Ann. For. Res. 68(1): 39-54, 2025
https://doi.org/10.15287/afr.2025.3433

ANNALS OF FOREST RESEARCH
www.afrjournal.org

Tree stem mean diameter reduction factor prediction through advanced modeling approaches

Maria J. Diamantopoulou

Diamantopoulou M.J., 2025. Tree stem mean diameter reduction factor prediction through advanced modeling approaches. Ann. For. Res. 68(1): 39-54.

Abstract Sustainable management of natural resources relies on accurate modelling of forest attributes to prevent degradation. This study explores advanced modelling techniques, including Artificial Neural Networks (ANN) and Support Vector Regression (SVR), for estimating the mean stem diameter reduction factor (taper) of standing fir trees (*Abies x borisii-regis* Matff.). These methods are compared against traditional non-linear regression model (NLR), developed using the Levenberg-Marquardt optimization algorithm. The ANN models employ cascade correlation, generalized regression, and Bayesian regularization back-propagation architectures, while the ε-SVR approach is assessed for its robustness. The results show that support vector regression (ε-SVR) achieved the lowest relative errors in model fitting, improving by 0.60% over cascade correlation and generalized regression and by 0.67% over Bayesian regularization. Regarding generalization ability, the ε-SVR model performed best, with a relative error of 4.90%, which was slightly lower than cascade correlation (by 0.1%), generalized regression (by 0.01%), and Bayesian regularization (by 0.04%). A comparative analysis between machine learning approaches and standard regression revealed that the ε-SVR model had the lowest mean error (0.0715), while the non-linear regression (NLR) model showed a higher mean error of 0.0955, which means 1.35 times greater. These findings highlight the strong capability of machine learning methods in accurately estimating and predicting the diameter reduction factor of trees, effectively capturing its non-linear behaviour compared to traditional regression models. Overall, this study underscores the potential of advanced machine learning techniques to enhance accuracy and adaptability in sustainable forest management.

Keywords: mean diameter reduction factor, cascade correlation, generalized regression, Bayesian regularization, support vector regression.

Addresses: School of Forestry and Natural Environment, Aristotle University of Thessaloniki, Greece.

✉ **Corresponding Author:** Maria J. Diamantopoulou (mdiamant@for.auth.gr).

Manuscript: received January 3, 2024; revised April 28, 2025; accepted May 26, 2025.

Introduction

A taper equation operates as a mathematical representation of the shape of a tree, allowing for the estimation of the volume of any height of the tree stem. This includes the total volume of the stem, and it can be applied with dimension limits of either height or diameter. The outer shape of a tree typically exhibits a clear curvature near its base, follows a linear track along the central section of the stem, and finally, shows variability along the upper stem (Shaw et al. 2003). Given

the challenges faced by the irregular shape of tree boles, there is a growing interest in achieving precise estimation and prediction of their volume and biomass production. The configuration of a tree bole resulted from the rotation of a generator curve, which is outwardly curved in the lower part of the tree stem and concave inward toward the upper section. Moreover, the shape of the tree stem falls between a cone and a parabolic form. Given the economic significance of accurate measurements in financial wood products, such as tree stem volume or its biomass, forest research is focused on exploring factors that could improve the precision of estimation in these products.

To this direction, one significant factor is the tree stem mean diameter reduction factor, known as the taper factor. This factor represents the average by which the diameter of a tree bole decreases per meter of the tree's height. Understanding the configuration of the values of the mean diameter reduction factor in a standing trees which can be derived from a relative model, is of utmost importance. This knowledge significantly contributes to the precise estimation of the tree stem volume, in the sense that the stem volume calculation particularly relying on the availability of known diameters at various heights. Additionally, this knowledge contributes to characterizing the structure of forest clusters (West 2015). At the same time, it offers essential information for the effective management of forested areas.

For several decades, despite its limitations, regression modelling (Draper & Smith 1998) has been extensively employed in forest research to address challenges in both research and practical applications. In the past two-three decades, there has been an exploration of the effectiveness of artificial neural networks (ANNs) (Patterson 1996, Swingler 1996, Gurney 1997) and support vector machines (Vapnik 1999, 2000, Montesinos López et al. 2022) in modelling regression tasks (ε-SVR), particularly within the realm of environmental modelling (Wang et al. 2009, Thomas et al. 2017). ANNs have gained significant attention primarily because the methodology bears similarities to statistical modelling. ANN models can be viewed as a complementary effort, avoiding the restrictive assumption of a specific statistical model, or as an alternative approach for fitting non-linear data. On the other hand, support vector regression (ε-SVR) algorithm is designed to address regression-type problems and has demonstrated the capability to mitigate both overfitting and local minima issues by minimizing the bound of generalization error, rather than focusing solely on the training error. These advantages position ε-SVR as a significant alternative algorithm in contrast to the commonly adopted non-linear regression (NLR) approach when constructing prediction models for trees attributes.

Given their advantages over NLR modelling, both ANN and ε-SVR have been applied in forestry modelling research. The aim is to construct accurate and reliable models with minimal reliance on ground-truth data. These approaches not only ensure the development of robust models but also contribute to saving time and effort in the field. Specifically, different artificial neural network algorithms have been tested for estimating forest and trees attributes. Diamantopoulou (2005) employed the cascade correlation algorithm in feedforward artificial neural networks (ANNs) to estimate tree stem diameters. In contrast, Özçelik et al. (2017) utilized the Levenberg-Marquardt algorithm for predicting tree biomass, Özçelik et al. (2019) tested the same neural network algorithm for pine stem diameter prediction, and Vieira et al. (2018) employed the Levenberg-Marquardt algorithm to estimate tree growth and height. Zhou et al. (2019) effectively estimated diameter at breast height using general regression neural networks. Furthermore, Ercanlı (2020) utilized deep learning to model the relationship between tree height and diameter at breast height, Sandoval & Acuña (2022) employed the Adam's optimization algorithm in back-propagation ANNs for stem taper estimations, while Heidarlou et al. (2023) explored forest cover changes in Iranian Zagros forests using ANNs for classification tasks. Owing to its capacity to effectively learn from data with noise by minimizing the generalization error, the support vector regression approach has been utilized in forest modelling, as well. Monnet et al. (2011) showcased the capacity of ε-SVR

models in predicting forest stand parameters like dominant height, basal area, mean diameter, and stem density. Jiao et al. (2013) employed support vector regression to estimate the volume of live tree timber, Diamantopoulou et al. (2018) utilized ε-SVR models for the prediction of tree-bark volume, Malek et al. (2019) estimated using ε-SVR models the stem diameter and biomass at the individual tree crown level, Hirigoyen et al. (2021) modelled leaf area index in Eucalyptus plantations using various machine learning methods, while Moradi et al. (2022) applied the support vector machines methodology to predict above-ground biomass of a oak forest.

To this direction, this study focuses on constructing accurate and reliable models for estimating the stem mean diameter reduction factor (taper factor), employing nonlinear regression, artificial neural network, and support vector machine learning techniques. Ultimately, considering the advantages and drawbacks of each modeling approach, the performance of each model was assessed and evaluated.

Materials and Methods

Ground-truth data

The primary data used in this study originated from the Pertouli University Forest in Greece, consisting of a nearly homogeneous stand of fir trees (*Abies x borisii-regis* Mattf.). A systematic sampling approach was employed, measuring a total of 728 fir trees to encompass variations across different sites. Nevertheless, upon conducting exploratory data analysis (Hoaglin et al. 2006, Myatt & Johnson 2014), it was observed that certain extreme data values were present, introducing significant noise into the initial dataset. For this reason, this study utilized a refined dataset, comprising measurements from 677 cleaned fir trees, to mitigate the impact of outliers and enhance the reliability of the analysis.

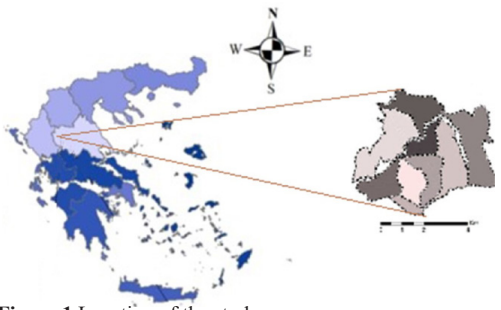

Figure 1 Location of the study area.

The dataset comprises variables of the following included measurements: a) of stump and diameter at breast height (0.3m and 1.3m from the ground, denoted as $d_{0.3}$, $d_{1.3}$, respectively) using a caliper, b) of the base diameter of the one meter top of the tree (denoted as d_{top}) using a Speigel Relaskop, and c) of total height (denoted as h_{total}) of the sampled trees using the Blume-Leiss hypsometer.

Data base construction

After the completion of the fieldwork, the measured variables which formed the initial database were the $d_{0.3}$, $d_{1.3}$, d_{top} and h_{total}. In order for the mean tree stem mean diameter reduction factor (*DRF*) to be calculated, the distance (L) between the stump diameter ($d_{0.3}$) and the base top diameter (d_{top}) for each tree was calculated, as well, using the formula (Matis 2004):

$$DRF = \frac{d_{0.3} - d_{top}}{L} \quad (1)$$

To estimate the stem mean diameter reduction factor (*DRF*), the diameter at breast height and total tree height were used as input (independent) variables. These variables were chosen because they have a strong correlation with the *DRF* and are among the most commonly measured tree attributes. Additionally, the diameter at breast height is the easiest to obtain due to its accessible position on the tree.

Models construction

In the construction of a machine learning model, having both training and testing datasets is crucial to check overfitting. This approach aims to capture not only general patterns but also sampling variations. In this regard, the available dataset was randomly divided into two distinct parts: a) the fitting

data sample, constituting 90% of the total data, that is the measurements from 609 trees and b) the testing sample, comprising the remaining 10% of the data, that is the measurements from the rest 68 trees. In order to facilitate the repeated validation of the constructed models, the fitting dataset was partitioned into training and validation subsets utilizing the k-fold cross-validation technique, where k was configured to 10 (Russell & Norvig 2020). The testing sample was employed to evaluate the predictive performance of the constructed models. It's important to highlight that the test dataset was kept separate from the fitting dataset to ensure there was no data leakage between them.

The developed models employed the tree bole mean diameter reduction factor (*DRF*) as the dependent variable, with the diameter at breast height ($d_{1.3}$) and the total height (h_{total}) of each tree serving as independent variables.

To employ regression analysis, the normality of the dependent variable (*DRF*) was assessed (Zuur et al. 2010) using Kolmogorov-Smirnov statistical test under the Lilliefors' significance correction (Lilliefors 1967) and a normal probability plot (IBM-SPSS 2021). To address the observed non-normality of the dependent variable, a combination of techniques was employed, which included bootstrapping (Efron & Tibshirani 1994) to estimate the standard errors of non-linear model parameters and the sequential quadratic optimization (Nocedal & Wright 2006) algorithm. The above techniques allowed the reliable estimation of both the bootstrapped 95% confidence intervals for the non-linear regression parameters and their 95% trimmed ranges, as well. Finally, the required effective initial values for the parameters estimation of the nonlinear models were acquired through the methods outlined in Draper & Smith (1998).

According to the artificial neural network modeling, three architectures were used, namely the cascade (CCANN) where the Levenberg-Marquardt algorithm was embedded, the feedforward generalized regression (GRNN) and the back-propagation Bayesian regularization (BRNN) artificial neural networks. All selected architectures and algorithms have shown their potential to regression type problems in forestry research (Diamantopoulou 2005, Özçelik et al. 2017, Özçelik et al. 2019, Vieira et al. 2018, Zhou et al. 2019). Therefore, applying these algorithms to estimate the stem mean diameter reduction factor (*DRF*) would be challenging. However, what makes this study particularly interesting is the comparative evaluation of their effectiveness in accurately estimating this important taper factor.

The cascade correlation algorithm, as described by Fahlman & Lebiere (1990) generates the cascade correlation Artificial Neural Network (ANN) of the feed-forward type networks. This supervised algorithm operates within multilayer feed-forward ANNs, utilizing the Levenberg-Marquardt algorithm (Levenberg 1944, Marquardt 1963) to adjust the ANN weights. This type of network possesses the capability to approximate any continuous function. The cascade correlation algorithm commences training without any hidden nodes. The introduction of each hidden node in the hidden layer occurs when the error between the network's realized output and the target is not sufficiently small. Each new added node establishes connections with all other nodes except the output nodes. The determination of the optimal number of hidden nodes involves a trial-and-error approach. This method begins with no hidden nodes, proceeds to train the network, and iteratively repeats the process by incrementally adding nodes until no further enhancement in network performance is observed. Owing to its dynamic expansion, persisting until the problem is effectively learned, the cascade correlation algorithm naturally emerges as a suitable solution for a given problem. This iterative procedure continues until the correlation between the hidden node's output and the network's residual

error is maximized (Equation 2):

$$S = \sum_{O} \left| \sum_{P} (v_p - \bar{v}) \cdot (e_{p,o} - \overline{e_O}) \right| \quad (2)$$

where o is the o_{th} node; p is the p_{th} training pattern; v_p is the value of the candidate node at pth training pattern; $\bar{v}$ is the mean of v; $e_{p,o}$ is the residual error observed at node o at p_{th} training pattern; $\overline{e_O}$ is the mean of $e_{p,o}$.

During each training step, a new hidden neuron is introduced, and its weights are fine-tuned to maximize the magnitude of the correlation (Equation 2) and then this connection is frozen. Subsequently, each hidden neuron undergoes training only once, after which its weights are fixed (Figure 2).

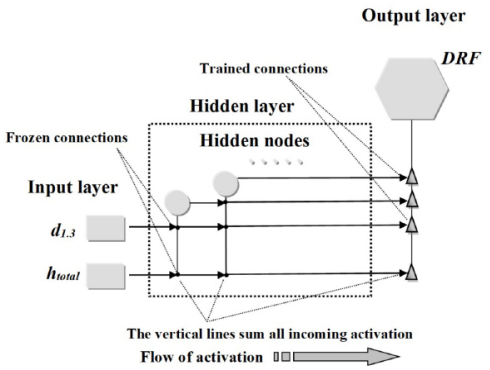

Figure 2 Cascade correlation neural network architecture.

Generalized regression neural networks (GRNN) which are known as Bayesian networks were formulated by Speckt (1991), transforming a statistical method of function approximation into the structure of a neural network. The regression network employs Bayesian techniques to predict the anticipated mean value of the output, given an input case, as illustrated below:

$$E[z/\bar{x}] = \frac{\int_{-\infty}^{+\infty} z f(x,z) dz}{\int_{-\infty}^{+\infty} f(x,z) dz} \quad (3)$$

where z is the output value, which is being estimated; x is the input case; $f()$ is the joint probability density function of the inputs and outputs.

Since the joint probability density function is unknown, it is approximated by the sum of Gaussian kernel functions. The GRNN architecture includes four layers (Figure 3).

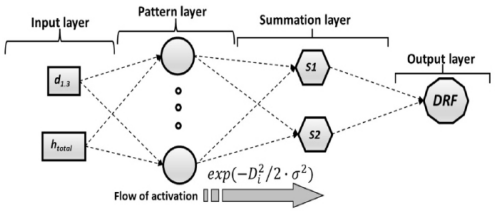

Figure 3 Generalized regression neural network architecture.

In each node of the pattern layer, the quantity of $exp\,(-D_i^2/2 \cdot \sigma^2)$ is the derived signal which is passed to the next layer (summation layer), where D_i is the distance between the training value and the point of prediction, which represents the degree of the adaptation of the predictions by the neural net to the actual training values and σ is the smoothing coefficient. That is the normal distribution is centred at each training value that fitted to the system. The signals reach at the summation layer are weighted with the estimated values of the training samples; the summation of these weighted signals represent the information of the first node (Figure 3 - S1) of the summation layer which is divided by the summation of the unweighted signals (Figure 3 - S2) reach to the second summation node. The estimated z value was derived as:

$$z = \frac{\sum_{i=1}^{n} \left(exp\left(\frac{-D_i^2}{2 \cdot \sigma^2} \right) \cdot z_i \right)}{\sum_{i=1}^{n} exp\left(\frac{-D_i^2}{2 \cdot \sigma^2} \right)} \quad (4)$$

where z_i are the observed values of the output variable for $i=1,..., nf$

From the above description, it is obvious that to control the magnitude of Di deviations, and finally the best GRNN model construction, the optimal smoothing coefficient (σ) value has to be selected. This optimal value was identified through an exhaustive grid-search approach (Belete & Huchaiah 2022) under a trial-and-error procedure, over the range [0, 10] with a step size of 0.001.

Bayesian Regularization Neural Networks

(*BRNN*) (Figure 4) have gained popularity for their robustness in contrast to multilayer perceptron backpropagation networks. Moreover, they can reduce the necessity for extensive cross-validation (Burden & Winkler 2008).

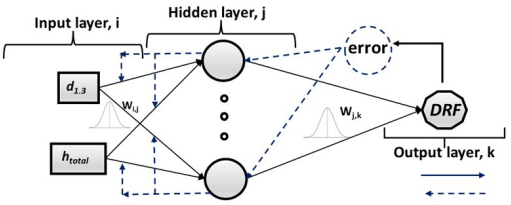

Figure 4 Bayesian regularization neural network architecture.

The Bayesian learning of this algorithm which lies on the Bayes' theorem, it has been extensively described by Titterington (2004), Kayri (2016), Sariev & Germano (2020). Indicatively, unlike conventional methods that seek point estimates for the parameters of the neural network, Bayesian neural networks treat the weight coefficients as random variables characterized by probability distributions known as prior distributions. That is, the Bayesian approach, relying on the probability distribution of the network weights, utilizes the Bayesian theorem to derive the probability distribution of the network predictions. During the training process, the mean square network error, incorporating the Bayesian regularization term, is minimized (Kayri 2016). Within the Levenberg-Marquardt algorithm, which is embedded to neural network training to minimize the objective function of the system, Bayesian regularization is integrated in order for any possible overfitting of the network to be prevented. This involves computing the Jacobian matrix (*J*), leading to the Hessian matrix (H_m) of the objective function. Jacobian matrix includes the first derivatives of the network errors with respect to the weights considered as random variables and biases. Ultimately, the training process continues until it converges while approaching the optimal number of nodes in the hidden layer.

As a subset of Support Vector Machines exploring the application of Support Vector Regression (*ε*-SVR) methodology (Vapnik 1999) holds promise in forest modelling. *ε*-SVR algorithm introduced and described in detail by Vapnik et al. (1997). Summarizing it can be stated that *ε*-SVR involves creating an initial space of (2*ε*) width, where $\varepsilon > 0$, encompassing the original data within the range *[-ε, +ε]*. By using an additional variable *ξi*, known as slack variable, the system tries to minimize the function (Vapnik 2000, Smola & Schölkopf 2004):

$$min\frac{1}{2}\|w\|^2 + C\cdot\sum_{i=1}^{n}(\xi_i^+ + \xi_i^-) \text{ subject to}$$
$$\begin{cases} y_i - w^T\varphi(x_i) - bc \le \varepsilon + \xi_i^+ \\ y_i - w^T\varphi(x_i) - bc \le -\varepsilon - \xi_i^- \\ \xi_i^+, \xi_i^- \ge 0,\ i = 1, \dots, n \end{cases} \quad (5)$$

where C represents a constant that governs the impact of each individual support vector, balancing model smoothness by considering the trade-off between prediction errors and model simplicity, *w* is the vector of the weights and *bc* is the system's bias.

The radial basis function (RBF) kernel:

$$K(x_i - x_j) = \exp\left(-\gamma\|x_i - x_j\|^2\right), \gamma > 0 \quad (6)$$

where, $\gamma = \left(1/_{2\sigma^2}\right)$ and $\|x_i - x_j\|$ and is the Euclidean distance between the support vectors (SV).

Euclidean distance is employed to transform the data into an m-dimensional super-space enabling the representation of complex nonlinear relationships through an optimal straight line (Williams 2011). Considering Equation (5) and (6), it can be concluded that the accuracy of estimation and the complexity of *ε*-SVR models are contingent upon three meta-parameters: (*ε*) which is governing the width of the *ε*-insensitive zone, gamma (*γ*) which is the free parameter of Gaussian radial basis function (*RBF*) kernels, and the cost parameter (*C*), which modulates the impact of

each support vector, striking a balance between mis-prediction and model simplicity.The optimal combination of these hyperparameters was selected through exhaustive grid-search approach (Belete & Huchaiah 2022), with values ranged a) for (ε) from 0.01 to 0.5 by 0.01, b) for (γ) from 0.01 to 1 by 0.01 and c) for (C) from 1 to 30 by 0.10.

The non-linear regression modelling was performed using the SPSS software (IBM-SPSS 2022), the cascade correlation and the Bayesian regularization neural networks modelling was accomplished with Matlab programming language and platform (Matlab 2022) both provided by Aristotle University of Thessaloniki, Greece, while the generalized regression and the ε-SVR modelling was implemented using the "Python" programming language (Python Software Foundation 2022).

Modeling evaluation metrics

The assessment of the reliability and accuracy of models constructed through both regression and machine learning approaches employed the following evaluation metrics:

$$AAE = \frac{\sum_{i=1}^{n}|y_i - \hat{y}_i|}{n} \quad (7)$$

$$RMSE = \sqrt{\left(\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}\right)} \quad (8)$$

$$R = \frac{\sum_{i=1}^{n}(\hat{y}_i \cdot y_i) - \frac{(\sum_{i=1}^{n}\hat{y}_i)\cdot(\sum_{i=1}^{n}y_i)}{n}}{\sqrt{\sum_{i=1}^{n}\hat{y}_i^2 - \frac{(\sum_{i=1}^{n}\hat{y}_i)^2}{n}}\cdot\sqrt{\sum_{i=1}^{n}y_i^2 - \frac{(\sum_{i=1}^{n}y_i)^2}{n}}} \quad (9)$$

$$RE\% = \left(\frac{\sqrt{\left(\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}\right)}}{\frac{(\sum_{i=1}^{n}y_i)}{n}}\right)\cdot 100 \quad (10)$$

where AAE is the absolute average error between observed (y_i) and estimated $\hat{y}_i$ by the model values; $RMSE$ is the root mean square error; R is the correlation coefficient; $RE\%$ is the relative efficiency of the model estimations.

Each of these metrics provides insight into model performance from different points of view. AEE indicates the typical size of errors, RMSE offers a more realistic measure of error size, and RE% enables direct comparison across models. Lower values for these metrics indicate better model performance. Conversely, the correlation coefficient (R) assesses the consistency of model predictions, with higher values denoting stronger agreement between estimated and observed values.

Results

The descriptive statistics such as the minimum (Min.) and maximum (Max.) values, the arithmetic mean (Mean) and standard deviation (Sd) of the mean diameter reduction factor (DRF), Sthe breast height diameter ($d_{1.3}$) and the total tree height (h_{total}), for the fitting and the test data sets are given in Table 1.

Given the non-linear behaviour observed in the stem mean diameter reduction factor (Figure 5), we employed a non-linear regression modelling approach (IBM-SPSS 2021).

Among the various linear and nonlinear regression models examined utilizing

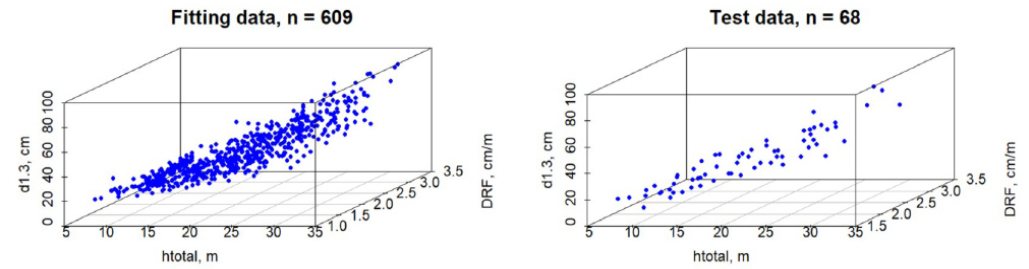

Figure 5 Scatterplots of the diameter reduction factor values (*DRF*) against diameter at breast height ($d_{1.3}$) and total height (h) values, for the fitting and the test data sets.

Table 1 Descriptive statistics of the mean diameter reduction factor (*DRF*), breast height diameter ($d_{1.3}$) and the total tree height (h_{total}), for the fitting and the test data sets.

Variable	Fitting data set, n_f = 609				Test data set, n_t = 68			
	Min.	Max.	Mean	S_d	Min.	Max.	Mean	S_d
DRF, cm/m	1.41	3.48	2.20	0.45	1.48	3.44	2.14	0.47
$d_{1.3}$, cm	12.00	95.50	39.60	17.26	13.00	87.00	38.62	18.05
h_{total}, m	6.00	35.00	19.55	7.07	7.00	34.00	19.23	7.45

bootstrapping and the sequential quadratic optimization algorithm, the form of the non-linear equation given below demonstrated the most effective fit to the available data:

$$a = 1 + e^{\left(b_0 - \left(\frac{b_1}{\sqrt{d_{1.3}}} + \frac{b_2}{d_{1.3}} + \frac{b_3}{h_{total}^2} + \frac{b_4}{d_{1.3}^2} + \frac{b_5}{d_{1.3}^3} + \frac{b_6}{\sqrt{h_{total}}}\right)\right)} \quad (11)$$

where *bi* are the asymptotic non-linear regression coefficients.

The sum of squared residuals was utilized as the loss function. As previously mentioned, a combination of methods was employed to address the observed non-normality of the dependent variable. The approach used aimed to ensure the reliable estimation of both the asymptotic 95% confidence intervals for the non-linear regression parameters and their 95% trimmed range. The regression coefficient values of Equation (11), their standard errors, the 95% asymptotic confidence intervals as estimated using the bootstrap technique along with their 95% trimmed range, are given in Table 2. All regression coefficients were found to be significantly different from zero, for significance level of a= 0.05 (Table 2).

The cascade correlation model that was chosen as the best option had a hidden layer with four input nodes. This specific number of hidden nodes yielded the optimal performance for both the fitting and test datasets (Figure 6a). Within the context of the generalized regression neural network model, the most notable adaptation to both datasets was observed in the model with an optimal smoothing coefficient (σ) of 1.4 (Figure 6b). As depicted in Figure 6c, the most effective Bayesian Regularization neural network model featured eight input nodes in the hidden layer. In the case of the chosen best support vector regression model, the ε-SVR model, which configured with hyperparameter values ε = 0.01, γ = 0.09, and C = 8.1, demonstrated the highest level of accuracy and reliability in adapting to both the fitting and test datasets (Figure 6d).

The machine learning models documented in Table 3 demonstrated the most effective alignment with the ground-truth data. They showcased robust generalization capabilities, as demonstrated by their prediction root mean

Table 2 Regression coefficient values, bootstrap standard errors, 95% bootstrap confidence intervals, and 95% trimmed ranges of Equation (11).

Coef.	Value	Bootstrap standard error	95% bootstrap conf. inter.		95% trimmed range	
			Lower limit	Upper limit	Lower limit	Upper limit
b_0	-7.51	0.658	-8.806	-6.214	-8.570	-6.140
b_1	-109.45	10.245	-129.597	-89.262	-124.460	-91.054
b_2	743.24	47.247	650.184	836.304	661.817	808.755
b_3	129.23	13.023	103.592	154.861	103.631	154.068
b_4	-7689.47	352.757	-8382.874	-6994.066	-7976.535	-7301.107
b_5	36405.06	1613.645	33228.598	39581.528	36330.275	36477.610
b_6	-22.33	1.285	-22.859	-19.802	-25.286	-20.010

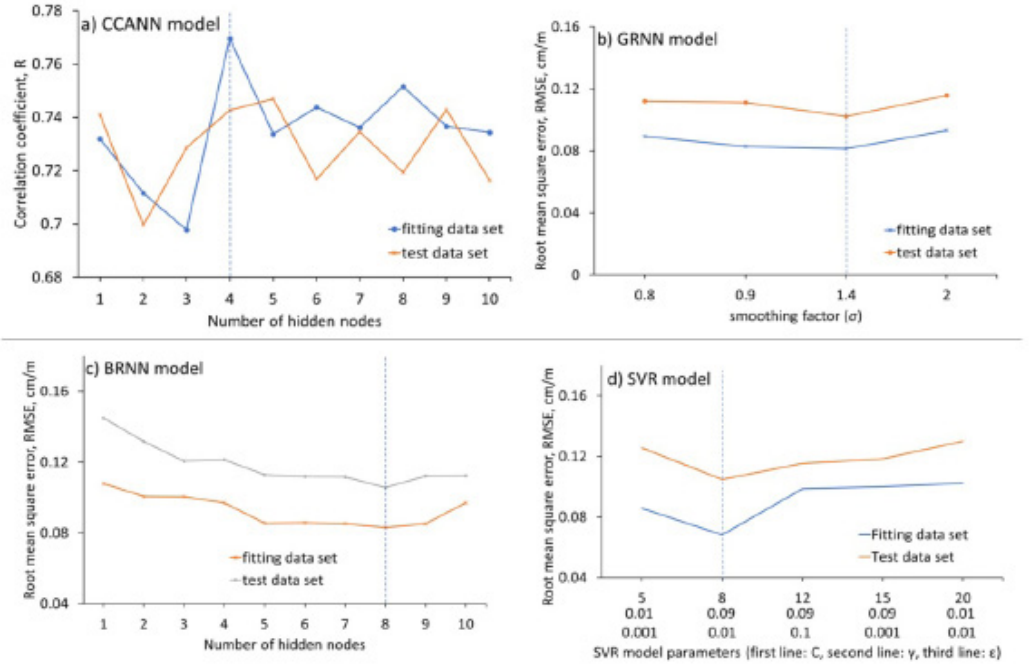

Figure 6 Parameter tuning for machine learning models, specifically: a) CCANN, b) GRNN, c) BRNN, and d) ε-SVR.

square errors, correlation coefficient and relative efficiency (%) when applied to entirely new and previously unobserved test datasets, which were not part of their construction phase.

Analysing the outcomes presented in Table 3, the relative efficiency of the machine learning models varied between 3.11% and 3.78% for the fitting dataset and between 4.90% and 5.00% for the test dataset. Notably, the ε-SVR model demonstrated the most favourable alignment with the data. Among the neural network models, it was the generalized regression architecture that yielded the most precise results, for both the fitting and the test data sets.

In order for a comparison along with an evaluation of the performance to be conducted between the non-linear regression modelling approach (NLR) against the machine learning

Table 3 Absolute average error, root mean square error, correlation coefficient and relative efficiency (%) of the mean diameter reduction factor for all machine learning constructed models, for the fitting and the test data sets.

Model/ structure*	Data set	*n*	*AAE*	*RMSE*	*R*	*RE%*
CCANN/	fitting	609	0.2158	0.0816	0.7618	3.71
2-4-1	test	68	0.2328	0.1070	0.7352	5.00
GRNN/	fitting	609	0.2136	0.0814	0.7736	3.71
2-609-2-1	test	68	0.2348	0.1022	0.7676	4.91
BRNN/	fitting	609	0.2156	0.0830	0.7643	3.78
2-8-1	test	68	0.2253	0.1056	0.7389	4.94
ε-SVR	fitting	609	0.1794	0.0683	0.8122	3.11
	test	68	0.2371	0.1048	0.7496	4.90

Note: * the notation 2-4-1 means a network with one input layer with 2 input nodes, one hidden layer with 4 input nodes and one output layer with one output node.

selected as the best fitted model, the results of Table 4 are provided below.

As shown in Table 4, the relative efficiency of all models in estimating the mean diameter reduction factor ranged from 3.26% for the ε-SVR model to 4.36% for the NLR model. Additionally, the root mean square error (*RMSE*) values for the NLR, CCANN, GRNN, and BRNN models were 1.337, 1.172, 1.169, and 1.187 times higher, respectively, compared to the *RMSE* obtained when applying the constructed ε-SVR model to the available dataset.

Table 4 Metrics of the mean diameter reduction factor for all constructed models, for all the available data set (n = 677).

	Models				
Metrics	**NLR**	**CCANN**	**GRNN**	**BRNN**	**ε-SVR**
AAE	0.2270	0.2262	0.2157	0.2166	0.1852
RMSE	0.0955	0.0837	0.0835	0.0849	0.0715
R	0.7293	0.7584	0.7724	0.7612	0.8049
RE%	4.36	3.82	3.81	3.87	3.26

The first row of Figure 7 (a1, b1, c1, d1, and e1) presents the comparisons between the mean diameter reduction factor estimates produced by all constructed models and the observed values. These models exhibited a generally similar trend, with their estimations closely aligned to the 45-degree line, albeit with a few points deviating noticeably from it. Notably, the ε-SVR model stands out as it outperformed others, with a substantial number of points falling directly on or very close to the 45-degree line, highlighting its superior performance. We can observe similar patterns from the residual histograms presented in Figure 7 (a2, b2, c2, d2, and e2) generated by all the constructed models. In these histograms, there is a peak in the distribution of residuals around zero, with the frequency decreasing as the residual values increase. This pattern suggests that the models are well-tuned, and their parameters are appropriately set, leading to a healthy and acceptable model performance. To gain insights into the misestimations at various breast height diameter levels, dot plots illustrating the residuals for each diameter level of the trees (Figure 7 - a3, b3, c3, d3, and e3) were generated. These plots reveal that the residuals are distributed within a band of approximately ±1 cm/m for each diameter level. This observation suggests that all models are able of producing mean *DRF* estimates with satisfactory accuracy for diameters ranging from 12 cm to approximately 100 cm.

Discussion

Understanding the mean diameter reduction factor of a tree stem is valuable as it offers insights into the taper of the tree stem. This knowledge enables precise estimation of stem diameters from the ground up to the tree's top, leading to reliable estimations of tree bole volume, through sectional volume calculation methods. This, in turn, serves as the initial step in accurately determining the weight of wood and finally the amount of carbon stored by a tree. Nevertheless, the challenge of identifying and measuring tree stem diameters outside the bark layer at various heights several meters above the ground is widely recognized. Consequently, the research for suitable estimation models for biologically challenging variables that are hard to measure has become a central point of extensive research within the field of forest science.

The primary objective of this study is to focus on employing advanced machine learning modelling techniques, specifically Artificial Neural Network modelling methods and the support vector machine technique. These methods are utilized to estimate the mean reduction factor (taper) of the stem in standing fir trees (*Abies x borisii-regis* Matff.) with the highest possible degree of accuracy. The machine learning approaches utilized considered as alternatives to traditional non-linear regression modelling. The development of non-linear regression models

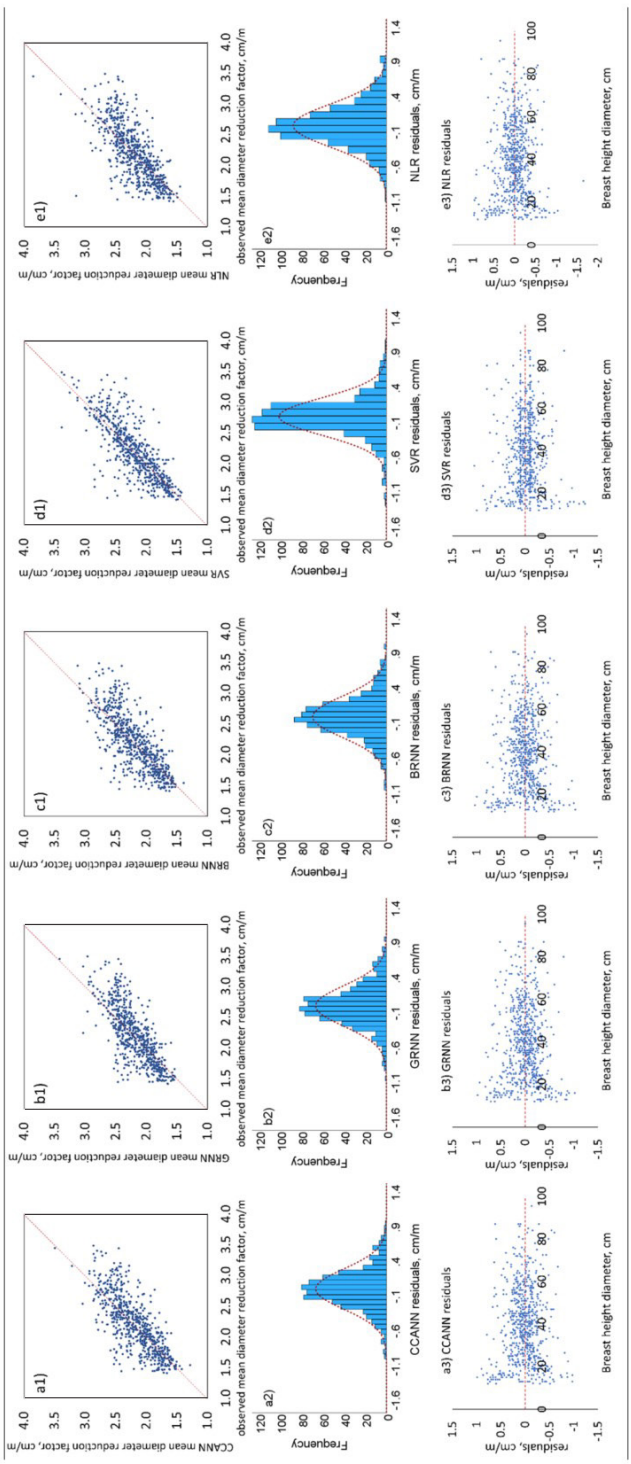

Figure 7 Comparison of mean diameter reduction factor estimates from the constructed models (a1, b1, c1, d1, e1) with observed values, residual histograms (a2, b2, c2, d2, e2) generated by the models, and dot plots illustrating model residuals across different levels of tree diameter at breast height (a3, b3, c3, d3, e3).

in this research involved the application of the Levenberg-Marquardt optimization algorithm. In addition, the construction of neural network models contained the utilization of various architectures, including cascade correlation, generalized regression, and Bayesian regularization back-propagation. Furthermore, the study also made use of the support vector regression approach, renowned for its exceptional properties (Vapnik 1999, 2000; Smola & Schölkopf 2004).

The need for implementing machine learning approaches for the accurate and reliable estimations of the mean stem diameter reduction arises from its capacity to address challenges met in forest data which are difficult to confront, including non-linear relationships, non-Gaussian distributions, and data noise. Machine learning approaches hold significant position as viable alternative to conventional regression models in various forest modelling applications. Additionally, it's essential to identify the precise form of the regression equation that can effectively represent the ground-truth forestry data, which is a challenging task, because it adds complexity and time to the application of regression methodology, and this aspect should not be disregarded. Moreover, achieving convergence in nonlinear regression models depends on having well-considered initial parameter values which requires additional modelling effort (Diamantopoulou 2022). In contrast, intelligent systems like the machine learning modelling approaches show the ability to handle ground-truth data without any prerequisites or specification of an equation form capable of accurately representing the data in hand. The system is able of independently generate the most suitable model, provided that its parameters' values were appropriately optimized. However, if the system is not handled correctly, it might become stuck in local minima or it might lead to over/under fitting (Kalkanlı Genç et al. 2023).

Because of its non-linear characteristics, modelling the mean stem diameter reduction factor presents a significant challenge. To the best of the authors' knowledge, research in this domain is limited, and artificial intelligence modelling methods have not be employed for this specific purpose. Our decision to choose among the cascade correlation, generalized regression, and Bayesian regularization back-propagation architectures was primarily motivated by the exploration of various architectures under different algorithms, with the goal of identifying the most flexible process. Additionally, we selected algorithms specifically designed to address regression-type problems. In the end, the support vector regression modelling approach, which relies on non-linear kernel functions (Boser et al. 1992), was chosen because of its ability to capture all patterns present in real-world data, including subtle ones. Furthermore, this approach has proved to be effective at overcoming local minima and handling outliers, as well.

The selected machine learning modelling approaches for the accurate and reliable estimation of the mean *DRF* offer several advantages. The cascade correlation model reduces the tuning effort by dynamically adding hidden nodes to the hidden layer as needed. By this way, the cascade architecture can effectively handle complex relationships (Diamantopoulou 2005). Generalized Regression, in contrast, requires only a single pass through the data. The optimization of the smoothing coefficient (σ), making GRNN an efficient and easy to apply modelling approach (Speckt 1991). According to the third approach, the Bayesian Regularization method (Sariev & Germano 2020), due to its Bayesian nature, provides robustness for noisy biological datasets, such as primary field measurements in forestry, ensuring smooth function approximation. Lastly, support vector regression (ε-SVR) has demonstrated its ability to balance model complexity and prediction accuracy (Vapnik 1999) while offering robustness to outliers, an issue frequently met in forestry data.

However, all machine learning modelling approaches require a deep understanding of the best optimization practices, particularly regarding hyperparameter tuning, as their successful training depends heavily, though not exclusively, on this process.

Based on the analysis outcomes, all examined models are acceptable options for estimating the mean diameter reduction factor. Nevertheless, it's essential to recognize the performance superiority of the ε-SVR modelling approach. The conventional regression method demanded substantial effort in constructing the final model, involving tasks such as managing prerequisites, determining the appropriate model structure, and specifying initial parameter values, especially in non-linear cases. In contrast, the non-parametric machine learning approaches did not require the handling of assumptions but did require effective management of their parameters. Specifically, the cascade correlation architecture required the determination of the number of hidden nodes, the generalized regression architecture demanded the proper selection of the smoothing coefficient, Bayesian Regularization called for specifying the number of hidden nodes in the hidden layer, and the support vector regression approach required precise and careful fine-tuning of three meta-parameters.

As shown in Figure 6, fine-tuning model parameters is essential across all modelling approaches. For GRNN, BRNN, and ε-SVR, as hyperparameters approach their optimal values, the root mean square error (*RMSE*) decreases for both the fitting and test datasets. The optimal hyperparameter values were selected when this reduction is followed by strong generalization ability, ensuring that both errors are low and close to each other. In contrast, for the cascade correlation model, the basic optimization metric is the correlation coefficient (Figure 6a) due to the system's ability to dynamically add hidden nodes. Increasing the number of hidden nodes significantly improves learning quality, leading to higher correlation coefficient values. In this case, the optimal number of hidden nodes was determined based on the highest achieved correlation value.

Taking the aforementioned points into account, the employed machine learning techniques present a trustworthy and precise solution for estimating the mean diameter reduction factor. From a practical forestry point of view, we believe that the findings obtained and enclosed in these models can serve as a valuable tool, enabling the accurate estimation of the mean diameter reduction factor for trees.

Based on the results of this research, the tested machine learning methodologies demonstrated strong abilities in accurately estimating and predicting the diameter reduction factor of trees. This resulted in effectively capturing the factor's non-linear behaviour, as compared to traditional non-linear regression models. However, a potential challenge lies in the lack of familiarity of machine learning models within forestry practice compared to the more established regression models. This challenge can be addressed through appropriate training for personnel who deal with the application of these models. It is worth noting that regression techniques which are now widely understood, were once considered complex and unfamiliar. Taking into consideration the advancements in computing, the personnel's growing improvement in programming skills and expertise, machine learning can similarly become an accessible and efficient tool for forestry practice. In this context, despite any initial hurdles, machine learning can be considered as a significant potential for professional and effective management of forest ecosystems.

Conclusions

The knowledge of the mean diameter reduction factor of trees can offer various benefits for the sustainable management of forests. For

example, it provides valuable insights into the taper of the tree stem, enabling precise calculations of tree volume. This paper conducts a comparative analysis with the goal of achieving precise and reliable estimation of the mean diameter reduction factor in trees. To achieve this objective, artificial neural network modelling techniques and the support vector machine technique were employed as alternatives to traditional non-linear regression modelling. The case study focuses on accurately estimating the mean stem diameter reduction factor (taper) for standing fir trees (*Abies x borisii-regis* Matff.).

Artificial intelligence can provide accurate, with generalization abilities, modelling solutions for addressing estimation challenges established by non-linear data with outliers and noise, which are commonly faced in forestry. Due to its non-parametric nature, this modelling approach is efficient and does not require prior assumptions or prerequisites.

The comparative analysis between the traditional non-linear model, artificial neural network models, and the support vector regression (ε-SVR) model concluded that the ε-SVR model achieved the best fit based on the evaluation metrics used. The artificial neural network models demonstrated similar performance, whereas the traditional non-linear model proved to be the least suitable for our data.

It's challenging to provide a conclusive answer regarding the superior method to recommend. This decision should be made considering significant factors such as the required effort for implementation in both field and office, along with the necessary skills.

Conflict of interest

The authors declare no financial or personal interests could influence the work presented in this paper.

References

Belete D.M., Huchaiah M.D., 2022. Grid search in hyperparameter optimization of machine learning models for prediction of HIV/AIDS test results. International Journal of Computers and Applications 44(9): 875-886. https://doi.org/10.1080/1206212X.2021.1974663.

Boser B., Guyon I., Vapnik V., 1992. A training algorithm for optimal margin classifiers. In: COLT '92, ACM Publications, Fifth annual workshop on Computational learning theory. Pittsburgh, Pennsylvania, USA, pp. 144–152. https://doi.org/10.1145/130385.130401.

Burden F., Winkler D., 2008. Bayesian Regularization of Neural Networks. In: Livingstone, D.J. (eds), Humana Press, Artificial Neural Networks. Methods in Molecular Biology™, vol 458. https://doi.org/10.1007/978-1-60327-101-1_3.

Diamantopoulou M.J., 2005. Predicting fir trees stem diameters using Artificial Neural Network models. Southern African Forestry Journal 205(1): 39-44. https://doi.org/10.2989/10295920509505236.

Diamantopoulou M.J., 2022. Simulation of over-bark tree bole diameters, through the RFr (Random Forest Regression) algorithm. Folia Oecologica 49(2): 93-101. https://doi.org/10.2478/foecol-2022-0010 .

Diamantopoulou M.J., Özçelik R., Yavuz H., 2018. Tree-bark volume prediction via machine learning: A case study based on black alder's tree-bark production. Computers and Electronics in Agriculture 151: 431-440. https://doi.org/10.1016/j.compag.2018.06.039.

Draper N.R., Smith H., 1998. Applied regression analysis, third ed., Wiley, NY, 716 p. https://doi.org/10.1002/9781118625590.

Efron B., Tibshirani R.J., 1994. An introduction to the bootstrap. Chapman & Hall/CRC, NY, 456 p. https://doi.org/10.1201/9780429246593.

Ercanlı İ., 2020. Innovative deep learning artificial intelligence applications for predicting relationships between individual tree height and diameter at breast height. Forest Ecosystems 7: 1-18. https://doi.org/10.1186/s40663-020-00226-3.

Fahlman S.E., Lebiere C., 1990. The cascade correlation learning architecture. In D.S. Touretsky (ed.), Advances in Neural Information Processing Systems 2, Morgan Kaufmann, San Mateo, CA, pp. 524-532.

Gurney K., 1997. An introduction to neural networks. CRC Press, UK, pp. 234. https://doi.org/10.1201/9781315273570.

Heidarlou H.B., Shafiei A.B., Tayyebi A., Borz S.A., 2023. Evaluating the effect of ongoing conservation policies and forest cover changes in Iranian Zagros forests based on a Land Transformation Model: transition to forest or deforestation? Annals of Forest Research 66(1): 79-97. https://doi.org/10.15287/afr.2023.2628.

Hirigoyen A., Acosta-Muñoz C., Salamanca A.J.A., Varo-Martinez M.A., Rachid-Casnati C., Franco J., Navarro-Cerrillo R., 2021. A machine learning approach to model leaf area index in Eucalyptus plantations using high-resolution satellite imagery and airborne laser

scanner data. Annals of Forest Research 64(2): 165-183. https://doi.org/10.15287/afr.2021.2073.

Hoaglin D.C., Mosteller F., Tukey J.W., 2006. Exploring data tables, trends and shapes. Wiley-Interscience, NY, 576 p.

IBM-SPSS 2021. IBM Corp. Released 2021. IBM SPSS Statistics for Windows, Version 29.0. IBM Corp, Armonk, NY.

Jiao Y., Zhao l., Deng O., Weihend X., Feng Z., 2013. Calculation of live tree timber volume based on particle swarm optimization and support vector regression. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE) 29(20): 160-167. https://doi.org/10.3969/j.issn.1002-6819.2013.20.022.

Kalkanlı Genç S., Diamantopoulou M.J., Özçelik R., 2023. Tree Biomass Modeling Based on the Exploration of Regression and Artificial Neural Networks Approaches. Forests 14: 2429. https://doi.org/10.3390/f14122429.

Kayri M., 2016. Predictive Abilities of Bayesian Regularization and Levenberg–Marquardt Algorithms in Artificial Neural Networks: A Comparative Empirical Study on Social Data. Mathematical and Computational Applications 21(2): 20. https://doi.org/10.3390/mca21020020.

Levenberg K., 1944. A method for the solution of certain non-linear problems in least squares. Quarterly of Applied Mathematics 2: 164-168. https://doi.org/10.1090/qam/10666.

Lilliefors H.W., 1967. On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown. Journal of the American Statistical Association 62(318): 399-402. https://doi.org/10.2307/2283970.

Malek S., Miglietta F., Gobakken T., Næsset E., Gianelle D., Dalponte M., 2019. Prediction of stem diameter and biomass at individual tree crown level with advanced machine learning techniques iForest-Biogeosciences and Forestry 12(3): 323-329. https://doi.org/10.3832/ifor2980-012.

Marquardt D.W., 1963. An algorithm for least-squares estimation of nonlinear parameters. Journal of the society for Industrial and Applied Mathematics 11(2): 431- 441. http://www.jstor.org/sTab./2098941.

Matis K.G., 2004. Forest Biometry II. Dendrometry, Pigasos Press, Thessaloniki, Greece. [In Greek].

Matlab 2022. MATLAB version 19.12.0, release name 2022a, num. 47. The MathWorks Inc., Natick, Massachusetts, available: https://www.mathworks.com/help/stats/index.html.

Monnet J.M., Chanussot J., Berger F., 2011. Support vector regression for the estimation of forest sand parameters using airborne laser scanning. IEEE Geoscience and Remote Sensing Letters 8(3): 580-584. https://doi.org/10.1109/LGRS.2010.2094179.

Montesinos López O.A., Montesinos López A., Crossa J., 2022. Support Vector Machines and Support Vector Regression. In: Multivariate Statistical Machine Learning Methods for Genomic Prediction. Springer, Cham. https://doi.org/10.1007/978-3-030-89010-0_9.

Moradi F., Sadeghi S.M.M., Heidarlou H.B., Deljouei A., Boshkar E., Borz S.A., 2022. Above-ground biomass estimation in a Mediterranean sparse coppice oak forest using Sentinel-2 data. Annals of Forest Research 65(1): 165-182. https://doi.org/10.15287/afr.2022.2390.

Myatt G.J., Johnson W.P., 2014. Making Sense of Data I. A Practical Guide to Exploratory Data Analysis and Data Mining. 2nd ed., John Wiley & Sons, Inc. New Jersey, 231p.

Nocedal J., Wright S.J., 2006. Numerical Optimization. 2nd ed., Springer Series in Operations Research and Financial Engineering, Springer-Verlag NY, 664p. https://doi.org/10.1007/978-0-387-40065-5.

Özçelık R., Diamantopoulou M.J., Eker M., Gürlevık N., 2017. Artificial neural network models: an alternative approach for reliable aboveground pine tree biomass prediction. Forest Science 63: 291-302. https://doi.org/10.5849/FS-16-006.

Özçelık R., Diamantopoulou M.J., Trincado G., 2019. Evaluation of potential modeling approaches for Scots pine stem diameter prediction in north-eastern Turkey Computers and Electronics in Agriculture 162: 773-782. https://doi.org/10.1016/j.compag.2019.05.033.

Patterson D., 1996. Artificial neural networks. Prentice Hall, Singapore, 477p.

Python Software Foundation 2022. © Copyright 2001-22, https://docs.python.org/3.9/index.html.

Russell S., Norvig P., 2020. Artificial Intelligence: A Modern Approach (Pearson Series in Artificial Intelligence), 4th ed., Pearson, 1136 p. https://books.google.gr/books?id=koFptAEACAAJ.

Sandoval S., Acuña E., 2022. Stem taper estimation using artificial neural networks for Nothofagus trees in natural forest. Forests 13: 2143. https://doi.org/10.3390/f13122143.

Sariev E., Germano G., 2020. Bayesian regularized artificial neural networks for the estimation of the probability of default. Quantitative finance, 20(2): 311-328. https://doi.org/10.1080/14697688.2019.1633014

Shaw D.J., Meldahl R.S., Kush J.S., Somers G.L., 2003. A tree taper model based on similar triangles and the use of crown ratio as a measure of form in taper equations for longleaf pine. Gen. Tech. Rep. SRS-66. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 8 p.

Smola A.J., Schölkopf B., 2004. A tutorial on support vector regression. Statistics and Computing 14: 199–222. https://doi.org/10.1023/B:STCO.0000035301.49549.88

Speckt D.F., 1991. A General Regression Neural Network. IEEE Transactions on Neural Networks, 2(6): 568-576. https://doi.org/10.1109/72.97934

Swingler K., 1996. Applying Neural Networks: A Practical Guide. 1st ed., Morgan Kaufmann, 303p.

Titterington D.M., 2004. Bayesian Methods for Neural Networks and Related Models. Statistical Science 19: 128-139. https://doi.org/10.1214/088342304000000099

Thomas S., Pillai G.N., Pal K., 2017. Prediction of peak

ground acceleration using ϵ-SVR, ν-SVR and Ls-SVR algorithm. Geomatics, Natural Hazards and Risk 8(2): 177-193. https://doi.org/10.1080/19475705.2016.1176604

Vapnik V., 1999. Three remarks on the support vector method of function estimation. In: Schölkopf B., Burges C.J.C., Smola A.J. (eds.), Advances in Kernel Methods—Support Vector Learning, MIT Press, Cambridge, MA, pp. 25–42.

Vapnik V., 2000. The nature of statistical learning theory. 2nd ed. Springer, New York, pp 314. https://doi.org/10.1007/978-1-4757-3264-1.

Vapnik V., Golowich S., Smola A., 1997. Support Vector Method for Function Approximation, Regression Estimation, and Signal Processing. In Mozer M., Jordan M., Petsche T. (eds.), Neural Information Processing Systems, Vol. 9. MIT Press, Cambridge, MA.

Vieira G.C., de Mendonça A.R., da Silva G.F., Zanetti S.S., da Silva M.M., Dos Santos A.R., 2018. Prognoses of diameter and height of trees of eucalyptus using artificial intelligence. Science of the Total Environment 619: 1473-1481. https://doi.org/10.1016/j.scitotenv.2017.11.138

Wang D., Wang M., Qiao X., 2009. Support vector machines regression and modeling of greenhouse environment. Computers and Electronics in Agriculture 66(1): 46-52. https://doi.org/10.1016/j.compag.2008.12.004

West P.W., 2015. Tree and Forest Measurement. 3rd ed, Springer, Cham, pp 214. https://doi.org/10.1007/978-3-319-14708-6

Williams G., 2011. Data Mining with Rattle and R: The Art of Excavating Data for Knowledge Discovery, use R. Springer Science+Business Media, LLC. https://doi.org/10.1007/9781441998-2

Zhou R., Wu D., Zhou R., Fang L., Zheng X., Lou X., 2019. Estimation of DBH at forest stand level based on multi-parameters and generalized regression neural network. Forests 10: 778. https://doi.org/10.3390/f10090778

Zuur A.F., Ieno E.N., Elphick C.S., 2010. A protocol for data exploration to avoid common statistical problems. Methods in Ecology and Evolution 1: 3–14. https://doi.org/10.1111/j.2041-210X.2009.00001.x