Nepal Journal of Science and Technology Vol. 14, No. 1 (2013) 147-152

Predicting Total Height from Diameter Using Nonlinear Models in *Pinus roxburghii*

T. B. Thapa[1], C. P. Upadhyaya[2], Y.P. Timilsina[2], S. K. Bhandari[2] and L. Puri[2]
[1]*District Forest Office, Kaski*
[2]*Institute of Forestry, Tribhuvan University, Pokhara*
e-mail: shesu15@yahoo.com

Abstract

Height of the trees is not only used to estimate the product available from trees and to determine the productive capacity of the particular land but also used in forest management and research activities. In this research, predictability of total height from diameter at breast height for chir-pine (*Pinus roxburghii*) by using nonlinear models was examined. Total height and diameter at breast height of 180 trees were measured by using Suunto Clinometer and diameter tape respectively. The measured data were subjected to previously developed robust and strong statistical models, and evaluated by using different test statistics. Among those, $H = e^{(-0.276)} D^{0.925}$ explained the greatest proportion of variations of height ($R^2_{adj.}$=0.910) and minimum root mean square error (RMSE=1.221) and showed desirable behavior of flexibility and robustness.

Key words: Chir-pine, diameter at breast height, statistical model, model evaluation, trees

Introduction

The relationship between height and diameter of plant is of wide concern to agronomists, foresters and agro-foresters interested in practical and economic problems, to ecologists and ecophysiologists interested in climate and environment, and to evolutionary theorists and ecosystem modelers to construct mathematical representations for general and specific applications (Thornley 1999). The tree height is commonly used to quantify the product from tree and to identify the productive capacity of the particular site on which the trees are growing (Lama *et al.* 2012). Tree height is also required for both forest management and research activities. Diameter at breast height (dbh) and total height are the commonly measured variables in an inventory. Unlike dbh, total height is less frequently used for construction or application of forest models because measurement of dbh is more cost effective, easy and accurate than total height (Sharma 2009). A prediction of total height from the given diameter might be the reliable option where such models are available. Many studies have presented models for the prediction of the height-diameter relationship of a stand. Most of these models use a representative sample of trees from the target stand (Curtis 1967, Arabatzis & Burkhart 1992, Huang *et al.* 1992, Lynch & Murphy 1995, Fang & Bailey 1998, Sharma & Portan 2007, Trincado *et al.* 2007, Newton & Amponsah 2007, Wagle 2007).

As only a few forest growth models have been developed in Nepal, the uncertainty of growth and yield estimates is often high. In order to safeguard against depletion of resources, community forests apply conservative estimates of productivity and allowable cut. A possible consequence of this is that forests are underutilized and provide less income to communities than could have been obtained with reliable information about annual increment. Hence, the potential value of preparing growth models to communities is likely to be high (Sapkota & Meilby 2009).

Chir pine (*P. roxburghii*) forest is located in a sub-tropical region with an altitude varying from 1000 m to 2000 m, and its standing volume is 6.3% of the total forest in the country (DFRS 1999). The economic contribution of Chir pine forest to national and local level development is valuable; and, therefore, its management is useful. For scientific management, species-specific individual tree or stand level models such as height-diameter models, site index models, growth models, and biomass and volume models need to be developed. Height-diameter models can be used as a sub-model in the more comprehensive models such as biomass models, growth and yield models or their simulation systems. Modeling works for Chir pine forests in the country include Joshi (1984), Joshi (1985), Rauntiainen (1992), Sharma & Pukkala (1990), Sharma 2009. But, only few of these are height diameter models. This study, therefore, aims at constructing height diameter models using data from middle hills of Nepal.

Methodology

Data: The data were collected from Galeshwor Thulo Salleri community forest and Banchare Dil community forest of Ghatan Village Development Committee of Myagdi district (Fig 1) (latitude 28^0 20' – 28^0 47' and longitude 83^0 08' – 83^0 53') of west Nepal. Size variation of the individual tree was identified from the previous record of the forest (Operational plan).

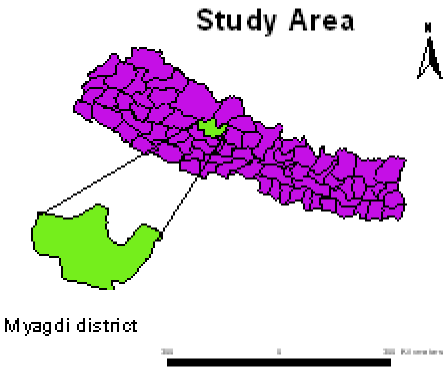


Fig. 1. Study area

Table 1. Descriptive statistics of tree variables

Tree variables	Tree Statistics	Diameter classes					
		0-10	10-20	20-30	30-40	40-50	>50
Height (m)	Number	30	30	30	30	30	30
	Mean	5.74	8.79	13.50	22.36	25.58	30.76
	Std. Deviation	1.72	2.65	2.21	4.63	3.45	2.500
	Minimum	2.50	4.30	7.50	14.00	16.80	24.00
	Maximum	8.50	15.50	19.00	30.00	32.00	35.00
Diameter (cm)	Number	30	30	30	30	30	30
	Mean	8.28	13.97	23.99	34.88	45.31	54.99
	Std. Deviation	1.45	2.88	2.52	2.79	2.674	4.03
	Minimum	5.57	10.51	20.06	30.40	40.19	50.00
	Maximum	9.97	19.75	29.94	39.81	49.68	61.94

Models: Nonlinear relationship between height and diameter was tested with a scattered plot diagram of height against dbh. Five different nonlinear models (Table 2) were used to fit height-diameter relationship. All these models possess few parameters, mathematically strong and therefore have commonly been used for modeling various tree and stand characteristics.

Table 2. Different models used to fit the diameter height relationship

Designation	Models
M1	$H = e^{\beta_0} D^{\beta_1} + \varepsilon_i$
M2	$H = \beta_0 + \beta_1 D + \beta_2 D^2$
M3	$H = e^{(\beta_0 + \beta_2 D)} D^{\beta_1} + \varepsilon_i$
M4	$H = e^{(\beta_0 + \beta_2 D2)} D^{\beta_1} + \varepsilon_i$
M5	$H = e^{(\beta_0 + \beta_2 D3)} D^{\beta_1} + \varepsilon_i$

H = total height (m); D = dbh (cm); $\hat{a}_0$, $\hat{a}_1$, $\hat{a}_2$ = parameters; and ln=natural logarithm; $\mathring{a}_{i\,s}$ are random and normally distributed errors.

Parameter estimation and model evaluation: The commonly used following two modeling approaches were utilized in this study also. First; fitting the candidate models; second; evaluation of the fitted models. In the first step, candidate models M1–M5 were fitted by regression analysis. The values of regression coefficients were estimated by the method of least square regression. The second step, i.e. the evaluation of the fitted models, was carried out using following criteria.

1. Adjusted coefficient of determination (R^2_{adj}): It shows a proportion of total variance explained by the model with the adjustment of the number of parameters, *p* and the number of non-missing observations, *n*. It is estimated as:

$$R^2_{adj} = 1 - (1 - R^2)\frac{(n-1)}{(n-p)}$$

2. Significance of the parameter values: Parameter estimates should be significantly different from zero ($p<0.05$).
3. Homogeneity of the residuals: Plotting of the residuals from the model over predicted values or independent variables should show a random, constant variance pattern around a residual value of zero (Clutter *et al.* 1983).
4. Distribution of residuals: Histograms of residuals were plotted to display the distribution (normal or abnormal) patterns of the residuals.
5. Root mean squared error (RMSE): RMSE determines the accuracy of model predictions and it is considered one of the most important model evaluation criteria. RMSE was calculated using following formula:

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

Where Y_i and $\hat{Y}_i$ are the observed and predicted values respectively; n is the total number of observations used to fit the model; and P is the number of parameters.

6. Visual examination of the fitted curves overlaid on the scattered plots of the observed data. It is the most important part in modeling.

Results and Discussion

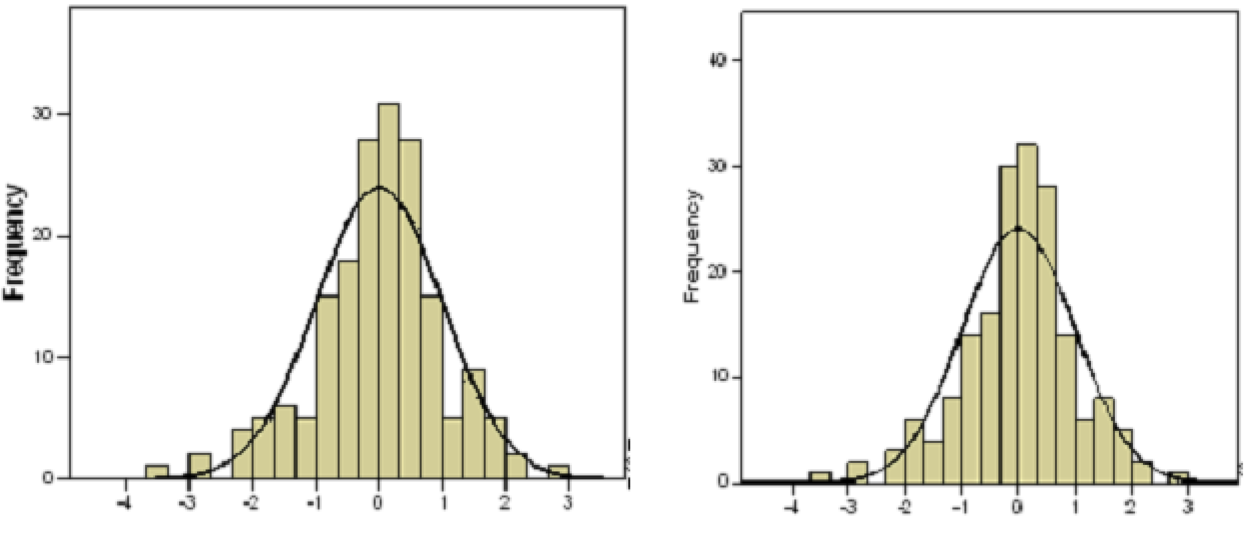
Parameters of all the models were significant ($P<0.05$) except the model M2 and M3. Except the model M3, M4 and M5, other models described more than 90% ($R^2_{adj.}$ >0.90) of height variability (Table 3). The models M1, M3, M4 and M5 showed the RMSE less than 2. However, M2 and M3 were excluded from the further analysis because one of its parameter estimates was non-significant ($P>0.05$). The model M4 was also excluded from further analysis because it demonstrated poor fit statistics. Out of two remaining models M1 demonstrated higher $R^2_{adj.}$ and lower RMSE than M5. However, the difference between these two models in fit statistics is very small. It is not a good decision to select a best model in such a case by considering only the value of $R^2_{adj.}$ and RMSE. Therefore, graphical analysis of residuals was also carried out. Since a residual may be viewed as the deviation between the data and fit, it is a measure of the variability not explained by the regression model. The residuals are the realized or observed values of the errors. Thus any departures from the underlying assumptions on the errors should show up in the residuals. Analysis of the residuals is an effective method for investigating several types of model deficiencies (Jayaraman 2000). The distribution of residuals was analyzed through the histogram to decide either the distribution is normal or abnormal. In both of the models (M1 and M5), the histograms are approximately normal. But, clear differences between these two models could not be observed (Fig 2). The visual examination and comparison of normal probability plot curves for model M1 showed more cluster of residuals points towards the line of equal distribution than M5 (Fig 3). Though some distinguishable differences were observed through normal probability plot curves, analysis of residuals through scatter plots was also carried out to draw a concrete conclusion (Fig 4). The scatter plot of residuals versus the corresponding fitted value is useful for detecting several common types of model inadequacies. The random distribution of the residuals against the predicted values with narrow horizontal band was observed in model M1 than that of model M5. Comparatively, the model M1 showed the homogeneous distribution of residuals against the predicted value than model M5 (Fig 4). The errors tested by the kolmogorov smirnov

($P>0.05$) shows the errors are normally distributed. Therefore, from the interpretation of fit statistics, graphical analysis of residuals and simplicity model M1 i.e. $H = e^{(-0.276)} D^{0.925}$ is considered best fitted model among the available ones. This model shows that unit increase in D (cm), average H (m) is increased by 2.53m in *P. roxburghii*. . The data for this study were collected from small area (from only two community forest areas of one VDC of Myagdi district), therefore, the results might not represent for outside the range of study area. Further researches are needed including large areas which might be applicable for the future also.

Table 3. Model parameter estimates and fit statistics

Model	R^2 adj.	RMSE	Parameter	Parameter value	P
M1	0.910	1.221	β_0	-0.276	0.000*
			β_1	0.925	0.000*
M2	0.906	2.933	β_0	-0.245	0.760
			β_1	0.691	0.000*
			β_2	-0.002	0.001*
M3	0.851	1.490	β_0	-0.333	0.069
			β_1	-0.001	0.039*
			β_2	0.954	0.000*
M4	0.880	1.299	β_0	-0.320	0.014*
			β_1	0.944	0.000*
			β_2	-0.000	0.046*
M5	0.890	1.266	β_0	-0.320	0.002*
			β_1	0.943	0.000*
			β_2	-0.000	0.039*

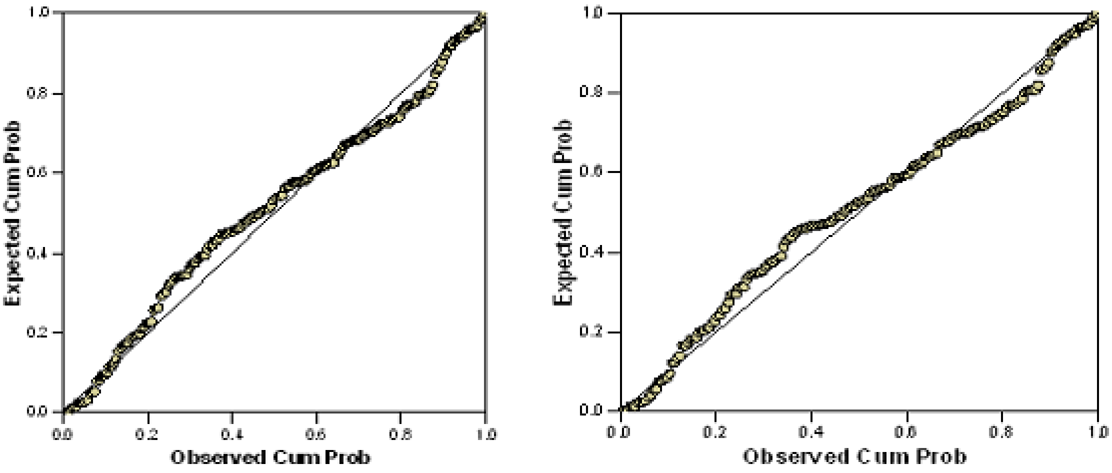
*significant ($P<0.05$)



Regression standaridized residual

Regression standardized residual

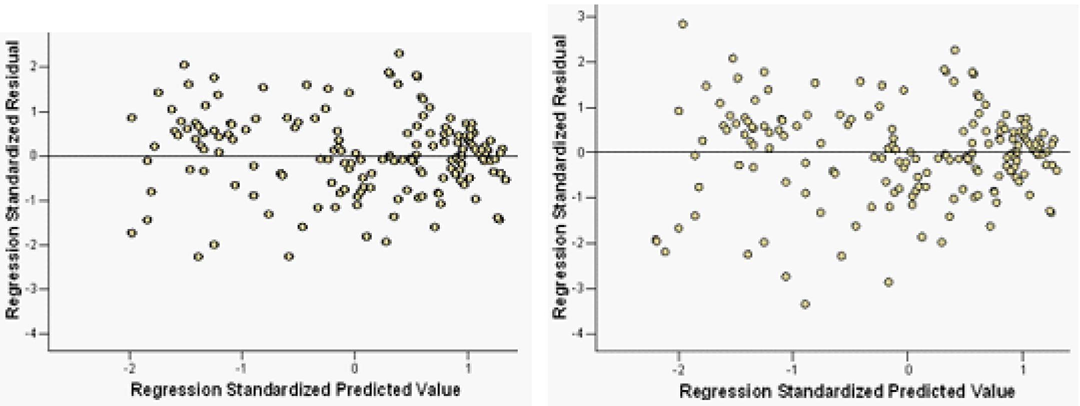
Fig. 2. Histogram of residuals for M1 and M5



M1

M5

Fig.3. Normal P-P curve of standardized residuals of M1 and M5



M1

M5

Fig. 4. Scatter plot of standardized predicted vs. standardized residual of M1 and M5

Acknowledgements

We are thankful to Pahalman Shermuja, Meghraj Poudel, Ananda Khadka and Rijan Tamrakar for their support during the data collection. We are also grateful to Livelihood and Forestry Program (LFP), Baglung for providing financial support. We also express our gratitude to the editor and two anonymous reviewers for their constructive comments on earlier version of the manuscript.

References

Adinugroho, W.C.D. and K. Sidiyasa. 2006. Biomass estimation model of above ground mahogany (*Swieteniamacrophylla king*) tree. *Jurnal Penelitian Hutandan Konservasi Alam* **3**(1):103 – 117.

Arabatzis, A.A. and H.E. Burkhart. 1992. Evaluation of sampling methods and model forms for estimating height-diameter relationships in loblolly pine plantations. *Forest Science* **38**:192–198.

Clutter, J. L., J.C. Fortson, L.V. Pienaar, G.H. Brister and R.L. Bailey. 1983. *Timber Management: a Quantitative Approach*. John Wiley & Sons, New York.

Curtis, R.O. 1967. Height-diameter and height-diameter-age equations for second-growth Douglas-fir. *Forest Science* **13**:365–375.

DFRS 1999. *National forest inventory report.* Department of forest research and survey. Ministry of forest and soil conservation, Kathmandu, Nepal.

Dorado, F.C., U. Dieguez-Aranda, M.B. Anta, M.S. Rodriguez and K. von Gadow. 2006. A generalized height-diameter model including random components for radiata pine plantations in northwestern Spain. *Forest Ecology and Management* **229**(1-3):202-213.

Edwards Jr, T.C., D.R. Cutler and N.E. Zimmermann, L. Geiser and G.G. Moisen. 2006. Effects of sample survey design on the accuracy of classification tree models in species distribution models. *Ecological Modelling* **199**(2):132-141.

Fang, Z. and R.L. Bailey. 1998. Height-diameter models for tropical forests on Hainan island in southern China. *Forest Ecology and Management* **110**:315–327.

Huang, S., S.J. Titus and D.P. Wiens. 1992. Comparison of nonlinear height-diameter functions for major Alberta tree species. *Canadian Journal of Forest Research* **22**:1297–1304.

Jayaraman, K. 2000. *Statistical manual for forestry Research.* Forestry Research Support Program for Asia (FORSPA), FAO, Bangkok. Publication No. 25/2000.

Joshi, M.R. 1985. *Prediction of biomass in a plantation stands of Chir pine (Pinus roxburghii Sarg.) in Nepal*. M.Sc. thesis, University of Oxford, UK.

Joshi, R.B. 1984. Total *and merchantable volume equations for natural silver fir and Chir pine of Nepal*. M. Sc. thesis, University of Athens, Georgia, USA, 35p.

Lama, R. P., S. K. Bhandari & Y.P. Timilsina. 2012. Height-diameter modeling in *Alnus nepalensis* D. Don. *Journal of Forest and Nature* (in press).

Lynch, T.B. and P.A. Murphy. 1995. Compatible height prediction and projection system for individual trees in natural, even-aged shortleaf pine stands. *Forest Science* **41**:194–209.

Newton, P.F. and I.G. Amponsah. 2007. Comparative evaluation of five height-diameter models developed for black spruce and jack pine stand types in terms of goodness-of-fit, lack-of-fit and predictive ability. *Forest Ecology and Management* **247**:149-166.

Rauntiainen, O. 1992. Observations on the growth of planted *Pinus roxburghii* on high quality sites in Kathmandu Valley. *Banko Janakari* **3** (3):37-42.

Sapkota, P. and H. Meilby. 2009. Modelling the growth of *Shorea robusta* using growth ring measurements. *Banko Jankari* **19**(2):25-32.

Sharma, M. and J. Portan. 2007. Height-diameter equations for boreal tree species in Ontario using a mixed-effects modeling approach. *Forest Ecology and Management* **249**:187-198.

Sharma, R.P. 2009. Modeling height-diameter relationship for chir pine trees. *Banko Jankari* **19**(2):3-9.

Sharma, E.R. and T. Pukkala. 1990. *Volume and biomass prediction equations of forest trees of Nepal.* Forest Survey and Statistical Division. Ministry of Forest and Soil Conservation, Kathmandu, Nepal.

Thornley, J. H.M. 1999. Modeling stem height and diameter growth in plants. *Annals of Botany* **84**:195-205.

Trincado, G., L. Curtis, V. Shaaf and H.E. Burkhart. 2007. Regional mixed effects height-diameter models for loblolly pine (*Pinus taeda* L.) plantations. *European Journal of Forest Research* **126**:253-262.

Wagle, B. 2007. *Growth of bluepine (Pinus wallichiana) in Lete and Kunjo of Mustang district*. M.Sc. thesis, Tribhuvan University, Nepal. 55p.