Trees, Forests and People 5 (2021) 100103

Contents lists available at ScienceDirect

Trees, Forests and People

journal homepage: www.elsevier.com/locate/tfp

Developing taper equations for planted teak (*Tectona grandis* L.f.) trees of central lowland Nepal

Anil Koirala[a,*], Cristian R. Montes[a], Bronson P. Bullock[a], Bishnu H. Wagle[b,1]

[a] Plantation Management Research Cooperative, Warnell School of Forestry and Natural Resources, University of Georgia, 180 E Green Street, Athens GA 30602, USA
[b] Institute of Forestry, Pokhara Campus, Tribhuvan University, Hariyokharka – 15, Pokhara, Nepal

ARTICLE INFO

Keywords:
Cross-validation
Destructive sampling
Dynamic taper model
Hardwood trees
Lowland Nepal

ABSTRACT

Foresters rely on taper models to estimate total and merchantable volume for commercial species. While there is an abundance of literature in taper modeling for many softwood and hardwood species, little is known of teak tree stem form. Taper modeling is in early stage for the forestry sector in Nepal. Currently, no publicly available taper equations exist for teak trees in Nepal. Therefore, the main goal of this study was to develop a taper equation for valuable teak trees of lowland Nepal. Destructive sampling was carried out in teak plantations in the Sagarnath Forestry Development Project. Five common taper equations (simple, segmented, and variable exponent types) from the literature as well as one dynamic taper equation were considered as candidate models. Our results showed that the nonlinear dynamic taper equation performed best in terms of fit and cross-validation statistics with least error and bias. This new taper model can be considered as a step forward for hardwood forest management in Nepal. It is expected that this equation will assist forest managers to predict stem volume and diameters to any merchantable limit for teak growing in similar site conditions.

1. Introduction

Teak (*Tectona grandis* L.f.) is a valuable hardwood species whose natural range encompasses regions of India, Myanmar, Thailand and Laos in the south and south East Asia (Kenzo et al., 2020; Tewari et al., 2014). Like other highly planted hardwoods (e.g. eucalyptus), teak is also planted around the world, particularly in the tropical areas of Caribbean, Africa, Asia and the Americas (Koirala et al., 2017). Teak is planted in about 70 countries around the world and has attracted large private investmentsIn global market, it is one of the highly demanded hardwood species for high-end furniture (Moya et al., 2014). Myanmar is the largest exporter of teak lumber accounting for about 50% of the overall exports in 2014, i.e. more than 1 million m^3 per year (Kollert and Walotek, 2015). Teak trees have desirable physical and esthetic qualities that are suitable for luxury outdoor and indoor furniture such as weathering resistance, seasoning without splitting and cracking, lightness and strength (Moya et al., 2014). Therefore, the species is often considered as the "king of woods" (Midgley et al., 2015). In Nepal, teak lumber is mainly utilized for high quality furniture manufacturing, indoor flooring, countertops, doors, windows and boatbuilding. In comparison to its other Nepali counterparts such as *Shorea robusta, Terminalia tomentosa* and *Dalbergia Sissoo* trees, teak grows faster and has a shorter rotation. Teak plantations are found scattered throughout the lowland region of Nepal, which has similar bioclimatic conditions to portions of neighboring India. The largest plantation of Teak in Nepal was established by the government of Nepal in the 1970s at Sagarnath Forestry Development Project (SFDP) (Thapa and Gautam, 2005).

As worldwide acreage of teak plantations has increased over past decades, many scientific studies have been published about different aspects of teak growth and yield from different parts of the world including China, Costa Rica, Ghana, India, Nepal (Fernández-Sólis et al., 2018; Koirala et al., 2017; Moya et al., 2020; Tewari et al., 2014; Víquez and Pérez, 2005; Yang et al., 2020). However, there are relatively fewer articles on taper modeling of teak trees (Adu-Bredu et al., 2008; Pérez and Kanninen, 2005; Warner et al., 2016). Generally, stem form of an excurrent tree resembles three solids of revolution: a neiloid for the bottom, a paraboloid for central bole, and a cone for treetops (Burkhart et al., 2018). The diameter of the bole decreases with an increase in tree height and eventually reaches zero at the tip of the tree. Teak trees generally have a large buttress, and they taper rapidly as height increases, therefore, traditional volume equations that are based only on total tree height and breast height diameter (dbh) might not address the volume loss incurred due to the rapid taper (Moya et al., 2020; Yang et al., 2020). This variation in diameter at any height can be modeled with the use of a taper equation. Basically, these equations estimate the diameter over bark of the tree stem at any given height up the stem. The

* Corresponding author.
E-mail addresses: anilk@uga.edu (A. Koirala), crmontes@uga.edu (C.R. Montes), bronsonbullock@uga.edu (B.P. Bullock), bishnu.wagle@maine.edu (B.H. Wagle).
1 Current Address (Bishnu H. Wagle):School of Forest Resources, University of Maine, 5755 Nutting Hall, Orono 04469, ME, USA.

https://doi.org/10.1016/j.tfp.2021.100103
Received 13 March 2021; Received in revised form 17 May 2021; Accepted 22 May 2021
Available online 25 May 2021

importance of tree taper models stems from their use to estimate total and merchantable volume and weight.

Most of the literature on taper modeling can be grouped into three types of models, single equation models, segmented models, and variable-exponent models. In single equation taper models, the taper of the stem is described by a single mathematical function, which can be either polynomial, trigonometric or a power function. One of the most popular taper models developed by Kozak et al. (1969) falls into this category. However, this single function sometimes fails to describe the entire stem profile, especially near the butt and upper portion of the stem (Jiang et al., 2005). The second group is a more complex segmented form of taper equations, first proposed by Max and Burkhart (1976). This type of taper equation models a tree stem as three segments: a neiloid, paraboloid and conoid that are joined to form the taper equation. In variable-exponent models, a single continuous function with an exponent changing from stump to top describes several intermediate forms like a neiloid and a paraboloid (Kozak, 2004, 1988; Newnham, 1992). In addition to these groups, other approaches such as dynamic modeling (García, 2015), mixed-effect modeling for either segmented or variable-exponent approaches (Garber and Maguire, 2003; Trincado and Burkhart, 2006; Yang et al., 2009), switching model (Valentine and Gregoire, 2001), and generalized additive models (GAM) including different splines (Robinson et al., 2011; Zapata-Cuartas et al., 2021) have been employed in developing taper equations.

Tree taper evaluation has traditionally been excluded from individual tree volume modeling process in Nepal (Gautam and Thapa, 2007; Koirala et al., 2017). Silwal et al. (2018) developed the first taper equation for *Shorea robusta* trees for the country and attempted to incorporate stem profile information into their volume models. Unfortunately, no publicly available taper equations for teak trees exists in Nepal. Therefore, the main objective of this study was to develop a taper equation for teak trees for the central lowland region of the country that will serve as a benchmark for future studies. This new taper equation is a step forward for growth and yield studies for the species in Nepal. Forest managers can easily follow the modeling approaches employed and apply the results to their teak plantations to improve forest management decision making.

2. Material and methods

2.1. Study area

The data utilized in this study came from the teak plantation stands managed by Sagarnath Forestry Development Project (SFDP) under the government of Nepal (Fig. 1). The project has engineered massive plantations of eucalyptus since its establishment on the preexisting forestland in 1978. Teak is the second most planted species after eucalyptus in this plantation development project. More than half of the nation's teak wood demand is fulfilled by the project. The total area of the project is 13, 512 ha, of which plantation area covers more than 8000 ha. Pure and mixed eucalyptus forests occupy about 90% of plantation area, while teak plantation occupy about 4% of the plantation area (i.e., about 500 ha.). Sagarnath project is situated in the lowland province (Province 2) of Nepal distributed among Rautahat, Sarlahi and Mahottari districts. The elevation of lowland Nepal ranges from 60 to 330 m above mean sea level. This region is characterized by hot summers (35° to 45 °C in April/May) with excess down pouring and dry winters (10° to 15 °C in January). The total annual precipitation ranges from 1130 mm to 2680 mm (FRA/DFRS, 2014). The native forest type in the region comprised of mixed hardwood tropical forest with the dominant species being *Shorea robusta*.

2.2. Destructive sampling and measurements

Twelve sub-compartments of teak plantations from the study sites spanning over a total of 100 ha from eastern sector (Sagarnath Division) were selected for conducting this research. A preliminary field inventory was carried out to examine the variation in breast height diameter (dbh) and total height of trees for the study sites. The average tree density and mean dbh for the sub-compartments were 344 trees ha $^{-1}$ and 35 cm, respectively. Forty-four trees without observable defects and abnormalities were selected for destructive sampling. The samples were selected to represent the variation within the forest and a range of sizes within each sub-compartment. There were six diameter classes in the sample with trees having diameters from 6 cm to 58 cm (Table 1). Six to nine trees in each diameter class were selected among the sampled trees which were felled by the project (Table 1). Candidate trees in the edge of a sub-compartment and those showing evident signs of disease were excluded from selection. Tree dbh was measured at 1.3 m in height, before felling. All branches were removed from the bole and its total length was measured. Diameter outside bark for each felled tree was measured at the ground level, breast height level, and at 3 m intervals up to the total height. Altogether, the data include 363 pairs of sectional measurements of diameter-height over all 44 trees destructively sampled.

2.3. Model fitting

For this study, six different commonly used and well-behaved taper models were used for fitting the dataset. They included: Kozak et al. (1969) single type model *M1*, Max and Burkhart (1976) segmented taper model *M2*, Kozak (2004) variable exponent model *M3* (referred to as 01 model in original article), Trincado and Burkhart (2006) non-linear mixed-effects segmented taper model *M4*, Sharma and Patron (2009) modified variable exponent model *M5*, and Garcia (2015) dynamic taper model *M6*. One thing to note is that model *M4* is the generalized version of *M2*, therefore, both have the same equation form. The difference is in the model fitting procedure; in *M2* all parameters were considered as fixed-effects while in *M4* three parameters were considered random-effects addressing the within tree variation. Out of six compared models, only *M4* was fitted with a mixed-effects modeling approach. Non-parametric and semi-parametric models such as GAMs and splines were excluded from model fitting process due to complexity in model comparisons with parametric models. Parameter estimations of all models except *M4* was carried out using non-linear least squares (*nls* function from *stats* package in R). For model M4, non-linear mixed effect approach was applied for parameter estimation (*nlme* package in R). Equation forms of these six models along with description of symbols and parameters are presented in Table 2.

Generally, taper data are obtained from repeated measurements along the tree bole. As a result of these repeated measurements, autocorrelation is typically present. The Durbin-Watson (DW) test, with null hypothesis of no autocorrelation in the data, was performed in order to examine the autocorrelation. The value of DW statistic was 1.025 and the p-value was less than 0.05 in the test. The null hypothesis of no autocorrelation in the data was rejected suggesting the presence of autocorrelation in our dataset. According to Li and Weiskittel (2010), this problem in taper modeling can be dealt with via two approaches. The first approach is to use mixed-effects modeling, and the other is to model the correlation structure directly during the modeling process. In this study, autocorrelations were modeled with first and second-order continuous autoregressive error structures, i.e. CAR(1) and CAR(2) (Diéguez-Aranda et al., 2005; Rojo et al., 2005). These error structures allow models to be easily used in irregular and unbalanced data (Gregoire et al., 1995) . The error term in the CAR(1) model is expanded as:

$$e_{ij} = d_1 \rho_1^{h_{ij} - h_{ij-1}} e_{ij-1} + \varepsilon_{ij} \tag{1}$$

where e_{ij} is the jth ordinary residual on the ith individual (i.e., the difference between the observed and the estimated diameters of tree i at height measurement j), $d_1 = 1$ for $j > 1$, and $d_1 = 0$ for $j = 1$, ρ_1 is the first-order autoregressive parameter that needs to be estimated, and $h_{ij} - h_{ij-1}$ is the distance separating the jth measurement from the jth

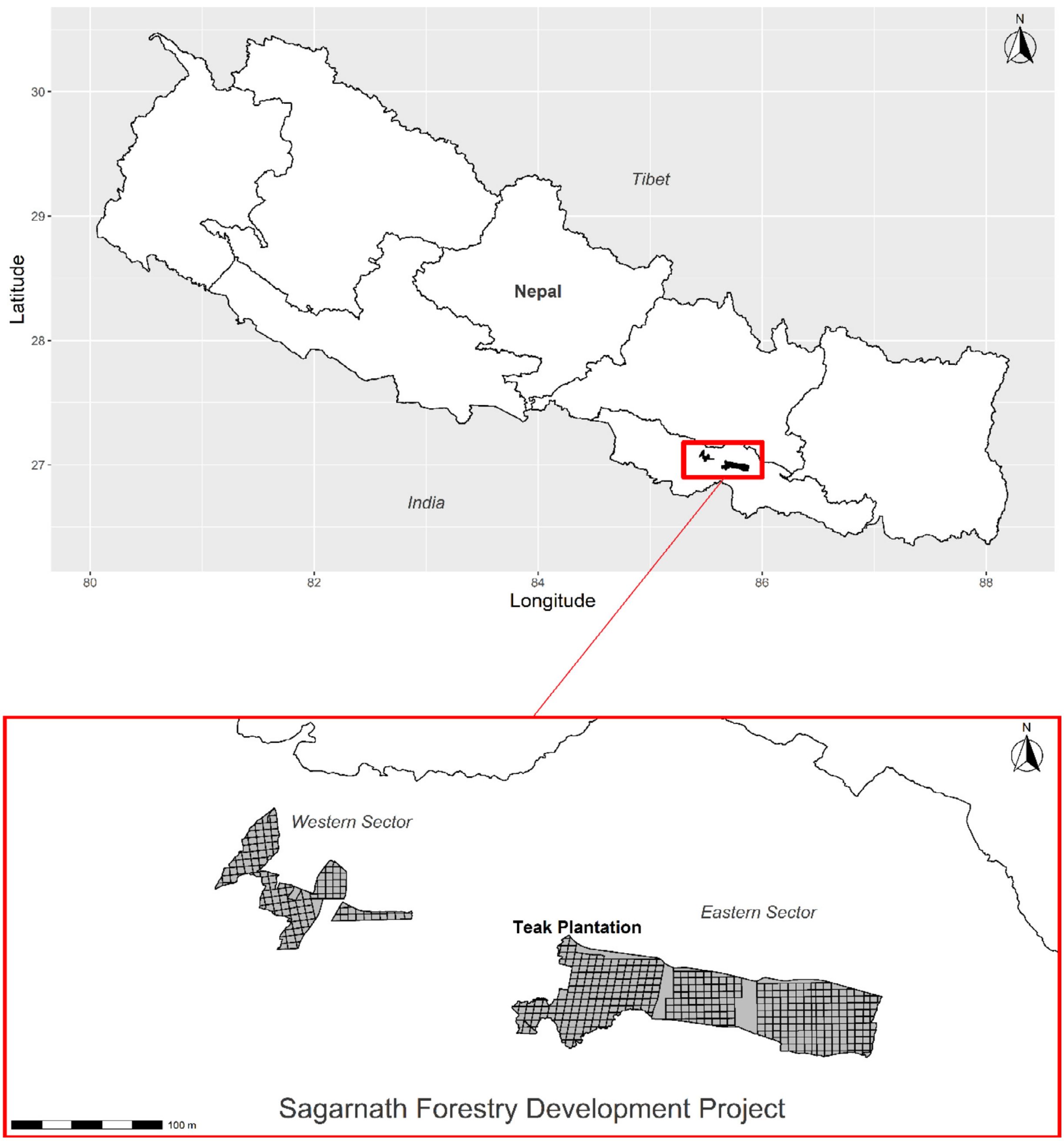

Fig. 1. Map of the study area showing the federal democratic republic of Nepal and Sagarnath Forestry Development Project. The project is in Province 2 of the country; the teak plantations is located on the eastern sector.

Table 1
Summary statistics for teak taper dataset for each diameter class. The diameter at breast height is given in centimeters while total tree height is presented in meter.

D-class	No. of Trees	Class Average Diameter (cm)	Class Minimum Diameter (cm)	Class Average Height (m)	Class Minimum Height (m)
0 –10	8	7.3	6.0	9.7	7.1
10- 20	9	14.7	10.2	13.5	9.8
20- 30	8	27.4	21.6	17.3	14.6
30- 40	7	36.9	31.5	23.5	20.2
40- 50	6	45.5	40.2	25.7	22.3
50–60	6	54.7	50.1	27.7	23.8

Table 2
Six taper models and their corresponding mathematical expressions.

Model	Equation forms
Kozak *M1*	$\frac{d_i^2}{D^2} = b_0 + b_1(\frac{h_i}{H}) + b_2(\frac{h_i^2}{H^2})$
Max and Burkhart *M2*	$\frac{d_i^2}{D^2} = b_1(\frac{h_i}{H} - 1) + b_2(\frac{h_i^2}{H^2} - 1) + b_3(a_1 - \frac{h_i}{H})^2 I_1 + b_4(a_2 - \frac{h_i}{H})^2 I_2$ where, $I_1 = 1\ if\ \frac{h_i}{H} \leq a_1, = 0$ otherwise; $I_2 = 1\ if\ \frac{h_i}{H} \leq a_2, = 0$ otherwise
Kozak Model 01 *M3*	$d_i = a_0 D^{a_1} X_i^{b_0 + b_1[1/e^{D/H}] + b_2 D^{X_i} + b_3 X_i^{D/H}}$ where, $X_i = \frac{(1-(\frac{h_i}{H})^{1/4})}{1-0.01^{1/4}}$
Trincado and Burkhart *M4*	$\frac{d_{ij}^2}{D_i^2} = b_{1i}(\frac{h_{ij}}{H_i} - 1) + b_{2i}(\frac{h_{ij}^2}{H_i^2} - 1) + b_{3i}(a_{1i} - \frac{h_{ij}}{H_i})^2 I_1 + b_{4i}(a_{2i} - \frac{h_{ij}}{H_i})^2 I_2$ I_1 and I_2 are same as model *M2* Within-tree variation: H_i = total tree height (m) for ith individual; d_i = diamter outside bark (cm) at stem height h_{ij} for the ith individual; h_{ij} = height (m) above ground for ith individual; and b_{1j}, b_{2i}, b_{3j}, and b_{4j} are paramters to be estimated
Sharma and Patron *M5*	$\frac{d_i}{D} = b_0(\frac{H-h_i}{H-1.3})(\frac{H}{1.3})^{b_1 + b_2(\frac{h_i}{H}) + b_3(\frac{h_i^2}{H^2})}$
Garcia *M6*	$d_i^2 = D^2(\frac{H-h_i-b_1+b_1 e^{[\frac{-(H-h_i)}{b_1}]} + b_2(H-h_i)e^{(-\frac{h_i}{b_3})}}{H-1.3-b_1+b_1 e^{[\frac{-(H-1.3)}{b_1}]} + b_2(H-1.3)e^{(-\frac{1.3}{b_3})}})$

d_i: diamter outside bark at any given height (cm); D: diamter outside bark at breast height (cm); h_i is height above ground (m); H: total tree height (m); b_0, b_1, b_2, b_3, b_4, a_0, a_1, and a_2: paramters to be estimated.

– *1* observation within each tree, $h_{ij} > h_{ijk}$. Likewise, the CAR(2) error structure is presented as:

$$e_{ij} = d_1 \rho_1^{h_{ij} - h_{ij-1}} e_{ij-1} + d_2 \rho_2^{h_{ij} - h_{ij-2}} e_{ij-2} + \varepsilon_{ij} \tag{2}$$

where $d_2 = 1$ for $j > 2$, and $d_2 = 0\ for\ j \leq 2$, ρ_2 is the second-order autoregressive parameter to be estimated, and $h_{ij} - h_{ij-2}$ is the distance separating the jth measurement from the jth – *2* observation within each tree, $h_{ij} > h_{ij-2}$. In these cases, ε_{ij} becomes independent and identically distributed. In order to evaluate the presence of autocorrelation and the effect of the CAR(1) and CAR(2) error structures used, scatterplots of residuals versus lagged residuals from the previous observations within each tree were examined visually.

2.4. Model evaluation

All of the six models fitted in this study were evaluated using leave-one-out cross-validation approach (Hastie et al., 2009). In this approach, a single observation is utilized as the validation set, and the remaining observations make up the training set. The statistical method follows fitting the models on the *n – 1* observation, which is almost as many as the number of observations in the full dataset. In a typical large dataset scenario, this approach could potentially be extremely time consuming. But for small dataset like this study, this approach is regarded better than other cross-validation schemes. The estimated models (*M1* to *M6*) were then applied to the left-out observation and the procedure was repeated for *n* times to calculate the goodness-of-fit statistics.

Models were compared based on four major goodness-of-fit statistics: root mean square error (RMSE), mean difference (MD) between observed and predicted diameter over bark, mean absolute difference (MAD), and adjusted R-squared (adj. R^2) Eqs. (3) to ((6) (Koirala et al., 2021). Mean difference and mean absolute difference are also referred as mean bias and mean absolute bias in some literatures e.g. Li and Weiskittel (2010). Lower values of RMSE, MD and MAD and higher value of adj. R^2 express better performance of selected equations. These statistics were calculated for full dataset at the beginning and later it was calculated for cross-validation. In order to better evaluate the models, a ranking technique was employed (Figueiredo-Filho et al., 1996). The best value in each statistic was assigned rank 1 while other values were ranked in ascending orders through rank 6 (the worst value). The model with the lowest total value of ranks was considered as the best model in terms of four goodness-of-fit statistics. All analyses for this study were performed in R 4.0.0 statistical software (R Core Team, 2020).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2}{n}} \tag{3}$$

$$MD = \frac{\sum_{i=1}^{n} (Y_i - \hat{Y}_i)}{n} \tag{4}$$

$$MAD = \frac{\sum_{i=1}^{n} \left| Y_i - \hat{Y}_i \right|}{n} \tag{5}$$

$$adj.\ R^2 = 1 - \left(\frac{(n-1)\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2}{(n-p)\sum_{i=1}^{n} (Y_i - \bar{Y}_i)^2} \right) \tag{6}$$

where, Y_i is the observed diameter over bark; $\hat{Y}_i$ is the predicted diameter; $\bar{Y}_i$ is the sample mean; n is the number of samples and p is the number of model parameters.

3. Results

The average diameter at breast height for all dataset was 31.41 cm with a standard deviation of 16.43 cm. The average total height for overall dataset was 18.91 m with standard deviation of 5.72 m. Fig. 2 illustrates a histogram of dbh distribution frequency of samples and the scatter plot of the relative diameter over relative height. The diameter distribution appears to be somewhat evenly distributed for all dbh classes except for smallest dbh class, which has two more trees than average. The plot of relative diameter versus relative height reflected the rate of decline in diameter with increasing height along tree bole.

At first, models were fit without considering the autocorrelation in the data. This resulted in autocorrelation, as seen in the panel (a) of Fig. 3, which is an example from the Max and Burkhart model, *M2*. However, this was an expected result for longitudinal data like ours. This trend in the residual disappeared in the CAR(1) model (Fig. 3, panel b). The CAR(2) model was able to deal with autocorrelation in the residuals as well, however, it did not perform well as the CAR(1) model (Fig. 3, panel c). After accounting for the CAR(1) error structure, the process of model fitting and cross-validation was carried out.

3.1. Model performance

Table 3 shows the goodness-of-fit statistics for both model fit and leave-one-out cross-validation. The range of root mean square error for all models was small, between 2.05 and 3.48 cm for overall dataset. Model *M1* showed the highest error of about 3.48 cm. Model *M3*, which is based on variable exponent approach, showed the lowest error of about 2.05 cm. The range of mean difference (MD) or bias was between −1.015 and 0.377 cm. The model with bias closer to zero was considered as the best model. Model *M3* showed MD close to zero. Model *M6*

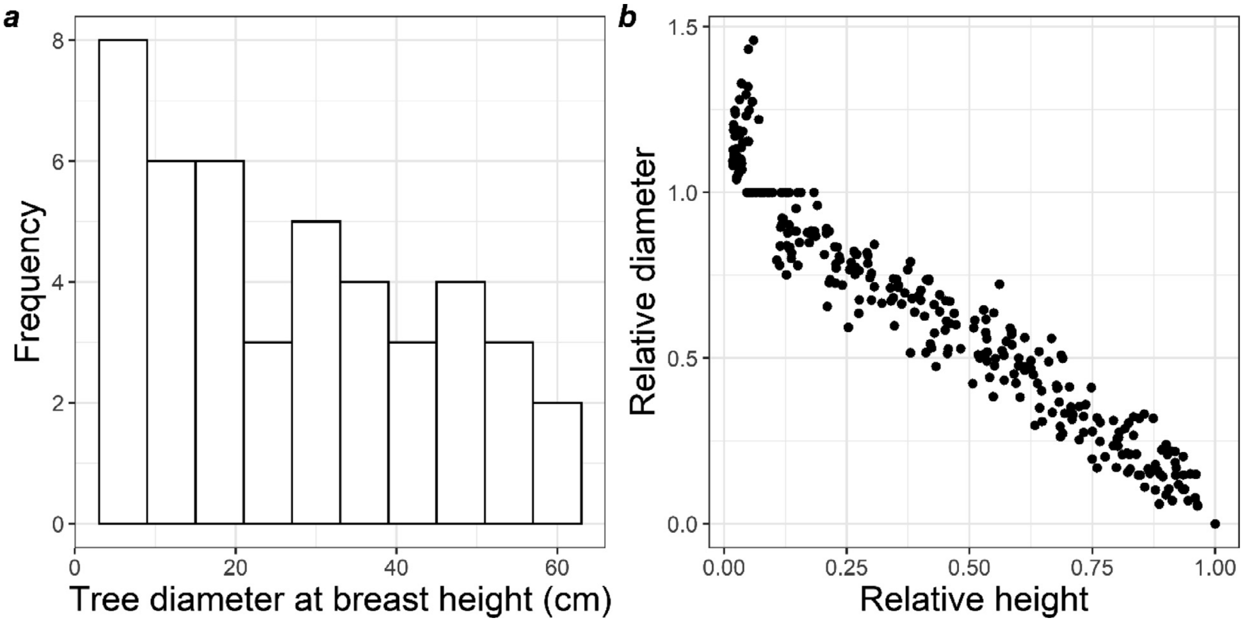

Fig. 2. Histogram showing the frequency of the sampled trees based on diameter at breast height (a) and scatter plot of the relative diameter (diameter over bark/diameter over bark at breast height) over relative height (stem observation height/total tree height) for all sampled trees (b).

Table 3
Fit and leave-one-out cross-validation statistics of models for prediction of stem diameter over bark (cm). The numbers in parenthesis indicate the ranking of each model for individual statistics.

Model	Fit				Leave-one-out cross validation				Total Score
	RMSE	MD	MAD	adj. R^2	RMSE	MD	MAD	adj. R^2	
M1	3.483 (6)	−1.015 (6)	2.455 (6)	0.949 (6)	3.507 (6)	−1.021 (6)	2.474 (6)	0.948 (6)	48
M2	2.549 (5)	−0.692 (5)	1.678 (5)	0.972 (5)	2.596 (5)	−0.704 (5)	1.709 (5)	0.971 (5)	40
M3	2.051 (1)	−0.116 (1)	1.503 (3)	0.980 (3)	2.095 (1)	−0.121 (1)	1.533 (3)	0.980 (3)	16
M4	2.376 (4)	−0.517 (4)	1.563 (4)	0.976 (4)	2.408 (4)	−0.584 (4)	1.611 (4)	0.973 (4)	32
M5	2.279 (3)	0.123 (2)	1.502 (2)	0.981 (2)	2.320 (3)	0.122 (2)	1.526 (2)	0.981 (2)	18
M6	2.138 (2)	0.377 (3)	1.442 (1)	0.982 (1)	2.173 (2)	0.373 (3)	1.462 (1)	0.982 (1)	14

returned the lowest mean absolute difference of 1.442 cm. In terms of adjusted R-squared, the lowest value of 94.9% was from model *M1* while the highest value of 98.2% was from the dynamic model *M6*. Similar values of RMSE, MD, MAD and FI were obtained for the leave-one-out cross-validation for all models. Based on the overall ranking, dynamic model *M6* got the lowest total score followed by model *M3* and model *M2* in terms of performance in both fit and leave-one-out cross-validation statistics. The differences in RMSE and bias between *M3* and other two models (*M6* and *M5*) was less than 1%. The adjusted R-squared for all three models were also higher than other models. Thus, these three models were selected as the best candidate models in the first round of model evaluation.

3.2. Final model selection

Further comparison of these three models were carried out by graphical analysis of residual plots and observed vs. fitted plots for all 44 trees (Fig. 4). The residual plots of all three models showed unbiasedness and relatively constant variance over the range of the data. Likewise, the observed vs. fitted plots of all three models depicted similar patterns and showed minimal deviance from the prediction line. These models were further examined in terms of their abilities to predict diameter over bark for three different sized (smallest, medium-sized and tallest) trees in the data.

Fig. 5 shows the predicted taper functions for three teak trees (one short, one middle-sized and one tall from the dataset). The selected short tree had total height of 7.10 m (shortest tree of the dataset), the middle-sized tree had total height of 19.6 m (middle-height tree from the dataset), and the taller tree had height of 28.03 m (tallest tree of the dataset). The dynamic model *M6* performed better in predicting upper stem diameter for the selected three trees than model *M3 and M5*. Overall, the dynamic model proposed by García (2015) was chosen to be the best model for predicting diameter over bark for teak trees of central lowland Nepal. The final model with parameters, standard errors (SE) and p-values is presented in Table 4.

4. Discussion

Taper equations are tools to predict the diameter of a tree stem, over or under bark, at any height from the ground. These equations are considered important forest management tools since they allow the estimation of stem volumes for any merchantable height and diameter and hence supersede the volume tables (McTague and Weiskittel, 2020). There are numerous forms of taper models available in the forestry literatures, which sometimes creates difficulties in appropriate model selection for a tree species from a newly studied area. Based on our extensive literature search, we believe that only a handful of taper modeling studies have been carried out for teak trees in south and southeast Asia (such as from Goodwin, 2007; Warner et al., 2016; Seppänen and Mäkinen, 2020). This is an interesting point, given that teak species are native to these regions. Even though teak is not technically native to Nepal, the growth conditions for teak in lowland Nepal resembles that of neighboring India, which has one of the world's largest natural teak forest areas. To date, the authors know of no taper equation addressing stem form

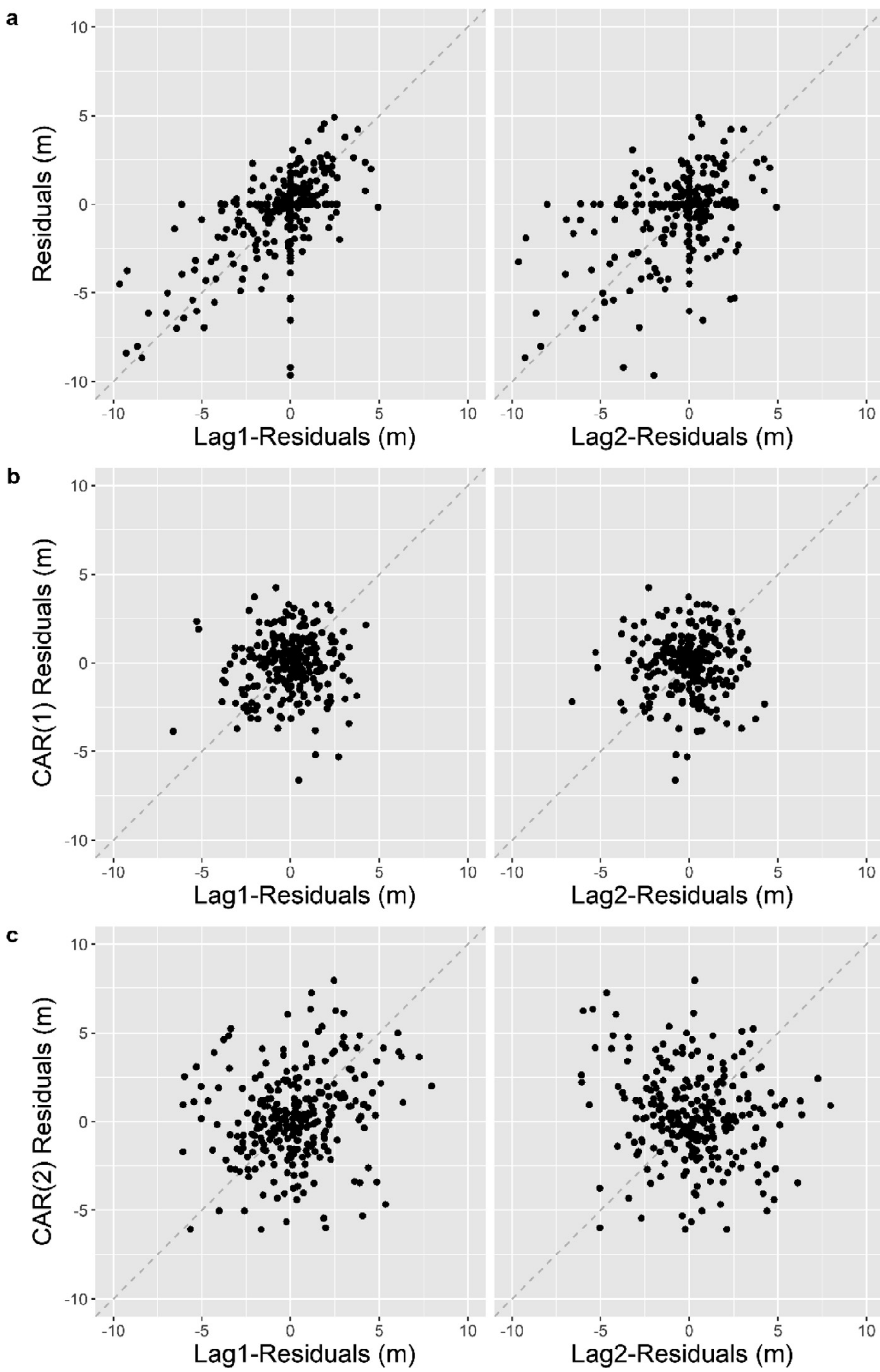

Fig 3. Residuals as a function of: a Lag1-residuals (left column), and Lag2-residuals (right column) for both fitting methods: without error structure (a), and assuming CAR(1) (b) and CAR(2) (c) error structures. Dashed line is a reference line.

Table 4
Final selected model with equation form, paramter estimates, standard error (SE) and p-value.

Equation	Paramter	Estimate	SE	Pr (> | t |)
$d_i^2 = D^2\left(\frac{H - h_i - b_1 + b_1 e^{\left[\frac{-(H-h_i)}{b_1}\right]} + b_2(H - h_i)e^{\left(-\frac{h_i}{b_3}\right)}}{H - 1.3 - b_1 + b_1 e^{\left[\frac{-(H-1.3)}{b_1}\right]} + b_2(H - 1.3)e^{\left(-\frac{1.3}{b_3}\right)}}\right)$ where, d_i: diamter outside bark at any given height (cm); D: diameter outside bark at breast height (cm); h_i is height above ground (m); and H: total tree height (m)	b_1	6.32107	1.34237	< 0.0001
	b_2	0.69215	0.09038	< 0.0001
	b_3	1.49947	0.11680	< 0.0001

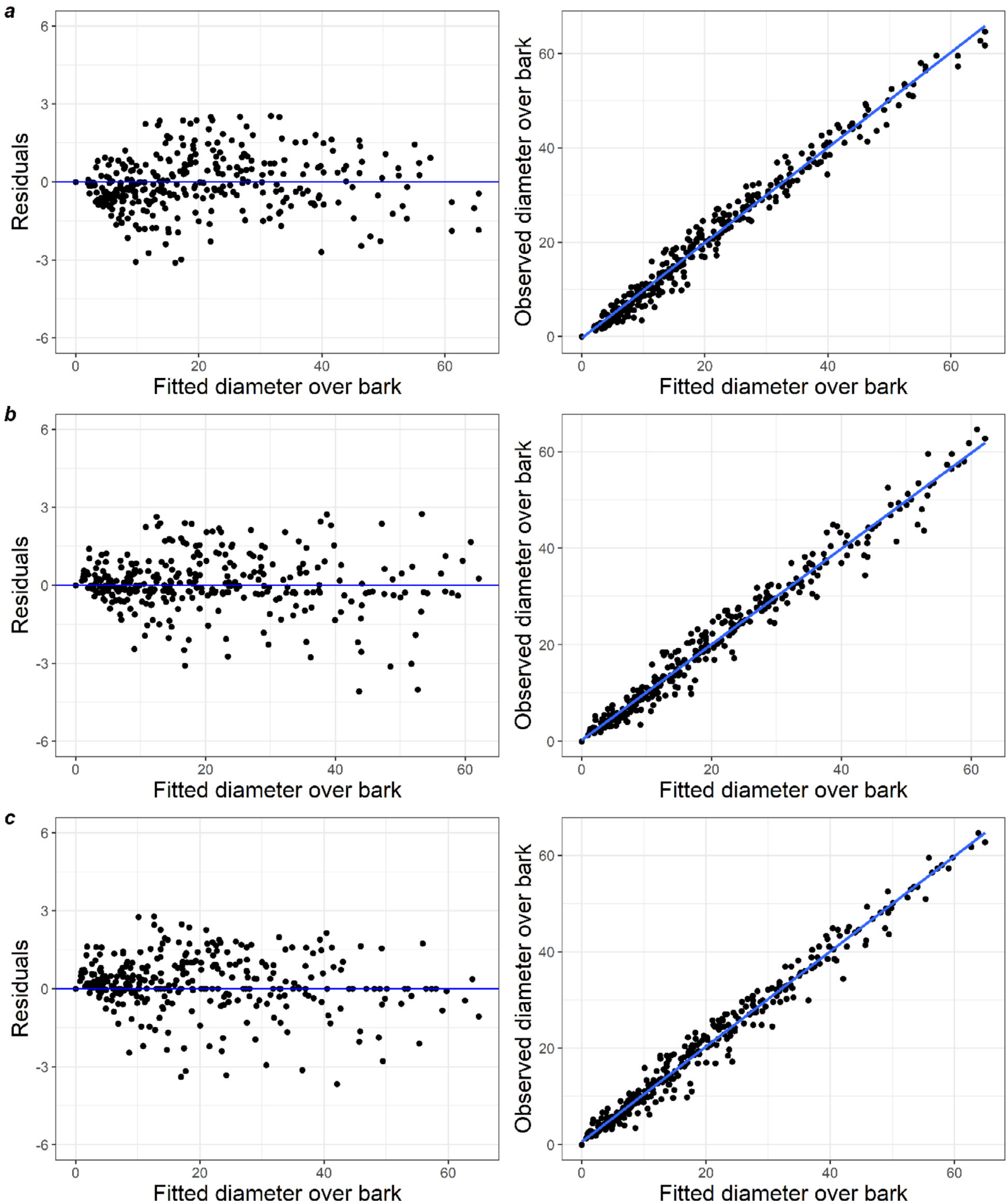

Fig 4. Residual plots and observed vs. fitted plots for models *M3* (a), *M5* (b) and *M6* (c) using data from all 44 trees. The unit for both residuals and fitted diameter over bark is in cm.

of teak trees in Nepal. Therefore, we relied on the popular taper equations available for other species in the literature. Five commonly used taper model forms (models *M1* to *M5*) and one distinct but promising taper equation (*M6*) were selected as candidate taper models for this study (Table 2). A leave-one-out cross-validation approach was utilized for model evaluation, as it is considered reliable when the sample size is small. It has been shown that this approach provides an approximately unbiased estimate of the test error (Allan, 1974; Cawley, 2006).

The Durbin-Watson test revealed autocorrelation in our dataset, therefore, CAR(1) and CAR(2) corrections were applied to the residuals. The trend in residuals disappeared with the use of these error structures. These results are comparable to other taper literature (Poudel et al., 2020; Rojo et al., 2005). The major purpose of using a correction for autocorrelation was to improve model properties such as unbiasedness and efficient parameter estimates error. It also prevents underestimation of the covariance matrix of the parameters, thereby

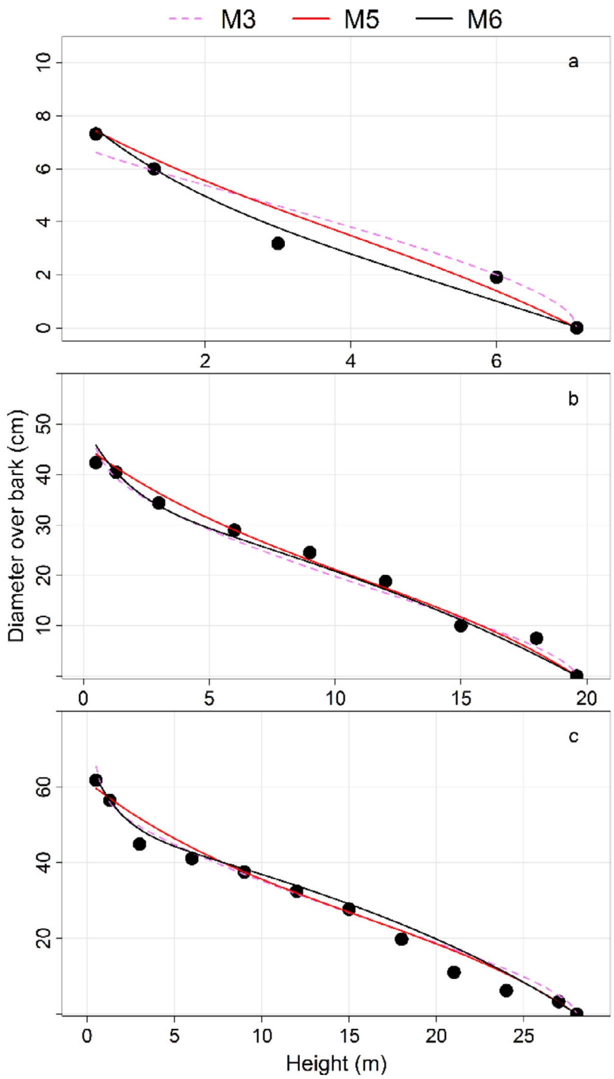

Fig 5. Prediction curves of three taper functions (*M3, M5* and *M6*) over actual measurement points for the shortest tree (a), medium-height tree (b) and the tallest tree (c) of the sample.

making it possible to carry out the usual statistical tests (West et al., 1984).

Overall, the dynamic model *M6* and the modified variable exponent models *M3* and *M5* had the lowest error and bias values. The adjusted R-squared for all models were greater than 94%, which implies the high predictive capabilities of all models investigated. However, the dynamic model outperformed other models with an adjusted R-squared value of 98.2%. Kozak et al. (1969) single type model *M1* performed poorly in all statistics. This model was one of the earliest models developed for modeling taper, which considered a single quadratic form for stem bole shape. The segmented model from Max and Burkhart (1976) *M2* also performed poorly and was second to the last in overall ranking. Nevertheless, this does not suggest that the Max and Burkhart (1976) model was not suitable for taper modeling. Model *M2* have been applied for many taper studies and regarded as reliable model (Cao and Wang, 2011; Cao, 2009; Sharma and Burkhart, 2003). The segmented model approach was originally developed for trees with excurrent forms like conifers, therefore, it does not come as a surprise that it is outperformed by the other models for teak trees, as they have very distinct stem forms when compared to conifers. Our results of Kozak (2004) model *M3* performing well did not coincide with the study carried out for primary conifer species in the Acadian region of the US (Li and Weiskittel, 2010), in which the Kozak model performed poorly. The authors are not aware of the dynamic taper equation being used to model taper of other species in the literature beyond that by Garcia in 2015. Models *M3, M5* and *M6* were selected as the best models in the initial evaluation.

The graphical analysis of residuals and predicted values showed that the performance all three models were somewhat similar (Fig. 5). For the last phase of model evaluation, prediction of diameters over bark of three trees (shortest, middle-sized, and tallest trees) from the sample was made using the three best models. The prediction curve from Garcia (2015) dynamic model *M6* was superior to the Kozak (2004) variable-exponent model *M3* and Sharma and Patron (2009) model *M5* for the smallest and tallest trees in the sample. In the dynamic equation, the butt-swell increment can be modeled by a decay function, which might have proven effective for teak that tends to have big buttress swell (Warner et al., 2018). In addition, the dynamic taper equation was derived on the basis of trees biological and physiological properties. The parsimony in the model and realistic representation of teak stem shape are additional advantages of the recommended model. New volume equations can be developed based on our taper model, which will help local foresters with assessing their standing volume before merchandising trees.

5. Conclusion

In order to construct taper models for teak tree species of central lowland Nepal, six different taper models were tested. Samples were collected from 44 felled teak trees in the Sagarnath Forestry Development Project, Nepal. Our analysis showed that the Garcia (2015) dynamic taper equation, *M6*, proved to be the best model for our dataset. The model resulted in lower root mean square error and mean bias for predicting diameter over bark of the stem. It also had the best adjusted R-squared value of all analyzed models. This stem taper equation can be used by researchers and foresters to calculate upper stem diameter as well as stem volume to any desirable merchantable limit. It is recommended that this model be applied to similar stand and site conditions. As this is the first model for teak stem taper in the country, it is advisable to test the model for other sites before application. Care must be taken when using this model on other sites and predicting beyond the observed range of tree size.

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

CRediT authorship contribution statement

Anil Koirala: Funding acquisition, Conceptualization, Methodology, Formal analysis, Validation, Writing – original draft. **Cristian R. Montes:** Funding acquisition, Writing – original draft. **Bronson P. Bullock:** Visualization, Writing – original draft. **Bishnu H. Wagle:** Visualization, Writing – original draft.

Acknowledgments

The authors would like to thank funding agencies for supporting this work. We would like to express our gratitude to Mr. Binod Kumar Singh for providing space and good working environment for field data collection. Finally, we would like to thank anonymous reviewers for their constructive comments and suggestions.

Funding

This project was supported by funding from Sagarnath Forestry Development Project – Students research grant acquired by the primary author. This project was also supported by funding and technical assistance from Plantation Management Research Cooperative of the University of Georgia.

References

Adu-Bredu, S., Bi, A.F.T., Bouillet, J.P., Mé, M.K., Kyei, S.Y., Saint-André, L., 2008. An explicit stem profile model for forked and un-forked teak (*Tectona grandis*) trees in West Africa. For. Ecol. Manag. 255, 2189–2203. doi:10.1016/j.foreco.2007.12.052.

Allan, D., 1974. The relationship between variable selection and prediction. Technometrics 16, 125–127.

Cao, Q.V., Wang, J., 2011. Calibrating fixed- and mixed-effects taper equations. For. Ecol. Manag. 262, 671–673. doi:10.1016/j.foreco.2011.04.039.

Burkhart, H.E., Avery, T.E., Bullock, B.P., 2018. Forest Measurements, 6th ed. Waveland Press, Inc.

Cao, Q.V., 2009. Calibrating a segmented taper equation with two diameter measurements. South. J. Appl. For. 33, 58–61. doi:10.1093/sjaf/33.2.58.

Cawley, G.C., 2006. Leave-one-out cross-validation based model selection criteria for weighted LS-SVMs. In: Proceedings of the 2006 IEEE International Joint Conference on Neural Network Proceedings, pp. 1661–1668. doi:10.1109/IJCNN.2006.246634.

Diéguez-Aranda, U., Burkhart, H.E., Rodríguez-Soalleiro, R., 2005. Modeling dominant height growth of radiata pine (*Pinus radiata* D. Don) plantations in north-western Spain. For. Ecol. Manag. 215, 271–284. doi:10.1016/j.foreco.2005.05.015.

Fernández-Sólis, D., Berrocal, A., Moya, R., 2018. Heartwood formation and prediction of heartwood parameters in *Tectona grandis* L.f. trees growing in forest plantations in Costa Rica. Bois Forets des Trop 335, 25–37. doi:10.19182/bft2018.335.a31499.

Figueiredo-Filho, A., Borders, B.E., Hitch, K.L., 1996. Taper equations for *Pinus taeda* plantations in Southern Brazil. For. Ecol. Manag. 83, 39–46. doi:10.1016/0378-1127(96)03706-1.

FRA/DFRS, 2014. Terai Forests of Nepal (2010 - 2012). Forest Resource Assessment (FRA) Nepal, Department of Forest Research and Survey.

Garber, S.M., Maguire, D.A., 2003. Modeling stem taper of three central Oregon species using nonlinear mixed effects models and autoregressive error structures. For. Ecol. Manag. 179, 507–522. doi:10.1016/S0378-1127(02)00528-5.

García, O., 2015. Dynamic modeling of tree form. Math. Comput. For. Nat. Sci. 7, 9–15.

Gautam, S., Thapa, H., 2007. Volume equation for *Populus deltoides* plantation in western Terai of Nepal. Banko Janakari 17, 70–73. doi:10.3126/banko.v17i2.2158.

Goodwin, A.N., 2007. A taper model for plantation-grown Teak. Consult. Rep. Fac. For. Kasetsart Univ. Bangkok, Thail.

Gregoire, T.G., Schabenberger, O., Barrett, J., 1995. Linear modeling of irregularly spaced, unbalanced, longitudinal data from permanent-plot measurements. Can. J. For. Res. 25, 137–156.

Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd ed. Springer New York.

Jiang, L., Brooks, J.R., Wang, J., 2005. Compatible taper and volume equations for yellow-poplar in West Virginia. For. Ecol. Manag. 213, 399–409. doi:10.1016/j.foreco.2005.04.006.

Kenzo, T., Himmapan, W., Yoneda, R., Tedsorn, N., Vacharangkura, T., Hitsuma, G., Noda, I., 2020. General estimation models for above- and below-ground biomass of teak (*Tectona grandis*) plantations in Thailand. For. Ecol. Manag. 457, 117701. doi:10.1016/j.foreco.2019.117701.

Koirala, A., Kizha, A.R., Baral, S., 2017. Modeling height-diameter relationship and volume of teak (*Tectona grandis* L. F.) in central lowlands of Nepal. J. Trop. For. Environ. 7. doi:10.31357/jtfe.v7i1.3020.

Kollert, W., Walotek, P.M., 2015. Global teak trade in the aftermath of Myanmar's log export ban (No. FP/49/E), Planted Forests and Trees Working Paper Series.

Kozak, A., 2004. My last words on taper equations. For. Chron. 80, 507–515. doi:10.5558/tfc80507-4.

Koirala, A., Montes, C. R., Bullock, B. P., 2021. Modeling dominant height using stand and water balance variables for loblolly pine in the Western Gulf, US. Forest Ecology and Management 479, 118610. doi:10.1016/j.foreco.2020.118610.

Kozak, A., 1988. A variable-exponent taper equation. Can. J. For. Res. 18, 1363–1368. doi:10.1139/x88-213.

Kozak, A., Munro, D.D., Smith, J.H.G., 1969. Taper functions and their application in forest inventory. For. Chron. 45, 278–283. doi:10.5558/tfc45278-4.

Li, R., Weiskittel, A.R., 2010. Comparison of model forms for estimating stem taper and volume in the primary conifer species of the North American Acadian Region. Ann. For. Sci. 67. doi:10.1051/forest/2009109.

Max, T.A., Burkhart, H.E., 1976. Segmented polynomial regression applied to taper equations. For. Sci. 22, 283–289.

McTague, J.P., Weiskittel, A., 2020. Evolution, history, and use of stem taper equations: a review of their development, application, and implementation. Can. J. For. Res. 1–26. doi:10.1139/cjfr-2020-0326.

Midgley, S., Somaiya, R.T., Stevens, P.R., Brown, A., Kien, N.D., Laity, R., 2015. Planted teak: global production and markets, with reference to Solomon Islands, ACIAR Technical Reports No. 85. Canberra.

Moya, R., Gaitán-Álvarez, J., Ortiz-Malavassi, E., Berrocal, A., Fernández, D., 2020. Equations for predicting heartwood merchantable volume and tradable sawlog in *Tectona grandis*. J. Trop. For. Sci. 32, 379–390. doi:10.2307/26940728.

Moya, Róger, Bond, Brian, Quesada, H., Moya, R., Bond, B., Quesada, Á.H., 2014. A review of heartwood properties of *Tectona grandis* trees from fast-growth plantations. Wood Sci. Technol. 48, 411–433. doi:10.1007/s00226-014-0618-3.

Newnham, R.M., 1992. Variable-form taper functions for four Alberta tree species. Can. J. For. Res. 22, 210–223. doi:10.1139/x92-028.

Pérez, D., Kanninen, M., 2005. Effect of thinning on stem form and wood characteristics of teak (*Tectona grandis*) in a humid tropical site in Costa Rica. Silva Fenn. 39, 217–225. doi:10.14214/sf.385.

Poudel, K.P., Ozçelik, R., Yavuz, H., 2020. Differences in stem taper of black alder (*Alnus glutinosa* subsp. barbata) by origin. Can. J. For. Res. 50, 581–588. doi:10.1139/cjfr-2019-0314.

R Core Team, 2020. R: a language and environment for statistical computing.

Robinson, A.P., Lane, S.E., Thérien, G., 2011. Fitting forestry models using generalized additive models: a taper model example. Can. J. For. Res. 41, 1909–1916. doi:10.1139/x11-095.

Rojo, A., Perales, X., Sánchez-Rodríguez, F., Álvarez-González, J.G., von Gadow, K., 2005. Stem taper functions for maritime pine (*Pinus pinaster* Ait.) in Galicia (Northwestern Spain). Eur. J. For. Res. 124, 177–186. doi:10.1007/s10342-005-0066-6.

Seppänen, P., Mäkinen, A., 2020. Comprehensive yield model for plantation teak in Panama. Silva Fenn. 54, 1–25. doi:10.14214/sf.10309.

Sharma, M., Burkhart, H.E., 2003. Selecting a level of conditioning for the segmented polynomial taper equation. For. Sci. 49, 324–330. doi:10.1093/forestscience/49.2.324.

Sharma, M., Parton, J., 2009. Modeling stand density effects on taper for jack pine and black spruce plantations using dimensional analysis. For. Sci. 55, 268–282. doi:10.1093/forestscience/55.3.268.

Silwal, R., Baral, S.K., Chhetri, B.B.K., 2018. Modeling taper and volume of Sal (*Shorea robusta* Gaertn. f.) trees in the western Terai region of Nepal. Banko Janakari 76–83.

Tewari, V.P., Álvarez-González, J.G., García, O., 2014. Developing a dynamic growth model for teak plantations in India. For. Ecosyst. 1, 1–10. doi:10.1186/2197-5620-1-9.

Thapa, H.B., Gautam, S.K., 2005. Growth performance of *Tectona grandis* in the western Terai of Nepal. Banko Janakari 15, 6–12. doi:10.3126/banko.v15i2.344.

Trincado, G., Burkhart, H.E., 2006. A generalized approach for modeling and localizing stem profile curves. For. Sci. 52, 670–682. doi:10.1093/forestscience/52.6.670.

Valentine, H.T., Gregoire, T.G., 2001. A switching model of bole taper. Can. J. For. Res. 31, 1400–1409. doi:10.1139/x01-061.

Víquez, E., Pérez, D., 2005. Effect of pruning on tree growth, yield, and wood properties of *Tectona grandis* plantations in Costa Rica. Silva Fenn. 39, 381–390.

Warner, A.J., Jamroenprucksa, M., Puangchit, L., 2018. Buttressing impact on diameter estimation in plantation teak (*Tectona grandis* L .f .) sample trees in northern Thailand. Agric. Nat. Resour. 51, 520–525. doi:10.1016/j.anres.2018.01.001.

Warner, A.J., Jamroenprucksa, M., Puangchit, L., 2016. Development and evaluation of teak (*Tectona grandis* L.f.) taper equations in northern Thailand. Agric. Nat. Resour. 50, 362–367. doi:10.1016/j.anres.2016.04.005.

West, P.W., Ratkowsky, D.A., Davis, A.W., 1984. Problems of hypothesis testing of regressions with multiple measurements from individual sampling units. For. Ecol. Manag. 7, 207–224.

Yang, B., Jia, H., Zhao, Z., Pang, S., Cai, D., 2020. Horizontal and vertical distributions of heartwood for teak plantation. Forests 11, 225. doi:10.3390/f11020225.

Yang, Y., Huang, S., Trincado, G., Meng, S.X., 2009. Nonlinear mixed-effects modeling of variable-exponent taper equations for lodgepole pine in Alberta, Canada. Eur. J. For. Res. 128, 415–429. doi:10.1007/s10342-009-0286-2.

Zapata-Cuartas, M., Bullock, B.P., Montes, C.R., 2021. A taper equation for loblolly pine using penalized spline regression. For. Sci. 67, 1–13. doi:10.1093/forsci/fxaa037.