Moisture Prediction in Maize using Three Term Back Propagation Neural Network

Said Jadid Abdulkadir, Siti Mariyam Shamsuddin, and Roselina Sallehuddin

Abstract— **The percentage of moisture content is one of the most important indexes in maize quality evaluation. Maize with high moisture content will not stay for extended periods; hence it is important to have accurate prediction of moisture content. As aflatoxin contamination in maize is of major concern, the necessity for suitable methods to predict moisture content with less time and higher accuracy assumes greater importance. Hence, Three Term Backpropagation network is proposed as a prediction tool for moisture content on maize. The new model is an improvement on Two Term Back propagation by the addition of an extra parameter, the proportional factor, which increases the convergence speed and reduces learning stalls in the conventional neural network. The experimental results are conducted using semi-annual datasets obtained from a maize thermal dryer. The results show that the proposed model outperforms Two Term Back propagation and other prediction methods like empirical correlation, analytical models (tank in series) and genetic algorithm which were used as prediction tools. Quantitatively, Three Term Back Propagation Neural Network obtained a higher precision result with a Mean Absolute Deviation (MAD) and Mean Absolute Percentage Error (MAPE) of 0.00145 and 0.00001 respectively.**

Index Terms — **Three term back propagation neural network, maize, moisture content, proportional factor.**

I. Introduction

Maize (Zea mays) is a cereal crop that is grown widely throughout the world in a range of agro ecological environments, rich in vitamins, carbohydrates, and essential minerals and contains 9% protein. All parts of the crop can be used for food and non-food products, but this study will concentrate on the food part which is the maize seed [1].

Moisture content is the level of water (moisture) in maize and is one of the most important factors in quality control of maize especially in controlling fungus infestation. Maize with high moisture content will not keep for extended periods in storage, so it is important, therefore, to have accurate determination of the moisture content and also for future storage planning after the harvesting season [2]. During this period aflatoxin contamination in maize is of major concern; the necessity for suitable methods to determine moisture content with less time and higher accuracy assumes greater importance. After shelling maize is stored in silos, the storage life of maize is greatly affected by the moisture content and temperature of the stored maize seeds. An increase in moisture results in respiration and enzymes production. This will lead to maize decay and greatly reduce the storage period of maize [3].

The level of moisture content is one of the most important marks in maize quality evaluation and overall in grain evaluation. The most commonly used method in measuring moisture content in maize is the direct method which involves the use of moisture meters before and after thermal drying. Not only is the measurement period time consuming but also there exists some defects of low measurement precision and instability due to the application of single sensor measurement [4]. Therefore the use of multi-sensor measurement and neural network methods of data processing, grain moisture measurement accuracy have been greatly improved.

With the purpose of improving the accuracy of moisture prediction, Artificial Neural Network (ANN) has been chosen as the basis of prediction in this study. Hence this study will investigate the use of Three Term Backpropagation (BP) Neural Network to predict the levels of moisture content of a particular agricultural product i.e. Maize, and by using datasets obtained from a multi-sensor moisture meter after thermal drying[5].

By the use of Three Term BP algorithm, it translates to stability and robustness in choice of initial weights, especially when relatively high values for the learning parameters are selected [6]. There are various methods for prediction of moisture levels for agricultural products. The simplest way is to use the available empirical correlations. However this approach generally fails for different agricultural products since such kind of correlations are not general since they are usually based on data obtained from a particular product [7]. That is, as the products changes, the equation alternates. In another study, [8] proposed to use Tanks-in-series (analytical model) to predict the average moisture content of solids. However, the solutions of these analytical models are very complicated and time consuming. Consequently, researchers have been using soft computing techniques among which artificial neural networks (ANN) and genetic algorithm (GA) have received much interest due to their ability to dynamic modeling of the drying characteristic of agricultural products [9]. However, they [9] pointed out that the ANN outperforms GA in predicting the experimental drying characteristics.

Standard Backpropagation (BP) algorithm has major limitations which are the existence of temporary, local minima resulting from the saturation behavior of the activation function and the slow rates of convergence

Manuscript received February 13, 2012; revised March 23, 2012. This work was supported in part by the Research Management Center (RMC), Universiti Teknologi Malaysia and Soft Computing Research Group (SCRG).

Said J. A., Siti Mariyam S. and Roselina S. are with the Soft Computing Research Group, Faculty of Computer Science and Information Systems, Universiti Teknologi Malaysia (e-mail: jadid86@gmail.com, mariyam@utm.my , roselina@utm.my).

especially for networks with more than one hidden layer [10]. Standard BP has been improved by adding an extra term, the proportional factor which is used to speed-up the weight adjusting process by increasing the convergence rate and decreasing learning stalls whilst maintaining the simplicity and efficiency of the standard two-term BP algorithm [11].

Adnan Topuz [9] proposed the prediction of moisture content of agricultural products using standard Back propagation neural network (BPNN). However, BPNN method was preferred due to its simplicity and reliability but the existence of drawbacks within the BP training algorithm hinders a more accurate prediction result [12]. Thus, there is need for a prediction model that will predict more accurate moisture content for agricultural products using Three Term BP without the existence of drawbacks within the BP training phase [13].

II. Background

A. Artificial Neural Network

Artificial Neural Network (ANN) is a model of thinking centered on the edifice of the human brain. It's made up of millions of neurons which act as processors which are interconnected by weighted links which are updated to obtain preferred outputs [14]. It uses a mathematical model for information processing which is based on the approach of computation inspired by the structure and operation of biological neurons organized into layers. Namely, there are three layers in a neural network; Input layer, Hidden layer and Output layer as shown in Fig. 1, where, and represents two input signals; drying time and drying temperature, represents weight from input to hidden node, represents weight from hidden unit to output node and represents output signal from the output node i.e. moisture content.

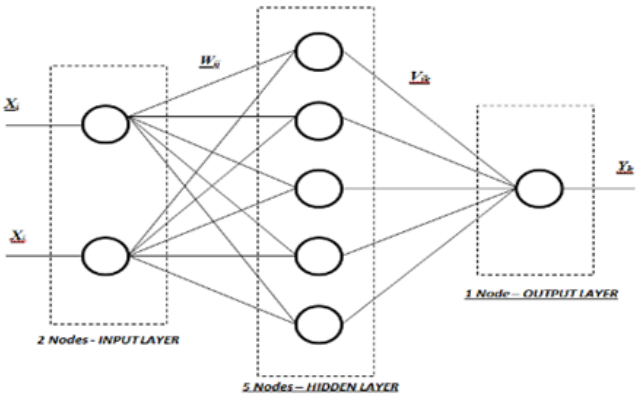

Fig.1. ANN model

There are different classifications of Artificial Neural Network (ANN), they can be classified as either feed forward, modular, recurrent, stochastic etc, the type of ANN is dependent on how data will be processed in that network. The simplest is the feed forward neural network, in this type of network the movement of information is only in a single forward direction. Data moves from the input to the hidden and lastly to the output nodes. The nodes in this layer are connected to other nodes in the next layer thus having a connection between every node in a given layer to the other next layer.

The advantage of Feed forward networks is that it can be trained to perform tasks on unfamiliar data, like classification tasks or identification tasks. ANN can also be classified by the learning process i.e. some use supervised while others use unsupervised learning. In Supervised learning both input and output values are provided and in unsupervised learning the network is not provided with desired outputs but only inputs.

Multilayer Perception (MLP) is the most common feed forward neural network. In order for the network to learn it requires desired output, therefore it is a supervised network. Correct mapping of inputs to the output is the main goal so that the model created can produce an actual output when the desired output was actually known.

There are many algorithms to fine-tune the weights, and Back propagation (BP) algorithm is widely used by the practitioners.

B. Back Propagation

Back propagation (BP) algorithm is a supervised learning method used for training artificial neuron networks (ANN). Training is usually carried out by iterative updating of weights based on the error signal. Then the error signal is back propagated to the lower layers. Back propagation is a descent algorithm which attempts to minimize the error rate at each iteration. By the use of 3-Term Back propagation algorithm it translates to stability and robustness in choice of initial weights, especially when relatively high values for the learning parameters are selected [15]. Learning is a fundamental and essential characteristic of ANN. It is capable of learning through the network experiences to improve their performance. When ANN is exposed to a sufficient number of samples, it can generalize well to other data that they have not yet encountered [16].

The input data is trained by the ANN and the results of the training process is an appropriate output according to the desired target. At first all weights are initialized to produce small random numbers. This prevents the network from being saturated by very large weight values. Back propagation (BP) is a gradient descent technique used to minimize the error for a particular training pattern; it is the most famous training algorithm for multi-layer perceptions [17].

From the training datasets a training input pattern is presented to the input layer, and then it is forward propagated from layer to layer until output pattern is determined from the output layer.

$$Output = \int(net_i), \tag{1}$$

$$net_i = \sum W_{ij} O_j + \vartheta_j, \tag{2}$$

where, W_{ij} is the weight connected between node i and j . θ_i is the bias of node i and O_j is the output of node j .

C. Two Term Back Propagation Parameters

The two learning parameters that involved in standard

Two Term BP are learning rate and Momentum factor, these parameters affect the convergence of the network and if the correct values are selected they separate the noise and avoid over-fitting of the signal [18].

The size of the weight adjustment made at every iteration is determined by the learning rate, from which it later affects the convergence rate. Studies have indicated that if small values are selected for the learning rate this finally leads to smaller weight changes and if large values are selected for the learning rate, the weight changes would be large. Thus the best value for the learning rate depends on the problem [19].

The value for the learning rate if is very small the gradient descent will progress in very small steps thus increasing the total convergence time. And if the value selected is very large the search path would oscillate about the ideal path [20].

If the learning rate selected is 0, then the network will not learn [21]. The normal range of learning rate values are between 0.25 and 0.35 [22]. The selection procedure in this study is through trial and error in this study until the optimum value is selected with the smallest error rate.

To prevent the network getting stuck in the shallow local minimum, the momentum factor is used [22]. Not only does it prevent the network getting stuck in shallow minimum but also it improves the rate of convergence [23].

The momentum rate allows for the movement in the same direction on successive steps, therefore it is also suppresses any oscillations that result from the changes in the slope of the error surface [23].

D. Three Term Back Propagation

Modification done by researchers to the two term BP algorithm was meant to improve efficiency and convergence rate of the algorithm. Zweiri [3] proposed a new term called the proportional factor (PF) which speeds up the weight adjustment process. This new proposal came to be known as Three Term Back propagation. This new approach showed that it out performs the standard BP algorithm in terms of both convergence rate and also escapes the local shallow minimum [24].

According to Zweiri [25], comparative test results were obtained from an XOR experiment which indicated that the inclusion of the proportional term on the standard back propagation algorithm has several positive effects such as increasing the convergence speed, robust to choice of initial weights and escaping from local minima, however the inclusion of the proportional factor does not have any effect on the complexity of the standard back propagation.

The BP algorithm is modified by adding an extra term called proportional term. It is as shown below:

$$\Delta W(k) = \alpha(-\Delta E(W(k))) + \beta \Delta W(k-1) + \gamma e(W(k)), \quad (3)$$

where α is learning rate, $\Delta E(W(k))$ is gradient of E at $W = W(k)$, with $k = 1,2,3,..,N$, being the iteration number, β is momentum term, $\Delta W(k-1)$ is a previous weight change, γ is proportional term and $e(W(k))$ is the difference between the output and the target at each iteration [26].

Note that BP algorithm given by Eq. (3) above has three terms. α is proportional to the derivative of $E(W(k))$, while β is proportional to the previous value of the incremental change of the weights and γ is the Proportional Factor that is proportional to $e(W(k))$[27]. The proportional term increases the learning speed of back propagation algorithm. This term is proportional to $e(W(k))$. This represents the difference between the target and output result at each iteration.

E. Cost Function

However, in our study, we use modified cost function as proposed by Shamsuddin [28] to improve speed of convergence of Three Term BP. This cost function will be used to calculate the error signal to measure the performance of the Three Term network. The selection of a good cost function helps in improving accuracy by iteratively updating weights thus minimizing error and giving a more accurate output. Also getting stuck into the local minima is a common problem and whereby researchers have stated that a good selection of cost functions can overcome this problem [28]. This has led to researchers coming up with cost functions so as to avoid the local minima. Hence, good cost function is important to ensure stable learning of the network [29]. The modified cost function (mm) is defined implicitly below as in [29]:

$$mm = \sum k \rho^{k}, \quad (4)$$

with,

$$\rho^{k} = E_k^2 \Big/ 2\, a_k (1 - a_k^2) \quad (5)$$

where,

$$E_k = t_k - a_k \quad (6)$$

And E_k error at output $unit_k$, t_k target value of output $unit_k$ and a_k an activation of $unit_k$.

III. Experimental Results

Maize datasets are normalized before the training process. After the process of data normalization, the data is divided into 10% for testing and 90% for training. Thus the number of instances for training was 90 and 39 for testing; this kind of selection is mainly done to avoid over-fitting problems. The network structure will consist of 2 input nodes, 5 hidden nodes and 1 output node.

A. Two Term Back Propagation using mm Cost Function

Using Two Term Backpropagation as a benchmark model, the selection of values for Learning and Momentum factor is done through trial and error (80 trials) and these values are increased in each case and the number of epochs. Hence the best parameters for learning and momentum rate are obtained from the test with the best error rate and lowest convergence speed.

The minimum number of iterations is 10 and the maximum

number is 1000 iterations. Two tests were performed, in Test 1, the minimum error obtained is 0.001662 at the 900 iteration with same values for learning and momentum rate of 0.9 and convergence time of 340ms. However, in Test 2 the minimum error obtained is 0.0016624 at the 1000 iteration with learning rate value of 1.0 and momentum rate of 0.1 and convergence time of 420ms. Thus the first test result becomes the best fit error value obtained using mm as cost function as shown in Fig.2

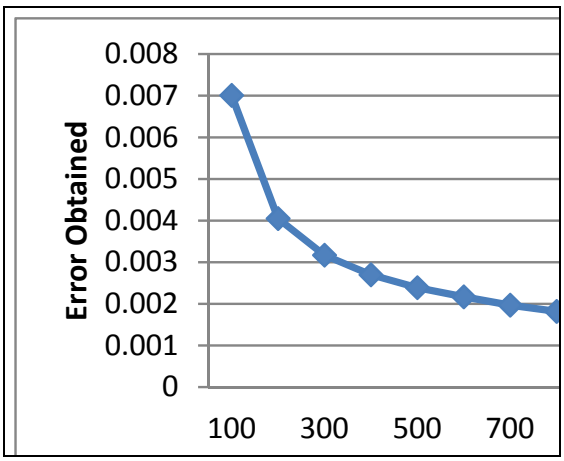

Fig. 2. Error obtained using two term back propagation

B. Three Term Back Propagation using mm Cost Function

In Three Term Backpropagation, an additional term i.e. proportional factor is required; hence the values for each parameter (momentum, learning and proportional factor) were identified after 80 trials with different combination of proportional, momentum and learning rate. The stopping criterion is the number of epochs which are executed between 10 to 1000 epochs in each of the experiments and an error threshold value of 0.005 [29]. The selection for initial weights is done randomly between weight values of -1 and 1 and the minimum and maximum number of iterations is 10 and 1000 iterations respectively. The predicted moisture content obtained from Three Term Backpropagation is shown in Table I.

TABLE I: ANALYSIS FOR THREE TERM BP

Samples	1	2	3	4	5
Actual	12.78	11.25	10.9	11.4	13.1
Predicted	12.77	11.24	10.89	11.39	13.09

Four tests were done to determine the best value for learning rate, momentum rate and proportional factor i.e. for Test 1, the values for learning, momentum rate and proportional factor were equal but only an increase in the number of epochs from 10 to 1000. In Test 2 the value for learning rate increases and momentum rate and proportional factor decreases as the number of epoch's increases. In Test 3, learning rate and proportional factor increases and momentum rate decreases as the number of epochs increases. After obtaining the best fit value with the smallest error rate, Test 4 is performed with a constant value for each of the three values but only an increase in the number of epochs. In each test, parameters i.e. error obtained, convergence time and number of epochs are used for comparative purposes.

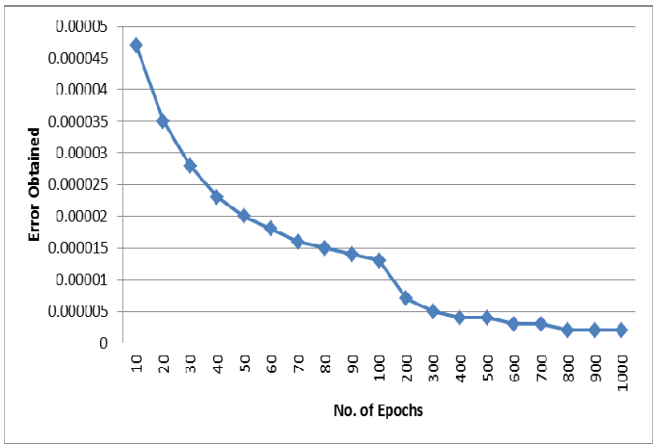

Fig. 3. Error obtained using Three term back propagation

C. Quantitative Analysis

To evaluate the performance of the model, three statistical tests are employed as given in Equations (7-9), namely;

Root Mean Square:

$$RMSE = \sqrt{\sum_{i=1}^{n} (de_i - o_i)^2}, \quad (7)$$

Mean Absolute Percentage Error:

$$MAPE = \sum_{i=1}^{n} \left[\frac{de_i - o_i}{de_i} \right] \times \frac{100}{n}, \quad (8)$$

Mean Absolute Deviation:

$$MAD = \sum_{i=1}^{n} \left(\frac{de_i - o_i}{n} \right), \quad (9)$$

where, n is the number of periods, de_i is actual values and o_i are forecasted values. The measurements are calculated using testing data. The error rate is assumed to be better if the value is very small or the value is near to zero. Statistical results tests which were obtained are given in Table II.

TABLE II: QUANTITATIVE ANALYSIS

	Two Term BP	Three Term BP
RMSE	0.03162	0.00118
MADE	0.01037	0.00145
MAPE	1.03692	0.00001

In overall performance, 3-Term Backpropagation outperforms 2-Term Backpropagation in all of the three commonly used quantitative error measurements as the best statistical value obtained is the one closer to zero or the smallest value.

Fig 2 and 3 shows the comparison between two and three term back propagation neural network when applied to a sample of moisture content datasets for the month of January, 2010.

TABLE III: ANALYSIS FOR TWO TERM BP

Samples	1	2	3	4	5
Actual	12.78	11.25	10.9	11.4	13.1
Predicted	11.38	9.85	9.5	10.0	11.7

Experimental results have shown that 3-Term Backpropagation using mm cost functions has produced better prediction results in terms of error rate. Quantitative analysis using 3-Term BP results outperformed the analytical result obtained by Adnan (2009) using 2-Term BP on agricultural products like carrots and cashewnuts. This can be seen in Table IV below:

TABLE IV: ANALYSIS WITH PREVIOUS WORKS

Error	Two Term BP	Three Term BP
MAD	0.044	0.00145
MAPE	0.112	0.00001

IV. CONCLUSION

The main purpose of the present study is to investigate the applicability of Three Term Back propagation neural network as a prediction tool for moisture content in maize. The experimental results shows that the newly proposed model outperforms Two Term Backpropagation in all of the three commonly used quantitative error measurements and in terms of both convergence rate and also escapes the local shallow minimum. Hence, it can be concluded that the Three Term Backpropagation model without the existence of drawbacks within the Backpropagation training phase [29] can be effectively utilized as a prediction tool for moisture prediction in maize as shown in Fig. 4.

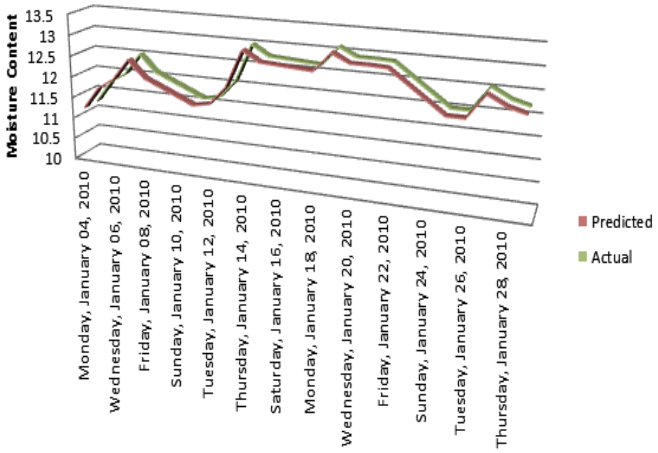

Fig. 4. Error obtained using three term back propagation

Also the proposed model can be effectively utilized as a prediction tool for determining moisture content in maize within the agricultural sector. This tool can be used as a planning tool for moisture prediction by farmers to prevent storage contamination due to aflatoxin, as a result of high moisture levels from long periods of storage during the harvesting period. Also it can be used by food production companies to schedule a plan for an increase or decrease of thermal drying for an annual period due to the forecasted moisture content from the previous harvesting season.

ACKNOWLEDGMENT

Authors would like to thank Research Management Center (RMC), Universiti Teknologi Malaysia and Soft Computing Research Group (SCRG) for the research activities and supports in making this study a success.

REFERENCES

[1] Copeland and McDonald, *Chapman & Hall New York, N.Y* 1995, pp. xvi, 409.
[2] X. L. Zhang and M. J. Gao, "The design of the new neural network model reference adaptive control systems," *System Engineering Transaction*, vol. 18, no. 2, pp.148-152, 2003.
[3] Y. H. Zweiri, J. F. Whidborne, K. Althoefer, and L. D. Seneviratne, "A new Three Term Backpropagation Algorithm with Convergence Analysis," *IEEE*, pp. 3882-3887.
[4] S. C. Ng, C. C. Cheung, S. H. Leung, and A. Luk, "Fast Convergence for Back-Propagation Network with Magnified Gradient Function," *IEEE*, pp. 1903-1908, 2003.
[5] K. Movagharneiad and M. Nikzad, "Modelling of tomato drying using artificial neural network," *Comput Electron Agr*, vol. 59, pp. 78-85, 2007.
[6] K. L. Babu and Y. P. Setty, *J Energy Heat Mass Transfer* 2001; 23:135.
[7] S. Erenturk and K. Erenturk, "Comparison of genetic algorithm and neural network approaches for the drying process of carrot," *J Food Eng* vol. 78, pp. 905-12, 2007.
[8] Y. H. Zweiri, J. F. Whidborne, K. Althoefer, and L. D. Seneviratne, "Stability Analysis of a Three-Term Backpropagation Algorithm," *Neural Networks* vol. 18, pp. 1341–1347, 2005.
[9] A. Topuz, "Predicting moisture content of agricultural products using artificial neural networks," *Advances in Engineering Software* vol. 41 pp. 464-470, 2010.
[10] S. M. Shamsuddin, M. N. Sulaiman, and M. Darus, "An Improved Error Signal For Bacpropagation Model For Classification Problems," *Intern. J.Computer Mathematics*. vol. 76(1-2), pp. 297-305.
[11] M. Rimer and T. Martinez, "CB3: An Adaptive Error Function for Backpropagation Training. Neural Processing Letters," vol. 24, pp. 81–92, 2006.
[12] S. C. Ng, C. C. Cheung, S. H. Leung, and A. Luk, "Fast Convergence for Back-Propagation Network with Magnified Gradient Function," IEEE. 1903-1908J.
[13] U. Duncombe, "Infrared navigation—Part I: An assessment of feasibility," *IEEE Trans. Electron Devices*, vol. ED-11, pp. 34-39, Jan. 1959.
[14] B. Guijarro-Berdinas, O. Fontenla-Romero, B. Perez-Sanchez, and P. Fraguela, "A Linear Learning Method for Multilayer Perceptrons Using Least-Squares," Lecture Notes in Computer Science 4881. Berlin Heidelberg: Springer, pp. 365–374.
[15] B. Guijarro-Berdi, O. Fontenla-Romero, B. Perez-Sanchez, and P. Fraguela, "A Linear Learning Method for Multilayer Perceptrons Using Least-Squares," Lecture Notes in Computer Science, pp. 365-374.
[16] H.-M. Lee, C.-M. Chen, T.-C. Huang, "Learning science improvement of back-propagation algorithm by error saturation prevention method," *Neurocomputing*, vol. 41, pp. 25-143, 2001.
[17] B. K. Humpert, "Improving Back Propagation With A New Error Function," *Neural Networks* vol. 7, no. 8, pp. 191-1192, 1997.
[18] M. E. Ibrahim and A. A. M. Al-Shams, "Transient stability assessment using artificial neural networks," *Electric Power Systems Research*, vol. 40, pp. 7-I6, 1997.
[19] I.-C. Kim, S.-I. Chien, "Speed-up of error ackpropagation algorithm with class-selective relevance," *Neurocomputing* vol. 48, pp. 1009–1014, 2002.
[20] M. Jiang, B. Deng, B. Wang, and B. Zhong, "A Fast Learning Algorithm Of Neural Networks By Changing Error Functions," *IEEE International Conference Neural Networks Signal Processing*. December 14-17, 2003, Nanjing. China, pp. 249-252
[21] N. Kandil, Khorasani, R.V. Patel, and V. K. Seed, "Optimum Learning Rate For backpropagation Neural Network," *IEEE*, pp. 465-468, 1993.
[22] T. Kathirvalavakumar and P. Thangavel, "A Modified

Backpropagation Training Algorithm for Feedforward Neural Networks," *Neural Processing Letters* vol. 23, pp. 111-119, 2006.

[23] E. Keogh. The UCR Time Series Data Mining Archive. Riverside CA. University of California - Computer Science & Engineering Department. [Online]. Available: http://www.cs.ucr.edu/~eamonn/TSDMA/index.html

[24] C-S. Liu and C-H. Tseng, "Quadratic optimization method for multilayer neural networks with local error-backpropagation," *International Journal of Systems Science* vol. 30, no. 8, pp.889 – 898, 1999.

[25] J. Lv and Z. Yi, *An Improved Backpropagation Algorithm Using Absolute Error Function*, ISNN 2005, LNCS 3496, pp. 585–590, 2005.

[26] M. Mandischer, "A comparison of evolution strategies and Backpropagation For neural network training," *Neurocomputing*, vol. 42, pp. 87–117, 2002.

[27] M. Kiyotoshi and J. Yi, "Backpropagation Based on the Logarithmic Error function and Elimination of Local Minima," *IEEE*, pp. 1117-1122, 2000.

[28] K. Movagharneiad and M. Nikzad, "Modelling of tomato drying using artificial neural network," *Comput Electron Agr* vol. 59, pp. 78-85, 2007.

[29] S. C. Ng, S. H. Leung, and A. Luk, "Fast Convergent Generalized Back-Propagation Algorithm with Constant Learning Rate," *Neural Processing Letters* vol. 9, pp. 13-23, 1999.