International Journal of Ecology and Environmental Sciences
www.ecologyjournal.in
Online ISSN: 2664-7133; Print ISSN: 2664-7125; Impact Factor: RJIF 5.28
Received: 28-08-2020; Accepted: 14-09-2020; Published: 29-09-2020
Volume 2; Issue 4; 2020; Page No. 01-11

Development and comparative diagnosis of conventional (linear/nonlinear) and artificial intelligence techniques-based predictive models for estimating timber volume of *Tectona grandis*

Peter T Birteeb[1*], Ajit[2], Cini Varghese[3], Seema Jaggi[4]
[1-4] ICAR-Indian Agricultural Statistics Research Institute, Pusa, New Delhi, India
[1] Faculty of Agriculture, University for Development Studies, Nyankpala Campus, Tamale, Ghana

Abstract
This study aimed to develop volume estimation models which will be robust and useful for predicting merchantable volume of teak trees in different teak growing regions of the world. The data was statistically simulated based on various published models for different teak growing conditions in different parts of the world. A total of thirteen models comprising nine conventional (linear and nonlinear) and four Artificial Intelligence (AI) techniques-based models, thus two Support Vector Machine (SVM) techniques and two Artificial Neural Network (ANN) techniques, were fitted to the data. Several statistical model selection criteria including Efron's pseudo R-squared, Root Mean Square Error, Mean Absolute Bias, Nash-Sutcliffe Efficiency, Index of Agreement and Akaike Information Criterion were used to evaluate and rank the models' performances from best to worst. All AI techniques-based models were superior over conventional models in performance, and the overall best model was SVM technique followed by an ANN technique. Among conventional models, allometric models generally fitted the data better than linear regression type models, with model M_5 being the best while M_2 was the worst. Combination of tree diameter at breast height (dbh) and height as predictors of tree volume was shown to improve model prediction accuracy for teak trees irrespective of the model involved. On the basis of the varied nature of the data used for model fitting, the developed models would be useful in making reliable predictions of teak timber volume for different teak growing regions across the world. The models have wide application potential and may be recommended for use in managing teak plantation inventory in different parts of the world.

Keywords: allometric model; artificial neural network; generic volume equations; linear regression; support vector machine; *Tectona grandis*

1. Introduction

Teak (*Tectona grandis*) is considered one of the most well-known and widely used timbers of the world. It is well recognized for its high-quality timber in the international market. Superior qualities of teak timber include durability, ease of seasoning without splitting and cracking, attractiveness in colour and grain, lightness with strength, ease of working and carving, resistance to termite, fungus, and weathering (Kaosa-ard 1998; Bermejo *et al*. 2004) [23, 6]. Its natural distribution ranged from India through Burma, Laos and Thailand (Kaosa-ard 1998; Hansen *et al*. 2015) [23, 18]. Recently, increasing demand of teak timber and its related products has resulted in shrinking of natural forests. To ensure sustainable supply of timber, teak trees have been successfully established as exotic timber species in many countries outside the natural distribution zones, a few major ones being Puerto Rico, Panama and Trinidad in Central America; Brazil and Ecuador in South America; Senegal, Ivory Coast, Ghana, Togo, Benin and Nigeria in West Africa; Sudan, Kenya and Tanzania in East Africa; and Nepal, China, Sri Lanka, Bangladesh and Indonesia in Asia (Hoare and Patanapongsa 1988; Perez and Kanninen 2003) [20, 40].

For teak or any timber species, sustainable timber production is fundamental to long-term success of forest management as land use system (Brienen and Zuidema 2006) [9]. The goal of proper management of timber plantation is to provide high quality timber in quantities and sizes that will ensure maximum satisfaction of owner(s) while resolving imposed constraints (Newberry 1984) [33]. Consequently, evaluating various management and utilization alternatives for timber resources require the use of accurate and flexible methods to estimate stand and tree growth and yield (Sharma *et al*. 2000) [43]. Statistical models become very useful tools in estimating tree growth and yield, when measurements of variables such as diameter at breast height (dbh), tree height and volume are available. The dbh (measured at 1.3 m above ground) and tree height are crucial in measurement of tree growth, and have been used to estimate the total and merchantable tree volume, site index, and other relevant variables in forest growth and yield, succession, and carbon budget models in many species (Peng *et al*. 2001) [39]. While diameter can be easily and directly measured on trees, the measurement of total tree height is relatively complex, costly and time consuming (Sharma and Parton 2007) [44]. Accurate measurement of height of standing trees is difficult especially in closed-canopy forests, so most researchers ignore it in carbon-accounting programs (Hunter *et al*. 2013; Larjavaara and Muller-Landau 2013) [21, 26]. Jiang and Li (2010) [22] developed mixed-effects models for estimating height from diameter. Site specific height-dbh model was found to be suitable in estimating tree height in the lowland forests of Tanzania (Mugasha *et al*. 2016) [29].

In timber production, it is desirable to obtain estimates of growing stock in terms of timber volume. The relationships among dbh, height and volume are exploited in developing volume equations as the most common procedure in estimating stand volumes (Bohre *et al*. 2013) [8]. Though dbh is mostly used as predictor in volume equations, inclusion of height generally provides better estimates (Hunter *et al*. 2013; Chave *et al*. 2014) [21, 12] because it helps account for variations in climate, soil and some cultural practices (Shamaki *et al*. 2011) [42]. Estimates of bole volume are very useful in forest inventory since the basic management unit of forests is the volume of timber (Cháidez 2009) [11]. Despite the increasing use of biomass and density, volume is the most extensively used traditional measure for tree quantity in forest management (Koirala *et al*. 2017) [24]. Bylin (1982) [10] developed equations for predicting tree volumes from stump diameter and stump height for 15 tree species. Nunifu and Murchinson (1999) [35] compared three methods of estimating teak stand volume per hectare *viz*. two-stage simple random sampling (SRS), two stage sampling with probability proportional to basal area at the second stage (PPG), and a standard volume equation developed using individual sub-sample tree dbh, height and volume measurements. They observed that volume estimates derived using the standard volume equation were more precise, hence they recommended its use to guarantee efficient yield estimates in standing trees. Shamaki *et al*. (2011) [42] reported that double log equation was best for estimating volume of teak trees in Nimbia Forest Reserve in Nigeria. Allometric model was found to be better than linear, logistic, gompertz and chapman-richards models in predicting biomass components of *Populus deltoides* in India (Ajit *et al*. 2011) [2].

The need to improve yield prediction accuracies has led to the deployment of data-driven models under Artificial Intelligence (AI) techniques, particularly Support Vector Machine (SVM) and Artificial Neural Network (ANN). The reason is that AI techniques are algorithm-based, do not require any assumptions about the form of a fitting function, and yet, they are able to learn from data and predict patterns more accurately (Diamantopoulou 2006) [16]. AI techniques outperformed traditional regression models in estimating the height of trees (Özçelik *et al*. 2013) [38], diameter and total height (Vieira *et al*. 2018) [52], tree volume (Diamantopoulou 2006; Lacerda *et al*. 2017) [16, 25] and biomass (Nandy *et al*. 2017) [31]. Özçelik *et al*. (2010) [37] employed ANNs to estimate dbh and tree volume from stump diameter for 3 economically important species in Turkey. SVM and ANN models were found to be better than nonlinear regression models in estimating inside-bark volume of *Eucalyptus globulus* (Nieto *et al*. 2012) [34], height of *Crimean juniper* tree (Özçelik *et al*. 2013) [38] and bark thickness of teak trees (Vendruscolo *et al*. 2019) [51].

If trees are to be felled for commercial purpose, then there is a need for volume models that are able to quantify tree volume well prior to tree felling (Mugasha *et al*. 2016) [29]. Tewari and Singh (2018) [49] argued that specific rather than generic volume equations should be preferred because of variations in growing rates, management regimens and production objectives of teak plantations across different tree growing regions. Specific volume models are expedient in estimation of the average volume for standing trees (Avery and Burkhart 2002) [5], nevertheless, their usefulness is limited to the growing region of the data under consideration. This is due to the fact that most volume equations are local, thus developed based on specific data from a locality under certain forest management practices.

Though there have been considerable research efforts to develop volume estimation models for teak, most developed models are region specific, and no particular model can be deemed appropriate for use to produce reliable volume estimates across different growing regions. Therefore, this research aims to develop generic volume estimation models for teak which will be robust and have wider application potential across different teak growing regions of the world. In this study, we have attempted (i) using statistical approaches to simulate data from region/climate specific individual published equations from different growing regions in six countries; (ii) combining individual simulated data sets from different regions of the world into one aggregated-data-set; and (iii) fitting conventional (linear/non-linear) and AI (SVM and ANN) models on the aggregated-data-set to present a generic model that may be used for predicting teak timber volume across different regions and climatic conditions.

2. Materials and methods

2.1 Study area

This study covers several teak growing regions in six countries. Eleven published articles based on primary data were selected, from which the general climatic conditions of the study areas are summarized in Table 1. Primary data refers to the data obtained by direct measurements (or estimation) of variables of interest from standing or felled trees in the field.

Table 1: Main characteristics of the study areas of teak plantations in different countries

Country	Study area(s)	Climatic conditions of study area(s)			Reference
		Rainfall (mm)	Temperature (°C)	Altitude (m)	
Ghana	Tamale, Savelugu, Yendi and Damongo forest districts	960 – 1200	28	151 – 217	Nunifu 1997 [36]
Costa Rica	Guanacaste, Puntarenas, Limon, Alajuela	1659 – 4200	25.9 – 27.1	25 – 300	Perez and Kanninen 2003 [40]
	Puerto Carrillo, Palo Arco and Moravia	1800 – 2450	26 – 29	25 – 400	Bermejo *et al*. 2004 [6]
	Carrillo, Garza, Jicaral, Tempisque, San Carlos, Parrita, Quepos, Palmar Norte and Buenos Aires	1659 – 3900	26.1 – 27.1	25 – 300	Perez 2008 [41]
Tanzania	Mtibwa, Longuza and Kilombero Valley Teak Company	1000 – 1400	19 – 31	160 – 560	van Zyl 2005 [50]
	Longuza Forest Plantation	1400	27	160 – 560	Mwangi 2015 [30]
Nigeria	Nimbia Forest Reserve, Kaduna State	1650 – 1700	17 – 35	600	Shamaki *et al*. 2011 [42]
	Agudu Forest reserve	1400 – 1500	16 – 44	287	Shuaibu 2016 [45]
Nepal	Sarlahi district	1130 – 2680	10 – 45	60 – 330	Koirala *et al*. 2017 [24]
India	Karnataka	1600 – 4500	11 – 38	0 – 1925	Tewari *et al*. 2013 [48]
	Godhara, Baria, Narmada, Vyara and Dang forest divisions in Gujarat	578 – 1107	12 – 49	198	Tewari and Singh 2018 [49]

2.2 Data simulation

The data generation process started with an extensive search of literature to identify key articles or researches, which had published data on volume prediction models of teak. The eleven articles which were selected from six countries are indicated in Tables 1. Those researches represent major teak growing regions across the world. Descriptive statistics (minimum, maximum, mean and standard deviation (SD) of the variables (dbh, height and volume)) from each research were tabulated (Table 2) and used for the simulation of data. In each of the selected researches, the authors had fitted several volume models and recommended some models as best based on well-defined model evaluation criteria. Parameter estimates of those best models, one from each article, were then used for simulation of data points for the present study.

For each of the published equations (representing different region and different climatic conditions), the minimum and maximum values of dbh (as depicted in Table 2) were used as the two parameters of a uniformly distributed random variable (X) to generate $N(=500)$ data points at random. Thus $X \sim U(a, b)$ where a and b are the minimum and maximum values respectively of the distribution. The data sets were generated in R using $runif(N, a, b)$ function. Height values were also generated similarly. However, the data set from Shamaki *et al*. (2011) [42] was generated using $rnorm(N, mean, sd)$ while the actual field data of Mwangi (2015) [30] was used. The randomly generated values of dbh or dbh and height were inputted into the published best equations (of the 11 respective research articles) to predict the corresponding volume values. However, these predicted volume values are static (in the sense of being generated from a fixed equation), accordingly there arises a need to associate a random (actually the error) component with each of these predicted volume values so that these pseudo generated-data set simulates the actual/original harvest-data set from which the published equation was derived. Thus to create randomness in the predicted volume values, the SDs associated with the original volume values (Table 2) were used to generate 500 raw random errors (say e_i, $i = 1, 2, \ldots, N$) for each data set, following a Normal-Distribution with MEAN=0 and SD=volume-standard-deviation-as-reported-in-article.

Since each published model had R^2 value, the difference of the R^2 value from one, denoted as $d_j = 1 - R_j^2$, (which is the portion of the total variance not accounted for by the model, *i.e.* error component) was obtained. Here, $j = 1, 2, \ldots, 11$ for the 11 equations.

Table 2: Descriptive statistics of teak data from published researches originating from different teak growing regions

Sample size (n)	Data use	Variable	Min	Max	Mean	SD	Source
368	Training	dbh (cm)	3.6600	33.5000	14.9200	8.2590	Nunifu 1997 [36]
		Height (m)	4.0000	26.4000	11.9780	5.7826	
		Vol (m^3)	0.0015	1.5413	0.4339	0.3740	
111	Training	dbh (cm)	2.4000	58.7000	17.8000	8.8000	Perez and Kanninen 2003 [40]
		Height (m)	3.8000	33.3000	16.5000	5.2000	
		Vol (m^3)	-	-	-	-	
285	Training	dbh (cm)	10.0000	27.2000	18.7000	2.4000	Bermejo *et al*. 2004 [6]
		Height (m)	12.0000	23.2000	17.9000	1.9000	
		Vol (m^3)	0.0199	0.3584	0.1403	0.0560	
222	Training	dbh (cm)	8.1000	79.4000	31.2719	17.3083	van Zyl 2005 [50]
		Height (m)	8.9800	33.8200	21.2837	6.8372	
		Vol (m^3)	0.0100	5.9400	1.0053	1.1369	
25	Training	dbh (cm)	9.4000	55.4000	23.3000	10.3000	Perez 2008 [41]
		Height (m)	12.4000	33.3000	20.6000	5.7000	
		Vol (m^3)	-	-	-	-	
364	Training	dbh (cm)	-	-	13.7240	1.6973	Shamaki *et al*. 2011 [42]
		Height (m)	-	-	5.5000	0.5661	
		Vol (m^3)	-	-	0.0785	0.0226	
91	Training	dbh (cm)	5.5000	36.0000	17.5700	5.6200	Tewari *et al*. 2013 [48]
		Height (m)	6.5000	21.4000	14.6900	3.4900	
		Vol (m^3)	0.0063	1.0357	0.2101	0.1651	
39	Validation	dbh (cm)	6.5000	35.2000	16.4800	6.6800	
		Height (m)	6.9000	21.3000	13.9700	3.4400	
		Vol (m^3)	0.0104	1.1037	0.2001	0.2133	
51	Training	dbh (cm)	1.0000	83.4000	37.3940	24.5329	Mwangi 2015 [30]
		Height (m)	1.5000	37.5000	25.6470	10.9960	
		Vol (m^3)	-	-	-	-	
70	Training	dbh (cm)	23.0000	36.0000	29.0000	3.0000	Shuaibu 2016 [45]
		Height (m)	11.6000	20.2000	15.8600	2.1700	
		Vol (m^3)	0.5100	1.4700	0.8900	0.1980	
30	Validation	dbh (cm)	21.7600	34.6800	29.1000	2.9000	
		Height (m)	12.6000	20.6000	16.7130	2.3980	
		Vol (m^3)	0.4982	1.4819	0.9410	0.2010	
31	Training	dbh (cm)	6.3700	58.9200	27.9100	16.0300	Koirala *et al*. 2017 [24]
		Height (m)	5.7000	26.1000	16.1300	5.2100	

		Vol (m^3)	0.0100	2.1300	0.6500	0.6400	
13	Validation	dbh (cm)	6.0500	57.3200	27.8500	18.7000	
		Height (m)	6.1000	23.3000	16.0800	5.6600	
		Vol (m^3)	0.0100	2.4400	0.7600	0.9200	
41	Training	dbh (cm)	7.3000	30.8000	18.0700	5.4600	Tewari and Singh 2018 [49]
		Height (m)	8.2000	22.0000	14.2700	3.5500	
		Vol (m^3)	0.0249	0.6589	0.1637	0.1388	

Min, minimum; Max, maximum; SD, standard deviation; dbh, diameter at breast height; Vol, volume.

Then d_j^{th} percentage of the random errors e_i (say, $e_{d_j} = d_j * e_i$) were calculated and added to the predicted volume values to give simulated volume values. So under each model, simulated volume = predicted volume + e_{d_j}. The variables in each simulated data set were grouped together using $cbind()$ function and saved to comma-separated values file using $write.csv()$ function in R. All the data sets were combined together to constitute one data set called aggregated data, containing 6511 final data points (observations) after data cleaning.

2.3 Fitting of volume models

A key step in development of statistical models is the need for two independent data sets so that one set is used for model estimation and the other set for model validation (Ajit 2010) [3]. However, if there are no two independent data sets, then the available data set may be judiciously partitioned into two, one used for model estimation and the other for model validation. In such a scenario, Geisser (1975) [17] suggested that a random sample (without replacement) of about 80% data points should be selected and utilized for model fitting while the remaining 20% data points is kept as the second independent data set to be used for model validation. In this study, the data set was randomly split into 80% and 20% for model training and validation respectively using $sample()$ function in R. The descriptive statistics of the randomly sampled data sets are given in Table 3. A careful look at these descriptive statistics in comparison to those in Table 2 suggests a similarity in pattern of the data points.

Table 3: Descriptive statistics of simulated data for teak

Sample size (n)	Data use	Variable	Minimum	Maximum	Mean	Standard deviation
5209	Training	DBH (cm)	1.4450	80.5000	26.2764	13.6375
		Height (m)	5.0000	37.5000	18.7277	5.4069
		Vol (m^3)	0.0009	6.9422	0.6750	0.8737
1302	Validation	DBH (cm)	1.6935	83.4000	25.7518	14.2987
		Height (m)	7.1215	36.6151	18.6345	5.5526
		Vol (m^3)	0.0011	6.6509	0.6730	0.9278

Nine conventional (linear and nonlinear) models and 4 AI technique-based models were fitted to the data. The 9 conventional models were chosen based on recommendations from literature (Nunifu 1997; Perez and Kanninen 2003; Bermejo *et al*. 2004; Perez 2008; van Zyl 2005; Shamaki *et al*. 2011; Shuaibu 2016; Tewari *et al*. 2013; Koirala *et al*. 2017, Tewari and Singh 2018) [36, 40, 6, 41, 50, 42, 45, 48, 24, 49]. We have denoted these 9 models as $M_i (i = 1, 2, \ldots, 9)$ and given their forms below.

M_1: $E(V) = \beta_0 + \beta_1 D + \beta_2 H$ (1)

M_2: $E(V) = \beta_0 + \beta_1 D$ (2)

M_3: $E(V) = (\beta_0 + \beta_1 D)^2$ (3)

M_4: $E(V) = \beta_0 D^{\beta_1}$ (4)

M_5: $E(V) = \beta_0 D^{\beta_1} H^{\beta_2}$ (5)

M_6: $E(V) = \beta_0 + \beta_1 D^2 H$ (6)

M_7: $E(V) = e^{\left(\beta_0 + \beta_1 ln(D^2 H)\right)}$ (7)

M_8: $E(\ln V) = \beta_0 + \beta_1 \ln D + \beta_2 \ln H$ (8)

M_9: $E(V) = \beta_0 + \beta_1 D + \beta_2 D^2$ (9)

Where $E()$ is expectation, V is volume of main stem (bole), D is diameter at breast height (dbh), H is height of main stem, ln is natural logarithm and β_0, β_1 and β_2 are parameter estimates to be determined when models are fitted. The models were fitted using either $lm()$ or $nls()$ functions of $nlstools$ package in R, depending on the type of model.

The 4 AI technique-based models comprised 2 ANN models (denoted as ANNdbh and ANNdbhht) and 2 SVM models (denoted as SVMdbh and SVMdbhht), all of which are based on some learning algorithms. ANN is an efficient computing system consisting of one or more layers, where each layer contains one or more simple processing units called neurons. A typical ANN has an input layer, a hidden layer (which may be absent in some cases) and an output layer. The input layer receives the input variables and transmits them to the hidden layer which processes them before transmitting to the output layer. There is an activation function(s) within the hidden layer which transform(s) the input received in order to ensure that the amplitude of the output of a neuron is within limit. A simple ANN model is defined by Haykin (2009) [19] as,

$$Y = \vartheta\left(\sum_{i=1}^{m} x_i w_{ji} + b_j\right) \quad (10)$$

where Y is the output variable, x_i is the i^{th} input variable, w_{ji} is the synaptic weight of j^{th} neuron assigned to x_i, b_j is bias of j^{th} neuron and $\vartheta(.)$ is the activation function.

To fit the 2 ANN models, the data set was processed by normalizing all variables due to the difference in units of measurements. The training data set was inputted into the ANN model for calibration, with dbh alone (ANNdbh) or dbh and height (ANNdbhht) as inputs and volume as output. Predictions were then made for both the training and validation data sets for computation of model evaluation statistics. The implementation of the models was done using feed forward neural network with multilayer perceptron architecture (having 2 hidden layers with

3, 2 neurons respectively) under the $neuralnet()$ function in R. A perceptron is the simplest form of a neural network (an artificial neuron) that does certain computations to detect features in an input data.

SVM is a type of algorithm that receives a training sample and constructs a hyperplane as the decision surface with the target of maximising the margin of separation between two patterns (Haykin 2009) [19]. It is used in solving both classification and nonlinear-regression problems. According to Cristianini and Shawe-Taylor (2000) [14], SVM uses a linear learning machine to learn a non-linear function in a kernel-induced feature space while the capacity of the system is controlled by a parameter that is independent of the dimensionality of the space. The technique uses a slack variable to measure the deviation of a data point from the ideal pattern-separating hyperplane. Thus, for any set of observations, the SVM technique finds a strip in the feature space that separates each class in a different semi space, maximizing the separation between classes (in a training sample) and minimizing some measure of the misclassification errors (Blanco *et al*. 2020) [7]. Haykin (2009) [19] provided detail exposition of the SVM techniques. A simple equation of the hyperplane that performs the pattern separation is,

$$w^T x + b = 0 \qquad (11)$$

Where $\boldsymbol{x}$ is an input vector, w is an adjustable weight vector, and b is a bias.

In this study, the 2 SVM models were fitted on the training data set with either dbh alone (SVMdbh) or dbh and height (SVMdbhht) as input variables respectively, and volume as output variable using $svm()$ function in R. Tuning of models was done using the $tune()$ function to obtain the best model.

2.4 Model evaluation and validation

Under every model, predictions of timber volume were done for both the training and testing data sets for computation of model evaluation statistics. This study employed several statistical criteria to assess model performance. These criteria are defined in the following: (i) *Efron's pseudo R-squared* (R^2) refers to the square of the correlation between the predicted values of a model and the actual values (Mangiafico 2016) [27]. It gives a relative measure of how well a model explains the data; (ii) *Root Mean Square Error* (RMSE) is a measure of average magnitude of the error of a model in predicting quantitative data. It is particularly useful when large errors are undesirable; (iii) *Mean Absolute Error* (MAE) is a measure of the average magnitude of the errors between the predicted and the observed values, without considering their direction; (iv) *Nash-Sutcliffe Efficiency* (NSE) is a normalized statistic that measures the relative magnitude of the residual variance compared to the observed data variance (Nash and Sutcliffe, 1970) [32]. It gives a measure of the predictive accuracy of a model; (v) *Index of Agreement* (IA) is a standardized measure of the amount of model prediction error (Willmott 1981) [53]. IA represents the ratio of the mean square error and the potential error. (vi) *Akaike Information Criterion* (AIC) is a measure of how well a model fits a data set without over-fitting it, relative to other model(s) fitted to the same data set. AIC of a model is useful only when compared to AICs of other models. The formulae, limits and interpretations of these statistics are presented in Table 4.

Table 4: Statistical criteria used for evaluating model performance

Statistic	Formula	Limits	Interpretation
R^2	$=1-\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2}$	0 to 1	Higher values indicate better fit
RMSE	$=\left(\frac{1}{n}\sum_{i=1}(y_i-\hat{y}_i)^2\right)^{\frac{1}{2}}$	0 to ∞	Lower values indicate better fit
MAE	$=\frac{1}{n}\sum_{i=1}\lvert y_i-\hat{y}_i\rvert$	0 to ∞	Lower values indicate better fit
NSE	$=1-\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2}$	-∞ to 1	1 = perfect fit; 0 = model predictions are as accurate as the observed mean; NSE < 0 = observed mean is a better than the model
IA	$=1-\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(\lvert\hat{y}_i-\bar{y}\rvert+\lvert\bar{y}_i-\bar{y}\rvert)^2}$	0 to 1	0 = no agreement, 1 = perfect agreement
AIC	$=2\ln\left(\frac{e^p}{\hat{L}}\right)$	-∞ to +∞	Lower values indicate better fit

where y_i is i^{th} observed data point, $\bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i$ is the mean of the observed data points, $\hat{y}_i$ is i^{th} predicted value of each model, n is the sample size, e is exponent, p is the number of estimated parameters in a model, and $\hat{L}$ is the maximum value of the likelihood function for the model.

Under each evaluation criterion, all the models were ranked for their performance following the procedure of Tewari and Singh (2018) [49]. Ranking was carried out in two ways, thus ranking all models (without including AIC) and ranking only conventional models (with AIC included). This was due to the fact that AIC values could not be computed for the AI technique-based models, as they are only approximations to an underlying model whereas AIC is based on assumption of asymptotic normality of maximum likelihood estimators (Anders and Korn 1999) [4]. Therefore, all the 13 models were assigned rank scores from best (1) to worst (13) without considering AIC. Rank scores were then summed for each model and the model with lowest total score declared the best. In the second case, by including AIC, the 9 conventional models were ranked in similar manner from best (1) to worst (9). Additional evaluation of models was done by regressing observed volume on predicted volume of training data set for each model using simple linear regression. The presence of autocorrelation in the residuals was tested by computing Durbin-Watson (DW) statistic for each regression (Montgomery *et al*. 2003) [28]. The statistic is given as,

$$DW = \frac{\sum_{i=2}^{n}(r_i - r_{i-1})^2}{\sum_{i=1}^{n} r_i^2}$$

Where $r_i (i = 1, 2, \ldots, n)$ are the residuals from an ordinary least squares analysis applied to the observed-predicted regression data? Values of DW range from 0 to 4, with values between 1.5 and 2.5 considered relatively normal and indicate no autocorrelation in the given data.

All models were validated by using each fitted model to predict tree volume for the test data set. All the evaluation statistics (except AIC) were computed for the test data set and ranked from best to worst.

3. Results and discussion

After fitting the models to the training data, the parameter estimates along with the respective t-test of significance and probabilities for each estimate are presented in Table 5. All parameter estimates for all models were highly significant (p < 0.001). This result agreed with Bohre *et al*. (2013) [8] who reported significant parameter estimates when different permutations of dbh and height were used as predictors of teak tree volume in Madhya Pradesh, India. Tewari *et al*. (2013) [48] and Tewari and Singh (2018) [49] also reported highly significant parameter estimates for several volume models of teak though, in both reports, few parameter estimates were insignificant. The estimates of intercepts (β_0) for models M_1, M_2 and M_3 were negatives (Table 5), which imply that these models may produce negative estimates of tree volume as they are linear regression type models. In a study of *Populus deltoides*, Ajit *et al*. (2011) [2] observed negative estimates of tree volume from a linear regression model. Avery and Burkhart (2002) [5] stated that negative intercepts are expected for merchantable volume prediction. Nevertheless, the possibility of producing negative volume estimates does not make biological sense as actual tree volume cannot be negative. Since the nonlinear models will rarely produce negative volume estimates, it means their predictions mimic actual field situation and should be preferred over the linear regression type models. The SE for parameter estimates of all models were generally low. This suggests that the models would predict teak volume with high accuracies, considering the fact that they were fitted on data that had high variability. There were no parameter estimates for ANN and SVM models (Table 5), because they are AI techniques which are based on algorithm and considered to be approximations of some true underlying models (Anders and Korn 1999) [4]. However, diagrams of the developed neural networks for ANNdbh and ANNdbhht models are given in Fig 1 and 2 respectively. In both networks, there were two hidden layers having 3 and 2 neurons respectively with the associated synoptic weights clearly indicated in the diagrams. These networks were chosen, based on their prediction accuracies, as the best among several networks fitted with varying number of hidden layers and neurons.

Table 5: Model parameter estimates for training data set for teak

Model	Parameter	Parameter estimate	Standard Error	t-value	Probability
M_1	β_0	-0.79460	0.01972	-40.30	<0.001
	β_1	0.06073	0.00065	92.89	<0.001
	β_2	-0.00674	0.00165	-4.09	<0.001
M_2	β_0	-0.86234	0.01071	-80.53	<0.001
	β_1	0.05851	0.00036	161.74	<0.001
M_3	β_0	-0.14831	0.00498	-29.76	<0.001
	β_1	0.03199	0.00011	293.37	<0.001
M_4	β_0	2.530e-04	8.401e-06	30.12	<0.001
	β_1	2.30300	8.308e-03	277.20	<0.001
M_5	β_0	1.962e-04	7.311e-06	26.83	<0.001
	β_1	2.16500	0.01370	158.07	<0.001
	β_2	0.24120	0.01863	12.95	<0.001
M_6	β_0	0.10280	3.776e-03	27.24	<0.001
	β_1	2.790e-05	1.033e-07	270.02	<0.001
M_7	β_0	-9.12146	0.03901	-233.90	<0.001
	β_1	0.88529	0.00344	257.50	<0.001
M_8	β_0	-8.64788	0.06021	-143.63	<0.001
	β_1	2.78806	0.01732	160.97	<0.001
	β_2	-0.44435	0.03323	-13.37	<0.001
M_9	β_0	0.08272	0.01141	7.247	<0.001
	β_1	-0.01306	7.517e-04	-17.38	<0.001
	β_2	1.067e-03	1.075e-05	99.27	<0.001

Note: Ye-x = Y×10^{-x}

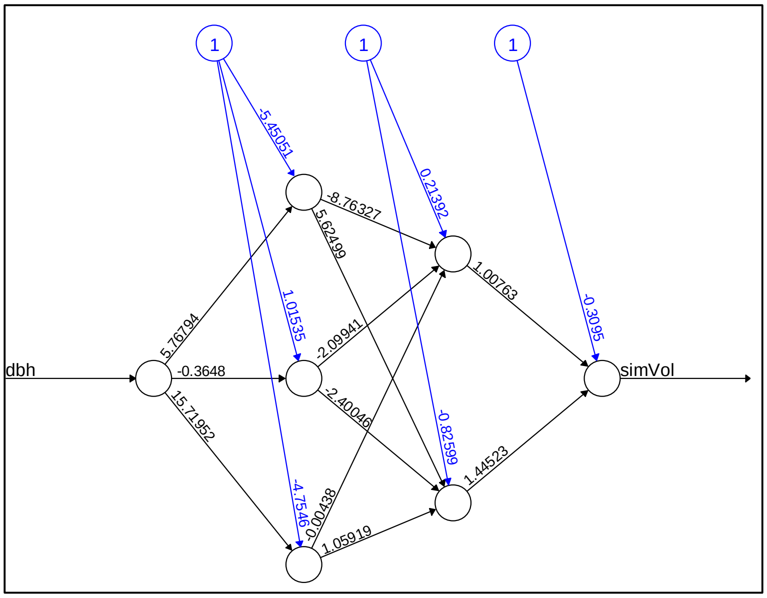

Fig 1: Artificial neural network with 2 hidden layers (3, 2 neurons) having dbh as input and tree volume (simVol) as output

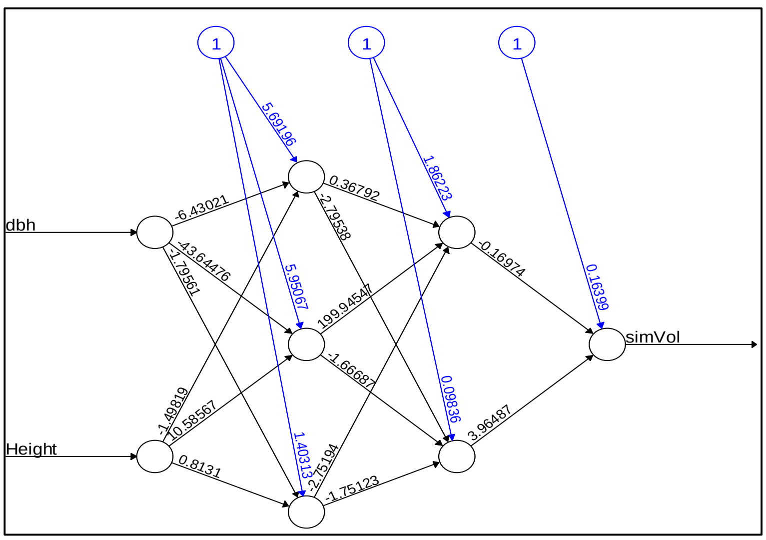

Fig 2: Artificial neural network 2 hidden layers (3, 2 neurons) having dbh and height as inputs and tree volume (simVol) as output

The R^2 values of the models were generally high except for models M_1 and M_2 which had moderate values (Table 6). This means that the models accounted for higher amount of variability in teak volumes. This finding is similar to the work of Tewari *et al*. (2013) [48], who reported that non-linear models were better in estimating teak volume. In a study of teak plantation from two different forest reserves in Nigeria, Adekunle (2007) [1] reported relatively moderate to high R^2 values (74.89 to 92.50%). The NSE is notably similar to R^2 in formula and the computed values too, except in few cases where there were few differences in values (Table 6). The very high values of NSE and IA indicates how well these models fitted the training data. The IA in particular, shows the degree to which the observed (actual) tree volumes are accurately estimated by the predicted tree volumes, and gives a measure of the degree to which a model's predictions are error free (Willmott 1981) [53]. Based on the rank scores, the overall best model was an SVM followed by ANN, both of which had dbh and height as independent variables. Clearly, all AI models outperformed the conventional models on all model evaluation criteria. This implied that AI models fitted the data better and would produce better predictions. The fitting patterns are corroborated in the scatter plots of observed against predicted volumes as well as the residual plots given in Appendix (Fig 3). In the observed-predicted plots, the data points appeared to line up along the main diagonals. Also, from the residual plots, the residuals of all AI models appeared to spread around zero quite well, to some extent, given the heterogeneous nature of the data under consideration.

In most studies involving comparison of AI and conventional models, the former has been noted to be phenomenal in performance (Özçelik *et al*. 2010; Özçelik *et al*. 2013; Lacerda *et al*. 2017) [37, 38, 25]. Socha *et al.* (2020) [46] attributed such phenomena to non-requirements of full a priori knowledge of a given data by AI models. Also, AI techniques are algorithm-based and so have enhanced ability to work with any noisy or low-quality data (Diamantopoulou *et al*. 2015) [15]. However, since AI models do not have parameter estimates for any fitting equations, they cannot be written out in an equation form for easy use by laymen. Therefore, their implementation may be hindered by lack of skilled personnel, appropriate computer software or both. In this regard, the use of conventional volume models may still be encouraged in many situations.

Table 6: Model evaluation statistics for models fitted on training data set of teak

Model	R^2	RMSE	MAE	NSE	IA	AIC	Total rank		Overall rank	
							No AIC	With AIC	No AIC	With AIC
M_1	0.8345 (12)	0.3554 (12)	0.2284 (12)	0.8345 (12)	0.9529 (12)	4011.82 (7)	60	67	11	8
M_2	0.8340 (13)	0.3559 (13)	0.2297 (13)	0.8340 (13)	0.9527 (13)	4026.53 (8)	65	73	12	9
M_3	0.9424 (8)	0.2098 (8)	0.1220 (6)	0.9423 (8)	0.9850 (8)	-1478.04 (4)	38	42	8	4
M_4	0.9459 (6)	0.2035 (6)	0.1200 (5)	0.9458 (6)	0.9860 (6)	-1800.14 (2)	29	31	6	2
M_5	0.9476 (5)	0.2001 (5)	0.1225 (7)	0.9476 (5)	0.9865 (5)	-1973.01 (1)	27	28	5	1
M_6	0.9333 (11)	0.2255 (10)	0.1570 (11)	0.9333 (10)	0.9825 (10)	-726.68 (6)	52	58	10	6
M_7	0.9356 (9)	0.2219 (9)	0.1418 (10)	0.9355 (9)	0.9832 (9)	-896.06 (5)	46	51	9	5
M_8	0.9334 (10)	0.2646 (11)	0.1391 (9)	0.9083 (11)	0.9793 (11)	4856.36 (9)	52	61	10	7
M_9	0.9426 (7)	0.2093 (7)	0.1249 (8)	0.9426 (7)	0.9850 (7)	-1504.75 (3)	36	39	7	3
ANNdbh	0.9665 (3)	0.1600 (3)	0.1076 (3)	0.9665 (3)	0.9914 (3)	n/a	15		3	
ANNdbhht	0.9768 (2)	0.1331 (2)	0.0822 (2)	0.9768 (2)	0.9941 (2)	n/a	10		2	
SVMdbh	0.9658 (4)	0.1616 (4)	0.1077 (4)	0.9658 (4)	0.9913 (4)	n/a	20		4	
SVMdbhht	0.9808 (1)	0.1211 (1)	0.0736 (1)	0.9808 (1)	0.9952 (1)	n/a	5		1	

R^2, Efron's pseudo R-squared; RMSE, root mean square error; MAE, mean absolute error; NSE, Nash-Sutcliffe efficiency; IA, Index of Agreement; AIC, Akaike Information Criterion; values in parentheses are rank scores; n/a, not applicable

The allometric model M_5 was the best among the conventional models, followed by M_4 (Table 6). These results are in agreement with the findings of Ajit *et al*. (2011) [2] who reported that allometric model was superior to standard linear and growth models in a study of *Populus deltoids*. Clearly, the outstanding models SVMdbhht, ANNdbhht and M_5 require the combined use of dbh and height as predictors of tree volume. Perhaps the inclusion of tree height as predictor of tree volume improves tree volume prediction accuracy of a model (Bohre *et al*. 2013; Chave *et al*. 2014) [8, 12]. It is important to note that SVMdbh, ANNdbh, M_4 and M_9 are also good models since they had quite high prediction accuracies and each have only dbh as predictor of tree volume so that, their use will not require height measurements. Measuring of tree height on standing trees is known to be quite tedious and expensive (Sharma and Parton 2007) [44]. Therefore, the use of these 4 models may be encouraged with the expectation that the marginal loss in prediction accuracies of these models is compensated for by an obvious reduction in the cost of data collection due to exclusion of height measurement.

From Table 7, the correlation coefficients between observed and predicted tree volumes were generally high with concomitant high values of R^{2*} resulting from the regression of observed volume on predicted volume. These results give a confirmation that all models were well fitted to the training data and AI technique-based models were superior in performance. The DW statistics ranged from 1.937 to 2.01 and implied that there was no problem of autocorrelation in the residuals. When autocorrelation is present in the residuals of a given model, it is an indication that the model structure is mis-specified (Montgomery *et al*. 2003) [28].

Table 7: Statistics for regressing observed on predicted values for each model

Model	Correlation	R^{2*}	DW	Probability
M_1	0.914	0.835	1.946	<0.001
M_2	0.913	0.834	1.946	<0.001
M_3	0.971	0.942	1.997	<0.001
M_4	0.973	0.946	2.003	<0.001
M_5	0.973	0.948	2.005	<0.001
M_6	0.966	0.933	2.016	<0.001
M_7	0.967	0.936	2.013	<0.001
M_8	0.966	0.933	2.004	<0.001
M_9	0.971	0.943	2.000	<0.001
ANNdbh	0.983	0.966	1.988	<0.001
ANNdbhht	0.988	0.977	1.986	<0.001
SVMdbh	0.983	0.966	1.981	<0.001
SVMdbhht	0.990	0.981	2.016	<0.001

DW, Durbin-Watson statistic; R^{2*}, coefficient of determination.

Model evaluation statistics computed for each model using the test or validation data set are given in Table 8. After ranking, the results followed a similar pattern to those obtained from the training data set, with the AI models still appearing superior. Since the data (Table 2) for the present study were generated under varied teak growing conditions across the world (Table 1), the developed models would be useful in estimating teak volume yields on a wider scale. They may be described as generic volume prediction models on the bases of their applicability over a wide range of teak growing regions in the world. Promoting the use of these generic models is in contrast with Tewari and Singh (2018) [49] who suggested that specific volume equations should be preferred over generic ones in estimating tree volume. The fitted models in this study are based on data with high variability and so would account for wider variations in teak growing environments (Shamaki *et al*. 2011) [42]. It is important to note that the input variables (especially dbh) are relatively easy to measure on standing trees, so using them as predictors imply that volume yields could be estimated through non-destructive sampling as suggested by Tackenberg (2007) [47] and Chen *et al*. (2009) [13]. The development and use of generic volume estimation models are very relevant in forest inventory management and should be encouraged. If reliable volume prediction models are available to provide accurate estimates of teak tree volumes of standing trees in several locations, they will contribute significantly in improving management of forest inventory, as foresters could make well-informed decisions regarding which particular forest section, tree and even time to harvest timber to meet specific market requirements as well as ensure sustainable use of timber resources.

Table 8: Model evaluation statistics obtained from using developed models to predict volume in validation data set

Model	R^2	RMSE	MAE	NSE	IA	Total rank	Overall rank
M_1	0.8378 (12)	0.3749 (12)	0.2434 (12)	0.8366 (12)	0.9530 (12)	60	12
M_2	0.8371 (13)	0.3756 (13)	0.2444 (13)	0.8360 (13)	0.9528 (13)	65	13
M_3	0.9447 (8)	0.2182 (8)	0.1256 (7)	0.9446 (8)	0.9855 (7)	38	8
M_4	0.9472 (6)	0.2132 (6)	0.1242 (5)	0.9471 (6)	0.9863 (6)	29	6
M_5	0.9501 (5)	0.2072 (5)	0.1252 (6)	0.9501 (5)	0.9870 (5)	26	5
M_6	0.9401 (10)	0.2275 (10)	0.1575 (11)	0.9398 (10)	0.9841 (10)	51	10
M_7	0.9431 (9)	0.2215 (9)	0.1396 (9)	0.9430 (9)	0.9850 (9)	45	9
M_8	0.9329 (11)	0.2817 (11)	0.1440 (10)	0.9078 (11)	0.9792 (11)	54	11
M_9	0.9448 (7)	0.2180 (7)	0.1293 (8)	0.9448 (7)	0.9855 (8)	37	7
ANNdbh	0.9666 (3)	0.1695 (3)	0.1073 (3)	0.9666 (3)	0.9914 (3)	15	3
ANNdbhht	0.9762 (2)	0.1432 (2)	0.0833 (2)	0.9762 (2)	0.9939 (2)	10	2
SVMdbh	0.9664 (4)	0.1700 (4)	0.1094 (4)	0.9664 (4)	0.9914 (4)	20	4
SVMdbhht	0.9768 (1)	0.1415 (1)	0.0771 (1)	0.9767 (1)	0.9941 (1)	5	1

R^2, Efron's pseudo R-squared; RMSE, root mean square error; MAE, mean absolute error; NSE, Nash-Sutcliffe efficiency; IA, Index of Agreement; values in parentheses are rank scores.

4. Conclusions

In this study, attempts were made to develop generic models for predicting volume of teak timber. The results showed that both conventional (linear and nonlinear) models and AI techniques-based (ANN and SVM) models fitted the observed data set well. Almost all the candidate models accounted for very high amount of variation in teak timber volume with very high prediction accuracies. The overall best model was SVMdbhht followed by ANNdbhht while M_5 ranked 5th among all models but 1st among only conventional models. The AI models generally outperformed conventional models, though the use of these AI models require skilled personnel and appropriate computer software. Combined use of dbh and height as predictors of tree volume was shown to enhance model prediction accuracy for teak trees irrespective of the model involved. The study was based on several published data from different teak growing regions of the world, hence the developed models have wide application potential and may be recommended for use in managing teak plantation inventory in different parts of the world.

5. Acknowledgment

Authors thank the PG School of Indian Agricultural Research Institute, New Delhi for providing research facilities for conduct of this research. Peter T. Birteeb thanks University for Development Studies for grant of study leave to undertake this research in India.

6. References

1. Adekunle VAJ. Non-linear regression models for timber volume estimation in natural forest ecosystem, Southwest Nigeria. Res. J For. 2007; 1:40-54.
2. Ajit, Das DK, Chaturvedi OP, Jabeen N, Dhyani SK. Predictive models for dry weight estimation of above and below ground biomass components of *Populus deltoides* in India: Development and comparative diagnosis. Biomass Bioenerg. 2011; 35:1145-1152.
3. Ajit. Estimation and validation methods in tree volume and biomass modeling: Statistical concept. National Research Centre for Agroforestry, Jhansi, India, 2010, 18. http://sscnars.icar.gov.in/Agro/1-Tree%20Growth%20Modelling-Statistical-Concepts.pdf. Accessed 12 July, 2020.
4. Anders U, Korn O. Model selection in neural networks. Neural Networks, 1999; 12:309-323. https://doi.org/10.1016/S0893-6080(98)00117-8
5. Avery TE, Burkhart HE. Forest Measurements, 5th edn. New York, 2002.
6. Bermejo I, Canellas I, Miguel AS. Growth and yield models for teak plantations in Costa Rica. Forest. Ecol. Manag. 2004; 189:97-110.
7. Blanco V, Japón A, Puerto J. Optimal arrangements of hyperplanes for SVM-based multiclass classification. Adv. Data Anal. Class if. 2020; 14:175-199. https://doi.org/10.1007/s11634-019-00367-6
8. Bohre P, Chaubey OP, Singhal PK. Biomass accumulation and carbon sequestration in *Tectona grandis* Linn. f. and *Gmelina arborea* Roxb. Int. J Bio-Sci. Bio-Tech. 2013; 5:153-172.
9. Brienen RJW, Zuidema PA. The use of tree rings in tropical forest management: Projecting timber yields of four Bolivian tree species. Forest Ecol. Manag. 2006; 226: 256-267.
10. Bylin CV. Volume Prediction from Stump Diameter and Stump Height of Selected Species in Louisiana. United States Department of Agriculture. Forest Service. Research Paper SO-182 11p Southern Forest Experiment Station: New Orleans, La, 1982.

11. Cháidez JN. Allometric equations and expansion factors for tropical dry forest trees of Eastern Sinaloa, Mexico. Trop. Subtrop. Agroecosystems. 2009; 10:45-52.
12. Chave J, Rejou-Mechain M, Burquez A, Chidumayo E, Colgan MS, Delitti WBC, *et al*. Improved allometric models to estimate the aboveground biomass of tropical trees. Glob. Chang. Biol. 2014; 20:3177–3190. doi:10.1111/gcb.12629
13. Chen W, Li J, Zhang Y, Zhou F, Koehler K, Leblanc S, *et al*. Relating biomass and leaf area index to non-destructive measurements in order to monitor changes in Arctic Vegetation. Arctic. 2009; 62:281-294.
14. Cristianini N, Shawe-Taylor J. An introduction to support vector machines and other kernel-based learning methods. Cambridge University Press, 2000, 204.
15. Diamantopoulou MJ, Özçelik R, Crecente-Campo F, Eler Ü. Estimation of Weibull function parameters for modelling tree diameter distribution using least squares and artificial neural networks methods. Biosyst. Eng. 2015; 133:33-45. doi:10.1016/j.biosystemseng. 2015.02.013
16. Diamantopoulou MJ. Tree-bole volume estimation on standing pine trees using cascade correlation artificial neural network models. Agricultural Engineering International: the CIGR Ejournal, 2006, 8: Manuscript IT 06 002.
17. Geisser S. The predictive sample re-use method with application. J Amer. Statist. Assoc. 1975; 70:320-328.
18. Hansen OK, Changtragoon S, Ponoy B, Kjær ED, Minn Y, Finkeldey R, *et al*. Genetic resources of teak (*Tectona grandis* Linn. f.) - strong genetic structure among natural populations. Tree Genet. Genomes, 2015, 11. doi:10.1007/s11295-014-0802-5
19. Haykin S. Neural Networks and Learning Machines, 3rd edn. Pearson Education, Delhi, 2009.
20. Hoare P, Patanapongsa N. Long-rotation, high value trees: an alternative strategy for private forestry. Commonw. For. Rev. 1988; 67:351-61.
21. Hunter MO, Keller M, Vitoria D, Morton DC. Tree height and tropical forest biomass estimation. Bio-geosciences Discuss. 2013; 10:10491-10529.
22. Jiang L, Li Y. Application of nonlinear mixed-effects modeling approach in tree height prediction. Journal of Computers. 2010; 5(10):1575-1581.
23. Kaosa-ard A. Overview of problems in teak plantation establishment. In: Kashio M, White K (ed). Teak for the Future. Bangkok: RAP Publication: 1998/5, FAO Regional Office for Asia and the Pacific, 1998, 49-60.
24. Koirala A, Kizha AR, Baral S. Modeling height-diameter relationship and volume of teak (*Tectona grandis* L. F.) in central lowlands of Nepal. J Trop. Forest. Environ. 2017; 7:28-42.
25. Lacerda THS, Cabacinha CD, Araújo Jr CA, Maia RD, Lacerda KWS. Artificial neural networks for estimating tree volume in the Brazilian savanna. Cerne. 2017; 23:483-491. doi:10.1590/01047760201723042347
26. Larjavaara M, Muller-Landau HC. Measuring tree height: a quantitative comparison of two common field methods in a moist tropical forest. Methods in Ecology and Evolution, 2013; 4:793-801.
27. Mangiafico SS. Summary and analysis of extension program evaluation in R, version 1.18.1. 2016; https://rcompanion. org/documents/RHandbookProgramEvaluation.pdf. Accessed 20 June, 2020.
28. Montgomery DC, Peck EA, Vining GG. Introduction to linear regression analysis, 3rd edn. John Wiley and Sons. Singapore, 2003, 672.
29. Mugasha WA, Mwakalukwa EE, Luoga E, Malimbwi RE, Zahabu E, Silayo DS, *et al*. Allometric models for estimating tree volume and aboveground biomass in lowland forests of Tanzania. Int. J. For. Res. 2016; 1-13. http://dx.doi.org/10.11 55/2016/8076271
30. Mwangi JR. Volume and biomass estimation models for *Tectona grandis* grown at Longuza forest plantation, Tanzania. MSc. Thesis. Sokoine University of Agriculture, Morogoro, Tanzania, 2015.
31. Nandy S, Singh R, Ghosh S, Watham T, Kushwaha SPS, Kumar AS, *et al*. Neural network-based modelling for forest biomass assessment. Carbon Manag. 2017; DOI: 10.1080/17583004.2017.1357402
32. Nash JE, Sutcliffe JV. River flow forecasting through conceptual models part I-A discussion of principles. J Hydrol. 1970; 10(3):282-290. doi:10.1016/0022-1694(70)90 255-6
33. Newberry JD. Methods for modeling whole stem diameter growth and taper. A Ph.D. Thesis. Virginia Polytechnic Institute and State University, Blacks, VA, 1984.
34. Nieto PJG, Torres JM, Fernández MA, Galán CO. Support vector machines and neural networks used to evaluate paper manufactured using *Eucalyptus globulus*. Appl. Math. Model. 2012; 36:6137-6145.
35. Nunifu TK, Murchinson HG. Provisional yield models of teak (*Tectona grandis* Linn F.) plantations in northern Ghana. For. Ecol. Manag. 1999; 120:171-178.
36. Nunifu TK. The growth and yield of teak (*Tectona grandis* Linn F.) plantations in northern Ghana. MSc Thesis. Lakehead University, Ontario, 1997.
37. Özçelik R, Diamantopoulou MJ, Brooks JR, Wiant HV. Estimating tree bole volume using artificial neural network models for four species in Turkey. J. Environ. Manage. 2010; 91:742-753.
38. Özçelik R, Diamantopoulou MJ, Crecente-Campo F, Eler U. Estimating *Crimean juniper* tree height using nonlinear regression and artificial neural network models. For Ecol Manag. 2013; 306:52-60.
39. Peng C, Zhang L, Liu J. Developing and validating nonlinear height-diameter models for major tree species of Ontario's boreal forests. North. J Appl. For. 2001; 18:87-94.
40. Perez CLD, Kanninen M. Provisional equations for estimating total and merchantable volume for *Tectona grandis* trees in Costa Rica. For. Trees Livelihoods. 2003; 13:345-359.
41. Perez D. Growth and volume equations developed from stem analysis for *Tectona grandis* in Costa Rica. J Trop. For. Sci. 2008; 20:66-75.
42. Shamaki SB, Akindele SO, Isah AD. Development of volume equations for teak plantation in Nimbia forest reserve in Nigeria using dbh and height. J Agric. Environ. 2011; 7:71-76.
43. Sharma M, Oerwald RG, Amateis RL. A consistent system of equations for tree and stand volume. North. J Appl. For. 2000; 54:110-125.

44. Sharma M, Parton J. Height–diameter equations for boreal tree species in Ontario using a mixed-effects modeling approach. For. Ecol. Manag. 2007; 249:187-198. doi:10.1016/j.foreco.2007.05.006
45. Shuaibu RB. Developing stem taper equation for *Tectona grandis* (teak) plantation in Agudu forest reserve, Nasarawa State, Nigeria. NSUK J. Sci. Tech. 2016; 5:199-206.
46. Socha J, Netzel LP, Cywicka D. Stem taper approximation by artificial neural network and a regression set models. Forests. 2020; 11:79. doi:10.3390/f11010079.
47. Tackenberg O. A new method for non-destructive measurement of biomass, growth rates, vertical biomass distribution and dry matter content based on digital image analysis. Ann. Bot. 2007; 99:777-783. doi:10.1093/aob/mcm009
48. Tewari VP, Mariswamy KM, Arunkumar AN. Total and Merchantable Volume Equations for *Tectona grandis* Linn. f. plantations in Karnataka, India. J Sustain. Forest. 2013; 32:213-229. DOI: 10.1080/10549811.2013.762187.
49. Tewari VP, Singh B. Total wood volume equations for *Tectona Grandis* Linn F. stands in Gujarat, India. J For. Environ. Sci. 2018; 34:313-320. https://doi.org/10.7747/JFES.2018.34.4.313
50. Van Zyl L. Stem form, height and volume models for teak in Tanzania. MSc Thesis, University of Stellenbosch, Stellenbosch, 2005.
51. Vendruscolo DGS, Cerqueira CL, e Carvalho SPC, Medeiros RA, da Silva RS. Thickness accuracy of teak bark by artificial intelligence. Floresta, Curitiba, 2019; 49:449-458. dio: 10.5380/rf.v49 i3.59106
52. Vieira GC, De Mendonça AR, Da Silva GF, Zanetti SS, Da Silva MM, dos Santos AR. Prognoses of diameter and height of trees of eucalyptus using artificial intelligence. Sci. Total Environ. 2018; 619-620:1473-1481.
53. Willmott CJ. On the validation of models. Phys. Geogr. 1981; 2(2):184-194. doi:10.1080/02723646.1981.10642213.