A Comparative study of log volume estimation by using statistical method

Siti Hajar Mohd Mushar[1], Sharifah Sakinah Syed Ahmad [2*], Nur Hajar Zamah Shari[3], Fauziah Kasmin[4]

[1,2,4] Department of Intelligent Computing and Analytics,
Faculty of Information and Communication Technology, Technical University of Malaysia
Malacca (UTeM), Durian Tunggal, Malaysia
[3] Forestry and Environment Division, Forest Research Institute Malaysia (FRIM),
Kepong, Malaysia

Corresponding author: sakinah@utem.edu.my

Abstract

Log volume estimation is a measurement of the amount of merchantable volume and precise estimation of log volume plays an important role in sustainable forest management. There are several log volume formula which commonly used in estimating the log volume, however, there are significant differences between each formula. Therefore, this paper evaluates the performance of three different log volume formula which are Smalian's, Huber's and Bruce's formula against several log sectional length. The performance of each log volume formula will be evaluated in terms of the bias, precision and accuracy of the estimation. The result shows that Huber's formula performs the best for log sectional length of 2 m, 4 m, 6 m and 8 m log sectional length. The log sectional length and prediction accuracy is inversely related whereby the shorter the log sectional length, the better the prediction accuracy is.

Keywords: Empirical formula, review, tree stem volume, volume table

INTRODUCTION

Previous study by [1] highlighted that there are several ways that are able to evaluate the feasibility of forests and part of it is the evaluation through the value of its volume, biomass and net primary production (NPP). Literally, volume of wood is the quantification of either an individual log, tree or group of trees in cubic value [2]. The urge in having a precise value of tree volume is highlighted by [3] where it plays a vital role in the forest management and administration. Besides, it is fundamental in designing a sustainable forest management plan and also has the capability in controlling the decision-making process related to the forest trade in activities [4, 5].

There are various methods that are used by researchers in estimating the log volume and log formula can be considered as one of the commonly-used methods by the researcher. It is highlighted by [6] that the selection of the log formula is influenced by several factors which are the comprehensibility in adopting the formula and the availability of tree sample dataset. Besides that, the researcher also looking at the precision and accuracy level of a particular log formula.

This study therefore set out to assess the effect of log sectional length on the accuracy of log volume estimation by using Smalian's, Huber's and Bruce's formula by using the statistical method.

BACKGROUND OF STUDY

Lists of log formulas can be found in the report of [7] whereby the listed formula is not only limited to the United States and Canada but other countries as well. It has been ruled out by [6] that the selection of the formula is influenced by the comprehensibility in adopting the formula and the availability of tree sample

dataset. Therefore, this study involves three types of log volume formula which are Smalian's, Huber's and Bruce's formula and the performance for each of the formulae in estimating the log volume by using different log sectional length is evaluated. The equation for these three formulae is tabulated in Table 1.

Table 1: Log Formula Used in Volume Computation

Log formula	Equation
Smalian's formula	$V_i = \left(\frac{1}{2}\right) * f * L * (dA^2 + dB^2)$ (1)
Huber's formula	$V_i = f * L * (dM)^2$ (2)
Bruce's formula	$V_i = f * L * (0.25(dB)^2 + 0.75\,(dA)^2)$ (3)

where:

$V_i =$ Log volume at ith (m³)
$L =$ Length of log (m)
$f =$ 0.00007854 (metric units)
$dA =$ Diameter at the small log end (cm)
$dB =$ Diameter at the large log end (cm)
$dM =$ Diameter at the mid-length log end (cm)

In brief, Smalian's formula is a formula that developed by Heinrich Ludwig Smalian in 1837 and it uses cross-sectional area at both logs ends in the calculation of log volume. Method of measurement for the application of Smalian's formula as shown in.

Figure 1. [8] as cited ([9]) claimed that Smalian's formula is widely-used in the USA. In an article written by [9], Smalian's formula has received considerable attention from the scholars in estimating the volume due to the simplicity in its adoption, especially for a condition where there is an absence of mid-length measurement due to the log stacked pile condition. Likewise, [10] asserts that this technique frequently used by log scalers in Canada due to that reason.
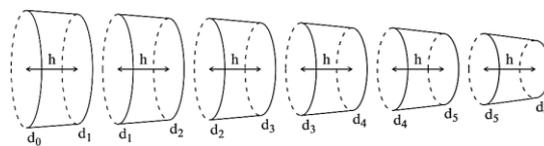

Figure 1: Tree data measurement of Smalian's formula (Jr. et al., 2016)

Meanwhile, volume estimation by using the average cross-sectional area of the tree is known as Huber's formula. Much of the available literature on Huber states that this method is the commonly-used method in Europe and some other tropical countries as well [8 ;11]. The measurement methods for volume estimation by using Huber's formula as shown in figure below.
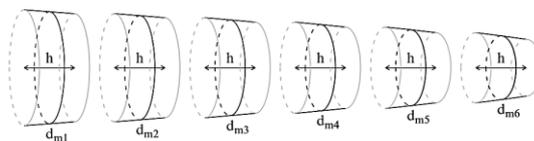

Figure 2: Tree data measurement of Huber's formula (Jr. et al., 2016)

Bruce's formula, a widely-used method in the Western United States is introduced by Bruce David in 1982 [11]. This formula takes into account both log ends which seems similar to that of Smalian's formula but

some modification is performed towards the original Smalian's formula by adding some coefficient in the calculation of its both log ends for better butt log volume estimation [5]. The measurement methods for volume estimation by using Bruce's formula is similar with Smalian's formula (referring to the **Figure 1**).

These three log formula will be compared to the Newton's formula which treats as "true" in this study. The complexity of the Newton's formula has made it hold the title as the best method. This is because, as can be seen in the equation below, it requires three kinds of diameter measurement which are small, midlength and large end diameter. However, in a condition where the number of diameter is even, the last interval need to be calculated by using Smalian's formula and then it is summed up with previously calculated volume by using Newton's formula [12]. Several previous studies like [13, 5] test the performance of Newton's formula with several other log volume formula found that Newton's is the most precise and accurate in estimating the volume. Therefore, the volume equation of Newton's formula is as below:

$$V_i = \left(\frac{1}{6}\right) * f * L * (dA^2 + 4dM^2 + dB^2) \qquad (4)$$

where:

$V_i =$		Log volume at ith (m³)
$L\ =$		Length of log (m)
$f\ =$		0.00007854 (metric units)
$dA =$		Diameter at the small log end (cm)
$dB =$		Diameter at the large log end (cm)
$dM =$		Diameter at the mid-length log end (cm)

MATERIAL AND METHODS

Data

This study adopts selective sampling methods. The number of trees is selected based on diameter classes and species group. The real tree data is collected at the compartment 37, Cherul Forest Reserve, Terengganu and the measurement were carried out on both, the standing and felled trees. Trees that are dead, dying, broken or even the bad form was excluded from measurements. Similarly, trees of a diameter at breast height (DBH) less than 15 cm. For the standing trees, handheld Laser Criterion 400 is used as an instrument to measure the diameter and height of the tree. Meanwhile, for the measurement on felled trees were taken by the forest officer by using the caliper and measuring tape before the log is transferred to the logger's lorry. Measurements on the diameter of each log section, merchantable length, total tree length and diameter at breast height (DBH) is taken and recorded.

The total tree sample data collected is 265, however, after going through the preprocessing procedure, only 241 tree sample data is used in this study. In short, the data preprocessing is important because the real world data is not clean and would have noise, inconsistent and sometimes incomplete due to human or instrumental error. Thus, it is highly important to ensure the use of quality data for a quality output.

Accuracy Assessment

Newton's formula has been recognized as the most accurate formula in calculating the tree volume as stated by [14, 15, 16]. Therefore, Newton's formula will set as the 'true volume' and the accuracy of Smalian's,

Bruce's and Huber's formula was evaluated by using several statistical performance measures. The performance of each formula will be evaluated in terms of its bias, precision and accuracy. The best log formula is the one that has the least biased, most accurate and precise in its prediction.

Bias is the average deviation between the estimated volume and true volume which obtained by using three log formula and Newton's formula. In the mensurational book written by [12], the occurrence of bias may due to the faulty measurement procedures, instrumental errors, flaws in the sampling procedure, errors in the computations or even mistakes in the recording. [17] classified the mean error (ME) as the common bias measure. Therefore, the bias is estimated by using the following formula:

$$Bias = \frac{1}{n}\sum_{i=1}^{n}(V_{mi} - V_{ni}) \qquad (5)$$

where:

$V_{mi} =$ Estimated log volume at ith (m³)
$V_{ni} =$ True log volume at ith (m³)
$n \quad =$ Number of sample

Precision is defined by [17] as the measure of consistency of the prediction and the precision may be affected by the limitations in the measurement or estimation technique which been used at different times and under several circumstances [12]. There are several statistical criteria that are considered by [12] as a precision measurement and one of it is the standard deviation (SD). The precision is estimated by using the following formula:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(V_i - \bar{V})^2}{(n-1)}} \qquad (6)$$

where:

$V_i =$ Log volume at ith (m³)
$\bar{V} =$ Mean volume (m³)
$n =$ Number of sample

Mean square error (MSE) is the most common accuracy measure because it comprises a combination concept of bias and precision. This can be seen where an accurate estimator would have a small value of variance and the least bias in the prediction. In short, MSE measures the closeness level of the estimation from the true value. The accuracy is estimated by using the following formula:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(V_i - \bar{V})^2 \qquad (7)$$

where:

$V_i =$ Log volume at ith (m³)
$\bar{V} =$ Mean volume (m³)
$n \quad =$ Number of sample

RESULTS AND DISCUSSION

Table 2 presents the result of bias, mean square error (MSE) and standard deviation (SD) of each log sectional length of 2 m, 4 m, 6 m and 8 m by using three log volume formula.

Table 2: Bias, mean square error (MSE) and standard deviation (SD) for different log sectional length.

	Bias (m³)	MSE (m³)	SD (m³)
2 m log interval			
Smalian	0.02225	0.00342	0.17064
Huber	**0.00989**	**0.00227**	0.17016
Bruce	-0.07535	0.01144	**0.16779**
4 m log interval			
Smalian	0.06658	0.01462	0.16508
Huber	**0.03844**	**0.00596**	0.16427
Bruce	-0.12524	0.03029	**0.15932**
6 m log interval			
Smalian	0.03351	0.65584	0.17060
Huber	**-0.01445**	**0.63849**	0.16889
Bruce	-0.24681	0.71479	**0.16176**
8 m log interval			
Smalian	0.15426	0.07258	0.16037
Huber	**0.08107**	**0.02349**	0.15734
Bruce	-0.22145	0.08319	**0.14763**

Based on

Table **2**, it can be clearly seen that Huber's formula is the least bias for all log sectional length. To be precise, the negative value of bias 6 m log sectional length indicates an underestimation in the log volume by 0.01445. Meanwhile, an overestimation in prediction is detected for 2 m, 4 m and 8 m log sectional length by 0.00989, 0.03844 and 0.08107 respectively.

Overall, the bias for all log sectional length is very low and can be considered as bias-free. Graphical representation of mean square error (MSE) in Figure 3 clearly shows that Smalian's and Huber's is almost the same for 2 m, 4 m and 8 m of log sectional length, while Huber's is the least for all log sectional length. This is supported by numerical values in Table 2 which shows that Huber's formula recorded the lowest mean square error (MSE) value for all log sectional length which followed by Smalian's and Bruce's. Mean square error (MSE) gives an indication of the prediction accuracy. Therefore, the lower the value of MSE, the higher the accuracy is.
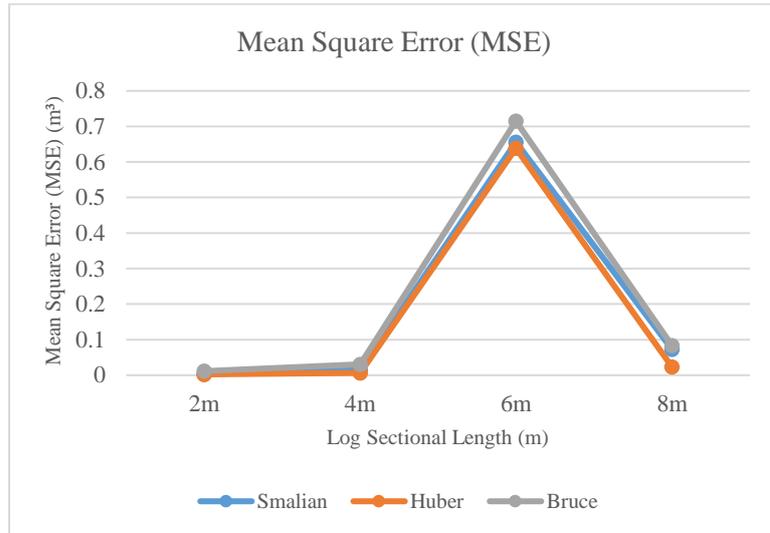
Figure 3: The variation of mean square error (MSE) for each log section.

Furthermore, Bruce's formula shows the lowest value of standard deviation (SD) for all log sectional length. A standard deviation measures the dispersion of data and indicates precision in the estimation. Hence, it can be said that Bruce's formula is the most precise method. However, higher in precision does not indicate a high in accuracy and *vice versa*. In this study, this applies to Bruce's formula and Smalian's formula. This is because Bruce's formula shows the highest value of MSE for all log sectional length. Not only that, Bruce's formula also reported the highest bias for all log sectional length. For the case of Smalian's formula, it records the highest value of standard deviation (SD) for all log sectional length tested in this study.

[17] identify the relationship between accuracy, bias and precision where, high prediction accuracy is contributed either by the least value of biasness or least value of variance or even both. The relationship between mean square error (MSE) with bias and variance has been proven in the previous study by Casella and Berger (1990) (as cited in [17]) where: $MSE = variance + bias^2$. This can be seen in Table **2** where, the log formula that shows the lowest value of mean square error (MSE) for each of the log sectional length, the biasness is also the least. In this study, this is referring to the Bruce's formula as it records the lowest value of MSE and bias for log sectional length of 2 m, 4 m, 6 m and 8 m.

CONCLUSION

A comparative study on the performance of Smalian's, Bruce's and Huber's formula in estimating the log volume is evaluated in terms of the bias, precision and accuracy towards volume estimation of several log sectional length. The result shows that Huber's formula is the least bias for all log sectional length while Bruce's formula is the most precise one as it records the lowest value of standard deviation for all log sectional length. In terms of the mean square error (MSE), Huber's formula record the lowest value for all log sectional length. Therefore, Huber's formula perform the best as it records the least bias and lowest value of MSE. For each log volume formula, the accuracy deteriorates as the log sectional length increases. Therefore, it can be concluded that the use of 2 m of log sectional length is recommended for the best result of log volume estimation and Huber's formula is the best.

REFERENCES

[1]   Van Tuyl, S., Law, B.E., Turner, D.P., and Gitelman, A.I., 2005. Variability in Net Primary Production and Carbon Storage in Biomass across Oregon Forests - An Assessment Integrating Data from Forest Inventories, Intensive Sites, and Remote Sensing. *Forest Ecol Manag*, 209 (3), pp.273–291.

[2]   Food and Agriculture Organization of the United Nations (FAO), 1997. *State of the World's Forests*. Rome, Italy.

[3]   Shari, N.H.Z., Ahmad, W.M.S.W., Musa, S., Mohd., W.R.W., and Ismail, H., 2010. Development of Local Volume Table for Second Growth Forests Using Standing Tree Measurements. *Malaysian Forester*, 73 (2), pp.163–170.

[4]   Davis, L.S., Johnson, K.N., Bettinger, P., and Howard, T.E., 2001. *Forest Management: To Sustain Ecological, Economic, and Social Values*. 4th ed. Long Grove, USA: Waveland Pr Inc.

[5]   Li, C., Barclay, H., Hans, H., and Sidders, D., 2015. *Estimation of Log Volumes: A Comparative Study*. Canada.

[6]   Alemdag, I.S., 1978. *An Analytical Look at the Log Volume Formulas*. Canada.

[7]   Freese, F., 1973. *A Collection of Log Rules*. U.S.D.A. Forest Service General Technical Report FPL 1. Madison, Wisconsin.

[8]   Grave, H.S., 1906. *Forest Mensuration*. New York: John Wiley & Sons, Inc.

[9]   León, G.C. de and Uranga-Valencia, L.P., 2013. Theoretical Evolution of Huber and Smalian Methods Applied to Tree Stem Classical Geometries. *Bosque*, 34 (3).

[10]   Ormerod, D.W., 1985. Scaling and Smalian's Formula: Time to Change? *Forest Chron*, 61 (1), pp.28–30.

[11]   Ozçelik, R., JR., H.V.W., and Brooks, J.R., 2006. Estimating Log Volumes of Three Tree Species in Turkey by Six Formulae. *Forest Prod J Journal*, 56 (November/December), pp.84–86.

[12]   Jr., J.A.K., Ducey, M.J., Beers, T.W., and Husch, B., 2016. Determination of Tree Volume, Weight and Biomass. In: *Forest Mensuration*. Wiley-Blackwell, pp.135–183.

[13]   Fonweban, J.N., 1997. Effect of Log Formula, Log Length and Method of Measurement on the Accuracy of Volume Estimates for Three Tropical Timber Species in Cameroon. *Commonw Forest Rev*, 76 (2), pp.114–120.

[14]   Chapman, H.H. and Meyer, W.H., 1949. *Forest Mensuration*. New York: McGraw Hill Book Co.

[15]   Loetsch, F., Zöhrer, F., and Haller, K.E., 1973. *Forest Inventory*. 2nd ed. Munich: BLV Verlagsgesellchaft.

[16]   Husch, B., Miller, C.I., and Beers, T.W., 1982. *Forest Mensuration*. 3rd ed. New York: John Wiley & Sons, Inc.

[17]   Walther, B.A. and Moore, J.L., 2005. The Concepts of Bias, Precision and Accuracy, and Their Use in Testing the Performance of Species Richness Estimators, with a Literature Review of Estimator Performance. *Ecography*, 28 (6), pp.815–829.