

Application of Artificial Neural Networks to Predict Cotton Production: A Case Study in Diyarbakır Province, Turkey

Nazire Mikail

Siirt University: Siirt Universitesi

Mehmet Fırat BARAN (✉ mfb197272@gmail.com)

Siirt University: Siirt Universitesi <https://orcid.org/0000-0002-7657-1227>

Research

Keywords: Crop fields, yield, artificial intelligence, regression, Diyarbakır, Turkey

Posted Date: September 18th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-76759/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Application of Artificial Neural Networks to Predict Cotton Production: A Case Study in Diyarbakır Province, Turkey

¹Nazire Mikail

^{2*}Mehmet Firat Baran

¹Department of Animal Science, Faculty of Agriculture, Siirt University, 56100 Siirt-TURKEY

²Department of Biosystem Engineering, Faculty of Agriculture, Siirt University, 56100 Siirt-TURKEY

^{2*}Corresponding Author: mfb197272@gmail.com

ABSTRACT

Cultivators are always curious about the factors affecting yield in plant production. Determining these factors can provide information about the yield in the future. The reliability of information is dependent on a good prediction model. According to the operating process, artificial neural networks imitate the neural network in humans. The ability to make predictions for the current situation by combining the information people have gained from different experiences is designed in artificial neural networks. Therefore, in complex problems, it gives better results than artificial neural networks.

In this study, we used an artificial neural network method to model the production of cotton. From a comprehensive datum collection spanning 73 farms in Diyarbakır, Turkey, the mean cotton production was 559.19 kg da⁻¹. There are four factors that are selected as pivotal inputs into this model. As a result, the ultimate ANN model is able to forshow cotton production, which is built on elements such as farm states (cotton area and irrigation periodicity), machinery usage and fertilizer consumption.

At the end of the study, cotton yield was estimated with 84% accuracy.

Keywords: Crop fields, yield, artificial intelligence, regression, Diyarbakır, Turkey

INTRODUCTION

Cotton is one of the most prominent products in the agricultural sphere, industry and trade because of its very different and important utilization areas. In addition to the escalating world population, the increasing needs of human beings for consumption raise the importance of this versatile plant day by day. Growing interest in natural fibers and rising living standards in the world increases the demand for cotton plants (Anonym 2020 a).

According to the data of the International Cotton Advisory Committee (ICAC), 32,825 million hectares of cotton were produced in the world in the 2018/19 production period. In this season, 37% of the 32,825 million hectares of cotton are cultivated in India. India is followed by the USA, China, Pakistan and Brazil in the width of the cultivation areas. As a result of the expansion of cotton acreage in African countries in recent years, despite its growth, Turkey has been ranked 11th in terms of world cotton cultivation area (Anonym 2020b).

It is an important industrial plant that constitutes raw materials with fiber in the textile industry, oil obtained from its core in the vegetable oil industry, achenium and pulp in the animal feed industry, linters in the paper, furniture and cellulose industries. Cotton is an important and strategic product that provides great benefits to our country's economy with this wide area of

use, added value and employment opportunities. Due to these features, it has contributed to the development of both agriculture and industry of the regions and countries grown (Anonym 2020c).

Areas where cotton production is intensive in Turkey; Aegean, Çukurova, Southeastern Anatolia Regions and Antalya. In the 2017/18 cotton season, 882 thousand tons of cotton fiber production was made in 502 thousand hectares, and approximately 1 million 571 thousand tons of cotton was consumed in Turkey. In the 2017/18 cotton production season, in return to 882 thousand tons of fiber cotton, 2,5 million tons of seed cotton was produced, and the fiber cotton yield was 1820 kg / ha. Şanlıurfa, Aydın, Bursa, Diyarbakır, Adana and İzmir are 6 provinces in Turkey that meet 88% of production, respectively. Şanlıurfa province alone meets 42% of all production. The share of the other 23 cotton-producing provinces in production is between 0.1% and 1.3% (Anonym 2020 d).

Since cotton is a selective plant in terms of climate characteristics, it can be grown in limited places in our country (Karademir et al. 2015). Cotton plants have great economic importance for humanity with their widespread and compulsory use and for producer countries with the added value and employment opportunities that they create. Increasing populations, growing interest in natural fibers and rising living standards raise the demand for cotton plants.

Artificial neural networks have been used with traditional statistical methods in recent years thanks to their highly accurate prediction and classification capability, and now, they are even more popular than statistical methods. However, ANN modeling also has its own challenges. We can practice many potential ways to form and train networks. To put it in different ways, with a sufficiently large number of independent parameters, neural networks can be trained to tightly suit data that are likely to include noise. Hence, it is very important for us to understand how to optimize the structure of networks and learning to develop reliable models for decision-making that generalize well to data that do not fit any distribution (Samarasinghe, 2006, p. 5). In recent years, ANN models have been applied to predict the production of different plants using artificial neural networks. As an example, a high-accuracy production estimation model has been developed for basil (Rostami et al., 2017) and wheat (Safa et al., 2015). In addition, many studies have created ANN models that measure the effect of potato production on the environmental quality index (Khoshnevisan et al., 2013), predict the maturity of cotton fibers (Farooq et al., 2018), and estimate the massiveness of energy ingested in production (Safa & Samarasinghe, 2011; Taki et al., 2012; Khoshnevisan et al., 2013; Nabavi-Pelesaraei et al., 2016; Khoshroo et al., 2018; Taki et al., 2018). With the same success, ANN is being used in other fields of agriculture. For example, high accuracy models can be created in the estimation of the temperature inside greenhouses (Saltuk & Mikail, 2019), in beef cattle production (Bozkurt et al., 2015), in the estimation of milk yield (Mikail et al., 2013; 2014; 2016) and in the diagnosis of diseases such as mastitis (Mammadova & Keskin, 2015).

In this study, we set down the elements affecting cotton production in Diyarbakır to create a model with the help of these factors and to make a model with multiple regression analysis and artificial neural networks. Additionally, it aims to estimate yield.

MATERIALS AND METHODS

Diyarbakır Province is in the central part of the southeastern Anatolia region and at the northern end of Mesopotamia. It is surrounded by provinces such as Siirt and Muş from the east; Mardin from the south; Şanlıurfa, Adıyaman, and Malatya from the west; and Elazığ and Bingöl from the north. Its area is 1.516.200,00 square kilometers, between 37.905199 and 40.231934 north latitudes and 40.37 and 41.20 east longitudes. It is surrounded by mountains that are not too high, and in the middle, there is concavity. It is covered with 37% mountains and 31% lowlands. The lowlands are fertile and suitable for agriculture. These fecund lands are irrigated by the

Tigris River and its tributaries. The city was founded on a horizontal surface on the eastern edge of the broad basalt plateau stretching between Karacadağ and the Tigris, above the Tigris valley and at the top of the river curve. Its altitude is 650 meters above sea level. This altitude changes between 640 m and 660 m in some places (Anonym, 2020e)

Material

The core material of this particular study is the data derived from cotton-producing agricultural enterprises in the Bismil district of Diyarbakır. The data covered by the research were obtained through 73 face-to-face surveys in 2019. However, the secondary data of the research were used from the publications of various national and international institutions and organizations related to the subject.

Method

Sampling Method

The Simple Random Sampling Method was used to determine the sampling frame and sample number in the study (Yamane, 1967).

$$n = \frac{N \times s^2 \times t^2}{(N-1)d^2 + (s^2 \times t^2)} \quad (1)$$

n = Sample Size

s = Standard Deviation

t = "t value" Related to the Selected Confidence Limit

N = Total Number of Units for Sampling Frame

d = Acceptable margin of error (%)

In the study, it was deemed appropriate to conduct a questionnaire on cotton production in 134 enterprises with a 95% confidence interval and 5% deviation from the mean.

The Farmer Registration System of Diyarbakır Provincial Directorate of the Ministry of Agriculture and Forestry records were used as sample villages and enterprises.

The primary data to be used in the study were the data obtained from the cotton producing enterprises in the sampling area. In determining the settlements within the scope of the survey application, the indicators of the ability to represent the research area were used, taking into account characteristics such as cotton cultivation area, number of enterprises, agricultural production potential, presence of tools - equipment, etc. The questionnaire forms used were arranged in accordance with the characteristics of the research field. Before the survey application, trial questionnaires were conducted in the selected settlements, after which we have given the final questionnaire forms considering the deficiencies observed.

These oblique details and cotton production were surveyed to design the model to be able to prognosticate cotton production. To be able to implement this in the ANN model, it is essential to pick a finite number of relevant and significant variables, voiding any sort of propensity. For that reason, all of the information was researched cautiously. Seventeen original variables participated to be presented as the most likely input in the final model.

Questionnaires containing unreliable information are not included in the process. For example, irrigation frequency entered as 0 is excluded. We used this information to draw the graphs and carry out statistical analysis operating with MINITAB Statistical Programme and MATLAB Software (Mathworks, 2009). In this context, 73 surveys conducted in Diyarbakır province Bismil central districts and villages were evaluated, and an estimation application was made.

Multiple Linear Regression

Repercussions Y and terms X_1, \dots, X_p will be applied in this general multiple linear regression model as

$$E(Y|X) = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p \quad (2)$$

In this equation, X in $E(Y|X)$ means that all the terms on the right side are being stipulated. Correspondingly, it happens while we are postulating on certain values for the predictors X_1, \dots, X_p henceforth, to be called X . The unknown parameters are indicated as β 's, and these particular parameters we need to account. In equation (2), the linear function of the parameters is stated, and for this reason, this is called linear regression. To obtain the simple regression problem, p should be equal to 1 so that X will have only one element. If $p = 2$, the mean function (2) complies with a plane in three dimensions, as shown in Fig. 1. However, if $p > 2$, the fitted mean function turns out to be a hyperplane, which is the universal meaning of a p dimensional plane in a $(p + 1)$ dimensional space (Weisberg, 2005).

Figure 1.

Artificial Neural Network

In Fig. 2, the ANN model with one input, one hidden neuron and one output was pictured in the input layer, hidden neuron layer and output layer, respectively. As a result, this ANN model has a_1 - one input-hidden layer weight and b_1 - one hidden-output layer weight. Here, X is the input and z is the network output. Bias input of + 1 has been noted in the hidden neuron and output neuron, respectively, with an associated weight of a_0 and weight of b_0 . Smith also used this notation for the weights.

Figure 2.

The network feeds with the input, and the hidden neuron computes the weighted total of inputs (also covering bias) and delivers this weighted total of inputs via the logistic function to be able to build the hidden-neuron output, y . However, the output neuron feeds by the output neuron y as input via the associated connection link. By this link, it is weighted. After that, the weighted input is transferred by the activation function of the neuron. Network output appears as a result of the output of this neuron's transformation. The hidden neuron supersedes a very significant part of processing.

The weighted input is shown below:

$$u = a_0 + a_1 x \quad (3)$$

After the first step, it follows by the idea that the weighted sum u should be passed through the logistic function by the hidden neuron. U is the logistic function's argument, and this function, without fail, is a standard function with $y=0.5$ at $u=0$. It will be more beneficial for the output y to be specified in terms of input x , so it will be possible to portray y as being laid out by x through u . Replacing u inside the logistic function, output y as the hidden neuron is also shown below:

$$y = \frac{1}{1 + e^{-(a_0 + a_1 x)}} \quad (4)$$

Model Assessment

In this model, we used the evaluation r- Pearson correlation coefficient, R^2 - coefficient of determination, RMSE - root mean squared error (Spiegel et al., 2009).

a) The coefficient of correlation

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x}_i)(y_i - \bar{y}_i)}{\sqrt{\sum_{i=1}^n (x_i - \bar{x}_i)^2 \sum_{i=1}^n (y_i - \bar{y}_i)^2}} \quad (5)$$

b) The coefficient of determination

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \tilde{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (6)$$

c) Root mean squared error

$$RMSE = \sqrt{\sum_{i=1}^n (y_i - \tilde{y}_i)^2 / n} \quad (7)$$

where y_i is the observed value, \hat{y}_i is the predicted value, \bar{y} is the arithmetic mean, and n is the total number of observations.

RESULTS AND DISCUSSION

In the present study, the maximum and minimum yields varied between 410 and 680 kg per da, and the average yield was 599.19 kg da⁻¹ (Table 1). The average cultivation area in the enterprises was found to be 87.64 da. The amount of fertilizer used in the enterprises was 87.84 kg per da, the tractor usage time was 1.8 hours and the irrigation frequency was approximately 6 times (Table 1).

Table 1.

In the study, no linear relationship was found between yield and cultivation area and tractor operating times. It was determined that the amount of fertilizer and the frequency of irrigation positively affect the yield, although it is weak. That is, the Pearson correlation coefficient between fertilizer amount and yield was $r = 0.29$ ($p < 0.05$), and the Pearson correlation coefficient between irrigation frequency and yield was calculated as $r = 0.38$ ($p < 0.01$) (Figure 3). Similar results were found in the study (Safa et al., 2015). In this study conducted in New Zealand, the Pearson correlation coefficient between N fertilizer consumption and wheat yield was determined as $r = 0.43$.

Figure 3.

Multiple Linear Regression Model

To estimate cotton production, the following estimation model was created as a result of multiple linear regression analysis. The determination coefficient was used as the model evaluation criterion. The coefficient of determination of the model was found to be $R^2 = \%18.3$

$$\text{Yield} = 346 + 0.039 \times \text{Cotton area} + 0.725 \times \text{Fertilizer} + 33.09 \times \text{Tractor usage} + 14.11 \times \text{Irrigation frequency}$$

Artificial Neural Network Model

According to this study, the ANN model consists of 20 neurons hidden layer, 4 neurons input layer and 1 output layer. (Figure 1). 4).

Figure 4. Generated ANN model

The Levenberg-Marquardt backpropagation learning method was used in the model. The data set is randomly divided into 3 shares: 70% training, 15% validation and 15% testing. The compatibility of the cotton production values estimated with each data set with the actual observation values was measured by the Pearson correlation coefficient (Figure 1). 5).

Figure 5. Correlation between observed and predicted cotton yield for a) training b) validation c) test and d) all data.

The model gave the best result after the 18th iteration. As seen from Figure 3, the model gave the best result with training data. Here, the degree of the linear relationship between the estimated yield values obtained by using the training data and the actual yield values was found to be 0.98. This is followed by validation (0.77) and testing (0.76) accordingly. When all data were evaluated together, this coefficient was found to be 0.91. However, the RMSE statistic of the model was found to be 20.37. This value was calculated as 44.81 in the estimates obtained by MLR (Table 2).

Table 2.

As a result of the comparison between the ANN model and the multiple linear regression models, it was seen that the multiple linear regression model cedes the correlation coefficient between the exact and predicted cotton yield of the ANN model, as the ANN model was much higher. In addition, the RMSE value was found to be much lower than the value obtained from the MLR estimation (Table 2).

Figure 6. Observed yield values, yield values estimated with MLR and estimated yield values with ANN

As shown in Figure 6, the ANN model can predict cotton yield with 84% accuracy. It was calculated as 91% in a similar study created for wheat production. The fact that there are weak relationships between the data leads to a very low accuracy as a result of the model made with MLP. In modeling with ANN, we can see that high accuracy predictions can be made even by using such data.

To increase the accuracy of the model in the future, it is envisaged to investigate other factors that affect efficiency and to create new models. In addition, we believe that more accurate results can be obtained by conducting research not only among cotton producers in Diyarbakır but also among all nearby producers.

CONCLUSION

In this study, we tried to predict cotton production with factors such as cotton area (da), fertilizer (kg da⁻¹), tractor usage (h da⁻¹) and irrigation frequency using the ANN model. The neural network model created with these factors can predict cotton production with an 84% accuracy. The results of this study demonstrated the capacity of ANN models to better forecast cotton production by applying different factors than a multiple regression model. In the future, it will be possible to develop high-accuracy ANN models with the addition of environmental and breeding factors affecting production and forbye by using more business data.

The results of this study are the start of developing suitable methods for estimating cotton production for the Diyarbakır region. Estimation models of different agricultural products can be created by using the ANN method.

Declarations

Ethics approval and consent to participate

The core material of this particular study is the data derived from cotton-producing agricultural enterprises in the Bismil district of Diyarbakır. The data covered by the research were obtained through 73 face-to-face surveys in 2019. The farmer's consent was obtained in the surveys conducted.

Consent for publication

I am very pleased to submit our manuscript entitled “Application of Artificial Neural Networks to Predict Cotton Production: A Case Study in Diyarbakır Province, Turkey” by Nazire Mikail and Mehmet Fırat Baran, for consideration of publication in Journal of Cotton Research.

In consideration of the publication, I hereby warrant and undertake:

1. This article is an original work and no portion of the study has been published.
2. None of the author has any potential conflict of interest related to this manuscript.
3. Author has contributed to the work and author has agreed to submit the manuscript.

Your consideration of the manuscript would be greatly appreciated.

Availability of data and material

The core material of this particular study is the data derived from cotton-producing agricultural enterprises in the Bismil district of Diyarbakır. The data covered by the research were obtained through 73 face-to-face surveys in 2019. However, the secondary data of the research were used from the publications of various national and international institutions and organizations related to the subject

Conflict of interest

Nazire Mikail and Mehmet Fırat Baran, declare that they have no competing interests.

Funding

No financial support was received for this study.

Authors' contributions

The authors' contribution rate to this study is 50%.

Acknowledgements

We would like to thank the cotton producers who supported us in the survey

Authors' information

First Author:

Asst. Prof. Dr. Nazire MİKAIL

Department of Animal Science, Faculty of Agriculture, Siirt University, 56100 Siirt-TURKEY

Education Background:

| Degree | Field | University | Year |
|------------------------|---|-----------------------|-------------|
| Bachelor | Faculty of Applied Mathematics and Cybernetics; Department of Applied Mathematics | Baku State University | 2000 |
| Post Graduate (Master) | Faculty of Applied Mathematics and Cybernetics; Department of Probability Theory and Statistics | Baku State University | 2002 |
| Doctorate (PhD) | Institute of Science, Department of Animal Sciences (Biometrics and genetics) | Selçuk University | 2012 |

Corresponding author

Assoc. Prof. Dr. Mehmet Fırat BARAN

Department of Biosystem Engineering, Faculty of Agriculture,

Siirt University, 56100 Siirt-TURKEY

E-mail: mfb197272@gmail.com

| Education Background | | |
|-----------------------------|------|---|
| Doctorate (PhD) | 2010 | Namık Kemal University, Graduate School of Natural and Applied Sciences Agricultural Machinery Mainscience Section- Tekirdağ/Turkey |
| Post Graduate (Master) | 2000 | Trace University, Graduate School of Natural and Applied Sciences Agricultural Machinery Mainscience Section- Tekirdağ/Turkey |
| Bachelor | 1997 | Trace University, Agricultural Faculty Agricultural Mechanization Mainscience Section – Tekirdağ /Turkey |

Research Interest

Energy Systems; Energy use in agriculture; Renewable Energy Technologies, Drying Technology, Agricultural Mechanization.

REFERENCES

- Anonym 2020a, Diyarbakır ve Şanlıurfa illerinde pamuk sektörü Envanteri, Access link: https://www.karacadag.gov.tr/Dokuman/Dosya/www.karacadag.org.tr_8_WH3D93PC_Diyarbakir_ve_Sanliurfa_illerinde_pamuk_sektoru_envanterinin_hazirlanmasi_projesi.pdf, Access date: 21.01.2020
- Anonym, 2020b, Cotton production report, April 2019, Access Link: <https://ticaret.gov.tr/data/5d41e59913b87639ac9e02e8/d0e2b9c79234684ad29baf256ae7dce.pdf> , Access date: 02.05.2020
- Anonym 2020c, Cotton report, 2018 Access Link: http://www.zmo.org.tr/genel/bizden_detay.php?kod=30467&tipi=17&sube=0, Access date: 21.01.2020
- Anonym 2020d, Cotton, Access link, <https://arastirma.tarimorman.gov.tr/tepge/Belgeler/PDF%20Tar%C4%B1m%20%C3%9Cr%C3%BCnleri%20Piyasalar%C4%B1/2019Ocak%20Tar%C4%B1m%20%C3%9Cr%C3%BCnleri%20Raporu/2019-Ocak%20Pamuk.pdf>, Access date: 21.01.2020
- Anonym, 2020e. Diyarbakır ili coğrafyası. <http://www.diyarbakirkulturturizm.gov.tr/TR-56881/cografya.html>, Erişim tarihi: 01.05.2020 (In Turkish)
- Bozkurt, Y., Aydoğan, T., Tuzun, C.G., Mikail, N., Varban, S., Dogan, C., Tatlı, M. (2015). Some applications of artificial neural networks used beef cattle production, 4th International Congress New Perspectives and Challenges of Sustainable Livestock Production, 07-09 EKİM 2015, Belgrade, Sırbistan.
- Farooq, A., Sarwar, M. I., Ashraf, M. A., Iqbal, D., Hussain, A., & Malik, S. (2018). Predicting Cotton Fibre Maturity by Using Artificial Neural Network. *Autex Research Journal*. <https://doi.org/10.1515/aut-2018-0024>
- Karademir, E , Karademir, C , Ekinci, R , Sevilmiş, U. (2015). İleri Generasyondaki Pamuk (*Gossypium hirsutum* L.) Hatlarında Verim ve Lif Kalite Özelliklerinin Belirlenmesi. *Türkiye Tarımsal Araştırmalar Dergisi* , 2 (2) , 100-107 . DOI: 10.19159/tutad.60964 (In Turkish)
- Khoshnevisan, B., Rafiee, S., Omid, M., Mousazadeh, H., & Sefeedpari, P. (2013). Prognostication of environmental indices in potato production using artificial neural networks. *Journal of Cleaner Production*. <https://doi.org/10.1016/j.jclepro.2013.03.028>
- Khoshnevisan, B., Rafiee, S., Omid, M., Yousefi, M., & Movahedi, M. (2013). Modeling of energy consumption and GHG (greenhouse gas) emissions in wheat production in Esfahan province of Iran using artificial neural networks. *Energy*. <https://doi.org/10.1016/j.energy.2013.01.028>
- Khoshroo, A., Emrouznejad, A., Ghaffarizadeh, A., Kasraei, M., & Omid, M. (2018). Sensitivity analysis of energy inputs in crop production using artificial neural networks. *Journal of Cleaner Production*. <https://doi.org/10.1016/j.jclepro.2018.05.249>

- Mikail, N., Altay, Y., Keskin, İ. (2013). A Sample Model Prediction of 305-Day Milk Yield of Holstein Cows Using Artificial Neural Networks, VIth International Balkan Animal Conference, BALNIMALCON, Tekirdağ, Turkey, 03-05 EKİM, 2013.
- Mikail, N., Keskin, İ., Altay, Y. (2014). Siyah Alaca ineklerin süt verimi tahmininde yapay sinir ağları ve destek vektör makineleri yöntemlerinin kullanımı, Uluslararası Mezopotamya Tarım Kongresi, 22-25 EYLÜL 2014, Diyarbakır, Türkiye.
- Mikail, N., Keskin, İ. (2015). Application of neural network and adaptive neuro-fuzzy inference system to predict subclinical mastitis in dairy cattle. *Indian J. Anim. Res.*, 49 (5) 2015 : 671-679.
- Mikail, N., Keskin, İ., Altay, Y., Dağ, B. (2016). A sample model prediction of milk yield in Akkaraman ewes using artificial neural networks. *International Human and Nature Sciences: Problems and Solution Seeking Congress*, 07-09 EKİM 2016, Bosna Hersek.
- Mathworks (2009). MATLAB - Mathworks - MATLAB & Simulink. <http://doi.org/2016-11-26>
- Nabavi-Pelesaraei, A., Rafiee, S., Hosseinzadeh-Bandbafha, H., & Shamsirband, S. (2016). Modeling energy consumption and greenhouse gas emissions for kiwifruit production using artificial neural networks. *Journal of Cleaner Production*. <https://doi.org/10.1016/j.jclepro.2016.05.188>
- Rostami, S., Choobin, S., Samani, B. H., Esmaceli, Z., & Zareiforoush, H. (2017). *Analysis and Modeling of Yield, CO2 Emissions, and Energy for Basil Production in Iran using Artificial Neural Networks*.
- Safa, M., & Samarasinghe, S. (2011). Determination and modelling of energy consumption in wheat production using neural networks: "A case study in Canterbury province, New Zealand." *Energy*. <https://doi.org/10.1016/j.energy.2011.06.016>
- Safa, M., Samarasinghe, S., & Nejat, M. (2015). Prediction of wheat production using artificial neural networks and investigating indirect factors affecting it: Case study in canterbury province, New Zealand. *Journal of Agricultural Science and Technology*.
- Samarasinghe, S. (2006). Neural Networks for Applied Sciences and Engineering. In *Neural Networks for Applied Sciences and Engineering*. <https://doi.org/10.1201/9781420013061>
- Saltuk, B., Mikail, N. (2019). Prediction of indoor temperature in a greenhouse: Siirt sample. *Fresen. Environ. Bull.* 28(4A), 3577-3585.
- Spiegel, M. R., Schiller, J. J., & Srinivasan, R. A. (2009). Probability and statistics. In *Schaum's outlines*.
- Taki, M., Ajabshirchi, Y., & Mahmoudi, A. (2012). Prediction of output energy for wheat production using artificial neural networks in Esfahan province of Iran. *International Journal of Agricultural Technology*, 8, 1229-1242.
- Taki, M., Rohani, A., Soheili-Fard, F., & Abdeshahi, A. (2016). Assessment of energy consumption and modeling of output energy for wheat production by neural network (MLP and RBF) and Gaussian process regression (GPR) models. *Journal of Cleaner Production*.

<https://doi.org/10.1016/j.jclepro.2017.11.107>

Weisberg, S. (2005). Applied Linear Regression: Third Edition. In *Applied Linear Regression: Third Edition*. <https://doi.org/10.1002/0471704091>

Yamane, T. (1967). Statistics: And Introductory Analysis, 2nd Ed., : In *Scottish Journal of Arts, Social Sciences and Scientific Studies*.

Figures

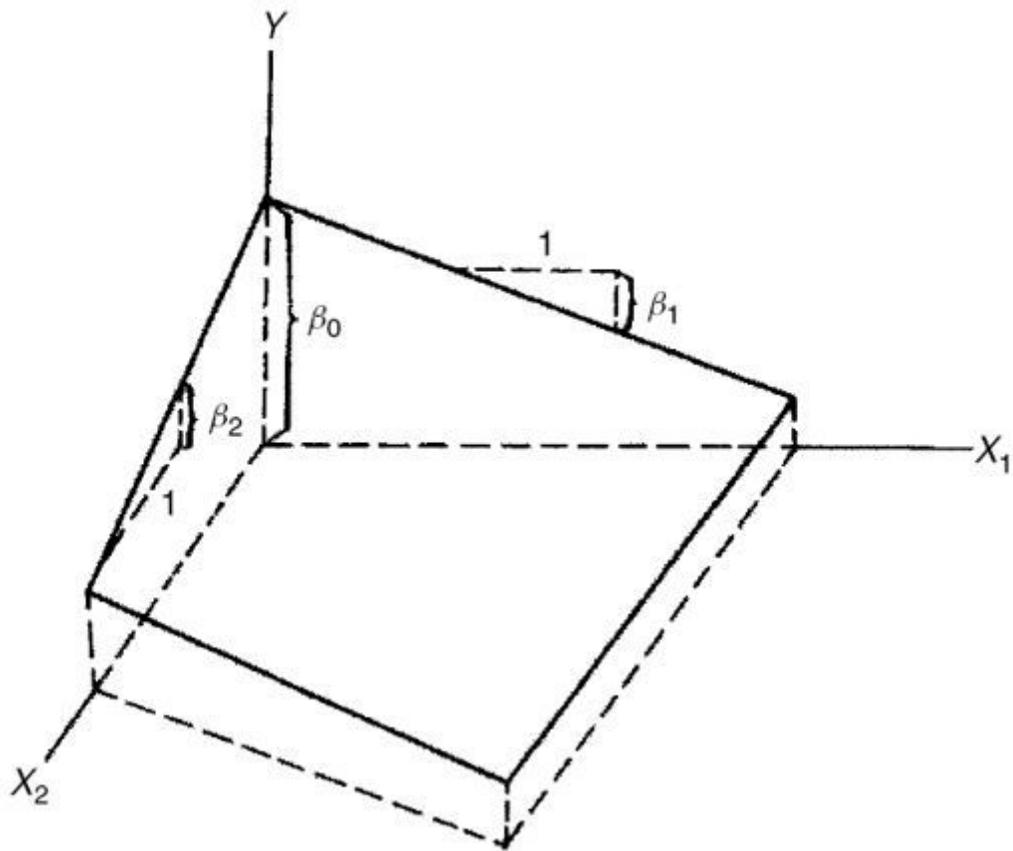


Figure 1

A linear regression surface if $p = 2$ predictors (Weisberg, 2005)

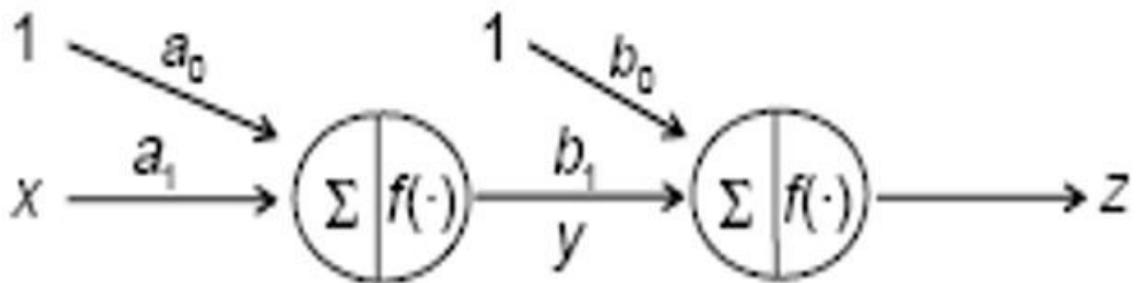


Figure 2

One-hidden neuron, one-input and one-output neural network

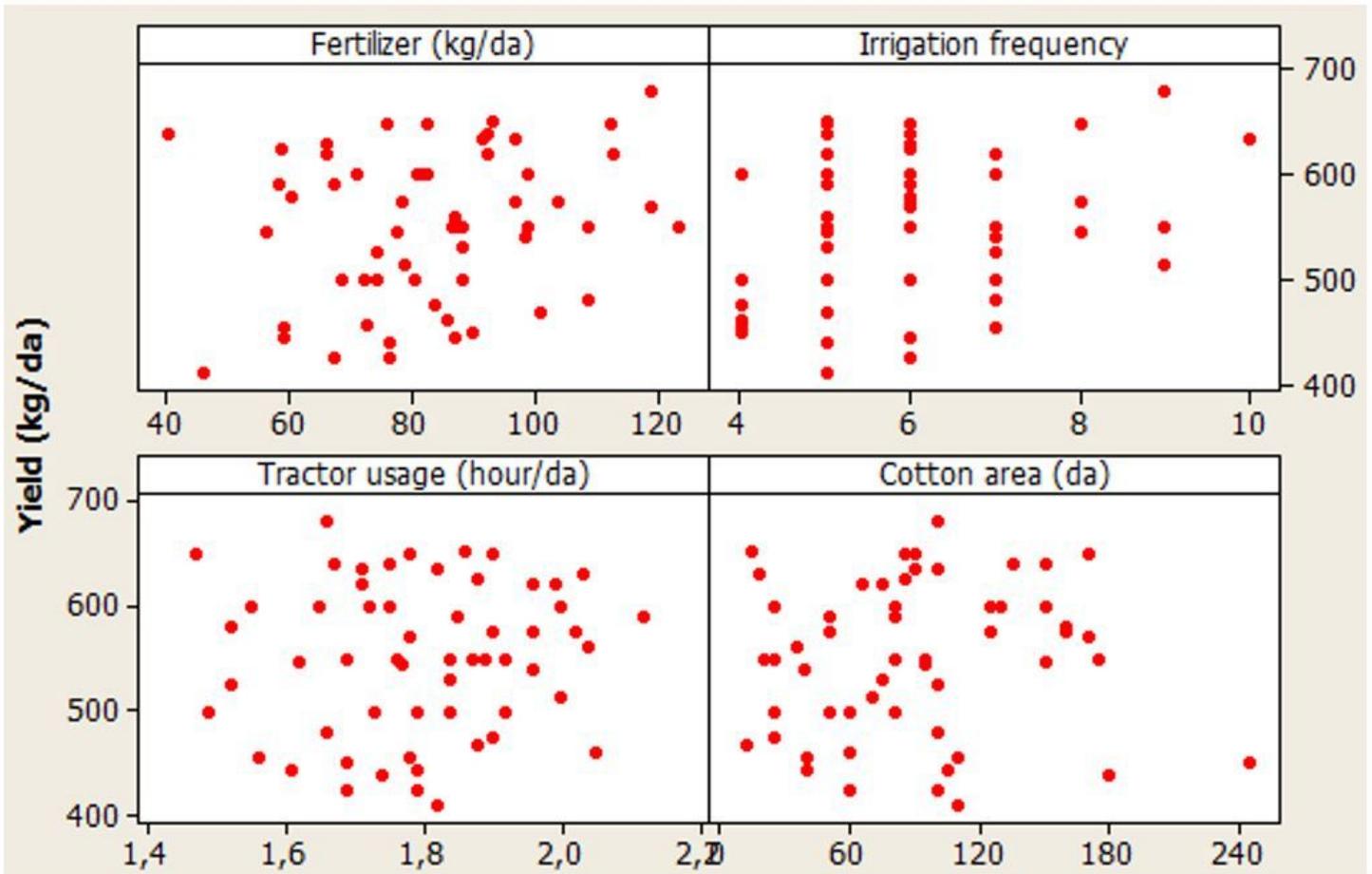


Figure 3

Correlations between cotton yield and input variables

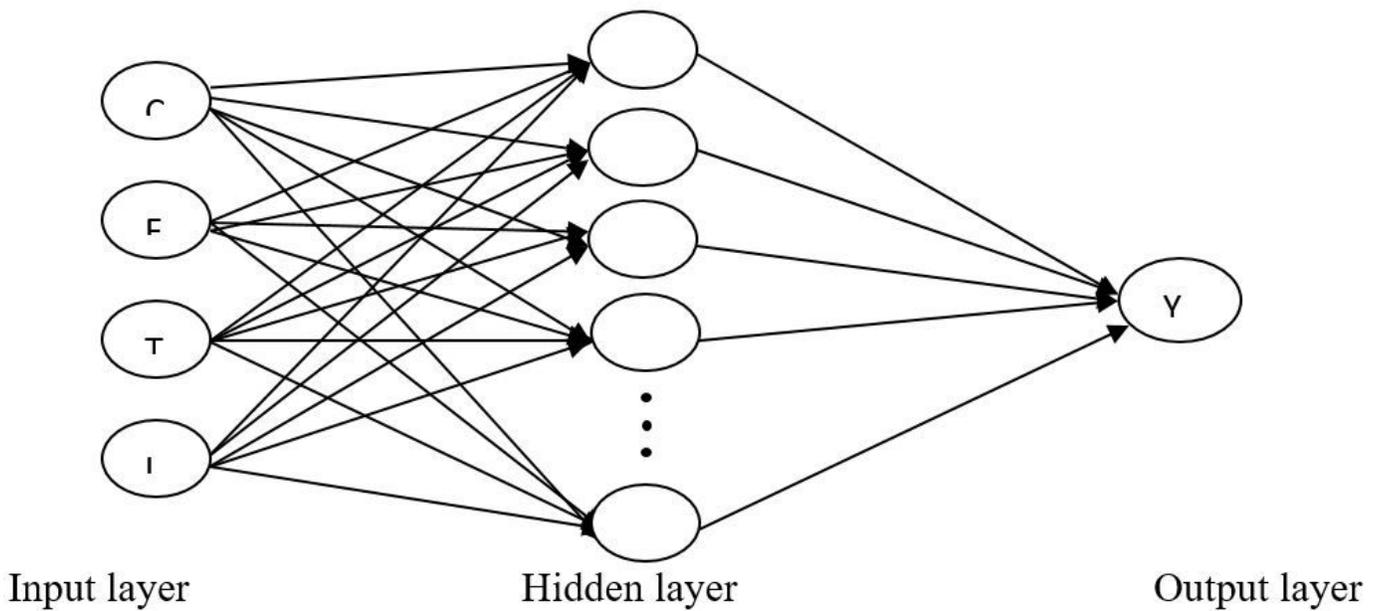
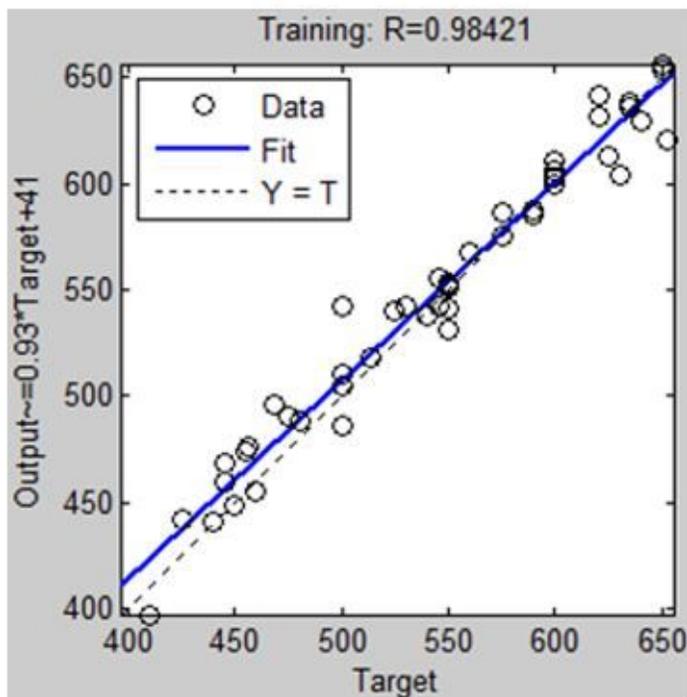
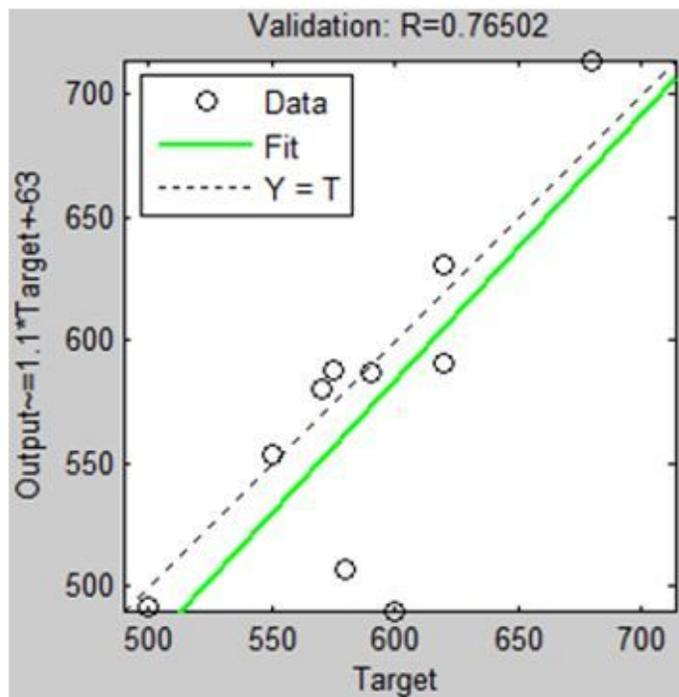


Figure 4

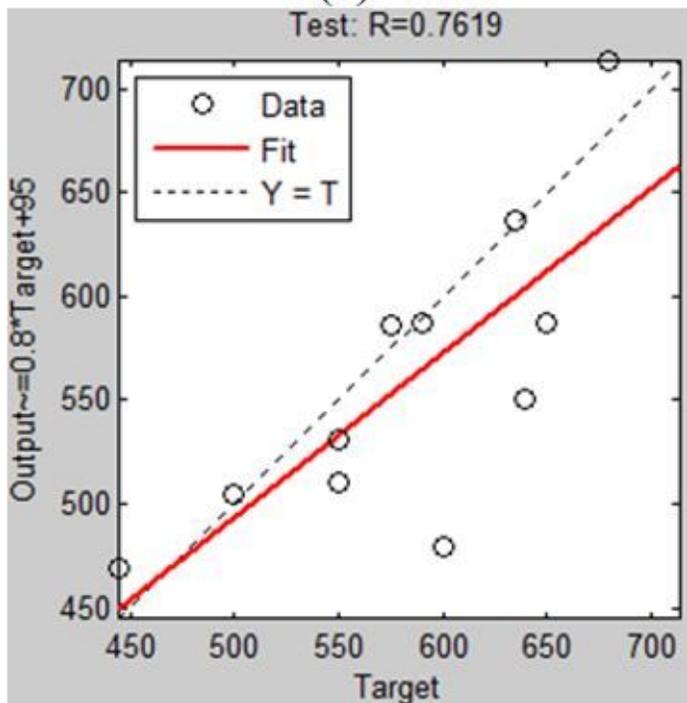
Generated ANN model



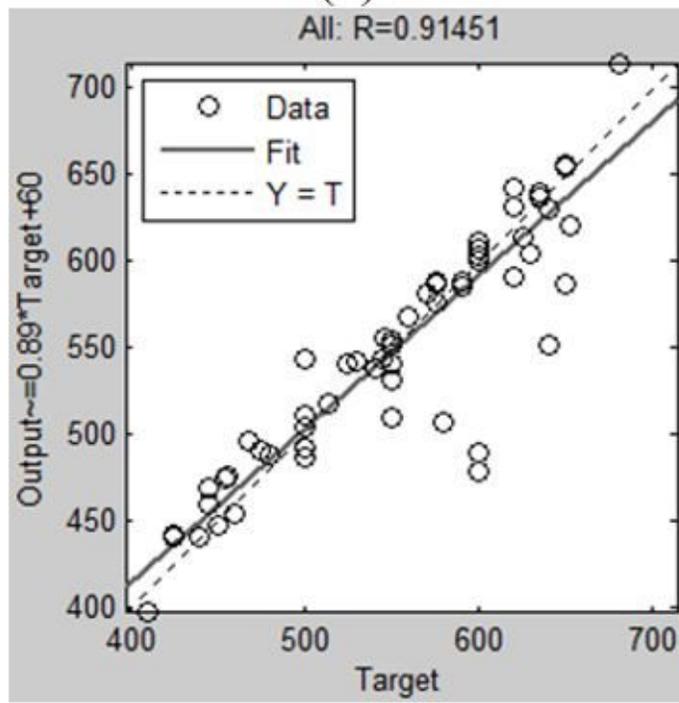
(a)



(b)



(c)



(d)

Figure 5

Correlation between observed and predicted cotton yield for a) training b) validation c) test and d) all data.

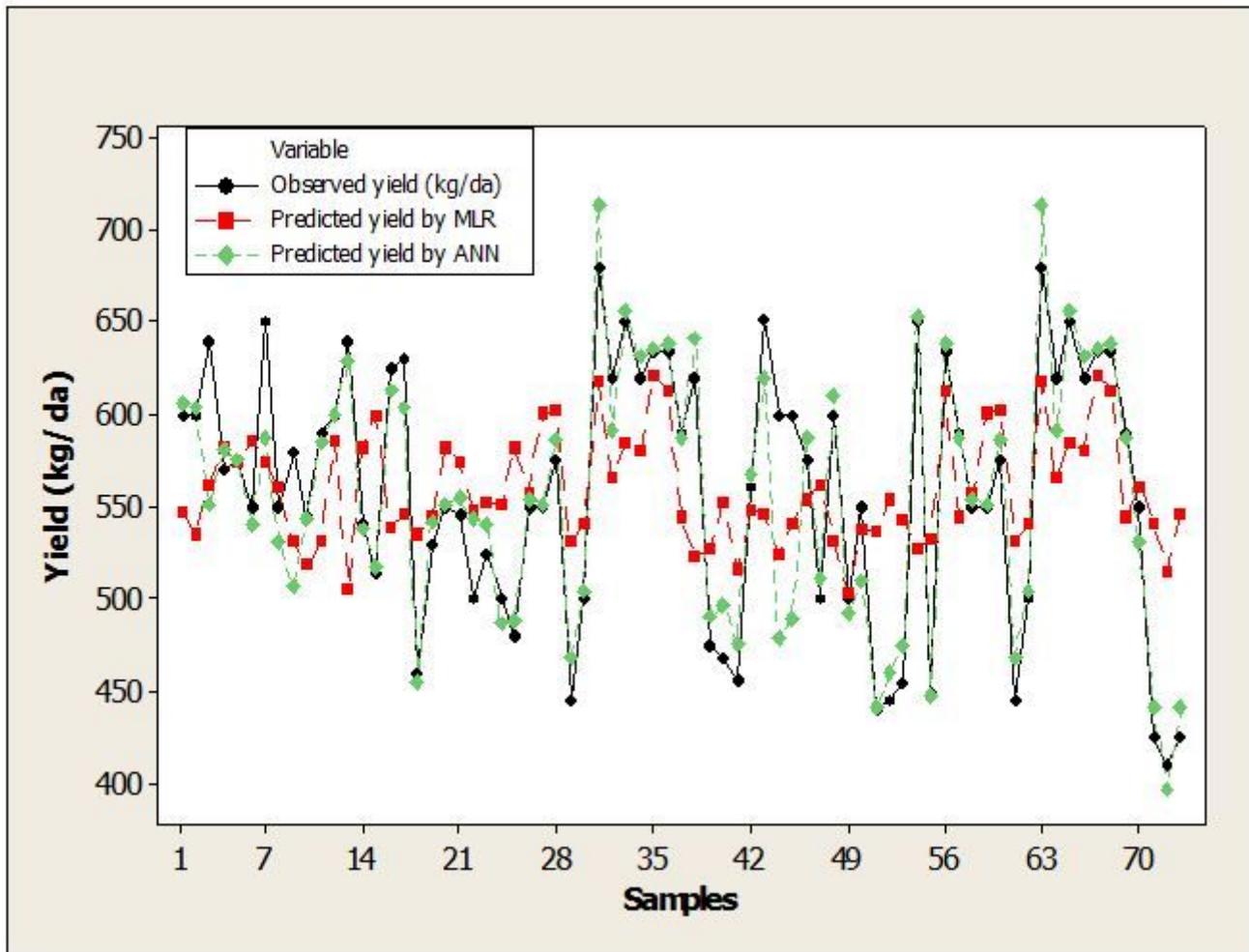


Figure 6

Observed yield values, yield values estimated with MLR and estimated yield values with ANN