Part II. Application to a Chemical Reactor

K. WATANABE and
D. M. HIMMELBLAU

Department of Chemical Engineering
The University of Texas
Austin, TX 78712

A two-level strategy for fault detection and diagnosis developed in Part I is applied to a chemical reactor in which heptane is converted to toluene. Simulation of various faults demonstrates that the proposed strategy is valid, and that it also represents an improvement over fault diagnosis via an extented Kalman filter.

SCOPE

A method of fault detection and diagnosis for nonlinear processes was proposed in Part I. The basic strategy comprises two-levels. The first-level consists of estimation of the state of a nonlinear process by a linear state estimator without any approximation. The second-level consists of identification of the parameters associated with faults in the nonlinear process by using the estimated state vector from the first-level. To be effective for fault detection (as contrasted with process control) methods for state estimation and parameter identification must without question yield unbiased estimates in the presence of abnormal operating conditions.

An example of the application of the proposed strategy is presented for an aromatization process for which the process model consists of coupled nonlinear ordinary differential equations. All of the details of the mathematical relations for the filter design and least squares estimation are illustrated for a reaction in which heptane is coverted to toluene. Results are compared with those from an extended Kalman filter.

CONCLUSIONS AND SIGNIFICANCE

Simulated causes of faults including change in the heat transfer coefficient, preexponential factor, and activation energy have been introduced into a chemical reactor. Simulated time records including the presence of noise were used to diagnose the causes of faults via the two-level proposed strategy.

We found that the extended Kalman filter needed substantially more computation time (more by a factor of 500) than did the proposed method hence the proposed method is much more favorable for implementations in real time on-line analysis. In addition, we observed that the extended Kalman filter on occasion led to biased estimates whereas the proposed method did not, hence the Kalman filter can be regarded as being less robust for incipient fault diagnosis even though it may be entirely adequate for process control. On the other hand, the proposed method does have some higher variances associated with its estimates than does the Kalman filter.

One contribution we have made is to show how the linear state space approach to estimation can be applied successfully to a class of nonlinear processes in the face of common understanding that the state space approach is not very satisfactory for processes with high degrees of nonlinearity of interest to chemical engineers.

INTRODUCTION

Process fault detection and diagnosis is important from both a theoretical and practical viewpoint. A number of modern state estimation and parameter identification techniques have been applied to unravel the pitfalls that exist but these techniques often involve a very heavy computational load for diagnosis or yield inaccurate diagnosis as discussed in Part I in the historical background.

In Part I, we developed a combined state estimation and parameter identification strategy. Our linear state estimation filter can reconstruct the correct state variables of nonlinear processes with unknown coefficients, an outcome that cannot be achieved by the Luenberger observer and/or the Kalman filter. Because we reconstruct all of the state variables, estimates of the fault coefficients can be made with minimum computation.

Here we apply the combined state estimation and parameter identification strategy to identify simultaneously more than one cause of a fault(s). As an example, we first model an aromatization process, institute one or two faults, and observe the deteriorating performance of the reactor as a function of time.

Typical fault modes of the reactor might be

(a) a decrease in the heat transfer coefficient caused by surface fouling,

(b) a change in flow rate caused by partial pipe plugging

K. Watanabe is on leave from Hosei University at Koganei, Tokyo, Japan.

0001-1541-83-6728-0250-$2.00.

(c) a change in effective reactor volume caused by malfunction of the agitator.

Pipe plugging and agitator malfunction most likely can be detected by observations external to the reactor whereas changes in the heat transfer coefficient ordinarily must be calculated approximately from an overall energy balance, hence are not easily and uniquely retrievable. The identification of the overall heat transfer coefficient is not easy because the coefficient is only weakly sensitive to the observed process variables so that its estimates are likely to be corrupted by measurement noise. Keep in mind that the fault diagnosis is to be carried out using the ordinary operating record of the reactor without the introduction of special perturbations in the process inputs.

In addition to changes in the heat transfer coefficient, we take into account changes in the preexponential factor and the energy of activation as fault modes. The difficulty of identifying these model coefficients and the reasons therefore are well known among chemical engineers.

In this paper, we carry out a fault diagnosis of the aromatization process following the procedure described in Part I. The step-by-step procedure is as follows:

(1) Build the process model to conform with the same structure as Eqs. 1 in Part I.

(2) Isolate the set of faults that cannot be detected by simple methods.

(3) Design the state estimation filter which reconstructs the correct state variables of the modelled nonlinear process including the set of unknown fault coefficients. The filter (Eqs. 2) in Part I can be designed by the procedure outlined by Eqs. 4, 5 or 7 in Part I if the two conditions 3a and 3b in Part I are satisfied. Alarm decisions can be made by checking whether or not all the state variables are within their specified bounds.

(4) Design the fault parameter estimator to diagnose the causes of faults following the design procedure given by Eqs. 15 and 16 in Part I. The estimator can be designed if the two conditions (14a) and (14b) of Part I are satisfied.

HEPTANE TO TOLUENE AROMATIZATION PROCESS

Although the theory for fault diagnosis developed in Part I was for a model nonlinear in the state and linear in the coefficients, we have applied the theory in the example below to a process model nonlinear in the coefficients by means of special formulation of the model.

Consider a heptane to toluene aromatization process via a catalyst, Figure 1. The reaction is

$$C_7H_{16} \xrightarrow{k} C_7H_8 + 4H_2 \tag{1a}$$

The reaction rate constant is assumed to follow the Arrhenius form

$$k(T) = k_o e^{-E_a/RT} \tag{1b}$$

where T is the reaction temperature, and the frequency factor and activation energy are determined by the main reaction 1a and the catalyst; the coefficients are as follows (Burnett et al., 1965):

k_o = frequency factor, 5.01×10^8 [1/h]
E_a = activation energy, 1.369×10^5 [J/gmol]
R = gas constant, 8.319 [J/(gmol)(K)]

The heat of reaction ΔH and the specific heat C_p are given as functions of temperature, but only ΔH will be considered as a function of temperature because the change of C_p with temperature is slight in the range of interest here:

$$\Delta H = 2.2026 \times 10^5 + 6.20244 \times 10^1 T - 5.536 \times 10^{-2} T^2 - 1.15 \times 10^{-6} T^3 + 3.1496 \times 10^{-7} T^4 \text{ [J/gmol]}$$
$$C_p = 490.7 \text{ [J/(gmol)(K)]}$$

The other process parameters are assumed to be constant

ρ = density, 593.0, [gmol/m^3]
h = overall heat transfer coefficient, 6.069×10^5 [J/(gmol)(K)]
a = area of heat exchange, 10.0 [m^2]
V = effective reactor volume, 30.0 [m^3]
q = inlet and outlet volumetric flow rate, 3.0 [m^3/h]

The process variables at a steady state operating point are given as follows

T_i = inlet temperature (preheated), 600.0 [K]
T_h = heating temperature, 850.0 [K]
T = reaction temperature, 742.0 [K]
Inlet concentration of C_7H_{16}, 1,000.0 [gmol/m^3]
Outlet concentration of C_7H_{16}, 463.0 [gmol/m^3]
Outlet concentration of C_7H_8, 537.0 [gmol/m^3]
Outlet concentration of H_2, 2,149.0 [gmol/m^3]

Under the normal operations, the above process variables are bounded as follows

$550.0 \le T_i \le 650.0$
$800.0 \le T_h \le 950.0$
$700 \le T \le 850.0$
$900.0 \le$ (inlet concentration of C_7H_{16}) $\le 1{,}100.0$
$500.0 \le$ (outlet concentration of C_7H_8) ≤ 700.0
$2{,}000.0 \le$ (outlet concentration of H_2) $\le 2{,}800.0$
$300.0 \le$ (outlet concentration of C_7H_{16}) ≤ 500.0

We assume that the process is operating near the steady state operating point as a normally operating process, and that all the state variables do not deviate too much from the listed values above. Then we are able to diagnose the process without using special test signals to excite the process strongly.

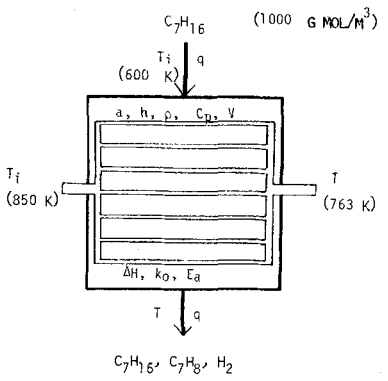

Figure 1. Heptane to toluene aromatization process.

PROCESS MODEL

Let

$x_1 \triangleq T$
x_2 = concentration of C_7H_8
x_3 = concentration of H_2
x_4 = concentration of C_7H_{16}
$u_1 \triangleq T_i$
u_2 = inlet concentration of C_7H_{16}
$u_3 \triangleq T_h$

and let w_1, w_2 and w_3 be system noises disturbing u_1, u_2 and u_3, respectively. Then a lumped model of the aromatization process is given as follows:

Energy Balance

$$\dot{x}_1 = \frac{q}{V}\{(u_1 + w_1) - x_1\} - \frac{\Delta H}{\rho C_p} k(x_1)x_4 + \frac{ah}{\rho C_p V}\{(u_3 + w_3) - x_1\}, \; x_1(0) = x_{10} \tag{2a}$$

$$\dot{x}_2 = -\frac{q}{V}x_2 + k(x_1)x_4,\ x_2(0) = x_{20} \tag{2b}$$

$$\dot{x}_3 = -\frac{q}{V}x_3 + 4k(x_1)x_4,\ x_3(0) = x_{30} \tag{2c}$$

$$\dot{x}_4 = -\frac{q}{V}x_4 - k(x_1)x_4 + \frac{q}{V}(u_2 + w_2),\ x_4(0) = x_{40} \tag{2d}$$

Not all the above equations are independent. Further define

$$\underline{x} = [x_1x_2x_3x_4]^T \tag{3a}$$

$$\underline{u} = [u_1u_2]^T \tag{3b}$$

$$\underline{w} = [w_1w_2]^T \tag{3c}$$

$$\underline{f} = \left[\frac{ah}{\rho C_p V}\{(u_3 + w_3) - x_1\} \quad \frac{\Delta H}{\rho C_p}k(x_1)x_4\right]^T \tag{3d}$$

$$A \triangleq \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \triangleq \left[\begin{array}{cc:cc} -q/V & 0 & 0 & 0 \\ 0 & -q/V & 0 & 0 \\ \hdashline 0 & 0 & -q/V & 0 \\ 0 & 0 & 0 & -q/V \end{array}\right] \tag{3e}$$

$$B \triangleq \begin{bmatrix} q/V & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & q/V \end{bmatrix},\ Q \triangleq \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} \triangleq \left[\begin{array}{cc} 1 & -\Delta H/\rho C_p \\ 0 & 1 \\ \hdashline 0 & 4 \\ 0 & -1 \end{array}\right] \tag{3f}$$

$$E \triangleq B$$

Let us assume that two output measurements, the reaction temperature and the outlet concentration of C_7H_8, each with zero mean observation noises $\underline{v}$, are available. Then the process model (Eqs. 2) can be expressed more compactly as

$$\dot{\underline{x}} = A\underline{x} + B\underline{u} + Q\underline{f} + E\underline{w},\quad \underline{x}(0) = \underline{x}_0 \tag{4a}$$

$$\underline{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = [I_2{:}0]\underline{x} + \underline{v} \tag{4b}$$

which has the same form as Eqs. 1a and 1b in Part I.

The sources of faults of the process considered here are presumed to be

(a) Incomplete reaction
(b) Deterioration of catalyst performance
(c) Fouling of the heat exchange surface

Detection of (a), incomplete reaction, can be categorized as an alarm decision and can be achieved by estimating the inaccessible state variables x_3 and x_4. Detection of (b), deterioration of catalyst performance, and (c), fouling of the heat exchange surface, are categorized as fault isolation and/or estimation and must be obtained by identifying the fault parameters in Eq. 4a. Other possible sources of faults that might be included for detection are

- Partial pipe plugging
- Misfeeding of the inlet component

but these can be detected relatively easily and inexpensively by measuring flow rates or concentrations via standard instruments. We shall describe how to detect items (a)–(c) assuming the other faults have been identified by other methods.

DESIGN OF STATE ESTIMATOR AND APPLICATION TO ALARM DECISIONS

First, we assume that two inputs and two outputs are measurable ($m = 2$), i.e.: (a) the concentration of the entering component C_7H_{16} and the inlet temperature are measurable; (b) the concentration of the product C_7H_8 and the temperature in the vessel (i.e., of the product) are measurable; (c) ΔH is constant; and (d) only the space velocity q/V of the process is known. We will now design a state estimator to provide estimates of the concentrations of C_7H_{16} and H_2 in the reactor.

From Eq. 3a in Part I, the rank of Q_1 is

$$\text{rank}\,[Q_1] = \text{rank}\begin{bmatrix} 1 & -\Delta H/\rho C_p \\ 0 & 1 \end{bmatrix} = 2(= l = m) \tag{5a}$$

Then the process (Eqs. 4) satisfies the condition 3a (or 5a or 7a) in Part I. Further from Eq. 7c in Part I, $A_{12} = 0$, so that

$$F = A_{22} = \begin{bmatrix} -q/V & 0 \\ 0 & -q/V \end{bmatrix} \tag{5b}$$

which is a stable matrix with all negative eigenvalues. Therefore process Eqs. 4 satisfies the conditions 7a–7c in Part I by which the state estimator can be designed.

Following Eqs. 4b–4g in Part I, we obtain the coefficient matrices of the state estimator Eq. 2 in Part I:

$$L = \begin{bmatrix} 0 & 4 \\ 0 & -1 \end{bmatrix},\ K = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},\ M = \begin{bmatrix} 0 & -4 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \tag{5c}$$

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 4 & -1 \end{bmatrix}^T,\ D = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}^T \tag{5d}$$

and we have the following state estimation filter

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} -q/V & 0 \\ 0 & -q/V \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & q/V \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix},\quad \begin{matrix} z_1(0) = z_{10} \\ z_2(0) = z_{20} \end{matrix} \tag{6a}$$

$$\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \\ \hat{x}_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 4 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \tag{6b}$$

which can be rewritten as

$$\dot{z}_1 = -(q/V)z_1,\qquad z_1(0) = z_{10} \tag{6c}$$

$$\dot{z}_2 = -(q/V)(-z_2 + u_2),\qquad z_2(0) = z_{20} \tag{6d}$$

$$\hat{x}_1 = y_1,\quad \hat{x}_2 = y_2,\quad \hat{x}_3 = z_1 + 4y_2,\quad \hat{x}_4 = z_2 - y_2 \tag{6e}$$

Note that in order to estimate just the outlet concentrations of C_7H_{16} and H_2 from Eqs. 6c–6e, the necessary measurements are the concentrations of the inlet component C_7H_{16} and the concentration of the product C_7H_8. Further note that the quantity ΔH is not involved in designing the state estimation filter. Thus, no matter what the value of ΔH may be as a function of temperature, the filter design does not change because the inaccessible state variables are reconstructed using only information about y_1, y_2, and u_2.

Fault (a) can be detected by checking whether the ratio of the inlet and outlet concentrations of C_7H_{16} is larger than a prespecified ratio r^* or not. In other words, if $(x_4/u_2) \geq r^*$, then some fault(s) have occurred; otherwise the condition is normal. The alarm decision is made using variables u_2 and x_4, and these variables can be obtained from continuous measurements of u_2 and x_1, and the one parameter q/V as described above.

The source of the faults cannot be identified from (x_4/u_2). This task is carried out by identification of the fault parameters to be described next.

DESIGN OF PARAMETER IDENTIFIER AND APPLICATION TO FAULT DIAGNOSIS

We first consider how the faults (b) and (c) influence the fault parameters. We will classify the deterioration of the catalyst performance into two cases:

- A decrease in the frequency factor k_o caused by loss of effective area or surface change (physical degradation)
- a decrease in k_o and an increase in the activation energy E_a caused by chemical degradation of the catalyst

The fouling of the heat exchange surfaces causes

- a decrease in the overall heat transfer coefficient h. It is known that because of Eq. 1b the simultaneous identification of valid values of E_a and k_o is quite difficult if the reaction temperature

does not change drastically. Keep in mind that for the case considered here the reaction temperature is reasonably constant (near the steady state), and we are using the operating process record without special disturbances so that we are treating a most difficult problem. What we do is to identify a parameter Δk_o and the overall fault parameter k_{of} which is a function of deviations of E_a and k_o. Let ΔE_a and Δk_o be deviations of E_a and k_o, respectively. Then the fault parameter k_{of} is obtained from Eq. 1b for a degraded catalyst as follows

$$k(T) = \{(k_o + \Delta k_o)e^{-\Delta E_a/RT}\} \cdot e^{-E_a/RT} \triangleq k_{of}e^{-E_a/RT} \tag{7}$$

Therefore, assuming E_a is held constant, the fault parameter k_{of} is considered to be an equivalent frequency factor for fault diagnosis with respect to changes in E_a and k_o.

Our formulation does not provide values of the parameter deviations of ΔE_a and Δk_o exept in cases in which it is known that $\Delta E_a = 0$. However, from the identified fault parameter k_{of} from Eq. 7, we can evaluate the overall reaction rate constant $k(T)$ if we know the value of the temperature. The deviation of $k(T)$ from its normal value informs us of the catalyst degradation. Furthermore, from Eq. 7 if $\Delta k_o < 0$ and $\Delta E_a = 0$, i.e., the fault is ascribed to physical degradation of the catalyst, the identified parameter is nothing more than the decreased frequency factor which is independent of changes in temperature. If $\Delta k_o < 0$ and $\Delta E_a > 0$, the fault is ascribed to physical and/or chemical degradation of the catalyst, and the identified parameter k_{of} changes as the reaction temperature changes. Therefore we can decide that physical degradation occurs if the identified parameter k_{of} is independent of changes in the reaction temperature; otherwise it is a chemical degradation.

Next we design the identifier for the two fault parameters discussed above assuming that the values of parameters under normal operation are known. Since all the state variables have been made accessible already by using the estimator Eqs. 6c–6e, identification of fault parameters is simplified.

Let

$$h = h^* + \Delta h \tag{8a}$$

$$k_{of} = k_o + \{\Delta k_o e^{-\Delta E_a/RT} + k_o(e^{-\Delta E_a/RT} - 1\} = k^*_{of} + \Delta k_{of} \tag{8b}$$

where h^* and k^*_{of} are known normal values of h and k_o respectively, and Δh and Δk_{of} are unknown deviations due to the faults. We identify the unknown deviations Δk_{of} and Δh for a process with faults.

Define

$$\hat{\underline{f}}^*(t) \triangleq \left[\frac{ah^*}{\rho C_p V}(u_3 - \hat{x}_1) \quad \frac{\Delta H}{\rho C_p} k^*_{of} \cdot e^{-E_a/R\hat{x}_1} \cdot \hat{x}_4\right]^T \tag{8c}$$

$$\Delta \underline{p} \triangleq [\Delta h \quad \Delta k_{of}]^T \tag{8d}$$

$$\underline{\hat{b}}(k) \triangleq [\{\underline{\hat{x}}(k+1) - \underline{\hat{x}}(k)\}/\tau] - A\underline{\hat{x}}(k) - B\underline{u}(k) - Q\hat{\underline{f}}^*(k) \tag{8e}$$

$$\hat{S}(k) \triangleq Q\begin{bmatrix} \frac{a}{\rho C_p V}(u_3 - \hat{x}_1) & 0 \\ 0 & \frac{\Delta H}{\rho C_p} e^{-E_a/R\hat{x}_1\hat{x}_4} \end{bmatrix} \tag{8f}$$

Then the discrete version of the process model (Eq. 4a) by using the approximation (Eq. 11a) of Part I is given as

$$\hat{S}(k)\Delta \underline{p} = \underline{\hat{b}}(k) + \underline{\epsilon}(k) \tag{8g}$$

where $\underline{\epsilon}(k)$ is an unknown vector. The linear least squares estimation of $\Delta \underline{p}$ is obtained by applying the recursive algorithm (Eqs. 16a–16d) of Part I to Eq. 8g.

SIMULATION RESULTS

Simulations were performed to compare the proposed method with the extended Kalman filter using a CDC Dual Cyber 170/750 at the University of Texas. The standard deviations of the system noises and observation noises in Eqs. 4a and 4b were as follows:

v_1 : 1.0 [K]
v_2 : 5.0 [gmol/m^3]
w_1 : 5.0 [K]
w_2 : 5.0 [gmol/m^3]
w_3 : 5.0 [K]

We assume the process inputs u_1, u_2 and u_3 were held constant. Eleven cases of the many faults that were simulated are listed in Table 1. Group (I) (short-term simulation) in Table 1 was devised to compare the performances of the proposed method and the extended Kalman filter. Group (II) (long-term simulation) in Table 1 was programmed to investigate how accurately the proposed method identifies the process faults. In the long term simulation of Group (II), equivalent simulations using the extended Kalman filter were omitted simply because of the very long computation times and correspondingly expensive computation costs of the Kalman filter (a short term simulation by the extended Kalman filter cost about $8 and the estimated cost for one long term sim-

TABLE 1. CASES OF SIMULATIONS (I) AND (II)

Case	Source of Fault	(I) Time Period 16 hour (h) Sampling Interval 0.04 hour	(II) Time Period 80 hour (h) Sampling Interval 0.2 hour
(A)	Decrease of overall heat transfer coefficient h from 6.069×10^5 to 4.20×10^5 [J/(m^2)(h)(K)] at (I) 4 hours and (II) 40 hours	(I-A) Figure 2	(II-A) Figure 7
(B)	Decrease of frequency factor k_o from 5.01×10^8 to 4.008×10^8 [1/h] at (I) 8 hours and (II) 40 hours	(I-B) Figure 3	(II-B) Table 2
(C)	Increase of activation energy E_a from 1.369×10^5 to 1.4375×10^5 [J/gmol] at (I) 8 hours and (II) 40 hours	(I-C) Figure 4	(II-C) Table 4
(D)	Decreases of h and k_o h: 6.069×10^5 to 4.20×10^5 [J/(m^2)(h)(K)] and k_o: 5.01×10^8 to 4.008×10^8 [1/h] at (a) (I), h: 4 hours and k_o: 8 hours and (b) (II) h: 48 hours and k_o: 40 hours	(I-D) Figure 5	(II-D) Table 4
(E)	Ramp increase of activation energy from 1.369×10^5 to 1.4395×10^5 [J/gmol] during 4 hours to 12 hours	(I-E) Figure 6	—
(F)	Decrease of k_o from 5.01×10^8 to 4.008×10^8 [1/h] at 8 hours and increase of inlet heating temperature from 850 to 950 [K] at 30 hours	—	(II-F) Figure 8(a) Table 5
(G)	Increase of E_a from 1.369×10^5 to 1.4275×10^5 [J/gmol] at 8 hours and increase of inlet heating temperature from 850 to 950 [K] at 30 hours		(II-G) Figure 8(b) Table 5

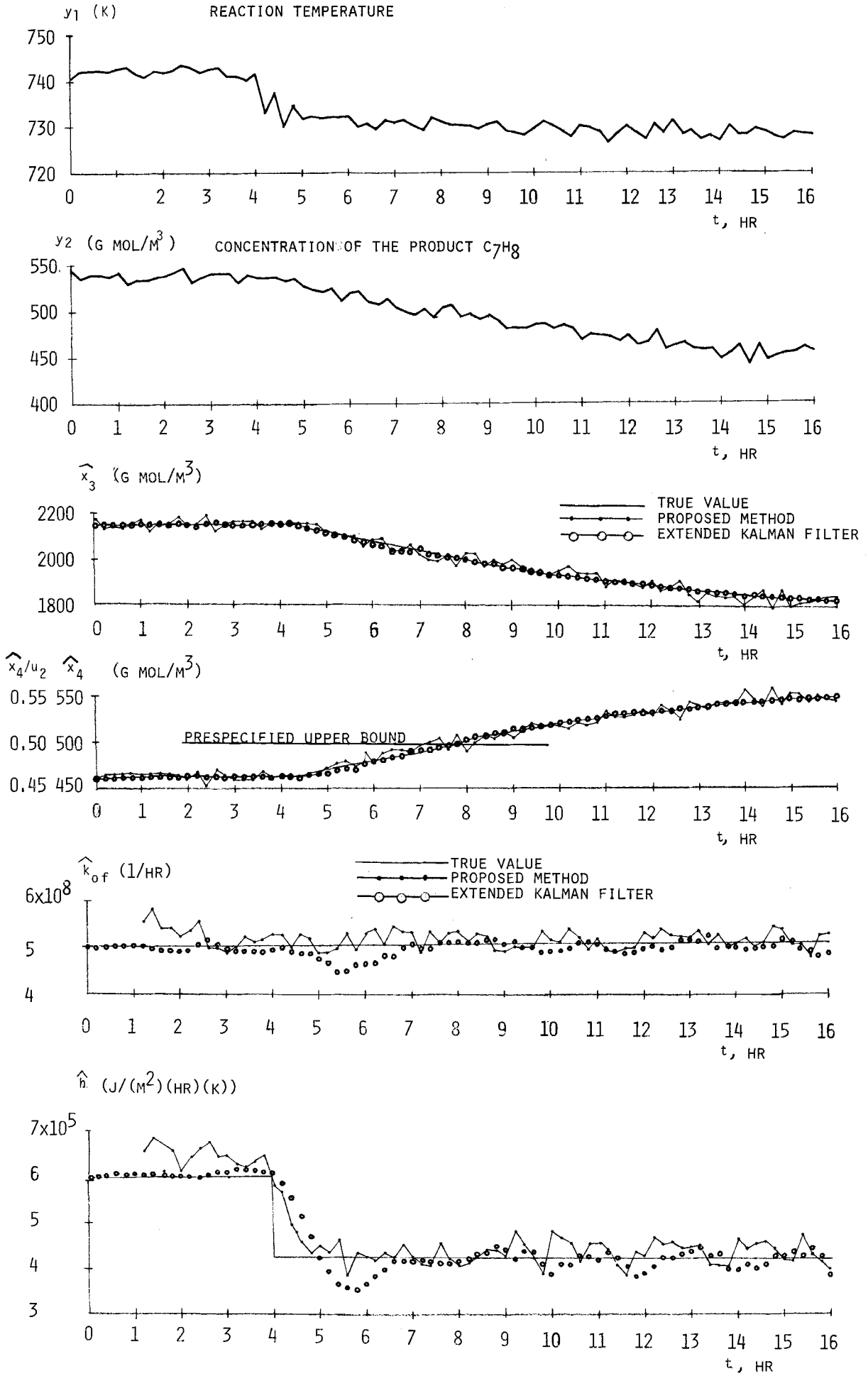

Figure 2. (a) and (b) Output measurements; (c) and (d) estimated state variables; and (e) and (f) trajectories of the recursively identified fault parameters k_{of} and h (physical degradation of the catalyst and fouling of the heat exchange surface).

ulation was about \$32, whereas the proposed method cost about \$0.02 for either a short term or a long term simulation).

In the results described below for the proposed two-level method, observer (6) was used to estimate the inaccessible state variables and the parameter identifier described by Eqs. 8 was used to identify the fault parameters. The time step for integration in the proposed method was same as the sampling interval of the measurements. The weighting factor λ in the identifier was chosen as $\lambda = 0.95$, a choice which led to slower convergence but smaller variance of the estimates than would be obtained for a bigger λ.

The extended Kalman filter described in Appendix A was designed assuming that the system noises and observation noises were uncorrelated each other and had the following statistics

$$R_c = \begin{bmatrix} 1 & 0 \\ 0 & 25 \end{bmatrix}, \quad G = \begin{bmatrix} 25 & & & & 0 \\ & 25 & & & \\ & & 25 & & \\ & & & 3.3 \times 10^{10} & \\ & 0 & & & 1.0 \times 10^{16} \end{bmatrix}$$

where R_c is the covariance matrix of observation noise and G is the covariance matrix of system noise of the augmented model in Appendix A. The values of the covariance matrices are determined by the standard deviations of the inputs and outputs measurements and those of the fault parameters. For our work we assumed that the standard deviation of h was equal to 0.3 × (normal value of h) and the standard deviation of k_{of} was equal to 0.2 × (normal value of k_o). A choice of smaller standard deviations leads to slower convergence but smaller variances for the estimates than would the choice of larger standard deviations. The selected standard deviations provided a sound balance between convergence speed and the magnitude of variances.

The initial values of the covariance matrix of the estimated state variables and the state variables themselves were

$$P(0) = \begin{bmatrix} 1 & & & & \\ & 25 & & 0 & \\ & & 25 & & \\ & & & 25 & \\ & 0 & & & 25 \end{bmatrix}, \quad z(0) = \begin{bmatrix} \underline{x}(0) \\ h(0) \\ k_{of}(0) \end{bmatrix} = \begin{bmatrix} 742.0 \\ 463.0 \\ 2{,}149.0 \\ 6.069 \times 10^5 \\ 5.01 \times 10^8 \end{bmatrix}$$

We chose the initial values of the state variables to be the values of the state variables at the normal steady state operating point. We employed a fourth order Runge Kutta method to integrate the equations in the extended Kalman filter, and used a time step of 0.002 hours which is almost the maximum permissible size that permits one to solve the filter equations with appropriate accuracy and stability. The noise level was 10^6 greater than the error generated by the integration formula. The integration time step had to be very small because of the stiffness of the filter.

The computation time necessary to estimate the complete time trajectory shown in the accompanying figures by the proposed two-level method was about 0.25 seconds whereas the corresponding computation time for the extended Kalman filter which was programmed quite straightforwardly without special procedures, was 120 seconds. Thus the proposed method is about 480 times faster than the extended Kalman filter. To implement the proposed two-level method on a microcomputer will involve only a small number of programming steps and a small amount of storage in contrast to the extended Kalman filter which will require substantial matrix operations and considerable storage.

Short-Term Simulations

We first examine the results for the cases in Group (I). Figure 2(a) shows the two output measurements for case (I-A). The trajectories of the unmeasured estimated state variables after the settling of the filter transients are shown in Figures 2(c) and (d). The standard deviation of the estimates $\hat{x}_3$ and $\hat{x}_4$ by the proposed method are about 5. and 20., respectively, whereas the estimates by the extended Kalman filter almost coincide exactly with the theoretical values.

Suppose the prespecified ratio of inlet and outlet concentrations of C_7H_{16} is given with the limit of $r^* = 0.5$ for an alarm decision. We can clearly reach a decision from the estimate $\hat{x}_4$ by either of the methods that the same fault has occurred at 7.6–8.2 hours. But we cannot determine causes of the fault from Figures 2(c) and (d) alone. In order to discover the causes we need to identify the values of the parameters k_{of} and h simultaneously. Figures 2(e) and (f) shows the trajectories of the identified parameters; h is obviously decreasing and k_{of} is stochastically stationary. The estimate $\hat{k}_{of}$ by the extended Kalman filter always has less variance than that of the proposed method, but the estimate immediately after the decrease of h has almost same variance as the proposed method. The latter avoids the undershoot and large perturbations of the extended Kalman filter.

Figure 3 shows the identified parameters both by the proposed method and the extended Kalman filter for case (I-B). Both methods identify the fault parameters k_{of} and h following a change

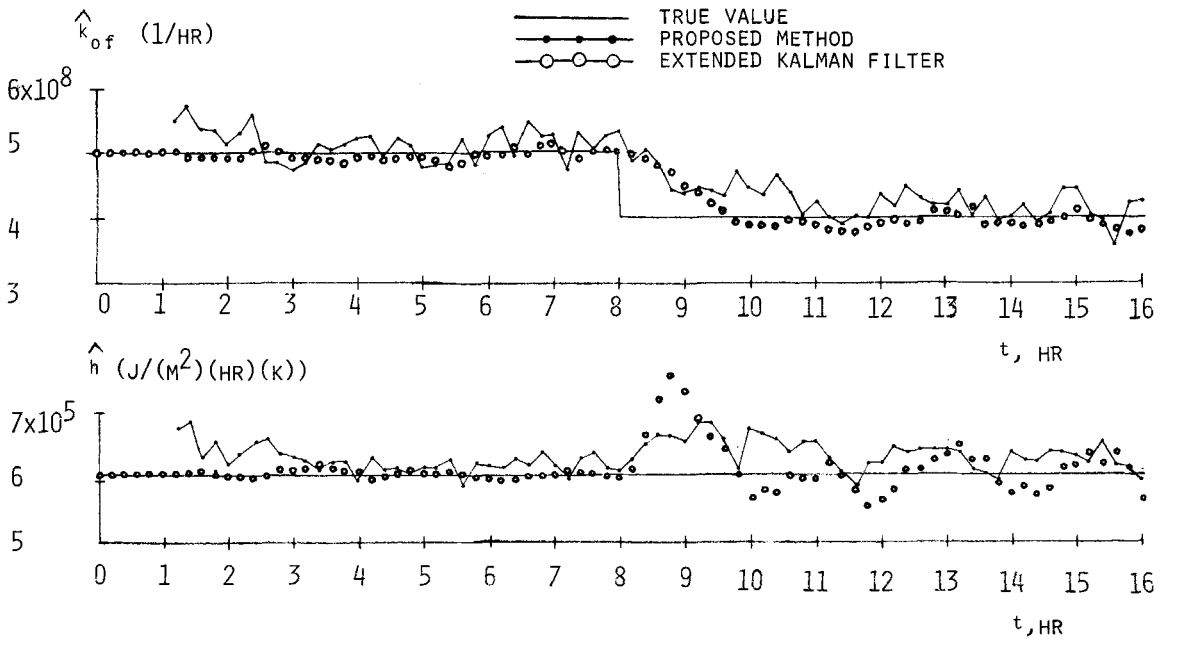

Figure 3. Trajectories of the recursively identified fault parameters k_{of} and h (physical degradation of catalyst).

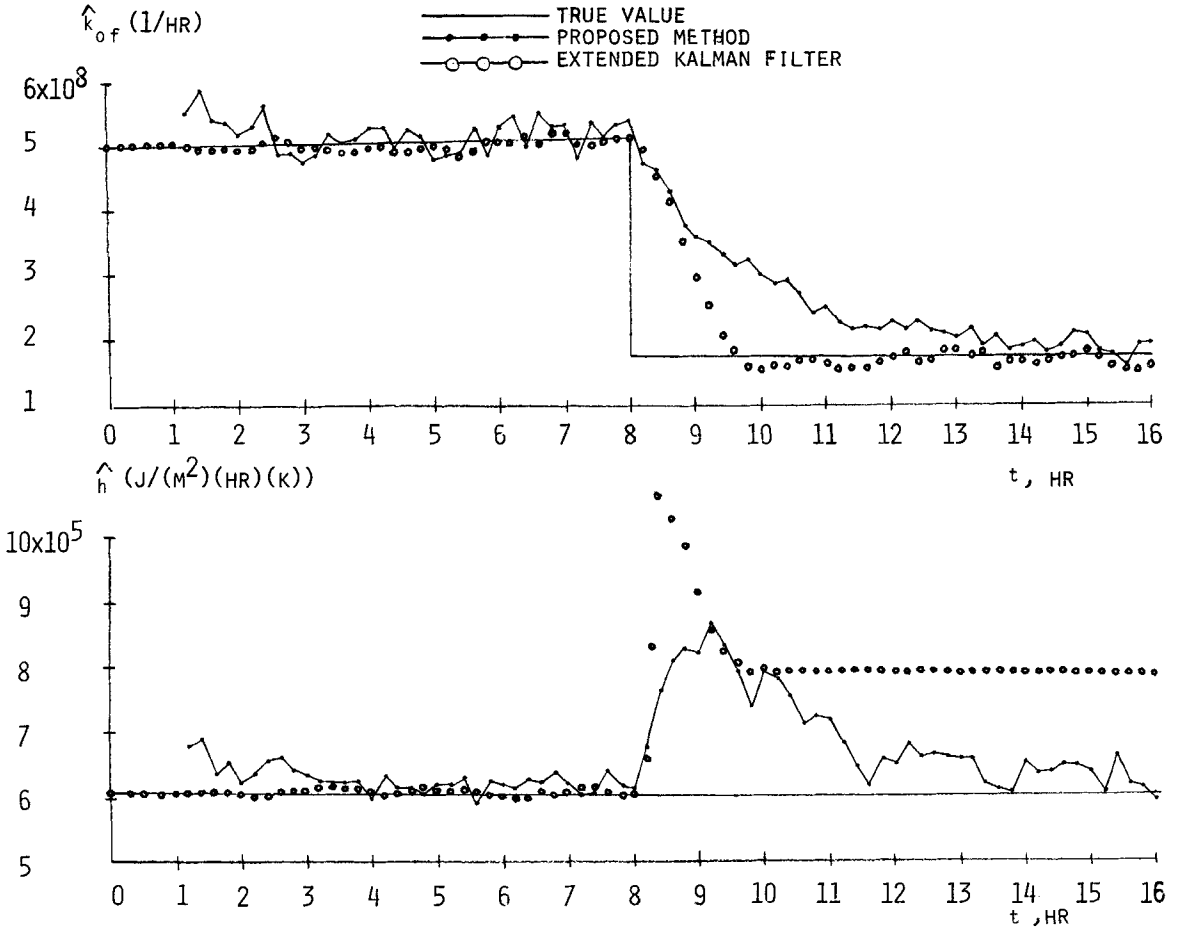

Figure 4. Trajectories of the recursively identified fault parameters k_{of} and h (chemical degradation of catalyst).

in k_{of}. The extended Kalman filter shows faster convergence of k_{of} whereas the estimates by the proposed method demonstrate slower convergence with larger variance. On the other hand, the estimate $\hat{h}$ by the extended Kalman filter method exhibits larger (perhaps unacceptable and certainly misleading) fluctuations after the change in k_{of} than does the proposed method.

Figure 4 shows the results for case (I-C). Again the estimate $\hat{k}_{of}$ by the extended Kalman filter has faster convergence than the proposed method. However the estimate $\hat{h}$ by the extended Kalman filter demonstrates some interesting behavior, behavior that would cause serious problems in practice, namely that the trajectory converges very nicely to a local estimate that is not the proper one, whereas the proposed method converges (somewhat slowly) to near the true value of h. The erroneous estimation probably occurs because of the use of the locally linearized state space equations in the filter equations.

Figure 5 shows the results for case (I-D). Both methods track changes in k_{of} and h with results similar to Figures 2 and 3. The variance of k_{of} by the proposed method is larger than that of the extended Kalman filter and the convergence speed is slow; however, again after the changes in the parameters, the variance of h by the extended Kalman filter is larger than that resulting for the proposed method.

Figure 6 shows the results for case (I-E), the case for a ramp increase of E_a. Because of the gradual increase of E_a, both methods track the change in the parameter. The statistics of the trajectories of the identified parameters have the same characteristics as the above cases.

From the above simulation results, we can summarize our observations:

(a) The proposed method is much faster in the computation and is numerically more stable than the extended Kalman filter.

(b) The extended Kalman filter yields slightly better estimates of k_{of} than the proposed method.

(c) The proposed method did not lead to misleading (erroneous) estimates of h as did the extended Kalman filter.

(d) Estimates by the Kalman filter are less sensitive than the proposed procedure to observation noise either with gradual parameter changes or after the transient of the filter settles down.

For this example, k_{of} could be estimated better than h because of the higher sensitivity to changes in k_{of} than h. Consequently, the proposed two-level method provides safer estimates at less cost than does the extended Kalman filter, even if in some cases the precision of the estimates is lower.

Long-Term Simulation

Next we examine the results in Group (II) for cases (A)–(G) in Table 1. The weighting factor was chosen as $\lambda = 0.85$ in all these cases. Figure 7 shows the trajectories of the identified parameters for case (II-A). The variance of the estimates is larger than for the case when $\lambda = 0.95$ but we note faster convergence. The average values of the trajectories of the identified parameters before and after the fault are also shown as in Figure 7. The average values of the estimates are very close to the theoretical values.

Table 2 shows the average values of the estimates before and after the faults for case (II-B). Table 3 shows the results for case (II-C), and Table 4 shows the results for case (II-D). The average values of all the identified parameters have values quite close to the theoretical values. From the average values of the identified parameters, we can note how the parameters change and can diagnose the causes of faults.

One problem with the above procedure is that in the estimations (A), (B), (C) and (D), we could not identify whether the degradation of the catalyst is due to chemical or physical degradation. However, from Eq. 7, it is known that the identified parameter k_{of} is a function of the reaction temperature. Then, if k_{of} is the result of a deviation of E_a, the identified parameter changes as the temperature changes. On the other hand k_{of} will be constant if the fault

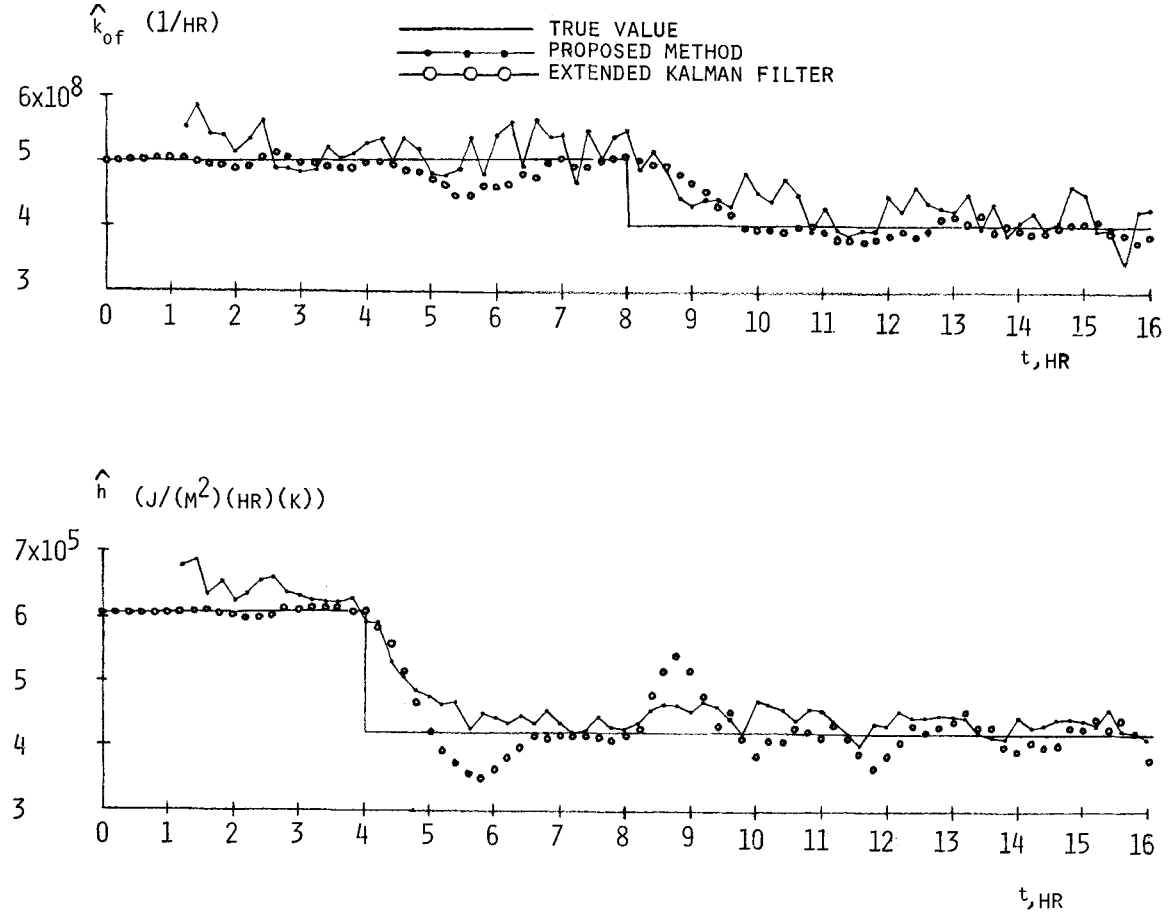

Figure 5. Trajectories of the recursively identified fault parameters k_{of} and h (physical degradation of catalyst and fouling of the surface at different times).

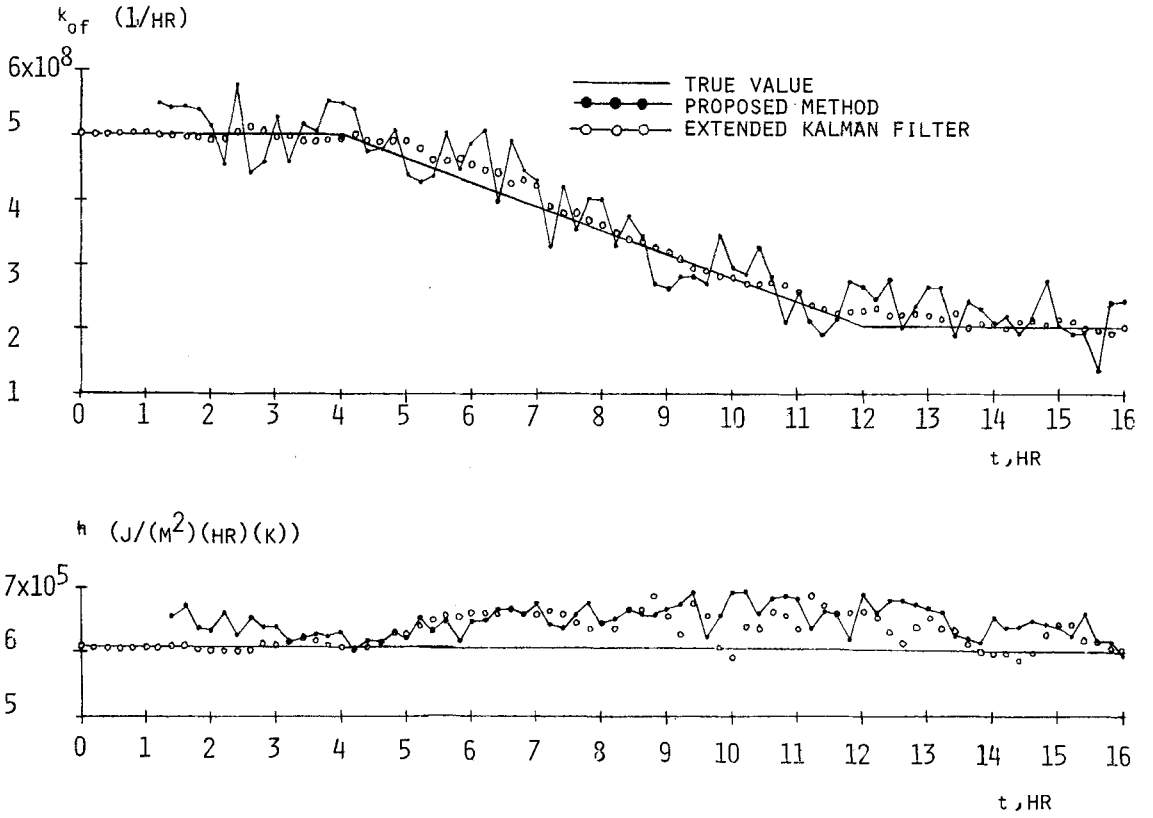

Figure 6. Trajectories of the recursively identified fault parameters k_{of} and h (ramp increase of the activation energy E_a).

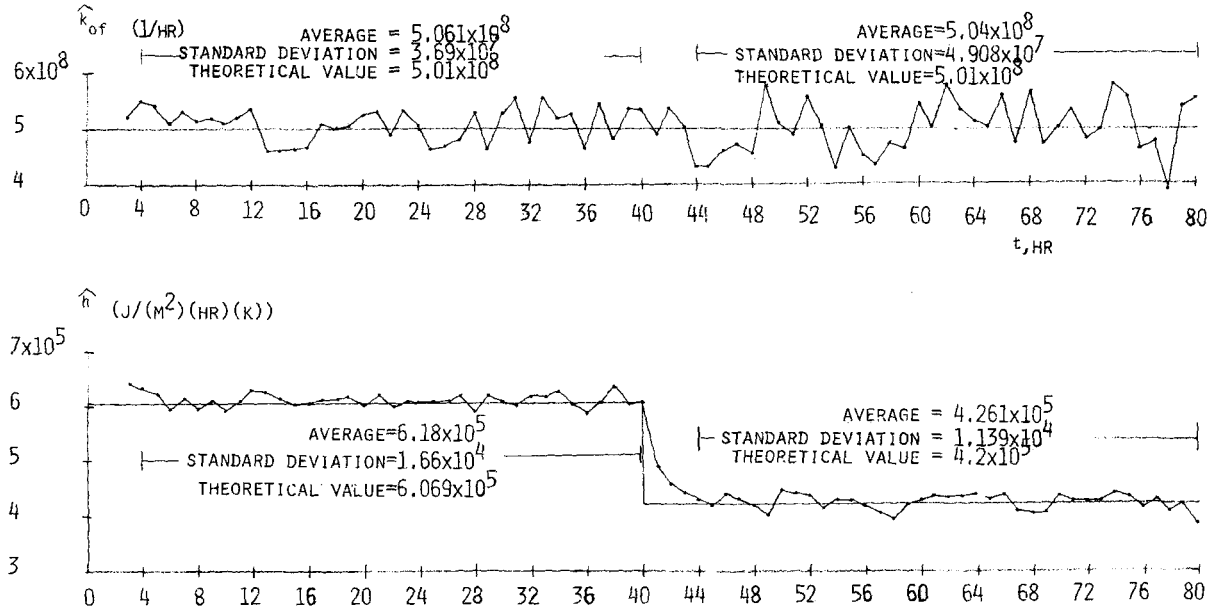

Figure 7. Trajectories of the recursively identified fault parameters k_{of} and h (fouling of the heat exchange surface).

TABLE 2. AVERAGE VALUES OF IDENTIFIED PARAMETERS: VALUES IN ( ) ARE TRUE VALUES

Time for Average	$\bar{k}_{of}$	$\bar{h}$	Avg. Reaction Rate Constant at T = 741 K
4–44 h	5.061×10^{8}	6.18×10^{5}	0.1146
	(5.01×10^{8})	(6.069×10^{5})	(0.1135)
44–80 h	4.035×10^{8}	6.167×10^{5}	0.09135
	(4.008×10^{8})	(6.069×10^{5})	(0.09078)

TABLE 3. AVERAGE VALUES OF IDENTIFIED PARAMETERS: VALUES IN ( ) ARE TRUE VALUES

Time for Average	$\bar{k}_{of}$	$\bar{h}$	Avg. Reaction Rate Constant at T = 741 K
4–44 h	5.061×10^{8}	6.18×10^{5}	0.1146
	(5.01×10^{8})	(6.069×10^{5})	(0.1135)
44–80 h (for $\hat{k}_{of}$)			
48–80 h (for $\hat{h}$)	1.706×10^{8}	6.178×10^{5}	0.03864
	(1.700×10^{8})	(6.069×10^{5})	(0.03851)

TABLE 4. AVERAGE VALUES OF IDENTIFIED PARAMETERS: VALUES IN ( ) ARE TRUE VALUES

Time for Average	$\bar{k}_{of}$	$\bar{h}$	Avg. Reaction Rate Constant at T - 741 K
4–40 h (for $\hat{k}_{of}$)	5.061×10^{8}	6.18×10^{5}	0.1146
	(5.01×10^{8})	(6.069×10^{5})	(0.1135)
4–48 h (for $\hat{h}$)			
44–80 h (for $\hat{k}_{of}$)	4.036×10^{8}	4.22×10^{5}	0.09142
	(4.008×10^{8})	(4.2×10^{5})	(0.09078)
52–80 h (for $\hat{h}$)			

TABLE 5. IDENTIFICATION OF k_{of} WHEN k_o CHANGES AND WHEN E_a CHANGES: VALUES IN ( ) ARE TRUE VALUES

Time for Average	$\bar{k}_{of}$ When k_o Changes	$\bar{k}_{of}$ When E_a Changes
12–31 h	4.035×10^{8}	1.704×10^{8}
	(4.008×10^{8})	(1.70×10^{8})
44–80 h	4.055×10^{8}	1.844×10^{8}
	(4.008×10^{8})	(1.828×10^{8})
Rate of Increase	0.51%	8.22%
	(0.0%)	(7.73%)

in k_{of} is only the result of the deviation of k_o. In other words, to discriminate among the causes of catalystic degradation, a test signal which changes the reaction temperature is needed. In the simulations (F) and (G), the inlet heating temperature was changed from 850 [K] to 950 [K] at 31 hours. Figures 8(a) and (b) shows the results when k_o changes and Figures 8(c) and (d) shows the results when E_a changes, each at 8 hours. The average values of k_{of} during the period 21 to 31 hours and 44 to 80 hours, when k_o and E_a change, are listed in Table 5. Obviously from the results shown in Table 5, we can identify that the fault shown in Figures 8(a) and (b) is only due to the change of k_o; i.e., it represents physical catalyst degradation, and the fault in Figures 8(c) and (d) is only due to change of E_a, i.e., it represents chemical catalyst degradation.

DISCUSSION

To assess the applicability of the proposed method to real processes, one should check whether or not the following points are satisfied:

(i) Can the real process be described by the process model Eqs. 1a and 1b (or instead of Eq. 1b, $\underline{y}(t) = c\underline{x}(t)$) in Part I?

(ii) Does the real process described by Eqs. 1 satisfy conditions 3a and 3b in Part I?

(iii) Does the real process described by Eqs. 1 satisfy the condition 12b or 14b in Part I?

(iv) Are the levels of observation and system noise relatively low?

If the real process satisfies the above four check points, one can apply our proposed method.

Process model (Eqs. 1) is more general than those models employed by Luenberger and/or Kalman for state estimation, but it is more restrictive than a general set of nonlinear ordinary differential equations. Nevertheless, the model (Eqs. 1) can describe a wide class of real nonlinear process with faults of interest to chemical engineers.

Condition 3a in Part I means that the number of output measurements must be larger or equal to the number of nonlinear elements and unknown coefficients in linear elements in the model. This condition may prove to be the most restrictive among all of the above conditions, but, of course, this constraint can be ameliorated by adding additional independent measurements. Condition 3b is not restrictive, because most of the process we deal with are stable.

Condition 12b or 14b also is not restrictive, because by adding

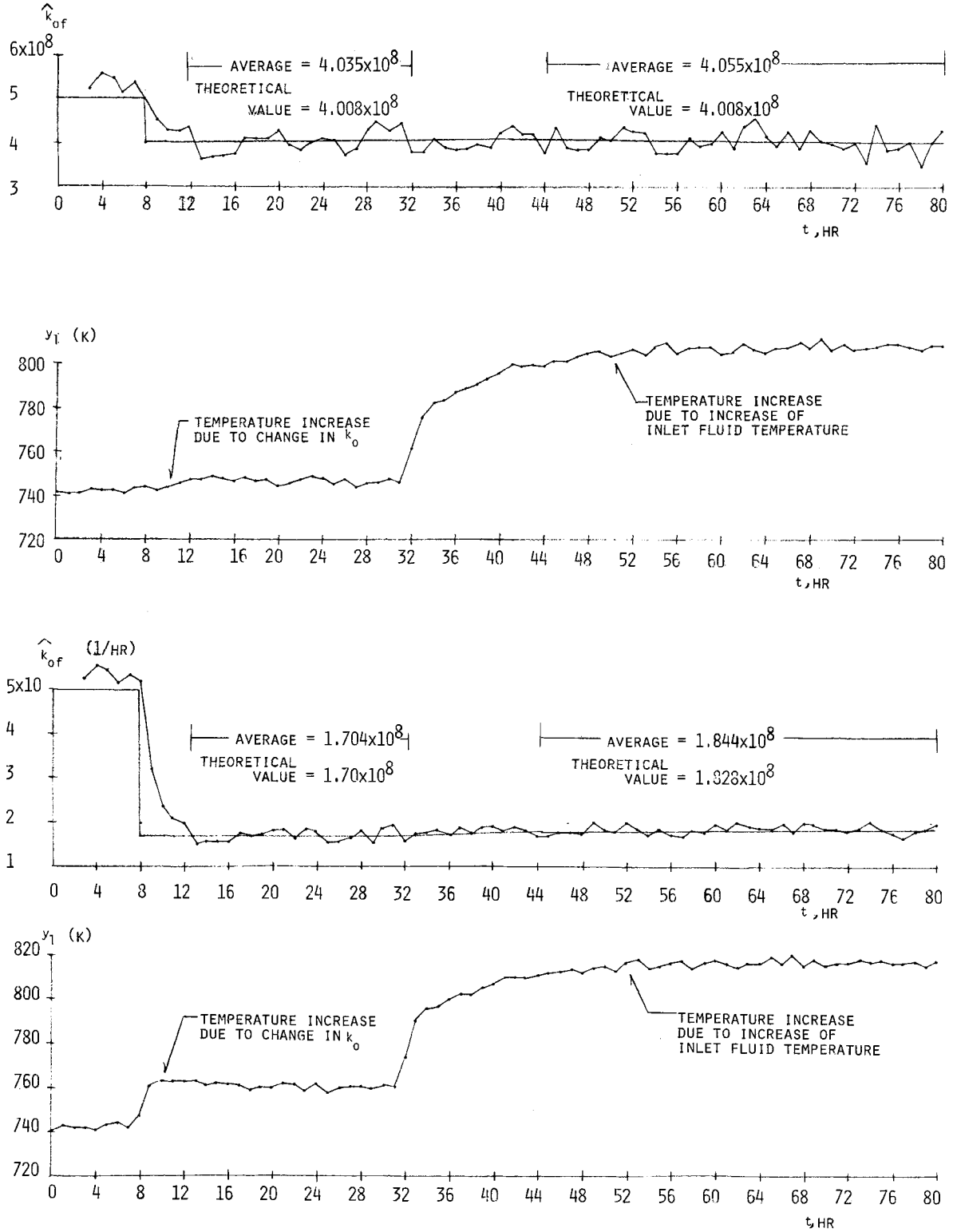

Figure 8. (a) and (b) Change of k_{of} following a change of reaction temperature; and (c) and (d) change of k_{of} following a change of reaction temperature.

test signals one can force the process to satisfy the condition. Point (iv) above depends on the given process. Insofar as the example we have used here is concerned, the estimates by our proposed method were more conservative and safer than those yielded the extended Kalman filter even though in the steady state the estimates were a little more sensitive than those of the extended Kalman filter.

Finally, suppose there exist coefficients in the process model whose normal values are not identifiable *a priori*. It is possible to identify both the fault parameters and the normal but unknown parameters simultaneously by our proposed procedure. However, the extra burden of identification of those normal but unknown parameters may require one to collect data at more measurement points than required for the fault parameters alone.

CONCLUSIONS

A new strategy for state estimation and parameter identification oriented to process fault diagnosis has been described. Simulations involving the conversion of heptane to toluene were carried out and the result of fault diagnosis for the proposed method were compared with those obtained using the extended Kalman filter.

We found that the extended Kalman filter needed substantially more computation time than the proposed method and in some instances provided incorrect estimates rendering the extended Kalman filter worthless for fault diagnosis. The proposed method yielded safe estimates even if in some cases the precision of the estimates was lower than those of the extended Kalman filter. The average values of the estimates yielded by the proposed method both after and before the initiation of faults were very close to the theoretical values, and hence could be used to diagnose correctly how the parameters changed.

The proposed state estimator can be applied not only to the estimation of inaccessible state variables but it can be applied to estimate unknown process inputs associated with process faults.

Individuals who wish to obtain a copy of the program listing for the proposed strategy should write the second author.

APPENDIX A: EXTENDED KALMAN FILTER

We used the augmented model of process (Eqs. 4) with the unknown fault parameters h and k_{of}, and develop the extended Kalman filter for the augmented model.

Define

$$\underline{z}(t) \triangleq [\underline{x}(t)^T \quad h \quad k_{of}]^T \tag{A1}$$

$$\underline{g}(\underline{x},\underline{u},\underline{w}) \triangleq A\underline{x} + B\underline{u} + Q\underline{f} + E\underline{w} \tag{A2}$$

$$\underline{y} \triangleq [x_1 \quad x_2]^T \tag{A3}$$

$$F \triangleq \begin{bmatrix} \dfrac{\partial \underline{f}}{\partial \underline{z}^T} \\ \hline 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{A4}$$

$$C \triangleq \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix} \tag{A5}$$

and let

$$S \triangleq E[\underline{w}\,\underline{w}^T] \tag{A6}$$

$$R_c \triangleq E[\underline{v}\,\underline{v}^T] \tag{A7}$$

Also let $P(t)$ be the estimate of the covariance matrix of the estimated variables. Then, the augmented model of the process (Eqs. 4) is given by

$$\dot{\underline{z}} = \begin{bmatrix} \underline{g}(\underline{x},\underline{u},\underline{w}) \\ \underline{0} \end{bmatrix}, \quad \underline{z}(0) = \underline{z}_0 \tag{A8}$$

and the extended Kalman filter for the augmented model is given by

$$\dot{\hat{\underline{z}}} = \underline{g}(\hat{\underline{z}},\underline{u},\underline{0}) + PC^T R_c^{-1}[\underline{y} - C\hat{\underline{z}}], \quad \hat{\underline{z}}(0) = \hat{\underline{z}}_0 \tag{A9}$$

$$\dot{P} = FP + PF^T + ESE^T - PC^T R_c^{-1} CP, \quad P(0) = P_0 \tag{A10}$$

NOTATION

A = process coefficient matrix for linear term
A_{22} = partitioned matrix of A
a = area of heat exchange of the process
B = process input matrix
$\underline{b}(k)$ = estimate of $\underline{b}(k)$ used in the parameter identification
C = process output matrix
C_p = specific heat of the process
C_7H_8 = toluene
C_7H_{16} = heptane
D = filter coefficient matrix
E = process noise matrix
E_a = activation energy of the reaction
ΔE_a = deviation of E_a due to fault
F = filter coefficient matrix
$\underline{f}(t,\underline{p})$ = vector of nonlinear functions or linear and/or nonlinear functions with unknown coefficients
G = covariance matrix of the system noise
$g(t)$ = process function
$\partial g/\partial \underline{z}^T$ = Jacobian matrix
H = filter coefficient matrix
H_2 = hydrogen
ΔH = heat of reaction
h = overall heat transfer coefficient of heat exchange
h^* = normal value of h
$\hat{h}$ = estimate of h
Δh = deviation of h due to fault
K = filter coefficient matrix
k = discrete time
$k(T)$ = reaction rate constant
k_o = frequency factor
k_{of} = fault parameter to be identified
k_{of}^* = normal value of k_{of}
$\hat{k}_{of}$ = estimate of k_{of}
Δk_o = deviation of k_o due to fault
L = filter coefficient matrix
M = filter coefficient matrix
$P(t)$ = covariance matrix of estimated state variables
$P(0)$ = initial value of $P(t)$
$\Delta \underline{p}$ = unknown fault parameter vector
Q = process coefficient matrix for nonlinear and/or linear terms with unknown parameters
Q_1, Q_2 = partitioned matrix of Q
q = inlet and outlet volumetric flow rate
R = gas constant
R_c = covariance matrix of observation noise
r^* = prespecified ratio of inlet and outlet concentrations
$\hat{S}(k)$ = estimate of $S(k)$ used in the parameter identification
T = reaction temperature of the process
T_i = inlet temperature of the process
T_h = heating temperature of the process
t = continuous time
$\underline{u}$ = input vector of process
$u_i (i = 1,2,3)$ = inputs of the process
V = effective reactor volume
$\underline{v}(t)$ = observation noise vector
$\underline{w}(t)$ = system noise vector
$\underline{x}(t)$ = state vector of process
$\underline{x}_0$ = initial value of $\underline{x}(t)$
$\dot{\underline{x}}(t)$ = derivative of $\underline{x}(t)$
$x_i (i = 1,2,3,4)$ = state variable of the process
$\hat{x}_i (i = 1,2,3,4)$ = estimates of x_i
$\underline{y}(t)$ = output vector of process
$y_i (i = 1,2)$ = outputs of the process
$\underline{z}(t)$ = state vector of state estimation filter
$\underline{z}_0$ = initial value of $\underline{z}(t)$
$\hat{\underline{z}}(t)$ = estimated state vector of the extended Kalman filter
$z_i (i = 1,2)$ = state variables of the filter

Greek Letters

$\epsilon(k)$ = error vector in identification
λ = weighting factor in identifier

ρ = density
τ = sampling interval

Mathematical Symbols

rank[ ] = rank of matrix
E[ ] = expectation

LITERATURE CITED

Burnett, R. L., H. L. Steinmertz, E. M. Blue, and J. J. Noble, "An analog computer model of conversion in a catalystic reformer," *The devision of Petroleum Chemistry, American Chemical Society*, Detroit, (Apr., 1965).

Manuscript received November 6, 1981; revision received April 21, and accepted April 30, 1982.

Reduced-Order Steady-State and Dynamic Models for Separation Processes

Part I. Development of the Model Reduction Procedure

One of the major difficulties with mathematical models of staged separation systems is the large dimensionality of the process model. This paper is concerned with simple (reduced-order) steady-state and dynamic models for processes such as distillation, absorption and extraction. The model reduction procedure is based on approximating the composition and flow profiles in the column using polynomials rather than as discrete functions of the stages. The number of equations required to describe the system is thus drastically reduced. The method is developed using a simple absorber system. In the second part of this paper, the application of the method to nonlinear multicomponent separation systems is demonstrated.

Y. S. CHO and B. JOSEPH

Chemical Engineering Department
Washington University
St. Louis, MO 63130

SCOPE

The literature on modeling of stage separation systems is quite extensive. A number of fast, efficient algorithms have been developed for solving the large set of equations resulting from a tray-by-tray model both in the steady state as well as in the dynamics. Yet there exists a need for reduced-order models for two reasons. In cases involving optimization of systems involving more than one column or a single large column, the computation involving the steady-state simulation of these columns can be extremely time-consuming. In the case of dynamic simulation, most of the integration packages deal with a single column. But in dynamic simulation the interaction with other units is very important and it is desirable to treat the column as part of a large network of units each described by a set of differential equations. Hence it is desirable to use a general purpose integration package for the dynamic simulation. In such instances, a reduced order model is preferable to making modifications to the integration package.

This paper is concerned with reduced-order approximations to the rigorous stage-by-stage models. Two approaches are developed in this paper. The first is based on approximating the tray-by-tray differential-difference equations by a set of partial differential equations. These partial differential equations are in turn solved by the orthogonal collocation method. In the second approach, the model reduction is achieved by a direct transition to polynomial representation of the composition and flow profiles in the column.

In this paper, we examine: (1) the accuracy of different approximation procedures used for generating the reduced-order models; (2) the significance of the type and degree of the polynomials used in the orthogonal collocation procedure; and (3) the ability of the reduced-order models to approximate the rigorous model both in the steady state and in the dynamics.

CONCLUSIONS AND SIGNIFICANCE

The number of equations required to describe the behavior of staged separation systems can be drastically reduced by the approximation procedure developed here. Preliminary experimentation using selected test examples show excellent agreement between the rigorous model and the simplified model both in the steady-state and in the dynamic response to identical step, sinusoidal and random input disturbances. The accuracy of the approximation can be improved by appropriate choice of the PDE approximation and the type of polynomials used in the model reduction.

The work is of considerable significance to practitioners involved in the modeling of staged separation systems. The reduced-order steady-state models are useful in the simulation and optimization of processes having large multicomponent distillation columns or interconnected separation trains. The reduced-order dynamic models would allow the study of dynamic systems hitherto considered difficult because of the large dimensionality of the equations involved. Application areas include the study of process interaction, analysis and troubleshooting of startup and shutdown operations, safety and reliability analysis, design and testing of control strategies, and process identification using transient responses.

Correspondence regarding this paper should be addressed to B. Joseph.
0001-1541-83-4518-0261-$2.00.