Non-destructive Estimation of Spinach Leaf Area: Image Processing and Artificial Neural Network Based Approach

Naveen Kumar Mahanti[1*], Subir Kumar Chakraborty[1] and V. Bhushana Babu[2]

[1]*Agro Produce Processing Division, ICAR-Central Institute of Agricultural Engineering, Bhopal-462038, India.*
[2]*Agriculture Farm Mechanization Division, ICAR-Central Institute of Agricultural Engineering, Bhopal-462038, India.*

Authors' contributions

This work was carried out in collaboration among all authors. Author NKM designed the study, performed the statistical analysis, wrote the protocol and wrote the first draft of the manuscript. Authors SKC and VBB managed the analyses of the study. Authors NKM managed the literature searches. All authors read and approved the final manuscript.

Original Research Article

ABSTRACT

Leaf area (L_A) measurement provides valuable key information in understanding the growth and physiology of a plant. Simple, accurate and non-destructive methods are inevitable for leaf area estimation. These methods are important for physiological and agronomic studies. However, the major limitations of existing leaf area measurement techniques are destructive in nature and time consuming. Therefore, the objective of the present work is to develop ANN and linear regression models along with image processing techniques to estimate spinach leaf area making use of leaf width (L_W) and length (L_L) and comparison of developed models performance based on the statistical parameters. The spinach leaves were grown under different nitrogen fertilizer doses (0,

_____

Corresponding author: E-mail: naveeniitkgp13@gmail.com;

50, 100, 150, 200, 250, 300, 350 and 400 kg N/ha). The morphological parameters length (L_L), width (L_W) and area (L_A) of leaves were measured using an image-processing software. The performance L_A= -0.66+0.64 (L_L × L_W) (R^2 = 0.98, RMSE = 3.25 cm^2) equation was better than the other linear models. The performance of the ANN model (R^2 = 0.99, RMSE = 3.10 cm^2) was better than all other linear models. Therefore, developed models along with image processing techniques can be used as a non-destructive technique for estimation of spinach leaf area.

1. INTRODUCTION

Spinach (*Spinacia oleracea* L.) is one of the major green leafy vegetable that belongs to the amaranth family and originated in Persia. It is widely cultivated in spring, autumn and winter seasons. Spinach, considered good for health and loaded with antioxidants and nutrients, is a rich source of vitamins A, C and K [1] and minerals such as calcium, iron, and magnesium [2].

A leaf is an integral part of a plant that is responsible for the rate of photosynthesis, crop growth and dry matter accumulation. Leaf area, one of the parameters, plays a vital role in assessing the leafy vegetative plant growth [3]. The conventional methods of leaf area measurement are blueprinting, tracing, photo-graphing and planimeter [4-5]. Such methods are time-consuming, and these methods require the excision of leaves from the plants [4-5]. Moreover, the accuracy in such a measurement process depends on the skills of the operator engaged in that work. Unlike destructive methods, non-destructive methods, which involves mathematical models to predict leaf area, allows for the repeated measurement of leaves over time while avoiding the biological alteration characteristic [6]. Furthermore, the use of simple linear measurement models for predicting leaf area (A_P) of crops eliminates the introduction of expensive leaf area meters [7].

Among the different methods of leaf area measurement techniques, the linear measurement of leaf length and width are more prevalent. Many researchers have said that the linear regression model developed with leaf length and width has good prediction accuracy to estimate leaf area [8-10,14-15]. This type of leaf area estimation is called the indirect method and it is non-destructive. Numerous researches have used linear regression models to estimate the area for a wide range of crops such as cherry [8], chestnut [9], sunflower [10], saffron [11], mango [12], rose [13], highbush blueberries [14] and cacao [15].

Artificial neural network is a machine learning method inspired by the biological neural network. It progressively utilize and is capable of performing many complex tasks and yields productive outcomes in comparison with other traditional predictive models like simple linear regression and correlation [16]. It is extensively used in hydrological forecasting [17], fruit weight prediction [18], corn and soybean crop yield prediction [19], leaf area estimation for corn [20] and tomato [21], water demand forecasting [22], soil parameters estimation [23] and estimation of rice leaf chlorophyll concentration [24].

It has been observed that the number of researchers has developed linear and non-linear models to predict the leaf area using length (L_L) and width (L_W) of leaves. Only limited work is published on development of the ANN model for predicting leaf area (A_P). Therefore the present work has been undertaken to develop ANN and linear regression models along with image processing techniques to estimate spinach leaf area making use of leaf width and length values and comparison of developed models performance based on the statistical parameters.

2. MATERIALS AND METHODS

2.1 Experimental location

The spinach plants were cultivated in plastic trays under different nitrogen fertilizer doses (0, 50, 100, 150, 200, 250, 300, 350 and 400 kg N/ha) during the winter (December-February) season. The experiment was carried out at the Precision Farming Development Centre (PFDC) field (23.3156°N, 77.4037°E) of the ICAR-Central Institute of Agricultural Engineering, Bhopal, Madhya Pradesh, India. The pictorial representation of spinach plants grown under different levels of nitrogen fertilizer is shown in Fig. 1.

2.2 Measurement of Leaf Morphological Properties

A total of 90 leaves without any physical damage were collected from all the treatments (10 leaves from each treatment) after 35 days of sowing. The spinach leaves were scanned using the CanoScan LiDE 110 Scanner in jpeg (Joint Photographic Experts Group) format. The scanned images were used to measure morphological properties of spinach leaves such as length, width and area using Image J software (Fig. 2a). Before analyzing the leaves, the scale was set using a known length. After that, the threshold was selected (Fig. 2b) and then leaf properties were analyzed using ROI (region of interest) manager.

Fig. 1. Plastic trays cultivation of spinach leaves under different levels of nitrogen fertilizer doses
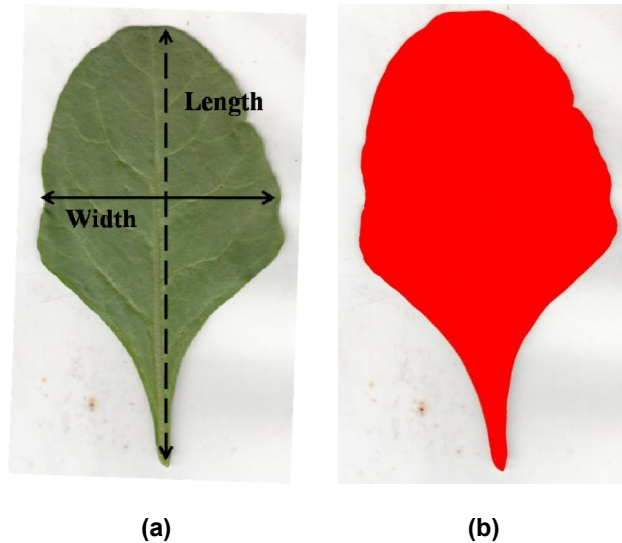

| (a) | (b) |

Fig. 2. Image processing operations to measure spinach morphological properties (a) spinach leaf scanned image and (b) colour threshold image

2.3 Linear Regression Model

Six linear regression models have been developed considering the different combinations of length and width (L_L, L_W, L_L^2, L_W^2, $(L_L+L_W)^2$, $L_L \times L_W$) as an independent variable and leaf area as a dependent variable. Even if the relationship between dependent and independent variables is curvilinear under this situation, the use of second-order polynomials yields best-fit regression models. The list of linear models used in this study is tabulated in Table 1. The coefficient of determination (R^2) and the mean square error (MSE) were used as summary statistics to check the best fitting of developed linear regression models and for comparison of models [25].

2.4 ANN Model

A feed-forward back propagation artificial neural network (FFBP-ANN) with topology comprising two layers empirically was found to be optimum in this investigation for the prediction of spinach *leaf area (L_A)*. The number of neurons in the input layer is equal to the number of independent variables, i.e. leaf length (L_L) and width (L_W), the number of output neurons was equal to the dependent variable, i.e., leaf area (L_A). There are no fixed, definite rules for determining the number of hidden neurons needed in a hidden layer [26]. The number of neurons in the hidden layer varied between 2 to 5, it was selected based on the trial and error method and the optimum number of neurons were finalized based on the statistical parameters, coefficient of determination (R^2) and root mean square error (*RMSE*). The pictorial representation of FFBP-ANN network adopted in the present study is depicted in Fig 3. The hidden layer maps with input X and output Y through a sequence of interconnected weights can be expressed mathematically as follows [27].

$$Y_j = \sum_{i=1}^{n} f(W_{ij} X_i) + b_j \qquad (1)$$

where, W_{ij} represents the weight of ith input vector which is connected to the jth neuron; n represents the number of inputs to the neuron; b_j is the bias associated with the jth neuron, which adds a constant term in the weighted sum to improve convergence; and f is the activation function that establishes the processing inside the neuron. The activation function may be a linear or non-linear (hyperbolic tangent or sigmoid) based on the topology of the network. In the present study tan sigmoid function was used as an activation function in the hidden layer and linear function in the output layer.

$$tansig(x) = \frac{2}{(1+e^{-2x})-1} \qquad (2)$$

A total of 90 experiments were conducted and the experimental data were used for training, testing and validation of the selected network. The data was first randomized, and then the data was divided into three parts. The first set was used for training the network (60%), the second set (20%) was used for validation of network and the third (20%) data set was used for testing the model. MATLAB2018a software was used for the development of neural network. Levenberg Marquardt's training algorithm was used for updating the weights of the input layer to the hidden layer and the hidden layer to output layer connections. The statistical parameters such as RMSE and R^2 were used to evaluate the performance of the developed network. The low RMSE and high R^2 represent good prediction accuracy.

$$R^2 = \frac{[\sum(Y_P - \overline{Y_P})(Y_O - \overline{Y_O})]^2}{\sum(Y_P - \overline{Y_P})^2 \sum(Y_O - \overline{Y_O})^2} \qquad (3)$$

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

Where,

y_i is the observed value for the ith observation and $\bar{y}_i$ is the predicted value.

Table 1. Coefficient of estimates and statistical results of different linear equations obtained from regression analysis to estimate the spinach leaf area

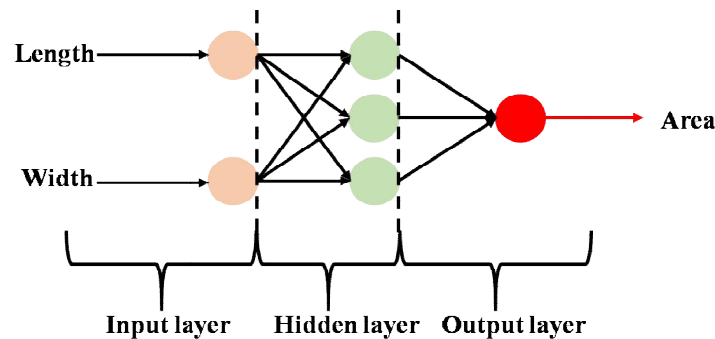| Model number	Form of model	Coefficients		R^2	RMSE(cm^2)
		a	b		
1	$L_A = a+b\ (L_L)$	-38.00	7.35	0.86	8.33
2	$L_A = a+b\ (L_W)$	-32.30	12.88	0.95	5.02
3	$L_A = a+b\ (L_L \times L_W)$	-0.66	0.64	0.98	3.25
4	$L_A = a+b\ (L_L + L_W)^2$	-1.45	0.15	0.96	4.14
5	$L_A = a+b\ (L_L^2)$	-0.83	0.34	0.88	7.58
6	$L_A = a+b\ (L_W^2)$	4.75	1.04	0.94	5.16

Fig. 3. Feedforward back propagation neural network adopted for the present study

3. RESULTS AND DISCUSSION

3.1 Linear Regression Models

The coefficient of estimates and statistical results obtained from the regression analysis of different linear equations to estimate the spinach leaf area (L_A) are tabulated in Table 1. The prediction accuracy of linear models was low when the models developed with leaf length (L_L) and leaf width (L_W) as an individual independent variable. The performance of model developed with leaf length (L_L) as an independent variable (Model-1) (R^2=0.86; RMSE=8.33) was lower than the model developed with leaf width (L_W) as an independent variable (Model-2) (R^2=0.95; RMSE=5.02). On the other hand, the performance of these models remains invariable with increasing the order of independent variables (*i.e* L_L^2: R^2=0.88; RMSE=7.58, and L_W^2: R^2=0.94; RMSE=5.16) (Model-5 & 6). The models developed with both L_L and L_W as an independent variable (Model-3 & 4) reported superior prediction accuracy than the remaining models. The model developed with both L_L and L_W (*i.e* $L_L \times L_W$-Model-3) as an independent variable performed better than the other linear models (R^2=0.98; RMSE=3.25). Similarly, some of the previous studies adopted different types of linear equations to predict the leaf area of different crops. The results obtained in the present study were better than the results reported for Solanumaethiopicum Shum (R^2=0.83-0.92) [28], saffron (*Crocus sativus* L.) (R^2=0.937) and corn (*Zea mays L.)* (R^2=0.88–0.97) [29]. This variance may be due to the variation in crop, type of linear model adopted and leaf morphology. A shape between an ellipse (0.78) and a triangle (0.5) of the equal length and maximum width can express the shape coefficient. The shape coefficient (regression coefficient of model-3) of spinach leaf (0.64) was close to walnut leaf crop (0.69) [30] but lower than the hazelnut leaf (0.74) [7] and kiwi fruit leaf (0.82) [31].

3.2 FFBP-ANN Model

The multilayer feed-forward back propagation neural network (FFBP-ANN) was developed using L_L and L_W as variables for the input layer and L_A as an output layer with one hidden layer with three neurons. The lower RMSE and high R^2 values were observed at three number of hidden neurons therefore it is considered as an optimum number of hidden neurons. The best model performance (i.e lower RMSE) was observed at the number of epochs equal to 16; it is shown in Fig 4a. From the figure, it is depicted that the performance of the model during training, validation and testing was almost the same. The statistical results of ANN model during training (R^2=0.996; RMSE=2.87), validation (R^2=0.994; RMSE=2.79) and testing (R^2=0.995; RMSE=3.10) reveals that the developed model was free from the over and underfitting. The plot between actual and predicted spinach L_A during training, validation and testing are shown in Fig 4b. It is evident that all the data points are close to the regression line, therefore, it can be concluded that the developed model was capable of predicting L_A with good prediction accuracy and no outliers were present in the data points.

The results obtained in the present study were better (R^2=0.995; RMSE=3.10) than the results reported for pepper (*Capsicum annuum* L.) leaf (R^2=0.987; RMSE=270.956 mm^2) [32], even though the ANN model developed with more number of hidden layers and neurons than the present study. Similarly, [21] also developed
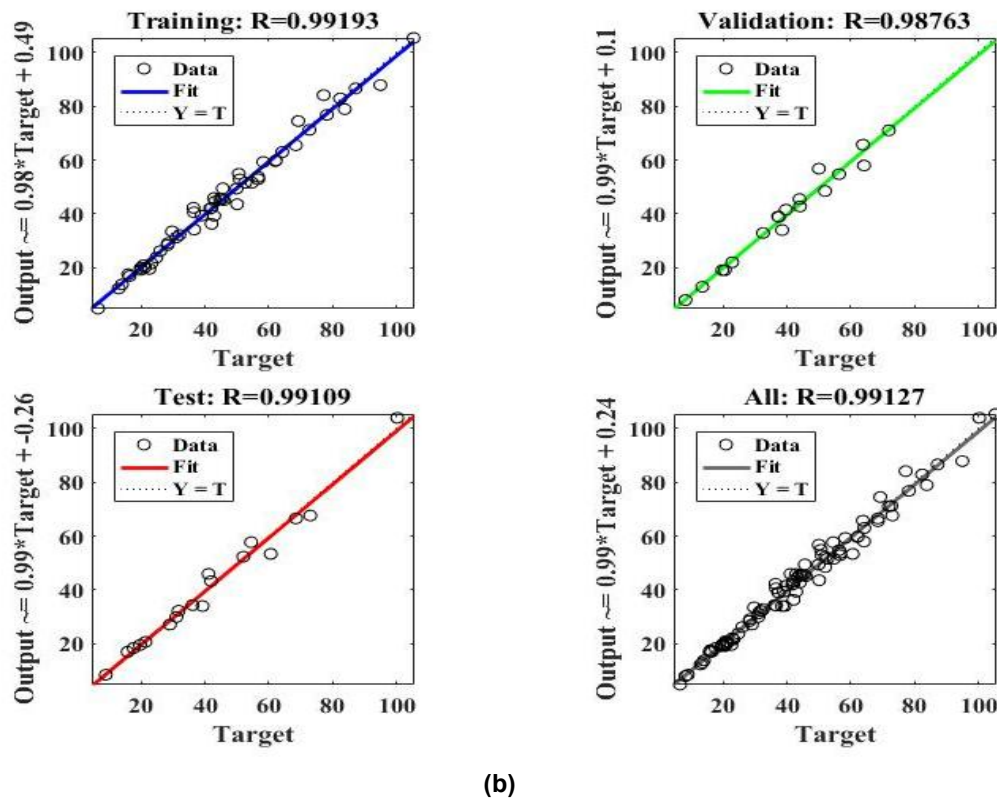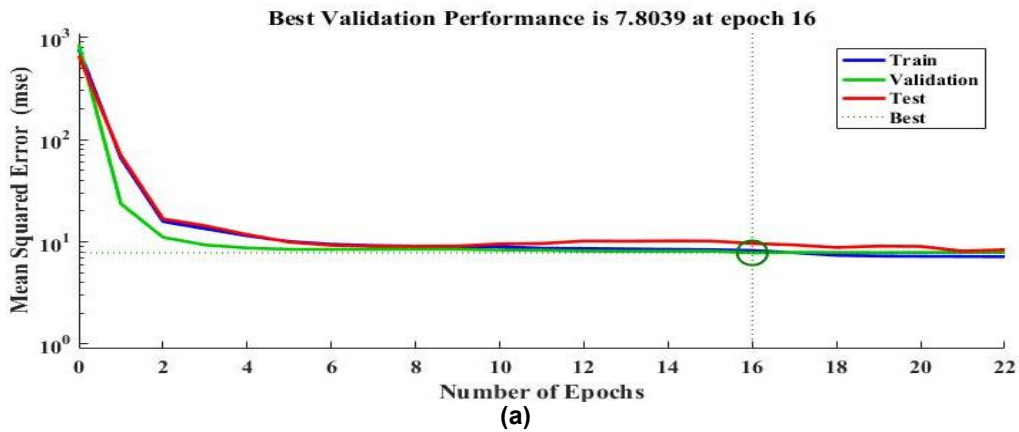
(a)

(b)

Fig. 4. The performance of FFBP-ANN during training, validation and testing, (a) performance of the model at a different number of epochs and (b) plot between actual and predicted spinach leaf area during model training, validation and testing

ANN model to predict the tomato leaf area. The prediction accuracy was lower (R^2= 0.93-0.98; RMSE=8.70-18.07 cm^2) than the present study [21].The prediction accuracy obtained in the present study were similar to the previous studies reported for corn leaves (R^2=0.98) [20], tomato leaves (R^2=0.96; RMSE=3.30 cm^2) [33] and durian (*Duriozibethinus*) leaves (R^2 = 0.94, RMSE = 4.81 cm^2) [11].

3.3 Comparison of ANN and LRM

The performance of linear models and FFBP-ANN model were compared based on the statistical indicators R^2 and RMSE. The performance of FFBP-ANN model (R^2=0.995; RMSE=3.10) was superior to all the linear models, but the notable distinction was not observed between the performance of ANN and

linear Model-3. The better prediction accuracy of FFBP-ANN model was due to its non-linear nature among input and output variables. The results obtained in the present study were similar to the results reported by [11] for durian (*Duriozibethinus*) leaves and for corn leaves (Odabas *et al.* 2013). Therefore, ANN can be used to predict the leaf area of spinach leaves with good prediction accuracy and minimum error.

4. CONCLUSION

This study was conducted with the aim of non-destructive estimation of spinach leaf area using linear and artificial neural network (ANN) models along with image processing techniques. The linear models developed with both L_L and L_W used as an independent variable yields good prediction accuracy with lower error (R^2=0.86-0.95; RMSE=5.03-8.22cm^2) than the models developed with L_L and L_W used as an independent variable individually (R^2=0.96-0.98; RMSE=3.25-4.14 cm^2). The performance of the ANN model was better (R^2 = 0.99, RMSE = 3.10 cm^2) than the linear models. The developed ANN model along with image processing techniques can be used as a substitute for sophisticated instruments for measurement of spinach leaf area.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Hedges LJ, Lister CE. Nutritional attributes of spinach, silver beet and eggplant. Crop Food Res Confidential Rep. 2007;1928.
2. Kawazu Y, Okimura M, Ishii T, Yui S. Varietal and seasonal differences in oxalate content of spinach. Scientia Hort. 2003;97:203-210.
3. Pandey SK, Singh H. A simple, cost-effective method for leaf area estimation. Journal of Botany. 2011;6.
4. Fallovo C, Cristofori V, Mendoza-de-Gyes E, Rivera CM, Rea R, Fanasca S, Bignami C, Sassine Y, Rouphael Y. Leaf area estimation model forsmall fruits from linear measurements. Hort Science. 2008;43(7): 2263–2267.
5. Radzali MH, Kamal NAM, Diah, NM. Measuring leaf area using Otsu segmentation method (LAMOS). Ind. J. Sci. Technol. 2016;9:1-6.
6. Swart ED, Groenwold R, Kanne HJ, Stam P, Marcelis LF, Voorrips RE. Non-destructive estimation of leaf area for different plant ages and accessions of *Capsicum annuum* L. The Journal of Horticultural Science and Biotechnology. 2004;79(5):764-770.
7. Cristofori V, Rouphael Y, Gyves EMD, Bignami C. A simple model for estimating leaf area of hazelnut from linear measurements. Scientia Horticulturae. 2007;13:221–225.
8. Demirsoy H, Demirsoy L. A validated leaf area prediction model for some cherry cultivars in Turkey. Pak. J. Bot. 2003;35(3): 361-367.
9. Serdar Ü, Demirsoy H. Non-destructive leaf area estimation in chestnut. Scientia Horticulturae. 2006;108(2):227-230.
10. Rouphael Y, Colla G, Fanasca S, Karam, F. Leaf area estimation of sunflower leaves from simple linear measurements. Photo-synthetica. 2007;45(2):306-308.
11. Kumar R. Calibration and validation of regression model for non-destructive leaf area estimation of saffron (*Crocus sativus* L.). Scientia Horticulturae. 2009;122(1): 142-145.
12. Ghoreishi M, Hossini Y, Maftoon M. Simple models for predicting leaf area of mango (L.). *Mangifera indica* Journal. 2012;45-53.
13. Fascella G, Darwich S, Rouphael Y. Validation of a leaf area prediction model proposed for rose. Chilean Journal of Agricultural Research. 2013;73(1):73-76.
14. Celik H, Odabas MS, Odabas F. Leaf area prediction models for highbush blueberries (*Vaccinium corymbosum* L.) from linear measurements. Adv Food Sci. 2011;33:16-21.
15. Salazar JCS, Melgarejo LM, Bautista EHD, Di Rienzo JA, Casanoves F. Non-destructive estimation of the leaf weight and leaf area in cacao (*Theobroma cacao* L.). Scientia Horticulturae. 2018;229:19-24.
16. Kucukonder H, Boyaci S, Akyüz A. A modeling study with an artificial neural network: Developing estimation models for the tomato plant leaf area. Turkish Journal of Agriculture and Forestry. 2016;40(2): 203-212.
17. Azadi S, Sepaskhah, AR. Annual precipitation forecast for west, southwest, and south provinces of Iran using artificial

neural networks. Theoretical and applied climatology. 2012;109(1-2):175-189.

18. Rad MRN, Fanaei HR, Rad MRP. Application of artificial neural networks to predict the final fruit weight and random forest to select important variables in native population of melon (*Cucumis melo* L.). Scientia Horticulturae. 2015;181:108-112.

19. Kaul M, Hill RL, Walthall C. Artificial neural networks for corn and soybean yield prediction. Agricultural Systems. 2005; 85(1):1-18.

20. Odabas MS, Ergun E, Oner F. Artificial neural network approach for the predicition of the corn (*Zea mays* L.) leaf area. Bulgarian Journal of Agricultural Science. 2013;19(4):766-769.

21. Vazquez-Cruz MA, Luna-Rubio R, Contreras-Medina LM, Torres-Pacheco I, Guevara-Gonzalez RG. Estimating the response of tomato (*Solanum lycopersicum*) leaf area to changes in climate and salicylic acid applications by means of artificial neural networks. Biosystems Engineering. 2012;112(4):319-327.

22. Pulido-Calvo I, Montesinos P, Roldán J, Ruiz-Navarro F. Linear regressions and neural approaches to water demand fore-casting in irrigation districts with telemetry systems. Biosystems Engineering. 2007; 97(2):283-293.

23. Moosavi AA, Sepaskha A. Artificial neural networks for predicting unsaturated soil hydraulic characteristics at different applied tensions. Archives of Agronomy and Soil Science. 2012;58(2):125–153.

24. Liu M, Liu X, Li M, Fang M, Chi W. Neural-network model for estimating leaf chlorophyll concentration in rice under stress from heavy metals using four spectral indices. Biosystems Engineering. 2010;106(3):223-233.

25. Salazar JCS, Melgarejo LM, Bautista EHD, Di Rienzo JA Casanoves F. Non-destructive estimation of the leaf weight and leaf area in cacao (*Theobroma cacao* L.). Scientia Horticulturae. 2018;229:19-24.

26. Mahanti, NK, Chakraborty, SK, Babu, VB. Sorption isotherms of ready-to-puff preconditioned brown rice: Development of classical models and artificial neural network approach. Journal of Food Process Engineering. 2019;42(6):e13220.

27. Panda BK, Chavan AN, Datta AK. Development of super-atmospheric oxidation chamber for orthodox tea processing and its validation through neural network approach. Journal of the Science of Food and Agriculture. 2019; 99(8):3752-3760.

28. Nakanwagi MJ, Sseremba G, Kabod NP, Masanza M, Kizito EB. Accuracy of using leaf blade length and leaf blade width measurements to calculate the leaf area of *Solanum aethiopicum* Shum group. Heliyon. 2018;4(12):e01093.

29. Oner F, Odabas MS, Sezer I, Odabas F. Leaf area prediction for corn (*Zea mays* L.) cultivars with multiregression analysis. Photosynthetica. 2011;49(4):637-640.

30. Keramatlou I, Sharifani M, Sabouri H, Alizadeh M and Kamkar B. A simple linear model for leaf area estimation in Persian walnut (*Juglans regia* L.). Scientia Horticulturae. 2015;184:36-39.

31. Mendoza-de Gyves E, Rouphael Y, Cristofori V, Mira FR. A non-destructive, simple and accurate model for estimating the individual leaf area of kiwi (*Actinidia deliciosa*). Fruits. 2007;62(3):171-176.

32. Ahmadian-Moghadam H. Prediction of pepper (*Capsicum annuum* L.) leaf area using group method of data handling-type neural networks. International Journal of AgriScience. 2012;2(11):993-999.

33. Küçükönder H, Boyaci S, Akyüz A. A modeling study with an artificial neural network: Developing estimation models for the tomato plant leaf area. Turkish Journal of Agriculture and Forestry. 2016;40(2): 203-212.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sdiarticle4.com/review-history/58371