STATE AND PARAMETER ESTIMATION THROUGH DYNAMIC BAYESIAN FORECASTING

Zhen Lu, Elaine Martin and Julian Morris

Centre for Process Analytics and Control Technology
School of Chemical Engineering and Advanced Materials,
University of Newcastle, Newcastle upon Tyne, NE1 7RU, England, UK

Abstract: Successful application of model based control depends on having good estimates for the system dynamic states and parameters. A multivariate dynamic linear model is developed for the estimation of the states from limited measurements in a non-linear system comprising model uncertainties. Since the noise statistics are rarely available *a priori*, the noise covariance matrix is treated as a tuning parameter and determined through repeated simulations. For non-linear, time varying processes, the assumption of a constant process noise covariance matrix does not realise accurate estimates. In this paper Monte Carlo simulations are used to obtain the time-varying noise covariance matrix. The methodology is demonstrated on a benchmark polymerisation process.

Keywords: State estimation; Parameter estimation; Dynamic linear model; Monte Carlo simulation; Extended Kalman filter

1. INTRODUCTION

In batch polymerisation processes, the operating objectives require the satisfaction of complex property requirements for the final polymer whilst reducing production costs. Most mechanical and rheological properties of polymer products are directly, or indirectly, linked to the molecular structural properties of the polymer chains that are not usually measured on-line. Average polymer molecular weight properties (e.g. number and weight average molecular weight), and particle size distribution which can be inferred from on-line measurements, are often selected as the major controlled variables that need to be maintained within well-determined limits so that the desired product quality criteria can be achieved.

Recursive stochastic state estimation techniques, such as the Extended Kalman filter have been traditionally used for state and parameter estimation especially for polymerization processes. The main bottleneck in the application of recursive stochastic state estimation techniques to real world situations is that the process noise statistics are rarely available *a priori*. In most applications, they serve as tuning parameters and are determined through a trial-and error procedure using repeated simulations.

Few techniques for determining the process noise covariance matrix have been developed for chemical engineering applications. Zhou and Luecke (1995) used maximum-likelihood estimation with linear regression to obtain the diagonal elements of the covariance matrices for linear systems. For non-linear systems, the use of innovation processes for estimating the noise statistics was proposed by Myers and Tapley (1976). However they assumed the covariance matrix to be constant. For batch processes with time-varying process dynamics that operate over a range of process conditions, this is not the case. The specification of a constant process noise covariance matrix may not be sufficient to provide sufficiently accurate estimation. Using a fixed value of noise statistic can lead to poor estimation or potentially result in filter divergence.

Valappil and Georgakis (2000) introduced two approaches to systematically estimate the process noise covariance matrix. The first method was based on Taylor series expansion whereas the second method used Monte Carlo simulations to calculate the time-varying values of the process noise

covariance matrix on-line. Both methods require information about the plant-model mismatch in the form of a parameter covariance matrix. The process noise covariance matrix is obtained from the parameter covariance matrix. If the user is not certain about the process-model mismatch, or the model uncertainty can not be represented by the parameter covariance matrix, it is difficult to apply these methods.

.

In this paper, a new approach is proposed where a multivariate dynamic model is constructed for the estimation of the states. Monte Carlo simulations are then used to calculate the time-varying process noise covariance matrix on-line from the prediction errors.

2. BAYESIAN DYNAMIC MODEL

2.1 Multivariate Dynamic Linear Model

The dynamic linear model (DLM) is a Bayesian forecasting tool based on a state space model that allows a variety of adaptive linear and generalised linear models to be fitted iteratively to univariate or multivariate time series data. A DLM incorporates information from any relevant source, including subjective expert views, leading to amended and updated model structures. The general multivariate dynamic linear model developed by West and Harrison (1997), is given by the following system of equations:

Observation equation:

$$\mathbf{Y}_t = \mathbf{F}_t \mathbf{x}_t + \mathbf{v}_t \qquad \mathbf{v}_t \sim N[0, \mathbf{V}_t] \tag{1}$$

System equation:

$$\mathbf{x}_t = \mathbf{G}_t \mathbf{x}_{t-1} + \mathbf{w}_t \qquad \mathbf{w}_t \sim N[0, \mathbf{W}_t] \tag{2}$$

Initial information:

$$(\mathbf{x}_0 \mid D_0) \sim N[\mathbf{m}_0, \mathbf{C}_0] \tag{3}$$

where $\mathbf{Y}_t$ is the observed vector of the series at time point t, $\mathbf{x}_t$ is the state vector, $\mathbf{v}_t$ is the observational error, $\mathbf{w}_t$ is the vector of process noise that is assumed to be independent and normally distributed and $\mathbf{D}_0$ is the initial prior information at $t = 0$. At any future time point, t, the available information set is:

$$D_t = \{\mathbf{Y}_t, D_{t-1}\} \tag{4}$$

At time t-1, for some mean $\mathbf{m}_{t-1}$ and variance matrix $\mathbf{C}_{t-1}$, the posterior is given by:

$$(\mathbf{x}_{t-1} \mid D_{t-1}) \sim N[\mathbf{m}_{t-1}, \mathbf{C}_{t-1}] \tag{5}$$

and the prior for the state vector at time t can be derived from the system equation:

$$(\mathbf{x}_t \mid D_{t-1}) \sim N[\mathbf{a}_t, \mathbf{R}_t] \tag{6}$$

where

$$\mathbf{a}_t = \mathbf{G}_t \mathbf{m}_{t-1} \quad \text{and} \quad \mathbf{R}_t = \mathbf{G}_t \mathbf{C}_{t-1} \mathbf{G}_t' + \mathbf{W}_t \tag{7}$$

According to the observation equation, the one step ahead forecast can be given by:

$$(\mathbf{Y}_t \mid D_{t-1}) \sim N[\mathbf{f}_t, \mathbf{Q}_t] \tag{8}$$

where

$$\mathbf{f}_t = \mathbf{F}_t' \mathbf{a}_t \quad \text{and} \quad \mathbf{Q}_t = \mathbf{F}_t' \mathbf{R}_t \mathbf{F}_t + \mathbf{V}_t \tag{9}$$

The feedback of information obtained at time t from vector $\mathbf{x}_t$ is achieved through the application of linear Bayes methods. As stated previously, the model at time (t-1) requires only the mean vector and covariance matrix of the posterior for $(\mathbf{x}_{t-1} \mid D_{t-1})$. Thus at time t, the corresponding moments:

$$(\mathbf{x}_t \mid D_t) \sim N[\mathbf{m}_t, \mathbf{C}_t] \tag{10}$$

are required in order to progress to time point, (t+1), and subsequent observations. The information obtained at time point, t, is used to update the prior moments to give:

$$\mathbf{m}_t = \mathbf{a}_t + \mathbf{A}_t \mathbf{e}_t \text{ and } \mathbf{C}_t = \mathbf{R}_t - \mathbf{A}_t \mathbf{Q}_t \mathbf{A}_t' \tag{11}$$

where

$$\mathbf{A}_t = \mathbf{R}_t \mathbf{F}_t' \mathbf{Q}_t^{-1} \quad \text{and} \quad \mathbf{e}_t = \mathbf{Y}_t - \mathbf{f}_t \tag{12}$$

2.2 Multivariate Non-linear Dynamic Model

For a non-linear model, the process can be expressed as:

$$\begin{aligned} \mathbf{Y}_t &= f_t(\mathbf{x}_t) + \mathbf{v}_t \quad \mathbf{v}_t \sim N[0, \mathbf{V}_t] \\ \mathbf{x}_t &= g(\mathbf{x}_{t-1}) + \mathbf{w}_t \quad \mathbf{w}_t \sim N[0, \mathbf{W}_t] \end{aligned} \tag{13}$$

Before the usual DLM updating procedure is applied, the model requires to be linearized. The most straightforward and easily interpreted linearization technique is the Taylor series approximation. Applying Taylor series expansion to the updating function about the mean $\mathbf{m}_{t-1}$:

$$\begin{aligned} \mathbf{g}(\mathbf{x}_t) = \mathbf{g}(\mathbf{m}_{t-1}) + \mathbf{G}_t(\mathbf{x}_{t-1} - \mathbf{m}_{t-1}) \\ + \text{quadratic and higher order terms} \end{aligned} \tag{14}$$

where $\mathbf{G}_t$ is the matrix derivative of the updating function evaluated at the estimate $\mathbf{m}_{t-1}$:

$$\mathbf{G}_t = \left[\frac{\partial g_t(\mathbf{x}_{t-1})}{\partial \mathbf{x}_{t-1}} \right]_{\mathbf{x}_{t-1}=\mathbf{m}_{t-1}} \tag{15}$$

$$\mathbf{a}_t = \mathbf{g}_t(\mathbf{m}_{t-1}) \text{ and } \mathbf{R}_t = \mathbf{G}_t \mathbf{C}_{t-1} \mathbf{G}_t' + \mathbf{W}_t \tag{16}$$

Proceeding to the observation equation, a similar approach applies. The non-linear regression function is linearized about the expected value $\mathbf{a}_t$ for $\mathbf{x}_t$, giving:

$$\mathbf{f}_t(\mathbf{x}_t) = \mathbf{f}_t(\mathbf{a}_t) + \mathbf{F}_t'(\mathbf{x}_t - \mathbf{a}_t) + \text{quadratic and higher order terms} \tag{16}$$

where $\mathbf{F}_t$ is the matrix derivative of $\mathbf{f}_t(.)$ evaluated at the prior mean $\mathbf{a}_t$:

$$\mathbf{F}_t = \left[\frac{\partial f_t(\mathbf{x}_t)}{\partial \mathbf{x}_t} \right]_{\mathbf{x}_t=\mathbf{a}_t} \tag{17}$$

The standard updating equations continue to apply.

Most process models constructed from limited experimental observations involve significant uncertainties. For batch and semi-batch processes, this is especially true. Model accuracy is obtained by tuning the covariance matrix ***Q*** of the process noise using a repeated simulation procedure.

If a covariance matrix whose entries are of small magnitude is selected, greater confidence will be expressed in the model and less on-line measurement information will be required to update the states. However, this may result in degraded estimates and possibly estimator divergence. Too much state compensation can cause the state estimates to be noisy and unreliable.

The process noise **w(t)** is mainly due to uncertainties in the model and can be either parametric or structural. Monte Carlo simulations are utilized to estimate a time-varying covariance matrix ***Q*** on-line.

3. ESTIMATION OF THE COVARIANCE MATRIX

The concept of applying Monte Carlo simulation to estimate the process noise covariance matrix is to capture the effect of uncertainties in the model through the statistics of **w(t)**. The key issue is how to derive the information from the measurements and model evolution to represent the process noise. In the algorithm described above, the prediction error vector $\mathbf{e}_t$ gathers the information about the errors that caused by process disturbance and model uncertainties. The updating information from prior to the posterior for the states is obtained by multiplying the prediction errors $\mathbf{e}_t$ with a gain $\mathbf{A}_t$. The idea in this paper is that, using the information drawn from the prediction errors $\mathbf{e}_t$ and the updating information to estimate the process noise covariance matrix.

At a desired time instance *t*-1, a set of samples $\{\mathbf{x}_{t-1}^k\}$ is randomly selected from the posterior distribution function, equation (5), of the state vector while the observation $\mathbf{Y}_{t-1}$ is given. For the k^{th} Monte Carlo simulation, a non-linear model is used to generate random samples of $\mathbf{x}_{t|t-1}$

$$\mathbf{x}_{t|t-1}^k = \mathbf{g}_t(\mathbf{x}_{t-1}^k) \tag{18}$$

Because $\mathbf{x}_{t-1}^k$ is randomly sampled directly from a probability distribution function, some samples may be located in the tails of the distribution. For non-linear systems, the presence of such samples will seriously affect the performance of the estimate. One solution is to reject these samples. The simplest way is to set a rejection bound. Those samples that have a probability larger than the rejection bound will be accepted, otherwise they are discarded and new samples generated. Once the new measurement $\mathbf{Y}_t$ is obtained, the prediction errors are calculated

$$\mathbf{e}_t^k = \mathbf{Y}_t - \mathbf{f}_t(\mathbf{x}_{t|t-1}^k) \tag{19}$$

and the information used to update the prior moments is calculated as follow:

$$\mathbf{?}_t^k = \mathbf{A}_t \mathbf{e}_t^k \tag{20}$$

The process noise is also obtained from the samples of the updating information:

$$\mathbf{w}_t^k = \mathbf{?}_t^k - \overline{\mathbf{?}}_t \tag{21}$$

where $\overline{\mathbf{?}}_t$ denotes the mean of the updating information. The process noise is normally distributed with zero mean. The process noise covariance matrix $\mathbf{Q}_t$ can be calculated from $\mathbf{w}_t^k$, and is a non-diagonal and time-varying matrix. Because the process measurements are available at discrete time instances, the preceding calculation of $\mathbf{w}_t^k$ and $\mathbf{Q}_t$ is performed for discrete time intervals.

For the model development described in the previous section, the process noise is assumed to be a white, Gaussian random process. Thus the approximation of normally distributed process noise needs to be tested using the values of the process noise data set $\{\mathbf{w}_t^k\}$ that were obtained from the Monte Carlo simulations.

For this, normal probability plots were used and the distribution was observed to be approximately normal. For the simulation case presented, 500 Monte Carlo simulations of the different state values were used, resulting in 500 evaluations of the process noise for each state.

4. RESULTS

4.1 MMA polymerization

The process studied is the free radical polymerisation reactor of methyl-methacrylate (MMA) (Mourikas *et al*, 2001). A mathematical model describes the dynamic behaviour of an experimental pilot scale system (Fig 1). Heating and cooling of the reaction mixture is achieved by controlling the flows of hot and cold water stream, through the reactor jacket. The polymerisation temperature is controlled by a cascade control system consisting of a primary PID and two secondary PI controllers. The polymerisation is highly exothermic and exhibits a strong acceleration in polymerisation rate due to gel-effects. Batch duration is 120 minutes.

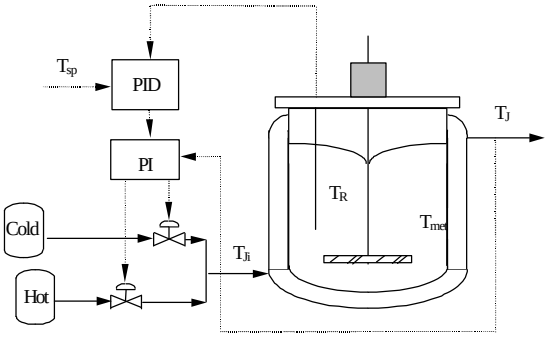


Fig. 1. Plant polymerisation reactor

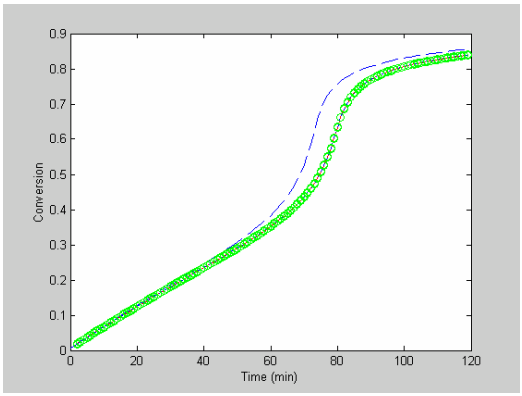
The MMA system consists of 11 states, which are monomer conversion; three moments of dead polymer that are used to calculate the molecular weight distribution; reactor and metal wall temperatures; and four jacket zone temperatures. In polymerization, frequent measurements of the reactor and the jacket inlet and outlet temperatures are usually available along with possibly jacket flow. Monomer conversion measurements can also be obtained from an on-line densitometer. The measurements in a real process environment will be corrupted with measurement noise. In the following simulation, Gaussian white noise is added to the measurements.

In this particular study, the process model mismatch is introduced in the form of time variation in a kinetic parameter. In practice important kinetic parameters such as k_p, the propagation rate constant, cannot be determined accurately and may vary during the polymerisation. In this study, the propagation rate constant is represented by $k_p = k_{p0} g_p g^s_{p,corr}$, where k_{p0} is an intrinsic chemical rate constant, g_p is a diffusion controlled function which includes a number of parameters that are often unknown. The stochastic correction term $g^s_{p,corr}$ is used to account for the imprecise knowledge of g_p. In the model, a random walk is assumed for the behaviour of the stochastic state. In the process, the actual value of $g^s_{p,corr}$ is assumed to decrease linearly from an initial value of 1.0 to 0.76.

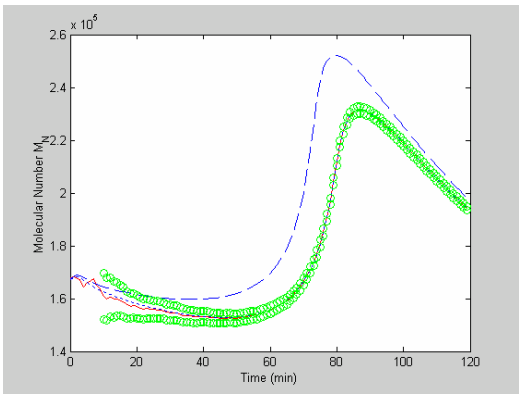
4.2 Discussion

The results of the estimation studies for the MMA polymerization reactor are shown in Fig. 2. It can be seen that the uncorrected model (dashed line), in which the stochastic correction term of the propagation rate constant is fixed at 1.0, differs significantly from the actual plant (dotted line), in which the propagation rate constant is time-varying. The estimates (solid line) closely match the actual process. These, results are compared with the estimation results from an EKF with fixed process noise covariance matrix, Fig. 3. The estimates of number average and weight average molecular weights in Fig. 2. match the actual process while the estimates in Fig. 3. show a discrepancy between the actual process and the estimates.

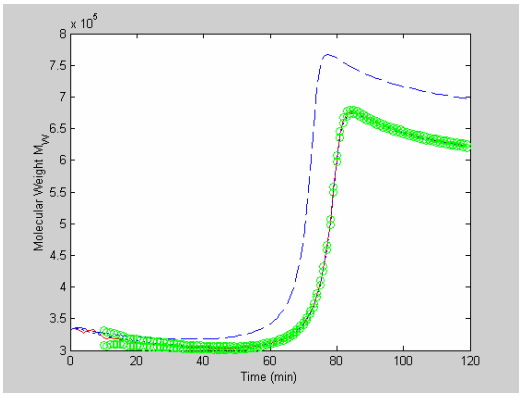
Comparing Fig. 2(e) and Fig. 3(e), the estimator involving Monte Carlo simulation tracks the decrease in the rate constant more closely and faster. This results in better state estimation performance. The 95% confidence bound (circle) of the estimate for the Bayesian approach is narrower than for the EKF, indicating that the estimates are more accurate and reliable. Since only a limited number of observation data can be used to update the estimate, the confidence bounds at the beginning are wide and hence the estimates are less reliable. As the observation data increases, the limits decrease in magnitude. Thus the structure of the confidence bounds for parameter estimation by the Bayesian approach are more reasonable than those of the EKF.



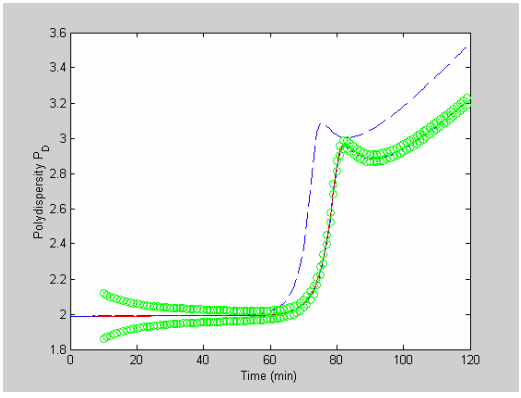
(a) Actual versus estimated (conversion)



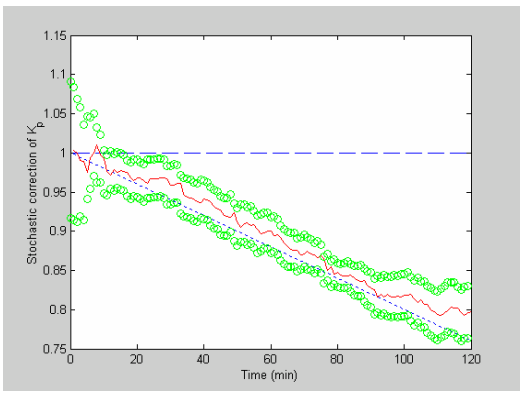
(b) Actual versus estimated (M_N)



(c) Actual versus estimated (M_W)



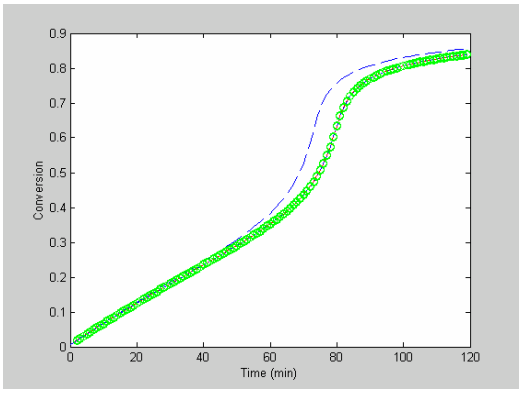
(d) Actual versus estimated (polydispersity)



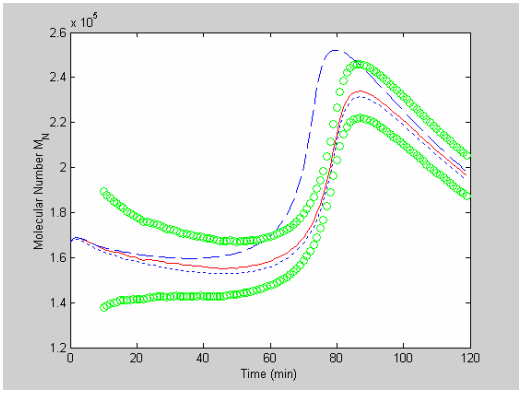
(e) Actual versus estimated (rate constant K_P)

Fig.2. Multivariate DLM state and parameter estimation.

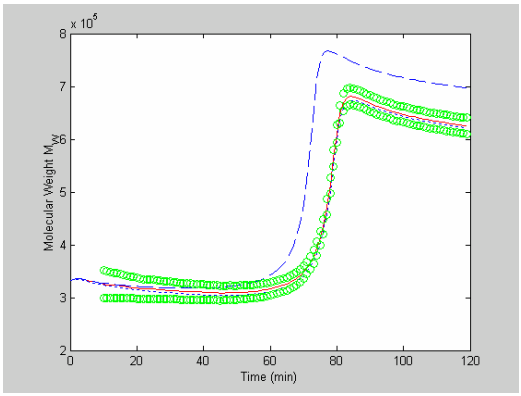
Key: Dotted line - actual plant; solid line - estimates; dashed line - process model with mismatch; circle - 95% confidence bounds



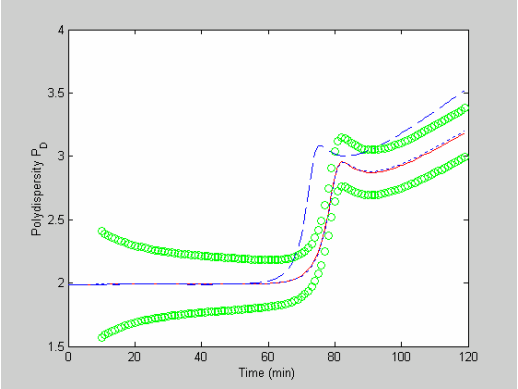
(a) Actual versus estimated (conversion)



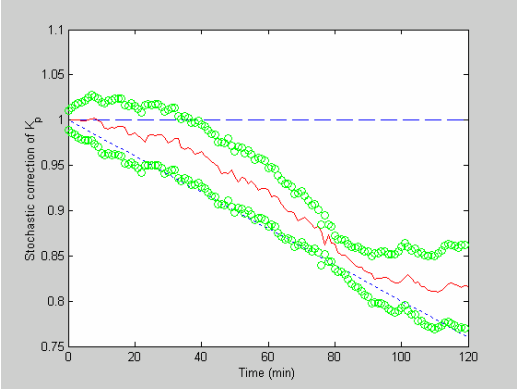
(b) Actual versus estimated (M_N)



(c) Actual versus estimated (M_W)



(d) Actual versus estimated (polydispersity)



(e) Actual versus estimated (rate constant K_P)

Fig.3. EKF state and parameter estimation with fixed process noise covariance matrix.

Key: Dotted line - actual plant; solid line - estimates; dashed line - process model with mismatch; circle - 95% confidence bounds

5. CONCLUSIONS

The feasibility of extending multivariate DLM to estimate the process states, monomer conversion and the molecular weights in MMA batch polymerizations has been demonstrated. The methodology provides a new approach to state estimation for possible application in on-line model-based optimising control. The estimator uses on-line measurements of key process variables including monomer conversion and reactor and jacket temperatures, to provide reliable estimates of the state variables. Monte Carlo simulation is used for the calculation of the process noise covariance matrix.

The results show that the approach presented can improve the performance of the state and parameter estimation. It also makes the design and the application of dynamic Bayesian forecasting more robust, since the methodology proposed eliminates the need for the tuning of the process noise covariance matrix. The non-diagonal and time-varying covariance matrix is obtained on-line in contrast to a diagonal and constant covariance matrix which is not able to adapt to non-linear systems with model uncertain. The algorithm can also be enhanced by Bayesian parameter estimation to provide a significantly enhanced overall state and parameter estimation methodology (Lu *et al*, 2001).

6. AKNOWLEDGMENTS

Zhen Lu acknowledges CPACT and the University of Newcastle for the financial support of his PhD and Prof. Kiparissides, CPERI, Thessaloniki, for the MMA reactor simulation.

7. REFERENCES

Crowley, T.J. and K.Y. Choi (1997). Discrete optimal control of molecular weight distribution in a batch free radical polymerization process. *Ind. Eng. Chem.*, **36**, pp. 3676-3684.

Crowley, T.J. and K.Y. Choi (1998). Experimental studies on optimal molecular weight distribution control in batch free radical polymerization process. *Chem. Eng. Sci.*, **53**, pp. 2769-2790.

Lu, Z., E.B. Martin and A.J. Morris (2002). Bayesian parameter estimation in batch polymerisation. *ESCAPE-12, European Symposium on Computer Aided Process Engineering-12*, 517-522.

Myers, K.E. and B.D. Tapley (1976). Adaptive sequential estimation with unknown noise statistics. *IEEE Trans. Automat. Contr.*, AC**21**, pp. 520-523.

Ruppen, D., D. Bonvin and D.W.T. Rippin (1997). Implementation of adaptive optimal operation for a semi-batch reaction system. *Comput. Chem. Eng.*, **22**, pp. 185-199.

Thomas, I. M., and C., Kiparissides (1984). Computation of the near-optimal policies for a batch polymerization reactor. *Can. J. Chem. Eng.*, **62**, pp. 284-291.

Valappil, J. and C. Georgakis (2000). Systematic estimation of state noise statistics for extended Kalman filters. *AIChE Journal.*, **46(2)**, pp. 292-308.

West, M., and P.J. Harrison (1997). *Bayesian forecasting and dynamic models.* Springer-Verlag, New York.

Zhou, J. and R. Luecke (1995). Estimation of the covariance of the process noise and measurement noise for a linear discrete dynamical system. *Comput. Chen. Eng.*, **19**, pp. 187-195.