Wheat Yield Prediction: Artificial Neural Network based Approach

Muhd Khairulzaman Abdul Kadir, Mohd Zaki Ayob, Nadaraj Miniappan

British Malaysian Institute, Universiti Kuala Lumpur, Malaysia

khairulzaman@unikl.edu.my

Abstract: Wheat yield prediction modeling is an important area of study because of its potential contribution to food security since it may be perceived to be a good indicator for global food availability. Many studies have been conducted in order to determine the best models for wheat yield prediction using various types of data which are available; these models include CERES-Wheat model, SIRIUS model and AFRCWHEAT2 model. In this study, our wheat yield prediction model is designed using a Multi-Layer Perceptron (MLP) backpropagation-based- feed forward artificial neural network (ANN). The data used was weather data including: sun, frost, rain and temperature as the input parameters from year 1997-2007. The output parameter of the model is using the wheat yield data for the years 1997 – 2007. The data is divided into three separate sets; – for training, validation and testing. Our MLP was able to predict, wheat yield with an accuracy of 98 %. Hence our MLP based wheat yield prediction model shows great promise as a tool which will be able to provide relatively accurate wheat yield prediction and may be applied to other crops.

Keywords: *Wheat yield; prediction; neural network; ANN; weather; intelligent systems*.

I. Introduction

Crop yield prediction is a very active area of current research interest and has been so since the 1980s. However, in the early days the work was mainly concerned with the study of linear systems models and hence was only concerned with the linear relationships among the various agricultural parameters. Therefore, most of the conventional or traditional models are not able perform well because they were not able to effectively deal with the complexity and non-linear nature of the data [1].

Basically, crop prediction models can be divided into two classes; statistical model and crop simulation model [2]. The early stage of the modeling usually involves statistical methods. This is where the systems use various regression techniques that compute crop yield empirically. On the other hand, simulation models involve the physiologically – based systems of either crops or plant which can affect growth either internally or externally; and is normally involve mathematical analysis in order to predict yield [3].

As an example, let us consider one of the wheat yield prediction models as in the case of the Crop Estimation through Resource and Environment Synthesis - Wheat (CERES) model. This type of simulation model uses a set of data which includes the weather, soil attributes and the detailed management practice which specific farm uses. The next model is representative of the more complex models where a very complex set of data is used to predict wheat yield. ECOSYS [4] and SIRIUS [5] are also categorized as complex models that use lots of different types of data and rely heavily on computer design to simulate the growth of wheat as in the CERES [6, 7].

Although simulation and statistical models have improved to become better crop prediction models, they still not able effectively with a complex data set. However, models based on Intelligent Systems (ISs) techniques are able to overcome this limitation. This type of technique can produce good results by manipulating raw and simple or complex data which they perform in competitively with the more complex models. The most popular ISs technique which has been used for crop or wheat prediction models is Artificial Neural Networks (ANNs) [1, 8].

The importance of wheat production is reflected in recent food security assessment survey conducted by the Department for Environment, Food and Rural Affairs (DEFRA) which concluded that crops, and specifically wheat, are the key indicators in determining global food availability [9]. The main objective of this paper is to study the ANN technique in context of wheat prediction models by using UK weather data as its inputs. Section 2 is concerned with the data acquisition, section 3 discusses the methodology used in the designed mode, section 4 the training of the system, section 5 the results and discussions and section 6 is the conclusion.

II. Data acquisition

In determining which data to use, a study of the factors that can affect wheat yield was conducted before deciding which the most reliable data to be utilized were potentially. This study established that the main factor in the growth of wheat is the weather or climate change which consists of various factors such sunlight, rain, frost and temperature [9, 10].

Temperature data can be divided into two groups. The first sets of temperature data are concerned with the minimum and maximum temperature, and the second set of data was from the seasonal temperature data. Here, the both sets of temperature data were taken from MET historic data collection from 1997 – 2007. The seasonal temperature data values are calculated from monthly data by averaging the three monthly values as seasons. So we have December to February (winter), March – May (spring), June to August (summer) and September to November (autumn). This dataset covered the period 1997 to 2007 on a monthly basis. The unit of temperature is degrees Celsius.

The other sets of weather data (sunlight, frost and rain) were also taken from Sheffield weather station. The readings for each parameter are as follows: frost was read in days, rain in millimeter (mm) and sunlight was taken in hours. This data was recorded on monthly basis from the year 1997 to 2007; as in the case of the temperature data.

Our objective was to predict wheat yield as accurately as possible as determined by the output of our ANN model. The wheat yield data which we used as the target output data was specified in terms of tone and relates to the same period of time as the weather data. The data was taken from the DEFRA statistical data collection site - [11].

Based on our research we understood that there are other factors which will affect the wheat yield and these included; soil condition, effect of pests or plant diseases and so on. These parameters are difficult to get hold of and they are not normally used to predict wheat yield, so we have not included such variables in our calculations. In the next section we will consider our methodology.

III. Methodology

ANN are a well established techniques and further details can for example be found in [12-15]. Generally, ANNs maybe based on single layer network consisting of 3 basic layers; an input layer, hidden layer and output layer.

A. Design network architecture

For our wheat yield prediction model we will be using a MLP network which has been shown to be effective at dealing with either linear or nonlinear data. In this model, seven (7) inputs are used. They are rain, sun and temperature as shown in figure 1. Each of the input will be explained in section 3.2.This architecture of the ANN we used is shown in figure 1 [8, 16].

In back propagation training algorithm, the first step is to initialize all weight and threshold (if any) to a random. After that, the network must activate its backpropagation by using the activation function. Then, the weight for each neuron will be updated accordingly based on the number of neurons for each layer. Finally, all of the above processes will be repeated until the sum square error is less than 0.001 [15].

Geman, et al.[17], Weigend [18], Tetko, et al.[19] and others indicate that there are a lot of ways in which to determine the number of neurons in the hidden layer and usually the optimum number is determined by trial and error. In this work, we experimented with 2, 4, 7, 10, 15, 20, 25 and 30 hidden neurons and the results were compared in order to determine which performed the best.

B. Input parameter

As shown in figure 1, our ANN has 7 input parameters which resulted in total of 924 (7x132) data points. This data set was based on weather or climate change data sets which were explained in the previous section. In this yield prediction model, the dataset was enriched by incorporating an additional 100 data point by generating plus and minus; 5% and 10% in order to improve generalization and to prevent over fitting of the data. After the addition of the enriched data, our dataset for this model consisted of 2772 (7x396) data points. The data was then divided randomly so that 60% was to be used for training, 20% for validation and the remaining 20% to be used for testing the system in order to determine optimum performance in the prediction of wheat yield.

C. Activation function

In ANN, 4 types of activation function which are regularly used are step function, sign function, sigmoid or hyperbolic function and linear function [15]. But, specifically in backpropagation multilayer network, sigmoid function and linear function are usually used as the activation function. This provides smooth and nonzero derivative with respect to input signals. Sometimes this function is called squashing function and is usually known as sigmoidal function [12, 13, 20].

In this wheat prediction model we use tangent sigmoid function in the hidden layer and output layer since as Negnevitsky [15] and others indicate they can accelerate the learning process.

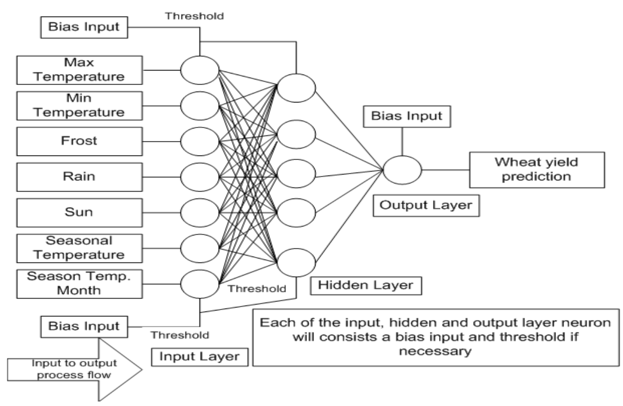

Figure 1: Design of network architecture

No.of Neuron	2	4	7	10	15	20	25	30	35
Training Regression	0.53	0.72	0.91	0.95	0.97	1.00	0.99	0.99	0.98
Validation Regression	0.44	0.65	0.81	0.87	0.91	0.98	0.94	0.91	0.87
Test Regression	0.54	0.61	0.80	0.90	0.93	0.99	0.93	0.80	0.92
All Regression	0.51	0.66	0.84	0.92	0.95	0.99	0.97	0.94	0.94
No. of iteration	20	16	28	40	57	107	58	69	28

Table 1: Comparisons for each hidden layer on number of neuron effect

IV. Training the ANN

In order to train the ANN to predict wheat yield using the wheat data, we use the Levenberg-Marquadt (LM) technique because it performs better than other approaches such as Gauss-Newton (GN) method and gradient descent algorithm [21, 22]. It also typically performs better in terms of, for example, fastest convergence and also it produces better results; for example the mean square error is lower for this yield prediction model. In other case, the main problem in this model is over fitting which may happen when the training set error becomes relatively small [16].

V. Results and discussions

We now apply our ANN model to our wheat yield data. After some experimentation we set the number of epochs to 150 since that seems to be sufficient for us to determine optimal performance for our wheat data; see table 1. In order to determine the maximum error this is based on the mean square error which is set to 8.

Our ANN was tested with the different numbers of hidden neurons; that is: 2, 4, 7, 10, 15, 20, 25, 30 and 35. Each result then based on the result of the regression analysis of the real wheat yield data and the output generated by our ANN model.

In regression analysis, the best result will be based on regression (R) value to 1 or the nearest to 1.The simulation results are shown in table 1 which are based on the enriched data. The best neuron quantity for this model was 20 neurons. It gave an accurate result which is around 0.9. No under training or over fitting to occur for training data, validation data and testing data.

In comparison to the use of original data with the data enrichment addition, the simulation was only being made for 20 neurons in hidden layer which shown in table 2. If we compare figure 2, figure 3 and table 2, it can be seen that the accuracy of the result is around 70% using only the original data, but in the case of the enriched data we see that the accuracy is around 98% for the yield prediction.

If we consider the mean square error (MSE) result for the data enriched data and the data which has been data enriched, the error is much higher for the simulation without using data enrichment as shown in figures 4 and 5. This result was due to fewer data being used for the training and especially for the testing and validation where the results become over fitting or under train. Data quantity gives major effect on this wheat yield prediction model. Greater data quantity will result better accuracy on the wheat yield prediction.

Figure 6, shows the simulation result for the original wheat yield data and prediction of wheat yield where it gives the accuracy of the system as around 98%. Here, it proves that the ANN wheat yield prediction model is a successful model if the model is given sufficient quantity of data input. As in this model, the data input being use is by having an enriched data based on -5%, +5%, -10% and +10% from the original data which total data become 2772 data.

	Number of neuron	Regression
With data enrichment	20	0.99
Real data only	20	0.77

Table 2: Comparisons on the effect of the data input

VI. Conclusion

The work presented in this paper shows that wheat yield can be accurately predicted using a Back Propagation Multi Layer Perceptron Artificial Neural Network; in this case using LM. The inputs applied to the ANN were weather data values; without needing to consider any other data input such as CERES (Crop Environment Resource Synthesis) – wheat model which used weather, soil condition, plant characteristics and crop management [7], AFRCWHEAT2 and SIRIUS model [23].

The architecture of our wheat yield prediction model is very simple. In terms of its accuracy it has been shown to be comparable with other wheat yield prediction models. Although this model is not as effective when the dataset is relatively small (less than 132 data) in which case its accuracy would be around 65%; with 924 data samples for wheat yield prediction. However, with larger quantities of data enriched ANNs, our wheat yield prediction model achieves accuracies of the order of 98%; with 2772 data samples.

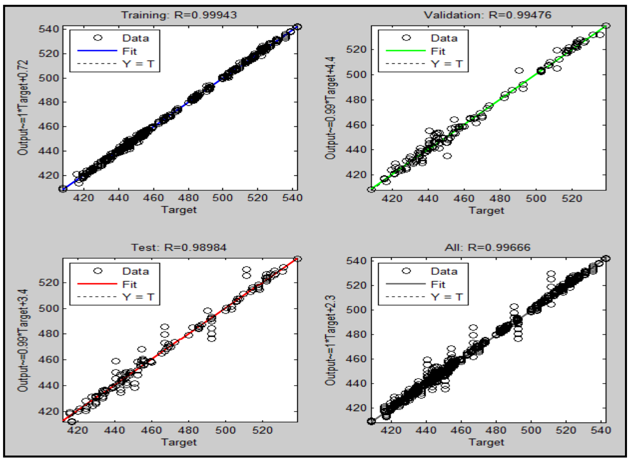

Figure 2: Simulation result for 20 hidden neuron with data enrichment

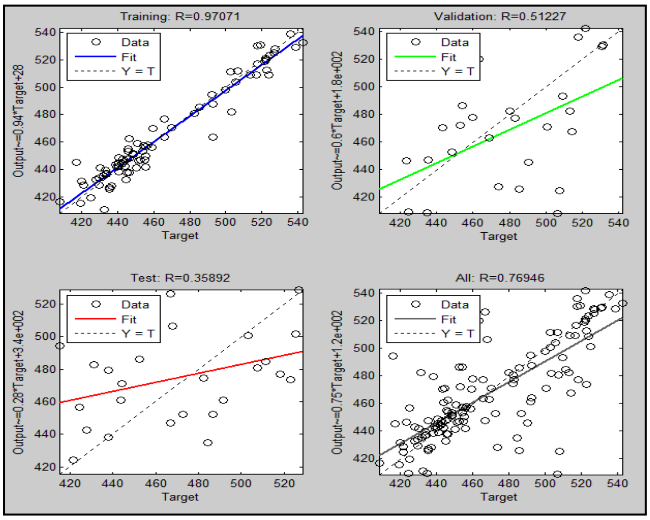

Figure 3: Simulation results for 20 hidden neuron with original data without data enrichment

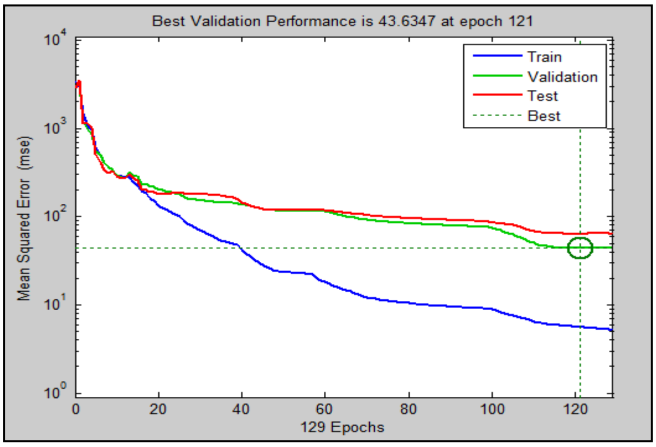

Figure 4: Performance on 20 hidden neuron with data enrichment

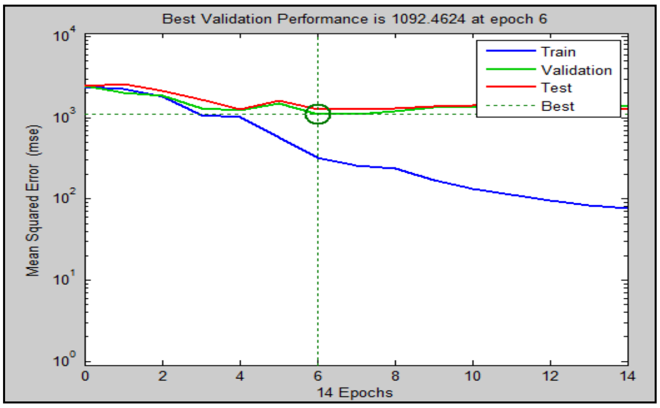

Figure 5: Performance on 20 hidden neuron without data enrichment

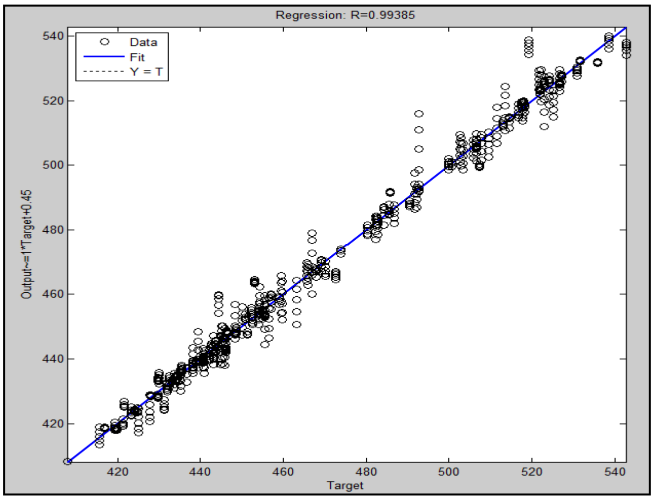

Figure 6: Comparisons on prediction result between the wheat yield original data (target – x-axis) and prediction result (output of the network –y-axis)

References

[1] X. Y. D. Jiang, N. Clinton, N Wang, "An artificial neural network model for estimating crop yields using remotely sensed information," *International Journal Remote Sensing,* vol. 25, pp. 1723 - 1732, 10 May 2004.

[2] A. K. B. Safa, M. Teshnehlab, A. Liaghat, "Artificial neural networks application to predict wheat yield using climatic data," presented at the International Conference, 2004.

[3] J. T. Ritchie, Otter, S., "Description and performance of CERES-Wheat: A user oriented wheat yield model," *Willis, W.O. (Ed), ARS Wheat Yield Project, Department of Agriculture, Agriculture Research Service,* vol. 38, pp. 159-175, 1985.

[4] R. F. Grant, "Mathematical modeling of root growth under different nitrogen and irrigation treatments: the ecosys approach," *Ecol. Modelling,* vol. 107, pp. 237 – 264, 1998.

[5] P. D. Jamieson, Semenov, M.A., Brooking, I.R., Francis, G.S., "Sirius: a mechanistic model of wheat response to environmental variation. ," *Eur. J. Agron.* , vol. 8, pp. 161-179, 1998.

[6] R. J. Brooks, *et al.*, "Simplifying Sirius: sensitivity analysis and development of a meta-model for wheat yield prediction," *European Journal of Agronomy,* vol. 14, pp. 43-60, 2001.

[7] N. M. J. C. a. G. H. M. Bannayana, "Application of the CERES-Wheat Model for Within-Season Prediction of Winter Wheat Yield in the United Kingdom," *Agronomy Journal,* vol. 95, pp. 114-125, 2003.

[8] T. F. B. S.A. Shearer, J.P Fulton, S.F. Higgins, "Yield Prediction Using A Neural Network Classifier Trained Using Soil Landscape Features and Soil Fertility Data.", presented at the Annual International Meeting, Midwest Express Center, Milwaukee, Wisconsin, 2000.

[9] F. a. R. A. Department for Environment, "UK Food Security Asessement: Detailed Analysis," F. a. R. A. Department for Environment, Ed., ed: DEFRA, 2010.

[10] R. L. Desjardins and C. E. Ouellet, "Determination of the importance of various phases of wheat growth on final yield," *Agricultural Meteorology,* vol. 22, pp. 129-136, 1980.

[11] F. a. R. A. Department for Environment. [Online]. Available: http://www.statistics.gov.uk/hub/

[12] J. W. Gardner and et al., "Application of artificial neural networks to an electronic olfactory system," *Measurement Science and Technology,* vol. 1, p. 446, 1990.

[13] J. W. Gardner, *et al.*, "Detection of vapours and odours from a multisensor array using pattern-recognition techniques Part 2. Artificial neural networks," *Sensors and Actuators B: Chemical,* vol. 9, pp. 9-15, 1992.

[14] L. Eduard and et al., "Non-destructive banana ripeness determination using a neural network-based electronic nose," *Measurement Science and Technology,* vol. 10, p. 538, 1999.

[15] M. Negnevitsky, *Artificial intelligence A guide to intelligent systems*: Addison-Wesly, 2005.

[16] M. B. H. Demuth, *Neural Network Toolbox: For use with Matlab*, 2004.

[17] S. a. B. Geman, E., "Neural networks and the bias / variance dilemma.," *Neural Computation,* vol. 4, pp. 1–58, 1992.

[18] A. Weigend, "On overfitting and the effective number of hidden units.", in *Proceedings of the 1993 Connectionist Models Summer School*, 1994, pp. 335-342.

[19] I. V. Tetko, Livingstone, D.J., and Luik, A.I., "Neural Network Studies. 1. Comparison of Overfitting and Overtraining. ," *J. Chem. Info. Comp. Sci.,* vol. 35, pp. 826-833, 1995.

[20] C.-T. S. J.-S.R. Jang, E. Mizutani, *Neuo-Fuzzy and Soft Computing*: Prentice Hall, 1997.

[21] B. G. Kermani, *et al.*, "Performance of the Levenberg-Marquardt neural network training method in electronic nose applications," *Sensors and Actuators B: Chemical,* vol. 110, pp. 13-22, 2005.

[22] M. I. A. Lourakis. (2005, February 11.). *A brief description of the Levenberg-Marquardt algorithm.* Available: http://www.ics.forth.gr/ lourakis/levmar

[23] S. Landau, *et al.*, "Testing winter wheat simulation models' predictions against observed UK grain yields," *Agricultural and Forest Meteorology,* vol. 89, pp. 85-99, 1998.