Use of Hammerstein Models in Identification of Nonlinear Systems

Esref Eskinat and Stanley H. Johnson
Dept. of Mechanical Engineering, Lehigh University, Bethlehem, PA 18015

William L. Luyben
Dept. of Chemical Engineering, Lehigh University, Bethlehem, PA 18015

The utility of the Hammerstein model, which is composed of a static nonlinear element in series with a linear dynamic part, was investigated to represent the dynamics of nonlinear chemical processes. Different methods to identify the parameters of Hammerstein models were tested. The methods were applied to the identification of simulated distillation columns and to an experimental heat exchanger process. The results show that the dynamics of such processes can be better represented by Hammerstein-type models than by linear models.

Introduction

Nonlinear behavior is the rule, rather than the exception, in the dynamic behavior of physical systems. Most physical devices have nonlinear characteristics outside a limited linear range. In most chemical processes, understanding the nonlinear characteristics is important for designing controllers that regulate the process. Chemical processes contain valves that saturate, connecting lines whose time delays vary with flow rate, reacting mixtures which obey a power law, and separation units which are very sensitive to input changes and disturbances. Besides those nonlinearities, there are phenomena such as incomplete mixing whose effects are not well understood. The combination of such effects can sometimes cause unpredictable behavior of a process. Even though most controllers used in process control are of the PID type, the understanding of process nonlinearities can be utilized to select manipulated variables, to locate sensors, and to use special effects such as antireset windup schemes, gain scheduling and override control.

One approach to understand the nonlinear behavior is to form a mathematical model of the process. To achieve this, a mathematical model of each unit operation has to be formed by making some simplifying assumptions, and then these models are combined to obtain a model representing the complete system. This approach has been utilized in modeling many processes and is a very useful one. However, there are some limitations to the mathematical modeling route:

- The assumptions made in deriving the model can be too restrictive or not very realistic. Therefore, the model might not capture the essentials of the behavior of the process.
- The parameters in the models such as overall heat transfer coefficients and reaction constants may not be accurate.
- The models derived can be very large: they can be made up of many differential equations. This limits the use of such models to derive analytical insights.

Another approach commonly used is to build the model directly from the observed behavior of the process itself. This kind of empirical model building is called system identification. It might be used along with mathematical modeling in some cases (i.e., estimating the parameters, such as heat transfer coefficients, from empirical data). However, the bulk of the work done in the field of system identification starts with representing the process as a black box. We may have access to the inputs and outputs, but the internal mechanisms are assumed to be totally unknown to us. The problem in system identification is to construct a model which would mimic the 'inner mechanisms' of the system, using the input/output data. The usual procedure is to select a model structure with some unknown parameters and then to estimate the model parameters. The last step is to check whether the model obtained is adequate. Some iterations on this procedure might be necessary to arrive at a model which is good enough for one's purposes.

The crucial step in the outlined method is the selection of the model structure. The vast majority of the work done on system identification is for linear model structures. This re-

Correspondence concerning this article should be addressed to S. H. Johnson.

striction is justified very well because linear systems are a specific class among the dynamic systems (which are nonlinear in general) for which the superposition principle is valid. This principle allows us to develop a very elegant and powerful theory for linear systems, which is not valid for nonlinear systems in general. Still, there is a need to obtain models of systems which behave nonlinearly. Some researchers have worked on the problem of constructing a nonlinear model from input/output measurements. The results they have obtained are usually specific to a certain model structure. A survey of those attempts is given by Billings (1980).

To simplify the general problem of identifying a nonlinear model from input/output data, model structures which contain a static nonlinearity in series with a linear dynamic system have been a subject of investigation. For such models two kinds of structure may be used, depending on the location of the static nonlinear (NL) element. Unlike linear systems, different locations for the static NL element give rise to different responses. When the NL element precedes the linear block, it is called the Hammerstein model, whereas if it follows the linear block, it is called the Wiener model. Our attention in this article will be focused on Hammerstein models.

A number of papers in the literature explore the identification of nonlinear systems using the Hammerstein model. Narendra and Gallman (1966) seem to be the first researchers who dealt with this problem. In this article, parameters of the Hammerstein model are obtained by separating the linear and static NL parts which are estimated iteratively. Then a number of papers which treat the Hammerstein model as a MISO linear model extended linear identification methods to such models. Chang and Luus (1971) used a simple least squares technique to estimate the system parameters. This method is later extended to include the colored noise effects (Haist et al., 1973). Also, Hsia (1976) has proposed a noniterative multistage least squares scheme which takes into account the presence of the colored output noise. A comparison of the simple least squares estimation vs. the Narendra-Gallman method (1966) is given by Gallman (1976). Haber (1979, 1988) used a combination of correlation analysis and least squares method to obtain the parameters of the Hammerstein models. Correlation analysis which estimates the model nonparametrically (i.e., impulse response coefficients for the linear part and a set of data points for the nonlinear part) was used by Billings and Fakhouri (1978). They mentioned that the method requires a white noise input, and therefore it is somewhat restrictive. Nonparametric identification using kernel regression estimate was considered by Greblicki and Pawlak (1987). Instrumental variable methods were used to identify Hammerstein models in (Bamberger, Isermann, 1978). The use of instrumental variable methods was further investigated by Stoica and Soderstrom (1982). Identification of Hammerstein models in the closed loop was investigated by Beyer et al. (1979). On-line identification of Hammerstein models was treated by Kortmann and Unbehauen (1987).

Most of the examples in the above papers were on estimating the parameters of a system which is already described by a Hammerstein model. Applications to real processes are rare.

Hammerstein Models

The Hammerstein model of a nonlinear system is shown in Figure 1. The static NL element scales the input $u(t)$ and

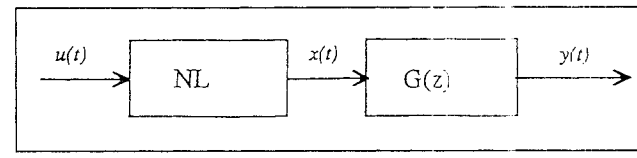


Figure 1. Hammerstein model.

transforms it to $x(t)$, and the dynamics are modeled by a linear discrete transfer function $G(z^{-1})$, whose output is $y(t)$. The Hammerstein model models the NL effects as an input-dependent gain nonlinearity. The slope of the nonlinearity at a certain operating point is the instantaneous gain of the system.

If the static NL function is assumed to be approximated by a finite polynomial expansion, the Hammerstein model can be described by the following equations:

$$\begin{aligned} y(t)+a_1y(t-1)+\ldots a_ny(t-n) &= b_1x(t-1)+b_2x(t-2)+\ldots+b_nx(t-n) \\ x(t) &= \gamma_1u(t)+\gamma_2u^2(t)+\ldots+\gamma_mu^m(t) \end{aligned} \tag{1}$$

The intermediate variable $x(t)$ cannot be measured, but it can be eliminated from the equations. Substituting the second equation into the first we get:

$$\begin{aligned} y(t)+a_1y(t-1)+\ldots+a_ny(t-n) &= b_1\gamma_1u(t)+\ldots+b_1\gamma_mu^m(t) \\ &\quad +\ldots+b_n\gamma_1u(t-n)+\ldots+b_n\gamma_mu^m(t-n) \end{aligned} \tag{2}$$

We can write the above equation in operator form as:

$$y(t)=\frac{B(q^{-1})}{A(q^{-1})}\sum_{i=1}^{m}\gamma_iu^i(t) \tag{3}$$

where the polynomials $A(q^{-1})$ and $B(q^{-1})$ are:

$$\begin{aligned} A(q^{-1}) &= 1+a_1q^{-1}+\ldots+a_nq^{-n} \\ B(q^{-1}) &= b_1q^{-1}+\ldots+b_nq^{-n} \end{aligned} \tag{4}$$

The orders of the polynomials are assumed to be the same, but this needs not be the case in general.

One way of identifying the Hammerstein-type models from input/output data is viewing the Hammerstein model as a multiinput/single-output (MISO) linear system. We can write Eq. 2 as:

$$\begin{aligned} y(t)+a_1y(t-1)+\ldots+a_ny(t-n) &= c_{11}u(t-1)+\ldots+c_{1m}u^m(t-1) \\ &\quad +\ldots+c_{n1}u(t-n)+\ldots+c_{nm}u^m(t-n) \end{aligned} \tag{5}$$

where

$$c_{ij}=b_i\gamma_j \tag{6}$$

This can be considered as a MISO system with inputs $u(t)$, $u^2(t),\ldots,u^m(t)$. Parameters of Eq. 5, a_i, c_{ij}, can be determined

from experimental data by applying the known identification methods for MISO linear models (Chang and Luus, 1971). There is a redundancy in the parameters of the Hammerstein model. Since the static gain of the Hammerstein model is determined by the static nonlinearity, the gain of the linear part is redundant and can be normalized to 1. Alternatively, the parameter γ_1, which determines the slope of the static nonlinearity at the origin, may be set to 1. In the latter case, the steady-state gain at the origin will be determined by the steady-state gain of the linear part and the NL function will act as a scaling relationship. This is easier than setting the steady-state gain of the linear part to 1, therefore we shall assume:

$$\gamma_1 = 1 \tag{7}$$

At first sight, it seems that Eq. 6 can be used to obtain b_i and γ_j, once the a_i's and c_{ij}'s are determined from the experimental data. From Eq. 6 and Eq. 7:

$$b_i = c_{i1} \quad i = 1,\ldots,n \tag{8}$$

Also, we can obtain n solutions for each γ_j from the above equations, which are given by:

$$\gamma_j^{(i)} = \frac{c_{ij}}{c_{i1}} \quad j = 1,\ldots,m \quad i = 1,\ldots,n \tag{9}$$

Because of experimental errors and noise and process-model mismatch, the values of $\gamma_j^{(i)}$ obtained from Eq. 9 will be different for each i. To obtain unique solutions for γ_j, two methods have been proposed in the literature. The n solutions obtained for $\gamma_j^{(i)}$ from Eq. 9 may be averaged, which gives:

$$\gamma_j = \frac{1}{n}\sum_{j=1}^{n}\gamma_j^{(i)} \tag{10}$$

Another approach is to select the $\gamma_j^{(i)}$ values which minimize the error between the measured output and the model output in a least squares sense (Chang and Luus, 1971).

With the methods explained above, however, it is not possible to extract the original parameters b_i and γ_j consistently from the knowledge of the estimated parameters c_{ij}. The values of b_i and γ_j obtained from the above-mentioned methods will not give back the original c_{ij} values obtained from the parameter estimation algorithm. Two different approaches are possible to avoid this inconsistency. We can either impose constraints on the parameters c_{ij} such that Eq. 9 gives the same value for $\gamma_j^{(i)}$, for all i, or estimate the parameters b_i and γ_j separately. In this work, the latter approach is followed because it is more direct and the formulation of the parameter estimation problem is simpler with this approach.

Parameter Estimation Methods for Hammerstein Models

Estimation of the parameters of the Hammerstein model separately (i.e., a_i, b_i and γ_j) is possible because of a special property of the Hammerstein model structure. Although the parameters enter the model nonlinearly, (b_i and γ_j multiplied) the predictor of $y(t)$ comes out to be linearly dependent on y and u^i. Therefore, it is possible to construct a predictor of $y(t)$ and the methods called prediction error methods are applicable. The first two algorithms used in this section depend on the prediction error formulation developed by Ljung (1987). The third algorithm also uses a prediction error method to estimate the parameters of the linear part, but the parameters of the NL part are estimated separately.

Prediction error methods

We shall assume that the output of the Hammerstein model is given by:

$$y(t) = \frac{B(q^{-1})}{A(q^{-1})}\sum_{i=1}^{m}\gamma_i u^i(t) + e(t) \tag{11}$$

where $e(t)$ is assumed to be independent (white) noise used to model the disturbances. Disturbances may also be modeled by filtered white noise, in which case the disturbances modeled by the $e(t)$ term have to be augmented with a linear filter. The parameters describing the noise filter have to be identified in this case. This would increase the complexity of the problem and is therefore omitted at this stage. The predictor of $y(t)$ is given by:

$$\bar{y}(t,\theta) = \frac{B(q^{-1})}{A(q^{-1})}\sum_{i=1}^{m}\gamma_i u^i(t) \tag{12}$$

To obtain the model parameters:

$$\theta = [a_1,\ldots,a_n,b_1,\ldots,b_n,\gamma_1,\ldots,\gamma_m]^T$$

we define the criterion:

$$V(\theta) = \frac{1}{N}\sum_{t=1}^{N}\epsilon^2(t,\theta) \tag{13}$$

N is the number of data points collected and $\epsilon(t,\theta)$ is the prediction error given by:

$$\epsilon(t,\theta) = y(t) - \bar{y}(t,\theta) \tag{14}$$

The model parameters θ can be obtained by minimizing the criterion given in Eq. 13. For the Hammerstein model, the derivative of the criterion is not a linear function of the parameters, so an analytical solution cannot be found. Therefore, iterative methods must be used to obtain a solution. The iteration steps for the parameters can be expressed as:

$$\theta_{k+1} = \theta_k + \delta H(\theta_k)^{-1} g(\theta_k) \tag{15}$$

where

δ = step length
$H(\theta)$ = Hessian of $V(\theta)$ or its approximation
$g(\theta)$ = Gradient of $V(\theta)$

For the Hammerstein model, the gradient can be computed as:

$$g(\theta)=\frac{dV}{d\theta}=\frac{2}{N}\sum_{i=1}^{N}\epsilon(t,\theta)\frac{\partial\epsilon}{\partial\theta} \tag{16}$$

The derivatives $\partial\epsilon/\partial\theta$ in the above equation can be computed from Eqs. 14 and 12 as:

$$\frac{\partial\epsilon(t)}{\partial a_j}=-\frac{1}{A(q^{-1})}\sum_{i=1}^{m}\gamma_i\frac{B(q^{-1})}{A(q^{-1})}u^i(t-j)$$

$$\frac{\partial\epsilon(t)}{\partial b_j}=\frac{1}{A(q^{-1})}\sum_{i=1}^{m}\gamma_i u^i(t-j)$$

$$\frac{\partial\epsilon(t)}{\partial\gamma_i}=\frac{B(q^{-1})}{A(q^{-1})}u^i(t) \tag{17}$$

Since the Hessian is difficult to compute, it is usually approximated by another function. The most common method, Levenberg-Marquardt method, approximates the Hessian of the quadratic criterion to be minimized with:

$$H(\theta)\approx\frac{1}{N}\sum_{t=1}^{N}\frac{\partial\epsilon(t,\theta)}{\partial\theta}\frac{\partial\epsilon(t,\theta)^T}{\partial\theta}+\mu I \tag{18}$$

As discussed previously, there is a redundancy in the parameters of the Hammerstein model. Since the steady-state gain of the model at the origin (i.e., $u=0$) can be expressed by the steady-state gain of the linear part, the parameter γ_1, which determines the slope of the NL part at the origin, can be normalized to 1. If this condition is not imposed, the redundancy in the parameters will show itself in the parameter estimation. In this case, if the number of parameters to be estimated is p, the Hessian will have a rank of $p-1$. The singularity in the Hessian makes the parameter estimation problem ill-conditioned, in which case the estimates of the parameters, if they can be obtained, will not be reliable. Therefore, the prediction error method was implemented by setting $\gamma_i=1$.

So, the prediction error method for the Hammerstein models can be summarized as follows:

Algorithm 1: Prediction Error Method

1. Start iterations with an initial estimate of parameters θ_0, and set $\gamma_1=1$.
2. Compute the prediction of the model and $V(\theta_k)$.
3. Pick a small value for μ (a typical choice is 0.0001).
4. Compute the gradient and the Hessian through Eqs. 17 and 18.
5. Update parameter estimates through Eq. 15, and calculate the new $V(\theta_k)$.
6. If $V(\theta_k)<V(\theta_{k-1})$, update solution and *decrease* μ by a factor (say) 10 and go to step 2.
7. If $V(\theta_k)>V(\theta_{k-1})$, *increase* μ by a factor (say) 10 and go to step 4.
8. Stop when the norm of the gradient is below a certain value or when the maximum number of iterations is reached.

The prediction error algorithm may also be implemented in a recursive manner. The recursive prediction error algorithm is defined by the equations (Kortmann and Unbehauen, 1987; Ljung and Soderstrom, 1983):

$$\theta(t)=\theta(t-1)+L(t)\epsilon(t) \tag{19}$$

$$\epsilon(t)=y(t)-\bar{y}(t) \tag{20}$$

$$L(t)=P(t-1)\Phi(t)[1+\Phi^T(t)P(t-1)\Phi(t)]^{-1} \tag{21}$$

$$P(t)=\frac{1}{\rho(t)}[P(t-1)-L(t)\Phi^T(t)P(t-1)] \tag{22}$$

$$\Phi(t)=\frac{\partial\epsilon(t)}{\partial\theta} \tag{23}$$

$\rho(t)$ is the forgetting factor, which is updated according to:

$$\rho(t)=\rho(t-1)+[1-\rho(t-1)]\Delta\rho \tag{24}$$

So, the recursive prediction error algorithm is as follows:

Algorithm 2: Recursive Prediction Error Method

1. Select a starting value for P (e.g., a diagonal matrix with large entries), ρ (i.e., 0.999), and θ_0 (set $\gamma_1=1$).
2. For $t=1$ to N, update the parameters θ according to Eqs. 19-23.

Narendra-Gallman algorithm

The two algorithms described above are based on the prediction error formulation. In the literature there is another algorithm which obtains the estimated parameters of Hammerstein models. The algorithm proposed by Narendra, Gallman (1966) obtains the parameters of the Hammerstein model by separating the estimation problem of the linear and static NL parts. It is based on the following observation:

Suppose the parameters of the linear part (a_i and b_i) are known. Then, we can solve for the parameters of the nonlinear part γ_i. Differentiating the loss function $V(\theta)$ given in Eq. 13 with respect to γ we get:

$$\frac{\partial V}{\partial\gamma}=\frac{2}{N}\sum_{t=1}^{N}\frac{B(q^{-1})}{A(q^{-1})}U(t)\left\{y(t)-\frac{B(q^{-1})}{A(q^{-1})}U^T(t)\gamma\right\} \tag{25}$$

where

$$U(t)=[u(t),u^2(t),\ldots,u^m(t)]^T$$

$$\frac{\partial V}{\partial\gamma}=\left[\frac{\partial V}{\partial\gamma_1}\cdots\frac{\partial V}{\partial\gamma_m}\right]^T$$

Equating Eq. 25 to zero, we can get the solution for γ as follows:

$$\gamma=\left[\sum_{t=1}^{N}\frac{B(q^{-1})}{A(q^{-1})}U(t)\frac{B(q^{-1})}{A(q^{-1})}U^T(t)\right]^{-1}\times\left[\sum_{t=1}^{N}\frac{B(q^{-1})}{A(q^{-1})}U(t)y(t)\right] \tag{26}$$

If the nonlinear part is "known," the input to the linear part $x(t)$ can be computed from:

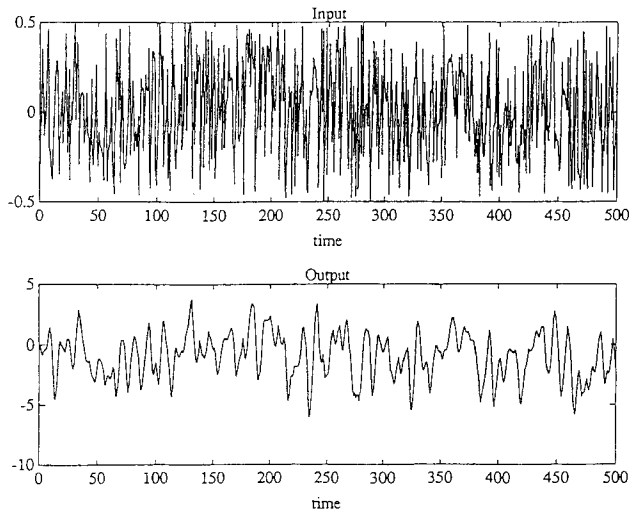


Figure 2. Input/output data for nonlinear system.

$$x(t) = \gamma_1 u(t) + \gamma_2 u^2(t) + \ldots + \gamma_m u^m(t) \qquad (27)$$

Using this input, we can improve the linear part of the model. So, this algorithm can be summarized as follows:

Algorithm 3: Narendra-Gallman Method

1. Calculate a linear model using the data in exactly the same manner as in algorithm 1. (All the parameters of the NL part except γ_1 are set to zero, γ_1 is set to 1.)
2. Solve for the nonlinear part using Eq. 26.
3. Calculate the new input for the linear part using Eq. 27 and go to 1.
4. Stop when convergence occurs.

Comparison of the algorithms

Three algorithms explained above were tested on an example with different noise levels added to the output. In this section, the results obtained from different algorithms will be compared based on these test cases.

The following examples were used to test the methods:

$$G(z^{-1}) = \frac{z^{-1} + 0.5z^{-2}}{1 - 1.5z^{-1} + 0.7z^{-2}}$$

$$N(u) = 5u^3 - 2u^2 + u$$

All three algorithms were tried on this example for both noisy and nonnoisy cases. The example model was simulated using uniform noise input varying between -0.5 and 0.5, and 500 data points were collected. The input and output data are plotted in Figure 2. Two different noise levels were tried besides the nonnoisy (NN) case. The standard deviations of the white noise disturbances that were added to the system output were selected as 0.1 (N1) and 1 (N2). N1 represents low-level noise (approximately 5% of output variance) and N2 is high-level noise (approximately 50% of output variance). The initial estimates of the parameters were taken as zero for algorithm 2, except for γ_1, which was set to 1. For the other two algorithms, the results of a linear estimation were taken as initial values for parameters estimates. The results of the parameter estimations are given in Table 1.

As can be observed from Table 1, the first and third algorithms give good results. The prediction error algorithm (algorithm 1) is slightly inferior to the Narendra-Gallman algorithm (algorithm 3) for high noise level (N2). Narendra-Gallman algorithm is very robust and gave good results in every case tried.

The results of the recursive prediction error algorithm (algorithm 2) are inferior compared to the other two algorithms. It can be seen that the convergence to the true parameter values is slow in the nonnoisy (NN) and low noise (N1) cases. In the high noise level case (N2), meaningful parameter estimates could not be obtained with this method. This algorithm may be useful for on-line applications where the parameter values

Table 1. Results for Three Different Algorithms

Parameters	A1 NN	A2 NN	A3 NN	A1 N1	A2 N1	A3 N1	A1 N2	A2 N2*	A3 N2
a1	−1.5000	−1.4836	−1.5000	−1.5012	−1.4869	−1.5013	−1.5084	—	−1.5118
a2	0.7000	0.6846	0.7000	0.7017	0.6946	0.7018	0.7154	—	0.7162
b1	1.0000	1.1566	1.0000	0.9834	0.7214	0.9985	0.8121	—	0.9776
b2	0.5000	0.6166	0.5000	0.4975	0.3729	0.5053	0.4929	—	0.5551
γ1	1.0000	1.0000	1.0000	1.0000	1.0000	0.9843	1.0000	—	0.8521
γ2	−2.0000	−1.6941	−2.0000	−2.0380	−2.8295	−2.0070	−2.4756	—	−2.0728
γ3	5.0000	3.3311	5.0000	5.1383	9.2564	5.0621	6.7029	—	5.6111

A1 = algorithm 1
A2 = algorithm 2
A3 = algorithm 3
NN = no noise
N1 = noise level 1
N2 = noise level 2

*The algorithm did not converge to meaningful results.

are known to a certain accuracy but have to be monitored due to slow process changes. In other cases, it should not be relied upon to estimate the unknown parameters because it may converge to wrong results.

Having seen that the Narendra-Gallman algorithm is more robust than the others, we extended the estimation of the Hammerstein models to multiple input/single output (MISO) case using this algorithm. The parameters of fully multivariable systems can be estimated by running the algorithm for each output and estimate parameters for each case.

Extension to the MISO Case

The multiple input/single output (MISO) Hammerstein model is more complicated than MISO linear models. The MISO case can give rise to more than one model structure. This is illustrated in Figure 3 for the two-input case. Depending on whether the static nonlinearity is combined or separate, we obtain two different models. The combined nonlinearity case is more general, but it can cause a very challenging parameter estimation problem because of the large number of parameters to be estimated.

For the two-input case, the MISO Hammerstein model can be described by:

$$y(t)=\frac{B_1(q^{-1})}{A_1(q^{-1})}x_1(t)+\frac{B_2(q^{-1})}{A_2(q^{-1})}x_2(t)$$

$$x_1(t)=N_1(u_1);\ x_2(t)=N_2(u_2)\text{: Separate NL}$$

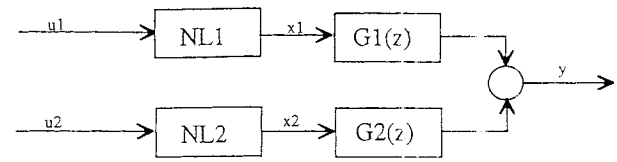
$$x_1(t)=N_1(u_1,u_2);\ x_2(t)=N_2(u_1,u_2)\text{: Combined NL} \quad (28)$$

For the separate parameterization, the functions N_1 and N_2 are given by expressions similar to Eq. 1. For combined parameterization, we obtain:

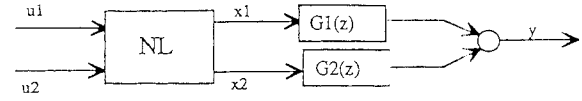
$$x_1(t)=\sum_{i=0}^{m}\sum_{j=0}^{n}\gamma_{1,ij}u_1^i(t)u_2^j(t)$$

$$x_2(t)=\sum_{i=0}^{p}\sum_{j=0}^{q}\gamma_{2,ij}u_1^i(t)u_2^j(t) \quad (29)$$

The Narendra-Gallman method estimates the parameters of the linear and nonlinear parts separately. At every iteration, assuming the linear part is known and following a procedure similar to SISO case (Eqs. 25-27), we may estimate the nonlinear parameters of a two-input MISO system from:



a. Separately parametrized NL



b. Commonly parametrized NL

Figure 3. Two types of MISO Hammerstein models.

We can express Eq. 30 as:

$$R\gamma=F \quad (32)$$

Solving for γ, we get:

$$\gamma=R^{-1}F \quad (33)$$

The intermediate variables x_1 and x_2 can be obtained from Eq. 29 once γ's are estimated. So a better linear model may be estimated using x_1 and x_2. This is continued until convergence occurs. The example given here is a two-input system, but the results can be generalized to more than two inputs as well.

The parameters of the MISO Hammerstein models may be estimated as explained above. However, when the nonlinearity is commonly parameterized, the number of parameters to be estimated can become excessive. For example, the distillation problem with a first-order linear model and third-order polynomial nonlinearity on both inputs requires the estimation of 34 parameters. Significant computer time is required to estimate such models. The large number of the parameters results from the many regressors generated by the multiplicative combinations of the inputs. Some of the regressors may contribute little to the problem. Deleting those parameters may give a parsimonious description of the system with negligible loss of accuracy.

Input Design

Selection of the input for the identification of nonlinear systems presents a different problem from that of linear systems. A standard signal for the identification of linear systems is the pseudorandom binary sequence (PRBS). However, Leontaritis and Billings (1987) noted that the use of binary signals may cause loss of identifiability for nonlinear systems. That

$$\begin{bmatrix} \sum_{t=1}^{N}\frac{B_1(q^{-1})}{A_1(q^{-1})}U_1(t)\frac{B_1(q^{-1})}{A_1(q^{-1})}U_1^T(t) & \sum_{t=1}^{N}\frac{B_1(q^{-1})}{A_1(q^{-1})}U_1(t)\frac{B_2(q^{-1})}{A_2(q^{-1})}U_2^T(t) \\ \sum_{t=1}^{N}\frac{B_2(q^{-1})}{A_2(q^{-1})}U_2(t)\frac{B_1(q^{-1})}{A_1(q^{-1})}U_1^T(t) & \sum_{t=1}^{N}\frac{B_2(q^{-1})}{A_2(q^{-1})}U_2(t)\frac{B_2(q^{-1})}{A_2(q^{-1})}U_2^T(t) \end{bmatrix}\begin{bmatrix}\gamma_1\\ \gamma_2\end{bmatrix}=\begin{bmatrix}\sum_{t=1}^{N}\frac{B_1(q^{-1})}{A_1(q^{-1})}U_1(t)y(t)\\ \sum_{t=1}^{N}\frac{B_2(q^{-1})}{A_2(q^{-1})}U_2(t)y(t)\end{bmatrix} \quad (30)$$

where

$$U_1(t)=[u_1(t),u_2(t),u_1(t)u_2(t),\ldots,u_1^m(t)u_2^n(t)]^T$$

$$U_2(t)=[u_1(t),u_2(t),u_1(t)u_2(t),\ldots,u_1^p(t)u_2^q(t)]^T \quad (31)$$

paper advises use of uniformly distributed noise signals when there is an amplitude constraint and normally distributed noise when there is a power constraint. In the simulations on distillation columns, it is observed that the use of random signals creates problems for the estimation algorithms. The input signal and the resulting output change quickly when the input is a random signal, and the estimation algorithm usually converges to a local minimum in those cases. To improve performance of the algorithms, pseudorandom signals with random amplitude have been used. This way the user has control over the duration of the input, and the output signal changes much more smoothly. The parameter estimation algorithm does not encounter any problems in this case.

Unlike the case of linear system identification, the selected maximum input amplitude does influence the model parameter estimates in the case of Hammerstein models. For example, for a small input amplitude, a model which is very close to the linearized model results; however, with larger input amplitudes the model changes. Therefore, the input amplitude should be representative of actual system operation to obtain a more useful model.

Application of Hammerstein Models to Distillation Examples

It seems reasonable to expect that Hammerstein models may be used to represent many chemical processes. The resulting models will not give an exact fit to the data, but the approximation is often good as will be shown in the distillation column and the heat exchanger examples below. Hammerstein models can give acceptable models where linear models fail to describe a process adequately. In these applications it is seen that the Hammerstein model approximates the steady-state operation curve fairly well, limited by the order of the polynomial assumed for the approximation of the static nonlinearity. The linear part gives an 'average' linear model which approximates the data. This 'average' linear model would be more useful for controller design than the linearized model (obtained from infinitesimal deviations from equilibrium) because most of the time the system will be better represented by this 'average' model compared to the linearized model.

Background on distillation

Distillation is one of the most important processes in the chemical industries, and its modeling and control can present serious problems to engineers. As the demanded product purities increase, the dynamic responses of distillation columns tend to become more nonlinear, and these columns are typically difficult to control.

In the literature there are several approaches to modeling of distillation columns. Linearization of the equations describing the behavior of the column is one of the basic approaches (Luyben, 1990). This approach is of limited value in high-purity columns because the linear region for such columns is extremely small. The steady-state gains rapidly approach zero as product purities increase. When there is a realistic disturbance to the column, apparent time constants are much smaller than the values predicted by linear analysis. These facts were observed and explained by Fuentes and Luyben (1983), Kapoor and McAvoy (1987), and Skogestad and Morari (1987). In the last paper, the inventory time constant concept, which was used by several authors before, is elaborated; and it is shown that the dominant time constant for changes between two steady states, which can be estimated by a simple formula, depends on the magnitude and direction of both the inputs (reflux flow and vapor flow) and the disturbances (feed flow and composition disturbances). The variations of the gains with the input magnitudes were also investigated and simple formulas were derived by extending the Fenske equation to the partial reflux case (Douglas et al., 1979).

Some authors have proposed the use of logarithmic functions of compositions in place of the composition measurements to compensate for the nonlinearities of a distillation column (Shinskey, 1987; Skogestad and Morari, 1988). The top and bottom compositions, x_d and x_b, are scaled to give new variables, X_D and X_B, in the following manner:

$$X_D = \log\left[\frac{1-x_d}{1-x_{dss}}\right]$$

$$X_B = \log\left[\frac{x_b}{x_{bss}}\right] \tag{34}$$

where x_{dss} and x_{bss} are the steady-state values of x_d and x_b.

Using logarithmic compositions is equivalent to assuming that the process gain varies in direct proportion to the impurities in the streams. For example, if the gain between bottom composition and vapor rate is:

$$K = \frac{dx_b}{dV} = kx_b \tag{35}$$

The process gain approaches zero as x_b approaches zero. Integrating Eq. 35 yields a logarithmic relationship between x_b and V. So, the use of logarithmic composition transformation captures one important aspect of the nonlinear behavior in a distillation column.

Use of Hammerstein models

The successful use of logarithmic transformations in distillation control shows that Hammerstein models may be appropriate for the representation of distillation column dynamics. Logarithmic transformation is a static data transformation. In industrial applications, it is common practice to logarithmically transform the data obtained from high-purity distillation columns before attempting to use linear identification methods. The use of Hammerstein models gives similar results. However, Hammerstein modeling is not application-specific as logarithmic transformation is and can be used for other nonlinear processes as well. The Hammerstein model gives the static nonlinearity as well as a linear model which portrays average input magnitude behavior.

The description of the static nonlinearity (i.e., the steady-state operation curve) in a distillation column can be obtained from a steady-state model, changing the input variables, reflux and vapor rate, and observing the steady-state values of overhead and bottoms compositions. Obviously, the problem is multivariable, as the compositions depend on more than one variable, but we shall deal with the SISO case for the time being for simplicity. The MISO case will be discussed later.

The properties of the distillation column used in the simu-

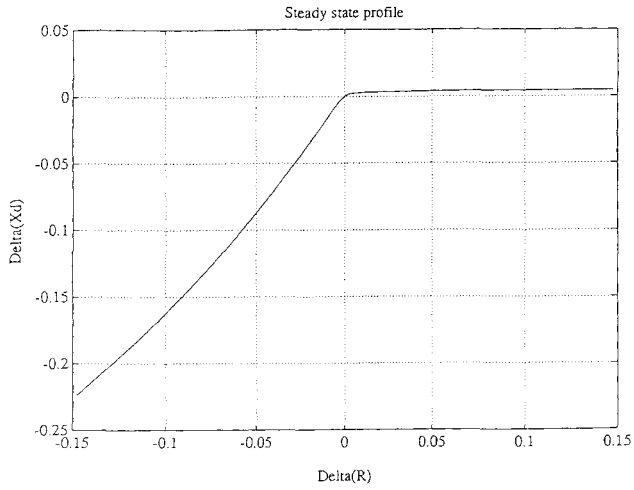


Figure 4. Steady-state operation curve for distillation column.

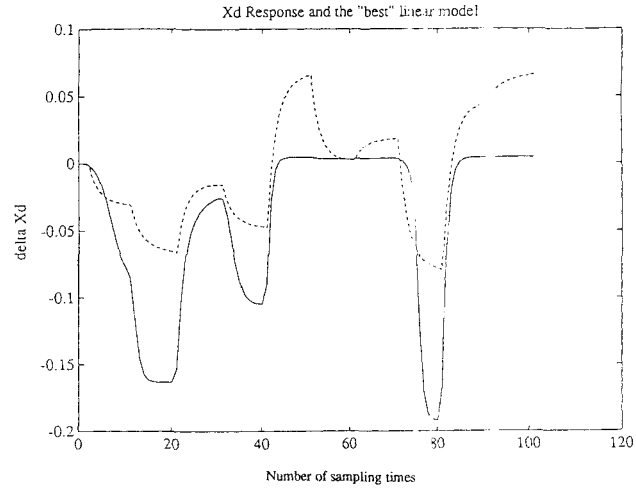


Figure 6. Simulated response of the 'best' linear model vs. actual data.

solid line, actual response; dashed line, linear model response

lations are given in the Appendix. This example has been used by other investigators (Wahl and Harriott, 1970; Skogestad and Morari, 1987). Figure 4 shows the steady-state variation of top composition, x_d, with respect to reflux, R. The steady-state gain is the slope of the curve at any operating point. It can be seen that the region of linear behavior around the origin is small. This is to be expected since the purity of x_d is high ($x_d = 0.995$). Another observation is the fact that the gains in the direction of low purity are close to the linearized gain, but in the high-purity direction they approach zero very quickly.

Results for the SISO case

The distillation column described in the Appendix was simulated, driven by a pseudorandom input with random amplitude which had a maximum amplitude that was 10% of the steady-state value. A set of 100 data points was collected with a sampling time of 10 min. In the simulations for the SISO case, the reflux flow, R, was changed, and changes in the top composition, x_d, were observed. The steady-state values of the input and output signals were subtracted from the data before using the estimation techniques described above.

The input and output data are plotted in Figure 5. As can be observed from this figure, the output is a nonlinear function of the input. A linear model would provide a poor fit to such data. The response of the 'best' linear model obtained using the actual input-output data and an output error algorithm to estimate the parameters is compared to that of the distillation column in Figure 6. It can be seen that the model is poor. Instead, the use of Hammerstein models gives reasonable fits to the data. The static nonlinearity was approximated by third- and fourth-order polynomials in two different cases. For the linear part, first-order models proved to be accurate enough. The use of second-order linear models does not increase the accuracy appreciably. The results obtained using the Hammerstein models are given in Table 2. The simulated responses of the models are compared with the actual response in Figure 7, and the comparison of the estimated static nonlinearities with the original steady-state curve is given in Figure 8.

As can be seen from these results, the Hammerstein model approximates the data reasonably well. The static nonlinearity approximates the steady-state operation curve well, and the linear part of the model converges to an 'average' linear model which approximates the data as closely as possible. Distillation column dynamics (i.e., apparent time constants) change as the response deviates from the steady state (Skogestad and Morari, 1987). The Hammerstein model, however, cannot take this

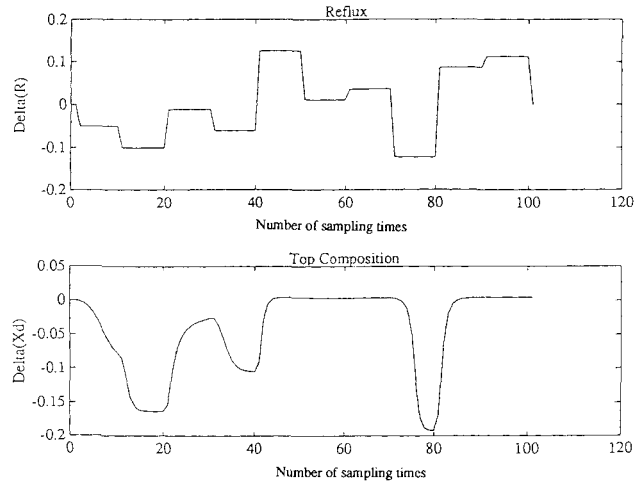


Figure 5. Input/output data for distillation column example.

Table 2. Estimated Parameters of Hammerstein Models for Distillation Example

	Model 1*	Model 2**
a_1	−0.796	−0.757
b_1	0.204	0.243
γ_1	1.01	1.04
γ_2	−7.13	−14.11
γ_3	−7.43	−16.72
γ_4	—	562.75

*Hammerstein model with 3'rd order nonlinearity
**Hammerstein model with 4'th order nonlinearity

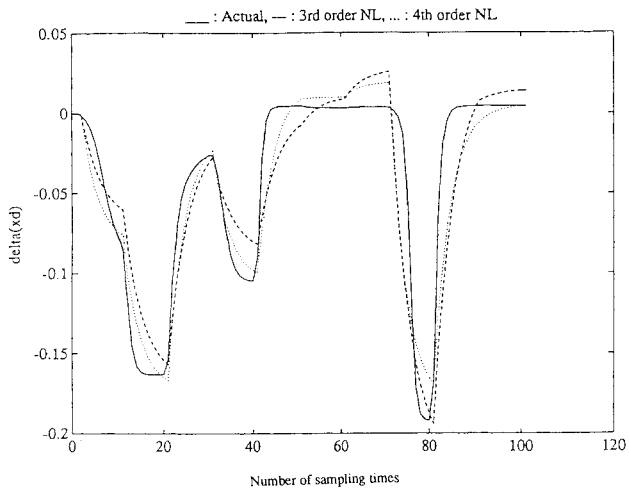


Figure 7. Simulated responses of estimated models vs. actual data.

solid line, actual response; dashed line, model with third-order NL; dotted line, model with fourth-order NL

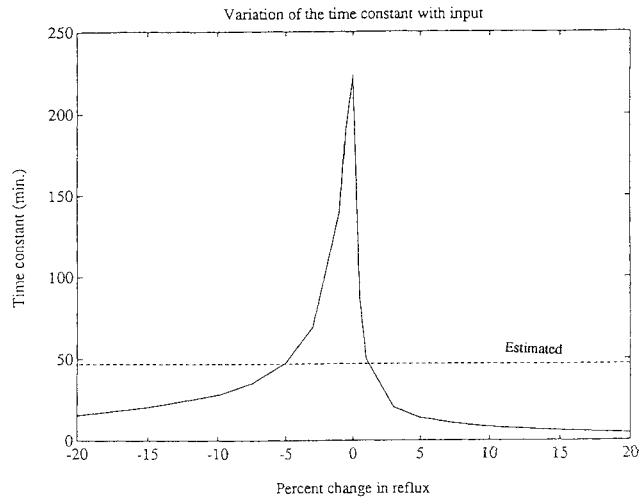


Figure 9. Variation of apparent time constant with change in reflux and time constant estimated by Hammerstein model (dashed line).

into account, and it estimates an 'average' time constant. The time constant of the response, assuming it is linear, changes with the reflux rate, R, as shown in Figure 9.

Figure 9 also shows the time constant estimated by the Hammerstein model. It can be seen that the estimated time constant is an 'average' in the actual range of operation.

For very high purity columns (i.e., $x_d = 0.999$ or higher), the steady-state operation curve changes very steeply and the Hammerstein models cannot follow this change with low-order polynomials. The use of high-order polynomials creates numerical problems, and the polynomials show unnecessary oscillations. To model very high-purity columns, it has been necessary to use the logarithmic transformations with Hammerstein models.

Use of log transformation

As stated before, logarithmic transformation of the output data can make it look more like the output of a linear system for moderately high-purity columns (i.e., $x_d = 0.99$–0.995), and linear models along with logarithmic transformation may give sufficiently accurate models. For very high purities, however, the nonlinear behavior cannot be corrected by logarithmic transformation alone. Use of Hammerstein models on top of the logarithmic transformation adds another nonlinear correction and can provide better models for high-purity columns. Note that such models would contain two nonlinearities.

As an example, a high purity column ($x_d = 0.9999$) was simulated. The input and the logarithmically transformed response of the process are plotted in Figure 10. The output data were first logarithmically transformed, and then both linear (first-order) and Hammerstein models (first-order linear dynamic, fourth-order NL parts) were fitted to the data. The model responses are plotted along with the original response in Figure 11. It can be seen that the Hammerstein model provides a better fit than the linear model. The fit of the Hammerstein model becomes increasingly better than the linear model as the product purity increases.

Noise sensitivity

The noise sensitivity of the estimation method is another

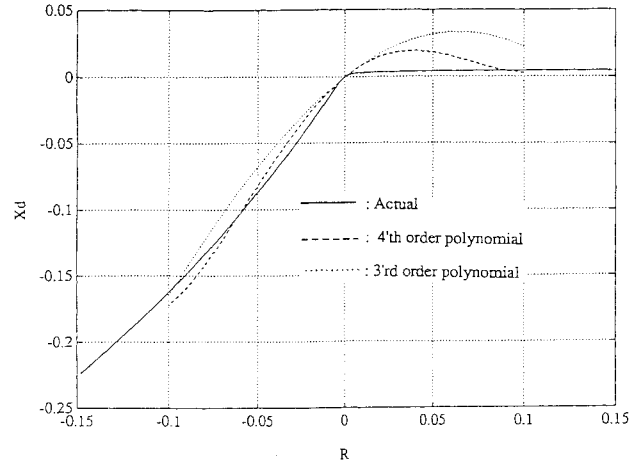


Figure 8. Estimated static NL vs. steady-state operation curve.

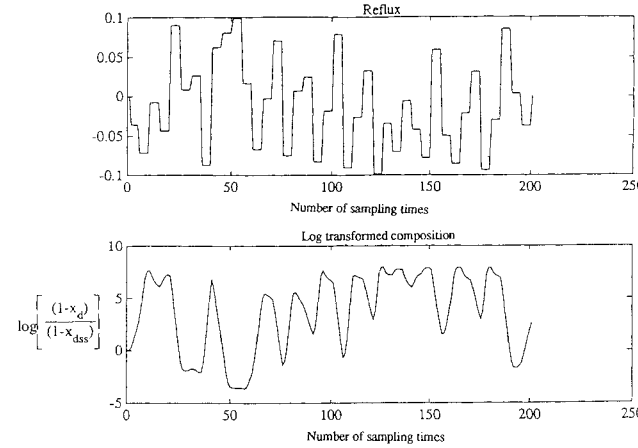


Figure 10. Very high-purity column: input and log-transformed output data.

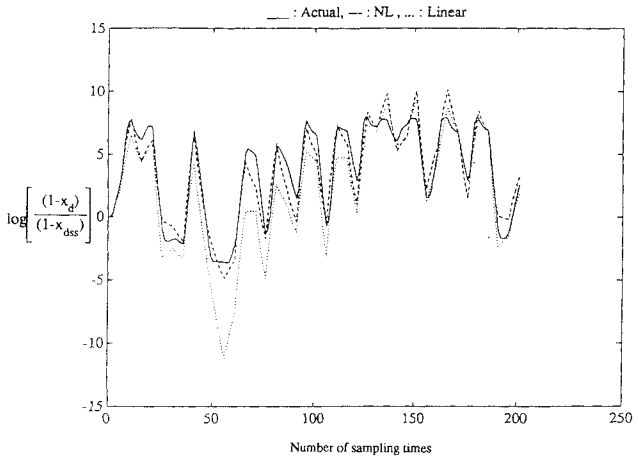


Figure 11. Simulated response of linear and nonlinear models vs. log-transformed output data.

solid line, actual response; dashed line, NL model response; dotted line, linear model response.

important question. If the algorithms diverge in the presence of noise, then they would not have much practical use because all the data obtained from process systems contain some noise. Simulations show that a realistic amount of noise added to the system output does not affect the results very much. A relevant example was given earlier. In that example, the model structure described the system correctly. In the case of distillation, however, there is some process model mismatch, as would be the case for many real processes. The effects of noise on such a model structure were investigated by simulation. The distillation column was simulated and some independent noise (whose standard deviation is approximately 10% of the maximum output amplitude) was added to the output data. The nonnoisy and noisy signals are plotted in Figure 12. The results of the estimation using the noisy and the nonnoisy signals are given in Table 3. It can be seen that the noise does not affect the parameter estimates much. The response obtained from the model which was estimated using the noise data is plotted

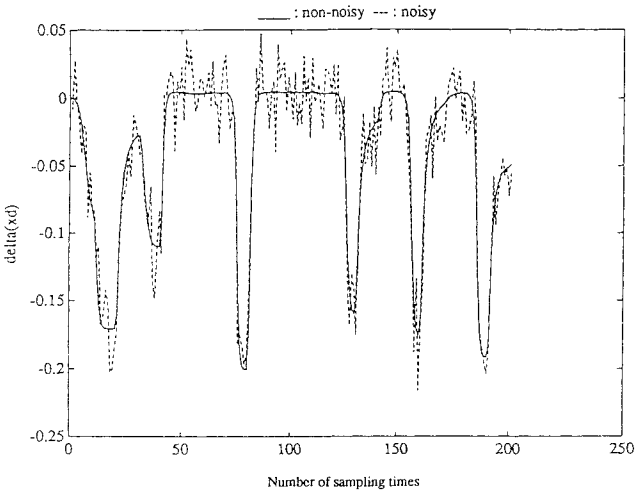


Figure 12. Nonnoisy and noisy output signals.

Solid line, nonnoisy signal; dashed line, noisy signal

Table 3. Estimated Parameters for Nonnoisy and Noisy Data

Data	Nonnoisy	Noisy
a_1	−0.728	−0.744
b_1	0.272	0.255
γ_1	1.08	1.16
γ_2	−8.52	−8.48
γ_3	−23.73	−27.65
γ_4	228.52	226.24

along with the nonnoisy output signal in Figure 13. It can be seen that the model replicates the output quite well despite the noise.

MISO models for distillation

Distillation columns and many other chemical processes are multivariable. Changes in one output variable depend on more than one input. For linear systems, this does not create much of a problem. Because of the superposition principle, the individual responses to each input can be added. Therefore, it is possible to identify transfer functions for linear systems element by element. However, this does not work for nonlinear systems. For nonlinear identification, all the inputs have to be changed simultaneously. Then, nonlinear MISO models may be fitted to each output. The nonlinear MISO models may be parameterized either separately or in a combined manner. The separate modeling is more restrictive and in the trials did not produce good models. Therefore, only the model structure in which the nonlinearity is commonly parameterized will be discussed below.

The distillation column previously described was simulated with the same steady-state conditions used in SISO simulations. In the distillation simulations, the reflux R and vapor rate V were changed, and top composition x_d was used as output. Pseudorandom signals with random amplitude, whose maximum is 5% of the steady state of R and V, were used to disturb the process. One set of input-output data is plotted in Figure 14. A model structure with a nonlinearity with combined parameterization, as in Figure 3, was fitted to the data. The linear parts are first-order and the nonlinearities contain up to the third power of each input. With such a model structure, the number of parameters to be estimated is 34. The MISO Narendra-Gallman method described previously was used to estimate the parameters of the model. The model response is plotted along with the actual response in Figure 15. It can be seen that the model represents the system quite well.

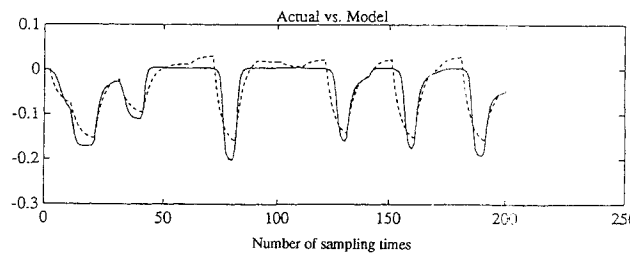


Figure 13. Output of the model estimated using noisy data.

solid line, actual response; dashed line, model response

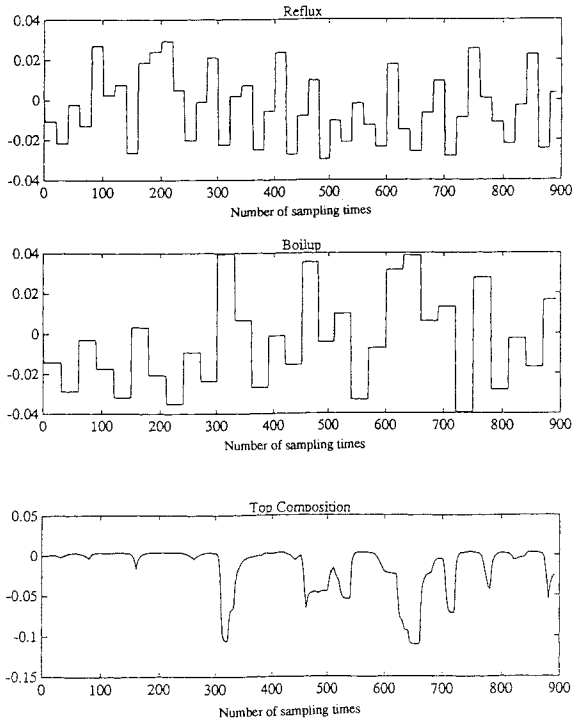


Figure 14. Input/output data for MISO example

Application to an Experimental Heat Exchanger

Hammerstein models were also applied to the identification of the nonlinear behavior of a steam-water heat exchanger. The experimental equipment is shown in Figure 16. The steam condenses in the two-pass shell and tube heat exchanger, thereby raising the process water temperature. The steam flow rate and process water flow rate may be controlled by pneumatic control values. Flow rates and temperatures are measured by flow sensors and thermocouples. The measurements are fed to a MICROMAC 6000 data acquisition and control system. The characteristics of the valves are nonlinear, but flow controllers are used to avoid the nonlinearity. The nonlinearity of the flow signal, which is obtained by measuring the pressure drop across the orifice, is avoided by taking the square root of the pressure transmitter signal.

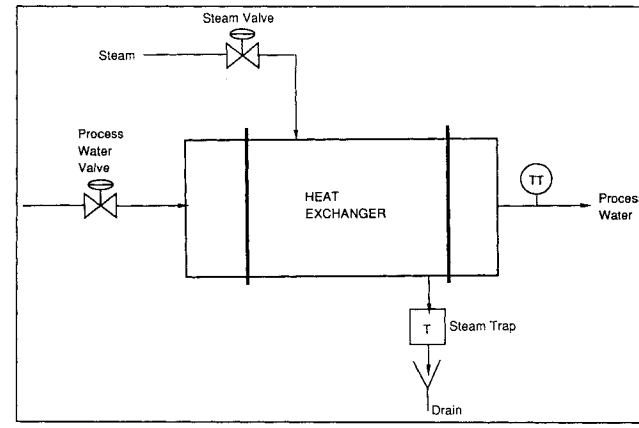


Figure 16. Heat exchanger.

The process water temperature changes are reasonably linear with respect to the changes in the steam flow rate. However, changes in the process water flow rate, keeping the steam flow constant, cause significant deviations from linearity for the measured process water exit temperature. In this section, we analyze and identify the nonlinear behavior of the steam exchanger with respect to the changes in the process flow rate. Hammerstein models will be used to identify the nonlinearities.

The exchanger displays an interesting nonlinear behavior when operated with a fixed steam flow rate. As can be seen in Figure 16, the condensed steam is drained through a steam trap which lets out only liquid. When the process water flow rate is high, the exit temperature of steam drops below the condensation temperature at atmospheric pressure. Therefore, the steam becomes subcooled liquid, which floods the exchanger. The flooding causes the heat transfer area to decrease; therefore, the heat transfer per unit mass of process water decreases. This is the main cause of the nonlinear behavior.

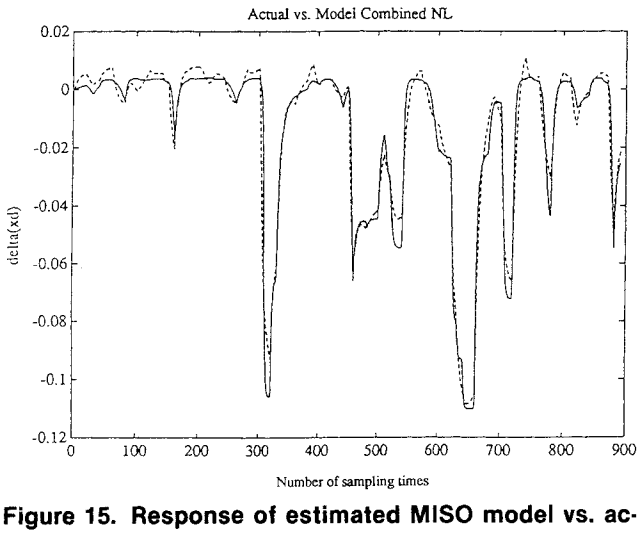


Figure 15. Response of estimated MISO model vs. actual data.

solid line, actual response; dashed line, model response

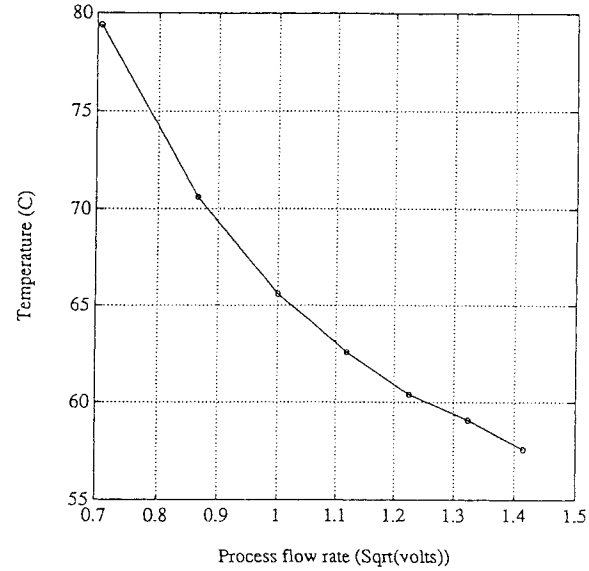


Figure 17. Steady-state operation points.

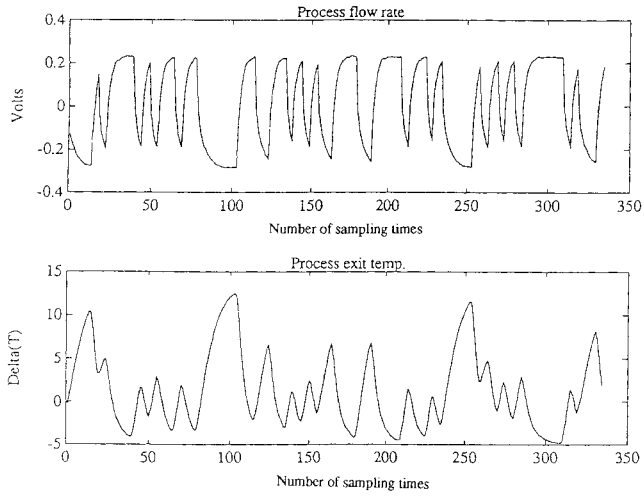


Figure 18. Heat exchanger input and output data.

The steady-state operation curve of water exit temperature vs. process water flow rate, Figure 17, shows this effect. There are two distinct regions of the exchanger: one not flooded (which corresponds to low process water flow rates); and the other the flooded region. It can be seen that the slope of the curve is less in the flooded region.

The dynamics of the steam exchanger were determined for the following steady-state conditions:

Steam flow rate:	62% of maximum
Water flow rate:	42% of maximum
Inlet water temperature:	30°C

To identify the nonlinearity, PRBS tests were performed on the heat exchanger. The process water flow rate was changed by changing the setpoint of the process water flow controller, and the process water exit temperature was observed. The input/output data are plotted in Figure 18. Due to loose tuning of the process flow controller, the flow rate responds slowly to the PRBS changes to the setpoint. The temperature response is nonlinear, which can easily be detected by the asymmetry of the response with respect to the starting value. The sampling time was 12 seconds, and the PRBS switching time was 1 minute.

To fit the data, both linear and nonlinear models were tried. The parameters of the linear model were estimated using an output error algorithm (Ljung, 1986). The nonlinear model is a Hammerstein model, with parameters estimated using the Narendra-Gallman algorithm. For the linear part, a second-order model with no time delay gives a good fit to the data. For the estimation of the Hammerstein model, the nonlinear part was assumed to be a fourth-order polynomial nonlinearity. The estimated models are as follows:

The Linear Model:

$$G_L(z^{-1}) = \frac{-5.7715z^{-1} + 5.673z^{-2}}{1 - 1.7608z^{-1} + 0.7661z^{-2}} \tag{36}$$

The Hammerstein Model:

$$G_H(z^{-1}) = \frac{0.207z^{-1} - 0.1764z^{-2}}{1 - 1.608z^{-1} + 0.6385z^{-2}}$$

$$N(u) = -31.549u + 41.732u^2 - 24.201u^3 + 68.634u^4 \tag{37}$$

The simulated responses of the above models to the process input are compared with the actual output in Figure 19. It can be seen that the Hammerstein model duplicates the process output very closely, whereas there is a considerable discrepancy between the linear model and the process output.

The estimated static nonlinearity is compared to the steady-state operation points in Figure 20. It can be seen that the trend in the static nonlinearity follows the operating points

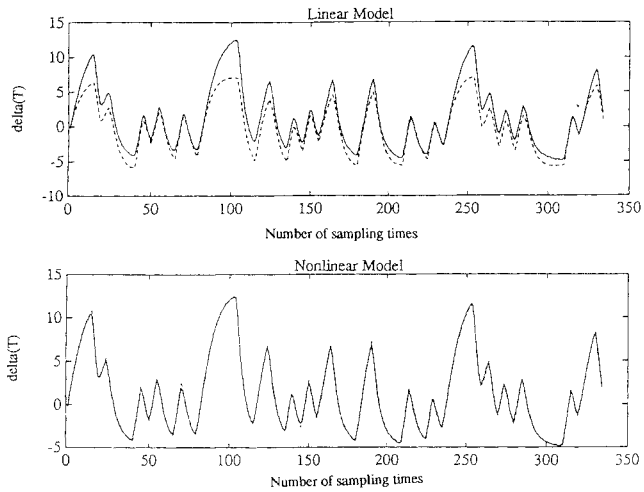


Figure 19. Responses of linear and nonlinear models vs. actual response.

solid line, actual response; dashed line, model response

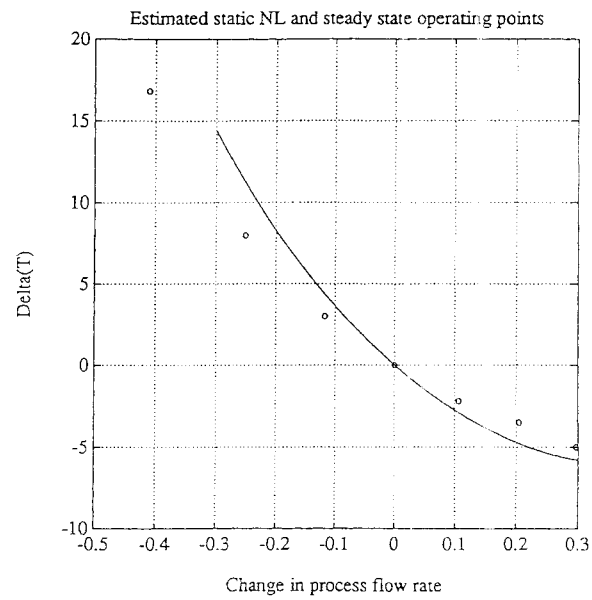


Figure 20. Estimated static nonlinearity and steady-state operation points.

closely. The nonlinear response in the steam exchanger is predicted by the Hammerstein model very well.

Conclusions

Identification of nonlinear processes using Hammerstein models was discussed in this article. Hammerstein models can be used in identifying many chemical processes where linear models are inadequate. The examples of distillation column and heat exchanger identification support this conclusion. For these examples, the estimated static nonlinearity approximates the actual steady-state operation curve, while the linear part of the model represents an 'average' linear model.

For the estimation of Hammerstein-type models, three different algorithms were tested. The Narendra Gallman algorithm was found to be most robust among those three, and it was extended to the MISO case. MISO identification of a distillation column shows that the proposed model structure represents the dynamics of the process accurately.

Acknowledgment

The many helpful suggestions of Dr. Randy McFarlane are gratefully acknowledged.

Notation

$A(z^{-1})$ = denominator of the linear part of the Hammerstein model
$B(z^{-1})$ = numerator of the linear part of the Hammerstein model
$e(t)$ = disturbances modeled by white noise
g = gradient of the loss function $V(\theta)$
H = Hessian of the loss function $V(\theta)$
m = number of parameters in the polynomial describing the nonlinearity
N = number of data points
n = number of parameters in $A(z^{-1})$ and in $B(z^{-1})$
NL = nonlinear
$PRBS$ = pseudorandom binary sequence
$u(t)$ = system input
$V(\theta)$ = criterion to be minimized in estimation
$x(t)$ = output of the NL part of the Hammerstein model
$y(t)$ = system output
$\bar{y}(t,\theta)$ = predictor of $y(t)$

Distillation

B = bottoms flow
D = distillate flow
F = feed flow
L_i = liquid flow rate leaving tray i
M_i = liquid holdup of the i'th tray
R = reflux flow
V = vapor rate
X_B = log transformed bottoms composition
X_D = log transformed top composition
x_b = bottoms composition
x_d = top composition
z_f = feed composition
α = relative volatility

Greek letters

$\epsilon(t,\theta)$ = prediction error at t
γ = parameters describing the static nonlinearity
θ = system parameters to be estimated
ρ = forgetting factor in recursive estimation

Literature Cited

Bamberger, W., and Isermann, R., "Adaptive On-Line Steady-State Optimization of Slow Dynamic Processes," *Automatic*, **14**, 223 (1978).

Beyer, J., G. Gens, and J. Wernstedt, "Identification of Nonlinear Systems in Closed Loop," *IFAC Symp. on Identification and System Parameter Estimation*, Pergamon Press, Darmstadt (1979).

Billings, S. A., "Identification of Nonlinear Systems—a Survey," *IEE Proc. Pt.D.*, **127**, 272 (1980).

Billings, S. A., and S. Y. Fakhouri, "Identification of a Class of Nonlinear Systems Using Correlation Analysis," *Proc. IEE*, **125**, 691 (1978).

Chang, F. H., and R. Luus, "A Noniterative Method for Identification Using the Hammerstein Model," *IEEE Trans. Automatic Control*, **AC-16**, 464 (1971).

Douglas, J. M., A. Jafarey, and T. J. McAvoy, "Short-Cut Techniques for Distillation Column Design and Control I: Column Design," *Ind. Eng. Chem. Proc. Des. Dev.*, **18**, 197 (1979).

Fuentes, C., and W. L. Luyben, "Control of High Purity Distillation Columns," *Ind. Eng. Chem. Process Des. Dev.*, **22**, 361 (1983).

Gallman, P., "A Comparison of Two Hammerstein Model Identification Algorithms," *IEEE Trans. Aut. Control*, **20**, 771 (1976).

Greblicki, W., and M. Pawlak, "Identification of Discrete Hammerstein Systems Using Kernel Regression Estimate," *IEEE Trans. on Aut. Control*, **31**, 74 (1987).

Haber, R., "Parametric Identification of Nonlinear Dynamic Systems Based on Correlation Functions," *IFAC Symp. on Identification*, Darmstadt, Germany (1979).

Haber, R., "Parametric Identification of Nonlinear Dynamic Systems Based on Nonlinear Crosscorrelation Functions," *IEE Proc. Pt. D.*, **135**, 405 (1988).

Haist, N. D., F. H. I. Chang, and R. Luus, "Nonlinear Identification in Presence of Correlated Noise Using a Hammerstein Model," *IEEE Trans. Aut. Control*, **18**, 552 (1973).

Hsia, T. C., "A Multistage Least Squares Method for Identifying Hammerstein Model Nonlinear Systems," *Proc. IEEE CDC*, 934 (1976).

Kapoor, N., and T. J. McAvoy, "An Analytical Approach to Approximate Dynamic Modeling of Distillation Towers," *Ind. Eng. Chem. Res.*, **26**, 2473 (1987).

Kortmann, M., and H. Unbehauen, "Identification of Nonlinear MISO Systems," *IFAC World Cong. Proc.*, Munich, Germany (1987).

Leontaritis, I. J., and S. A. Billings, "Experimental Design and Identifiability for Nonlinear Systems," *Int. J. Systems Sci.*, **18**, 189 (1987).

Ljung, L., *System Identification Toolbox: User's Guide*, The Mathworks Inc., Sherborn, MA (1986).

Ljung, L., *System Identification: Theory for the User*, Prentice Hall, Englewood Cliffs, NJ (1987).

Ljung, L., and T. Soderstrom, *Theory and Practice of Recursive Identification*, MIT Press, Cambridge, MA (1983).

Luyben, W. L., *Process Modeling, Simulation, and Control for Chemical Engineers*, 2nd ed., McGraw Hill, New York (1990).

Narendra, K. S., and P. G. Gallman, "An Iterative Method for the Identification of the Nonlinear Systems Using the Hammerstein Model," *IEEE Trans. Aut. Control*, **12**, 546 (1966).

Shinskey, F. G., *Process Control Systems*, 3rd ed., McGraw Hill, New York (1987).

Skogestad, S., and M. Morari, "The Dominant Time Constant for Distillation Columns," *Comput. Chem. Eng.*, **1**, 607 (1987).

Skogestad, S., and M. Morari, "Understanding the Dynamic Behavior of the Distillation Columns," *Ind. Eng. Chem. Res.*, **27**, 1842 (1988).

Stoica, P., and T. Soderstrom, "Instrumental Variable Methods for Identification of Hammerstein Systems," *Int. J. Control*, **35**, 459 (1982).

Wahl, E. F., and P. Harriott, "Understanding and Prediction of Dynamic Behavior of Distillation Columns," *Ind. Eng. Chem. Proc. Des. Dev.*, **9**, 396 (1970).

Appendix

The properties of the distillation column used in the simulations in consistent set of units are:

No. of trays	= 25	Relative volatility	= 2
Feed tray	= 12	Distillate flow	= 0.5
Feed composition	= 0.5	Bottoms flow	= 0.5
Top composition	= 0.995	Tray holdup	= 0.5
Bottom composition	= 0.005	Condenser and	
Reflux	= 1.477	Bottoms holdups	= 0.5
Boilup	= 1.977		

The assumptions made in deriving the dynamic model are:

- Constant relative volatility
- 100% stage efficiency
- Immediate vapor and liquid response
- Perfect level control on reflux drum and column base.

With those assumptions, the dynamic model is described by the following equations:

Bottom and Reboiler

$$M_b\frac{dx_b}{dt}=L_1x_1-Bx_b-Vy_b$$

Condenser and Accumulator

$$M_d\frac{dx_d}{dt}=Vy_{nt}-Rx_d-Dx_d$$

Feed Tray

$$M_{nf}\frac{dx_{nf}}{dt}=V(y_{nf-1}-y_{nf})+L_{nf+1}x_{nf+1}-L_{nf}x_{nf}+Fz_f$$

Top Tray

$$M_{nt}\frac{dx_{nt}}{dt}=V(y_{nt-1}-y_{nt})+Rx_d-L_{nt}x_{nt}$$

Other Trays

$$M_i\frac{dx_i}{dt}=V(y_{i-1}-y_i)+L_{i+1}x_{i+1}-L_ix_i$$

where

$$L_i=\begin{cases}R & i>nf\\ R+F & i\le nf\end{cases}$$

$$y_i=\frac{\alpha x_i}{1+(\alpha-1)x_i}$$

Manuscript received Sept. 13, 1990.

Corrections

In the article titled "Two-Phase Concurrent Flow in Packed Beds" (1961, Vol. 7, p. 231) by Larkin et al., Eq. 23 on p. 237 contains a typographical error. Instead of -0.774, the first term should be -0.744 as correctly given with the upper curve in Figure 7. (Proof: At $\chi=10$ and 12.33, the correct equation using -0.744 as its first term gives $R_\ell=0.4699$ and 0.5001 which agrees with the values read from the solid curve in the upper half of Figure 7. If -0.774 were used as the first term in Eq. 23, the calculated values will be $R_\ell=0.4385$ and 0.4669 which fall below the solid curve.)

This article is a significant contribution and is very useful for packed-bed reactor designs in refineries.