Diameter distribution modeling based on Artificial Neural Networks for Kunduz Forests

İlker Ercanli[a,*], Ferhat Bolat[a]

[a] Çankırı Karatekin University, Faculty of Forestry, Çankırı, Turkey
[*] Corresponding author: ilkerercanli@karatekin.edu.tr

Abstract: The diameter distribution models are tools for providing more detailed knowledge on the forest structure, detailed predictions for stand volume, basal area and number of trees on diameter classes. A wide range of probability density functions have been used in forestry to model tree diameter distributions (e.g., normal, log-normal, gamma, Weibull, beta, Johnson-SB), although the three-parameter Weibull is possibly the most frequently used in forest applications due to their ability to describe flexibly various diameter distributions. Alternative techniques for modelling diameter distributions, Artificial Neural Network Analysis (ANN) has been introduced in forest literature for modeling different individual and stand attributes. Although different statistical modeling techniques based on the probability density functions have been proposed to model the diameter distributions, only a few studies concerning an applications of Artificial neural networks (ANNs) to predict tree frequency in diameter class. The data used in this study were collected by Turkey Forest Management Directorate as a part of local forest inventory works from even-aged pine forest stands located in the Kunduz Forests. In these stands, 637 sample plots were used to model diameter distributions by using the 3-parameters Weibull probability density function based some different parameter prediction methods and various Artificial neural network type such as the feed-forward backprop, Cascade Correlation, Elman backprop, Layer Recurrent and NARX network. As application of probability density function, the parameter of Weibull pdf were estimated by using five methods based on 25th, 31th, 50th, 63th and 95th percentiles obtained from data including 637 sample plots. Also, Applications of Artificial neural networks (ANNs) were carried out to model diameter distributions in Kunduz forests. Multiple layer network structures such as the feed-forward backprop, Cascade Correlation, Elman backprop, Layer Recurrent and NARX network with training function of Levenberg-Marquardt and transfer function of Log-sigmoid transfer function were used to obtain relative frequency predictions. Also, these artificial neural networks (ANNs) structures such as the feed-forward backprop, Cascade Correlation, Elman backprop, Layer Recurrent and NARX network with five Weibull parameter prediction methods were compared based on evaluations of the magnitudes and distributions of models' residual and six goodness-of-fit statistics. As training ANNs using these 25th, 50th and 95rd percentiles, ANN based on the feed-forward backprop gave better fitting ability with AIC (34360.8), RMSE (27.28), R^2 (0.784) than the Weibull parameter prediction methods and the other studied ANNs. This ANNs model based feed-forward backprop accounted for more than 78 % of total variance in number of trees in diameter classes with these diameter percentiles values.
Keywords: Artificial Neural Network, Diameter distribution, Probability density Functions, Weibull function

1. Introduction

The predictions for diameter distribution of trees is important to define forest structure and for different forestry calculations (Loetsch et al., 1973). These predictions can present vital information concerning the size-class distribution of a forest stand, especially in the arrangement of a tabulation of numbers of trees by diameter class (Wang and Rennolls, 2005). The biometricians who studied about diameter distribution modeling have long made effort with the probability density functions (pdfs) to describe diameter distributions of forest stands. At the beginning of literature for diameter distribution modeling, 1883, Gram proposed "the normal distribution" for the distribution oriental beech stands, de Liocourt introduced a technique based on the geometric progression for defining diameter distributions from uneven-aged forests in 1898 (Meyer and Stevenson, 1943). In 1930's, some mathematical series were used to define the diameter distributions. As first probability density function (pdf)'s applications to forestry, different probability density functions (pdf's) such as the log-normal (Bliss and Reinker, 1964), gamma (Nelson, 1964), beta (Clutter and Bennett, 1965), Weibull (Bailey and Dell, 1973), and Johnson's SB distributions (Harley and Schreuder, 1977) were used in literatures. Within these functions, the 3-parameters Weibull probability density function have been favored widely in forestry due to their capability to describe compliantly various diameter distributions (Cao, 2004; Mateus and Tomé, 2011; Lima, et al., 2014). The weibull probability density function has gained importance due to the simplicity in estimating its parameters and it flexibility in fitting wide varieties of unimodal shapes.

Artificial neural network (Ann) has been increasingly used as an alternative method and efficient tools for fitting diameter distributions, which is regardless of any distribution function where a suitable statistical function must first be established. Artificial neural network correspondingly studies to being mathematical model of information processing, which is like to the structure of the synapses of the brain, and it is constituted with a large number of interconnected nodes (or neurons) (Chaoui et. al., 2009; Ghosal and Chaki, 2010). ANN have nonlinear connections of natural systems to have advantage for especially describing the relationships between nonlinear tree and forest attributes in forestry. While several studies have developed the model providing some prediction and forecasting in a number of areas, including finance, power generation, medicine, water resources and environmental science, only a few studies concerning with modeling diameter distributions using Artificial neural network modeling approach exist, e.g. Leduc et. al., 2001; Abbasi et. al. 2008; Cai et. al. (2012); Diamontopoulou et. al., 2015. Therefore, the objectives of the study are (1) to evaluate various artificial neural network type such as the feed-forward backprop, Cascade Correlation, Elman backprop, Layer Recurrent and NARX network in prediction of diameter distributions for Kunduz forests and (2) compared these artificial neural network's predictions with the 3-parameters Weibull probability density function based some different parameter prediction methods.

2. Material and methods

The data used in this study were collected by Turkey Forest Management Directorate as a part of local forest inventory works from even-aged pine forest stands located in the Kunduz Planning Unit, Vezirköprü Forest Enterprise, Samsun Forest District Directorate, northeast Turkey (longitude 35°48′–35°01′W; latitude 41°00′–41°19′N). The study area was characterized geomorphologically as a steep terrain land with moderate and steep slopes ranging from 20% to 50%, with an average of 45%. The average annual temperature reaches a maximum of 32.5°C in the summer and a minimum of 6.2°C in winter, with an average annual temperature of 10.6°C. The average annual precipitation of the study area is 500 mm. The climatic regime is a typical Black Sea climate, characterized by a mild winter and a cool summer.

In these stands, 637 sample plots were used to model diameter distributions by using the 3-parameters Weibull probability density function based some different parameter prediction methods and various Artificial neural network type such as the feed-forward backprop, Cascade Correlation, Elman backprop, Layer Recurrent and NARX network. These sample plots have circular shape and the size of these were assessed as 400 m^2, 600 and 800 m^2 by considering crown closure. These sample plots are obtained by based on standard Turkey Forest Management Inventory system. In each sample plot, DBH was measured using calipers for every living tree with a DBH > 8 cm and total tree height (h) was not measured.

As application of probability density function, the parameter of Weibull pdf were estimated by using five methods based on 25^{th}, 31^{th}, 50^{th}, 63^{th} and 95^{th} percentiles obtained from data including 637 sample plots. The 3-parameters Weibull probability density function and these five method based on diameter percentiles were given in Equation 1-6, respectively.

The 3-parameters Weibull probability density function:

$$F(x,\alpha,\beta,\gamma) = \frac{\alpha}{\beta}\cdot\left(\frac{x-\gamma}{\beta}\right)^{\alpha-1}\cdot exp\left(-\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)$$

The method 1 including 31^{th} and 63^{rd} percentiles:

$$\alpha = 0.5\cdot d_{min} \qquad \gamma = \frac{Ln\left(\frac{Ln(1-0.63)}{Ln(1-0,31)}\right)}{Ln(d_{\%63}-\alpha)-Ln(d_{\%31}-\alpha)} \qquad \beta = \frac{d_{\%63}-\alpha}{(-Ln(1-0.63))^{\frac{1}{\gamma}}}$$

The method 2 including 50^{th} and 95^{rd} percentiles:

$$\alpha = 0.5\cdot d_{min} \qquad \gamma = \frac{Ln\left(\frac{Ln(1-0.95)}{Ln(1-0,50)}\right)}{Ln(d_{\%95}-\alpha)-Ln(d_{\%50}-\alpha)} \qquad \beta = \frac{d_{\%50}-\alpha}{(-Ln(1-0.50))^{\frac{1}{\gamma}}}$$

The method 3 including 25^{th}, 50^{th} and 95^{rd} percentiles:

$$\alpha = 0.5\cdot d_{min} \qquad \gamma = \frac{Ln\left(\frac{Ln(1-0.95)}{Ln(1-0,25)}\right)}{Ln(d_{\%95}-\alpha)-Ln(d_{\%25}-\alpha)} \qquad \beta = \frac{d_{\%50}-\alpha}{(-Ln(1-0.50))^{\frac{1}{\gamma}}}$$

The method 4 including 31^{th}, 50^{th} and 63^{rd} percentiles:

$$\alpha = 0.5\cdot d_{min} \qquad \gamma = \frac{Ln\left(\frac{Ln(1-0.63)}{Ln(1-0,31)}\right)}{Ln(d_{\%63}-\alpha)-Ln(d_{\%31}-\alpha)} \qquad \beta = \frac{d_{\%50}-\alpha}{(-Ln(1-0.50))^{\frac{1}{\gamma}}}$$

The method 5 including minimum, quadradic mean diameter (dg), 25^{th}, 50^{th} and 95^{rd} percentiles:

$$\alpha = \frac{n^{0.3333}\cdot d_{min}-d_{\%50}}{n^{0.3333}-1} \quad \gamma = \frac{2.343088}{Ln(d_{\%95}-\alpha)-Ln(d_{\%25}-\alpha)} \qquad \beta = \frac{\alpha\cdot\Gamma_1}{\Gamma_2} + \sqrt{\left(\frac{\alpha}{\Gamma_2}\right)\cdot\left(\Gamma_1^{\,2}-\Gamma_1\right)+\left(\frac{d_g^{\,2}}{\Gamma_2}\right)}$$

In the neural network model building, both training, verification and testing data sets that randomly partitioned into training (75% of all data), verification (15% of all data) and test (the remaining 10% of all data) data sets were used for taking general relationships between input variables and target variable. In ANN training process, input variables were diameter central value for each diameter class, best predictive percentiles values within these five method and number of tree for sample plots. Target variable is relative frequencies calculated as ratio of number of trees in diameter classes to total number of trees in each sample plots. Multiple layer network structures such as the feed-forward backprop, Cascade Correlation, Elman backprop, Layer Recurrent and NARX network with training function of Levenberg-Marquardt and transfer function of Log-sigmoid transfer function were used to obtain relative frequency predictions. In ANNs training process, the number of neurons is used as 10 with number of two layers including hidden and output layers, since this structure are the most frequently chosen values in ANNs. All these applications for ANN was carried out using MATLAB-nntool module. Also, these artificial neural networks (ANNs) structures such as the feed-forward backprop, Cascade Correlation, Elman backprop, Layer Recurrent and NARX network with five Weibull parameter prediction methods were compared based on evaluations of the magnitudes and distributions of models' residual and six goodness-of-fit statistics: sum of squared errors (SSE), Akaike's information criterion (AIC), Bayesian information criterion (BIC), Root Mean Square Error (RMSE), Mean Squared Error (MSE) and Adjusted Coefficient of Determination (R^2_{adj}).

3. Result and discussions

The values of goodness-of-fit statistics, such as SSE, AIC, BIC, RMSE, MSE and R^2_{adj}, for these ANNs, including the feed-forward backprop, Cascade Correlation, Elman backprop, Layer Recurrent and NARX network types, and five Weibull parameter prediction methods based on diameter percentiles were given in table 1. The AIC was between 35677.2 and 36349.1, RMSE between 30.97 and 33.04, and R^2 between 0.684 and 0.722 in Weibull parameter prediction methods. However, The AIC was between 34360.8 and 35167.2, RMSE between 27.28 and 29.48, and R^2 between 0.748 and 0.784 in ANNs. Within Weibull parameter prediction methods, the method 3 with 25th, 50th and 95rd percentiles produced more predictive results than other methods. As training ANNs using these 25th, 50th and 95rd percentiles, ANN based on the feed-forward backprop gave better fitting ability with AIC (34360.8), RMSE (27.28), R^2 (0.784) than the Weibull parameter prediction methods and the other studied ANNs. This ANNs model based feed-forward backprop accounted for more than 78 % of total variance in number of trees in diameter classes with these diameter percentiles values.

Table 1. The goodness-of-fit statistics of number of trees predictions for the ANNs types and Weibull parameter prediction methods

Technique	SSE	AIC	BIC	RMSE	MSE	R^2_{adj}
The method 1 with 31th and 63rd percentiles	5337121	36033.9	36049.0	32.05	1027.16	0.702
The method 2 with 50th and 95rd percentiles	5670836	36349.1	36364.2	33.04	1091.38	0.684
The method 3 with 25th, 50th and 95rd percentiles	4983020	35677.2	35692.3	30.97	959.01	0.722
The method 4 with 31th, 50th and 63rd percentiles	5355450	36051.7	36066.9	32.10	1030.69	0.701
The method 5 with minimum, quadradic mean diameter, 25th, 50th and 95rd percentiles	5227325	35925.9	35941.0	31.72	1006.03	0.709
ANN based on the feed-forward backprop	3867748	34360.8	34375.9	27.28	744.37	0.784
ANN based on Cascade Correlation	4325355	34941.8	34956.9	28.85	832.44	0.759
ANN based on Elman backprop	4517125	35167.2	35182.3	29.48	869.35	0.748
ANN based on Layer Recurrent	4000311	34535.9	34551.0	27.75	769.88	0.777
ANN based on NARX	4279939	34886.9	34902.0	28.70	823.70	0.761

4. Conclusions

These results underlined that the ANN are able to predict the number of trees in diameter classes describing diameter distributions, and to generate more accurate predictions than parameter prediction method for the 3-parameters Weibull probability density function. These empirical results for ANN helped to develop successfully for modeling diameter distributions. Based on the results obtained from the research, forest managers could use ANN for predicting number of trees in diameter classes and this information would also be beneficial for evaluating different management strategies and developing forest management plans. It may produce unbiased diameter distribution predictions using different ANN structure in specific sample plots. However, these predictions will be improved by the addition of permanent sample plots comparing alternative silvicultural treatments and alternatives. However, these results should be generalized to outside the studied forest area was further analyzed and evaluated at other forest sites since the results of this tree growth predictions are specific to this species and the type of forest structure that was studied. It is expected that this study will make a valuable contribution to forestry literature.

Acknowledgment

We would like to thank to for support to the Head of Forest Management and Planning Department, General Directorate of Forestry, Republic of Turkey.

References

Abbasi, B., Rabelo, L., Hosseinkouchack M., 2008. Estimating parameters of the three-parameter Weibull distribution using a neural network. European Journal of Industrial Engineering, 2: 428-445.

Bailey, R. L., Dell, T. R., 1973. Quantifying Diameter Distributions with The Weibull Function. Forest Science, 19: 97–104

Bliss, C. L., Reinker, K. A., 1964. A Lognormal Approach to Diameter Distributions in Even-Aged Stands. Forest Science, 10: 350–360.

Cai, S., Kang, X., Zhang, L., 2012. Simulation of Trunk diameter distribution for three broadleaved trees based on artificial neural networks. International Journal of Advancements in Computing Technology, 4: 520-527.

Cao, Q. V., 2004. Predicting Parameters of A Weibull Function for Modelling Diameter Distribution. Forest Science, 50: 682 – 685.

Chaoui, H., Sicard, P., Gueaieb, W., 2009. ANN-based adaptive control of robotic manipulators with friction and joint elasticity. IEEE Trans Ind Electron, 56: 3174–3187

Clutter, J. L., Bennett, F. A., 1965. Diameter Distributions in Old-Field Slash Pine Plantation, Georgia Forest Research Council, Report No.13.

Diamantopoulou, M. J., Özçelik, R., Crecente-Campo, F., Eler, Ü., 2015. Estimation of Weibull function parameters for modelling tree diameter distribution using least squares and artificial neural networks methods. Biosystem enginering, 133: 33-45.

Ghosal, S., Chaki, S., 2010. Estimation and optimization of depth of penetration in hybrid CO2 LASER-MIG welding using ANN-optimization hybrid model. International Journal of Advanced Manufacturing Technology, 47:1149-1157.

Hafley, W.L., Schreuder, D. H. T., 1977. Statistical distributions for fitting diameter and height data in even-aged stands. Canadian journal forest Research, 7: 481–487.

Lima, R. A. F., Batista J. L. F., Prado, P. I., 2014. Modeling tree diameter distributions in Natural Forests: an evaluation of 10 statistical models. Forest Science, 61, 2: 320-327.

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

Loetsch, F., Zöhler, F., Haller, K. E., 1973. Forest inventory 2. BLV Verlagsgeschellschaft, Munic, Germany. 469 p.

Nelson, T. C., 1964. Diameter Distribution and Growth of Loblolly Pine. Forest Science, 10: 105–115.

Mateus A., Tomé M., 2011. Modelling the diameter distribution of eucalyptus plantations with Johnson's S_B probability density function: parameters recovery from a compatible system of equations to predict stand variables. Annals of Forest Science, 68: 325-335.

Meyer H. A., Stevenson D. D., 1943. The structure and growth of virgin beech-birch-maple-hemlock forests in northern Pennsylvania. Journal of Agricultural Research, 67: 465–484.

Wang, M., Rennolls. K., 2005. Tree diameter distribution modelling: Introducing the logit-logistic distribution. Canadian journal forest Research, 35: 1305–1313.