

# The use of tree crown variables in over-bark diameter and volume prediction models

Ramazan Özçelik<sup>(1)</sup>, Maria J Diamantopoulou<sup>(2)</sup>, John R Brooks<sup>(3)</sup>

Linear and nonlinear crown variable functions for 173 Brutian pine (*Pinus brutia* Ten.) trees were incorporated into a well-known compatible volume and taper equation to evaluate their effect in model prediction accuracy. In addition, the same crown variables were also incorporated into three neural network (NN) types (Back-Propagation, Levenberg-Marquardt and Generalized Regression Neural Networks) to investigate their applicability in over-bark diameter and stem volume predictions. The inclusion of crown ratio and crown ratio with crown length variables resulted in a significant reduction of model sum of squared error, for all models. The incorporation of the crown variables to these models significantly improved model performance. According to results, non-linear regression models were less accurate than the three types of neural network models tested for both over-bark diameter and stem volume predictions in terms of standard error of the estimate and fit index. Specifically, the generated Levenberg-Marquardt Neural Network models outperformed the other models in terms of prediction accuracy. Therefore, this type of neural network model is worth consideration in over-bark diameter and volume prediction modeling, which are some of the most challenging tasks in forest resources management.

**Keywords:** Crown Variables, Taper, Back-Propagation ANNs, Levenberg-Marquardt ANNs, Generalized Regression Neural Networks

## Introduction

Crown characteristics are an important component of growth and yield models. Tree crown research contributes to several key forest ecosystem attributes: biodiversity, productivity, forest management, forest environment, and wildlife (Avery & Burkhart 2002). The crown of a tree has a strong influence on stem shape, as foliage provides carbohydrates for tree growth and development of the whole tree and their vertical distribution influences stem shape (Leites & Robinson 2004, Li & Weiskittel 2010). Crown ratio (CR) is considered as an expression of the tree's photosynthetic potential, and therefore commonly included as a key variable in growth and yield models.

Tree stem shape has been commonly modeled using taper models (Muhairwe 1994). Taper models are used to estimate diameter

along the bole at any given height, so tree volume can then be determined based on these diameters and corresponding heights. The auxiliary variables used to increase the accuracy of existing taper equations include: (1) crown dimensions; (2) stand and site variables; and (3) upper stem diameter measurements (Trincado & Burkhart 2006). Larson (1963) reported that within the crown, stem diameters at particular heights are generally smaller when compared with trees of the same dimensions but shorter crowns. As a result, tree boles cannot be completely described as a function of bole length and diameter. In attempts to describe tree taper, numerous models of varying complexity have been advanced. In most mathematical models, taper is modeled in terms of dbh (diameter at breast height) and total height. A few researchers have considered using crown

variables (e.g., crown length - CL, CR, and crown height - CH) as covariates (Newnham 1992, Leites & Robinson 2004, Jiang et al. 2007) for describing tree profiles because of the relationship between crown and stem form development, but previous studies have shown mixed results on the benefit of adding crown variables in taper models. The main crown variable utilized in taper and volume models was CR (Petersson 1999, Jiang et al. 2007, Li & Weiskittel 2010, Jiang & Liu 2011). However, CL is an interesting variable, which may influence the prediction of diameter and volume in combination with CR (Mäkela 2002). For this reason, some forms of CR and CL functions were incorporated into the tree-stem taper and volume prediction models in this study.

For stem taper and volume predictions using regression analysis, an appropriate nonlinear function must first be identified, which is a very difficult task. The main reason that artificial neural network (ANN) applications have received attention is that the methodology is comparable to statistical modeling and ANNs can be seen as a complementary effort (without the restrictive assumption of a particular statistical model) or as an alternative approach to fitting nonlinear models to data. Due to the fact that Neural Networks (NNs) attempt to find the best nonlinear function based on the network's complexity, without the constraint of pre-specified nonlinearity, we investigated their applicability in over-bark diameter and stem volume predictions using the same crown variables previously described.

ANNs have been successfully applied in the field of forest modeling. Among others, ANNs have been used for: (a) prediction of diameter distribution (Leduc et al. 2001); (b) forest attributes prediction (Corne et al. 2004); (c) bark volume prediction for standing trees (Diamantopoulou 2005); and (d) prediction of inside-bark diameter and heartwood diameter (Leite et al. 2011). While ANNs have been applied to the prediction of tree volume (Özçelik et al. 2010), inclusion of crown variables into ANN models has not been reported.

The objective of this study was to investigate the level of improvement in diameter over bark (*dob*) and stem volume predictions through the incorporation of crown variables. An additional purpose of this study was to test the performance of different ANNs that can be employed for diameter and volume predictions through the incorporation of crown variables. For this purpose, a modified segmented polynomial taper equation (Clark et al. 1991) and ANN models were utilized with Brutian pine (*Pinus brutia* Ten.) data collected from southern Turkey.

□ (1) Faculty of Forestry, Süleyman Demirel University, East Campus, TR-32260, Isparta (Turkey); (2) Faculty of Forestry and Natural Environment, Aristotle University of Thessaloniki, GR-54124 Thessaloniki (Greece); (3) Division of Forestry and Natural Resources, West Virginia University, 322 Percival Hall, 26506-6125 Morgantown (WV - USA)

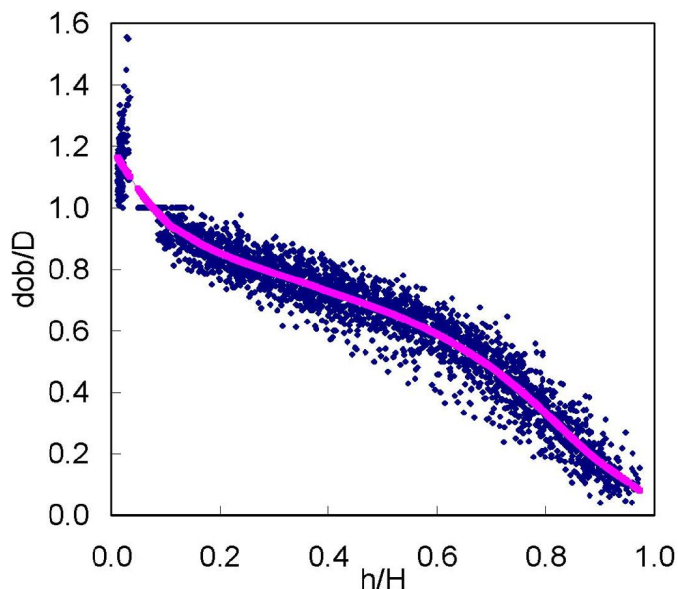
@ Maria J Diamantopoulou (papamich@agro.auth.gr)

Received: Nov 12, 2012 - Accepted: Aug 08, 2013

**Citation:** Özçelik R, Diamantopoulou MJ, Brooks JR, 2014. The use of tree crown variables in over-bark diameter and volume prediction models. iForest 7: 132-139 [online 2014-01-13] URL: <http://www.sisef.it/forest/contents/?id=ifor0878-007>

Communicated by: Luca Salvati

**Fig. 1** - Plot of relative height (h/H) vs. relative diameter (dob/D) over-bark for the Brutian pine trees studied.



**Material and methods**

*Data*

One hundred seventy three sample trees of Brutian pine were selected from natural even-aged managed stands in Bucak Forest Enterprise, southern Turkey, on lands owned by the Forest Service. Trees were felled through the clear-cutting of areas of the Bucak Forest Enterprise and were systematically sampled to cover the range of diameters within a stand, with emphasis on dominant and codominant individuals. Trees possessing multiple stems, broken tops, obvious cankers or crooked boles were not included in the sample. Total height was measured to the nearest 0.05 m. Diameter over-bark (D) at breast height (1.3 m) was measured and recorded to the nearest 0.1 cm using digital calipers. Diameter over-bark was measured at 0.3, 1.3, 2.3 m and then at intervals of 1 m along the remainder of the stem. In each section, two perpendicular diameters over-bark were measured and then arithmetically averaged. The height to base of the live crown was determined by identifying that point along the bole where the lowest live branch or branch whorl was attached to the main

bole as indicated by Jiang et al. (2007). Finally, CL and CR were derived from crown measurements. Crown ratio was defined as the ratio between length of the crown and total tree height. Actual volumes and sectional volumes in cubic meters were calculated using the overlapping bolts method as described by Bailey (1995).

A scatter plot of relative diameter against relative height was examined visually to detect possible anomalies in data. Extreme data points were observed; therefore the systematic approach proposed by Bi (2000) for detecting abnormal data points was applied to increase the efficiency of the process. For this reason, a nonparametric taper curve was fitted by local regression, using the LOESS procedure. This involved local quadratic fitting with a smoothing parameter of 0.25 for the dataset, which was selected after iterative fitting and visual examination of the smoothed taper curves overlaid on the data. The number of extreme values accounted for less than 0.2% for dataset. The plot of relative height against relative diameter used in this study, together with the LOESS regression line, is shown in Fig. 1.

All trees with total height less than 5.3 m

were eliminated, as they could not be used to fit the modified Clark et al. (1991) model. Sample trees were selected to adequately represent the distribution of these trees in the population in terms of their respective diameter and height classes. Approximately 25% of sample trees were selected at random and used as the validation data set, while the rest was used for model fitting. Summary statistics for both data sets are shown in Tab. 1.

*Taper and volume equations*

The modified form of the segmented polynomial model published by Clark et al. (1991) was used for this study, taking into account the conclusions of the work by Jiang et al. (2005). As noted by Jiang et al. (2005) segmented polynomial models appear to be more accurate than other model formulations for estimating diameter, height, and volume.

Two crown variables, CL and CR were included into the best fitted model identified for Brutian pine. Leites & Robinson (2004) incorporated both linear and nonlinear CR and CL functions. The forms of CR and CL functions used in this study include (eqn. 1 to 5):

$$\lambda_1 + \lambda_2 (CR)$$

$$\lambda_1 + \lambda_2 (CL)$$

$$\lambda_1 + \lambda_2 (CL) + \lambda_3 (CR)$$

$$\lambda_1 + \lambda_2 (CR)^{\lambda_3}$$

$$\lambda_1 + \lambda_2 (CL)^{\lambda_3}$$

where  $\lambda_i$  are the parameters to be estimated from data. Eqn. 1 through 5 were incorporated into the existing taper and volume models (Jiang et al. 2005, 2007) for parameters  $b_1, b_2, b_3,$  and  $b_4$  to ascertain the effects of incorporating crown variables into the existing model forms for Brutian pine. The parameters ( $b_1, b_2, b_3$  and  $b_4$ ) in the taper and volume equations were replaced with all combinations of CR, CL and CR with CL functions and were evaluated for model improvement.

In addition to the evaluation of the entire stem, model performance was examined using sectional relative height classes from 10% to 90% of total height. For the taper and volume model forms used in this study, the upper stem diameter at 5.30 m is a required input variable. Diameters at 5.30 m were obtained through actual field measurements.

*Neural network models*

Known advantages of ANNs over traditional approaches (Basheer & Hajmeer 2000, Swingler 2001, Jena et al. 2009, Diamantopoulou 2010) accelerated their use in this re-

**Tab. 1** - Summary of Brutian pine tree attributes for model fitting and for the validation data. (SD): Standard deviation.

Data set	Parameter	Unit	Mean	SD	Min	Max
Fitting data (measurements from 131 trees, n = 2173)	Over-bark dbh (D)	cm	39.85	12.16	9	64
	Total height (H)	m	18.77	3.9	8.8	26.8
	Disk diameter (dob)	cm	26.34	13.27	2	73
	Disk height (h)	m	8.63	5.57	0.3	24.3
	Stem volume (V)	m <sup>3</sup>	1.06	0.77	0.02	3.31
Validation data (measurements from 42 trees, n = 729)	Over-bark dbh (D)	cm	42.39	13.8	11	72
	Total height (H)	m	19.35	4.13	9.5	26.6
	Disk diameter (dob)	cm	27.72	14.13	2	80
	Disk height (h)	m	8.95	5.72	0.3	24.3
	Stem volume (V)	m <sup>3</sup>	1.23	0.98	0.03	4.14

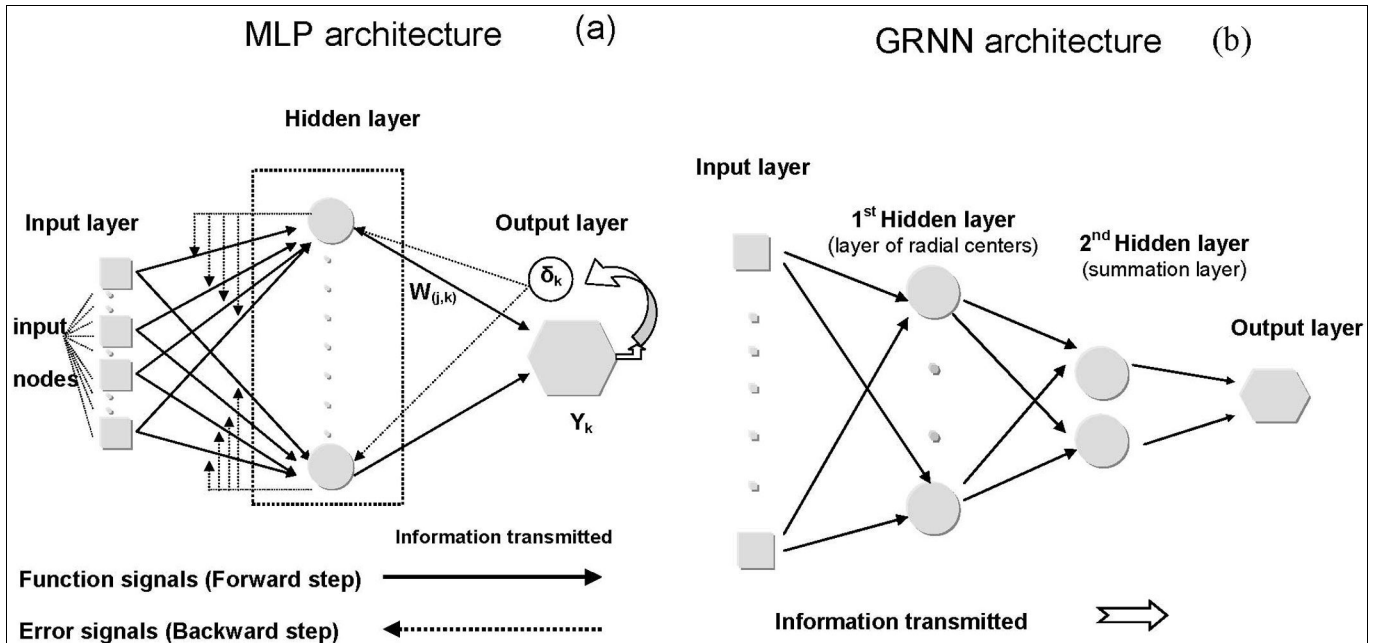


Fig. 2 - (a) The multilayer-perceptron (MLP) and (b) the generalized regression neural network (GRNN) architectures.

search as an alternative approach to regression analysis for fitting nonlinear data. Due to their advantages and limitations (Basheer & Hajmeer 2000, Diamantopoulou 2010), two different NN architectures (Fig. 2) were used: (a) the multilayer-perceptron (MLP); and (b) the generalized regression neural network (GRNN).

Furthermore, in the multilayer-perceptron learning step two different optimization algorithms were used: (1) the back-propagation (BP), which produces the back-propagation artificial neural network (BPANN) models (Rumelhart et al. 1986, Fausett 1994, Haykin 1994, Patterson 1996); and (2) the Levenberg-Marquardt (LM) algorithm which produces the Levenberg-Marquardt artificial neural network (LMANN) models (Levenberg 1944, Marquardt 1963). The latter algorithm was presented as an intermediate optimization technique between Gauss-Newton and gradient descent methods in order to address the limitations of each of these algorithms.

The generalized regression neural network (GRNN) is one type of NN which was devised by Speck (1991). This regression network architecture uses Bayesian techniques to estimate the expected mean value of the output variable.

Appropriate input variables to the NN models can be selected in advance based on *a priori* knowledge of the physical problem, an approach that is commonly utilized in the field of geotechnical engineering (Maier & Dandy 2000). In order for the generated NN models to be comparable to the modified form-class segmented taper equations of Clark et al. (1991), the developed neural net-

work models (BPANN, LMANN and GRNN models) used diameter at breast height over-bark (D), total height of the tree (H), the stem height above the ground to the measurement point (h), the diameter over-bark at 5.3 m above the ground (F) as input variables, and tree diameter (*dob*) over-bark at any given height (h) as the output variable. Moreover, in order to investigate the level of improvement in *dob* and stem volume predictions, the additional crown variables, CL and CR, were incorporated as input variables to the NN models.

For the development of the BPANN models, the effectiveness and convergence of training depends significantly on the values of learning rate (LR) and momentum factor (M). The numbers of neurons in the hidden layer of the ANNs were finalized after a trial and error procedure using different combinations of learning rates and momentum factors. Each combination of LR and M was tested for different numbers of hidden neurons. For the LMANN models development, the effectiveness and convergence of training depends significantly on the adjustment of the damping factor ( $\mu$ ). This was achieved by starting with the value of 0.1 and then using an adjustment factor ( $\nu$ ), which when multiplied by  $\mu$  provides an increment or when  $\mu$  is divided by  $\nu$  provides a decrement, until the lowest sum of square errors value was obtained. For the GRNN models, the network was trained for smoothing coefficient values ( $\sigma$ ) equal to 0.001, 0.002, ..., 1.0, ..., 2.0. The best combinations of all parameters that provide for the best learning of the BPANN, LMANN, and GRNN models are given in Tab. 2.

#### Criteria of model evaluation

The statistics used to compare the models included the average bias ( $B$ ), the standard error of the estimate ( $SEE$ ), the mean absolute error ( $MAE$ ), and a fit index ( $FI$ ). These evaluation statistics are defined as follows (Schlaegel 1981 - eqn. 6 to 9):

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

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

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i|$$

$$FI = 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2}$$

where  $Y_i$  is the observed value for the  $i$ -th observation,  $\hat{Y}_i$  is the predicted value for the  $i$ -th observation,  $\bar{Y}$  is the mean of the  $Y_i$ ,  $df$  are the degrees of freedom of the model,  $n$  is the number of observations in the dataset, and  $SEE$  is the standard error of the estimate.

To concurrently minimize taper and volume errors, both equations were fitted simultaneously using SAS PROC MODEL (SAS Institute Inc 2002). All parameters were shared by both the taper and volume equations. Correlated error structure in the data was not taken into account in the SAS MODEL procedure. Prediction accuracy is little affected by the correlated error struc-

**Tab. 2** - The best combinations of all parameters that conduct to the best learning of the BPANN, LMANN and GRNN models for the prediction of the diameter over-bark (dob) to a measurement point, and the prediction of the stem volume (V) over-bark. (OM): the model without incorporation of the crown variables; (MCR): the model with the CR variable inclusion; (MCRCL): the model with CR and CL variables inclusion; (D): diameter at breast height over-bark (cm); (dob): diameter over-bark (cm) to measurement point at height h; (H): total tree height (m); (h): height above the ground to the measurement point (m); (F): diameter over-bark (cm) at 5.3 m above ground; (V): stem volume over-bark from stump (m<sup>3</sup>); (CL): crown length; (CR): crown ratio.

*BPANN models that resulted to the best learning*

Model	Number of nodes			Number of Epochs	Learning rate	Momentum factor
	Input layer	Hidden layer	Output layer			
OM	4 : (D, H, h, F)	8	1 : (dob)	1000	0.10	0.30
MCR	5 : (D, H, h, F, CR)	10	1 : (dob)	1000	0.10	0.30
MCRCL	6 : (D, H, h, F, CR, CL)	13	1 : (dob)	1000	0.10	0.30
OM	5 : (D, H, dob, h, F)	6	1 : (V)	1000	0.09	0.20
MCR	6 : (D, H, dob, h, F, CR)	10	1 : (V)	1000	0.07	0.30
MCRCL	7 : (D, H, dob, h, F, CR, CL)	11	1 : (V)	1000	0.05	0.30

*LMANN models that resulted to the best learning*

Model	Number of nodes			Number of Epochs	Initial (μ) value	Adjustment factor (ν)
	Input layer	Hidden layer	Output layer			
OM	4 : (D, H, h, F)	8	1 : (dob)	3000	0.1	10
MCR	5 : (D, H, h, F, CR)	10	1 : (dob)	1000	0.1	10
MCRCL	6 : (D, H, h, F, CR, CL)	13	1 : (dob)	1000	0.1	10
OM	5 : (D, H, dob, h, F)	6	1 : (V)	2000	0.1	10
MCR	6 : (D, H, dob, h, F, CR)	10	1 : (V)	2000	0.1	10
MCRCL	7 : (D, H, dob, h, F, CR, CL)	11	1 : (V)	2000	0.1	10

*GRNN models that resulted to the best learning*

Model	Number of nodes			Smoothing coefficient (σ)	
	Input layer	1 <sup>st</sup> Hidden layer	2 <sup>nd</sup> Hidden layer		
OM	4 : (D, H, h, F)	1956	2	1 : (dob)	0.041
MCR	5 : (D, H, h, F, CR)	1956	2	1 : (dob)	0.049
MCRCL	6 : (D, H, h, F, CR, CL)	1956	2	1 : (dob)	0.049
OM	5 : (D, H, dob, h, F)	1956	2	1 : (V)	0.041
MCR	6 : (D, H, dob, h, F, CR)	1956	2	1 : (V)	0.041
MCRCL	7 : (D, H, dob, h, F, CR, CL)	1956	2	1 : (V)	0.039

ture, even when the correlated errors structure is accounted for in the equation fitting process (Kozak 1997). The effects of autocorrelation were ignored.

**Results**

*Approximation through taper and volume equations system*

Crown ratio (CR) and crown length (CL) were incorporated into the existing taper and volume equations. The parameters ( $b_1, b_2, b_3$  and  $b_4$ ) for the existing taper and volume equations (Clark et al. 1991, Jiang et al. 2005) were replaced utilizing both the linear and non-linear functions that incorporate CR, CL, and both crown variables together (eqns. 1 - 5). The full model failed to con-

verge when employing either the linear (eqns. 1 - 3) or nonlinear form of the crown variables (eqns. 4 - 5). All combinations of CR, CL and CR with CL functions with each parameter were tested, but only the replacement of  $b_4$  with the linear eqn. 1 and eqn. 3 resulted in significant parameter estimates ( $P < 0.0001$ ). Parameters estimates for over-bark (dob) taper and volume equations (Clark et al. 1991, Jiang et al. 2005) with and without the linear CR and linear CR with CL are listed in Tab. 3.

Model OM represents the original model forms without the addition of the crown variable functions, while models MCR and MCRCL represent the modified model after incorporating eqn. 1 and eqn. 3 for the  $b_4$  parameter, respectively. Based on the fit sta-

tistics for the OM, MCR, and MCRCL model forms (Tab. 4), inclusion of the linear CR (eqn. 1) and linear CR with CL (eqn. 3) functions improved the fit for over-bark taper equations for Brutian pine.

For *dob* prediction, the inclusion of crown variables (both CR and CR with CL) had a positive effect for all fit statistics except average bias (Tab. 4) when using the actual upper diameter measurement at 5.30 m. A lower bias does not guarantee good model performance, since large positive and negative values may algebraically counterbalance. Since it indicates the spread of the biases (residuals), the overall standard error of estimate is a better single indicator of goodness of fit (Kozak & Smith 1993). The additions of linear functions of CR and CR with CL improved model performance by reducing SEE by 1.18-3.28 % and 1.39-2.78 % for diameter and volume estimation, respectively.

For stem volume prediction, significant improvements were observed in the modified Clark et al. (1991) equation using actual upper stem diameter measurements due to the inclusion of only CR and CR with CL, for

**Tab. 3** - Parameter estimates for the compatible taper and volume equations based on the model fitting data. (OM): the original model forms (Clark et al. 1991); (MCR): the original model form with CR; (MCRCL): the original model form with CR and CL.

Model	$b_1$	$b_2$	$b_3$	$b_4$	$\lambda_1$	$\lambda_2$	$\lambda_3$
OM	85.9076	6.7407	0.6977	2.3021	-	-	-
MCR	84.8311	6.7349	0.6973	-	2.9278	-1.4063	-
MCRCL	85.90755	6.7201	0.7034	-	2.913	0.0988	-3.3829

**Tab. 4** - Stem fit statistics for the compatible volume and taper equation systems for Brutian pine based on the model fitting data. (OM): the original model forms (Clark et al. 1991); (MCR): the original model form with CR; (MCRCL): the original model form with CR and CL.

Model	Taper (cm)				Volume (m <sup>3</sup> )			
	Bias	SEE	MAE	FI	Bias	SEE	MAE	FI
OM	0.045	1.7535	1.1907	0.9825	0.001	0.0072	0.0043	0.9846
MCR	0.0163	1.7328	1.1815	0.983	0.0009	0.0071	0.0043	0.9848
MCRCL	0.0872	1.696	1.1602	0.9837	0.0009	0.007	0.0042	0.985

**Tab. 5** - Fit statistics for the BPANN, LMANN and GRNN models for diameter over-bark (*dob*) to a measurement point and the prediction of stem volume (V) over-bark, based on model fitting data (n = 2173). (OM): the model without incorporation of the crown variables; (MCR): the model with the CR variable inclusion; (MCRCL): the model with CR and CL variables inclusion.

Model	<i>dob</i> (cm)				V (m <sup>3</sup> )			
	Bias	SEE	MAE	FI	Bias	SEE	MAE	FI
BPANN_OM	-0.0089	1.588	1.1418	0.9857	0.00079	0.00279	0.002	0.9977
BPANN_MCR	-0.0072	1.534	1.1252	0.9865	0.00073	0.00259	0.0018	0.998
BPANN_MCRCL	-0.005	1.5045	1.0867	0.9871	-0.00015	0.00247	0.0017	0.9981
LMANN_OM	-0.0154	1.5196	1.0859	0.9869	-1.0·10 <sup>-6</sup>	0.00227	0.0016	0.9985
LMANN_MCR	-0.0184	1.496	1.08	0.9872	1.5·10 <sup>-7</sup>	0.00212	0.0015	0.9987
LMANN_MCRCL	0.0025	1.444	1.0336	0.9881	-4.8·10 <sup>-6</sup>	0.00204	0.0014	0.9988
GRNN_OM	-0.0221	1.5254	1.0712	0.9868	5.7·10 <sup>-5</sup>	0.0028	0.0018	0.9977
GRNN_MCR	-0.0269	1.496	1.012	0.9873	7.5·10 <sup>-5</sup>	0.00268	0.0015	0.9979
GRNN_MCRCL	-0.0211	1.4661	0.9833	0.9878	8.7·10 <sup>-5</sup>	0.0026	0.0014	0.9979

**Tab. 6** - Fit statistics for the compatible volume and taper equation systems and for the BPANN, LMANN and GRNN models for diameter over-bark (*dob*) to a measurement point and the prediction of stem volume (V) over-bark, based on model validation data (n = 729). (OM): the model without incorporation of the crown variables; (MCR): the model with the CR variable inclusion; (MCRCL): the model with CR and CL variables inclusion.

Model	<i>dob</i> (cm)				V (m <sup>3</sup> )			
	Bias	SEE	MAE	FI	Bias	SEE	MAE	FI
OM	0.0751	1.9673	1.3238	0.9806	0.0006	0.0075	0.0048	0.9883
MCR	0.0686	1.9236	1.2885	0.9813	0.0007	0.0071	0.0046	0.989
MCRCL	0.0841	1.9635	1.3307	0.9814	0.0004	0.0076	0.005	0.989
BPANN_OM	0.0935	1.871	1.3066	0.9824	0.0043	0.0055	0.0044	0.9933
BPANN_MCR	0.0707	1.863	1.3291	0.9826	0.0043	0.0054	0.0043	0.9936
BPANN_MCRCL	-0.0178	1.857	1.3188	0.9827	0.0038	0.005	0.0038	0.9945
LMANN_OM	0.0788	1.8093	1.2721	0.9836	0.0035	0.0047	0.0036	0.995
LMANN_MCR	-0.0787	1.7908	1.2976	0.9839	0.0035	0.0046	0.0036	0.9952
LMANN_MCRCL	0.0552	1.7873	1.2717	0.984	0.0033	0.0044	0.0034	0.9956
GRNN_OM	0.5644	2.7244	1.9538	0.9628	0.0032	0.0058	0.0041	0.9926
GRNN_MCR	0.4131	2.7221	2.0848	0.9628	0.0026	0.0056	0.0034	0.9931
GRNN_MCRCL	0.3122	2.7211	2.113	0.9629	0.002	0.0054	0.0034	0.9934

the species studied. For the volume function, the MCRCL model performed slightly better than the MCR model (Tab. 4).

#### Approximation through neural network models

The BPANN, LMANN and GRNN model fit statistics for *dob* and cubic meter volume prediction, with and without the crown variables, are provided in Tab. 5.

Results indicate that the inclusion of both crown variables (CR and CL) had positive effects in *dob* and volume predictions, since their use reduced the standard error of estimate, the bias and the mean absolute errors, and increased the values of the fit index for all types of NN. According to the fitness capability of the different types of NN mo-

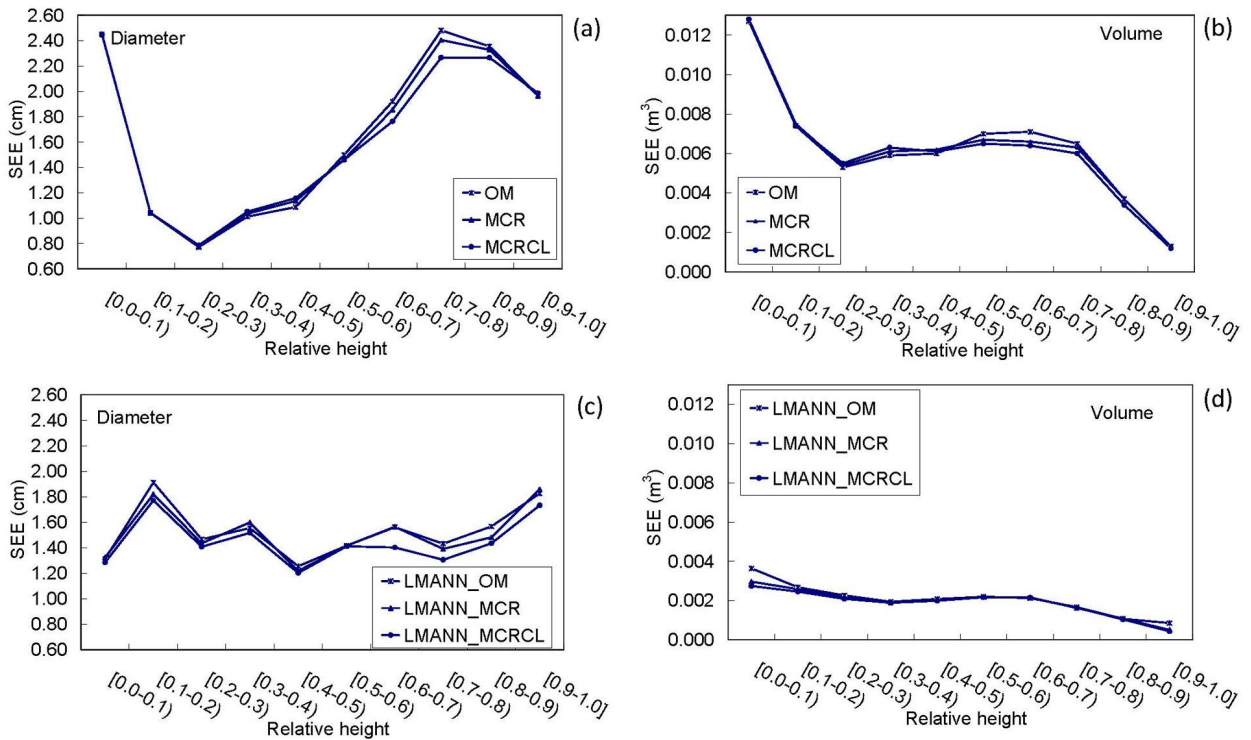
del used, the Levenberg-Marquardt models of Tab. 2 gave the most accurate results (Tab. 5). Specifically, with regards to the *dob* predictions, the inclusion of the CR and CL crown variables into the LMANN model resulted in the reduction of SEE values by 4.02% and 1.51%, according to the BPANN\_MCRCL and GRNN\_MCRCL models, respectively. Further, reductions of the SEE values were observed for volume predictions, as well. Namely, SEE values were reduced by 17.41% and 21.54% according to the BPANN\_MCRCL and GRNN\_MCRCL models, respectively (Tab. 5). In order to validate the fitted NN models, the same fit statistics (*B*, *SEE*, *MAE* and *FI*) were calculated for the predictions of *dob* and cubic meter volume based on the validation data

set (Tab. 6).

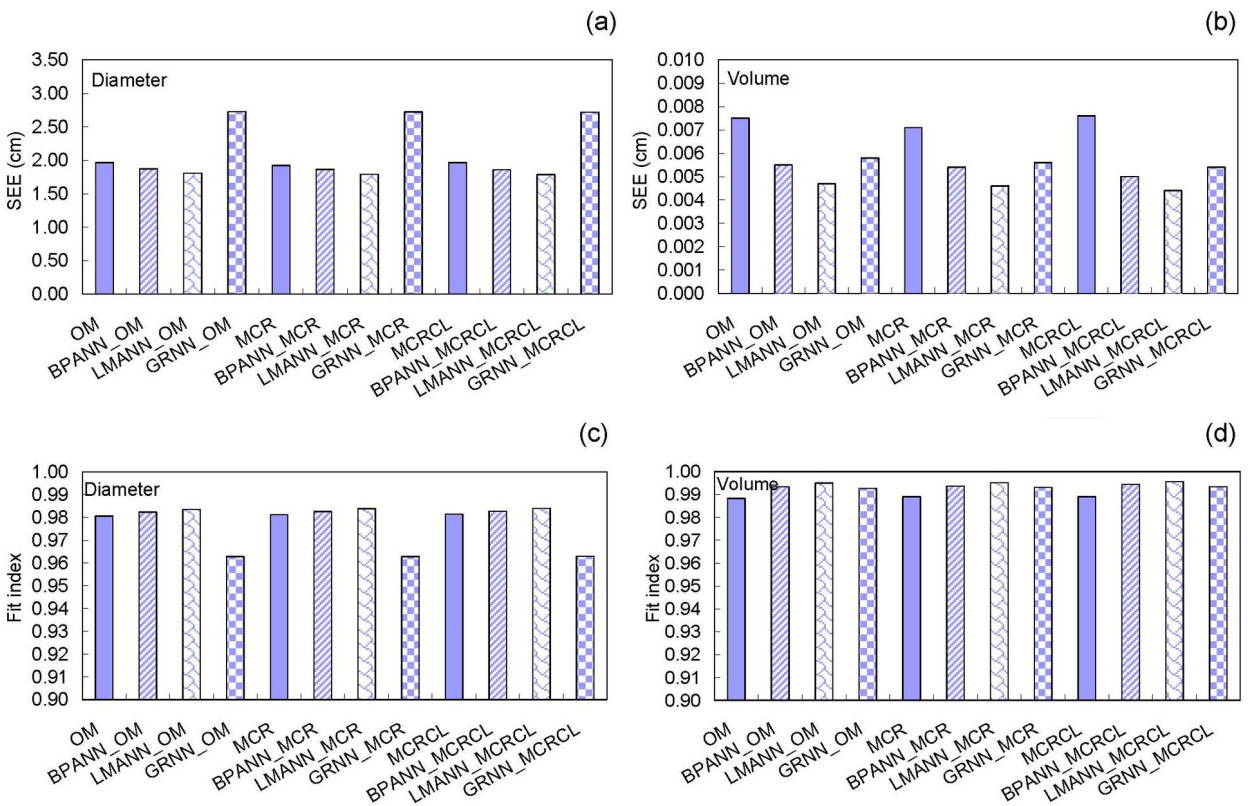
Similarly to what noted with the model fitting data, the inclusion of both crown variables as input variables resulted in the most accurate *dob* and cubic meter volume predictions for all the NN models tested (Tab. 6). Further, the LMANN\_MCRCL model showed consistent performance, with better *dob* and volume predictions than the predictions obtained from the BPANN and GRNN models based on the validation data set.

The SEE for predicting *dob* and cubic meter volume for the 10 relative height classes for both the compatible taper and volume equation system and for the LMANN models are shown in Fig. 3.

As seen in Fig. 3a and Fig. 3c, the *dob* SEE values are significantly lower for the



**Fig. 3** - The standard errors of estimate (SEE) for estimating diameter over-bark for the taper compatible volume system (a) and for the Levenberg-Marquardt models (c) and volume over-bark along the stem for the taper compatible volume system (b) and for the Levenberg-Marquardt models (d), by relative height classes, using the fitting data.



**Fig. 4** - The standard errors of estimate (SEE) and fit indexes (FI) for predicting diameter over-bark and volume over-bark along the stem for the taper and compatible volume system (OM) and for the back-propagation (BPANN\_OM, BPANN\_MCR and BPANN\_MCRCL), Levenberg-Marquardt (LMANN\_OM, LMANN\_MCR and LMANN\_MCRCL), and the generalized regression (GRNN\_OM, GRNN\_MCR and GRNN\_MCRCL) models, using the validation data set.

LMANN\_MCRCCL model across all relative height class except for 10-50% of relative height, where SEE values are slightly lower than the corresponding values of the OM taper model. For volume estimations, SEE values show the same pattern as the *dob* errors and are clearly lower for the LMANN models across all relative height classes than the corresponding values of the nonlinear regression models (Fig. 3b and Fig. 3d).

Validation results for all the models are shown in Fig. 4. The models were found to generalize for data that was not used in the fitting process and showed consistent performance for both *dob* and volume prediction.

The results obtained for the validation data set for all models were in agreement with their performance for the model fitting data set (Tab. 4, Tab. 5 and Tab. 6).

## Discussion

One of the underlying goals for efficient timber resources management is that of optimizing the prediction accuracy of constructed forest-data models. The performance of nonlinear regression models and NN models (BPANN, LMANN and GRNN models) were compared for the estimation of diameter over-bark and cubic meter volume based on Brutian pine data from southeast Turkey.

The incorporation of crown variables to regression and NN modeling procedures showed improvements in the accuracy for diameter and stem volume predictions in Brutian pine. Slightly better results were obtained for estimating stem form than stem volume, when employing taper and volume equations. As indicated by Weiskittel et al. (2011), crown variables often explain very little of the variation in stem volume. The level of improvement is likely a function of the natural variation in crown variables. Evaluation of sectional performance showed an improvement in stem form predictions for relative height classes over 50% of total height. The relative size of these improvements was similar to those found by Burkhart & Walton (1985) and Valenti & Cao (1986). As Jiang et al. (2007) pointed out, this is not a surprising result given the fact that only the  $b_4$  parameter was changed.

For environmental issues, such as forest modeling where the complexity of the natural problem is faced, it is very difficult to suggest a specific approach for a given problem. As pointed out by Özçelik et al. (2010), convenience and economics play the most important role when choosing the method to be used for forest inventory. The determination of the proper approach should take into account both the advantages and limitations of each method. Because of the ability for NNs to automatically fit complex nonlinear models when the complexity of the problem cannot be fully examined, or when

the prediction accuracy is the most important element in a survey, NN models appear to be the best option. Our findings are consistent with previous studies by Diamantopoulou (2005), Pao (2008), Özçelik et al. (2010), Leite et al. (2011), where ANN models generated a better fit when compared to regression models.

In practical forestry, the application of NN models by practitioners can be achieved through the use of trained models that have been constructed by experts for this purpose. However, their use by practitioners requires computational skills but not *a priori* analytical knowledge when the constructed NN models are being provided. Under the above limitation, NN models are accurate and easy to apply. On the other hand, although regression analysis is based on rules or equations that must be explicitly programmed, nonlinear regression models give reasonable and accurate results. Furthermore, regression models are a tested methodology that provide acceptable results which can easily be applied, and thus worth considering as a solution to a given estimation problem. In this case there is a trade-off between selecting a model which is quite simple to understand and easy to apply (regression model), and one which is more accurate but more difficult to comprehend (NN model).

## Conclusions

Accurate estimation of over-bark diameter and stem volume is crucial for the efficient management of forest resources. The inclusion of linear CR and linear CR with CL functions in existing segmented taper and cubic meter volume equations for Brutian pine in Turkey resulted in significant reduction of model sum of squared error. Prediction improvements for upper stem diameter and volume were greater for model forms with CR and CL than model forms with CR alone, though overall improvements were small. Similar results were obtained using the back-propagation, Levenberg-Marquardt and generalized regression neural network models. The incorporation of the crown variables to these models also exhibited improved performance.

Our results indicate that the nonlinear regression model had larger SEE and smaller FI values than the three types of NN models tested, when evaluating both *dob* and volume predictions (Tab. 4 and Tab. 5). Moreover, the performance of the fitted Levenberg-Marquardt artificial neural network models, where both crown variables were embedded, provided superior performance when compared to nonlinear regression, BPANN and GRNN models for both the fitting and the validation data sets (Tab. 4, Tab. 5 and Tab. 6).

Implementation of the NN approaches does offer a number of advantages over the more

traditional regression method of forest-data modeling and should be viewed as a useful alternative to this technique (Özçelik et al. 2010). The major advantage of NNs for over-bark diameter and stem volume modeling is that when they are used, the underlying relationships between the input and output variables are automatically assimilated into the connection weights of the network. Therefore, they are able to fit complex nonlinear models not specified in advance, unlike other nonlinear modeling techniques such as regression analysis. In spite of NN's advantages over regression modeling techniques, the discussed accuracy-convenience trade-offs have to be seriously considered in order to determine the best method to apply. When prediction accuracy is the most important element in a survey, then NN models seem to be the best option. However, when convenience is the limiting factor, or if additional accuracy is not most important, nonlinear models can be utilized.

## Funding

This study was supported by The Scientific and Technological Research Council of Turkey (TUBITAK-BIDEB).

## References

- Avery TE, Burkhart HE (2002). Forest measurements. McGraw-Hill Inc, New York, USA, pp. 480.
- Bailey RL (1995). Upper stem volumes from stem analysis data: an overlapping bolts method. Canadian Journal of Forest Research 25: 170-173. - doi: 10.1139/x95-020
- Basheer IA, Hajmeer M (2000). Artificial neural networks: fundamentals, computing, design, and application. Journal of Microbiological Methods 43: 3-31. - doi: 10.1016/S0167-7012(00)00201-3
- Bi H (2000). Trigonometric variable-form taper equations for Australian eucalyptus. Forest Science 46 (3): 397-409.
- Burkhart HE, Walton S (1985). Incorporation crown ratio into taper equations for loblolly pine trees. Forest Science 31 (2): 478-484. [online] URL: <http://www.ingentaconnect.com/content/saf/fs/2000/00000046/00000003/art00009>
- Clark A, Souter RA, Schlaegel BE (1991). Stem profile equations for southern tree species. Research Paper no. SE-282, Southeastern Forest Experiment Station, USDA Forest Service, Asheville, NC, USA.
- Corne SA, Carver SJ, Kunin WE, Lennon JJ, Van Hees WWS (2004). Predicting forest attributes in southeast Alaska using artificial neural networks. Forest Science 50 (2): 259-276. [online] URL: <http://www.ingentaconnect.com/content/saf/fs/2004/00000050/00000002/art00010>
- Diamantopoulou MJ (2005). Artificial neural networks as an alternative tool in pine bark volume estimation. Computers and Electronics in Agriculture 48: 235-244. - doi: 10.1016/j.compag.2005.04.002

- Diamantopoulou MJ (2010). Filling gaps in diameter measurements on standing tree boles in the urban forest of Thessaloniki, Greece. *Environmental Modelling & Software* 25: 1857-1865. - doi: [10.1016/j.envsoft.2010.04.020](https://doi.org/10.1016/j.envsoft.2010.04.020)
- Fausett L (1994). *Fundamentals of neural networks architectures, algorithms and applications*. Prentice-Hall, Englewood Cliffs, NJ, USA.
- Haykin S (1994). *Neural networks: a comprehensive foundation*. Prentice Hall, NJ, USA. [online] URL: <http://dl.acm.org/citation.cfm?id=541500>
- Jena RK, Aqel MM, Srivastava P, Mahanti PK (2009). Soft computing methodologies in bioinformatics. *European Journal of Scientific Research* 26: 189-203. - doi: [10.1016/j.foreco.2005.04.006](https://doi.org/10.1016/j.foreco.2005.04.006)
- Jiang L, Liu R (2011). Segmented taper equations with crown ratio and stand density for Dahurian Larch (*Larix gmelinii*) in northeastern China. *Journal of Forestry Research* 22: 347-352. - doi: [10.1007/s11676-011-0178-4](https://doi.org/10.1007/s11676-011-0178-4)
- Jiang L, Brooks JR, Wang J (2005). Compatible taper and volume equations for yellow-poplar in West Virginia. *Forest Ecology and Management* 213: 399-409. - doi: [10.1016/j.foreco.2005.04.006](https://doi.org/10.1016/j.foreco.2005.04.006)
- Jiang L, Brooks JR, Hobbs GR (2007). Using crown ratio in yellow-poplar compatible taper and volume equations. *Northern Journal of Applied Forestry* 24 (4): 271-275. [online] URL: <http://www.ingentaconnect.com/content/saf/njaf/2007/00000024/00000004/art00005>
- Kozak A, Smith JHG (1993). Standards for evaluating taper estimating systems. *Forestry Chronicle* 69: 438-444. - doi: [10.5558/tfc69438-4](https://doi.org/10.5558/tfc69438-4)
- Kozak A (1997). Effects of multicollinearity and autocorrelation on the variable-exponent taper functions. *Canadian Journal of Forest Research* 27: 619-629. - doi: [10.1139/x97-011](https://doi.org/10.1139/x97-011)
- Larson PR (1963). Stem form development of forest trees. *Forest Science Monograph* 5 (4): 21-21. [online] URL: <http://www.ingentaconnect.com/content/saf/fs/1963/00000009/a00504s5/art00001>
- Leduc DJ, Matney TG, Belli KL, Baldwin VC (2001). Predicting diameter distributions of longleaf pine plantations: a comparison between artificial neural networks and other accepted methodologies. *Research Paper SRS-25*, Southern Research Station, USDA Forest Service, Asheville, NC, USA, pp. 24.
- Leite HG, Marques da Silva ML, Binoti DHB, Fardin L, Takizawa FH (2011). Estimation of inside-bark diameter and heartwood diameter for *Tectona grandis* Linn. trees using artificial neural networks. *European Journal of Forest Research* 130: 263-269. - doi: [10.1007/s10342-010-0427-7](https://doi.org/10.1007/s10342-010-0427-7)
- Leites LP, Robinson AP (2004). Improving taper equations of loblolly pine with crown dimensions in mixed-effects modeling framework. *Forest Science* 50: 204-212. [online] URL: <http://www.ingentaconnect.com/content/saf/fs/2004/00000050/00000002/art00006>
- Levenberg K (1944). A method for the solution of certain problems in least squares. *Quarterly of Applied Mathematics* 2: 164-168.
- Li R, Weiskittel AR (2010). Comparison of model forms for estimating stem taper and volume in the primary conifer species of North American Acadian region. *Annals of Forest Science* 67: 302-520. - doi: [10.1051/forest/2009109](https://doi.org/10.1051/forest/2009109)
- Maier HR, Dandy GC (2000). Neural networks for the prediction and forecasting of water resources variables: a review of modeling issues and applications. *Environmental Modelling and Software* 15: 101-124. - doi: [10.1016/S1364-8152\(99\)00007-9](https://doi.org/10.1016/S1364-8152(99)00007-9)
- Marquardt D (1963). An algorithm for least squares estimation of non-linear parameters. *Journal of the Society for Industrial and Applied Mathematics* 11: 431-441. - doi: [10.1137/0111030](https://doi.org/10.1137/0111030)
- Muhairwe CK (1994). Tree form and taper variation over time for interior lodgepole pine. *Canadian Journal of Forest Research* 24: 1904-1913. - doi: [10.1139/x94-245](https://doi.org/10.1139/x94-245)
- Mäkela A (2002). Derivation of stem taper from the pipe theory in a carbon balance framework. *Tree Physiology* 22: 891-905. - doi: [10.1093/treephys/22.13.891](https://doi.org/10.1093/treephys/22.13.891)
- Newnham RM (1992). Variable-form taper function for four Alberta tree species. *Canadian Journal of Forest Research* 22: 210-223. - doi: [10.1139/x92-028](https://doi.org/10.1139/x92-028)
- Özçelik R, Diamantopoulou MJ, Wiant HV, Brooks JR (2010). Estimating tree bole volume using artificial neural network models for four species in Turkey. *Journal of Environmental Management* 91: 742-753. - doi: [10.1016/j.jenvman.2009.10.002](https://doi.org/10.1016/j.jenvman.2009.10.002)
- Pao HT (2008). A comparison of neural network and multiple regression analysis in modeling capital structure. *Expert Systems with Applications* 35: 720-727. - doi: [10.1016/j.eswa.2007.07.018](https://doi.org/10.1016/j.eswa.2007.07.018)
- Patterson DW (1996). *Artificial neural networks*. Prentice Hall, Englewood Cliffs, NJ, USA.
- Petersson H (1999). A segmented stem profile model for *Pinus sylvestris*. *Forest Ecology and Management* 124: 13-26. - doi: [10.1016/S0378-1127\(99\)00052-3](https://doi.org/10.1016/S0378-1127(99)00052-3)
- Rumelhart DE, Hinton GE, Williams RJ (1986). Learning internal representations by error propagation. In: "Parallel distributed processing vol. 1" (Rumelhart DE, McClelland JL, Corporate PDP Research Group eds). MIT Press, Cambridge, MA, USA, pp. 318-362.
- SAS Institute Inc (2002). *SAS/ETS User's Guide, Version 9.0*. SAS Institute Inc, Cary, NC, USA.
- Schlaegel BE (1981). Testing, reporting, and using biomass estimation models. In: *Proceedings of the "3rd Meeting of the Southern Forest Biomass Working Group"* (Gresham CA ed). Georgetown (SC - USA), 11-12 June 2001. Bell W Baruch Forest Science Institute, Clemson University, NC, USA, pp. 95-112.
- Spekt DF (1991). A generalized regression neural network. *IEEE Transactions on Neural Networks* 2: 568-576. - doi: [10.1109/72.97934](https://doi.org/10.1109/72.97934)
- Swingler K (2001). *Applying neural networks, a practical guide* (3<sup>rd</sup> edn). Morgan Kaufman Publishers Inc, S. Francisco, CA, USA, pp. 300.
- Trincado G, Burkhart HE (2006). A generalized approach for modeling and localizing stem profile curves. *Forest Science* 52: 670-682.
- Valenti MA, Cao QV (1986). Use of crown ratio to improve loblolly pine taper equations. *Canadian Journal of Forest Research* 16: 1141-1145. - doi: [10.1139/x86-201](https://doi.org/10.1139/x86-201)
- Weiskittel AR, Hann DW, Kershaw JA, Vanclay JK (2011). *Forest growth and yield modelling*. Wiley and Blackwell, London, UK, pp. 415. - doi: [10.1002/9781119998518](https://doi.org/10.1002/9781119998518)