PROCESS DESIGN AND CONTROL

An Exponentially Weighted Moving Average Method for Identification and Monitoring of Stochastic Systems

Shyh-Hong Hwang,* Ho-Tsen Chen, and Chuei-Tin Chang

Department of Chemical Engineering, National Cheng Kung University, Tainan, Taiwan 70101, R.O.C.

To identify parametric models for stochastic systems, the standard least-squares method tends to yield biased parameter estimates owing to correlated residuals resulting from unknown stochastic disturbances. Although the consistency properties of parameter estimates could generically be secured by instrumental variable methods, the inadequate choices of instruments and prefilters would render them much less efficient. This article establishes a method to identify an ARARX (AutoRegressive AutoRegressive with eXogenous input), an ARMAX (AutoRegressive Moving Average with eXogenous input), or a BJ (Box–Jenkins) model based on the process output data smoothed by the EWMA (Exponentially Weighted Moving Average). The major advantages of the method are 2-fold. First, the proposed off-line and online algorithms often acquire unbiased, efficient, and consistent parameter estimation from identification tests operating in open loop or closed loop. Second, the resultant process plus disturbance model can be easily employed to remove the autocorrelation in process data for accurate statistical process monitoring. Monte-Carlo simulation studies demonstrate that the proposed method provides reliable parametric models for a wide variety of noise characteristics and is highly robust with respect to the sampling period, sample size, and noise-to-signal ratio.

1. Introduction

Chemical and industrial processes are often subject to stochastic disturbances or noises that influence process outputs randomly. With the advances in process monitoring and control techniques, it is required that models for stochastic systems be acquired in a more accurate and efficient manner.[1−5] Discrete system identification is a useful means for this purpose, despite the fact that most physical systems are continuous in nature. A great number of identification methods have been available for finding discrete-time parametric models in noisy situations.[6−10]

The reliability of parametric identification methods can be evaluated by their statistical properties such as bias, efficiency, and consistency.[6,11] The standard least-squares (LS) method, a special case of the prediction error method, describes the stochastic system by an ARX (AutoRegressive with eXogenous input) model structure. It can then be shown that the residual is uncorrelated (an independent sequence), rendering the least-squares estimator (LSE) unbiased and consistent. On the other hand, when the actual noise characteristics violate the ARX assumption, the residual becomes autocorrelated and the LSE would be severely biased.

Clarke[12] presented the generalized least-squares (GLS) method to deal with the case of correlated residuals. The method assumes an ARARX (AutoRegressive AutoRegressive with eXogenous input) model structure in which the residual is characterized by an AR noise filter. Hence, if the choice of the noise filter is adequate, the residual constructed from the filtered input and output signals becomes uncorrelated. The GLS algorithm can be accomplished in an off-line iterative or online recursive manner[13,14] and belongs in the class of methods based on the whitening of the prediction error.[10] Other methods include the extended least-squares method and the maximum likelihood method.

The IV (Instrumental Variable) methods mitigate the effect of correlated residuals by creating new instrumental variables that are highly correlated with the uncontaminated process variables but uncorrelated with the noise disturbances.[7,15] Söderstrom and Stöica[16] discussed and compared the consistency aspects of various IV methods. The IV methods belong in the class of methods based on the uncorrelation of the observation vector and the prediction error.[10] Other methods consist of the output error with fixed compensator method and the output error with filtered observations method.

Ljung[8] presented the four-step IV (IV4) method that is available in the commercial MATLAB software. The IV4 method is a combination of the IV and GLS algorithms. As a result, it avoids the iteration procedure and possesses better statistical properties for a wider range of noise characteristics than the GLS method. Among others, the IV4 method generically secures the consistency properties of parameter estimates as the sample size N tends to infinity. However, the parameter estimates must be efficient or reliable for finite or possibly small N since the experiment time is often limited in practice. In this regard, the GLS and IV4 methods are not reliable enough to deal with system identification under widely different noise levels and characteristics.

Although identification in open-loop operation is simpler, there exist situations where the system must be identified in closed loop. For instance, a feedback controller is already operating, and it is not possible to open the loop to acquire data for identification. However, the feedback control would introduce a correlation between the noise and the process input. This may impose a bias on many open-loop estimation techniques.[10] To overcome this problem, Landau and Zito[10] proposed several closed-loop identification methods by incor-

* To whom correspondence should be addressed. Tel: 886-6-2757575 ext. 62661. Fax: 886-6-2344496. E-mail: shhwang@mail.ncku.edu.tw.

10.1021/ie0707218 CCC: $40.75
Published on Web 10/03/2008

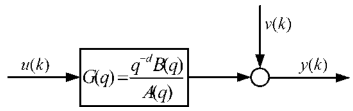

Figure 1. Noisy discrete-time system.

porating the controller design, including the closed-loop output error (CLOE) method, the filtered closed-loop output error (FCLOE) method, and the adaptive filtered closed-loop output error (AFCLOE) method.

There are conflicting issues between identification and monitoring of a stochastic system. For example, a white measurement noise induces a highly correlated residual to complicate identification, whereas it simplifies monitoring by causing uncorrelated sample data. Another issue regards the choice of the sampling period T. To enhance the ability of monitoring and/or regulating disturbances, it is advantageous to sample data as fast as possible. However, building discrete-time models with very small T compared to the time constants of the system would make the model fit concentrated to the high-frequency band and might not allow for much noise reduction.[8] It was reported that a sampling period could be selected widely from one percent to one-half of the effective process time for different control or monitoring purposes.[17,18]

In this work, an EWMA (Exponentially Weighted Moving Average) method is established to obtain reliable parameter estimates for an ARARX, an ARMAX (AutoRegressive Moving Average with eXogenous input), or a BJ (Box−Jenkins) model. The underlying estimation algorithms are based on the process output data smoothed by the EWMA. The off-line algorithm is accomplished by repetitive use of a four-step procedure. The identified process plus disturbance model could be employed to remove the autocorrelation in process data for accurate statistical process monitoring. The online algorithm entails two recursive routines for updating alternately the process parameters and the disturbance parameters. Both algorithms work well for data acquisition in open loop or closed loop. Monte-Carlo simulation has been carried out to investigate thoroughly the effects of the sampling period, sample size, and noise-to-signal ratio on parameter estimates. It is demonstrated that the proposed method can cope with a wide range of noise characteristics and is superior to the GLS, IV4, and AFCLOE methods in view of bias, efficiency, and consistency.

2. The Stochastic System Description

Consider a stochastic system as depicted in Figure 1, where $u(k)$ and $y(k)$ denote observations of process inputs and outputs in discrete time, $v(k)$ the stochastic disturbance (or noise), d the integer number of pure time delay, and $G(q)$ the process transfer function. The system can be described by the following difference equation:

$$y(k) = \frac{B(q)}{A(q)}u(k-d) + v(k) \tag{1}$$

$$A(q) = 1 + a_1q^{-1} + \cdots + a_{n_A}q^{-n_A}$$

$$B(q) = b_1q^{-1} + b_2q^{-2} + \cdots + b_{n_B}q^{-n_B}$$

where q^{-1} is the backward shift operator and n_A and n_B denote, respectively, the orders of the polynomials $A(q)$ and $B(q)$.

With the orders and delay given, the standard LS method has been developed to estimate the process model parameters a_i and b_i based on input-output observations. However, the accuracy of the standard LS method stipulates that the noise $v(k)$ should be described by the following AR structure:

$$v(k) = \frac{1}{A(q)}e(k) \tag{2}$$

where $e(k)$ is a realization of a white-noise process. Equation 1 is then expressed as an ARX model structure:

$$A(q)y(k) = B(q)u(k-d) + e(k) \tag{3}$$

It follows that the residual (or equation error) is uncorrelated.

Introduce the prediction error $\varepsilon(k,\theta)$ as a linear regression:

$$\varepsilon(k, \theta) = y(k) - \phi^T(k)\theta \tag{4}$$

where the observation vector (or regressor) $\phi(k)$ and the parameter vector θ are

$$\phi(k) = [-y(k-1) \quad \cdots \quad -y(k-n_A) \quad u(k-1-d) \quad \cdots \quad u(k-n_B-d)]^T \tag{5}$$

$$\theta = [a_1 \quad \cdots \quad a_{n_A} \quad b_1 \quad \cdots \quad b_{n_B}]^T \tag{6}$$

The least-squares estimator, $\hat{\theta}_{LS}$, can be obtained by minimizing the quadratic error criterion as

$$\hat{\theta}_{LS} = \arg\min_{\theta} \sum_k \varepsilon^2(k, \theta) = \arg\min_{\theta} \sum_k [y(k) - \phi^T(k)\theta]^2 \tag{7}$$

The above LSE is unbiased and consistent because the sequence indicated in eq 4 is zero-mean and independent. On the other hand, a large number of noise characteristics do not fall into the category of eq 2. As a result, the standard LS method would generally acquire biased parameter estimates. The GLS and IV4 methods were proposed to deal with such situations by postulating explicitly or implicitly a more complicated AR noise model as

$$v(k) = \frac{1}{A(q)C(q)}e(k) \tag{8}$$

$$C(q) = 1 + c_1q^{-1} + \cdots + c_{n_C}q^{-n_C}$$

This gives rise to an ARARX model structure:

$$A(q)y(k) = B(q)u(k-d) + \frac{1}{C(q)}e(k) \tag{9}$$

On the basis of the postulation in eq 8, the GLS and IV4 methods resolve partially the stochastic identification difficulty that the noise characteristics may vary widely in practice. Nevertheless, the two methods could be neither unbiased nor efficient if the actual noise dynamics differ significantly from eq 8. For example, the noise consists of significant zeros or is much faster than the process dynamics (such as measurement noise). Moreover, such deficiencies in the statistical properties of parameter estimates are deteriorated with an increase in the sampling rate and/or a decrease in the sample size.

3. EWMA Parameter Estimators

The EWMA is a well-known and popular statistic used for smoothing and forecasting time series.[19−21] To facilitate system identification in the face of multifarious noise characteristics, we introduce the EWMA as

$$\zeta(k) = (1-\lambda)\zeta(k-1) + \lambda y(k) \tag{10}$$

where $0 < \lambda \le 1$ is the weighting constant and the starting value (the arbitrary k_0 represents a starting instant) is

$$\zeta(k_0) = 0$$

The EWMA $\zeta(k)$ denotes a weighted average of all past and current output observations. The value of λ dictates the weight assigned to the current observation $y(k)$. When λ approaches zero, a large number of past data points are being effectively employed for calculations of the EWMA $\zeta(k)$. In the other extreme of λ approaching one, virtually no averaging is being performed and $\zeta(k)$ is essentially $y(k)$.

Substituting eq 1 into the $y(k)$ term in eq 10 and ignoring the effect of noise $v(k)$ could arrive at a linear regression equation:

$$\zeta(k) = (1-\lambda)^{(k-k_0)}\zeta(k_0) - \sum_{i=1}^{n_A} a_i \sum_{j=0}^{k-k_0-1} \lambda(1-\lambda)^j y(k-j-i) + \sum_{i=1}^{n_B} b_i \sum_{j=0}^{k-k_0-1} \lambda(1-\lambda)^j u(k-j-i-d) \quad (11)$$

With an appropriate λ, the EWMA-LS estimator for the parameters a_i and b_i is given by

$$\hat{\theta}_{EWMA}^{LS} = [\sum_{k=k_0+1}^{k_0+N} \varphi(k)\varphi^T(k)]^{-1} \sum_{k=k_0+1}^{k_0+N} \varphi(k)\zeta(k) \quad (12)$$

where $\zeta(k)$ is computed by eq 10. The regressor

$$\varphi(k) = [-Y_1(k) \quad \cdots \quad -Y_{n_A}(k) \quad U_1(k) \quad \cdots \quad U_{n_B}(k)]^T \quad (13)$$

can be obtained by means of the direct formulas:

$$Y_i(k) = \sum_{j=0}^{k-k_0-1} \lambda(1-\lambda)^j y(k-j-i) \quad (14a)$$

$$U_i(k) = \sum_{j=0}^{k-k_0-1} \lambda(1-\lambda)^j u(k-j-i-d) \quad (14b)$$

or the recursive formulas with $Y_i(k_0) = U_i(k_0) = 0$:

$$Y_i(k) = (1-\lambda)Y_i(k-1) + \lambda y(k-i) \quad (15a)$$

$$U_i(k) = (1-\lambda)U_i(k-1) + \lambda u(k-i-d) \quad (15b)$$

The EWMA-LS estimator (eq 12) can be further improved by using the IV method. Supposing a signal $w(k)$ is introduced to create the instrumental variables $\psi(k)$, the EWMA-IV estimator turns out to be

$$\hat{\theta}_{EWMA}^{IV} = [\sum_{k=k_0+1}^{k_0+N} \psi(k)\varphi^T(k)]^{-1} \sum_{k=k_0+1}^{k_0+N} \psi(k)\zeta(k) \quad (16)$$

where

$$\psi(k) = [-W_1(k) \quad \cdots \quad -W_{n_A}(k) \quad U_1(k) \quad \cdots \quad U_{n_B}(k)]^T \quad (17)$$

$$W_i(k) = (1-\lambda)W_i(k-1) + \lambda w(k-i); W_i(k_0) = 0 \quad (18)$$

The effectiveness of the EWMA-IV estimator is attributed to the construction of $\psi(k)$, which counteracts the influence of noises on parameter estimation.

4. Selection of Weighting Constant λ

The preceding EWMA estimators demand the selection of λ. In general, the value of λ should be set close to one if the noise characteristics resemble the AR structure (eq 8). When the noise exhibits a significant deviation from that structure, a smaller value of λ should be employed to alleviate its loathsome effect on parameter estimates. To find the best λ value, we develop two error criteria. The first is the output error criterion:

$$J_{OE} = \sum_{k=k_0+1}^{k_0+N} [OE(k)]^2 = \sum_{k=k_0+1}^{k_0+N} [y(k) - y_M(k)]^2 \quad (19)$$

where OE(k) denotes the output error at instant k and $y_M(k)$ is the model predicted outputs by passing the inputs through the identified model. The second is the filtered output error criterion, composed of a prefilter $F(q)$ that converts the possibly autocorrelated signal OE(k) into a white sequence $e(k)$, i.e.,

$$F(q)OE(k) = e(k) \quad (20)$$

$$F(q) = 1 + f_1 q^{-1} + \cdots + f_{n_F} q^{-n_F}$$

Equation 20 can be expressed as a linear regression as

$$OE(k) = [-OE(k-1) \quad \cdots \quad -OE(k-n_F)]\theta_F \quad (21)$$

$$\theta_F = [f_1 \quad \cdots \quad f_{n_F}]^T$$

Estimate $F(q)$ using the LS method and denote the result by $\hat{F}(q)$. The filtered output error criterion thus takes the form:

$$J_{FOE} = \sum_{k=k_0+1}^{k_0+N} [FOE(k)]^2 = \sum_{k=k_0+1}^{k_0+N} [\hat{F}(q)OE(k)]^2 \quad (22)$$

In the evaluation of the output error OE(k), the model predictions $y_M(k)$ should be calculated according to

$$y_M(k) = \frac{\hat{B}(q)}{\hat{A}(q)} u(k-d) = G(q, \hat{\theta})u(k)$$

where $\hat{A}(q)$ and $\hat{B}(q)$ denote, respectively, the estimated polynomials of $A(q)$ and $B(q)$. This requires knowing the data of $y(k)$ for $k \le 0$ (the so-called initial conditions). However, these data are often unknown or noise-corrupted, so that the simplest way would be to replace the unknown initial values by zeros. On the contrary, the proposed EWMA estimators involve only known data if the starting instant k_0 is chosen to be a sufficiently larger value.

The best λ value is sought within the interval from 0 to 1 via minimizing the filtered output error criterion (eq 22). It will be shown that the proposed method with λ given in this way yields excellent parameter estimates for almost all types of noise characteristics. Moreover, the method is hardly affected by the presence of unknown initial conditions.

The use of the optimal λ is best suited to off-line identification experiments. For online applications or for simplicity, a good value for λ can be assigned a priori in accordance with the following guidelines:

(1) Choose $0.05 \le \lambda \le 0.25$ if the sampling period (T) is smaller than one-twentieth of the settling time (t_{set}), i.e., $T < t_{set}/20$. Here, the settling time denotes 95% complete response time for a step change in the process input.

(2) Choose $0.75 \le \lambda \le 1$ if the sampling period is larger than one-tenth of the settling time, $T > t_{set}/10$.

(3) The larger the sampling period is, the larger the λ should be chosen.

(4) The more the noise characteristics are deviated from the AR structure (eq 8), the more the λ value should be set smaller than one.

5. Off-Line Identification Algorithm

The sole use of the EWMA-LS estimator (eq 12) or the EWMA-IV estimator (eq 16) cannot guarantee the best efficiency for parameter estimates of the process model in the face of diverse noise characteristics. Besides, it is sometimes desirable to obtain not only an AR noise approximation but also the noise model in the accurate form of

$$v(k) = \frac{H(q)}{D(q)}e(k) \tag{23}$$

where

$$D(q) = 1 + d_1 q^{-1} + \cdots + d_{n_D} q^{-n_D}$$

$$H(q) = 1 + h_1 q^{-1} + \cdots + h_{n_H} q^{-n_H}$$

Equation 23 together with eq 1 is the so-called Box–Jenkins (BJ) model. The ARARX model can be obtained for $D(q) = A(q)C(q)$ and $H(q) = 1$, whereas the ARMAX model corresponds to $D(q) = A(q)$.

We thus propose the following four-step identification procedure with the weighting constant λ given a priori:

Step 1. Apply the EWMA-LS estimator to find the parameters a_i and b_i. Denote the parameter estimates by $\hat{\theta}_{LS}^{(1)}$ and the corresponding transfer function by

$$\hat{G}^{(1)}(q) = \frac{q^{-d}\hat{B}^{(1)}(q)}{\hat{A}^{(1)}(q)}$$

Step 2. Create the instrumental variables as in eqs 17 and 18:

$$w^{(1)}(k) = \hat{G}^{(1)}(q)u(k)$$

$$\psi^{(1)}(k) = [-W_1^{(1)}(k) \quad \cdots \quad -W_{n_A}^{(1)}(k) \quad U_1(k) \quad \cdots \quad U_{n_B}(k)]^T$$

and obtain the EWMA-IV estimator. Denote the parameter estimates by $\hat{\theta}_{IV}^{(2)}$ and the corresponding polynomials by $\hat{A}^{(2)}(q)$ and $\hat{B}^{(2)}(q)$.

Step 3. Postulate an AR noise model as given by eq 8 with $n_C = n_A + n_B$. Let

$$\hat{\varepsilon}(k) = \hat{A}^{(2)}(q)y(k) - \hat{B}^{(2)}(q)u(k-d)$$

and rearrange eq 9 as

$$C(q)\hat{\varepsilon}(k) = e(k)$$

This results in the following regression equation:

$$\hat{\varepsilon}(k) = [-\hat{\varepsilon}(k-1) \quad \cdots \quad -\hat{\varepsilon}(k-n_C)]\theta_C \tag{24}$$

$$\theta_C = [c_1 \quad \cdots \quad c_{n_C}]^T$$

Estimate the parameters c_i by the LS method and denote the results by the prefilter $\hat{C}(q)$.

Step 4. Calculate three filtered signals based on $\hat{C}(q)$ and $\hat{G}^{(1)}(q)$ as

$$y_F(k) = \hat{C}(q)y(k)$$

$$u_F(k) = \hat{C}(q)u(k)$$

$$w_F(k) = \hat{C}(q)w^{(1)}(k)$$

Apply the EWMA-IV estimator again based on the filtered signals to provide the final parameter estimates as

$$\hat{\theta}_{IV}^{(4)} = [\sum_{k=k_0+1}^{k_0+N} \psi_F(k)\varphi_F^T(k)]^{-1} \sum_{k=k_0+1}^{k_0+N} \psi_F(k)\zeta_F(k) \tag{25}$$

where

$$\varphi_F(k) = [-Y_1^F(k) \quad \cdots \quad -Y_{n_A}^F(k) \quad U_1^F(k) \quad \cdots \quad U_{n_B}^F(k)]^T$$

$$\psi_F(k) = [-W_1^F(k) \quad \cdots \quad -W_{n_A}^F(k) \quad U_1^F(k) \quad \cdots \quad U_{n_B}^F(k)]^T$$

In the above expressions, $\zeta_F(k)$, $Y_i^F(k)$, $U_i^F(k)$ and $W_i^F(k)$ are evaluated in the same way as in eqs 10, 15, and 18 with y, u, and w replaced by y_F, u_F, and w_F, respectively.

The presented four-step procedure is a substantial improvement over the IV4 method. In addition to the use of the EWMA, another distinct feature from the IV4 method is that the instrumental variables $\psi_F(k)$ employed in the final step are generated by the transfer function model obtained in step 1, $\hat{G}^{(1)}(q)$, rather than that in step 2. It is found that instrumental variables created in this way could mitigate the effect of noises much better for higher-order ($n_A \geq 3$) systems.

The proposed off-line algorithm entails repeated use of the preceding four-step procedure to find the optimal λ value that minimizes the error criterion (eq 22), thus giving the best final estimates of model parameters $\hat{a}_i$, $\hat{b}_i$, and $\hat{c}_i$ (from eqs 24 and 25). These parameter estimates constitute an ARARX model of eq 9. Note that with the assumption of the ARARX structure, the noise of unknown dynamics is approximated by a high-order AR structure (eq 8). The efficiency of such an approximation has been discussed by Wahlberg.[22]

To convert the identified ARARX model to an ARMAX model, we simply generate an output signal $v_w(k)$ by passing a simulated white sequence $e_w(k)$ through the identified AR disturbance part, i.e.

$$v_w(k) = \frac{1}{\hat{A}(q)\hat{C}(q)}e_w(k)$$

Assuming $D(q) = \hat{A}(q)$ in eq 23, the numerator $H(q)$ can then be secured by applying the LS method to the following regression equation:

$$\hat{A}(q)v_w(k) - e_w(k) = \sum_{j=1}^{n_H} h_j e_w(k-j) \tag{26}$$

If the exact process plus disturbance model is of the BJ structure, then the direct conversion from the identified ARARX model is not feasible because the incorrect component of $\hat{A}(q)$ existing in the AR disturbance approximation would impose a bias on the ARMA disturbance estimate. On the other hand, the relative error balanced model reduction (REBMR) technique presented by Wahlberg[22] could secure an efficient ARMA model from a good AR estimate. Hence we employ the REBMR technique to estimate the ARMA part of the BJ model as follows:

Step 1. Calculate the output error OE(k) defined in eq 19, which represents the noise $v(k)$. The prefilter $F(q)$ estimated by eq 21 constitutes an alternative high-order AR disturbance estimate:

$$v(k) = \frac{1}{\hat{F}(q)}e(k) \tag{27}$$

Step 2. Find the ARMA disturbance model by applying the REBMR technique to the above AR estimate.

6. Model-Based Monitoring

The standard assumptions for most conventional control charts used in process monitoring are that the process (noise) data when it is in statistical control are normally and independently distributed.[21] Both the mean and standard deviation are considered fixed and unknown. An out-of-control condition is a change or shift in the mean (or standard deviation) to some different value. Unfortunately, the dynamic feature of eq 23 implies that the process data tend to be autocorrelated over time, causing the conventional control charts to make too many false alarms.

One can use the identified disturbance model to remove the autocorrelation from the noise data $v(k)$ and apply control charts to the residuals or prediction errors. Suppose that eq 23 for the noise disturbance is available. The one-step-ahead prediction for the noise at time k made at time $k - 1$ is given by

$$\hat{v}_{k-1}(k) = \left[1 - \frac{D(q)}{H(q)}\right]v(k) \tag{28}$$

The sequence of one-step-ahead prediction errors

$$x(k) = v(k) - \hat{v}_{k-1}(k) \tag{29}$$

is independently and identically distributed with mean zero. When the process is under automatic control, the noise observation data can be calculated by

$$v(k) \approx y(k) - y_M(k) \tag{30}$$

where $y_M(k)$ is the predicted process outputs by passing the controller outputs $u(k)$ through the identified process model $G(q, \hat{\theta})$. For convenience, we recommend using the AR disturbance model of eq 27 to generate the sequence $x(k)$ for statistical monitoring:

$$x(k) \approx \hat{F}(q)[y(k) - y_M(k)] \tag{31}$$

The EWMA control chart has been frequently applied to monitor small shifts in the process mean.[21,23] The EWMA $\varpi(k)$ for the prediction errors is defined as

$$\varpi(k) = \eta x(k) + (1 - \eta)\varpi(k - 1) \tag{32}$$

The EWMA control chart is constructed by plotting $\varpi(k)$ versus the sample number k (or time). Inasmuch as $x(k)$ approximates an independent random sequence with mean zero and variance σ^2, the upper and lower control limits are known as

$$\mathrm{UCL} = \bar{x} + L\sigma\sqrt{\eta/(2 - \eta)} \tag{33a}$$

$$\mathrm{LCL} = \bar{x} - L\sigma\sqrt{\eta/(2 - \eta)} \tag{33b}$$

where the center line is the process target (or mean) $\bar{x}$ and L is the width of the control limits. A point that plots outside of the control limits indicates an out-of-control condition.

7. Online Identification Algorithm

In this section, we develop a recursive algorithm for online applications by incorporating the EWMA in eq 10 and the prefilter in eq 24. Two recursive routines are derived for updating $\hat{\theta}$ and $\hat{\theta}_C$, alternately. The first routine is based on the EWMA-LS estimator given by eq 12 with the filtered inputs and outputs, $u_F(k)$ and $y_F(k)$, as depicted in step 4 of the off-line algorithm. With fresh data continuously in supply, the estimator for the model parameters θ at time k can be updated by

$$\hat{\theta}(k) = \hat{\theta}(k-1) + \mu_F(k)P_F(k-1)\varphi_F(k)[\zeta_F(k) - \varphi_F^T(k)\hat{\theta}(k-1)] \tag{34}$$

where the covariance matrix P_F and the correction factor μ_F are

$$P_F(k) = \frac{1}{\alpha}[P_F(k-1) - \mu_F(k)P_F(k-1)\varphi_F(k)\varphi_F^T(k)P_F(k-1)] \tag{35}$$

$$\mu_F(k) = 1/[\alpha + \varphi_F^T(k)P_F(k-1)\varphi_F(k)] \tag{36}$$

In the above formulas, the target $\zeta_F(k)$ and the regressor $\varphi_F(k)$ are calculated as

$$\zeta_F(k) = (1 - \lambda)\zeta_F(k-1) + \lambda y_F(k); \zeta_F(k_0) = 0$$

$$\varphi_F(k) = [-Y_1^F(k) \ \cdots \ -Y_{n_A}^F(k) \ \ U_1^F(k) \ \cdots \ U_{n_B}^F(k)]^T$$

$$Y_i^F(k) = (1 - \lambda)Y_i^F(k-1) + \lambda y_F(k-i); Y_i^F(k_0) = 0, i = 1, \cdots, n_A$$

$$U_i^F(k) = (1 - \lambda)U_i^F(k-1) + \lambda u_F(k-i-d); U_i^F(k_0) = 0, i = 1, \cdots, n_B$$

$$u_F(j) = \hat{C}^{(k-1)}(q)u(j), j = k - n_B - d, \cdots, k - 1 - d$$

$$y_F(j) = \hat{C}^{(k-1)}(q)y(j), j = k - n_A, \cdots, k$$

The above routine is followed by the second recursive routine based on eq 24 for updating the prefilter parameters θ_C as

$$\hat{\theta}_C(k) = \hat{\theta}_C(k-1) + \mu_C(k)P_C(k-1)\xi(k)[\hat{\varepsilon}(k) - \xi^T(k)\hat{\theta}_C(k-1)] \tag{37}$$

where

$$P_C(k) = \frac{1}{\alpha}[P_C(k-1) - \mu_C(k)P_C(k-1)\xi(k)\xi^T(k)P_C(k-1)]$$

$$\mu_C(k) = 1/[\alpha + \xi^T(k)P_C(k-1)\xi(k)]$$

$$\xi(k) = [-\hat{\varepsilon}(k-1) \ \cdots \ -\hat{\varepsilon}(k-n_C)]^T$$

$$\hat{\varepsilon}(j) = \hat{A}^{(k)}(q)y(j) - \hat{B}^{(k)}(q)u(j-d), j = k - n_C, \cdots, k$$

The two routines can be started by assigning arbitrary initial values to $\hat{\theta}$ and $\hat{\theta}_C$ as well as letting $P_F(k_0) = \gamma_F I_F$ and $P_C(k_0) = \gamma_C I_C$. Note that γ_F and γ_C are large positive numbers, and I_F and I_C are identity matrices of appropriate dimensions. The forgetting factor α is introduced in the two routines to ensure a faster adaptive capability for time-varying systems. A value between 0.95 and 1 is recommended to avoid the loss of estimation accuracy. For constant parameter systems, we suggest setting α equal to one.

8. Model Structure Determination

The identification method discussed so far assumed that the system orders n_A and n_B and the delay d were given a priori. In practice, it is useful to make tests on the adequacy of the model structure, i.e., the orders and the delay. In the present framework, we propose a simple testing technique based on the goodness of fit of the model indicated by the output error criterion J_{OE} in eq 19. The testing technique is to compare the goodness of fit of the model for possible orders n_A and delays d on the assumption that n_B is equal to n_A. In general, J_{OE} decreases with an increase in n_A. As n_A is less than or equal to the true order, a significant reduction in J_{OE} with n_A is often observed. On the other hand, J_{OE} would cease to decrease significantly when n_A becomes greater than the true order. Moreover, if the model order is chosen correctly, the actual delay would exhibit an appreciable drop in J_{OE}.

The testing technique can then proceed as follows. First, apply the identification method to observed data for each n_A and d and calculate the corresponding error function J_{OE}. Second, plot a curve of logarithm of J_{OE} against n_A for each d. Finally, determine the best estimates of n_A and d in light of the above rules of thumb.

9. Simulation Results

We intend to compare the reliability of the proposed method against the GLS, IV4, and AFCLOE methods. The reliability can be evaluated by virtue of statistical properties such as bias, efficiency, and consistency. We are particularly concerned with the bias and efficiency properties in the face of various noise characteristics as well as different choices of sampling period T and sample size N. No general answer can be given to this concern. Monte-Carlo studies have been performed, in which

Table 1. Various Types of Stochastic Disturbances $G_v(s)$

type I	$1/(2s+1)(3s+1)^2(4s+1)$
type II	$(-2s+1)/(2s+1)(3s+1)$
type III	$(-2s+1)/(4s+1)$
type IV	$1/(2s+1)$
type V	$(s+1)(s^2+s+1)/(2s+1)(3s+1)(4s+1)$
type VI	$(s^2+3s+6)/(2s^2+3s+1)$
type VII	1
type VIII	$(0.2s+1)^3/(0.3s+1)(0.4s+1)^2$

stochastic systems have been simulated many times with different realizations of noise sequences. To investigate the effect of T, we described the discrete-time systems as resulting from the sampling of a continuous-time process transfer function, $G(s)$, connected with a zero-order hold. The noise disturbances $v(k)$ with versatile dynamic characteristics were simulated by the outputs of continuous-time disturbance transfer functions, $G_v(s)$, excited by a zero-mean, white sequence $e(k)$. Table 1 enumerates various types of disturbances to be considered. For open-loop identification, test data were generated with the process input $u(k)$ as a white binary ± 1 signal. For closed-loop identification, the system was under feedback control and test data were generated by a white binary ± 1 external excitation superposed to the reference. The input $u(k)$ was the controller output. The variance of $e(k)$ was adjusted so as to attain the desired noise-to-signal ratio (NSR), defined as the ratio of the standard deviation of the noise to the standard deviation of the signal. For each case, the simulation run was repeated 500 times with different realizations of $e(k)$. The mean ($\bar{a}_i$, $\bar{b}_i$) and standard deviation ($\bar{\sigma}_{a_i}$, $\bar{\sigma}_{b_i}$) of each parameter estimate were calculated from the 500 simulation runs. We thus define the following two composite quantities for evaluating the parameter estimates of the process model as a whole:

$$\text{percent error in mean} = \frac{\sum_{i=1}^{n_A} |\bar{a}_i - a_i| + \sum_{i=1}^{n_B} |\bar{b}_i - b_i|}{\sum_{i=1}^{n_A} |a_i| + \sum_{i=1}^{n_B} |b_i|} \times 100$$

$$\text{percent standard deviation} = \frac{\sum_{i=1}^{n_A} \bar{\sigma}_{a_i} + \sum_{i=1}^{n_B} \bar{\sigma}_{b_i}}{\sum_{i=1}^{n_A} |a_i| + \sum_{i=1}^{n_B} |b_i|} \times 100$$

Following the definition of the efficiency by Hsia,[6] the estimation algorithm is said to be efficient if it is unbiased (the percent error in mean is close to zero) and the percent standard deviation is as small as possible for finite N. The estimation algorithm is consistent if both the percent error in mean and percent standard deviation tend to be zeros as N approaches infinity.

Example 1

$$G(s) = \frac{(-2s+1)e^{-2Ts}}{(2s+1)(3s+1)(4s+1)}$$

The first example concerns a third-order process with a settling time of about 21. Given that $n_A = n_B = 3$, $n_C = n_F = 6$, and $d = 2$, three open-loop cases were postulated to investigate the effects of the sampling period (T), number of data points (N), and noise-to-signal ratio (NSR):

Case A: $N = 500$, NSR = 20%

Case B: $T = 0.5$, NSR = 20%

Case C: $T = 0.5$, $N = 500$

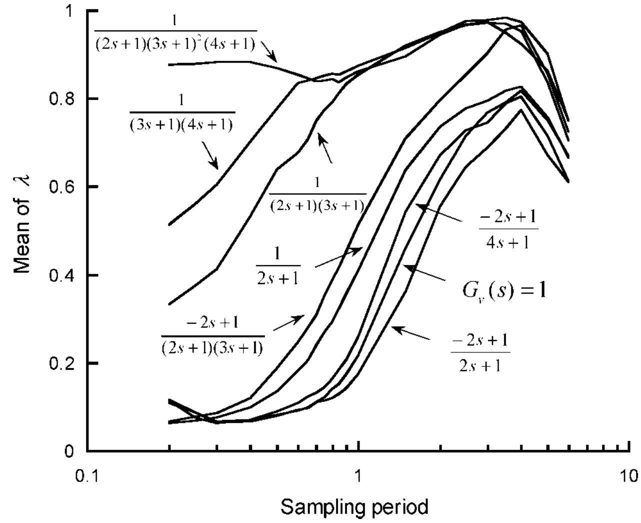

Figure 2. Mean of weighting constant λ versus sampling period for Example 1 subject to various types of stochastic disturbances.

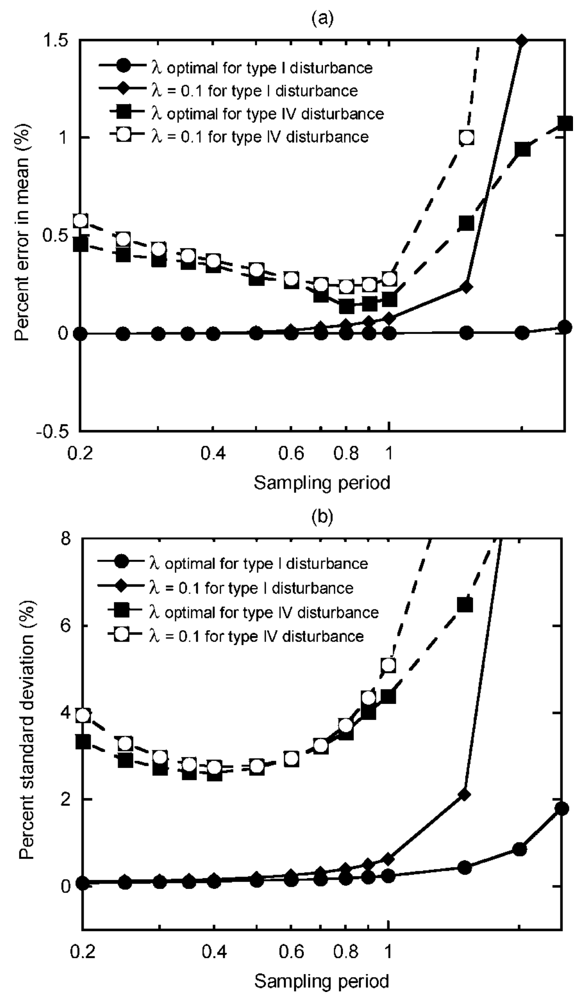

Figure 3. Comparison of identification results of Example 1 by the proposed off-line algorithm using the optimal λ and $\lambda = 0.1$ by virtue of (a) percent error in mean versus sampling period; (b) percent standard deviation versus sampling period.

For Case A with a specific type of stochastic disturbance, 500 simulation runs were performed for each sampling period between 0.2 and 6. The proposed off-line algorithm with the weighting constant λ determined by minimizing the criterion (eq 22) is applied to get the optimal parameter estimates for

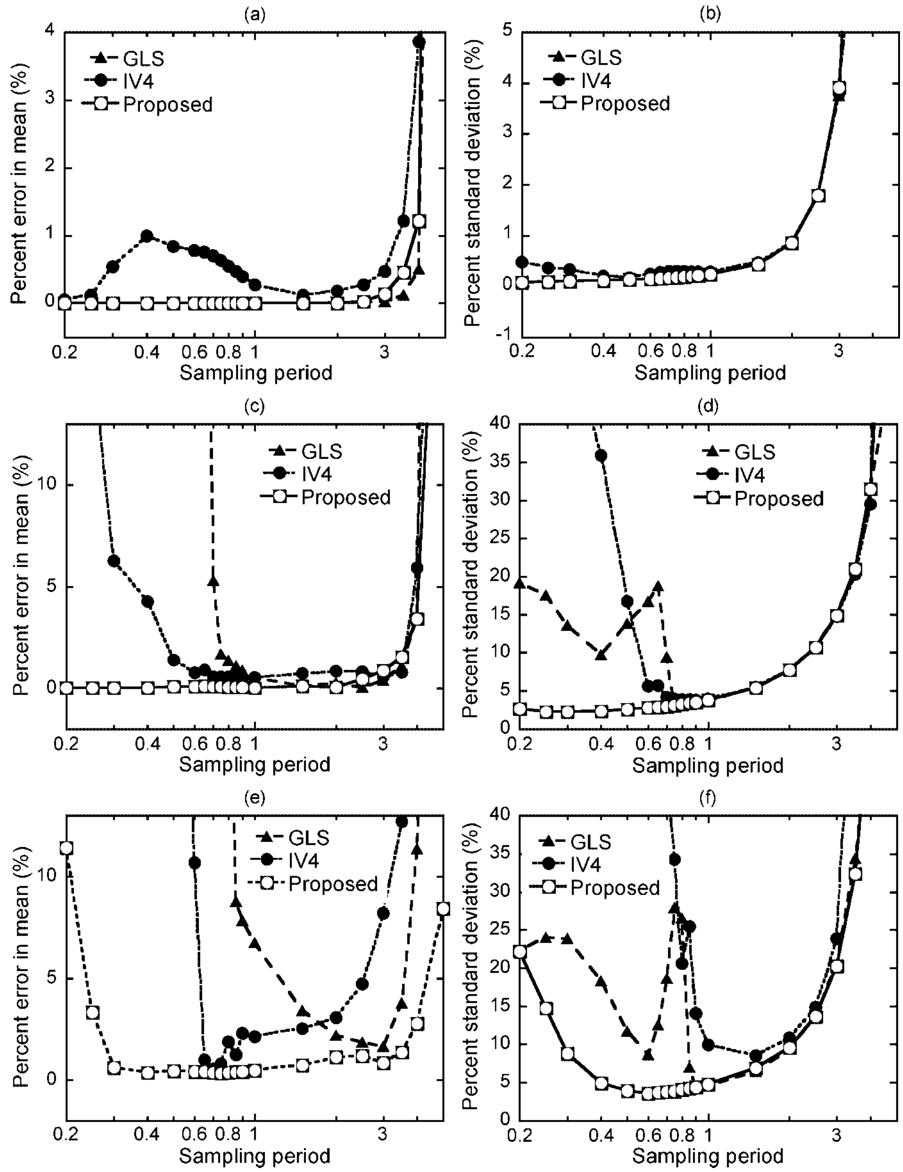

Figure 4. Effect of sampling period on identification results by three algorithms for Example 1 subject to (a, b) type I disturbance; (c, d) type II disturbance; (e, f) type III disturbance.

each run. Figure 2 plots the mean of the weighting constants (out of 500 runs) versus the sampling period subject to different noise disturbances. Note that the trends shown in these curves confirm the preceding guidelines (3 and 4) for choosing λ. Figure 3 compares the identification results delivered by the optimal λ with those by $\lambda = 0.1$ for the type I and type IV disturbances, which exhibit distinct dynamic features as seen in Table 1. Figure 2 reveals that the optimal λ varies widely with the sampling period and noise characteristics. On the contrary, Figure 3 indicates that if the sampling period is less than one, using a fixed constant of $\lambda = 0.1$ would yield the identification results comparable to the optimal ones for both types of disturbances. This event coincides with guideline (1) for choosing λ. It should be emphasized that this size of T, say, less than one-twentieth of the settling time, is perhaps the most important for control purposes.

It is interesting to know how the proposed off-line, GLS, and IV4 algorithms perform against the sampling period under Case A with the type I to type III disturbances, which exhibit the decreasing degree of resemblance to the AR structure (eq 8). Two issues are studied: bias and efficiency. First, Figure 4 shows clearly that all the three algorithms become poor if the sampling period is greater than 3. This is not surprising because the dynamic features of the process are concealed from the sampled data for a sampling period larger than about one-seventh of the settling time. For the type I disturbance that exhibits the highest degree of resemblance to the AR structure (eq 8), all three algorithms arrive at unbiased and efficient parameter estimates for a wide range of T, as seen in Figures 4a and 4b. Note that the percent error in mean is less than 1% and the percent standard deviation is less than 4% for $0.2 < T < 3$. For the type II disturbance, Figures 4c and 4d indicate that the proposed off-line algorithm is valid for the widest range of the sampling period ($0.2 < T < 3$), whereas the GLS and IV4 algorithms are valid for a modest range of the sampling period ($0.7 < T < 3$). For the type III disturbance, the proposed algorithm is still valid for the widest range of the sampling period ($0.3 < T < 3$), whereas the other two algorithms essentially fail for almost all sampling periods as shown in Figures 4e and 4f. As a consequence, the proposed algorithm is undoubtedly the best in that it always yields unbiased and the most efficient parameter estimates for the widest range of the sampling period. For the

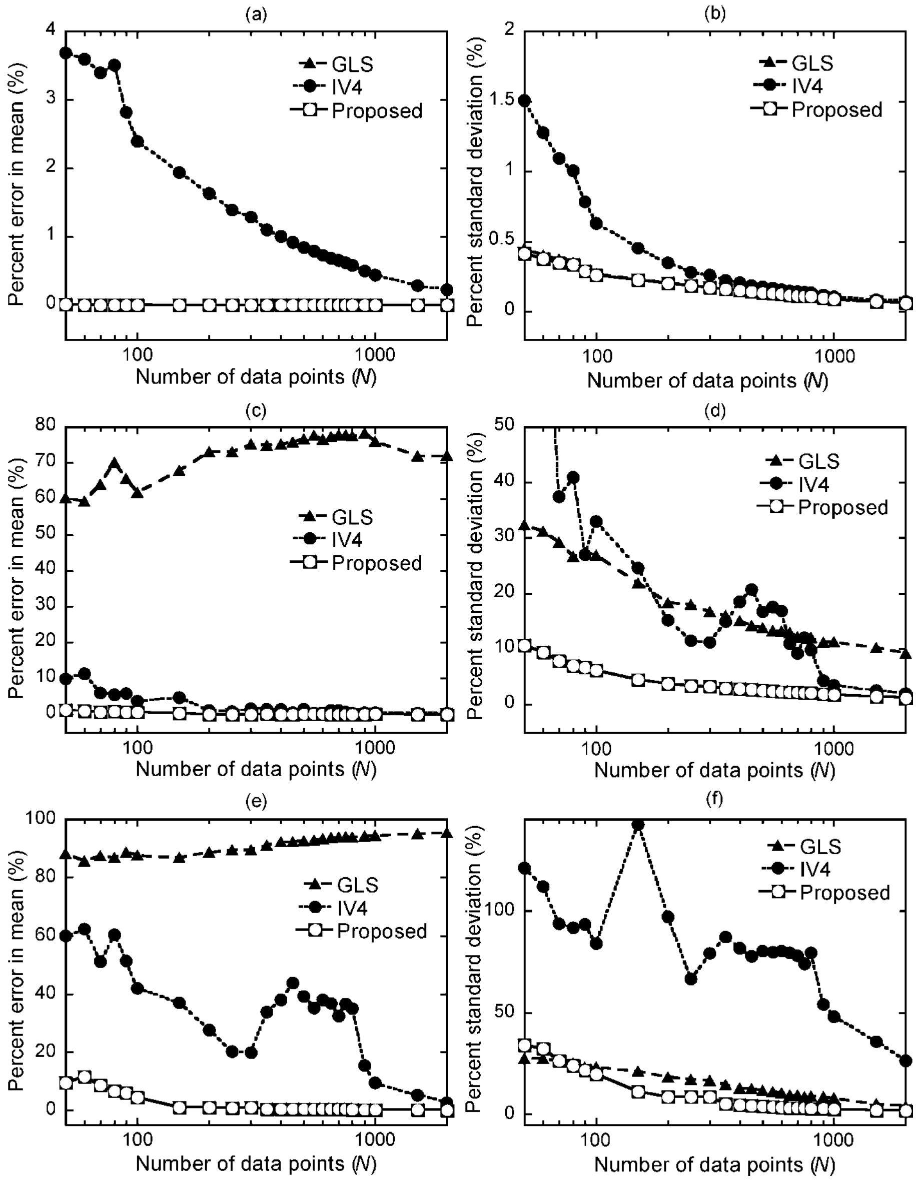

Figure 5. Effect of N on identification results by three algorithms for Example 1 subject to: (a, b) type I disturbance; (c, d) type II disturbance; (e, f) type III disturbance.

GLS and IV4 algorithms, the more the noise characteristics are deviated from the AR structure (eq 8), the narrower the valid range of the sampling period becomes.

We further investigate the effect of the number of data points (N) for Case B subject to the type I to type III disturbances. It follows from Figure 5 that the proposed algorithm gives rise to a consistent estimator for all three types of disturbances since both the percent error in mean and the percent standard deviation are close to zeros as N is increased to 2000. The parameter estimator obtained by the GLS algorithm, however, is consistent only for the type I disturbance. Note that the resulting percent error in mean is large and does not show any tendency to decrease with N for the other disturbances. The parameter estimator given by the IV4 algorithm does show some tendency to be consistent; nevertheless, it is still much inferior to the proposed algorithm in practical applications, which is reliable even for small N.

The reliability of an identification algorithm would eventually fail as the noise-to-signal ratio (NSR) increases to a large value. The question is how large the NSR can be tolerated for reliable estimation under different noise characteristics. Figure 6 compares the three algorithms against NSR between 5% and 100% for Case C with the type I to type III disturbances. It appears that the proposed off-line algorithm is the most robust with respect to NSR because it can lead to unbiased and efficient parameter estimates at least for NSR less than 50% despite different noise characteristics. On the contrary, the GLS and IV4 algorithms yield accurate parameter estimates only for very small NSR when the noise disturbance exhibits a distinct feature from the AR structure (eq 8) as evidenced in Figures 6c−f.

Up to now, it was assumed that order n_A and delay d were known a priori. We now apply the proposed testing technique to determine their exact values in a noisy environment. Figure 7 plots the calculated J_{OE} curves for $n_A = 1,...,5$ and $d = 1, 2, 3$ for a test run subject to the type III disturbance with NSR = 20%, $T = 1$, and $N = 500$. We observe that the J_{OE} curve for $d = 2$ ceases to drop significantly at $n_A = 3$, and the corresponding J_{OE} value is the smallest. Namely, the exact values of n_A and d are thus determined.

The preceding simulation work was focused on the accurate estimation of the process model under open-loop operation and subject to zero initial states. Indeed, the high-order AR

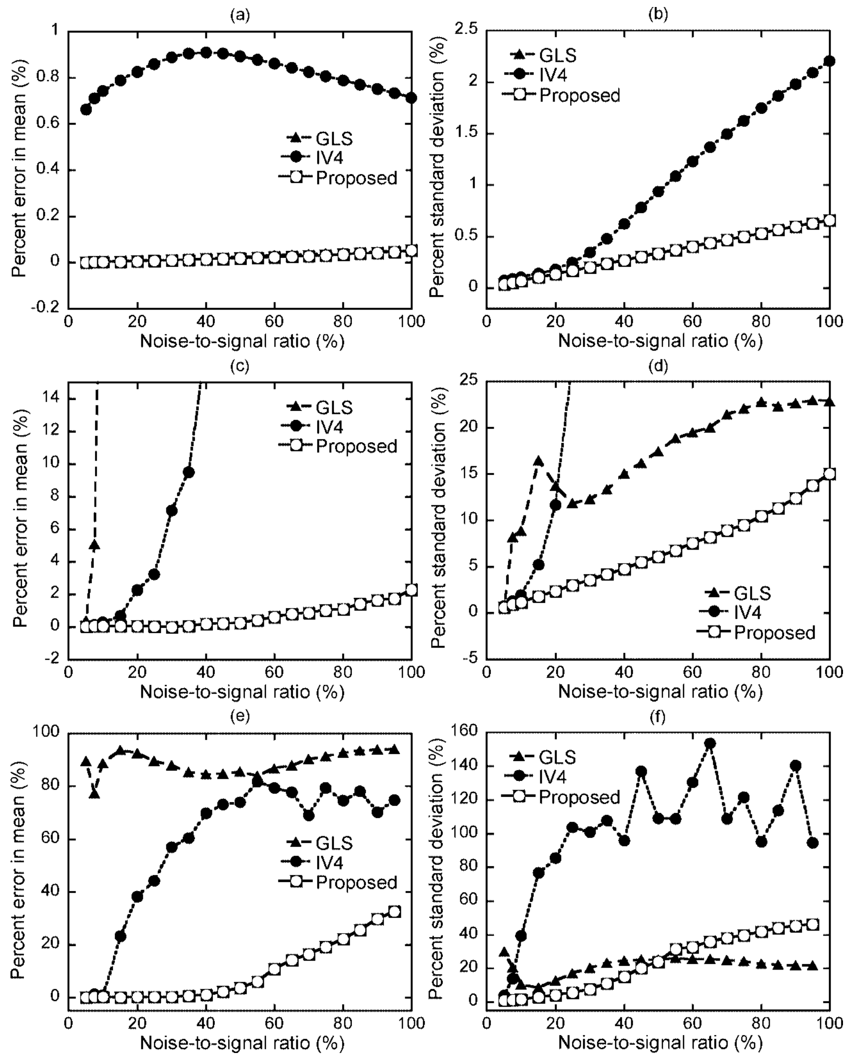

Figure 6. Effect of NSR on identification results by three algorithms for Example 1 subject to (a, b) type I disturbance; (c, d) type II disturbance; (e, f) type III disturbance.

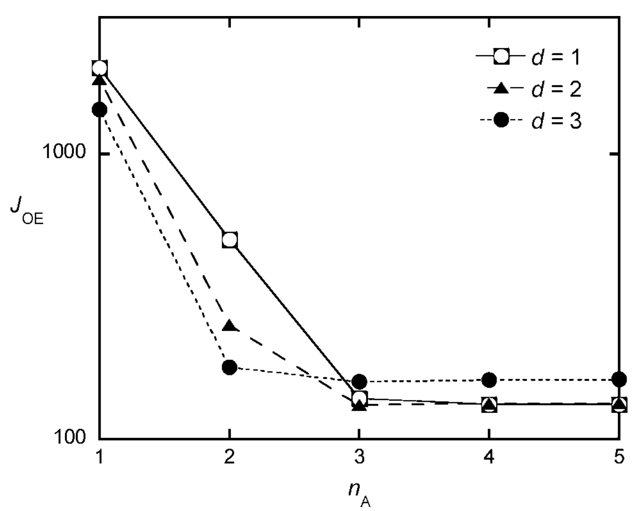

Figure 7. Curves of J_{OE} versus n_A for Example 1.

disturbance approximation of eq 8 together with the optimal weighting constant λ can cope with the multifarious features of noise dynamics. Furthermore, the off-line algorithm is suited to open-loop and closed-loop operation with unknown initial states and can provide a process plus disturbance model other than the ARARX structure. To avoid the minimization procedure, the λ value can be predetermined if the sampling period is such that the foregoing selection guidelines are applicable.

As a verification, the off-line algorithm with $T = 0.5$, $N = 500$, and NSR = 20% was applied to the open-loop and closed-loop identification of the third-order process under the type V disturbance (an ARMAX system) and then the type VI disturbance (a BJ system). For open-loop identification, a white binary $\pm$ 1 external excitation was introduced to the process input and the λ value was set to 0.1 according to guideline (1) because of $T < t_{set}/20$. For closed-loop identification, a white binary $\pm$ 1 external excitation was superposed to the reference of the proportional-integral control system and the optimal λ was sought to give the best parameter estimation. Table 2 enumerates the parameter estimates using eq 26 for the ARMAX model subject to nonzero initial conditions ($y(0) = 0.5$, $y'(0) = 0.5$, $y''(0) = -0.5$). Each estimate is expressed in terms of the mean and standard deviation (in the parenthesis) for 500 simulation runs. Table 3 lists the parameter estimates using the REBMR technique with $n_F = 6$ for the BJ model subject to zero initial conditions. The reduced order of the ARMA disturbance part of the BJ model is chosen to be 2, as inferred from the dominant number of the resultant six singular values. For instance, the singular values of one realization are 3.5948, 0.4535, 0.0575, 0.0448, 0.0385, and 0.0000, revealing that merely the first two states are dominant. The identification

Table 2. Open-Loop and Closed-Loop Parameter Estimates for the ARMAX System Subject to Nonzero Initial Conditions

	exact	open loop with $\lambda = 0.1$	closed loop with optimal λ
$\hat{A}(q)$	$1 - 2.508q^{-1} + 2.094q^{-2} - 0.582q^{-3}$	$1 - 2.506(\pm0.0212)q^{-1} + 2.089(\pm0.0416)q^{-2} - 0.580(\pm0.0208)q^{-3}$	$1 - 2.506(\pm0.0170)q^{-1} + 2.090(\pm0.0330)q^{-2} - 0.580(\pm0.0163)q^{-3}$
$\hat{B}(q)$	$-0.0079q^{-1} + 0.0041q^{-2} + 0.0078q^{-3}$	$-0.0079(\pm0.0008)q^{-1} + 0.0039(\pm0.0018)q^{-2} + 0.0079(\pm0.0012)q^{-3}$	$-0.0080(\pm0.0006)q^{-1} + 0.0041(\pm0.0012)q^{-2} + 0.0079(\pm0.0008)q^{-3}$
$\hat{D}(q)$	$A(q)$	$\hat{A}(q)$	$\hat{A}(q)$
$\hat{H}(q)$	$1 - 1.971q^{-1} + 1.429q^{-2} - 0.362q^{-3}$	$1 - 1.931(\pm0.0503)q^{-1} + 1.389(\pm0.103)q^{-2} - 0.332(\pm0.113)q^{-3}$	$1 - 1.933(\pm0.0496)q^{-1} + 1.394(\pm0.101)q^{-2} - 0.336(\pm0.112)q^{-3}$

Table 3. Open-Loop and Closed-Loop Parameter Estimates for the BJ System Subject to Zero Initial Conditions

	exact	open loop with $\lambda = 0.1$	closed loop with optimal λ
$\hat{A}(q)$	$1 - 2.508q^{-1} + 2.094q^{-2} - 0.582q^{-3}$	$1 - 2.507(\pm0.0266)q^{-1} + 2.091(\pm0.0517)q^{-2} - 0.581(\pm0.0254)q^{-3}$	$1 - 2.502(\pm0.0307)q^{-1} + 2.083(\pm0.0589)q^{-2} - 0.577(\pm0.0286)q^{-3}$
$\hat{B}(q)$	$-0.0079q^{-1} + 0.0041q^{-2} + 0.0078q^{-3}$	$-0.0079(\pm0.0008)q^{-1} + 0.0040(\pm0.0012)q^{-2} + 0.0079(\pm0.0009)q^{-3}$	$-0.0078(\pm0.0008)q^{-1} + 0.0040(\pm0.0014)q^{-2} + 0.0079(\pm0.0009)q^{-3}$
$\hat{D}(q)$	$1 - 1.385q^{-1} + 0.472q^{-2}$	$1 - 1.385(\pm0.101)q^{-1} + 0.475(\pm0.0952)q^{-2}$	$1 - 1.386(\pm0.101)q^{-1} + 0.474(\pm0.0966)q^{-2}$
$\hat{H}(q)$	$1 - 0.330q^{-1} + 0.375q^{-2}$	$1 - 0.339(\pm0.0960)q^{-1} + 0.354(\pm0.0443)q^{-2}$	$1 - 0.338(\pm0.0969)q^{-1} + 0.355(\pm0.0434)q^{-2}$

results demonstrate that the off-line algorithm works well for both systems under open-loop or closed-loop operation. Furthermore, using the fixed λ would still arrive at unbiased and efficient estimation, though the efficiency, as evaluated by the standard deviation, is slightly worse than the use of the optimal λ. The effect of nonzero initial conditions seems negligible.

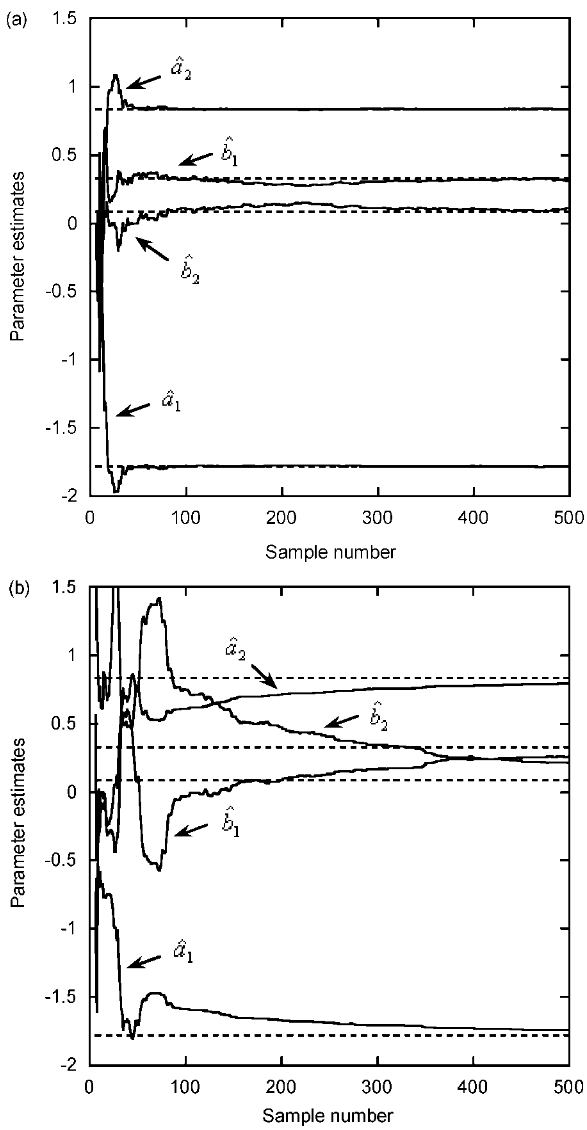

Figure 8. Real-time estimation results for Example 2 under feedback control by (a) the proposed online algorithm; (b) the recursive AFCLOE method.

Example 2

$$G(s) = \frac{0.254s + 1.8198}{s^2 + 0.3567s + 0.2426}$$

This second-order example has a settling time of 17. When sampled at $T = 1$, the continuous-time system gives rise to the discrete-time system employed by Ljung[8] for verifying the IV4 method. We first identify this example online using the developed recursive EWMA algorithm, as given in eqs 34−37. This is useful for applications such as adaptive control. The algorithm with $n_A = n_B = 2$, $n_C = 4$, $d = 0$, $T = 0.5$, and $\lambda =$

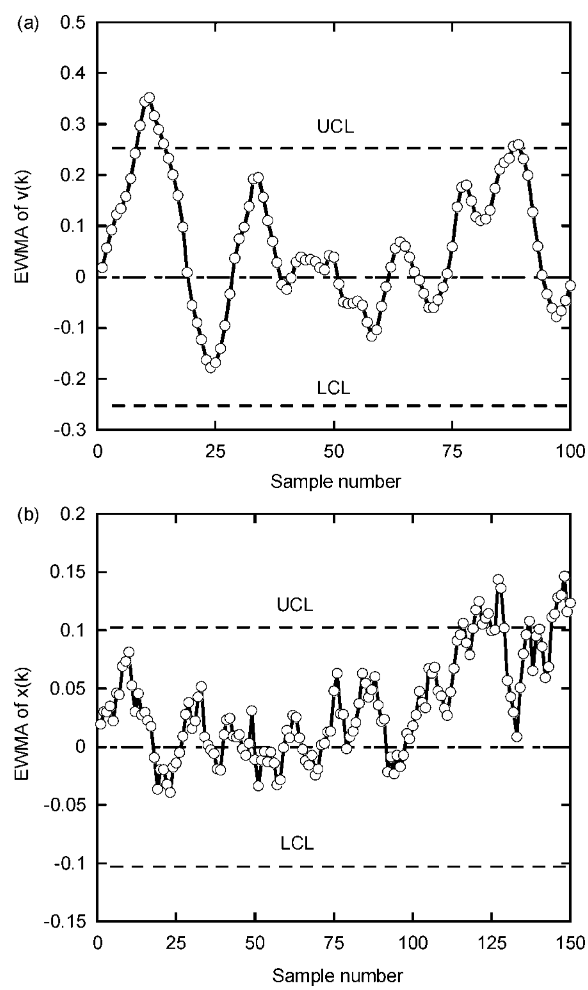

Figure 9. Monitoring of Example 2 based on EWMA control charts of (a) noise $v(k)$; (b) prediction errors $x(k)$.

0.1 leads to real-time estimation results based on a closed-loop simulation run corrupted by the type VII disturbance with NSR = 20%. The test run under proportional-integral control was excited by a white binary ± 1 input signal with the transition time interval $2T$. It follows that the four parameter estimates given by the algorithm adapt fast to their true values as shown in Figure 8a, whereas convergence of the recursive AFCLOE algorithm seems rather slow and inaccurate as seen in Figure 8b.

Next we apply the EWMA control charts based on eqs 31−33 with $\eta = 0.1$ and $L = 2.7$ to the data samples from Example 2 subject to the type VIII disturbance. A test run under proportional-integral control was performed to obtain the process model and the prefilter $F(q)$ with $n_F = 6$ using the off-line algorithm with $T = 0.25$, $N = 200$, NSR = 50%, and $\lambda = 0.1$. A small T was chosen because of fast disturbance dynamics. Thereafter, the process was switched to generalized minimum variance control at sample number zero and a shift of small magnitude in the mean was introduced at sample number 100. The EWMA control chart based directly on the noise data calculated by eq 30 is depicted in Figure 9a. Because the noise data are highly autocorrelated, it is not surprising to see that the control chart causes many false alarms when the process is in statistical control. On the other hand, the EWMA control chart based on the prediction errors calculated by eq 31 signals at sample number 116 as shown in Figure 9b, so we would conclude correctly that the process is out of control.

10. Conclusions

It has been shown that the proposed EWMA method tends to preserve good statistical properties of unbiasedness, efficiency, and consistency for a wide variety of process and noise dynamics in open-loop or closed-loop operation. In particular, the method is very robust with respect to the noise level as well as different choices of the sampling period and sample size. Consequently, it can identify a reliable ARARX, ARMAX, or BJ model from limited data acquisition. The resultant process plus disturbance model has proved to be very effective in removing the autocorrelation in process data for accurate statistical process monitoring.

The proposed method outperforms the conventional GLS and IV4 techniques in that the two techniques cannot secure satisfactory statistical properties if the disturbance dynamics consist of significant zeros or are much faster than the process dynamics. Moreover, if the experiment time is limited, the two techniques could be neither unbiased nor efficient because of the inevitable selection of a small sampling period or a small sample size.

A reliable recursive algorithm is provided for real-time applications. The λ value can be predetermined in accordance with the suggested guidelines. In the presence of significant noise, the recursive algorithm converges much faster to the true parameter estimates than the AFCLOE algorithm.

In practice, the information content in the external signals, input signal design, knowledge of model structure, and controller tuning would play a big role in the effectiveness of the proposed identification method. These issues have not been thoroughly resolved and merit further study.

Acknowledgment

This work was supported by the National Science Council of ROC under Grant NSC-96-2221-E006-144.

Literature Cited

(1) Notohardjono, B. D.; Ermer, D. S. Time Series Control Charts for Correlated and Contaminated Data. *ASME J. Eng. Ind.* **1986**, *108*, 219.

(2) Clarke, D. W.; Mohtadi, C.; Tuffs, P. S. Generalized Predictive Control−Part I. The Basic Algorithm. *Automatica* **1987**, *23*, 137.

(3) MacGregor, J. F. On-Line Statistical Process Control. *Chem. Eng. Prog.* **1988**, *84*, 21.

(4) Janakiram, M.; Keats, J. B. Combining SPC and EPC in a Hybrid Industry. *J. Qual. Technol.* **1998**, *30*, 189.

(5) Gultekin, M.; Elsayed, E. A.; English, J. R.; Hauksdottir, A. S. Monitoring Automatically Controlled Processes Using Statistical Control Charts. *Int. J. Prod. Res.* **2002**, *40*, 2303.

(6) Hsia, T. C. *System Identification−Least-Squares Methods*; Lexington Books: MA, 1977.

(7) Söderstrom, T.; Stöica, P. G. *Instrumental Variable Methods for System Identification*; Springer: Berlin, 1983.

(8) Ljung, L. *System Identification: Theory for the User*; Prentice Hall: Upper Saddle River, NJ, 1999.

(9) Hwang, S. H.; Lin, M. L. A Method for Identification of Discrete Parametric Models with Unknown Orders and Delays. *J. Chin. Inst. Chem. Eng.* **2002**, *33*, 373.

(10) Landau, I. D.; Zito, G. *Digital Control Systems: Design, Identification and Implementation*; Springer: Berlin, 2006.

(11) Ljung, L. *System Identification: Theory for the User*; Prentice Hall: Englewood Cliffs, NJ, 1987.

(12) Clarke, D. W. Generalized Least Squares Estimation of Parameters of a Dynamic Model. Presented at the First IFAC Symposium on Identification in Automatic Control Systems, Prague, 1967.

(13) Hasting-James, R.; Sage, M. W. Recursive Generalised-Least-Squares Procedures for On-Line Identification of Process Parameters. *Proc. IEE* **1969**, *116*, 2057.

(14) Söderstrom, T. Convergence Properties of the Generalized Least Squares Identification Method. *Automatica* **1974**, *10*, 617.

(15) Wong, K. Y.; Polak, E. Identification of Linear Discrete Time Systems Using the Instrumental Variable Approach. *IEEE Trans. Automat. Contr.* **1967**, *AC-12*, 707.

(16) Söderstrom, T.; Stöica, P. G. Comparison of Some Instrumental Variable Methods - Consistency and Accuracy Aspects. *Automatica* **1981**, *17*, 101.

(17) Smith, C. A.; Corripio, A. B. *Principles and Practice of Automatic Process Control*; John Wiley & Sons: New York, 1985.

(18) Seborg, D. E.; Edgar, T. F.; Mellichamp, D. A. *Process Dynamics and Control*; John Wiley & Sons: New York, 1989.

(19) Montgomery, D. C.; Mastrangelo, C. M. Some Statistical Process Control Methods for Autocorrelated Data. *J. Qual. Technol.* **1991**, *23*, 179.

(20) Box, G. E.; Jenkins, P.; Reinsel, G. M. *Time Series Analysis: Forecasting and Control*, 3rd ed.; Prentice Hall: Englewood Cliffs, NJ, 1994.

(21) Montgomery, D. C. *Introduction to Statistical Quality Control*, 4th ed.; John Wiley: New York, 2001.

(22) Wahlberg, B. Estimation of Autoregressive Moving-Average Models via High-Order Autoregressive Approximations. *J. Time Ser. Anal.* **1989**, *10*, 283.

(23) Hunter, J. S. The Exponentially Weighted Moving Average. *J. Qual. Technol.* **1986**, *18*, 203.

Received for review May 21, 2007
Revised manuscript received August 8, 2008
Accepted August 19, 2008

IE0707218