12-1-2000

Time-Varying System Identification Using Modulating Functions and Spline Models With Application to Bio-Processes

Sridhar Ungarala
Cleveland State University, s.ungarala@csuohio.edu

Tomas B. Co
Michigan Technological University

Publisher's Statement

NOTICE: this is the author's version of a work that was accepted for publication in Computers & Chemical Engineering. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in Computers & Chemical Engineering, 24, 12, (December 1, 2000) DOI 10.1016/S0098-1354(00)00624-4

Original Citation

Ungarala, S., , & Co, T. B. (2000). Time-varying system identification using modulating functions and spline models with application to bio-processes. Computers and Chemical Engineering, 24(12), 2739 - 2753.

Repository Citation

Time-varying system identification using modulating functions and spline models with application to bio-processes

Sridhar Ungarala [a,*], Tomas B. Co [b]

[a] *Department of Chemical Engineering, Cleveland State University, Cleveland, OH 44115, USA*
[b] *Department of Chemical Engineering, Michigan Technological University, Houghton, MI 49931, USA*

1. Introduction

Process modeling and parameter estimation together have become an integral part of the study of dynamic processes, and the design of advanced control and optimization schemes. Many real processes are modeled in continuous-time by nonlinear differential equations with characteristic parameters motivated by the physical processes occurring in the system. These parameters are indications of complex relationships between various physical characteristics of the process and may often exhibit time dependency due to the inherent non-stationary nature of the process or due to model inaccuracies. A knowledge of the process parameters along with the process state variables is highly desirable in real-time for three important reasons. First, tracking the trajectory of a time-varying parameter provides a better understanding of the process response to aid in model development and verification (Bastin & Dochain, 1990; Font & Fabregat, 1997). Second, the design of model based predictive control algorithms requires determination of the process parameters accurately in real-time (Goodwin & Mayne, 1987; Hecker, Knapp & Isermann, 1994; Feng, Genceli & Nikolaou, 1996). And third, process monitoring and supervision are closely related to the identification of changes or drifts in process parameters which may indicate the inception and growth of fault modes in a process (Iserman, 1984; Höfling & Isermann, 1996).

Many different approaches are available for the identification of continuous-time nonlinear models (Young, 1981; Unbehauen & Rao, 1990; Unbehauen, 1996). The modulating functions method belongs to a class of identification techniques, which apply an integral transform to the measured signals over a finite period of time. It is an efficient approach for continuous-time system identification because it avoids the approximation of time derivatives of measured signals and does not require a knowledge of the initial conditions. The

method was first suggested by Shinbrot (1957) as a means of parameter estimation in nonlinear dynamic systems by converting the original differential equations into a set of algebraic equations. Several versions of the method are currently in use with a variety of modulating functions, including spline type functions (Maletinsky, 1979; Preisig & Rippin, 1993a), Hermite functions (Takaya, 1968; Jalili, Jordan & Mackie, 1992), Poisson moment functionals (Saha & Rao, 1982), Hartley modulating functions (Patra & Unbehauen, 1995; Daniel-Berhe & Unbehauen, 1998), sinusoidal functions (Shinbrot, 1957; Pearson & Lee, 1985; Benhadj-Braiek & Rotella, 1990; Co & Ydstie, 1990; Co & Ungarala, 1997) and wavelets (Schoenwald, 1993; Carrier & Stephanopoulos, 1998).

Although the modulating functions method has been available in several versions for over three decades, surprisingly few applications can be found in the literature. Applications are almost nonexistent in chemical and biochemical engineering disciplines, which offer a rich variety of identification problems involving physically meaningful time-varying parameters. Catalyst deactivation is a prime example. Catalyst activity is generally a complex function of time, temperature and other physico-chemical properties. It cannot be measured directly and the relevant kinetic parameters must be inferred from measurements of state variables such as temperature and species concentrations (Cheng, Abi & Kershenbaum, 1996; Tatiraju & Soroush, 1998). In large industrial processes like refineries, tracking catalyst deactivation rate constants in the fluid catalytic cracking process offers substantial economic advantages (Ancheyta-Juárez, López-Isunza, Anguilar-Rodríguez & Moreno-Mayorga, 1997). Measurement of heat transfer coefficients, changes in which may be early indications of fouling, is another potential application (Chen & Yuan, 1996). In microbial growth processes, the specific growth rate and yield factors are generally time-varying in nature, which need to be estimated for control and optimization purposes (Bastin & Dochain, 1986).

Applications of trigonometric and spline type modulating functions have been reported earlier for parameter estimation in nonlinear CSTR and observer schemes in batch reactors (Preisig, 1988, 1989; Co & Ydstie, 1990; Preisig & Rippin, 1993b). Recently reported applications include the estimation of kinetic parameters and reaction order (Ungarala & Co, 1996; Co & Zhong, 1997; Zhong, 1997), parameter estimation in microbial growth systems (Ungarala & Co, 1998a,b), and Hartley modulating function for bilinear models of electrical and robotic systems (Daniel-Berhe & Unbehauen, 1997, 1999).

Time-varying parameters have been dealt with previously by assuming a functional form for the time dependency of the parameter and obtaining the estimates of the constant coefficients in the functional form (Puchkov & Chinayev, 1973; Benhadj-Braiek & Rotella, 1990; Preisig & Rippin, 1993a). The modulating functions method is essentially an off-line method that is implemented on a batch of time series data to estimate constant parameters. The possibility for real-time applications of the method was suggested by Pearson and Lee (1985) through sequential processing of data. The main bottleneck for on-line estimation of time-varying parameters is the large amount of computation required. To address the issue of indeterminate and time-varying parameter estimation, in an earlier paper (Co & Ungarala, 1997) the authors developed a batch scheme recursive implementation of the Pearson–Lee method. The approach avoids the use of FFT at every sampling time by using efficient matrix update relations. Operating as a filter on a data window, the batch scheme introduces a lag and estimates the parameters as average values for the batch. The present work focuses on adapting the batch scheme recursive method for estimating parameters as polynomials of arbitrary degree. This approach eliminates the lag introduced by the windowing operation.

In the following sections the theoretical development of the modulating functions method is presented for time-varying system identification by modeling the parameters as polynomial splines and a sliding window recursive implementation of the method is derived for on-line parameter estimation. The recursion relations use shifting properties and commutative ring properties of trigonometric functions. A batch of time series data are processed at each sampling instant as the data window is moved forward. The estimation algorithm is fast and computationally efficient as it does not require sequential use of FFT. The method addresses the issue of on-line implementation of the modulating functions method. An application of the on-line estimation algorithm is illustrated with a simulation example involving a continuous-time model of a fermentation processes with endogenous metabolism. Finally, a few related computational issues are discussed.

2. Modulating functions method

A modulating function $\phi(t) \in C^K$, defined over a finite time interval [0, T], satisfies the following terminal conditions,

$$\mathrm{d}^k\phi(0) = 0 = \mathrm{d}^k\phi(T) \quad k = 0, 1, \ldots, K-1 \tag{1}$$

where d^k is the differentiation operator $\mathrm{d}/\mathrm{d}t$ of order k. A function $f(t) \in L^1$ over [0, T] is modulated by taking the inner product with a modulating function $\phi(t)$,

$$\langle f, \phi \rangle = \int_0^T f\phi \, \mathrm{d}t \tag{2}$$

The terminal constraints of Eq. (1) essentially make the boundary conditions of the function $f(t)$ irrelevant after modulation. The following adjoint property transfers the differentiation operation from the function $f(t)$ on to the modulating function $\phi(t)$, thereby eliminating the need to approximate time derivatives from noisy measurement data.

$$\langle \mathrm{d}^k f, \phi \rangle = (-1)^k \langle f, \mathrm{d}^k \phi \rangle \tag{3}$$

Pearson and Lee (1985) utilized sinusoidal functions to construct orthogonal sets of modulating functions and took advantage of FFT/DFT methods. A family of r trigonometric modulating functions $\Phi(t)$, are constructed as linear combinations of cosine and sine functions of the first $1+L$ frequency modes,

$$\Phi(t) = CF(\omega, t) \tag{4}$$

$$F(\omega, t) = (1, \cos \omega t, -\sin \omega t, \ldots, \cos L\omega t, -\sin L\omega t)^t \tag{5}$$

where L is an integer valued filtering parameter and $\omega = 2\pi/T$. The combination matrix C, of size $r \times (1+2L)$, is formulated to satisfy the terminal conditions set in Eq. (1) and guarantees the linear independence of the r modulating functions. There are several procedures for obtaining C, including the one by Maletinsky (1979), Pearson and Lee (1985) who used Vandermonde matrix identities and an explicit formulation given in Co and Ungarala (1997).

Sinusoidal functions are smooth and infinitely differentiable, which facilitates the use of a simple block diagonal matrix operator to perform differentiation on trigonometric modulating functions,

$$\mathrm{d}^k \Phi(t) = (-1)^k C \mathbf{D}^k F(\omega, t) \tag{6}$$

$$\mathbf{D} = \omega \mathrm{diag}(0, r, 2r, \ldots, Lr); \qquad r = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{7}$$

Consider a general nonlinear differential equation possessing the affine structure shown below (Co & Zhong, 1997),

$$\sum_{j=1}^{p} \mathrm{d}^{k_j} \zeta_j(t, u, y) = \sum_{j=1}^{q} \alpha_j(t) \mathrm{d}^{k_j} \psi_j(t, u, y) \tag{8}$$

where ζ_j and ψ_j are nonlinear functions of time t, input u, and output y, which do not contain any unknown parameters and α_j are the unknown time-varying parameters to be determined. For $k_q = 0$ and $\psi_q = 1$, the parameter α_q serves as a bias term if the process model requires one. Modulation of Eq. (8) with a set of linearly independent modulating functions converts it into a set of algebraic equations,

$$\sum_{j=1}^{p} \langle \Phi, \mathrm{d}^{k_j} \zeta_j \rangle = \sum_{j=1}^{q} \langle \Phi, \alpha_j \mathrm{d}^{k_j} \psi_j \rangle \tag{9}$$

where Φ is a column vector of r modulating functions of order $K = \max\{k_j\}$, which is the order of the system. The number of parameters to be determined is q, hence the number of linearly independent algebraic equations generated must be at least equal to q, i.e. $r \geq q$.

The unknown parameters α_j are time-dependent, consequently they cannot be moved out of the integral sign,

$$\int_0^T \phi \alpha_j(t) \, \mathrm{d}^{k_j} \psi_j \, \mathrm{d}t \neq \alpha_j(t) \int_0^T \phi \, \mathrm{d}^{k_j} \psi_j \, \mathrm{d}t \tag{10}$$

In cases where a model of the parameter's functionality with time and process variables is previously determined, the embedded parameters belonging to the function can be estimated, provided the differential equation remains linear in parameters (Unbehauen & Rao, 1990; Preisig & Rippin, 1993a). Such an approach is limited to fixed and predetermined functional forms of time dependencies. It may not be suitable for processes in which previous knowledge of the behavior of the parameter is not available or when the process moves into previously unknown fault modes or when a failure event causes drastic changes in the process.

The trajectory of the unknown time-varying parameter may be modeled by a piecewise polynomial spline of appropriate degree over a finite time interval and the constant coefficients of the polynomial can be estimated subsequently. Benhadj-Braiek and Rotella (1990) combined the modulating functions method with polynomial approximations for off-line identification of vectorial time-variant differential equations. They considered time-varying parameter matrices comprising elements of polynomial functions and modulating functions of the kind $t^n \psi$, where ψ is a trigonometric function. Zhong (1997) employed the method to estimate rate parameters in catalyst deactivation kinetics. Obtaining a polynomial fit to a parameter over a finite time interval facilitates the tracking of the behavior of the parameter with time and eliminates the complexities associated with establishing the functionality of the parameter in terms of other variables. Once the trends in parameters are accurately estimated, the factors influencing the parameter's time dependency can be investigated to find a model for the behavior of the parameter.

Let us assume that a time-varying parameter $\alpha(t)$, can be modeled by a polynomial spline of degree M on a finite time interval (Ungarala & Co, 1996),

$$\alpha = a_0 + a_1 \tau + a_2 \tau^2 + \ldots + a_M \tau^M \quad (0 \leq \tau \leq T) \tag{11}$$

It is evident from the structure of the differential equation (Eq. (8)), that the unknown parameters multiply pure derivatives of a function of the input and output signals, in other words, the nonlinearities are integrable. A polynomial spline model of the parameters allows us to extract the coefficients of the polyno-

mial out to multiply the derivatives of a new set of functions and preserve the structure of the differential equation. Lemma 1 provides the basis for rearranging the product of a parameter and the derivative of a nonlinear function in the original differential equation.

Lemma 1. *Let the functions $f(t)$ and $g(t)$, belonging to a class of functions C^k, be defined over $[0, T]$, then*

$$f\mathrm{d}^k g = \sum_{i=0}^{k} (-1)^i \binom{k}{i} \mathrm{d}^{k-i}(g\,\mathrm{d}^i f) \tag{12}$$

Proof: see Appendix A.

Lemma 2 makes use of Lemma 1 on the inner product $\langle \phi,\ \alpha \mathrm{d}^k \psi \rangle$, to move the coefficients of the polynomial out of the inner product.

Lemma 2. *If $\alpha(t)$ is approximated as a polynomial spline in τ of degree M on $[0, T]$, modulation is applied to a new set of functions $\mathrm{d}^{k-1}(\psi\tau^{n-1})$ instead of $\tau^n \mathrm{d}^k\psi$,*

$$\langle \phi, \alpha \mathrm{d}^k \psi \rangle = \sum_{n=0}^{M} a_n \sum_{i=0}^{\min\{k,n\}} \beta_{k,i,n} \langle \phi, \mathrm{d}^{k-i}(\psi \tau^{n-i}) \rangle \tag{13}$$

where a_n are the coefficients of the polynomial spline and,

$$\beta_{k,i,n} = (-1)^i \frac{k!n!}{i!(k-i)!(n-i)!} \tag{14}$$

Proof: see Appendix B

To illustrate the use of Lemma 2, consider the Van der Pol oscillator,

$$\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} = a \frac{\mathrm{d}y}{\mathrm{d}t}(1 - by^2) - cy \tag{15}$$

where the parameters a and b are time-varying in nature. Eq. (15) can be transformed into affine form as shown below,

$$\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} = \alpha_1 \frac{\mathrm{d}y}{\mathrm{d}t} + \alpha_2 \frac{\mathrm{d}(y^3)}{\mathrm{d}t} + \alpha_3 y \tag{16}$$

where the parameters α_j are combinations of a, b and c. Suppose α_1 is modeled by a linear polynomial spline, α_2 is modeled by a quadratic spline and α_3 is a constant over a given time interval $0 \le \tau \le T$,

$$\frac{\mathrm{d}^2 y}{\mathrm{d}\tau^2} = (a_{10} + a_{11}\tau)\frac{\mathrm{d}y}{\mathrm{d}\tau} + (a_{20} + a_{21}\tau + a_{22}\tau^2)\frac{\mathrm{d}(y^3)}{\mathrm{d}\tau} + a_{30} y \tag{17}$$

Applying the result of Lemma 2,

$$\frac{\mathrm{d}^2 y}{\mathrm{d}\tau^2} = a_{10}\frac{\mathrm{d}y}{\mathrm{d}\tau} + a_{11}\frac{\mathrm{d}(\tau y)}{\mathrm{d}\tau} - y \quad + a_{20}\frac{\mathrm{d}(y^3)}{\mathrm{d}\tau} + a_{21}\frac{\mathrm{d}(\tau y^3)}{\mathrm{d}\tau} - y^3 \quad + a_{22}\frac{\mathrm{d}(\tau^2 y^3)}{\mathrm{d}\tau} - 2\tau y^3 \quad + a_{30} y \tag{18}$$

Using Pearson–Lee method with trigonometric modulating functions, the constant coefficients a_{ij}, in the above equation can be obtained by least squares (Co & Ungarala, 1997). See Appendix C for the extension of Pearson–Lee method for a general differential equation in affine form with q unknown parameters modeled as polynomial splines of different degrees.

3. On-line implementation

The parameter estimation as outlined in the previous section is implemented on a batch of time series data spanning a time interval $[0, T]$. Pearson and Lee (1985) suggested the applicability of the method to on-line identification by sequentially solving the least squares estimation problem for $[0, T]$, $[\Delta t, T + \Delta t]$, $[2\Delta t, T + 2\Delta t]$ and so on. Although the idea is conceptually appealing, in practice it requires a large amount of computation because FFT is used at every sampling instant and the data sets over the interval of duration T are required to be of length 2^m.

Pearson–Lee method, like the other versions of the modulating functions method, is an off-line procedure despite possessing significant advantages over many system identification techniques. Co and Ungarala (1997) developed computationally efficient recursion formulae for Pearson–Lee method applied to sequential estimation of time-varying parameters. However, the procedure introduced a lag of $T/2$ time duration in the estimates. The lag is a direct consequence of the assumption that the parameters are constant inside a given interval $[k\Delta t, T + k\Delta t]$, thus yielding average values of the parameters over the length of the interval.

In this section we will generalize the treatment of time varying parameter estimation based on the polynomial spline models for parameters introduced in the previous section. A set of sliding window recursion relations with simple matrix update equations will be constructed that do not require the use of FFT for sequential estimation. A batch of data spanning T time duration is processed at each sampling instant to determine the constant coefficients of the parameter polynomials.

At the current sampling instant k, the integrals $C^n_{m,k}(f)$ and $S^n_{m,k}(f)$ resulting from the modulation of $\tau^n f(\tau)$ with trigonometric functions are,

$$C^n_{m,k}(f) = \int_0^T \tau^n f_{(k\Delta t + \tau - T)} \cos\left(\frac{2\pi}{T} m\tau\right) \mathrm{d}\tau \tag{19}$$

$$S^n_{m,k}(f) = -\int_0^T \tau^n f_{(k\Delta t + \tau - T)} \sin\left(\frac{2\pi}{T} m\tau\right) \mathrm{d}\tau \tag{20}$$

These integrals can be numerically computed using a quadrature formula such as the rectangular rule, the trapezoidal rule or the Simpson's rule. Co and Ungarala (1997) contains a comparison between rectangular and trapezoidal methods. We will use the trapezoidal rule to approximate $C^n_{m,k}(f)$ and $S^n_{m,k}(f)$ as shown below,

$$C^n_{m,k}(f) \approx (\Delta t)\bar{C}^n_{m,k}(f) \tag{21}$$

$$S^n_{m,k}(f) \approx (\Delta t)\bar{S}^n_{m,k}(f) \tag{22}$$

where,

$$\bar{C}^n_{m,k}(f) = \sum_{j=0}^{N-1} (j\Delta t)^n f_{(j+k-N)} \cos\left(\frac{2\pi}{N} mj\right) + \frac{1}{2}\gamma^n_{N,k}(f) \tag{23}$$

$$\bar{S}^n_{m,k}(f) = -\sum_{j=0}^{N-1} (j\Delta t)^n f_{(j+k-N)} \sin\left(\frac{2\pi}{N} mj\right) \tag{24}$$

$$\gamma^n_{N,k}(f) = (N\Delta t)^n f_{(k)} - (0)^n f_{(k-N)} \tag{25}$$

with $\gamma^0_{N,k}(f) = f(k) - f_{(k-N)}$ when $n = 0$ and $\gamma^n_{N,k}(f) = (N\Delta t)^n f_{(k)}$ for $n > 0$.

Using shifting properties of trigonometric functions, efficient recursion formulae can be developed for evaluating $\bar{C}^n_{m,k}$ and $\bar{S}^n_{m,k}$ for successive sampling instants. Lemma 3, which forms the basis for online implementation of the extended Pearson–Lee method, is a matrix equation that relates $\bar{C}^n_{m,k-1}$ and $\bar{S}^n_{m,k-1}$ to $\bar{C}^n_{m,k}$ and $\bar{S}^n_{m,k}$, respectively. The essence of the lemma is as follows, if the integrals are computed at the present sampling instant k with N data points from the past, successive values of the integrals are calculated by using the next available sample and discarding the oldest sample. Thus the recursion formulae allow us to evaluate the integrals sequentially for $[\Delta t,\ T+\Delta t]$, $[2\Delta t,\ T+2\Delta t]$ and so on as a window of N data points is moved one sampling time forward, without using FFT.

Lemma 3. *A matrix equation relates* $\bar{C}^n_{m,k-1}(f)$ *and* $\bar{S}^n_{m,k-1}(f)$ *to* $\bar{C}^n_{m,k}(f)$ *and* $\bar{S}^n_{m,k}(f)$, *respectively, by shifting a window of N data points one sample time forward*:

$$\begin{matrix}\bar{C}^n_{m,k}(f) \\ \bar{S}^n_{m,k}(f)\end{matrix} = \theta_m \left[\begin{pmatrix} \sum_{l=0}^n h_{n,l}\bar{C}^l_{m,k-1}(f) \\ \sum_{l=0}^n h_{n,l}\bar{S}^l_{m,k-1}(f) \end{pmatrix} + \begin{pmatrix} \frac{1}{2}\delta^n_{N,k-1}(f) \\ 0 \end{pmatrix} \right] + \begin{pmatrix} \frac{1}{2}\gamma^n_{N,k}(f) \\ 0 \end{pmatrix} \tag{26}$$

where,

$$\theta_m = \begin{matrix} \cos((2\pi/N)m) & -\sin((2\pi/N)m) \\ \sin((2\pi/N)m) & \cos((2\pi/N)m) \end{matrix} \tag{27}$$

$$h_{n,l} = \begin{matrix} n \\ l \end{matrix} (-\Delta t)^{n-l} \tag{28}$$

$$\delta^n_{N,k-1}(f) = [(N-1)\Delta t]^n f_{(k)} - (-\Delta t)^n f_{(k-N)} \tag{29}$$

Proof: see Appendix D

A similar version of the moving discrete Fourier transform was used for model reduction in Lilly (1991). Lemma 3 can be used to formulate a straight forward matrix update equation to recursively compute $Z^n_{L,k}(f)$ which is a matrix of the integrals $C^n_{m,k}(f)$ and $S^n_{m,k}(f)$ for successive k as in the following Corollary 4. Define the following update matrices with respect to the edges of a data window of length N.

1. Update matrix for a new data point

$$\Delta_k(f) = (\delta^0_{N,k}(f), \ldots, \delta^n_{N,k}(f)) = f_{(k)}\eta_n - f_{(k-N)}\eta^0_n \tag{30}$$

$$\eta_n = (1, (N-1)\Delta t, \ldots, \{(N-1)\Delta t\}^n) \tag{31}$$

$$\eta^0_n = (1, -\Delta t, \ldots, (-\Delta t)^n) \tag{32}$$

2. Update matrix for trapezoidal rule

$$\Gamma_k(f) = (\gamma^0_{N,k}(f), \ldots, \gamma^n_{N,k}(f)) = f_{(k)}\rho_n - f_{(k-N)}\rho^0_n \tag{33}$$

$$\rho_n = (1, N\Delta t, \ldots, (N\Delta t)^n) \tag{34}$$

$$\rho^0_n = (1, 0, \ldots, 0^n) \tag{35}$$

Corollary 4. *Matrix formulation for recursive update of* $Z^n_{L,k}(f)$

$$Z^n_{L,k}(f) = \Theta\ Z^n_{L,k-1}(f)H_n + \frac{1}{2} w\Delta_{k-1}(f) + \frac{1}{2} w\Gamma_k(f) \tag{36}$$

where,

$$Z^n_{L,k}(f) = \begin{pmatrix} C^0_{0,k}(f) & C^1_{0,k}(f) & \ldots & C^n_{0,k}(f) \\ C^0_{1,k}(f) & C^1_{1,k}(f) & \ldots & C^n_{1,k}(f) \\ S^0_{1,k}(f) & S^1_{1,k}(f) & \ldots & S^n_{1,k}(f) \\ & & \ddots & \vdots \\ C^0_{L,k}(f) & C^1_{L,k}(f) & \ldots & C^n_{L,k}(f) \\ S^0_{L,k}(f) & S^1_{L,k}(f) & \ldots & S^n_{L,k}(f) \end{pmatrix} \tag{37}$$

$$\Theta = \mathrm{diag}(1, \theta_1, \ldots, \theta_L) \tag{38}$$

$$H_n(i,j)=\begin{cases}h_{n,i-j} & i\geq j\\ 0 & i<j\end{cases}\quad i,j=1,2,\ldots,n+1 \tag{39}$$

$$w=(1,1,0,\ldots,1,0)^t \quad \text{(vector of length } 1+2L) \tag{40}$$

The following lemma states a commutative relationship between D^k and Θ which will be used to improve the computational efficiency of the algorithm.

Lemma 5. *Let* $\Theta=diag(1,\theta_1,\ldots,\theta_L)$ *and* $\mathbf{D}=\omega\, diag(0, r, 2r, \ldots, Lr)$ *then the following commutativity relationship holds true,*

$$\mathbf{D}^k\Theta=\Theta\mathbf{D}^k \tag{41}$$

Proof: see Co and Ungarala (1997)

Lemma 6. *The recursion formulae to sequentially solve the least squares problem resulting from the modulation of the affine differential equation with polynomial spline models for time-varying parameters are:*

$$G_P(k)=\Theta\left\{G_P(k-1)H+\frac{1}{2}V_P\odot(w\Delta_P(k-1))\right\}+\frac{1}{2}V_P\odot(w\Gamma_P(k)) \tag{42}$$

$$G_Q(k)=\Theta\left\{G_Q(k-1)H+\frac{1}{2}V_Q\odot(w\Delta_Q(k-1))\right\}+\frac{1}{2}V_Q\odot(w\Gamma_Q(k)) \tag{43}$$

where $\odot$ *is the Haddamard product and*

$$G_P(k)=(\mathbf{D}^{k_1}Z^0_{L,k}(\zeta_1),\ \mathbf{D}^{k_2}Z^0_{L,k}(\zeta_2),\ \ldots,\ \mathbf{D}^{k_p}Z^0_{L,k}(\zeta_p)) \tag{44}$$

$$G_Q(k)=(\Psi_1,\Psi_2,\ldots,\Psi_q) \tag{45}$$

$$\Psi_j=(\mathbf{D}^{k_j}Z^{M_j}_{L,k}(\psi_j),\ \mathbf{D}^{k_j-1}Z^{M_j-1}_{L,k}(\psi_j),\ \ldots,\ \mathbf{D}^{k_j-\mu_j}Z^{M_j-\mu_j}_{L,k}(\psi_j)) \tag{46}$$

$$V_P=(\mathbf{D}^{k_1}w,\ldots,\mathbf{D}^{k_p}w) \tag{47}$$

$$V_Q=(\mathbf{D}^{k_1}w,\ldots,\mathbf{D}^{k_q}w) \tag{48}$$

$$H=\text{diag}(\cdots,H_{M_j},\ldots,H_{M_j-\mu_j},\ldots) \tag{49}$$

$$\Delta_P(k)=w(\Delta_k(\zeta_1),\ \Delta_k(\zeta_2),\ \ldots,\ \Delta_k(\zeta_p)) \tag{50}$$

$$\Delta_Q(k)=w(\Delta_k(\psi_1),\ \Delta_k(\psi_2),\ \ldots,\ \Delta_k(\psi_q)) \tag{51}$$

$$_P(k)=w(\ _k(\zeta_1),\ \ _k(\zeta_2),\ \ldots,\ \ _k(\zeta_p)) \tag{52}$$

$$_Q(k)=w(\ _k(\psi_1),\ \ _k(\psi_2),\ \ldots,\ \ _k(\psi_q)) \tag{53}$$

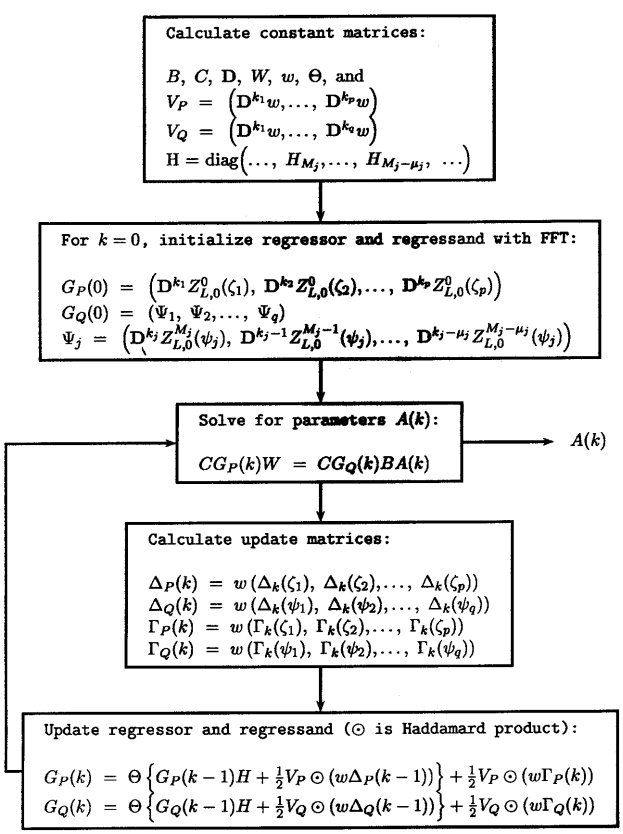

Fig. 1. Algorithm of on-line identification of no-stationary continuous-time systems.

Proof: use Corollary 4

Fig. 1 summarizes the on-line implementation of Pearson–Lee modulating functions method with spline models of time-varying parameters. For a given window size N, filtering parameter L and spline models for the parameters, the constant matrices C, B, Θ, $\mathbf{D}$, W, w, V_P, V_Q and H are calculated. The on-line identification algorithm is initialized by computing the regressor and regressand using Pearson–Lee method with FFT and solving the resulting least squares problem for the polynomial coefficients. The regressor and regressand are updated for the next available sample by calculating the update matrices. Thus, the least squares problem is sequentially solved at each sampling instant to obtain the parameter vector $A(k)$.

4. Bio-process model identification

On-line identification of microbial growth processes has been receiving increasing attention in recent years, primarily due to the growing interest in the application of modern systems theory to bio-process control and optimization schemes. Modeling and identification of bio-processes are known to be the main bottlenecks in the design of bio-process control methodologies. Apart from being highly nonlinear in nature, bio-processes are generally non-stationary. Variability in model parame-

ters may be induced by the sensitivity and selectivity exhibited by microorganisms to their environmental conditions.

The typical unstructured model of a fermentation process comprises a set of mass balance equations of the state variables coupled with one or more kinetic rate models. A multitude of analytical expressions are available to represent the complex relationships between kinetic rate parameters and the state variables (Bastin & Dochain, 1990). Regardless of the type of model employed, the model parameters are often assumed to have biological significance and are treated as characteristic of the process, although the identification of a unique set of physically meaningful parameters may not be feasible for many systems. While any such interpretation is limited by the validity of the model, changes in model parameters may indicate physicochemical changes inside the process (Holmberg & Ranta, 1982). The rational design of advanced control and optimization schemes for bio-processes is therefore intricately linked to the identification of kinetic rate parameters accurately and in real-time (Van Impe & Bastin, 1995; Thatipamala, Hill & Rohani, 1996).

The problems of identifiability of kinetic rate models such as the Monod law with limited and noisy experimental data have been well documented (Nihtilä & Virkkunen, 1977; Holmberg, 1982; Andrews, 1984; Stephanopoulos & San, 1984). Several methodologies are currently employed for on-line identification of the kinetic rate models. Adaptive estimation techniques employing recursive least squares (RLS) methods have been frequently used to track the Monod law parameters. Aborhey and Williamson (1978) reported the estimation of Monod parameters and yield coefficient as functions of culture temperature and Golden and Ydstie (1989) used recursive least squares with forgetting factors. The RLS algorithms suggested in the literature are adaptations of schemes developed for time-invariant systems and they fare poorly when used for monitoring parameters that drift with time, besides being very sensitive to noise (Holmberg & Ranta, 1982). Estimation of time varying parameters in a linearized discrete fermentation model are also reported in the literature (Zhou & Cluett, 1996). However, the dynamic behavior of biological systems cannot be adequately described by a linear model and the parameter drifts of the linearized discrete model may not correspond to changes in actual physical parameters.

In general, the functionality of the kinetic rate parameters with respect to process variables are inadequately known. Hence, it is difficult to define and validate these functions. The analytical expressions available for kinetic rate parameters are limited in scope because they are often system specific and extensions to wider operating regions may not be feasible (Stephanopoulos & San, 1984). Motivated by the need to find estimation schemes that are independent of kinetic rate models, several researchers proposed nonlinear adaptive schemes for estimating the kinetic rates themselves as time-varying parameters. The extended Kalman filter (EKF) is widely applied to this effect (Stephanopoulos & San, 1984; Shimizu, Takamatsu, Shioya & Suga, 1989). The EKF is derived from the application of Kalman filter to the linearized model of a nonlinear system. These algorithms may result in biased estimates if they are not properly initialized, and analyzing the stability and convergence properties of the EKF schemes over wide ranges of operation is known to be extremely difficult.

Bastin and Dochain (1986) developed an adaptive observer to estimate the unmeasurable state variables and kinetic rate parameters on-line. Despite the simple structure of the proposed approach, the tuning of these estimators has been known to be difficult (Oliveira, Ferreira, Oliveira & Feyo de Azevedo, 1996; Farza, Busawon & Hammouri, 1998). Assuming that the oxygen uptake rate (OUR) data and its derivative are available on-line, Lubenova (1996) suggested an adaptive algorithm for the estimation of specific growth rate. Although these algorithms are simple in structure and their stability is proven, they tend to be highly sensitive to initial guesses and the OUR data from off gas analysis can have a high signal-to-noise ratio, particularly in animal cell cultures where the oxygen demand can be extremely small (Singh, 1996).

4.1. Fermentation model

The growth dynamics of a general single organism-single rate limiting substrate fermentation process in a chemostat are described by the following set of mass balance equations with endogenous metabolism and maintenance requirements included (Blanch & Clarck, 1996),

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \mu x - Dx - K_{\mathrm{e}}x \tag{54a}$$

$$\frac{\mathrm{d}s}{\mathrm{d}t} = D(s_{\mathrm{i}} - s) - \frac{\mu}{Y}x - mx \tag{54b}$$

where x is the concentration of the biomass, s is the concentration of the rate limiting substrate, D is the rate of dilution, s_{i} is the concentration of limiting substrate in the nutrient feed, μ is the specific growth rate, Y is the biomass on limiting substrate yield factor and K_{e} and m are endogenous metabolism rate constant and maintenance coefficient, respectively. Additional mass balance equations are used to account for the formation of synthesis products, if any.

The process parameters in the model μ, Y, K_{e} and m represent the state of growth of the biomass and substrate consumption in the chemostat. The specific

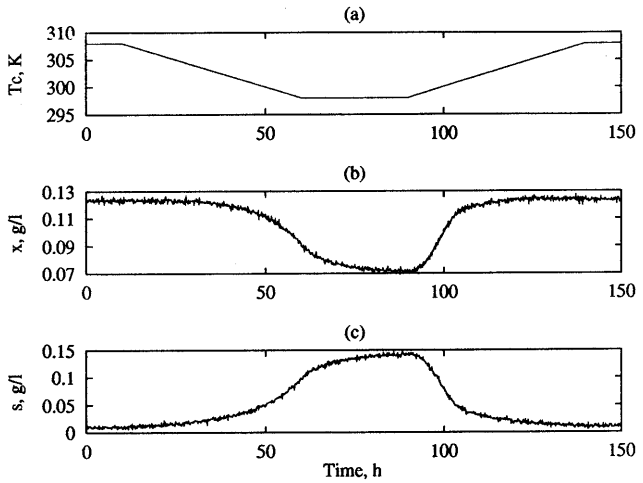

Fig. 2. Simulation data (a) culture temperature, (b) biomass concentration and (c) limiting substrate concentration.

growth rate is one of the most important parameters of the process because maximum μ is synonymous with maximum biomass cultivation. In order to optimize the biomass production, the specific growth rate needs to be observed and controlled. Generally, μ is a nonlinear function of the state variables, environmental conditions, operating conditions and other physico-chemical factors such as temperature, pH, light intensity and cell age. In practice, μ is not directly measurable and thus, has to be estimated from measured data such as biomass and substrate concentrations. One of the most simple and frequently used expressions for the specific growth rate is the Monod law,

$$\mu = \frac{\mu^* s}{K_s + s} \tag{55}$$

where μ is represented as a function of the substrate concentration only, μ^* is the maximum growth rate and K_s is the half saturation constant which is termed as a measure of the organism's affinity to the substrate. The functionality of μ with respect to s is not important for the asymptotic solution of the model Eqs. (54a) and (54b), as long it is a monotonously increasing function. The Monod law fails to predict some experimentally observed phenomena like transient oscillations and hysteresis in biomass growth (Tang, Sitomer & Jackson, 1997). Assuming that the Monod law is valid, Topiwala and Sinclair (1971) studied the effects of temperature on specific growth rate and found that the Monod law parameters μ^* and K_s are strong functions of temperature.

The yield coefficient Y relates the amount of biomass produced per unit of limiting substrate consumed. A knowledge of Y is essential for the formulation of the nutrient feed and to make the desired substrate rate limiting. Cells consume some substrate to provide energy for maintaining chemical gradients, transport of nutrients across the cell membrane and DNA repair and replication. Maintenance energy requirements are accounted for in the substrate mass balance, Eq. (54b). Cells may also consume part of their own mass for providing maintenance energy, termed as endogenous metabolism, which is accounted for in the cell mass balance, Eq. (54a), Both the endogenous metabolism rate K_e, and the maintenance coefficient m, vary with biomass growth rate, temperature and pH (Topiwala & Sinclair, 1971). Hence, the observed yield coefficient Y_{obs} exhibits a strong dependency on growth rate and environmental variables. The potential yield Y, remains a constant for a wide range of operating conditions until, at high temperatures and extreme pH values, the cells begin to die.

4.2. Simulation example

Consider a fermentation process characterized by the state space model given in Eqs. (54a) and (54b). Let us assume that the 'true' specific growth rate is governed by the Monod law. Furthermore, the Monod parameters μ^* and K_s and the endogenous metabolism rate K_e, will be assumed to be governed by the Topiwala–Sinclair model for the organism *Aerobacter aerogenes* with glucose as the rate limiting substrate as shown below,

$$\mu^* = 2.45 \times 10^{10} e^{-14230/RT_c} - 1.39 \times 10^{23} e^{-32900/RT_c} \tag{56}$$

$$\frac{1}{K_s} = 2.96 \times 10^{10} e^{-11800/RT_c} \tag{57}$$

$$K_e = 2.71 \times 10^{5} e^{-9000/RT_c} \tag{58}$$

where T_c is the culture temperature. The maintenance coefficient $m = 0$ and the yield coefficient is assumed constant at $Y = 0.5$ in the range $25°C \le T_c \le 45°C$.

The model equations were integrated for 150 h of chemostat operation with a trapezoidal signal for the culture temperature T_c, as shown in Fig. 2a.

$$T_c(t) = \begin{cases} 35°C & t \le 10 \\ 35 - 0.2(t-10) & 10 < t < 60 \\ 25°C & 60 \le t \le 90 \\ 25 + 0.2(t-90) & 90 < t < 140 \\ 35°C & t \ge 140 \end{cases} \tag{59}$$

The operating conditions and initial values for integration are, $D = 0.65$ per h, $s_i = 0.3$ g/l and $x_0 = 0.1235$ g/l, $s_0 = 0.0107$ g/l. The output sequence shown in Fig. 2b and c was generated with a sampling time $\Delta t = 0.15$ h. The biomass concentration x, was corrupted with white noise of Gauss (0, 0.001) and the substrate concentration s, was corrupted with Gauss (0, 0.002).

The first step towards using modulating functions method is to represent the model equations in the prescribed affine form. Eq. (54a) is rearranged to obtain

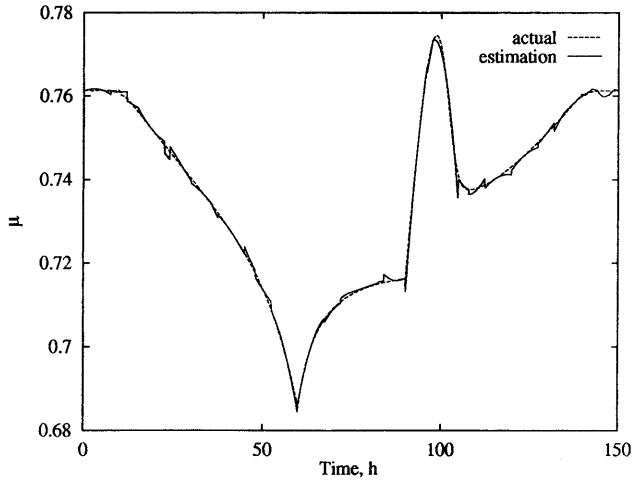

Fig. 3. Estimation of specific growth rate μ.

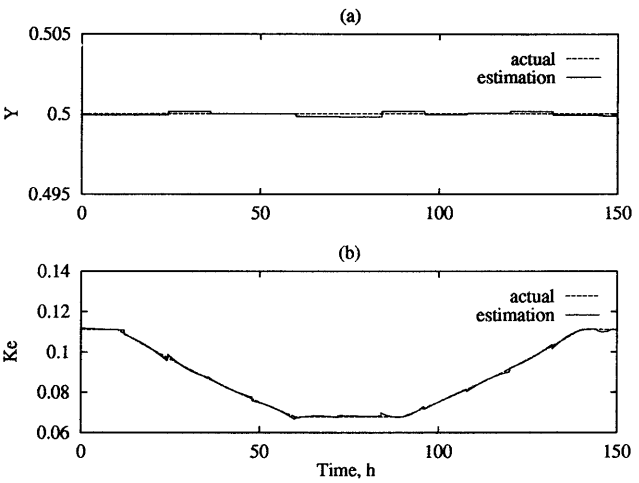

Fig. 4. Estimation of (a) yield coefficient Y and (b) endogenous metabolism rate constant K_e.

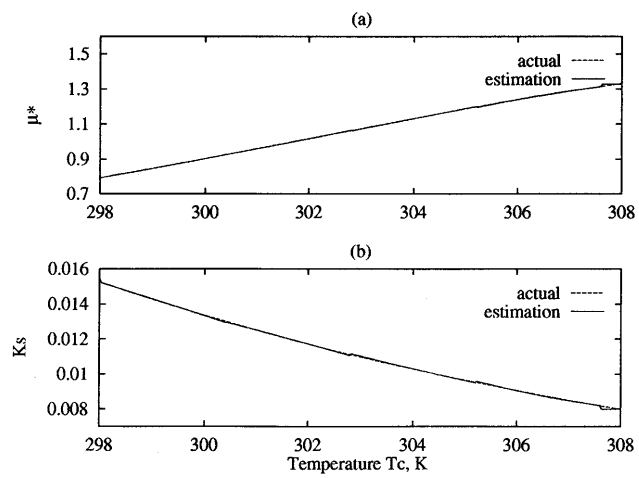

Fig. 5. Estimation of Monod law parameters (a) maximum specific growth rate μ^* and (b) half saturation constant K_s as functions of temperature.

an affine form (Eq. (60b)) for the estimation of $(\mu - K_e)$. Elimination of μ from Eqs. (54a) and (54b) results in the second affine equation (Eq. (60b)) for the estimation of Y and K_e.

$$\frac{dx}{dt} + Dx = (\mu - K_e)x \tag{60a}$$

$$\frac{dx}{dt} + Dx = Y\ D(s_i - s) - \frac{ds}{dt}\ - K_e x \tag{60b}$$

The parameter $(\mu - K_e)$ was estimated by assuming that it is modeled by a quadratic polynomial spline inside a data window of $N = 50$ samples using a filtering parameter $L = 3$. From the second affine form Eq. (60b), parameters Y and K_e were estimated by assuming that Y is a constant and K_e is approximated by a quadratic spline with $N = 80$ and $L = 6$. As the data window moves forward at each sampling time, the parameters are given by evaluating the polynomial spline models at the leading edge ($\tau = T$) of the sliding data window. Once the window slides over and processes N samples, the previous N estimates of the parameters were smoothed out over the length of the window by their respective spline models. The smoothing eliminates the small errors caused by the spline models at the edges of the data window. The resulting estimate of the specific growth rate μ as a function of time is shown in Fig. 3. Estimates of the yield coefficient Y, and endogenous metabolism rate K_e are shown in Fig. 4. The variations in the parameters were tracked faithfully with good accuracy. Note that the parameter $(\mu - K_e)$ was estimated using the biomass concentration only; in the absence of endogenous metabolism, the estimation of specific growth rate does not require the substrate concentration measurements.

The estimates of μ and K_e were obtained without using the models of their dependencies on time and temperature. Thus, the specific growth rate can be estimated as a time-varying parameter, independent of the analytical expression chosen to model it. Nonetheless, in cases where μ is known to be modeled by the Monod law fairly well, the Monod parameters μ^* and K_s can also be estimated on-line. Substituting Monod law into the substrate mass balance Eq. (54b), results in an affine equation from which both μ^* and K_s are readily computed.

$$\frac{1}{2}\frac{ds^2}{dt} - Ds(s_i - s) = K_s\ \frac{ds}{dt} - D(s_i - s)\ \ - \frac{\mu^*}{Y} sx \tag{61}$$

Fig. 5 shows the results of the estimation of μ^* and K_s, plotted against the culture temperature T_c, with quadratic spline approximations, $N = 80$ and $L = 6$. The Arrhenius parameters in the Topiwala–Sinclair models of Eqs. (56)–(58), can be estimated subsequently by employing a batch modulating functions method or other nonlinear optimizers, utilizing the

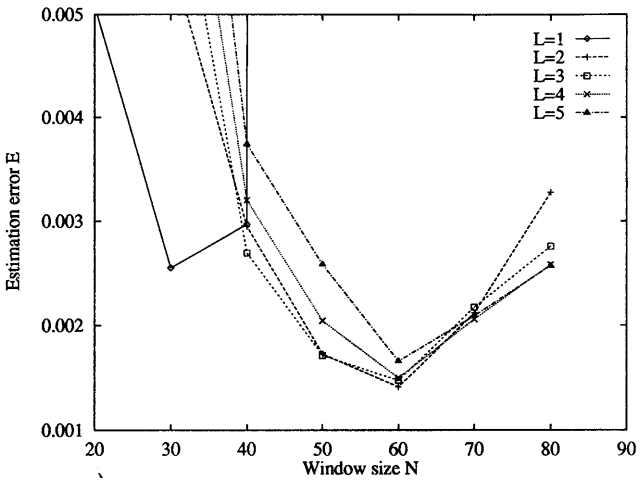

Fig. 6. Effect of window size N and filtering parameter L on the estimation error for the estimation of $\mu - K_e$ modeled by a quadratic spline ($M = 2$).

smooth estimates of μ^*, K_s, K_e and T_c measurements. Pearson–Lee method was used on Eqs. (57) and (58) after taking the logarithm on both sides to express them in affine form. A batch of 70 estimates of K_s and K_e and the corresponding T_c (noise-free) was used to compute the Arrhenius constant and activation energy for the half saturation constant (K_s), $A_1 = 2.96 \times 10^{10}$, $E_1 = 11800$, and for the endogenous metabolism rate (K_e), $A_2 = 2.71 \times 10^5$, $E_2 = 9000$. Eq. (56) cannot be converted to affine form, hence, the Levenberg–Marquardt method was used with initial guesses (15000, 30000). The fitted values of the Arrhenius parameters of maximum specific growth rate (μ^*) are $A_3 = 2.45 \times 10^{10}$, $A_4 = -1.39 \times 10^{23}$ and $E_3 = 14230$, $E_4 = 32900$. These are the exact values used in the simulation.

5. Computational issues

In the derivation of the recursion formulae for the online implementation of Pearson–Lee method, it was assumed that signals from a continuous system are sampled at a uniform rate. The sampling time interval Δt, plays an important role in parameter identification schemes (Sinha & Puthenpura, 1985). If the process is sampled too fast or Δt is too small, the computational burden on the algorithm increases, which is not desirable especially in the case of on-line identification. A general rule of thumb for selecting the sampling time is, $\Delta t < 0.1\tau_{max}$ where τ_{max} is the dominant time constant of the process. Another related factor in terms of computational load is the variation of the parameters with time. If the parameters are changing relatively slowly, even though samples are available, the least squares problem need not be solved for every Δt. The parameters may be estimated after N updates to the recursion formulae by using least squares once every N samples, which in turn reduces the computational load. The resulting estimates are the same as the smoothed out estimates mentioned above. It should be noted that the computational burden is independent of the length of the data window.

The algorithm contains three critical variables to be selected by the user, the window size $T = N\Delta t$, the degree of polynomial model M_j and the filtering parameter L. In order to assess the influence of these variables on the accuracy of the estimates obtained, define the following expression for estimation error,

$$E = \sqrt{\frac{1}{T_e}\int_0^{T_E} [\alpha(t)_{true} - \alpha(t)_{est}]^2 \, dt} \tag{62}$$

where α is the parameter and T_e is the total time of the estimation.

As a general rule, the window should be sufficiently wide to contain the characteristic nonlinearity information of the system. If the window is too small, the data may not be sufficient to distinguish the system dynamics, which will result in large fluctuations around true parameters. A small window is usually suitable for parameters that vary slowly enough so that they can be treated as essentially constants or at best ramp functions inside the window. On the other hand a wider window will require increased complexity of the parameter's functionality with time.

Simulations were performed for different values of N, L and M for the estimation of $(\mu - K_e)$. The results are summarized in Fig. 6; there exists an optimum window width, which adequately captures the dynamics of the process, the estimation error passes through a minimum as the window size increases. This feature is consistent with other recursive moving window estimation methods. The filtering parameter L must be chosen such that the terminal conditions for the modulating functions are not violated. A general rule is $2L \geq K$, the order of the system (Pearson & Lee, 1985). By definition, modulation results in the filtering of signals through the Fourier series coefficients. If L is too large, the higher frequency noise accompanying the signals receives emphasis resulting in deteriorated estimates. The negative effects of a large L can be compensated for by choosing a wider window, but a large window may require a more involved expression to approximate the time dependency of the parameter.

In general, parameter trends inside a window that captures the dynamics adequately, can be approximated by second, third or at most fourth degree polynomials; using higher degrees may result in errors because of the tendency of piecewise splines to stray away near the ends of the interval. Fig. 7 shows the effect of the degree of polynomial approximation M. A zero degree assumption shifts the parameter estimates by a delay proportional to the window size, since the estimates are

assumed constant and are averaged over the length of the window (Co & Ungarala, 1997). On the other hand, using increasingly higher degree results in undesirable oscillations around the true parameters at the edges of the data window.

6. Conclusions

The modulating function methods offer two distinct advantages over other identification methods, they allow for arbitrary initial conditions and avoid the approximation of time derivatives from noisy signals. The Pearson–Lee method is a computationally efficient off-line estimation scheme that employs trigonometric modulating functions and takes advantage of fast Fourier transform techniques. However, the method was not suitable for on-line application because of the considerable computational burden it imposes with the use of FFT for every sampling instant. In this paper, the Pearson–Lee method was extended for the estimation of time-varying parameters by approximating them as polynomial splines over finite time intervals. A sliding window recursive estimation algorithm is derived to sequentially determine the coefficients of the parameter model polynomials as a data window of fixed length moves forward at each sampling time. The algorithm is initiated with the Pearson–Lee formulation using FFT on a batch of data and then the regressor and regressand are updated by a simple matrix update equation. The algorithm is derived in a matrix formulation so that efficient matrix manipulation abilities of commercial software programs like MATLAB can be utilized.

The recursive estimation algorithm was utilized for time dependent parameter tracking in a fermentation process model. The specific growth rate and endogenous metabolism rate were estimated as time-varying parameters without utilizing the associated kinetic rate models. The estimates yielded were accurate and faithful to their true variations with time in the presence of noise. The estimation of the embedded parameters in Monod law and Arrhenius parameters is also illustrated. The simulation study concludes that model independent estimation of specific growth rate and yield factor in bio-processes facilitates further investigations into possible correlations between parameters and the various factors that influence them.

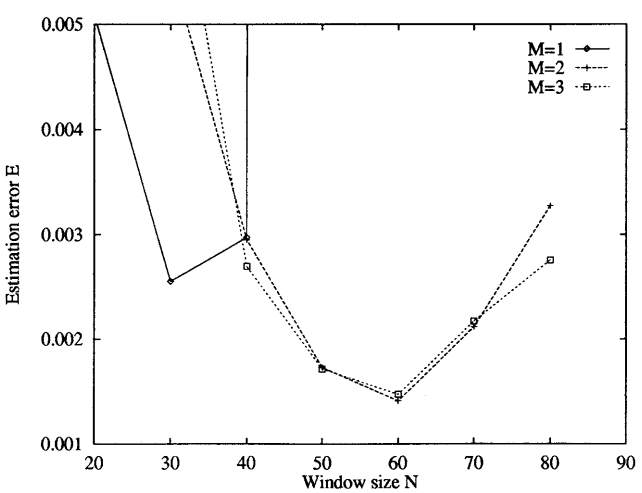

Fig. 7. Effect of the degree of spline model on the estimation error for the estimation of $\mu - K_e$ with $L = 3$.

The effects of three important user selected variables in the estimation algorithm were investigated. It was found that an optimum data window size exists to capture the dynamics information adequately and the estimation error is dependent on the filtering parameter and the degree of polynomial used for modeling the parameters. The recursive version of the modulating functions method was shown to be fast and computationally efficient for real time implementation. It proves to be a practical tool for time dependent parameter tracking and will be a good alternative to existing methods of parameter estimation in continuous time systems.

Finally, it must be noted that at present the window size T, is fixed for the algorithm. It would be desirable to have an adaptive window size that varies with the frequency content of the signals so that a small window size can be employed at high frequencies and larger window sizes at low frequencies. Schoenwald (1993), Carrier and Stephanopoulos (1998) proposed a wavelet modulating function to address the issue. Future research will be centered around implementing a variable window size, including forgetting factors to discount past data and approximation of the parameters with adaptive degrees of polynomials.

7. Nomenclature

Latin symbols

C	combination matrix
$\bar{C}^n_{k,m}$	summation as defined in Eq. (23)
D	dilution rate, per h
D	matrix differentiation operator
K_e	endogenous metabolism rate g/l
K_s	half saturation constant, per h
L	filtering parameter
M	degree of polynomial
N	number of data points in [0, T]
$\bar{S}^n_{k,m}$	summation as defined in Eq. (24)
R	gas constant, cal/K mole
T	size of the sliding window
T_c	culture temperature, K
Y	yield factor
$Z^n_{L,k}$	matrix of summations defined in Eq. (37)
a_j	coefficients of polynomials
c_m	$\cos(2\pi m/N)$

d	differentiation operator $\mathrm{d}/\mathrm{d}t$ of order k_j
r	matrix rotation operator
s_m	$\sin(2\pi m/N)$
s	substrate concentration, g/l
s_i	inlet substrate concentration, g/l
x	biomass concentration, g/l

Greek symbols

$_P$, $_Q$	update matrices
Δ_P,Δ_Q	update matrices
$_k$	difference vector
Δ_k	difference vector
Δt	sampling time
Φ	family of r modulating functions of order K
α_j	unknown parameter
^n_N	difference defined in Eq. (25)
^n_N	difference defined in Eq. (29)
ζ	nonlinear function of u, y, and t
η_n,η^0_n	vectors defined in Eqs. (31) and (32)
θ	matrix defined in Eq. (27)
μ	specific growth rate, per h
μ^*	maximum specific growth rate, per h
ρ_n, ρ^0_n	vectors defined in Eqs. (34) and (35)
ϕ	modulating function
ψ	nonlinear function of u, y, and t

Appendix A. Lemma 1

Proof (by induction):

$$f\,\mathrm{d}g = \mathrm{d}(gf) - g\,\mathrm{d}f \tag{A.1}$$

satisfies the lemma for $k=1$. Consider the differentiation of $f\,\mathrm{d}^k g$,

$$f\,\mathrm{d}^{k+1}g = \mathrm{d}(f\,\mathrm{d}^k g) - \mathrm{d}f\,\mathrm{d}^k g \tag{A.2}$$

Suppose the lemma is true for any k. Differentiating the lemma yields,

$$\mathrm{d}(f\,\mathrm{d}^k g) = \sum_{i=0}^{k}(-1)^i \begin{matrix} k \\ i \end{matrix} \mathrm{d}^{k-i+1}(g\,\mathrm{d}^i f) \tag{A.3}$$

Applying the lemma to $\mathrm{d}f\,\mathrm{d}^k g$,

$$\mathrm{d}f\,\mathrm{d}^k g = \sum_{i=0}^{k}(-1)^i \begin{matrix} k \\ i \end{matrix} \mathrm{d}^{k-i}(g\,\mathrm{d}^{i+1}f) = \sum_{i=1}^{k+1}(-1)^{i-1} \begin{matrix} k \\ i-1 \end{matrix} \mathrm{d}^{k-i+1}(g\,\mathrm{d}^i f) \tag{A.4}$$

Substituting Eqs. (A.3) and (A.4) into Eq. (A.2),

$$f\,\mathrm{d}^{k+1}g = \sum_{i=1}^{k}(-1)^i \left[\begin{pmatrix} k \\ i \end{pmatrix} + \begin{pmatrix} k \\ i-1 \end{pmatrix} \right] \mathrm{d}^{k-i+1}(g\,\mathrm{d}^i f) + (-1)^0 \begin{matrix} k \\ 0 \end{matrix} \mathrm{d}^{k+1}(g\,\mathrm{d}^0 f) + (-1)^{k+1} \begin{matrix} k \\ k \end{matrix} \mathrm{d}^0(g\,\mathrm{d}^{k+1}f) \tag{A.5}$$

but,

$$\begin{matrix} k \\ i \end{matrix} + \begin{matrix} k \\ i-1 \end{matrix} = \begin{matrix} k+1 \\ i \end{matrix} \quad \text{and} \quad \begin{matrix} k \\ 0 \end{matrix} = \begin{matrix} k \\ k \end{matrix} = 1 \tag{A.6}$$

Thus the lemma is valid for $k+1$,

$$f\,\mathrm{d}^{k+1}g = \sum_{i=0}^{k+1}(-1)^i \begin{matrix} k+1 \\ i \end{matrix} \mathrm{d}^{k+1-i}(g\,\mathrm{d}^i f) \tag{A.7}$$

Appendix B. Lemma 2

Proof: using Lemma 1on $\alpha\,\mathrm{d}^k\psi$ with,

$$\alpha = a_0 + a_1\tau + a_2\tau^2 + \ldots + a_M\tau^M \tag{B.1}$$

$$\alpha\,\mathrm{d}^k\psi = \sum_{i=0}^{k}(-1)^i\frac{k!}{i!(k-i)!}\mathrm{d}^{k-i}(\psi\mathrm{d}^i\alpha) = \sum_{i=0}^{k}(-1)^i \frac{k!}{i!(k-i)!}\mathrm{d}^{k-i}\ \psi\sum_{n=0}^{M}a_n\frac{n!}{(n-i)!}\tau^{n-i} = \sum_{j=0}^{M}a_n\sum_{i=0}^{\min\{k,n\}}(-1)^i\frac{k!n!}{i!(k-i)!(n-i)!}\mathrm{d}^{k-i}(\psi\tau^{n-i}) \tag{B.2}$$

Note that,

$$\mathrm{d}^i\tau^n = \begin{cases} \dfrac{n!}{(n-i)!}\tau^{n-i} & n \geq i \\ 0 & n < i \end{cases} \tag{B.3}$$

Modulation of Eq. (B.2) results in,

$$\langle\phi,\ \alpha\mathrm{d}^k\psi\rangle = \sum_{n=0}^{M}a_n\sum_{i=0}^{\min\{k,n\}}\beta_{k,i,n}\langle\phi,\ \mathrm{d}^{k-i}(\psi\tau^{n-i})\rangle \tag{B.4}$$

where,

$$\beta_{k,i,n} = (-1)^i\frac{k!n!}{i!(k-i)!(n-i)!}$$

Appendix C. Spline models and Pearson–Lee method

The general nonlinear differential equation in the affine form,

$$\sum_{j=1}^{p} \mathrm{d}^{k_j}\zeta_j(t, u, y) = \sum_{j=1}^{q} \alpha_j \mathrm{d}^{k_j}\psi_j(t, u, y) \tag{C.1}$$

If k_j is the order of d operator on ψ_j and M_j is the degree of the spline model for α_j, then define $\mu_j = \min\{k_j, M_j\}$. The nonlinear functions ζ_j and ψ_j of the affine equation can be expanded and grouped into row vectors P and Q, respectively,

$$P = (\mathrm{d}^{k_1}\zeta_1,\ \mathrm{d}^{k_2}\zeta_2, \ldots,\ \mathrm{d}^{k_p}\zeta_p) \tag{C.2}$$

$$Q = (\bar{\psi}_1,\ \bar{\psi}_2, \ldots,\ \bar{\psi}_q) \tag{C.3}$$

where,

$$\bar{\psi}_j = [\mathrm{d}^{k_j}(\psi_j),\ \mathrm{d}^{k_j}(\tau\psi_j), \ldots,\ \mathrm{d}^{k_j}(\tau^{M_j}\psi_j),\ \mathrm{d}^{k_j-1}(\psi_j),\ \mathrm{d}^{k_j-1}(\tau\psi_j), \ldots,\ \mathrm{d}^{k_j-1}(\tau^{M_j-1}\psi_j), \ldots,\ \mathrm{d}^{k_j-\mu_j}(\psi_j),\ \mathrm{d}^{k_j-\mu_j}(\tau\psi_j), \ldots,\ \mathrm{d}^{k_j-\mu_j}(\tau^{M_j-\mu_j}\psi_j)] \tag{C.4}$$

Define a matrix B such that,

$$B = \begin{matrix} b_1 & 0 & \cdots & 0 \\ 0 & b_2 & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & b_q \end{matrix} \qquad b_j = \begin{matrix} \lambda_0 \\ \lambda_1 \\ \\ \lambda_{\mu_j} \end{matrix} \tag{C.5}$$

$$\lambda_i = 0_{(M_j - i + 1)\times i} \begin{matrix} \beta_{k_j,i,i} & 0 & \cdots & 0 \\ 0 & \beta_{k_j i,i+1} & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & \beta_{k_j,i,M_j} \end{matrix} \tag{C.6}$$

The affine equation is now represented in a matrix formulation,

$$PW = (QB)A \tag{C.7}$$

where W is a column vector of ones of length p and A is a column vector of length $(q + \Sigma_1^q M_j)$ containing the coefficients of the polynomial splines which are the new parameters to be estimated.

$$A = (a_{1,0}, \ldots, a_{1,M_1}, \ldots, a_{q,0}, \ldots, a_{q,M_q})^t \tag{C.8}$$

Modulation with a set of r trigonometric modulating functions Φ, yields a set of r linear algebraic equations.

$$\langle \Phi, P\rangle W = \langle \Phi, Q\rangle BA \tag{C.9}$$

Note that the size of the parameter vector has increased and consequently, $r \geq (q + \Sigma_1^q M_j)$. The parameter estimation problem is now reduced to a linear regression and the parameter vector A is obtained by using regular least squares method.

$$(CG_P W) = (CG_Q B)A \tag{C.10}$$

where,

$$G_P = (D^{k_1}Z_L^0(\zeta_1), D^{k_2}Z_L^0(\zeta_2), \ldots, D^{k_p}Z_L^0(\zeta_p)) \tag{C.11}$$

$$G_Q = (\Psi_1, \Psi_2, \ldots, \Psi_q) \tag{C.12}$$

$$\Psi_j = (D^{k_j}Z_L^{M_j}(\psi_j), D^{k_j-1}Z_L^{M_j-1}(\psi_j), \ldots, D^{k_j-\mu_j}Z_L^{M_j-\mu_j}(\psi_j)) \tag{C.13}$$

$$Z_L^n(f) = \begin{matrix} C_0^0(f) & C_0^1(f) & \ldots & C_0^n(f) \\ C_1^0(f) & C_1^1(f) & \ldots & C_1^n(f) \\ S_1^0(f) & S_1^1(f) & \ldots & S_1^n(f) \\ & & \ddots & \\ C_L^0(f) & C_L^1(f) & \ldots & C_L^n(f) \\ S_L^0(f) & S_L^1(f) & \ldots & S_L^n(f) \end{matrix} \tag{C.14}$$

$$C_m^n(f) = \int_0^T \tau^n f \cos m\omega\tau \, \mathrm{d}\tau \tag{C.15}$$

$$S_m^n(f) = -\int_0^T \tau^n f \sin m\omega\tau \, \mathrm{d}\tau \tag{C.16}$$

Appendix D. Lemma 3

Proof: define the following differences,

$$\Delta_{N,k}^n(f) = (N\Delta t)^n f_{(k)} - (0)^n f_{(k-N)} \tag{D.1}$$

$$\Delta_{N,k}^n(f) = \{(N-1)\Delta t\}^n f_{(k)} - (-\Delta t)^n f_{(k-N)} \tag{D.2}$$

The integral $C_{m,k}^n(f)$ is numerically approximated with the trapezoidal rule as follows,

$$\begin{aligned}
\bar{C}_{m,k}^n(f) &= \sum_{j=0}^{N-1} (j\Delta t)^n f_{(j+k-N)} \cos\frac{2\pi}{N}mj + \frac{1}{2}\Delta_{N,k}^n(f) \\
&= \sum_{j=1}^{N} (j-1)^n (\Delta t)^n f_{(j+k-1-N)} \\
&\quad \cos\frac{2\pi}{N}m(j-1) + \frac{1}{2}\Delta_{N,k}^n(f) \\
&= \sum_{j=1}^{N}\sum_{l=0}^{n} \binom{n}{l} j^l (-1)^{n-l} (\Delta t)^n f_{(j+k-1-N)} \\
&\quad \left[\cos\frac{2\pi}{N}mj \cos\frac{2\pi}{N}m + \sin\frac{2\pi}{N}mj \sin\frac{2\pi}{N}m\right] + \frac{1}{2}\Delta_{N,k}^n(f) \\
&= c_m \sum_{l=0}^{n} h_{n,l} \sum_{j=1}^{N} (j\Delta t)^l f_{(j+k-1-N)} \cos\frac{2\pi}{N}mj \\
&\quad + s_m \sum_{l=0}^{n} h_{n,l} \sum_{j=1}^{N} (j\Delta t)^l f_{(j+k-1-N)} \sin\frac{2\pi}{N}mj + \frac{1}{2}\Delta_{N,k}^n(f)
\end{aligned} \tag{D.3}$$

where,

$$h_{n,l} = \binom{n}{l}(-\Delta t)^{n-l} \qquad c_m = \cos\frac{2\pi}{N}m$$

$$s_m = \sin\frac{2\pi}{N}m \tag{D.4}$$

$$\bar{C}^n_{m,k}(f) = c_m \sum_{l=0}^{n} h_{n,l} \sum_{j=0}^{N-1} (j\Delta t)^l f_{(j-k-1-N)} \cos\frac{2\pi}{N}mj$$

$$+ (N\Delta t)^l f_{(k-1)} - (0)^l f_{(k-1-N)}] + s_m \sum_{l=0}^{n} h_{n,l}$$

$$\sum_{j=0}^{N-1} (j\Delta t)^l f_{(j-k-1-N)} \sin\frac{2\pi}{N}mj \quad + \frac{1}{2}\,{}^{n}_{N,k}(f) \tag{D.5}$$

but,

$$\bar{S}^l_{m,k-1}(f) = -\sum_{j=0}^{N-1} (j\Delta t)^l f_{(j+k-1-N)} \sin\frac{2\pi}{N}mj$$

$$\bar{C}^n_{m,k}(f) = c_m \sum_{l=0}^{n} h_{n,l} \sum_{j=0}^{N-1} (j\Delta t)^l f_{(j-k-1-N)} \cos\frac{2\pi}{N}mj$$

$$+ \frac{1}{2}((N\Delta t)^l f_{(k-1)} - (0)^l f_{(k-1-N)}) \quad + \frac{1}{2} c_m$$

$$\sum_{l=0}^{n} \binom{n}{l} (-\Delta t)^{n-l} (N\Delta t)^l f_{(k-1)} - \sum_{l=0}^{n} \binom{n}{l}$$

$$(-\Delta t)^{n-l} (0)^l f_{(k-1-N)} \quad - s_m \sum_{l=0}^{n} h_{n,l} \bar{S}^l_{m,k-1} + \frac{1}{2}\,{}^{n}_{N,k}(f)$$

$$= c_m \sum_{l=0}^{n} h_{n,l} \bar{C}^l_{m,k-1} + \frac{1}{2} c_m\,{}^{n}_{N,k-1}(f) - s_m \sum_{l=0}^{n} h_{n,l} \bar{S}^l_{m,k-1}$$

$$+ \frac{1}{2}\,{}^{n}_{N,k}(f) \tag{D.6}$$

Similarly for the integral $S^n_{m,k}(f)$,

$$\bar{S}^n_{m,k}(f) = -\sum_{j=0}^{N-1} (j\Delta t)^n f_{(j+k-N)} \sin\frac{2\pi}{N}mj$$

$$= s_m \sum_{l=0}^{n} h_{n,l} \bar{C}^l_{m,k-1} + \frac{1}{2} s_m\,{}^{n}_{N,k-1}(f) + c_m \sum_{l=0}^{n} h_{n,l} \bar{S}^l_{m,k-1} \tag{D.7}$$

References

Aborhey, S., & Williamson, D. (1978). State and parameter estimation of microbial growth processes. *Automatica*, *14*, 493–498.

Ancheyta-Juárez, J., López-Isunza, F., Anguilar-Rodríguez, E., & Moreno-Mayorga, J. C. (1997). A strategy for kinetic parameter estimation in the fluid catalytic cracking process. *Industrial & Engineering Chemistry Research*, *36*, 5170–5174.

Andrews, G. F. (1984). Parameter estimation from batch culture data. *Biotechnology and Bioengineering*, *26*, 824–825.

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

Bastin, G., & Dochain, D. (1990). *On-line estimation and adaptive control of bioreactors*. Amsterdam: Elsevier.

Benhadj-Braiek, E., & Rotella, F. (1990). Identification of non-stationary continuous systems using modulating functions. *Journal of Franklin Institute*, *327*, 831–840.

Blanch, H. W., & Clarck, D. S. (1996). *Biochemical Engineering*. New York: Marcel Dekker.

Carrier, J. F., & Stephanopoulos, G. (1998). Wavelet-based modulation in control-relevant process identification. *American Institute of Chemical Engineering Journal*, *44*, 341–360.

Chen, Z., & Yuan, W.-K. (1996). Simultaneous estimation of kinetic and heat transfer parameters of a wall-cooled fixed-bed reactor. *Chemical Engineering Science*, *51*, 4791–4800.

Cheng, Y. S., Abi, C. F., & Kershenbaum, L. S. (1996). On-line estimation for a fixed-bed reactor with catalyst deactivation using nonlinear programming techniques. *Computers and Chemical Engineering*, *20*, S793–S798.

Co, T. B., & Ydstie, B. E. (1990). System identification using modulating functions method and fast Fourier transforms. *Computers and Chemical Engineering*, *14*, 1051–1066.

Co, T. B., & Ungarala, S. (1997). Batch scheme recursive parameter estimation of continuous-time systems using the modulating functions method. *Automatica*, *33*, 1185–1191.

Co, T. B., & Zhong, K. (1997). Lie transformations for parameter estimation of continuous-time nonlinear systems. *Proceedings of the ACC*, *5*, 3053–3057.

Daniel-Berhe, S., & Unbehauen, H. (1997). Physical parameter estimation of the nonlinear dynamics of a single link robotic manipulator with flexible joint using HMF-method. *Proceedings of the ACC*, *3*, 1504–1508.

Daniel-Berhe, S., & Unbehauen, H. (1998). Bilinear continuous-time systems identification via Hartley based modulating functions. *Automatica*, *34*, 301–318.

Daniel-Berhe, S., & Unbehauen, H. (1999). Physical parameters estimation of the nonlinear continuous-time dynamics of a DC motor using Hartley modulating functions method. *Journal of Franklin Institute*, *336*, 481–501.

Farza, M., Busawon, K., & Hammouri, H. (1998). Simple nonlinear observers for on-line estimation of kinetic rates in bioreactors. *Automatica*, *34*, 301–318.

Feng, W., Genceli, H., & Nikolaou, M. (1996). Constrained model predictive control with simultaneous identification using wavelets. *Computers and Chemical Engineering*, *20*, S1011–S1016.

Font, J., & Fabregat, A. (1997). Testing a predictor–corrector integral method for estimating parameters in complex kinetic systems described by ordinary differential equations. *Computers and Chemical Engineering*, *21*, 719–731.

Golden, M. P., & Ydstie, B. E. (1989). Adaptive extremum control using approximate process models. *American Institute of Chemical Engineers Journal*, *35*, 1157–1169.

Goodwin, G. C., & Mayne, D. Q. (1987). A parameter estimation perspective of continuous time model reference adaptive control. *Automatica*, *23*, 57–70.

Hecker, O., Knapp, T., & Isermann, R. (1994). Robust adaptive control of a time varying process using parallel recursive estimators. *Automatica*, *30*, 599–604.

Höfling, T., & Isermann, R. (1996). Fault detection based on adaptive parity equations and single-parameter tracking. *Control Engineering Practice*, *4*, 1361–1369.

Holmberg, A. (1982). On the practical identifiability of microbial growth models incorporating Michaelis–Menten type nonlinearities. *Journal of Mathamatical Bioscience*, *62*, 23–43.

Holmberg, A., & Ranta, J. (1982). Procedures for parameter and state estimation of microbial growth process models. *Automatica*, *18*, 181–193.

Iserman, R. (1984). Process fault detection based on modeling and estimation methods — a survey. *Automatica*, *20*, 387–404.

Jalili, S. A., Jordan, J. R., & Mackie, R. D. L. (1992). Measurement of the parameters of all-pole transfer functions using shifted Hermite modulating functions. *Automatica*, *28*, 613–616.

Lilly, J. H. (1991). Efficient DFT-based model reduction for continuous systems. *IEEE Transactions on Automatic Control*, *AC-36*, 1188–1193.

Lubenova, V. (1996). On-line estimation of biomass concentration and non stationary parameters for aerobic bioprocesses. *Journal of Biotechnology*, *46*, 197–207.

Maletinsky, V. (1979). Identification of continuous dynamical systems with spline type modulating functions method. In R. Iserman, *Identification and system parameter estimation*. Oxford: Pergamon Press.

Nihtilä, M., & Virkkunen, J. (1977). Practical identifiability of growth and substrate consumption models. *Biotechnology and Bioengineering*, *19*, 1831–1850.

Oliveira, R., Ferreira, E. C., Oliveira, F., & Feyo de Azevedo, S. (1996). A study on the convergence of observer-based kinetics estimators in stirred tank bioreactors. *Journal of Process Control*, *6*, 367–371.

Patra, A., & Unbehauen, H. (1995). Identification of a class of nonlinear continuous-time systems using Hartley modulating functions. *International Journal of Control*, *62*, 1431–1451.

Pearson, A. E., & Lee, F. C. (1985). On the identification of polynomial input–output differential equations. *IEEE Transactions on Automatic Control*, *AC-30*, 778–782.

Preisig, H. A. (1988). The use of differential information for batch reactor control. *Proceedings of the ACC*, 671–676.

Preisig, H. A. (1989). On-line observation of the composition in non-isothermal batch recators with non-linear reactions. *Proceedings of the ACC*, 1549–1554.

Preisig, H. A., & Rippin, D. W. T. (1993a). Theory and applications of the modulating functions method-I. Review and theory of the method and theory of the spline-type modulating functions. *Computers and Chemical Engineering*, *17*, 1–16.

Preisig, H. A., & Rippin, D. W. T. (1993b). Theory and applications of the modulating functions method-III. Application to industrial process, a well-stirred tank reactor. *Computers and Chemical Engineering*, *17*, 29–39.

Puchkov, V. F., & Chinayev, P. I. (1973). The method of sliding modulation functions for identifying plants with delay. *Soviet Automatic Control*, *6*, 5–11.

Saha, D. C., & Rao, G. P. (1982). A general algorithm for parameter identification in lumped continuous systems. *IEEE Transactions on Automatic Control*, *AC-27*, 223–225.

Schoenwald, D. A. (1993). System identification using a wavelet-based approach. *Proceedings of the IEEE Decision and Control*, *4*, 3064–3065.

Shimizu, H., Takamatsu, T., Shioya, S., & Suga, K.-I. (1989). An algorithmic approach to constructing the on-line estimation system for the specific growth rate. *Biotechnology and Bioengineering*, *33*, 354–363.

Shinbrot, M. (1957). On the analysis of linear and nonlinear systems. *Transactions of the American Society of Mechanical Engineers*, *79*, 547–552.

Singh, V. (1996). On-line measurement of oxygen uptake in cell culture using the dynamic method. *Biotechnology and Bioengineering*, *52*, 443–448.

Sinha, N. K., & Puthenpura, S. (1985). Choice of the sampling interval for the identification of continuous-time systems from samples of input/output data. *IEE Proceedings*, *132-D*, 231–244.

Stephanopoulos, G., & San, K.-Y. (1984). Studies on on-line bioreactor identification 1: theory. *Biotechnology and Bioengineering*, *26*, 1176–1188.

Takaya, K. (1968). The use of Hermite functions for system identifications. *IEEE Transactions on Automatic Control*, *AC-13*, 446–447.

Tang, B., Sitomer, A., & Jackson, T. (1997). Population dynamics and competetion in chemostat models with adaptive nutrient uptake. *Journal of Mathamatical Bioscience*, *35*, 453–479.

Tatiraju, S., & Soroush, M. (1998). Parameter estimator design with application to chemical reactors. *Industrial & Engineering Chemistry Research*, *37*, 455–463.

Thatipamala, R., Hill, G. A., & Rohani, S. (1996). On-line estimation and adaptive optimization using state equations for continuous production of bioethonology. *Journal of Biotechnology*, *48*, 179–190.

Topiwala, H., & Sinclair, C. G. (1971). Temperature relationships in continuous culture. *Biotechnology and Bioengineering*, *13*, 795–813.

Unbehauen, H. (1996). Some new trends in identification and modeling of nonlinear dynamic systems. *Applied Mathematics and Computation*, *78*, 279–297.

Unbehauen, H., & Rao, G. P. (1990). Continuous-time approaches to system identification — a survey. *Automatica*, *26*, 23–35.

Ungarala, S., & Co, T. B. (1996). Recursive parameter estimation of time-varying nonlinear systems. *American Institute of Chemical Engineers Journal Annual Meeting*, Chicago, IL.

Ungarala, S., & Co, T. B. (1998a). Direct estimation of microbial specific growth rates. *American Institute of Chemical Engineers Journal Annual Meeting*, Miami Beach, FL.

Ungarala, S., & Co, T. B. (1998b). Model parameter tracking in microbial growth processes. *American Institute of Chemical Engineers Journal*, *44*, 2129–2134.

Van Impe, J. F., & Bastin, G. (1995). Optimal adaptive control of fed-batch fermentation processes. *Control Engineering Practice*, *3*, 939–954.

Young, P. C. (1981). Parameter estimation for continuous-time models — a survey. *Automatica*, *17*, 23–29.

Zhong, K. (1997). On-line parameter estimation of catalyst deactivation kinetics. M.S. thesis, Michigan Technological University, Houghton.

Zhou, Q., & Cluett, W. R. (1996). Recursive identification of time-varying systems via incremental estimation. *Automatica*, *32*, 1427–1431.