STUDY OF ENVIRONMENTAL FACTORS AFFECTING THE GROWTH AND PRODUCTION OF BARLEY PLANT BY USING ARTIFICIAL NEURAL NETWORK TECHNOLOGY

Shaimaa M. Baraka, El-Awady M.N., Zeinab H. Behairy and Genaidy M.A.
Agric. Engineering Dept., Fac. of Agric., Ain Shams Univ., P.O. Box 68, Hadyek Shoubra 11241, Cairo, Egypt
*Corresponding author: shaymaa_baraka@agr.asu.edu.eg
Received 4 July, 2019 Accepted 29 July, 2019

ABSTRACT

The impact of climate on crop production has vital importance. Climate variables affect the different crops during different stages of the growth and the development.

This research aims to study the environmental factors affecting the growth and production of barley (Hordeum Sp., Gramineae) in a hydroponic system, to provide information to farmers and decision makers by using Artificial Neural Network (ANN) Model for production prediction.

Multilayer feed-forward ANN (fully connected) was used in supervised manner and the training method was the back-propagation algorithm by using MATLAB program.

The inputs in the ANN model of barley were: seeds density (kg/m²), lighting duration (h/day), light intensity (Lux), temperature (ºC), relative humidity (%), and growing period (days). The outputs were: plant length (cm), yield (kg/m²), protein (%), dry matter (%), and conversion factor.

Results revealed that the optimal configuration for the ANN model consisted of four layers (6-25-30-5). The hidden layers had 25 and 30 nodes in the first and second hidden layers respectively for the ANN model. Hyperbolic tangent transfer function was employed in hidden and output layers of the ANN model. The learning rate and the momentum parameter were 0.005 and 0.9 respectively for the ANN model. Iterations were 10000 epochs during training process for the ANN model. The results showed that the variation between target and predicted outputs was small while the correlation coefficient (R) was 0.99.

Also, the results revealed that the major parameters affecting on all the outputs were seeds density and the duration of the lighting followed by the other factors i.e. temperature (ºC), relative humidity (%), growing period (days) and light intensity (Lux). Seeds density has a higher percent relative importance, on yield, plant length, protein (%), DM (%), and conversion factor equal to 22.8%, 24%, 25%, 24% and 22.8% respectively.

The developed ANN model was beneficial tool for barley production prediction. The barley yield prediction could be helpful for farmers, decision makers and planning to manage their crop better by providing a series of recommendations about crops planting and clarifying its impact on changes to these factors under the study in order to avoid losses and reach the best benefit (maximization of yield).

Keywords: Artificial Neural Network (ANN), Barley plant, Environmental factors, Growth and production, Hydroponic system.

INTRODUCTION

The impact of climate on crops production has vital importance. Climate variables affect the different crops during different stages of the growth and the development.

In the agricultural sector, it is important to study the impacts of climate change on crops production in North African countries as in Egypt (Abou-Hadid, 2006).

The modern agricultural researches should be required to provide information to farmers and decision makers on how to accomplish sustainable agriculture over the wide variations in climate
around the world (Murthy, 2004). Also, there is a need to develop the product-climate models to analyze plant reaction to climate changes. (Lamba and Dhaka, 2014).

Artificial Neural Networks (ANNs) are one of the most important techniques used to solve these problems because they have the ability to process data that is very complex compared to traditional computing techniques (Hamoda et al. 1999). ANNs have been applied successfully to hundreds of applications. 

Lamba and Dhaka (2014) represent the forecasting models such as Statistics, Simulation, Remote Satellite Sensed and Mathematical in the field of Agriculture (Wheat crop). It showed a compact combination of all these models and shows why Neural Network Model is important from other models for nonlinear data behavior system like wheat crop yield prediction.

Mokarram and Bijanzadeh (2016) analyzed and compared Multiple Linear Regression (MLR), and (ANN) including Multi-Layer Perceptron (MLP) and Radial Basis Function (RBF) models to predicting Biological Yield (BY) and Yield (Y) of barley. Among the MLR, MLP and RBF models, MLP model had the highest R² values for the prediction of BY (R²=0.894) and Y (R²=0.922). Overall, ANN models can be used to successfully estimate BY and Y from data.

This research aims to study the environmental factors affecting the growth and the production of barley to provide information to farmers and decision makers by using ANNs.

Therefore, objectives of this research are

1. Recording the environmental factors that affecting on the growth and the production of barley.
2. Clustering data related to the environmental factors affecting the growth and the production of barley from literatures.
3. Dividing the data set to 3 groups: training, testing and validation.
4. Construct, train, and valid ANN model using the clustered data to determine the optimal model.
5. Testing the developed ANN model using data collected from literatures that are not used in training the developed ANN model.
6. Depicting and obtaining relationships among the affecting factors and the plant production.

MATERIALS AND METHODS

- Procedures

The general multilayer Artificial Neural Network (ANN) design is used. It has seven primary steps:

1. Preparing data set for the network
2. Constructing the network
3. Configuring the network
4. Training the ANN model
5. Verifying the ANN model
6. Testing the ANN model
7. Sensitivity analysis

- Preparing data set for the network

Before beginning the network design process, the sample data were prepared first. It might happen outside the framework of ANN software, but this step is critical to the success of the design process, therefore, the network can only be as accurate as the data that are used to train the network.

- Recording the environmental factors that affect the growth and the production of barley

The used ANN model was implemented under 6 inputs and 5 outputs. The inputs are: seed density (kg/m²), lighting duration (h/day), light intensity (Lux), temperature (°C), relative humidity (%) and growing period (days). But the outputs are: vegetative yield (kg/m²), plant length (cm), protein (%), dry matter (%) and conversion factor as shown in Fig (1).
Study of Environmental Factors Affecting the Growth and Production of Barley Plant by Using Artificial Neural Network Technology

AUJASCI, Arab Univ. J. Agric. Sci., 27(2), 2019

- Clustering data related to the environmental factors

  Source of clustered data


- Constructing the network

  After the data have been collected, the next step is to construct the network. To define a problem, arrange a set of input parameters as columns in a matrix, then arrange another set of target parameters (the correct output vectors for each of the input vectors) into a second matrix.

- Dividing the data set

  All data were divided randomly into three sets. The first assigned for the training process (70%), the second for the validation process (15%), and the third for the testing process (15%).

  MATLAB (R2015a) that includes an “Artificial Neural Network toolbox” (nntool) was used to construct, train, and simulate neural networks.

  For each data set, the optimal value of ANN variables such as: number of hidden layers, number of neurons in each hidden layers, the type of transfer function, learning runs and iterations need to be determined.

  The optimum values were based on minimizing the difference between the ANN output and the desired output (RMSE) and maximization correlation coefficient (R).

  For visualization the optimal ANN architecture, the trial and error procedure was used in the present study.

  Multilayer feed-forward ANN was chosen as one hidden layer and two hidden layers with a trial number of nodes and different transfer function in the hidden layers.

  The ANN was trained for a trial number of epochs with different learning rates and a fixed momentum, and the error gradient was observed over these epochs. Then, increasing or decreasing the number of hidden nodes changed the ANN architecture. The training process was repeated for the new architecture. This procedure was continued for several different architectures. Eventually, the ANN architecture that resulted in the high correlation coefficient (R) and less root mean square error (RMSE) over the training epochs was adopted as the optimal ANN architecture.

  In the present study, the performance of the proposed ANN model is examined for two distinct types of transfer functions of the hidden layer namely; the sigmoid and hyperbolic tangent function but transfer function for the output layer was hyperbolic tangent. The number of nodes ranged...
from 2 to 30 in the first hidden layer and ranged from 4 to 30 assigned for the second hidden layer. The number of epochs was 10000, 20000, and 40000. The learning rate ranged from 0.001 to 0.01. In addition, momentum parameter of 0.9 was kept the same during this test.

Once the optimal configuration with respect to the number of neurons and learning runs was found and the performance of the ANN was tested on different sizes of data sets chosen in a random fashion.

In the present study, multi-layer feed-forward ANN (fully connected) was used in a supervised manner and the training method is the back-propagation algorithm.

Basically back-propagation consisted of the presentation of a set of examples and the desired outputs. The ANN is then presented with a set of training patterns; each consisting of an example of the problem to be solved (the input) and the desired solution to this problem (the outputs). These training patterns are presented repeatedly to the ANN model and weights and thresholds are adjusted by small amounts that are dictated by the general delta rule. This adjustment is performed after each iteration, when the ANN has computed output is different from the desired output. This process continues until weights converge to the desired error level or the output reaches an acceptable level.

- Sensitivity analysis

A sensitivity analysis is performed to investigate the effect of each input parameter on the output of the weights method as a procedure for partitioning the connection weights to determine the relative importance (RI %) of the various inputs (Nourani and Fard, 2012).

Connection weights algorithm (CW):

\[ Q_{hi} = \sum_{n=1}^{nh} W_{hi} \cdot W_{nh} \]  
\[ RI(\%)_i = \frac{\sum_{h=1}^{nh} Q_{hi} \cdot 100}{\sum_{l=1}^{ni} \sum_{h=1}^{nh} Q_{hi}} \]

Where:

- \( W_{hi} \): Weights of the connection from input neuron to hidden neurons, and
- \( W_{ho} \): Weights of the connection from hidden neurons to output neuron.

RESULTS AND DISCUSSION

- The optimal ANN model

From the previous experience was obtained the best ANN model with one hidden layer which architecture unit was (No. of nodes of 25, epoch of 10000, learning rate of 0.003 with Momentum parameter 0.9 and transfer function for the hidden and output layer was hyperbolic tangent). The computed errors converged to a minimum value for all predicted outputs with two hidden layers and 25 neurons in the first hidden layer and 30 neurons in the second hidden layer. It was achieved at 10000 learning runs (epochs). The hidden layers and output layer has a hyperbolic tangent activation function. The learning rate and momentum parameters were 0.005 and 0.9 respectively.

Figs. (2, 3 and 4) showed that architecture of the unit of the ANN model and the best performance at the maximization correlation coefficient and the minimization RMSE.

**Fig. 2.** Architecture of unit of the ANN model of barley

AUJASCI, Arab Univ. J. Agric. Sci., 27(3), 2019
Study of Environmental Factors Affecting the Growth and Production of Barley Plant by Using Artificial Neural Network Technology

Fig. 3. The best performance of the optimal ANN model

Fig. 4. The correlation coefficient of the optimal ANN model of barley

- Sensitivity analysis

Sensitivities were determined to computing the effect on the ANNs output response. The relative importance (RI %) of the various inputs showed the effect of each input parameter on the outputs.

The results are presented in Table (1) showed that the major parameters affecting on all the outputs are the seeds density and the duration of the lighting. Seeds density had higher percent relative importance, on yield, plant length, protein %, DM% and conversion factor equal to 22.8%, 24%, 25%,
24% and 22.8% respectively. And the least important factor was the intensity of the lighting on yield, plant length and protein by (RI) 12.15%, 11% and 12% respectively. As well as temperature on dry matter by (RI) 11% and relative humidity on conversion factor by (RI) 13%.

Table 1. The relative importance (RI %) of the input parameters on the outputs for the ANN model of the barley plant under a hydroponic system

<table>
<thead>
<tr>
<th>Input</th>
<th>yield</th>
<th>length</th>
<th>Protein</th>
<th>DM</th>
<th>Conversion factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>seed density</td>
<td>22.8</td>
<td>24</td>
<td>25</td>
<td>24</td>
<td>22.8</td>
</tr>
<tr>
<td>(kg/m²)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lighting</td>
<td>21.72</td>
<td>19</td>
<td>20</td>
<td>20</td>
<td>22.3</td>
</tr>
<tr>
<td>(h/day)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lux</td>
<td>12.15</td>
<td>11</td>
<td>12</td>
<td>14</td>
<td>13.5</td>
</tr>
<tr>
<td>Temp. (cº)</td>
<td>14.57</td>
<td>14</td>
<td>13</td>
<td>11</td>
<td>14.9</td>
</tr>
<tr>
<td>RH (%)</td>
<td>14.34</td>
<td>17</td>
<td>16</td>
<td>18</td>
<td>13</td>
</tr>
<tr>
<td>Growing days</td>
<td>14.42</td>
<td>15</td>
<td>14</td>
<td>12</td>
<td>13.6</td>
</tr>
</tbody>
</table>

- Application of ANN-based equation

The architecture of the network developed in this study is a multilayer perceptron (6-25-5). It has six inputs (Xᵢ, i=1 to 6), five output (z) and twenty five neurons in the hidden layer. The transfer function is the hyperbolic tangent. This mathematical definition is given in Eq. (3):

\[ f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \]  

Each of these input layer’s neurons receives one input (Xᵢ, i=1 to 6) and broadcasts such signal to each one of the hidden layer’s neurons. Each hidden neuron computes its transfer function and sends its result (Yᵢ, j=1 to 25) to the output layer’s neuron which finally produces the response of the network (Z).

The output signal of each hidden neuron (Yᵢ) is calculated as Eq. (4):

\[ Y_j = f \left( \sum_{i=1}^{6} W_{ji} X_i + b_j \right) \]

Where:

- \( W_{ji} \): Weights of the connections between the input and hidden neurons.
- \( X_i \): The input variables.
- \( b_j \): The bias on hidden neuron.

While the output of the network is given by Eq. (5):

\[ Z = f \left( \sum_{j=1}^{25} W_{j1} Y_j + b_1 \right) \]  

Where:

\[ W_{ji}, X_i, Y_j, b_j \] are the parameters of the network.

The mathematical formula for predicting using the ANN approach is given in Eq. (5).

The developed ANN model was beneficial tool for barley production prediction; it was used to predict the yield by sensing variable parameters of seed density (kg/m²), lighting duration (h/day), light intensity (Lux), temperature (cº), relative humidity (%) and growing period (days) related to vegetative yield (kg/m²).

The barley yield prediction could be helpful for farmers, decision makers and planning to manage their crops better by providing a series of recommendations about crop planting and clarifying its impact on changes to these factors under the study in order to avoid losses and reach the best benefit (maximization of yield).

Furthermore, these systems can readily be used in commercial greenhouses so the derived ANN models are relatively easy to deploy to a commercial setting where they can subsequently be improved over time.

In order to obtain the highest productivity, we must provide the appropriate conditions from the other factors. El-Deeba et al (2009) showed that the lighting operating hours of about 12-16 hours are the most suitable for fodder production of barley in intensity of light 2000 Lux on 8 days growing at a temperature of 20-25°C and R.H. of 60-70%.

In a comprehensive review, Ehret et al (2011) introduced all crops attributes responded in much the same way to individual climatic factors. Radiation and temperature generally induced strong positive responses while RH produced a negative response. In the ANN models, radiation and temperature were still prominent, but the importance of CO₂ in predicting a crop response increased.

AUJASCI, Arab Univ. J. Agric. Sci., 27(3), 2019
Application of ANN in predicting wheat crop yield considered by Khairunniza-Bejo et al (2014) who concluded that the best model for prediction of wheat yield was ANN models compared to other models. ANN has become a good method because of its ability to predict, forecasting and classification in biological science fields.

According to the results obtained, it was found that the most important factors affecting the production are the seeds density and lighting and so agreed with (Elsoury et al 2015) who showed a green fodder yield increased when the seeding rate increased from 4.01 to 5.5 kg/m² then decreased when the seeding rate increased to 6.47 kg/m². this finding is agreed with the result of (Sneath and McIntosh, 2003).

SUMMARY AND CONCLUSION

Artificial Neural Networks (ANNs) were applied to study the effect of the environmental factors on the growth and the production of barley to predict plant yield. The data needed to develop a model with minimum error which gathered from literatures. Multilayer feed-forward back-propagation ANN was used in a supervised manner.

The inputs were: seeds density (kg/m²), lighting duration (h/day), light intensity (Lux), temperature (°C), relative humidity (%) and growing period (days). But the outputs are: yield (kg/m²), plant length (cm), protein %, dry matter % and conversion factor. The results showed the following:

- The optimal ANN model (6-25-30-5) consisted of 4 layers, one for inputs (6), two hidden layers, one for outputs (5), neurons in the first hidden layer (25), neurons in the second hidden layer (30), learning rate (0.005), epoch (10000), transfer function (tan-sig) and momentum parameter (0.9).
- The performance index for the optimal ANN model maximization correlation coefficient (R=0.99) and minimization Root Mean Square Error (RMSE = 1.58)
- The most important factors affecting the production and growth of barley as a hydroponic fodder are the seeds density and the duration of the lighting and the least important factor is the intensity of the lighting.
- Seeds density has a higher percent relative importance, on yield, plant length, protein %, DM % and conversion factor equal to 22.8%, 24%, 25%, 24% and 22.8% respectively.

RECOMMENDATION

- The success of the modifying the ANN model performance in the network development and selection sets indicates that it may be a valuable tool for further research.
- Artificial Neural Network (ANN) model was used to supply information to farmers and decision makers for plant production prediction.
- The combination of artificial intelligence and agriculture will become more important in the field of researches. It has the ability to solve the problems of the complex agricultural system which deals with many factors that affect the target outputs and interpret the relationships between these inputs and outputs.
- Undeniably, the application of ANN to precision agriculture plays a crucial role in future evaluation of the concept of precision agriculture as a sustainable means of meeting world’s food demands.

REFERENCES

Abou-Hadid A.F. 2006. Assessment of impacts, adaptation, and vulnerability to climate change in North Africa: food production and water resources. A Final Report Submitted to Assessments of Impacts and Adaptations to Climate Change (AIACC), Project No. AF 90, pp. 1-37.


دراسة العوامل البيئية المؤثرة علي نمو وإنتاج نبات الشعير المستنبت باستخدام تقنية الشبكات العصبية الاصطناعية

المحقق

شيماء محمد بركه - محمد نبيل العوضي - زينب حسين بحيري - محمد عبد المجيد جنيدي
قسم الهندسة الزراعية - كلية الزراعة - جامعة عين شمس - القاهرة - مصر

*Corresponding author: shaymaa_baraka@agr.asu.edu.eg

الموجز

يتأثر الإنتاج النباتي بشكل كبير بالظروف المناخية ويتباين تأثير الطقس علي المحصول خلال مراحل النمو المختلفة.

يُقدم البحث دراسة عن العوامل البيئية المؤثرة علي نمو وانتاج نبات الشعير المستنبت باستخدام تقنية الشبكات العصبية الاصطناعية، لذلك من الضروري دراسة وتحليل العلاقة بين التغيرات المناخية والإنتاج النباتي، وإمداد المزارعين وصانعي القرار بالمعلومات التي تتعلق بأن كيفية تحقيق الزراعة المستدامة على مدى التباينات الواسعة في المناخ.

لتحقيق أهداف الدراسة تم استخدام الشبكة العصبية الاصطناعية (ANN) متعددة الطبقات ذات التغذية الأمامية والتعليم الموجه وطريقه الانتشار العكسي.

كانت مدخلات النموذج لنبات الشعير كالآتي:
- كثافة التقاوي (كجم/م2)
- وطول فترة الإضاءة (ساعة/يوم)
- وشدة الإضاءة (لكس)
- ودرجة الحرارة (م°)
- ونسبة الرطوبة (%%)
- وطول فترة النمو (يوم).

وأدت النتائج أن النموذج الأمثل الذي له إمكانية في التنبؤ عباره عن (6 - 25 - 30 - 5) وهي عدد العناصر في طبقات المدخلات والطبقات المختفية وطبقات المخرجات على التوالي، وكان معدل التعليم (0.05) وعدد المكررات (10000) وكان النتيجة صغر بين المخرجات المتوقعة والشاملة، وعمر الأرباط بساوي 0.98 والجزء الترميز لمتوسط مربع الأخطاء 1.58.

وأخيراً أن كثافة التقاوي وطول فترة الإضاءة هما أكثر العوامل تأثيرا علي النمو والحصاد، بينما كانت شدة الإضاءة هي العامل الأقل تأثيرا.

وبناءً على النماذج التي تم الحصول عليها، يمكن استخدام النماذج المستنبت في التنبؤ بنتائج النمو بدوامات لفترات زمنية مختلفة.

والمصطلحات الدالة: الشبكات العصبية الإصطناعية، الشعير المستنبت، العوامل البيئية.