Prediction of Dough Rheological Properties Using Neural Networks[1]

R. RUAN,[2] S. ALMAER,[3] and J. ZHANG[2]

Cereal Chem. 72(3):308–311

ABSTRACT

A neural network was designed to predict the rheological properties of dough from the torque developed during mixing. Dough rheological properties were determined using traditional equipment such as farinograph and extensigraph. The back-propagation neural network was designed and trained with the acquired mixer torque curve (input) and the measured rheological properties (output). The trained neural network accurately predicted the rheological properties (>94%) based on the mixer torque curve. The ability to measure the rheology of every batch of dough enables online process control by modifying subsequent process conditions. This development has significant potential to improve product quality and reduce cost by minimizing process variability during dough mixing.

Dough rheological properties are important for both product quality and process efficiency. Dough rheological properties, indicated by parameters such as the farinograph peak, extensibility, and maximum resistance, can be related to product specific volume and textural attributes. These parameters subsequently determine consumer acceptance. Therefore, accurate prediction of dough rheology could realize many benefits to the baking industry. However, measuring rheology of every batch is impractical, while predicting these rheological properties has historically proved to be complex. Therefore, most plant operations measure the rheological properties of only a few batches of dough per production shift. This makes online and intime process adjustment impossible.

Neural networks are new information processing techniques offering solutions to problems that have not been explicitly formulated. Much of the excitement surrounding neural networks is their unique ability to learn by experience. In the past few years, neural networks have shown increased power over many other statistical methods when solving nonlinear prediction problems (Bochereau et al 1992).

Neural network technology has been inspired by biological models. The building blocks of neural networks are neurons or processing elements. In biological systems, neurons operate by receiving input from individual dendrites. This input is weighted according to the synapses, and the resulting quantities are summed. If the sum is greater than the neuron threshold, the neuron executes a transfer function on the weighted sum, and passes the value onto the next neuron. Figure 1 illustrates a processing element (PE), the artificial analog of a neuron. Transfer function maps a PE's possibly infinite summation of input to a predefined range, the output. The operation of a processing element parallels its biological equivalent with synapses being replaced by connection weights.

In artificial neural networks, PEs are combined into layers. The parallel structure of the neural networks distinguishes them from traditional serial processing computers and results in some of the fundamental properties of neural networks. Neural networks can solve of problems that are traditionally difficult or impossible using alternative computing techniques. These problems can be characterized as involving complex, nonlinear processes, and noisy or incomplete data. The capability of neural networks to solve such problems suggests that neural networks can become valuable tools for food and agricultural industry since complex, nonlinear processes and noisy data are commonplace, and most food and agricultural processing involves estimation, prediction and control. Furthermore, the structure of neural networks provides not only structural parallelism, but also processing parallelism. This enables very fast decisions to be made in real time.

The learning or training phase of a neural network typically requires paired input-output data. The input is fed into the network, transferred through the network layers, and ultimately calculates a predicted output. This predicted output is subsequently compared with the actual output, and the connection weights between the PEs are modified to minimize the deviation between the predicted and actual output. This process continues until a defined accuracy has been reached. This is the concept of back-propagation. During this training phase, many factors of a neural network structure, such as the number of hidden nodes, and the number of layers, are varied by a trial-and-error approach to obtain the optimum network. At this point, the network can be fed input data alone, and the model will accurately calculate the predicted output.

Two of the key neural network variables studied in this research were learning rate and momentum. Learning rate controls the degree at which connection weights are modified during the training phase. The larger the learning rate, the larger the weight changes, and the faster the learning will proceed. However, if learning rates are set too high, the neural network will not converge to its true optimum. Momentum weights the importance of the previous iterations to the next connection weight modification.

Application of neural networks in food, agricultural, and bio-

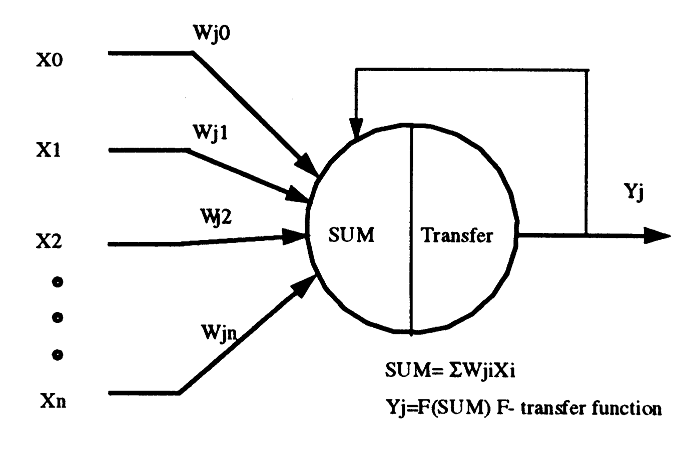

Fig. 1. Schematic diagram of a processing element. X = input; W = weight; Y = output.

[1]Journal paper 21,381 of the Minnesota Agricultural Experimental Station, St. Paul, MN.

[2]Assistant professor and visiting scholar, respectively. Biosystems and Agricultural Engineering Department, University of Minnesota, St. Paul, MN.
[3]Research scientist, Pillsbury Company, Minneapolis, MN.

logical industries is still in its infancy. However, some, such as the examples explained below, have already shown great promise.

Bochereau et al (1992) applied multilayer neural networks to predict apple quality from near infra-red (NIR) spectra. They used classical data analysis to extract principal components from the NIR spectra. These components were used as the input to successfully predict the sugar content of the fruit.

Thai et al (1992) used neural networks to determine the dominant wavelengths from the reflectance spectrum data of tomatoes when the color changed from green to red. Thai et al (1990) also evaluated the performance of neural networks in the evaluation of human preferences for honeydew melon.

Murase and Koyama (1991) built a three-layer neural network to predict the growth rate of radish sprouts with excellent accuracy. Ambient temperature, concentration of nutrient solution measured by electrical conductivity, and the time after seeding were used as input. Growth rate of hypocotyl length and the average radish sprout weight were the output.

Qian and Sejnowski (1988) and Zhang et al (1992) presented neural network models for predicting the secondary structure of globular proteins. The models learned from existing protein structures how to predict the secondary structure of local sequences of amino acids. The average success rate on a testing set of proteins nonhomologous with the corresponding training set was >64.3%. Seginer and Sher (1992) demonstrated the potential usefulness

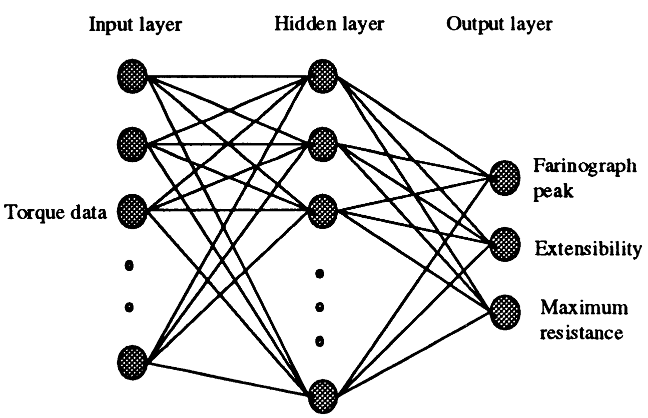

Fig. 2. Three-layer neural network structure.

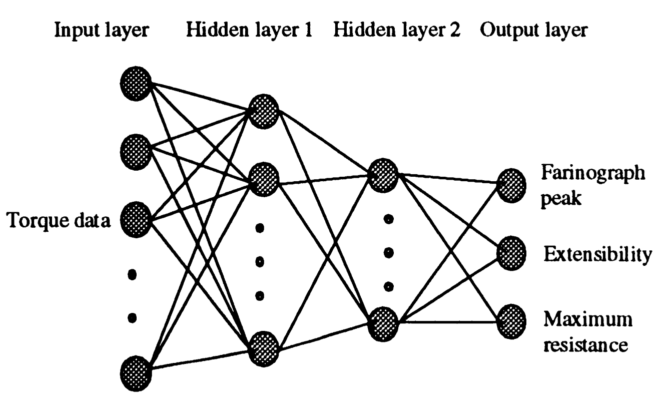

Fig. 3. Four-layer neural network structure.

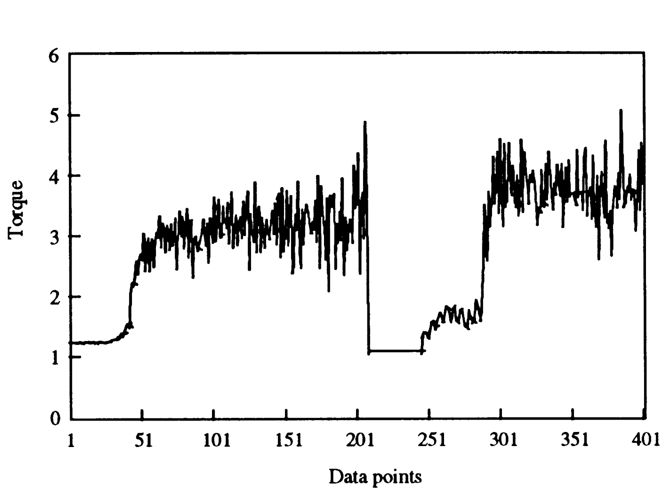

Fig. 4. Typical mixer torque curve generated during dough mixing.

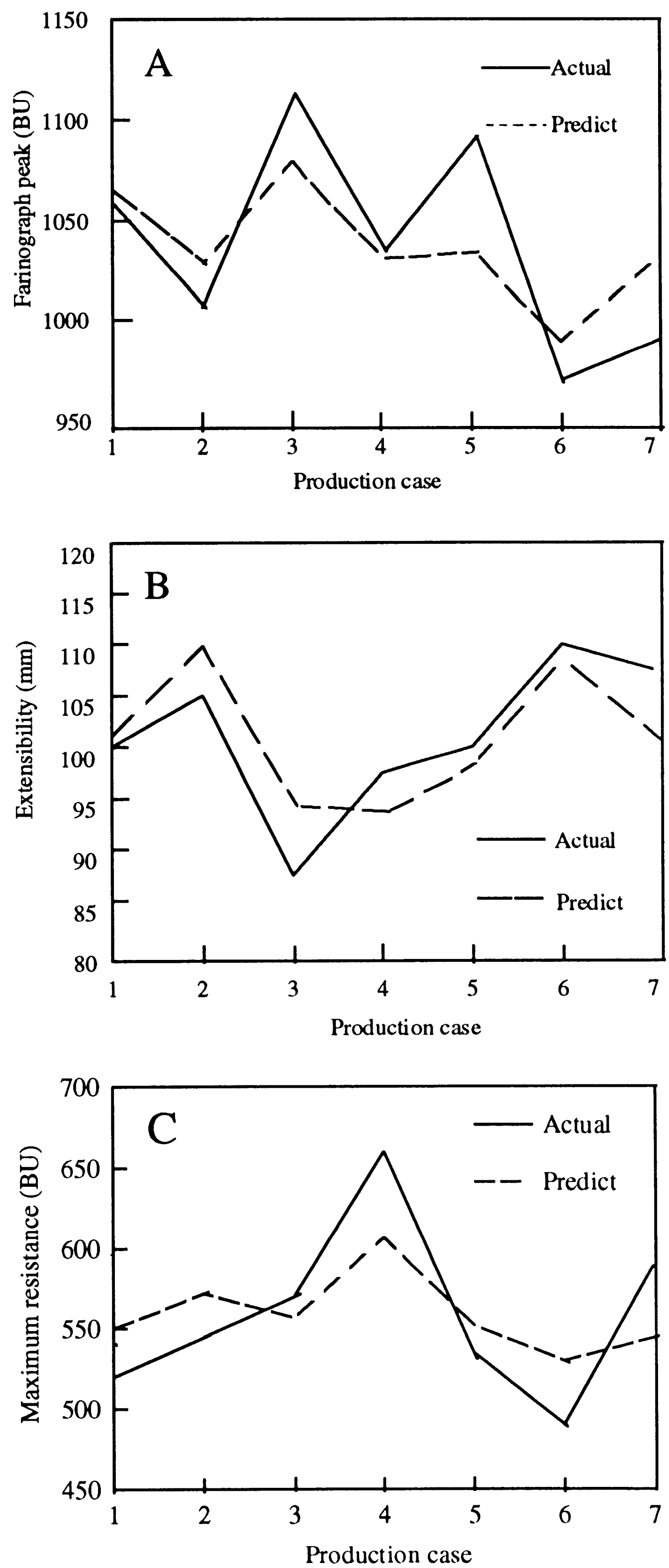

Fig. 5. Comparison of actual and prediction results. **A**, farinograph peak; **B**, extensibility; **C**, maximum resistance.

of the neural network approach for greenhouse environmental control, using days from transplanting, state of crop, next day's weather, unit of energy, and unit market price of crop as the input to predict next day's optimal temperature setpoint. Linko and Zhu (1992) have successfully applied neural networks for real time variable estimation and prediction in the control of glucoamylase fermentation. Input included oxygen consumption rate, carbon dioxide evolution rate, nitrogen utilization rate, pH, and agitation rate. The neural network was trained to predict the biomass and enzyme activity. Other potential applications for neural networks in the food, agricultural, and biological industry have been discussed (Bullock et al 1992, Rehbein et al 1992, Eerikainen et al 1993).

As discussed previously, dough rheology prediction is a complex problem. Work input, captured by the mixer torque curve, relates to dough quality; however, the precise relationship is unclear. There is no fundamental theory to relate work input during mixing to dough rheology. Many statistical methods have been studied, but an accurate model does not exist due to the complex, nonlinear nature of the problem, and the noisy data. The objective of this study was to develop a neural network to predict dough rheology using the work input during mixing.

MATERIALS AND METHODS

Experimental Design and Procedure

Sixty-two batches of dough were mixed using commercial formulations. An automatic data acquisition system based on a current transformer was developed to collect the work input, or torque, during mixing. The data acquisition rate was set at one reading per second, and a total of 400 points were acquired. Three rheological parameters were measured during this study: farinograph peak, extensibility, and maximum resistance. Extensibility and maximum resistance represent the width and height of the extensigraph curve respectively. AACC standard procedures were followed.

Neural Network Architecture

A total of 400 input PEs represent the 400 values from the mixer torque curve. There are three output PEs: farinograph peak, extensibility, and maximum resistance. A three-layer (one hidden layer) (Fig. 2) and a four-layer (two hidden layers) (Fig. 3) neural network structure were compared. All PEs in the neural networks were fully connected. Different numbers of hidden layer PEs, learning rates, and momentums were also evaluated.

Forty-nine of the 62 cases were used as the training set for the neural network, while six of remaining batches were used to test the network during training. Seven batches were used in a production set to evaluate network predictability. NeuroShell 2 software (Ward Systems Group, Frederick, MD) was used in neural network development.

RESULTS AND DISCUSSION

An example of a mixer torque curve is presented in Figure 4. The actual and predicted rheological properties of the production set are presented in Figure 5. The average predicted error was used to evaluate the neural network performance.

$$\text{Average prediction error} = 1/7 \times \sum_{i=1}^{7} \left| \frac{Actual_i - Predict_i}{Actual_i} \right| \times 100\%$$

where $Actual_i$ and $Predict_i$ are ith actual and predicted values in production set respectively

The effect of the number of hidden layer PEs on neural network is shown in Table I. The optimal prediction results were obtained with 50 hidden layer PEs, and the learning rate and momentum both set at 0.1. Learning rate and momentum did not have significant effects on the network performance as shown in Table II. This was expected, since the system converged well in all settings, except when learning rate and momentum were both set at 1. However, these settings do significantly affect the training speed. Therefore, as long as the system converges well, higher learning rate and momentum parameters can be used. In this case, a learning rate and momentum of 0.2 and 0.4, respectively, provided the optimum between convergence efficiency and prediction accuracy.

There is no theoretical reason to ever use more than two hidden layers. It is strongly recommended that one hidden layer be the first choice for any practical feedforward network design. More hidden layers may cause overfitting, since the network focuses excessively on the idiosyncrasies of individual samples. If using a large number of hidden PEs does not satisfactorily solve the

TABLE I
Comparison of Network Performance with Different Numbers of Hidden Layer Processing Elements (PEs) (n = 7)

	Farinograph Peak (BU)		Extensibility (mm)		Maximum Resistance (BU)	
Hidden Layer PEs	Average Absolute Error	Average Relative Error (%)	Average Absolute Error	Average Relative Error (%)	Average Absolute Error	Average Relative Error (%)
25	28.5	2.7	2.6	2.5	48.5	9.1
50	23.6	2.3	2.5	2.5	41.6	7.8
100	31.0	3.0	3.7	3.7	45.5	8.5
150	32.7	3.2	4.1	4.0	47.3	8.8

TABLE II
Network Prediction Results Using Different Learning Rates and Momentums

	Farinograph Peak (BU)		Extensibility (mm)		Maximum Resistance (BU)	
Learning Rate/ Momentum	Average Absolute Error	Average Relative Error (%)	Average Absolute Error	Average Relative Error (%)	Average Absolute Error	Average Relative Error (%)
0.1/0.0	24.8	2.4	2.8	2.7	49.6	9.3
0.1/0.2	25.0	2.4	2.6	2.5	48.9	9.2
0.2/0.4	23.9	2.3	2.3	2.3	40.6	7.6
0.4/0.6	24.7	2.4	2.6	2.6	43.1	8.1
0.6/0.9	26.6	2.6	5.0	4.9	38.8	7.3
1.0/1.0	39.5	3.9	3.6	3.6	40.9	7.2

TABLE III
Comparison of Four-Layer Network Performance Using Different Number of Hidden Layer Processing Elements (PEs)

Back Propagation Network		Farinograph Peak (BU)		Extensibility (mm)		Maximum Resistance (BU)	
First Middle Layer PEs	Second Middle Layer PEs	Average Absolute Error	Average Relative Error (%)	Average Absolute Error	Average Relative Error (%)	Average Absolute Error	Average Relative Error (%)
40	10	30.9	3.0	2.8	2.7	30.0	5.8
25	25	33.4	3.3	4.0	3.9	34.5	6.6
20	5	24.6	2.4	2.1	2.0	34.7	6.4
10	5	27.5	2.7	2.6	2.5	36.7	6.8

problem, then it may be worth using a second hidden layer and possibly reducing the total number of hidden PEs (Masters 1993). Since a single hidden layer did not produce optimal prediction results for extensibility (Table I), the second hidden layer was added. Table III shows the prediction results using two hidden layers with different PEs. The average predicted error improved using two hidden layers, especially for maximum resistance and extensibility.

A training set of 49 batches is very small for neural network development. We expect that a larger training set will greatly improve the reliability of the neural network. However, since the measurement error for farinograph peak, extensibility, and maximum resistance are typically 5, 1, and 5%, respectively, the prediction results obtained are very encouraging.

The immediate benefit of predicting dough rheology is the ability to measure the rheology of every batch of dough. This will enable process control to be used because the first step of any control application requires measurement of the key parameters. Being able to measure dough rheology at the mixer provides an exciting opportunity to reduce process through modifying key parameters such as flour-water ratios and mix times after each and every batch. By predicting rheology at the mixer, there is also an opportunity to use subsequent operations during production to correct or minimize variability.

CONCLUSION

A back-propagation neural network has been developed to accurately predict the farinograph peak, extensibility, and maximum resistance of dough using the mixer torque curve. We have determined that two hidden layers are beneficial for this neural network application. Other neural network variables, such as the number of hidden layer PEs, learning rate, and momentum did not significantly impact network predictability. This development has significant potential to improve product quality by minimizing process variability. The ability to measure the rheology of every batch of dough will enable online process control through modifying process conditions. This study illustrates that neural network technology holds great promise for the food, agricultural and biological industry.

LITERATURE CITED

BOCHEREAU, L., BOURGINE, P., and PALAGOS, B. 1992. A method for prediction by combining data analysis and neural networks: Application to prediction of apple quality using near infra-red spectra. J. Agric. Eng. Res. 51:207-216.

BULLOCK, D., WHITTAKER, D., BROWN, J., and COOK, D. 1992. Neural networks for your toolbox. Agric. Eng. 73(4):10-31.

EERIKAINEN, T., LINKO, P., LINKO, S., SIIMES, T., and ZHU, Y.-H. 1993. Fuzzy logic and neural network applications in food science and technology. Trends Food Sci. Tech. 4:237-242.

LINKO, P., and ZHU, Y. 1992. Neural network modeling for real-time variable estimation and prediction in the control of glucoamylase fermentation. Process Biochemistry 27:275-283.

MASTERS, T. 1993. Practical Neural Network Recipes in C++. Academic Press. Boston.

MURASE, H., and KOYAMA, S. 1991. Application of neural networks to agricultural engineering problems. Am. Soc. Agric. Eng. Pap. 917504. The Society: St. Joseph, MI.

QIAN, N., and SEJNOWSKI, T. J. 1988. Predicting the secondary structure of globular protein using neural network models. J. Mol. Biol. 202:865-884.

REHBEIN, D. A., MAZE, S. M., and HAVENER, J. P. 1992. The application of neural networks in the process industry. ISA. 31(4):7-13.

SEGINER, I., SHER, A. 1992. Neural-Net for Greenhouse Climate Control. Am. Soc. Agric. Eng. Pap. 927013. The Society: St. Joseph, MI.

THAI, C. N., RESURRECCION, A. V. A., DULL, G. G., and SMITTLLE, D. A. 1990. Modeling consumer preferences with neural networks. Am. Soc. Agric. Eng. Pap. 907550. The Society: St. Joseph, MI.

THAI, C. N., SHEWFELT, R. L., and LATIMER, J. G. 1992. Neural network analysis of tomato reflectance spectra. Am. Soc. Agric. Eng. Pap. 927057. The Society: St. Joseph, MI.

ZHANG, X., MESIROV, J. P., and WALTZ, D. L. 1992. Hybrid system for protein secondary structure prediction. J. Mol. Biol. 225:1040-1063.

[Received October 7, 1994. Accepted January 24, 1995.]