نشریه مکانیزاسیون کشاورزی/ دوره ۶ شماره ۳/ سال ۱۴۰۰/ صفحه‌های ۶۵-۵۹

# Prediction of Chlorophyll Content of Tomato Plant by Artificial Neural Networks and Adaptive Nero-Fuzzy Inference System

Vali Rasooli Sharabiani[1]*, Asma Kisalaei[1], Ebrahim Taghinezhad[2]

Received: 13 Feb 2021 Accepted: 28 Aug 2021

[1] Department of Biosystems Engineering, Faculty of Agriculture and Natural Resources, University of Mohaghegh Ardabili, Ardabil, Iran
[2]Department of Agricultural Technology Engineering, Moghan College of Agricultural and Natural Resources, University of Mohaghegh Ardabili, Ardabil, Iran
*Corresponding authors: vrasooli@uma.ac.ir

**Abstract**

Approximately three-quarters of harvested tomatoes are freshly used. Good quality is an important factor in distributing of fresh tomato. Chlorophyll is the green chemicals to provide required food of plants and ensure plant growth and productivity. The main function of chlorophyll is to absorb blue and red lights and perform photosynthesis. In recent years, the tendency to use of prediction methods such as soft computing and artificial intelligence for growth of plans has increased. The main aim of this study was to investigate the relationship between height and chlorophyll content in the leaves of tomato plants using modeling and predicting techniques and compare the accuracy of these methods. In this study, some cultivated plants of tomato were randomly selected for height and SPAD measurements. The results showed the relationship between Chlorophyll content and height of plants was very low ($R^2$ = 0.276). However using the modelling of ANN and ANFIS improved the prediction power up to ($R^2$=0.982 and 0.913), respectively.

**Keywords:** ANFIS, Chlorophyll Content, Modeling, Neural Networks, Tomato

# پیش‌بینی محتوی کلروفیل گیاه گوجه‌فرنگی با استفاده از شبکه‌های عصبی و انفیس

ولی رسولی شربیانی[۱]*، اسما کیسالایی[۱] و ابراهیم تقی نژاد[۲]

تاریخ دریافت: ۹۹/۱۱/۲۵ تاریخ پذیرش: ۱۴۰۰/۶/۶

۱- گروه مهندسی بیوسیستم، دانشکده کشاورزی و منابع طبیعی، دانشگاه محقق اردبیلی
۲- گروه مهندسی فناوری کشاورزی، دانشکده کشاورزی و منابع طبیعی، دانشگاه محقق اردبیلی
* مسئول مکاتبه: vrasooli@uma.ac.ir

**چکیده**

**تقریبا سه چهارم کل گوجه فرنگی تولید شده در جهان بصورت تازه مورد مصرف قرار می گیرد. کیفیت خوب برای توزیع گوجه فرنگی معیار مهمی است. کلروفیل یک ماده شیمیایی سبز رنگ برای تامین غذای مورد نیاز گیاه و تضمین رشد و بهره وری گیاه است. وظیفه اصلی کلروفیل جذب نورهای آبی و قرمز و انجام فتوسنتز است. در سال های اخیر تمایل به استفاده از روش های پیش بینی مانند محاسبات نرم و هوش مصنوعی برای پایش رشد گیاهان ا افزایش یافته است. هدف اصلی این تحقیق بررسی رابطه ارتفاع و محتوای کلروفیل در برگ گیاه گوجه فرنگی با استفاده از تکنیک های مدل سازی و پیش بینی و مقایسه دقت این روش ها بود. در این تحقیق تعدادی از بوته های گیاه گوجه فرنگی برای اندازه گیری ارتفاع و SPAD به طور تصادفی انتخاب شدند. نتایج نشان داد که ارتباط بین میزان کلروفیل و ارتفاع گیاه بسیار کم است ($R^2$=۰/۲۷۶). با این حال، استفاده از مدل‌سازی ANN و ANFIS، قدرت پیش‌بینی را به ترتیب تا ($R^2$=۰/۹۸۲ و $R^2$= ۰/۹۱۳) افزایش داد.**

***واژه‌های کلیدی:*** **محتوی کلروفیل گیاه، گوجه فرنگی، مدل سازی، شبکه های عصبی، انفیس**

**How to cite:**
Rasooli Sharabiani, V., Kisalaei, A., Taghinezhad, E. 2021. *Prediction of Chlorophyll Content of Tomato Plant by Artificial Neural Networks and Adaptive Nero-Fuzzy Inference System.* Journal of Agricultural Mechanization 6 (3): 59-65.

## Introduction

Tomato with scientific name Lycopersicon esculentum, is produced in large quantities in the world. In 2012, the total production of tomato was estimated at 162 Mtons (FAO, 2012). The most of tomato is consumed freshly and also processed in some products such as extract, nymphs, sauces and various canned beans (Akanbi, Adeyemi *et al.*, 2006). In addition, dried tomato is used as a component in pizza and varieties of vegetables and spices. Several studies have been done on the physical properties of this product. These include the study conducted to construct and control of tomato harvesting robots. Li et al. (2011) have investigated the physical properties of tomato such as height, diameter, surface area of product, product volume, product mass, density or product density, shell density or gelatin material, shape factor, porosity and curvature of two tomato cultivars by visual analysis and displacement method water. In another study, multi-product engineering properties related to harvesting simulation and tissue evaluation of two tomato cultivars were studied simultaneously in six stages of the study period. The product yield potential scale was proposed based on the R: G: B ratio for a single stage of processing (Li *et al.,* 2015). Chlorophyll is a green chemical substance that provides the food needed by plants. The main function of the costume is to absorb blue and red lights and perform the process of photosynthesis by presence of chlorophyll (Home and Goldman, 1994). All data and information about the properties of cultivated products and food are valuable and important, because the information are an input for quality prediction and product behavior models which could be related with any measured factor (Arazuri, *et al.,* 2007). The tendency to predictive methods, including software calculations and artificial intelligence, has increased in recent years. One of these methods is the fuzzy method with benefits of looking at ambiguous data. The fuzzy method provides a linguistic logic for modeling advanced systems (Soyguder *et al.*, 2009). Another method is artificial neural network a computer designed programs for modeling and investigating the relationship between dependent and independent variables. Both methods are combined in ANFIS method. ANFIS uses a hybrid learning algorithm to identify the parameters of the sugen fuzzy inference system. This system uses partial least squares and back propagation algorithms methods to train the membership parameters uses to simulate a set of training data (Soyguder and Alli, 2009). Also the temperature of the sea surface, the air temperature and daylight hours were the input of the model and the only output of the model was chlorophyll content. The network was trained using 20 neurons in the hidden layer. In another study by Nagamani *et al.* (2007). In this study, the aim of the study was to determine the relationship between tomato plant height and the amount of chlorophyll present in tomato leaves using modeling and prediction methods as well as comparing the accuracy of these methods. The present study consists of three parts. The first part in the measurement and analysis of required data. In the second part, modeling and forecasting methods are designed using required data. In the third part, the analysis and comparison of these methods and suggestion of the best method are carried out using comparative parameters.

## Material and methods

Tomato plants were initially transplanted and after sufficient growth transferred into experiment pots. The number of pots was 18 out of which 31 plants were randomly selected and the chlorophyll was measured. For this measurement a chlorophyll meter (SPAD 502) was used at 3 different heights per plant and the result was recorded for each height. MATLAB 2013 software was used to develop a radial basis function neural network modeling. Some of recorded data were removed as irrelevant data to improve the accuracy of the models. After the data gathering step, predictive methods were used for the modeling.

### Artificial neural network method (Radial base function)

In recent years, the power of the algorithm for learning the neural network has led engineers to tend to find this soft computing method in most of their applications (Wen *et al.*, 2012). Among the different types of neural networks, radial base has become very popular due to its ability to directly estimate complex nonlinear mappings (Moody and Darken, 1989). The purpose of this research is modeling the estimation of chlorophyll versus tomato plant height.

Neural network with radial base function is a three-layer feedforward neural network consisting of input, hiding and output layers (Fig 1).

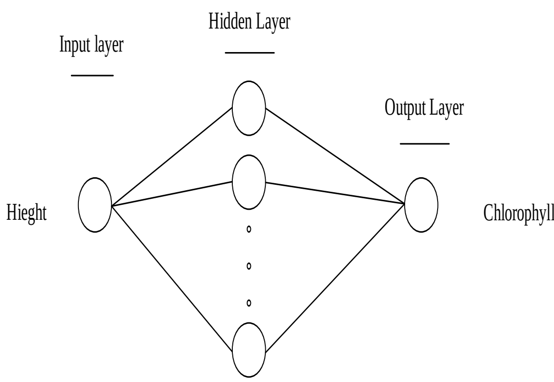


**Fig 1. The structure of RBF network**

Each of the neurons in the hidden layer is associated with a Gaussian function. Each of the output neurons calculates a simple weighted sum of the hidden layer responses for their respective inputs. Each of the output neurons calculates a simple weighted sum of the hidden layer responses for their respective inputs. The input layer in the RBF network is a non-linear mapping of X to the corresponding formula presented in equation (1):

$$\phi_i(x) = \exp\left[ -\frac{\| x - c_i \|^2}{2\sigma^2} \right]_{i=1,2,\dots,p} \qquad (1)$$

Where: $c_i$ and $\sigma_i$ are the center and width of the i^th^ latent layer, respectively.

Neural network output with radial base function for linear mapping realizes from ϕ to y. This mapping is calculated as equation (2) and outputs the network:

$$Y = \sum_{i=1}^{r} w2_{ik}\, \phi_i(x)_{;k=1,2,\dots,m} \qquad (2)$$

K is output layer nodes, the weight refers to the relationship between the hidden layer and the output layer and the hidden layer response is for a node of the output layer. In fact, the neural network with radial base function performs the following operations for learning the network:

1. Learning algorithm without supervision to learn the centers and width of the main functions,
2. Assigning weights related to the relationship between hidden and output units.

The corresponding inputs and outputs obtained in the data entry stage were used to train the network. MATLAB 2013 software was used to develop a radial basis function neural network. To determine the optimal number of neurons in the hidden layer, the network was trained with a neuron in the hidden layer. After training with a neuron in the hidden layer, the network performance charts were extracted. The errors have been reduced by adding neurons to the hidden layer. This continued until there was no change in the network error. The number of neurons as the optimal number of neurons was 15.

## Adaptive neuro-fuzzy inference system (ANFIS)

ANFIS is a neural-fuzzy network with 5 layers as shown in Fig 2. Each of the nodes in each layer is defined by the function corresponding to that node. This network is a Feed forward neural network used for non-linear predictions. In this network, Takagi-Sugeno model uses fuzzy inputs to predict completely complex nonlinear problems. ANFIS uses the least squares algorithm to identify and apply learning-related parameters.

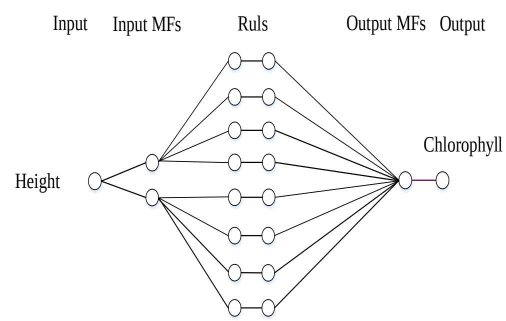


**Fig 2. The structure of Neuro-Fuzzy system (ANFIS)**

The main purpose of using ANFIS is to optimize the parameters of the fuzzy inference system by using the learning method available in the input-output pairs. Data from data-gathering process was used to teach the neural-fuzzy network. It was in a way that the network with best estimates and predictions close to real values was designed. The specifications and features related to ANFIS which are the type and number of membership functions and network training algorithm, are the main determinant of the performance of this network. This network was trained with different types of membership functions with different numbers. Each MSE instruction (relationship 1) which is provided by the system itself is considered to select the best network structure. The best result (lowest MSE) was obtained in the Gaussian type membership function with 2 numbers in the input and trained with the same network structure.

In order to compare the results of the network output and the actual output as well as accuracy and performance designed networks comparative parameters including root mean square error (RMSE), correlation coefficient (R) and mean absolute error (MAE) were used (equations 3-6).

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(A-P)^2 \quad (3)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(A-P)^2} \quad (4)$$

$$R = \left(1-\left(\frac{\sum_{i=1}^{n}(A-P)^2}{\sum_{i=1}^{n}A_i^2}\right)\right)^{1/2} \quad (5)$$

$$MAE = \frac{\sum_{i=1}^{N}|A-P|}{N} \quad (6)$$

Where: A represents the actual value, P is the predicted value, and N is the number of data.

The use of these parameters allows us to find out how close the amount predicted by the network is to the actual value.

## Results and discussion

The purpose of this study was to find the relationship between the plant height of tomato crop and the amount of chlorophyll content in the plant leaves. According to a report by Zhang *et al.* (2014) there is a very strong relationship between chlorophyll and nitrogen amount. The amount of SPAD-value is always used to determine the amount of nitrogen and chlorophyll and to assess the health of the plant. In a study commissioned by Azizi *et al.* (2011) on soy beans, it was concluded that there is a direct relationship between chlorophyll content, photosynthesis rate and nitrogen content of the plant. The results of the data processing operation are shown in Fig 3. In this graph, the vertical axis represents the chlorophyll values and the horizontal axis shows the plant height values. Accordingly, the correlation between chlorophyll and plant height is 0.52 (with a coefficient of determination equal to 0.276), increasing the height increases the chlorophyll content of the plant, which increases with a slope of 0.4478.

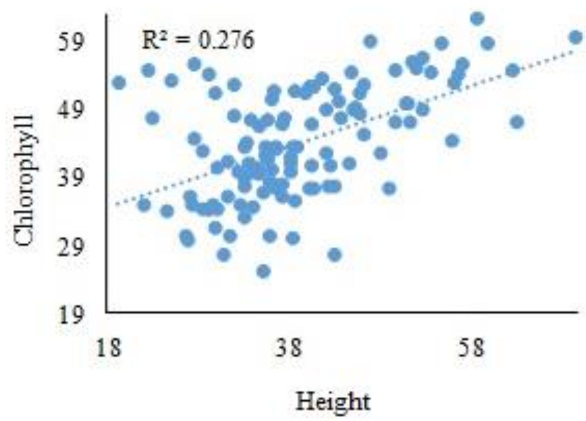


**Fig 3. Data graph of tomato plant chlorophyll based on plant height**

With further study on the relationship between chlorophyll and plant height of tomatoes by using linear regression analysis, Table 1 was extracted. According to this table, the relationship between height and chlorophyll was significant at a probability level of 1% with a standard deviation of 8.87 and a mean error of 0.8166.

**Table 1. The relationship between chlorophyll and tomato plant height using linear regression analysis**

| Parameter | Correlation Coefficient | Standard deviation | Mean absolute error | F |
|---|---|---|---|---|
| Height-Chlorophyll | 0.52 | 8.87 | 0.8166 | 44.174 |

Several studies have been done to estimate the chlorophyll content of plants. In a study by Samli *et al.* (2014), a feed forward neural network model was used to estimate the amount of chlorophyll using input data to the network. Multi-Layer Perceptron (MLP) was used to estimate chlorophyll using satellite data obtained. In this study, the network was trained using reflection spectra of known chlorophyll content. After the training process, all weights belonging to neurons were frozen and data related to the testing stage and network training were applied to the network in order to extract the output of the model that were applied to the network.

The power of network generalizability is by evaluating the network output relative to the data not applied in the network's training phase. Since these methods are non-parametric methods, performing this step to estimate the power of generalizability of networks is an inevitable way. The networks were trained using input and output with ANFIS and neural network with radial base function. The test phase of the networks created after the training of the networks. Network test data applied as network input and network outputs extracted. The relationship between the data extracted from the networks, which refers to output data from the test stage. And actual data were investigated using comparative parameters including root mean square error (RMSE), correlation coefficient (R) and mean absolute error (MAE). The Fig 4 shows the results of the modeling using a radial base neural network.

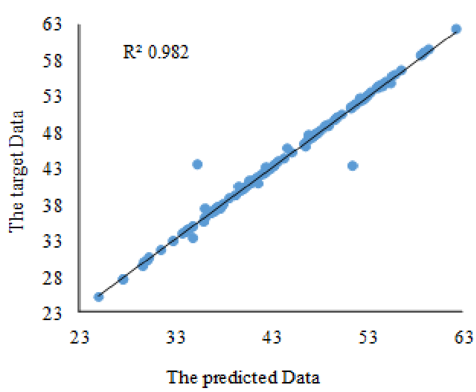


**Fig 4. The output results of RBF network**

In these charts, the data on the horizontal axis are predicted data or the same network output, and the data on the vertical axis are actual values or, in fact, the desired test values. In Fig 5, the output of the neural-fuzzy network is plotted against actual data. The most important indicators that can be measured at this stage are mean squared error (RMSE), coefficient of correlation (R) and mean square error (MSE) and the results are presented in Table 2. In a study by Samli *et al.* (2014), the correlation coefficient between predicted and actual data was 0.91.

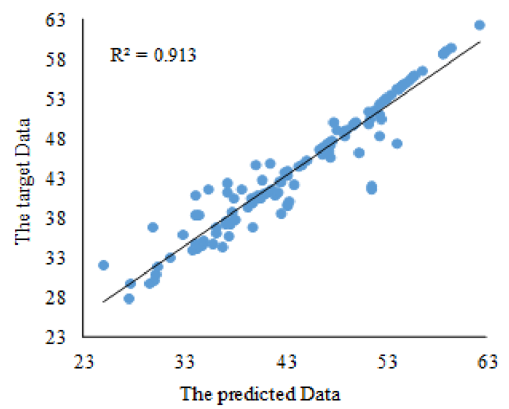


**Fig 5. The obtained results of the output of Neuro-Fuzzy system**

Nagamani *et al.* (2007) also reported that the coefficient of determination of training data was equal to 0.99 and the root mean square error was 0.07, respectively, these values for test data were reported to be 0.86 and 0.13. In the present study, the coefficient of determination for both methods is greater than 0.9.

**Table 2. The obtained results of comparison parameters**

| | RBF | ANFIS |
|---|---|---|
| R | 0.99 | 0.95 |
| RMSE | 1.104 | 2.44 |
| MAE | 0.242774 | 1.3008 |

Thus, it can be claimed that there is a high correlation between the target values and the predicted values for the dependent variables in the proposed networks. According to the graphs, the linear relationship between the predicted and measured values with a slope of 45 degrees and a width from the origin is approximately equal. The parameter R represents the Pearson correlation coefficient which measures the linear correlation between the two random variables. The RMSE parameter represents the root mean square error which is the difference between the predicted value of the model and the actual value, and is a good tool for comparing the predicted errors by the dataset. MAE also represents the mean absolute error. The higher $R^2$ values and the lower RMSE, showed the higher accuracy for prediction of the actual values (Faizollahzadeh Ardabili *et al*, 2016). According to Table 2 we can conclude that the neural network with the radial base function has the lowest RMSE and MAE (1.104 and 0.2427, respectively) and the highest R (0.99) compared to ANFIS. According to the results reported by other researchers, the use of artificial neural networks increases the system's predictive power and gives high accuracy in output generation.

## Conclusion

The purpose of this study was to model the amount of chlorophyll in tomato plants versus plant height. Modeling and prediction operations were performed using neural network models with radial and invisible base function based on input and output data. The results obtained from the network output were compared with comparative parameters such as R, RMSE and MAE with actual results. The results indicated that the neural network with a radial base function has high R values and low RMSE and MAE values. Therefore, it can be concluded that a neural network with a radial base function in the design of the network predicts the chlorophyll content, which is directly related to the amount of nitrogen, better than the plant height. Unlike other methods, the neural network allows even independent parameters to be included in the processing, in addition to the more limited factors in the estimation of participation. Network testing for the data that has not even learned the network is an important criterion for network verification. The correlation for data that the network has not even learned them in the training stage, and are only participated in the testing stage, can be an acceptable coefficient for network generalization.

## References

Akanbi, C.T., Adeyemi, R.S. and Ojo, A. (2006). *Drying characteristics and sorption isotherm of tomato slices.* Journal of Food Engineering, 73: 141-146.

Arazuri, S., Jaren, C., Arana, J. I., Perez, and Ciriza, J. J. (2007). *Influence of mechanical harvest on the physical properties of processing tomato* (Lycopersicon esculentum Mill.). Journal of Food Engineering. 80(1): 190-198.

Azizi, G., Alimardani, L. and Siahmargoei, A. (2011). *Evaluation of Relation of chlorophyll meter's number with chlorophyll content, photosynthesis and nitrogen content of soybean's leaf.* (6)23. 34-40.

Faizollahzadeh Ardabili, S., Mahmoudi, A. and Mesri Gundoshmian, T. (2016). *Modeling and simulation controlling system of HVAC using fuzzy and predictive (radial basis function, RBF) controllers.* Journal of Building Engineering. (6): 301–308.

FAO. (2010). *Pyrrolizidine alkaloids in foods and animal feeds. FAO Consumer Protection Fact Sheets.* (2): 1-6.

FAO, W. (2012). *The state of food insecurity in the world.* 8-11.

Li, Z. (2011). *Physical and mechanical properties of tomato fruits as related to robot's harvesting.* Journal of Food Engineering 103(2): 170-178.

Li, Z., Ly, K., Wang, Y., Zhao, B. and Yang, Z. (2015). *Multi-scale engineering properties of tomato fruits related to harvesting, simulation and textural evaluation. Food Science and Technology.* 61(2): 444-451.

Moody, J. and Darken, C. J. (1989). *Fast learning in networks of locally-tuned processing units.* Neural computation. 1(2): 281-294.

Nagamani, P. V., Chauban, P. and Dwivedi, R. M. (2007). *Estimation of chlorophyll-A concentration using an artificial neural network (ANN) based algorithm with oceansat-I OCM data. Journal of the Indian Society of Remote Sensing.* V: 35, Issue 3, pp 201-207.

Samli, R., Sivri, N., Sevgen, S. and Kiremetci, V. Z. (2014). *Applying Artificial Neural Networks for the estimation of Chlorophyll-A concentrations along the Istanbul coast.* Pol. J. Enviro. Stu. V: 23. No 4. 1281-1287.

Soyguder, S. and Alli, H. (2009): *An expert system for the humidity and temperature control in HVAC systems using ANFIS and optimization with Fuzzy Modeling Approach.* Energy and Buildings. 41(8): 814-822.

Soyguder, S. and Alli, H. (2009). *Design and simulation of self-tuning PID-type fuzzy adaptive control for an expert HVAC system. Expert Systems with Applications.* 36(3): 4566-4573.

Wen, X. L., Wang, H.T. and Wang, H. (2012). *Prediction model of flow boiling heat transfer for R407C inside horizontal smooth tubes based on RBF neural network.* Procedia Engineering. 31: 233-239.

Zhang, J., Huang, W. and Zhou, Q. (2014): *Reflectance Variation within the In-Chlorophyll Centre Waveband for Robust Retrieval of Leaf Chlorophyll Content.* PLoS ONE 9(11): e110812. doi:10.1371/journal.,pone.0110812