RESEARCHERID
THOMSON REUTERS

[Çelik , 6(9): September 2019] **ISSN 2348 – 8034**
DOI- 10.5281/zenodo.3455868 **Impact Factor- 5.070**

GLOBAL JOURNAL OF ENGINEERING SCIENCE AND RESEARCHES

ESTIMATION OF THE ORANGE PRODUCTION IN TURKEY BY MEANS OF ARTIFICIAL NEURAL NETWORKS

Şenol Çelik
Bingol University Faculty of Agriculture, Department of Biometry and Genetics, Bingol, Turkey

ABSTRACT

Citrus fruits which are grown in tropical and subtropical climate areas; is a plant including Citrus fruit tree species of the family Rutaceae such as citrus, orange, tangerine, grapefruit and lemon. The reason is that the orange is the most preferred citrus fruit by the juice industry. According to the FAO (Food and Agriculture) statistics of the year 2017, Brazil is the most important country with 17 459 908 tons of production and Chine follows the country with 8 685 812 tons of production. India ranks the third with 7 647 000 tons of production. Turkey ranked the seventh in the world with 1 950 000 tons of orange production annually.

Keywords: *Artificial neural network, production, orange, prediction.*

I. INTRODUCTION

Citrus fruits which are grown in tropical and subtropical climate areas; is a plant including Citrus fruit tree species of the family Rutaceae such as citrus, orange, tangerine, grapefruit and lemon. Citrus fruits, of which the origin is China, Southeast Asia and India, can be widely produced in Turkey because of the fact that the ecological conditions are suitable (1). With 55 % shares of the total citrus fruits production in the world, orange is the most produced citrus fruit species. The reason is that the orange is the most preferred citrus fruit by the juice industry. According to the FAO (Food and Agriculture) statistics of the year 2017, Brazil is the most important country with 17 459 908 tons of production and Chine follows the country with 8 685 812 tons of production. India ranks the third with 7 647 000 tons of production. Turkey ranked the seventh in the world with 1 950 000 tons of orange production annually. Top 10 largest orange producing countries in the world are shown in Table 1 (2).

Table 1. Countries with the highest orange production

1	Brazil	17459908
2	China	8685812
3	India	7647000
4	Mexico	4629758
5	United States of America	4615760
6	Spain	3357163
7	Egypt	3013758
8	Indonesia	2295325
9	Turkey	1950000
10	Pakistan	1585090

1 900 000 tons of oranges have been produced in Turkey in 2018. With 525 821 tons and 27.67% share of production, Antalya has the highest share in Turkey. Antalya is followed by Adana with 416 102 tons and 21.90 share of production and Hatay with 319 026 tons and 16.79 % share of production (3).

There are some statistical studies conducted on orange production. In a study, regression analysis between the price and production amount according to the weekly periods has been carried out. It has been stated that there is an inverse relationship between the variable 'week' and the total amount of domestic orange per week; and between the unit price and the total amount of domestic orange per week. It has been determined that there is a high level and significant relationship between the variable 'unit price of domestic orange' and the scores of total amount of domestic orange per week (4). In another study, chi square test of independence has been conducted to determine the strategies the export companies in Hatay practice in the international markets. As a result of the study, the businesses adapted differentiated marketing strategy and the pricing in the foreign markets is based on the market prices (5). It has been stated that the citrus fruits producers applying good agricultural practice in Mersin are more educated and they are more open to innovations compared to the conventional producers. The authors determined that the relative profit of the producers applying good agricultural practice are higher than that of conventional producers when it comes to the production of lemon, orange and mandarin (6).

The aim of this study is to model the orange production in Turkey with Artificial Neural Networks and present and estimation for the future.

II. MATERIAL AND METHOD

Material

The material of the study consists of the dependent and independent variables to be used in the projection of orange production amount in Turkey. The dependent variables are the figures of the orange production covering the years 1965-2018, and the independent variable is the 'series of years'. These variables have been selected in order to estimate in a correct manner with the help of the model to be established with ANN. These variables have been taken from the web site of Turkish Statistical Institute (TSI) (7, 8).

Method

Artificial neural network (ANN), in many over complex mathematical problems which imply complicated non-linear equations, a multilayer perceptron network can be simply used by defining the weights and appropriate functions. Be based on the type of the problem in neurons, various activation functions can be used. An input layer for introducing problem inputs, a hidden layer, and another output layer which ensures the solution to the problem are used in such networks. This kind of networks are mostly trained through Back Propagation method (9). Neural networks are machine learning models used principally to classify or predict data. The model architecture is performed starting from data and learning rules (10).

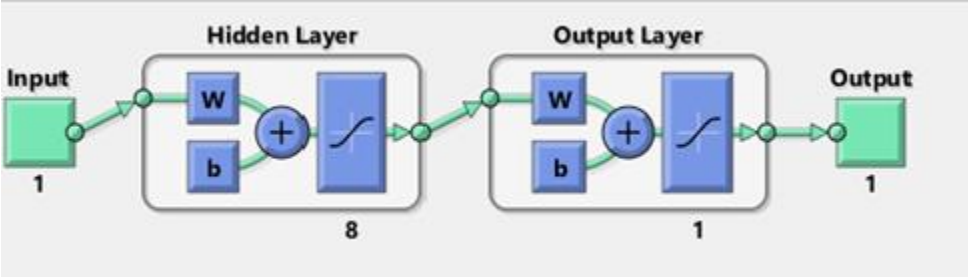

Figure 1. Input and output parameters for training ANN

Artificial neural network was used to predict hardness and relative density using 'nntool' in MATLAB software package. In this research a feed-forward back propagation algorithm with training function (*trainlm*) and adaption learning function (*learngdm*) were used in the prediction. Figure 1 displays artificial neural network with 1 hidden layer and 8 neurons. Years variable was used as the input parameter while the orange production was employed as a target or output value in the artificial neural network models. Input experimental data and target values were used to train the ANN algorithm. In order to prevent biased response normalizing processing were realized toward maximum or minimum value of input data set using the following equation.

ISSN 2348 – 8034
Impact Factor- 5.070

$$X' = \frac{X_i - X_{max}}{X_{max} - X_{min}}$$

where X' is the normalized value, X is the training data set, X_i is the values of each input data in training set, i = 1,2,3,.... The normalized value lies (0, 1 range) and to obtain original value reverse normalization operations were performed (11).

To measured the performance of ANN model such as mean squared error (MSE) and determination coefficient (R2) are calculated. The MSE is calculated as follows (12).

$$MSE = \frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{n}$$

where, n= Number of observation data, Y_i=Target value (Original value), $\hat{Y}_i$ = Predicted value or Forecast value.

Performance of trained model measured in terms of mean squared error (MSE) and correlation factor (R). The correlation factor R was generated by MATLAB software package while MSE values was measured from above. The trials were accomplished until correlation coefficient R approaches closer to the value one. The correlation factor R value close to one ensures better regression fit for given training set (11).

III. RESULTS

'Years' are used as independent variables and 'production amount' is used as dependent parameters. MATLAB program has been used in the analysis of data. The data has been divided into three parts: 70 percent was used for the training, 15 percent for validity and 15 percent for the test stage. In case that there are various data set, artificial neural network models can make estimations for future periods in the light of the long-term available data. Moreover, it could be adopted by training without any counselor, in the case that there is no output variable amongst the data sets, and it can cluster the sample group. The estimation by the means of simulation function has been used in the study. Levenberg-Marquardt Back Propagation has been used to train the developed artificial neural networks. After training the network, the error difference between the real situation and the results produced by the network has been examined. In the multi layered Artificial Neural Networks, the transmission function amongst the input, output and hidden layers makes difference. The network with 8 neural hidden layers is operated based on TANSIG function. The network with 8 neural hidden layers reached the highest value of performance in the usage of TANSIG function. The performance graphic is shown in Figure 1. That means there is a graphic indicating the change of the error values belonging to the training, validation and test clusters in each iteration at the end of the training of the network. As seen in the graphic, iteration number in training of the network is given as 36 and the best performance has been achieved in the 30th iteration.

ISSN 2348 – 8034
Impact Factor- 5.070

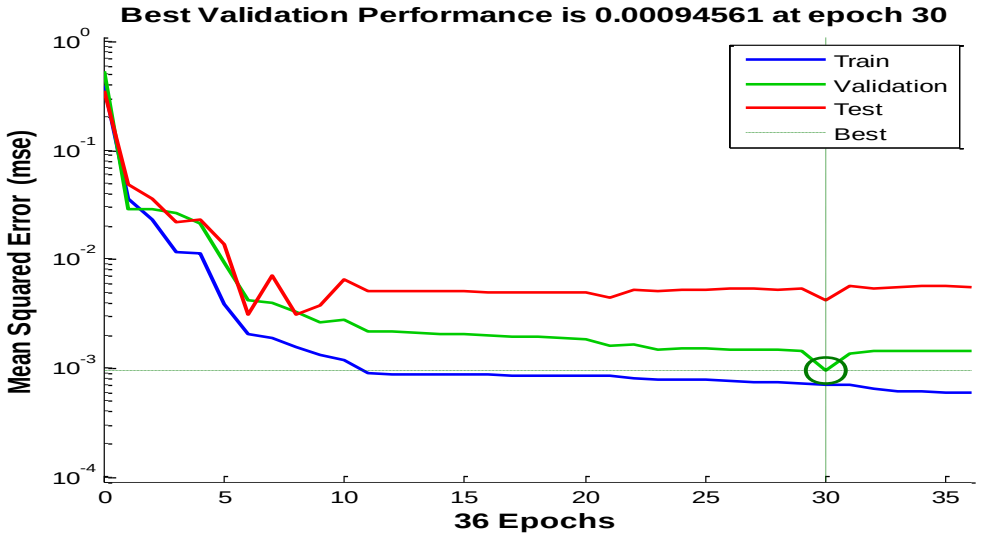

Figure 1. Network performance graph

In Figure 2, regression of the data for training, validation and the test have been given. When the graph is examined, it is obvious that the values of training, validity and all other values are higher than 0.99. The test value has been found as 0.96059. The training process of the network has been resulted in great success. In other words, the variables which we determined as 'independent variables' effect the orange production at the rate of 99%.

ISSN 2348 – 8034
Impact Factor- 5.070

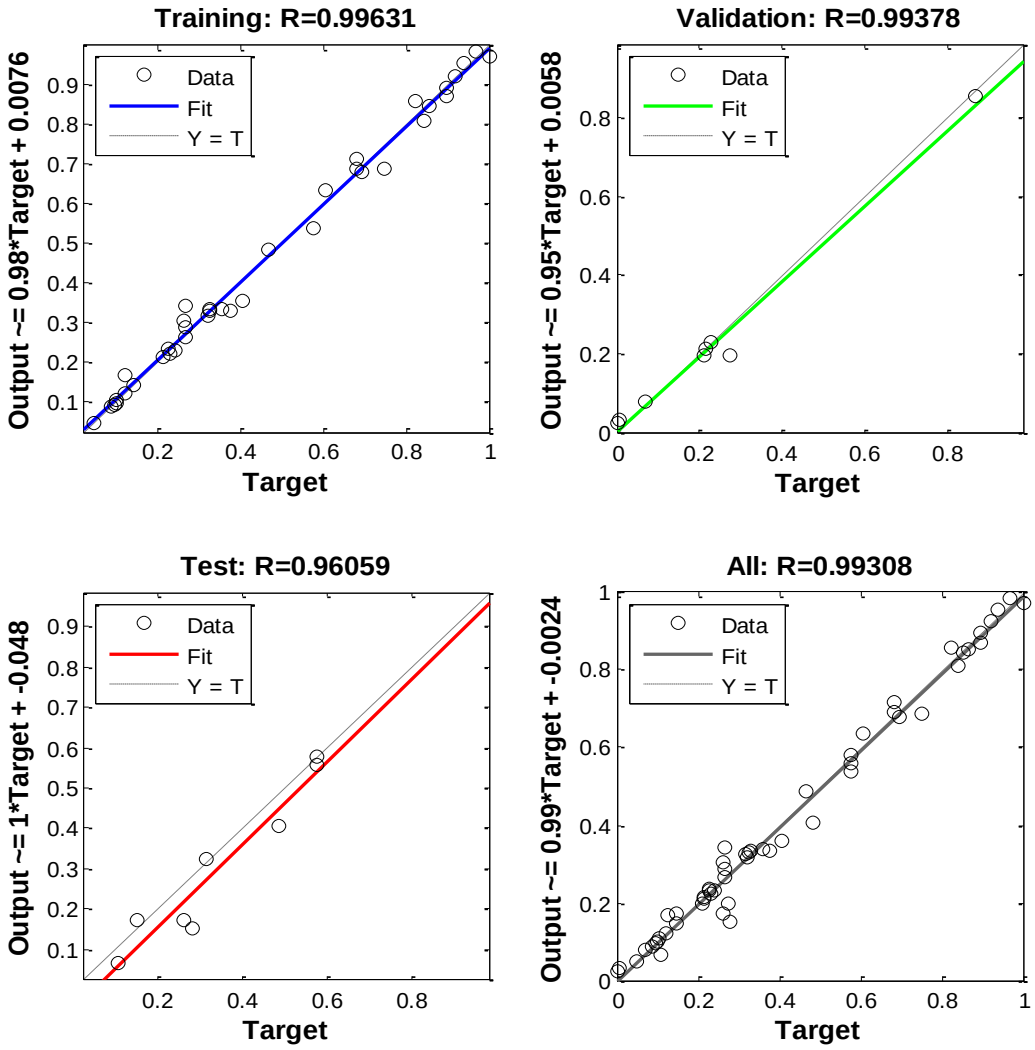

Figure 2. Regression fit and R values for training testing and validation

In Table 2, the results of orange production estimations respecting the years 2019-2025, along with the established ANN model have been given. MSE values stated here is the mean squared error of the observation values between the years 1965 and 2018 used in comparing the real values with the test output values given by the network inferentially after the test. MSE value has been found as 0.00124 and the coefficient of determination (R^2) as 0.986. It has been estimated that the orange production will be between 1 941 106 and 3 522 957 tons during 2019-2025.

Table 2. Prediction of orange production in Turkey

Years	Prediction
2019	1 941 106
2020	2 094 997
2021	2 204 013

ISSN 2348 – 8034
Impact Factor- 5.070

2022	2 431 168
2023	2 739 108
2024	3 259 138
2025	3 522 957

The estimated values obtained by training the Artificial Neural Network with data related to the years 1965-2018 and being tested, are close to the real observed values. This fact demonstrates that the established Artificial Neural Network Model is able to make estimations for the next term over the input variable '1'. In other words, how much production values to be obtained from input variables can be determined with the existing model. Graphical display of the values of the years 2009-2018 and the estimated values with ANN of the years 2019-2025 has been presented in Figure 3. When the graph has been examined, orange production, being 1 689 921 tons in 2009 is expected to increase over 3 500 000 tons in 2025.

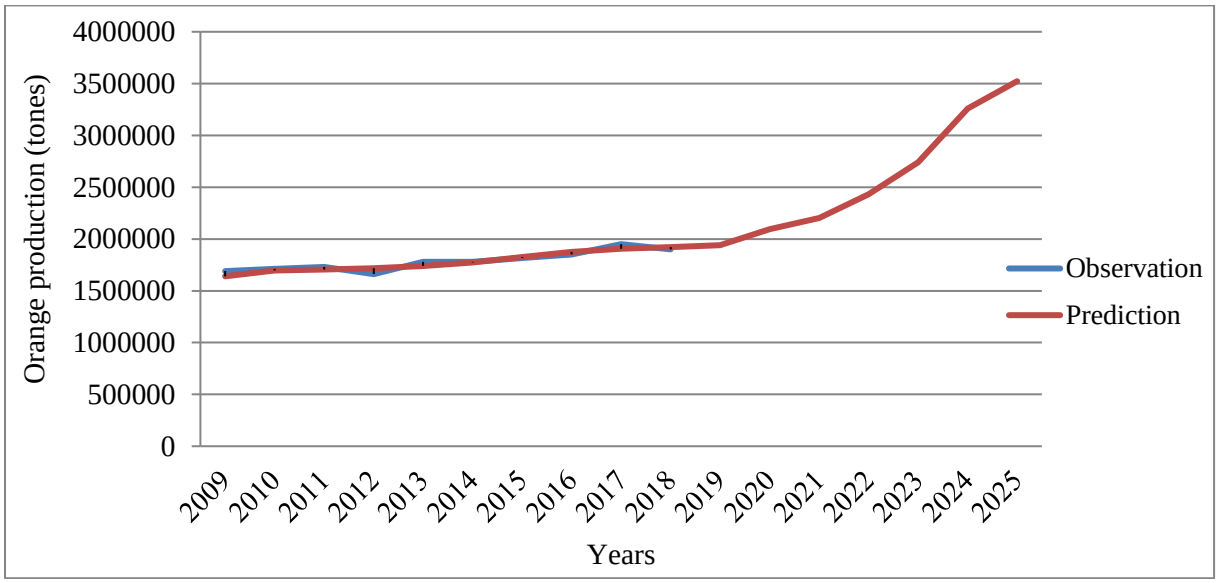

Figure 3. ANN forecast chart

IV. CONCLUSION

The amount of orange production in Turkey has been estimated by means of Artificial Neural Networks in the study. 'Years' (1965-2018) are used as input variable and 'orange production amount' is used as output variable. Afterwards, training, testing and validation processes of the network were performed, and estimations have been fulfilled.

The results indicated that the established ANN model and the estimation process were successful. High regression and low MSE (Mean Squared Error) values in the stages of training, testing and validation also indicate such achievement.

When the orange production estimations were examined, the production being 1 900 000 tons in 2018 is expected to increase by 85.42% and reach 3 522 957 tons in 2025.

It has been seen that Artificial Neural Networks model is very successful in terms of estimating the existing data generally. It is thought that comparative examination of estimated performances will give much more efficient

results by combining Artificial Neural Networks with different alternative techniques in estimations studies for the future.

REFERENCES

1. *Kayabaşı N., Etikan S. 2015. Subjective and Objective Evaluation of Colors Obtained from some Citrus (Rutaceae) Leaves. SDU Journal of Technical Sciences, 5(1): 20-28.*
2. *FAO, 2017. Food and Agriculture Organization of the United States. Production statistics, crops. http://www.fao.org/faostat/en/#data/QC.*
3. *TSI, 2018. Turkey Statistics Institute. Bitkisel üretim istatistikleri. Meyveler, içecek ve baharat bitkileri. https://biruni.tuik.gov.tr/medas/?kn=92&locale=tr.*
4. *Polat, Ö. 2010. Analysis of Adana city wholesale prices of fresh fruits. Cukurova University Institute of Natural and Applied Sciences, Msc Thesis, Adana, Turkey.*
5. *Kızıltuğ, T. 2017. Orange production and foreign trade business marketing strategies in Hatay province. Ankara University Graduate School of Natural and Applied Sciences, Master Thesis, Ankara, Turkey.*
6. *Subaşı, O. S., Uysal, O., Aktaş, E., Aydın, B. 2016. Comparative economic analysis of good agricultural practices in citrus production: The case of Mersin. 12th Ulusal Tarım Ekonomisi Kongresi, 25-27 Mayıs, Isparta, Turkey.*
7. *TSI, 2018. Turkey Statistics Institute. Citrus fruits, 1988-2018. http://www.tuik.gov.tr/PreTablo.do?alt_id=1001*
8. *TSI, 2018. TurkStat, Statistical Indicators, 1923 - 2013. Production of citrus fruit.*
9. *Manhaj, M. 2002. Principles of Artificial Neural Networks. Published by Industrial University of Amirkabir (Tehran Polytechnic), Tehran.*
10. *Kim, P. 2017. MATLAB Deep Learning With Machine Learning, Neural Networks and Artificial Intelligence. ISBN-13: 978-1484228449.*
11. *Yang, B., Wang, F., Zhang, J. S. 2003. Microstructural characterization of in situ TiC/Al and TiC/Al–20Si–5Fe–3Cu–1Mg composites prepared by spray deposition. Acta Mater, 51:4977–4989.*
12. *Eidgahee, D. R., Haddad, A., Naderpour, H. 2018. Evaluation of shear strength parameters of granulated waste rubber using artificial neural networks and group method of data handling. Scientia Iranica, 25. doi.org/10.24200/sci.2018.5663.1408.*