International Journal of Statistical Sciences *ISSN 1683-5603*
Vol. 20(2), 2020, pp 267-276

Prediction of Paddy Germination Stored in Different Technologies Using Artificial Neural Network

Mohammad Afzal Hossain*[1], M. A. Awal[2], M. R. Ali[2], Machbauddin[3], and M. M. Alam[2]

[1]Senior Scientific Officer, Bangladesh Rice Research Institute, Gazipur
[2]Professor, Dept. of Farm Power and Machinery,
[3]Asst. Professor, Dept. of Computer Science and Mathematics,
Bangladesh Agricultural University, Mymensingh
* Correspondence should be addressed to M. Afzal Hossain
(Email: engr.afzal@yahoo.com)

[Received November 01, 2019; Accepted January 25, 2020]

Abstract

Artificial neural network (ANN) offers machine learning and prediction in performing agricultural task smartly. It can identify subtle patterns in input training data which may be missed by conventional statistical analysis. A multilayer feed forward (with back propagation of error learning mechanism) was developed in MATLAB 2015 software to predict paddy quality stored in different technologies using ANN. The real time observations were used to train the ANN model obtained from storing paddy in different technologies. Collected data were separated into two sets training (75%) and test (25%). The ANN model was trained by backpropagation algorithm so that application of a set of input would produce the desired set of output. A model of hermetically stored paddy was developed for germination prediction of stored paddy. The developed model was 4-layered feed-forward neural network. Input layer of the model consists of 6-7 neurons, which corresponds to the input variables. The output layer has one neuron, which represents the predicted germination of the stored product. These showed a set of inputs, which self-adjust to produce consistent responses. Developed ANN models can predict germination (R^2=0.674) of stored paddy, a finding similar to that of experimental data.

Keywords: Prediction, germination, paddy, quality, stored, technologies and model

AMS Classification: 93B05; 93B52; 49J30.

1. Introduction

Artificial neural network (ANN) is designed to mimic the learning functions of the human brain to recognize patterns and predict. ANNs are formed from simulated neurons that are analogous to functions of the human brain for numerous reasons. In the brain, a neuron sends out an electrical signal through a strand known as an axon, which splits into many branches. At the end of each branch, there is an area called a synapse (Dahikar *et al*., 2015). The ANN structure is aparallel system based on human brain's biological neural process used to solve complex problems which it tries to imitate into mathematical models. The number of hidden nodes depend on the specific problem understudy and can easily extend to more hidden layers. The input contains nodes that correspond to input variables while the output contains nodes that correspond to output variables (Kaul and Hill, 2009). The ability to predict the future enables the farmer to make the most appropriate decision in anticipation of that future. (ANN) offers exciting possibilities to perform machine learning and prediction, and should be abundantly utilized in performing agriculture prediction task (Zhange, 1998), (Patuwo *et al*., 1993). ANN is a computational mechanism that is able to acquire, represent, and compute a weighting or mapping from one multivariate space of information to another, given a set of data to represent that mapping (Garret J H., 1994). It can identify subtle patterns in input training data which may be missed by conventional statistical analysis. In contrast to regression models, neural networks do not require knowledge of the functional relationships between the input and the output variables (Turban *et al*., 1998).

Many researchers have developed several predicting crop yield models in relation to different parameters as influencing factors by applications of ANN together with statistical techniques such as linear regression technique. ANN is increasingly being applied to process control and other areas, including the dynamic modeling of process operations, process prediction, optimizing, non-linear transformation, remote sensing technology and parameter estimation for the design of controllers (Yang *et al*., 2009).

ANN is inspired by how the human's central nervous system works, such as how the brain processes information. It is fundamentally a 'black box' approach. In this kind of technology the black box has the ability to learn the input-output correlation by training the input to produce the expected output (Ismael *et al*., 2011). Prediction is an important application of theANN model. ANN model makes prediction by smartly analyzing the trend from an already existing voluminous historical set of data. Many mathematical or statistical models are very accurate in calculation but not in prediction, because they cannot adapt to irregularly varying patterns of data which can neither be written as a function nor deduced from a formula. In ANN model the artificial neurons can learn from experience (that isby back-propagation of errors) to make the next guess and so on; hence, ANN is a better interpreter of real life situations (Kumar *et al*., 2012). ANN consists of neurons which have been related together with special arrangement. Neurons are in layers and every network consists of some 2 R. AMIRI-CHAYJAN and M. ESNA-ASHARI neurons in input layer, one or more neurons in output layer, and neurons in one or more hidden layers. The learning purpose in artificial neural networks is weights updating, so that when presenting a set of inputs, desired outputs are obtained. The most common type of artificial neural networks is feed forward back propagation (FFBP) (Jam and Fanelli, 2000).

Each network is trained with presented patterns. During this process, the connection weights between layers are changed until the differences between predicted values and the target (experimental) are reduced to reach permissible limit (Heristev, 1998).The number of iteration used by a machine learning model in the training dataset to reach on the optimum stage is called an epoch.In every epoch, the weight and other internal parameters are adjusted and an epoch consisted of a set of batches. Again, the gradient descent algorithm is used to evaluate the learning stage in each epoch. It is useful to set an epoch, since certain instruments (such as RMS error graph) update their calculations at the end of an epoch. FFBP network consists of one input layer, one or several hidden layers and one output layer. For learning this network, a back propagation (BP) learning algorithm is usually used. In case of BP algorithm, first output layer weights are updated. A desired value exists for each neuron of output layer. Based on this value and the learning rules, the weight coefficient is updated, which indicates that the model is properly trained. The ability to predict the future enables the farmer to take the most appropriate decision in anticipation of that.

The feed-forward neural networks are the most popular architectures due to their structural flexibility and good representational capabilities (Salehi *et al.,* 2011). Any ANN model contains an input layer, an output layer and one or more hidden layers. The number of neurons in the input and output layers are equal to the number of system inputs and outputs, respectively. ANN models find relationships by observing a large number of input and output examples to develop a formula that can be used for predictions (Snehal *et al.* 2015). A minimum of three layers is required in an ANN model: the input, hidden and output layers. Feed-forward networks may be based on linear or non-linear transfer functions that affect the output from the input and hidden layers. Non-linear networks may be trained using supervised learning, learning by example with outputs, or unsupervised learning, self-organizing without outputs. Supervised learning uses known outputs to train the ANN and is more commonly used than unsupervised learning. Back-propagation is a form of supervised learning where the error rate is sent back through the network to alter the weights to improve prediction and decrease error. The general process to build a neural network model including the creation of data sets for training and testing, training multiple networks with varied parameters, analyzing network results, and testing the models (Broner & Comstock, 1997).

Manual control of rice processing systems can decrease the ultimate quality of rice. Conversely, intelligent control of processing systems can prevent it. Some investigations on the moisture content changes of stored rice with air temperature at the duration of aeration have been carried out for different cultivars using models with three or more coefficients represented (Jindal and Siebenmorgen, 1994). (Hossain, 2019) showed stored paddy quality, measured by discoloration, insect infestation, germination capacity, etc., deteriorates in traditionally stored paddy. Researchers from many scientific disciplines are designing ANN to solve a variety of problems in pattern recognition, prediction, optimization, associative memory, and control. ANN might help the farmers ~~to~~ identify the suitable storage technology through prediction. Hence, the study was carried out to predict appropriate paddy storage technology for farmers of Bangladesh.

2. Methodology

ANN was used to develop a model to predict the quality of paddy stored in different technologies. A multilayer feed forward with back propagation of error learning mechanism was generated in MATLAB 2015 software to develop the model.

2.1. Structure of ANN model of hermetically stored paddy

Input data used to train the neural network model were obtained from the studies of hermetically stored paddy both at on-station and on-farm household condition. ANN model of hermetically stored paddy was developed creating ANN program for prediction of germination of stored paddy. The model was a 4-layered feed-forward neural network. As shown in Figure 1, the input layer of the first model consists of 6 neurons, which corresponds to the 6 input variables, and the output layer has one neuron, which represents the final moisture content of the dried product. A total of 61 training pairs were used to train the model, which were prepared from four independent experimental data. The input variables of the first model are as follows:

i. moisture content (MC)
ii. air temperature (AT)
iii. relative humidity (RH)
iv. insect infestation (II)
v. grain color (GC) and
vi. storage technologies (ST)

Germination percentage (GP) was used as the output variable. Six inputs together with one output are considered as one set of training pair. The selection of number of neurons for each hidden layer is optional. Larger number of neurons yields more accurate result but complication arises to attain proper training. Also it takes time to build and compute the model. The number of neurons for hidden layer-1 of the first model was nine and that of the hidden layer-2 was four which performed best in model.

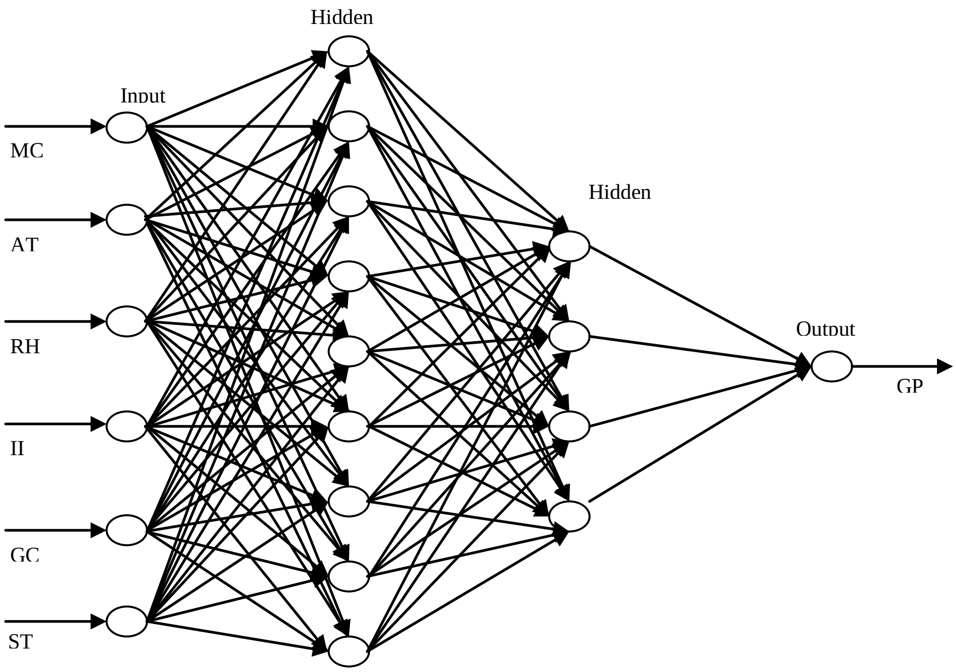

Figure 1: ANN structure for germination prediction of hermetically stored paddy

2.2. Data preparation

Collected data were separated randomly into training and test sets. 75% of the data were utilized for training the predictive model and 25% of the data were reserved for testing.

2.3. Build the prediction model

Prediction model was developed using artificial neural network.

2.4. Model validation

To compare the performance of prediction models,root mean square error (RMSE) and coefficient of determination (R^2)were evaluated.

3. Results and Discussion

3.1. Performance of developed model for predicting germination (%) of stored paddy

The developed model for predicting germination percentage of stored paddy was trained with experimental data. After 6 iterations of training, the squared sum of differences (errors) between the observed and predicted outputs reached a significantly low level. Then the training performance of the model was checked by simulation. All input variables of the training pairs were used in the simulation to make comparison between the tested and simulated outputs of the model. The experimental and thesimulated germination percentage (%) of stored paddy are shown in Figure 2. The highest germination (%) was counted in the stored paddy in PICS bag followed by Grain Pro bag and Plastic drum which were within the Bangladesh seed standard (80%). It is found that predicted and actual germination (%) of stored paddy in hermetic storage i.e. Grain Pro Bag, PICS Bag and Plastic Drum was very close. Since input parameters were maintained in these technologies predicted output was found very close to other. In case of traditional technologies i.e. *Dole*, *Motka* and Plastic Bag, these were lower (65, 71); (62, 66) and (69, 64), respectively which varied largely. Because of the nature of construction materials of these technologies (porous behavior) input parameters could not be controlled. The best validation performance of developed model was found at epoch 6 as shown in Figure 3. Predicted germination (%) of stored paddy was consistent with actual data (R^2=0.827).Variation between the predicted and observed germination percentage of hermetically stored paddy was almost same. A little difference was observed betweenpredicted and observed values in traditional storage technologies (*Dole*, *Motka* and Plastic Bag), which indicates variation of input data affects simulated values. It was largedue to wide variation of weather conditions. If training data differs from the actual data, the trained model may yield behaviordifferent fromthe actual system. However, in spite of these insignificant differences, the model can predict the variation of germination percentage as good as the actual system.

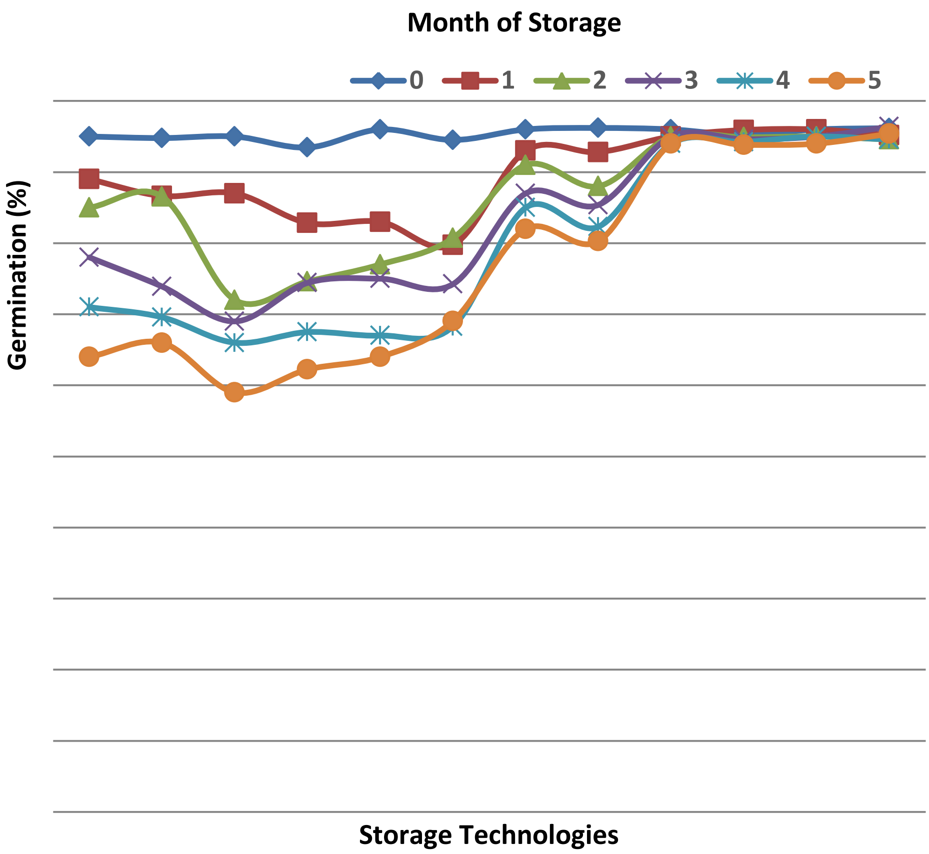

Figure 2: Experimental and simulated germination (%) of stored paddy

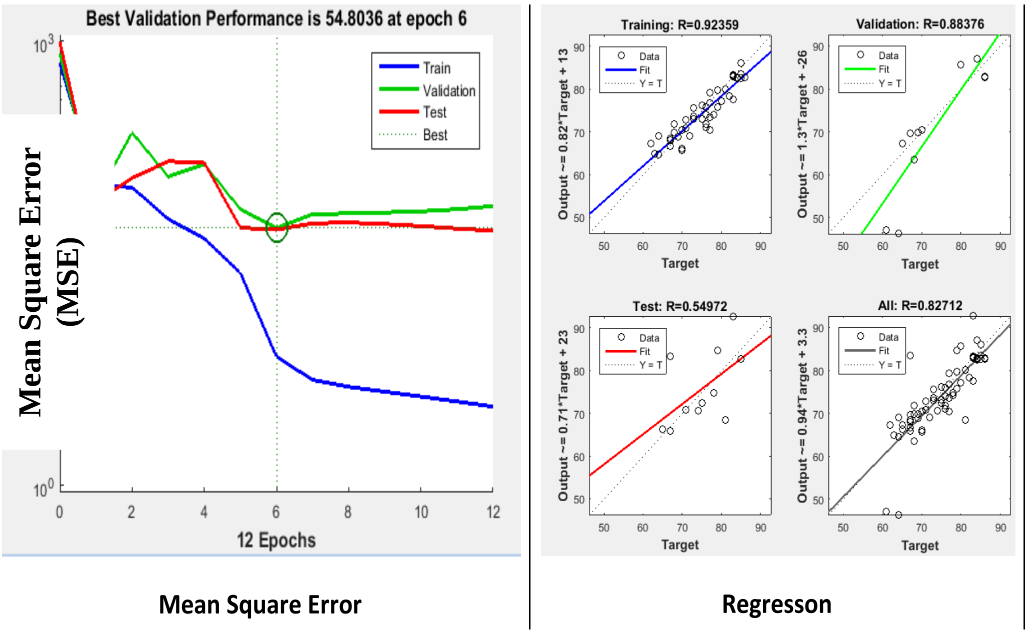

Figure 3: Regression in germination prediction of stored paddy

4. Conclusion

ANN is used inprocess control, dynamic modeling, prediction, optimizing, non-linear transformation, and remote sensing technology. A multilayer feed forward with back propagation of error learning mechanism was developed using ANN to predict stored paddy quality in different technologies. It trained the hidden layers by propagating the output error back through the network layer by layer, and adjusting weights at each layer to produce the desired set of output accurately. The developed ANN model of hermetically stored paddy was used to predict germination percentage prediction. The model was obtained as a 4-layered feed-forward neural network. Input layer of the models consists of neurons, which corresponds to the input variables, and the output layer has 1 neuron, which represents the predicted germination percentage of the stored product. It showed a set of inputs self-adjusted to produce consistent responses fruitfully. Developed ANN model can predict germination percentage (R^2=0.674) of stored paddy as good as actual experiments. Therefore, developed artificial neural network model is compatible in prediction of stored paddy quality with real time observations.

Acknowledgement: This research was performed as part of PHLIL project ‘‘USAID Post-Harvest Loss Reduction Innovation Lab-Bangladesh Component’’ funded by the USAID and ADM Institute for the Prevention of Post Harvest Loss, University of Illinois, USA (Grant Number: UReRA 2015-05296-01-00).

References

[1] Broner, I. and Cosmsock, C. R., (1997). Combining expert systems and neural networks for learning site-specific conditions. Computers and Electronics in Agriculture 19: 37–53.

[2] Dahikar, S. Rode, S. V. and Deshmukh, P., (2015). An artificial neural network approach for agricultural crop yield prediction based on various parameters. International Journal of Advanced Research in Electronics and Communication Engineering 4 (1): 94-98.

[3] Garret J. H. (1994). Where and why artificial neural networks are applicable in civil engineering. Journal of Computing in Civil Engineering, 8 (2): 129-130.

[4] Heristev, R. M. (1998). The ANN Book GNNU Published License available on http/://ftp.funet.fi/pub/sci/neural/books.

[5] Hossain M. A, (2019). Appropriate Paddy Storage Technology for Small Farmers in Bangladesh. PhD Thesis. Dept. of Farm Power and Machinery, Bangladesh Agricultural University, Mymensingh 2202.

[6] Ismael, M., Ibrahim, R., (2011). Adaptive neural network prediction model for energy consumption, IEEE.

[7] Jam L. and Fanelli A. M., (2000). Recent Advances in Artificial Neural Networks Design and Applications. CRC Press, Boca Raton, FL, USA.

[8] Jindal V. K. and Siebenmorgen T. J., (1994). Moisture transfer in blended long-grain rough rice. Trans. ASAE, 37, 195-201.

[9] Kaul, M., Hill, R. L. and Walthall, C. (2005). Artificial neural network for corn and soybean prediction, Agricultural System 85: 1-18.

[10] Kumar, A., Kumar, A., Rajeev, R., Sarthak, K., (2012). A rainfall prediction model using Artificial Neural Network, IEEE Control and system graduate research colloquium.

[11] Patuwo E., M. Y. Hu, and M. S. Hung, (1993). Two-group classification using neural networks, Decision Science, vol. 24, no. 4, pp.825–845.

[12] Salehi H., Zeinali Heris S., Koolivand Salooki M. K., and Noei S. H., (2011). Designing a neural network for closed ermosyphon with nanofluid using a genetic algorithm. Braz. J. Chem. Eng., 28, 157-168.

[13] Snehal S. Dahikar, Sandeep V. Rode, Pramod Deshmukh, (2015). An Artificial Neural Network Approach for Agricultural Crop Yield Prediction Based on Various Parameters. International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE) Volume 4, Issue 1, January.

[14] Turban E., (1998). Decision Support System and Expert Systems, New Jersey, USA: Prentice Hall, pp. 694.

[15] Yang L., Peng L., ZhangL., and Yang S., (2009). A prediction model for population occurrence of paddy stem borer (*Scirpophagaincertulas*), based on back propagation artificial neural network and principal components analysis. Comput. Electron. Agr., 68(2), 200-206.

[16] Zhange, (1998). Predicting with artificial neural network, International Journal of Predicting, vol. 14, pp.35-62.

[17] Zhang, N. R., (2010). Lecture 5: Multiple linear regression. Statistics 203, Stanford University.http://statweb.stanford.edu/~nzhang/203_web/lecture4.pdf.