Tree-Bole Volume Estimation on Standing Pine Trees Using Cascade Correlation Artificial Neural Network Models

M.J. Diamantopoulou

Aristotle Univ. of Thessaloniki, Faculty of Forestry and Natural Environment, Lab. of Forest Biometrics, N. Vamva 7-9, GR-54 250 Thessaloniki, Greece.

papamich@agro.auth.gr

ABSTRACT

Total tree volume estimation is an integral part of forest growth and yield forecasting. Complex formulae are used to estimate bole volume by section, based on relationships proposed by Huber, Smalian and Newton. All these relationships require many measurements of bole diameters at certain heights that are difficult to obtain on standing trees especially when diameter measurements have to be taken several meters above ground. The common practice used till now days to face the problem is the application of regression analysis for tree-bole estimation, but there are many problems to be solved and assumptions to be carefully selected etc. In this paper an attempt was made to overcome the above difficulties by indirect tree volume estimation using the necessary values of the diameters at certain heights and the Cascade Correlation Artificial Neural Network models (CCANNs). The cascade correlation algorithm accomplished the training of the ANNs, which is a feed-forward and supervised learning algorithm. Adaptive gradient and Kalman's learning rules were used to modify the artificial neural networks weights. Kalman's learning rule was found superior for the estimation of diameter values at certain heights of the tree-bole. The networks are designed to adapt weights of the synapses, by using the hyperbolic-tangent function of training. The reliability of the developed CCANNs is assessed by validation on independent testing data set. Paired t-test and 45-degree line test were also used for validation of the selected CCANNs. The system proposed in this paper, can be applied in forest inventory calculations producing an accurate estimate of any bole section volume. For example, total tree-bole volume estimation resulted to a root mean square error value of 0.0054 m^3 (9.2%). This tree-bole volume estimation is based only on two diameter measurements (stump diameter, $d_{0.3}$ and diameter at breast height, $d_{1.3}$) and an estimate of total tree height (h), and is accurate enough to replace many standard forestry measurement procedures.

Keywords: Cascade correlation, neural networks, adaptive gradient, Kalman filtering, diameters, volume, pine trees

1. INTRODUCTION

Financial exploitation of forests composes an important part of man activity. Volume is the most widely used measure of wood quantity and tree volume can be estimated from empirical relationships between certain tree-bole dimensions and the tree volume. There have been numerous approaches for modeling the bole volume of a tree (Cao *et al.*, 1980; Bailey, 1994;

Philip, 1994). A well-known and accepted procedure in finding relationships between predictable quantities and predictors is regression analysis (Draper and Smith, 1998). Artificial neural networks (Dowla and Rogers, 1995; Gurney, 1999) are becoming a very popular estimation tool, because they do not require no assumptions about the form of a fitting function. Data-driven models, such as artificial neural network models, which can discover relationships from data without having the complete physical understanding of the system, are preferable, due to their superior ability to resolve the nonlinear nature of the relationships, given a sufficient training data set.

In recent years, Artificial Neural Network (ANN) models have become extremely popular for prediction and forecasting in a number of areas, including finance, power generation, medicine, water resources and environmental science (Maier and Dandy, 2000; Diamantopoulou *et al*., 2005). Artificial neural networks are gaining wide acceptance in control and management applications in agriculture (Sigrimis *et al.,* 1999). With the ANN models new opportunities have emerged to enhance the tools we use to process data. When combined with the data storage and processing capabilities of Geographic Information Systems (GIS), ANN models promise to provide new analytical capabilities in modeling management operations and decision support systems (Ayala *et al.,* 1999; Cros *et al.,* 2003; Satake *et al.,* 2003). A number of researchers have investigated the applicability of artificial neural network models to the field of forest modeling (Guan and Gertner, 1991; Blackard and Dean, 1999; Zhang *et al*., 2000; Hasenauer *et al*., 2001; Leduc *et al.*, 2001; Liu *et al.*, 2003; Corne *et al.*, 2004; Diamantopoulou, 2005a, 2005b). In general, artificial neural networks are appropriate in modeling situations (Bailey and Thompson, 1990; Leduc *et al.*, 2001): 1. where the application is data intensive and dependent on multiple interacting parameters; 2. where the problem area is rich in historical data or examples for their using as training data; 3. where the available data are incomplete and contain errors for filling missing values, and 4. when the function to determine solutions is unknown or expensive to discover.

Most of these conditions are to some degree met by pine trees stem diameter prediction functions investigated in this paper. So, an attempt was made to identify a neural network procedure that will be available for accurate estimation of stem diameters at certain heights as an alternative to a typical regression relationship. Such neural network models could replace many standard forestry mensuration procedures difficult to apply in a forest environment, in order to use them for estimating not only the total stem volume, but also the volume of many parts of the tree-bole.

2. MATERIALS AND METHODS

A total of 94 pine trees from the Seich – Sou urban forest of Thessaloniki, Greece, were measured. This forest is an almost pure pine (*Pinus brutia*) forest. Systematic sampling was used to ensure that all different site classes would be included. Tree measurements included: 1. stump diameter (0.3m height from ground, $d_{0.3}$), and diameter at breast height (1.3m height from ground, $d_{1.3}$), both measured by Finnish caliper (Philip, 1994) and 2. all diameters at one-meter height interval above breast height ($d_{2.3}$, $d_{3.3,}$ $d_{4.3,}$ $d_{5.3,}$ $d_{6.3}$, $d_{7.3}$, $d_{8.3}$ and $d_{9.3}$) and total height (h) of the sampled trees, both measured by Speigel Relaskop (Philip, 1994).

2.1 Artificial Neural Networks

Artificial neural networks are simply a system of interconnected computational units, or nodes. A simple neural network consists of an input, a hidden and an output layer. The input layer has a certain number of nodes (input nodes) that receive data from sources external to the network and send them to the second, the hidden layer. The hidden layer consists of hidden nodes, which send and receive data only from other nodes in the network. The third (output) layer consists of output nodes that receive data generated by the network and produce outputs (Figure 1).

The mathematical form of a three layer feed-forward neural network having n_I input nodes, n_H hidden and n_K outputs, is given by (Gupta *et al.*, 2000):

$$y_j(t) = f(s_j^h(t)) = f\left(\sum_{i=0}^{n_I} W_{ji}^h x_i(t)\right), \quad \text{for } j = 1, \ldots, n_H, \text{ number of the hidden nodes} \tag{1}$$

$$z_k(t) = f(s_k^o(t)) = f\left(\sum_{j=0}^{n_H} W_{kj}^o y_j(t)\right), \quad \text{for } k = 1, \ldots, n_K, \text{ number of the output nodes} \tag{2}$$

where $x_i(t)$ is the input value of node i of the input layer, $y_j(t)$ is the quantity computed by the node j of the hidden layer and $z_k(t)$ is the output computed by the node k of the output layer, W_{ji}^h is the adjustable weight that controls the strength of the connection between input node i and hidden node j, and W_{kj}^o is the adjustable weight that controls the strength of the connection between hidden node j and output node k.

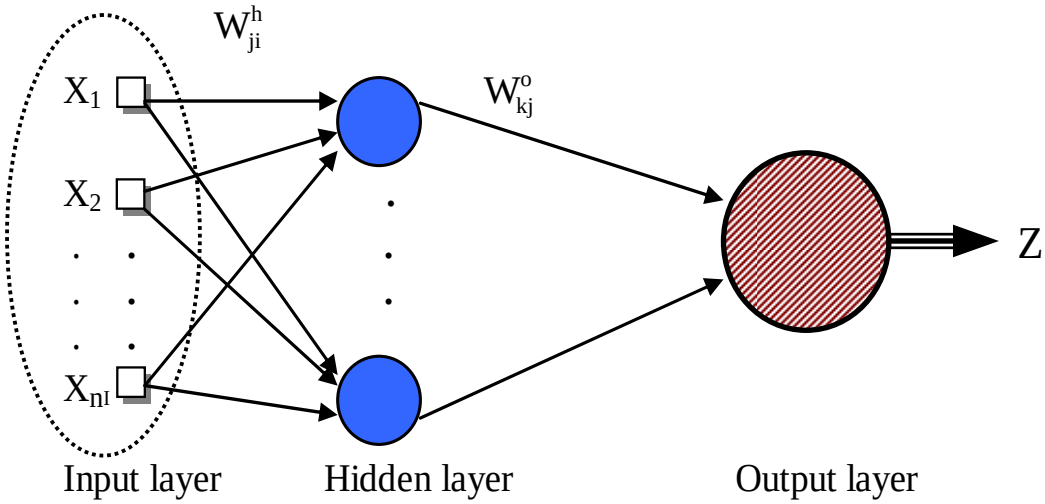

Figure 1. A three-layer feed-forward neural network architecture.

In this study, the training of artificial neural network models suitable for the current application is the cascade correlation algorithm (Fahlman and Lebiere, 1990), which produces the Cascade Correlation Artificial Neural Network (CCANN) that is a feed-forward type with a supervised training algorithm.

For the training of a neural network model, it is very important to split the data into a training and a testing data set to ensure against overfitting and in an effort capturing not just general patterns but sampling variations, as well (Leahy, 1994). For CCANN models construction, data were randomly partitioned into training (90% of all data) and test (the remaining 10% of all data) data sets were used.

As mentioned above, the type of training of CCANNs selected for evaluation is supervised feed-forward. That is the process where the information flows in one direction from the input layer to the hidden layer to the output layer and adjusting the weights in the neural net using a learning algorithm, as well. Adaptive gradient descent algorithm (Jacob, 1988; Charalambous, 1992; Gupta *et al.*, 2000) and Kalman's learning algorithm (Kalman, 1960; Brown and Hwang, 1992; Grewal and Andrews, 1993; Demuth and Beale, 2001) were used. The adaptive gradient method is relatively simple to understand and implement. Begins with a guess for the values of the weights and the weights are iteratively updated by adjustments in a direction that tends to most rapidly reduce the sum of squared-output errors. Training is considered to be completed when values of the weights have been found for which the gradient of the sum squared-output errors function is almost zero. For a specified number of hidden nodes the best values of the network weights can be found by minimizing the sum of squared-output errors. During training, the adaptive gradient learning rule modifies the weights in response to the training data. The weights of a processing element can latch onto spurious information in the training data, such as data that does not represent a general trend in the input data (overtraining). Only the general trends remain encoded in the weights.

Kalman's learning rule is acceptable to regression type problems in which the number of inputs is not too large. In addition, for these applications, since it is not always possible or desirable to measure every variable that one wishes to control, a Kalman filter is used, because it provides a means for inferring the missing information from indirect and noisy measurements (Grewal and Andrews, 2001). This filter combines all available measurement data in order to produce an estimate of the desired variables in such a way that the error is statistically minimized. A Kalman filter is an optimal recursive data processing algorithm. It includes two phases: Predict and Update. The predict phase uses the estimate from the previous step to produce an estimate of the current state. In the update phase, measurement information from the current step is used so as to refine this prediction and arrive at a new, more accurate estimate. The feed-forward ANN models can be trained by adjusting their weights using a stream of input-output observations. The objective is to obtain a set of weights, so that the neural network accurately predicts future outputs. The architecture that was used is the Cascade method of training, based on the Cascade Correlation (Fahlman and Lebiere, 1990). This architecture represents a kind of meta-algorithm, in which classical learning algorithms, like adaptive gradient or Kalman's, are embedded. The Cascade part refers to the architecture and its mode of construction entails adding hidden units one at a time and always connecting all the previous units to the current unit. The Correlation part refers to the way hidden units were trained by trying to maximize the correlation between

output of the hidden unit and the desired output of the network across the training data.

Generally, the objective of the training algorithm needed by the network is to reduce the global error e (Deo and Thirumalaiah, 2000) by adjusting the weight and biases:

$$e = (1/P) \cdot \sum_{p=1}^{P} e_p \tag{3}$$

where: P is the total number of training patterns and e_p is the error for the training pattern p defined by:

$$e_p = (1/2) \cdot \sum_{i}^{n} (O_i - d_i)^2 \tag{4}$$

where: n_K is the total number of the output nodes, O_i is the network output at the i^{th} output node and the di is the desired target output.

The critical question is how many hidden nodes are required with concern to overfitting, undertraining and training efficiency. The cascade correlation algorithm starts the training without any hidden nodes. If the error between the network realized output and the target is not small enough, it adds one hidden node. This node is connected to all other nodes except the output nodes. Because of its dynamic expansion that continues until the problem is successfully learned the cascade correlation algorithm automatically consists a suitable algorithm for a given problem. This procedure goes on until the correlation between the hidden node's output and the residual error of the network, is maximized (Deo and Thirumalaiah, 2000):

$$S = \sum_{O} \left| \sum_{P} (v_p - \bar{v}) \cdot (e_{p,o} - \bar{e}_o) \right| \tag{5}$$

where: o is the o^{th} node, p is the p^{th} train pattern, v_p is the candidate node's value at p^{th} training pattern, $\bar{v}$ is the average of v over all patterns, $e_{p,o}$ is the residual error observed at node o at p^{th} training patern and $\bar{e}_o$ is the average of $e_{p,o}$ over all patterns.

The architecture of the neural networks used is three-layer feed-forward neural network architecture. There is one input layer composed of three input variables ($d_{0.3}$, $d_{1.3}$, h), one output layer composed of one output variable (d_i, different in each CCANN model) and between them one hidden layer composed of a number of nodes that has to be specified (Yuan *et al.* 2003). The optimal number of hidden nodes is commonly determined by trial and error. The usual approach is to begin with a small number of hidden nodes and train the network, iteratively repeating the process for an increasing number of nodes till no further improvement in network performance is obtained. The geometry of CCANNs, which determines the number of connection weights and how these are arranged, depends on the number of hidden layers and the number of the hidden nodes in these layers. In the developed

CCANNs, maximizing the correlation between output of the hidden unit and the desired output of the network across the training data optimizes the number of the hidden nodes. However, the final network architecture and geometry are tested to avoid over-fitting as suggested by Maier and Dandy (2000). The activation function selected to be a continuous and bounded nonlinear transfer function form (Fausett, 1994):

$$f(s) = \tanh(s) = \left(1 - e^{(-2s)}\right) \Big/ \left(1 + e^{(-2s)}\right) \quad (6)$$

where: $s = \sum w_i x_i$ .That is the information the node transmits, in which w_i are the weights and xi are the input values with $s \in [-\infty, +\infty]$ and $f(s) = \tanh(s) \in (-1, +1)$.

The indicators calculated for total data set, training and test set were the correlation coefficient (R), the average absolute error (AVABE), the maximum absolute error (MAE), the root mean square error (RMSE) and the % root mean square error (RMSE%) of the mean of each diameter used as the output node, of each trained network.

In addition, using train and test data sets validation of the chosen CCANN models was also made. For this, the measured values of diameters at 2.3, 3.3, 4.3, 5.3, 6.3, 7.3, 8.3 and 9.3 meters above ground were compared with the corresponding values predicted by the chosen CCANN models, for each diameter separately. The comparisons were made with the help of paired t-test and 45–degree line test. In addition, using train and test data sets the chosen CCANN models were also validated.

3. RESULTS AND DISCUSSION

The descriptive statistics of stump diameter ($d_{0.3}$), diameter at breast height ($d_{1.3}$), all diameters at one-meter height intervals above breast height ($d_{2.3}$, $d_{3.3}$, $d_{4.3}$, $d_{5.3}$, $d_{6.3}$, $d_{7.3}$, $d_{8.3}$ and $d_{9.3}$) and total height (h) of the sampled trees, are given in Table 1.

Table 1. Descriptive statistics of pine trees diameters, and total height

Variable	Min.	Max.	Mean	Variance	Number of trees
$d_{0.3}$, cm	9.00	39.00	17.87	33.90	94
$d_{1.3}$, cm	6.00	38.50	13.69	35.20	94
$d_{2.3}$, cm	4.14	30.90	10.72	23.20	94
$d_{3.3}$, cm	2.54	23.88	8.65	17.67	94
$d_{4.3}$, cm	0.32	21.34	6.65	16.77	93
$d_{5.3}$, cm	1.27	17.52	5.58	12.97	72
$d_{6.3}$, cm	0.95	14.33	4.94	11.27	47
$d_{7.3}$, cm	1.27	12.10	4.75	10.05	25
$d_{8.3}$, cm	1.27	9.55	4.63	7.18	13
$d_{9.3}$, cm	2.87	6.67	4.28	2.53	7
h, m	3.3	12.0	6.44	2.76	94

Different networks structures using using Kalman's and adaptive gradient learning algorithms were tested in order to determine the weights and the optimum number of hidden nodes in each hidden layer (Table 2).

Adaptive gradient learning rule uses back-propagated gradient information to guide an iterative line search algorithm. Kalman's filter learning rule considers the weights to be states and the desired outputs to be the observations within a discrete space transition framework.

Standard non-linear Kalman filter theory via algorithms coded in Matlab (Grewal and Andrews, 2001) is used to obtain the best estimate of the weights based on the stream of training data. The best solution given by the network that composed of one input layer with three input variables, one hidden layer with specific number of nodes for each output variable (Table 2) and one output layer with one output variable. The number of nodes in this one hidden layer was determined based on the maximum value of correlation coefficient. Table 2 clearly shows that the correlation coefficients (R) of the selected CCANN models by using the adaptive gradient learning rule are less than those using the Kalman's learning algorithm, for all output variables. Consequently, for this study, better results have been reached by using Kalman's learning algorithm.

Table 2. The correlation coefficient and the optimum number of nodes in the hidden layers of the selected CCANN model, for each output variable for the training data, using Kalman's and adaptive gradient learning algorithms

Output variable (in cm)	$d_{2.3}$	$d_{3.3}$	$d_{4.3}$	$d_{5.3}$	$d_{6.3}$	$d_{7.3}$	$d_{8.3}$	$d_{9.3}$
Kalman's learning algorithm								
Optimum number of nodes in the hidden layer	3	2	2	1	1	1	1	1
Maximum value of Correlation Coefficients	0.982	0.974	0.984	0.983	0.991	0.986	0.988	0.992
Adaptive gradient learning algorithm								
Optimum number of nodes in the hidden layer	1	6	3	4	1	2	1	1
Maximum value of Correlation Coefficients	0.949	0.948	0.957	0.934	0.909	0.899	0.886	0.986

The hyperbolic transfer function (eq. 6) was chosen as the activation function of all nodes and networks tested.

The correlation coefficient (R), the average absolute error (AVABE), the maximum absolute error (MAE), the root mean square error (RMSE) and the % root mean square error (RMSE%) of the mean of the variable used as output in each network, between the output of each hidden unit and the desired output of each network, for total data set, training and test set using Kalman's learning algorithm are given in Table 3.

Table 3. R, AVABE, MAE and RMSE (%) for total, training and test data set for all output variables, using Kalman's learning algorithm

Variable	Data set	R	AVABE	MAE	RMSE (%)
$d_{2.3}$	Total	0.9823	0.709	3.479	0.9338 (8.71%)
	Train	0.9819	0.710	3.479	0.9480 (8.85%)
	Test	0.9920	0.703	1.282	0.8046 (7.59%)
$d_{3.3}$	Total	0.9757	0.708	4.161	0.9261 (10.71%)
	Train	0.9744	0.736	4.161	0.9582 (11.06%)
	Test	0.9902	0.452	1.354	0.5910 (6.87%)
$d_{4.3}$	Total	0.9836	0.577	1.989	0.7377 (11.09%)
	Train	0.9838	0.579	1.989	0.7459 (11.19%)
	Test	0.9824	0.565	1.414	0.6656 (10.24%)
$d_{5.3}$	Total	0.9839	0.496	1.719	0.6403 (11.48%)
	Train	0.9834	0.509	1.719	0.6574 (11.72%)
	Test	0.9887	0.390	0.908	0.4824 (9.04%)
$d_{6.3}$	Total	0.9903	0.255	2.449	0.4810 (9.74%)
	Train	0.9910	0.237	2.449	0.4732 (9.83%)
	Test	0.9860	0.403	0.859	0.6004 (10.03%)
$d_{7.3}$	Total	0.9822	0.424	1.360	0.5528 (11.64%)
	Train	0.9862	0.425	1.360	0.5403 (11.21%)
	Test	0.9944	0.422	1.089	0.6369 (14.99%)
$d_{8.3}$	Total	0.9878	0.365	0.853	0.4671 (10.09%)
	Train	0.9883	0.352	0.853	0.4443 (9.474%)
	Test	0.9999	0.433	0.813	0.5764 (13.41%)
$d_{9.3}$	Total	0.9755	0.338	0.939	0.2807 (6.56%)
	Train	0.9924	0.394	0.939	0.2905 (5.05%)
	Test	----	----	----	----

The corresponding values of the root mean square error (RMSE) and the % root mean square error (RMSE%) of the mean of the variable used as output in each network, using the adaptive gradient learning rule, are given in Table 4.

Table 4. RMSE and RMSE (%) for total, training and test data set for all output variables, using adaptive gradient learning rule

Output var.	$d_{2.3}$	$d_{3.3}$	$d_{4.3}$	$d_{5.3}$	$d_{6.3}$	$d_{7.3}$	$d_{8.3}$	$d_{9.3}$
Total	1.510	1.310	1.247	1.297	1.297	0.934	0.870	0.597
	(14%)	(15%)	(19%)	(23%)	(23%)	(20%)	(19%)	(14%)
Train	1.531	1.352	1.237	1.329	1.329	1.841	1.800	0.932
	(14%)	(16%)	(18%)	(24%)	(24%)	(17%)	(17%)	(11%)
Test	1.318	0.877	1.333	1.006	1.006	0.986	0.54	---
	(12%)	(10%)	(20%)	(19%)	(19%)	(24%)	(13%)	

In order to estimate all diameter values of tree bole at one-meter height intervals above 1.3 meters from the ground, the traditional multiple linear regression (MLR) method was also used for an indicative comparison. In this study, the total data sets were used for the regression model building work. MLR technique in statistical package SPSS (Norusis, 2000) was used to develop estimation equations for the same output diameters as dependent variables ($d_{2.3}$, $d_{3.3,}$ $d_{4.3,}$ $d_{5.3,}$ $d_{6.3,}$ $d_{7.3,}$ $d_{8.3}$ and $d_{9.3}$) and the corresponding to the selected CCANN models input parameters ($d_{0.3,}$ $d_{1.3}$ and h) as independent variables. The corresponding values of the root mean square error (RMSE), used as the basic indicator, of the MLR models, for all dependent variables, for the total data sets, are given in Table 5.

Table 5. Root mean square error (RMSE) of the MLR models, for the total data sets

Dependent var.	$d_{2.3}$	$d_{3.3}$	$d_{4.3}$	$d_{5.3}$	$d_{6.3}$	$d_{7.3}$	$d_{8.3}$	$d_{9.3}$
Total data set	94	94	93	72	47	25	13	7
RMSE	1.72	1.73	1.94	1.88	2.08	2.28	1.84	1.01

According to the results of Tables 3, 4 and 5, it is clear that for the same output variable, the MLR models precision could not be considered sufficient as for all dependent variables the values of RMSE were higher.

Comparing the RMSE and RMSE (%) values of Table 3 with those of Table 4, it is clear that the CCANN models trained using the adaptive gradient learning rule are considered of inferior accuracy.

Finally, the best prediction solution for all diameter data sets has been given by the network that composed of one input layer with three input variables ($d_{0.3}$, $d_{1.3}$ and h), one hidden layer with specific number of nodes for each output variable (Table 2) and one output layer with one output variable, trained using the Kalman's learning algorithm. The number of nodes in this one hidden layer was determined based on the maximum value of correlation coefficient (Table 2).

The computed t-values and slopes for the selected CCANN models for all diameters, using train and test data sets, are given in Table 5. The computed t-values for the CCANN models were less than two-sided tabular t-values (a=0.05), for both train and test data sets. These imply that there were no significant differences between the measured and the predicted values. Also, it can be observed that the CCANN models tended to make an angle of 45 degrees with the axes, meaning there is no significant difference between the measured and the predicted values.

As mentioned above data randomly partitioned into training (90% of all data) and test (the remaining 10% of all data) data sets, were used for the CCANN models training and gradual build. It is very important to mention that the test data sets were never seen by the selected models in the training procedure. After the confirmation of the validity of the chosen CCANN models, we can get easily the diameter values at one meter height intervals from breast height of single pine trees of $9\text{cm} \le d_{0.3} \le 39\text{cm}$, $6\text{cm} \le d_{1.3} \le 38.5\text{cm}$ and $3.3\text{m} \le h \le 12\text{m}$.

Table 5. t-values and slopes for the CCANN models, for all diameters, using train and test data sets

Model	Variable	Data set	|t – value|	Two-sided tabular value (a=0.05)	Slope (^{o})
CCANN	$d_{2.3}$	Train	1.364	1.989	45.8
CCANN	$d_{2.3}$	Test	2.112	2.262	46.5
CCANN	$d_{3.3}$	Train	0.210	1.989	44.8
CCANN	$d_{3.3}$	Test	0.768	2.262	44.0
CCANN	$d_{4.3}$	Train	0.390	1.989	45.4
CCANN	$d_{4.3}$	Test	0.402	2.262	45.2
CCANN	$d_{5.3}$	Train	0.209	1.999	44.8
CCANN	$d_{5.3}$	Test	0.027	2.365	45.1
CCANN	$d_{6.3}$	Train	2.012	2.019	46.0
CCANN	$d_{6.3}$	Test	0.015	2.776	45.8
CCANN	$d_{7.3}$	Train	0.950	2.080	45.7
CCANN	$d_{7.3}$	Test	1.242	4.303	43.0
CCANN	$d_{8.3}$	Train	1.175	2.228	43.5
CCANN	$d_{8.3}$	Test	1.141	12.71	41.0
CCANN	$d_{9.3}$	Train	1.742	2.571	47.1
CCANN	$d_{9.3}$	Test	----	----	----

The diameter values produced by CCANN models could be used for several purposes. For example, the values of these diameters are necessary for efficient estimation of a single tree volume using the well-known Smalian's sectional method (Philip, 1994):

$$v = \sum_{i=1}^{k} v_i + v_t = (\pi/4)\cdot\left[\left(d_{0.3}^2 + d_{1.3}^2\right)/2\right] + \ldots + (\pi/4)\cdot\left[\left(d_{k-1}^2 + d_k^2\right)/2\right] + (\pi/12)\cdot d_k^2 \cdot l_t \qquad (7)$$

where: v_i is the i_{th} volume of the k sections of the bole; v_t is the top section volume; $d_{1.3,}$ $d_{2.3,}$ …, d_k are all diameters at one-meter height intervals above breast height and l_t is the length of the top bole section.

The average absolute error (AVABE), the maximum absolute error (MAE), the root mean square error (RMSE) and the % root mean square error (RMSE%) of the mean between the total volume values (v_{CCANN}) as calculated by Smalian's sectional method using diameter values estimated by the CCANN models and the measured total volume values (v), for the 94 pine trees, are given in Table 6 and are shown in Figure 2.

Table 6. R, AVABE, MAE and RMSE (%) for the total volume values (v_{CCANN})

Variable	Data set	AVABE	MAE	RMSE	RMSE (%)
v_{CCANN}	Total	0.00365	0.02103	0.00538	9.217

The measured (v) values and the estimated (v_{CCANN}) values yielded slope equal to 44.04^o. The proximity of each point to the 45-degree line (Figure 2a) and the similar height of the columns (Figure 2b), throughout the range of the measured volume indicate that the ANN model is reasonable accurate. Moreover, the data of Figure 2 were tested using the paired t-test and the percentage estimation error. The computed t-value = 0.2211 was less than the

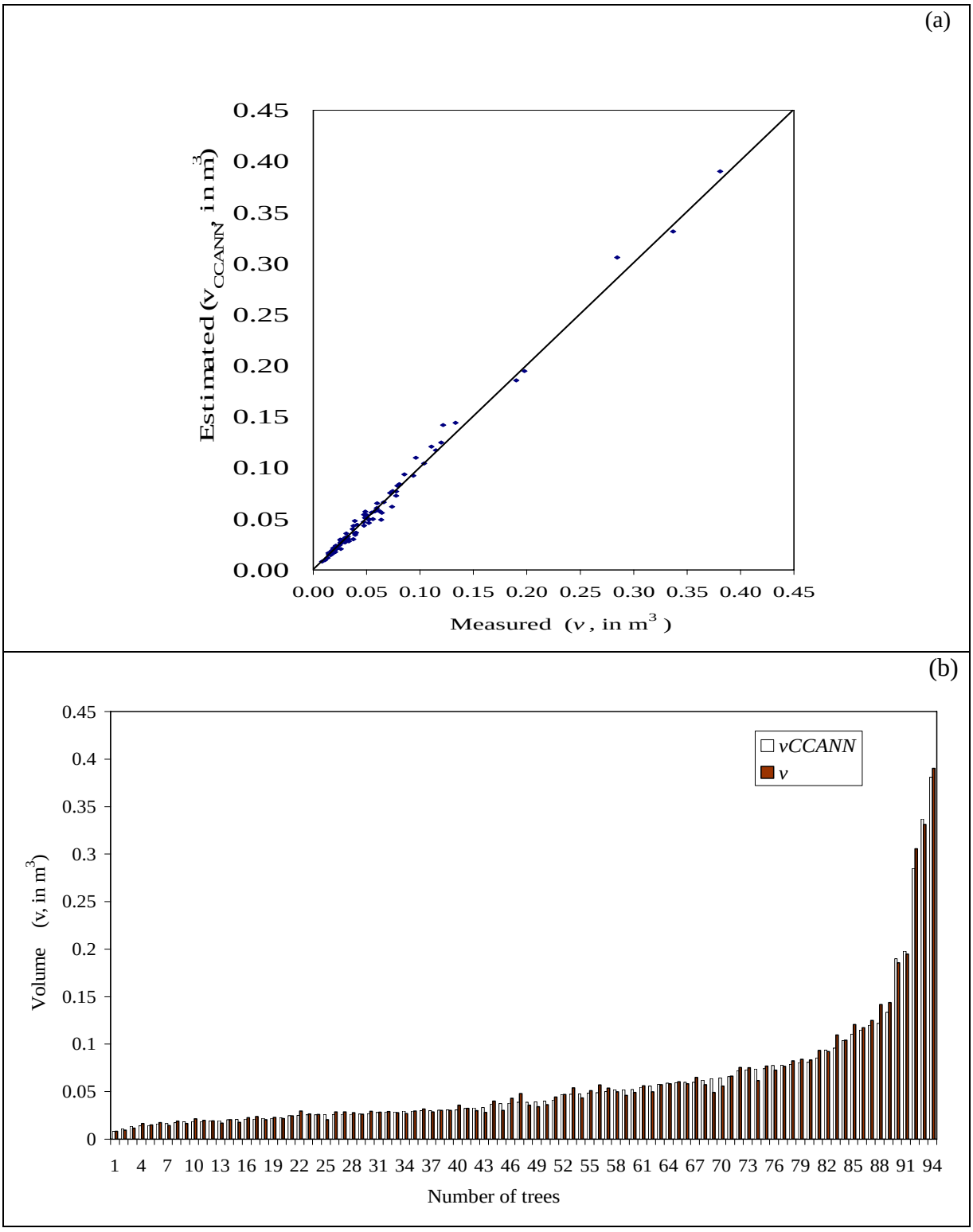

Figure 2. 45-degree line plot (a) of estimated total volume (v_{CCANN}) values vs the corresponding measured values (v) and histogram (b) of estimated and the corresponding measured total volume values (v).

two-sided tabular t-value = 1.986 (a=0.05), for the total volume data set and the percentage estimation error in minimum and maximum total volume estimates was equal to +2.44% and -2.37%, respectively. These imply that there was no significant difference between the measured total volume values and the calculated by Smalian's sectional method using the diameter values estimated by the CCANN models.

4. CONCLUSIONS

An artificial neural network modeling technique has been successfully applied for predicting all diameters at one-meter height intervals above breast height ($d_{2.3}$, $d_{3.3}$, $d_{4.3}$, $d_{5.3}$, $d_{6.3}$, $d_{7.3}$, $d_{8.3}$ and $d_{9.3}$) of pine trees. The method is accurate enough and promise of replacing many standard forestry mensuration procedures.

Neural networks are becoming a popular estimation tool, because of the absence of assumptions that free the modeler from searching on parametric approximating functions that may fit the observed data. Although a model does not have to be specified, the number of hidden nodes and layers and the transfer function to be used still must be determined. The cascade correlation algorithm achieved the successful configuration and training of neural networks. Kalman's and adaptive gradient learning rules were used to modify the artificial neural networks weights. Kalman's learning rule was found to be superior and it was adapted for the modification of the artificial neural networks weights. The networks are trained by adjusting weights between neurons and using the hyperbolic-tangent function for activation. The neural network model that was found to be appropriate for all diameters prediction was a three layer network, that is one input layer composed of three nodes, one output layer composed of one node and between them one hidden layer composed of specified number of nodes, different for each diameter with values of correlation coefficients between 0.9744 to 0.9999 and values of mean square errors between 5.05% to 14.99%.

Based on the results obtained by the comparative study between regression models and the CCANN models, where both, Kalman's and adaptive gradient learning algorithms, are embedded, the ANN modeling technique seems to be a promising alternative to regression analysis.

The system proposed, can be applied in forest inventory to produce an accurate estimate of the total volume. Total volume values (v_{CCANN}) as calculated by Smalian's sectional method using diameter values estimated by the CCANN models, gave results with root mean square error equal to 0.00538 m^3 (9.2%).

5. REFERENCES

Ayala, R., A. Becerra, L. F. Iribarne, A. Bosch and, J. R. Diaz. 1999. GIS System as a Decision Support Tool for Agricultural Planning in Arid Zones of Spain. *Agricultural Engineering International: the CIGR Journal of Scientific Research and Development.* Vol I: 1-11.

Bailey, R. L. 1994. A compatible volume-taper model based on the Schumacher and Hall generalized constant form factor volume equation. *Forest Science* 40(2): 303-313.

Bailey, R. L. and D. Thompson. 1990. How to develop neural-networks applications. *AI Expert* S(6): 38-47.

Blackard, J. and D. Dean. 1999. Comparative accuracies of artificial neural networks and discriminant analysis in predicting forest cover types from cartographic variables. *Computers and Electronics in Agriculture* 24: 131-151.

Brown, R. G. and P. Y. C. Hwang. 1992. *Introduction to Random Signals and Applied Kalman Filtering*. 2nd ed. New York: Wiley.

Cao, Q. V., H. E. Burkhart and T. A. Max. 1980. Evaluation of two methods for cubic-volume prediction of loblolly pine to any merchantable limit. *Forest Science* 26: 71-80.

Charalambous, C. 1992. Conjugate gradient algorithm for efficient training of artificial neural networks. In *Proc. IEE* 139(3): 301-310.

Corne, S. A., S. J. Carver, W. E. Kunin, J. J. Lennon and W. W. S. Van Hees. 2004. Predicting forest attributes in southeast Alaska using artificial neural networks. *Forest Science* 50(2): 259-276.

Cros, M-J., F. Garcia, R. Martin-Clouaire and, J-P. Rellier. 2003. Modeling Management Operations in Agricultural Production Simulators. *Agricultural Engineering International: the CIGR Journal of Scientific Research and Development.* Manuscript IT 02 004. Vol V: 1-11.

Demuth, H. and M. Beale. 2001. *Neural Network Toolbox. For Use with Matlab®*. User's Guide. Ver. 4. pp 1-239.

Deo, M. C. and K. Thirumalaiah. 2000. Real time forecasting using neural networks. In: *Artificial neural networks in hydrology*, eds R. S. Govindaraju and A. Ramachandra Rao, 53-71. Netherlands: Kluwer.

Diamantopoulou, M. 2005a. Artificial neural networks as an alternative tool in pine bark volume estimation. *Computers and Electronics in Agriculture* 48: 235-244.

Diamantopoulou, M. 2005b. Total tree volume estimation using Artificial Neural Network models. In *Proc. XXXI CIOSTA – CIGR V International Congress*, 140–147, Hohenheim, Germany, 19-21 September.

Diamantopoulou, M., D. Papamichail and V. Antonopoulos. 2005. The use of a neural network technique for the prediction of water quality parameters. *Operational Research. An International Journal* 5(1): 115-125.

Dowla, U.F. and L. Rogers. 1995. *Solving Problems in Environmental Engineering and Geosciences with Artificial Neural Networks.* MIT Press.

Draper, N. R. and H. Smith. 1998. *Applied Regression Analysis*. 3rd ed. New York: Wiley.

Fahlman, S. E. and C. Lebiere. 1990. The Cascade Correlation learning architecture. *Advances in Neural Information Processing Systems* 2: 524-532.

Fausett, L. 1994. *Fundamentals of Neural Networks Architectures. Algorithms and Applications.* USA:Prentice Hall.

Grewal, M. S. and A. P. Andrews. 1993. *Kalman filtering: theory and practice.* USA:Prentice Hall.

Grewal, M. S. and A. P. Andrews. 2001. *Kalman filtering. Theory and practice using MATLAB®*, 2nd ed. New York: Wiley.

Gupta, H., K. Hsu and S. Sorooshian. 2000. Effective and efficient modeling for streamflow forecasting. In: *Artificial neural networks in hydrology*, eds R. S. Govindaraju and A. Ramachandra Rao, 7-22. Netherlands: Kluwer.

Gurney, K. 1999. *An Introduction to neural networks.* UK.:Prentice Hall.

Guan, B. T. and G. Gertner. 1991. Modeling red pine tree survival with an artificial neural network. *Forest Science* 37(5): 1429-1440.

Hasenauer, H., D. Merkl and M. Weingartner. 2001. Estimating tree mortality of Norway spruce stands with neural networks. *Advances in Environmental Research* 5(4): 405-414.

Jacob, R. A. 1988. Increased rates of convergence through learning rate adaptation. *Neural Networks* 1: 295-307.

Kalman, R. E. 1960. A New Approach to Linear Filtering and Prediction Problems. Transaction of the ASME. *Journal of Basic Engineering* 35-45.

Leahy, K. 1994. The overfitting problem in perspective. *Al Expert* 9(IO): 35-36.

Leduc, D. J., T. G. Matney, K. L. Belli and V. C. Baldwin. 2001. Predicting Diameter Distributions of Longleaf Pine Plantations: A Comparison between Artificial Neural Networks and other accepted methodologies. *Southern Research Station RS-25:* 1-17.

Liu, C., L. Zhang, C. J. Davis, D. S. Solomon, T. B. Brann and L. E. Caldwell. 2003. Comparison of Neural Networks and Statistical Methods in Classification of Ecological Habitats Using FIA Data. *Forest Science* 49(4): 619-631.

Maier, H. R. and G. C. Dandy. 2000. Neural Networks for the prediction and forecasting of water resources variables: a review of modeling issues and applications. *Environmental Modeling and Software* 15: 101-124.

Norusis, M. J. 2000. *SPSS for Windows*. Chicago: SPSS Inc.

Philip, M. S. 1994. *Measuring Trees and Forests.* 2nd ed. Wallingford, UK: CAB International

Satake, T., O. Sakata, Y. Ohta and, T. Furuya. 2003. Optimal Layout Design for Agricultural Facility Using Simulated Annealing. *Agricultural Engineering International: the CIGR Journal of Scientific Research and Development.* Manuscript IT 02 002. Vol V: 1-11.

Sigrimis, N., Y. Hashimoto, A. Munack and, J. De Baerdemaeker. 1999. Prospects in Agricultural Engineering in the Information Age – Technological Developments for the Producer and the Consumer. *Agricultural Engineering International: the CIGR Journal of Scientific Research and Development.* Vol I: 1-21.

Yuan, H. C., F. L. Xiong and X. Y. Huai. 2003. A method for estimating the number of hidden neurons in feed-forward neural networks based on information entropy. *Computers and Electronics in Agriculture* 40: 57-64.

Zhang, Q. B., R. I. Hebda, Q. J. Zhang and R. I. Alfaro. 2000. Tree-Ring Growth Responces to Climatic Variables Using Artificial Neural Networks. *Forest Science* 46(2): 229-239.