

Wiener-Hammerstein Model Identification-Recursive Algorithms*

Hosam E. EMARA-SHABAIK**, Mohammed S. AHMED**
and Khaled H. AL-AJMI**

Recursive algorithms for parameter estimation of Wiener-Hammerstein (W-H) models are developed. These algorithms are derived on the basis of minimizing cost functions of the output errors, the equation errors, and the prediction errors. Convergence conditions of the developed algorithms are established. Numerical results, to test the algorithms and illustrate their estimation accuracy, are given.

Key Words: Nonlinear Models, Wiener-Hammerstein Models, Block-Cascade Models, Linear-Nonlinear-Linear Models, Identification, Parameter Estimation, Recursive Algorithms, Output Error, Equation Error, Prediction Error

1. Introduction

The Wiener-Hammerstein (W-H) models are among the widely used models for nonlinear systems representation, Haber and Unbehauen⁽¹⁾. Such models are composed of cascaded dynamic linear parts with a static nonlinearity in between. These models are also referred to in the literature as sandwich, or Linear-Nonlinear-Linear (LNL) models, Palm⁽²⁾, and as the general or block-cascade models, Haber and Unbehauen⁽¹⁾. Mathematically, they are parameterized in terms of a finite number of parameters. Other models available for nonlinear systems representation include the NARMAX, and the Volterra series models, which, despite of their generality, have difficulties when applied to strongly nonlinear systems, Vandersteen et al.⁽³⁾. Wiener-Hammerstein models, on the other hand, are capable of representing any nonlinear system to any arbitrary degree of accuracy, Palm⁽²⁾. There are two subclasses of the Wiener-Hammerstein models, namely, the Hammerstein and the Wiener models subclasses. The Hammerstein

model is composed of a dynamic linear block following a static nonlinearity, i.e. it represents a linear dynamic system with actuator static nonlinearity. The Wiener model, on the other hand, is composed of a linear dynamic system followed by a static nonlinearity, i.e. it represents a linear dynamic system with sensor static nonlinearity. Wiener-Hammerstein models and their subclasses have found applications in many fields. They are used to model problems of fluid-structure dynamic interactions, Bendat⁽⁴⁾. Also, these models are used to model the human ankle stiffness dynamics, Kearney, et al.⁽⁵⁾, and used to model biological systems, Korenberg⁽⁶⁾. In the process control field, Eskinat, et al.⁽⁷⁾ used Hammerstein models with static polynomial type nonlinearity for modeling a distillation column and a heat exchanger. Kalafatis, et al.⁽⁸⁾ used Wiener models to represent a pH neutralization process. They reported that although the process is not in the model class, successful results have been obtained for all the examined cases.

The identification of Wiener-Hammerstein models and their subclasses is an important research area, Haber and Unbehauen⁽¹⁾, and Palm⁽²⁾, which has been treated by a number of authors. It is possible to recast a single-input single-output Hammerstein model as a linear model with multi-input and single-output, and then apply standard linear estimation

* Received 25th April, 2001

** Systems Engineering Department, KFUPM, Dhahran 31261, Saudi Arabia. E-mail: shabaik@ccse.kfupm.edu.sa

techniques to estimate an extended set of parameters, Ljung and Soderstrom⁽⁹⁾, Soderstrom and Stoica⁽¹⁰⁾, Goodwin and Sin⁽¹¹⁾. Billings and Fakhori⁽¹²⁾ used correlation techniques to estimate the impulse response of the linear subsystems and the parameters of the static nonlinearity. More recently, Kalafatis, et al.⁽⁸⁾ identified Wiener models in terms of the frequency response of the linear subsystem and the inverse of the static nonlinearity using batch least squares estimation technique. Yoshine and Ishii⁽¹³⁾ developed kernel representation, which can be used to estimate the impulse responses of the linear subsystems and the coefficients of the static nonlinearity of a W-H model. Emara-Shabaik, et al.⁽¹⁴⁾, developed offline parameter estimation algorithms for W-H models. Other authors, e.g., Chen and Fassios⁽¹⁵⁾, and Boutayeb and Darouach⁽¹⁶⁾ also investigated the problem of identifying Wiener-Hammerstein models.

In this paper, recursive algorithms based on the output error, the equation error, and the prediction error are developed to simultaneously estimate the parameters of the linear and nonlinear parts of the (W-H) model. The paper is organized as follows. The Wiener-Hammerstein model structure and the problem formulation are presented in Section 2. The derivations of the algorithms and their convergence conditions are given in Section 3. The developed algorithms are verified via numerical examples in Section 4.

Nomenclature

- t : time
 $u(t)$: system input sequence, known
 $y(t)$: the true System's measured output
 $e(t)$: output measurements noise
 $\hat{y}(t)$: the W-H model output corresponding to the input ' $u(t)$ '
 q^{-1} : backward time shift operator, i.e.
 $q^{-1}y(t)=y(t-1)$
 θ : system's parameters vector
 $\hat{\theta}(t)$: parameter estimates at time ' t '
 θ_o : the true system's parameters vector
 $E\{x(t)\}$: expected value of $x(t)$
 $E\{x(t)/z(t)\}$: expected value of $x(t)$ conditioned on $z(t)$
 $\eta_i[A]$: the i th eigen value of a matrix A
 $\eta_{\max}[A]$: maximum eigen value of a matrix A
 $\eta_{\min}[A]$: minimum eigen value of a matrix A
 I : the identity matrix
 $\|x\|$: the euclidean norm of a vector ' x '

2. The Wiener-Hammerstein Model Structure and Problem Formulation

A Wiener-Hammerstein system is composed of a

cascade of linear dynamic block, static nonlinear block, and a second linear dynamic block. In discrete time formulation, such a system with noise corrupted output is mathematically represented by,

$$y(t)=\frac{B(q^{-1})}{A(q^{-1})}w(t)+e(t) \quad (1)$$

$$w(t)=f(v(t)) \quad (2)$$

$$v(t)=\frac{D(q^{-1})}{C(q^{-1})}u(t) \quad (3)$$

Here in, $y(t)$ and $u(t)$ are the system output and input at time instant t , while $w(t)$ and $v(t)$ are inaccessible intermediate signals. Also, $A(q^{-1})$, $B(q^{-1})$, $C(q^{-1})$, $D(q^{-1})$, are polynomials in the backward time shift operator, (q^{-1}) , with known orders n_a , n_b , n_c , and n_d respectively as follows,

$$A(q^{-1})=1+a_1q^{-1}+\dots+a_{n_a}q^{-n_a} \quad (4.a)$$

$$B(q^{-1})=1+b_1q^{-1}+\dots+b_{n_b}q^{-n_b} \quad (4.b)$$

$$C(q^{-1})=1+c_1q^{-1}+\dots+c_{n_c}q^{-n_c} \quad (4.c)$$

$$D(q^{-1})=1+d_1q^{-1}+\dots+d_{n_d}q^{-n_d} \quad (4.d)$$

We assume the system to be time invariant. Further, we consider the case where the nonlinearity is represented by a polynomial of a known degree m as follows,

$$f(v(t))=g_1v(t)+g_2v^2(t)+\dots+g_mv^m(t) \quad (5)$$

The above model is also referred to as the plant or the true system. Furthermore, The $n_\theta \times 1$ vector of the unknown parameters of the above system is defined as follows,

$$\theta=(a_1\dots a_{n_a}b_1\dots b_{n_b}c_1\dots c_{n_c}d_1\dots d_{n_d}g_1\dots g_m)^T \quad (6)$$

The problem at hand is to estimate the parameter vector (θ) using record of measured output sequence corresponding to a known input sequence. The input signal $u(t)$ is taken as a stationary white Gaussian process. Such an input is known to provide sufficient excitation to enable the identification of a nonlinear system, Leontaritis and Billings⁽¹⁷⁾.

Let $\varepsilon(t, \theta)$ be the error between the true system's measured output and the W-H model output. Specific definitions of this error will be given later. The model parameter vector (θ) is to be estimated such that the following cost function, $V_t(\theta)$, is minimized.

$$V_t(\theta)=\frac{1}{2}\sum_{k=1}^t\lambda^{t-k}\varepsilon^2(k, \theta) \quad (7)$$

Where, $\lambda \leq 1$ is an arbitrary weighting factor that can be used to put more emphasis on the recent data when taken less than one. Thus, in effect, the initial data can be discarded once the estimation algorithm is under way. This has the effect of preventing possible deterioration in parameter estimation accuracy for nonlinear systems, Goodwin and Sin⁽¹¹⁾. The above cost function is adopted in Ljung and Soderstrom⁽⁹⁾, Soderstrom and Stoica⁽¹⁰⁾, and Goodwin and Sin⁽¹¹⁾. It is to be noticed that such a cost function is nonlinear in the W-H model parameters. Therefore, its minimi-

zation will require a numerical procedure.

3 Recursive Algorithms for W-H Parameter Estimation

To develop the recursive parameter estimation algorithms, we adopt the general methodology outlined in Ljung and Soderstrom⁽⁹⁾, and Soderstrom and stoica⁽¹⁰⁾. We assume there is a unique parameter vector (θ_0) for which the cost function given by Eq. (7) attains its global minimum value. Further, assuming the true system generating the data to belong to the model set then (θ_0) will correspond to the true system's parameter vector. Now, consider the second-order Taylor's series approximation of $V_i(\theta)$ around the estimate $\hat{\theta}(t-1)$, which is considered to minimize $V_{i-1}(\theta)$, as follows,

$$V_i(\theta) \approx V_i(\hat{\theta}(t-1)) + [V_i'(\hat{\theta}(t-1))]^T \times (\theta - \hat{\theta}(t-1)) + \frac{1}{2}(\theta - \hat{\theta}(t-1))^T V_i''(\hat{\theta}(t-1)) \times (\theta - \hat{\theta}(t-1)), \quad (8)$$

Where, $V_i'(\hat{\theta}(t-1))$, and $V_i''(\hat{\theta}(t-1))$ are the first and the second order derivatives of $V_i(\theta)$ with respect to (w.r.t.) (θ) evaluated at $\hat{\theta}(t-1)$. The minimizing solution $\hat{\theta}(t)$ of the right hand side of Eq. (8) is then given by,

$$\hat{\theta}(t) = \hat{\theta}(t-1) - [V_i''(\hat{\theta}(t-1))]^{-1} [V_i'(\hat{\theta}(t-1))] \quad (9)$$

The expression above is the well-known Newton-Raphson formula. It is known to have a local quadratic convergence to the true value (θ_0). If the formula, in Eq. (9) above, is to be implemented as given, it requires the evaluation of $V_i'(\hat{\theta}(.))$, and $V_i''(\hat{\theta}(.))$ at every iteration. Such a situation is not practical. Therefore, recursive forms for evaluating these two terms are essential. Equation (7) can be written in a recursive form as follows,

$$V_i(\theta) = \lambda V_{i-1}(\theta) + \frac{1}{2} \varepsilon^2(t, \theta) \quad (10)$$

Differentiating the above relationship w.r.t. (θ), one gets,

$$V_i'(\theta) = \lambda V_{i-1}'(\theta) + \varepsilon(t, \theta) \varepsilon'(t, \theta) \quad (11)$$

And,

$$V_i''(\theta) = \lambda V_{i-1}''(\theta) + \varepsilon'(t, \theta) [\varepsilon'(t, \theta)]^T + \varepsilon(t, \theta) \varepsilon''(t, \theta) \quad (12)$$

Where, $\varepsilon'(t, \theta)$ is the Gradient vector and $\varepsilon''(t, \theta)$ is the Hessian matrix of the scalar error $\varepsilon(t, \theta)$ w.r.t. (θ). Furthermore, we make the approximation, $V_{i-1}''(\hat{\theta}(t-1)) \cong V_{i-1}''(\hat{\theta}(t-2))$, and neglect $\varepsilon(t, \theta) \varepsilon''(t, \theta)$ as in Ljung and Soderstrom⁽⁹⁾, and Soderstrom and Stoica⁽¹⁰⁾. Neglecting the third term on the right hand side of Eq. (12), means that a Gauss-Newton algorithm rather than a Newton-Raphson algorithm is used. Also, since $\hat{\theta}(t-1)$ is assumed to minimize $V_{i-1}(\theta)$, it then follows from Eqs. (11) and (12) that.

$$V_i'(\hat{\theta}(t-1)) = \varepsilon(t, \hat{\theta}(t-1)) \varepsilon'(t, \hat{\theta}(t-1)) \quad (13)$$

And

$$V_i''(\hat{\theta}(t-1)) = \lambda V_{i-1}''(\hat{\theta}(t-2)) + \varepsilon'(t, \hat{\theta}(t-1)) [\varepsilon'(t, \hat{\theta}(t-1))]^T \quad (14)$$

At this point, and without any confusion, the following simplified notations are adopted.

$$\varepsilon(t) = \varepsilon(t, \hat{\theta}(t-1)) \quad (15)$$

$$\phi(t) = -[\varepsilon'(t, \hat{\theta}(t-1))]^T \quad (16)$$

$$P(t) = [V_i''(\hat{\theta}(t-1))]^{-1} \quad (17)$$

Then, Eq. (14) can be rewritten as,

$$P^{-1}(t) = \lambda P^{-1}(t-1) + \phi^T(t) \phi(t) \quad (18)$$

Invoking the matrix inversion lemma with regard to Eq. (18) above, and substituting into Eq. (9), we get the following recursive parameter estimation algorithm.

$$\begin{aligned} \hat{\theta}(t) &= \hat{\theta}(t-1) + K(t) \varepsilon(t) \\ K(t) &= P(t) \phi^T(t) \\ &= P(t-1) \phi^T(t) / [\lambda + \phi(t) P(t-1) \phi^T(t)] \\ P(t) &= \frac{1}{\lambda} \left\{ P(t-1) - \frac{P(t-1) \phi^T(t) \phi(t) P(t-1)}{\phi(t) P(t-1) \phi^T(t) + \lambda} \right\} \end{aligned} \quad (19)$$

Expressions for $\varepsilon(t)$ and $\phi(t)$ based on the output error, the equation error, and the prediction error are given in the following sections. Detailed derivations are given in Appendix A.

3.1 The Recursive Output Error (ROE) algorithm

In this method $\varepsilon(t, \theta)$ is equated to the output error $e_{OE}(t, \theta)$, which is defined by,

$$\varepsilon_{OE}(t, \theta) = y(t) - \hat{y}(t) \quad (20.a)$$

$$\hat{y}(t) = \frac{B(q^{-1})}{A(q^{-1})} w(t) \quad (20.b)$$

$$w(t) = g_1 v(t) + g_2 v^2(t) + \dots + g_m v^m(t) \quad (20.c)$$

$$v(t) = \frac{D(q^{-1})}{C(q^{-1})} u(t) \quad (20.d)$$

The output error generation is as illustrated in Fig. 1. And, it follows that,

$$\phi(t) = - \left[\frac{\partial e_{OE}(t, \theta)}{\partial \theta} \right]^T = \left[\frac{\partial \hat{y}(t)}{\partial \theta} \right]^T \quad (21)$$

Therefore, the algorithm in Eq. (19) is augmented with the following expressions for computing the elements of $\phi(t)$, which are evaluated at the previous parameter

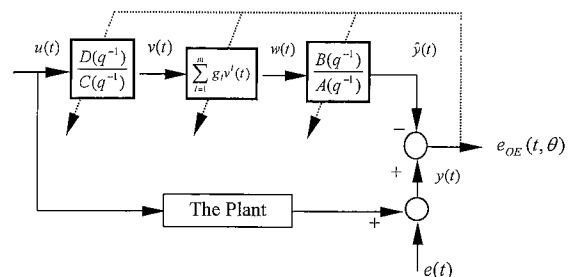


Fig. 1 Output error method for the Wiener-Hammerstein model identification

ter estimates $\hat{\theta}(t-1)$, to give the recursive output error (ROE) algorithm for estimating the parameters of the W-H model.

$$\frac{\partial \hat{y}(t)}{\partial a_i} = -\frac{q^{-i}B(q^{-1})}{A^2(q^{-1})} \left[\sum_{l=1}^m g_l \left(\frac{D(q^{-1})}{C(q^{-1})} u(t) \right)^l \right],$$

$$i=1, 2, \dots, n_a \quad (22.a)$$

$$\frac{\partial \hat{y}(t)}{\partial b_i} = -\frac{q^{-i}}{A(q^{-1})} \left[\sum_{l=1}^m g_l \left(\frac{D(q^{-1})}{C(q^{-1})} u(t) \right)^l \right],$$

$$i=1, 2, \dots, n_b \quad (22.b)$$

$$\frac{\partial \hat{y}(t)}{\partial c_i} = -\frac{B(q^{-1})}{A(q^{-1})} \left[\sum_{l=1}^m l \times g_l v^{l-1}(t) \right] \frac{q^{-i}D(q^{-1})}{C^2(q^{-1})} u(t),$$

$$i=1, 2, \dots, n_c \quad (22.c)$$

$$\frac{\partial \hat{y}(t)}{\partial d_i} = \frac{B(q^{-1})}{A(q^{-1})} \left[\sum_{l=1}^m l \times g_l v^{l-1}(t) \right] \left(\frac{q^{-i}}{C(q^{-1})} u(t) \right),$$

$$i=1, 2, \dots, n_d \quad (22.d)$$

$$\frac{\partial \hat{y}(t)}{\partial g_i} = \frac{B(q^{-1})}{A(q^{-1})} \left(\frac{D(q^{-1})}{C(q^{-1})} u(t) \right)^i,$$

$$i=1, 2, \dots, m \quad (22.e)$$

3.2 The Recursive Equation Error (REE) algorithm

In this method $\varepsilon(t, \theta)$ is equated to the equation error $e_{EE}(t, \theta)$ as illustrated in Fig. 2, and is given by,

$$e_{EE}(t) = A(q^{-1})y(t) - B(q^{-1})w(t) \quad (23.a)$$

$$w(t) = g_1 v(t) + g_2 v^2(t) + \dots + g_m v^m(t) \quad (23.b)$$

$$v(t) = \frac{D(q^{-1})}{C(q^{-1})} u(t) \quad (23.c)$$

The recursive equation error (REE) algorithm is obtained by augmenting equation (19) with the following expressions for computing the elements of $\phi(t)$ which are to be evaluated at the previous parameter estimates $\hat{\theta}(t-1)$.

$$\phi(t) = - \left[\frac{\partial e_{EE}(t, \theta)}{\partial \theta} \right]^T \quad (24)$$

Therefore, the various elements of $\phi(t)$ are obtained as follows,

$$\frac{\partial e_{EE}(t)}{\partial a_i} = q^{-i} y(t), \quad i=1, 2, \dots, n_a \quad (25.a)$$

$$\frac{\partial e_{EE}(t)}{\partial b_i} = -q^{-i} \left[\sum_{l=1}^m g_l \left(\frac{D(q^{-1})}{C(q^{-1})} u(t) \right)^l \right],$$

$$i=1, 2, \dots, n_b \quad (25.b)$$

$$\frac{\partial e_{EE}(t)}{\partial c_i} = B(q^{-1}) \left[\sum_{l=1}^m l \times g_l v^{l-1}(t) \right] \frac{q^{-i}D(q^{-1})}{C^2(q^{-1})} u(t),$$

$$i=1, 2, \dots, n_c \quad (25.c)$$

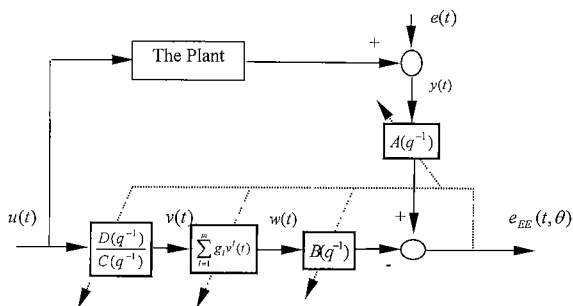


Fig. 2 Equation error method for the Wiener-Hammerstein model identification

$$\frac{\partial e_{EE}(t)}{\partial d_i} = -B(q^{-1}) \left[\sum_{l=1}^m l \times g_l v^{l-1}(t) \right] \frac{q^{-i}}{C(q^{-1})} u(t),$$

$$i=1, 2, \dots, n_d \quad (25.d)$$

$$\frac{\partial e_{EE}(t)}{\partial g_i} = -B(q^{-1}) \left(\frac{D(q^{-1})}{C(q^{-1})} u(t) \right)^i,$$

$$i=1, 2, \dots, m \quad (25.e)$$

3.3 The Recursive Prediction Error (RPE) algorithm

In this method the model parameters are augmented with the parameters of the assumed noise filter $F(q^{-1})$ as follows.

$$F(q^{-1}) = 1 + f_1 q^{-1} + \dots + f_{n_f} q^{-n_f} \quad (26)$$

Also, $\varepsilon(t, \theta)$ is equated to the prediction error $e_{PE}(t, \theta)$ as illustrated in Fig. 3, and is given by,

$$e_{PE}(t) = F^{-1}(q^{-1}) [A(q^{-1})y(t) - B(q^{-1})w(t)] \quad (27.a)$$

$$w(t) = g_1 v(t) + g_2 v^2(t) + \dots + g_m v^m(t) \quad (27.b)$$

$$v(t) = \frac{D(q^{-1})}{C(q^{-1})} u(t) \quad (27.c)$$

The recursive prediction error (RPE) algorithm is obtained by augmenting the algorithm in Eq.(19) with the following expressions for computing the elements of $\phi(t)$ which are to be evaluated at the previous parameter estimates $\hat{\theta}(t-1)$.

$$\phi(t) = - \left[\frac{\partial e_{PE}(t, \theta)}{\partial \theta} \right]^T \quad (28)$$

Therefore, the elements of $\phi(t)$ are obtained as follows,

$$\frac{\partial e_{PE}(t)}{\partial a_i} = \frac{q^{-i}}{F(q^{-1})} y(t), \quad i=1, 2, \dots, n_a \quad (29.a)$$

$$\frac{\partial e_{PE}(t)}{\partial b_i} = -\frac{q^{-i}}{F(q^{-1})} \left[\sum_{l=1}^m g_l \left(\frac{D(q^{-1})}{C(q^{-1})} u(t) \right)^l \right],$$

$$i=1, 2, \dots, n_b \quad (29.b)$$

$$\frac{\partial e_{PE}(t)}{\partial c_i} = \frac{B(q^{-1})}{F(q^{-1})} \left[\sum_{l=1}^m l \times g_l v^{l-1}(t) \right] \frac{q^{-i}D(q^{-1})}{C^2(q^{-1})} u(t),$$

$$i=1, 2, \dots, n_c \quad (29.c)$$

$$\frac{\partial e_{PE}(t)}{\partial d_i} = -\frac{B(q^{-1})}{F(q^{-1})} \left[\sum_{l=1}^m l \times g_l v^{l-1}(t) \right] \frac{q^{-i}}{C(q^{-1})} u(t),$$

$$i=1, 2, \dots, n_d \quad (29.d)$$

$$\frac{\partial e_{PE}(t)}{\partial f_i} = -\frac{q^{-i}}{F(q^{-1})} [A(q^{-1})y(t) - B(q^{-1})w(t)]$$

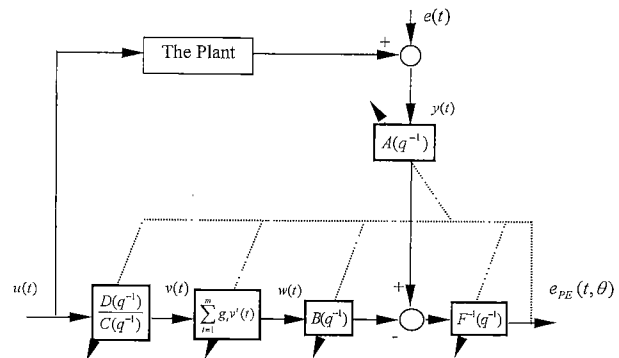


Fig. 3 Prediction error method for the Wiener-Hammerstein model identification

$$\times \sum_{i=1}^m g_i \left(\frac{D(q^{-1})}{C(q^{-1})} u(t) \right)^i, i=1, 2, \dots, n_f \quad (29.e)$$

$$\frac{\partial e_{PE}(t)}{\partial g_i} = -\frac{B(q^{-1})}{F(q^{-1})} \left(\frac{D(q^{-1})}{C(q^{-1})} u(t) \right)^i, \quad i=1, 2, \dots, m \quad (29.f)$$

3.4 Convergence conditions

Conditions for convergence of the developed recursive parameters estimation algorithms are given by the following theorem.

Theorem :

Let the true system belong to the model set. Let, $\frac{B(q^{-1})}{A(q^{-1})}$, $\frac{B(q^{-1})}{F(q^{-1})}$, $\frac{1}{A(q^{-1})}$, $\frac{1}{C(q^{-1})}$, $\frac{1}{F(q^{-1})}$, $B(q^{-1})$ be stable, and the nonlinearity of the model be bounded with bounded derivatives. Also, let, $\mathfrak{F}_t$, $t=1, 2, \dots$ be an increasing sequence of sigma algebras generated by $[u(0), u(1), \dots, u(t), y(0), y(1), \dots, y(t-1)]$; and let the measurements noise satisfy the following conditions:

$$(C1) \quad E\{e(t)/\mathfrak{F}_t\}=0$$

$$(C2) \quad E\{e^2(t)/\mathfrak{F}_t\}=\sigma^2 \leq \infty$$

Then, the ROE, REE, and RPE algorithms converge in the mean square sense to the true parameters, under the following conditions.

$$\frac{\eta_{\max}[P(t)]}{\eta_{\min}[P(t)]} < \infty \quad (30)$$

$$\eta_{\max}[P(t)] \rightarrow 0 \quad (31)$$

Proof :

The proof of the above theorem is given in Appendix B.

4. Results

In this section, the performance of the developed recursive parameter estimation algorithms is illustrated by numerical examples. The examples are such that the linear blocks have 2nd or 3rd order dynamics and the nonlinearity is of up to a third order which is representative of many physical systems. A

white Gaussian input sequence with zero mean and variance 0.5 is used. The output measurements noise $e(t)$ is a zero-mean Gaussian sequence with variance 0.05. In the first example the true system is described by,

$$y(t) = \frac{1+0.3q^{-1}}{1-0.4q^{-1}+0.25q^{-2}} w(t) + e(t)$$

$$w(t) = 0.5v(t) - 1.8v^2(t) + 2.5v^3(t)$$

$$v(t) = \frac{1+0.7q^{-1}}{1+0.5q^{-1}-0.25q^{-2}} u(t)$$

The parameter estimates obtained by the different algorithms are as shown in Table 1.

The time histories of the estimates for this example, as obtained by the RPE algorithm, are as shown in Fig. 4.

In the second example the true system is given by,

$$y(t) = \frac{1+0.3q^{-1}}{1-0.4q^{-1}+0.25q^{-2}} w(t) + e(t)$$

Table 1 Parameter estimates of the first example

Param.	Algo.	True Value	ROE	REE	PE
a ₁		-0.04	-0.3903±0.0095	-0.3968±0.0021	-0.3993±0.0006
a ₂		0.25	0.2488±0.0021	0.2475±0.0008	0.2491±0.0007
b ₁		0.30	0.3033±0.0041	0.2854±0.0031	0.3018±0.0027
c ₁		0.50	0.4951±0.0625	0.5111±0.0511	0.4989±0.0515
c ₂		0.25	-0.2478±0.0057	-0.2462±0.0038	-0.2487±0.0041
d ₁		0.70	0.6931±0.0302	0.7123±0.0379	0.6974±0.0402
g ₁		0.50	0.5123±0.0209	0.4794±0.0061	0.4984±0.006
g ₂		-1.80	-1.8008±0.1310	-1.8104±0.1606	-1.7992±0.1542
g ₃		2.5	2.4942±0.1824	2.8045±0.1592	2.5003±0.1687

Table 2 Parameter estimates of the second example

Param.	Algo.	True Value	ROE	REE	PE
a ₁		-0.4	-0.4012±0.0072	-0.40729±0.0017	-0.4015±0.0006
a ₂		0.25	0.2494±0.0032	0.2864±0.0005	0.2502±0.0004
b ₁		0.3	0.3009±0.0055	0.2272±0.0016	0.2976±0.0019
c ₁		-0.6	-0.5883±0.0327	-0.6296±0.0342	-0.6003±0.0254
c ₂		0.25	-0.2545±0.0064	-0.2382±0.0045	-0.2512±0.0049
c ₃		0.15	0.1495±0.0007	0.1400±0.0004	0.1494±0.0003
d ₁		0.55	-0.5400±0.0590	-0.5797±0.0411	-0.5492±0.0304
d ₂		0.2	0.2036±0.0057	-0.1936±0.0043	-0.2027±0.0053
g ₁		1.00	1.0037±0.0061	1.0056±0.0265	1.0047±0.0023
g ₂		0.5	2.4922±0.2842	2.4988±0.1152	2.4966±0.1133

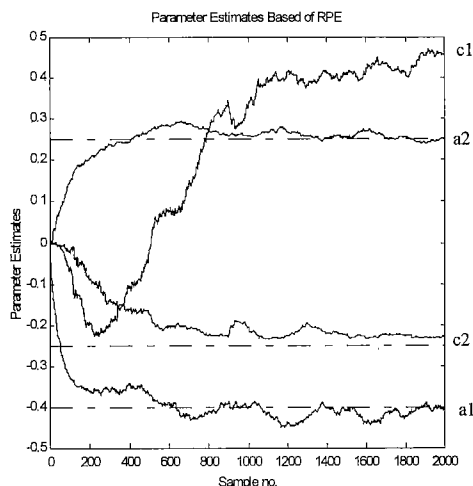
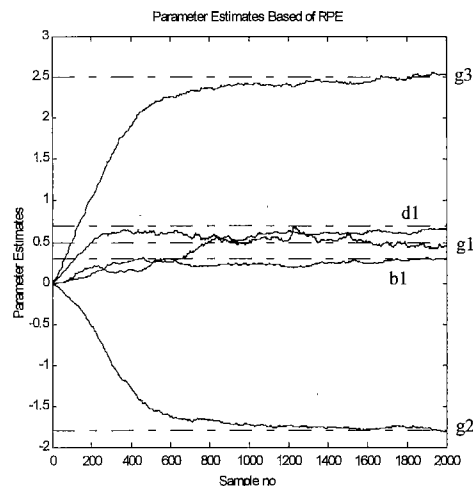


Fig. 4 Parameter estimates based on RPE of the first example



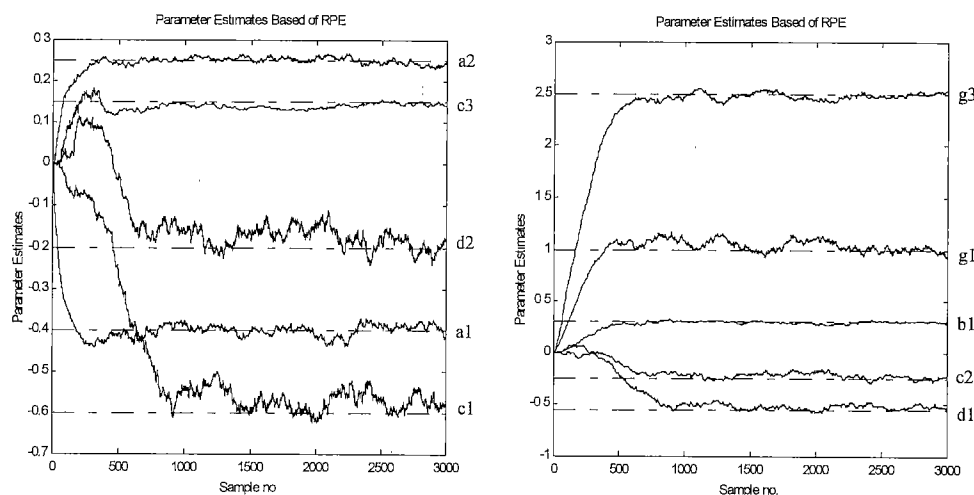


Fig. 5 Parameter estimates based on RPE of the second example

$$w(t) = v(t) + 2.5v^3(t)$$

$$v(t) = \frac{1 - 0.55q^{-1} - 0.2q^{-2}}{1 - 0.6q^{-1} - 0.25q^{-2} + 0.15q^{-3}} u(t)$$

The parameter estimates of this example are as shown in Table 2.

Figure 5 shows the traces of the parameters estimates for this example as obtained by the RPE algorithm.

In the above examples the following values are used, $\hat{\theta}(0)=0$, $\lambda=0.99$, $P(0)=\alpha I$, $\alpha=10^{-3}$. The obtained numerical results indicate the high accuracy of the parameters estimates. The REE provided the least accurate estimates among the three algorithms. Both, the ROE, and the RPE are close in accuracy with the RPE being slightly more accurate than the ROE.

5. Discussions

The numerical results show that among the three algorithms, developed in this paper, the (REE) algorithm provided the least accurate estimates. This can be attributed to the effect of the noise correlation that is introduced by the formulation of the equation errors. This introduced noise correlation is treated in the (RPE) algorithm by including a noise filter whose parameters are to be estimated along with the system parameters. Although, the recursive prediction error (RPE) algorithm gives the most accurate estimates, it was only slightly better than the recursive output error (ROE) algorithm. This is due to the fact that in both formulations, either OE or PE, the parameter estimation problem remains highly nonlinear in the case of W-H models. On the other hand, the (RPE) algorithm requires more computations than the (ROE) algorithm due to the need to estimate the noise filter parameters. Therefore, on this basis the (ROE) algorithm might be favored in the case of W-H

models.

6. Conclusions

In this paper recursive identification algorithms, for Wiener-Hammerstein models, are derived based on the minimization of the output error (OE), the equation error (EE), and the prediction error (PE). Convergence conditions of the developed algorithms are established. Numerical examples are given which demonstrate the effectiveness of the developed algorithms.

Acknowledgements

The authors wish to acknowledge the support provided by King Fahd University of Petroleum and Minerals.

Appendix A : Algorithms Derivatives Expressions

In all the expressions derived here the index ‘ i ’ runs over its full range for each case as defined in the text. We give the derivatives expressions w.r.t. the components of the parameter vector θ keeping in mind that the expressions are to be evaluated at the previous estimates $\hat{\theta}(t-1)$.

A.1 The Recursive Output Error (ROE) algorithm

The output error is defined as follows,

$$e_{OE}(t, \theta) = y(t) - \hat{y}(t)$$

Therefore,

$$\phi(t) = - \left[\frac{\partial e_{OE}(t)}{\partial \theta} \right]^T = \left[\frac{\partial \hat{y}(t)}{\partial \theta} \right]^T$$

According to Fig. 1 we have,

$$A(q^{-1})\hat{y}(t) = B(q^{-1})w(t) \quad (A.1)$$

$$w(t) = g_1v(t) + g_2v^2(t) + \dots + g_mv^m(t) \quad (A.2)$$

$$C(q^{-1})v(t) = D(q^{-1})u(t) \quad (A.3)$$

Derivatives of $\hat{y}(t)$ w.r.t. the ‘ a_i ’ parameters :

Taking the derivatives of both sides of Eq. (A.1)

w.r.t. to each of the 'a_i' parameters, followed by substitutions for unmeasured signals in terms of the input according to Eqs. (A.2), and (A.3), we get.

$$\frac{\partial}{\partial a_i} [A(q^{-1}) \hat{y}(t)] = A(q^{-1}) \frac{\partial \hat{y}(t)}{\partial a_i} + q^{-i} \hat{y}(t) = 0$$

Therefore,

$$\frac{\partial \hat{y}(t)}{\partial a_i} = \frac{q^{-i} B(q^{-1})}{A^2(q^{-1})} \left[\sum_{l=1}^m g_l \left(\frac{D(q^{-1})}{C(q^{-1})} u(t) \right)^l \right]$$

Derivatives of $\hat{y}(t)$ w.r.t. the 'b_i' parameters:

Taking the derivatives of both sides of Eq. (A.1) w.r.t. to each of the 'b_i' parameters, followed by substitutions for unmeasured signals in terms of the input according to Eqs. (A.2), and (A.3), we get.

$$\begin{aligned} \frac{\partial}{\partial b_i} [A(q^{-1}) \hat{y}(t)] &= A(q^{-1}) \frac{\partial \hat{y}(t)}{\partial b_i} \\ &= \frac{\partial}{\partial b_i} [B(q^{-1}) w(t)] = q^{-i} w(t) \end{aligned}$$

Therefore,

$$\frac{\partial \hat{y}(t)}{\partial b_i} = \frac{q^{-i}}{A(q^{-1})} \left[\sum_{l=1}^m g_l \left(\frac{D(q^{-1})}{C(q^{-1})} u(t) \right)^l \right]$$

Derivatives of $\hat{y}(t)$ w.r.t. the 'c_i' parameters:

Taking the derivatives of both sides of Eq. (A.1) w.r.t. to each of the 'c_i' parameters, followed by substitutions for unmeasured signals in terms of the input according to Eqs. (A.2), and (A.3), we get.

$$\begin{aligned} \frac{\partial}{\partial c_i} [A(q^{-1}) \hat{y}(t)] &= A(q^{-1}) \frac{\partial \hat{y}(t)}{\partial c_i} \\ &= \frac{\partial}{\partial c_i} [B(q^{-1}) w(t)] = B(q^{-1}) \frac{\partial w(t)}{\partial c_i} \\ \frac{\partial w(t)}{\partial c_i} &= \frac{\partial}{\partial c_i} \left[\sum_{l=1}^m g_l v^l(t) \right] = \left[\sum_{l=1}^m l \times g_l v^{l-1}(t) \right] \frac{\partial v(t)}{\partial c_i} \\ \frac{\partial}{\partial c_i} [C(q^{-1}) v(t)] &= C(q^{-1}) \frac{\partial v(t)}{\partial c_i} + q^{-i} v(t) \\ &= \frac{\partial}{\partial c_i} [D(q^{-1}) u(t)] = 0 \end{aligned}$$

Hence,

$$\frac{\partial v(t)}{\partial c_i} = -\frac{q^{-i}}{C(q^{-1})} v(t)$$

Therefore,

$$\begin{aligned} \frac{\partial \hat{y}(t)}{\partial c_i} &= -\frac{B(q^{-1})}{A(q^{-1})} \left[\sum_{l=1}^m l \times g_l v^{l-1}(t) \right] \\ &\times \frac{q^{-i} D(q^{-1})}{C^2(q^{-1})} u(t) \end{aligned}$$

Derivatives of $\hat{y}(t)$ w.r.t. the 'd_i' parameters:

Taking the derivatives of both sides of Eq. (A.1) w.r.t. to each of the 'd_i' parameters, followed by substitutions for unmeasured signals in terms of the input according to Eqs. (A.2), and (A.3), we get.

$$\begin{aligned} \frac{\partial}{\partial d_i} [A(q^{-1}) \hat{y}(t)] &= A(q^{-1}) \frac{\partial \hat{y}(t)}{\partial d_i} \\ &= \frac{\partial}{\partial d_i} [B(q^{-1}) w(t)] = B(q^{-1}) \frac{\partial w(t)}{\partial d_i} \\ \frac{\partial w(t)}{\partial d_i} &= \sum_{l=1}^m g_l v^l(t) = \sum_{l=1}^m l \times g_l v^{l-1}(t) \frac{\partial v(t)}{\partial d_i} \\ \frac{\partial}{\partial d_i} [C(q^{-1}) v(t)] &= C(q^{-1}) \frac{\partial v(t)}{\partial d_i} \end{aligned}$$

$$= \frac{\partial}{\partial d_i} [D(q^{-1}) u(t)] = q^{-i} u(t)$$

Hence,

$$\frac{\partial v(t)}{\partial d_i} = -\frac{q^{-i}}{C(q^{-1})} u(t)$$

Therefore,

$$\frac{\partial \hat{y}(t)}{\partial d_i} = -\frac{B(q^{-1})}{A(q^{-1})} \left[\sum_{l=1}^m l \times g_l v^{l-1}(t) \right] \frac{q^{-i}}{C(q^{-1})} u(t)$$

Derivatives of $\hat{y}(t)$ w.r.t. the 'g_i' parameters:

Taking the derivatives of both sides of Eq. (A.1) w.r.t. to each of the 'g_i' parameters, followed by substitutions for unmeasured signals in terms of the input according to Eqs. (A.2), and (A.3), we get.

$$\begin{aligned} \frac{\partial}{\partial g_i} [A(q^{-1}) \hat{y}(t)] &= A(q^{-1}) \frac{\partial \hat{y}(t)}{\partial g_i} \\ &= \frac{\partial}{\partial g_i} [B(q^{-1}) w(t)] = B(q^{-1}) \frac{\partial w(t)}{\partial g_i} \\ \frac{\partial w(t)}{\partial g_i} &= \frac{\partial}{\partial g_i} \left[\sum_{l=1}^m g_l v^l(t) \right] = v^i(t) \end{aligned}$$

Hence,

$$\frac{\partial \hat{y}(t)}{\partial g_i} = \frac{B(q^{-1})}{A(q^{-1})} \left(\frac{D(q^{-1})}{C(q^{-1})} u(t) \right)^i$$

Following similar procedure the expressions for the recursive equation error (REE) algorithm, and the recursive prediction error (RPE) algorithm as given in the text can be derived.

Appendix B: Proof of theorem on convergence

Referring to Eq. (19), the parameter estimates are given by,

$$\hat{\theta}(t) = \hat{\theta}(t-1) + P(t) \phi^T(t) \varepsilon(t) \quad (\text{B.1})$$

Let the parameter estimation error be,

$$r(t) = \hat{\theta}(t) - \theta_0$$

For the output error method we have,

$$\varepsilon(t, \theta) = e_{oe}(t, \theta) = y(t) - \hat{y}(t)$$

Since the true system belongs to the model set, then to a first order approximation we have,

$$r(t) = \lambda P(t) P^{-1}(t-1) r(t-1) + P(t) \phi^T(t) e(t) \quad (\text{B.2})$$

Let, $\bar{r}(t) = E\{r(t)/\mathfrak{F}_t\}$, and taking the conditional expectation of both sides of Eq. (B.2) w.r.t. $\mathfrak{F}_t$ and noting that $\mathfrak{F}_{t-1} \subset \mathfrak{F}_t$ we get,

$$\bar{r}(t) = \lambda P(t) P^{-1}(t-1) \bar{r}(t-1) \quad (\text{B.3})$$

From Eq. (18) it follows that,

$$\eta_j[P^{-1}(t)] \geq \lambda \eta_j[P^{-1}(t-1)], \forall j \quad (\text{B.4})$$

Hence, Eqs. (B.3) and (B.4) imply,

$$\|\bar{r}(t)\| \leq \left(\frac{\eta_{\max}[P(t)]}{\eta_{\min}[P(t)]} \right) \|\phi(t)\|$$

Therefore, under the conditions stated in the theorem and given by Eqs. (30) and (31) it follows that,

$$\lim_{t \rightarrow \infty} E\{r(t)/\mathfrak{F}_t\} = 0$$

i.e. $\lim_{t \rightarrow \infty} E\{\hat{\theta}(t)/\mathfrak{F}_t\} = \theta_0$

Let, $\bar{r}(t) = r(t) - \bar{r}(t)$ then, according to Eqs. (B.2) and (B.3) we have,

$$\hat{r}(t) = \lambda P(t)P^{-1}(t-1)\hat{r}(t-1) + P(t)\phi^T(t)e(t) \quad (\text{B.5})$$

And, let the parameter estimation error variance be,

$$R(t) = E\{\hat{r}^T(t)\hat{r}(t)/\mathfrak{F}_t\}$$

Then it follows from Eqs.(B.4) and (B.5) and the properties of the measurements noise that,

$$R(t) \leq \left(\frac{\eta_{\max}[P(t)]}{\eta_{\min}[P(t)]} \right)^2 R(t-1) + (\eta_{\max}[P(t)])^2 \sigma^2(\|\phi(t)\|)^2 \quad (\text{B.6})$$

Under the stability of the filters $\frac{B(q^{-1})}{A(q^{-1})}$, $\frac{B(q^{-1})}{F(q^{-1})}$,

$\frac{1}{A(q^{-1})}$, $\frac{1}{C(q^{-1})}$, $\frac{1}{F(q^{-1})}$, $B(q^{-1})$ and the boundedness of the system nonlinearity as described in the statement of the theorem, it follows that $\|\phi(t)\|$ is also bounded. Therefore, under the conditions given by Eqs.(30) and (31) it follows that,

$$\lim_{t \rightarrow \infty} R(t) = 0$$

Similarly, one can show the same convergence results to hold for the REE, and the RPE algorithms after making the following substitutions:

For the REE,

$$r(t) = \lambda P(t)P^{-1}(t-1)r(t-1) + P(t)\phi^T(t)A(q^{-1})e(t)$$

And, for the RPE,

$$r(t) = \lambda P(t)P^{-1}(t-1)r(t-1) + P(t)\phi^T(t)F^{-1}(q^{-1})A(q^{-1})e(t) \quad \square$$

References

- (1) Haber, R. and Unbehauen, H., Structure Identification of Nonlinear Dynamic Systems—A Survey on Input/Output Approaches, *Automatica*, Vol. 26, No. 4 (1990), pp. 651-677.
- (2) Palm, G., On Representation and Approximation of Nonlinear Systems Part II: Discrete Time, *Biological Cybernetics*, Vol. 34, No. 1 (1979), pp. 49-52.
- (3) Vandersteen, G., Rolain, Y. and Schoukens, J., Non-Parametric Estimation of The Frequency-Response Functions of The Linear Blocks of a Wiener-Hammerstein Models, *Automatica*, Vol. 33, No. 7 (1997), pp. 1351-1355.
- (4) Bendat, J.S., *Nonlinear System Analysis and Identification from Random Data*, (1990), John Wiley & Sons.
- (5) Kearney, R.E., Stein, R.B. and Parameswaran, L., Identification of Intrinsic and Reflex Contributions to Human Ankle Stiffness Dynamics, *IEEE Transactions on Biomedical Engineering*, Vol. 44, No. 6 (1997), pp. 493-504.
- (6) Korenberg, M.J., Parallel Cascade Identification and Kernel Estimation for Nonlinear Systems, *Annals of Biomedical Engineering*, Vol. 19 (1991), pp. 429-455.
- (7) Eskinat, E., Johnson, S.H. and Luyben, W.L., Use of Hammerstein Models in Identification of Nonlinear Systems, *AIChE Journal*, Vol. 37, No. 2 (1991), pp. 255-268.
- (8) Kalafatis, A.D., Wang, L. and Cluett, W.R., Identification of Wiener-type Nonlinear Systems in a Noisy Environment, *International Journal of Control*, Vol. 66, No. 6 (1997), pp. 923-941.
- (9) Liung, L. and Soderstrom, T., *Theory and Practice of Recursive Identification*, (1983), The MIT Press.
- (10) Soderstrom, T. and Stoica, P., *System Identification*, (1989), Prentice-Hall, Inc.
- (11) Goodwin, G.C. and Sin, K.S., *Adaptive Filtering, Prediction and Control*, (1984), Prentice-Hall, Inc.
- (12) Billings, S.A. and Fakhouri, S.Y., Identification of Nonlinear Systems Using Correlation Analysis and Pseudorandom Inputs, *International Journal of Systems Science*, Vol. 11, No. 3 (1980), pp. 261-279.
- (13) Yoshine, K. and Ishii, N., Nonlinear Analysis of a Linear-Non-Linear-Linear System, *International Journal of Systems Science*, Vol. 23, No. 4 (1992), pp. 261-279.
- (14) Emara-Shabaik, H.E., Ahmed, M.S. and Al-Ajmi, K.H., Parameter Estimation of Wiener-Hammerstein Models, *JSME Int. J., Ser. C*, Vol. 44, No. 1 (2001), pp. 118-124.
- (15) Chen, C.H. and Fassios, S.D., On The Estimation of Stochastic Wiener-Hammerstein Type Systems with Non-Smooth Non-linearity, *Proceedings of the American Control Conference*, Albuquerque, New Mexico, (1997).
- (16) Boutayeb, M. and Darouach, M., Recursive Identification Method for MISO Wiener-Hammerstein Model, *IEEE Transactions on Automatic Control*, Vol. 40, No. 2 (1995), pp. 287-291.
- (17) Leontaritis, I.J. and Billings, S.A., Experimental Design and Identifiability for Nonlinear Systems, *International Journal of Systems Science*, Vol. 18, No. 1 (1987), pp. 189-202.