A NEW APPROACH TO MODELING AND CONTROL OF A
FOOD EXTRUSION PROCESS USING ARTIFICIAL
NEURAL NETWORK AND AN EXPERT SYSTEM

OTILIA POPESCU¹, DIMITRIE C. POPESCU² and JOSEPH WILDER¹

¹Center for Advanced Information Processing
Rutgers University
P.O. Box 1390
Piscataway, NJ 08855-1390

²Department of Electrical and Computer Engineering
Rutgers University
94 Brett Road
Piscataway, NJ 08854-8058

AND

MUKUND V. KARWE³

Center for Advanced Food Technology
Rutgers University
63 Dudley Road
New Brunswick, NJ 08901-8520

Accepted for Publication November 10, 2000

ABSTRACT

The paper presents a new approach to the modeling of the start-up part of
a food extrusion process. A neural network model is proposed and its parameters
are determined. Simulation results with real data are also presented. The inputs
and outputs of the model are among those used by the human operator during
the start-up process for control. An intelligent controller structure that uses an
expert system and "delta-variations" to modify inputs is also proposed.

INTRODUCTION

Extrusion cooking is used to manufacture food products like breakfast
cereals, expanded snacks, pasta, or bases for drinks and soups. Modified starch
and some types of pet foods are also manufactured by extrusion cooking.

¹ Author to whom all correspondence should be addressed.
Mathematical modeling of extrusion processes began in the field of plastic and polymer engineering and many of the results developed are applicable in food extrusion. However, the properties of food materials are more complex and more difficult to measure than those of synthetic polymers, which makes the mathematical modeling task more challenging. Modeling of quality changes during food extrusion is even more difficult. The functions of a cooker extruder include pumping, mixing, shaping, and working both as a heat exchanger and as a biochemical reactor. Therefore, for accurate understanding of extruder behavior, material flow, and physico-chemical changes during extrusion cooking, several models must be considered together. Because of the large number of variables involved in these models, and also because of nonlinear relationships among them, several approximations have to be made, which limits the results of the mathematical models making it difficult to be used over a wide range of conditions and materials (Mercier et al. 1989).

The interest in dynamic modeling and control of the food extrusion processes has increased in the last decade, especially due to the improvements in hardware and software. Mathematical models describing food extrusion process can be found in (Yaco 1985; Chang and Tan 1993) and extensive simulations have been performed (Karwe and Jaluria 1990; Della Valle et al. 1993; Sastrohartono et al. 1994) to study the extrusion processes. Mathematical models were also developed for step changes in extruder operating parameters (screw speed, feed rate, and barrel temperatures) in a twin-screw extruder (Lu et al. 1992, 1993). Also, new modeling tools like artificial neural networks (ANN) have been developed. ANNs are used in a wide range of practical applications from speech and character recognition, to modeling and control of industrial processes. In food processing, ANNs have been used to model the wine making process (Rattaray and Floros 1999). In food extrusion, ANNs have been used with some success in identification and control of the steady-state part of the extrusion process (Eerikainen et al. 1994). Zhu et al. (1995) have applied ANN methodology to enzymology. Haley and Mulvaney (1995) point out that due to the fact that ANNs can be trained to "learn" patterns associated with a particular product they can be also used for process control applications in the food industry.

However, transients in the food extrusion process, which include changes in operating conditions, start-up, and shutdown, are still a subject of research, since during transients the product deviates from specifications. Start-up and shutdown procedures, in particular, are very complex and waste both time and raw material, therefore appropriate models and control tools are required, to minimize losses. During the start-up period the objective is to reach steady state as fast as possible, and this part of the food extrusion process is more complex because the number of parameters that change and need to be taken into account
is larger than during the steady state operation. No mathematical models have been established for the start-up part of the food extrusion process yet.

This paper deals with modeling the start-up of a food extrusion process using artificial neural networks. After establishing the model, a "delta variations" controller is proposed for this part of the process, in an attempt to emulate the behavior of a human operator.

ANNs for Modeling and Control

ANNs were developed in the early 1940s using biological neural networks as a model. They consist of simple interconnected processing elements, called neurons. A neural network is a massively parallel distributed processor that resembles the brain in two respects: it acquires knowledge through a learning process; knowledge is stored by means of inter-neuron connection strengths known as synaptic weights.

The power of a neural network comes from its massively parallel distributed structure, as well as from its ability to learn and generalize. Generalization means that the neural network can produce reasonable outputs even for inputs not encountered during the training process. These capabilities enable neural networks to solve complex problems that are very difficult or even impossible to solve with conventional deterministic or statistical methods.

In addition to their parallel structure and adaptation properties, ANNs are also universal approximators, which makes them very promising for modeling and control of nonlinear systems (De Wilde 1996; Suykens et al. 1996). Also, the multiple input-multiple output structure makes ANNs very well suited for application to multivariable systems.

The most widely used type of ANNs are the feedforward neural networks, which consist of a set of source nodes (the sensory units) that constitute the input layer, one or more hidden layers and an output layer which perform the computation. In a feedforward neural network, the input signal propagates through the network in a forward direction, on a layer-by-layer basis. When the history of the process must also be taken into account feedback connections from the output variables to the network input can also appear. However, the feedforward characteristic of the network remains unchanged since there are no feedback connections from the hidden layers and the input signal still goes in the forward direction.

The network is trained using the so-called "error back-propagation process" which consists of two passes through the different layers of the network: a forward pass and a backward pass. In the forward pass, a known input vector

---

¹One can regard steady state as a particular case of start-up.
is applied to the input nodes of the network, and its effect propagates through the network, layer by layer and a set of outputs is produced as the actual response of the network. During the forward pass, the synaptic weights of the network are all fixed. The actual response of the network is subtracted from the desired response to produce an error signal which is then propagated backward through the network. This constitutes the backward pass, in which the synaptic weights are adjusted in accordance with the error-correction rule. The weights are adjusted so as to make the actual response of the network move closer to the desired response. A version of the back-propagation algorithm which updates weights after each pattern by cycling through the training data can be found in (Haykin 1994).

The Artificial Neural Network Model

The target of our research was the development of an intelligent controller for the start-up phase of the food extrusion process that uses an expert system which must incorporate decision rules for situations generated by the process. In general when working with food materials, the system behaves slightly different during start-up due to inherent differences in properties of the same material (e.g. corn meal coming from different manufacturers). Therefore, it is not possible to program the input variables for start-up under different situations that may occur. Hence, we decided to use an expert system to analyze the start-up process and make decisions about how the variables must be modified. The expert system must incorporate decision rules for situations generated by the process and should be similar to the way a human operator makes decisions.

In order to develop such an expert system, we need first an accurate model of the process. We consider that an ANN model is well suited for start-up of the extrusion process, because the process is very complex and highly nonlinear. The expert system and the model together form a simulator of the controlled process on which extensive testing can be performed until a desired level of performance is reached, at which point the model can be substituted with the real process. This proposed structure is presented in Fig. 1.

Simulation of extruder start-up requires initial and final values for input variables. These values depend on the product and they are known to the operator that controls the process (water flow rate, flour mass flow rate, screw speed, barrel temperatures, etc.). The accumulator unit (ACC) modifies input variables according to the decisions of the expert system and passes them to the process model which will calculate the corresponding output. The expert system incorporates a comprehensive set of rules which test both input and output variables deciding whether steady-state has been reached, or how the input variables must be modified the next time slot. It also monitors both input and output values to be within allowed ranges to avoid potential extreme situations.
The neural network model of the food extrusion process presented in Fig. 2 consisted of the input layer with 36 neurons, the output layer with 4 neurons, and a hidden layer with 6 neurons. The model structure was established and trained using the methodology described in next section.

Optimum Structure and Training of the ANN Model

In order to gather data necessary to train the neural network model for the start-up part of the extrusion process, experiments were carried out and process inputs and outputs were recorded till the process reached steady-state. All experiments were conducted on a co-rotating, self-wiping, twin-screw extruder (ZSK-30, Krupp Werner & Pfleiderer Corp., Ramsey, NJ). The extruder had two 30.7 mm diameter screws and a processing length of 878 mm (L/D of 28.6). The extruder barrel consisted of a feed zone and five independent zones.
as shown schematically in Fig. 3. Each zone was made from a separate steel block. Zones 1 to 3 were about 180 mm long each. Zones 4 and 5 were 90 mm long each. Each zone was equipped with two resistive heaters, separate cooling water channels and a thermowell with a control thermocouple (type J). All zones were held together tightly by four tie rods. A die zone which was immediately after zone 5 did not have any heating or cooling arrangements. A die with two holes (3 mm diameter, 5 mm long) was used in all experiments.

Corn meal (Conolly Calhoun Conolly, Pennsauken, NJ) was the raw material used in all experiments. The moisture content of corn meal was determined prior to each run by heating a known mass of the sample at 135°C for 2 h in an oven (AACC method 44-19).

FIG. 2. NEURAL NETWORK MODEL FOR FOOD EXTRUSION PROCESS
FIG. 3. SCHEMATIC DESCRIPTION OF THE TWIN-SCREW EXTRUDER AND PROCESS VARIABLES

Process Variables

$T_i [^\circ C]$ - temperature inside zone $i$, $i = 1 \ldots 5$

$T_{\text{die}} [^\circ C]$ - temperature at the die

$F [\text{g/min.}]$ - flour mass flow rate

$W [\text{g/min.}]$ - water flow rate

RPM [rpm] - screw speed

$T_q [\% \text{ of maximum}]$ - torque

$P [\text{kPa}]$ - pressure at the die

$M$ - overall acceptability of the product
A K-Tron twin-screw feeder (Model T-20, K-Tron Corporation, Pitman, NJ), was used to meter corn meal into the feed section of the extruder. It is equipped with a hopper and a rotary agitator to prevent bridging of corn meal. The feeder was calibrated prior to each extrusion run.

Water was injected into the feeding section immediately after the flour feed using a triple action piston pump with adjustable frequency and stroke (U.S. Electric Co., Milford, CT). Air in the water line from the pump was purged and the water pump calibrated before each run.

Pressure and temperature at the die were measured using Dynisco (Sharon, MA) TPT 463E transducers. Torque and screw speed were displayed on the extruder control panel. Torque was indicated as a percentage of maximum allowable torque. Die pressure, die temperature and torque were recorded during the run on a PC using Keithley Metrabyte data acquisition system.

The input variables considered were the flour mass flow rate ($F$), the water flow rate ($W$), the screw speed (RPM), and the temperatures of the three barrel zones closest to the die section (zones 3, 4, and 5; see Fig. 3), since these were the process variables modified by the operator during the start-up process. Temperatures in zones 1 and 2 were not included in the model since they were modified only occasionally during start-up and do not bring enough information to justify the increase in complexity of the neural network model.

The output variables considered were the total torque ($T_D$), the pressure at the die ($P$), the die temperature ($T_{die}$), and the overall acceptability of the product ($M$). The product was graded on a 10 to 1 scale by the experienced supervisor; 10 representing the worst overall acceptability of the product at the beginning of the start up process and 1 representing the best overall acceptability of the product that should be reached at steady state. We would like to note here that product moisture, texture, crispness, etc., were considered by the supervisor in grading the overall acceptability of the product.

In order to define the neural network we must specify the number of layers and neurons on each layer, as well as a set of weights for each neuron. These make up the network topology, but for training and dynamic simulation, several other parameters like learning coefficients and momentum constants (Haykin 1994) must also be specified. All these parameters constitute degrees of freedom in the selection of the neural network model. Topology of the network can be varied by changing the number of input, hidden, and output neurons, as well as the number of time delays for different inputs and outputs. Also, initial weight values can affect the convergence of the backpropagation algorithm. This is why optimization of the neural network parameters is a difficult and time consuming process.

An original approach was taken in the selection of the activation function of the neurons, which determines the range of the output values of the network. Instead of using the logistic sigmoid function whose output is between -1 and 1
with scaled values of the process inputs and outputs, a scaled and delayed version of it was used. This was chosen so that it can be well approximated by a linear function of the range of values taken by the process inputs and outputs. Its equation is

\[ f(x) = \frac{M-m}{1+e^{-a(x-\tau)}} + m \]  

where parameters

- \( m \) is minimum value of \( f(x) \)
- \( M \) is maximum value of \( f(x) \)
- \( a \) is slope parameter of \( f(x) \)
- \( \tau \) delay parameter of \( f(x) \)

are determined by trials, in which Eq. 1 is plotted for different values of the parameters, and we select those values for which the resulting curve is closest to a straight line within the range of values taken by process parameters. The function selected for our neural network and its parameters are presented in Fig. 4.

As a performance criterion in estimating the quality of the neural network model the cumulated squared error over a whole training sequence of data was used

\[ E = \sum_{n=0}^{N} \sum_{i=1}^{L} [y_{ai}(n)-y_{ni}(n)]^2 \]  

where \( y_{ai}(n) \) is the actual value of the \( i \)-th output, \( y_{ni}(n) \) is the target (desired) value of the \( i \)-th output, \( L \) is the number of network outputs, \( N \) is the length of the training sequence. As the backpropagation training algorithm is run and the parameters of the neural network (weights and biases) are adjusted, \( E \) decreases and after some number of training epochs will reach a point from where it is no longer decreased by continuing the training. This is the point where the parameters of the network have been adjusted to match the training data, and further training using the same training data will bring nothing new for the neural network. It is also possible that continuing the training beyond this point will increase \( E \) as the backpropagation algorithm continues to change network parameters. This measure will be the criterion used in selecting an optimal network architecture.
FIG. 4. SCALED LOGISTIC SIGMOID FUNCTION USED AS ACTIVATION FUNCTION (EQ. (1))

The network architecture is determined by the number of hidden neurons and the number of delays that will account for the history of the extrusion variables. The number of hidden neurons is very important for the neural network model: too few will hinder the learning of nonlinearity, while too many may result in overlearning of training data or memorizing, which affects the generalization capabilities of the network. Since in extrusion, the history of variables is also very important to model the current values of all variables, a series of delay experiments had to be performed to find out a proper configuration of input and output layers. For example, when the number of delays is 3, each variable \( x \) has its \( x(t) \), \( x(t-1) \), \( x(t-2) \), and \( x(t-3) \) terms in the network.

It can be observed from Fig. 5 and 6 that the smallest cumulated squared error was obtained after training a neural network with 6 hidden neurons and three delays on each input and output variable.

After these trials, a neural network was chosen that can be used as a simulator for the start-up phase of the extrusion process (within the ranges of
input and output training values). It had 6 inputs and 4 outputs, 36 input neurons, 6 hidden neurons, and 4 output neurons, and its structure is presented in Fig. 2.

![Graph showing effect of hidden neuron number on neural network performance](image)

**FIG. 5. EFFECT OF HIDDEN NEURON NUMBER ON NEURAL NETWORK PERFORMANCE**

We must also note that process data used for training was intentionally and artificially corrupted with a small amount of noise and then presented to the backpropagation training algorithm. This was done to ensure better generalization and prevent memorizing, as well as to diversify the training data set.

**Simulation and Experimental Results**

The structure of the programs used in simulations is presented in Fig. 7. Programming was done in C and MATLAB on Sun Sparcstation using Solaris operating system. The routines for neural network initialization, training using
the back propagation algorithm, and simulation, were written in C, while MATLAB was used for preparing the files with training and test data, as well as for result plotting and data visualization.

![Graph](https://via.placeholder.com/150)

**FIG. 6. EFFECT OF INPUT AND OUTPUT DELAYS ON NEURAL NETWORK PERFORMANCE**

Neural network data, consisting of number of inputs and outputs, number of neurons on each layer, and weights and biases for each neuron, was stored in ASCII files which are very easy to manipulate and can also be used by other programs for data visualization and result plotting.

An initial simulation was carried out on the same data that has been selected for training. Since the actual training data was a noisy version of data collected during the experiment, using the real experimental data for simulation helps in deciding if the neural network model performs well on data close to the training set, and is also a step towards model validation. Results of this simulation are presented in Fig. 8 and 9.
Next, simulation on a data set different from the one chosen for training was done. Actual validation of the neural network model is done by comparing the response of the neural network to input data that is no longer close to the training set with the response of the real system. Results of this simulation are presented in Fig. 10 and 11.

In order to see how well the model performed relative errors between model outputs and measured process outputs for both simulations described above are shown in Fig. 12 and 13. The relative error in variable $x$ is defined as the ratio of the error between the simulated and real output to the real output

$$
\varepsilon = \frac{x_{\text{sim}} - x_{\text{real}}}{x_{\text{real}}}
$$

(3)
Although at the beginning, the relative errors are high, they decrease and settle down within ±10%. Only the output corresponding to overall acceptability of the product in the model displays high relative errors in the end. However, even though the 100% relative error in the end seems high, one should keep in mind that this represents only a 1-point error on the 10 to 1 grading scale used. Also, the apparent running away of the relative error for this parameter is due to the choice of its definition, since it can be seen from Fig. 9 and 11 that the predicted overall acceptability of the product has stabilized.
Proposed Start-up Controller

One approach in designing a controller for the start-up phase of the extrusion cooking process is to use an expert system that modifies control variables using "delta variations". This is a version of discrete-time control in which, at each step, the value of a control variable is either incremented or decremented with a fixed amount $\Delta$, or is not modified. The decision "increment/decrement/do not modify" is taken based on the current and possibly past values of the measured (output) variables according to the set of rules incorporated in the expert system. To have a good process start-up controller these rules must be refined and extended before being incorporated into the expert system, such that the final set of rules includes all possible situations that may occur during start-up. The block diagram of such a system was presented in Fig. 1. The proposed expert system shown in Fig. 1 tests, at regular intervals, the output variables of the model (in this case they are torque, die pressure and die temperature, which is what a human operator looks at). These
variables are compared at each interval with the desired values and also tested for extreme situations. The rules that are incorporated in the expert system must cover as many situations as can be expected under factory operating conditions. The control decisions taken should generate a smooth running extruder that reaches steady state as quickly as possible. Based on the decisions taken by the expert system, model inputs are updated by the accumulator which will increase or decrease them with the prescribed values. The neural network model is used to test the expert system for all possible situations. When all the rules in the expert systems are very well formulated, and simulations meet the required specifications, the expert system will be switched to the real extruder.

![Graphs showing input variables](image)

**FIG. 10. INPUT VARIABLES DIFFERENT FROM THE TRAINING SET**
This is data collected from different experiments than the data used for training.

We should emphasize here that the model and the expert system must be developed for every product, because different food materials behave differently and generate very different situations. However, the basic procedure remains the
same, and after developing the system for one product it should be possible to modify for another product.

FIG. 11. COMPARISON OF REAL (--) AND SIMULATED OUTPUTS (—) FOR INPUTS DIFFERENT FROM THE TRAINING SET

Future work needs to be done to determine and set up a comprehensive set of rules to be used by the expert system. Considerably expanded training and testing data sets are required in order to cover virtually all situations that can occur. Future research should also include reliable on-line moisture measurement devices, flavor and color sensors, residence time and distribution sensors. When these are combined into a robust ANN model and an expert system the most desirable control system for a food extruder can be developed.
CONCLUSIONS

A new approach to modeling and control of a food extrusion process is presented. The work dealt with the transient part of the process for which the procedure of developing a neural network model is presented. The design steps, which included choosing the network topology, training, validation and testing, are illustrated with data taken from a real process that uses corn meal as raw material. The neural network model approximated the transient part of the process well even when input variables were quite far away from training data. It was found that for the extruder system presented in the paper, 6 hidden neurons and 3 input and output delays gave optimum performance.
FIG. 13. RELATIVE ERRORS BETWEEN MODELED OUTPUTS AND MEASURED PROCESS OUTPUTS FOR INPUTS DIFFERENT FROM THE TRAINING SET

ACKNOWLEDGMENTS

This is publication no. D10550-1-00 of the New Jersey Agricultural Experimental Station supported by state funds and the Center for Advanced Food Technology, (CAFT) of Rutgers, The State University of New Jersey.

REFERENCES


