacing d^2 with the function:

$$V = kD^2\int_o^H \frac{Hh^{-0.1074}}{0.972215373H-1.2638}-\frac{h^{0.8926}}{0.972215373H-1.26388}dh$$

Integrating

$$V = kD^2\left(\frac{Hh^{0.8926}}{0.867799466H-1.128139288}-\frac{h^{1.8926}}{1.840015434H-2.392019288}\right)$$