

Approximate Maximum Likelihood Parameter Estimation for Nonlinear Dynamic Models: Application to a Laboratory-Scale Nylon Reactor Model

M. Saeed Varziri, Kim B. McAuley,* and P. James McLellan

Department of Chemical Engineering, Queen's University, Kingston, ON, Canada K7L 3N6

In this article, parameters and states of a laboratory-scale nylon 612 reactor model (Schaffer et al. *Ind. Eng. Chem. Res.* **2003**, *42*, 2946–2959; Zheng et al. *Ind. Eng. Chem. Res.* **2005**, *44*, 2675–2686; and Campbell, D. A. Ph.D. Thesis, Department of Mathematics and Statistics, McGill University, Montreal, Quebec, Canada, 2007) are estimated using a novel approximate maximum likelihood estimation (AMLE) algorithm (Poyton et al. *Comput. Chem. Eng.* **2006**, *30*, 698–708; Varziri et al. *Comput. Chem. Eng.*, published online, <http://dx.doi.org/10.1016/j.compchemeng.2008.04.005>; Varziri et al. *Ind. Eng. Chem. Res.* **2008**, *47*, 380–393; and Varziri et al. *Can. J. Chem. Eng.*, accepted for publication). AMLE is a method for estimating the states and parameters in differential equation models with possible modeling imperfections. The nylon reactor model equations are represented by stochastic differential equations (SDEs) to account for any modeling errors or unknown process disturbances that enter the reactor system during experimental runs. In this article, we demonstrate that AMLE can address difficulties that frequently arise when estimating parameters in nonlinear continuous-time dynamic models of industrial processes. Among these difficulties are different types of measured responses with different levels of measurement noise, measurements taken at irregularly spaced sampling times, unknown initial conditions for some state variables, unmeasured state variables, and unknown disturbances that enter the process and influence its future behavior.

1. Introduction

Parameter estimation in dynamic models that are described by a combination of nonlinear algebraic and differential equations is a challenging problem. The complexity of the problem increases significantly if it is acknowledged that there are two different types of random errors that influence the measurements obtained from dynamic processes: measurement errors and process disturbances. Measurement errors are problematic because they can make it difficult for modelers to obtain reliable parameter estimates, but random process disturbances can be even more problematic because they influence the future behavior of the process and therefore future measurements of process outputs. For example, consider an unknown disturbance that influences the temperature in a chemical reactor. The change in temperature can alter the rates of chemical reactions and can influence several different types of process measurements and how they change over time. Modelers often have knowledge about the quality of the measurements that are available for parameter estimation (e.g., good estimates of measurement variance from repeated measurements or from sensor suppliers), but they do not have a priori knowledge about the quality of their model equations, which are only approximate representations of the true physical process because of disturbances that are not included in the model equations and simplifying assumptions that are made during model development.

Approximate maximum likelihood parameter estimation (AMLE) is a novel parameter estimation algorithm that we recently developed to address the problem of parameter estimation in continuous-time nonlinear dynamic models, in which model discrepancies are significant.^{4–7} A convenient way to account for modeling errors and process disturbances is to include Gaussian noise terms on the right-hand side of the state equations, thereby converting ordinary differential equation models into stochastic differential equation (SDE) models (as

shown in eq 1, below). Until now, the AMLE algorithm has been tested only using simulated data and simple dynamic models of single-phase continuous stirred-tank reactors.^{4–7} The purpose of this article is to examine and demonstrate the application of AMLE to parameter and state estimation for a two-phase laboratory-scale nylon reactor model, originally developed by Schaffer et al.¹ The nylon reactor model has four states, which are described by one algebraic equation and three nonlinear differential equations. Only two of the four states were measured, and these measurements were made at irregular sampling times that were convenient for the experimenters.^{1,2}

We show how AMLE can address frequently encountered parameter estimation difficulties such as working with multi-response models with different levels of measurement accuracy; extracting information from multiple experimental runs with nonuniform sampling times; and dealing with unmeasured states, unknown initial conditions, and unknown levels of modeling error (due to disturbances and structural imperfections). It is also shown that the AMLE framework readily facilitates the integration of additional steady-state or dynamic information that might be available from different sources. Our objective in developing AMLE was to produce a straightforward parameter estimation algorithm that can help modelers to obtain more-reliable parameter estimates and model predictions that can be used in nonlinear model-based control and optimization schemes.

In section 2, the AMLE algorithm is briefly reviewed. In section 3, the laboratory-scale nylon reactor model is introduced. AMLE is then used for parameter estimation in the proposed overall dynamic model. Conclusions are presented in section 4.

2. Review of the AMLE Fitting Criterion

Maximum Likelihood (ML) estimation is a very popular method for parameter estimation in a wide variety of model types because of its desirable asymptotic properties.^{9,10} Unfortunately, ML estimation in nonlinear stochastic differential equations (SDEs) is generally very difficult. This difficulty arises because, when the initial state condition with an assumed

* To whom correspondence should be addressed. E-mail: kim.mcauley@chee.queensu.ca. Tel.: +1 (613) 533 6637. Fax: +1 (613) 533 6637.

probability distribution function (PDF) is propagated through a nonlinear mapping, calculating the propagated PDF can be very complex and computationally intensive.¹¹ Available parameter estimation algorithms try to approximate the mapped PDF using techniques such as local linearization and extended Kalman filtering, deterministic sampling techniques, or ensemble averaging.^{11–18}

AMLE is an approximate ML-based method that maximizes the conditional joint density function of the states and measurements, given the model parameters, while assuming a piecewise polynomial discretization scheme for the time evolution of the states of the dynamic model. AMLE transforms the problem of state and parameter estimation in SDEs into a nonlinear minimization problem. In the following paragraphs, we briefly review the AMLE algorithm. We refer the reader to Varziri et al.^{5,7} for a more detailed description. To keep the notation simple, a single-input single-output (SISO) model with a known initial condition is used; extension to multi-input multi-output (MIMO) systems with unknown initial conditions is straightforward.⁵

Consider the following continuous-time stochastic dynamic model

$$\frac{dx(t)}{dt} = f[x(t), u(t), \theta] + \eta(t) \quad x(t_0) = x_0 \quad y(t_{mj}) = x(t_{mj}) + \varepsilon(t_{mj}) \quad (1)$$

$x \in \mathcal{R}$ is the state variable, $u \in \mathcal{R}$ is the input variable, and $y \in \mathcal{R}$ is the output variable. $\theta \in \mathcal{R}^p$ is the vector of unknown model parameters and $f: \mathcal{R} \times \mathcal{R} \times \mathcal{R}^p \rightarrow \mathcal{R}$ is a nonlinear function of the state variables, the input variables, and the parameters. We assume that f satisfies some regularity conditions,¹⁹ so that eq 1 has a unique solution. ε is a discrete zero-mean uncorrelated normal random variable with variance σ_m^2 . $\eta(t)$ is a continuous zero-mean stationary white-noise process with covariance matrix $E[\eta(t) \eta(t + \tau)] = Q\delta(\tau)$, where Q is the corresponding power spectral density and $\delta(\cdot)$ is the Dirac delta function. The random noise trajectory, $\eta(t)$, is a series of random steps with a switching time of Δt , where $\Delta t \rightarrow 0$. We also assume that the process disturbance, $\eta(t)$, and the measurement noise, $\varepsilon(t)$, are not correlated. The set of times at which the measurements are available is denoted by t_{mj} ($j = 1-n$). The measurement times, t_{mj} , do not need to be uniformly spaced. The vector of outputs at observation times $y(t_{mj})$ ($j = 1-n$) and its corresponding state vector of true values $x(t_{mj})$ ($j = 1-n$) are denoted by \mathbf{y}_m and \mathbf{x}_m , respectively. {Note that this dynamic model can be more rigorously written as $dx(t) = f[x(t), u(t), \theta] dt + d\omega(t)$, where $d\omega(t)$ is a Wiener process.²⁰}

In engineering applications, reasonable estimates for the measurement noise variance are usually available either from repeated experimental observations or from the manufacturer of the measurement device. Obtaining a reasonable estimate for the process disturbance intensity, Q , however, is very difficult. Assuming known measurement noise variance, σ_m^2 , but unknown process disturbance intensity Q , Varziri et al.⁷ used an ML argument²¹ to propose a two-step optimization scheme that allows for estimating the process disturbance intensities along with model states and parameters. The idea is to select Q to ensure that the estimated measurement noise variance, $\hat{\sigma}_m^2$, is close to the known value of σ_m^2 . The two-step optimization scheme proposed by Varziri et al.⁷ can be summarized as follows

$$\text{Outer optimization problem: } \hat{Q} = \arg \min_Q \left(\frac{\hat{\sigma}_m^2(Q)}{\sigma_m^2} - 1 \right)^2 \quad (2)$$

Inner optimization problem: $\hat{\theta}, \hat{x}(t) = \arg \min_{\theta, x(t)}$

$$\left(\frac{(\mathbf{y}_m - \mathbf{x}_m)^T (\mathbf{y}_m - \mathbf{x}_m)}{2\sigma_m^2} + \frac{1}{2\hat{Q}} \int_{t_0}^{t_f} \left\{ \frac{dx(t)}{dt} - f[x(t), u(t), \theta] \right\}^2 dt \right) \quad (3)$$

where $[t_0, t_f]$ is the time span over which the measurements are taken. The outer optimization minimizes the discrepancy between the estimated and known measurement variances, and the inner optimization minimizes the criterion in eq 3 using the value of Q obtained from the outer optimization.

Because $x(t)$ is unknown, minimizing eq 3 over $x(t)$ and θ is an infinite-dimensional optimization problem (a calculus of variations problem) that is generally hard to solve. To turn the problem into a finite-dimensional problem, the state trajectory, $x(t)$, in AMLE is assumed to be sufficiently accurately approximated by a basis function expansion. B-splines provide a convenient basis because of their compact support and other favorable properties^{4,22,23}

$$x(t) \approx x_-(t) = \sum_{i=1}^c \beta_i \phi_i = \boldsymbol{\varphi}^T(t) \boldsymbol{\beta} \quad \frac{dx_-(t)}{dt} = \frac{d}{dt} \left[\sum_{i=1}^c \beta_i \phi_i(t) \right] = \sum_{i=1}^c \beta_i \dot{\phi}_i(t) = \dot{\boldsymbol{\varphi}}^T(t) \boldsymbol{\beta} \quad (4)$$

where β_i , $i = 1-c$, represents B-spline coefficients; $\boldsymbol{\beta}$ is a vector of c B-spline coefficients; $\phi_i(t)$, $i = 1-c$, represents B-spline basis functions;²⁴ and $\boldsymbol{\varphi}(t)$ is a vector containing the c basis functions. Note that other basis functions could readily be used instead of B-splines.²³

By substituting eq 4 into eq 3, we obtain the following finite-dimensional inner optimization problem

$$\hat{\theta}, \hat{\boldsymbol{\beta}} = \arg \min_{\theta, \boldsymbol{\beta}} \left(\frac{(\mathbf{y}_m - \mathbf{x}_m)^T (\mathbf{y}_m - \mathbf{x}_m)}{2\sigma_m^2} + \frac{1}{2\hat{Q}} \int_{t_0}^{t_f} \left\{ \frac{dx_-(t)}{dt} - f[x_-(t), u(t), \theta] \right\}^2 dt \right) \quad (5)$$

The objective function in eq 5 can be minimized either simultaneously over θ and $\boldsymbol{\beta}$ or in an iterative way as described by Varziri et al.⁵ Minimizing eq 5 provides point estimates for the model parameters and the spline coefficients, given the current estimate of the process disturbance intensity \hat{Q} . The estimated spline coefficients can then be used to determine the estimated state trajectory, $x(t)$, which, in turn, is used to evaluate $\hat{\sigma}_m^2$ from eq 6 below. The variance estimate, $\hat{\sigma}_m^2$, in the objective function of eq 8 is the approximate ML estimator developed by Heald and Stark²¹

$$\hat{\sigma}_m^2 = \frac{[\mathbf{y}_m - \mathbf{x}_m(Q)]^T [\mathbf{y}_m - \mathbf{x}_m(Q)]}{n} + \frac{\text{trace}(\mathbf{A}^{-1})}{n} \hat{\boldsymbol{\theta}} \quad (6)$$

where \mathbf{A} is the Hessian matrix of the objective function in eq 3 with respect to the discrete state variables, evaluated at the converged parameter and state estimates. Varziri et al.⁵ showed that if the SDE model in eq 1 is discretized using Euler's method with a very small Δt as the discretization interval, minimizing eq 3 (and approximately minimizing eq 5) is equivalent to minimizing $-\ln p(\mathbf{y}_m, \mathbf{x} | \theta)$, where \mathbf{x} is the vector of discretized states and p denotes the probability density function. Because the state trajectory is approximated in the inner optimization step in AMLE using a B-spline expansion, matrix \mathbf{A} can also be approximated in terms of the Hessian matrix of the objective function in eq 5 with respect to the B-spline coefficients $\boldsymbol{\beta}$ evaluated at $\hat{\theta}$ and $\hat{\boldsymbol{\beta}}$. This information is often available from

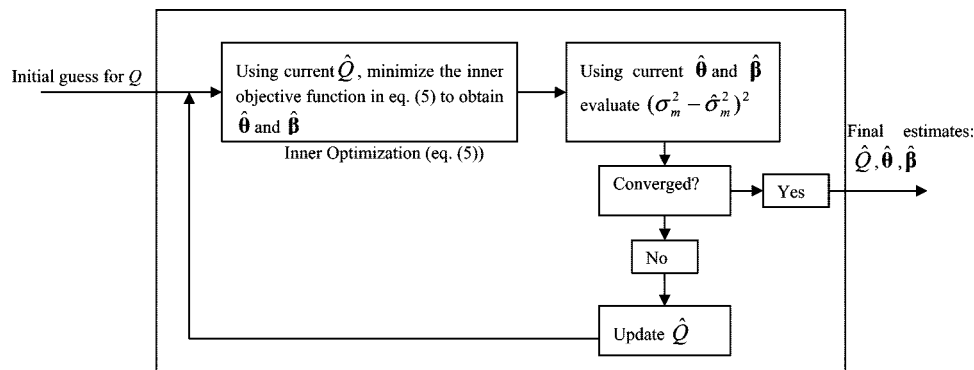


Figure 1. AMLE algorithm when Q is unknown.

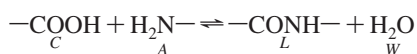
solving the inner optimization problem in eq 5 and does not need to be recalculated.

The converged results of this overall optimization problem (eqs 2 and 5) provide the modeler with $\hat{\theta}$, $\hat{\beta}$, and \hat{Q} . $\hat{\theta}$ is the desired estimate for the fundamental model parameters. The estimated spline coefficients $\hat{\beta}$, can be used to obtain the estimated state trajectory \tilde{x} , and \hat{Q} provides information about the magnitude of the model uncertainty and process disturbances. A diagram summarizing the algorithm appears in Figure 1.

Note that the outer optimization (eq 3) is an ad hoc way to ensure that the estimated intensity results in an estimate for the measurement noise variance that is consistent with its known value. Therefore, even though the parameters in the inner objective function are obtained by maximizing a likelihood function (conditioned on Q), the estimated process disturbance intensity is not a maximum likelihood estimate.

3. Case Study: Laboratory-Scale Nylon 612 Reactor Model

3.1. Nylon 612 Mathematical Model. Nylons are widely used polymers that are produced according to the following reaction between carboxylic acid end groups (C) and amine end groups (A) to produce amide linkages (L) and water (W)



The forward reaction is a polyamidation reaction in which carboxyl and amine end groups are consumed to form amide links and water. In the reverse hydrolysis reaction, amide links are broken to form carboxyl and amine end groups.

The purpose of the experimental study and modeling of nylon reactions conducted by Shaffer et al.¹ and Zheng et al.² was to gain quantitative knowledge of the kinetics and equilibrium of the polycondensation reaction at the high temperatures and low water contents that are experienced in the final stages of commercial polyamidation processes.

The following equations describe the dynamic behavior of the liquid contents of a well-stirred melt-phase nylon polymerization reactor (see Figure 2), through which a gaseous mixture of nitrogen and steam is bubbled. Refer to Schaffer et al.¹ and Zheng et al.² for detailed information about the reactor equipment and the development of the material-balance equations.

$$\frac{dL}{dt} = -\frac{dA}{dt} = -\frac{dC}{dt} = k_p \left(CA - \frac{LW}{K_a} \right) = f_1(\theta, C, A, L, W) \quad (7)$$

$$\frac{dW}{dt} = k_p \left(CA - \frac{LW}{K_a} \right) - k_m(W - W_{\text{eq}}) = f_2(\theta, C, A, L, W, W_{\text{eq}}) \quad (8)$$

In eqs 7 and 8

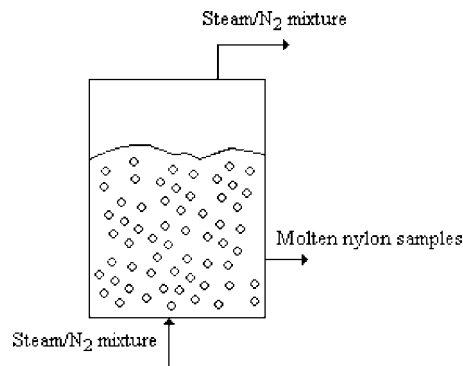


Figure 2. Nylon reactor schematic.

$$k_p = k_{p0} \exp \left[-\frac{E}{R} \left(\frac{1}{T} - \frac{1}{T_0} \right) \right] \quad (9)$$

is the temperature-dependent polycondensation rate constant, and k_{p0} is the polycondensation rate constant at the reference temperature, $T_0 = 549.15$ K, which was chosen to be in the middle of the temperature range over which the experiments were conducted. K_a is the apparent equilibrium constant, and $k_m = 24.3 \text{ h}^{-1}$ is a mass-transfer coefficient that was previously estimated¹ for the conditions encountered in the laboratory-scale reactor.

The concentration of water in the polymer melt, W_{eq} , that would be in equilibrium with the water vapor in the gas bubbles is determined using a Flory–Huggins-based expression²⁶

$$W_{\text{eq}} = 5.55 \times 10^4 \frac{P_w}{P_w^{\text{sat}}} \exp \left(-9.624 + \frac{3613}{T} \right) \quad (10)$$

where the saturation vapor pressure of the water in the gas bubbles, P_w^{sat} , can be calculated using the Wagner equation

$$\ln(P_w^{\text{sat}}/P_c) = [-7.77224(1 - T/T_c) + 1.45684(1 - T/T_c)^{1.5} - 2.71492(1 - T/T_c)^3 - 1.41336(1 - T/T_c)^6]/(T/T_c) \quad (11)$$

where P_c and T_c are the critical pressure and temperature of water, respectively.

The concentration of amide linkages, L , in the molten nylon 612 can be obtained from the material balance equation¹

$$155.23L = 10^6 - 115.15C - 58.10A - 18.02W \quad (12)$$

Note that, because the concentration of the amide linkages L can be computed algebraically from eq 12, only the differential equations for A , C , and W need to be solved.

To allow for possible modeling errors and process disturbances, stochastic terms can be added to the differential equations

$$\begin{aligned}\frac{dA}{dt} &= -f_1(\boldsymbol{\theta}, C, A, L, W) + \eta_A \\ \frac{dC}{dt} &= -f_1(\boldsymbol{\theta}, C, A, L, W) + \eta_C \\ \frac{dW}{dt} &= f_2(\boldsymbol{\theta}, C, A, L, W, W_{\text{eq}}) + \eta_W\end{aligned}\quad (13)$$

where η_A , η_C , and η_W are continuous-time Gaussian disturbances with intensities Q_A , Q_C , and Q_W [(mol Mg⁻¹)²/h], respectively.

Schaffer et al.¹ noticed that the apparent polycondensation equilibrium constant, K_a , depends on the water concentration as well as the temperature of the molten nylon. Zheng et al.² were concerned that activity coefficients for the end groups and amide links might also be influenced by temperature and therefore modified semiempirical K_a model proposed by Schaffer et al.¹ Six parameters were estimated in their nylon reactor model using a weighted nonlinear least-squares (WNLS) approach. Initial values of the states, which are uncertain, were assumed to be perfectly known and were not estimated. In this article, it will be shown that the AMLE algorithm naturally facilitates estimation of unknown initial state conditions, while properly accounting for different measurement variances for measured states. Many of the parameter values estimated from Zheng et al.'s model and the available data showed very high correlations.

Campbell³ and Ramsay et al.²² considered parameter estimation in a simplified version of Zheng et al.'s nylon reactor model using a generalized smoothing (GS) approach. Like AMLE, the GS approach solves a two-step minimization problem. However, there are some important differences that distinguish AMLE from GS. The main differences are in the form of the outer optimization objective function. In the GS approach, the outer objective function is a WNLS objective function that is minimized over the model parameters $\boldsymbol{\theta}$. The inner objective function in GS contains a tuning parameter that is adjusted manually by the user to account for possible model imperfections. A major advantage of the AMLE algorithm over the GS approach is that AMLE eliminates the requirement for manual tuning. In AMLE, the disturbance intensities, Q , are estimated using the outer optimization so that the estimated measurement noise variance is consistent with prior knowledge about the quality of the measurements.

To resolve concerns about possible overparameterization in Zheng et al.'s expression for the apparent equilibrium constant, Varziri²⁶ used additional steady-state data²⁷ and selected the following semiempirical expression for K_a

$$K_a = \left[\frac{1 + a\sqrt{W_{\text{eq}}}}{\gamma_W/\gamma_{W_0}} \right] K_{a0} \exp\left[-\frac{\Delta H}{R} \left(\frac{1}{T} - \frac{1}{T_0} \right) \right] \quad (14)$$

where the activity coefficient for water in the molten nylon²⁵ is given by $\gamma_W = \exp(9.624 - 3613/T)$ and $\gamma_{W_0} = 20.97$. The empirical term $1 + a\sqrt{W_{\text{eq}}}$ in the numerator of eq 14 accounts for the influence of water on the activity coefficients of amine ends, carboxyl ends, and amide links in the polymer melt.

Equation 14 and additional steady-state data²⁷ are used in this article. The model selection scheme presented by Varziri²⁶ is briefly described below.

3.2. Model Selection for K_a . The strategy used by Varziri²⁶ for K_a submodel selection is based on the definition of the apparent polycondensation equilibrium constant²

$$K_a = \frac{L_{\text{eq}} W_{\text{eq}}}{C_{\text{eq}} A_{\text{eq}}} \quad (15)$$

where the subscript eq indicates equilibrium concentrations, which are available from steady-state data. From the steady-state data, K_a was calculated and treated as a measured response (y_{K_a}). The variance $\sigma_{K_a}^2$ of the error, ε_{K_a} , was estimated to be 94.9. Varziri²⁶ denoted any postulated semiempirical submodel for K_a by

$$y_{K_a} = f_{K_a}(\boldsymbol{\theta}_{K_a}, W_{\text{eq}}, T) + \varepsilon_{K_a} \quad (16)$$

Eight possible semiempirical submodels for K_a were considered, including those previously proposed.¹⁻³ For each of these candidate submodels, the nonlinear regression problem in eq 16 was solved, using the available steady-state data (see section 3.3), so that estimates for empirical parameters $\boldsymbol{\theta}_{K_a}$ could be obtained. Based on the sum of squared residuals, qualitative analysis of residual plots, and also approximate individual confidence intervals for parameter estimates, the model in eq 14 was concluded to be the most suitable. The estimated parameters are $a = 0.60 \text{ mol}^{-0.5} \text{ Mg}^{0.5}$, $K_{a0} = 22.01$, and $\Delta H = -39.62 \text{ kJ mol}^{-1}$, respectively.

3.3. Experimental Data. Note that all of the data used in this article result from previous experiments conducted by Schaffer et al.¹ and Zheng et al.^{2,27} Recent steady-state data that were not considered in previous parameter estimation studies¹⁻³ arise from experimental runs conducted to study the effects of sodium hypophosphite catalyst on the polycondensation kinetics.²⁷

Of the six dynamic experimental runs that are available for parameter estimation, the first three were conducted at temperatures of 263, 271, and 281 °C, respectively, whereas the last three runs were conducted at 284 °C. The additional steady-state data arise from experimental runs at 290 °C. During each experimental run, the concentrations of A and C are measured at several nonuniformly spaced times. The standard deviations of the A and C concentration measurements are known to be $\sigma_A = 0.6 \text{ mol Mg}^{-1}$ and $\sigma_C = 2.4 \text{ mol Mg}^{-1}$.¹ Because the measured concentration of A is more accurate, more A concentration measurements than C concentration measurements were made during some of the runs. The water concentration, W, was not measured; therefore, W is an unmeasured dynamic state. The equilibrium water concentration, W_{eq} , in the molten nylon can be computed from the input variable P_W using eq 10.

3.4. Parameter Estimation Results. In this section, we use the objective function of the form given in eqs 2 and 3 to estimate five parameters, $\boldsymbol{\theta} = [k_{p0}, K_{a0}, E, a, \Delta H]^T$, along with three process disturbance intensities, $\mathbf{Q} = [Q_A, Q_C, Q_W]^T$, and three state trajectories in the nylon 612 reactor model described by eqs 9–14.

To form the AMLE objective function, we use B-spline expansions to approximate state trajectories. For the i th experimental run for A, C, and W, from eq 4

$$A_{-i}(t) = \boldsymbol{\varphi}_{A_i}^T(t) \boldsymbol{\beta}_{A_i} \quad C_{-i}(t) = \boldsymbol{\varphi}_{C_i}^T(t) \boldsymbol{\beta}_{C_i} \quad W_{-i}(t) = \boldsymbol{\varphi}_{W_i}^T(t) \boldsymbol{\beta}_{W_i} \quad (17)$$

Because W is not measured, there is no SSE term associated with this component. The AMLE objective function corresponding to the i th run for this multiresponse model with an unmeasured state⁶ becomes

$$\begin{aligned}
J_i = & \frac{1}{2\sigma_C^2} \sum_{j=1}^{N_{C_i}} [y_{C_i}(t_j) - C_{-i}(t_j)]^2 + \frac{1}{2Q_C} \int_{t_{C_0}}^{t_{C_i}} \left\{ \frac{dC_{-i}(t)}{dt} + \right. \\
& \left. f_1[A_{-i}(t), C_{-i}(t), W_{-i}(t), L_i(t), \boldsymbol{\theta}] \right\}^2 dt + \\
& \frac{1}{2\sigma_A^2} \sum_{j=1}^{N_{A_i}} [y_{A_i}(t_j) - A_{-i}(t_j)]^2 + \frac{1}{2Q_A} \int_{t_{A_0}}^{t_{A_i}} \left\{ \frac{dA_{-i}(t)}{dt} + \right. \\
& \left. f_1[A_{-i}(t), C_{-i}(t), W_{-i}(t), L_i(t), \boldsymbol{\theta}] \right\}^2 dt + \\
& \frac{1}{2Q_W} \int_{t_{W_0}}^{t_{W_i}} \left\{ \frac{dW_{-i}(t)}{dt} - f_2[A_{-i}(t), C_{-i}(t), W_{-i}(t), W_{\text{eq}}(t), L_i(t), \boldsymbol{\theta}] \right\}^2 dt
\end{aligned} \quad (18)$$

where N_{A_i} and N_{C_i} are the numbers of A and C measurements, respectively, in the i th experimental run.

Because initial conditions for A and C are assumed to be unknown, they are included in the corresponding SSE terms, and the B-spline expansions are not constrained to the initial conditions as they would be if the true initial values were known.

Because only six experimental runs without the catalyst are used in the overall parameter estimation, the inner AMLE objective function for these six runs becomes

$$J_{\text{inner}} = \sum_{i=1}^6 J_i \quad (19)$$

Using the objective function in eq 19, the extra steady-state information available from the three experimental runs with catalyst would be disregarded. To include this extra information, the objective function in eq 19 can be modified as follows

$$J_{\text{inner}} = \sum_{i=1}^6 J_i + \frac{1}{2} \sum_{i=1}^3 [y_{K_a} - f_{K_a}(\boldsymbol{\theta}_{K_a}, W_{\text{eq}}, T)]^2 \quad (20)$$

where $f_{K_a}(\boldsymbol{\theta}_{K_a}, W_{\text{eq}}, T)$ is defined by the right-hand side of eq 14. This modification arises naturally from the ML development when y_{K_a} is considered as an additional nondynamic measured response.²⁶

To form the objective function for the outer optimization problem, note that the approximate ML estimators for the measurement variances for A and C can be expressed as

$$\hat{\sigma}_x^2 = \frac{\sum_{i=1}^6 \sum_{j=1}^{N_{x_i}} [y_{x_i}(t_j) - x_{-i}(t_j)]^2}{\sum_{i=1}^6 N_{x_i}} + \frac{\text{trace}(\mathbf{A}_x^{-1})}{\sum_{i=1}^6 N_{x_i}} \quad x = A, C \quad (21)$$

where \mathbf{A}_x^{-1} is the inverse of the Hessian of the inner objective function, J_{inner} , with respect to the discrete state variables. Denoting $\nu_x = 1/\hat{\sigma}_x^2 \text{trace}(\mathbf{A}_x^{-1})$, the second term on the right-hand side of eq 21 can be written as $\hat{\sigma}_x^2 \nu_x / \sum_{i=1}^6 N_{x_i}$. By moving this term to the left-hand side of eq 21 and factoring out $\hat{\sigma}_x^2$, eq 21 can be rearranged²¹ to

$$\hat{\sigma}_x^2 = \frac{\sum_{i=1}^6 \sum_{j=1}^{N_{x_i}} [y_{x_i}(t_j) - x_{-i}(t_j)]^2}{\sum_{i=1}^6 N_{x_i} - \nu_x} \quad x = A, C \quad (22)$$

Table 1. Process Disturbance Intensity Estimates^a

intensity	estimate
Q_A	4.37
Q_C	15.56
Q_W	0.88

^a All intensities have units of $(\text{mol Mg}^{-1})^2/\text{h}$.

Table 2. Point Estimates and Approximate 95% Confidence Intervals for the Nylon Reactor Model Parameters

parameter	initial guess	estimate	lower bound	upper bound
k_{p0}	0.019	0.0129	0.0106	0.0152
K_{a0}	22.01	36.605	27.247	45.963
E	45.9	2E-4	-53.434	53.435
a	0.6	0.286	0.170	0.402
ΔH	-39.62	-51.012	-60.208	-41.815

In this article, the known values of the measurement noise variance, σ_x^2 , is used in place of $\hat{\sigma}_x^2$ to calculate ν_x . The denominator in eq 22 is the degrees of freedom, $\text{DOF}_x = \sum_{i=1}^6 N_{x_i} - \nu_x$.

The outer optimization problem can then be written as

$$\hat{\mathbf{Q}} = \arg \min_{\mathbf{Q}} \left\{ \text{DOF}_A \left[\frac{\hat{\sigma}_A^2(\mathbf{Q})}{\sigma_A^2} - 1 \right]^2 + \text{DOF}_C \left[\frac{\hat{\sigma}_C^2(\mathbf{Q})}{\sigma_C^2} - 1 \right]^2 \right\} \quad (23)$$

Note that the terms in the objective function of eq 23 are weighted by DOF_A and DOF_C to account for the different numbers of available measurements for the amine and carboxyl end groups.

Three of the parameters, namely, a , K_{a0} , and ΔH , were estimated by Varzir²⁶ using the steady-state data (section 3.2). These parameters were estimated again using AMLE, and the previously estimated values of $a = 0.60 \text{ mol}^{-0.5} \text{ Mg}^{0.5}$, $K_{a0} = 22.01$, and $\Delta H = -39.62 \text{ kJ mol}^{-1}$ were used as initial parameter guesses. For the rest of the parameters, estimated values reported by Zheng et al.² were used as initial values, $\boldsymbol{\theta}_0$, as shown in Table 2, below.

The initial values for the spline coefficients, β_{A_i} and β_{C_i} ($i = 1-6$) were obtained by fitting a smoothing spline²³ to the measured data. Because the water concentration is not measured, the initial guesses for the spline coefficients β_{W_i} ($i = 1-6$) were obtained by fitting a smoothing spline to W_{eq} , which was calculated from eq 10. An alternative would be to solve the model differential equations and fit a smoothing spline to the solution for W . Based on our experience, obtaining good (nonzero) initial guesses for B-spline coefficients is not necessary but can reduce the convergence time. The initial values of the process disturbance intensities were arbitrarily set to 1. The estimated intensities, $\hat{\mathbf{Q}}$, are shown in Table 1. The estimated parameters, $\hat{\boldsymbol{\theta}}$, are reported in Table 2. The results in Tables 1 and 2 were obtained using the following stopping conditions: For the inner problem, the optimizer stopped when the objective function changed by less than 1×10^{-8} ; the outer optimizer stopped when the value of the objective function changed by less than 1×10^{-2} .

The final B-spline fits using the estimated parameters and coefficients are shown in Figure 3. The numerical solution of the model equations, which does not account for the stochastic process disturbances, is shown in Figure 4. Initial values for the output trajectories in Figure 4 were determined using the AMLE algorithm. Note that the B-spline curves pass very close to the A measurements, and they smooth out some of the noise associated with the noisier C measurements because we prescribed that the standard deviation of the known noise

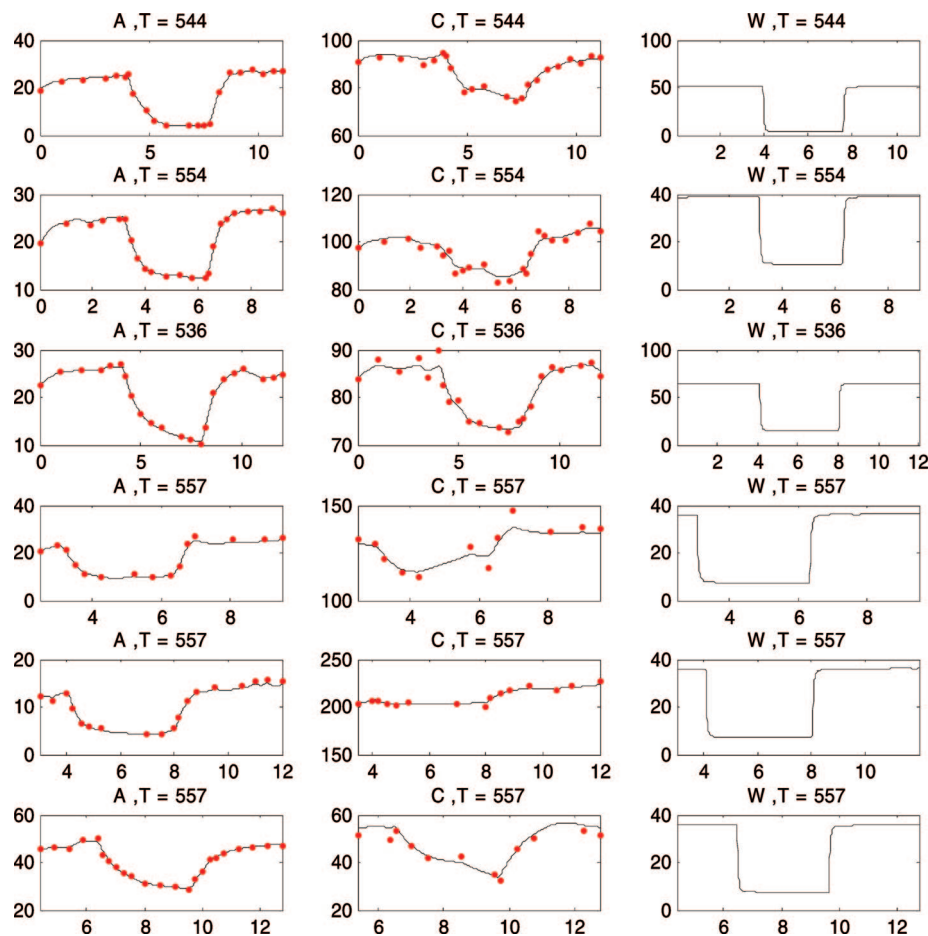


Figure 3. Optimal B-spline trajectories of A , C , and W for six experimental runs (—, B-spline fit; ●, measured data).

variance associated with C is larger than that of A ($\sigma_A = 0.6$ mol Mg⁻¹ and $\sigma_C = 2.4$ mol Mg).

To obtain approximate confidence intervals for the model parameters, the inverse of the Fisher information matrix can be used as an approximation to the covariance matrix of the combined vector of states and parameters.^{6,8} We approximate the Fisher information matrix using the Hessian (\mathbf{H}) of the inner objective function with respect to the model parameters θ evaluated at the converged values. Approximate $100(1 - \alpha)\%$ confidence intervals for the model parameters can be obtained as follows

$$\theta = \hat{\theta} \pm z_{\alpha/2} \sqrt{\text{diag}[\mathbf{H}^{-1}(\hat{\theta})]} \quad (24)$$

Care should be taken in interpreting the confidence intervals obtained from eq 24; these intervals are approximate and also do not take into account the uncertainty in the estimates of the process disturbance intensities.

Approximate 95% confidence intervals for the parameters are presented in Table 2, and a correlation matrix for the parameter estimates is presented in Table 3. The confidence interval corresponding to the activation energy E contains zero. This is consistent with the findings of Campbell³ and Ramsay et al.²² Although this parameter could reasonably be zero from a purely statistical point of view, physical insight tells us that the activation energy is a positive constant. Unfortunately, the dynamic data are not rich enough to provide sufficient information to estimate E reliably. Table 3 indicates that correlations between the parameter estimates are small, except for between K_{a0} and a . These two parameters are nearly perfectly negatively correlated.

To examine the robustness of the proposed algorithm to poor initial parameter guesses, the AMLE parameter estimation was repeated using the arbitrary value of 1 as an initial guess for the five model parameters: $\theta_0 = [1, 1, 1, 1, 1]^T$. Despite these poor parameter guesses, the algorithm converged to point estimates that are nearly identical to those shown in Tables 1 and 2 (identical up to the second decimal place) for the model parameters and the disturbance intensities.

For comparison, we also used a standard WNLS approach, which does not account for model imperfections and process disturbances, to estimate the model parameters. To have a fair comparison, we included the additional steady-state data (from the experimental runs with catalyst) in our WNLS analysis, and we also estimated 18 initial conditions for A , C , and W for the six experimental runs. First, we used the same good initial parameter guesses that were used for AMLE. The estimated parameters did not move very far from the initial guesses, producing the parameter estimates in Table 4. This result is not surprising, because the parameter estimation was started from the optimal WNLS parameter estimates obtained by Zheng et al.² The main differences between the WNLS approach reported here and that of Zheng et al.² is that Zheng et al. assumed known initial conditions for A , C , and W and did not use steady-state information from the experiments performed with catalyst. In a second trial, when we used the poor initial parameter guesses, the WNLS algorithm converged to unrealistic values possibly corresponding to a local minimum (Table 4). The results in Table 4 were obtained using a relative tolerance value of 1×10^{-12} for changes in the objective function and for changes in the norm of the parameter values. Attempts to obtain confidence

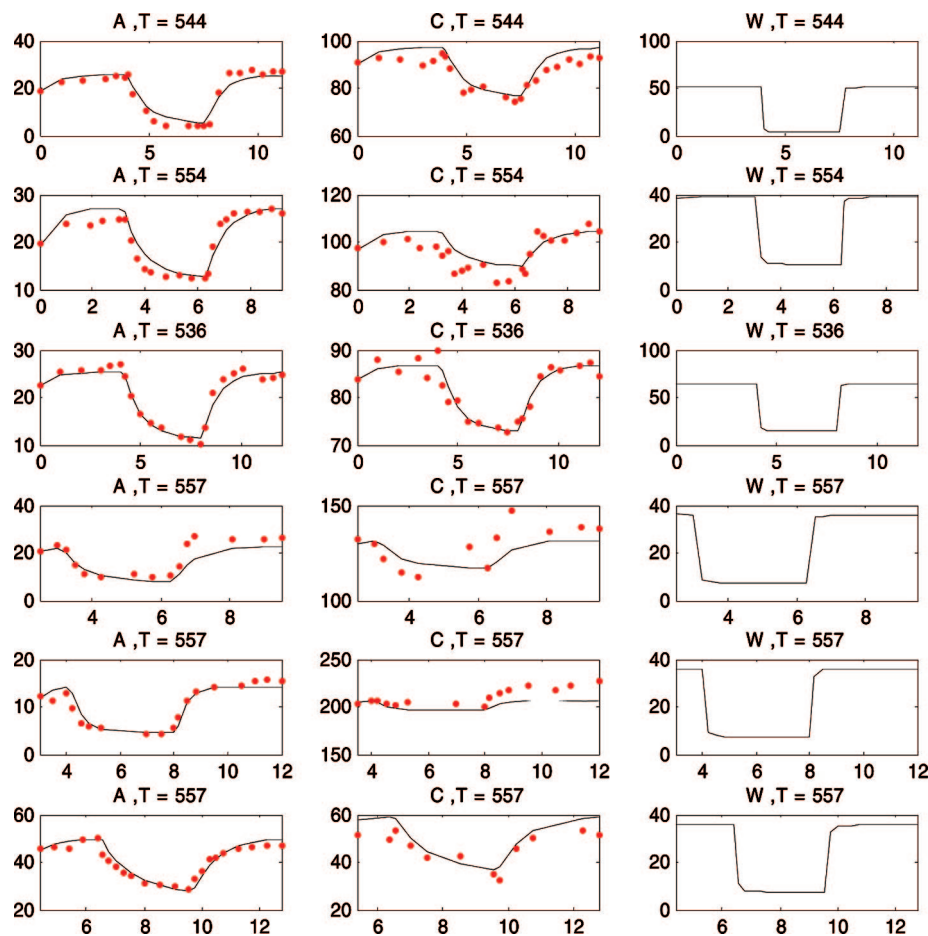


Figure 4. Numerical solution of model equations (without stochastic terms) for A , C , and W trajectories for six experimental runs using the estimated parameters (—, numerical model solution; ●, measured data).

Table 3. Correlation Matrix for AMLE Parameter Estimates

	K_{a0}	k_{p0}	E	ΔH	a
K_{a0}	1	-0.2478	-0.0564	-0.4858	-0.9935
k_{p0}		1	-0.1309	0.0542	0.2807
E			1	0.1061	0.0480
ΔH				1	0.4659
a					1

Table 4. Point Estimates for the Nylon Reactor Model Parameters Using the WNLS Method

parameter	estimation results using good initial guesses		estimation results using poor initial guesses	
	initial guess	estimate	initial guess	estimate
k_{p0}	0.019	0.019	1	1
K_{a0}	22.01	22.0104	1	10.62
E	45.9	45.9	1	1
a	0.6	0.611	1	1.123
ΔH	-39.62	-39.62	1	17.81

intervals for the WNLS parameters were unsuccessful because the Jacobian, evaluated at the converged estimates, was very ill-conditioned, indicating severe correlations among the parameter estimates. Note that Zheng et al.² were able to compute approximate confidence intervals for their parameter estimates, presumably because they assumed perfectly known initial conditions.

3.5. Implementation Considerations. Choice of the Knot Sequence. For three state trajectories and six experimental runs ($3 \times 6 = 18$), B-spline expansions are required; we used

a fourth-order B-spline basis (third-degree). Rich knot sequences are required to ensure that the state trajectories are flexible enough to capture all of the features in the dynamic response. To set up the bases, 60 knots were uniformly placed along the time horizon of each state trajectory for each experimental run. Because the partial pressure of the water in these experimental runs was adjusted using step changes, the water concentration in the molten polymer underwent fast changes that were almost perfect steps. To accommodate the sharp transitions in the W response, 20 extra knots were uniformly placed in the neighborhood (within 0.5 h) of the times at which step changes occurred. These extra knots helped remove ripples in the final B-spline expansion of the W trajectory⁴ that were apparent using the initial coarse knot sequence. In addition to the refined knot sequence, coincident knots were placed at the times of step changes so that first- and higher-order derivatives of the state trajectories were not continuous at the times of the step changes. Overall, 108 B-spline coefficients were used for each of the 18 B-spline expansions.

If the process disturbances and structural model imperfections are small, the values of the B-spline coefficients are determined by requiring the estimated trajectory to closely follow the solution to the differential equations. However, if the process disturbances and model imperfections are large, the measured data have a larger effect on the estimated value of the B-spline coefficients, and having too many coefficients, compared to the number of data points, can lead to large variances for the estimated coefficients.

Calculating the Integral in the Inner Minimization. To calculate the integral in eq 18, a quadrature rule was used; four collocation points were placed between every two knots. As noted by Campbell³ and Ramsay et al.,²² the step input in W_{eq} results in discontinuous derivatives for all three model outputs at the times of step changes. As a consequence, a small neighborhood (within 1×10^{-4} h) around each of the times at which the step changes occur was removed during calculation of the integral in eq 18.

Minimization Routines. As discussed above, overall, the inner objective function in this problem should be optimized over $108 \times 18 = 1944$ B-spline coefficients as well as five model parameters θ . Every time that the variables in the outer minimization problem (i.e., the process disturbance intensities) are updated, the inner problem should be solved again. Therefore, it is essential for the AMLE algorithm to take advantage of fast and efficient state-of-the-art minimization routines. Based on our experience, IPOPT,²⁸ which is a nonlinear solver that can be used with AMPL,²⁹ provides an excellent tool for solving nonlinear optimization problems. AMPL endows IPOPT with automatic differentiation capability, which eliminates the requirement of providing the nonlinear solver with an analytical or a numerical Jacobian. The Hessian matrix that is used in eq 21 was obtained using the automatic differentiation feature in AMPL.

Unfortunately, we were not able to implement the complete two-level minimization problem as appears in eqs 20 and 23 using AMPL, because complicated user-defined functions (such as the solution to an optimization problem) are not permitted in AMPL in a straightforward fashion. Note that each iteration of the outer optimization problem requires the solution of another (inner) minimization problem. Instead, we opted to use *nlscon*,³⁰ which is a very efficient nonlinear solver in Matlab, to solve the outer optimization problem and the combination of AMPL and IPOPT to solve the larger inner minimization problem by calling AMPL from within Matlab.

Another issue regarding the inner minimization problem is whether the initial values of the parameter guesses, θ_0 , for each iteration should be set to the converged values from the previous iteration or the same initial guesses as used in the first iteration. We tested both methods for this problem and found that both methods led to the same point and interval estimates for the model parameters and disturbance intensities.

4. Summary and Conclusions

In this article, the AMLE parameter estimation algorithm was applied to a laboratory-scale nylon 612 reactor model, which is a difficult and practical chemical engineering example, to estimate the states, parameters, and process disturbance intensities. We showed that using AMLE allows the modeler to take modeling errors into account and to obtain a measure of these discrepancies by estimating the corresponding process noise intensities.

This parameter estimation problem involved several challenges. The first challenge was to exploit all of the available data, which consisted of data from six dynamic experimental runs without a catalyst and also steady-state data from three additional experimental runs with a catalyst. It was shown that the AMLE objective function can easily be modified to include the additional steady-state information. The nylon reactor is a multiresponse model with three outputs. Two of the outputs, namely, the concentration of carboxylic acid end

groups and the concentration of amine end groups, were measured at different and nonuniform sampling times, with different levels of accuracy. The third output, water concentration, was not measured. It was demonstrated that AMLE can easily cope with these challenges, because the form of its inner objective function allows for the incorporation of nonuniform observation times and unmeasured states. Another difficulty encountered in the nylon reactor parameter estimation problem is that initial state conditions are not perfectly known and should be treated as unknowns that need to be estimated. These unknown initial conditions were obtained as a byproduct of estimating the state trajectories using AMLE, which is facilitated by using B-spline expansions to represent the state trajectories.

AMLE was successfully applied to the nylon reactor problem, and parameters, states, and process disturbance intensities were estimated. Approximate confidence intervals were obtained for model parameters. The interval estimate corresponding to the activation energy parameter, E , contained zero, indicating that there was insufficient dynamic information available to obtain a reliable estimate.

To investigate how sensitive the outcome of AMLE was to initial parameter guesses, we repeated the parameter estimation algorithm using arbitrary initial parameter guesses that were far from the estimated values. Despite these poor initial guesses, the algorithm converged to the same point estimates as it had for the good initial guesses. A standard weighted nonlinear least-squares algorithm failed to converge to reasonable parameter estimates using the poor initial guesses and converged to parameter estimates that were very close to the corresponding initial values for the good initial guesses.

Acknowledgment

The authors thank Cybernetica, E.I. du Pont de Nemours and Company, Hatch, Matrikon, SAS, MITACS (Mathematics of Information Technology and Complex Systems), the Natural Sciences and Engineering Research Council of Canada (NSERC), the Government of Ontario, and Queen's University for financial support and Dr. Tom Harris from Queen's University for technical advice.

Nomenclature

AMLE = approximate maximum likelihood parameter estimation
 DOF = degrees of freedom
 GS = generalized smoothing
 MIMO = multi-input multi-output
 ML = maximum likelihood
 NLS = nonlinear least-squares
 PDF = probability density function
 SDE = stochastic differential equation
 SISO = single-input single-output
 SSE = sum of squared errors
 WNLS = weighted nonlinear least-squares
 A = concentration of amine end groups, mol Mg⁻¹
 A = Hessian of the inner objective function with respect to the system states
 C = concentration of carboxylic acid end groups, mol Mg⁻¹
 E = activation energy, kJ mol⁻¹
 $E\{\cdot\}$ = expectation operator
 H = Hessian matrix of the inner objective function with respect to model parameters
 ΔH = apparent enthalpy of polycondensation, kJ mol⁻¹

J_i = i th component (i th experimental run) of the inner objective function
 J_{inner} = inner objective function
 K_a = apparent polycondensation equilibrium constant
 K_{a0} = apparent polycondensation equilibrium constant at T_0 and low water content
 L = concentration of amide links
 N_{A_i} = number of observations for A in the i th experimental run
 N_{C_i} = number of observations for C in the i th experimental run
 P_c = critical pressure of water, kPa
 P_w = partial pressure of water in the gas phase, kPa
 P_w^{sat} = saturation pressure of water in the gas phase, kPa
 Q = process disturbance intensity
 Q_A = process disturbance intensity for SDE corresponding to A , (mol Mg⁻¹)²/h
 Q_C = process disturbance intensity for SDE corresponding to C , (mol Mg⁻¹)²/h
 Q_W = process disturbance intensity for SDE corresponding to W , (mol Mg⁻¹)²/h
 Q = vector containing process disturbance intensities
 R = ideal gas law constant, 8.3145×10^{-3} kJ mol⁻¹ K⁻¹
 T = temperature, K
 T_0 = reference temperature, 549.15 K
 T_c = critical temperature of water, K
 W = concentration of water, mol Mg⁻¹
 X = exponent in the kinetic expression for the reaction rate
 a = empirical model parameter, (mol Mg⁻¹)^{0.5}
 f = nonlinear function
 f_{K_a} = nonlinear function for K_a model
 k_m = volumetric liquid-phase mass-transfer coefficient for a nylon/water system, h⁻¹
 k_p = apparent polycondensation rate constant, Mg mol⁻¹ h⁻¹
 k_{p0} = apparent polycondensation rate constant at the reference temperature, 549.15 K
 t_{mj} = time of the j th measurement, h
 $u(t)$ = input function
 x_0 = initial state value
 $x(t)$ = state of the system
 $\tilde{x}(t)$ = spline approximation of the system state
 \mathbf{x}_m = vector of state values at measurement times
 $y(t_{mj})$ = measured value at time t_{mj}
 y_{K_a} = K_a calculated from the measurements
 \mathbf{y}_m = vector of the measurements
 $z_{\alpha/2}$ = normal random deviate corresponding to an upper tail area of $\alpha/2$
 α = significance level for confidence intervals
 β_i = i th B-spline coefficient
 $\boldsymbol{\beta}$ = vector of B-spline coefficients
 $\delta(\cdot)$ = Dirac delta function
 $\varepsilon(t_{mj})$ = normally distributed measurement noise for state i
 ε_{K_a} = approximate error in K_a calculated using the measurements
 η_A = Gaussian process disturbance for differential equation of state A , mol Mg⁻¹/h
 η_C = Gaussian process disturbance for differential equation of state C , mol Mg⁻¹/h
 η_W = Gaussian process disturbance for differential equation of state W , mol Mg⁻¹/h
 $\boldsymbol{\theta}$ = vector of model parameters
 $\boldsymbol{\theta}_0$ = initial value for the vector of model parameters
 $\boldsymbol{\theta}_{K_a}$ = vector of K_a model parameters
 σ_C^m = measurement noise variance
 σ_A^2 = measurement noise variance for A , (mol Mg⁻¹)²
 σ_C^2 = measurement noise variance for C , (mol Mg⁻¹)²

$\sigma_{K_a}^2$ = approximate noise variance for K_a calculated from measurements

φ_i = i th B-spline basis function

$\boldsymbol{\varphi}$ = matrix containing all φ_i values

Literature Cited

- (1) Schaffer, M. A.; McAuley, K. B.; Cunningham, M. F.; Marchildon, E. K. Experimental study and modeling of nylon polycondensation in the melt phase. *Ind. Eng. Chem. Res.* **2003**, *42*, 2946–2959.
- (2) Zheng, W.; McAuley, K. B.; Marchildon, E. K.; Yao, K. Z. Effects of end group balance on melt-phase nylon 612 polycondensation: Experimental study and mathematical model. *Ind. Eng. Chem. Res.* **2005**, *44*, 2675–2686.
- (3) Campbell, D. A. Bayesian Collocation Tempering and Generalized Profiling for Estimation of Parameters for Differential Equation Models. Ph.D. Thesis, Department of Mathematics and Statistics, McGill University, Montreal, Quebec, Canada, 2007.
- (4) Poyton, A. A.; Varziri, M. S.; McAuley, K. B.; McLellan, P. J.; Ramsay, J. O. Parameter estimation in continuous-time dynamic models using principal differential analysis. *Comput. Chem. Eng.* **2006**, *30*, 698–708.
- (5) Varziri, M. S.; Poyton, A. A.; McAuley, K. B.; McLellan, P. J. Selecting optimal weighting factors in iPDA for parameter estimation in continuous-time dynamic models. *Comput. Chem. Eng.*, published online, <http://dx.doi.org/10.1016/j.compchemeng.2008.04.005>.
- (6) Varziri, M. S.; McAuley, K. B.; McLellan, P. J. Parameter estimation in continuous-time dynamic models in presence of unmeasured states and nonstationary disturbances. *Ind. Eng. Chem. Res.* **2008**, *47*, 380–393.
- (7) Varziri, M. S.; McAuley, K. B.; McLellan, P. J. Parameter estimation in nonlinear stochastic continuous-time dynamic models with unknown disturbance intensity. *Can. J. Chem. Eng.*, accepted for publication.
- (8) Seber, G. A. F.; Wild, C. J. *Nonlinear Regression*; John Wiley and Sons, Inc.: New York, 1989.
- (9) Bates, D. M.; Watts, D. G. *Nonlinear Regression Analysis and Its Applications*; John Wiley and Sons, Inc.: New York, 1988.
- (10) Timmer, J. Parameter estimation in nonlinear stochastic differential equations. *Chaos Sol. Frac.* **2000**, *11*, 2571–2578.
- (11) Jazwinski, A. H. *Stochastic Processes and Filtering*; Academic Press: New York, 1970.
- (12) Maybeck, P. S. *Stochastic Models, Estimation, and Control*; Academic Press: New York, 1982; Vol. 2.
- (13) Nielsen, J. N.; Madsen, H.; Young, P. C. Parameter estimation in stochastic differential equations: An overview. *Annu. Rev. Cont.* **2000**, *24*, 83–94.
- (14) Roweis, S. T.; Ghahramani, Z. Learning Nonlinear Dynamical Systems Using the Expectation–Maximization Algorithm. In *Kalman Filtering and Neural Networks*; Haykin, S., Ed.; John Wiley & Sons: New York, 2001; pp 175–220.
- (15) Evensen, G. The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dyn.* **2003**, *53*, 343–367.
- (16) Julier, S. J.; Uhlmann, J. K. Unscented filtering and nonlinear estimation. *Proc. IEEE* **2004**, *92* (3), 401–422.
- (17) Voss, H. U.; Timmer, J.; Kurths, J. Nonlinear dynamical system identification from uncertain and indirect measurements. *Int. J. Bifurcation Chaos.* **2004**, *14* (6), 1905–1933.
- (18) Kristensen, N. R.; Madsen, H.; Jorgensen, S. B. Parameter estimation in stochastic grey-box models. *Automatica* **2004**, *40*, 225–237.
- (19) Kloeden, P. E.; Platen, E. *Numerical Solution of Stochastic Differential Equations*; Springer-Verlag: New York, 1992.
- (20) Maybeck P. S., *Stochastic Models, Estimation, and Control*; Academic Press: New York, 1979; Vol. 1.
- (21) Heald, J. P. M.; Stark, J. Estimation of noise levels for models of chaotic dynamical systems. *Phys. Rev. Lett.* **2000**, *84*, 2366–2369.
- (22) Ramsay, J. O.; Hooker, G.; Campbell, D.; Cao, J. Parameter estimation for differential equations: A generalized smoothing approach. *J. R. Stat. Soc.: Ser. B* **2007**, *69* (5), 741–796.
- (23) Ramsay J. O.; Silverman B. W., *Functional Data Analysis*, 2nd ed.; Springer: New York, 2005.
- (24) de Boor C., *A Practical Guide to Splines*; Springer: New York, 2001.
- (25) Schaffer, M. A.; Marchildon, E. K.; McAuley, K. B.; Cunningham, M. F. Prediction of water solubility in nylon melts based on Flory–Huggins theory. *Polym. Eng. Sci.* **2003**, *43*, 639–646.
- (26) Varziri, M. S. Parameter estimation in nonlinear continuous-time dynamic models with modeling errors and process disturbances. Ph.D. Thesis, Queen's University, Kingston, Canada, in preparation.

(27) Zheng, W.; McAuley, K. B.; Marchildon, E. K.; Yao, K. Z. Melt-phase nylon 612 polycondensation kinetics: Effects of sodium hypophosphite catalyst. *Can. J. Chem. Eng.* **2007**, *85*, 180–187.

(28) Wächter, A.; Biegler, L. T. On the implementation of a primal-dual interior point filter line search algorithm for large-scale nonlinear programming. *Math. Prog.* **2006**, *106* (1), 25–57.

(29) Fourer, R.; Gay, D. M.; Kernighan, B. W. *AMPL: A Modeling Language for Mathematical Programming*, 2nd ed.; Thomson Brooks/Cole: Florence, KY, 2003.

(30) Nowak, U.; Weimann, L. *a Family of Newton Codes for Systems of Highly Nonlinear Equations*; Report TR-91-10; Konrad-Zuse-Zentrum für Informationstechnik: Berlin, Germany, 1991.

Received for review March 31, 2008

Revised manuscript received June 2, 2008

Accepted July 9, 2008

IE800503V