PARAMETER ESTIMATION FOR BATCH PROCESSES USING A BAYESIAN APPROACH

Zhen Lu, Julian Morris and Elaine B. Martin

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

Abstract: In a changing operational environment, a major challenge that still exits is the assured state and parameter estimation of dynamic processes. The values or expressions of important parameters can be difficult to determine and initial errors may be present in some parameters as a result of changes in the initial operating conditions. Furthermore as a result of variations in the environmental and operational conditions or the dynamic characteristics of the process, parameters may be time-varying. In this paper, an on-line Bayesian parameter estimator is developed and evaluated on a simulation of a batch methyl methacrylate process.

Keywords: Bayesian Analysis; Parameter estimation; Polymerization.

1. INTRODUCTION

The operating objectives in many batch polymerisation processes must satisfy complex property requirements in terms of the final polymer whilst simultaneously achieving the best operational performance of the batch. Most properties of the polymer products are directly or indirectly linked with the molecular structural properties of polymer chains (e.g. molecular weight distribution, copolymer composition distribution, chain sequence length distribution, etc.), which are difficult (sometimes impossible) to measure on-line. Average polymer molecular weight properties (e.g. number and weight average molecular weight), which can be indirectly inferred from the on-line measurement of the solution viscosity or melt index of the polymer, are commonly selected as the control variables that need to be maintained within well-determined limits so that product quality criteria can be satisfied.

Control strategies require that pre-determined trajectories for key process variables (e.g. reactor temperature) are followed during batch operation (Thomas and Kiparissides, 1984). However, the operational performance of a batch polymerisation reactor is affected by process disturbances potentially resulting in changing operating conditions. A consequence of this is the introduction of process-model mismatch and hence the need for time-varying model parameter values and the regular updating of the optimal control trajectories during batch operation, otherwise the control strategy may fail to meet the product quality specifications and the operating requirements (Kiparissides, *et al,* 2002).

Despite the large number of papers discussing the control of polymer properties using an Extended Kalman Filter (EKF), very few tackle the issue of robust parameter updating (Gagnon and MacGregor, 1991; Kozub and MacGregor, 1992). MacGregor and co-workers (1991, 1992) analysed the parameter updating issue and emphasised the need to update as many parameters as possible in order to incorporate all possible changes that could affect the process.

The general problem of parameter estimation is associated with the fitting of a model to a set of measurements. Given a model with some unknown parameters, a parameter estimator calculates those

parameter values that result in the model predicted values of the process outputs being closest to the corresponding measured values of the process outputs. In off-line parameter estimation, a model is fitted optimally to the process measurements from one or more completed process runs. In contrast, in on-line parameter estimation, a model is fitted to the past and present process measurements whilst the process is in operation (Narendra and Annaswamy; 1988, Bastin and Dochain, 1986; Yin, 1993).

Schuler and Schmidt (1992) discussed the effect of unmeasured and time varying parameters such as reactive heat flow and heat transfer coefficients, and presented a parameter adaptive EKF with additional stochastic states to approximate these parameters. Scali, *et al,* (1997) applied a parameter adaptive EKF to a methyl methacrylate (MMA) polymerisation in a continuous solution CSTR reactor. Sirohi and Choi (1996) applied the EKF and an optimisation based non-linear dynamic parameter estimator to a continuous olefin polymerisation reactor. They combined the two algorithms to estimate the uncertain kinetic parameters on-line.

Parameter estimation through on-line optimisation has also been widely applied. For this approach, parameter estimates are obtained by solving on-line a minimisation problem such as the sum of squared errors (Muske and Rawlings, 1995; Robertson, *et al*, 1996). As in state estimation, this parameter estimator can either have an increasing or constant horizon (Muske and Rawlings, 1995). Although this methodology is computationally expensive, it allows the inclusion of constraints in the estimation. With the advancement of computers in terms of speed, the computational problem will soon cease to be an issue.

An alternative approach to implementing parameter estimation is to make use of Bayesian technologies. Here prior information about the parameters is combined with the process observations.

2. BAYESIAN PARAMETER ESTIMATION

2.1 Basic Framework of Bayesian Parameter Estimation

Consider a non-linear process. It is assumed that the process has an $(n\times 1)$ state vector, $\mathbf{x}(t)$, that can be described by the continuous time process model:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \boldsymbol{\theta}, t) \qquad (1)$$
$$\mathbf{x}(0) = \mathbf{x}_0$$

where the measurement model takes the form:

$$\mathbf{y}(t) = \mathbf{h}(\mathbf{x}(t), \boldsymbol{\theta}, t) + \mathbf{v}_t \qquad (2)$$

$\mathbf{f}(.)$ and $\mathbf{h}(.)$ are non-linear system and observation functions respectively. $\mathbf{y}(t)$ is an $(m\times 1)$ vector of observations over the time period $t = 1,2,\ldots$. For convenience, the measurement information $\mathbf{y}(t)$ is represented by $\mathbf{y}_t$. Let $\mathbf{D}_t$ be the information set available prior to time t. Thus given the initial prior information $\mathbf{D}_0$ at time point $t = 0$, at any future time t, the available information set is:

$$\mathbf{D}_t = \{\mathbf{y}_t, \mathbf{D}_{t-1}\} \qquad (3)$$

$\boldsymbol{\theta}$ is a vector of unknown parameters about which there may be some prior beliefs that can be expressed as a probability density function, $p(\boldsymbol{\theta})$. This prior distribution can be defined using either previously monitored data, or an expert's knowledge. $\mathbf{v}_t$ is the vector of observational error, and is defined as a normal random variable with zero mean:

$$\mathbf{v}_t \sim N[0, \mathbf{V}_t] \qquad (4)$$

Equations (1) and (2) define the model relating $\mathbf{y}_t$ to the parameter vector $\boldsymbol{\theta}_t$ at time t, and has a probability distribution that is a function of the unknown parameters. The dependence of $\mathbf{y}_t$ on $\boldsymbol{\theta}_t$ can be expressed as the conditional probability density function $p(\mathbf{y}_t \mid \boldsymbol{\theta}_t)$:

$$p(\mathbf{y}_t \mid \boldsymbol{\theta}_t) = p(\mathbf{h}(\mathbf{x}_t, \boldsymbol{\theta}_t) + \mathbf{v}_t \mid \boldsymbol{\theta}_t) \qquad (5)$$

Once $\boldsymbol{\theta}_t$ has been defined, the following characteristics can be obtained:

$$\begin{aligned} &E(\mathbf{y}_t \mid \boldsymbol{\theta}_t) = E(\mathbf{h}(\mathbf{x}_t, \boldsymbol{\theta}_t) \mid \boldsymbol{\theta}_t) + E(\mathbf{v}_t \mid \boldsymbol{\theta}_t) \\ &= \mathbf{h}(\mathbf{x}_t, \boldsymbol{\theta}_t) \\ &\operatorname{cov}(\mathbf{y}_t \mid \boldsymbol{\theta}_t) \\ &= \operatorname{cov}(\mathbf{h}(\mathbf{x}_t, \boldsymbol{\theta}_t) \mid \boldsymbol{\theta}_t) + \operatorname{cov}(\mathbf{v}_t \mid \boldsymbol{\theta}_t) = \mathbf{V}_t \\ &p(\mathbf{y}_t \mid \boldsymbol{\theta}_t) \propto \\ &\exp\{-(\mathbf{y}_t - \mathbf{h}(\mathbf{x}_t, \boldsymbol{\theta}_t))' \mathbf{V}_t^{-1} (\mathbf{y}_t - \mathbf{h}(\mathbf{x}_t, \boldsymbol{\theta}_t)) / 2) \end{aligned} \qquad (6)$$

Thus the observation equation can be expressed as a conditional distribution:

$$p(\mathbf{y}_t \mid \boldsymbol{\theta}_t) \sim N(\mathbf{h}(\mathbf{x}_t, \boldsymbol{\theta}_t), \mathbf{V}_t) \qquad (7)$$

To update the probability density of the unknown $\boldsymbol{\theta}_t$ after new process data has been obtained, Bayes theorem is applied:

$$p(\boldsymbol{\theta}_t \mid \mathbf{y}_t) = \frac{p(\mathbf{y}_t \mid \boldsymbol{\theta}_t) p(\boldsymbol{\theta}_t)}{p(\mathbf{y}_t)} \qquad (8)$$

The denominator, $p(\mathbf{y}_t)$ is the unconditional probability density of the observation data and from the law of total probability:

$$p(\mathbf{y}_t) = \int p(\mathbf{y}_t \mid \boldsymbol{\theta}_t) p(\boldsymbol{\theta}_t) d\boldsymbol{\theta}_t = c \tag{9}$$

This value acts as a normalising constant. To construct the Bayesian parameter estimator, the posterior density is expressed as:

$$p(\boldsymbol{\theta}_t \mid \mathbf{y}_t) = c^{-1} e^{\log p(\mathbf{y}_t \mid \boldsymbol{\theta}_t) + \log p(\boldsymbol{\theta}_t)} \tag{10}$$

A point estimate of the unknown parameters $\boldsymbol{\theta}_t$, $\hat{\boldsymbol{\theta}}_t$ is the value that maximises the a-posteriori density. That is, $\hat{\boldsymbol{\theta}}_t$ maximises the probability that the estimate is correct. To maximise the posterior probability density $p(\boldsymbol{\theta}_t \mid \mathbf{y}_t)$, the function $J(\boldsymbol{\theta}_t \mid \mathbf{y}_t)$ is minimised:

$$J(\boldsymbol{\theta}_t \mid \mathbf{y}_t) = -\log p(\mathbf{y}_t \mid \boldsymbol{\theta}_t) - \log p(\boldsymbol{\theta}_t) \tag{11}$$

This function is associated with the new measured process data and the prior information. The incorporation of measurement data and prior information provides a weighting based on knowledge of the statistical error. Consequently the parameters can be identified from measurements and constrained by the prior density. Thus, a working expression for $J(\boldsymbol{\theta}_t \mid \mathbf{y}_t)$ can be derived from equation (11):

$$J(\boldsymbol{\theta}_t \mid \mathbf{y}_t) = \tag{12}$$
$$\frac{1}{2}[\mathbf{y}_t - \mathbf{h}(\mathbf{x}_t, \boldsymbol{\theta}_t)]^T \mathbf{V}_t^{-1} [\mathbf{y}_t - \mathbf{h}(\mathbf{x}_t, \boldsymbol{\theta}_t)]$$
$$+ \frac{1}{2}\log(2\pi \mid \mathbf{V} \mid) - \log p(\boldsymbol{\theta}_t)$$

and the Bayesian parameter estimation problem is defined as a constrained minimization problem:

$$\min_{\boldsymbol{\theta}_t} J(\boldsymbol{\theta}_t \mid \mathbf{y}_t) \tag{13}$$
$$\text{s.t.} \quad \boldsymbol{\theta}_t \in \Omega_{\boldsymbol{\theta}}$$

where the constraints $\Omega_{\boldsymbol{\theta}}$, in the simplest case, is a set determined by the lower and upper bounds of the elements of $\boldsymbol{\theta}_t$ and is solved using SQP.

2.2 Bayesian Parameter Estimation for Dynamic Processes

Uncertain or unknown parameters can result in model-plant mismatch that can lead to the failure of model based control and optimization strategies. For some of these parameters, it is difficult to determine the accurate value since operational conditions are constantly changing. Others may be time-varying and thus need to be estimated on-line at every sampling interval.

Performing discretisation of the dynamic model with respect to the state variables $\mathbf{x}_t$ gives:

$$\mathbf{x}_t = \mathbf{f}_t(\mathbf{x}_{t-1}, \boldsymbol{\theta}_t) + \mathbf{w}_t \quad \mathbf{w}_t \sim N[0, \mathbf{W}] \tag{14}$$
$$\mathbf{y}_t = \mathbf{h}_t(\mathbf{x}_t, \boldsymbol{\theta}_t) + \mathbf{v}_t \quad \mathbf{v}_t \sim N[0, \mathbf{V}]$$

The error sequences $\mathbf{v}_t$ and $\mathbf{w}_t$ are zero mean, white Gaussian random sequences with:

$$E[\mathbf{v}(t)\mathbf{w}^T(t)] = 0 \tag{15}$$
$$E[\mathbf{v}(t)\mathbf{v}^T(t+\tau)] = 0;$$
$$E[\mathbf{w}(t)\mathbf{w}^T(t+\tau)] = 0$$

2.3 Estimation for Non-time-varying Parameters

For those parameters that are not known accurately and that are not time-varying, such as the initiator concentration, or those parameters that change in magnitude only after a change in operational conditions occur, updated parameter estimates at every time instance is not necessary. After the process starts-up or a change in operational condition occurs, parameter estimation is necessary once sufficient observational data become available.

A Bayesian parameter estimator for non-time-varying parameters can be derived and has the following objective function:

$$J(\boldsymbol{\theta} \mid \sum_{k=1}^{n} \mathbf{y}_k) =$$
$$\frac{1}{2}[\sum_{k=1}^{n}(\mathbf{y}_k - \mathbf{h}_k(\mathbf{x}_k, \boldsymbol{\theta}_k))]^T (n\mathbf{V})^{-1} [\sum_{k=1}^{n}(\mathbf{y}_k - \mathbf{h}_k(\mathbf{x}_k, \boldsymbol{\theta}_k))]$$
$$+ \frac{1}{2}\log(2\pi \mid n\mathbf{V} \mid) - \log p(\boldsymbol{\theta}) \tag{16}$$

2.4 On-line Estimation for Time-varying Parameters

Time-varying parameters need to be estimated on-line to minimise model-plant mismatch and ensure acceptable performance of model based control and optimization strategies. The estimator structure derived in equation (16) is not suitable for estimating time-varying parameters. For time-varying parameters, at any time instant, *t*, the unknown parameters are estimated by:

$$\hat{\boldsymbol{\theta}}_t = \min J_t(\boldsymbol{\theta}_t \mid \mathbf{y}_t) \tag{17}$$
$$= \frac{1}{2}[\mathbf{y}_t - \mathbf{h}_t(\mathbf{x}_t, \boldsymbol{\theta}_t)]^T \mathbf{V}^{-1} [\mathbf{y}_t - \mathbf{h}_t(\mathbf{x}_t, \boldsymbol{\theta}_t)]$$
$$+ \frac{1}{2}\log(2\pi \mid \mathbf{V} \mid) - \log p(\boldsymbol{\theta}_t)$$

The Bayesian updating process is an iterative procedure. The prior distribution represents the analyst's state of knowledge prior to the last set of

evidence, not prior to all evidence. In many cases therefore, when new observational data are obtained, the prior distributions are updated using Bayes theorem thereby becoming posterior distributions. The informative prior distribution is simply the posterior distribution from the last updating calculation. That is the prior distribution function $p(\theta_t)$ can be the posterior distribution from the last updating calculation, $p(\theta_{t-1} \mid \mathbf{y}_{t-1})$.

For a Bayesian normal model, the means of the posterior distribution of parameters at time *t-1* are the estimates of the parameters. At any time point *t*, the measurement vector is $\mathbf{y}_t$ and the estimation vector $\hat{\theta}_{t-1}$ is the mean of the distribution $p(\theta_t)$. Hence:

$$p(\theta_t) = p(\theta_{t-1} \mid \mathbf{y}_{t-1}) \text{ and } \bar{\theta}_t = \hat{\theta}_{t-1} \tag{18}$$

Thus the new estimates at time point *t* are $\hat{\theta}_t$.

3. BAYESIAN PARAMETER ESTIMATION FOR A BATCH MMA POLYMERISATION PROCESS

The process studied is a pilot scale polymerisation reactor (Achilias and Kiparissides, 1992). The batch reaction is the free-radical batch polymerisation of methyl-methacrylate (MMA) with a water solvent and benzoyl peroxide initiator.

The jacketed reactor is provided with a stirrer for thorough mixing of the reactants. Heating and cooling of the reaction mixture is achieved by controlling the flows of two water streams (a hot and a cold water stream) at an appropriate temperature 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 reactor temperature is fed back to the primary controller whose output is taken as the set-point of the two secondary controllers. The manipulated variables for the two secondary controllers are hot and cold water flow rates. The hot and cold water streams are mixed before entering the reactor jacket and provide heating or cooling for the reactor. The jacket outlet temperature is fed back to the two secondary controllers.

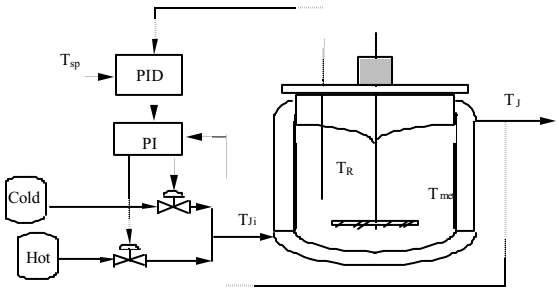

Fig. 1. Schematic of pilot plant polymerization reactor.

A detailed dynamic mathematical model of the reactor covering reaction kinetics, heat and mass balances has been described by Mourikas (1998). This simulation model serves as a realistic test bed for investigating parameter estimation.

The variables that potentially could be used for estimation include: reactor temperature, jacket inlet temperature, jacket outlet temperature, monomer conversion, coolant flow rate and reaction time. In this study it is assumed that monomer conversion and reactor and jacket temperatures are selected as the available measurement set for the estimation procedure. The measurements are corrupted by white Gaussian noise with standard deviations of 0.1%, 0.1K and 0.1K respectively. The sampling rate is 1 minute with batch duration being 120 minutes.

3.1 Erroneous Initial Initiator Concentration and Time-varying Initiator Decomposition Rate

Consider the case where two key parameters for the initiation reaction are incorrectly estimated, initiator concentration and initiator decomposition rate constant. The initiator concentration I is represented by $I = I_0 I_s$ with stochastic correction term I_s, while the initiator decomposition rate constant k_d is represented by $k_d = k_{d0} g^s_{d,corr}$ with the stochastic correction term $g^s_{d,corr}$. I_0 and k_{d0} are computed from kinetic equations (Mourikas, 1998). I_s and $g^s_{d,corr}$ describe the variation associated with the initiator concentration and initiator decomposition rate constant respectively. If there is no model mismatch then the calculated values I_s and $g^s_{d,corr}$ will be equal to the actual values I and k_d, that is the parameters I_s and $g^s_{d,corr}$ will be equal to unity. Uniform prior and normal prior distributions were formulated for the unknow parameters, $\theta = [I_s, g^s_{d,corr}]$:

$$p(\theta_i) \sim U[0.01,10] \qquad (19)$$
$$p(\theta_i) \sim N[1,0.25]; i = 1,2$$

Table 1 Parameter estimation results using different prior distributions and sizes of measurement data set.

Prior Distribution	Data set	Estimated Value	Error $(\hat{\theta} - \theta)/\theta$
$U[0.01,10]$	10	0.725; 0.640	44.9%;28.9%
$N[1,0.25]$	10	0.659; 0.686	31.8%;23.7%
$U[0.01,10]$	70	0.502; 0.896	0.34%;0.43%
$N[1,0.25]$	70	0.543; 0.844	8.5%; 6.17%
$U[0.01,10]$	120	0.499; 0.901	0.26%;0.12%
$N[1,0.25]$	120	0.500; 0.900	0.02%;0.02%

The results in Table 1 illustrate the poor performance of the estimator when only 10 observations are used to calculate the parameter estimates for both the uniform and normal prior distributions. Since two parameters are estimated simultaneously, the complexity of the estimation accordingly increases. The estimation errors for the unknown parameters using a uniform prior distribution decrease significantly (Table 1) when the size of the measurement set is increased to 70. In contrast, the errors of the estimates using the normal prior distributions are larger for the same data set size.
When the size of the measurement data set is increased to 120, the errors for the unknown parameters using both the normal and uniform prior distribution are small thus indicating that both prior distributions are appropriate.

Comparison of the results in this section illustrate that when the prior information about the unknown parameters is described by a uniform prior distribution, the unknown parameters can be estimated accurately using a small number of samples. However for the normal prior distribution, a larger sample size is needed to obtain estimates with a level of accuracy comparable to that attained with the uniform distribution.

3.2 Time-varying Termination Rate Constant

In this second study, process model mismatch is introduced in the form of a time variation in a kinetic parameter. In practice, important kinetic parameters such as k_t, the termination rate constant, cannot be determined accurately and may vary during the polymerisation process. In this study, the termination rate constant is represented by $k_t = k_{t0} g_t g^s_{t,corr}$, where k_{t0} is the rate constant, g_t is a diffusion controlled function, which includes a number of parameters that are often unknown. The stochastic correction term $g^s_{t,corr}$ is used to account for the imprecise knowledge of g_t. In the model, a random walk is assumed for the behaviour of the stochastic state. In the process, the actual values of $g^s_{t,corr}$ are assumed to vary linearly from an initial value of 0.9 to 0.66. The time-varying parameter to be estimated is set as $\theta_t = g^s_{t,corr,t}$

It is assumed that the mean of the parameter, θ_0, is unity and the standard deviation is 0.05. Using maximum entropy, the prior distribution of the unknown parameter is $p_{\theta,0}(\theta_0) \sim N(1,0.05)$. According to the discussion in section 2.3, the prior distribution at time point t is:

$$p_{\theta,k}(\theta_k) \sim N(\mu_k, \sigma_k) \qquad (20)$$
$$\theta_{k=0} = 1; \mu_k = \theta_{k-1}; \sigma_k = 0.05$$

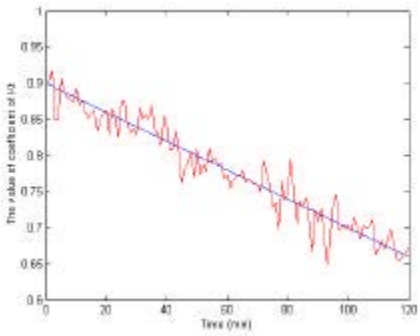

Fig. 2. Estimation of the time-varying correction term for the termination rate constant.

The parameter is assumed to decrease linearly with 10% initial error whilst the parameter in the model is fixed at unity. The parameter estimate is shown in Fig. 2. The Bayesian parameter estimator identifies the starting value of the time-varying parameter and follows the changes closely.

3.3 Time-varying Reactor Heat Transfer Coefficient

A common effect in polymerisation reactors is reactor fouling It is well known that the heat transfer coefficient is strongly correlated with the viscosity of the reacting mixture, which varies during the polymerization. Thus the reactor heat transfer coefficient is an important parameter to monitor. However, it is difficult to identify the overall heat transfer coefficient of the reactor wall.

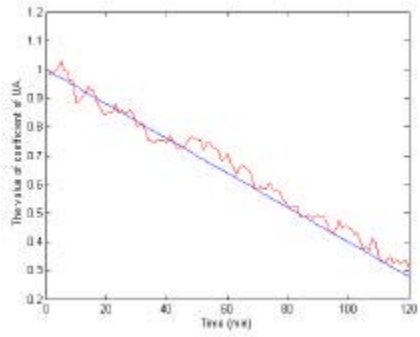

Fig. 3. Estimation of the time-varying correction term for the heat transfer coefficient.

In this study, the heat transfer coefficient is represented by $U = U_0 U_s$, where U_0 represents the experimental value of the heat transfer coefficient. U_s is introduced as a stochastic correction term for the heat transfer coefficient. If there is no reactor fouling, the stochastic correction term is equal to unity. In this study, the heat transfer coefficient is assumed to decrease. It is assumed that, at time point $t = 0$, the mean of the time-varying parameter, θ_0, is unity and the standard deviation is 0.05. The prior distribution of the time-varying parameter is:

$$p_{\theta,k}(\theta_k) \sim N(\mu_k, \sigma_k) \qquad (21)$$
$$\theta_{k=0} = 1; \mu_k = \theta_{k-1}; \sigma_k = 0.05$$

Fig. 3 shows that the Bayesian parameter estimator can track the changes tightly for a time-varying heat transfer coefficient so that the impact of reactor fouling can be identified and the model accordingly updated.

4. CONCLUSIONS

This paper has presented a study on parameter estimation based on Bayesian theory. A basic Bayesian parameter estimation framework is derived from Bayesian Maximum a Posteriori. A Bayesian parameter estimator is constructed for unknown non time-varying parameters, which use historical measurements to update the model before implementing the model for model based optimisation and control. An on-line estimator was then derived to deal with unknown time-varying parameters.

Unlike conventional estimation methods, Bayesian parameter estimation allows expert knowledge and prior beliefs, in the form of a prior probability distribution, to be formally incorporated into the statistical analysis. The prior distribution summarises all available information and expert opinion relating to the parameter of interest before the data have been observed. All knowledge about the underlying parameter is then contained in the posterior probability distribution after observing the data. Thus Bayesian methods use the information to get the best estimate.

The construction of a prior distribution and the data set size affect the performance of the estimator. The impact of different prior distributions for the estimation of the parameters is evaluated by applying the method to a MMA process. It is shown that for a small data set, constructing an appropriate prior distribution is very important for parameter estimation.

The estimation of time-varying parameter was also evaluated. The simulation shows that the Bayesian parameter estimator can estimate the initial error of the parameter and track the time-varying parameter over the duration of the batch.

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. The authors acknowledge the EU Research Training Network, BatchPro, No. HPRN-CT-2000-00039.

REFERENCES

Achillas, D.S. and C. Kiparissides (1994). On the validity of the steady state approximations in high conversion diffusion-controlled free-radical copolymersiations reactions. *Polymer*, **35(8)**, pp. 1714-1721,

Bastin, G. and D. Dochain (1986). On-line estimation of microbial specific growth rates. *Automation*, **22**, pp. 705 – 709.

Gagnon, L. and J.F. MacGregor (1991). State estimation for continuous emulsion polymerization. *Can. J. Chem. Eng.*, **69**, pp. 648-656.

Kozub, D.J. and J.F. MacGregor (1992). State estimation for semi-batch polymerization reactors. *Chem. Eng. Sci.*, **47(5)**, pp. 1047-1061.

Kiparissides, C., G.P. Seferlis, G. Mourikas and A. J. Morris (2002). On-line optimizing control of molecular weight properties in a batch free radical polymerization reactors, *Ind. Eng. Chem. Res.*, **41**, pp. 61620-6131.

Mourikas, G. (1998). *Modelling, estimation and optimisation of polymerisation processes*. PhD Thesis, University of Newcastle, England.

Muske, K.R. and J.B. Rawlings (1995). Non-linear moving horizon state estimation. In: Berber R. (Ed.), *Methods of model based process control.* Netherlands: Kluwer Academic Publisher.

Narendra, K.S. and A.M. Annaswamy (1988). *State adaptive systems.* NJ: Englewood Cliffs, Prentice-Hall.

Robertson, D.G., Lee, J.H. and Rawlings, J.B. (1996). A moving horizon-based approach for least squares state estimation. *AIChE.* **42**, pp. 2209-2224.

Scali, C., M. Morretta and Semino, C. (1997). 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)** 357-369.

Schuler, H. and Ch-U Schmidt (1992). Calorimetric state estimators for chemical reactor diagnosis and control: review of methods and applications. *Chem. Engng Sci.*, **47**, pp. 899-915.

Sirohi, A. and K.Y. Choi (1996). On-line parameter estimation in a continuous polymerization process. *Ind. Engng Chem. Res.*, **35**, 1332-1343.

Thomas, I.M. and C. Kiparissides (1984). Computation of the near-optimal temperature and initiator policies for a batch polymerization reactor. *Can J Chem Eng* **62 (2)**, pp. 284-296.

Yin, K.W. (1994). Parameter estimation and related problems for external input-output models. *Chem. Engng Sci.*, **48**, pp. 3870-3875.