International Journal of Applied Biology and Environmental Sciences
E-ISSN: 2667-6540 5(1): 30-41, 2023
Doi: 10.5505/ijabes.2023.61687
Research Article

FORECASTING OF TURKEY'S GRAIN YIELD AN ANNUAL BASIS BY USING MULTI-LAYER PERCEPTRON NEURAL NETWORK: WHEAT, BARLEY, RYE AND OAT

Zekiye Budak Başçiftçi[1,a], Nazife Gözde Ayter Arpacıoğlu[1,b*], Eyyüp Gülbandılar[2,c], Murat Olgun[1,d]

[1]*Osmangazi University, Faculty of Agriculture, Field Crop Department, Eskişehir*
[2]*Osmangazi University , Faculty of Engineering and Arthitecture, Department of Computer Engineering,Eskişehir*

**Corresponding Author:*
E-mail: gaytergu.edu.tr

(Received 16th February 2023; accepted 17th April 2023)

a: ORCID0000-0002-4034-2537, b: ORCID 0000-0002-5121-4303, c: ORCID 0000-0001-5559-5281,
d: ORCID 0000-0001-6981-4545

ABSTRACT. The objective of this study was to analyze acreage, production, and yield on cereals (wheat, barley, rye and oats), cereals in Turkey for long-term period, taking into account the changes occurred, forecasts in crop production for the future population was determined by used Multi-Layer Perceptron Neural Network Model (MLPNN). The model application was performed with software by using the time (year), population, wheat, barley, rye and oat production, croplands of wheat, barley rye and oat contents as the inputs/output variables for MLPNN, and yields of wheat yield, barley rye and oat contents. The performance of the proposed MLPNN model is very high and its use for forecasts in crop production for the future population (between R^2 = 0.97 and R^2=0.99 for training model).

Keywords: *Cereal production, forecasting, multi-layer perceptron neural network, population, yield*

INTRODUCTION

Production, consumption and trade balances are changing rapidly in basic agricultural products that are rapidly increasing in the world. Increasing population and increasing incomes and consumption habits of people, the fact that agriculture has a globally increasing commercial value, and the fact that it is a factor affecting the economies of the countries in terms of added value and employment increases the importance of the agricultural sector. As long as humanity exists, agriculture will continue to exist as a strategic sector and food as a strategic product. Agriculture and food are not a matter of the present, but of our future. Cereals are an important input not only in human nutrition but also in livestock and agriculture-based industry. In Turkey, 23.4 million hectares of land have been used for crop production, 66,4 % (about 16 million ha) of crop production areas, excluding fallow areas, have been allocated to field crops. Wheat with 58% share in cereal production, is followed by barley, oat and rye, respectively. World wheat production, acreage and yield are 565 million tons, 220 million ha and 2,56 ton /ha, respectively. Acreage, production and yield in wheat are 9,8 million ha, 29 million tons 2,60 ton/ha, respectively. Turkey in terms of total wheat area and production Although the seventh place in the world, almost all the wheat produced is used for domestic consumption and does not have a voice in world wheat trade. Turkey with an average of 285 kg of wheat consumption per person is at the top and 50% of daily calorie needs are met from wheat. In addition to wheat, barley, rye and halter have an important potential both in human use and in animal husbandry. For a healthy animal production, the cultivation area and production of barley, oats and rye, which have an important place in animal husbandry, should be increased. In future, not only in the world but in Turkey, with the rapidly increasing population, the need for wheat, barley, rye and oats production will tremendously increase. As long as these needs are not met by production or if they are not sufficiently produced, these needs will be met by importation. In wheat Today's production, consumption meets the total population in the near future will be much more than today's population, considering that the production of wheat in the same period should be increased significantly. In this study, cereals in Turkey acreage, production, and yield in cereals (wheat, barley, rye and oats) for long-term period were analyzed, taking into account the changes occurred, forecasts in crop production for the future population was determined.

Recently, artificial intelligence methods such as artificial neural networks are widely used in agriculture as in many fields. Alvarez was investigated at the regional scale the effects of soil properties and climate on wheat yield in the Argentine Pampas in order to generate models suitable for accounting spatial and inter annual yield variability [1]. Rahman and Bala were reported a study to the use of ANN models as a new approach for predicting growth on three different locations of jute such as leaf dry matter, root dry matter and bark dry matter [2]. Safa et al. (2015) developed two model to an artificial neural network (ANN) approach and multiple linear regression to estimation for the wheat production [3]. Zeng et al. aimed to test the efficiency of statistical models established using partial least squares regression and artificial neural network (ANN) in predicting seed yields of sunflower [4] . Gos et al. (2020), predicted with combined Trend and Seasonal Components (TBATS) and Support Vector Machines (SVM) model to estimate time series of the minimum and maximum daily air temperatures in a period of six years for various climatic localizations in Europe [5]. Chmielewski and Köhn were studies that show how strong and in which months the grain yield of spring barley and oats, as well as their yield components, are affected by weather using factor analysis [6] . In recently, some studies were produced that it was forecast of wheat yield by using optical and radar satellite images [7] [8] . As we summarized above, we could not find artificial intelligence modeling studies for the production of grain production for the coming years as a result of our literature review. In this study, we planned a study on the forecasting of grain yield in the coming years by using multi-layer perceptron neural network (MLPNN).

ARTIFICIAL NEURAL NETWORKS

w_1, Artificial Neural Networks (ANN) is the mathematical modeling of biological structure and function of the nervous system. Popular interest in neural networks first began in 1943 with a simple introduction to the structure of neurons and then emerged hundreds of different ANN model. The differences among these models arise from their functions, applied values, topologies and learning algorithms [9] . An artificial neural network consists of simple processing units, neurons in other words, and weighted connections between neurons [10].

In the mathematical model of the neuron, the signals are multiplied and summed with the weights of the connections they pass through. The sum is fed into a transfer function, which is called activation. The learning ability of an artificial neuron depends on the success of the selected algorithm in regulating weights. A typical artificial neuron is shown and a model of a multilayer neural network in Figure 1 and Figure 2, respectively. Neuron inputs are represented by x_1, x_2,.., x_n; neuron output is represented by y; the weights of the inputs are represented by w_2,...,w_n; and the activation (transfer) function is represented by f [11] [12].

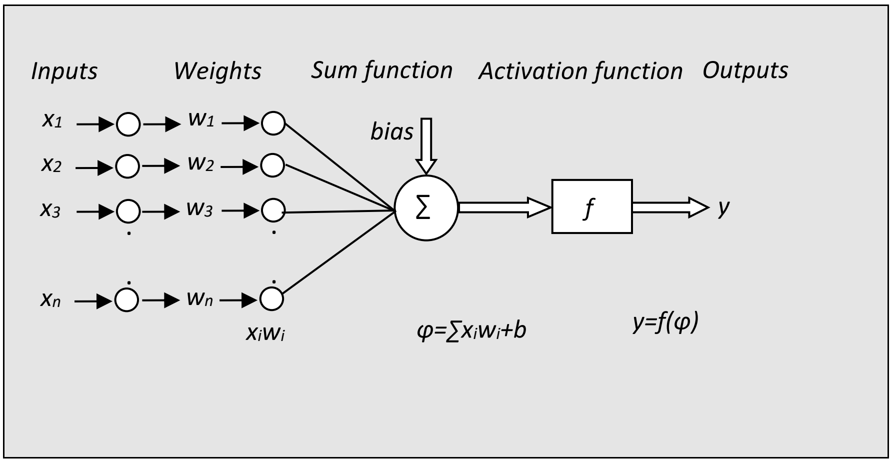

Fig. 1. *Artificial neuron structure [12]*

An ANN model consists of three perceptual units: the input layer, the intermediate layer (hidden layer) and the output layer (Figure 2). There may be a single intermediate layer or multiple intermediate layers. For this reason, such networks are called multi-layer perceptron neural network (MLPNN). MLPNN have three main distinguishing features. First, each neuron in the network has a nonlinear activation function. Second, the intermediate layer may contain one or more layers that are not part of the input or output layers. Third, the network has a high degree of connectivity due to synapses [13].

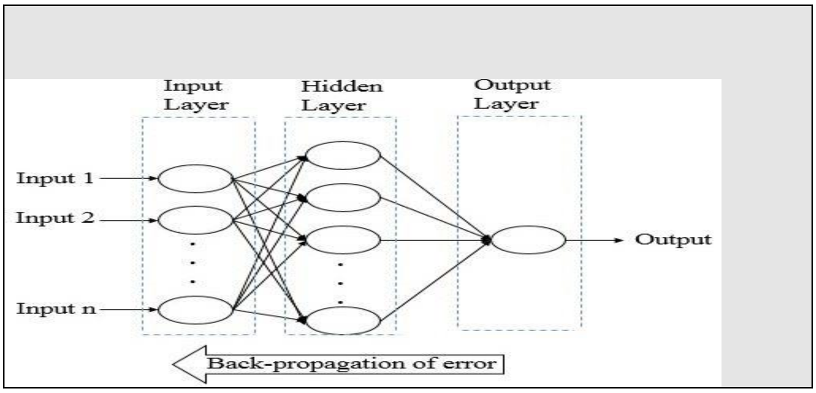

Fig. 2. *Multi-Layer perceptron neural network model [14]*

Artificial neural networks are designed to calculate the closeness of the outputs of the output layer to the actual outputs, or the targets in other words. This difference between the target and output is called error and calculated with mean square. This process of approximation is called network training. In this process, the weight values are adjusted and updated to reach the target using mean square. Various algorithms such as Hebbian learning and perception learning have been developed for this. The model presented in this study focuses especially on the back-propagation algorithm because it utilizes this algorithm.

An artificial neural network is trained using two techniques: online mode and batch mode. The number of weight updates for the same data set is quite different in these two methods. Online training updates a weight for each data point, but batch training updates a weight for the entire data set. Here, in the process, there is an expression called epoch (i.e., number of iterations), which refers to the number of times the data set will be updated. Therefore, updating weights once for each epoch is called batch mode, and updating weights for each data point is called online mode [11].

The back-propagation algorithm works in multilayer ANN based on the network's learning what it needs to do through examples. The examples of what the network is required to do are fed into the algorithm and the network weights are updated so that the network can reach the target. The network will be provided by the system with the outputs for the inputs determined when learning completed. Back-propagation networks are more suitable for use in pattern recognition and mapping [13].

APPLICATION

Data Collection And Artificial Neural Network Modeling

In this study, a MLPNN model was designed by using the results of 91 sample analyzes obtained through the experimental process output and government report. All data were normalized between 0 and 1 using the standard equation. A total of 91 samples were reserved for 53 sample to training (government report between 1960 and 2012 years) and 38 data to testing (linear regression (LR) prediction value between 2013 and 2050 years).

Firstly, a reliability analysis was performed by using SPSS.16, assuming that the target is wheat yield, barley yield, rye yield and oat yield contents and the inputs are the time, population, wheat production, barley production, rye production, oat production, wheat cropland, barley cropland, rye cropland and oat cropland contents. As can be seen in Fig. 3 on the importance values of the output variables, the impact scores of grain yields, which was the most effective parameter, wheat production, oat cropland, barley production, oat production, wheat cropland, rye production, rye cropland, population, time and barley cropland were 0.167, 0.156, 0.146, 0.117, 0.082, 0.081, 0.079, 0.061, 0.055 and 0.055, respectively. Therefore, these parameters were selected as inputs of the MLPNN application.

The network model of the developed MLPNN back-propagation algorithm by using MatLab Toolbox is shown in Fig. 4. The model application was performed with software by using the time (year), population, wheat production, barley production, rye production, oat production, wheat cropland, barley cropland, rye cropland and oat cropland contents as the inputs for MLPNN, and wheat yield, barley yield, rye yield and oat yield contents as output variables for MLPNN. The designed MLPNN consisted of feed-forward back propagation, two hidden layers, training function (Levenberg-Marquardt), adaptation learning function (Gradient descent learning function), transfer function (Hyperbolic Tangent Sigmoid Transfer Function) and performance function (MSE-mean squared error) as demonstrated in Figure 4. Momentum rate and learning rate values were determined, and the model was trained through iterations. The parameter values obtained from MatLab Toolbox-based MLPNN back-propagation algorithm was given in Table 1.

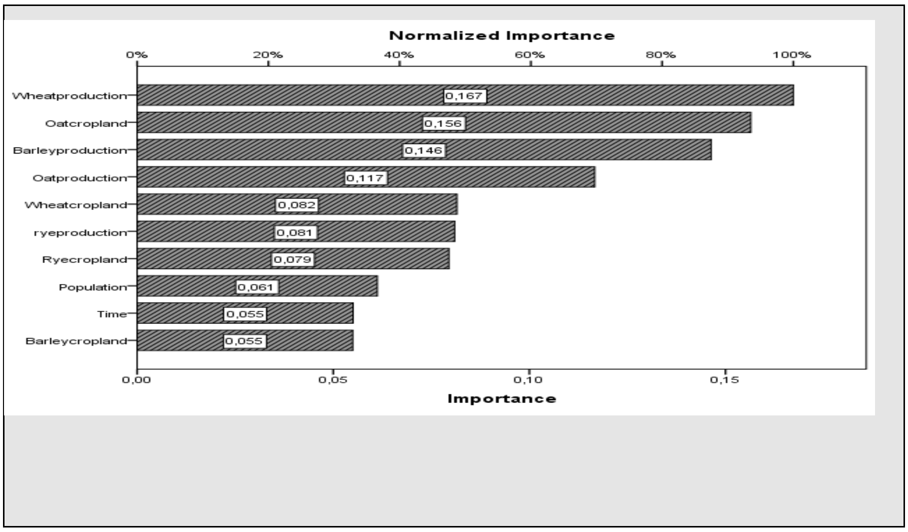

Fig. 3. *Reliability analysis of government grain yields results*

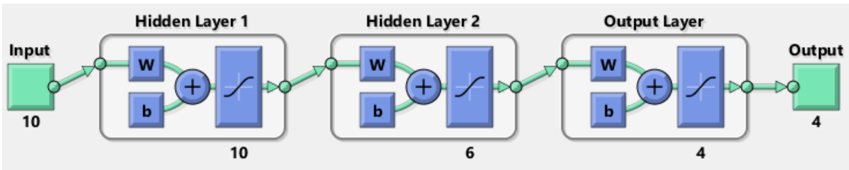

Fig. 4. *Block schema of network model developed in application*

Table 1. *The values of parameters used in developed model*

Parameters	**Value**
Number of input layer neurons	10
Number of hidden layers	2
Number of first hidden layer neurons	10
Number of second hidden layer neurons	6
Number of output layer neuron	4
Error after learning	1.92×10^{-4}
Learning rate	0.7
Epoch	1000
Momentum rate	0.5

The network models tried to be compared according to the absolute fraction of variance (R^2), mean absolute percentage error (MAPE) and a root-mean squared (RMS) error criteria. These criteria are defined by Eqs. (1), (2) and (3) respectively.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left|t_i - o_i\right|^2} \quad (1)$$

$$R^2 = 1 - \left(\frac{\sum_{i=1}^{N}(t_i - o_i)^2}{\sum_{i=1}^{N}\left(t_i - \overline{o_{all}}\right)^2} \right) \quad (2)$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\left(\frac{t_i - o_i}{t_i}\right)\right| * 100 \quad (3)$$

Here ***t*** is the target value, ***o*** is the network output value, ***N*** is the total number of patterns [15].

Application Of Developed MLPNN

After the execution of the MLPNN application, the process was terminated when the acceptable error value was reached in the training process. The real grain yield values of government report were compared to the grain yield values resulting from the training of the MLPNN application (Figure 5, Figure 6, Figure 7 and Figure 8). There was a significant relationship between R^2, MAPE and RMSE values on the figures and the data. For the IBM SPSS packet software program analyses, it was performed on the training data did not find any statistically significant difference (C_{wheat} =0.995, C_{barley}=0.987, C_{rye}=0.987, C_{oat}=0.992, $P<0.001$).

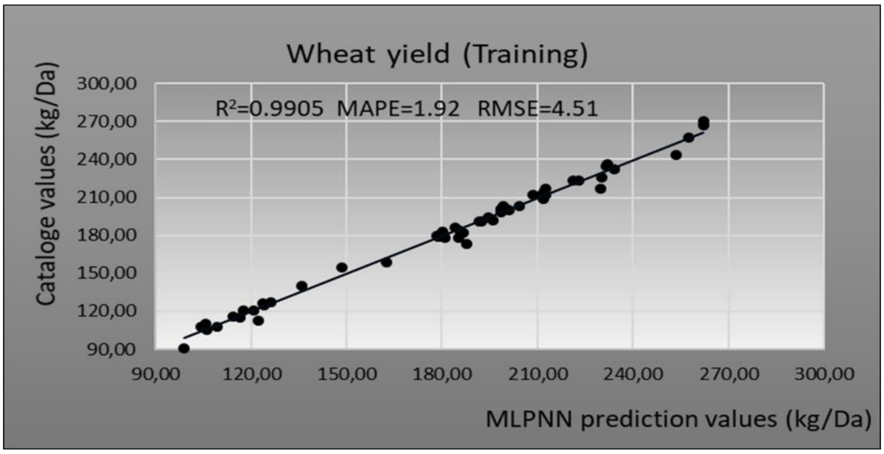

Fig. 5. *Scatter plot of government report results with training resultsof model for wheat yield.*

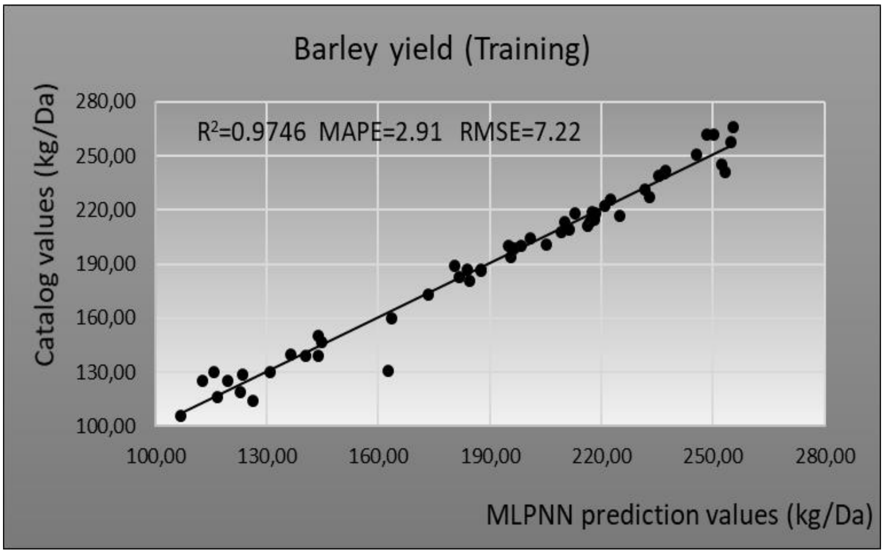

Fig. 6. *Scatter plot of government report results with training results of model for barley yield.*

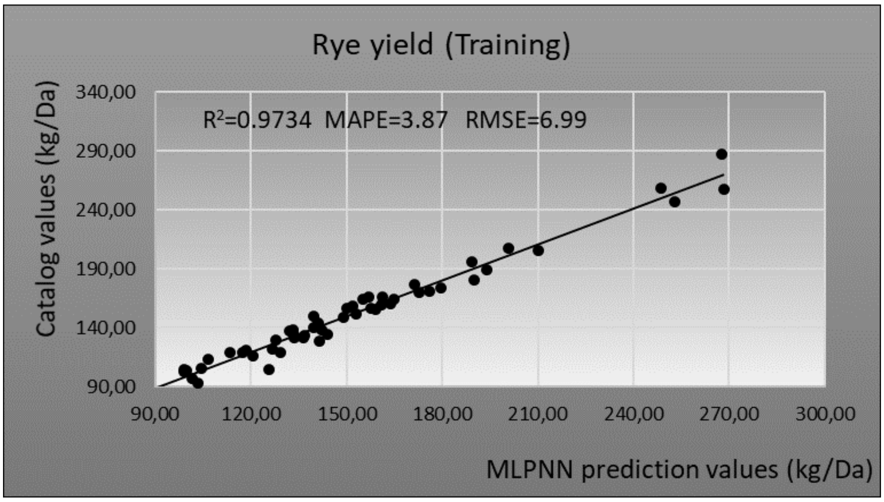

Fig. 7. *Scatter plot of government report results with training results of model for rye yield.*

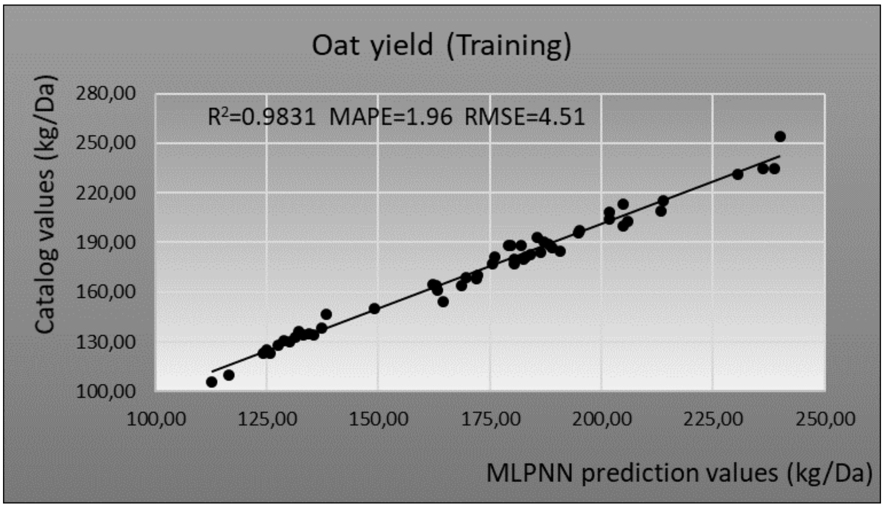

Fig. 8. *Scatter plot of government report results with training results of model for oat yield.*

38 data samples randomly selected from the dataset for the trained MLPNN model were used to test the model. The grain yield values obtained from the test data and experimental results of the trained MLPNN model were compared (Figure 9, Figure 10, Figure 11 and Figure 12). As can be seen in the scatter plot, there was a significant relationship among the data for R^2, MAPE and RMSE results. In addition, the analyses performed on the test data did not find any statistically significant difference (C_{wheat} =0.507, C_{barley}=0.956, C_{rye}=0.807, C_{oat}=0.712, P<0.001).

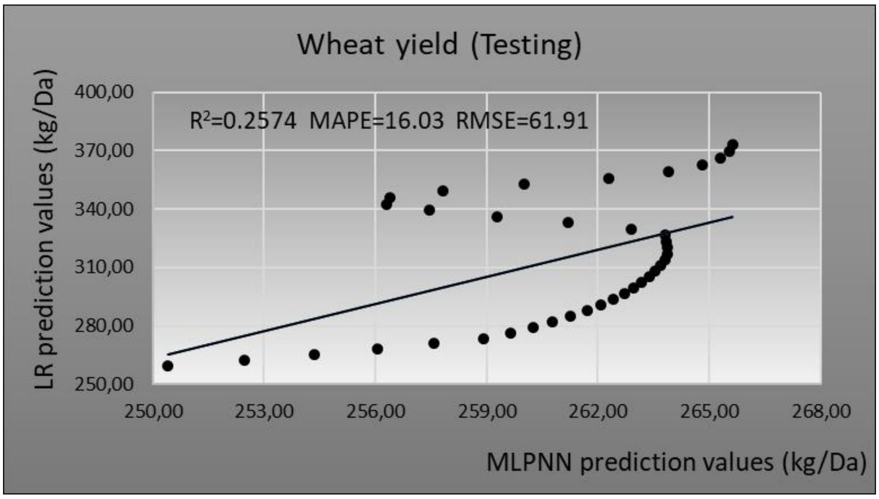

Figure 9. *Comparison between government report results and testing results of model for wheat yield (LR: Linear Regression).*

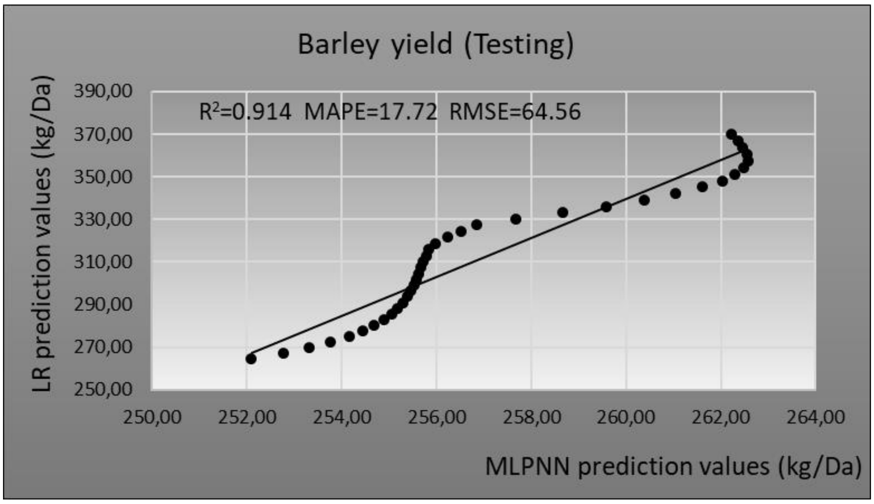

Fig. 10. *Comparison between government report results and testing results of model for barley yield (LR: Linear Regression).*

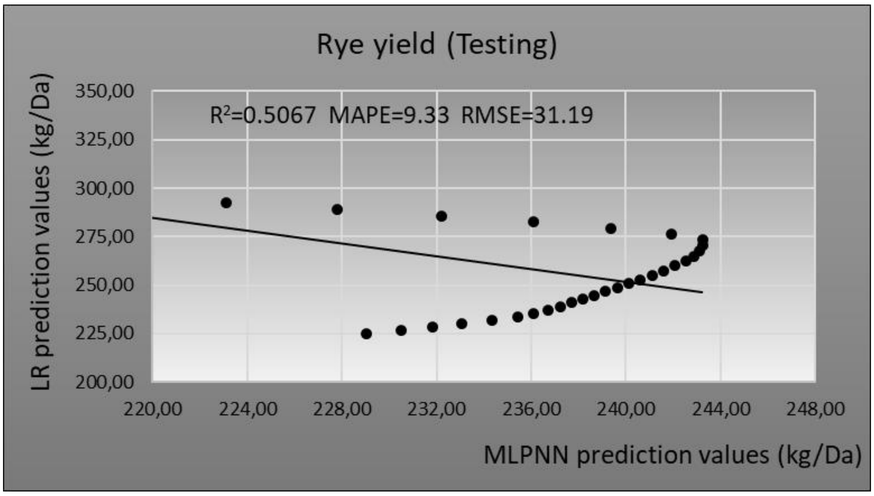

Fig. 11. *Comparison between government report results and testing results of model for rye yield (LR: Linear Regression).*

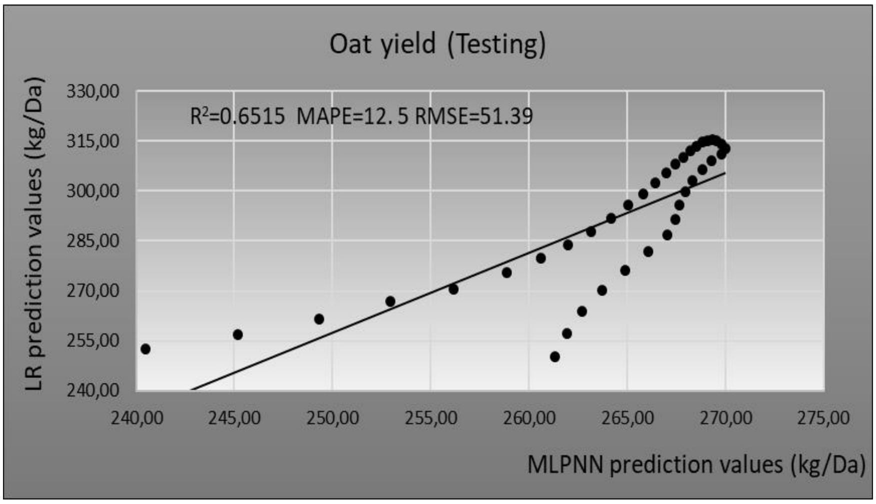

Fig. 12. *Comparison between government report results and testing results of model for oat yield (LR: Linear Regression).*

CONCLUSION AND RECOMMENDATIONS

Wheat yield, barley yield, rye yield and oat yield as the output parameters, a reliability analysis was performed in SPSS for the time, population, wheat production, barley production, rye production, oat production, wheat cropland, barley cropland, rye cropland and oat cropland contents as the input parameters. For the importance values of the output parameters, wheat production, oat cropland, barley production, oat production, wheat cropland, rye production, rye cropland, population, time and barley cropland as an input parameters were 0.167, 0.156, 0.146, 0.117, 0.082, 0.081, 0.079, 0.061, 0.055 and 0.055, respectively. It is not a coincidence for both the positive results obtained in the training levels of the developed MLPNN model and high degree results in the importance levels of the selected input variables. It shows how positive both the selection of the developed model parameters and the selection of input variables.

As mentioned above, the MLPNN application performed in MatLab Toolbox, wheat production, oat cropland, barley production, oat production, wheat cropland, rye production, rye cropland, population, time and barley cropland were used as the input parameters. The MLPNN was trained using a total of 91 data sets and tested with 38 data sets.

The linear regression analysis for the real grain yield values of government report were compared to the grain yield values resulting from the training of the MLPNN application, We did not find any statistically significant difference between R^2 = training data (C_{wheat} =0.995, C_{barley}=0.987, C_{rye}=0.987, C_{oat}=0.992, P<0.001). Similar success was achieved during the MLPNN testing stage, too.

The results showed that the performance of the proposed MLPNN model is very high and its use for forecasts in crop production for the future population. Data sets in the MLPNN model proposed and implemented in this study were obtained from government report between 1960 and 2012 years and prediction value between 2013 and 2050 years. Cereals consisting of wheat, barley, oats and rye have an important place in nations' nutrition in the world. Agricultural production and agriculture-based industry are not independent from other industrial practices in countries with national economies. Therefore, changes and fluctuations in agricultural production affect the country's economy. In terms of feeding the increasing population in the future, agricultural production should also be increased. For this purpose, extending acreage and increasing barley, oats and rye, as well as wheat, are mandatory for both the food industry and animal husbandry.

Cereals, consisting of wheat, barley, oats and rye, have an important place in human nutrition and livestock in the world. Agricultural production couldn't be considered independently of the economy, economic developments, and supports for agriculture depend on the development of economies. Cereals play an important role in agricultural production that is so vital in the country's economy. The production of cereals, an indispensable plant for agricultural production, where dry farming is dominant, should be increased. Increasing the production of cereals, including wheat, barley, rye and oats is imperative not only for human nutrition but for the food industry and livestock.

REFERENCES

[1] Alvarez, R. (2009): Predicting average regional yield and production of wheat in the Argentine Pampas by an artificial neural network approach. European Journal of Agronomy, 30, 70–77.

[2] Rahman, M.M. & B.K., Bala (2010): Modelling of jute production using artificial neural networks, Biosystems Engineering, 105, 350–356

[3] Safa, M., S. Samarasinghe & M. Nejat (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 17(4): 791-803.

[4] Zeng, W., Xu, C., Zhao, G., Wu, J. & Huang, J. (2018): Estimation of sunflower seed yield using partial least squares regression and artificial neural network models, Pedosphere, 28(5), 764-774.

[5] Gos, M., Krzyszczak, J., Baranowski, P., Murat, M. & Malinowska, I. (2020):Combined TBATS and SVM model of minimum and maximum air temperatures applied to wheat yield prediction at different locations in Europe, Agricultural and Forest Meteorology,281,107827,https://doi.org/10.1016/j.agrformet. 2019. 107827

[6] Chmielewski, F.M. & Köhn, W. (1999): Impact of weather on yield components of spring cereals over 30 years, Agricultural and Forest Meteorology, 96, 49-58.

[7] Fieuzal, R. & Baup, F. (2017): Forecast of wheat yield throughout the agricultural season using optical and radar satellite images, International Journal of Applied Earth Observation and Geoinformation , 59, 147–156.

[8] Lurstwut, B. & Pornpanomchai, C. (2017): Image analysis based on color, shape and texture for rice seed (Oryza sativa L.) germination evaluation, Agriculture and Natural Resources, 51(5), 383-389. https://doi.org/10.1016/j.anres.2017.12.002

[9] Gershenson, C. (2003): Artificial Neural Networks for Beginners, Sussex, UK, pp. 3-4. https://arxiv.org/ftp/cs/papers/0308/0308031. pdf (12.01.2011)

[10] Kriesel, D. (2005): Brief Introduction to Neural Networks, pp. 37-58. http://www.dkriesel.com/_media/science/neuronalenetze-en-zeta2-2col-dkrieselcom.pdf (22.05.2017)

[11] Abraham, A. (2005): Artificial Neural Networks, Oklahoma State University, Stillwater, Oklahoma, USA, pp. 901-906.

[12] Chakraborty, R. (2010): Fundamentals of Neural Networks: AI Course Lecture, 37-38, pp. 4-14.

[13] Eluyode, O.S. & Akomolafe, D.T. (2013): Comparative study of biological and artificial neural networks. European Journal of Applied Engineering and Scientific Research, Vol. 2. No.1. pp. 36-46.

[14] Jadidi, A., Menezes, R., Souza, N. & Lima, A.C.C. (2019): Short-term electric power demand forecasting using NSGA II-ANFIS model. Energies, 12(10):1891, DOI: 10.3390/en12101891

[15] Doğan, A. & Gönüllü, G. (2018): Current Debates in Sustainable Architecture Urban Design & Environmental Studies, (Chapter Editors: Kara, Ç., Özden, G., Karacasu, M., Sarıçiçek, İ.: Artificial Neural Network Modelling For Mode Choice Prediction In Transportation Planning), IJOPEC Publication Limited, İstanbul, pp.102-110.