Modeling and Simulation of Apple Drying, Using Artificial Neural Network and Neuro -Taguchi’s Method

M. Mousavi1,*, and S. Javan1

ABSTRACT

Important parameters on apple drying process are investigated experimentally and modeled employing artificial neural network and neuro-Taguchi’s method. Experimental results show that the apple drying curve stands in the falling rate period of drying. Temperature is the most important parameter that has a more pronounced effect on drying rate than the other two parameters i.e. air velocity and the thickness of apple slices. In order to model the drying process, a software was developed which uses the error back propagation algorithm for training. At first, the software was used to simulate the time-dependent variations of moisture content using neural network. Then in order to model the time derivation of moisture ratio in break point, the software was utilized in two ways. First, it was used with no use of any optimization method for modeling the process. In the other approach, the software in a hybrid fashion with Taguchi’s method as an optimization method is utilized to correct weight matrix entries. The results demonstrate that the use of neuro-Taguchi’s method can give some improvements over neural network accuracy as compared with conventional neural networks approach. By using neuro-Taguchi’s method, error is reduced by about 46.4%.

Keywords: Apple drying, Artificial neural network, Modeling, Neuro-Taguchi, Simulation.

INTRODUCTION

Decreasing the moisture content of food materials is one of the oldest and most important ways of food preservation. With decrease in moisture content, the microbiological and enzymatic decay is retarded and rate of harmful reactions and nonenzymatic color change into brown is as well decreased to a large extent. In addition, a decrease in weight and volume of product amounts to reduction in transport and storage costs.

Drying of food materials leads to new more easily handled and consumed products. Apple is among the most important fruits in dried fruit production industry.

Artificial neural networks (ANNs) are optimization algorithms in which it is attempted to mathematically model the learning process. The model is a simple approximation of such a complex process, but it utilizes the basic foundations and concepts inherent in the learning processes of humans and animals. ANNs are universal function approximators that typically work much better than the more traditional (polynomial) function approximation methods.

The first artificial neural network model was presented by McCulloch and Pitts (1943) and later extended by others (Rumelhart and McClelland, 1986). Indeed, ANNs have attracted a lot of interest in the past decade and in certain processes. Artificial neural networks are emerging as a promising tool in process identification and control owing to their capacity to accurately model processes (Hussain and Rahman, 2009).
Artificial neural networks can be employed to simulate the non-linear input/output dynamics of a process based on a time history of process data (Venkatasubramanian and McAvoy, 1992; Menhaj, 1998; Razavi et al., 2003).

A number of ANNs applications have been reported in the area of prediction of food properties. Examples are the prediction of dough rheological properties (Ruan et al., 1995), fruit ripening (Morimoto et al., 1997), inactivation kinetics (Geeraerd et al., 1998), viscosity (Bouchard and Grandjean, 1995), smell by electronic nose (Payne, 1998), food quality (Ni and Gunasekaran, 1998), loaf volume during baking (Horimoto et al., 1995), sensory prediction (Bomio, 1998), classification of nut (Casasent et al., 1998), thermal conductivity (Hussain and Rahman, 1999; Sablani and Rahman, 2003), periodic thermal treatments for surface decontamination of strawberries (Scheerlinck et al., 2004), and cross flow milk ultrafiltration (Razavi et al., 2004).

Neural network modeling is essentially black box in nature. The capability of neural network to learn non-parametric or structure-free approximations is its strength, but this is also its weakness. In order to promote neural networks, many different ways have been employed. Many new learning algorithms have been designed and developed to train neural networks. The certain ways focus mainly on the size of the neural network, namely the number of hidden layers and its corresponding number of neurons. For example, Maza (1991) has developed a heuristic algorithm, called the SPLIT net, for the dynamic adjustment of the number of hidden neurons in a neural network. In his approach, neurons of a neural network are regarded as feature detectors. They can take three states, i.e. yes, no and maybe. Each state corresponds to the output value of a hidden neuron when a particular pattern of inputs is presented. Each hidden neuron carries an evaluation function that calculates the relative frequency of the occurrence of a particular state. A hidden neuron will be split into two neurons with identical connection weights when the neuron’s evaluation function exceeds a predetermined threshold value.

Wang and Hsu (1991) have introduced a new algorithm called the “self Grow Learning Algorithm” for determining the appropriate number of hidden neurons. This method consists of a set of rules for either adding or deleting a hidden neuron. This algorithm allows the neural network to learn and to reach the global minimum in finite time.

Conventional artificial neural networks (multi-layer perceptrons) do not attempt to precisely model the vagueness or fuzziness of data. This often culminates in poorly trained networks where the problem becomes more significant as the uncertainty in the data increases while the size of the training set decreased. Fuzzy neural networks (FNNs) make use of fuzzy logic to model fuzzy data. Fuzzy neural networks had been applied to coffee and tainted-water data from an electronic nose. The results showed that the fuzzy neural model had about half the number of misclassified patterns as contrasted to their non-fuzzy counterparts, in addition to the model converge in less time and with much reduced error. The accuracy was improved to 93% for data on coffees and 85% for data on water by making use of the fuzzy neural model as compared to the figures of 86% and 75% before, respectively (Singh et al., 1996).

Taguchi is another method. This method is a variance reduction technique which can improve quality of a neural network at a minimum cost. In Taguchi’s method the combinations of process parameters are selected by orthogonal array. The results of simulation corresponding to orthogonal array are employed as training data for the artificial neural network to attain optimal conditions. Georgilakis et al. (2001) demonstrate the capability of the Taguchi’s technique to accurately characterize and
successfully optimize the transformer core production process with a minimum of experiments. In order to optimize the annealing process of cores, five controllable variables were identified for each of which two possible levels were considered. Ko et al. (1998) describe a new method of implemental design in multi-stage metal forming processes while considering workability as limited by ductile fracture. Artificial neural network using Taguchi’s method has been implemented in minimizing objective function relevant to the forming process.

The design of a neural network involves the selection of an optimal set of design parameters to achieve fast convergence speed during training as well as the required accuracy during recall. Khaw et al. (1995) described an innovative application of Taguchi’s method to determine design parameters to meet the required training speed and accuracy.

In the present research, the objectives were to identify important parameters on apple drying process and to determine the effect of each parameter, using neural network and neuro–Taguchi’s methods.

**Drying Rate**

Moisture transfer in materials during the drying process can be classified into two stages (periods):

1. From the inside of materials to surface (internal moisture transfer),
2. From surface to air through evaporation (surface moisture transfer).

If in the drying process the rate of moisture transfer from inside to the surface is less than the rate of surface moisture transfer, this process will be called falling–rate drying, and if the two are equal, the process will be called as constant–rate drying. Corresponding to the nature of material and drying air, one can observe both constant–rate and falling–rate drying in the drying kinetics.

For almost all food materials, constant–rate drying is generally a very short period in the drying process and mechanism and it will soon shift into the falling rate period.

During the falling rate drying, the evaporation area gradually recedes deeper into the food material. Drying happens in two falling rate periods. The first falling rate drying period will appear if the surface is partially wetted and as the dry surface appears, the second falling rate drying period will start. One can ignore the intermediate period between these two stages or periods (Keey, 1978; Treybal, 1980).

If drying happens in the first falling rate period, the internal resistance will have an effective predomination over the rate of drying. The turbulency of air on the surface can increase the rate of drying. During the second falling rate period, the rate of drying is controlled only by the internal resistance (Demirel and Turhan, 2003).

The point at which the drying process changes from the first stage to the second period is called “break point”.

The following equation has been employed in this research to find out Moisture Ratio (MR):

\[
MR = \frac{X - X_e}{X_0 - X_e}
\]  

In this equation \(X\) is the average moisture content with respect to time, \(X_0\) is initial moisture content, \(X_e\) equilibrium moisture content, and \(MR\) stands for moisture ratio. All moisture contents are expressed on a dry basis, i.e. in kg moisture/kg of dry solid.

If in the falling rate period \(MR\) is plotted versus time, a point will be found at which both the slope of curve as well as an drying mechanism change. This point is the aforesaid “break point”.

**Taguchi Method**

Taguchi’s method was developed as a process optimization technique by Genuchi Taguchi during the 1950’s (Ranjit, 1990).
The method is based on the statistical analysis of data and offers a simple means of analysis as well as an optimization of complex systems. The combinations of process parameters are selected through orthogonal array. Orthogonal array is a table of integers whose column elements show minimum, mean, and maximum levels of parameters. Each row of orthogonal array specifies one experiment. On the other hand, each row of the table specifies a special group of parameter levels which are to be evaluated.

Table 1 represents an orthogonal array $L_8(2^3)$ in which the first column specifies the number of experiments and the first row denotes parameters. The rest of the table is filled by 1s and 2s. Numbers 1 and 2 indicate minimum and maximum levels of each parameter, respectively.

If there are $m$ factors and $L$ levels for each parameter, the number of total experiments ($N$) will be equal to $L^m$.

With $N$ experiments, one can find the effect of each level on each factor by averaging the results containing that level and that factor. Through an application of method, the best composition of factors that can generate the best answer can be found.

**MATERIALS AND METHODS**

**Experimental**

In this research a pilot batch-through circulation dryer was initially made. This dryer involves a drying chamber, a blower, a hot air chamber, a gas burner, a temperature and humidity control board, sieve trays, as well as a hot air distributor (Figure 1). In order to carry out the experiments in similar conditions, Urmia apples obtained through purchasing Daniels company were used.
from local market were manually sorted and those with diameters ranging from 7-9 cm selected. They were put in the refrigerator to establish similar storage conditions for the samples before being dried.

To start each experiment, some apples were removed from refrigerator and placed at room temperature for 8 hours until similar environmental conditions established. At each stage a 25 g sample of the apples was put into an oven of 105ºC until the variation of its moisture content reached its minimum possible level. The dried samples were placed in desiccator until cold. The samples were weighed for the determination of their initial moisture content. Samples were cored, peeled, and cut into the desired thickness by a STIEBEL ELTRON cutting machine. Similar size pieces were weighed. To prevent changing the color of apple pieces into brown, they were immersed into a solution of 0.2% of metha bi sulfite for 10 minutes. This might also preserve vitamins C and A of the samples. Then the samples were removed from the solution and left until they reached to their initial weight.

Samples, after being put in a dryer, were removed from the dryer every 5 minutes to be weighed by use of an FX 3200 digital balance. After 45 minutes past, the samples were weighed every 15 minutes, until no change observed in their weight used finding equivalent moisture content. Main aim of determination of amounts of moisture content (X), initial moisture content (X₀), and equivalent moisture content (Xₑ), is calculation of amounts of MR according to Equation (1).

\[ X = \frac{M - S_s}{S_s} \]  

where \( M \) is the weight of wet solid in each time, \( S_s \) is weight of dry solid and \( X \) is moisture content on a dry basis.

In this dryer, temperature was measured by both PT 200 digital thermometer and PT 100 thermocouple, moisture content was measured by use of SUNWAN humidistat, and air velocity measured through TESTO 435 digital flowmeter.

For an application of the array \( L_s \) of Taguchi’s method, three parameters of temperature, thickness, and air velocity were varied each at two levels. These levels for temperature were \( T_1 = 70ºC \) and \( T_2 = 80ºC \); for thickness they were \( Z_1 = 4 \text{ mm} \), and \( Z_2 = 6 \text{ mm} \); and the levels for velocity were \( V_1 = 0.75 \text{ m s}^{-1} \) and \( V_2 = 1.5 \text{ m s}^{-1} \). Hence, according to this experimental design, the number of experiments would be equal to 8.

**Modeling and Simulation**

**Artificial Neural Network**

An artificial neural network similar to a biological neural system is composed of simple processing elements called neurons that are connected to each other by weights. Artificial neural networks include an input layer, one or more hidden layers and an output layer.

Each connection between two neurons has an associated value called weight. Number of neurons at input and output layers is equal to the number of system’s inputs and outputs, respectively.

According to the universal function approximation, a neural network which has one hidden layer and sufficient number of neurons in the hidden layer will be able to map every input onto every output.

Numerous types of the artificial neural network exist such as Multi-Layer Perceptron (MLP), Radial Basis Function (RBF), and Recurrent Neural Networks (RNN), but each type consists of the same basic features, namely: neurons, layers and connection. Multi-layer perceptron networks constitute one of the most popular and successful neural network architectures, suited to a wide range of applications. A simple example of such network structure is presented in Figure 2. The net input \( \left( S_j \right) \) to node \( j \) in the hidden and output layers is of the form:
Figure 2. A simple presentation of connections among inputs and output of neural network.

\[ S_j = \sum_{i=1}^{k} w_{ji} x_i + b_j, \]  

where, \( S_j \) is the net input of each neuron in the hidden and output layers, \( x_i \) denotes the inputs, \( w_{ji} \) represents the weights associated with each input/node connection and \( b_j \) is the bias associated with node \( j \).

Each neuron consists of a transfer function expressing its internal activation level. Output from a neuron is determined by transforming its input using a suitable transfer function. Generally, the transfer functions are sigmoidal, hyperbolic tangent and linear functions. Linear function is usually used in the output layer, while sigmoidal function widely used to establish non-linear relationship (Dornier et al., 1995; Niemi et al., 1995; Bowen et al., 1998 a, b; Delgrange et al., 1998; Teodosiu et al., 2000).

If \( P \) is the number of patterns, the relationship between inputs of network and input values of transfer function can be written as:

\[
\begin{bmatrix}
  x_1^{(1)} & \cdots & x_k^{(1)} \\
  \vdots & & \vdots \\
  x_1^{(p)} & \cdots & x_k^{(p)}
\end{bmatrix}
\begin{bmatrix}
  w_1 \\
  \vdots \\
  b
\end{bmatrix}
= \begin{bmatrix}
  \text{net}^{(1)} \\
  \vdots \\
  \text{net}^{(p)}
\end{bmatrix}
\]  

where \( (1), \ldots, (P) \), indicate patterns. The size of matrix is \( (p \times k + 1) \) and \( k \) is equal to number of inputs.

In order to have an artificial neural network of acceptable prediction ability, the neural network should find values of weights and biases which minimize error between actual and predicted outputs.

The aim of training process is to determine the weights that produce, from the inputs, the best fit of the predicted outputs over the entire training data set. At the training process, weights and biases are optimized in order to minimize a suitable error function. This form of training is termed "supervised training". The error function is determined by the following equation:

\[ E = \frac{1}{2} \sum_{i=1}^{P} (y_i' - y_i)^2 \]  

where \( y_i \) and \( y_i' \) are the computed and actual outputs, respectively.

By using one pattern such as \( p \) from training data set and determining the weights and biases related to one node such as \( i \), the output from node \( i \) (produced by using pattern \( p \)) can be calculated.

The difference between the computed output vector and the target vector must be compared with the intended desired error
and if the difference was greater than the desired one, weights and biases should be duly corrected. In order to correct weight and bias matrices, the following equation is employed,

$$\Delta w_i = \eta \left[ y_i - y_i \right]_p, \quad p = 1, 2, ..., P_i$$

(6)

where $\eta$ is a number between 0 and 1, called “learning rate".

The process of correction of weight and bias matrices and calculation of the value of $E$ will continue until the value of $E$ becomes smaller than the permissible desired error.

Normally, the back propagation method is employed to train multi-layer perceptron. To simplify mathematical relationships, the variation in weights is initially calculated for one pattern, and then the variation values are summed up together for all the patterns.

Generally the training process of a multi-layer perceptron can be summarized as follows:

1) Weights are randomly considered at a limited range.
2) One pattern is used by network with its output(s) being calculated.
3) To minimize error, the required variations of output layer weights are calculated.
4) According to the following equations, the required variations of the before last layer weights are calculated.

$$\Delta w_{mn} = \eta \delta_n y_m$$

(7)

where:

$$\delta_n = \left( y'_n - y_n \right) f'(net_n)$$

(8)

5) Steps 2, 3 and 4 are repeated for other patterns.
6) Total variations which are calculated for all patterns are used by neural network as variations of weights.
7) Steps 2 to 6 are continued until reaching either a minimum possible or an acceptable error.

Finally the best matrices of weights and biases are found.

For modeling of the drying process, a software was developed. The back propagation algorithm, which is most widely used for training multi layer perceptron, is applied for correction of weights in this software.

After developing the software, four variables are considered as the inputs for modeling the drying process. These are temperature, fruit thickness, air velocity, and time. In this study, values of final moisture at different time stages, temperatures, apple thicknesses, and air velocities were considered as network’s output.

The data set was separated in the three sets of: training, validation, and testing. Almost 60, 20, and 20% of data were considered for training, validation, and testing respectively.

It was found by trial and error that in order to minimize the intended error only one hidden layer including four neurons would be sufficient, and the suitable transfer functions for hidden and output layers would be sigmoidal and linear functions, respectively. Learning rate is equal to 0.01 and the number of training cycles needed to get to this error is about 1,000 epochs.

**Neuro–Taguchi’s Method**

For modeling a process by artificial neural network, no prior knowledge concerning the process is required. In fact neural network modeling is essentially a black box in nature with any existing prior knowledge being ignored. This feature is an indication of neural network method strength, but it also shows its weakness and this resulting in poor generalization. These weaknesses of neural network have encouraged recent researchers to focus on combining neural network with other such methods as fuzzy logic, Kalman filtering and Taguchi (Dimitris and Lyle, 1992; Michael and Mark, 1994; Thompson and Kremer, 1992; Hussain et al., 2002).
Figure 3. The results of dynamic simulation of moisture during the drying process in terms of temperature, thickness, and air velocity; ANN: 4/4/1 used.
In this paper Taguchi’s method as an optimization technique for training neural network is employed.

In neural network one can randomly consider a matrix as weight matrix for primary assumption to train in the conventional way.

Now it seems that the value of error which is calculated using a primary logic assumption for the first trial of weight matrix is significantly decreased as compared to the value of error which is calculated using a random weight matrix, with a similar number of training cycles.

Hence, for using Taguchi’s method in the software which is used for drying process modeling, a primary assumption for weights matrix is required. In order to find a primary assumption for weights matrix, every one of the two levels of each factor is considered for an evaluation of either the average weight and the error value, in a definite number of learning cycles.

Then with a comparison among obtained average errors, suitable values as primary assumption for weight matrix entries (related to each one of input factors) are considered.

In this way, in a definite number of learning cycles, the value of error is significantly decreased.

In this research, just values of weights between input and hidden layers are corrected and the values of weights between hidden and output layers at first assumption are quite randomly selected.

Of course for using neuro-Taguchi’s method, time as an input was omitted yielding the input parameters, for modeling drying process, limited to 3. The value of \( \frac{d(MR)}{dt} \) was calculated at break point for each one of the eight experimental sets and was considered as network’s output.

**RESULTS AND DISCUSSION**

The results of modeling while using neural network and software and while considering time as one of the inputs are shown in Figure 3. This figure shows the experimental data and the modeling results using neural network for moisture at different thicknesses, air velocities, temperatures, and times increments. Increasing temperature results in a substantial increase in the rate of drying due to water evaporation increase. Another factor exerting an important effect on drying rate is the thickness of solid to be dried. Increasing thickness decreases the rate of drying as a result of increase in resistance of water transfer from solid inside (Etemadi and Mousavi, 2004). Increasing air velocity...
also increases the drying rate, but it exerts little effect on the drying rate. This is because the basic way is more effective in the constant-rate period, due to increasing transfer coefficient rather, than in the falling-rate period.

It can be seen, with regard to Figure 3, that there is an excellent agreement between the predictions of neural network model (solid lines) and the experimental data for different conditions.

Slow drying is one of the characteristics of materials which fall in the falling-rate period. An investigation of the experimental data shows that there are two falling-rate periods (Figure 4), where the slope of curve is changed at the break point. For each one of the eight experiments, one can find curves similar to those in Figure 4. The calculated values of \( \frac{d(MR)}{dt} \) at break point are presented in Table 2. The variations of \( \frac{d(MR)}{dt} \) in terms of different parameters is the same as the variations of \( X \) in Figure 3. The software can be applied in process modeling without time considered as an input parameter. In this case, temperature, thickness, and air velocity are the inputs while the value of \( \frac{d(MR)}{dt} \) at break point being considered as output.

Two methods of neural network and neuro-Taguchi are employed for modeling. The difference between the two methods is just in the training process of the network. In the first method, weights between input and hidden layers are quite randomly selected. But in the case of using Taguchi’s method, in the training process, the values (as primary assumption for weights) between input and hidden layers are taken into consideration.

In order to use Taguchi’s method as an optimization technique for training neural network in neuro-Taguchi’s method, none of network’s parameters such as transfer functions, number of hidden layers, number of neurons in the hidden layer, number of learning cycle, and learning rate change as contrasted to neural network method.

Considered function for error (E) is the same as that presented by Equation (5). Then according to the following relation, the error matrix (M) with the same dimensions as output matrix is found as:

\[
M = \left( \frac{1}{p} \times E \right)
\]  

When one adds the entries of error matrix and divides the sum by the number of matrix entries, he can find a value which can be used as a suitable factor to compare the result of each obtained from these two methods.

The error values for neural network and hybrid neural network, both with eight patterns, are equal to 0.075 and 0.040, respectively.

### Table 2. Value of inputs and calculated results of \( \frac{d(MR)}{dt} \) as output value.

<table>
<thead>
<tr>
<th>Temperature</th>
<th>Thickness</th>
<th>Velocity</th>
<th>( \frac{d(MR)}{dt} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>70</td>
<td>4</td>
<td>0.75</td>
<td>(-9.27 \times 10^{-5})</td>
</tr>
<tr>
<td>70</td>
<td>4</td>
<td>1.5</td>
<td>(-9.7 \times 10^{-5})</td>
</tr>
<tr>
<td>70</td>
<td>6</td>
<td>0.75</td>
<td>(-6.09 \times 10^{-5})</td>
</tr>
<tr>
<td>70</td>
<td>6</td>
<td>1.5</td>
<td>(-6.29 \times 10^{-5})</td>
</tr>
<tr>
<td>80</td>
<td>4</td>
<td>0.75</td>
<td>(-1.52 \times 10^{-4})</td>
</tr>
<tr>
<td>80</td>
<td>4</td>
<td>1.5</td>
<td>(-1.7 \times 10^{-4})</td>
</tr>
<tr>
<td>80</td>
<td>6</td>
<td>0.75</td>
<td>(-1.39 \times 10^{-4})</td>
</tr>
<tr>
<td>80</td>
<td>6</td>
<td>1.5</td>
<td>(-1.5 \times 10^{-4})</td>
</tr>
</tbody>
</table>
With a comparison between these two values, it becomes evident that error value is improved by about 46.4%, using neuro-Taguchi’s method in the software mentioned in the section on modeling and simulation. Considered values as primary assumptions for weights matrix cause this difference to occur between error values. Using neuro-Taguchi’s method, the network bears a suitable assumption for weight matrix. Since in these two methods of training, the number of learning cycles as well as the other parameters of network are the same, the result shows that neuro–Taguchi’s method is a significantly more suitable training technique than the conventional neural network training method.

CONCLUSIONS

This study shows that temperature is a more important parameter, in apple drying rate, than either the velocity of air or thickness of apple slice. To model the drying process, a software was developed. This software uses back propagation algorithm for the training process and correction of weights similar to a multi-layer perceptron. Dynamic modeling of apple drying shows that neural network is a proper tool for modeling moisture variations in terms of time and other parameters.

In another modeling, time was omitted from among the inputs of the model and the values of temperature, air velocity, and thickness considered as the only inputs, while the value of \( \frac{d(MR)}{dt} \) at break point taken as output. When applying neuro-Taguchi’s method in the training process, a logical assumption (instead of a random one for the first trial of weight matrix) causes a significant improvement in the accuracy of the neural network. By using Taguchi’s in combination with neural network method, the weaknesses of neural network as related to the required accuracy are obviated.

REFERENCES


