Predictive Control of Quality in Batch Polymerization Using Hybrid ANN Models

Andy Yen-Di Tsen, Shi Shang Jang, and David Shan Hill Wong
Chemical Engineering Dept., National Tsing Hua University, Hsin-Chu, Taiwan, 30043 R.O.C.

Babu Joseph
Chemical Engineering Dept., Washington University, St. Louis, MO 63130

Two issues involving the methodology used for on-line control of product quality in batch manufacturing processes are addressed: the generation of fast, data-driven process models and the use of such process models for on-line feedback control of product quality. The methodology is investigated using the example of the control of dispersity and molecular weight distribution in a batch reactor for emulsion polymerization of vinyl acetate. An artificial neural network (ANN) is used as a model to predict the quality as a function of the manipulated variables and on-line measurements. This model is constructed using an augmented dataset that integrates experimental information and knowledge from a mathematical model. The proposed model is compared with other types such as a theoretical model whose key parameters are fitted to experimental data. The hybrid ANN is superior to the parameter-fitting approach for this case. Experimental and simulation studies confirm the advantage of using the proposed model and the predictive control algorithm.

Introduction

Objective

Batch processes are especially suited for manufacturing of small amounts of high value added products. Quality control in batch processes presents a uniquely challenging problem because the product quality is not known until the batch processing is completed. One is naturally forced to use predictive models to compensate for undesirable disturbances detected through intermediate measurements while the batch is in progress. Due to the complex nonlinear nature of batch-process dynamics, a model capable of accurately predicting the final product quality is usually difficult to generate. Feedback control therefore requires significant operator experience and input. Statistical process control (SPC) and statistical quality control (SQC) provide systematic ways of monitoring process status (Klein, 1991; Levinson, 1992) but does not address the issue of on-line feedback corrections when the quality measurement is not directly available. Due to the strong nonlinearity exhibited by most batch processes, there is still much room for development of automated decision-making tools for batch quality control.

Correspondence concerning this article should be addressed to S. S. Jang.

Artificial neural networks have been suggested as an alternative for generation of process models. These models are data-driven and have the advantages of computational efficiency and ease of construction. They also provide convenient means for on-line adaptation as shown by Joseph and Wang Hanratty (1993). Since batch processes are often poorly understood physically, the artificial neural network (ANN) seems to be a promising modeling tool for modeling such systems.

Joseph and Wang Hanratty (1993) developed a "shrinking horizon" predictive control scheme applied to an autoclave curing process for manufacturing composites using an ANN model. In this approach, ANN predictive models were developed from simulation runs. Corrective control actions were based on predictions of the ANN model and intermediate measurements. They showed that tight quality control can be achieved using model-based on-line feedback corrections to the batch recipe. Their studies were limited to simulated processes. They also did not address an important issue of generating the data required for training the ANN model. Normal operating data cannot be used for fitting models be-

cause of the lack of variability and correlation among the variables. On the other hand, the design and execution of a large number of experiments to generate data for modeling is expensive and time-consuming and may not be acceptable in a production environment.

The objective of this work is to evaluate the ANN model-based predictive control approach for the product quality control in a batch reactor for the emulsion polymerization of vinyl acetate. A new approach for creating a *hybrid* ANN model that combines scarce experimental data with prior knowledge encapsulated in approximate fundamental, first-principles models is developed and tested. The model is obtained by "training" an ANN model based on an augmented data set. The augmented data set combines the information from the experimental data set and the "trends" predicted by a rigorous model built from fundamental physical and chemical phenomena taking place in the system. This approach is applied to a bench-scale unit for the batch polymerization of vinyl acetate.

Polymerization model

The vinyl acetate polymerization reactor is an interesting paradigm for studying batch process control. Considerable past work has been done on trying to model this process. A complex first-principle model for the batch polymerization of vinyl acetate was derived by Penlidis et al. (1985). The model was implemented to compute the optimal control profile by Jang and Yang (1989). The experimental results show that the model is acceptable in predicting monomer conversion but unreliable in predicting quality of the product, that is, number average molecular weight and dispersity. Penlidis et al. (1988) found that impurities affect the performance of the reactor in two ways. Impurity dissolved in water will determine the induction period before the polymerization starts to take off. Impurities dissolved in the monomer may affect such final conditions of the batch as end time and molecular weights. In general it is difficult to develop mathematical models to quantify the precise effects of these impurities on product quality. Hence it is desirable to seek empirical or semiempirical models that can capture the actual behavior of a given reactor system. Even if precise mathematical models are available on-line, estimation and adaptation may be needed since critical measurements are often not available. In such cases, tools such as the extended Kalman filter must be used in conjunction with the nonlinear model of the process (Kozub and MacGregor, 1992).

In this work, the model developed by Penlidis et al. (1985) is used as the basis for developing a hybrid ANN model. The purpose is to demonstrate how experimental data can be used to augment model accuracy when prior knowledge is incomplete. The hybrid ANN model is, in turn, used to perform feedback predictive control based on the strategy proposed by Joseph and Wang Hanratty (1993).

Hybrid ANNs

Recently, the use of an ANN to capture nonlinear dynamics and serve as a "black-box" model has been extensively researched. ANNs have been applied to diverse areas of chemical process engineering such as fault diagnosis (e.g., Hoskins and Himmelblau, 1988; Venkatasubramanian and Chan, 1989; Kramer and Leonard, 1990), data reconciliation (Whitely and Davis, 1992), and process control (Pao, 1989; Bhat and McAvoy, 1990; Hernandez and Arkun, 1992; Psichogios and Ungar, 1991, 1992). Sarle (1994) and Ripley (1993) interpret that ANNs are nothing more than nonlinear regression and discriminate models.

A number of workers have addressed the issue of combining fundamental mathematical models with data-driven regression models such as ANN. Thompson and Kramer (1994) classified the approaches into two categories: the process model "in series" to ANN and the model "in parallel" to ANN.

Psichogios and Ungar (1992) implemented an "in series" hybrid ANN to perform tracking of a biochemical reactor. In their approach, a neural network is used to estimate parameters of a first-principles model. It was found that this hybrid ANN requires less training data, and extrapolates and interpolates better than standard "black-box" models. However, in this approach, it is implicitly assumed that all essential characteristics of the process had been captured in the first-principles model.

Kramer et al. (1992) classified prior knowledge into two forms: "hard constraints" (such as mass balances) that the network response must adhere to, and "default models," which are idealized theoretical models of system behavior. Hard constraints are incorporated by formulating the training problem into a semiinfinite programming problem. A connectionist model is trained to learn the residual between training targets and default model predictions. In this case, the ANN model is "in parallel" to the process model. Radial basis function networks are used to ensure that the combined model extrapolates according to the default model for novel cases outside the training data set. This hybrid ANN model was used to predict the behavior of a simulated vinyl acetate polymerization reactor.

The Batch Process Quality Control Problem

We start with a general statement of the batch product quality control problem, and then consider a simplified version that was used in the application study on hand. These simplifications are not necessary but used for convenience only.

Consider a batch process with an operating period $0 \le t \le t_f$. Its operating objective $\Psi[t_f, \boldsymbol{x}_f; \boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{d}(t)]$ would be a function of the end time t_f, state vector of the system at the end of the batch $\boldsymbol{x}_f$, the state vector trajectory $\boldsymbol{x}(t)$, the manipulated variable vector, $\boldsymbol{u}(t)$, and some unknown disturbance vector $\boldsymbol{d}(t)$. Assume that the batch operation only considers the quality constraints $\boldsymbol{q}[\boldsymbol{y}(t_f)] \ge 0$, where $\boldsymbol{y}$ are the measurements and functions of state vector $\boldsymbol{x}$ only. A general optimal control problem for the batch process is to compute the control policy $\boldsymbol{u}(t)$ such that:

$$\min_{\mathbf{u}(t)} \{\Psi(t_f, \boldsymbol{x}_f; \boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{d}(t))\}, \qquad \text{(OP1)}$$

subject to system dynamics:

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{d}(t)), \qquad \text{(OP1.C1)}$$

with the end product quality

$$y(t_f) = g(x_f) \qquad \text{(OP1.C2)}$$

and quality constraints

$$q(y(t_f)) \geq 0. \qquad \text{(OP1.C3)}$$

In addition some intermediate measurements of the process are also available:

$$z = h(x(t), u(t), d(t)). \qquad \text{(OP1.C4)}$$

Model predictive control could be performed only if accurate models for the system dynamics, f, the relation between measured variables and the quality of the final product q, as well as the relation between measured variables and the state of the system, g, were known. However, such models are generally unavailable.

To evaluate the model predictive control strategy we consider a simplified version of the preceding problem adapted to the vinyl polymerization process. In batch polymerization, usually only discrete manipulative actions are taken. We consider a one-shot control policy whereby the feedback control is based on a single intermediate conversion measurement, z. Further, only one primary disturbance in the feed is considered in this study. The method can be extended to the case of multiple control actions, multiple measurements, and multiple disturbances. The manipulative action is classified into two classes, those taken before the intermediate measurement $\{u_0\}$ and those taken after $\{u_1\}$. Thus, the on-line predictive control problem can be rewritten in the following form:

$$\min_{u_1} \{\Psi(t_f, x_f, x_0, u_0, d, u_1)\}, \qquad \text{(OP2)}$$

subject to the identification of

$$d = d(z), \qquad \text{(OP2.C1)}$$

the system model:

$$x_f = x(t_f, x_0, u_0, d, u_1) \qquad \text{(OP2.C2)}$$

and the quality constraints:

$$q(x_0, u_0, d, u_1) \geq 0. \qquad \text{(OP2.C3)}$$

A further assumption made here is that the disturbance is observable, that is, we can estimate it using z. Thus any disturbances entering the system that do not affect z cannot be corrected by this control strategy. If there are other disturbances on the process, then additional measurements will be necessary so that the effect of these disturbances can also be included in the prediction of product quality.

Neural network model predictive control

In many cases, the effect of disturbances, d, cannot be fully understood. Even worse, many key parameters in the physical model cannot be evaluated accurately. Hence predictions based on first principles only are generally not quantitatively accurate enough for use in predictive control. It is very useful if the product quality model can be augmented by a data-driven model such as ANN. The most extreme case is to develop the following ANN models:

$$\Psi = \text{ANN}_\Psi(t_f, x_f, x_0, u_0, d, u_1) \qquad (1)$$

$$d = \text{ANN}_d(x_0, u_0, z) \qquad (2)$$

$$q = \text{ANN}_q(x_0, u_0, d, u_1) \qquad (3)$$

based entirely on experimental data. Other approaches are also feasible. For example, one could consider building a single ANN model to predict the quality and the objective function at the same time. Using the preceding approach, the optimal control problem (OP2) can be rewritten as

$$\min_{u_1} \Psi = \text{ANN}_\Psi(t_f, x_f, x_0, u_0, d, u_1), \qquad \text{(OP3)}$$

subject to the identification of

$$d = \text{ANN}_d(t_f, x_f, x_0, u_0, d, u_1) \qquad \text{(OP3.C1)}$$

and the quality constraints

$$q = \text{ANN}_q(t_f, x_f, x_0, u_0, d, u_1). \qquad \text{(OP3.C2)}$$

The model predictive control strategy is implemented as follows. We start with a nominal control policy. After some time has elapsed, a sample is taken from the polymerization reactor and the conversion is measured. This measurement is used in conjunction with the models just given to decide on the corrective action necessary. The accuracy of the corrective action can only be evaluated after the batch is completed.

The main drawback of the preceding approach is the large number of experimental runs required to train such ANNs. Hundreds of data points are needed to train a network with five or six input variables. This is generally not feasible in industrial processes. If a first-principles model that can predict the "trend" of the objective function and the constraints is available, then it can be used to generate "augmented data points" for training. The next section describes how this augmented data set is created and used to train what we call a "augmented data ANN" model.

Augmented Data ANN Model

A rigorous fundamental model for the vinyl polymerization process is available in Penlidis et al. (1985). Details of the model are available in their original work, and hence are not reproduced here. This is a fairly complex model and its use for on-line control would require substantial computational resources.

Although a qualitatively sound first-principles model may not produce predictions that are accurate in absolute terms, it is usually able to predict the relative trends in process behavior. Mathematically this can be interpreted to mean that while the model predictions may not be quantitatively pre-

cise, the model captures the gradients with respect to processing variables at least qualitatively. This leads to the concept for building the hybrid ANN built using data extrapolated and interpolated from experimental points using the first-principles model. Hundreds of new points can be generated this way from a small experimental data set and the ANN can then be trained using this augmented data set.

The basic idea is illustrated in Figure 1. Let $m_1 \ldots m_M$ represent M independent variables in the process model, f represents the output variable, $(m^e_{11}, m^e_{21}, \ldots)$ $(m^e_{12}, m^e_{22}, \ldots)$, and $(m^e_{13}, m^e_{23}, \ldots)$ represent three experimental data points (dark circles in the figure). The extrapolation equation from an experimental data $(m^e_{11}, m^e_{21}, \ldots)$ to an augmented data point $(m^a_1, m^a_2, \ldots)$ (open circles in the figure) can be written as follows:

$$f^a(m^a_1, m^a_2, \ldots) = f^e(m^e_{11}, m^e_{21}, \ldots) + \left.\frac{\partial f}{\partial m_1}\right|^{\text{model}}_{m^e_{21}, \ldots} (m^a_1 - m^e_{11}) + \left.\frac{\partial f}{\partial m_2}\right|^{\text{model}}_{m^e_{11}, \ldots} (m^a_2 - m^e_{21}) + \ldots.$$

Since the mathematical model is often encapsulated in a simulation, extrapolation is done using numerically generated gradients of the output function.

Due to measurement errors and other noise, extrapolation from different experimental points will give different results. The average of these values, weighted by the proximity of experimental conditions to the input grid position (see Figure 1), is used as the training data at the particular grid point. The augmented data set can be given as:

$$f^a_i = \sum_{k=1}^{N} W_{ik} \left[f^e(m^e_{1k}, m^e_{2k}, \ldots) + \sum_{j=1}^{M} \left.\frac{\partial f}{\partial m_j}\right|^{\text{model}}_{m^e_{j'}, j' \neq j} (m^a_j - m^e_{jk}) \right], \quad (4)$$

with w_{ik} being a weighting factor that is inversely proportional to the distance of the augmented data point from the experimental data:

$$W_{ik} = \frac{\left(\sqrt{\sum_{j=1}^{M} (m^a_j - m^e_{ji})^2} \right)^{-1}}{\sum_{i'=1}^{N} \left(\sqrt{\sum_{j=1}^{M} (m^a_j - m^e_{ji'})^2} \right)^{-1}}. \quad (5)$$

Equations 4 and 5 represent a simple intuitive way of integrating experimental data and first-principle model predictions into a neural network model. There is a question of which experimental data points to use in the extrapolation. If the data points are widely scattered, then it is probably more appropriate to use only the closest experimental points for the interpolation (nearest neighbor policy). There are many possibilities here and the engineers will have to use their judgment in deciding what is best for a given application. In this study the data points were all spaced sufficiently close, and hence all of the experimental data points were employed in the generation of each augmented data point.

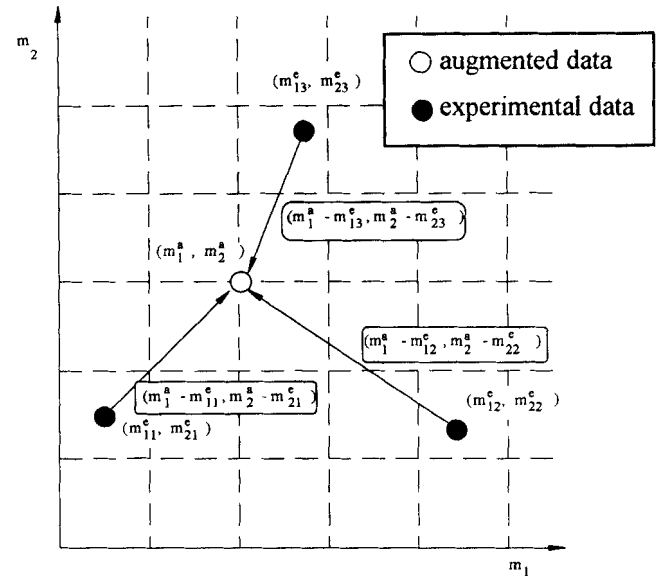


Figure 1. Extrapolation from different experimental points.

Development of ANN Models for Batch Emulsion Polymerization of Vinyl Acetate

The mechanism of batch emulsion polymerization of vinyl acetate had been extensively investigated by Lindemann (1967), Hamielec and coworkers (Friis et al., 1974; Penlidis et al., 1985, 1988), among others. As mentioned earlier, the mathematical model developed by Penlidis et al. is used in this work. This model is able to track conversion/time history of the batch process fairly accurately. However, it predicts a sharp rise in molecular weight of product at high conversion (see, e.g., Jang and Yang, 1989). Therefore, prediction of the quality of the product is not reliable.

For this application we have taken the control objective as: *achieve a desired monomer conversion (92% in our case) in the least possible time.* The manipulative actions are the reaction temperature, T, and amount of initiator added, I. The product quality constraints are number average molecular weight, M_n, and dispersity, D, the ratio of weight average to number average molecular weights ($D = M_w/M_n$). The initial conditions—monomer charge $M_0 = 400$ mL, $I_0 = 40$ mL, and $T_0 = 50$°C—are kept constant in all experiments of this study for simplicity, though this assumption is not necessary.

For a given set of I_0, M_0, and T_0, the monomer conversion depends primarily on the unknown disturbance, the amount of inhibitor present in the initial monomer charge. The monomer conversion is a monotonically decreasing function of the inhibitor at fixed time, plus it is a function of monomer soluble impurity (Penlidis et al., 1988). Therefore, an intermediate measurement of monomer conversion can be used to identify the unknown disturbance, which includes the lumped effect of the inhibitors and impurities. Ideally, the intermedi-

ate measurement should be taken as early as possible. Since conversion can only be accurately measured when it exceeds about 20% using an on-line gas chromatograph (GC), the earliest time intermediate measurement for reliable estimate of the disturbance variable is about 45 min (y_{45}).

The completion of intermediate measurement using a GC requires another 20 minutes. Therefore, manipulative action is taken at the 65th minute, with a one-shot addition of initiator I_{65} and the temperature of the reactor controlled at T_{65} throughout the later stage of the process. Since I_0, M_0, and T_0 are kept constant in this study, ANNs for predicting the end time, the number average molecular weight (M_n), and dispersity (D) have the following simplified forms:

$$t_f = \mathrm{ANN}_{t_f}(y_{45}, I_{65}, T_{65}) \tag{6}$$

$$M_n = \mathrm{ANN}_{M_n}(y_{45}, I_{65}, T_{65}) \tag{7}$$

$$D = \mathrm{ANN}_{D}(y_{45}, I_{65}, T_{65}). \tag{8}$$

An on-line density meter can be used to measure the conversion continuously and faster than an on-line GC. The availability of such an instrument will definitely improve the controllability of the process. However, in many batch processes, only a limited set of intermediate measurements is available and the situation we have described is typical and represents a more challenging control problem.

The parameters chosen for the ANN structure (number of hidden layers, number of hidden nodes, activation function, etc.) were chosen after some initial experimentation. There are no strict guidelines provided in the literature for choosing these parameters. It is possible to improve the efficiency and accuracy of the networks by fine-tuning these parameters. An ANN model with two hidden layers of 10 nodes was chosen to represent all three network models.

The following gradients $\frac{\partial t_f}{\partial T_{65}}$, $\frac{\partial M_n}{\partial T_{65}}$, $\frac{\partial D}{\partial T_{65}}$, $\frac{\partial t_f}{\partial I_{65}}$, $\frac{\partial M_n}{\partial I_{65}}$, $\frac{\partial D}{\partial I_{65}}$ were computed numerically from a computer simulation of the mathematical model. Finite difference approximation was used at each experimental data point.

Model Predictive Control of the Batch Latex Reactor

Given the preceding hybrid ANN model, the model predictive control problem can be formulated as follows:

$$\min_{I_{65}, T_{65}} \left\{\mathrm{ANN}_{t_f}(y_{45}, I_{65}, T_{65})\right\}, \tag{OP4}$$

subject to

$$0.8\times10^5 \le \mathrm{ANN}_{M_n}(y_{45}, I_{65}, T_{65}) \le 2.8\times10^5 \quad \text{(OP4.C1)}$$

$$2 \le \mathrm{ANN}_{D}(y_{45}, I_{65}, T_{65}) \le 3 \quad \text{(OP4.C2)}$$

$$0\ \mathrm{mL} \le I_{65} \le 40\ \mathrm{mL} \quad \text{(OP4.C3)}$$

$$50°\mathrm{C} \le T_{65} \le 62°\mathrm{C}. \quad \text{(OP4.C4)}$$

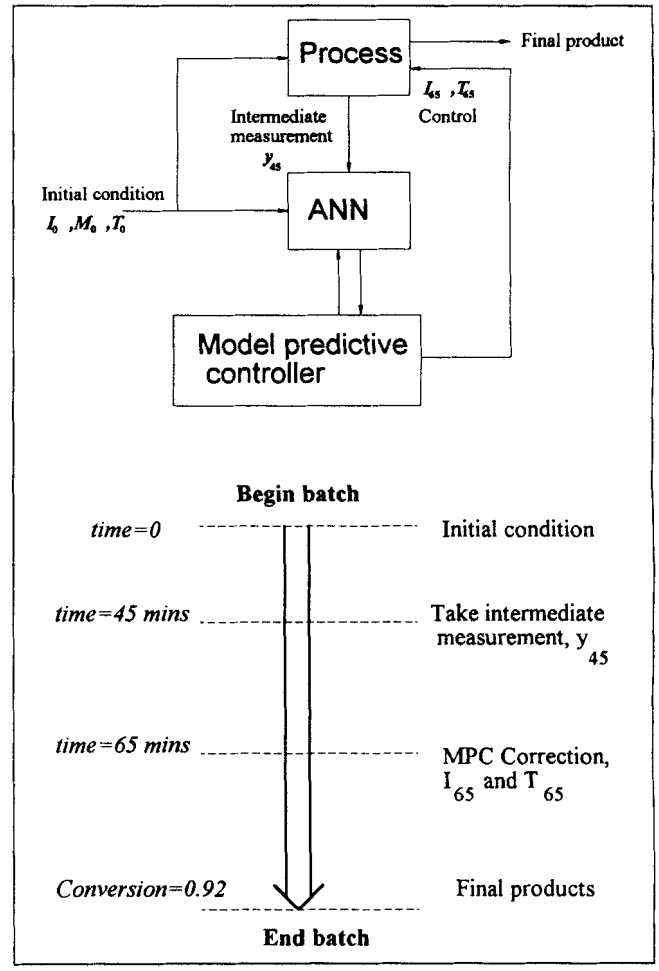


Figure 2. Control policy.

The upper and lower bounds of the molecular weight and dispersity are chosen based on typical industrial practice. The upper bound of the initiator added is due to the limitation of the electrolyte composition in a latex system. The lower bound of T_{65} is implemented because, below that temperature, the polymerization reaction is too slow. The upper bound is fixed by the limitations of the apparatus on hand.

A plot of the control policy is shown in Figure 2. As the temperature increases, the exothermic reaction rate increases. It becomes increasingly difficult to control the temperature of the reaction due to the limitation of heat removal equipment. Although experiments can be conducted up to 65°C, a more conservative limit of 62°C is used in optimal control.

This control problem can be solved using any one of numerous optimization codes. For this work we used a generalized reduced gradient (GRG) method. For a more detailed discussion of the nonlinear model predictive control algorithm and the modifications necessary for batch process control problems, see Joseph and Wang Hanratty (1993).

Experiments

The experimental setup is illustrated in Figure 3. The reactor is a 1-liter glass vessel placed in a constant temperature

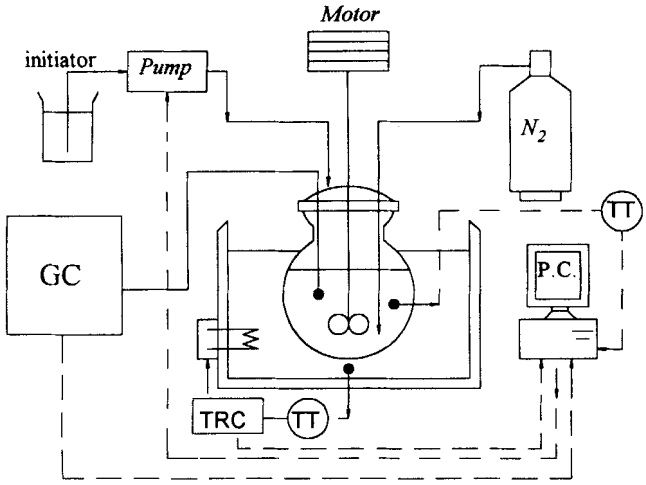


Figure 3. Experimental setup.

bath. The content is kept well mixed using a motor-driven stirrer (speed 120 rpm). Nitrogen is used to maintain the pressure inside the reactor. Initiator is added using a metering pump with a maximum capacity of 150 mL/min.

The monomer vinyl acetate used is of industrial grade. It was purified by extraction using sodium hydroxide, and washing with deionized water several times. The product was then distilled under subatmospheric pressure at 70°C. Reagent-grade sodium dodecylsulfate and potassium peroxodisulfate were used as emulsifier and initiator, respectively.

At the beginning of an experiment, fixed quantities of monomer, deionized water, and emulsifier were added to the reactor and stirred under nitrogen for about one hour until the mixture was completely emulsified. The reactor temperature was controlled initially at 50°C. Initiator was then added.

Monomer concentration was measured 45 minutes after the addition of the initiator. About 1 g of the reactor content was sampled, to which 5 g of sodium chloride solution was added to destroy the emulsion. The solution was then extracted with 12 g of dimethyl ether. Six g of the organic phase was taken out, to which 0.8 g methanol was added as the internal standard. The sample was then analyzed by a gas chromatograph using a Porapak Q column, a FID detector, and nitrogen as the carrier gas. The entire procedure takes about 20 min.

After analyzing monomer conversion at 45 min and identifying the amount of inhibitor in the charge, initiator was added again and the temperature of the reactor was raised according to the predictive control strategy. Samples were again taken at 5-min intervals, starting about 20 min before the predicted end time. Monomer concentration was measured and the actual end time (92% monomer conversion) was interpolated. The molecular weight and dispersity of the samples taken at three intervals closest to the end time were measured using gel permeation chromatography.

Analytical experimental errors in measurement of the end time, number average molecular weight, and dispersity are estimated at 10%, 5%, and 10%, respectively.

Results

In order to identify the ANN models to be used in predictive control, sixteen experiments were performed to establish the training set. The training set was designed to sample the expected range of the operating variables, T_{65} and I_{65}, reasonably well. In the training data set, y_{45} ranges from 0.24 to a high of 0.49. In addition, four sets of control experiments were used as the test set. The conditions and results of these experiments are tabulated in Table 1.

Based on the number of parameters that should be fitted in the ANN model, 1,000 augmented data sets were generated at evenly spaced input grid points to train the ANNs in Eqs. 6, 7, and 8. Back-propagation ANN (Rumelhart et al., 1986) with two hidden layers of 10 nodes were used. The ANN was trained to a precision level such that the total sum of the squared differences between ANN output and training data is less than 10^{-3}. As stated earlier, the hybrid ANN utilized the gradients computed from the theoretical model for the effects of T_{65} and I_{65}. A smaller ANN that uses a single 3-node hidden layer network, trained directly by the experimental data, was implemented to predict the "gradient" of the intermediate measurement y_{45}. This approach was found to be superior to using the theoretical model for this particular variable.

Comparison of the augmented data neural network with other models

The conventional approach to empirical modeling is to obtain a much smaller ANN, termed ANN-EXP, by fitting only experimental data. The number of nodes must be kept small to avoid overfitting the data. An ANN with one hidden layer containing three nodes was used in this study.

On the other hand, the conventional approach to modeling using prior knowledge is to adapt the theoretical model by fitting some of the key model parameters to the experimental data. In this article we refer to this third model as the *fitted theoretical model.* It is natural to assume that the polymeriza-

Table 1. Conditions and Results of all Experimental Runs

Run	y_{45}	I_{65}	T_{65}	t_f	$MW \times 10^{-5}$	D
Training Set						
1	0.46	0	60	89	2.01	2.8
2	0.33	0	67	94	2.13	2.9
3	0.30	40	67	82	1.49	2.7
4	0.28	20	50	117	2.06	2.3
5	0.33	0	50	126 [*125*]	1.64 [*1.67*]	2.7 [*2.7*]
6	0.34	40	58	96	1.98	3.7
7	0.31	40	55	105	1.60	3.3
8	0.49	0	50	118	1.74	2.6
9	0.28	0	55	116	1.83	2.4
10	0.31	40	55	102	1.69	2.6
11	0.28	40	61	93	1.48	2.9
12	0.24	40	50	125	1.73	2.6
13	0.30	30	50	123	1.87	2.5
14	0.44	20	55	114	1.89	3.0
15	0.28	0	60	102	2.09	3.1
16	0.41	30	66	88	2.10	2.9
Test Set						
17*	0.32	39	58	106 [*99*]	1.92 [*1.8*]	3.2 [*3.0*]
18*	0.28	37	57	101 [*102*]	2.00 [*1.8*]	3.2 [*3.0*]
19*	0.31	39	57	103 [*100*]	1.82 [*1.8*]	2.9 [*3.0*]
20†	0.26	40	62	93 [*96*]	1.98 [*1.7*]	3.8 [*3.7*]

* ANN model predictive control.
† Aggressive control.
[ ] Predictive values.

Table 2. Updated Kinetics Parameters

	Original Value	Updated Value
A_p	0.89×10^7	2.27×10^7
E_p	$-5{,}650$	$-5{,}563$
A_{fm}	0.43×10^6	0.41×10^6
E_{fm}	$-9{,}020$	$-11{,}146$
A_{fp}	3.55×10^6	4.04×10^6
E_{fp}	$-9{,}950$	$-9{,}168$

tion reaction and chain transfer reaction constants are incorrect and should be updated. There are six such constants:

1. A_p and E_p for the polymer propagation rate constant: $k_p = A_p \times e^{-E_p/RT}$.
2. A_{fm} and E_{fm} for the rate constant of chain transfer to monomer: $k_{fm} = A_{fm} \times e^{-E_{fm}/RT}$.
3. A_{fp} and E_{fp} for the rate constant of chain transfer to polymer: $k_{fp} = A_{fp} \times e^{-E_{fp}/RT}$.

The following optimization problem was solved to "update" these parameters:

$$\min_{\substack{A_p, A_{fm}, A_{fp}, \\ E_p, E_{fm}, E_{fp}}} \sum_{j=1}^{16} \left[\left(\frac{t_{f,j} - t^*_{f,j}}{t_{f,j}} \right)^2 + \left(\frac{M_{n,j} - M^*_{n,j}}{M_{n,j}} \right)^2 + \left(\frac{D_j - D^*_j}{D_j} \right)^2 \right], \quad \text{(OP5)}$$

subject to

$$t^*_{f,j} = f_{tf}(A_p, A_{fm}, A_{fp}, E_p, E_{fm}, E_{fp}, I_{65}, T_{65}, d) \quad \text{(OP5.C1)}$$

$$M^*_{n,j} = f_{M_n}(A_p, A_{fm}, A_{fp}, E_p, E_{fm}, E_{fp}, I_{65}, T_{65}, d) \quad \text{(OP5.C2)}$$

$$D^*_j = f_D(A_p, A_{fm}, A_{fp}, E_p, E_{fm}, E_{fp}, I_{65}, T_{65}, d) \quad \text{(OP5.C3)}$$

$$\left\{ \frac{y_{45,i} - f_{y_{45}}(A_p, A_{fm}, A_{fp}, E_p, E_{fm}, E_{fp}, I_{65}, T_{65}, d)}{y_{45,i}} \right\}^2 \le 0.01, \quad i = 1, \ldots, 16 \quad \text{(OP5.C4)}$$

The "updated" parameters are shown in Table 2.

Figure 4 shows the error in the predictions of the three models. Table 3 gives the mean and standard deviation of the correlation and prediction error of the three models. It was found that the ANN-EXP model fits the training set very well, but error for the test set is substantially higher. This is natural because there are only 16 data points in the training set, and no guarantee of successful extrapolation and interpolation is incorporated in the model development. On the other hand, the fitted theoretical model performs poorly for the training set, since only partial knowledge has been captured in the model. No matter how we adjust the key parameters in the model, agreement with data can only be limited. However, error for the test set is of the same order as that for the training set, indicating that the fitted theoretical model extrapolates and interpolates better than the ANN-EXP, as expected. The augmented data ANN achieve a correlation accuracy in the training data set that is only slightly inferior to

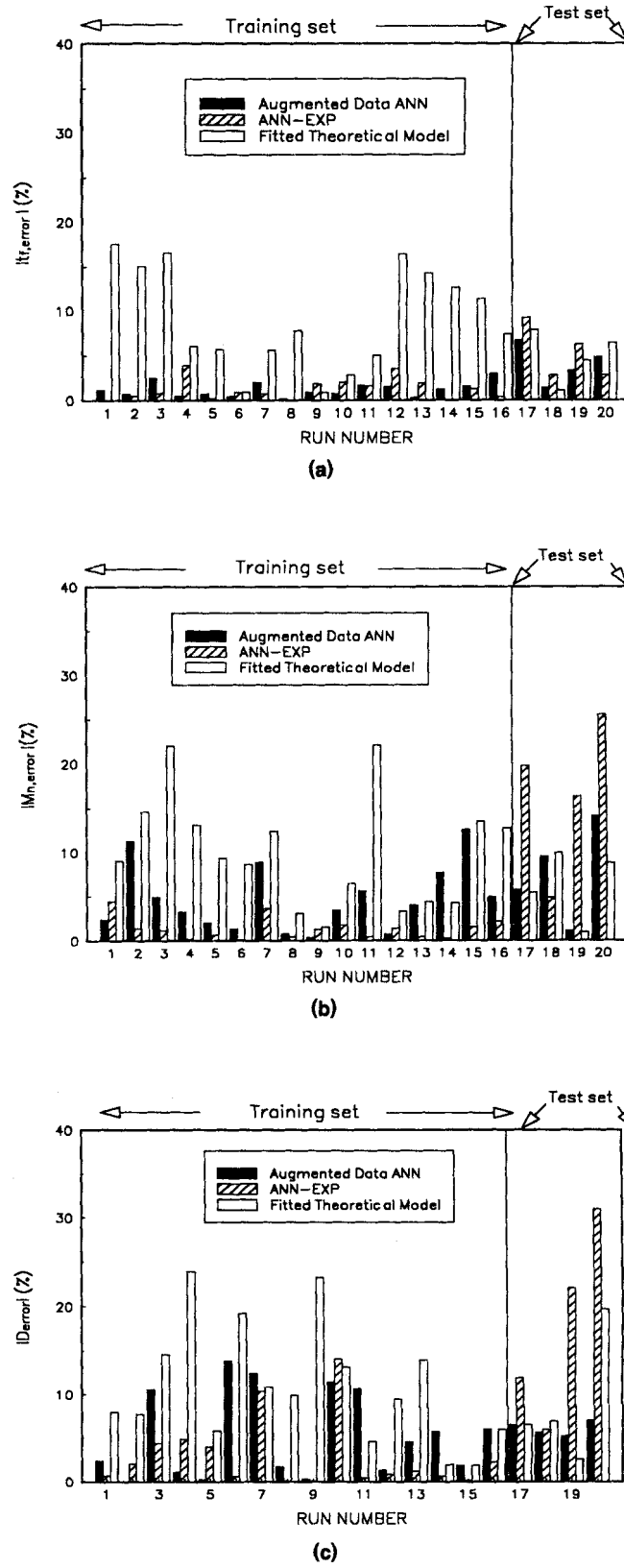


Figure 4. (a) Absolute values of percent error in predicted end time; (b) absolute values of percent error in predicted number average molecular weight; (c) absolute values of percent error in predicted dispersity.

Table 3. Mean and Standard Deviations of the Absolute Values of Percent Errors in Model Prediction

		%	Augmented Data ANN	ANN-EXP	Fitted Theoretical Model
Run 1–16 [training set]	t_f	μ	1.16	1.20	9.11
		σ	0.78	1.15	5.53
	M_n	μ	4.62	1.30	10.03
		σ	3.65	1.20	6.10
	D	μ	5.58	2.92	10.88
		σ	4.97	3.88	6.59
Run 17–20 [test set]	t_f	μ	4.04	5.25	4.92
		σ	1.95	2.70	2.54
	M_n	μ	7.64	16.64	6.27
		σ	4.80	7.58	3.51
	D	μ	5.15	17.74	8.97
		σ	1.25	9.57	6.47
Run 1–20	t_f	μ	1.74	2.01	8.27
		σ	1.60	2.27	5.35
	M_n	μ	5.23	4.37	9.28
		σ	4.09	7.09	5.87
	D	μ	5.49	5.89	10.50
		σ	4.48	8.09	6.61

μ: mean value; σ: standard deviation.

ANN-EXP. However, by using the knowledge contained in the theoretical model, the augmented data ANN is able to extrapolate and interpolate better for the test data set.

Comparison with other types of hybrid model

Two other types of hybrid ANN model proposed separately by Psichogios and Ungar (1992), that is, "the series hybrid ANN model," and Kramer et al. (1992), "the parallel hybrid ANN model," were also adapted to this system and compared with the augmented data ANN.

In the "series hybrid ANN," we implement the fitted theoretical model as the default model. We further assume that the frequency factors are functions of d, or alternatively y_{45}. A back propagation neural network (BPN) with one input node and a hidden layer of three nodes and three output nodes is used to represent this relation. The weights of this

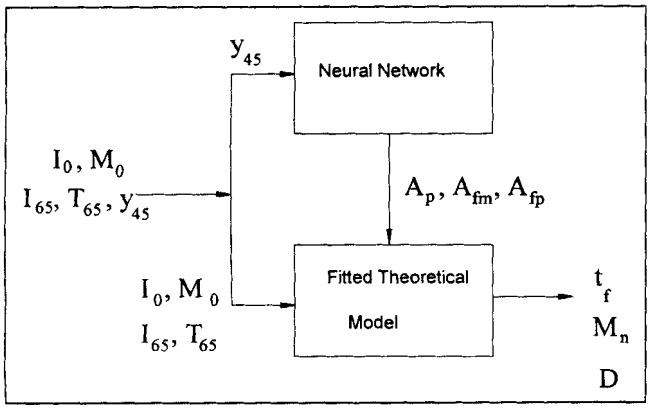


Figure 5. Series hybrid ANN.

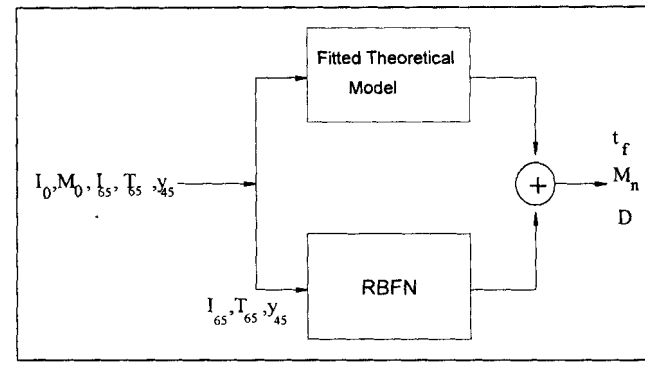


Figure 6. Parallel hybrid ANN.

neural network are obtained by fitting the training data set. This neural network is then used in series with the default model to give prediction of end-product quality as is illustrated in Figure 5. In the "parallel hybrid ANN," the residual of experimental data and default-model predictions is used to train an ANN. Radial basis transfer functions (RBFNs) are used for the ANN model to ensure that the extrapolation is dominated by the default model. The plot of the implementation of "the parallel hybrid ANN model" is given in Figure 6.

The results are shown in Table 4 and Figure 7. Although both the series and parallel hybrid ANNs improve over the fitted theoretical model, the performance is not as good as the hybrid approach using augmented data points. Both the series and parallel hybrid ANNs assume that all essential characteristics of the system have been captured by the default model, which is not correct in this example. The augmented data hybrid ANN seems to be a better reconciliation

Table 4. Mean and Standard Deviations of the Absolute Values of Percent Errors in Model Prediction Using Different Types of Hybrid ANN Models

		%	Augmented Data ANN	Series Hybrid ANN	Parallel Hybrid ANN
Run 1–16 [training set]	t_f	μ	1.16	8.39	3.20
		σ	0.78	7.48	4.26
	M_n	μ	4.62	7.78	3.58
		σ	3.65	6.51	3.78
	D	μ	5.58	11.15	4.46
		σ	4.97	5.55	4.82
Run 17–20 [test set]	t_f	μ	4.04	4.58	3.01
		σ	1.95	1.60	2.66
	M_n	μ	7.64	5.03	16.26
		σ	4.80	2.80	6.85
	D	μ	5.15	13.01	14.70
		σ	1.25	9.67	8.20
Run 1–20	t_f	μ	1.74	7.63	3.16
		σ	1.60	6.90	3.99
	M_n	μ	5.23	7.23	6.12
		σ	4.09	6.06	6.82
	D	μ	5.49	11.52	6.51
		σ	4.48	6.63	6.99

μ: mean value; σ: standard deviation.

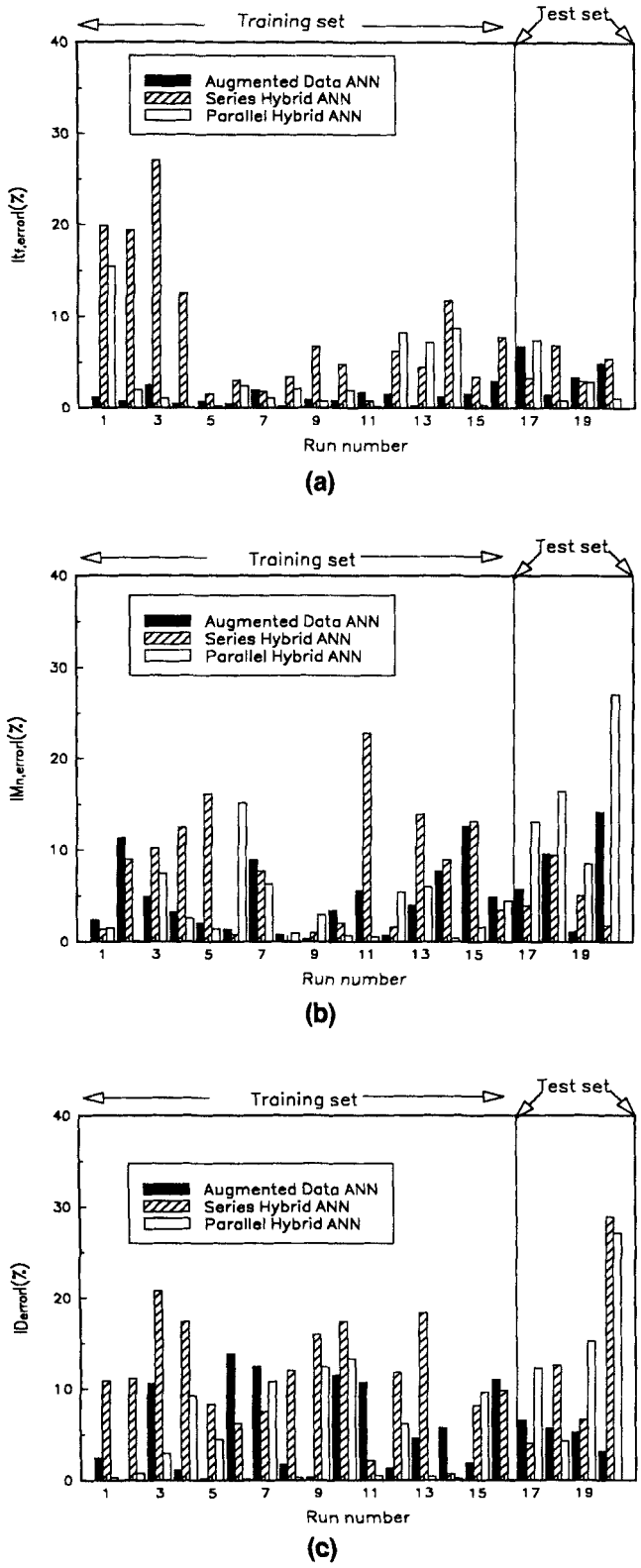


Figure 7. (a) Absolute values of percent error in predicted end time with different hybrid ANN models; (b) absolute values of percent error in predicted number average molecular weight with different hybrid ANN models; (c) absolute values of percent error in predicted dispersity with different hybrid ANN models.

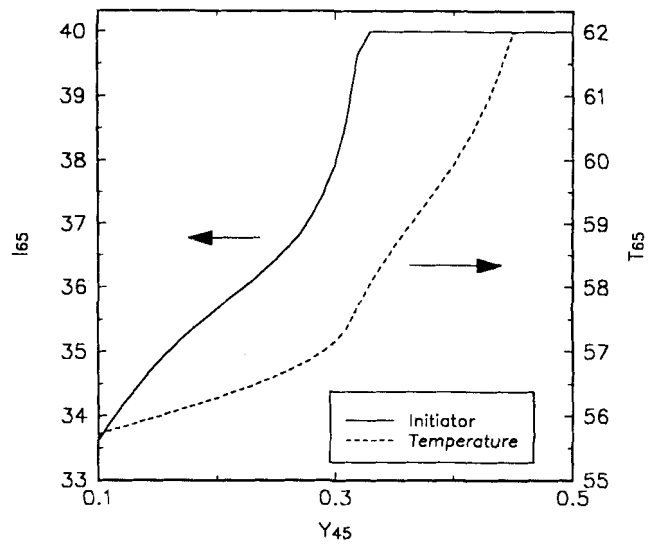


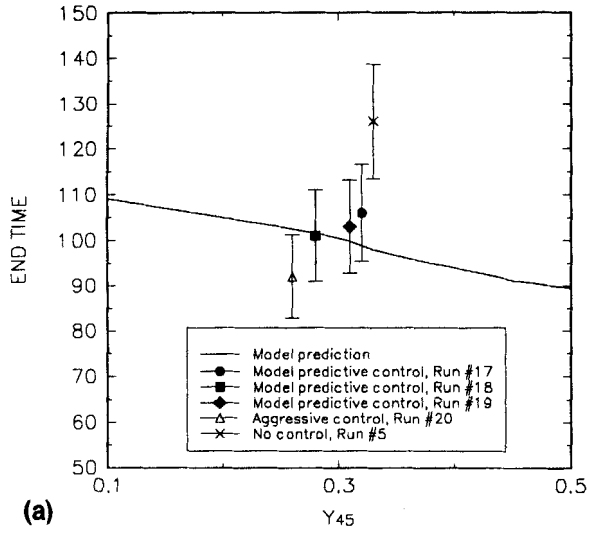
Figure 8. Control actions vs. intermediate measurement.

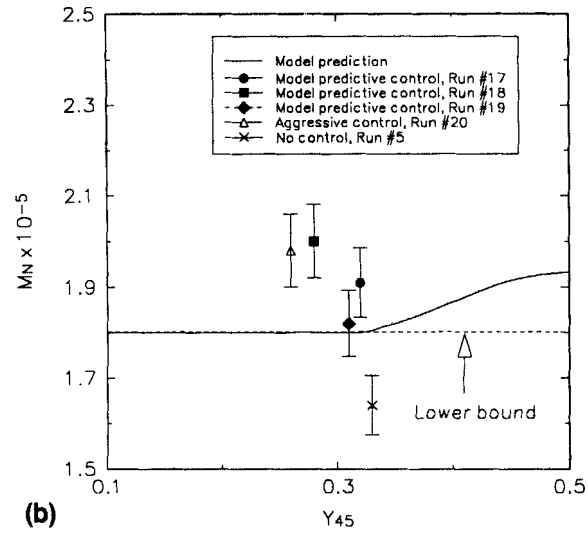
between empirical data and theoretical knowledge when both abundance of data and adequate prior knowledge are lacking.

Model predictive control of latex reactor

In this work, a one-shot policy, as shown in Figure 2, is implemented to perform the model predictive control. It is possible to obtain the control profiles of T_{65} and I_{65} as functions of y_{45} based on the hybrid ANN model described earlier by solving the model predictive problem OP4. The result is shown in Figure 8, which shows that the control actions can be more aggressive to shorten the batch time if the intermediate measurement shows that the amount of feed impurity is small (high conversion), and the product qualities can still be guaranteed as shown in Figure 9. In polymerization reaction the dispersity of the polymer product will be larger with higher reactor temperature and higher concentration of free radicals (initiator). Figure 8 shows that if y_{45} is low, that is, inhibitor is high, then the control action I_{65} and T_{65} cannot be aggressive to shorten the batch time since the constraint on dispersity will be violated (see also Figure 9c).

Based on the results shown in Figure 8, three test batch runs (runs 17, 18 and 19 in Table 1) were carried out to evaluate the predictions of the hybrid ANN model. The results are given in Figure 9. It shows that the control actions taken do shorten the batch time while maintaining the product quality within its tolerances. If one simply wants to shorten the batch time and ignore the existence of the impurity by setting the control actions into their upper bounds (62°C for T_{65} and 40 mL for I_{65}, called *aggressive control*), the upper bound on the dispersity of the polymer product is violated as predicted by the model. This is confirmed by run 20. However, if no predictive control action is taken, and a fixed recipe is followed, the batch time can be very long and the number average molecular weight of the polymer product is out of the lower bound as shown in Figure 9 and Table 1 (run 5).





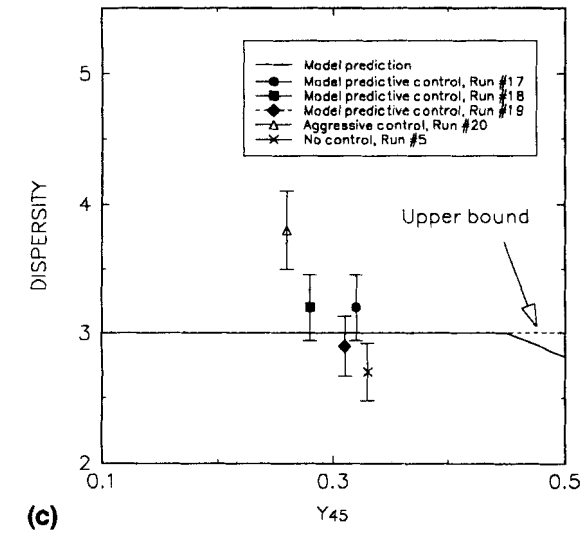


Figure 9. (a) End time: Experimental and simulation results; (b) number average molecular weight: experimental and simulation results; (c) dispersity: experimental and simulation results.

Conclusion

A model predictive control scheme for batch processes using an augmented data hybrid ANN has been developed and applied to the latex polymerization process. The control strategy uses this augmented data hybrid ANN to identify the unknown and unmeasured disturbances in the initial charge of the batch. In the application to the batch polymerization of vinyl acetate, the amount of inhibitor present in the monomer charge was estimated by measuring monomer conversion. Based on the information of the intermediate measurement, the augmented data hybrid ANN is used to predict and hence control the batch: end time and the molecular weight and dispersity of the final polymer product.

The advantages of using the ANN-based predictive control strategy are:

1. The effect of unmeasured disturbances can be incorporated easily using intermediate measurements.
2. The ANN structure allows us to capture nonlinear relationships among the variables.
3. The predictive capability of the model allows us to correct for changing raw material properties and output specifications.

The augmented data hybrid ANN is trained using a combination of actual experimental data and augmented data points generated through extrapolation and interpolation using gradient information from the theoretical model. This procedure of augmenting the data allows development of ANN models when experimental data are scarce and expensive to obtain.

The augmented data hybrid ANN was also shown to be superior in terms of its correlative capability to other hybrid ANN approaches. It also performed better than a theoretical model whose parameters were fitted to the experimental data.

Although we restricted ourselves to specific types of measurements, models, and control actions, the methodology presented is quite general and can be applied to the modeling and control of other batch processes where product quality control is difficult to achieve using conventional techniques and algorithms.

Acknowledgment

Financial support provided by the National Science Council, Taiwan, through grant NSC82-0402-E007-249, and by the National Science Foundation, USA-Taiwan Cooperative Research, through grant INT-93-11390, are gratefully acknowledged.

Notation

A_{fm} = frequency factor of the rate constant of chain transfer to monomer
A_{fp} = frequency factor of the rate constant of chain transfer to polymer
A_p = frequency factor of the polymer propagation rate constant
D_{error} = relative error of the predictive dispersity
D_j = dispersity of the final product at run j
D_j^* = predictive dispersity of the final product at run j
E_{fm} = activation energy of the rate constant of chain transfer to monomer
E_{fp} = activation energy of the rate constant of chain transfer to polymer
E_p = activation energy of the polymer propagation rate constant
f^a = augmented data function
f^e = experimental result
h = intermediate states of the system

k_{fm} = rate constant of chain transfer to monomer
k_{fp} = rate constant of chain transfer to polymer
k_p = polymer propagation rate constant
$M_{n,\text{ error}}$ = relative error of the predictive number average molecular weight
$M_{n,j}$ = number average molecular weight of the final product at run j
$M^*_{n,j}$ = predictive number average molecular weight of the final product at run j
N = number of experimental data points
t = time
$t_{f,\text{ error}}$ = relative error of the predictive end time
$t_{f,j}$ = end time of run j
$t^*_{f,j}$ = predictive end time of run j
$\boldsymbol{x}_0$ = initial state variables
y_{45} = monomer conversion at 45th minute
$y_{45,j}$ = monomer conversion at 45th minute at run j
$y^*_{45,j}$ = predictive monomer conversion at 45th minute at run j

Literature Cited

Bhat, N., and T. J. McAvoy, "Use of Neural Nets for Dynamic Modeling and Control of Chemical Process System," *Comput. Chem. Eng.*, **14**, 573 (1990).

Friis, N., D. Goosney, J. D. Wright, and A. E. Hamielec, "Molecular Weight and Branching Development in Vinyl Acetate Emulsion Polymerization," *J. Appl. Poly. Sci.*, **18**, 1247 (1974).

Hernandez, E., and Y. Arkun, "Study of the Control Relevant Properties of Back Propagation Neural Network Models of Nonlinear Dynamical Systems," *Comput. Chem. Eng.*, **16**, 227 (1992).

Hoskins, J. C., and D. M. Himmelblau, "Artificial Neural Network Models of Knowledge Representation in Chemical Engineering," *Comput. Chem. Eng.*, **12**, 881 (1988).

Jang, S. S., and W. L. Yang, "Dynamic Optimization of Batch Emulsion Polymerization of Vinyl Acetate—An Orthogonal Polynomial Initiator Policy," *Chem. Eng. Sci.*, **44**, 515 (1989).

Joseph, B., and F. Wang Hanratty, "Predictive Control of Quality in a Batch Manufacturing Process Using Artificial Neural Network Models," *Ind. Eng. Chem. Res.*, **23**, 1951 (1993).

Klein, R. A., "Achieve Total Quality Management," *Chem. Eng. Process*, **87**, 83 (1991).

Kozub, D., and J. F. MacGregor, "State Estimation for Semi-Batch Polymerization Reactors," *Chem. Eng. Sci.*, **47**(5), 1047 (1992).

Kramer, M. A., and J. A. Leonard, "Diagnosis Using Back Propagation Neural Networks—Analysis and Criticism," *Comput. Chem. Eng.*, **14**, 1323 (1990).

Kramer, M. A., M. L. Thompson, and P. M. Bhagat, "Embedding Theoretical Model in Neural Networks," *ACC/WA14*, 475 (1992).

Levinson, W. A., "Make the Most of Control Charts," *Chem. Eng. Process*, **88**, 86 (1992).

Lindemann, M. A. "The Mechanism of Vinyl Acetate Polymerization," in *Kinetics and Mechanisms of Polymerization*, Vol. 1, *Vinyl Polymerization*, Part I, Marcel Dekker, New York, p. 207 (1967).

Pao, Y., *Adaptive Pattern Recognition and Neural Networks*, Addison-Wesley, Reading, MA (1989).

Penlidis, A., J. F. MacGregor, and A. E. Hamielec, "A Theoretical and Experimental Investigation of the Batch Emulsion Polymerization of Vinyl Acetate," *Polym. Process Eng.*, **13**, 185 (1985).

Penlidis, A., J. F. MacGregor, and A. E. Hamielec, "Effect of Impurities on Emulsion Polymerization: Case I Kinetics," *J. Appl. Poly.*, **35**, 2023 (1988).

Psichogios, D. C., and L. H. Ungar, "A Hybrid Neural Network—First Principles Approach to Process Modeling," *AIChE J.*, **38**, 1499 (1992).

Psichogios, D. C., and L. H. Ungar, "Direct and Indirect Model Based Control Using Artificial Neural Networks," *Ind. Eng. Chem. Res.*, **30**, 2564 (1991).

Ripley, B. D., "Statistical Aspects of Neural Networks," *Networks and Chaos: Statistical and Probabilistic Aspects*, Chapman & Hall, London, p. 40 (1993).

Rumelhart, D. E., G. E. Hinton, and R. J. Willams, "Learning Internal Representation by Error Propagation," *Parallel Distributed Processing*, Vol. 1, MIT Press, Cambridge, MA, p. 318 (1986).

Sarle, W. S., "Neural Networks and Statistical Model," *Proc. of the SAS Users Group Int. Conf.*, Dallas, TX (Apr., 1994).

Thompson, M. L., and M. A. Kramer, "Modeling Chemical Processes Using Prior Knowledge and Neural Networks," *AIChE J.*, **40**, 1328 (1994).

Venkatasubramanian, V., and K. A. Chan, "A Neural Network Methodology for Process Fault Diagnosis," *AIChE J.*, **35**, 1993 (1989).

Whitely, J. R., and J. Davis, "Knowledge-Based Interpretation of Sensor Patterns," *Comput. Chem. Eng.*, **16**, 329 (1992).

Manuscript received Aug. 22, 1994, and revision received Apr. 7, 1995.