Power System Technology
ISSN:1000-3673

Received: 16-05-2024 Revised: 12-06-2024 Accepted: 17-07-2024

# Evaluating the Performance of Machine Learning Approaches in Estimating Reforested Populus (Populus Deltoids) Trunk Weights

**Sina Pourrajabali Moafi**[1]*

(Email: sinapourrajabalimafi@gmail.com; ORCID: https://orcid.org/0009-0003-7561-9238)

**Vahid Hemmati**[1]

(Email: vahidhemmatiliau@gmail.com)

**Alireza Eslami**[2]

(Email: eslami117@gmail.com)

**Seyed Armin Hashemi**[1]

(Email: seyedarminhashemi@gmail.com)

**Seyed Yousef Torabian**[1]

(Email: drtorabian@yahoo.com)

1) Department of Forest science and engineering, Lahijan Branch, Islamic Azad University, Lahijan, Iran.

2) Department of Horticultural Science, Rasht Branch, Islamic Azad University, Rasht, Iran

*Corresponding author: Vahid Hemmati )

**Abstract**

Measuring the weight of trees for the users of this industry is always possible after cutting the trees and measuring it directly through a scale. This method has many difficulties and problems. Hence, researchers have always looked for indirect methods of measuring the weight of tree trunk wood to overcome these limitations. The current research aimed at providing a numerical model to estimate the weight of Populus Deltoids tree trunk. It examined 400 trees in the afforested areas of the west of the Gilan province in northern Iran, in a humid-moderate climatic region. Then 11 variables of each tree were measured before cutting them. These variables are diameter at height 1.30 meters (D1.3), diameter at height 1 meter, diameter at height 2 meters, diameter at height 3 meters (D3), diameter at height 4 meters (D4), collar diameter (DC), stump diameter, crown height, trunk height (Htr) and stump height; these were independent variables (input) in modeling. Then, the weight of the trunk of the trees was got after cutting the trees by direct measurement through a scale. Pearson's correlation test showed that variables D1.3, D3, D4, DC and Htr are the most effective variables. The combination of these five variables gave the arrangement of the models' input scenarios. After dividing the trees into calibration phases (including 300 trees [75%]) and validation (including 100 trees [25%]), the research tested three models including multiple linear regression (MLR), multilayer

Power System Technology
ISSN:1000-3673
Received: 16-05-2024 Revised: 12-06-2024 Accepted: 17-07-2024

perceptron (MLP) and hybridized form MLP with Genetic Optimization Algorithm (MLP-GA). It reported in comparison the MLR linear model as superior to the two artificial intelligence models MLP and MLP-GA. Comparing the performance of MLP and MLP-GA models showed that the combination of GA with MLP model can increase in average the accuracy of MLP model in estimating tree trunk weight by 16.6%. This superiority showed that the linear relationship between the tree variables and the trunk weight variable is stronger than the non-linear relationship between them. The evaluation criteria of the best estimation of tree trunk weight are equal to root-mean-square error = 65.98 kg and $R^2$ = 91.93%, which belongs to the MLR model. The current approach can provide information on the amount of wood of tree planting to gardeners and users of wood industry before cutting trees. It is probably a valuable research source for other similar and different climatic regions and other species of economic trees.

**Keywords:** Populus Deltoids, Tree Weight, Numerical Modeling, Trunk Weight Estimation, Artificial Neural Networks

## 1. Introduction

Trees as the most valuable natural resources play a very important role in human life. They contribute to the continuity of human life and the health of the natural cycle by absorbing toxic gases, including carbon dioxide, and producing oxygen, which is a most basic need of human and animal life. Nowadays, the need for wood has caused the afforestation of fast-growing trees to be very important in the world. Fast-growing and deciduous trees such as poplar, which mostly have large, stout and single trunks can be because of their short exploitation period a very suitable option for meeting the wood needs of human societies [1,2,3,4]. Poplars are among the trees, which have always been the focus of villagers and local gardeners because of their unique features such as the possibility of cultivation in different climatic conditions, the cultivatability in small, large, single tree, mass and row levels, and cultivation as a windbreaker[5,6]. It has also interested wood producers, owners of various wood industries and other relevant users because of the ease of propagation and the possibility of its cultivation together with agricultural products, the usefulness of its leaves in animal feed, the possibility of exploitation in short periods of time, etc.[7,8,9].

Researchers have introduced many methods to estimate the amount of wood. They have used many models to estimate the volume of the tree trunk such as Newton, Ismailin, Pressler and Huber models[10,11] . But all the gardeners of poplar trees, as well as other trees, sell the wood of their poplar cultivation based on its weight component. Therefore, the weight of the trees is supposedly a much more important component than the volume for the poplar cultivators, from which the relevant professionals can find out their income. This component is very important for poplar cultivators, wood contractors, manufacturers, managers, natural resource planners and wood industry. Measuring the tree weight for the users of this industry is always possible after cutting the trees and measuring it directly through a platform scale. This method is associated with many difficulties and problems, and cannot provide information on the amount of the wood of tree planting to gardeners and wood industry users in the period before cutting trees. Hence, researchers have always looked for indirect methods to measure the weight of tree trunks to overcome these limitations.

The numerical methods of artificial intelligence and machine learning have expanded and performed in recent years very well in estimating variables in forestry science and engineering.

Power System Technology
ISSN:1000-3673
Received: 16-05-2024 Revised: 12-06-2024 Accepted: 17-07-2024

Neural networks are among these models such as multilayer perceptron (MLP) and radial basis function (RBF). Investigated two neural networks, RBF and MLP during a study on a 60-hectare educational and breeding forest of Gorgan, regarding the relationship between the number per hectare of forest and topographical characteristics. As the results showed, both neural network models have excellent performance in this field[12]. Multiple linear regression (MLR) analysis was performed to compare its results with artificial intelligence models, and the results showed the superiority of neural networks over MLR in estimating the number of trees per hectare. The results also showed that the neural network technique can accurately estimate 65% of changes in the number per hectare of forest through topographical characteristics. Another study selected the marked trees of the educational research forest of Tarbiat Modares University, and measured with high accuracy the diameter at the breast, the diameter at the height of the log, the diameter at the end of the trunk, the height of the trunk and the height of the whole tree[10]. Then, they estimated the volume of tree trunks by using these variables as inputs of RBF and MLP models. The results showed the RBF model is relatively more accurate than the MLP model in estimating the tree trunk volume. A study examined the possibility of predicting the width of earthworks of Surdar-Vatashan forest roads. It made MLP and MLR models in MATLAB and R software, and carried out the correlation tests and analysis of variance in the SPSS environment[13]. According to the evaluation of the modeling results, the MLP model can provide more successful estimates than the MLR. A study used the MLP model to estimate the industrial and firewood volumes of trees. The results showed the acceptable accuracy of the MLP model in estimating both types of volumes[14]. Used the MLP model to estimate the survival and mortality of trees in northern Iran[11]. The estimates of this model were compared with the MLR model, and the results revealed the superiority of the neural network in estimating the survival probability of trees. tested and evaluated the performance of MLP and MLR models in estimating bark thickness and effective variables in forestry. The results also made known the superiority of the accuracy of the MLP model compared to the MLR in estimating the bark thickness of trees[15].

Previous studies have provided no method for indirect measurement of the tree weight component, and the tested models have often aimed at estimating components such as tree trunk volume, tree bark thickness, survival and mortality of trees, etc. The current research evaluates for the first time the numerical models in estimating the weight of tree trunks. Some meta-heuristic models also emerged in recent years in engineering sciences; these result from integrating complex optimization algorithms (such as genetic algorithm, particle swarm optimization algorithm, etc.) with neural network models. These models can significantly improve the accuracy of the estimates by optimizing the parameters of the neural networks. This increase in accuracy is desirable and confirmable in many scientific studies, such as solar radiation modeling[16,17] evaporation-transpiration modeling [18,19]lake water level modeling , snow modeling (, groundwater modeling ,drought etc[20,21]. However, no researcher has used so far such hybrid models in forestry science and engineering. Thus, we have used the meta-heuristic form of the MLP model (MLP merged with genetic algorithm, as MLP-GA) besides the MLR and MLP models, and compared and evaluated at the end their performance, which is another aspect of innovation in current research. The cultivation of poplar in the Gilan province, Iran, is a financial support for every farmer so that he can have hope for a favorable income from the exploitation of its wood product that is the subject of the current research.

Power System Technology
ISSN:1000-3673

Received: 16-05-2024 Revised: 12-06-2024 Accepted: 17-07-2024

## 2. Materials and Methods

### 2.1. Study region

The region under study is the 19th series of Pilembera in the 9th district of Shafarood, hand-planted forests that are in the north of Rezvanshahr and 4 km from Shafarood Company. It is under the supervision of the Department of Natural Resources from the protection area of Rezvanshahr, under the supervision of Gisoom plans, and is in the range of Y=41641225, X=356600, Y=4157928, X=331941 according to the UTM system.

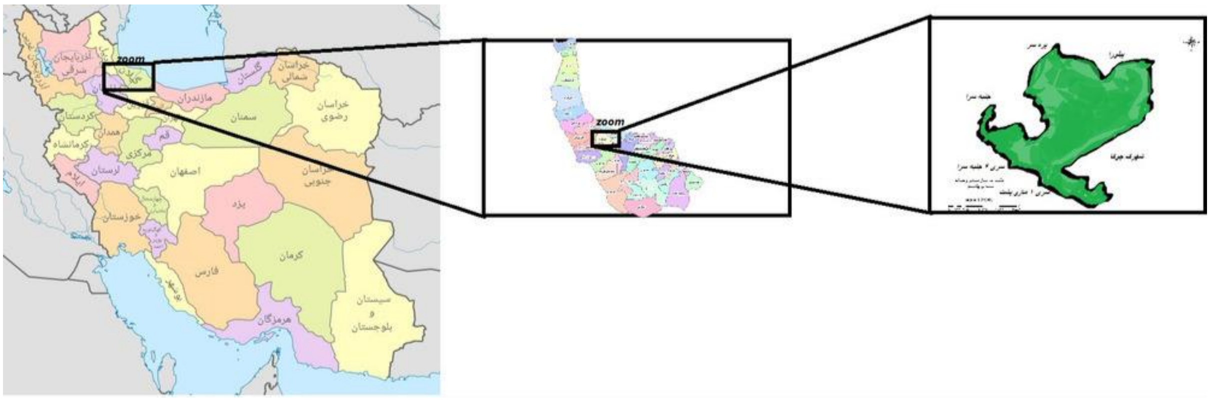


Figure 1. Situation of the study region

The desired series is limited from the north to Pareh Sar and village Abansara, from the south to Chuka town, from the east to the seashore, and from the west to 18 series Janbesara. The total area of this series is about 1594.6 hectares, of which 1281.5 hectares are usable, 61.5 hectares are vacant forest areas, 10.1 hectares are cultivated areas and occupations, 65.9 hectares are protected areas, 5/ 71 hectares are existing roads, 1.1 hectares are high-voltage power lines, 10 hectares are research afforestation, 14.5 hectares are fire line and 78.5 hectares are annual afforestation which are not hidden in the usable area. The average annual rainfall of the region is 1484 mm. The parent rocks of this region are a combination of Quaternary deposits. They are mostly deposits and debris. The parent rocks of this area are sandy, which belong to the first Mesozoic period. Its upper layer is composed of sandstones and its lower layer is composed of alternating layers of shale and sand. This region has brown and ash-brown soil, which has acid characteristic; the climate is Caspian region.

The annual afforestation range of the series 19 of Pilembera is part of the steppe parts of the Hyrcanian forests. Naturally, the main population of these regions in the past was oak grove, and was accompanied by species of mummers, alder, parrotia persica, white pellet, and ulmus minor. In the last three decades, it has been reforested with species such as spruce, summer alder, winter alder, Taeda pine, maple, van, etc. because of the destructions.

Power System Technology
ISSN:1000-3673

Received: 16-05-2024 Revised: 12-06-2024 Accepted: 17-07-2024

Figure 2. A view of spruce plantations in the studied area

**2.2. Data collection**

400 Populus deltoids were randomly selected from the afforestation of the region under study, after a forest excursion and finding the desired sample plot. The mentioned researches used 400 deltoids for modeling; but since a larger number of samples always leads to the presentation of a more valid model, this research attempted to consider a larger number of samples for this purpose. This study has used the following measurements:

1-Measurement of breast diameter (1.30 cm), diameter in one meter, two meters, three meters and four meters using a two-armed ruler (caliper).

2- Measurement of the log diameter and the diameter of the collar through a caliper.

3- Measurement of the total height, crown height and trunk height before the cutting operation by Bloomless device, and Measurement of the log height after the cutting operation by a tape measure.

4- Measurement of the weight of each tree after the cutting operation by a scale as a dependent variable in determining relationships.

Power System Technology
ISSN:1000-3673
Received: 16-05-2024 Revised: 12-06-2024 Accepted: 17-07-2024

Figure 3. Moments of tree weight measurement steps

Table 1 has shown the results of the statistical summary of the data.

Table 1. statistical characteristics of the measured variables

| Variable (abbreviation) | Mean | StDev | CoefVar | Minimum | Maximum | Skewness |
|---|---|---|---|---|---|---|
| Total height (Htot) | 29.030 | 102.080 | 351.56 | 2.48 | 2057.00 | 19.72 |
| Diameter at height 1.30 (D1.3) | 26.731 | 4.832 | 18.08 | 0.90 | 47.60 | 0.11 |
| Diameter at height 1 (D1) | 49.100 | 258.200 | 525.67 | 16.30 | 3630.00 | 12.10 |
| Diameter at height 2 (D2) | 33.340 | 148.910 | 446.66 | 3.50 | 3010.00 | 20.02 |
| Diameter at height 3 (D3) | 24.973 | 4.455 | 17.84 | 15.20 | 44.50 | 0.47 |
| Diameter at height 4 (D4) | 23.905 | 4.383 | 18.33 | 14.30 | 43.10 | 0.50 |

Received: 16-05-2024 Revised: 12-06-2024 Accepted: 17-07-2024

| | | | | | | |
|---|---|---|---|---|---|---|
| Collar diameter (Dc) | 33.427 | 7.762 | 23.22 | 17.90 | 74.90 | 1.17 |
| Stump diameter (Ds) | 36.180 | 116.370 | 321.65 | 17.00 | 2360.00 | 19.96 |
| Crown height (Hc) | 11.528 | 5.842 | 50.67 | 3.60 | 121.58 | 16.73 |
| Trunk height (Htr) | 12.288 | 2.582 | 21.01 | 7.14 | 20.00 | 0.47 |
| Stump height (Hs) | 9.582 | 5.711 | 59.61 | 3.50 | 90.40 | 10.20 |
| Weight (W) | 558.600 | 203.700 | 36.48 | 98.90 | 1256.00 | 0.19 |

**2.3. Multiple linear regression (MLR)**
Multivariate linear regression is a linear relationship between several series of independent variables (input) to model a dependent variable (target). The basis of this model is the optimization of regression coefficients for each input variable through the Least Squares algorithm. The mathematical form of this model is:

$$y = b_1x_1 + b_2x_2 + \cdots + b_nx_n + c \quad \text{(Eq. 1)}$$

Where y is the target variable, $x_i$ is input variable, $b_i$ is regression coefficient and c is the constant coefficient of the model[22] .

**2.4. Multilayer perceptron (MLP) neural network**
The MLP model is a type of artificial neural network the law of back propagation error completes its structure. Each neuron in the MLP network performs two calculations, the first is the functional signal, and the second is the instantaneous estimation of the gradient of the error curve regarding the parameters that connect the input of the neuron to the neuron itself. These gradients are needed to propagate the error signal in the network [23]. Figure 4 shows the MLP network convention with a 2-4-2-1 structure (2 inputs; 2 hidden layers with 4 neurons in the first and 2 neurons in the second hidden layer; 1 output).

Power System Technology
ISSN:1000-3673
Received: 16-05-2024 Revised: 12-06-2024 Accepted: 17-07-2024

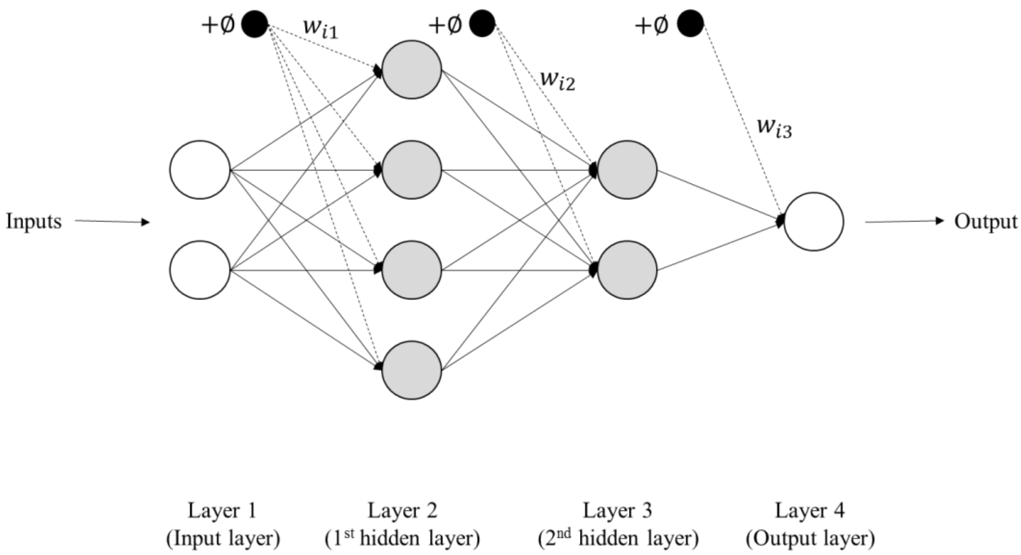


Figure 4. Structure of a MLP neural network with two hidden layers

The matrix of weights ($w_i$) establish the internal relations of the neural network. The input pulse of each neuron is a function of the matrix of weights and information received by that neuron ($x_i$); equation (3) expresses it. ∅ in this equation is the amount of bias.

$$net = \sum_{i=1}^{n} w_i x_i + \emptyset \qquad \text{(Eq. 2)}$$

The goal of each training algorithm is to minimize the network error (equation 4).

$$E = \frac{1}{PN_{output}} \sum_{p}^{p} \sum_{i=1}^{N_{output}} (t_i - o_i)^2 \qquad \text{(Eq. 3)}$$

P in this equation is the number of training patterns, N is the number of output neurons, $o_i$ is the observed value of i-th neuron, $t_i$ is the calculated value of i-th neuron by the network. This research uses the Levenberg-Marquardt algorithm to train the designed network because of its high speed in the training process. For additional information about the MLP model, see references [24].

**2.5. Genetic algorithm (GA)**

Genetic algorithm is a meta-heuristic method for solving optimization problems. presented it as a part of the set of evolutionary algorithms. This algorithm uses methods derived from natural evolution such as selection and displacement, and Mutation on a population of acceptable solutions results in the optimal answer for problems. The key component of GA is a chromosome, which represents a solution in the search space of the optimization problem. Each chromosome is composed of genes, each of which explains a parameter of the problem [25]. The search in this algorithm starts with an initial population of strings and in each iteration, unique strings are evaluated according to the efficiency condition and a fitness value is assigned to them.

Received: 16-05-2024 Revised: 12-06-2024 Accepted: 17-07-2024

Selection is an action in which the chromosomes of the next generation are determined based on the survival law of the current generation, which is the most important part in this algorithm. The general method of chromosome selection is to determine the probability of their selection based on the value of the target function. Thus, some random numbers are generated and compared with their cumulative probability. To know the details and mathematical relationships of this algorithm, see reference articles[25]. The adapted value of GA operators in MLP neural network optimization is observable in Table 2.

Table 2. Arrangement of genetic algorithm's operators for optimizing MLP model

| Parameter | Value |
|---|---|
| Population | 100 |
| Maximum Number of Iterations | 200 |
| Mutation Rae | 0.5 |
| Crossover percentage | 0.7 |
| Selection pressure | 0.8 |

**2.6. Performance evaluation metric**

We must compare the model outputs with their actual values to measure the validity of the prediction provided by the models, for which evaluation criteria are appropriate. The criteria of this research are: Root Mean Square Error (RMSE), Normalized RMSE (NRMSE), Coefficient of Determination ($R^2$), Willmott Index (WI) and Nash Sutcliff (NS) efficiency whose equations. The equations:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - f_i)^2} \qquad \text{(Eq. 4)}$$

$$NRMSE = \frac{RMSE}{y_{max} - y_{min}} \qquad \text{(Eq. 5)}$$

$$R^2 = \left[\frac{\sum_{i=1}^{n}(y_i - \bar{y})(f_i - \bar{f})}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2} * \sqrt{\sum_{i=1}^{n}(f_i - \bar{f})^2}}\right]^2 \qquad \text{(Eq. 6)}$$

$$WI = 1 - \frac{\sum_{i=1}^{N}(y_i - f_i)^2}{\sum_{i=1}^{N}(|f_i - \bar{y}| + |y_i - \bar{y}|)^2} \qquad \text{(Eq. 7)}$$

$$NS = 1 - \frac{\sum_{i=1}^{n}(y_i - f_i)^2}{\sum_{i=1}^{n}(y_i - \bar{f})^2} \qquad \text{(Eq. 8)}$$

Power System Technology
ISSN:1000-3673

Received: 16-05-2024 Revised: 12-06-2024 Accepted: 17-07-2024

$y_i$ and $\bar{y}$ in the above equations are the observed data and their average, respectively; $f_i$ and $\bar{f}$ are predicted data and their average; $y_{max}$ and $y_{min}$ are respectively the maximum and minimum observation data, and n are the number of data. Values closer to zero for RMSE and NRMSE and values close to one for NS, WI and $R^2$ coefficients show better performance for the model. NRMSE has 4 qualitative intervals for performing models: 1) NRMSE > 0.3 for poor model performance, 2) 0.2 < NRMSE < 0.3 for average model performance, 3) 0.1 < NRMSE < 0.2 for good model performance and 4) 0 < NRMSE < 0.1 for excellent model performance [26].

## 3. Results

After the field measurement of tree components and data sorting, tree components were evaluated in terms of their correlation with tree weight to select input combinations of the models. Pearson's correlation test performed this in SPSS software. Figure 5 has shown the results.

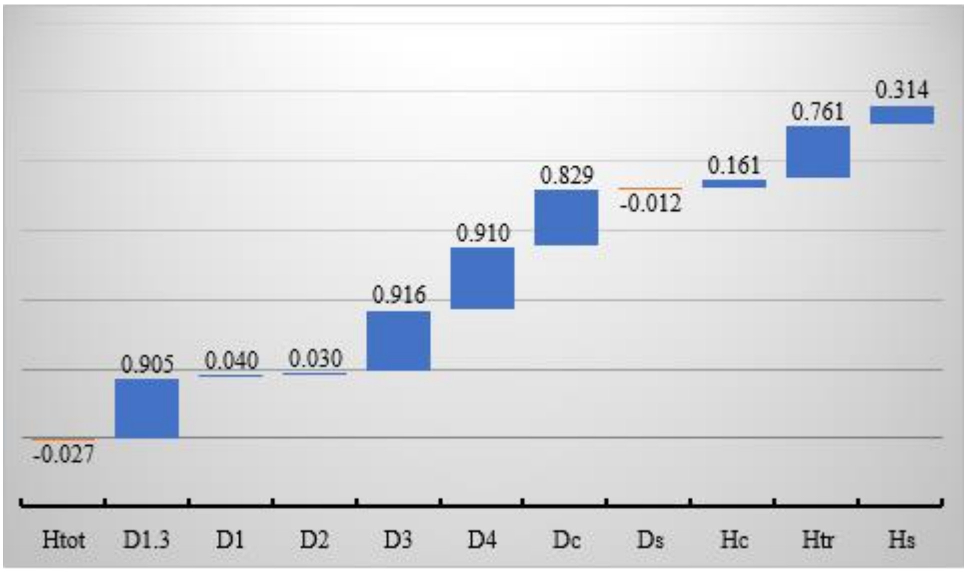


Figure 5. Correlation between the tree's components and their trunk weight

As the evaluations show, the correlation of Htot, D1, D2 and Ds variables with the weight of tree trunks is very weak and non-significant (-0.027, 0.040, 0.030 and -0.012, respectively); so, it is not reasonable to use them as input variables in modeling. Other tree variables, including D1.3, D3, D4, Dc, Hc, Htr and Hs, have a significant correlation at the 0.01 level with the tree trunk weight. The different intensity of correlation of variables with tree trunk weight reveals that Hc and Hs variables have relatively low correlation intensity with tree trunk weight (0.161 and 0.314, respectively), despite their significant correlation. Therefore, the use of these variables as input will definitely reduce the accuracy of models in estimating the weight of tree trunks. Only five variables D1.3, D3, D4, Dc and Htr, with correlations equal to 0.905, 0.916, 0.910, 0.829, and 0.761 are supposedly effective inputs in modeling. Therefore, the input combinations of the models are adjusted through these variables (Table 3).

Power System Technology
ISSN:1000-3673
Received: 16-05-2024 Revised: 12-06-2024 Accepted: 17-07-2024

Table 3. The input scenarios of the models

| Input scenario | Description | Name | Input variables | Target variable |
|---|---|---|---|---|
| All variables (SC1) | All of the related variables considered as the models' inputs | SC1 | $D_{1.3}, D_3, D_4, D_c, H_{tr}$ | W |
| Simplicity (SC2) | Variables that are easier to measure considered as the models' inputs | SC2 | $D_{1.3}, D_c, H_{tr}$ | W |
| Parsimony (SC3) | The least number of variables (two variables) considered as the models' inputs | SC3-1 | $D_c, H_{tr}$ | W |
| | | SC3-2 | $D_{1.3}, H_{tr}$ | W |
| | | SC3-3 | $D_3, H_{tr}$ | W |
| | | SC3-4 | $D_4, H_{tr}$ | W |

Table 3 shows three scenarios under study. The first scenario (SC1) applied all effective input variables as input of the model. The second scenario (SC2) considered variables that were easier to measure in the field. For example, variables such as D3 and D4 (omitted in the SC2 scenario), require a ladder for sampling and cause a waste of time for the researcher. The third scenario (SC3) has parsimony for its basis. This scenario considered only two variables as input of the model so that the model can estimate the weight of the tree trunk based on the minimum effective variables. It includes 4 different binary combinations of effective input variables (SC3-1, SC3-2, SC3-3 and SC3-4) (Table 3).

### 3.1. Modeling trunk weight using multiple linear regression

The MLR model, after adjusting the scenarios and their inputs, was first implemented on the data. Table 4 shows the evaluation results of this model.

Table 4. Evaluation of multiple linear regression

| Input scenario | Calibration phase | | Validation phase | |
|---|---|---|---|---|
| | RMSE (kg) | WI | RMSE (kg) | WI |

Power System Technology
ISSN:1000-3673
Received: 16-05-2024 Revised: 12-06-2024 Accepted: 17-07-2024

| | | | | |
|---|---|---|---|---|
| SC1 | 67.730 | 0.968 | 67.407 | 0.974 |
| SC2 | 67.895 | 0.968 | 67.983 | 0.973 |
| SC3-1 | 97.612 | 0.927 | 108.702 | 0.910 |
| **SC3-2*** | **71.267** | **0.964** | **65.980** | **0.973** |
| SC3-3 | 75.155 | 0.960 | 78.536 | 0.963 |
| SC3-4 | 77.487 | 0.957 | 79.779 | 0.961 |

* The bold row refers to the best performance of the model

The MLR model could show relatively good accuracy in estimating the trunk weight in all the tested scenarios, because WI is reportedly higher than 0.9 in all scenarios (between 0.927 to 0.968 in calibration phase and 0.910 to 0.974 in validation phase). The amount of RMSE in this evaluation shows the model error in different scenarios, which varies between 67.730 to 97.612 kg in the calibration phase, and 65.980 to 108.702 kg in the validation phase. Since the validation phase is the most important period to identify the actual performance of the models, the best performance of the model should refer to this period. Therefore, the SC3-2 scenario has provided the best performance for the MLR model in estimating the weight of tree trunks both among the parsimony scenarios (see Table 3) and compared to the two scenarios SC1 (all efficient variables) and SC2 (simplicity). The following equations are for scenarios SC1, SC2, and SC3 to estimate the trunk weight of poplar trees in this region, and the researchers use these equations and measure the input components required for each of these equations to calculate the trunk weight (W) before tree cutting:

$$W = -606.8 + 42.69\ D1.3 + 12.2\ D3 - 8.44\ D4 - 7.50\ Dc + 13.59\ Htr \quad \text{(Eq. 9)}$$

$$W = -601.1 + 46.49\ D1.3 - 7.58\ Dc + 13.51\ Htr \quad \text{(Eq. 10)}$$

$$W = -522.6 + 34.07\ D1.3 + 13.61\ Htr \quad \text{(Eq. 11)}$$

**3.2. Modeling trunk weight using the artificial intelligence algorithms MLP and MLP-GA**

MLP and MLP-GA models were also implemented regarding the input scenarios the Table 3 describes. It is noteworthy that only SC3-2 scenario is supposedly SC3 scenario (parsimony) for artificial intelligence models because of its remarkable superiority among the four SC3 scenarios (because of modeling by MLR model). As for MLP and MLP-GA models, the hidden layers and the type of transfer function were selected by trial and error. The number of a hidden layer and the transfer function of saturated linear (satlin) had the best compatibility with the data under study in this research. Table 5 gives the evaluation results of these models.

Power System Technology
ISSN:1000-3673
Received: 16-05-2024 Revised: 12-06-2024 Accepted: 17-07-2024

Table 5. Evaluation of the machine learning models MLP and MLP-GA

| Input scenario | Model | Calibration phase | | Validation phase | |
|---|---|---|---|---|---|
| | | RMSE (kg) | WI | RMSE (kg) | WI |
| SC1 | MLP | 59.704 | 0.976 | 84.704 | 0.958 |
| | MLP-GA | 54.090 | 0.980 | 73.205 | 0.969 |
| SC2 | MLP | 59.526 | 0.976 | 85.060 | 0.956 |
| | MLP-GA | 56.240 | 0.979 | 72.942 | 0.968 |
| SC3 | MLP | 67.916 | 0.968 | 82.052 | 0.960 |
| | **MLP-GA*** | **59.624** | **0.976** | **69.810** | **0.971** |

* The bold rows refer to the best performance of the model

The MLP and MLP-GA models have relatively good accuracy in estimating trunk weight in all three scenarios, because WI is reportedly higher than 0.9 in all scenarios. The amount of RMSE in the two models reveals the superiority of the MLP-GA model in all scenarios. This superiority for SC1, SC2 and SC3 scenarios is respectively 10.4%, 5.8% and 13.9% in the calibration phase and 15.7%, 16.6% and 17.5% in the validation phase. The best performance of both AI models was achieved in the parsimony input scenario (SC3). The validation of these two models reveals that the weakest performance in trunk weight estimation belongs to the MLP model in the SC2 scenario with RMSE = 85.060 kg and WI = 0.956, and the most accurate performance belongs to the MLP-GA model in the SC3 scenario, with RMSE = 69.810 kg. and WI = 0.971.

**3.3. Comparison of the models' performances**

This section compares graphically the performance of the models. These comparisons are for the validation phase. First, the estimations of the models in the three input scenarios were drawn as scatter plots against the observational values of tree trunk weight (Figure 6).

Power System Technology
ISSN:1000-3673

Received: 16-05-2024 Revised: 12-06-2024 Accepted: 17-07-2024

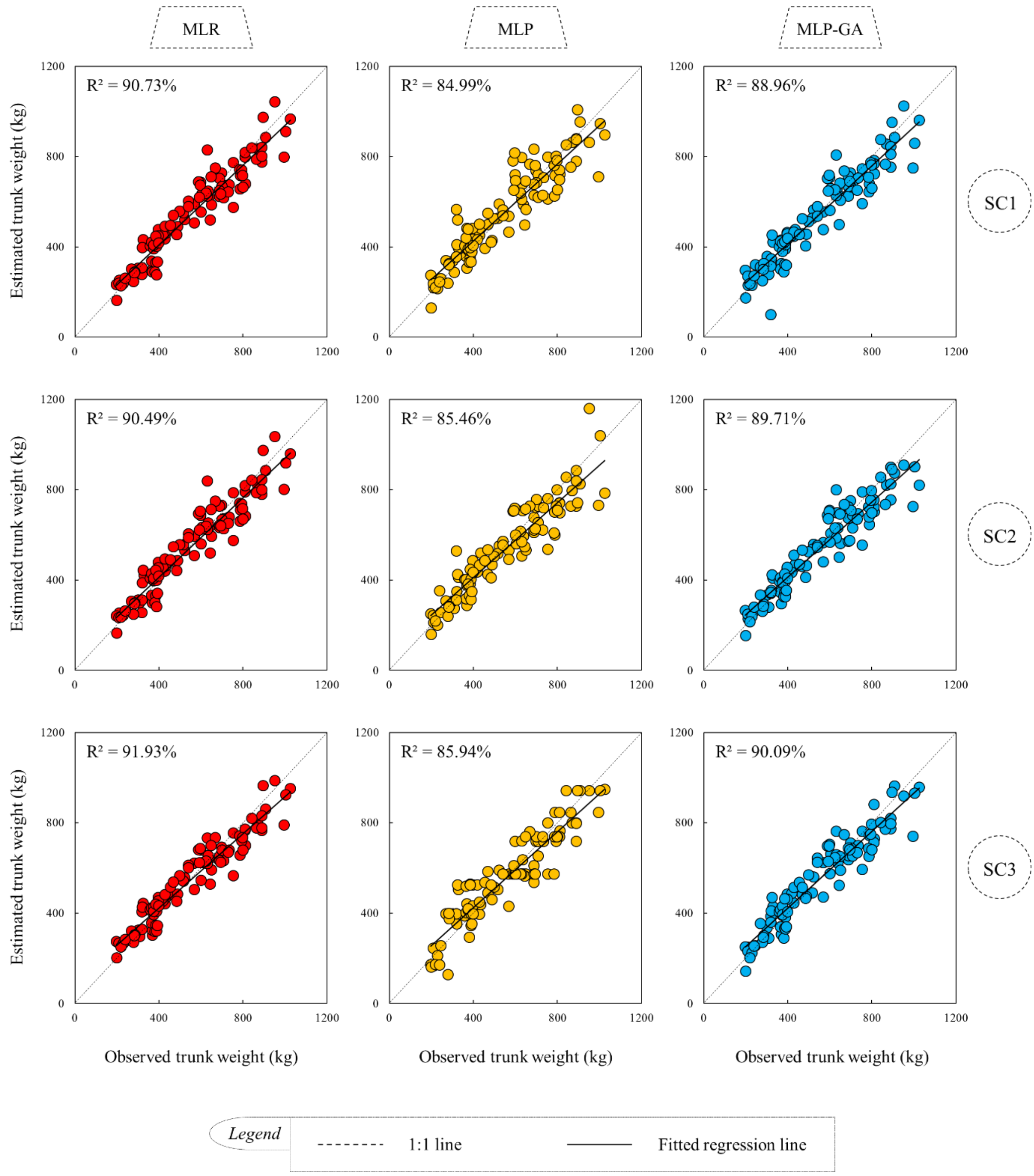


Figure 6. Regression plots of the observational trunk weight and the models' estimations in different scenarios

As Figure 6 shows, the fitted regression lines between estimates and observations have a tiny slope compared to the X=Y line (1:1 line); it reveals the acceptable performance of all three models in all three input scenarios in estimating the weight of tree trunks. This slope in the SC3 scenario is lower than the other two scenarios. The regression line in all scatter plots is slightly lower than the 1:1 line in high values, and higher than the 1:1 line in low values. Thus, the models underestimated the weight of heavier trees slightly, and overestimated slightly the

Power System Technology
ISSN:1000-3673
Received: 16-05-2024 Revised: 12-06-2024 Accepted: 17-07-2024

weight of lighter trees. The points of the diagrams have a significant concentration around the axis of the fitted regression line, which shows a high correlation between estimates and observations. The weakest concentration of points around the regression line is observable in the estimates of the MLP model, and the highest concentrations in the outputs of the MLR model. The $R^2$ coefficient of all three models is the highest in the SC3 scenario, the highest of which belongs to the MLR model with $R^2 = 91.93\%$. The lowest $R^2$ also belongs to the MLP model in the SC1 scenario, which is reportedly 84.99%. Taylor diagram is usable to examine the performance of the models in each scenario separately. Taylor diagram as a method compares and examines simultaneously in a graphic form the several simulated series of a target variable. This diagram can simultaneously evaluate estimates and observations from the point of view of error, correlation, and standard deviation [27] (Figure 7).

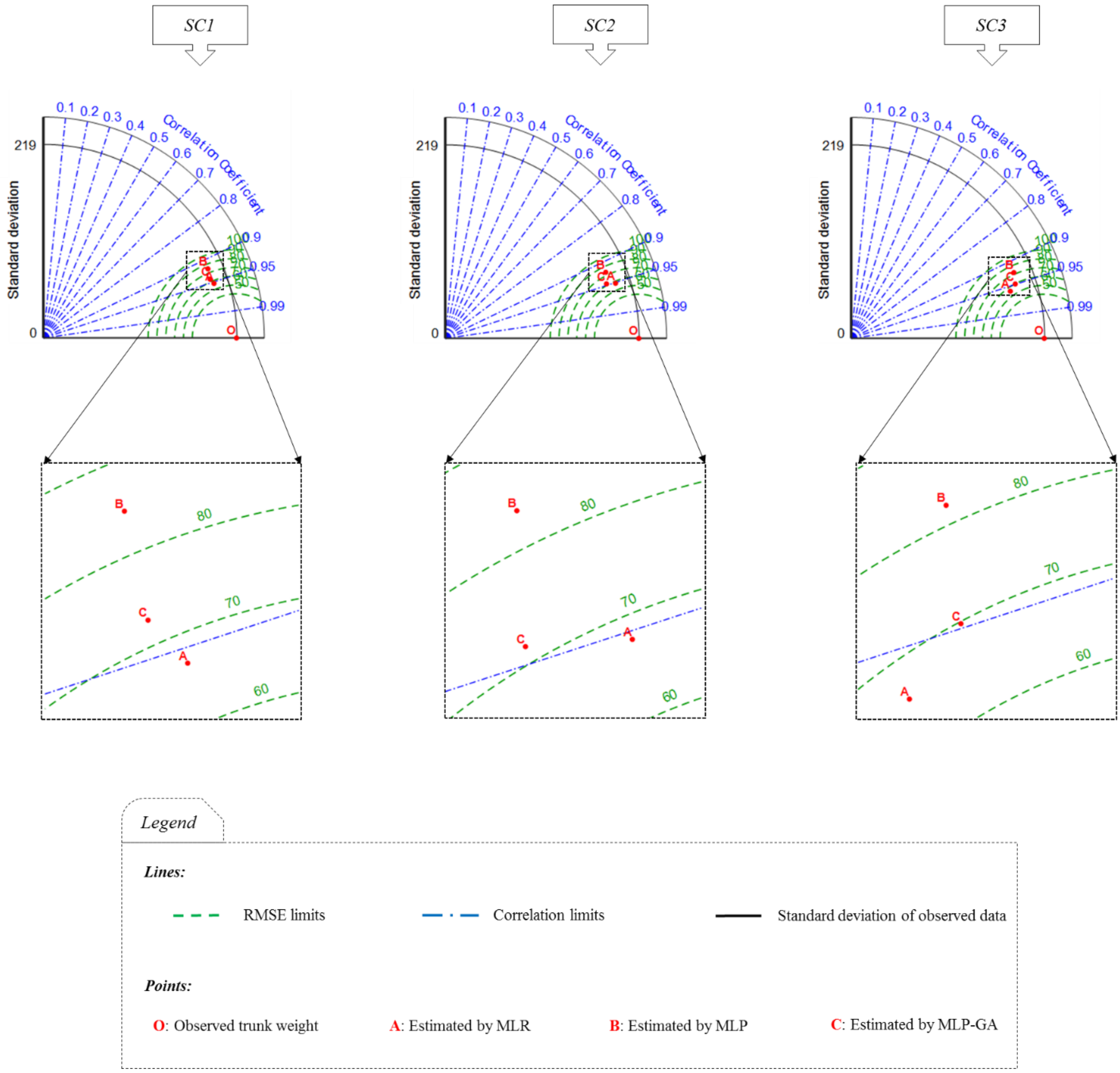


Figure 7. Taylor diagrams to evaluate the models' performances in different scenarios

Power System Technology
ISSN:1000-3673
Received: 16-05-2024 Revised: 12-06-2024 Accepted: 17-07-2024

A first look at Figure 7 can reveal that in all three scenarios, all three models negligibly underestimated the standard deviation of observational data. The point associated with the MLR model estimates (point A) is closer to the observations (point O) in all three scenarios, and the point associated with the MLP estimates (point B) is the farthest from point O in all three scenarios (above the dotted line RMSE = 80 kg). This means that in all three scenarios, the MLP model has the weakest performance, and the MLR model has the most accurate performance. The point of the MLP-GA model (point c) in the SC1 and SC2 scenarios is between the two dotted lines of RMSE = 70 kg and RMSE = 80 kg, but this point in the SC3 scenario is below the RMSE = 70 kg and almost coincides with the line of R = 0.95. This shows the better compatibility of this model with the SC3 scenario, which, of course, is like the MLR and MLP models. Thus, point A in the SC3 scenario, which is the closest to the dotted line of RMSE = 60kg and the line of R = 0.99, shows the best estimate among the models and scenarios. We examine qualitatively below the performance of the models based on the two criteria of NRMSE and NS (Figure 8).

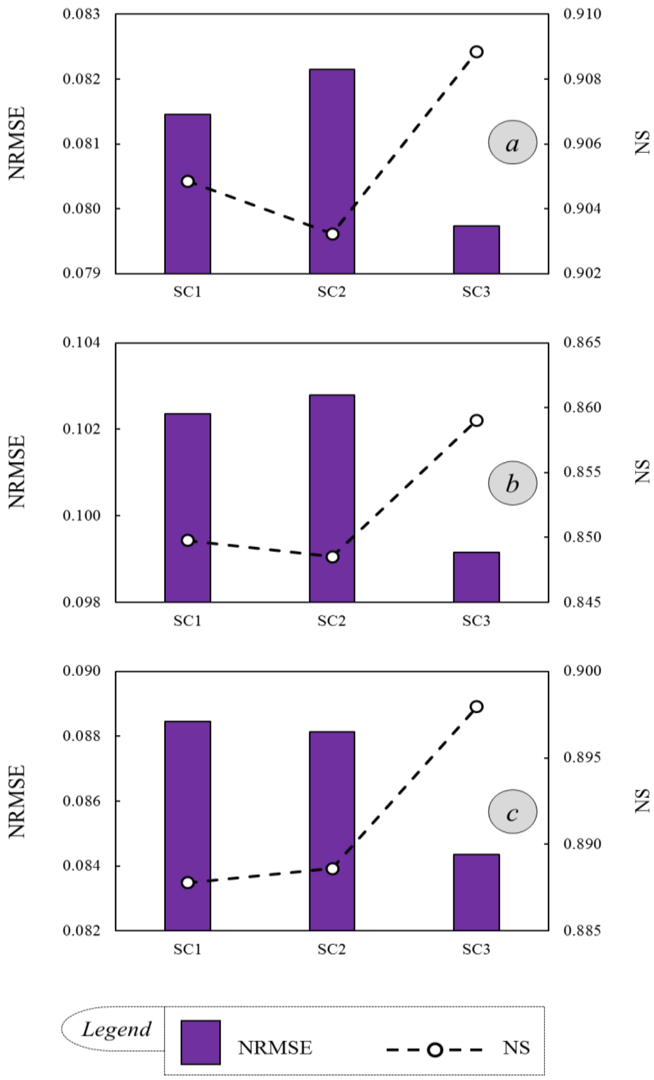


Figure 8. Combo-graphs for qualitative comparison between the models in different scenarios based on the criteria NS and NRMSE (a: MLR b: MLP; c: MLP-GA)

Power System Technology
ISSN:1000-3673

Received: 16-05-2024 Revised: 12-06-2024 Accepted: 17-07-2024

The NS value in all diagrams of Figure 8 is higher than 0.84. This reveals that all three models have performed very well in each input scenario. The amount of NS in the MLR model (Figure 8-a) is between 0.903 and 0.909, which shows the relative superiority of this model compared to the models MLP (0.848 < NS < 0.861) and MLP-GA (0.889 < NS < 0.897). The amount of NRMSE qualitatively has four ranges, which are NRMSE > 0.3, 0.2 < NRMSE < 0.3, 0.1 < NRMSE < 0.2, and NRMSE < 0.1, based on which the performance quality of the models is poor, moderate, good, and excellent [28]. The amount of NRMSE of the MLP model, in the current research, in the two input scenarios of SC1 and SC2 is between 0.1-0.2, which is qualitatively good. NRMSE in other models and scenarios is smaller than 0.1, which shows that the used models have an excellent performance in estimating the weight of tree trunks. Figure 9 shows the distribution of the errors of the models in estimating the trunk weight as evaluated by violin plots.

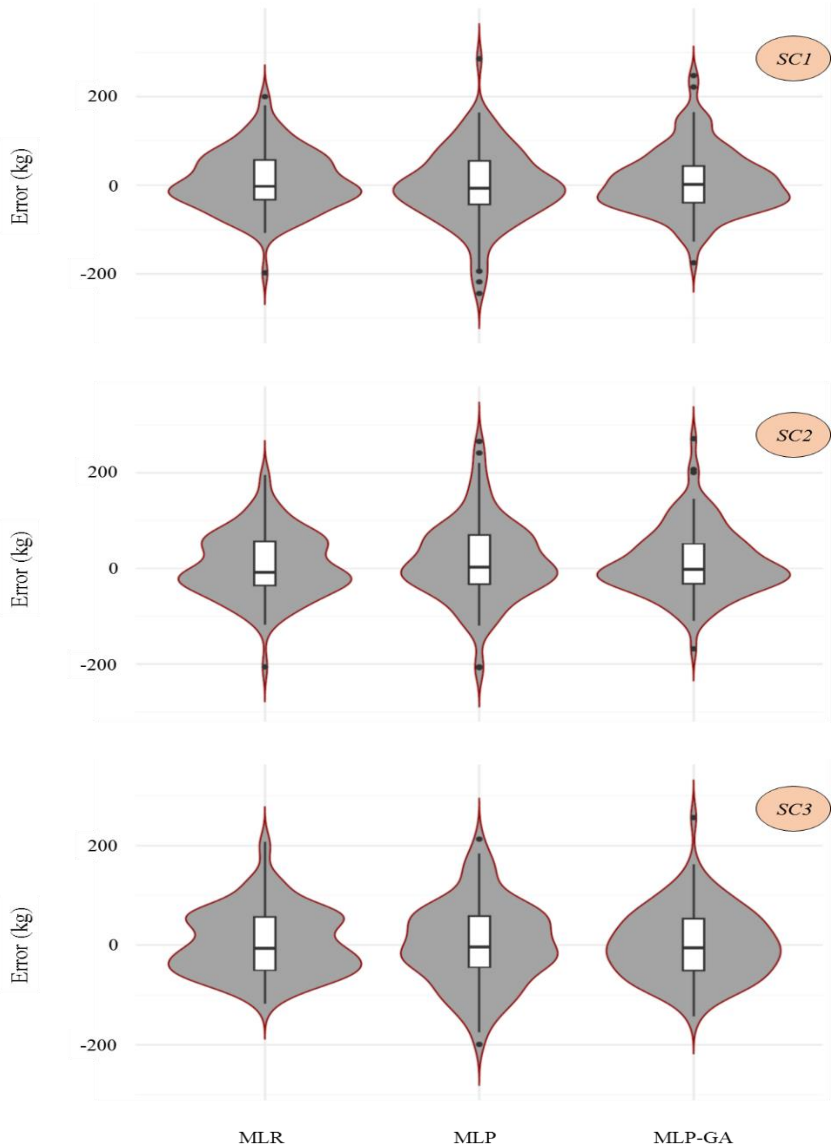


Figure 9. Violin plots to investigate the models’ distribution

Power System Technology
ISSN:1000-3673

Received: 16-05-2024 Revised: 12-06-2024 Accepted: 17-07-2024

As violin plots reveal (Figure 9), the greater the elongation of the violins along the vertical axis, the greater the dispersion of errors and more errors far from zero. In all three input scenarios, this issue is more visible in violins of the MLP model. The reverse of this issue shows more errors close to zero and shows a better performance of the model. The diagrams show that this occurred more for MLR and MLP-GA models, whose violins have more protrusion along the axis of Error = 0 kg. All three scenarios reveal that the MLP-GA model has an outlier above the Error = 200 kg line in its upper tail. Thus, the MLP-GA model had significant underestimations sometimes, what was seen in the MLR model in only one scenario (SC1) and confirm its superior performance over MLP-GA. Finally, the MLR model in the SC3 scenario has more errors around the axis of Error = 0 kg and it is the only function without outliers. We can introduce this as the best performance presented in this research.

## 4. Discussion

The MLP model in estimating tree trunk volume in Tehran, Iran (cold dry climate) and got $R^2$ = 0.86. This research, like ours (R2 ≈ 0.85), shows the effectiveness of the MLP model in estimating the economic components of the tree; but [10.11] aimed at modeling the tree trunk volume with the samples of 100 trees and did not specify the tree species. In modeling the trunk volume of black pine trees in Crimea by [29] the MLP model achieved $R^2$ = 0.95, which is in line with the current research (R2 ≈ 0.85), and shows the efficiency of this model in estimating the economic components of the tree. The current research was about the weight estimation, but the research of was on volume estimation[30]. The tree species under study and the climatic conditions of the regions are also different in the two researches, which can be the reasons for this slight difference in performing the MLP model. However, there was no research that tested the performance of numerical models in estimating the economic weight of tree trunks. Thus, the aforementioned researches were the closest ones that are comparable with the results of the current study.

## 5. Conclusion

The current research carried out the numerical modeling of the trunk weight of poplar trees. It examined 400 Populus Deltoids and measured 11 other variables from each tree, which were used as input of the models. Examinations showed that variables D1.3, D3 and D4 are the most effective variables in estimating the weight of tree trunks. The research compared three models of MLR, MLP and MLP-GA, and the linear model of MLR was superior to the two artificial intelligence models of MLP and MLP-GA. This superiority shows that the linear relationship between the tree variables and the trunk weight variable is stronger than the non-linear relationship between them. As the comparison of performing MLP and MLP-GA models shows, the combination of GA with MLP model can increase the accuracy of the MLP model in estimating tree trunk weight by 16.6%. Our suggestion is: testing the combined form of MLP model with other evolutionary optimization algorithms such as firefly, particle swarm, ant colony, etc. to increase the accuracy of estimations. Other artificial intelligence models such as adaptive neuro-fuzzy inference system (ANFIS), support vector machine (SVM), group method of data handling (GMDH) and radial basis function (RBF) can also be evaluated in estimating the weight of tree trunks. We suggest their simple and hybrid form to future researchers. It is noteworthy that the current study was conducted in the temperate and humid region of Gilan and on afforested fir trees, which has valid results in this region. Therefore, it has research value for other similar and different climatic regions and other species of economic trees.

Power System Technology
ISSN:1000-3673
Received: 16-05-2024 Revised: 12-06-2024 Accepted: 17-07-2024

**Ethics approval**

This research does not include biological and human data. The data of this research does not belong to any organization and was measured in a field survey of a poplar garden under the supervision of Islamic Azad University, Lahijan branch.

**Consent to Participate**

All the authors of this study declare their free and informed consent to participate in this manuscript and submit it in this journal.

**Consent to Publish**

The authors agree that the datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

**Authors contributions**

Conceptualization [Sina Pourrajabali] and [Vahid Hemmati]; Methodology [Sina Pourrajabali]; Soft-ware [Sina Pourrajabali]; Data curation [Sina Pourrajabali]; Writing-Original draft preparation [Sina Pourrajabali], [Vahid Hemmati] and [Alireza Eslami]; Investigation [Sina Pourrajabali] and [Vahid Hemmati]; Validation [Sina Pourrajabali]; Visualization [Sina Pourrajabali]; Writing-Reviewing and Editing [Sina Pourrajabali], [Vahid Hemmati], [Alireza Eslami], [Seyed Armin Hashemi] and [Seyed Yousef Torabian]; Discussion [Sina Pourrajabali]; Supervision [Vahid Hemmati] and [Alireza Eslami]. All authors have read and agreed to the published version of the manuscript.

**Competing Interests**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Funding**

There is no institutional/governmental funding received for this study.

**References**

[1]Malik, A., Tikhamarine, Y., Sammen, S. S., Abba, S. I., & Shahid, S. (2021). Prediction of meteorological drought by using hybrid support vector regression optimized with HHO versus PSO algorithms. Environmental Science and Pollution Research, 28, 39139-39158.

[2]Poplars, F. A. O. (2016). Other Fast-Growing Trees—Renewable Resources for Future Green Economies. Synthesis of Country Progress Reports. Proceedings of the 25th Session of the International Poplar Commission, Berlin, Germany, 13-16.

[3]Guo, X. Y., & Zhang, X. S. (2010). Performance of 14 hybrid poplar clones grown in Beijing, China. biomass and bioenergy, 34(6), 906-911.

[4]Chavan, S. B., & Dhillon, R. S. (2019). Doubling farmers' income through Populus deltoides-based agroforestry systems in northwestern India. Current Science, 117(2), 219-226.

[5]Cleugh, H.A., 2004 Effects of Windbreak on Air Flow, Micro Climate & Crop Yield. USA: Springer Link, Inc.

Power System Technology
ISSN:1000-3673
Received: 16-05-2024 Revised: 12-06-2024 Accepted: 17-07-2024

[6]Woodruff, N. P., & Zingg, A. W. (1952). Wind-tunnel studies of fundamental problems related to windbreaks (Vol. 112). US Department of Agriculture, Soil Conservation Service.
[7]Lefroy, E. (1991). Trees and shrubs as sources of fodder in Western Australia. In The role of trees in sustainable agriculture conference, Albury NSW. September (pp. 19-32).
[8]Sirohi, C., Dhillon, R. S., Chavan, S. B., Handa, A. K., Balyan, P., Bhardwaj, K. K., ... &
[9]Fadhillah, M. F., Lee, S., Lee, C. W., & Park, Y. C. (2021). Application of support vector regression and metaheuristic optimization algorithms for groundwater potential mapping in Gangneung-si, South Korea. Remote Sensing, 13(6), 1196.
[10]Bayat, M. (2017). Predicting and assessing the tree species survival and determining Physiographic factors affecting on it in Mazandaran province Forests using artificial neural networks. Journal of Natural Environment, 70(4), 771-782.
[11]Bayati, H., & Najafi, A. (2011). Application of artificial intelligence in trees stem volume estimation.
[12]Ghanbari, F., SHATAEI, J. S., DEHGHANI, A. A., & AYOUBI, S. A. (2009). Tree density estimation of forests by terrain analysis and artificial neural network.
[13]Peyrov, S., Najafi, A., & Alavi, S. J. (2014). Prediction of forest roadway using artificial neural network and multiple linear regressions. Forest Sustainable Development, 1(3), 285-296.
[14]Gorzin, F., Namiranian, M., Omid, M., & Bayat, M. (2017). Prediction commercial and cordwood volume of broadleaves using Artificial Neural Networks (Case study: Gorazbon distric of Kheyrood forest, Nowshahr). Iranian Journal of Forest and Poplar Research, 25(3), 517-527.
[15]Mosaffaei, Z., & Jahani, A. (2021). Modeling of ash (Fraxinus excelsior) bark thickness in urban forests using artificial neural network (ANN) and regression models. Modeling Earth Systems and Environment, 7(3), 1443-1452.
[16]Khosravi, A., Nunes, R. O., Assad, M. E. H., & Machado, L. (2018). Comparison of artificial intelligence methods in estimation of daily global solar radiation. Journal of cleaner production, 194, 342-358.
[17]Mohammadi, B., & Aghashariatmadari, Z. (2020). Estimation of solar radiation using neighboring stations through hybrid support vector regression boosted by Krill Herd algorithm. Arabian Journal of Geosciences, 13, 1-16.
[18]Aghelpour, P., & Varshavian, V. (2020). Evaluation of stochastic and artificial intelligence models in modeling and predicting of river daily flow time series. Stochastic Environmental Research and Risk Assessment, 34(1), 33-50.
[19]Aghelpour, P., & Varshavian, V. (2021). Forecasting different types of droughts simultaneously using multivariate standardized precipitation index (MSPI), MLP neural network, and imperialistic competitive algorithm (ICA). Complexity, 2021, 1-16.
[20]Ahmadi, F., Mehdizadeh, S., Mohammadi, B., Pham, Q. B., Doan, T. N. C., & Vo, N. D. (2021). Application of an artificial intelligence technique enhanced with intelligent water drops for monthly reference evapotranspiration estimation. Agricultural Water Management, 244, 106622.
[21]Mohammadi, B., & Mehdizadeh, S. (2020). Modeling daily reference evapotranspiration via a novel approach based on support vector regression coupled with whale optimization algorithm. Agricultural Water Management, 237, 106145.
[22]Ober, P. B. (2013). Introduction to linear regression analysis.

Power System Technology
ISSN:1000-3673

Received: 16-05-2024 Revised: 12-06-2024 Accepted: 17-07-2024

[23]Howard, D., & Mark, B. (2001). User's guide for neural network toolbox with Matlab. Version, 6, 1.

[24]Asl, S. F. Z., Karami, A., Ashari, G., Behrang, A., Assareh, A., & Hedayat, N. (2011). Daily global solar radiation modeling using multi-layer perceptron (MLP) neural networks. International Journal of Environmental and Ecological Engineering, 5(7), 403-405.

[25]Jaramillo, J. H., Bhadury, J., & Batta, R. (2002). On the use of genetic algorithms to solve location problems. Computers & Operations Research, 29(6), 761-779.

[26]Aghelpour, P., Bahrami-Pichaghchi, H., Varshavian, V., Graf, R., Gholami Sefidkouhi, M. A., & Khoshravesh, M. (2022a). Evaluating the impact of large-scale climatic indices as inputs for forecasting monthly river flow in Mazandaran Province, Iran. Pure and Applied Geophysics, 179(4), 1309-1331.

[27]Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. Journal of geophysical research: atmospheres, 106(D7), 7183-7192.

[28]Bahrami-Pichaghchi, H., & Aghelpour, P. (2023). An estimation and multi-step ahead prediction study of monthly snow cover area, based on efficient atmospheric-oceanic dynamics. Climate Dynamics, 60(3-4), 743-765.

[29]Şenyurt, M., & Ercanli, I. (2019). A comparison of Artificial Neural Network models and regression models to predict tree volumes for Crimean Black Pine trees in Cankiri Forests. Journal of Forestry Society of Croatia/Sumarski List Hrvatskoga Sumarskoga Drustva, 143.