Journal of Medicinal Plants and By-products (2024) 1: 123-135
https://doi.org/10.22034/jmpb.2023.355310.1384

Original Article

The Effect of Soil Physical and Chemical Properties on the Performance Indices of Artichoke's Leaf using Artificial Neural Network (ANN)

Azadeh Alizadeh[1], Azim Ghasemnezhad[1]*, Aboutaleb Hezarjaribi[1], Mohammad Zaman Aladdin[2]

[1]*Faculty of Plant Production, Gorgan University of Agriculture and Natural Resources, Gorgan, Iran*
[2]*Department of soil Sciences, Gorgan University of Agricultural Sciences and Natural Resources*

Article History

Received: 20 July 2021
Accepted: 18 February 2023

Keywords
Cynara scolymus
Easily Accessible Soil Parameters
Performance

***Corresponding author**
Ghasemnezhad@gau.ac.ir

ABSTRACT

The present study aims to estimate the performance of artichoke via physic-chemical parameters of soil including soil texture, pH, and bulk density using the artificial neural network (ANN) method. Thus, the soils of sixty points across croplands and forests of Golestan province, Iran were sampled, and soil parameters were measured in the lab. Based on the obtained parameters the different models were performed. The experiment was conducted as a randomized complete block design with three replications. The results showed that ANN models were more efficient than the multivariate regression models (MR model). All ANN models were better to estimate plant weight performance compared with the MR model. Plants grown in the soil samples of the "Ahangar Mahalleh area" showed the highest level of yield performance. Based on the findings, model number 5 with a minimum input parameter was selected as an optimal model. All ANN models were better than the multivariate regression models in the estimation of plant weight. As model 5 had almost similar performance with a minimal number of inputs compared with the other models, this model can be selected as the best model.

INTRODUCTION

Artichoke (*Cynara scolymus* L.) of the compositae family is an upright columnar perennial, which grows up to 1.5 m. This plant has large jagged leaves, with a light green adaxial surface. The abaxial surface is white and opaque due to the presence of trichomes [1]. Capitol is large with blue-violet tubular blossoms and is surrounded by blunt husk leaves. Involucral bracts are broad and contain nutrient reserves. Seeds are light brown and small with dark brown reticulates.

An artificial neural network (ANN) is an imitation of the human nervous system. This network tries to develop a brain-like structure able to learn, generalize and decide [2]. The objectives of these structures are to introduce a dynamic system to teach the model, save the model mechanism in its memory and use it for instances which the model has not faced before. The application of these models in Iran and agriculture research is relatively new. However, due to their ability to model intricate processes, it is possible to use them widely in agricultural sciences. Prediction of time and place of rainfall [3], rainfed wheat performance [4], evapotranspiration [5] and CO_2 flux in the ecosystem are among many applications of ANN in agricultural sciences. To this day, no studies are available on the application of ANN in the determination of input parameters required to simulate the effect of the physic-chemical properties of soil on the qualitative parameters of artichoke leaves. The present research was carried out to determine the minimum number of input parameters affecting the qualitative properties of artichoke leaves and the estimation of easily accessible soil parameters using ANN in Golestan province, Iran [6]. Keravner and Rosh [7] used ANN and linear regression to predict the amino acid content in cattle feed and reported that ANN described the relationship between amino acids and other nutritional elements more efficiently.

The objective of this study was to determine the easily accessible soil parameters required to estimate antioxidant content and performance of artichoke, assess the efficiency of easily accessible soil parameters on artichoke performance using ANN and determine the best model to predict the easily accessible soil parameters affecting artichoke performance.

MATERIALS AND METHODS

The present research was carried out in 2012 at the research greenhouse of Gorgan University of Agricultural Sciences and Natural Resources, Iran. The experiment was conducted as a randomized complete block design with three replications in pots. The treatment was the soils collected from 60 different regions of Golestan province (Table 1), Iran, which were used to fill the pots. The soils were collected from a depth of 30 cm and were transferred to the lab to determine their physic-chemical properties (Table 2). The measured soil parameters included soil texture, organic carbon percentage, neutralizing material percentage (lime percentage), pH, EC, CEC, NPK and bulk density [8]. Page *et al*. [9] method was used to measure organic carbon, EC, pH, and soil texture was measured by hydrometer method [10].

Table 1 The statistical description of chemical properties of the soils

Parameter	Mean±SD
Organic carbon (OC)	1.21±0.05
TNV%	15±0.69
CEC	10.85±0.26
EC	0.47±0.04
pH	7.67±0.04
N%	0.12±0.09
P%	16.7±1.04
K%	303.47±40.0

Table 2 The statistical description of physical properties of the soils

Parameter	Mean	CV%	Skewness
Sand%	32.05	0.43	0.55
Silt%	44.27	0.31	0.17
Clay%	23.33	0.38	-0.35
BD	1.66	0.16	-1.38

Measurement of nitrogen percentage was carried out by ammonia nitrogen (NH_4^+) and nitrate nitrogen (NO_3^-). Ammonia nitrogen was measured using Bremner and Mulvaney [11] method, and Page *et al*. [9] method was used to measure soil sodium and potassium.

Preparation of Plant Samples

The pots were filled with each of the collected soils and perlite (10:1) with three replications. Plastic pots with 35 cm height × 20 cm diameter were used in this study. The seeds of the artichoke were provided by the horticulture laboratory, department of horticulture, Gorgan University of Agricultural Sciences and Natural Resources and Natural Resources, Golestan Province, Iran. The pots were first irrigated to field capacity; then two seeds were sown in each pot. A two mm layer of perlite then covered the sown seeds. The seeds emerged seven days later. The plants were thinned at the 4-leaf stage so that only one healthy plant was present in each pot. Hand weeding and irrigation were carried out throughout the growing season until harvest [8], which was 120 days.

Measurement of morphological attributes and performance components

The plants were sampled 120 days after sowing. The number of healthy and unhealthy (chlorotic, necrotic and infested) leaves, plant weight and height and root length were measured.

Harvest

The plants were harvested when the leaf margins changed from smooth and thorn-free conditions to a jagged form in approximately 98 percent of the plants. The leaves were separated from shoots after a preliminary air-drying and were then placed for 48 h at 45 °C in an oven, and were eventually ground [8].

Modeling with ANN

The development of an artificial neural network requires devising its technical components. To fulfill the objectives of the present study, ANNs with various structures such as perceptron was used to select the best and most efficient network by calculating their errors [12]. In addition, a sensitivity analysis was done to determine performance-affecting factors. Finally, the coefficient of determination (R^2) and root mean square error (RMSE) was used to determine the best model. Easily accessible soil parameters (soil texture, organic carbon percentage, neutralizing material percentage, pH, EC, CEC, NPK and bulk density) were considered as inputs and performance (plant weight) was the output of the model.

Data Standardization

Entering the raw data may lead to lower network speed and precision. To avoid such conditions and uniform the value of the data, the data should be standardized before neural network training. This will prevent the excessive reduction of weights [12]. In addition, by regulating the entry data in a specific range, neurons will be placed in an optimum range and protect them against early-saturation. Furthermore, the data will be converted to values between 0 and 1, as the output of numerous functions are between 0 and 1 and conversion of the data plays a crucial role in network training. The changes in the weight of neurons will be minimal for the values close to 0, whereas, for the values close to 0.5, the rate of neuron response to signals will be more rapid. The following formula was used to standardize the data:

$$X_n = 0.5 + 0.5\left(\frac{X - X_{mean}}{X_{max} - X_{min}}\right) \quad (1)$$

Where X_n is the normalized data, X is the observed data, X_{mean}, X_{min} and X_{max} are the average, minimum and maximum observed data, respectively.

Data classification

Artificial neural networks require a series of input and output data for development and training so they can deduct their non-linear relationships by a logical analysis of this data and simulate similar cases. Artificial neural network models require three data sets training, validation and testing. Training data are used to find a relationship between the input and output observed data. Validation data are used to control and monitor proper training of the network, and test data are used to evaluate the performance of the suggested network. In this study, 60, 20 and 20% of data were allocated to training, validation and testing of the model.

Network design

Clay, silt and sand percentage, organic carbon percentage, soil pH, salinity, NPK, CEC and bulk density were considered as input parameters and the logarithm of plant weight (performance) was considered as output. Then, 60 (60 soil samples), 20 (12 soil parameters) and 20 (12 soil parameters) percent of the data were chosen as training, validation and test, respectively. Matlab software V. 7.9 and MLP network were used to train the ANN. The training process, which includes changes in the different layers throughout the training period- was done to minimize the difference between the observed (for testing) and predicted data. Levenberg-Marquardt algorithm and a hidden layer with the Logsig threshold function and Tansig threshold function for the output layer were used during the training process. Finally, the best network structures for performance and antioxidant content parameters were determined based on the lowest RMSE and the highest R^2.

Evaluation of model precision

In the present study, R^2 and RMSE of the observed and predicted data were utilized to evaluate the model performance. The mathematical expression of RMSE is as follows:

$$RMSE = \frac{\sqrt{\sum(t-a)^2}}{N} \quad (2)$$

In which a and t are the observed and predicted values of performance and N is the number of data. The value of RMSE indicates the extent that the model has predicted the measurements lower or higher than the actual. In case the predicted and observed values are equal, RMSE=0. The coefficient of determination is obtained by fitting a line to the predicted vs observed data scatter graph.

Sensitivity Analysis

Sensitivity analysis provides the modeler with useful information about the sensitivity of the model to input variables. Less effective variables may be omitted from the network by determining the influence extent of the input variables on the model precision and thus, develop a simpler model. In other words, sensitivity analysis is used to determine which of the 12 parameters (Clay, sand and silt percentage, organic carbon percentage,) has a higher impact on the performance, and its changes are more sensitive. In this study, sensitivity coefficient (dimensionless) was used [13] as follows: First, 12 parameter model (without changes in the inputs) was entered into the model and the output was obtained (control). Then, one of the variables was changed by 10% and entered into the model while the others were constant and the output was calculated. The difference between these two outputs (control and the changed output) was calculated using equation (3):

$\delta\hat{y}_i = \hat{y}(i+0.1) - \hat{y}$ (3)

Where $\hat{y}(i+0.1)$ is the output with one of the variables changed by 10%, and $\hat{y}_i$ is the control output (without changes). Sensitivity coefficient was calculated using equation (4), which represents the extent of model sensitivity to the parameter in the j^th^ observation:

$$ss_{ij} = \left(\frac{\delta\hat{y}_i}{\delta\beta_i}\right)\beta_j \quad (4)$$

Where j represents the j^{th} input variable.

$$\delta\beta_j = 0.1\beta_j \quad (5)$$

$\delta\beta_j$ is the changed output which is calculated using equation (5) (the variables changed by 10% in this study). These steps will be done for each variable, i.e. each variable will be changed by 10% while the others are kept constant. Composite-scaled sensitivity (CSS) coefficient was used to calculate model sensitivity for all observations.

Hill [13] defined this coefficient for the j^{th} parameter as follows:

$$CSS_i = \left(\frac{1}{N}\sum_{i=1}^{N} ss_{ij}\right)^{\frac{1}{2}} \quad (6)$$

Equation (7) is in fact, the mean sensitivity coefficient for each variable. To simplify the comparison of CSS values for different variables, relative CSS values called relative sensitivity coefficient (γ) was used as follows:

$$\frac{CSS_i}{max(CSS)} \quad (7)$$

In which max (CSS) is the maximum value for CSS for all variables entered into the model. The maximum value is one and is related to the parameter with the maximum CSS.

Statistical Analysis

SPSS 20 statistical software (IBM Corp., Armonk, NY, USA) was used for mean comparison. In addition, MATLAB 7.00 software (The MathWorks, Inc.) was used for the estimation of performance by the leaf weight of artichoke using ANN.

RESULTS AND DISCUTION

Description of Variables

The statistical description of the physico-chemical properties of soil is presented in Tables 1 and 2. As may be observed in Table 1, pH has the lowest coefficient of variation (CV) (0.038%) among the physico-chemical properties, whereas the CV of P is the highest with 1.04%. The values of the skewness coefficient presented in Tables 1 and 2 indicate that except for lime, K, sand, silt, clay and BD that have normal distributions and skewness values of -1 to +1, other parameters have Log-normal distribution.

Artificial Neural Network Modelling

In this section, the results obtained from the best ANN structure with 12 parameters for 60 soil samples, as well as the results of the sensitivity analysis, will be presented. In this study, the perceptron network and multiple layers were used, and transition functions in the hidden and output layers and the number of neurons in the hidden layer were tested. Finally, the best network structure for plant weight was obtained by trial and error. Selection of the best network for the prediction of plant weight was made based on the lowest RMSE and the highest R^2. The best arrangement of the hidden layer with the Levenberg-Marquardt training algorithm was on a hidden layer, 34 neurons and the Logsig threshold function for the hidden layer and the Tansig function for the output layer. Tables 3 and 4 present the estimated parameters for the model in the training, validation, test and total for plant weight, respectively. Useful information on the model performance may be deducted from the slope of the line fitted to the predicted vs observed data graph. If the slope is 1, the predicted and observed values are equal. Figure 1 shows R^2 and the function of the line fitted to the predicted vs observed plant weight in the training and testing stages. As shown in Figure 1, the slope of the fitted line for the plant weight data is 0.99, which indicates that the predicted and observed values are close to each other. Since R^2 in training and validation stages are 0.92 and 0.88, respectively, the estimation of the plant weight by the model has good precision, as shown in Figure 3.

In this step, a very precise model was obtained to estimate the plant weight using the 12 easily accessible soil parameters. However, this study aimed to rapidly and easily estimate plant weight and antioxidant content using easily accessible soil parameters. Thus, the application of this model contradicts its objectives, as the measurement of these 12 parameters is costly and time-consuming. Thus, it is possible to determine the parameters sensitive to plant performance using the sensitivity analysis and use the minimal parameters as input to estimate plant weight.

Table 3 Estimated statistical parameters for the stages of training, validation, testing and total in the 12-parameter model for single plant weight (g)

Level	R^2	RMSE
Training	0.92	0.030
Validation	0.88	0.040
Test	0.86	0.041
Total	0.89	0.035

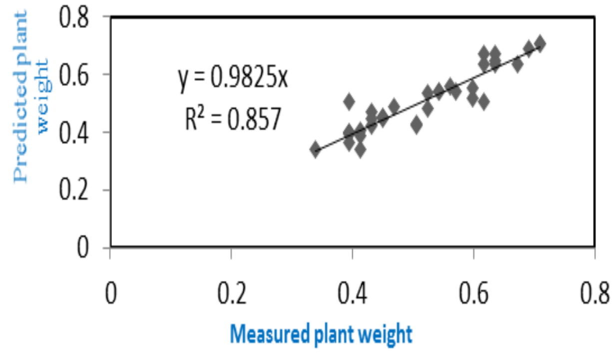

Fig. 1 Line fitted to predicted versus measured data of plant weight in the training stage in the 12-parameter model

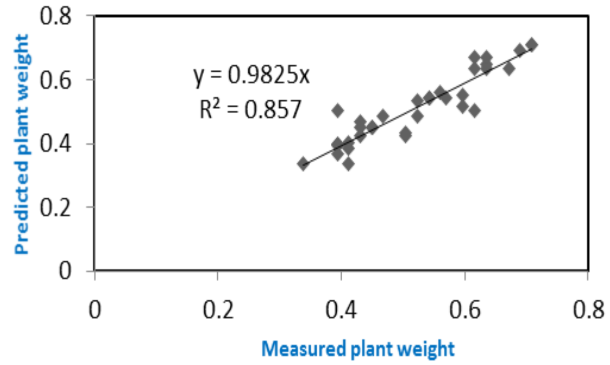

Fig. 2 Line fitted to predicted versus measured data of plant weight in testing stage in the 12-parameter model

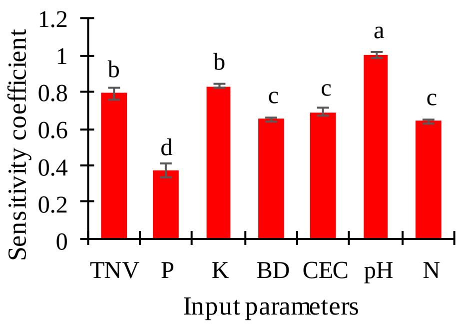

Fig. 3 Histogram of the sensitivity coefficient of plant weight using the Hill method

Sensitivity Analysis

Chlorophyll, carotenoid, antioxidant, phenol, flavonoid and plant weight were measured, but modeling was only done for plant weight (performance) due to their importance. The most sensitive parameter for chlorophyll was salinity, for carotenoid was lime percentage, for phenol was organic carbon percentage and for flavonoid was pH. After modeling performance using the 12 parameters with ANN and obtaining the best network, a sensitivity analysis was done using a dimensionless sensitivity coefficient [13] to determine the most sensitive parameters. Table 4 and Figure 3 show the results of the sensitivity analysis for plant weight performance. Hill [13] reported that if the sensitivity coefficient of a parameter were more than 0.1, that parameter would be among the sensitive parameters of the model. In this study, plant weight (performance) was sensitive to all parameters. The base of the modeling for both output parameters of plant weight (performance) should be the order of the sensitivity coefficient for the 12 soil parameters. However, the aim of this study was to estimate the plant weight using a minimum number of experiments and parameters. Thus, the most sensitive parameters were determined, respectively (According to Figure 2), and the modeling was done with these parameters. As shown in Figure 3, there is a decline in the sensitivity coefficient (a decrease in the sensitivity of the parameters). The results indicate that the plant weight (performance) of artichoke has the highest and lowest sensitivity to pH and soil phosphorus (respectively).

Table 4 Results of the sensitivity analysis of the easily accessible soil parameters of plant weight performance

Parameters	Relative sensitivity coefficient (γ)
pH	1
OC, %	0.97
K	0.83
TNV, %	0.79
CEC	0.69
EC	0.68
BD	0.65
N	0.64
Clay	0.57
Silt	0.51
Sand	0.44
P	0.37

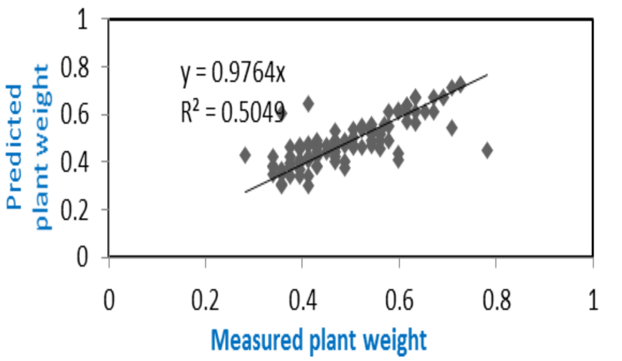

Fig. 4 Line fitted to predicted versus measured data of plant weight in the training phase of the model No.1

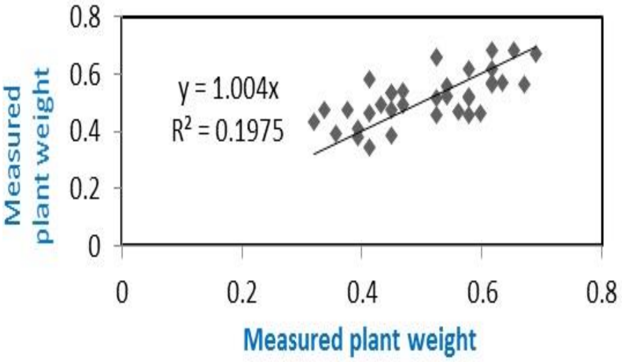

Fig. 5 Line fitted to predicted versus measured data of single plant weight in the testing phase of the model No.1

Designing Different ANN Models using the Sensitive Parameters

As mentioned earlier, the objective of the present study is to estimate the performance (plant weight) using minimal available parameters and at a lower cost. Therefore, four easily accessible soil parameters with the highest sensitivity coefficients were respectively entered into the model to develop the ANN models (Table 5).

Table 6 Statistical parameters estimated for the stages of training, validation, test and total in the model (1) for plant weight performance

Level	R^2	RMSE
Training	0.50	0.070
Validation	0.29	0.086
Test	0.19	0.078
Total	0.42	0.075

Results of the ANN Models with one Experiment to Obtain Model Inputs

Results of a model (1): In model (1), plant weight was estimated based on the organic carbon parameter. The suitable arrangement for the hidden layer with the Levenberg-Marquardt algorithm was selected as one hidden layer, 34 neurons, Logsig threshold function for the hidden layer and Tansig function for the output layer. Table 6 shows the statistical parameters for the training, validation, test and total stages in the model (1) for the plant weight. Figures 4 and 5 show R^2 and the equation of the line fitted to the predicted vs observed data for plant weight in the training stages for model (1).

Table 5 Different ANN models based on the results of the sensitivity analysis and the minimum number of experiments required obtaining the model inputs

Model	Input parameter	Number of experiments
Model 1	Organic carbon	1
Model 2	TNV, %	1
Model 3	pH	1
Model 4	Texture	1
Model 5	Organic carbon + lime percentage	2
Model 6	pH + Organic carbon	2
Model 7	pH + lime percentage	2
Model 8	Texture+ Organic carbon	2
Model 9	Organic carbon+ pH + lime percentage	3
Model 10	Texture+ Organic carbon+ pH + lime percentage	4

In model (1), only the organic carbon parameter was developed as the ANN model due to the high sensitivity coefficient.
In model (2), ANN is developed using lime percentage parameter.
In model (3), ANN is developed using pH parameter.
In model (4), ANN is developed using clay, silt and sand percentage parameters.
In model (5), the lime percentage parameter is added to model (1).
In model (6), the pH parameter is added to model (1).
In model (7), the pH parameter is added to model (2).
In model (8), the texture parameter is added to model (1).
In model (9), the pH parameter is added to the model (5) and was investigated.
In model (10), the texture parameter is added to model (9) and was investigated.
It is expected that increasing the number of parameters (model 1 to model 10) will lead to higher R^2 and lower RMSE values.

As shown in Figure 3, the slope of the line in the training and test stages are 0.97 and 1, respectively, which indicate that the predicted and observed data are close to each other. However, since R^2 in the training and test stages are 0.50 and 0.19, respectively, the model does not estimate the plant weight precisely. In addition, R^2 for the plant weight was more precise compared with the results of multivariate regression in the test stage.

Results of the model (2): In model (2), plant weight was estimated based on the lime percentage parameter. The best arrangement for the hidden layer with the Levenberg-Marquardt algorithm was selected as one hidden layer, 34 neurons, Logsig threshold function for the hidden layer and Tansig function for the output layer. Table 7 shows the statistical parameters for the training, validation, test and total stages in the model (2) for the plant weight. Figure show R^2 and the equation of the line fitted to the predicted vs observed data for plant weight in the training stages for the model (2). As evident in Figure 4, the slope of the line in training and test stages are 0.97 and 0.94, respectively, which indicate that the predicted and observed data in model (2) are close to each other.

Furthermore, R^2 in training and test stages are 0.25 and 0.18, respectively. Although the slope of the line in model (2) is close to 1 in the training and test stages and R^2 is lower compared to model (1), the model is still not precise enough to estimate the plant weight performance. Also, R^2 for the plant weight was more precise compared with the results of multivariate regression in the test stage.

Table 7 Statistical parameters estimated for the stages of training, validation, test and total in the model No. 2 for plant weight performance

Level	R^2	RMSE
Training	0.25	0.068
Validation	0.20	0.075
Test	0.18	0.068
Total	0.23	0.070

Results of the model (3): In model (3), plant weight was estimated based on pH parameter. The best arrangement for the hidden layer with the Levenberg-Marquardt algorithm was selected as one hidden layer, 34 neurons, Logsig threshold function for the hidden layer and Tansig function for the output layer.

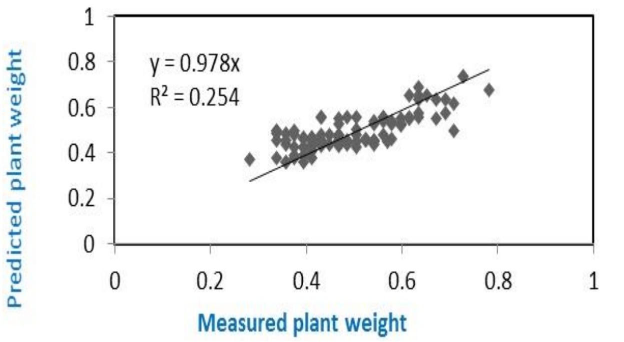

Fig. 6 Line fitted to predicted versus measured data of single plant weight in the training phase of the model No. 2

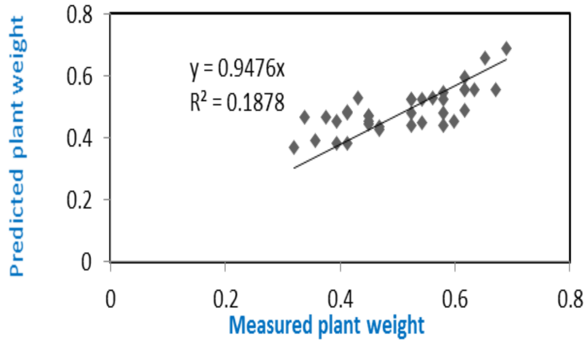

Fig. 7 Line fitted to predicted versus measured data of single plant weight in the testing phase of the model No.2

Table 8 shows the statistical parameters for the training, validation, test and total stages in the model (3) for the plant weight and antioxidant content. Figures 8 and 9 show R^2 and the equation of the line fitted to the predicted vs observed data for plant weight and antioxidant content in the training stages for model (3). As shown in Figures 8 and 9, the slope of the line in training and test stages are 0.98 and 0.95 for plant weight, and 0.98 and 1.02 for antioxidant content, respectively, which indicate that the predicted and observed data are close to each other. Furthermore, R^2 in training and test stages are 0.55 and 0.25 for plant weight and 0.18 and 0.10 for antioxidant content, respectively. The slope of the line is not much different, and R^2 in the training and testing stages has increased for the plant weight and decreased for antioxidant content compared with the previous model. However, R^2 Is still low in both training and testing stages and the model is still not precise enough to estimate the plant weight and antioxidant content performance. In addition, R^2 for the plant weight was more precise compared with the results of multivariate regression in the test stage.

Table 8 Statistical parameters estimated for the stages of training, validation, test and total in the model No. 3 for plant weight performance

Level	R^2	RMSE
Training	0.55	0.059
Validation	0.35	0.068
Test	0.25	0.065
Total	0.47	0.062

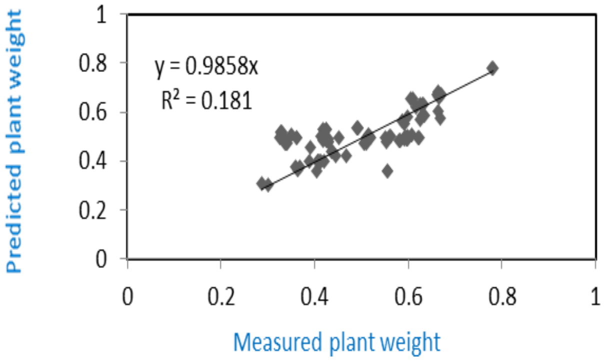

Fig. 8 Line fitted to predicted versus measured data of single plant weight in the training phase of the model No. 3

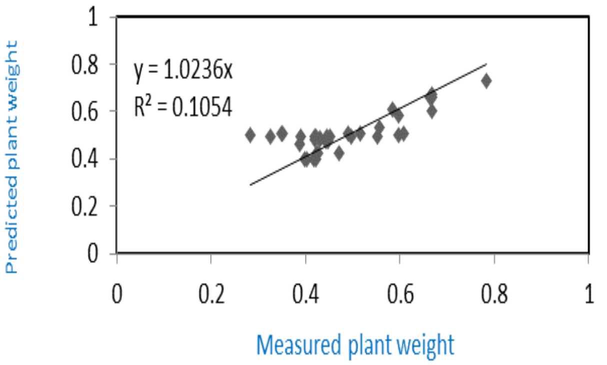

Fig. 9 Line fitted to predicted versus measured data of single plant weight in the testing phase of the model No. 3

Results of the model (4): In model (4), plant weight and antioxidant content were estimated based on soil texture parameters. The best arrangement for the hidden layer with the Levenberg-Marquardt algorithm was selected as one hidden layer, 34 neurons, Logsig threshold function for the hidden layer and Tansig function for the output layer. Table 9 shows the statistical parameters for the training, validation, test and total stages in the model (4) for the plant weight and antioxidant content. Figures 10 and 11 show R^2 and the equation of the line fitted to the predicted vs observed data for plant weight and antioxidant content in the training stages for model (4). As shown in Figures 10 and 11, the slope of the line in training and test stages are 0.99 and 1.01 for plant weight and 0.99 and 1.01 for antioxidant content, respectively, which indicate that the predicted and observed data close to each other. Furthermore, R^2 in training and test stages are 0.77 and 0.58 for plant weight and 0.79 and 0.37 for antioxidant content, respectively. The slope of the line is not much different, and R^2 in the training and testing stages has increased for the plant weight and antioxidant content compared with the previous model. Furthermore, R^2 has increased significantly in both training and testing stages, and the model shows an acceptable precision to estimate the plant weight and antioxidant content performance. In addition, R^2 for the plant weight was more precise compared with the results of multivariate regression in the test stage.

Table 9 Statistical parameters estimated for the stages of training, validation, test and total in the model No. 4 for plant weight performance

Level	R^2	RMSE
Training	0.77	0.045
Validation	0.62	0.073
Test	0.58	0.074
Total	0.69	0.058

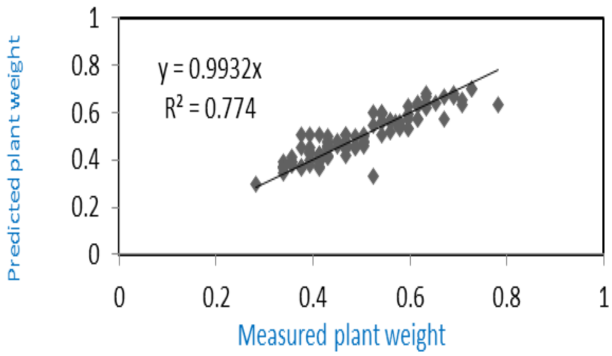

Fig. 10 Line fitted to predicted versus measured data of single plant weight in the training phase of the model No. 4

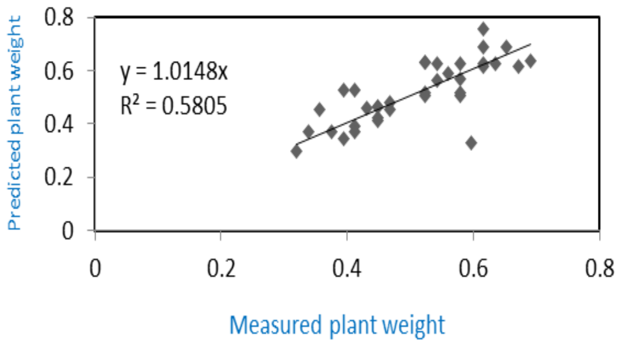

Fig. 11 Line fitted to predicted versus measured data of single plant weight in the testing phase of the model No. 4

Results of the ANN Models with two Experiments to Obtain Model Inputs

Results of the model (5): In model (5), plant weight and antioxidant content were estimated based on organic carbon and lime percentage parameters. The best arrangement for the hidden layer with the Levenberg-Marquardt algorithm was selected as one hidden layer, 45 neurons, Logsig threshold function for the hidden layer and Tansig function for the output layer. Tables 10 show the statistical parameters for the training, validation, test and total stages in the model (5) for the plant weight and antioxidant content. Figures 10 and 11 show R^2 and the equation of the line fitted to the predicted vs observed data for plant weight and antioxidant content in the training stages for model (5). The slope of the line in training and test stages are 0.98 and 0.97 for plant weight and 0.98 and 0.97 for antioxidant content, respectively, which indicate that the predicted and observed data close to each other. Furthermore, R^2 in training and test stages are 0.80 and 0.62 for plant weight, respectively. The slope of the line is not much different; however, R^2 in the training and testing stages for antioxidant content is acceptable, and thus, the model shows an acceptable precision to estimate antioxidant content performance. In addition, R^2 for the plant weight was more precise compared with the results of multivariate regression in the test stage.

Table 10 Statistical parameters estimated for the stages of training, validation, test and total in the model No. 5 for plant weight performance

Level	R^2	RMSE
Training	0.80	0.046
Validation	0.79	0.050
Test	0.62	0.063
Total	0.76	0.051

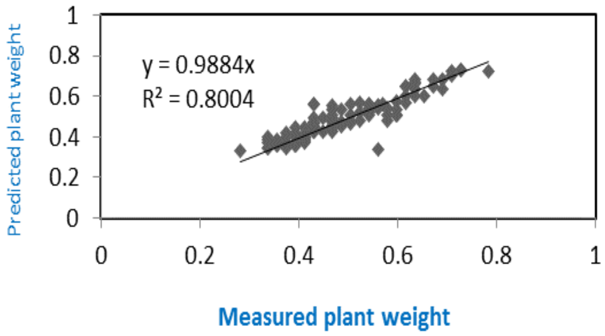

Fig. 12 Line fitted to predicted versus measured data of single plant weight in the training phase of the model No. 5

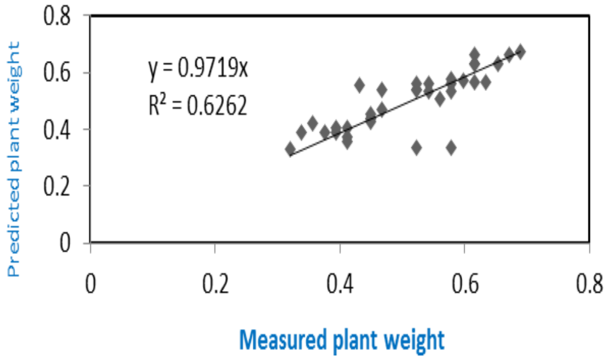

Fig. 13 Line fitted to predicted versus measured data of single plant weight in the testing phase of the model No. 5

Results of the model (6): In model (6), plant weight and antioxidant content were estimated based on organic carbon and lime percentage parameters. The best arrangement for the hidden layer with the Levenberg-Marquardt algorithm was selected as one hidden layer, 34 neurons, Logsig threshold function for the hidden layer and Tansig function for the output layer. Tables 11 show the statistical parameters for the training, validation, test and total stages in the model (6) for the plant weight and antioxidant content. Figures 14 and 15 show R^2 and the equation of the line fitted to the predicted vs observed data for plant weight and antioxidant content in the training stages for model (6). The slope of the line in training and test stages are 0.99 and 0.95 for plant weight and 0.97 and 1 for antioxidant content, respectively, which indicate that the predicted and observed data close to each other. Furthermore, R^2 in training and test stages are 0.73 and 0.21 for plant weight and 0.66 and 0.50 for antioxidant content, respectively. According to the results, this model does not have an acceptable precision to estimate the plant weight and antioxidant content. In addition, R^2 for the plant weight was more precise compared with the results of multivariate regression in the test stage.

Table 11 Statistical parameters estimated for the stages of training, validation, test and total in the model No.6 for plant weight performance

Level	R^2	RMSE
Training	0.73	0.050
Validation	0.28	0.088
Test	0.21	0.083
Total	0.53	0.066

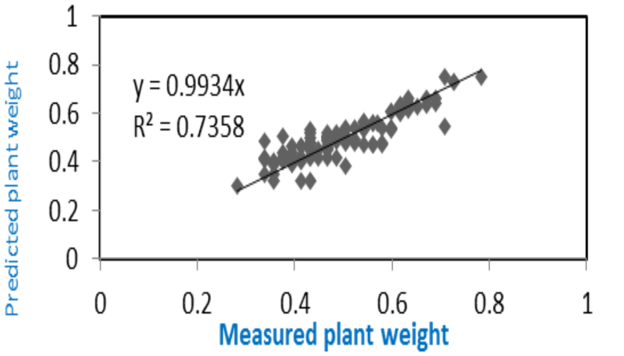

Fig. 14 Line fitted to predicted versus measured data of single plant weight in the training phase of the model No. 6

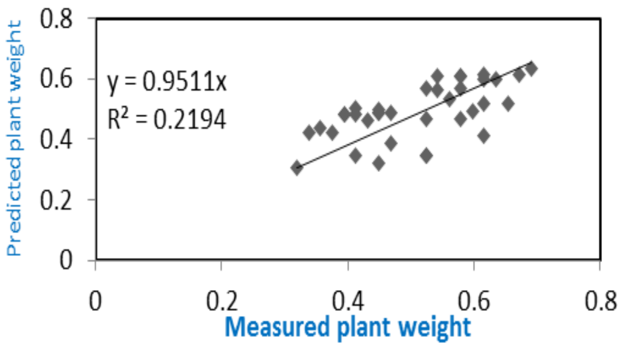

Fig. 15 Line fitted to predicted versus measured data of single plant weight in the testing phase of the model No. 6

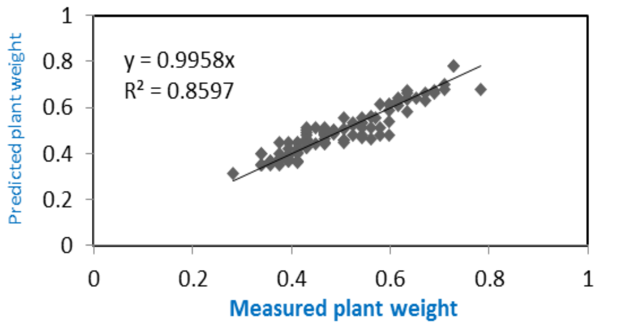

Fig. 16 Line fitted to predicted versus measured data of single plant weight in the training phase of the model No. 7

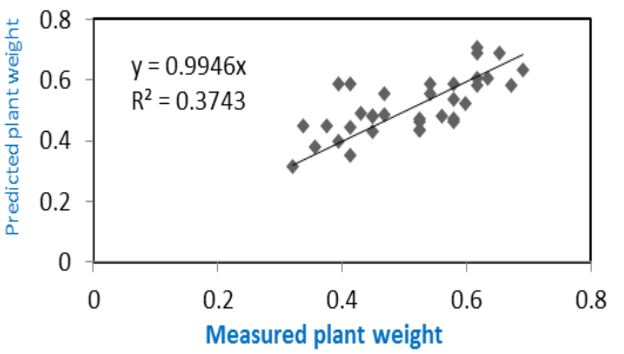

Fig. 17 Line fitted to predicted versus measured data of single plant weight in the testing phase of the model No. 7

Results of the model (7): In model (7), plant weight and antioxidant content were estimated based on organic carbon and lime percentage parameters. The best arrangement for the hidden layer with the Levenberg-Marquardt algorithm was selected as one hidden layer, 45 neurons, Logsig threshold function for the hidden layer and Tansig function for the output layer. Figures 16 and 17 show R^2 and the equation of the line fitted to the predicted vs observed data for plant weight and antioxidant content in the training stages for model (7). The slope of the line in training and test stages are 0.99 and 0.99 for plant weight and 0.98 and 1 for antioxidant content, respectively, which indicate that the predicted and observed data close to each other. Furthermore, R^2 in training and test stages are 0.85 and 0.37 for plant weight and 0.84 and 0.61 for antioxidant content, respectively. The model shows an acceptable precision to estimate only performance. In addition, R^2 for the plant weight was more precise compared with the results of multivariate regression in the test stage.

Table 12 Statistical parameters estimated for the stages of training, validation, test and total in the model No. 7 for plant weight performance

Level	R^2	RMSE
Training	0.85	0.038
Validation	0.66	0.068
Test	0.37	0.073
Total	0.73	0.053

Table 13 Statistical parameters estimated for the stages of training, validation, test and total in the model No. 8 for plant weight performance

Level	R^2	RMSE
Training	0.96	0.018
Validation	0.67	0.077
Test	0.58	0.082
Total	0.80	0.052

Results of the model (8): In model (8), plant weight and antioxidant content were estimated based on organic carbon and soil texture parameters. The suitable arrangement for the hidden layer with the Levenberg-Marquardt algorithm was selected as one hidden layer, 34 neurons, Logsig threshold function for the hidden layer and Tansig function for the output layer. Tables 13 show the statistical parameters for the training, validation, test and total stages in the model (8) for the plant weight. Figures 18 and 19 show R^2 and the equation of the line fitted to the predicted vs observed data for plant weight and antioxidant content in the training stages for the model (8).

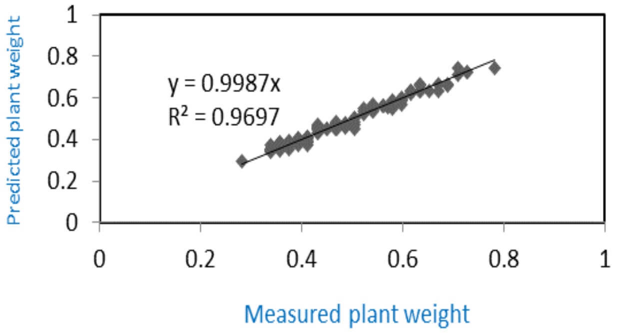

Fig. 18 Line fitted to predicted versus measured data of single plant weight in the training phase of the model No. 8

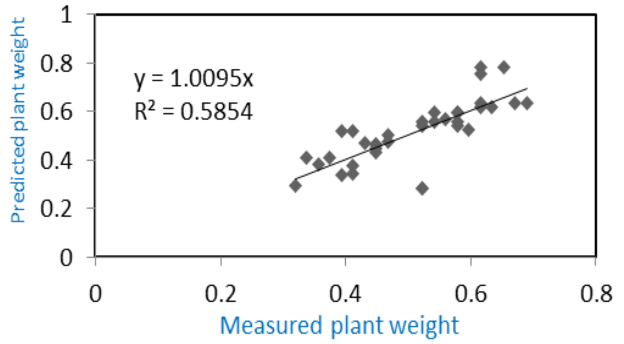

Fig. 19 Line fitted to predicted versus measured data of single plant weight in the testing phase of the model No. 8

The slope of the line in training and test stages are 0.99 and 1 for plant weight and 0.99 and 1.01 for antioxidant content, respectively, which indicate that the predicted and observed data close to each other. Furthermore, R^2 in training and test stages are 0.96 and 0.58 for plant weight and 0.99 and 0.69 for antioxidant content, respectively. The model shows an acceptable precision to estimate plant weight and antioxidant content performance. In addition, R^2 for the plant weight was more precise compared with the results of multivariate regression in the test stage.

Results of the ANN Model with Three Experiments to Obtain Model Inputs

In model (9), plant weight and antioxidant content were estimated based on organic carbon, soil texture and pH parameters. The best arrangement for the hidden layer with the Levenberg-Marquardt algorithm was selected as one hidden layer, 34 neurons, Logsig threshold function for the hidden layer and Tansig function for the output layer. Tables 14 show the statistical parameters for the training, validation, test and total stages in the model (9) for the plant weight. Figures 20 and 21 show R^2 and the equation of the line fitted to the predicted vs observed data for plant weight and antioxidant content in the training stages for the model (9). Figures 20 show R^2 and the equation of the line fitted to the predicted vs observed data for antioxidant content in the training stages for model (9).

Table 14 Statistical parameters estimated for the stages of training, validation, test and total in the model (9) for plant weight performance

Level	R^2	RMSE
Training	0.79	0.047
Validation	0.76	0.050
Test	0.66	0.059
Total	0.76	0.050

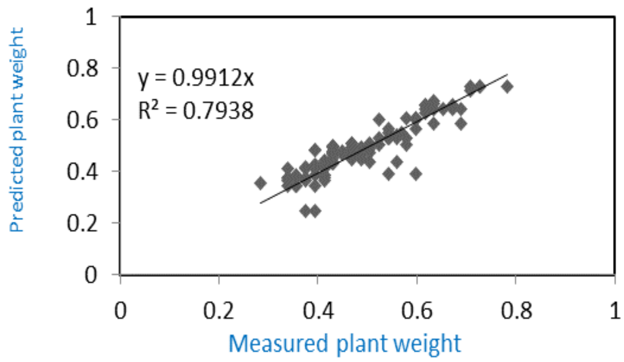

Fig. 20 Line fitted to predicted versus measured data of single plant weight in the training phase of the model No. 9

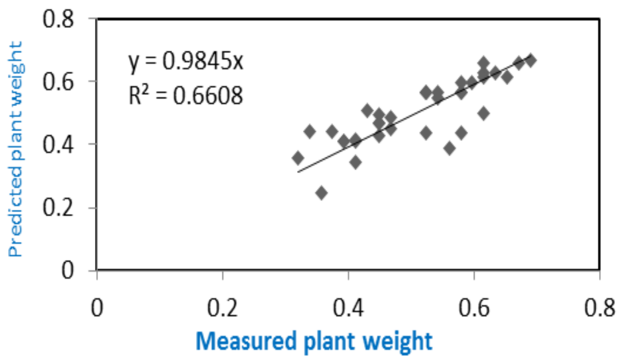

Fig. 21 Line fitted to predicted versus measured data of single plant weight in the testing phase of the model No. 9

The slope of the line in training and test stages are 0.99 and 0.98 for plant weight and 0.99 and 1.01 for antioxidant content, respectively, which indicate that the predicted and observed data close to each other. Furthermore, R^2 in training and test stages are 0.79 and 0.66 for plant weight, respectively. In addition, R^2 for the plant weight was more precise compared with the results of multivariate regression in the test stage.

Table 15 Statistical parameters estimated for the stages of training, validation, test and total in the model No. 10 for plant weight performance

Level	R^2	RMSE
Training	0.96	0.018
Validation	0.73	0.065
Test	0.63	0.074
Total	0.83	0.046

Table 16 Comparison of RMSE and R^2 in different models at the test stage for plant weight performance

Model	RMSE	R^2
1	0.078	0.19
2	0.068	0.18
3	0.062	0.58
4	0.074	0.58
5	0.063	0.62
6	0.083	0.21
7	0.073	0.37
8	0.082	0.58
9	0.059	0.66
10	0.074	0.63
Multivariate regression	0.99	0.12

Results of the ANN Model with four Experiments to Obtain Model Inputs

Results of the model (10): In model (10), plant weight and antioxidant content were estimated based on organic carbon, soil texture and pH parameters. The best arrangement for the hidden layer with the Levenberg-Marquardt algorithm was selected as one hidden layer, 34 neurons, Logsig threshold function for the hidden layer and Tansig function for the output layer. Tables 15 show the statistical parameters for the training, validation, test and total stages in the model (10) for the plant weight. Figures 22 and 23 show R^2 and the equation of the line fitted to the predicted vs observed data for plant weight and antioxidant content in the training stages for the model (10). Figures 22 shows R^2 and the equation of the line fitted to the predicted vs observed data for antioxidant content in the training stages for the model (10). The slope of the line in training and test stages are 0.99 and 1.04 for plant weight and 1, respectively, which indicate that the predicted and observed data close to each other. Furthermore, R^2 in training and test stages are 0.96 and 0.63 for plant weight, respectively. The model shows an acceptable precision to estimate plant weight. In addition, R^2 for the plant weight was more precise compared with the results of multivariate regression in the test stage.

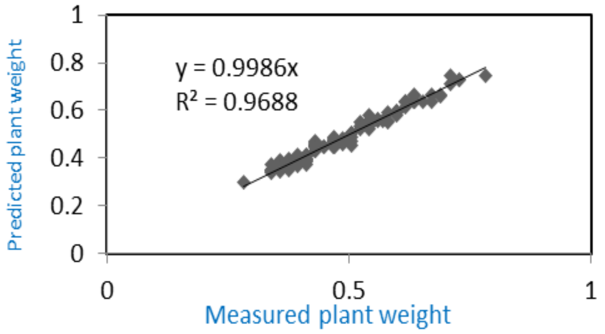

Fig. 22 Line fitted to predicted versus measured data of single plant weight in the training phase of the model No. 10

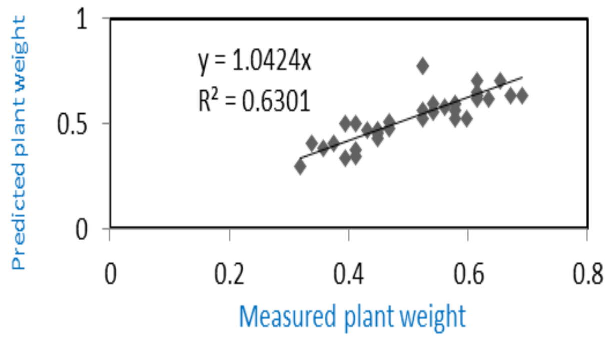

Fig. 23 Line fitted to predicted versus measured data of single plant weight in the testing phase of the model No. 10

Comparison of the Designed Model Results with the Sensitive Parameters

The suitable model was determined by comparison of the ten developed models and the multivariate regression model. According to Tables 16 and 17, increasing the number of input parameters will lead to higher R^2 and lower RMSE in the training, validation and testing stages, which indicate the increased precision of the model in the estimation of artichoke plant weight [14]. This is in accordance with the results of Shop *et al*. [15], Shop and Lich

[16], and Moazenzadeh *et al*. [17]. All ANN models were better than the multivariate regression models in the estimation of plant weight. Models 1-4 had almost similar performances. In a way, model 5 can be selected as the best model, as this model had almost similar performance with a minimal number of inputs compared with the other models. However, model 4 had a higher R^2 value, and is cheaper, as all texture parameters may be measured with a single experiment. The results indicate promising ANN application in the estimation of plant weight using soil parameters. However, further research is recommended to reach fully certain results.

CONCLUSION

The results showed that the ANN method has high precision in the estimation of artichoke plant weight so that in 7 out of 10 designed models (R^2 in the testing stage), it describes the variation of plant weight of antioxidant in the studied region using the 12 easily accessible soil parameters. Plant weight is highly dependent on soil pH, organic carbon, potassium and lime percentage. This study showed that the pH parameter is the most important factor affecting artichoke plant weight. In addition, silt percentage was determined as the most important factor-influencing yield. The results of the present study are only applicable to the studied region and other regions with similar topography, climate, soil and management practices. However, similar studies may be carried out in other regions using ANN. Based on the findings of the mentioned model, Organic carbon and the percentage of lime in soil are the two most important soil readily available parameters for estimating artichoke biomass production.

REFERENCES

1. Ebrahimi-Mameghani M., Asghari-Jafarabadi M., Rezazadeh K. TCF7L2-rs7903146 polymorphism modulates the effect of artichoke leaf extract supplementation on insulin resistance in metabolic syndrome: a randomized, double-blind, placebo-controlled trial. J. Integr. Med. 2018;16:329–334.
2. Sabater C., Sabater V., Olano A., Montilla A., Corzo N. Ultrasound-assisted extraction of pectin from artichoke by-products. An artificial neural network approach to pectin characterization. Food Hydrocoll. 2020;98:105238.
3. Bergmeir C., Benitez J.M. Neural networks in R using the stuttgart neural network simulator: RSNNS. J. Stat. Softw. 2012;46:1–26.
4. Sabater C., Molina-Tijeras J.A., Vezza T., Corzo N., Montilla A., Utrillab P. Intestinal anti-inflammatory effects of artichoke pectin and modified pectin fractions in the dextran sulfate sodium model of mice colitis. Artificial neural network modelling of inflammatory markers. Food & Function. 2019;10(12):7793–7805
5. Kumar N., Raghuwanshi N.S., Singh R., Wallender W.W., Pruitt W.O. Estimating evapotranspiration using artificial network. J. Irrig. Drain. Eng. 2002;128(4):224-233.
6. Drummond S.T., Sudduth K.A., Joshi A., Birrell SL, Kitchen NR. Statistical and neural methods for site-specific yield prediction. Transactions of the ASAE. 2003;46(1):5-14.
7. Cravener T.L., Roush W.B. Prediction of amino acid profiles in feed ingredients:genetic algorithm calibration of artificial neural networks. Anim. Feed Sci. Technol. 2001;90:131-141.
8. Sabourifard H., Ghasemnezhad A., Hemmati K., Hezarjeribi A., Bahrami M., Nosrati M. Assessment of summer savory (Satureja hortensis L.) biomass by easily-attainable soil parameters and artificial network. J. Hortic. Sci. 2017;31(3):448-456.
9. Page A., Miller R., Keeney D. Methods of soil analysis.2th ed. Part 2: chemical and biological properties. Soil Sci. Soc. Am. Inc. publisher. 1987.
10. Movahedi Naeini S.A., Rezaei M. Soil physics (fundamentals and applications). Gorgan University of Agricultural Sciences and Natural Resources Publications, 474 p. 2008. (In Persian)
11. Bremner J.S., Mulvaney C.S. Nitrogen-total. In A. L. Page (Ed.), Methods of soil analysis, Part 2. American Society of Agronomy (pp. 595-624). Madison, Wisconsin. 1982.
12. Sajikumara N., Thandaveswra B. A non-linear rainfall-runoff model using an Artificial Neural Network. Hydrol. J. 1999;216:32-55.
13. Hill M. Methods and guidelines for effective model calibration. U.S. Geological survey Water- Resources Investigations Rep. 1998;98-4005.
14. Sabourifard H., Ghasemnezhad A., Hemmati K., Hezarjeribi A., Bahrami M., Nosrati M. Assessment of summer savory (Satureja hortensis L.) biomass by easily-attainable soil parameters and artificial network. J. Plant Prod. Sci. 2017;26(2):47-58.
15. Schaap M., Leij F., van Genuchten M. Neural network analysis for hierarchical prediction of soil hydraulic properties. Soil Sci. Soc. Am. J. 1998;62:847-855.
16. Schaap M., Leij F. Using neural networks to predict soil water retention and soil hydraulic conductivity. Soil Tillage Res. 1998;47:37-42.
17. Moazenzadeh R., Ghahraman B., Fathalian F., Khoshnoodiyazdi A. Effect of type and number of input variables on estimation of moisture retention curve and saturated hydraulic conductivity. Journal of Water and Soil. 2009;23(3):57-70.