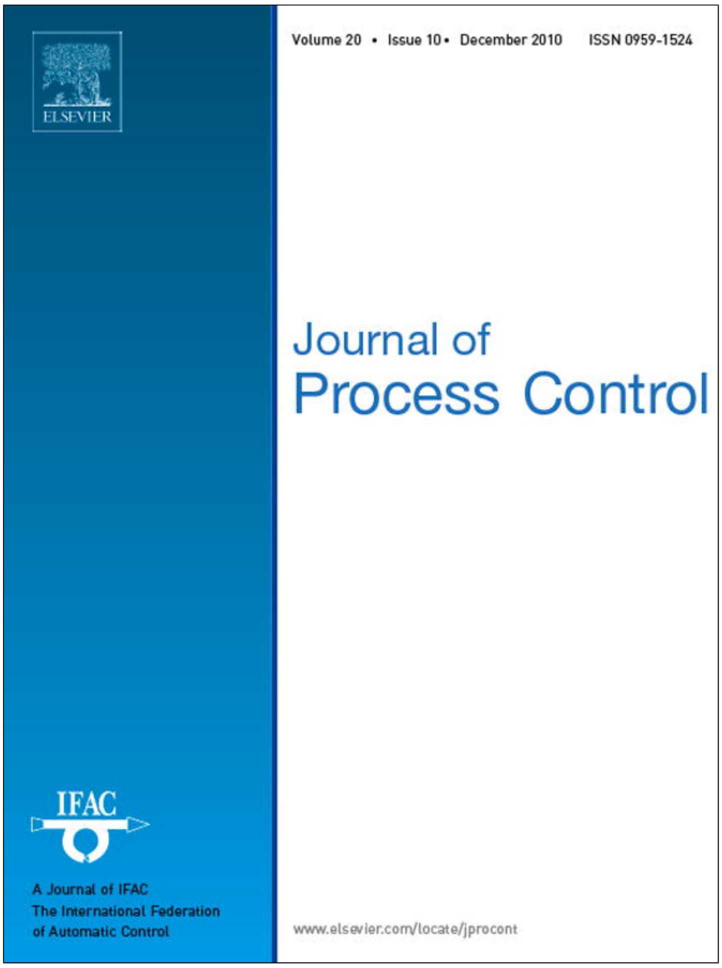
Volume 20 • Issue 10 • December 2010 ISSN 0959-1524
ELSEVIER
Journal of
Process Control
IFAC
A Journal of IFAC
The International Federation
of Automatic Control
www.elsevier.com/locate/jprocont

Journal of Process Control 20 (2010) 1235–1242

Contents lists available at ScienceDirect

Journal of Process Control

journal homepage: www.elsevier.com/locate/jprocont

Control performance assessment for nonlinear systems

W. Yu [a,*], D.I. Wilson [b], B.R. Young [c]

[a] Industrial Information & Control Centre, The University of Auckland, Auckland, New Zealand
[b] Electrical & Electronic Engineering, Auckland University of Technology, Auckland, New Zealand
[c] Department of Chemical & Materials Engineering, The University of Auckland, Auckland, New Zealand

ARTICLE INFO

Article history:
Received 22 June 2010
Received in revised form 24 August 2010
Accepted 3 September 2010

Keywords:
Control performance assessment
Nonlinear systems
ANOVA-like variance decomposition

ABSTRACT

Assessing the quality of existing industrial control loops, or comparing between two alternative controller designs is becoming an important routine auditing task for the control engineer. While most of the research and commercial activity in CPA has been applied to linear systems to date, those researchers investigating nonlinear systems fall into one of two groups. The first group focussed on the diagnosis of a common specific nonlinearity, namely valve stiction [1–3], while the second group tried to establish the minimum variance performance lower bound (MVPLB) [4–8]. In this paper we will propose a new CPA performance index for general nonlinear models based on an ANOVA-like variance decomposition method. The results of two simulation examples illustrate that the proposed methodology is efficient and accurate.

1. Introduction

On a recent site visit to a major dairy processing plant, we found over 500 control loops were under the maintenance supervision of a single instrument engineer. At that ratio, which from anecdotal evidence is by no means unusual in the industry, a realistic period between loop inspections is in the order of years. Consequently it is not surprising that engineers are overwhelmed by the sheer number of loops that need attention on any typical industrial processing plant as noted by control audits [9,10].

Control performance assessment, or CPA, is a technology to diagnose and maintain operational efficiency of control systems developed in a direct response to address this increasingly important economic problem. CPA is routinely applied in the refining, petrochemicals, pulp and paper and the mineral processing industry as noted by [11–14], although these and many related publications, are primarily restricted to linear systems.

In practice, industrial control loops invariably include nonlinearites from the control valve, the measurement, or the process itself. The estimates of the minimum variance performance lower bound (MVPLB) and the performance index using the linear CPA techniques may be distorted by these nonlinearities. For example, for a dynamic linear system with an additive linear disturbance, if this system has a valve stiction problem, it is shown that the estimates of the performance index using linear CPA techniques will provide over-estimation [8]. To deal with this situation, recent research has proposed several methods to extend CPA into nonlinear systems [6–8].

In the case of nonlinear systems, [4] superimpose a nonlinear dynamic model to an additive linear or partially nonlinear disturbance. It is shown that a minimum variance feedback invariant exists for a class of nonlinear models and the MVPLB can be estimated from routine operating data. Continuing this idea, estimations of the MVPLBs for the moderate valve stiction cases are proposed by [6–8]. These applications are based on one general nonlinear structure: nonlinearity in the dynamics or nonlinearity caused by a static nonlinearity from manipulated variables plus an additive disturbance which is in form of an ARMA model.

In this paper, we will propose a new performance index which can be used for general single-input single-output (SISO) nonlinear systems. While for linear systems, the MVPLB can be estimated through the impulse response functions (IRF) since there is a direct relationship between the impulse response and the variance, this is not true for nonlinear systems. The general form of the MVPLB for nonlinear systems may be very complex and numerically difficult to estimate.

Conversely the performance index proposed in this work is based on an ANOVA-like variance decomposition method. Historically the ANOVA-like variance decomposition method was used to provide variance analysis for nonlinear systems with the multi-disturbance sources [15]. However instead of different disturbance sources, the ANOVA-like variance decomposition will be used in this paper with respect to the time horizon for the SISO nonlinear system with one disturbance.

The layout of the paper is as follows. In Section 2, a general class of nonlinear model, the NARMAX model, is introduced for the study of CPA. Section 3 provides the new performance index

* Corresponding author.
E-mail addresses: w.yu@auckland.ac.nz (W. Yu), david.i.wilson@aut.ac.nz (D.I. Wilson), b.young@auckland.ac.nz (B.R. Young).

0959-1524/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jprocont.2010.09.002

and the ANOVA-like variance decomposition method. Section 4 outlines a numerical Monte Carlo method used to estimate this performance index. In Section 5, two simulations are used to illustrate the proposed methodology. This is followed by a discussion and conclusions highlighting both the limitations and potential of the proposed methods.

2. Process description

Stochastic nonlinear dynamic systems are difficult to identify both because of the paucity of accurate mechanistic models and the fact that full state variables are rarely available. Therefore when one is forced to use only external data [16–21], then the flexibility of nonlinear autoregressive moving average with exogenous input (NARMAX) models proposed by [19,20] are attractive. NARMAX models also provide a unified representation for a wide class of discrete time nonlinear dynamic/stochastic systems suitable for the types of processes found in industry in CPA applications. Consequently, in this paper, the general SISO discrete nonlinear system will be represented by a NARMAX model

$$y_t = f(y^*_{t-1}, u^*_{t-b}) \tag{1}$$

where y_t is the deterministic output of the system in response to the inputs, u_t. The integer b represents the number of whole periods of delay in the system and is the number of sampling intervals that elapse between making a change in the process input and first observing its effect. The superscript $*$ is used to represent the vector collecting the immediate historical values, i.e. $y^*_{t-1} \overset{\text{def}}{=} (y_{t-1}, \cdots, y_{t-n_z})$.

Given that the relation $f(\cdot)$ is often complex, one typically approximates the nonlinearity with a simpler generic function, often polynomials consisting of summations of terms involving $y_{t-i}, y_{t-i}y_{t-j}, \ldots, u_{t-b-i}, u_{t-b-i}u_{t-b-j}, \ldots$, and cross terms $y_{t-i}u_{t-b-j}, \ldots$. The flexible and popular Hammerstein and Wiener systems are included in this framework [22,23].

In this work, we are also interested in systems affected by disturbances. The most elementary representation is

$$y_t = f_2(y^*_{t-1}, u^*_{t-b}) + a_t \tag{2}$$

where a_t is white noise with zero mean and constant variance σ_a^2.

Eq. (2) provides the simplest stochastic nonlinear system since only additive uncorrelated noise is considered. In reality, the disturbance might include autoregressive and moving average terms or cross-products between the disturbance and the inputs and outputs. The most general form of the nonlinear stochastic system is the NARMAX model,

$$y_t = f_3(y^*_{t-1}, u^*_{t-b}, a^*_t) \tag{3}$$

3. Alternative performance indices for nonlinear systems

3.1. Existing performance indices for linear systems

The performance index proposed for SISO linear systems in [24,25] was based on the concept of minimum variance control [26,27]. This performance index was defined as the ratio of the best achievable variance to the variance of the controlled variable under assessment. What made this concept useful was the insight of [24] which showed that these performance indices can be estimated directly from the routing operating data by fitting the controlled variable into a ARIMA time series model.

Many important industrial processes can be represented by a linear transfer function with an additive disturbance

$$y_t = \frac{\omega(q^{-1})}{\delta(q^{-1})} q^{-b} u_t + d_t \tag{4}$$

where y_t is the deviation of the process variable from its setpoint, u_t is the manipulated variable, b is the number of whole periods of process delay, and $\omega(q^{-1})$ and $\delta(q^{-1})$ are polynomials in the backward shift operator q^{-1}. The disturbance d_t is usually represented by the Autoregressive-Integrated-Moving-Average (ARIMA) models [27,28] as,

$$d_t = \frac{\theta(q^{-1})}{\phi(q^{-1})\nabla^d} a_t = \Psi(q^{-1}) a_t \tag{5}$$

where the a_t's are a sequence of independently and identically distributed (iid) random variables with mean zero and constant variance σ_a^2. The difference operator is defined as $\nabla \overset{\text{def}}{=} (1 - q^{-1})$. We assume that all the polynomials are assumed to be stable and δ, θ and ϕ are monic.

The minimum variance control first derived by [26,27] is a feedback controller which achieves minimum output variance. To derive the minimum variance controller, we need to know the b-step ahead minimum-mean-square-error forecast for the y_{t+b},

$$y_{t+b} = \frac{\omega(q^{-1})}{\delta(q^{-1})} u_t + d_{t+b} = \frac{\omega(q^{-1})}{\delta(q^{-1})} u_t + d_{t+b|t} + e_{t+b|t} = y_{t+b|t} + e_{t+b|t} \tag{6}$$

where $d_{t+b|t}$ and $y_{t+b|t}$ are the b-step ahead minimum-mean-square-error forecast for the disturbance and y_{t+b}, respectively and the prediction error, $e_{t+b|t}$, is a moving average process of order $b-1$ as

$$e_{t+b|t} = (1 + \psi_1 q^{-1} + \cdots + \psi_{b-1} q^{-(b-1)}) a_{t+b} \tag{7}$$

where the ψ weights are identical with the first b impulse coefficients of the disturbance transfer function in (5).

The control signal which results in the minimum achievable variance in the output can be obtained by solving

$$\frac{\omega(q^{-1})}{\delta(q^{-1})} u_t + d_{t+b|t} = 0 \tag{8}$$

from which it is evident that the process output under minimum variance control y^{MV}_{t+b}, will depend on only the most recent b past disturbances, i.e.,

$$y^{MV}_{t+b} = e_{t+b|t} \tag{9}$$

The key observation is no matter which feedback controller is used, as long as closed-loop stability is preserved, the prediction error, $e_{t+b|t}$ is unaffected. The minimum variance performance lower bound (MVPLB), as measured in the mean square sense, can be written as

$$\sigma^2_{MV} = \text{var}\{y^{MV}_{t+b}\} = (1 + \psi_1^2 + \ldots + \psi_{b-1}^2)\sigma_a^2 \tag{10}$$

The most attractive feature of CPA using minimum variance control as a benchmark is that the performance lower bounds can be estimated by using only routine closed-loop operating data with *a priori* knowledge of time delay [24]. The performance index, often termed the Harris Index, η, is defined as

$$\eta = \frac{\sigma^2_{MV}}{\sigma^2_y} \tag{11}$$

where $0 < \eta < 1$, σ_y^2 is the actual operating variance of the controlled variable under assessment, and σ^2_{MV} is the theoretical minimum variance lower bound. A performance index closes to 0 implies that there is a high potential for reducing the output variance by retuning the existing controller or implementing an improved control algorithm.

3.2. MVPLB-based performance indices for nonlinear systems

The MVLPB-based performance index reviewed in Section 3.1 designed for linear systems can be extended to a class of nonlinear systems which includes a nonlinear dynamic model and an additive linear or partially nonlinear disturbance as proposed in [4]. There it was shown that a minimum variance feedback invariant still exists meaning that the MVPLB can be also estimated from routine operating data. Continuing this idea of a MVPLB-based index, estimations of the MVPLBs for the moderate valve stiction cases are proposed by [6–8].

However there are significant theoretical difficulties in extending the MVPLB-based index to general nonlinearities such as NARMAX models. Quite apart from the fact that the MVPLB may be very difficult to estimate or even not exist, idea of variance decomposition using the impulse response function and the concept of feedback invariance is not necessarily valid for nonlinear systems. To illustrate these problems let us try to develop the MVPLB for the NARMA models in Eq. (3). With a feedback controller, $u_t = g(y_t, \cdots, y_{t-n_y})$, the closed-loop of the system in Eq. (3) can be written as,

$$y_t = f_4(y_{t-1}, \cdots, y_{t-n_y}, a_t, \cdots, a_{t-n_a}) \tag{12}$$

Using the concept of MVPLB, we are interested in y_{t+b} and its prediction $y_{t+b|t}$. $y_{t+b|t} = E\{y_{t+b|t}\}$ is the optimal predictor that minimizes the variance of the b-step ahead prediction error. It is also known as the conditional mean, and is obtained by taking the expectation of future values of y_{t+b}, $b > 0$, using available information up to and including time t.

Associated with the predictor, are the prediction errors,

$$\begin{aligned} e_{t+b|t} &= y_{t+b} - y_{t+b|t} = f_4(y_{t+b-1}, \cdots, y_{t+b-n_y}, a_{t+b}, \cdots, a_{t-n_a}) \\ &\quad - E\{f_4(y_{t+b-1}, \cdots, y_{t+b-n_y}, a_{t+b}, \cdots, a_{t-n_a})|t\} \end{aligned} \tag{13}$$

where the variance of the prediction error is,

$$\sigma_e^2(b) = E\{e_{t+b|t}^2\} \tag{14}$$

if the process is stationary, then $\lim_{b\to\infty}\sigma_e^2(b)$ is equal to the variance of y, $V(y)$ [29].

It is clear that in general the b-step ahead prediction error $e_{t+b|t}$ not only depends on the disturbance a_{t+b-i}, $i = 0, 1, \cdots, b-1$ but also the past output values y_{t-i}, $i = 1, 2, \ldots$ and the past disturbances a_{t-i}, $i = 0, 1, \ldots$. So the variance of $e_{t+b|t}$ is not the minimum variance performance lower bound for the general case. Furthermore, the interpretation of the variance of $e_{t+b|t}$ becomes more difficult given the complexity of the nonlinear systems, so we need another method which can provide us the explicit variance decomposition for the complex nonlinear systems. For this reason we propose an ANOVA-like variance decomposition method to provide the new performance index.

3.3. ANOVA-based performance indices

For the output of a static system considered represented as an analytic function of p input variables, e.g., $Y = f(X_1, X_2, \cdots, X_p)$, the relative importance of the independent inputs can be quantified by the fractional variance which is defined as the fractional contribution to the output variance due to the uncertainties in inputs. This can be calculated using an ANOVA-like decomposition formula for the total output variance $V(Y)$ of the output Y [30,31] as

$$V(Y) = \sum_i V_i + \sum_i \sum_{j>i} V_{ij} + \cdots + V_{12\cdots p} \tag{15}$$

where

$$\begin{aligned} V_i &= V(E(Y|X_i = x_i)) \\ V_{ij} &= V\left(E(Y|X_i = x_i, X_j = x_j)\right) - V(E(Y|X_i = x_i)) - V(E(Y|X_j = x_j)) \end{aligned} \tag{16}$$

and so on, where $E(Y\,|\,X_i = x_i)$ denotes the expectation of Y conditional on X_i having a fixed value x_i, and V stands for variance over all the possible values of x_i.

For dynamic systems, we can take an analogous approach where we separate the disturbance entering the system in Eq. (12) after time 0, say $[a_{t+b}, a_{t+b-1}, \cdots, a_1]$, into two groups: $x_1 = [a_{t+b}, \cdots, a_{t+1}]$ and $x_2 = [a_t, a_{t-1}, \cdots, a_1]$. The first group includes all the disturbances entering the system after time t and the second group includes all the disturbances entering the system up to and including time t and including time t starting from the initial time $t = 0$.

To decompose the variance of the nonlinear process using the ANOVA-like decomposition method in Eq. (15), we use the following assumptions:

- Eq. (12) can be computed subject to the initial condition, I_0, to give y_{t+b-i}, $i = 1, 2, \ldots, t+b-1$ for any choice of $a_{t+b}, a_{t+b-1}, \ldots, a_1$.
- The initial condition I_0 is a random vector with probability density function $P(I_0)$ uncorrelated and independent from subsequent disturbances entering the system.

We are interested in determining the sensitivity of output y_{t+b} in Eq. (12) to variations of two vector series x_1 and x_2. Since the future behavior of y_{t+b} is dependent on initial conditions due to the nonlinearity, we cannot use the ANOVA-like decomposition in Eqs. (15) and (16) directly since the initial condition must be considered within the variance decomposition. Instead using the well-known variance decomposition theorem [32], we can decompose the variance of y_{t+b} into two terms

$$V[y_{t+b}] = E_{I_0}[V_x[y_{t+b}|I_0]] + V_{I_0}[E_x[y_{t+b}|I_0]] \tag{17}$$

where $x = [x_1, x_2]$ denotes all of disturbances entering the system from time 1 to time t, $E_{I_0}[\cdot]$ denotes the expectation of $[\cdot]$ with respect to I_0 and $V_{I_0}[\cdot]$ denotes the variance of $[\cdot]$ with respect to I_0. Given that Eq. (17) is the sum of positive numbers, it follows that $V[y_{t+b}] \geq E_{I_0}[V_x[y_{t+b}|I_0]]$. The special situation where $V[y_{t+b}] = E_{I_0}[V_x[y_{t+b}|I_0]]$, will be discussed in the following paragraphs.

The first term in Eq. (17) is the fractional contribution to the variance of y from the disturbance signal and the interaction between disturbance and the initial condition. The second term is the fractional contribution to the output *solely* due to the uncertainties in the initial condition. The conditional variance given initial condition I_0, $V_x[y_{t+b}\,|\,I_0]$, can be decomposed directly using the ANOVA-like decomposition method as:

$$V_x|I_0 = V_x[y_{t+b}|I_0] = V_1|I_0 + V_2|I_0 + V_{12}|I_0 \tag{18}$$

where

$$\begin{aligned} V_1|I_0 &= V_{x_1}[E_{x_2}[y_{t+b}|(x_1, I_0)]] \\ V_2|I_0 &= V_{x_2}[E_{x_1}[y_{t+b}|(x_2, I_0)]] \\ V_{12}|I_0 &= V_x[E_x[y_{t+b}|(x, I_0)]] - V_1|I_0 - V_2|I_0 \end{aligned} \tag{19}$$

The variance decomposition with consideration of the initial condition can be obtained by simply calculating the expectation of the conditional variance decomposition in Eq. (18) with respect to the initial condition I_0. This procedure is not necessary if the initial condition has (can be approximately assumed to have) a linear relationship with the output Y_t. The variance decomposition can be calculated with the results of the conditional variance decomposition in Eq. (18) based on the mean values of initial condition. Further information about this topic can be found in [33]. $E_{I_0}[V_1|I_0]$ denotes the main effect of x_1 on the $V[y_{t+b}]$ and $E_{I_0}[V_{12}|I_0]$ is the interaction contributing to the $V[y_{t+b}]$ that is not accounted for the main effects of x_1 and x_2. Consequently we propose a suitable performance index as

$$\eta_t = \frac{E_{I_0}[V_1|I_0]}{V[y_{t+b}]} \tag{20}$$

While this index is applicable for any nonlinearity, the index has little worth for those processes that are strongly non-ergodic where the second term in Eq. (17) dominates. However such cases are more pathological than common in industry.

If the NARMA model is stationary, the distribution of $\lim_{t\to\infty} y_{t+b}$ reaches an equilibrium. For stationary linear time series, this limiting distribution is independent of initial condition but for a stationary nonlinear model, the limiting distribution, and hence, $V[y]$, may depend on the initial condition. Therefore, the performance index in Eq. (20) will depend on the initial condition. If the distribution of $\lim_{t\to\infty} y_{t+b}$ does in fact not depend on the initial conditions, the process is termed ergodic [29]. Sufficient conditions to determine the ergodicity of a nonlinear time series are given in ([29], p. 127). For an ergodic nonlinear time series, $V_{I_0}[E_x[Y_{t+b}|I_0]]$ in Eq. (17) will be zero for $t\to\infty$, the variance decomposition can be expressed when $t\to\infty$ as

$$V[y_{t+b}] = E_{I_0}[V_1 + V_2 + V_{12}|I_0] = V_1 + V_2 + V_{12} \tag{21}$$

where

$$\begin{aligned} V_1 &= V_{x_1}[E_{x_2}[y_{t+b}|x_1]] \\ V_2 &= V_{x_2}[E_{x_1}[y_{t+b}|x_2]] \\ V_{12} &= V[y_{t+b}] - V_1 - V_2 \end{aligned} \tag{22}$$

The performance index associated to the ergodic nonlinear system will be,

$$\lim_{t\to\infty} \eta_t = \lim_{t\to\infty} \frac{V_1}{V[y_{t+b}]} \tag{23}$$

We will approximate the infinite limit in Eq. (23) by some suitably large value, η_M.

However this performance index is not practical since the nonlinear system may not be ergodic or even stationary, and even if the nonlinear time series is ergodic, the ergodicity is also very difficult to be determined. Finally calculating the approximation of η_t for $t\to\infty$, η_M, requires enormous computational cost. We suggest to pick up several η_t, $t=c_1, c_2, \cdots, c_m$, $c_j \ll M$, $j=1, \cdots, m$ and to use the average of η_i as the performance index,

$$\eta = \frac{1}{m} \sum_{t=c_i, i=1}^{m} \eta_t \tag{24}$$

Remarks.

- This new performance index defined in Eq. (24) may depend on the controller, initial condition and the length of time t.
- The partial variance $E_{I_0}[V_1|I_0]$ is equal to the minimum variance performance lower bound if the closed loop system is linear and stationary.
- The partial variance $E_{I_0}[V_1|I_0]$ is equal to the minimum variance performance lower bound for some nonlinear systems such as those discussed in [4].
- This performance index is strictly bounded in [0, 1]. If η reaches 1, it means that the variance of outputs is contributed mostly by the x_1, so the system controller is close to the minimum variance controller.
- If we build the performance index defined in Section 3.1 for the general NARMAX model, we will encounter two problems: (i) the MVPLB may not exist; (ii) the MVPLB is very difficult to estimate. The new performance index in Eq. (24) provides the more reliable solution.

We know that the variance of output y_{t+b}, $V[y_{t+b}]$, can be decomposed into two parts: $E_{I_0}[V_x|I_0]$ and $V_{I_0}[E_x[Y_{t+b}|I_0]]$ (see Eq. (17)). If the second term dominates the $V[y_{t+b}]$, the performance index in Eq. (24) will be meaningless. For this circumstance, we suggest to use the variance decomposition in Eq. (18) directly to obtain a performance index as,

$$\tilde{\eta}_t = \frac{V_1|I_0}{V_x|I_0} \tag{25}$$

Unlike performance indices based on linear systems, this index is only strictly applicable for a given initial condition. Consequently in practice one may re-compute the index over a range of anticipated initial conditions.

4. Computing the performance index

The practical computation of the performance index Eq. (25) requires two steps. First one must estimate the closed loop model defined in Eq. (12), and then use a Monte Carlo strategy to approximate performance index.

For the first identification step, several strategies have been proposed such as orthogonal least squares (OLS) methods [34] and fast orthogonal search (FOS) methods [35,36], and approximations based on Artificial Neural Network (ANN) models are discussed in [37,38].

Since the disturbance term in Eq. (12) is generally unmeasured, the identification will require an iterative approach. The identification procedures will be: (i) set the initial sequence a_t by fitting a linear model or setting the a_t to zero, (ii) identify the NARMA model, (iii) replace the initial sequence a_t by the prediction errors or residuals, (vi) repeat the steps (ii) and (iii) until a certain identification criteria is achieved.

One suitable criteria is Akaike's Information Criterion AIC(s) [34],

$$\mathrm{AIC}(\lambda) = K \ln \hat{\sigma}^2_{\Xi} + N\lambda \tag{26}$$

where N is the number of the model parameters, K is the number of outputs and $\hat{\sigma}^2_{\Xi}$ is the residual error. The tuning parameter λ is a positive value chosen to provide a penalty for model complexity and using statistical arguments, a value of $\lambda = 4$ is recommended in [34,39].

4.1. Estimating the performance index

Once the parameters of the closed loop are estimated, we may try to obtain analytical solutions using the methods proposed in Section 3.2. However a more practical way is to use a Monte Carlo (MC) method which while computationally demanding due to the potentially high dimensionality [40], can be offset somewhat by employing alternative strategies such as the Fourier Amplitude Sensitivity Test (FAST) [41,42] and Sobol's method [43,44].

For a model of the form given by Eq. (12), the MC estimates of the mean and variance of y_{t+b} given the initial condition I_0 can be calculated by,

$$\hat{\bar{y}}_{t+b}|I_0 \simeq \frac{1}{N} \sum_{k=1}^{N} f_4(x_1^{(k)}, x_2^{(k)})|I_0 \tag{27}$$

$$\hat{V}_x|I_0 \simeq \frac{1}{N} \sum_{k=1}^{N} (f_4(x^{(k)})|I_0)^2 - (\hat{\bar{y}}_{t+b}|I_0)^2 \tag{28}$$

where $x^{(k)} = (x_1^{(k)}, x_2^{(k)})$ is a set of N simulations of multidimensional inputs that have the requisite distribution. The partial variance

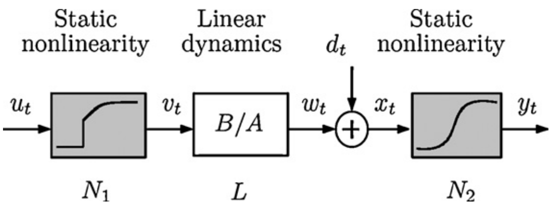

Fig. 1. The Hammerstein–Wiener model structure used in this paper.

$V_1\,|\,I_0$ can be estimated as [30,45],

$$\hat{V}_1|I_0 = \frac{1}{N}\sum_{k=1}^{N} f_4(x_1^{(k)}, x_2^{(k)}) f_4(x_1^{(k)}, \vec{x}_2^{(k)}) - (\hat{\bar{y}}_{t+b}|I_0)^2 \tag{29}$$

$\vec{x}_2^{(k)}$ is independent of the set of $x_2^{(k)}$. Hence Eq. (29) means that for computing $V_1\,|\,I_0$ we multiply values of f_4 corresponding to a set $(x_1^{(k)}, x_2^{(k)})$ by values of f_4 from a different set $(x_1^{(k)}, \vec{x}_2^{(k)})$ which includes the same partial set $(x_1^{(k)})$. To calculating the $\hat{V}_1|I_0$ with the different initial conditions, the average of these values can be used as the estimates of $E_{I_0}[V_1|I_0]$ in Eq. (20). The other two partial variances: $E_{I_0}[V_2|I_0]$ and $E_{I_0}[V_{12}|I_0]$ can be computed in the same manner.

5. Simulation experiments

The purpose of this section is to demonstrate the proposed method to estimate the new performance index for the nonlinear systems. We will select one class of block-oriented nonlinear models, Hammerstein–Wiener models (shown in Fig. 1) as the simulation examples. Nonlinear block-oriented models consist of the interconnection of a linear time invariant (LTI) systems with static, or memoryless, nonlinearities. This class includes Hammerstein models, Wiener models and combinations of the two [18,22]. Such block-oriented nonlinear descriptions are very useful modelling input/output nonlinearities and have been implemented many industrial processes (i.e. [46–49]). In the first example, a special case of the HW model, a Wiener model, is used to test the proposed approach. In the second example, a HW model will be used.

5.1. One theoretical performance index

The main reason to select the HW model is that we can obtain a theoretical performance index as the basis which can be used to evaluate the proposed new performance index.

We assume the plant can be adequately modelled by a Hammerstein–Wiener model (shown in Fig. 1) as,

$$y_t = N_2(x_t) \tag{30}$$

$$x_t = w_t + d_t \tag{31}$$

$$w_t = \frac{B(q^{-1})}{A(q^{-1})} q^{-b} v_t \tag{32}$$

$$v_t = N_1(u_t) \tag{33}$$

where $A(q^{-1})$ and $B(q^{-1})$ are polynomials in the backshift operator q^{-1}, and b is the time delay of the system. The signals u_t and y_t are the process input and output, respectively, and the internal signals v_t, w_t and x_t are assumed nonmeasurable. The functions N_1 and N_2 represents the static nonlinearities for input and output, respectively. The disturbance d_t is modelled as the same ARIMA model defined in Eq. (5).

The derivation of the minimum variance controller with respective to x_t for a process described by Eqs. (30)–(33) is straightforward [50,4]. The series x_{t+b} can be written as:

$$\begin{aligned} x_{t+b} &= \frac{B(q^{-1})}{A(q^{-1})} N_1(u_t) + x^{ss} + d_t = \frac{B(q^{-1})}{A(q^{-1})} N_1(u_t) + d_{t+b|t} + x^{ss} + e_{t+b|t} \\ &= x_{t+b|t} + x^{ss} + e_{t+b|t} \end{aligned} \tag{34}$$

where x^{ss} is a constant steady-state value (if there is no offset and disturbance, $N_2(x^{ss}) = y_{sp}$), $d_{t+b|t}$ is the b-step ahead prediction in form of,

$$d_{t+b|t} = \frac{P_b(q^{-1})}{\phi(q^{-1})\nabla^h} a_t \tag{35}$$

and $e_{t+b|t}$ is the b-step ahead prediction error as,

$$e_{t+b|t} = (1 + \psi_1 q^{-1} + \cdots + \psi_{b-1} q^{-(b+1)}) a_{t+b} \tag{36}$$

$P_b(q^{-1})$ is a polynomial in the backshift operator obtained by solving the Diophantine equation:

$$\frac{\theta(q^{-1})}{\phi(q^{-1})\nabla^h} = 1 + \psi_1 q^{-1} + \cdots + \psi_{b-1} q^{-b+1} + \underbrace{q^{-b}\frac{P_b(q^{-1})}{\phi(q^{-1})\nabla^h}}_{\text{remainder}} \tag{37}$$

From Eq. (34), we can find that the b-step prediction error, $e_{t+b|t}$, is the control invariant. The control signal which results in the minimum achievable variance in the x_t can be obtained by solving the following relation:

$$\frac{B(q^{-1})}{A(q^{-1})} N_1(u_t) + d_{t+b|t} = 0 \tag{38}$$

Therefore, x_t under minimum variance control, x^{MV}, will depend on only the most recent b past disturbances,

$$x^{MV} = x^{ss} + e_{t+b|t} \tag{39}$$

The process output under this MVC is

$$y^{\lozenge} = N_2(x^{MV}) \tag{40}$$

Associating this $y^{\lozenge}$, the theoretical performance index is defined as,

$$\eta^{\lozenge} = \frac{\sigma^2_{y^{\lozenge}}}{\sigma^2_y} \tag{41}$$

5.2. A Wiener model

A pH neutralization system modelled as a Wiener process shown in Fig. 2 from [46] is adopted as a simulation test. The linear plant is

$$\frac{B(q^{-1})}{A(q^{-1})} = \frac{0.0049 - 0.0094q^{-1} + 0.0045q^{-2}}{1 - 2.9160q^{-1} + 2.8339q^{-2} - 0.9179q^{-3}} \tag{42}$$

with time delay $b = 3$ controlled using a PI feedback controller

$$G_c = \frac{0.1 - 0.5q^{-1}}{1 - q^{-1}} \tag{43}$$

This plant is subjected to an additive disturbance of

$$d_t = \frac{a_t}{1 - 0.8q^{-1}} \tag{44}$$

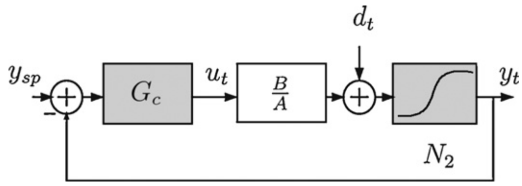

Fig. 2. A Wiener model structure in closed loop.

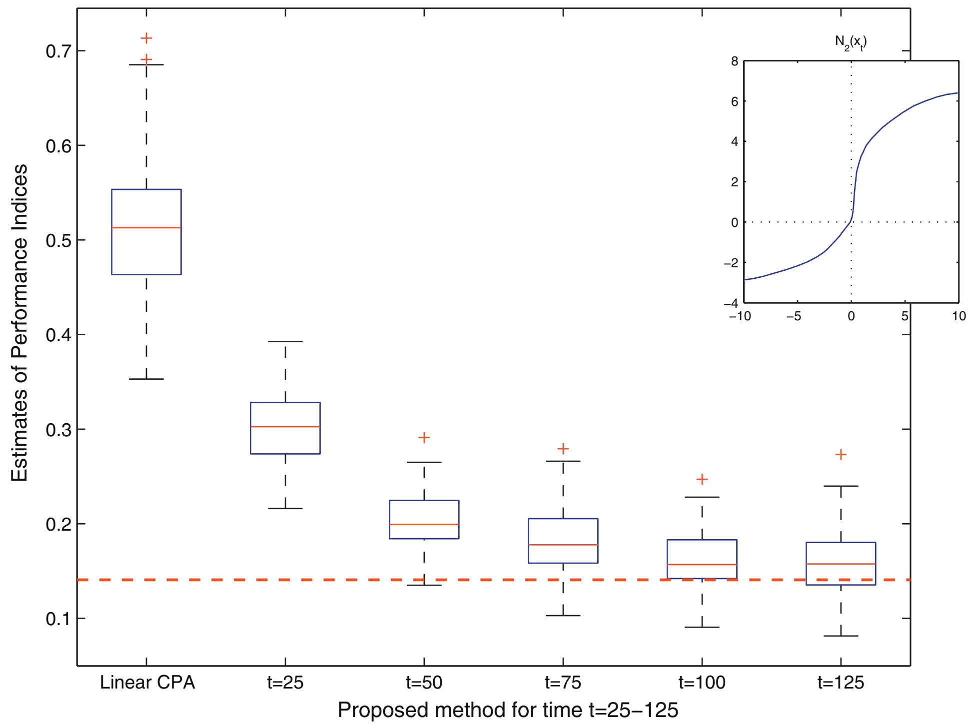

Fig. 3. Comparative box plots of the quality estimates of the performance index for a Wiener model. The horizontal dashed line is the theoretical performance index of 0.1404.

where a_t is a sequence of independent and identically distributed Gaussian random variables with zero mean and nominal variance $\sigma_a^2 = 0.01$. The static nonlinearity N_2 is plotted in the right-hand corner of Fig. 3.

We use the program `nlarx` from the System Identification Toolbox to estimate the NARMA model. In this function, the dynamic structure and the nonlinearity estimators are the two main design choices. The choice of nonlinearity estimators is very often arbitrary and needs several trials before getting a satisfactory result. Readers are referred to [51] or the help manual for implementation details.

Since the time t may affect the results of performance index in Eq. (20), we select 5 time lengths of 25, 50, 75, 100 and 125. Using the Monte Carlo method in Section 4.1, the estimates of the performance index in Eq. (20) are plotted in Fig. 3 where we also

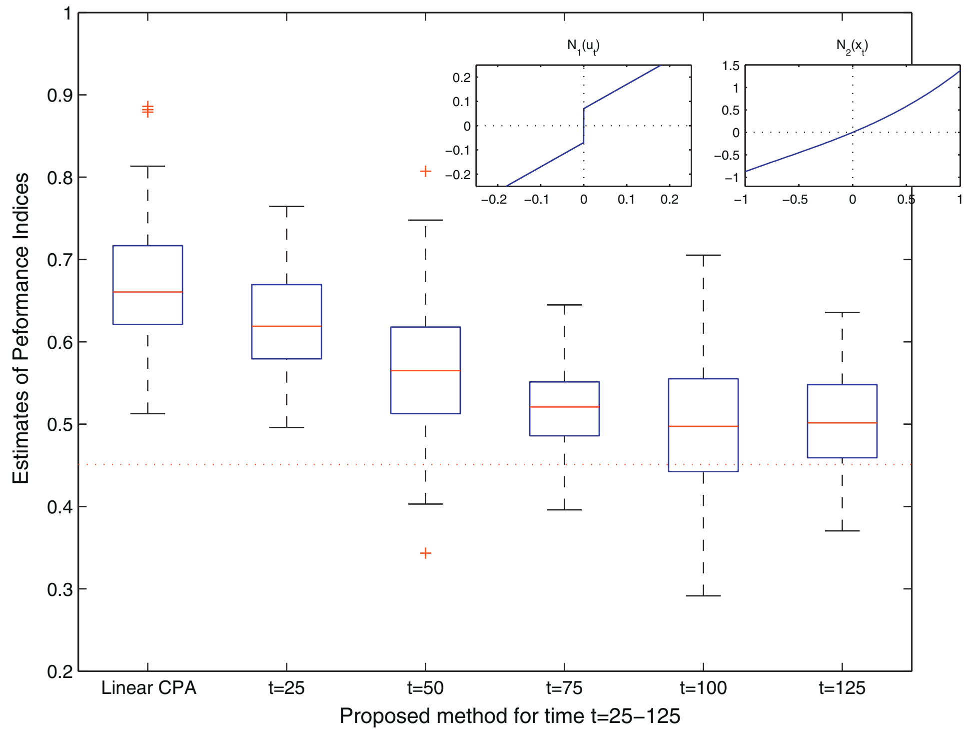

Fig. 4. Comparative box plots of the quality estimates of the performance index for a HW model.

compare the performance index estimated using linear CPA techniques. The theoretical performance index which is used as the basis is $\eta^\star = 0.1404$.

5.3. A Hammerstein–Wiener model

In this section, a HW model with linear dynamics

$$\frac{q^{-b}B(q^{-1})}{A(q^{-1})} = \frac{q^{-3}(1-0.5q^{-1})}{1-1.5q^{-1}+0.7q^{-2}} \tag{45}$$

under a feedback control with a PI controller,

$$G_c = \frac{0.05-0.02q^{-1}}{1-q^{-1}} \tag{46}$$

The first static nonlinearity, N_1, is a coulombic and viscous friction nonlinearity, $v_t = \text{sgn}(u_t)(|u_t| + 0.07)$, where 0.07 is the offset, and the trailing nonlinearity, N_2, is a third order polynomial, $y_t = x_t + 0.25x_t^2 + 0.125x_t^3$. Both N_1 and N_2 are inserted in Fig. 4. The same disturbance model structure used in the Wiener example is adopted for this simulation, the only difference being that the variance of a_t, $\sigma_a^2 = 0.05$. The estimates of the performance index using the proposed method are plotted in Fig. 4.

6. Discussion

From the simulations for Wiener and HW models, we observe that the estimates of the performance index using linear CPA techniques (using a linear ARMA model to fit the process output) have significant biases. The new performance indices converge to the theoretical performance indices as time t increases. Our proposed method can provide more reliable information for control performance assessment.

An obvious question regarding the application of our proposed methods is when is it necessary to apply nonlinear, as opposed to the simpler linear, CPA strategies. The answer is simple: one just needs to check the nonlinearity of the output perhaps using the Hinich test proposed by [52] and discussed further in [29,53]. In a related application, these tests have been used for diagnosing the valve stiction in [54] and estimating MVPLB for the valve stiction problem [8].

7. Conclusions

The contribution of this work is to propose a new performance index for the general nonlinear systems. This new performance index is not based on the MVPLB which is difficult to estimate for general nonlinear systems. It is bounded in the region [0, 1], so one can use it in the same manner as the traditional performance index based on the MVPLB.

The proposed strategy is valid for almost any nonlinearity because the ANOVA approach is model independent. However the system identification portion is more problematic for severe nonlinear systems which consequently effects the quality of the analysis. We chose NARMA illustrative examples for this paper simply because it is convenient for simulation.

This algorithm requires only observable signals and crude estimates of the plant dominant time constants and plant delay. The proposed method does not require one to identify the process (linear dynamic and static nonlinearity) and disturbance structure, but one does need to estimate the closed-loop nonlinear model. The simulation results for two classes of nonlinear models show that our approach can provide reliable estimates for CPA of nonlinear systems in contrast to simply ignoring the nonlinearity.

The identification of the closed-loop model will directly affect the estimates of the new performance index. Due to the complexity of nonlinear systems, our method cannot provide a guaranteed CPA solution. For some specific nonlinear systems, we may need more information instead of only output data and plant delay.

Acknowledgments

Financial support for this project from the Industrial Information and Control Centre, Faculty of Engineering, The University of Auckland, New Zealand is gratefully acknowledged.

References

[1] A. Horch, A simple method for detection of stiction in control valves, Control Engineering Practice 7 (1999) 1221–1231.
[2] M.A.A.S. Choudhury, S.L. Shah, N.F. Thornhill, D.S. Shook, Automatic detection and quantification of stiction in control valves, Control Engineering Practice 14 (2006) 1395–1412.
[3] N.F. Thornhill, A. Horch, Advances and new directions in plant-wide disturbance detection and diagnosis, Control Engineering Practice 15 (2007) 1196–1206.
[4] T.J. Harris, W. Yu, Controller assessment for a class of nonlinear systems, Journal of Process Control 17 (2007) 607–619.
[5] Y.F. Zhou, F. Wan, A newral netwark approach to control performance assessment, International Journal of Intelligent Computing and Cybernetics 1 (4) (2008) 1617–1633.
[6] W. Yu, D.I. Wilson, B.R. Young, Control performance assessment in the presence of valve stiction, in: K. Grigoriadis (Ed.), The Eleventh IASTED International Conference on Intelligent Systems and Control, ISC 2008, Orlando, FL, USA, November 16–18, 2008, pp. 379–384.
[7] W. Yu, D.I. Wilson, B.R. Young, Eliminating valve stiction nonlinearities for control performance assessment, in: International Symposium on Advanced Control of Chemical Processes, ADCHEM 2009, Istanbul, Turkey, July 12–15, 2009, pp. 526–531, International Federation of Automatic Control.
[8] W. Yu, D.I. Wilson, B.R. Young, Nonlinear control performance assessment in the presence of valve stiction, Journal of Process Control 20 (6) (2010) 754–761.
[9] W.L. Bialkowski, Dreams versus reality: a view from both sides of the gap, Pulp & Paper Canada 94 (11) (1998) 19.
[10] L. Desborough, R. Miller, Increasing customer value of industrial control performance monitoring: Honeywell's experience, in: AIChE Symposium Series, volume 98, 2002, pp. 153–186.
[11] S. Joe Qin, Control performance monitoring —a review and assessment, Computers in Chemical Engineering 23 (2) (1998) 173–186.
[12] T.J. Harris, A review of performance monitoring and assessment techniques for univariate and multivariate control systems, Journal of Process Control 9 (1) (1999) 1–17.
[13] B. Huang, S.L. Shah, Performance Assessment of Control Loops: Theory and Applications, Springer, 1999.
[14] M. Jelali, An overview of control performance assessment technology and industrial applications, Control Engineering Practice 14 (5) (2006) 441–466.
[15] T.J. Harris, W. Yu, Variance decompositions of nonlinear dynamic stochastic systems, Journal of Process Control 20 (2) (2010) 195–205.
[16] F.J. Doyle, R.K. Pearson, B.A. Ogunnaike III, Identification and Control Using Volterra Models, Springer, London, 2002.
[17] H. Diaz, A.A. Desrochers, Modeling of nonlinear discrete-time systems from input–output data, Automatica 5 (1988) 629–641.
[18] R. Haber, H. Unbehauen, Structural identification of nonlinear dynamic systems—a survey on input/output approaches, Automatica 26 (4) (1990) 651–677.
[19] I.J. Leontaritis, S.A. Billings, Input–output parametric models for nonlinear systems, part I: deterministic nonlinear systems, International Journal of Control 41 (4) (1985) 303–328.
[20] I.J. Leontaritis, S.A. Billings, Input–output parametric models for nonlinear systems, part II: stochastic nonlinear system, International Journal of Control 41 (4) (1985) 329–344.
[21] R.K. Pearson, B.A. Ogunnaike, Nonlinear process identification, in: M.A. Henson, D.E. Seborg (Eds.), Nonlinear Process Control, Prentice Hall, Upper Saddle River, NJ, 1997, pp. 11–102.
[22] R.K. Pearson, Discrete-time Dynamic Models, Oxford University Press, New York, 1999.
[23] K.R. Sales, S.A. Billings, Self-tuning control of non-linear ARMAX models, International Journal of Control 51 (1990) 753–769.
[24] T.J. Harris, Assessment of control loop performance, Canadian Journal of Chemical Engineering 67 (1989) 856–861.
[25] N. Stanfelj, T.E. Marlin, J.F. MacGregor, Monitoring and diagnosing process control performance: the single-loop case, Industrial & Engineering Chemistry Research 32 (1993) 301–314.
[26] K.J. Åström, Introduction to Stochastic Control Theory, Academic Press, New York, 1970.
[27] G.E.P. Box, G.M. Jenkins, Time Series Analysis Forecasting and Control, Holden-Day, San Francisco, 1970.
[28] L. Ljung, System Identification: Theory for the User, Prentice-Hall, New York, 1987.

[29] H. Tong, Non-linear Time Series, Oxford University Press, New York, 1990.
[30] G.E.B. Archer, A. Saltelli, I.M. Sobol', Sensitivity measures, Anova-like techniques and the use of bootstrap, Journal of Statistical Computations and Simulations 58 (1997) 99–120.
[31] D.C. Cox, An analytical method for uncertainty analysis of nonlinear output functions, with applications to fault-tree analysis, IEEE Transactions on Reliability R-31 (5) (1982) 465–468.
[32] E. Parzen, Stochastic Processes, Holden Day, San Francisco, 1962.
[33] W. Yu, T.J. Harris, Parameter uncrertainty effects on variance-based sensitivity analysis, Reliability Engineering and System Safety 94 (2) (2009) 596–603.
[34] S. Chen, S.A. Billings, W. Luo, Orthogonal least squares methods and their application to nonlinear system identification, International Journal of Control 50 (5) (1989) 1873–1896.
[35] K.H. Chon, M.J. Korenberg, N.H. Holstein-Rathlou, Application of fast orthogonal search to linear and nonlinear stochastic systems, Annals of Biomedical Engineering 25 (1997) 793–801.
[36] M.J. Korenberg, Identifying nonlinear difference equation and functional expansion representations: the fast orthogonal algorithm, Annals of Biomedical Engineering 16 (1988) 123–142.
[37] T. Terasvirta, D.V. Digk, M.C. Medeiros, Linear models, smooth transition autoregressions, and neural networks for forecasting macroceconomic time series: a re-examination, International Journal of Forecasting 21 (2005) 755–774.
[38] S. Chen, S.A. Billings, Neural networks for nonlinear dynamic system modeling and identification, International Journal of Control 56 (2) (1992) 319–346.
[39] I.J. Leontaritis, S.A. Billings, Model selection and validation methods for nonlinear systems, International Journal of Control 45 (1987) 311–341.
[40] H. Rabitz, O.F. Alis, J. Shorter, K. Shim, Efficient input–output model representations, Computer Physics Communications 117 (1999) 11–20.
[41] R.I. Cukier, C.M. Fortuin, K.E. Shuler, A.G. Petschek, J.H. Schaibly, Study of the sensitivity of coupled reaction system to uncertainties in rate coefficients. I theory, Journal of Chemical Physics 59 (8) (1973) 3873–3878.
[42] A. Saltelli, S. Tarantola, K. Chan, A quantitative model-independent method of global sensitivity analysis of model output, Technometrics 41 (1) (1999) 39–56.
[43] I.M. Sobol', Sensitivity estimates for nonlinear mathematical models, Matematicheskoe Modelirovanie 2 (1993) 112–118 (in Russian), translated in Mathematical Modeling and Computational Experiments, 1, 407–414.
[44] T. Homma, A. Saltelli, Importance measures in global sensitivity analysis of nonlinear models, Reliability Engineering and System Safety 52 (1996) 1–17.
[45] I.M. Sobol', Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates, Mathematics and Computer in Simulation 55 (2001) 271–280.
[46] J.C. Gomez, E. Baeyens, Identification of block-oriented nonlinear systems using orthonormal bases, Journal of Process Control 14 (2004) 685–697.
[47] Q. Zheng, E. Zafiriou, Volterra–Laguerre models for nonlinear process identification with application to a fluid cracker unit, Industrial and Engineering Chemistry Research 43 (2004) 340–348.
[48] G.R. Averin, The Hammerstein–Wiener Model for Identification of Stochastic Systems, Automation and Remote Control 64 (9) (2003) 1418–1431.
[49] S.W. Sung, C.H. Je, J. Lee, D.H. Lee, Improved system identification method for Hammerstein–Wiener processes, Korean Journal of Chemical Engineering 25 (2008) 631–636.
[50] M.J. Grimble, Non-linear generalized minimum variance feedback, feedforward and tracking control, Automatica 41 (2005) 957–969.
[51] L. Ljung, Q. Zhang, P. Lindskog, A. Louditski, R. Singh, An integrated system identification toolbox for linear and non-linear models, in: 14th IFAC Symposium on System Identification, Newcastle, Australia, June 14, 2007, pp. 379–384.
[52] M.J. Hinich, Testing for Gaussianity and linearity of a stationary time series, Journal of Time Series Analysis 3 (13) (1982) 169–176.
[53] R. Haber, L. Keviczky, Nonlinear System Identification—Input–Output Modeling approach, Kluwer Academic Publishers, Dordrecht, The Netherland, 1999.
[54] M.A.A.S. Choudhury, S.L. Shah, N.F. Thornhill, Diagnosis of poor control-loop performance using higher-order statistics, Automatica 40 (10) (2004) 1719–1728.