American Journal of Environmental Sciences

Research Article

Tree Volume Prediction Model of *Cupressus lusitanica* Species: in North-Western Highlands of Ethiopia

[1]Yishak Adgo Kassie and [2]Abebaw Alene Yallew

[1]*Department of Natural Forest and Climate Change, Ethiopian Forestry Development, Dire Dawa Centre, Dire Dawa, Ethiopia*
[2]*Department of Plantation Forest and Forest Protection, Ethiopian Forestry Development, Addis Ababa, Ethiopia*

Article history
Received: 04-10-2024
Revised: 21-12-2024
Accepted: 14-01-2025

Corresponding Author:
Yishak Adgo Kassie
Department of Natural Forest and Climate Change, Ethiopian Forestry Development, Dire Dawa Centre, Dire Dawa, Ethiopia
Email: yishakadgo21@gmail.com

Abstract: A significant basis of raw materials for the forest products industry is provided by *Cupressus lusitanica,* one of the commercial timber tree species found in the forest. Accurate tree and stand volume development models are necessary for intensive forest management to yield timber. The goal of this study was to develop and validate volume equations for *Cupressus lusitanica* tree species in North-western highlands, Ethiopia. Four tree volume estimation models were developed and validated for *Cupressus lusitanica* plantation forest at Tsarikan. Tsarikan plantation forest was selected purposively, and systematic sampling technique was employed to select four square plots having an area of 100 m^2 each. A total number of 46 trees were observed and diameter at breast height and tree height were measured. Among the computed estimation models, an equation that used both tree *Diameter at Breast Height (*DBH) and height as independent predictors (V = -5.6*10-17+1.25*10-5(DBH2*h)) was selected to be the best fit and suitable model for individual tree volume prediction of *Cupressus lusitanica* monoculture forest depending on the respective standard error of estimation and coefficient of determination values. This volume estimation model can be used by forest growers and managers for sustainable production and utilization of *Cupressus lusitanica* plantation forest at North-Western highlands of Ethiopia, as well as other areas of the country having similar agroecology. The use of species- and site-specific models are strongly recommended for the countries having a varied range of topographical and biophysical situations like Ethiopia.

Keywords: *Cupressus lusitanica*, Model Fitting, Standard Error, Volume Equations

Introduction

Cupressus lusitanica, one of the forest's commercial timber tree species, offers a substantial supply of raw materials to the forest products sector (Faedo de Almeida *et al.*, 2016). For intensive forest management to produce timber, accurate models for tree and stand volume development are required (Shamaki and Akindele, 2014). An equation that enables the accurate estimation of total stem volume is one of the essential parts of a forest growth and yield modelling system (Wang *et al.*, 2017). To develop a volume equation, precise information on the stem volume and relevant predictor factors of the sample trees are required (Silva *et al.*, 2020). Tree volume estimation is essential for forest management tasks like stock assessment, wood valuation, delineating the boundaries of the forest and estimating growth and production (Levick *et al.*, 2016). For volume estimation, allometric models including parameters such as Diameter at Breast Height (DBH), total tree height and occasionally even some measurements of tree form are frequently employed (Adekunle *et al.*, 2013). Numerous allometric tree volume models have been developed for different kinds of tree species and forests in Europe (Forrester *et al.*, 2017).

The earliest and most popular type of models are growth and yield models, which forecast the alterations that tree stands would experience over time (Pretzsch *et al.*, 2015). Tree number estimation has historically been given a lot of weight in forest management (Kumar *et al.*, 2020). A set of equations that allow for an accurate estimation of the total stem volume is one of the most crucial parts of a system for modelling forest growth and yield (Tsega *et al.*, 2019). To construct a volume equation, precise information on the stem volume and relevant predictor factors of the sample trees is required (Yao *et al.*, 2012). Studies on growth and yield, assessments of the forest growing stock, timber value estimation and harvest zone selection all require it (Machado *et al.*, 2015). In volume equations, Height (H) and Diameter at Breast Height (DBH) are frequently used to determine stem volume (Chukwu *et al.*, 2020). By developing an allometric formula and applying field measurements of the diameter at breast height (DBH) and total height, the stem volume can be ascertained (Seo *et al.*, 2015). However, in this study, the Diameter at Breast Height (DBH) and height of each tree were measured and the volume of each tree was calculated using the 0.5 form factor that is the default (Mandal *et al.*, 2020).

These models can be used to investigate the interactions between various elements and analyse the vertical structure of forest communities, although they might not always be reliable predictors unless more exact height values are needed (Vibrans *et al.*, 2015). For instance, determining the stem volume of standing trees is necessary to evaluate the biomass of the forest. Moreover, stem volume estimations are necessary for sustainable forest management (Njana *et al.*, 2016). While there are several tree-sectional volumes, total tree volume models make up the majority of volume models created for sub-Saharan Africa (Mauya *et al.*, 2014). A tree species' rate of diameter growth is mostly dictated by the quality of the site, soil type and the quantity of trees in the stand (Erkan and Aydin, 2016). According to Masota (2014), there is a strong correlation between the diameter at breast height and the overall height of the tree and environmental parameters such as soil nutrients, climate, disturbance, successional status, topographic position, tree species and genetic factors. When calculating volume, this variation presents a problem that necessitates the use of models calibrated for particular circumstances, tree species, or environmental factors (Henry *et al.*, 2015).

In addition to producing timber, in the plantation forest offers the surrounding inhabitants certain ecosystem services. In this study to develops, compares, tests and validates various allometric models to estimate stem tree volume (stem plus branches) for individual *et al.*, 2011). It would be challenging to collect enough data to develop a single universal model that takes into consideration all variables and species, though, because tropical forests involve a wide range of environmental conditions and species (Harrison *et al.*, 2018). Developing species- and climate site-specific models may be the most effective way to address these problems. According to Pathmanathan (2015), argue that as there isn't a single volume model that works for all applications, volume equations need to improve in terms of accuracy, flexibility, validity and normality of their predictions. This study can be a portion of the solution, as the volume estimation equation for *C. lusitanica* trees at Tsarikan plantation forest. Many allometric tree volume models have been developed in Australia (Mauya *et al.*, 2014) for a variety of tree species and forest types. However, very few models have been developed for sub-Saharan African tropical forest types (Henry *et al.*, 2011). Furthermore, questions concerning the accuracy of models used to estimate forest volume in tropical forests are raised by the absence of a wide range of species, tree sizes and geographic locations (Henry *C. lusitanica* in North-Western highlands of Ethiopia has not yet been generated. The aims of this study were to develop an appropriate tree volume estimation model for *C. lusitanica* tree species in Awi zone, Ethiopia.

Materials and Methods

Study Area Description

The study area was conducted in Tsarikan plantation forest, found in Amhara region, Fagita Lekoma District. This district is geographically located between 10^0 57′ 23″-$11^0$11′ 21″ N and 36^0 40′ 01″-37^0 50′ 21″E. The average annual temperature and rainfall are 24 ^{o}C and 2874.5 mm, respectively and the elevation ranges from 2630-2750 m above sea level (m.a.s.l). Tsarikan' state forest has an area of about 72 hectares and is located 12 km north of Injibara. The forest was established around 1972 E.C. artificially by order of the state in order to cover the mountainous areas with forest rather than cultivated land for environmental purposes and firewood. The forest has no forest roads, which are much needed infrastructure for forest management (Fig. 1).

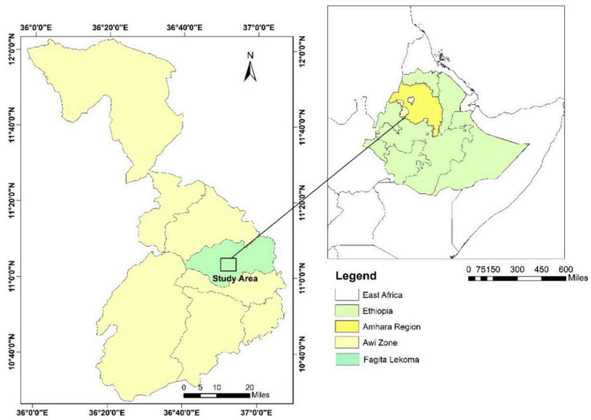

Fig. 1: Map of the study area

Climate and Topography

The forest site is found in dega climatic zone, mean annual temperature of the area is11-22°C and the altitude of the forest area ranges from 2630-2750 m above sea level. The rainy season starts from May and ends at October which is unimodal (one rainy season). The amount of mean annual rain fall is estimated about 1300-1500 mm. The major land form of the plantation forest Area is dominated by slopes up to 40%. Whether the classification of terrain is of macro or micro scale there are primarily three terrain features which are of great importance for the tree harvesting work: ground condition, ground roughness, slope. The carrying capacity of the ground that is, its capacity to resist physical pressure, primarily depends on type and moisture of the soils. Therefore, Tsarikan forest ground condition is good especially during dry season. The surface structure influencing a particular set of operations depends on the nature, size and number of the obstructions. Humic Nistisols are the most common soil types in this area.

Sampling Design and Technique

In the study area a systematic sampling technique was employed. For plantation forests with well-defined spacing and shape, a square plot is relatively more suitable than a circular plot. Therefore, a square plot within 10□10 m (100 m^2) was considered for data collection. The interval between consecutive sample plots was 80 meters apart based on the length of the forest which is actually about 300 meters. In order to reduce the boarder effect, the sample plots were laid- out 20 m away from the edge of the forest both sides. To gather the volume data, five trees that were closest to the sample plots' centre were chosen for stem examination. After measuring each of the 46 sample trees' Diameters at Breast Height (DBH), the trees were felled and divided into five equal-length sections that were spaced 30 cm from the tree's tip to the stump height. The actual tree volume for *C. lusitanica* was determined using volume formula.

A tape was used to measure the length of each segment and a calliper was used to measure the mid-diameter over bark in two opposing directions perpendicular to the longitudinal axis of the tree's bole. The total height (m) from stump to tip height (m) was then calculated by adding the segment lengths. Using Huber's formula, the total volume over bark (without stump volume) of each sample tree (V m^3) was determined. Because of its higher accuracy and ease of use, the Huber method was employed to estimate the cubic volume of logs and trees (Schikowski *et al.*, 2018). The volume of each tree was determined using its own diameter, height and using the common form factor (0.5). For 35 (75%) data's the initial model development and 11 (25%) data for model validation, all of the data set was used. four volume estimation equation forms were used.

Data Analysis

The most popular technique for locating estimators of an equation's parameters is the least squares method. Because the variation of the tree volume is not homogeneous, it is challenging to fit stem volume data using least squares (Cunia, 1964). Using the weighted least squares approach to estimate the regression parameters is one strategy to account for the variance's non-homogeneity (WLS). Four (4) tree volume estimation models were considered and the best suitable model was selected. 75% of trees were used for model development by using the general form factor for this specific agroecology (0.5), while the remaining 25% were used for the validation of the developed model (Fischer *et al.*, 2021). The parameters of these models were estimated using the Statistical Analysis System non-linear procedure. In order to evaluate the performance of these models, the standard error of estimate (SEE) and coefficient of determination (R^2) were determined. To identify the best fit model(s), the model with the lowest values of SEE and the model with the R^2 value closest to 1 were considered. The tree volume equations' exponent values are examined using the t-test under the assumption that the residuals are normally distributed (Teshome, 2005). In order to determine whether the weighted residuals from the four equations were normal, the (Duchnowski and Wyszkowska, 2020) test was used (Table 1).

Table 1: Volume equations used for model development and validation

Model code	Model form
1	$V = a\ dbh^b$
2	$V = a+b(dbh)$
3	$V = a+b(dbh^2)$
4	$V = a+b(dbh^2*h)$

V= volume (m 3); h = tree total height (m); DBH = tree diameter at breast height (cm); "a "and "b" are the estimated parameters

Results

Model Development and Evaluation

From this result, four volume model forms were chosen and further tested in the *C. lusitanica* species using the extensive mensuration literature based on evaluations of tree volume models in various forest types. Three of the models have considered only DBH as an independent variable, whereas both parameters (DBH and H) are considered independent variables in the fourth model. The estimated volume and measured volume in m^3 were then computed. The value of DBH was measured in centimetres and then converted to metres in order to calculate tree volume. It was also noticed that the standard errors for each of the predictive parameters varied between models. The performance of the models was assessed using the four-fit statistics shown below (Table 2). The coefficients for model parameters were identified and the standard errors for each of the predicting parameters were also observed to be variable from model to model. Using the four-fit statistics, the performance of the models was evaluated as below:

$$\textit{Model 1}\left(V = a\ DBH^b\right) \tag{1}$$

The given non-linear or exponential model was changed into a linear equation by taking both sides into the natural logarithm that is indicated below (Table 2).

Table (2), the parameters a and b are determined as follows: the intercept 'a' = -10.36 and 'b' = 2.70, but the value of 'a' was converted to an exponent. i.e., '10.37 = 0.00000314. Therefore, the fitted allometric equation to obtain the tree volume of *C. lusitanica* tree species by using DBH (m) as an explanatory variable was developed as V= $0.00000314dbh^{2.7037}$ it shows in the (Table 2):

$$Model\ 2\,(V = a + b\,(DBH)) \tag{2}$$

Table 2: Regression statistics for the first model

Regression Statistics					
Multiple R	0.98				
R Square	0.96				
Adjusted R Square	0.96				
Standard Error	0.18				
Observations	31				
ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	21.59	21.59	684.56	1.01266E-21
Residual	29	0.91	0.03		
Total	30	22.50			

Table (3), shows that the values of parameters 'a' and 'b' are -0.19595 and 0.016112, respectively and the fitted allometric equation to obtain the tree volume of *C. lusitanica* tree species by using DBH as an explanatory variable was V = -0.19595 + 0.016112 (DBH) OR V = 0.016112 (DBH) - 0.19595:

$$Model\ 3\,(V = a + b\,(DBH2)) \tag{3}$$

Table 3: Regression statistics for the second model

Regression Statistics					
Multiple R	0.95				
R Square	0.90				
Adjusted R Square	0.89				
Standard Error	0.03				
Observations	31				
ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.31	0.31	262.66	4.49E-16
Residual	29	0.03	0.001		
Total	30	0.34			

Table (4), shows that the values of parameters 'a' and 'b' are also estimated as -0.03869 and 0.000378, respectively and the fitted allometric equation to obtain the tree volume of *C. lusitanica* tree species by using DBH as an explanatory variable was V = 0.000378 (DBH) - 0.03869:

$$Model\ 4\,(a + b\,(DBH2 * h)) \tag{4}$$

Table (5), shows that the values of parameter 'a'= $-5.6*10^{-17}$ and parameter 'b'=10.0000125 and the fitted allometric equation to estimate tree volume of *C. lusitanica* tree species by using DBH and height as an explanatory variable was V = $-5.6*10^{-17}+1.25*10^{-5}(dbh^2*h)$.

Table 4: Regression statistics for the third model

Regression Statistics					
Multiple R	0.97				
R Square	0.95				
Adjusted R Square	0.95				
Standard Error	0.02				
Observations	31				
ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.32	0.32	529.67	3.54E-20
Residual	29	0.02	0.00		
Total	30	0.34			

Table 5: Regression statistics for the fourth model

Regression Statistics					
Multiple R	1				
R Square	1				
Adjusted R Square	1				
Standard Error	3.12442E-17				
Observations	31				
ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.34	0.34	3.49E+32	0
Residual	29	2.83E-32	9.76E-34		
Total	30	0.34			

Discussion

Model Goodness of Fit

Standard Error of Estimate (SEE) and coefficient of determination (R^2) are the most commonly used methods to measure goodness of fit (Chicco *et al*., 2021). As this results the value of the coefficient of determination (R^2) closes to 1 and as the value of the standard error of estimate becomes low, the model is said to be well fitted (Table 6).

Table 6: Allometric equations and respective error of estimate and coefficient of determination

No.	Allometric Equations	Error of Estimate	Coefficient of Determination
1	V = $0.00000314dbh^{2.7037}$	0.177	0.9595
2	V = 0.016112(DBH) - 0.19595	0.0342	0.9
3	V = 0.000378(DBH) - 0.03869	0.025	0.948
4	V=$-5.6*10^{-17}+1.25*10^{5}(dbh^2*h)$	3.12*10-17	1.0

Table (6), shows that in the fourth model, the Standard Error of Estimate (SEE) is nearly negligible and the coefficient of determination (R^2) is equal to 1, it indicating that the model is relatively the best fit to use for the estimation of *C. lusitanica* tree species in this study area.

Model Validation

Model goodness of fit alone cannot justify the utilization of a model, whereas it must be validated whether the fitted model is good or not (Barrett, 2007). A T-test-tested for testing the goodness of fit (Table 7).

Model 1 ($V = aDBH^b$)

The computed t-statistical test and respective p value for the first model are 4.4159 and 0.00084, respectively. If the value of P is greater than 0.05 at a 95% confidence interval, then the statistic test is non-significant and vice versa (Amrhein *et al.*, 2017). In this case, the value of P= 0.00084, which is less than 0.05, indicating that there is a significant difference between the observed value and the predicted value. Therefore, the fitted allometric equation cannot be used to predict the individual tree volume of the *C. lusitanica* tree species (Table 7).

Table 7: t-Test Paired two sample for means

	Variable 1	Variable 2
Mean	0.25	0.02
Variance	0.03	0.00
Observations	10	10
Pearson Correlation	0.97	
Hypothesized Mean Difference	0	
Df	9	
t Stat	4.42	
P(T<=t) one-tail	0.00	
t Critical one-tail	1.83	
P(T<=t) two-tail	0.00	
t Critical two-tail	2.26	

Model 2 ($V = a+b(DBH)$)

From this result the computed t-statistical test is 1.345 and the p-value is 0.1057. The computed statistical test indicated that there is no significant difference between the observed tree volume and the predicted tree volume. Therefore, this fitted allometric equation can be used to predict individual tree volumes in the *C. lusitanica* tree species (Table 8).

Table 8: t-Test Paired two sample for means

	Variable 1	Variable 2
Mean	0.25	0.22
Variance	0.03	0.02
Observations	10	10
Pearson Correlation	0.94	
Hypothesized Mean Difference	0	
Df	9	
t Stat	1.35	
P(T<=t) one-tail	0.11	
t Critical one-tail	1.83	
P(T<=t) two-tail	0.21	
t Critical two-tail	2.26	

Model 3 ($V = a+b(DBH^2)$)

From this finding the computed t-statistic and P value are 0.932 and 0.1878, respectively. The computed statistical test indicated the difference is not significant between observed and predicted tree volume and then the fitted allometric model can be used to predict individual tree volume in the *C. lusitanica* tree species (Table 9).

Table 9: t-Test paired two sample for means

	Variable 1	Variable 2
Mean	0.25	0.23
Variance	0.03	0.02
Observations	10	10
Pearson Correlation	0.96	
Hypothesized Mean Difference	0	
Df	9	
t Stat	0.93	
P(T<=t) one-tail	0.19	
t Critical one-tail	1.83	
P(T<=t) two-tail	0.38	
t Critical two-tail	2.26	

Model 4 ($a+b(DBH^2*h)$)

Table (10), shows that the computed t-test for the fourth allometric model is 1.38479 with a respective P-value of 0.09974, indicating a non-significant difference between observed and predicted tree volume and hence the fitted allometric model can be used to predict individual tree volume in the *C. lusitanica* tree species. Among the different models used, the fourth model 4 ($a+b(dbh^2*h)$) estimates individual tree volume using DBH and total height as predictor variables. The standard error estimate of this model is almost negligible (3.12*10-17) when compared to the other three models. Additionally, the coefficient of determination is greater than the rest of the models usually = 1. The lowest estimate of Standard Error (SEE) and the highest coefficient of determination (R^2) make the model more suitable and recommended for predicting the individual tree volume of *C. lusitanica* species in this study.

Table 10: t-Test Paired Two Sample for Means

	Variable 1	Variable 2
Mean	0.24	0.25
Variance	0.03	0.03
Observations	10	10
Pearson Correlation	0.99	
Hypothesized Mean Difference	0	
Df	9	
t Stat	1.38	
P(T<=t) one-tail	0.09	
t Critical one-tail	1.83	
P(T<=t) two-tail	0.19	
t Critical two-tail	2.26	

As a result, the best tree volume estimation model was found to be V = a + b (DBH^2*h), if both DBH and tree height were used as predictors. Moreover, the model (V = 0.000378 (DBH) - 0.03869) can also be used for tree volume estimation using tree DBH only as a predictor. This study area shown how air temperature is the main factor influencing *C. lusitanica* growth. This outcome is in line with a productivity model formed for young, densely stocked *C. lusitanica*, which discovered that the mean annual air temperature had the strongest correlation with volume mean annual increment (Watt *et al.*, 2008). Additionally, this result supports anecdotal reports that the species favors warm locations and climates (Hoey *et al.*, 2016). The temperature optimum for *C. lusitanica* is not reached in this region's temperate environment, according to the positive linear association between mean minimum air temperature and site index.

The variances of the residuals were homogeneously distributed, as demonstrated by the best model scatter plots of weighted values (Stöckl *et al.*, 2014). In general, it was discovered that the three models that solely used DBH as a predictor performed worse than the models that included both height and DBH. Despite this, some models performed worse when it came to *C. lusitanica* volume prediction (Isaac, 2018). This study's productivity determinants significantly improve knowledge of how the environment controls *C. lusitanica* growth. The findings clearly show that the species prefers warm, fertile locations with sufficient root depth. The study's findings regarding the relationship between productivity and root depth indicate that the species needs deep soils. One of the main physical characteristics of soil that has been identified as affecting forest production is soil depth. Shallow soils change the availability of water in addition to limiting production by limiting access to nutrients (Calvaruso *et al.*, 2017). Although these conclusions are in line with anecdotal reports, they go beyond them by outlining the functional forms of driving variables and their proportional significance to total output.

Since *C. lusitanica* is sensitive to site fertility there is a lot of room to improve the model's representation of site fertility (Burdon and Moore, 2018). In contrast to deforest/mixed scrubland sites, the results reported here clearly demonstrate a significant increase in stands developed on ex-pasture/cropland/grassland sites, which are more likely to have had a history of fertilization. According to earlier studies, the commonly planted plantation species P. radiata is more productive on ex-pasture sites (Beets *et al.*, 2019).Therefore, future studies should concentrate on generating geographical layers for soil chemical characteristics such the soil C:N ratio, which have been shown to affect *C. lusitanica* productivity in the past (Mauritsson, 2018). Predicting the site index for this species is probably going to be better with the use of layers for continuous factors characterizing soil fertility. Despite the model's applicability across wide gradients, there wasn't enough information available for about low rainfall and extremely low temperatures both have a significant negative impact on productivity, as demonstrated by a prior model formed using data from young, heavily stocked plots. In order to ascertain how rainfall affects productivity on dryland sites and to validate the shape of the air temperature relationship at low values, more measurements from mature trees from cold and dry locations should be made.

Model Testing, Fitting and Validation

The selected equations were evaluated for prognostication, accuracy and precision using test data that was suppressed. Three factors were used by Özçelik and Cao (2017) to evaluate the equations in terms of R: (a) bias (the mean of the variations between the projected and observed volumes); (b) mean absolutes differences; and (c) standard deviation differences, which are also referred to as standard error of estimation (SEE). As previously reported in other investigations, the study was able to fully fit the association with a high model fit (Kuria *et al.*, 2019). Two techniques for evaluating regression models based on statistical fitness or prediction errors derived from ordinary residuals were developed by (Alexander *et al.*, 2015). The first technique compares models using statistics obtained directly from models derived from entire data sets, whereas the second process uses a validation data set, which usually comprises less than or half of the full data set (Cawley and Talbot, 2010). The results are compared with the real observations and the model is used to forecast the behaviour of the forest where the test data were obtained (Blanco *et al.*, 2007).

The behaviour of the stands that generated the test data is predicted by the model and the outcomes are compared to the actual observations (Pretzsch, 2009). From this result based on its R^2 and standard error value, the top model (model 4) was chosen after the first model was formed. It demonstrates that model 4 was the top-performing model and performed the best in practically every fit criterion (Sharma *et al.*, 2021). Dependable models and input data are necessary for reliable predictions (Blischke and Murthy, 2000). To ascertain a yield prediction model's accuracy and validity, testing is necessary (Haboudane *et al.*, 2004). Nonetheless, the model's precision is dependent on the sample's ability to accurately reflect the forest, the quantity and duration of remeasurements, the predictor variables' covariances and the model's coefficients (Babcock *et al.*, 2016). One method of fitting field data to a pre-established regression is one way to fit a yield prediction model from forest data (Palanivel and Surianarayanan, 2019). Comparing model simulation with growth and yield measurements is the most effective method of validating a process-based model (Zhou *et al.*, 2005).The easiest way to test a model is to use another set of forest data that wasn't used to produce the yield model (Weiskittel *et al.*, 2011). A model's validity and precision must be assessed once it has been built and fitted to data (Steyerberg *et al.*, 2010). The standard error value and coefficient of determination from the first model's development in this study were utilized to establish that model 4 was the best.

Conclusion and Recommendation

Developing forest management plan is basic activity for sustainable forest production and utilization of forest products. Forest inventory is very challenging and resource demanding activity in forestry and it is a common census that the use of precise forest yield prediction models is preferable rather than direct measurement. Accurate model was developed to estimate volume for *C. lusitanica* species in the highlands of Amhara region. A model that used both DBH and height as independent variables is the most suitable for individual tree volume prediction. Moreover, an alternative model that used DBH alone as independent variable is also developed. These findings demonstrate how useful theme spatial layers are as motivating factors for assembly productivity models. When compared to national models that were previously formed using observed point data for *C. lusitanica*, the accuracy of the model provided in this paper was favorable.

The use of spatial layers to generate models is highly supported by this accuracy and the compatibility of important driving variables with those previously reported for *C. lusitanica*. As more factors, like the chemical characteristics of the soil, become available, models formed from these layers should get better. This method significantly lowers the cost of developing the model and the comprehensive maps offer crucial decision assistance for identifying the best locations for species like *C. lusitanica*. General equations used for yield prediction can lead either to over-estimation or under-estimation. However, for countries having a wide range of geographical and biophysical conditions and diverse tree species like Ethiopia, the use of species- and site-specific models are strongly recommended for precise prediction. For forest management choices concerning *C. lusitanica* plantations in Ethiopia, the designed models must be properly documented and included into forest information systems. Further research is recommended to investigate the application of the developed models to other forest areas where this tree species is grown. A better approach would be to develop region-specific models for all forest plantation areas in Ethiopia.

Acknowledgment

We would like to thank Almighty God, the son of St. Mary, who creates everything from nothing and who gave us the mind to think and helped us in all our duties to accomplish this study.

Funding Information

The authors have not received any financial support or funding for this study

Author’s Contributions

Both the authors have equally contributed to this manuscript.

Ethics

The authors declares that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this study and that this study adheres to ethics

References

Adekunle, V. A. J., Nair, K. N., Srivastava, A. K., & Singh, N. K. (2013). Models and form factors for stand volume estimation in natural forest ecosystems: a case study of Katarniaghat Wildlife Sanctuary (KGWS), Bahraich District, India. *Journal of Forestry Research*, *24*(2), 217-226. https://doi.org/10.1007/s11676-013-0347-8

Alexander, D. L. J., Tropsha, A., & Winkler, D. A. (2015). Beware of R2: Simple, Unambiguous Assessment of the Prediction Accuracy of QSAR and QSPR Models. *Journal of Chemical Information and Modeling*, *55*(7), 1316-1322. https://doi.org/10.1021/acs.jcim.5b00206

Amrhein, V., Korner-Nievergelt, F., & Roth, T. (2017). he earth is flat (p > 0.05): significance thresholds and the crisis of unreplicable research. *PeerJ*, *5*, e3544. https://doi.org/10.7717/peerj.3544

Babcock, C., Finley, A. O., Cook, B. D., Weiskittel, A., & Woodall, C. W. (2016). Modeling forest biomass and growth: Coupling long-term inventory and LiDAR data. *Remote Sensing of Environment*, *182*, 1-12. https://doi.org/10.1016/j.rse.2016.04.014

Beets, P. N., Kimberley, M. O., Garrett, L. G., Paul, T. S. H., & Matson, A. L. (2019). Soil productivity drivers in New Zealand planted forests. *Forest Ecology and Management*, *449*, 117480. https://doi.org/10.1016/j.foreco.2019.117480

Blanco, J. A., Seely, B., Welham, C., Kimmins, J. P. (Hamish), & Seebacher, T. M. (2007). Testing the performance of a forest ecosystem model (FORECAST) against 29 years of field data in a*Pseudotsuga menziesii*plantation. *Canadian Journal of Forest Research*, *37*(10), 1808-1820. https://doi.org/10.1139/x07-041

Blischke, W. R., & Murthy, P. (2000). *Reliability: modeling, prediction and optimization*.

Burdon, R. D., & Moore, J. R. (2018). Adverse Genetic Correlations and Impacts of Silviculture Involving Wood Properties: Analysis of Issues for Radiata Pine. *Forests*, *9*(6), 308. https://doi.org/10.3390/f9060308

Calvaruso, C., Kirchen, G., Saint-André, L., Redon, P.-O., & Turpault, M.-P. (2017). Relationship between soil nutritive resources and the growth and mineral nutrition of a beech (Fagus sylvatica) stand along a soil sequence. *CATENA*, *155*, 156-169. https://doi.org/10.1016/j.catena.2017.03.013

Cawley, G. C., & Talbot, N. L. (2010). On over-fitting in model selection and subsequent selection bias in performance evaluation. *The Journal of Machine Learning Research*, *11*, 2079-2107.

Chicco, D., Warrens, M. J., & Jurman, G. (2021). The Coefficient of Determination R-Squared is More Informative Than SMAPE, MAE, MAPE, MSE and RMSE in Regression Analysis Evaluation. *PeerJ Computer Science*, *7*, 623. https://doi.org/10.7717/peerj-cs.623

Chukwu, O., Ogana, F. N., & Nwatu, J. U. (2020). Comparison of Volumes Estimated from Breast Height and Stump Diameters: Application to Tectona grandis Data. *Forest Science*, *66*(5), 551-555. https://doi.org/10.1093/forsci/fxaa015

Cunia, T. (1964). Weighted least squares method and construction of volume tables. *Forest Science*, *10*(2), 180-191. https://doi.org/10.1093/forestscience/10.2.180

Duchnowski, R., & Wyszkowska, P. (2020). Testing normality of chosen R-estimates used in deformation analysis. *Journal of Geodetic Science*, *10*(1), 7-13. https://doi.org/10.1515/jogs-2020-0002

Erkan, N., & Aydin, A. C. (2016). Effects of spacing on early growth rate and carbon sequestration in Pinus brutia Ten. plantations. *Forest Systems*, *25*(2), e064. https://doi.org/10.5424/fs/2016252-09290

Faedo de Almeida, C. C., D'Angelo Rios, P., Bayestorff da Cunha, A., Melo Ampessan, C. G., & Spanhol, A. (2016). Applicability evaluation of Cupressus lusitanica for pulp production. *Maderas. Ciencia y Tecnología*, *18*(4), 651-662. https://doi.org/10.4067/s0718-221x2016005000057

Fischer, G., Nachtergaele, F., van Velthuizen, H., Chiozza, F., Francheschini, G., Henry, M., Muchoney, D., & Tramberend, S. (2021). Global Agro-Ecological Zones (GAEZ v4) Model Documentation. *FAO & IIASA*. https://pure.iiasa.ac.at/id/eprint/17175/1/GAEZ4_model_documentation+app_20210415_withFAOstyles_covers.pdf

Forrester, D. I., Tachauer, I. H. H., Annighoefer, P., Barbeito, I., Pretzsch, H., Ruiz-Peinado, R., Stark, H., Vacchiano, G., Zlatanov, T., Chakraborty, T., Saha, S., & Sileshi, G. W. (2017). Generalized biomass and leaf area allometric equations for European tree species incorporating stand structure, tree age and climate. *Forest Ecology and Management*, *396*, 160-175. https://doi.org/10.1016/j.foreco.2017.04.011

Haboudane, D. (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sensing of Environment*, *90*(3), 337-352. https://doi.org/10.1016/j.rse.2003.12.013

Harrison, X. A., Donaldson, L., Correa-Cano, M. E., Evans, J., Fisher, D. N., Goodwin, C. E. D., Robinson, B. S., Hodgson, D. J., & Inger, R. (2018). A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ*, *6*, e4794. https://doi.org/10.7717/peerj.4794

Henry, M., Cifuentes Jara, M., Réjou-Méchain, M., Piotto, D., Michel Fuentes, J. M., Wayson, C., Alice Guier, F., Castañeda Lombis, H., Castellanos López, E., Cuenca Lara, R., Cueva Rojas, K., Del Águila Pasquel, J., Duque Montoya, Á., Fernández Vega, J., Jiménez Galo, A., López, O. R., Marklund, L. G., Milla, F., de Jesús Návar Cahidez, J., … Saint-André, L. (2015). Recommendations for the use of tree models to estimate national forest biomass and assess their uncertainty. *Annals of Forest Science*, *72*(6), 769-777. https://doi.org/10.1007/s13595-015-0465-x

Henry, M., Maniatis, D., Gitz, V., Huberman, D., & Valentini, R. (2011). Implementation of REDD+ in sub-Saharan Africa: state of knowledge, challenges and opportunities. *Environment and Development Economics*, *16*(4), 381-404. https://doi.org/10.1017/s1355770x11000155

Hoey, A., Howells, E., Johansen, J., Hobbs, J.-P., Messmer, V., McCowan, D., Wilson, S., & Pratchett, M. (2016). Recent Advances in Understanding the Effects of Climate Change on Coral Reefs. *Diversity*, *8*(2), 12. https://doi.org/10.3390/d8020012

Isaac, K. (2018). *Spatio-Temporal Degradation Detection And Modeling Future Scenarios Of Embobut Forest In Elgeyo Marakwet County*.

Kumar, I., Gupta, R. K., & Kumar, A. (2020). Construction of Two-Way Volume Table of Eucalyptus Tereticornis for Mid-Himalayan Region of Himachal Pradesh. *International Journal of Bio-Resource and Stress Management*, *11*(1), 046-050. https://doi.org/10.23910/ijbsm/2020.11.1.2024b

Kuria, C. N., Kirongo, Balozi B., & Kipkore, W. (2019). Growth and yield models for plantation-grown cupressus lusitanica for central Kenya. *African Journal of Education, Science and Technology*, *5*(2), 34-58.

Levick, S. R., Hessenmöller, D., & Schulze, E.-D. (2016). Scaling wood volume estimates from inventory plots to landscapes with airborne LiDAR in temperate deciduous forest. *Carbon Balance and Management*, *11*(1), 7. https://doi.org/10.1186/s13021-016-0048-7

Machado, R. R., Conceição, S. V., Leite, H. G., Souza, A. L. de, & Wolff, E. (2015). Evaluation of forest growth and carbon stock in forestry projects by system dynamics. *Journal of Cleaner Production*, *96*, 520-530. https://doi.org/10.1016/j.jclepro.2013.09.049

Mandal, R. A., Khadka, G. B., Shrestha, M., Sah, S. P., & Lamichhane, A. (2020). Modeling the diameter at breast height (DBH) with height and volume of Shorea robusta using destructive method: A study from Banke District, Nepal. *Discovery*, *56*(293), 239-253.

Masota, A. M. (2014). Volume Models for Single Trees in Tropical Rainforests in Tanzania. *Journal of Energy and Natural Resources*, *3*(5), 66-76. https://doi.org/10.11648/j.jenr.20140305.12

Mauritsson, K. (2018). *Ecological and Edaphic Correlations of Soil Invertebrate Community Structure in Dry Upland Forests of Eastern Africa*.

Mauya, E. W., Mugasha, W. A., Zahabu, E., Bollandsås, O. M., & Eid, T. (2014). Models for estimation of tree volume in the miombo woodlands of Tanzania. *Southern Forests: A Journal of Forest Science*, *76*(4), 209-219. https://doi.org/10.2989/20702620.2014.957594

Njana, M. A., Meilby, H., Eid, T., Zahabu, E., & Malimbwi, R. E. (2016). Importance of tree basic density in biomass estimation and associated uncertainties: a case of three mangrove species in Tanzania. *Annals of Forest Science*, *73*(4), 1073-1087. https://doi.org/10.1007/s13595-016-0583-0

Özçelik, R., & Cao, Q. V. (2017). Evaluation of Fitting and Adjustment Methods for Taper and Volume Prediction of Black Pine in Turkey. *Forest Science*, *63*(4), 349-355. https://doi.org/10.5849/fs.2016-067

Palanivel, K., & Surianarayanan, C. (2019). An Approach for Prediction of Crop Yield Using Machine Learning and Big Data Techniques,. *International Journal of Computer Engineering and Technology*, *10*(3), 110-118. https://doi.org/10.34218/ijcet.10.3.2019.013

Pathmanathan, H. (2015). *Strengthening Resilience and Food Security Through the Development of Agroforestry in Post-War Kilinochchi, Sri Lanka*.

Pretzsch, H. (2009). Forest Dynamics, Growth, and Yield. *Forest Dynamics, Growth and Yield: From Measurement to Model*, 1-39. https://doi.org/10.1007/978-3-540-88307-4_1

Pretzsch, H., Forrester, D. I., & Rötzer, T. (2015). Representation of species mixing in forest growth models. A review and perspective. *Ecological Modelling*, *313*, 276-292. https://doi.org/10.1016/j.ecolmodel.2015.06.044

Schikowski, A. B., Corte, A. P. D., Ruza, M. S., Sanquetta, C. R., & Montaño, R. A. N. R. (2018). Modeling of stem form and volume through machine learning. *Anais Da Academia Brasileira de Ciências*, *90*(4), 3389-3401. https://doi.org/10.1590/0001-3765201820170569

Seo, Y. O., Lumbres, R. I. C., Won, H. K., Jung, S. C., & Lee, Y. J. (2015). Evaluation and validation of stem volume models for Quercus glauca in the subtropical forest of Jeju Island, Korea. *Journal of Ecology and Environment*, *38*(4), 485-491. https://doi.org/10.5141/ecoenv.2015.051

Shamaki, S. B., & Akindele, S. O. (2014). Modelling diameter distribution for a seven year Eucalypus camaldulensis stand in Afaka Forest Reserve-Nigeria using Weibull function. *Journal of Agriculture and Environment*, *10*(1 & 2), 45-53.

Sharma, P. N., Shmueli, G., Sarstedt, M., Danks, N., & Ray, S. (2021). Prediction-Oriented Model Selection in Partial Least Squares Path Modeling. *Decision Sciences*, *52*(3), 567-607. https://doi.org/10.1111/deci.12329

Silva, V. S. da, Silva, C. A., Mohan, M., Cardil, A., Rex, F. E., Loureiro, G. H., Almeida, D. R. A. de, Broadbent, E. N., Gorgens, E. B., Dalla Corte, A. P., Silva, E. A., Valbuena, R., & Klauberg, C. (2020). Combined Impact of Sample Size and Modeling Approaches for Predicting Stem Volume in Eucalyptus spp. Forest Plantations Using Field and LiDAR Data. *Remote Sensing*, *12*(9), 1438. https://doi.org/10.3390/rs12091438

Steyerberg, E. W., Vickers, A. J., Cook, N. R., Gerds, T., Gonen, M., Obuchowski, N., Pencina, M. J., & Kattan, M. W. (2010). Assessing the Performance of Prediction Models. *Epidemiology*, *21*(1), 128-138. https://doi.org/10.1097/ede.0b013e3181c30fb2

Stöckl, D., Van Uytfanghe, K., Van Aelst, S., & Thienpont, L. M. (2014). A statistical basis for harmonization of thyroid stimulating hormone immunoassays using a robust factor analysis model. *Clinical Chemistry and Laboratory Medicine (CCLM)*, *52*(7), 965-972. https://doi.org/10.1515/cclm-2013-1038

Teshome, T. (2005). Analysis of individual tree volume equations for*Cupressus lusitanica*in Munessa forest, Ethiopia. *The Southern African Forestry Journal*, *203*(1), 27-32. https://doi.org/10.2989/10295920509505215

Tsega, M., Guadie, A., Teffera, Z. L., Belayneh, Y., & Niu, D. (2019). Development and validation of a stem volume equation for*Cupressus lusitanica*in Gergeda Forest, Ethiopia. *Southern Forests: A Journal of Forest Science*, *81*(1), 79-84. https://doi.org/10.1080/02571862.2018.1512786

Vibrans, A. C., Moser, P., Oliveira, L. Z., & Maçaneiro, J. P. de. (2015). Height-Diameter Models For Three Subtropical Forest Types In Southern Brazil. *Ciência e Agrotecnologia*, *39*(3), 205-215. https://doi.org/10.1590/s1413-70542015000300001

DOI: 10.3844/ajessp.2025.79.88

Wang, X., Bi, H., Ximenes, F., Ramos, J., & Li, Y. (2017). Product and Residue Biomass Equations for Individual Trees in Rotation Age Pinus radiata Stands under Three Thinning Regimes in New South Wales, Australia. *Forests*, *8*(11), 439. https://doi.org/10.3390/f8110439

Watt, M. S., Davis, M. R., Clinton, P. W., Coker, G., Ross, C., Dando, J., Parfitt, R. L., & Simcock, R. (2008). Identification of key soil indicators influencing plantation productivity and sustainability across a national trial series in New Zealand. *Forest Ecology and Management*, *256*(1-2), 180-190. https://doi.org/10.1016/j.foreco.2008.04.024

Weiskittel, A. R., Hann, D. W., Kershaw, J. A., & Vanclay, J. K. (2011). *Forest growth and yield modeling*.

Yao, W., Krzystek, P., & Heurich, M. (2012). Tree species classification and estimation of stem volume and DBH based on single tree extraction by exploiting airborne full-waveform LiDAR data. *Remote Sensing of Environment*, *123*, 368-380. https://doi.org/10.1016/j.rse.2012.03.027

Zhou, X., Peng, C., Dang, Q.-L., Chen, J., & Parton, S. (2005). Predicting forest growth and yield in northeastern Ontario using the process-based model of TRIPLEX1.0. *Canadian Journal of Forest Research*, *35*(9), 2268-2280. https://doi.org/10.1139/x05-149