Journal of Applied Sciences

ISSN 1812-5654

science alert

ANSInet
an open access publisher
http://ansinet.com

Journal of Applied Sciences 8 (2): 346-351, 2008
ISSN 1812-5654

Estimation of Paddy Equilibrium Moisture Sorption Using ANNs

[1]R. Amiri Chayjan and [2]Y. Moazez
[1]Department of Agricultural Machinery Engineering, Faculty of Agriculture,
Bu-Ali Sina University, Hamedan, Iran
[2]University of Applied Science and Technology, Agricultural Jahad Ministry, Miandoab, Iran

Abstract: In this research, Artificial Neural Networks (ANNs) used for prediction of Equilibrium Moisture Content (EMC) of three varieties of paddy (Sadri, Tarom and Khazar) as a new method. Feed forward back propagation and cascade forward back propagation networks with Levenberg-Marquardt and Bayesian regularization training algorithms used for training of input patterns. Optimized trained network has the ability of EMC prediction to test patterns at thermal boundary of 20-40°C and relative humidity boundary of 13.5-87% with $R^2 = 0.9929$ and mean absolute error 0.0229. Comparison between optimized ANN result and empirical model of Henderson showed that artificial neural network not only can simultaneously predict the EMC of samples of all varieties but also has better coefficient of determination and less mean absolute error.

Key words: Rice, equilibrium moisture content, Levenberg-Marquardt algorithm, Henderson model, back propagation network

INTRODUCTION

Equilibrium Moisture Content (EMC) is the function of water activity and environmental air temperature x = f (aw, T). This variable as durability criterion and quality changes of food and agricultural products at the duration of storing and packaging has the special importance (Veltchev and Menkov, 2000). Sorption isotherms represent the fundamental relationship between EMC and relative humidity of food products. Sorption characteristics of agricultural products use for designing, modeling and optimization of some process such as drying, aeration and storage (Bala, 1997; Palipane and Driscoll, 1992; Pahlevanzadeh and Yazdani, 2005).

Paddy is a hygroscopic material and at the different conditions of environmental weather gain or loss moisture depending on the grain-air interactions that take place. Aeration that relates the air relative humidity and grain moisture content is essential to optimize the grain quality. Rice is a very delicate cereal and because it is consumed as whole kernel and the formation of fissure and cracking provoked by ventilating with a high relative humidity air affects negatively on the final quality (Siebenmorgen *et al*., 1998). Rice has the second degree of main food of world people after wheat. Manual control of rice processing systems can decrease the ultimate quality of rice, so the intelligent control of processing systems can prevent it.

Some investigations related to the moisture content changes of stored rice with air temperature at the duration of aeration for different varieties have been done, so that models with three or more coefficients have been represented (Jindal and Siebenmorgen, 1994; Khankari *et al*., 1994).

Artificial Neural Networks (ANNs) were used for some industrial applications such as performance prediction of an industrial paper dryer (Huang and Mujumdar, 1993), predictive control of a drying process (Jay and Oliver, 1996), modeling the moisture content of thin layer corn drying during process for wet milling quality at constant air flow rate and absolute humidity and variable temperature (Trelea *et al*., 1997), air heater plant for a dryer (Thyagurajan *et al*., 1997), sorption isotherm of black tea (Pancharyia *et al*., 2002).

The precise prediction of EMC can not only decrease the storage losses of rice but also can more effect in processing systems. In this investigation, artificial neural network was used for prediction of water activity of three varieties of paddy (Sadri, Tarom and Khazar) for simulation of sorption isotherm and EMC at thermal boundary of 20-40°C and relative humidity of 13.5-87%, in the other word a three dimensional mapping was created for prediction of EMC, using variables of temperature, relative humidity and variety.

MATERIALS AND METHODS

Sorption isotherm model: The most common physical model for deriving EMC of agricultural products is Henderson and Kachru models. Zomorodian (2001) in a

Corresponding Author: Reza Amiri Chayjan, Department of Agricultural Machinery Engineering, Faculty of Agriculture, Bu-Ali Sina University, Hamedan, Iran Tel:+98 912 4186340 Fax:+98 811 4227012

research, predicted EMC of three Iranian rice varieties of Sadri, Tarom and Khazar, using empirical models. The best result derived using Henderson model because the highest value of R^2 and the lowest value of error derived by this model. This model is one of the best models, has the empirical basis and is as below (Henderson, 1952; Henderson and Perry, 1977):

$$1 - RH = e^{-CTMe^{n}} \quad (1)$$

Where:

RH = Environmental air relative humidity (decimal)
T = Environmental absolute temperature (K)
Me = Equilibrium moisture content
c and n = Constants for different materials that calculate by experimental method

Supremacy of Henderson model for prediction of EMC expressed by coefficient of determination (R^2) and Mean Absolute Error (MAE).

Artificial Neural Networks (ANNs): ANN is one of the artificial intelligence methods. It uses simple processing elements named neuron. ANNs tries to discover the inherent relationship between parameters through learning process. It creates a mapping between input space (input layer) and target space (output layer). Hidden layer/layers process the input data from input layer and produce answer in output layer. Training is a process that finally results in learning. Each network is trained with presented patterns. During this process, the connection weights between layers is changed until the differences between predicted values and the target (experimental) is reduce to be in permissible limit. Weights interpret the memory and knowledge of network. With the aforementioned conditions, learning process took place. Trained ANN can be used for prediction of outputs of new unknown patterns (Zhang *et al*., 2002).

The Advantage Using of ANN are; high computation rate, Learning ability through pattern presentation, Prediction of unknown pattern, Flexibility affront the noisy patterns. In this research, feed forward and cascade forward networks were used. In addition, several learning algorithms utilized.

Feed Forward Back Propagation (FFBP) consists of one input layer, one or several hidden layers and one output layer. Usually for learning of this network, back propagation learning algorithm (BP) is used. In the case of BP algorithm, first output layer weights updated. For each neuron of output layer, desired value exists. By this value and learning rules, weight coefficient updated. BP algorithm for further problems, present suit results but for some other problems gives an improper result. In some

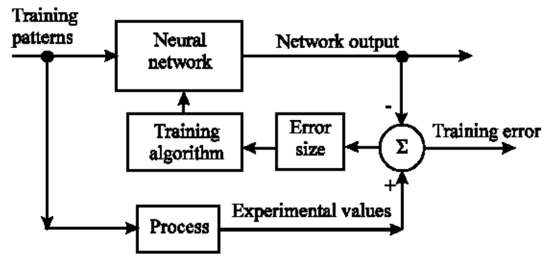

Fig. 1: Training process of back propagation networks

cases, because of local minimum, the learning process was upset. It is because of lying of answer at the smooth part of threshold function. Figure 1 shows the training process of BP algorithm for updating of weights and biases.

During training of this network, calculations were done from input of network toward output and then values of error propagated to prior layers. Output calculations were done layer to layer and output of each layer will be input of next layer.

Cascade Forward Back Propagation (CFBP) like FFBP network uses the BP algorithm for updating of weights, but the main symptom of this network is that each layer neurons related to all previous layer neurons.

Two training algorithms used for updating of network weights. These algorithms are Levenberg-Marquardt and Bayesian regulation back propagation algorithms.

Levenberg-Marquardt algorithm (LM) is a Hessian based algorithms and allows the network to learn features of a complicated mapping more suitable. Gradient-based training algorithms, such as back propagation, are most commonly used by researches. They are not efficient because the gradient vanishes at the solution. In a Hessian based algorithm the training process converges quickly as the solution is approached, because the Hessian does not vanish at the solution. To benefit from the advantages of Hessian based training, we use Levenberg-Marquardt algorithm. The LM algorithm is a non-linear least squares optimization.

Bayesian Regularization Algorithm (BR) initializes with random distribution of initial weights and biases. After presentation of input patterns to networks, updating of initial weight begins to obtain final distribution using algorithm. This procedure is robust to high noise level and has good approximation with arbitrary accuracy for training and it can improve generalization performance. In this algorithm, instead of the Sum of Squared Error (SSE) on the training set, a cost function, which is the SSE plus a penalty term, is automatically adjusted (Girosi *et al*., 1995).

Structural learning with forgetting is the main technique used for regularization (Girosi *et al*., 1995; Kozma *et al*., 1996). It has good approximation with

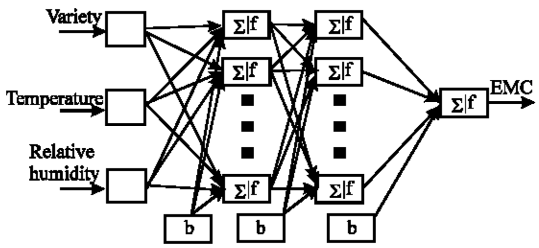

Fig. 2: Used ANNs topology

Table 1: Input parameters for ANNs and their boundaries

Parameters	Minimum	Maximum	No. of levels
Variety	-	-	3
Air temperature (°C)	20.0	40	3
Relative humidity (%)	13.5	87	6

arbitrary accuracy for training and it can improve generalization performance. Sensitivity analysis was performed order to decrease the chance of being trapped in a local minimum and to find stable results.

ANNs design and optimization methods: Considering and applying of three inputs in all experiments, the EMC value derived for different conditioning with these constraints, Networks with three neurons in input layer (variety, relative humidity and temperature) and one neuron in output layer (EMC) were designed. Figure 2 shows the considered neural network topology, Input and output parameters. Boundaries and levels of input parameters have been mentioned in Table 1. Neural network toolbox (ver. 4.1) of MATLAB software used for this research.

For obtaining of desired answer, FFBP and CFBP neural networks utilized. Training process by these networks is iterative. When the error between desired value and predicted value became minimum, training process towards to stability. The increasing method used for selection layers and neurons for evaluation of various topologies. By this method, when the network trapped into the local minimum, new neurons gradually add to the network. This method has practical potential for detecting the optimum size of network. advantages of this method are: network complexity increase gradually with increasing of neurons b) the optimum size of network always obtain by adjustments c) monitoring and evaluation of local minimum are done during the training process. Various threshold functions used to obtain the optimized status (Demuth and Beale, 2003):

$$Y_j = \frac{1}{1+\exp(-X_j)} \qquad \text{(LOGSIG)} \quad (2)$$

$$Y_j = \frac{2}{(1+\exp(-2X_j))-1} \qquad \text{(TANSIG)} \quad (3)$$

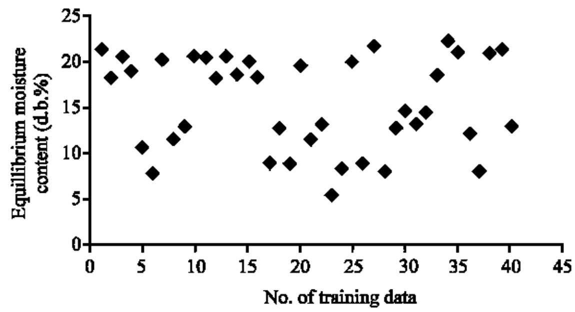

Fig. 3: Distribution of training data

Where:

X_j = Sum of weighted inputs for each neuron in jth layer and computed as below:

$$X_j = \sum_{i=1}^{m} W_{ij} \times Y_i + b_j \qquad (4)$$

Where:

m = Number of output layer neurons
W_{ij} = Weight of between ith and jth layers
Y_i = Ith neuron output and
b_j = Bias of jth neuron for FFBP and CFBP networks

About 75% of all data randomly selected for training of network with suitable topology and training algorithm (Fig. 3).

The following criterion of mean square error has defined to minimize the training error (Heristev, 1998)

$$MSE = \sum_{p=1}^{M}\sum_{i=1}^{N}(S_{ip} - T_{ip})^2 \qquad (5)$$

Where:

MSE = Mean Square Error
S_{ip} = Network output in ith neuron and pth pattern
T_{ip} = Target output at ith neuron and pth pattern
N = Number of output neurons
M = Number of training patterns

For optimizing the selected network from prior stage, the secondary criteria were used as fallow:

$$R^2 = 1 - \frac{\sum_{k=1}^{T}[S_k - T_k]}{\sum_{k=1}^{T}[S_k - T_m]}, \ (T_m = \frac{\sum_{k=1}^{T} S_k}{T}) \qquad (6)$$

$$MAE = \frac{1}{T}\sum_{k=1}^{T}\left|S_k - T_k\right| \qquad (7)$$

$$SD_{MEA} = \sqrt{\frac{\sum_{k=1}^{T}\left|S_k - T_k\right| - \left|\overline{S_k - T_k}\right|}{T-1}} \qquad (8)$$

Where:

R^2 = Coefficient of determination
MEA = Mean absolute error
SD_{MEA} = Standard deviation of mean absolute error

For increasing of accuracy and processing velocity of network, input data normalized at boundary of [0, 1].

RESULTS AND DISCUSSION

For solving of this problem, FFBP and CFBP neural networks have been used. Two strategy utilized for investigation of different threshold functions effects on network optimization that include of unique threshold function for all layers (Table 2) and various threshold functions for layers (Table 3). Both of two strategies used for FFBP and CFBP networks with learning algorithms of LM and BR. The best results of used networks and algorithms for first and second strategies showed in Table 2 and 3, respectively.

The best results of CFBP network with LM algorithm in first strategy with TANSIG threshold function is relate to 3-6-6-1 topology. This composition produce MSE = 0.0035, R^2 = 0.9842 and MAE = 0.0405 and converge in 28 epochs. In addition, the best result of this network with BR algorithm is related to first strategy with TANSIG threshold function and 3-6-5-1 topology. This composition has MSE = 0.0075, R^2 = 0.9855 and MAE = 0.0321 and converged in 21 epochs. Furthermore, in this stage, application of BR algorithm has better result than LM algorithm because it produces less MAE value and more R^2 value. It is obvious, in both strategies for CFBP network, TANSIG threshold function has the better performance in equal initial condition for producing of error and R^2 values.

The results for FFBP network in first strategy shows the LM algorithm has the better performance related to BR algorithm, because the best result for LM algorithm derived for 3-5-5-1 topology that has MAE = 0.0229 and R^2 = 0.9875 at 55 training epochs, but for BR algorithm, the best results is for 3-5-2-1 topology. It produce MAE = 0.0239, R^2 = 0.9840 at epoch of 22. With respect to results of first strategy, the best result is for FFBP network with LM algorithm and TANSIG threshold function and by 3-5-5-1 topology. Its R^2 showed in Fig. 4A and real error depicted in Fig. 5A.

The best result for second strategy and CFBP with LM algorithm is for 3-6-6-1 topology and LOGSIG-TANSIG-TANSIG threshold functions. This composition has MAE = 0.0324, R^2 = 0.9802, also for BR algorithm, the best topology is 3-5-4-1 with LOGSIG-TANSIG-LOGSIG threshold functions. It produce MAE = 0.0245, R^2 = 0.9874 in 11 epochs. Furthermore, for CFBP network, BR algorithm has the better result related to LM algorithm.

FFBP network for second strategy and LM algorithm for 3-3-2-1 topology and threshold functions of LOGSIG-TANSIG-LOGSIG with MAE = 0.0229 and R^2 = 0.9929 has the suit performance. In addition, the BR algorithm with FFBP network and 3-5-2-1 topology with LOGSIG-LOGSIG-TANSIG threshold functions, produce MAE = 0.0297 and R^2 = 0.9809.

With attention to results, for second strategy of FFBP network, LM algorithm has the best performance; also this result is better than first strategy, because the MAE and R^2 are the better values. The best R^2 showed in Fig. 4B and its MAE for testing patterns depicted in Fig. 5B. Results show that its MAE is less than that of Fig. 4A related to first strategies. This network selected as an optimized network. Results of optimized network is better than statistical results that presented by

Table 2: Training algorithm for different neurons and hidden layers for several networks at the uniform threshold function for layers

Network	Training algorithm	Threshold function	No. of layers and neurons	MSE	R^2	MAE	SD_{EMA}	Epoch
CFBP	LM	TANSIG	3-6-6-1	0.0035	0.9842	0.0405	0.0277	28
		LOGSIG	3-3-3-1	0.0069	0.9774	0.0362	0.0214	14
	BR	TANSIG	3-6-5-1	0.0075	0.9855	0.0321	0.0190	21
		LOGSIG	3-4-3-1	0.0046	0.9795	0.0314	0.0203	23
FFBP	LM	TANSIG	3-5-5-1	0.0011	0.9875	0.0229	0.0174	55
		LOGSIG	3-4-2-1	0.0011	0.9864	0.0259	0.0181	19
	BR	TANSIG	3-5-2-1	0.0067	0.9840	0.0239	0.0187	22
		LOGSIG	3-5-3-1	0.0085	0.9747	0.0318	0.0206	14

Table 3: Training algorithm for different neurons and hidden layers for several networks at the uniform threshold function for layers

Network	Training algorithm	Threshold function	No. of layers and neurons	MSE	R^2	MAE	SD_{EMA}	Epoch
CFBP	LM	TANSIG-TANSIG-LOGSIG	3-6-6-1	0.00093	0.9802	0.0324	0.0208	16
	BR	LOGSIG-TANSIG-LOGSIG	3-5-4-1	0.00069	0.9874	0.0245	0.0186	11
FFBP	LM	LOGSIG-TANSIG –LOGSIG	3-3-2-1	0.00049	0.9929	0.0229	0.0165	42
	BR	TANSIG-LOGSIG – LOGSIG	3-5-2-1	0.00075	0.9809	0.0297	0.0160	17

Table 4: MAE and mean R^2 values of Henderson model and optimized ANNs

	Parameter	
Method	R^2	MAE
Henderson Model	0.9923	0.3694
ANN	0.9929	0.0229

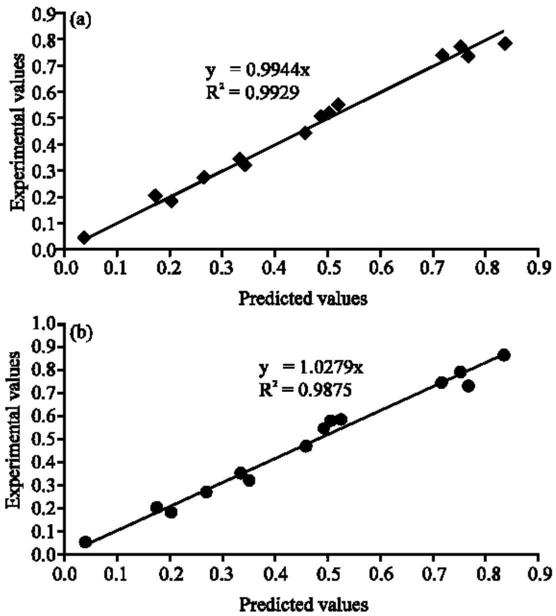

Fig. 4: Predicted values of EMC using ANNs versus experimental values (a) FFBP network, LM algorithm, TANSIG threshold function and 3-5-5-1 topology (b) FFBP network, LM algorithm, LOGSIG-TANSIG-LOGSIG threshold function and 3-3-2-1 topology

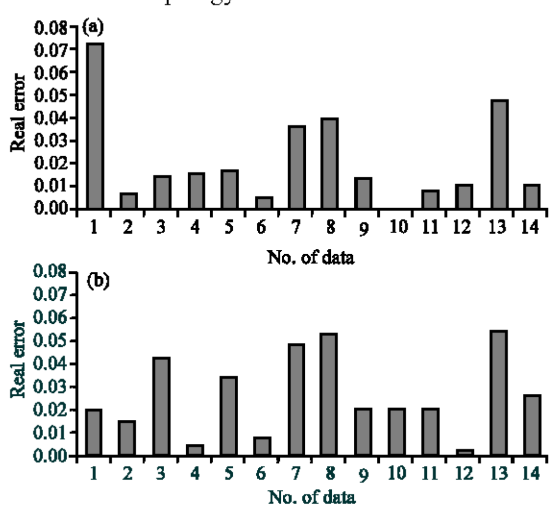

Fig. 5: Real error of paddy EMC predicting by optimized ANN (a) FFBP network, LM algorithm, TANSIG threshold function and 3-5-5-1 topology (b) FFBP network, LM algorithm, LOGSIG-TANSIG-LOGSIG threshold function and 3-3-2-1 topology

Zomorodian (2001). Henderson model and artificial neural network with optimized topology have not significant difference in R^2 value, but the absolute error of ANNs (0.0229) is less than that of Henderson model (Table 4). Also Henderson model presented 9 separately models (for three varieties and three temperatures levels, for prediction of EMC. The average value of 9 statistical models was 0.9923 but ANNs presented unique value for all varieties and temperatures with R^2 = 0.9929.

CONCLUSION

Artificial neural networks used as a new method for nonlinear mapping for EMC prediction of three varieties of paddy by two independent parameters of air temperature and relative humidity. The best ANN for data training was FFBP with LM algorithm and LOGSIG-TANSIG-LOGSIG threshold functions, four layers, three neurons for first hidden layer and two neurons for second hidden layer. Results showed that the EMC of three varieties of paddy could be predicted with less mean absolute error and more coefficient of determination compared to Henderson model.

REFERENCES

Bala, B.K., 1997. Drying and Storage of Cereal Grains. Oxford and IBH Publishing Co. Pvt. Ltd, New Delhi

Demuth, H. and M. Beale, 2003. Neural Network Toolbox for Matlab-Users Guide Version 4.1. The Mathworks Inc., Natrick, USA.

Girosi, F., M. Jones and T. Poggio, 1995. Regularization theory and neural network architectures. Neural Comp., 7: 219-269.

Henderson, S.M., 1952. A basic concept of equilibrium moisture. Agric. Eng., 331: 29-33.

Henderson, S.M. and R.L. Perry, 1977. Agricultural Process Engineering. AVI Publisher, New York.

Heristev, R.M., 1998. The ANN Book. GNU Public License. (ftp://ftp.funet.fi/pub/sci/ neural/books/).

Huang, B. and A.S. Mujumdar, 1993. Use of neural network to predict industrial dryer performance. Drying Tech., 11 (3): 525-541.

Jay, S. and T.N. Oliver, 1996. Modeling and control of drying process using neural networks. In: Proceeding of 10th International Drying Symposium (IDS 96). Krakow. Poland, 13: 1393-1400.

Jindal, V.K. and T.J. Siebenmorgen, 1994. Moisture transfer in blended long-grain rough rice. Trans. ASAE., 37 (1): 195-201.

Khankari, K.K., R.V. Morey and S.N. Patankar, 1994. Mathematical model for moisture diffusion in stored grain due to temperature gradients. Trans. ASAE., 37 (5): 1591-1604.

Kozma, R., M. Sakuma, Y. Yokoyama and M. Kitamura, 1996. on the accuracy of mapping back propagation with forgetting. Neurocomp, 13 (2-4): 295-311.

Pahlevanzadeh, G. and H.M. Yazdani, 2005. Moisture adsorption isotherms and isosteric energy for almond. J. Food Process Eng., 28 (4): 331-345.

Palipane, K.B. and R.H. Driscroll, 1992. Moisture sorption characteristics of in-shell macodamid nut. J. Food Eng., 18 (1): 63-76.

Pancharyia, P.C., D. Popovic and A.L. Sharma, 2002. Desorption isotherm modeling of black tea using artificial neural networks. Drying Tech., 20 (2): 351-362.

Siebenmorgen, T.J., A.A. Perdon, X. Che and A. Mauromoustakos, 1998. Relating rice milling changes during adsorption to individual kernel moisture content distribution. Cereal Chem., 75 (1): 129-136.

Thyagurajan, T., R.C. Panda, Y. Shanmugam and P.G. Rao, 1997. Development of ANN model for non-linear drying process. Drying Tech., 15 (10): 2527-2540.

Trelea, I.C., F. Countrios and E. Trystram, 1997. Dynamic models for drying and wet-milling quality degradation of corn using neural networks. Drying Tech., 15 (3-4): 1095-1102.

Veltchev, Z.N. and N.D. Menkov, 2000. Desorption isotherm of apples at several temperatures. Drying Tech., 18 (4-5): 1127-1137, 2000

Zhang, Q., S.X. Yang, G.S. Mittal and S. Yi, 2002. Prediction of performance indices and optimal parameters of rough rice drying using neural network. Biosyst. Eng., 83 (3): 281-290.

Zomorodian, A., 2001. Investigation on drying property of thin layer Iranian paddy varieties for determination of equilibrium moisture content (in Farsi). Agric. Eng. Res. J., 7 (2): 27-40.