Determination of the tomato (*Solanum lycopersicum* L)’s area foliar through artificial intelligence technique

Determinacion del área foliar de jitomate (*Solanum lycopersicum* L) mediante técnica de inteligencia artificial

MARTINEZ-RUIZ, Antonio†, QUINTANAR-OLGUIN, Juan*, PÉREZ-JIMÉNEZ Genaro and FLORES-DE LA ROSA Felipe Roberto

Instituto Nacional de Investigaciones Forestales Agrícolas y Pecuarias, Mexico.

ID 1st Author: *Antonio, Martinez-Ruiz* / **ORC ID**: 0000-0001-6555-4651, **CVU CONACYT ID**: 364739

ID 1st Co-author: *Juan, Quintanar-Olguin* / **ORC ID:** 0000-0003-2388-5027, **CVU CONACYT ID**: 203741

ID 2nd Co-author: *Genaro Pérez-Juménez* / **ORC ID:** 0000-0003-0403-0189

ID 3rd Co-author: *Felipe Roberto, Flores-De la Rosa* / **ORC ID:** 0000-0002-1435-936X, **CVU CONACYT ID**: 535996

DOI: 10.35429/JANRE.2021.9.5.1.6 Received: July 30, 2021; Accepted December 30, 2021

Abstract

Studying leaf area (LA) is of vital importance in study of the interception of light, since it has an influence on respiration, dry matter production and directly with consumption of water and nutrients of a plant. Applying artificial neural networks (ANN) as an indirect method to estimate LA has proven to be an accurate tool. In present work, a backpropagation multilayer artificial neural network (ANN) was trained and evaluated to predict LA in a hydroponic tomato crop in a greenhouse. Data collection was carried out by means of a random sampling every 15 days of four plants, which the length and width of each leaf were measured, and by means of an integrator LI-3100, LICOR leaf area was obtained. With this information, the most efficient network structure was searched, the best ANN was trained, validated, and tested. The best neural network structure was obtained with an input variable, a hidden layer with 5 neurons, applying a non-linear activation function of sigmoidal tangent type, where input variable combining characteristic dimensions of leaf was the most efficient.

Machine learning, Artificial neural networks, Simulation models

Resumen

Estudiar el área foliar es de vital importancia en el estudio de la intercepción de la luz, ya que tiene influencia en la respiración, producción de materia seca y directamente con el consumo de agua y nutrimentos de una planta. Aplicar las redes neuronales artificiales (ANN) como método indirecto para estimar el área foliar, ha mostrado ser una herramienta precisa. En el presente trabajo se entrenó y evaluó una red neuronal artificial multicapa (ANN) de retro propagación de los errores, para predecir el área foliar (AF) en un cultivo de jitomate hidropónico en invernadero. La toma de datos se realizó mediante un muestreo aleatorio cada 15 días de cuatro plantas, a las que se midió el largo y ancho de cada hoja, y mediante un integrador LI-3100, LICOR se obtuvo el área foliar. Con esta información se buscó la estructura de red más eficiente, se entrenó, valido y probó la ANN. La mejor estructura de red neuronal se obtuvo con una variable de entrada, una capa oculta con 5 neuronas, aplicando una función de activación no lineal del tipo tangente sigmoidal, donde la variable de entrada combinando las dimensiones características de la hoja resultó la más eficiente.

Aprendizaje de máquina, Redes neuronales artificiales, Modelos de simulación

Citation: MARTINEZ-RUIZ, Antonio, QUINTANAR-OLGUIN, Juan, PÉREZ-JIMÉNEZ Genaro and FLORES-DE LA ROSA Felipe Roberto. Determination of the tomato (*Solanum lycopersicum* L)’s area foliar through artificial intelligence technique. Journal-Agrarian and Natural Resource Economics. 2021. 5-9:1-6.

* Correspondence to Author (Email: quintanar.juan@inifap.gob.mx)
† Researcher contributing as first author.

Introduction

The leaves are the active frontier for the exchange of carbon and water between plants, canopy and the atmosphere, this green organ helps in feeding plants through photosynthesis and evapotranspiration. The leaf area index (LAI) is a dimensionless variable, it represents a structural attribute of the components of the leaves, it has vital importance in the interception of light which has an influence on respiration, dry matter production, yield and directly with the consumption of water and nutrients, plant growth and development (Firouzabadi *et al*. 2015; Öztürk *et al*., 2019). The LA stimulates the amount of light that enters through the canopy and influences the microclimate; therefore it is used as an indicator of the health of the plant and its development. Therefore, the study of LA or LAI is essential and highly significant in horticultural and agronomic physiological studies. Studies have been carried out that focus on the evaluation of LA using direct and indirect techniques. Although the direct methods are precise, the samplings involve destroying the plants, it is a hard work activity (Colaizzi *et al*., 2017), and complicated as the experimentation area increases. Therefore, nowadays the attention of researchers is focused on looking for indirect methods that avoid the destruction of plants (Hossain *et al*., 2017). Some methods used are, for example, images analysis (Campillo *et al*., 2010), mapping PAR radiation (Zhang *et al*., 2015), use of ceptometers (Mendoza-Pérez *et al*., 2017), others have applied allometric models relating characteristic dimensions of the leaves (Astegiano *et al*., 2001; Colaizzi *et al*., 2017).

In recent years artificial neural networks (ANNs) have been tested within agricultural systems with successful results. Yuan *et al*. (2017) evaluated a neural network that can be used to estimate the LA in soybean crops. On the other hand, Kumar *et al*. (2017) developed an artificial neural network (ANN) based on the characteristic lengths of the leaves as they are; length (L) and width (W) which was compared with regression models, finding greater efficiency of the network compared to regression models. On the other hand, Ahmadian-Moghadam *et al*. (2012) looked for an ANN to calculate the LA of the cucumber crop, finding good adjustments.

Küçükönder *et al*. (2016) used an ANN and analysis of regression techniques to develop the best regression model and conclude that ANN models are an excellent alternative to estimate LA and Shabani *et al*. (2017) used an ANN to estimate the LA of different plants and found that ANN gave acceptable results. The objective of this work was to find the best structure of a multilayer backpropagation artificial neural network (ANN). Training, testing and validation with this network was carried out to predict the LA of a hydroponic tomato crop (*Solanum lycopersicum* L.) in a greenhouse.

Materials and method

Establishment of the experiment

The experiment was carried out in a saw-type greenhouse, with overhead ventilation, located at the Universidad Autónoma Chapingo, with coordinates 19º 29’ North latitude, 98º 53 ’West longitude, latitude 2240 m. The greenhouse has an N-S orientation with a 700-gauge white plastic cover treated against ultraviolet radiation, with three section (8.5 mx 76 m) a total area of 1938 m^2, three zenith windows, two lateral and two frontals, all with anti-aphid mesh of 25 x 40 threads / in^2. The opening of the vents was carried out in an automated way. The test was carried out in a hydroponic tomato crop (*Solanum lycopersicum* L.), *cv*. "*Rafaelo*", which was germinated on April 7 and transplanted on May 7 with a density of 2.6 plants / m^2, 13-liter pots were used with a combination of tezontle substrate and coconut fiber (70:30). A drip irrigation system with inserted self-compensating drippers was installed, with an injection system and automatic irrigation programming.

Data collection

The LA variable of the crop was measured every 15 days, for this purpose a destructive method was applied, consisting of selecting 4 plants randomly and they were taken to the laboratory, where the characteristic dimensions like length (L) and width (A) of each leaf were measured at the same time with the use of a model leaf area integrator (LI3100, LICOR) the leaf area of each leaf was measured.

ISSN 2524-2091
RINOE® All rights reserved.

Structure of the neural network

With the data of length, width and leaf area, the best structure of the neural network (Figure 1) of the backpropagation was searched, having as input variables the characteristic dimensions of the leaves; length (L) and width (W) and the different combinations between these variables, length^2 (L^2), width^2 (W^2), length × width (L × W), length^2 × width (L^2 × W) and length × width^2 (L×W^2). The data was organized as follows; For the training of the network, 50% of the data was used, for the test 25% and the validation 25%, for several data (n = 299). A multilayer network of back propagation of the errors (back propagation neural network) was used. To perform the data analysis, they were normalized to a value between 0 - 1 following Equation (1). The procedure consisted of finding the most efficient network that could make the best forecast of the LA variable once trained with a set of data, for which an independent data block (n = 110) was used to perform the validation with the purpose of corroborating the prediction capacity of the network structures. The combination of input with one and two hidden layers were tested, varying the number of neurons per layer from 1 to 7 neurons.

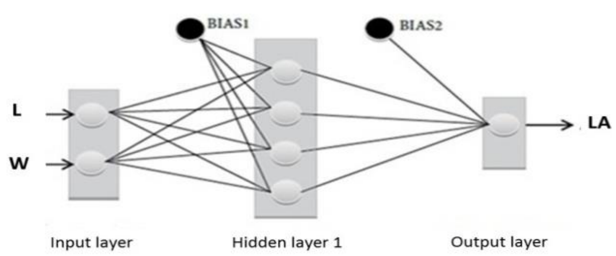

Figure 1 Architecture of a neural network to estimate the leaf area, for two input variables length (L) and width (W), a hidden layer and an output variable (leaf area, LA).

$$Xnorm = \frac{X - Xmin}{Xmax - Xmin} \quad (1)$$

Where $Xnorm$ is the value of the normalized variable (LA), X is the original variable to normalize (LA), $Xmax$ y $Xmin$ are the minimum and maximum original variables (LA). During network training the input values to the network are calculated as follows:

$$Net = \sum_{i=1}^{N} x_i w_i + \quad \theta_i \quad (2)$$

In equation (2), θ_i is the bias value, x_i are the values of the input variable (s), w_i are the values of the weights corresponding to the ith value. To calculate the output values of the network, the sigmoidal function was used as the activation function described in Equation (3).

$$F(net) = \frac{1}{1+e^{(-Net)}} \quad (3)$$

To decide which network structure was the best during training, validation and testing, the correlation coefficient (R) was estimated, to validate the structure of the networks that resulted best, the adjustment statistics were evaluated in addition to the correlation coefficient, the root mean square of error (RMSE) and bias (BIAS) between values measured in the laboratory and those estimated by the neural network.

Results and discussion

As mentioned in the previous section, for the training of the neural network, one and two hidden layers were evaluated, progressively varying the number of neurons per layer, and combining the input variables of the characteristic dimensions of the leaves of tomato cultivation. Where it was found that the most efficient network and with the highest value of the correlation coefficient was for when the structure of the network was formed by a hidden layer with 5 neurons, the adjustments of the measured data versus estimated during the training, validation and test are presented in Figure 2. Where it is evidenced that the input variable that best adjusts the values during the forecast is length × width^2 (L × W^2), considering this combination as an input variable, it is important to mention that the network was also evaluated considering two inputs: the length and the width. However, no good fits were found, so it was decided to use an input variable, and this is given by the different combinations of the characteristic dimensions. The correlation coefficient values in this research (Figure 2) were slightly lower for training (R = 0.98), test (R = 0.97) and all (R = 0.98), found by Küçükönder *et al*. (2016).

ISSN 2524-2091
RINOE® All rights reserved.

Table 1 shows the results found during the validation of the different structures of the neural network with different neurons, with one and two hidden layers modifying the combinations in the input variable. Where the best fit statistics resulted for when the network was tested with an input variable f (L × W^2) ** and {5} neurons, because the bias values (BIAS) and RMSE were lower and the value of the correlation coefficient (R) was higher in all cases, in descending order, it was for the structure of the network in which the input variable was given by f (L × W) with {7} neurons and finally the third option was when two input variables given by; f (W, L × W^2), with {5} neurons.

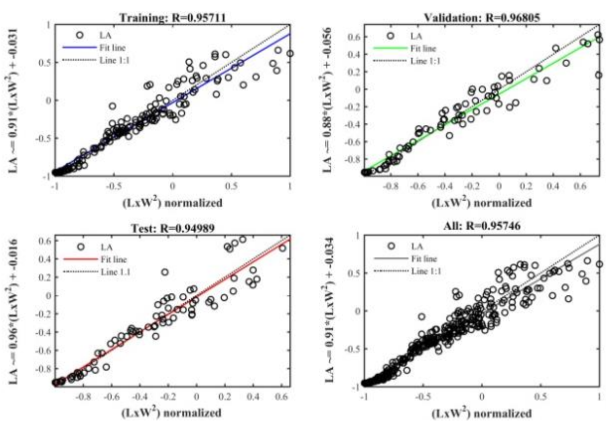

Figure 2 Training, validation and testing of the artificial neural network for the prediction of the leaf area of the tomato crop for a network structure of a hidden layer with five neurons {5}, f (L * W^2)

The predictions for LA given by the best structure of the network were found that the values were similar for the correlation coefficient and very high for the RMSE values (0.11) to those found by Küçükönder *et al*. (2016) for a network of a hidden layer with 4 neurons, in tomato leaves. On the other hand, Ahmadian-Moghadam *et al*. (2012) found that a network of a hidden layer with 2 neurons was the network that best predicted the leaf area in the pepper crop (*Capsicum annuum* L.).

Inputs variables	One hidden layer (neurons)	Statistics of fiting		Two hidden layer (neurons)	Statistics of fiting	
f(L)	{5}	RMSE	89.784	{1 x 7}	RMSE	90.669
		BIAS	-46.845		BIAS	-47.597
		R	0.940		R	0.941
f(W)	{6}	RMSE	69.766	{1 x 6}	RMSE	72.833
		BIAS	-25.450		BIAS	-23.915
		R	0.960		R	0.954
f(L×W)	{7}	RMSE	80.263	{1 x 5}	RMSE	59.084
		BIAS	-13.642		BIAS	-18.168
		R	0.936		R	0.969
f(L×W^2) **	{5}	RMSE	56.890	{1 x 3}	RMSE	56.516
		BIAS	-8.255		BIAS	-9.151
		R	0.967		R	0.968
f(L^2×W)	{6}	RMSE	69.962	{1 x 3}	RMSE	67.619
		BIAS	-29.821		BIAS	-28.532
		R	0.963		R	0.966
f (A, L×W^2)	{5}	RMSE	57.201	{1 x 6}	RMSE	58.447
		BIAS	-13.769		BIAS	-9.046
		R	0.968		R	0.967

Table 1 Statistics of goodness of fit of the validation's performance of the artificial neural network (ANN) for predicting the LA of the tomato crop

When the network was structured using two hidden layers, good adjustments were obtained, although slightly lower than when a single layer was used, the best combination for this case were; in which two input variables f (W, L × W^2) were used with {1 x 6} neurons in the first and second hidden layer, f (L × W^2) with {1 x 3} neurons, In this research work, the evaluations of the network with two hidden layers were carried out just to know if it was possible to find better adjustments since, as is known, as a neural network has more inputs or outputs or more than two hidden layers and number of neurons the complexity of the network structure increases, surely when the problem to be solved is more complex it is justifiable to increase the number of hidden layers and / or the number of neurons. It has been found that using the characteristic lengths improves the prediction of the variable in question, according to the results presented by Astegiano *et al*. (2001) and Küçükönder *et al*. (2016) when evaluating allometric models to evaluate the LA for tomato crop. In Figure 3, the scatter plot of the estimated data of the 6 best configurations of neural networks is shown with an input variable and a hidden layer, which resulted with the best fit in the training, validation and test, in addition the values are presented measured versus values that resulted from the forecast, where it is observed that the data from the neural network that most closely approximates the 45º line (line 1: 1) are those that are similar to the values measured in the laboratory, therefore a lower bias value (BIAS) and those that are concentrated around this 45º line have a lower error (RMSE). Where Figure 3C), 3D) and 3F) correspond to f (L × W^2) ** y {5} neurons, f (L × W) with {7} neurons and f (W, L × W^2), with {5} neurons, respectively.

Öztürk *et al*. (2019) also found excellent fits to using a similar procedure to predict LA in 13 commercial species. While Shabani *et al*. (2017) report surprising results when applying artificial intelligence techniques to predict different crops with different leaf shapes, using the characteristic dimensions of each species.

ISSN 2524-2091
RINOE® All rights reserved.

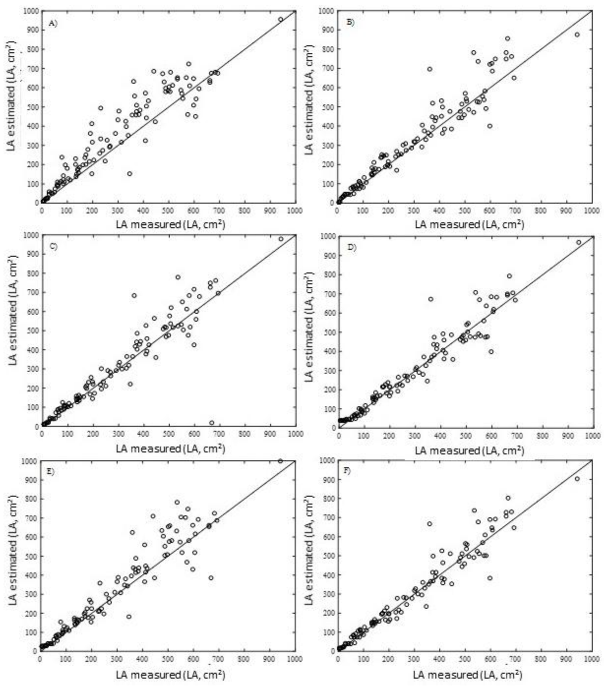

Figure 3 Prediction of the leaf area (FA) of the tomato crop by validating the following configurations of the artificial neural network: A) *f* (Length), B) *f* (Width), C) *f* (Length × Width), D) *f* (Length × Width^2), E) *f* (Length^2 × width), F) *f* (Width, Length × Width^2) and line 1: 1

Conclusions

Artificial neural networks (ANN) are a useful tool to estimate the leaf area of a crop with an acceptable precision, since in most cases they have devices such as leaf area integrators or direct meters of the leaf area index in the field. It is not possible due to how expensive they can be and another disadvantage of these devices is that they are easily decalibrated, and sometimes using them leads to knowing several parameters that are often unknown for a particular crop.

However, it is necessary to test the efficiency of the ANN networks that were better in this work with information from another crop cycle and with other tomato varieties. In the case where the networks with two hidden layers were evaluated and the number of neurons changed, good adjustments were also achieved, however, in order to solve a problem that lacks considerable complexity since it is only interesting to estimate only one subject output variable to a small number of inputs, it is justifiable to use neural networks with a hidden layer.

References

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

Astegiano, E. D., Favaro, J. C., & Bouzo, C. A. 2001. Estimación del área foliar en distintos cultivares de tomate (Lycopersicon esculentum Mill.) utilizando medidas foliares lineales. Investigación Agraria: Producción y Protección Vegetales, 16(2): 249-256.

Campillo, C., Garcia, M. I., Daza, C., & Prieto, M. H. 2010. Study of a non-destructive method for estimating the leaf area index in vegetable crops using digital images. Hortscience, 45(10): 1459-1463.

Colaizzi, P. D., Evett, S. R., Brauer, D. K., Howell, T. A., Tolk, J. A., & Copeland, K. S. 2017. Allometric method to estimate leaf area index for row crops. Agronomy Journal, 109(3): 883-894.

Ghadami Firouzabadi, A., Raeini-Sarjaz, M., Shahnazari, A., & Zareabyaneh, H. 2015. Nondestructive estimation of sunflower leaf area and leaf area index under different water regime managements. Archives of Agronomy and Soil Science, 61(10): 1357-1367.

Hossain, S. A. A. M., Lixue, W., Taotao, C., & Zhenhua, L. 2017. Leaf area index assessment for tomato and cucumber growing period under different water treatments. Plant, Soil and Environment, 63(10): 461-467.

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

Mendoza-Pérez, C., Ramírez-Ayala, C., OjedaBustamante, W., & Flores-Magdaleno, H. 2017. Estimation of leaf area index and yield of greenhouse-grown poblano pepper. Ingeniería agrícola y biosistemas, 9(1): 37-50.

Öztürk, A., Cemek, B., Demirsoy, H., & Küçüktopcu, E. 2019. Modelling of the leaf area for various pear cultivars using neuro computing approaches. Spanish Journal of Agricultural Research, 17(4).1-11.

RINOE® All rights reserved.

Shabani, A., Ghaffary, K. A., Sepaskhah, A. R., & Kamgar-Haghighi, A. A. 2017. Using the artificial neural network to estimate leaf area. Scientia Horticulturae, 216: 103-110.

Zhang, J., Zhang, Q., & Whiting, M. D. 2015. Mapping interception of photosynthetically active radiation in sweet cherry orchards. Computers and Electronics in Agriculture, 111: 29-37.

ISSN 2524-2091
RINOE® All rights reserved.