Block-Oriented NARMAX Models with Output Multiplicities

Martin Pottmann and Ronald K. Pearson
E.I. du Pont de Nemours and Company, Wilmington, DE 19880

Block-oriented modeling approaches to the development of NARMAX models for processes exhibiting output multiplicities are introduced. "Feedback block-oriented" model structures are proposed that can force the resulting NARMAX model to have a particular steady-state behavior and reduce the model-identification problem to a constrained, linear model identification. These models are particularly suitable for gray-box modeling approaches where it is assumed that the steady-state behavior of the process is known and a NARMAX model is sought to approximate its dynamic response. The proposed approaches are applied to a CSTR model with three steady states to illustrate their performance.

Introduction

There has been much recent interest in nonlinear model-based control within the chemical engineering community. A critical step in the application of these methods is the development of a suitable model of the process dynamics. Sjoberg et al. (1995) describe the following four approaches to model development:

1. White-box modeling
2. Gray-box, physical modeling
3. Gray-box, semiphysical modeling
4. Black-box modeling

At one extreme, white-box models are based entirely on fundamental process understanding and they make no use of empirical data. These models have the advantage of possessing a clear physical interpretation; however, from the perspective of typical model-based process control schemes, they suffer two major disadvantages. First, detailed fundamental models tend to be highly complex in structure, typically involving dozens or hundreds of coupled differential algebraic equations, partial differential equations, and/or integrodifferential equations. The second disadvantage of white-box models is that they are generally continuous-time models, while process measurements and control actions are typically taken at discrete times. For both of these reasons, white-box models are inherently incompatible with many popular model-based control strategies.

At the other extreme, black-box models are purely empirical, based entirely on input/output data. Consequently, these models generally lack the physical interpretation possible for white-box models, but in return, we have the option of choosing "convenient" model structures that facilitate formulating and solving the associated control problem. One of the most popular classes of black-box models is the class of Nonlinear autoregressive moving average models with exogeneous inputs (NARMAX), defined by the equation (Billings and Voon, 1986)

$$y(k) = F[y(k-1), \ldots, y(k-p), u(k-d), \ldots, u(k-q), e(k-1), \ldots, e(k-r)] + e(k), \quad (1)$$

where $\{u(k)\}$ is the scalar input (manipulated variable) sequence, $\{y(k)\}$ is the scalar output (response variable) sequence, $\{e(k)\}$ is a sequence of model prediction errors, and p, q, r, and d are integer order parameters satisfying $p \geq 0$, $q \geq d \geq 0$, and $r \geq 0$. Here we consider prediction models of the form:

$$y(k) = \hat{y}(k) + e(k), \quad (2)$$

$$\hat{y}(k) = F[\hat{y}(k-1), \ldots, \hat{y}(k-p), u(k-1), \ldots, u(k-q)], \quad (3)$$

where $\{e(k)\}$ is assumed to be zero-mean (i.e., there is no systematic bias in the model predictions). Note that under

Correspondence concerning this article should be addressed to M. Pottmann.
Present address of M. Pottmann: DuPont DACRON R Intermediates, Cape Fear Site, P.O. Box 2042, Wilmington, NC 28402.

these assumptions, Eqs. 2 and 3 correspond to a special case of Eq. 1, with $d = 1$ and $r = p$. In what follows, we focus on the qualitative behavior of the prediction model Eq. 3, so we write $y(k)$ instead of $\hat{y}(k)$ for simplicity.

Johansen (1996) argues that pure black-box modeling is an ill-posed problem, in the sense that the solution will generally be nonunique, and it need not depend continuously on the available data. To overcome these difficulties, he proposes incorporating prior knowledge to constrain the empirically determined model and thereby obtain a better-posed identification problem. One of the specific constraints Johansen considers is that the identified NARMAX model be consistent with a steady-state mass balance for the physical process. While he does not refer to it as such, this approach is an example of the semiphysical approach to gray-box modeling listed earlier. In contrast, if we choose the model structure on the basis of physical intuition, and consequently obtain a model whose parameters must be estimated from empirical data (e.g., kinetic rate constants, heat-transfer coefficients, etc.), the result may be viewed as a physical gray-box model.

Here, we focus on a special case of semiphysical gray-box modeling that generalizes Johansen's steady-state mass-balance example. In particular, steady-state process characterizations are often more readily available than dynamic ones, in part because steady-state models are simpler and easier to develop than detailed dynamic models. Thus, we consider the identification of semiphysical gray-box models that combine prior knowledge of steady-state behavior with empirical data describing process dynamics. More specifically, we are interested here in constructing NARMAX models that are constrained to match a known steady-state characteristic exactly, while approximating the behavior of an observed input–output data set as well as possible. Further, the processes of interest here exhibit *output multiplicities*: a given steady-state input variable u_s may correspond to more than one steady-state response y_s. The canonical example of this behavior is the CSTR (Uppal et al., 1974) that can, under physically reasonable assumptions, exhibit three steady states. Other examples of processes that can exhibit output multiplicities include ideal binary distillation (Jacobson and Skogestad, 1991), homogeneous azeotropic distillation (Bekiaris et al., 1993), and fluidized catalytic crackers (Arbel et al., 1995).

To proceed, it is necessary to impose structural constraints on the NARMAX nonlinearity $F(\cdot)$ in Eq. 3, since even in the case of relatively simple polynomial NARMAX models, the qualitative behavior of the steady-state solutions (in particular, both input multiplicity and output multiplicity) is difficult to classify. For example, chapter 2 of Von Seggern (1990) gives approximately 50 pages of algebraic curves, all of which may be viewed as possible steady-state loci (u_s, y_s) for low-order polynomial NARMAX models. As an example, the equation (Von Seggern, 1990, p. 54);

$$ay_s^2 + bu_s y_s^2 - c^2 u_s^3 = 0,$$

is shown, for certain values of the parameters a, b, and c, to consist of three disconnected branches. The point of this example is that these three nonlinear terms are precisely the ones that appear in the heat exchanger NARMAX model obtained by Billings and Voon (1986) via stepwise regression.

To develop a practical semiphysical gray-box modeling approach for processes with known steady-state behavior, we restrict consideration to the feedback block-oriented model structures described in the next section, for which we can specify the steady-state locus *a priori*. We then develop a sequential modeling procedure that consists of the following steps:

1. Based on fundamental process understanding, specify a scalar nonlinear function $N(\cdot)$ that yields the correct steady-state behavior.
2. Given this nonlinear function, solve a constrained optimization problem for the NARMAX model parameters that best approximate the empirically observed process dynamics.

Alternatively, note that we can easily obtain a purely empirical modeling procedure if we estimate the nonlinear function $N(\cdot)$ from observed steady-state process data instead of obtaining it from a first-principles model; in either case, it is extremely convenient to be able to separate the steady-state and dynamic modeling into two independent steps.

Feedback–Block-Oriented Models

It is not difficult to show that nonlinear autoregressive terms are necessary to achieve output multiplicity in a NARMAX model (Pearson, 1997). Here, we consider four model structures, designated Structure I through Structure IV, that are all closely related to each other and all capable of exhibiting output multiplicity. More importantly, it will be shown that these model structures permit the steady-state behavior to be specified independently of the dynamics. The remainder of this section defines these four model structures, demonstrates that they can exhibit output multiplicity, and presents some general characteristics of these models and their interrelationships. In the third section, we give a detailed description of the gray-box model identification procedure presented earlier, while the fourth section illustrates its application to a CSTR example with output multiplicity.

Structure I

The basic model structure considered here—Structure I—is defined by the block diagram shown in Figure 1, and consists of a linear *ARMAX*(p, q) dynamic model in the forward path with the static nonlinearity $N(\cdot)$ in the feedback path. Note that feedback systems of this structure (in both discrete and continuous time) have been studied extensively (e.g., in the context of the well-known Lur'e problem (Vidyasagar, 1978)), and numerous results are available concerning the stability of such systems. However, we are not aware of any systematic use of this model structure for nonlinear system identification and, in particular, for gray-box modeling. Furthermore, this structure appears to be a quite natural addition to the family of block-oriented models, of which the two most popular members are the Wiener and Hammerstein models.

The equivalent NARMAX representation for this model is

$$y(k) = \sum_{i=1}^{p} a_i y(k-i) + \sum_{j=1}^{q} b_j u(k-j) + \sum_{j=1}^{q} b_j N[y(k-j)]. \quad (4)$$

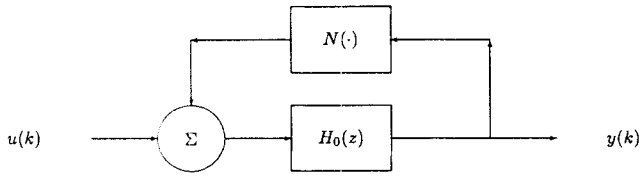

Figure 1. Feedback block-oriented model.

Note that the third term on the righthand side of this equation provides the nonlinear autoregression necessary for the existence of output multiplicities.

The steady-state behavior of this model is determined by setting $y(k) = y_s$ and $u(k) = u_s$ for all k, leading to the following general equation:

$$y_s\left(1 - \sum_{i=1}^{p} a_i\right) - u_s \sum_{j=1}^{q} b_j = N(y_s) \sum_{j=1}^{q} b_j. \tag{5}$$

It is clear from this expression that, given u_s, the steady-state response y_s will depend on both the functional form of $N(\cdot)$ and the linear model parameters $\{a_i\}$ and $\{b_j\}$. To simplify this result, define the constant

$$g_0 = \frac{\sum_{j=1}^{q} b_j}{1 - \sum_{i=1}^{p} a_i} \tag{6}$$

and note that g_0 is the steady-state gain of the linear block in Figure 1. With these definitions, Eq. 5 can then be rewritten more simply as

$$y_s - g_0 u_s = g_0 N(y_s). \tag{7}$$

Structures II, III and IV

Given the brief overview just presented of Structure I, we now consider three closely related variations on this theme—Structures II, III, and IV. Since all of these model forms are special cases of Structure IV, we begin the discussion by defining this structure, which is shown in Figure 2. This model is called the "sandwich" extension of the basic model because the static nonlinearity $N(\cdot)$ in Structure I is replaced with the cascade combination of $N(\cdot)$ with linear dynamic models $H_1(z)$ and $H_2(z)$ preceding and following it, respectively. This model may be redrawn as indicated in Figure 3, provided

$$H_*(z) = H_2(z) H_0(z) H_1(z). \tag{8}$$

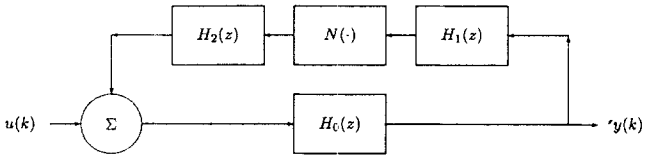

Figure 2. "Sandwich model" variation of the original structure.

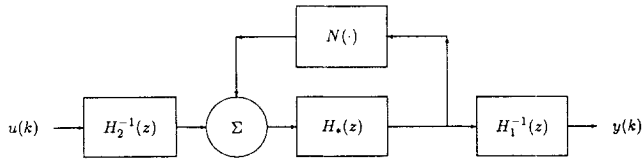

Figure 3. Equivalent representation of the "sandwich model" variation.

This result follows from the standard rules for manipulating linear block diagrams, noting that the position of the static nonlinearity $N(\cdot)$ has not been changed relative to either $H_1(z)$ or $H_2(z)$. The advantage of this block diagram representation is that it represents Structure IV as the cascade connection of Structure I with linear dynamic models both preceding and following it. *In particular, note that the presence or absence of output multiplicities in this model is determined entirely by their presence or absence in the Structure I model formed from $N(\cdot)$ and $H_*(z)$.*

If we denote the steady-state gain of the linear models H_0, H_1, and H_2 by g_0, g_1, and g_2, respectively, Eq. 7 describing the steady-state behavior of Structure I generalizes to the following result for Structure IV:

$$y_s - g_0 u_s = g_0 g_2 N(g_1 y_s). \tag{9}$$

Structures II and III are special cases of Structure IV obtained by setting $H_1(z) = 1$ or $H_2(z) = 1$, respectively. Structure II thus represents a variation on Structure I in which the static nonlinearity $N(\cdot)$ has been replaced by a Hammerstein model (Greblicki and Pawlak, 1989), while in Structure III, the static nonlinearity has been replaced by a Wiener model (Greblicki, 1992).

Structure II'

Not all of these structures have convenient NARMAX representations, but one that does is a special case of the Hammerstein structure (Structure II), which will be designated Structure II'. For the general Hammerstein structure, the output $y(k)$ is given by

$$y(k) = \sum_{i=1}^{p} a_i y(k-i) + \sum_{j=1}^{q} b_j u(k-j) + \sum_{j=1}^{q} b_j x(k-j), \tag{10}$$

where $x(k)$ is the output of the Hammerstein model in the feedback path, given by

$$x(k) = \sum_{n=1}^{s} d_n x(k-n) + \sum_{l=0}^{r-1} \tilde{c}_l N[y(k-l)]. \tag{11}$$

Note that Structure I can be represented as a special case of Structure II by taking $s = 0$ (i.e., omitting the first sum in Eq. 11, $r = 1$, and $\tilde{c}_0 = 1$, yielding $x(k) = N(y(k))$. In particular, note that it is necessary to take the lower limit of the second sum in this expression as $l = 0$ to avoid introducing an *extra* delay in the feedback path, beyond that caused by the standard linear model $H_0(z)$ in the forward path. While "delay-

free" models of this form are a little unusual, note that the linear dynamics represented by the model in Eq. 11 are realizable since they do not require *future* values of $y(k)$. Also, it follows from Eq. 10 that the overall model obtained for Structure II exhibits the usual one-sample processing delay, corresponding to the fact that all of the sums in Eq. 10 have lower limits of 1. Also, the notation $\tilde{c}_l$ has been introduced here to simplify the final expression for the NARMAX representation of this model.

To obtain this NARMAX representation, first note that if $b_1 \neq 0$, the last sum in Eq. 10 may be rewritten as

$$\sum_{j=1}^{q} b_j x(k-j) = b_1 \left[x(k-1) + \sum_{j=1}^{q-1} \left(\frac{b_{j+1}}{b_1} \right) x(k-1-j) \right]. \tag{12}$$

Structure II′ is defined by imposing the following two restrictions:

$$s = q - 1 \tag{13}$$

$$d_j = -b_{j+1}/b_1 \quad \text{for} \quad j = 1, 2, \ldots, q-1, \tag{14}$$

for which Eq. 12 reduces to

$$\sum_{j=1}^{q} b_j x(k-j) = b_1 \left[x(k-1) - \sum_{j=1}^{s} d_j x(k-1-j) \right]$$

$$= b_1 \left\{ \sum_{l=0}^{r-1} \tilde{c}_l N[y(k-l-l)] \right\} = \sum_{j=1}^{r} c_j N[y(k-j)], \tag{15}$$

where $c_j = b_1 \tilde{c}_{j-1}$ for $j = 1, 2, \ldots, r$. Substituting this result into Eq. 10 then yields the general NARMAX representation for Structure II′:

$$y(k) = \sum_{i=1}^{p} a_i y(k-i) + \sum_{j=1}^{q} b_j u(k-j) + \sum_{j=1}^{r} c_j N[y(k-j)], \tag{16}$$

where $r \neq q$ and $c_j \neq b_j$ in general. Finally, note that the steady-state gain of the *linear dynamics* included in the Hammerstein model (Eq. 11) is given by

$$g_2 = \frac{\sum_{j=0}^{r-1} \tilde{c}_j}{1 - \sum_{i=1}^{s} d_i}, \tag{17}$$

By substituting Eqs. 13 and 14 and the definition of c_j given before into Eq. 17, this ratio may be reduced to

$$g_2 = \frac{\sum_{j=1}^{r} c_j}{\sum_{j=1}^{q} b_j}. \tag{18}$$

Some properties of model Structures I and II′

Because Structures I and II′ provide a particularly convenient basis for the gray-box modeling procedures described in the next section, it is useful to conclude this section with a brief discussion of some of their important general characteristics. Because Structure I is a special case of Structure II′, obtained by setting $r = q$ and $c_j = b_j$ for $j = 1, 2, \ldots, q$, the following observations made for Structure II′ are also applicable to Structure I.

First, note that the NARMAX representation for Structure II′ (Eq. 16) contains only linear moving-average terms. Thus, it follows that the zero dynamics for this model (i.e., solutions of Eq. 16 to the condition $y(k) = 0$ for all k) are linear. Thus, while this model can exhibit either minimum-phase or nonminimum-phase (e.g., inverse response) behavior, it cannot exhibit more complex nonlinear zero dynamics. For this reason, the model structures considered here may not be flexible enough to describe processes with severely nonlinear zero dynamics. Another direct consequence of the inclusion of only linear moving-average terms in these model structures is that they cannot exhibit *input multiplicity*. In contrast, both the Hammerstein and Wiener models can exhibit input multiplicities if the static nonlinearities on which they are based are nonmonotonic.

Another feature of the Structure II′ family of NARMAX models is that these models can be explicitly inverted, making them attractive for use in developing IMC controllers. Specifically, note that if $b_1 \neq 0$, the control input $u(k)$ required at time step k to achieve model response $y(k+1)$ at time step $k+1$ is

$$u(k) = \frac{1}{b_1} \left\{ y(k+1) + \sum_{i=1}^{p} a_i y(k+1-i) - \sum_{j=1}^{r} c_i N[y(k+1-j)] - \sum_{l=2}^{q} b_l u(k+1-l) \right\}. \tag{19}$$

Finally, the feedback–block-oriented models introduced here can also be used for modeling nonlinear processes without steady-state multiplicities. In these cases, the model structures proposed here represent additional options beyond the more familiar Hammerstein and Wiener models for the development of simple nonlinear models. In addition, note that many chemical processes exhibit physical feedback structures (e.g., recycle streams) suggestive of the mathematical model structures considered here.

Identification Procedures

The nonlinearity in Structures I through IV necessary to achieve a particular steady-state behavior may be determined independently of the dynamics of the other elements appearing in these structures. This observation provides the basis for the following gray-box identification strategy when the steady-state behavior is known. If the physical system of interest exhibits output multiplicity but does not exhibit input multiplicity, the relationship between any given steady-state output value y_s and its associated steady-state input value u_s is unique. Here, we assume this relationship is specified by the known function:

$$u_s = f(y_s). \tag{20}$$

If we substitute this expression into Eq. 9, we obtain a general expression for the nonlinearity $N(\cdot)$, which is applicable to all four of the model structures considered in the second section:

$$N(x) = \frac{x - g_0 g_1 f(x/g_1)}{g_0 g_1 g_2}. \tag{21}$$

This equation defines the nonlinearity $N(\cdot)$, which is required to match the steady-state behavior of the process, given the steady-state gains g_0, g_1, and g_2 of the linear blocks in the model. Here, we assume all of these gains are fixed at unity for convenience.

Given this result, we are left with the task of estimating the remaining model parameters. Here, we consider this problem for Structures I and II′ discussed in the second section. For Structures III and IV and the general case of Structure II, the problem appears to be somewhat messy practically, but not intractible.

For Structure I, once we have specified the nonlinearity $N(\cdot)$, the remaining model identification task is to estimate the unknown parameters $\{a_i\}$ and $\{b_j\}$ in the model:

$$y(k) = \sum_{i=1}^{p} a_i y(k-i) + \sum_{j=1}^{q} b_j \{u(k-j) + N[y(k-j)]\}, \tag{22}$$

subject to the steady-state gain constraint $g_0 = 1$, implying

$$\sum_{j=1}^{q} b_j = 1 - \sum_{i=1}^{p} a_i. \tag{23}$$

Similarly, for Structure II′, the model-identification task is to estimate the unknown parameters $\{a_i\}$, $\{b_j\}$, and $\{c_j\}$ in the model:

$$y(k) = \sum_{i=1}^{p} a_i y(k-i) + \sum_{j=1}^{q} b_j u(k-j) + \sum_{j=1}^{r} c_j N[y(k-j)], \tag{24}$$

subject to the steady-state gain constraints $g_0 = 1$, $g_2 = 1$. These constraints imply Eq. 23 as before, but also require

$$\sum_{j=1}^{r} c_j = \sum_{j=1}^{q} b_j. \tag{25}$$

We use a constrained least-squares parameter estimation algorithm (Draper and Smith, 1981) for estimating the dynamic model parameters for both Structure I and Structure II′. We emphasize that the choice of gains of the linear models is indeed arbitrary: it is straightforward to show that for any finite nonzero gains g_0 and g_2 the resulting models will exhibit identical input/output behavior, although the model parameters will be different for different choices of these gains. Note that extreme gains (either very large or very small) should be avoided in practice to prevent poor numerical conditioning of the resulting estimation problem.

Application to a CSTR Example

The following example illustrates the ideas described earlier, demonstrating the steps required for a simple CSTR example for which the steady-state analysis is straightforward. No additional fundamental complexities are introduced in considering more realistic examples, so long as an accurate representation of the steady-state nonlinearity $u_s = f(y_s)$ can be obtained.

CSTR model

The specific model considered here assumes a first-order exothermic reaction, leading to the following dimensionless state and output equations (Uppal et al., 1974):

$$\dot{x}_1 = -x_1 + Da(1-x_1)e^{x_2/(1+x_2/\varphi)} \tag{26}$$

$$\dot{x}_2 = -x_2 + BDa(1-x_1)e^{x_2/(1+x_2/\varphi)} + C(u - x_2) \tag{27}$$

$$y = x_1, \tag{28}$$

where x_1 and x_2 are the conversion and the dimensionless reactor temperature, respectively. For the following choice of system parameters,

$$Da = 0.072, \quad \varphi = 20.0, \quad B = 8.0, \quad C = 0.3, \tag{29}$$

the CSTR exhibits output multiplicities, as shown in Figure 4.

To develop models for this process, we first generated four 1,500-point input data sequences $\{u(k)\}$ consisting of random steps, uniformly distributed in amplitude on the interval $[-1.4, 1.4]$. At each step k, $u(k)$ had a fixed probability p of assuming a new, independent value drawn from this distribution and a probability $1-p$ of remaining at the previous value $u(k-1)$. Here, we considered the four cases $p = 100\%$, $p =$

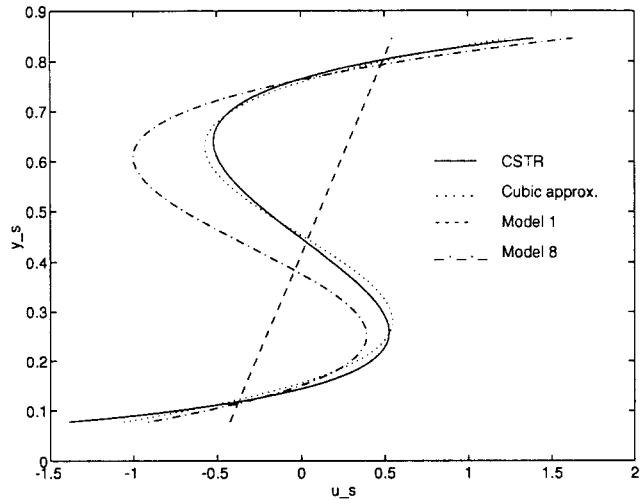

Figure 4. Steady-state process and model characteristics.

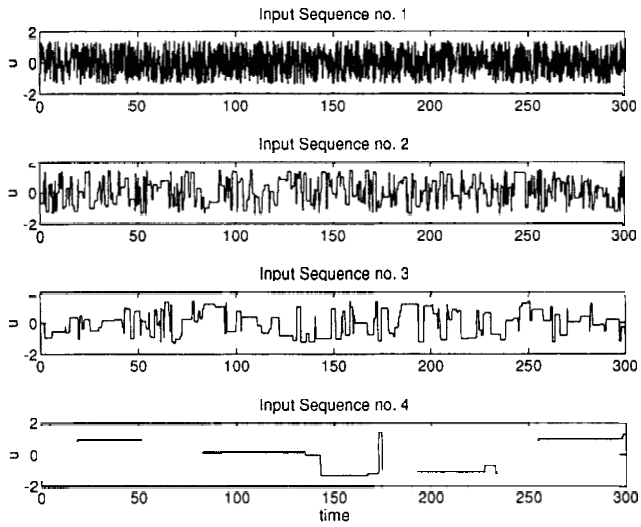

Figure 5. Input sequences for model identification.

30%, $p = 10\%$, and $p = 1\%$. Note that input sequence no. 1 ($p = 100\%$) therefore represents a "uniform white noise" sequence. The exact model responses were then computed for these sequences by numerically integrating the model equations, using a sampling period of 0.2 dimensionless time units. The four input and output data sets are shown in Figures 5 and 6. Ultimately, data set no. 3 was found to contain both reasonable amounts of information about the local process dynamics around both stable steady-states, and about the transition dynamics between the steady-state regimes. For this reason, the results presented here were obtained using data set no. 3 as the input sequence.

Gray-box models, exact nonlinearity

To develop block-oriented models, we begin by computing the steady-state characteristic $u_s = f(y_s)$, from which we obtain the feedback nonlinearity $N(y) = y - f(y)$. Setting the time derivatives to zero in the CSTR model equations and solving for $f(\cdot)$ yields

$$N(y) = y - x_2(y)\left(1 + \frac{1}{C}\right) + \frac{BDa}{C}(1-y)e^{(x_2(y))/(1+x_2(y)/\varphi)}, \tag{30}$$

where

$$x_2(y) = \frac{\log\left(\dfrac{y}{Da(1-y)}\right)}{1 - \log\left(\dfrac{y}{Da(1-y)}\right)/\varphi}. \tag{31}$$

Given this static nonlinearity, we then estimate the dynamic model parameters from input/output data (data set no. 3, as discussed above). Here, we consider Structures I and II′, leading to the two constrained linear optimization problems discussed in the third section.

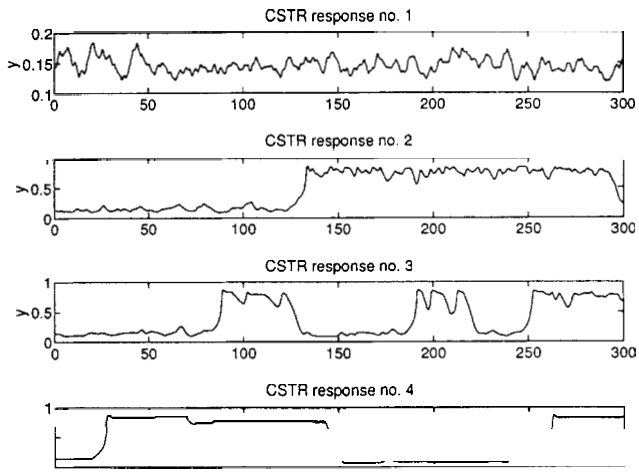

Figure 6. CSTR responses to the input sequences of Figure 5.

Gray-box models with cubic nonlinearity

Next we consider the case in which the static nonlinearity is not known exactly, but is approximated by the cubic nonlinearity:

$$\tilde{N}(x) = v_0 + v_1 x + v_2 x^2 + v_3 x^3. \tag{32}$$

The best approximation of this form to the exact nonlinearity $N(y)$ defined in Eqs. 30 and 31 was determined by minimizing the RMS (root mean square) steady-state error:

$$\text{RMS}\, E_{ss} = \left[\frac{1}{M}\sum_{k=1}^{M}(\hat{u}_k - u_k)^2\right]^{1/2},$$

where $u_k = N(y_k)$ and $\hat{u}_k = \tilde{N}(y_k)$; here, $\{y_k\}$ is a sequence of steady-state output values regularly spaced at $M = 100$ points between $y_{ss} = 0.07$ and $y_{ss} = 0.845$. The resulting approximation $\tilde{N}(x)$ is plotted in Figure 4.

Simulations

Models of Structure I and Structure II′ were identified, both for the exact nonlinearity $N(x)$ and the cubic approximation $\tilde{N}(x)$ defined earlier. In addition, two other cases were also considered for comparison: an unconstrained cubic NARMAX model and an affine model. Specifically, the cubic NARMAX model was of the form:

$$y(k) = y_0 + \sum_{i=1}^{p} a_i y(k-i) + \sum_{i=1}^{q} b_i u(k-i) + \sum_{i=1}^{r} c_i y^2(k-i) + \sum_{i=1}^{s} d_i y^3(k-i), \tag{33}$$

where no constraints were imposed on the parameters y_0, a_i, b_i, c_i, and d_i. Note that the cubic gray-box model structure described before corresponds to a restricted special case of

this more general model structure. For this reason, we would expect this more general model to fit the available input/output data with greater accuracy, but describe the steady-state process behavior less accurately. In particular, the primary difference between the feedback-block-oriented models and this cubic NARMAX model is that no *independent* source of data is being used to obtain a reasonable approximation to the steady-state response: steady-state behavior must be estimated from the same response data as that used to model the dynamic response of the process. Note that the steady-state behavior of the cubic NARMAX model defined in Eq. 33 is defined by the roots of the cubic equation:

$$\left(\sum_{i=1}^{s} d_i\right) y_s^3 + \left(\sum_{i=1}^{r} c_i\right) y_s^2 + \left(\sum_{i=1}^{p} a_i - 1\right) y_s + \left(y_0 + u_s \sum_{i=1}^{q} b_i\right) = 0. \quad (34)$$

In addition, because it does represent a common standard of comparison, we have also considered an affine model, corresponding to Eq. 33 with $r = s = 0$.

Altogether, eleven different models were compared, and these models are summarized in Table 1. This table specifies the structure of each model and lists the order parameters p, q, r, and s appearing in the defining equations (Eq. 4 for Structure I, Eq. 16 for Structure II′, and Eq. 33 for the affine and cubic empirical models). To facilitate comparisons, the number of degrees of freedom available to describe both steady-state and dynamic behavior is also listed in this table. Note that the use of the exact nonlinearity $N(x)$ in Structure I and Structure II′ corresponds to an infinite number of degrees of freedom, while the four parameters of the cubic approximation v_0, v_1, v_2, and v_3 represent four degrees of freedom when we use the approximate nonlinearity $\tilde{N}(x)$. The number of dynamic degrees of freedom in the Structure I and Structure II′ models is determined by both the order parameters and the number of steady-state gain constraints. Specifically, the number of degrees of freedom available to describe the process dynamics for Structure I is $p + q - 1$, while the number of dynamic degrees of freedom for Structure II′ is $p + q + r - 2$. For the cubic model defined in Eq. 33, it may be argued that the four coefficients enclosed in parentheses in the steady-state equation (Eq. 34) also represent four static degrees of freedom, even though they involve all of the model parameters. The total number of parameters in this model is $p + q + r + s + 1$, due to the presence of the constant term y_0; consequently, the number of dynamic degrees of freedom assigned to this model will be $p + q + r + s - 3$. In the affine special case of this model ($r = s = 0$), since this constant term influences only the steady-state behavior, it is reasonable to count it as one static degree of freedom. In addition, the steady-state gain represents a second static degree of freedom, leaving $p + q - 1$ dynamic degrees of freedom. The key point of this table is to emphasize that while we cannot argue that the complexity of the models considered is exactly the same in all cases, it is arguably comparable, so we are not attempting to compare models of radically differing complexity.

Table 1. Descriptions of the Eleven NARMAX Models

Model No.	Description	Order Parameter p	q	r	s	Dynamic d.o.f.	Static d.o.f.
1	Affine	4	4	0	0	7	2
2	Structure I, exact	4	4	—	—	7	∞
3	Structure I, cubic	4	4	—	—	7	4
4	Structure II′, exact	3	3	3	—	7	∞
5	Structure II′, cubic	3	3	3	—	7	4
6	Empirical A	3	2	3	2	7	4
7	Empirical B	2	3	2	3	7	4
8	Empirical C	2	2	3	3	7	4
9	Empirical D	3	3	2	2	7	4
10	Empirical E	3	3	3	3	9	4
11	Structure I, low order	2	2	—	—	3	∞

Comparative results

All eleven models were determined for data set no. 3, and a comparison of several of these models is shown in Figure 7. In particular, Figure 7 contains the model responses of a Structure I model (Model 2), a Structure II′ model (Model 4), and two empirical models: the affine model (Model 1) and one of the cubic NARMAX models (Model 8). With the exception of Model 1, all models appear to provide reasonable approximations to the process behavior for this particular input sequence.

In order to establish a systematic basis for model comparison, a thorough model validation was performed next. For each of the eleven model structures listed in Table 1, 450 independent model validation runs of a duration of 250 sampling intervals were performed, each based on an input sequence $\{u(k)\}$ of the type used for data set no. 3 (switching probability of 10%). The overall results for these simulations are summarized in Figure 8. There, box plots (Hoaglin et al., 1991) of the RMS model prediction error (RMSE) are plotted vs. the model number listed in Table 1. The minimum RMSE observed in each of the 450 simulations is indicated by the small horizontal line at the bottom of each box plot, and the maximum RMSE observed is indicated by the small horizontal line at the top of each box plot. The larger hori-

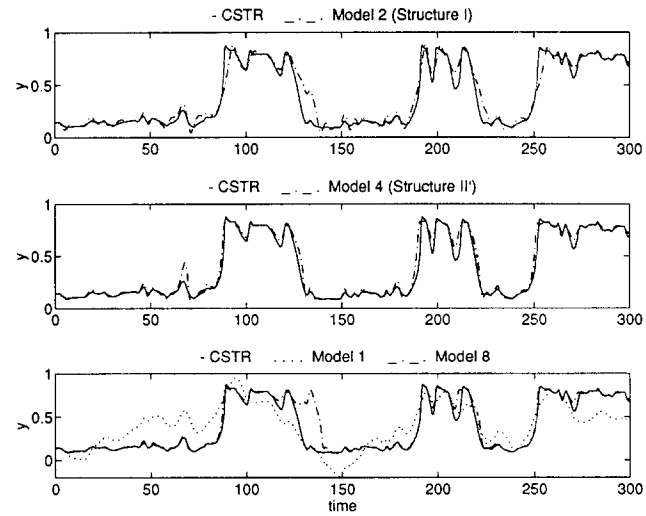

Figure 7. Representative model-identification results.

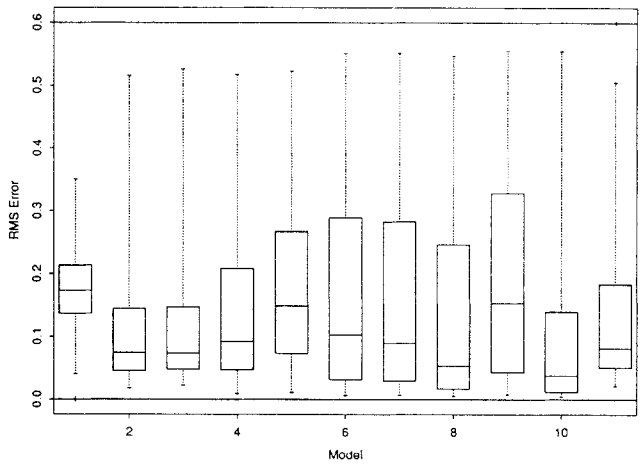

Figure 8. Summary of all 450 model validation runs.

zontal line at the center of each box represents the median (50th percentile) RMSE for the 450 simulations, while the larger lines at the top and bottom of each box represent the upper and lower quartiles (75th and 25th percentiles), respectively.

Initially, these results appear quite puzzling and highly variable. Viewed in terms of median RMS error, all of the nonlinear models considered here outperform the affine model (Model 1), with Model 10 (the cubic NARMAX model with 13 total degrees of freedom) performing best. On the other hand, based on the interquartile distances (i.e., the height of the central boxes in each plot), all of the nonlinear models appear to be more highly variable in their performance than the affine model. Even more troubling, the worst-case errors for each of the nonlinear models is significantly larger than for the affine model. Closer examination of the simulation results reveals a fundamental difficulty that is likely to arise in any attempt at empirical modeling of systems exhibiting output multiplicity: Figure 9 shows a plot of the input and output data generated by one of the CSTR simulation runs, along with the predicted outputs $\hat{y}(k)$ generated by Models 2 and 8 for the same input sequence $\{u(k)\}$.

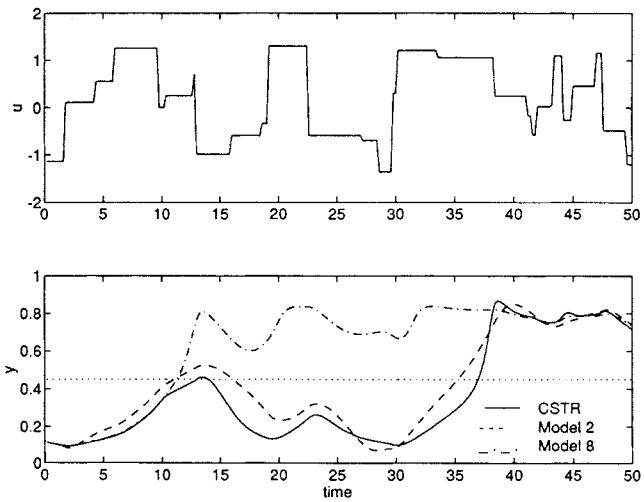

Figure 9. Transition between steady-state operating regions.

Note that at $t \sim 11$, the Model 8 response $\hat{y}(k)$ apparently "switches" from a lower stable steady state to an upper one, but the CSTR simulation as well as the Model 2 response do not "switch" until much later ($t \sim 37$). Consequently, the prediction errors for Model 8 are quite large during this "early transition," and these errors dominate the overall prediction error for this simulation run. It is important to emphasize that this behavior is in no way limited to the Structure I or Structure II′ models, as demonstrated by the fact that *all* of the nonlinear model structures considered here sometimes exhibited this type of behavior.

In light of this observation, we partitioned the simulation runs into "valid" and "invalid" runs, based on whether the model predictions $\hat{y}(k)$ did or did not follow all of the steady-state transitions exhibited by the CSTR simulation for the given input sequence. Specifically, a run was declared "invalid" if the model predictions remained in the incorrect steady state for more than 30 sample times k. Here, both the CSTR simulation and the model predictions were declared to lie in the domain of the lower stable steady state if $y(k)$ or $\hat{y}(k)$ were at most 0.45, and they were declared to be in the domain of the upper steady state otherwise. Note that $y(k) = 0.45$ approximately corresponds to the midpoint of the unstable steady-state branch of Figure 4. The box plots obtained from the "valid" runs only are plotted in Figure 10. It is clear from this plot that all of the nonlinear models considered here provide better predictions than the affine model (Model 1) does, by almost any criterion we could consider.

These results are summarized in more detail in Table 2, which ranks the eleven models considered here by four different criteria. The first of these criteria is the RMS steady-state error, summarized for each model type in the second and third columns of this table. The model structures introduced in the second section exhibit the best possible performance with respect to this criterion, since their steady-state error is identically zero by construction.

The fourth and fifth columns of Table 2 give an analogous rank ordering in terms of the RMS error observed for the "valid" runs only. In terms of this ranking, Model 10 (Empirical E) performs the best, consistent with the fact that it has the largest number of parameters (13) and no explicit constraints. In fact, the five unconstrained cubic NARMAX

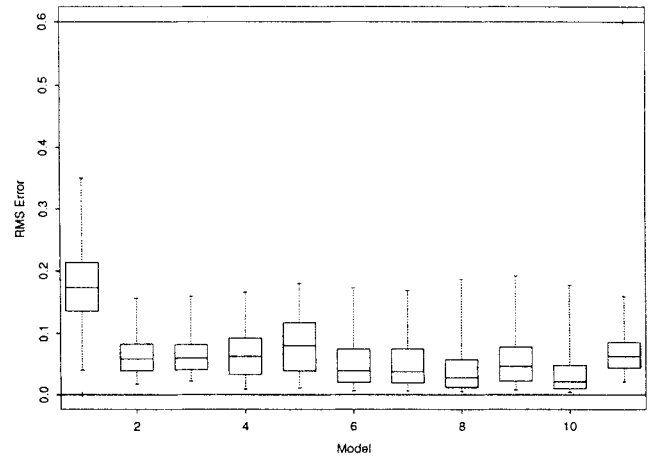

Figure 10. Summary of "valid" model validation runs.

Table 2. Four Rankings of the Eleven NARMAX Models

	$RMSE_{ss}$, (Steady-State)		Median *RMSE*, Valid Runs		Percent Invalid Runs		Median *RMSE*, Invalid Runs	
Rank	Model	Value	Model	Value	Model	Value	Model	Value
1	2	0	10	0.021	1	0.00	1	0.172
2	4	0	8	0.027	10	23.6	2	0.265
3	11	0	7	0.037	2	25.6	3	0.276
4	3	0.072	6	0.039	3	26.7	5	0.277
5	5	0.072	9	0.046	11	30.9	4	0.282
6	10	0.253	2	0.058	8	31.1	11	0.283
7	8	0.315	3	0.059	4	33.6	10	0.330
8	6	0.334	11	0.061	7	38.0	9	0.333
9	7	0.342	4	0.062	6	40.0	7	0.335
10	9	0.385	5	0.080	5	44.7	6	0.338
11	1	0.559	1	0.172	9	48.4	8	0.342

models considered here (Models 6 through 10) as a group offer the best prediction-error performance, followed by the feedback–block-oriented models as a group, with the affine model exhibiting the poorest performance. These results are exactly what we expect: the feedback–block-oriented models have approximately four fewer degrees of freedom to capture the process dynamics than the unconstrained empirical models do, leading to about a factor of 2 in performance degradation. In exchange, the feedback–block-oriented models offer significantly better steady-state performance. Further, note that the Structure I models considered here (Models 2, 3 and 11) appear to perform slightly better than the Structure II′ models, suggesting that the greater structural flexibility of Structure II′ is not a significant advantage in this particular application. As expected, all of the nonlinear models considered here greatly outperform the affine model in terms of dynamic behavior *provided the nonlinear model does not miss transitions between stable steady states.*

This point is emphasized in the last two columns of Table 2, which ranks the models considered here in terms of RMS prediction error for those "invalid" runs where transitions were missed. Here, the opposite ordering is observed as for the "valid" runs just considered. That is, the affine model performs the best in this case, followed by the feedback–block-oriented models, followed by the unconstrained cubic NARMAX models. In fact, these results should not be surprising: the affine model cannot capture the output multiplicity of the CSTR, and can at best adopt a "middle ground," attempting to approximate both upper and lower stable steady states with a nonexistent intermediate one. Thus, while the affine model cannot approximate either steady state well (consistent with its poor performance for the "valid" simulation runs), neither can it remain too long in the domain of the *wrong* steady state. The dynamic performance seen in the "invalid" runs is therefore worse for the nonlinear models—which *can* remain too long in the domain of the wrong steady state—than for the affine model. Further, because the unconstrained cubic models (Models 6 through 10) are inherently poorer approximations to these steady states, the degree of mismatch when a transition is missed can be even greater than in the feedback–block models that accurately describe the steady-state behavior.

Closely related to this last point, another important measure of nonlinear model performance is its tendency to miss transitions between steady states. The fraction of simulation runs that were declared "invalid" by the criteria described earlier represents one quantitative measure of this tendency. A model ranking based on these results is given in columns 6 and 7 of Table 2. Adopting the view taken before that the affine model does not model either stable steady state well but does not miss transitions between steady states, this model is ranked first. The best of the nonlinear models is Model 10, the unconstrained cubic NARMAX model with 13 total parameters. This model is arguably the most complex one considered here, and this result suggests it is able to use these parameters to capture *transition dynamics*—the qualitative behavior of transitions between steady states—better than the other unconstrained cubic NARMAX models. Following Model 2, the next best transition behavior is seen in the Structure I models, particularly Models 2 and 3, which perform almost as well as Model 10 with respect to missed transitions.

Summary and Conclusions

This article has described a semiphysical gray-box modeling procedure applicable to processes with output multiplicities. This modeling procedure assumes the steady-state behavior is known *a priori* and constructs a NARMAX model whose steady-state behavior agrees exactly with that of the process, and whose dynamic response provides a best fit to available input/output data. The key to this procedure is the class of feedback–block-oriented NARMAX models described in the second section, which permit the steady-state behavior to be specified independently of the dynamic behavior. This general approach was discussed in the third section, including detailed identification algorithms for two specific structures (Structures I and II′). These discussions exploit the fact that these two model structures have simple NARMAX representations; in fact, it is interesting to note that both of these representations are special cases of the NARMAX class of models proposed by Chen and Tsay to facilitate the use of nonparametric techniques (Chen and Tsay, 1993; Pearson, 1995).

The fourth section presented a summary of 450 model validation runs for each of the 11 NARMAX model structures, including examples of Structures I and II′, along with unconstrained cubic NARMAX models and an affine model for comparison. An important result to emerge from these comparisons was the significance of *transition dynamics*—the

dynamic behavior seen in transition between stable steady states—as a figure of merit for empirical models of processes that can exhibit output multiplicity. In particular, one of the effects of plant–model structural mismatch in this example (i.e., CSTR dynamic model vs. NARMAX model) was that of missed transitions between stable steady states: even if both the steady-state behavior is completely correct and the local dynamics around these steady states are approximated reasonably well, a sequence $\{u(k)\}$ that causes the CSTR model to switch from one steady state to another at time k_s need not cause the prediction model to follow this transition. In retrospect, this result should not be surprising: Transitions between stable steady states occur rather infrequently in the data sets considered here, so there is little information from which to accurately model transition dynamics. Because input sequences with different characteristics can exhibit great differences in their tendency to cause such transitions in the CSTR model, it is clear that input sequence design is an important issue here.

In view of these observations, the issue of transition dynamics appears to be both a fundamental one in modeling systems with multiple steady states, and an inherently difficult one to address. A possibility suggested by one of the reviewers of this article is to extend the "two-step" methodology proposed here (i.e., steady-state and dynamic characterization) to a "three-step" methodology:

1. Steady-state characterization
2. Local dynamic characterization
3. Transition dynamic characterization

In particular, Steps 2 and 3 would be based on different input sequences: those used in Step 2 would be designed to excite local dynamics (possibly by restricting the range of excitation), while those used in Step 3 would be designed to cause transitions between steady states. In any case, it is important to emphasize that, if we are attempting to develop a single empirical model (or partially empirical model) over a wide enough operating range to include multiple steady states, it is necessary to drive the process "hard enough" (i.e., over a wide enough range of inputs) to exhibit its full range of dynamic behavior. If such excitation of the process is not possible in a given practical situation, we must be content with alternative approaches like linear multimodeling (Bannerjee et al., 1995; Johanson and Foss, 1993). Even in this case, however, the issue of transition dynamics remains: How are the local models to be identified and coordinated to adequately approximate transitions between local model regimes?

We propose to explore a number of extensions of the ideas presented here, including the following. First, it is possible to extend the four model structures considered here by replacing the linear dynamic models on which they are based with linear multimodels (Bannerjee et al., 1995). Other extensions will address the multivariable modeling problem. Finally, while the emphasis of this article has been on gray-box modeling problems in which the steady-state operating characteristics are known, it would also be interesting to consider extensions to cases where $F(\cdot)$ is not known and must also be estimated from data, particularly using nonparametric approaches (Greblicki and Pawlak, 1989; Greblicki, 1992).

Literature Cited

Arbel, A., I. H. Rinard, and R. Shinnar, "Dynamics and Control of Fluidized Catalytic Crackers: 2. Multiple Steady States and Instabilities," *Ind. Eng. Chem. Res.*, **34**, 3014 (1995).

Bannerjee, A., Y. Arkun, R. K. Pearson, and B. A. Ogunnaike, "H_∞ Control of Nonlinear Processes Using Multiple Linear Models," European Control Conf., Rome, Italy (1995).

Bekiaris, N., G. E. Meski, C. M. Radu, and M. Morari, "Multiple Steady States in Homogeneous Azeotropic Distillation," *Ind. Eng. Chem. Res.*, **32**, 2023 (1993).

Billings, S. A., and W. S. F. Voon, "A Prediction-Error and Stepwise-Regression Estimation Algorithm for Non-Linear Systems," *Int. J. Control*, **44**, 803 (1986).

Chen, R., and R. S. Tsay, "Nonlinear Additive ARX Models," *J. Amer. Stat. Assoc.*, **88**, 995 (1993).

Draper, N. R., and H. Smith, *Applied Regression Analysis*, Wiley, New York (1981).

Greblicki, W., and M. Pawlak, "Nonparametric Identification of Hammerstein Systems," *IEEE Trans. Inform. Theory*, **IT-35**, 409 (1989).

Greblicki, W., "Nonparametric Identification of Wiener Systems," *IEEE Trans. Inform. Theory*, **IT-38**, 1487 (1992).

Hoaglin, D. C., F. Mosteller, and J. W. Tukey, *Fundamentals of Exploratory Analysis of Variance*, Wiley, New York (1991).

Jacobsen, E. W., and S. Skogestad, "Multiple Steady States in Ideal Two-Product Distillation," *AIChE J.*, **37**, 499 (1991).

Johansen, T. A., and B. A. Foss, "Constructing NARMAX Models Using ARMAX Models," *Int. J. Control*, **58**, 1125 (1993).

Johansen, T. A., "Identification of Non-Linear Systems Using Empirical Data and Prior Knowledge—An Optimization Approach," *Automatica*, **32**, 337 (1996).

Pearson, R. K., "Nonlinear Input/Output Modeling," *J. Process Control*, **5**, 197 (1995).

Pearson, R. K., *Discrete-Time Dynamic Models*, in preparation.

Sjoberg, J., Q. Zhang, L. Ljung, A. Benveniste, B. Delyon, P. Glorennec, H. Hjalmarsson, and A. Juditsky, "Nonlinear Black-Box Modeling in System Identification: A Unified Overview," *Automatica*, **31**, 1691 (1995).

Uppal, A., W. H. Ray, and A. B. Poore, "On the Dynamic Behavior of Continuous Stirred Tank Reactors," *Chem. Eng. Sci.*, **29**, 967 (1974).

Vidyasagar, M., *Nonlinear Systems Analysis*, Prentice Hall, Englewood Cliffs, NJ (1978).

Von Seggern, D. H., *CRC Handbook of Mathematical Curves and Surfaces*, CRC Press, Boca Raton, FL (1990).

Manuscript received Feb. 13, 1997, and revision received July 21, 1997.