

Systematic Estimation of State Noise Statistics for Extended Kalman Filters

Jaleel Valappil and Christos Georgakis

Chemical Process Modeling and Control Research Center and Dept. of Chemical Engineering,
Lehigh University, Bethlehem, PA 18015

The successful application of model-based control depends on the information about the states of the dynamic system. State-estimation methods, like extended Kalman filters (EKF), are useful for obtaining reliable estimates of the states from a limited number of measurements. They also can handle the model uncertainties and the effect of unmeasured disturbances. The main issue in applying EKF remains that one needs to specify the confidence in the model in terms of process noise covariance matrix. The information about the model uncertainties can effectively and systematically calculate the process noise covariance matrix for an EKF. Two systematic approaches are used for this calculation. The first is based on a Taylor series expansion of the nonlinear equations around the nominal parameter values, while the second accounts for the nonlinear dependence of the system on the fitted parameters by Monte Carlo simulations that can easily be performed on-line. The value of the process noise covariance matrix obtained is not limited to a diagonal form and depends on the current state of the dynamic system. Thus the a-priori information regarding the uncertainty in the model is utilized and the need for extensive tuning of the EKF is eliminated. The application of these techniques to example processes is also discussed. The accuracy of this methodology is compared very favorably with the traditional methods of trial-and-error tuning of EKF.

Introduction

Model-based predictive control techniques have become increasingly popular in chemical-process applications over the last decade. The successful implementation of these techniques relies on the information about the states of the dynamic system to make predictions into the future. Most of the models used in chemical-process applications are developed from a combination of physical and empirical knowledge. Inevitably, these models exhibit structural and parametric mismatch with the process resulting from fitting a model of limited complexity to a finite set of data for a complex process. Ignoring these uncertainties can result in the failure to achieve the benefits of model-based control. This prompts the use of state estimators to estimate unmeasured states and the effect of disturbances on-line while accounting for the process-model mismatch. Traditional predictive con-

trol methods, like dynamic matrix control (DMC), have tended to use a constant disturbance model, which is estimated as the difference between current process measurements and model predictions. As pointed out by Lee and Ricker (1994), there are limitations to this method, including its inability to handle unstable and integrating processes. The use of a state estimator can overcome many of the problems associated with the previous approach and can make the closed-loop system more robust to process-model mismatch.

The purpose of state estimation is to optimally use the information available in the process model and the measurements to obtain estimates of the unmeasured states of the dynamic system. The information about the states can be used for monitoring (Ku et al., 1994) and control of the process. Deterministic approaches to this problem typically involve state observers, while stochastic state-estimation techniques consider the stochastic characteristics of the dynamic system and the measurements in estimating the states. Recursive stochastic state-estimation techniques, like extended Kalman

Correspondence concerning this article should be addressed to C. Georgakis.

filters, have been traditionally used for state estimation in many of the chemical processes. Recently, more rigorous techniques for chemical-process applications have been proposed. These include nonlinear observers (Soroush, 1997), moving-horizon estimation (Jang et al., 1986), and statistically rigorous techniques designed to handle the nonnormal and non-Gaussian features of the process (Terwiesch and Agarwal, 1995). The local linearization of the model and the assumption of Gaussian and white noise processes used in EKF have been quoted as its main drawbacks, and has been used to justify more rigorous techniques like nonlinear observers. Even if these two limitations of the EKF are accepted, a critical step in its application remains the estimation of the process noise covariance matrix, Q . But the fact remains that the recursive nature of EKF and its ability to account for model uncertainties in a transparent way makes it practical, and hence quite attractive for different chemical-process applications. Further, the amount of information and computations required for EKF is nominal compared to other rigorous estimation techniques that require full probability distribution functions of the noise processes or solution of a nonlinear programming problem on-line.

EKF have been applied to a wide range of chemical processes, both batch and continuous. The variety of applications include chemical-reaction systems (De Valliere and Bonvin, 1989), polymerization processes (Jo and Bankoff, 1976; Kozub and Macgregor, 1992; Crowley and Choi, 1998), and bioreactors (Gudi et al., 1995). Polymerization processes have especially benefited from EKF, as many of the important states in these processes are not measured on-line and need to be estimated to ensure proper process operation. The application of EKF requires that the system be locally detectable with the available measurements and stabilizable with the noise model used (Lewis, 1986). Further, the application of Kalman filter requires the user to specify the relative amounts of confidence in the model and the measurements. These are specified in the form of process-noise and measurement-noise covariances, Q and R respectively. The performance of the filter is very dependent on specifying the right noise levels, and incorrect specification of these statistics can even cause the filter to diverge. While the specification of the measurement covariance matrix, R , can be directly derived from the accuracy of characteristics of the measurement device (R is thus assumed to be known in this study), specification of the Q is often attempted in a trial-and-error approach. For a system with n states, this requires the specification of $n(n+1)/2$ elements in Q , taking into account the fact that it is a symmetric matrix. For time-varying processes like batch and semibatch and for nonlinear ones, the specification of a constant matrix Q might not be sufficient in order to provide a sufficiently accurate filter performance.

As mentioned earlier, in most of the process models, there are unavoidable plant-model mismatches resulting from fitting a simple model to a complex process (Terwiesch et al., 1994). If the process-model mismatch is significant, it can result in biased state estimates. Further, unmeasured disturbances that enter the process on-line introduce biases in the estimates. A solution to this is the use of a parameter-adaptive EKF, in which nonstationary stochastic states are included in the filter and estimated along with the original

system states. This does not eliminate the need for proper tuning of the filter; the problem of specifying the right noise levels for the process and measurement noises is important when one is left with a less than perfect model. Also, it has been suggested that all the uncertain parameters be estimated with the parameter-adaptive EKF. But in most cases, the number of uncertain model parameters is larger than the number of measurements. Then the augmented system is not observable, and the parameters cannot be updated using the EKF.

The main bottleneck in the application of EKF to real-world situations is the lack of a robust design methodology, as demonstrated in an industrial application study (Wilson et al., 1998). The process-noise statistics are rarely available *a priori*, and in most of the Kalman filter applications, they are viewed as tuning parameters and selected by a trial-and-error procedure using repeated simulations. Adaptive filtering techniques have been used for linear and nonlinear systems, and these techniques estimate the noise statistics along with the state estimates (Mehra, 1972). Covariance matching and residual whitening have been utilized to determine the noise statistics that would be in agreement with the statistics of the filter (Maybeck, 1982). A few techniques for determining the process and measurement noise covariances for chemical engineering applications have been discussed in the literature. An approach to the selection of Q and R for linear systems was discussed by Zhou and Luecke (1995). In this case, the process and measurement noise models are combined with state estimates and a constant state-estimator error covariance in order to determine the cumulative error covariances. The maximum-likelihood method with linear regression techniques is used to obtain the diagonal elements of the covariance matrices. The use of residuals from the filter for the selection of the process and measurement covariances have been reported. For nonlinear systems, the use of innovations process for estimating the noise statistics was given by Myers and Tapley (1976). The innovations process (difference between actual measurement vector and the predicted measurement vector) can be used to adapt the covariances on-line, as reported in the experimental study for an emulsion polymerization reactor by Dimitratos et al. (1989). This makes the filter more open to the incoming data and avoid the divergence of filter that might occur due to inexact *a priori* statistics.

Most of the methods just mentioned assume constant noise characteristics and the availability of data required to obtain a true representation of noise statistics. The assumption of time-invariant process noise is more appropriate for a continuous process that operates at a steady state. But for continuous or batch processes with time-varying process dynamics and operating at a range of process conditions, these noise statistics are time varying. Using a fixed value of noise statistics can lead to poor filter performance and even result in filter divergence, as is demonstrated for some example processes later. Even in cases where the fixed noise statistics are acceptable, finding the appropriate values can be a tedious task. One is left with the task of selecting a number of parameters, at least equal to the number of states for the case where the matrix is assumed to be diagonal. These realities prompt one to consider alternate schemes to improve the estimate of the process-noise statistics.

Two approaches to systematically estimate the process-noise covariance matrix are discussed in this article. Both approaches involve straightforward on-line computations and utilize the current states from the filter. What is needed to apply these techniques is a linear or nonlinear process model and a parametric quantification of its uncertainty. The model uncertainty is specified by the parameter covariance matrix, obtained during the model development. The proposed two methods calculate time-varying values of the process-noise covariance matrix, \mathbf{Q} , on-line that are used by the filter. The first method involves linear approximation of the dependence of the model predictions on the model parameters and hence, is computationally straightforward. For nonlinear processes, the second method is more rigorous in the sense that it fully considers the nonlinear dependence of model predictions on the parameters and involve statistical Monte Carlo simulations. These computations involve repeated algebraic evaluations and can be handled on-line without much computational burden, as shown in the case studies presented later in this article. The application of these techniques to some chemical processes is presented here. It is demonstrated that improved filter performance is realized using these approaches compared to the conventional trial-and-error tuning techniques.

Extended Kalman Filter: A Review

The EKF computes the optimal state estimate at each time instant using a linearized model of the nonlinear system and linear estimation principles. For the case of sampled data measurements of a continuous process, a hybrid EKF is the most appropriate. It uses the continuous-time model of the process and discrete observation for update. The hybrid Kalman filter algorithm is used in this study for state estimation and is described briefly below. For a full description, the reader should refer to the books in this area (such as Jazwinski, 1970; Maybeck, 1982).

Consider a nonlinear system as shown below:

$$\dot{\mathbf{x}}(t) = \mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t, \mathbf{p}(t)] + \mathbf{w}(t) \quad (1)$$

$$\mathbf{z}(t) = \mathbf{h}[\mathbf{x}(t), \mathbf{u}(t), t] + \mathbf{v}(t), \quad (2)$$

where \mathbf{u} denotes the deterministic inputs, \mathbf{p} denotes the model parameters, and \mathbf{G} represents a matrix that defines how the process noise affects the state derivatives. The process-noise vector, $\mathbf{w}(t)$, is assumed to be a white Gaussian random process with zero mean and covariance $\mathbf{Q}(t)$:

$$E[\mathbf{w}] = 0 \quad (3)$$

$$E[\mathbf{w}(t)\mathbf{w}^T(\tau)] = \mathbf{Q}(t)\delta(t-\tau). \quad (4)$$

The measurement error, $\mathbf{v}(t)$, is also assumed to be a white Gaussian random process, with

$$E[\mathbf{v}] = 0 \quad (5)$$

$$E[\mathbf{v}(t)\mathbf{v}^T(\tau)] = \mathbf{R}(t)\delta(t-\tau). \quad (6)$$

The EKF formulation uses linearized models of the nonlinear system for state estimation. Here, the system is linearized

at each time step to obtain the local state-space matrices as below:

$$\mathbf{F}(t) = \left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right)_{\mathbf{x}(t), \mathbf{u}(t), t, \mathbf{p}_{\text{nom}}}$$

$$\mathbf{H}(t) = \left(\frac{\partial \mathbf{h}}{\partial \mathbf{x}} \right)_{\mathbf{x}(t), \mathbf{u}(t), t, \mathbf{p}_{\text{nom}}},$$

where \mathbf{p}_{nom} denotes the nominal parameter values. The propagation of the states and the state covariances is achieved by integrating the nonlinear model equations and the covariance propagation equations into the next time step. The Kalman gain is then computed. The measurement update equations are then used to estimate the state and the covariance updates. The equations that compose the different steps in the EKF are given below.

State Propagation Equation:

$$\hat{\mathbf{x}}_k^- = \hat{\mathbf{x}}_{k-1}^+ + \int_{k-1}^k \mathbf{f}(\hat{\mathbf{x}}, \mathbf{u}, \tau) d\tau. \quad (7)$$

Covariance Propagation Equation:

$$\mathbf{P}_k^- = \mathbf{P}_{k-1}^+ + \int_{k-1}^k \left[\mathbf{F}(\tau) \mathbf{P}(\tau) + \mathbf{P}(\tau) \mathbf{F}(\tau)^T + \mathbf{G}(\tau) \mathbf{Q}(\tau) \mathbf{G}^T(\tau) \right] d\tau. \quad (8)$$

Kalman Gain Equation:

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}_k^T [\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T + \mathbf{R}_k]^{-1}. \quad (9)$$

State Update Equation:

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K}_k [\mathbf{z}_k - \mathbf{h}(\hat{\mathbf{x}}_k^-, k)]. \quad (10)$$

Covariance Update Equation:

$$\mathbf{P}_k^+ = [\mathbf{I}_n - \mathbf{K}_k \mathbf{H}_k] \mathbf{P}_k^-. \quad (11)$$

In the preceding equations, the superscript $-$ indicates the values before the measurement update and the superscript $+$ indicates the values after measurement update has occurred; \mathbf{K} is the Kalman gain; and \mathbf{P} denotes the state covariance matrix. The interest here is to obtain the state estimate, $\hat{\mathbf{x}}_k^+$, which are reasonably certain and are thus characterized by reasonably small values of the covariance matrix, \mathbf{P}_k^+ . The state detectability condition indicates that the \mathbf{P}_k^+ will not explode to infinite values, but the values of $\mathbf{Q}(t)$ used will certainly affect the calculated values of \mathbf{P}_k^+ .

Most of the process models that can be developed from limited experimental observations and with a reasonable investment in effort and time are far from perfect and involve significant uncertainties. This is especially true for batch and semibatch processes, where the models are substantial abstractions of a complex process. Thus using these models requires judicious specification of their accuracy to avoid biases in model prediction. This model accuracy is specified by tun-

ing the Kalman filter that involves the selection of the process-noise covariance matrix, Q . If this matrix is guessed low, the filter will believe the model excessively and will not use the on-line measurements properly to correct the states. This can lead to poor performance or even filter divergence. On the other hand, if the matrix Q is guessed higher than the actual value, the state estimates will be noisy and uncertain, as this would lead to increased values of the state covariance matrix, P . Choosing the right value of the tuning parameters is very important for successful application of EKF.

Proposed Approaches

The approaches in this article aim at making the selection of the Q matrix more systematic. This also helps to improve the performance of the filter when there is significant process-model mismatch. The proposed methods need the information about the model uncertainty in the form of a parameter covariance matrix. Usually this information is available during the parameter estimation step. The determination of the parameter covariance matrix during model development is demonstrated in one of the examples later. Both techniques utilize the information about the process in the form of a nonlinear model. Consider the nonlinear system as shown below:

$$\dot{\mathbf{x}}(t) = \mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t, \mathbf{p}] + \mathbf{w}(t), \quad (12)$$

where \mathbf{x} denotes the states, \mathbf{u} the deterministic inputs, \mathbf{p} denotes the model parameters, and $\mathbf{w}(t)$ the process noise due to modeling inaccuracies and other unmodeled effects. In the present studies, $\mathbf{w}(t)$ is considered to be an additive term solely due to the model uncertainties, as they often dominate the process-noise component. The model uncertainties can be either structural, parametric, or both. The structural uncertainties can in many cases be adequately captured as uncertainty in model parameters. The estimated parameters in

the model are assumed to be normally distributed with the means and covariances given by:

$$E[\mathbf{p}] = \mathbf{p}_{\text{nom}} \quad (13)$$

$$E[(\mathbf{p} - \mathbf{p}_{\text{nom}})(\mathbf{p} - \mathbf{p}_{\text{nom}})^T] = \mathbf{C}_p, \quad (14)$$

where \mathbf{p}_{nom} denotes the nominal parameter values and \mathbf{C}_p is the parameter covariance matrix.

The main idea is to use the available information about the model parametric uncertainties and translate this information to the process-noise covariance matrix, Q . Both approaches described below are executed on-line and utilize the latest information from the filter. Thus the estimates of the states from the measurement update at current time is used to perform the calculation of the process noise for the next time instant. The procedure is shown in Figure 1.

Consider the actual plant represented in the nonlinear state space form as given below.

$$\dot{\mathbf{x}}(t) = \mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t, \mathbf{p}(t)], \quad (15)$$

where \mathbf{p} denotes the parameters of the plant. An approximation to the plant is obtained using the nonlinear model with the nominal values of the parameters, as it will be used for the on-line estimation:

$$\dot{\mathbf{x}}_{\text{nom}}(t) = \mathbf{f}[\mathbf{x}_{\text{nom}}(t), \mathbf{u}(t), t, \mathbf{p}_{\text{nom}}]. \quad (16)$$

To make these model predictions identical to the real plant's evolution in time, we add to the righthand side of Eq. 16, the noise term $\mathbf{w}(t)$ with, as yet, unspecified noise statistics. If we are able to specify a $\mathbf{w}(t)$ for the model prediction to be exact, then Eq. 16 can be written as

$$\dot{\mathbf{x}}_{\text{nom}}(t) = \mathbf{f}[\mathbf{x}_{\text{nom}}(t), \mathbf{u}(t), t, \mathbf{p}_{\text{nom}}] + \mathbf{w}(t). \quad (17)$$

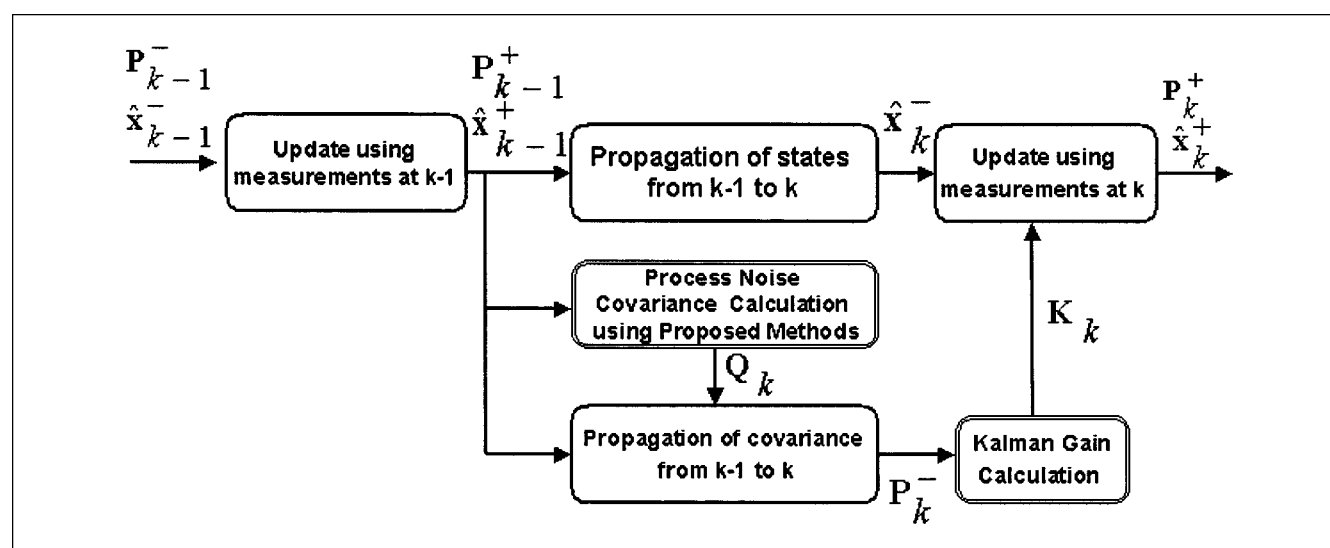


Figure 1. Determination of process noise covariance, Q on-line.

The method utilizes the states after the measurement update to calculate the Q for the next time update of state covariance.

Observing that the lefthand side of both Eq. 15 and Eq. 17 are desired to be the same, we can calculate the value of $w(t)$ that will yield such a result:

$$w(t) = f[x(t), u(t), t, p] - f[x_{\text{nom}}(t), u(t), t, p_{\text{nom}}]. \quad (18)$$

The two methods that we propose differ in the way the $w(t)$ statistics are calculated from the known statistics of the p variables.

Linearized approach

The first approach is called the linearized approach to calculate the $w(t)$ statistics, because it utilizes a Taylor series expansion of the righthand side of Eq. 18. The nonlinear dependence of $w(t)$ on p can be linearized around the nominal parameter values and states as shown below:

$$\begin{aligned} f[x(t), u(t), t, p] &= f[x_{\text{nom}}(t), u(t), t, p_{\text{nom}}] \\ &+ \left(\frac{\partial f}{\partial x} \right)_{x_{\text{nom}}(t), u(t), t, p_{\text{nom}}} (x - x_{\text{nom}}) \\ &+ \left(\frac{\partial f}{\partial p} \right)_{x_{\text{nom}}(t), u(t), t, p_{\text{nom}}} (p - p_{\text{nom}}), \end{aligned} \quad (19)$$

with the higher-order terms neglected as small. Consequently, $w(t)$ can be approximated by

$$\begin{aligned} w(t) &= \left(\frac{\partial f}{\partial x} \right)_{x_{\text{nom}}(t), u(t), t, p_{\text{nom}}} [x(t) - x_{\text{nom}}(t)] \\ &+ \left(\frac{\partial f}{\partial p} \right)_{x_{\text{nom}}(t), u(t), t, p_{\text{nom}}} (p - p_{\text{nom}}) \\ &= J_{x, \text{nom}}(t)[x(t) - x_{\text{nom}}(t)] + J_{p, \text{nom}}(t)(p - p_{\text{nom}}). \end{aligned} \quad (20)$$

As nominal values of the states, we take the values of the current filter state estimates and we further assume that they are reasonably close to the actual state values. Thus the first term in the preceding expansion that involves the states equals zero, $[x(t) - x_{\text{nom}}(t) = 0]$ and the process noise is calculated by

$$w(t) = J_{p, \text{nom}}(t)(p - p_{\text{nom}}). \quad (21)$$

Calculating the expected value of both sides yields

$$\bar{w}(t) = J_{p, \text{nom}}(t) \Delta \bar{p} = 0, \quad (22)$$

indicating that the noise sequence $w(t)$ has zero mean if the linearization in the parameters employed is accurate. Then the desired calculation of the covariance $Q(t)$ of $w(t)$ is given by

$$Q(t) = J_{p, \text{nom}}(t) C_p J_{p, \text{nom}}^T(t), \quad (23)$$

where C_p is the parameter covariance matrix, which, as explained earlier, is obtained during the parameter estimation

step of the modeling task. For the implementation of this procedure, it is necessary to calculate the sensitivity matrix $(\partial f / \partial p)$, which can be determined on-line either analytically or numerically. Then the on-line calculation of $Q(t)$ is performed using the nominal parameter values, p_{nom} , and the current state estimates at each time step, both of which affect $J_{p, \text{nom}}$. (For the case where this linearization is accurate enough, the preceding approach is equivalent to using $G(t) \hat{w}(t)$ instead of $w(t)$ in Eq. 1. Then one needs to set $G(t) = J_{p, \text{nom}}(t)$ and select the covariance of $\hat{w}(t)$ as equal to C_p .)

The main advantage of this linearized approach is that it involves very simple algebraic calculations and can easily be executed on-line. The method is approximate, it assumes linearity in the effect of parameters, and may not work very well for strongly nonlinear cases. This is demonstrated in some example cases later.

Monte Carlo approach

The technique proposed here aims to remedy the dependence on the linearization approximation of the approach in the previous section and is a generalization of the methodology proposed by Fotopoulos et al. (1998). The prior study was directed toward batch processes and involved Monte Carlo simulations off-line to estimate a time-varying $Q(t)$. The main disadvantage was that the noise statistics would be different once the batch trajectory was different from the one used in the Monte Carlo simulations. A more appropriate technique would be to do the simulations on-line, which is proposed here. The measurement error covariance matrix, R , is usually known from the error statistics of the measurement device and is readily available. The process noise, $w(t)$, is mostly due to the uncertainties in the model that can be either parametric or structural. As mentioned earlier, the structural uncertainties are difficult to characterize, and their effects will be examined later in one of the example processes.

Monte Carlo methods (Kleijnen, 1974) can be applied to the simulations of stochastic systems with parameters that are described by certain probability distribution functions. Once the probability distribution function (PDF) of the parameters is known, Monte Carlo simulations are done with random sampling from this PDF. A number of simulations can be performed and the result is taken as an average from all the observations. In many cases, it is possible to predict the variance in the average result, and this can be used to decide the number of simulations that are needed to achieve a given error in the result. In the application discussed here, this is not utilized, as it was found that the values of interest do not vary significantly beyond a specified number of simulations (500 simulations). The major components of a Monte Carlo simulation include (1) the probability distribution function that characterizes the system; (2) random-number generator; (3) sampling rule; (4) estimation of the quantity of interest from the observations; (5) error estimation as a function of number of simulations; and (6) possible variance reduction techniques. Monte Carlo simulation is routinely used in diverse fields.

Monte Carlo simulations are utilized here to capture the effect of parameter uncertainties on the process-noise statistics. For this, the parameters can be assumed to be normally distributed with mean value equal to the best parameter esti-

mate and standard deviation obtained from the parameter covariance matrix. If a more elaborate probability distribution function is available for the estimated model parameters, it can easily be used for the simulations. The values of the residuals from Monte Carlo simulations are used to obtain an estimate of the process noise covariance, Q .

The central idea is to capture the effect of uncertainties in the parameters through the statistics of the $w(t)$ variable. At a desired time instance, Monte Carlo simulations are run with parameter values sampled from the assumed normal distribution, with means and covariances as given in Eq. 14. For the Monte Carlo simulation number k , the nonlinear model with the parameter values, p_{nom} , is used to calculate the process noise as

$$w^k(t) = f[\hat{x}(t), u(t), t, p^k] - f[\hat{x}(t), u(t), t, p_{nom}], \quad (24)$$

where k denotes the index of the simulation. These noise values are used to find the process-noise means and covariances for that particular time instant. The process noise is supposed to have zero mean for a system where the parameter effects are linear, as was shown in the previous section. But for nonlinear systems, this is not quite valid. Thus the nonzero mean property is to be accounted for and the deviation from the mean is defined as below:

$$\tilde{w}^k(t) = w^k(t) - \bar{w}(t), \quad (25)$$

where $\bar{w}(t)$ denotes the mean of the process noise. The matrix Q is obtained as the covariance of these $\tilde{w}^k(t)$ values, assuming a normally distributed data set. The process-noise covariance matrix, Q , thus calculated from $\tilde{w}^k(t)$ is a nondiagonal and time-varying matrix. Because the process measurements are available at discrete time instances, the preceding calculation of $w^k(t)$ and $Q(t)$ will be done at the same discrete time instances. To keep the on-line calculations reasonable, we assume that their values are constant during the sample interval.

The mean of the process noise obtained is utilized during the state propagation step

$$\hat{x}_{k+1}^- = \hat{x}_k^+ + \int_k^{k+1} f(\hat{x}, u, \tau) d\tau + \bar{w}^k(t) * T_s, \quad (26)$$

where T_s is the sample time of the filter. The process-noise covariance is utilized in the covariance propagation Eq. 8, where $Q(\tau)$ is substituted by the Q_k value that is constant through this interval.

In the determination of the process covariance Q earlier, the process noise is assumed to be white, Gaussian random process. The approximation of normally distributed process noise was tested using the values of the process noise, $w(t)$, obtained from the Monte Carlo simulations. For this, the normal probability plots were used and the distribution was observed to be approximately normal. The nonzero mean characteristics of the process noise are accounted for here. For the case studies presented later, 500 Monte Carlo simulations of the different parameter values were used, resulting in 500 evaluations of the process noise. Beyond this number, the process-noise statistics of interest were found to be unaf-

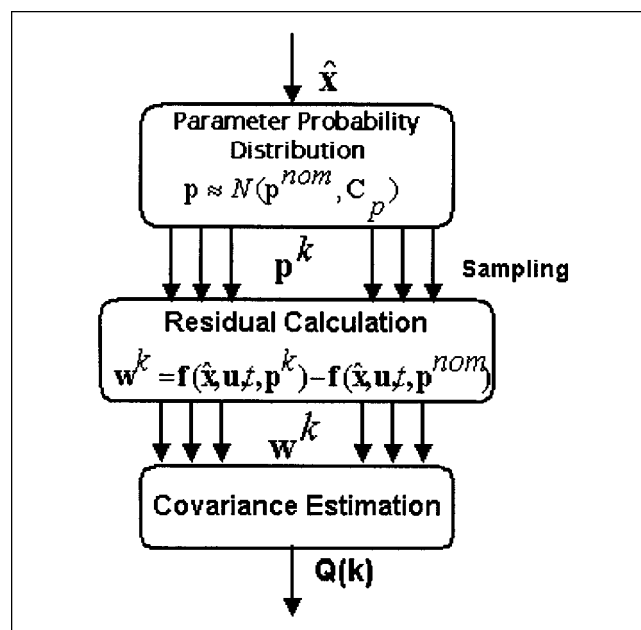


Figure 2. Monte Carlo method for the calculation of the covariance matrix $Q(t)$.

ected by the number of parameter samples. Because the Monte Carlo simulations do not entail any iterative calculations, the additional computational burden is minimal for sampling times that are not excessively short. The Monte Carlo technique showing the three major steps is given in Figure 2.

Case Studies

The efficiency of the techniques just described is studied with three different example systems. A variety of plant-scale chemical processes are studied, which include both batch and continuous. Brief descriptions of these processes are given below.

Transesterification Process. A transesterification reaction taking place in CSTR or in a semibatch reactor.

Methyl Methacrylate Polymerization. A two-time-scale filter approach is studied, in which two EKF's are used in parallel to handle the problem of frequent and infrequent measurements with measurement delay.

Semibatch System Using Tendency Models. In this case study, the model used in the estimator is simpler than the simulation model and is derived by use of the tendency modeling approach with the data collected from the simulated process. This is the more realistic case studied, as it resembles the approach that one will undertake with a real experimental unit.

The performance of EKF's with different approaches in the calculation of Q is compared for the example processes discussed earlier. The following choices for calculating Q are examined.

Monte Carlo Approach. Introduced in this article.

Linearization Approach. Introduced in this article.

Time-Average Approach. Where Q is obtained as the mean of the values from the Monte Carlo case.

Diagonal Approach. Where Q is equal to the diagonal values of the matrix from the time-average case. The off-diagonal terms are set to zero.

In presenting the EKF performance, the results of the diagonal approach are not shown in the plots, as the performance in this case is found to be poor for most of the cases. But the values of performance indices are presented for comparative purposes. It is possible that changing the diagonal values of the matrix Q can improve performance, but this would involve trial and error to find the right values. What is also important to note is that none of the preceding four cases realistically represents the difficult situation that a designer faces in order to select the appropriate values of Q by a trial-and-error approach.

Many chemical processes are characterized by a number of intrinsic features that are never completely captured in a model. A typical chemical process has its inherent variability. Apart from that, the estimation of parameters result in an error in the parameter values, which causes systematic process-model mismatch. In a realistic case, both the inherent process variability and the parameter errors are present to an extent. This is especially true for many complex batch processes, where all the intrinsic process features are rarely captured in a practical model. To include the preceding effects, different types of uncertainties are studied for the cases just listed. The types of uncertainties used in the simulations are briefly discussed below.

1. **Fixed Parametric Uncertainty.** In this case, the parameters for the simulated plant and the model are assumed to differ by a fixed value as below. For the parameter i , this is given as

$$p_{i,\text{plant}} = p_{i,\text{nom}} \pm \sigma_i, \quad (27)$$

where $p_{i,\text{plant}}$ denotes the parameter values for the plant, $p_{i,\text{nom}}$ stands for the nominal values used in the nonlinear model used by the EKF, and σ_i denotes the standard deviation for the parameter. The parameter covariance matrix used for the process-noise covariance calculation is assumed to be diagonal, with the diagonal values given by Eq. 28

$$C_{p_{ii}} = \sigma_i^2. \quad (28)$$

2. **Random Variation in Plant Parameters.** Many of the batch and semibatch processes are characterized by structural uncertainties that are time varying. To capture this effect, the parameters in the plant are assumed to vary with time, taking values at each sample interval from a normal distribution. Thus,

$$p_{i,\text{plant}} \approx N(\bar{p}_{i,\text{plant}}, \sigma_{i,\text{plant}}), \quad (29)$$

where $\bar{p}_{i,\text{plant}}$ denotes the mean value of plant parameters, and $\sigma_{i,\text{plant}}$ is the standard deviation for the same. The mean value of the varying plant parameter is assumed to be different from the nominal value of the model parameter by a fixed amount

$$\bar{p}_{i,\text{plant}} = p_{i,\text{nom}} \pm \sigma_i. \quad (30)$$

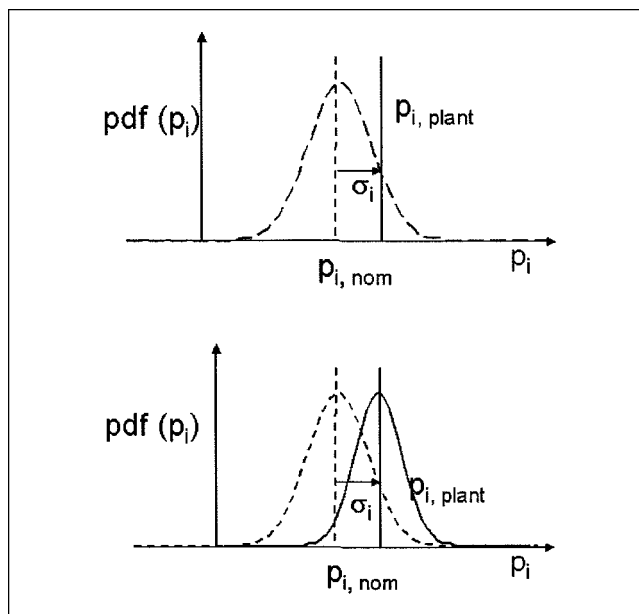


Figure 3. Different types of process-model mismatch utilized for the case study simulations.

In the first case, there is a fixed mismatch between the process and the model parameters, and the envelope with dotted lines shows the parameter distribution utilized for the proposed methods. In the second case shown in the bottom figure, the solid envelope shows the random variation in the plant parameters and the dotted envelope shows the distribution used for the proposed methods.

The parameter covariance matrix used in the EKF is still given by the Eq. 28, and the two parameter distributions are depicted in Figure 3.

3. **Structural Model Uncertainties.** The case study that utilizes the tendency model involves a very realistic scenario where the EKF model is structurally different from that of the simulated plant.

A relative measure of performance is selected in which the possible difference between the estimated and the true value of a state from the EKF are compared to the same differences of an open-loop estimator. Open-loop estimation is taken here to imply the estimation of the system states by use of the dynamic model without utilizing the benefits of the process measurements

$$J_i = \frac{\|\hat{z}_i - z_i\|}{\|z_{i,ol} - z_i\|} = \frac{\sqrt{\sum_{j=1}^N [\hat{x}_i(j) - x_i(j)]^2}}{\sqrt{\sum_{j=1}^N [x_{i,ol}(j) - x_i(j)]^2}}, \quad (31)$$

where the notation z_i denotes the vector with the true values of the state i for all the sample instances. Thus, \hat{z}_i denotes the EKF state estimates and $z_{i,ol}$ is the open-loop estimate. The L_2 norm is used for this calculation, and N denotes the number of data points in the simulation. Thus this performance index gives a measure of improvement obtained by using the EKF compared to the case where the model is used

in open loop. This quantifies the benefit obtained by using the closed-loop filter, which can then be compared to the time needed for tuning the filter. The smaller the value of this index, the better the state estimation performance.

The initial values of the state estimates, $\hat{x}(0)$, and the initial value of the state covariance matrix, $P(0)$, have an important effect on the performance of the filter. However, it is not appropriate to consider them, and in particular, $P(0)$, as a tuning parameter. To obtain the best performing filter, one needs to select an $\hat{x}(0)$ that is as close to the true value $x(0)$. Only our knowledge about the specific process can help us make this choice as accurate as possible. Then we need to make an assessment as to how confident we are about this choice and quantify such an assessment in the values of $P(0)$. While we wish to keep $P(0)$ as small as possible, we should never allow it to be too small. A small $P(0)$ value implies to the filter that we are quite confident about our $\hat{x}(0)$ values. In case this is not really true, the filter is provided false information and there should be no surprise if the filter diverges. For the case studies examined here, the $P(0)$ is assumed diagonal with its i th entry equal to

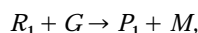
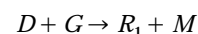
$$p_{ii}(0) = [x_i(0) - \hat{x}_i(0)]^2. \quad (32)$$

Case Study 1: transesterification process

This case study pertains to an industrial process for the transesterification stage of polyethylene terephthalate (PET) production. PET is the basic raw material for the production of films, fibers, and plastic objects. The production of PET is carried out in two stages: (1) the transesterification step, and (2) the polycondensation step. The first stage is studied here and involves the production of bis(2-hydroxyethyl) terephthalate by reacting dimethyl terephthalate with ethylene glycol. It has been reported that the presence of diethylene glycol (DEG) even in small concentrations can cause decreased polymer crystallinity, leading to unacceptable properties in the final polymer (Choi and Khan, 1988). Furthermore, the amount of the different species in the transesterification product has an impact on the final product quality. Thus the on-line composition monitoring of the different species in this process becomes very important.

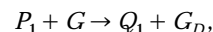
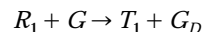
Both CSTR and semibatch reactors have been used for the transesterification reaction, and the state estimation in both cases is examined here. In the CSTR process, molten DMT, ethylene glycol, and the metal acetate catalyst are fed and the reactor, which is kept under an inert atmosphere in a temperature range of 180–200°C. The vapors, which are composed mostly of methanol, some ethylene glycol and trace amounts of other compounds, are rectified in a distillation column and the ethylene glycol is recycled back to the reactor. The methanol vapors are condensed and removed.

A detailed reaction mechanism for this process involves a number of intermediate species and reaction steps. A less detailed, yet workable reaction network for transesterification, was given by Choi and Khan (1988) and involves the following important reactions:



where D denotes dimethyl terephthalate, G denotes ethylene glycol, R_1 denotes methyl hydroxyl ethyl terephthalate, P_1 stands for bis-hydroxyethyl terephthalate, and M denotes methanol. The reverse reactions are insignificant as long as the continuous methanol removal is performed efficiently.

There are two important side reactions that lead to the formation of the diethylene glycol as shown below:



where T_1 is methyl terephthalate acid and G_D denotes diethylene glycol. The second side reaction is not considered in this study, as it is not significant under the conditions of interest. Its elimination reduces the dimensions of the problem.

The model for the transesterification process consists of six material balance equations, two heat balance equations for the reactor and coolant, and a volume equation (total mass balance with constant density assumption) for the case of semibatch operation. The equations corresponding to these reactions used to simulate the process here are given by Choi and Khan (1988) with only a small change in the way the catalyst effect is accounted. This catalyst effect on the reaction rate is modeled as a modifier of the reaction frequency factor (Ravindranath and Mashelkar, 1981). For the CSTR operation, it is assumed that the reactor is operated at a constant volume. Thus the effluent flow rate can be related to the methanol-vapor flow rate and the feed rate as below:

$$q_1 = q_0 - V_m(r_1 + r_2)V, \quad (33)$$

where q_1 is the effluent flow rate, q_0 is the inlet flow rate, r_1 and r_2 are the rate of the reactions that produce methanol, and V_m is the molar volume of methanol. The kinetic parameters and the physical properties are obtained from the literature (Choi and Khan, 1988).

In the application of EKF to the transesterification process reported by Choi and Khan (1988), both frequent (temperature) and infrequent, but delayed, measurements (combination of concentrations) were used. For such cases, it is possible to use a two-time-scale filter to achieve the maximum accuracy in the state-estimation performance. To simplify the procedure here, we assume that all the measurements are available with the same sampling rate. Besides the temperature measurement, we assume that an additional measurement, related to linear combination of concentrations, is available at each time, with a sampling interval of 1 min. Such a measurement could be available from an on-line spectroscopic measurement system. Another important measurement is the amount of methanol produced per unit time. It contains the information about the rates of two reactions. All the measurements available in this process are listed below with the amount of measurement noise assumed in parentheses:

- Reactor temperature (1%).
- Jacket temperature (1%).
- Flow rate of methanol drawoff (5%).
- Infrared spectroscopy measurement proportional to the sum of concentrations of all species ($\sum_{i=1}^6 C_i$) (5%). This is not a very realistic assumption, as the different species in the

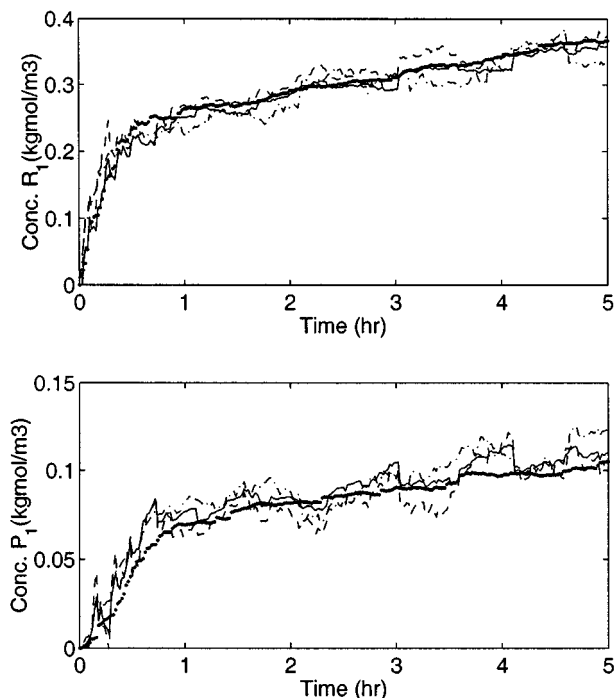


Figure 4. State estimation using EKF for the transesterification reactor startup case.

Actual plant (thick line), Monte Carlo method (solid line), linearized approach (dashed line), and time-average approach (dotted line).

measurement are weighted differently by the infrared spectroscopy.

- For semibatch operation, the reactor volume is also measured (5%).

Performance of Filter-Design Methods for the Startup of the Reactor. The state estimates of the EKF for a reactor startup case is shown in Figure 4 (feed flow rate = 300 L/h, feed concentrations of $D = 0.5$ kmol/m³, and $G = 1.0$ kmol/m³, respectively). The mismatch between the plant and the model is assumed to be in the form of both a fixed and randomly varying parametric uncertainty (5% of the parameter value). The plant values for the parameters E and A are taken to be 5% larger than the corresponding values used in the model. In the randomly varying parameter case, the standard deviation of the plant parameter is 5% higher than that of the model parameters. The comparison between the three methods (Monte Carlo, linearized, and time-average) are shown in Figure 4 for the estimation of the two important product concentrations in the reactor. The concentration of reactants, R and G , are estimated well, as the measurements are rich in information about these states and the uncertainties do not bias the predictions of these states. The main challenge is in estimating the product concentrations, R_1 and P_1 . The concentration of P_1 is not estimated quite well due to the fact that the uncertainties in the first reaction are compounded by the uncertainties in the second reaction. Further, a comparison of performance measure for different tuning procedures is also presented in Table 1. The plots do not show the results for the case of fixed and diagonal Q , as the performance in this case is very poor.

Table 1. Quantitative Performance Measure, J_i for State Estimation Using EKF for Transesterification Reactor Startup Case

Case	D	G	R_1	P_1
Diagonal, fixed Q	1.116	0.685	0.756	1.228
Nondiagonal, fixed Q	0.436	0.151	0.548	1.074
Linearization Approach	0.230	0.158	0.306	0.816
Monte Carlo Approach	0.229	0.154	0.261	0.677

Performance of Filter Methods for a Step Change in Feed Flow to the Reactor. The performance of different EKF tuning techniques for a step change in the feed (from 300 L/h to 450 L/h) to the transesterification reactor is studied next. For this, the reactor is allowed to come to a steady state from the startup case shown before. Here there is no initial mismatch between the filter states and the process states. To initialize the filter, the state covariances are kept diagonal at very low values. The model uncertainty in this case is considered to be a fixed 5% parametric error in the process parameter values.

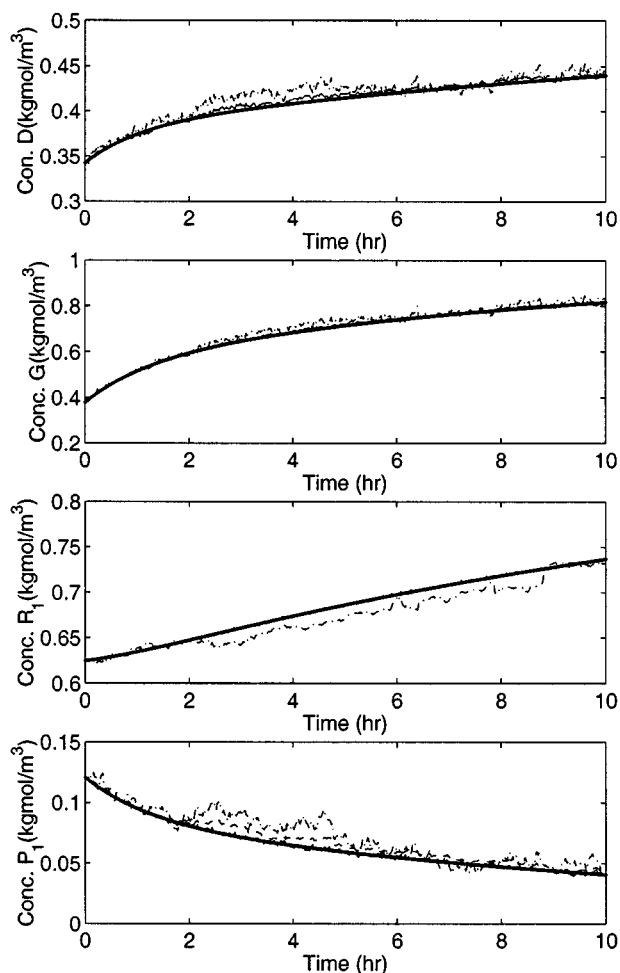


Figure 5. State estimation using EKF for a change in feed rate to transesterification reactor.

Actual plant (thick line), Monte Carlo method (solid line), linearized approach (dashed line), and time-average approach (dotted line).

The state estimates obtained using EKF are shown in Figure 5. For the reactant concentrations (D and G) and one of the product concentrations, both the Monte Carlo and linearization approaches perform equally accurately. Only in estimating the concentration of states R_1 and P_1 is there a noticeable difference in the performance. There we see a much inferior performance of the filter with the diagonal Q matrix. A comparison of the time variation in the diagonal values of Q calculated by the Monte Carlo and linearization approaches showed that the difference in the average values for the two cases is small, but there is a good difference in variability.

Performance of Filter-Design Methods for Semibatch Operation. The transesterification reaction can also be carried out in semibatch mode, as discussed in Ravindranath and Mashelkar (1981). In this case, the DMT is charged initially in the batch and the ethylene glycol is fed into the reactor continuously. The remaining features including the methanol removal and measurements remain the same as in the CSTR

operation. A comparative plot of the plant and the open-loop model prediction is shown in Figure 6. (Semibatch feed rate = $300 \text{ m}^3/\text{h}$, concentration of G in feed = 0.5 kmol/m^3 .) It is to be noted that the initial states for the concentration of D and G are off from the actual values by $+0.3 \text{ kmol/m}^3$ and -0.3 kmol/m^3 , respectively. The process-model mismatch is introduced by allowing the plant parameters to take random values around the mean values that were 5% different from the parameter values used in the process model. The random variation of the plant parameters follows a normal distribution with a standard deviation equal to 5% of the actual parameter values. The state-estimation results are given in Figure 7. The filter is able to converge to the right values subsequently. As can be seen from the plots and the table, the EKF using Monte Carlo simulations performs better than the other methods. The results for individual states are quite similar to the case study of reactor startup. The corresponding Q values are also shown in Figure 8. It is clear from this figure that the linearization approach leads to slightly reduced val-

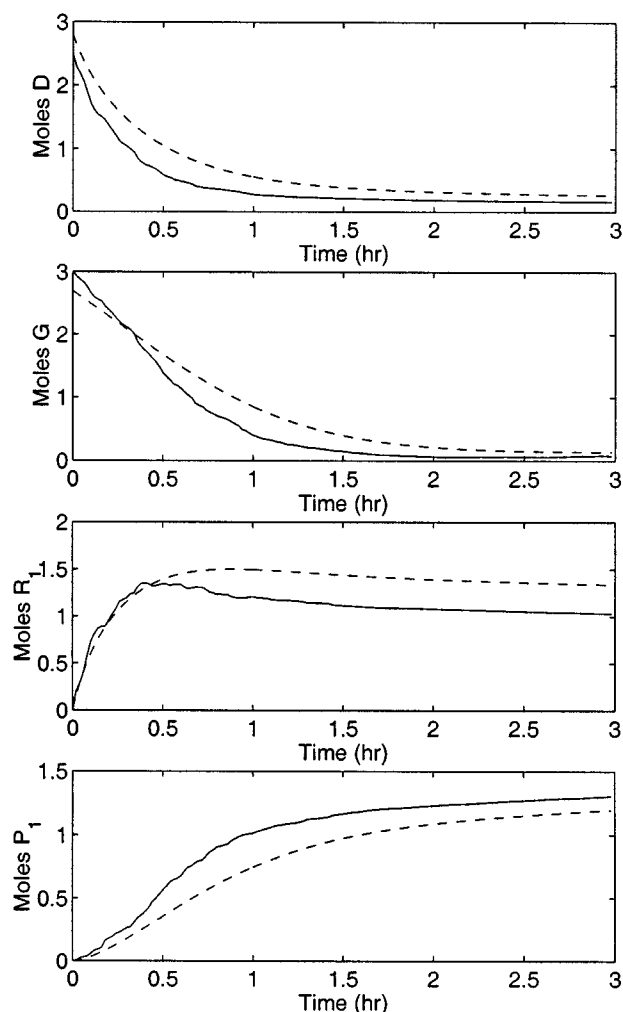


Figure 6. Open-loop model predictions vs. plant's behavior for the semibatch operation of transesterification process.

The solid lines show the plant and dashed lines denote the model predictions.

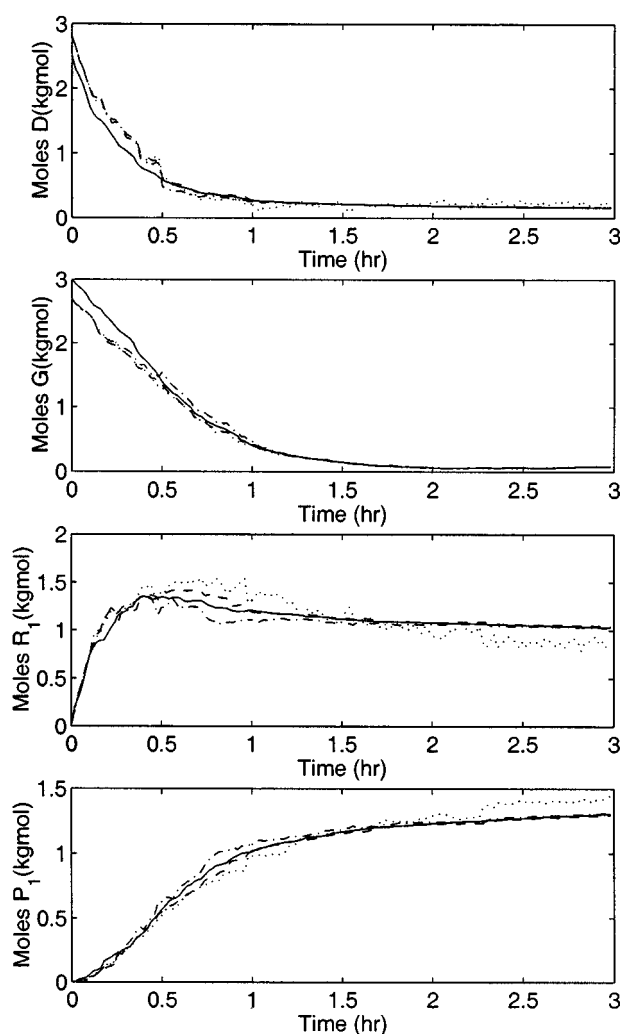


Figure 7. Performance of different tuning techniques in estimating the four concentrations.

The main difference is in the estimation of concentration of R_1 .

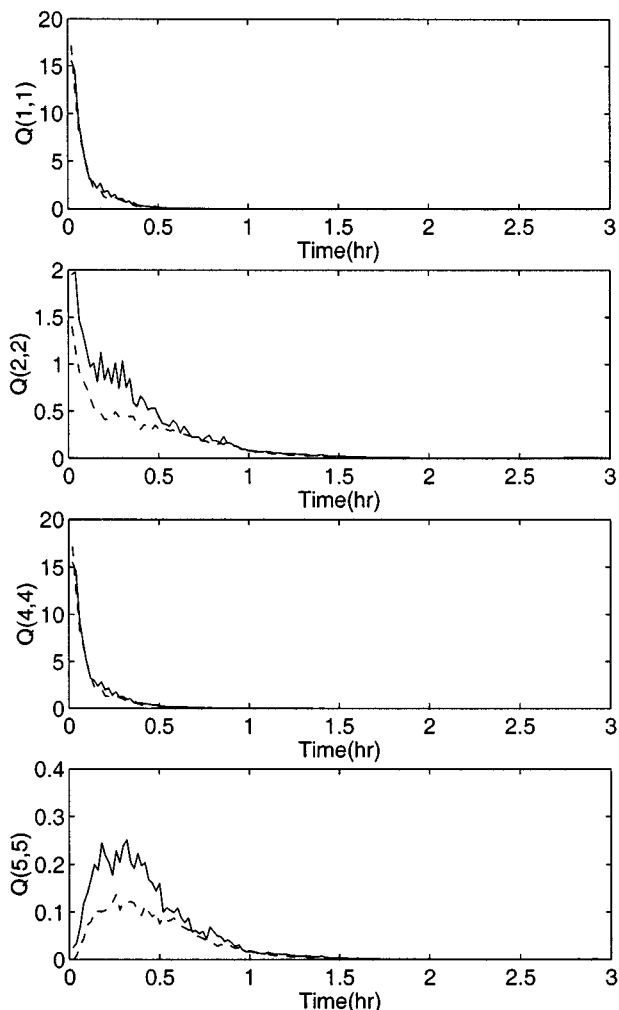


Figure 8. Q for Monte Carlo (solid line) vs. linearization case (dashed line).

ues of the process noise covariance compared to the Monte Carlo case.

Performance of Filter-Design Methods for Pure Batch Operation. The transesterification reaction carried out in a batch reactor where both the reactants are charged initially and methanol is stripped off in a continuous manner (initial charge of $D = 2.5 \text{ kmol/m}^3$ and $G = 6.0 \text{ kmol/m}^3$). The values of initial states for the concentration of D and G in the filter differ from the actual values by 1.0 kmol/m^3 for both the reactants. The process-model mismatch is again introduced in the form of a fixed 5% difference in the model parameters and the average value of randomly varying plant parameters, as discussed before. In this case, the filter is able to converge to the right values in a very short amount of time (Figure 9). The open-loop model predictions show significant mismatch comparable to the previous case of semibatch operation. The EKF performs well in estimating this state in spite of the significant process-model mismatch (Table 2).

The Monte Carlo and the linearization approaches work significantly better than the case where Q is kept constant, diagonal, or nondiagonal. The performance of the Monte Carlo method compared to the linearization approach is no-

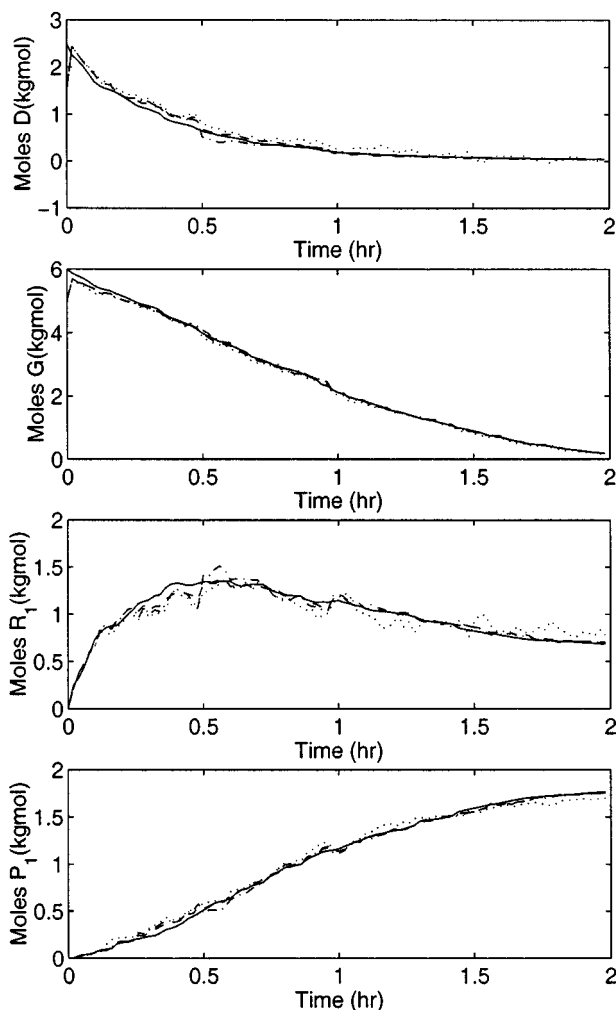


Figure 9. Performance of different filter designs in estimating the four concentrations.

Thick line (plant), solid line (Monte Carlo approach), dashed line (linearization approach), and dash-dot line (nondiagonal, fixed Q).

Table 2. Quantitative Performance of Different State Estimation Designs for the Batch Transesterification Process

EKF Design Case	D	G	P_1	R_1
Diagonal, fixed Q	0.730	1.965	1.311	1.180
Nondiagonal, fixed Q	0.554	0.249	0.274	0.106
Linearization Approach	0.444	0.213	0.178	0.066
Monte Carlo Approach	0.418	0.208	0.121	0.043

ticeably better for the states P_1 and R_1 , which are significantly biased by the process-model mismatch.

Case Study 2: Methyl methacrylate polymerization

The solution polymerization of methyl methacrylate (MMA) to form PMMA has been studied by different authors. In these processes, the important final-product properties are strongly dependent on the molecular weight distribution (MWD) of the polymer. The main issue is still that the

MWD is usually not measurable on-line and thus needs to be estimated from other available on-line measurements. A number of studies are reported on state estimation for MMA polymerization process (Tatiraju and Soroush, 1997; Adekun and Schork, 1989; Crowley and Choi, 1998). The attempt here is to study the efficacy of the tuning procedures on the state estimation of this process. Thus, many of the previously reported concepts are used here.

The MMA polymerization is usually carried out as a free-radical reaction using AIBN as an initiator and toluene as solvent. The equations used to simulate the process are given by Schmidt and Ray (1981) and are not reproduced here. The MMA system consists of eight states, which are initiator, monomer, and solvent concentrations; reactor and jacket temperatures; and the three moments of molecular-weight distribution. The kinetic parameters reported in Schmidt and Ray (1981) are used in the simulations.

The following measurements are considered to be available in the process:

- Reactor and jacket fluid temperature with 1% measurement noise.
- Monomer concentration in the reactor with 5% measurement noise.
- Delayed measurements of number-average and weight-average molecular weights (30-min delay) with 5% measurement noise.

In MMA polymerization, frequent measurements of the temperatures and density of the reactor contents are usually available. Density measurements can be used to estimate the conversion from MMA to PMMA. Thus, one has temperatures and the monomer concentration as the three on-line measurements with which the initiator concentration can be estimated. It can be seen from the system equations that the three moments of molecular-weight distribution are not observable with the available frequent measurements, and hence cannot be estimated in closed loop using an EKF. In such a case, one would just estimate the observable states and perform an open-loop prediction on the unobservable states. This could lead to significant error in the molecular-weight predictions. Alternately, one can infrequently, and with a certain time delay, measure the number-average and weight-average molecular weights off-line. A two-time-scale filter is used in such situations, and a number of such studies are reported in the literature (Scali et al., 1997; Liotta, 1996; Liotta et al., 1997). When only the frequent measurements are available, the filter performs closed-loop estimation with only the observable states, while the states related to molecular-weight distribution are estimated in open loop. When the delayed measurements of these states are available, a full-order filter is used to re-update all the states at the measurement time. This update is then used along with the later measurements to update the state values to the present. The additional details on two-time scale filter are available in Liotta (1996) and Liotta et al. (1997).

The process-model mismatch is introduced in the form of both a fixed parametric difference and random variation in the kinetic parameters (3% of the parameter value). These very infrequent measurements are not able to compensate for the process-model mismatch, and the estimation of moments of the molecular-weight distributions are biased. The information about the model uncertainty is utilized in the calcula-

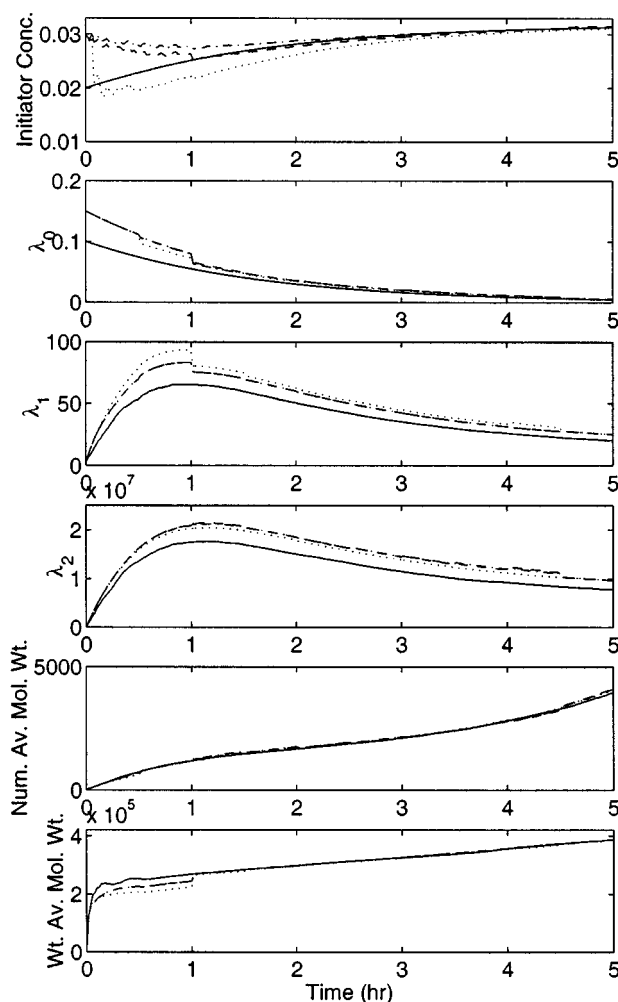


Figure 10. State estimation using EKF for a startup case.

(1) Thick line—actual plant; (2) solid line—Monte Carlo method; (3) dashed line—linearization approach; and (4) dash-dot line—nondiagonal, fixed Q .

tion of Q on-line using the two proposed approaches. Both the sampling frequency and the measurement delay for the molecular-weight states are taken to be 30 min.

The results of state-estimation studies for the startup of the MMA polymerization reactor are shown in Figure 10 (feed flow rate = 0.6 m³/h, inlet conc. of monomer = 5 kmol/m³, inlet conc. of initiator = 0.032 kmol/m³), and quantitative results are given in Table 3. The estimates of initiator concentration, moments of molecular-weight distribution (MWD) and the number-average and weight-average molecular weights are shown in Figure 10. Both the linearization and

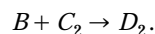
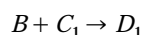
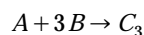
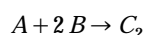
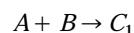
Table 3. Quantitative Performance Measure for the Initiator Conc. and Molecular Weights for MMA Reactor

Case	Initiator Conc.	Num. Av. Mol. Wt.	Wt. Av. Mol. Wt.
Diagonal, fixed Q	0.022	0.089	0.231
Nondiagonal, fixed Q	0.016	0.079	0.177
Linearization Approach	0.018	0.077	0.143
Monte Carlo Approach	0.017	0.077	0.146

Monte Carlo approaches are found to perform equally well in this case. The first delayed measurement sample is assumed to be taken at the start of the run at $t = 0$ min. The EKF is able to converge to the right values of the number-average and weight-average molecular weights at $t = 1$ h, when the delayed measurement becomes available. The constant and nondiagonal value of Q performs better than in the previous case studies. An interesting observation is the estimates of initiator concentration, where the convergence using a fixed value of Q seems faster initially. This is due to the fact that the values of Q are smaller in this case, and thus the filter believes the model more, which leads to the sudden drop of initiator concentration. But it takes much longer for this state to converge to the right values.

Case Study 3: Semibatch system using tendency models

A semibatch reaction network that is representative of fine and speciality chemical environments is studied as an example here. The simulated system has five reactions in series and parallel and involves seven species. Reactants A and B are initially fed to the reactor along with the solvent. The reactant B and additional solvent are added in semibatch mode. The desired product in the system, C_2 , and C_1 , C_3 , D_1 , and D_2 are byproducts. The five simulated chemical reactions are given below:



The rate law for these reactions can be expressed in general as

$$r = Ae^{-E/RT}f(C). \quad (34)$$

The values of A and E , as well as the function f describing the reaction rates for different reactions, are given in Table 4. The preceding reaction network and the kinetic model will be used to simulate the process, but the estimator will assume it is unknown. Consequently, we will need to develop the model of the process and, for this purpose, we will use the tendency modeling approach (Fotopoulos, 1996; Rastogi et al. 1992).

The data for development of the tendency model are obtained by simulating the process using the PC-based HYSYS simulator. These parameters are chosen to closely resemble

the epoxidation of oleic acid process that our group studied experimentally in the past (Rastogi et al., 1992). Sixteen simulated runs of the process, characterized by the two different values of the initial concentrations of A , B , batch temperature, and semibatch feed rate, were performed.

The tendency modeling approach involves the determination of the stoichiometric and kinetic models from the available batch data. The stoichiometry is obtained using singular value decomposition and structured target factor analysis (Fotopoulos et al., 1994). The singular value decomposition for the preceding system data yields four significant single values and a minimum of four reactions needed to adequately represent the process. The structured target factor analysis yielded 11 vectors, representing physically meaningful reactions, that lie in the stoichiometric space. These vectors are then grouped together so that all possible reaction networks with four linearly independent reactions are enumerated. The extent of the reactions are then calculated from the data for each of these reaction networks and are checked whether they are positive and monotonically increasing. If negative and decreasing extents are calculated for a reaction, its direction is reversed. The reaction networks that satisfy these criteria are accepted as possible stoichiometries. For the example process given above, six possible stoichiometries were obtained following this procedure. One of these stoichiometries composes all but the third reaction from the five reactions in the actual process.

It is to be noted that the third reaction $A + 2B \rightarrow C_3$ in the original network does not significantly occur at the conditions at which the process model is identified. Thus the tendency-modeling procedure is able to determine the actual stoichiometry taking place in the simulated process along with five other equally plausible stoichiometries. When one does not know the details of the simulated, or more appropriately, the experimental process, all six stoichiometries are equally acceptable.

Kinetic parameter estimation is the next step, and it consists of estimating the reaction orders, the frequency factors, and the activation energies. The Gauss Newton method was used to do the least-square fitting of the kinetic parameters. This involves determining the sensitivity matrix of the outputs with respect to the parameters, which is calculated by integrating the sensitivity equations along with the system equations using the ODESSA package (Leis and Kramer, 1988). The sensitivity matrix is also useful for estimating the confidence intervals of parameters. The stoichiometric and kinetic components of the model are coupled with dynamic material and energy balances to obtain the complete tendency model given below.

Complete Tendency Model.

$$V = \sum_{i=1}^7 \frac{N_i M_i}{\rho_i} + V_s + \frac{m_s t}{\rho_s} \quad (35)$$

$$\frac{d(VC_i)}{dt} = \sum_{j=1}^4 V_{r_j} \nu_{ij} + F_i^s \quad (36)$$

$$\begin{aligned} \frac{d(V\rho C_p T)}{dt} = & \sum_{j=1}^4 V_{r_j} (-\Delta H_j) \\ & + m_s C_{p_s} (T_s - T) + UA(T_j - T) \end{aligned} \quad (37)$$

Table 4. Kinetic Parameters and the Rate Expressions for the Actual Process

Reaction No.	A	E (KJ/Kgmol.)	ΔH (KJ/Kgmol.)	f
1	$1 \cdot 10^7$	41800	-832000	$C_A C_B$
2	$6 \cdot 10^{10}$	62800	-75100	$C_A C_B^5$
3	$2.93 \cdot 10^8$	37000	-110000	$C_A C_B^3$
4	$7 \cdot 10^6$	41800	-106000	$C_B C_{C_1}^5$
5	$4 \cdot 10^{13}$	83700	-90900	$C_B C_{C_2}$

$$V_c \rho_c C_{p_c} \frac{d(T_j)}{dt} = m_c C_{p_c} (T_{in_c} - T_j) + UA(T - T_j), \quad (38)$$

where V is the batch volume; V_s , m_s , and ρ_s are the volume, mass flow rate, and density of the semibatch feed; t is the time; C_i , N_i , M_i , ρ_i , F_i^s are the molar concentration, number of moles, molecular weight, density, and the semibatch molar feed rate of component i in the system; ρ , C_p , and T are the density, specific heat capacity, and temperature of the batch; C_{p_s} , T_s are the specific heat capacity and temperature of the semibatch feed; m_c , ρ_c , C_{p_c} , and T_c are the mass flow rate, density, specific heat capacity, and temperature of the jacket fluid; r_i and ΔH_i are the reaction rate and heat of reaction; U is the heat-transfer coefficient; and A is the heat-transfer area. The main features of the preceding model are that it is nonlinear and is only approximate. Within the adaptive character of the tendency modeling approach, the tendency model can be updated to improve its predictive capabilities as further data become available from subsequent batch runs.

Quantification of Uncertainty in Tendency Models. The tendency models are developed from a restricted number of plant data and there are unavoidable uncertainties in the model. Since any model is an abstraction of reality, both the structural and parametric uncertainties are present to some degree in most real situations. In the study presented here, the structural uncertainty is in the rate laws. In many cases, the structural uncertainties are partially and indirectly reflected in the uncertainty of the model parameters.

Different techniques have been proposed in the literature for estimating the confidence intervals for the model parameters. These include Monte Carlo simulations and bootstrap and jackknife approaches (Caceci, 1989). It is to be noted that these techniques involve estimating parameters for a number of cases and are computationally burdensome. An extension to the parameter confidence interval for linear systems can be used to approximately characterize the parameter uncertainty (Fotopoulos, 1996) in nonlinear systems. The parameter-estimation step involves the determination of the sensitivity matrix, S , which is given as $\partial f / \partial p$. This can be used to obtain an estimate of the parameter covariance matrix at the converged parameter values, as shown below:

$$C_p = s^2 (S^T S)^{-1}, \quad (39)$$

where S is the sensitivity matrix and s^2 is the residual mean square. The diagonal values of the C_p can be used to provide a rough estimate of the confidence limits for each of the estimated parameters. It is to be noted that the method just described is derived from the linear estimation theory and can result in an underestimate for highly nonlinear systems. The

estimated parameters and their standard deviations are given in Table 5.

State Estimation Using Tendency Models. When all the states of the system are not observable, the application of EKF demands judicious selection of the states to estimate. If the process is detectable, it is still possible to apply a full order filter and estimate all the states. This arises from the fact that Kalman filters for linear systems demand detectability for convergence. Such a scheme of state estimation for a polymerization system was given by Kozub and Macgregor (1992). It has been reported, however, that the full-order filter is sensitive to initial conditions and does not offer any apparent benefits (Scali et al., 1997). A different approach to the preceding situation is to estimate only the subsystem containing observable states. The unobservable subsystem can then be estimated using open-loop prediction.

The batch-reaction system under study is not completely observable, since the components D_1 and D_2 do not appear in any of the rate equations. Thus these states are affected by other states, but no feedback exists and the measurements do

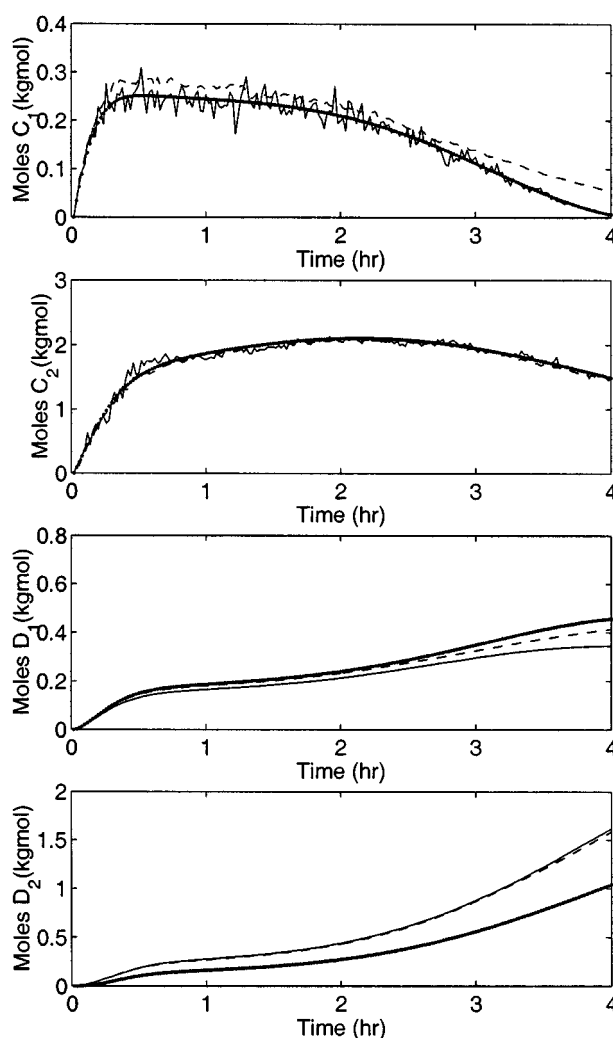


Figure 11. EKF for estimating C_1 and C_2 and open-loop predictions for D_1 and D_2 .

(1) Thick line—actual plant; (2) solid line—Monte Carlo method; (3) dashed line—linearization approach.

Table 5. Kinetic Parameters and their Standard Deviations for the Tendency Model

Reaction No.	A	Standard Deviations	E (KJ/Kgmol.)	Standard Deviations	f
1	$2 \cdot 10^5$	$\pm 5.145 \cdot 10^4$	27100	± 1901.3	$C_A C_B$
2	$6.12 \cdot 10^{12}$	$\pm 1.81 \cdot 10^{10}$	69080	± 1507.2	$C_A C_B$
3	$1.66 \cdot 10^8$	$\pm 4.53 \cdot 10^6$	45980	± 1946.6	$C_B C_{C1}$
4	$4 \cdot 10^{12}$	$\pm 4.266 \cdot 10^{11}$	76167	± 2033.9	$C_B C_{C2}$

not yield any information regarding these states. The analysis of observability matrix yields seven observable states, and the system is divided into two parts, the observable and the unobservable ones. The observable states include the concentrations of A, B, C_1, C_2 , volume V , and temperatures T and T_j . It is also to be noted that since there are only four reactions, only four components are needed to be estimated to determine the extent of the reactions. These components can then be used to determine the composition of other components in the system, provided we know the initial compositions of the additional species.

Thus the objective here is to estimate the concentrations of products C_1 and C_2 using EKF, and use this to calculate the concentrations of D_1 and D_2 , utilizing the extent of reaction or open-loop predictions. The following measurements are considered to be available in the process:

- Reactor and jacket fluid temperature with 1% measurement noise.
- Concentrations of species A and B with 10% measurement noise.
- Batch volume with 5% noise.

The results of state estimation for a semibatch case is shown

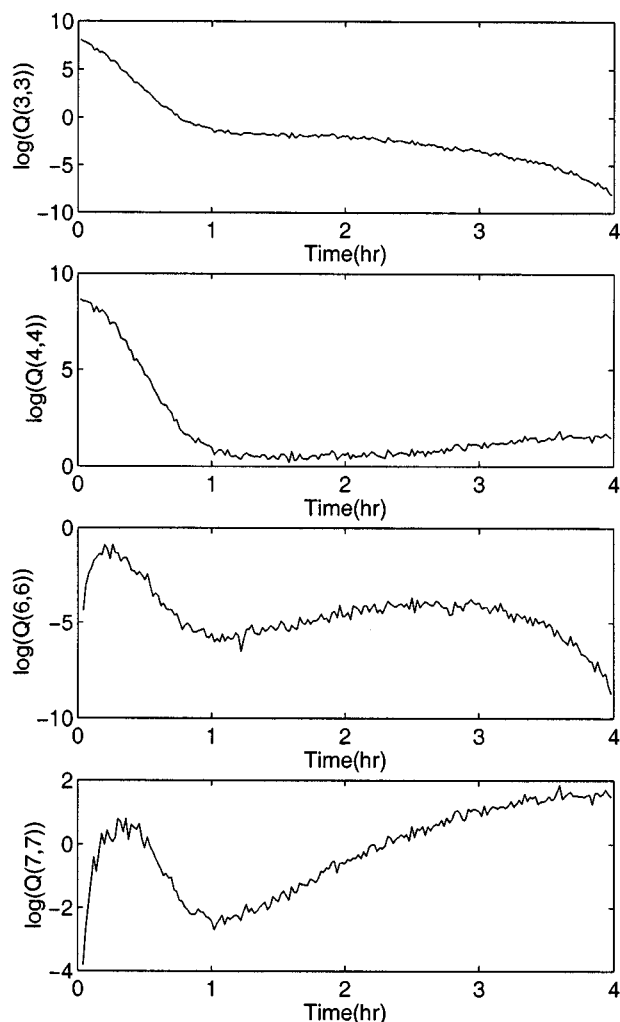


Figure 12. Diagonal values of Q for the four concentrations (C_1 , C_2 , D_1 , and D_2).

in Figure 11 (initial charge of $A = 3$ kmol and $B = 7$ kmol, and feed rate of $B = 1.0$ kmol/h). The model performs reasonably well under these conditions. Here two of the concentrations are estimated using the closed-loop EKF (C_1 and C_2), while the estimates of concentrations D_1 and D_2 are obtained using open-loop predictions. The concentration of the main product, C_2 , is estimated quite well. The estimation of the concentration of C_1 is not very good as the model predictions of this state are quite biased. The two states that are calculated by open-loop prediction using estimated states also show bias due to the process-model mismatch. The values of the diagonal terms of Q are shown in a logarithmic plot in Figure 12. The values undergo significant change with time in this case. Initially, the formation of C_1 and C_2 from A and B is significant. After about 1 h, C_1 and C_2 further react to D_1 and D_2 until the end of the batch.

The operation of this process with a different recipe is shown in Figure 13. Here the batch is started from an initial

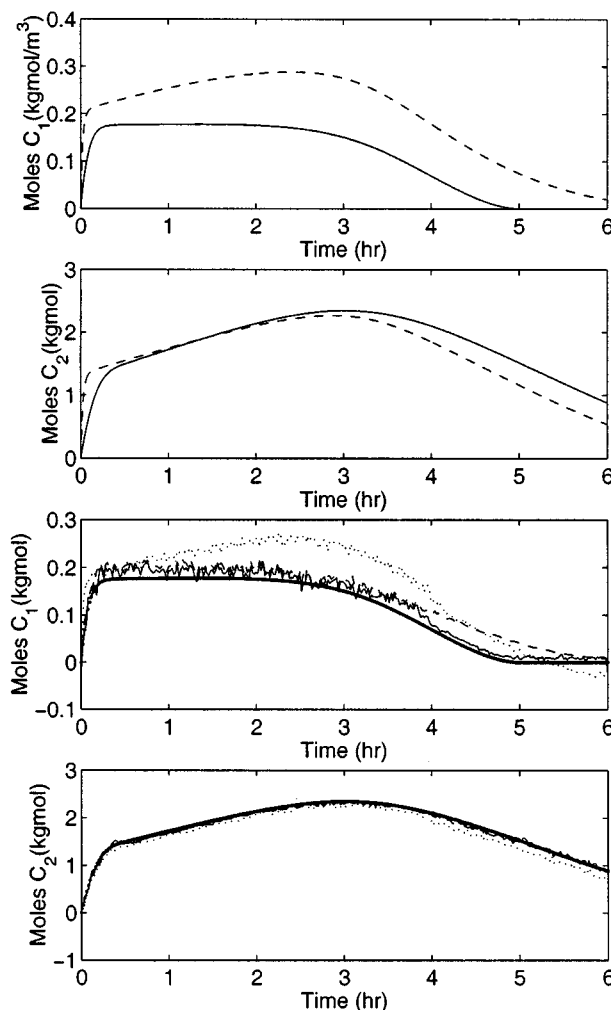


Figure 13. Open-loop predictions (top two) and closed-loop estimation (bottom two).

The plots on the bottom show the performance of linearization (dashed line); fixed, diagonal Q (dotted line), and Monte Carlo approaches (solid line); and in estimating the two product concentrations. Solid line represents true values in top two plots.

Table 6. Quantitative Performance Measure for the Two Product Concentrations Estimated Using EKF

Case	Moles C_1	Moles C_2
Diagonal, fixed Q	2.223	1.821
Nondiagonal, fixed Q	0.684	0.553
Linearization Approach	0.314	0.159
Monte Carlo Approach	0.168	0.190

point that was not covered when the runs were conducted during the development of the model. The batch is started with equal amounts of A and B (3 kmol) and a feed rate of $B = 1.5$ kmol/h. The concentration of C_1 is not estimated quite well due to the mismatch that affects this state significantly. In estimating this state, the Monte Carlo tuning pro-

cedure performs better compared to the linearization method. Both methods estimate the concentration of C_2 equally well. The performance of the other two tuning techniques is not very good and is given in Table 6. The estimation of concentrations using EKF for a pure batch case is also studied here. All the runs conducted for identifying the tendency model were done in semibatch with continuous addition of the reactant B . Thus the performance of EKF using the tendency model for a pure-batch operation is of interest (initial charge of $A = 3.0$ kmol, $B = 8$ kmol). The open-loop predictions using the model and the closed-loop EKF performance for the two product concentrations, C_1 and C_2 , are shown in Figure 14. The methods are able to estimate the concentrations quite well even though the model was not identified under these conditions.

The example given earlier was different from the other examples in one aspect: it involves using models that were developed from plant or experimental data using a systematic procedure, tendency modeling. Further, the covariance matrix for the parameters was also obtained from the data during the development of the model. The EKF performs quite adequately in estimating the states for this process.

Conclusions

In this article, we have presented two systematic approaches for the calculation of the process-noise covariance matrix, and we have examined the efficacy of these techniques in improving the performance of the EKF for different chemical processes. The results show that these approaches can go a long way toward making the design and the application of EKF easier, since the need to tune the filter is eliminated. The first method uses a linearization approach, whereas the second method utilizes Monte Carlo simulations to determine the process-noise covariance matrix. Both of these methods use information from the filter and are thus executed on-line. The calculations to be done on-line are nominal and are handled easily with the computing power available at present. (For the MMA polymerization system, the calculation of Q using the Monte Carlo method takes about 1.1 on an SGI-IRIX workstation.) The methodologies described earlier require the information about the process-model mismatch in the form of a parameter covariance matrix. In many cases, the parameter covariance matrix obtained during the identification is not very exact. Our experience shows that these techniques are quite robust, and their performance is not degraded seriously with a parameter covariance matrix that is not very exact. Thus there is a tolerance for the information about the model uncertainty within which the estimator will deliver reliable state estimates. If the user is not certain about the process-model mismatch, trials with different parameter covariance matrices might be attempted. However, the most rational approach is to have the best estimate of the parameter covariance and to introduce a single tuning coefficient, η , in front of the Q matrix with a nominal value of unity. When the filter diverges due to an underestimation of the process-model mismatch, the value of η can be easily increased. If, on the other hand, the model proves more accurate than initially considered, the value of η can be slowly reduced, keeping in mind that $\eta = 0$ corresponds to open-loop estimation.

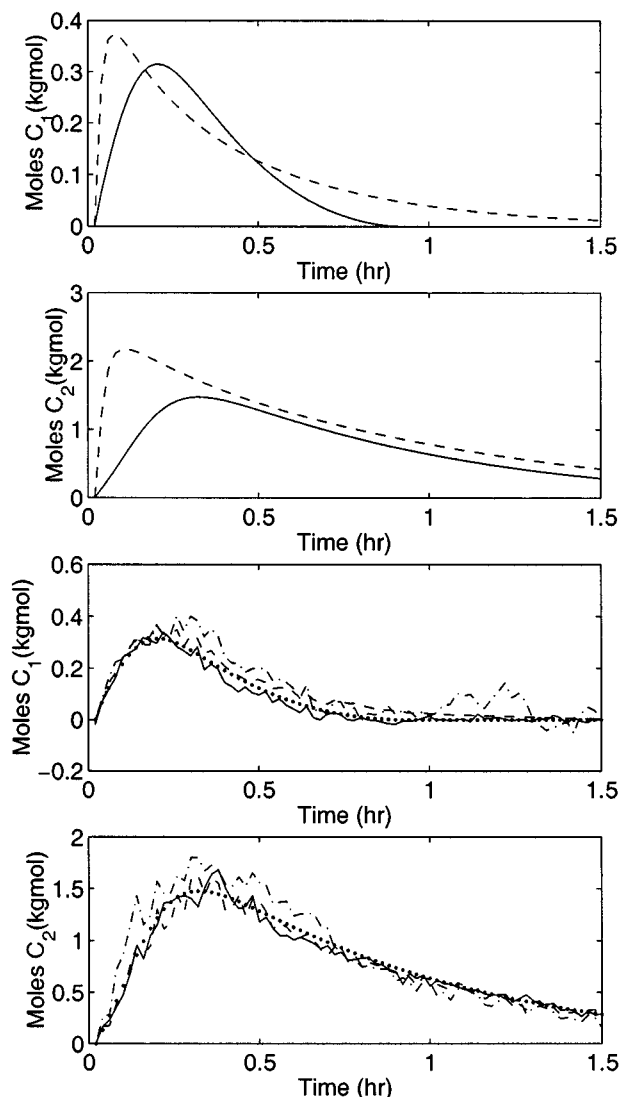


Figure 14. Open-loop predictions (top two) and closed-loop estimation (bottom two).

The plots on the bottom show the performance of different tuning techniques (dotted: actual plant; solid: Monte Carlo; dashed: linearization; dash-dotted: fixed, diagonal Q) in estimating the two concentrations. Solid lines represent true values in top two plots.

It is noted that if the model uncertainties bias the prediction of states significantly, the methods can result in very noisy estimates, especially for the Monte Carlo case. This is due to the fact that in these cases, the values of Q are very large and the filter believes the measurements to a large extent, thus capturing the noise in the measurements. Thus there is a limit to the process-model mismatch for which these techniques can deliver good estimates of the states. The adequacy of using the linearization approach or the fixed Q depends on the process under study. For cases that are only mildly nonlinear, the linearization method is quite sufficient and the introduction of Monte Carlo simulations does not change the state estimator performance significantly. Furthermore, the simulation studies clearly indicate that for the case where the Q matrix is not time varying, significant gains are achieved when using a nondiagonal Q compared to a diagonal one. This is due to the fact that the systematic parametric uncertainties cause significant cross-correlations between the process noises for different states. Hence, the use of a diagonal Q , which does not capture this effect, can result in poor filter performance. One needs to note that in most trial-and-error approaches for the tuning of the EKF, it is the diagonal form of Q that is used.

The main advantage of the techniques presented in this article is that they eliminate the need for the tuning of the EKF by using the information available about the parametric model uncertainty. The proposed techniques have been applied to several case studies related to systems of batch, semi-batch, and continuous processes operating at different regimes. These cases are characterized by significant variation in the calculated process noise covariance matrix, $Q(t)$, demonstrating that a constant diagonal form that the trial-and-error approaches use is inadequate. The case studies clearly show that the proposed methodologies hold significant potential for use in a variety of chemical as well as in other processes where reliable information about the states is important for monitoring or model-based control.

Literature Cited

- Adebekun, D. K., and F. J. Schork, "Continuous Solution Polymerization Reactor Control. 2. Estimation and Nonlinear Reference Control During Methyl Methacrylate Polymerization," *Ind. Eng. Chem. Res.*, **28**(12), 1846 (1989).
- Caceci, M., "Estimating Error Limits in Parametric Curve Fitting," *Anal. Chem.*, **61**, 2324 (1989).
- Choi, K., and A. Khan, "Optimal State Estimation in the Transesterification Stage of a Continuous Polyethylene Terephthalate Condensation Polymerization Process," *Chem. Eng. Sci.*, **43**, 749 (1988).
- Crowley, T. J., and K. Y. Choi, "Experimental Studies on Optimal Molecular Weight Distribution Control in a Batch Free-Radical Polymerization Process," *Chem. Eng. Sci.*, **53**, 2769 (1998).
- De Valliere, P., and D. Bonvin, "Application of Estimation Techniques to Batch Reactors—3. Modeling Refinements which Improve the Quality of State and Parameter Estimation," *Comput. Chem. Eng.*, **14**, 799 (1989).
- Dimitratos, J., C. Georgakis, M. S. El-Aasser, and A. Klein, "Dynamic Modeling and State Estimation for an Emulsion Copolymerization Reactor," *Comput. Chem. Eng.*, **13**, 21 (1989).
- Fotopoulos, J., "Quantification of Model Uncertainty and Its Impact on the Optimization and State Estimation of Batch Processes via Tendency Modeling," PhD Thesis, Lehigh Univ., Bethlehem, PA (1996).
- Fotopoulos, J., C. Georgakis, and H. G. Stenger, "Use of Tendency Models and Their Uncertainty in the Design of State Estimators for Batch Reactors," *Chem. Eng. Process.*, **37**, 545 (1998).

- Fotopoulos, J., C. Georgakis, and H. Stenger, "Structured Target Factor Analysis for the Stoichiometric Modeling of the Batch Reactors," *Proc. Amer. Control. Conf.*, IEEE, New York, Part Vol. 1, 495 (1994).
- Gudi, R. D., S. L. Shah, and M. R. Gray, "Adaptive Multirate State and Parameter Estimation Strategies with Application to a Bioreactor," *AIChE J.*, **41**, 2451 (1995).
- Jang, S. S., B. Joseph, and H. Muka, "Comparison of Two Approaches to Online Parameter and State Estimation of Nonlinear Systems," *Ind. Eng. Chem., Process Des. Dev.*, **25**, 809 (1986).
- Jazwinski, A. H., *Stochastic Processes and Filtering Theory*, Academic Press, New York (1970).
- Jo, J. H., and S. G. Bankoff, "Digital Monitoring and Estimation of Polymerization Reactors," *AIChE J.*, **22**, 361 (1976).
- Kleijnen, J. P. C., *Statistical Techniques in Simulation*, Dekker, New York (1974).
- Kozub, D., and J. F. Macgregor, "State Estimation for Semi-Batch Polymerization Reactors," *Chem. Eng. Sci.*, **47**, 1047 (1992).
- Ku, W., R. H. Storer, and C. Georgakis, "Uses of State Estimation for Statistical Process Control," *Comput. Chem. Eng.*, **18**, S571 (1994).
- Lee, J. H., and N. L. Ricker, "Extended Kalman Filter Based Nonlinear Model Predictive Control," *Ind. Eng. Chem. Res.*, **33**, 1530 (1994).
- Leis, R. J., and M. Kramer, "The Simultaneous Solution and Sensitivity Analysis of Systems Described by Ordinary Differential Equations," *ACM Trans. Math. Softw.*, **14**(1), 45 (1988).
- Lewis, F. L., *Optimal Estimation with an Introduction to Stochastic Control Theory*, Wiley, New York (1986).
- Liotta, V., "Control of Relative Particle Growth in Emulsion Polymerization," PhD Thesis, Lehigh Univ., Bethlehem, PA (1996).
- Liotta, V., C. Georgakis, and M. S. El-Aasser, "Real-time Estimation and Control of Particle Size in Semi-Batch Emulsion Polymerization," *Proc. 1997 Amer. Control Conf.*, American Autom. Control Council, Evanston, IL, Part Vol. 2, 1172 (1997).
- Maybeck, P. S., *Stochastic Models, Estimation and Control*, Academic Press, New York (1982).
- Mehra, R. K., "Approaches to Adaptive Filtering," *IEEE Trans. Automat. Contr.*, **AC-17**, 693 (1972).
- Myers, K. E., and B. D. Tapley, "Adaptive Sequential Estimation with Unknown Noise Statistics," *IEEE Trans. Automat. Contr.*, **AC-21**, 520 (1976).
- Rastogi, A., C. Georgakis, and H. G. Stenger, "The Identification of Kinetic Expressions and the Evolutionary Optimization of Specialty Chemical Batch Reactors, Using Tendency Models," *Chem. Eng. Sci.*, **47**(9–11), 2487 (1992).
- Ravindranath, K., and R. A. Mashelkar, "Modeling of Poly(ethylene Terephthalate) Reactors: 1 A Semibatch Ester Interchange Reactor," *J. Appl. Polym. Sci.*, **26**, 3179 (1981).
- Scali, C., M. Morretta, and D. Semino, "Control of the Quality of Polymer Products in Continuous Reactors: Comparison of Performance of State Estimators with and without Updating of Parameters," *J. Process Control*, **7**(5), 357 (1997).
- Schmidt, A. D., and W. H. Ray, "The Dynamic Behavior of Continuous Polymerization Reactors: 1," *Chem. Eng. Sci.*, **36**, 1401 (1981).
- Soroush, M., "Nonlinear State-Observer Design with Application to Reactors," *Chem. Eng. Sci.*, **52**, 387 (1997).
- Tatiraju, S., and M. Soroush, "Nonlinear State Estimation in a Polymerization Reactor," *Ind. Eng. Chem. Res.*, **36**, 2679 (1997).
- Terwiesch, P., and M. Agarwal, "A Discretized Nonlinear State Estimator for Batch Processes," *Comput. Chem. Eng.*, **19**, 155 (1995).
- Terwiesch, P., M. Agarwal, and W. T. Rippin, "Batch Unit Optimization with Imperfect Modeling: A Survey," *J. Process Control*, **4**, 238 (1994).
- Wilson, D. I., M. Agarwal, and D. W. T. Rippin, "Experiences in Implementing the Extended Kalman Filter in an Industrial Batch Reactor," *Comput. Chem. Eng.*, **22**, 1653 (1998).
- Zhou, J., and R. Luecke, "Estimation of the Covariances of the Process Noise and Measurement Noise for a Linear Discrete Dynamical System," *Comput. Chem. Eng.*, **19**, 187 (1995).

Manuscript received May 17, 1999.