A Comparison of Nonlinear Control Performance Assessment Techniques for Hammerstein-Wiener Processes

Wei Yu,[*] David Wilson[†] and Brent Young[‡]

Abstract

Assessing the quality of industrial control loops is an important routine auditing task for the control engineer. However there are complications when considering nonlinear control loops, where one must consider both the type of nonlinearity, and the precise structure of the loop. This paper shows that certain nonlinear CPA strategies fail when faced with disturbance signals that are immediately passed through a nonlinearity, whereas others, admittedly more computationally demanding, are immune to this noise structure.

1. INTRODUCTION

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. Linear CPA is routinely applied in the refining, petrochemicals, pulp and paper and the mineral processing industries as noted by [1, 2, 3].

In practice though, industrial control loops invariably include nonlinearites due to the control valve, the measurement transducer, or the process itself. 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 [4] shows one tends to overestimate the performance index for linear systems with an additive linear disturbance affected by valve stiction, when using linear CPA techniques. To deal with this situation, recent research has proposed several methods to extend CPA into nonlinear systems [5].

This paper compares three such techniques to quantify a controller performance index for a class of nonlinear systems. The first method is the standard linear approach, the second is a semi-parametric method which was originally proposed to deal with linear systems in presence of valve stiction [5] and more general block-oriented nonlinear systems [6]. The third method is based on the technique of ANOVA-like variance decomposition [7]. In this paper, these CPA techniques are applied to linear dynamic systems with static nonlinearities.

The layout of the paper is as follows. Section 2 introduces the Hammerstein-Wiener model used for the comparison study. Section 3 briefly summarizes the three techniques and in section 4, a simulation example is used to quantify the comparison. This is followed by a discussion and conclusions highlighting both the limitations and potential of the proposed methods.

2. Process description

Many industrial plants can be adequately modelled by a Hammerstein-Wiener (HW) model as shown in Fig. 1 with,

$$y_t = N_2(x_t) + \text{noise} \tag{1}$$

$$x_t = w_t + \text{noise} \tag{2}$$

$$w_t = \frac{B(q^{-1})}{A(q^{-1})} q^{-b} v_t, \text{ with } v_t = N_1(u_t) \tag{3}$$

where $A(q^{-1})$ and $B(q^{-1})$ are polynomials in the backward shift operator q^{-1}, b is the time delay of the system, 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 unmeasured. The functions N_1 and N_2 represent static nonlinearities. In a process control loop, the HW model structure can be justified by considering the input nonlinear block N_1 to represent nonlinearities such as equal percentage valve characteristics or quantisation due to pulse-width modulated controllers, while the out-

[*]W. Yu is a postdoctoral fellow with the Industrial Information & Control Centre, Department of Chemical & Materials Engineering, The University of Auckland, New Zealand wyu048@aucklanduni.ac.nz

[†]D.I. Wilson is an associate professor with the Department of Electrical & Electronic Engineering, Auckland University of Technology, New Zealand david.i.wilson@aut.ac.nz

[‡]B.R. Young is an associate professor with the Department of Chemical & Materials Engineering, The University of Auckland, New Zealand b.young@auckland.ac.nz

put nonlinear block N_2 could be modelling a thermocouple or thyristor transducer calibration curves.

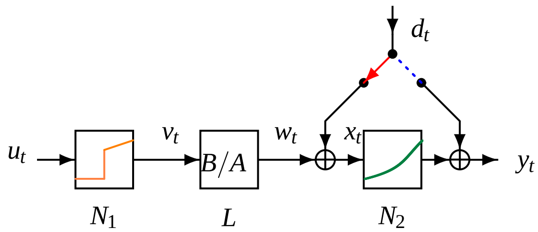

Figure 1. The Hammerstein-Wiener model for this paper showing the two possible noise signal entry points considered in this paper.

The disturbance, d_t, is modeled as the output of a linear Autoregressive-Integrated-Moving-Average (ARIMA) filter driven by white noise a_t of zero mean and variance σ_a^2 of the form,

$$d_t = \frac{\theta(q^{-1})}{\phi(q^{-1})\nabla^h} a_t \tag{4}$$

where $\nabla \stackrel{\text{def}}{=} (1-q^{-1})$ is the difference operator and h is a non-negative integer, typically less than 2. The polynomials $\theta(q^{-1})$ and $\phi(q^{-1})$ are monic and stable.

As indicated in Fig. 1, we consider two different locations where this additive disturbance enters the system; the first location is immediately upstream of the static nonlinearity N_2, and the second is downstream of the static nonlinearity N_2.

3. Techniques of estimating the performance index for nonlinear systems

3.1. Linear CPA technique

The performance index proposed for SISO linear systems in [8] was based on the concept of minimum variance control [9]. This performance index, often termed the Harris Index, η, was defined as the ratio of the best achievable variance to the variance of the controlled variable under assessment,

$$\eta = \frac{\hat{\sigma}_{MV}^2}{\sigma_y^2}. \tag{5}$$

What made this concept useful was the insight of [8] which showed that η can be estimated directly from routine operating data by fitting the controlled variable to an ARIMA time series model.

3.2. A semi-parametric method of CPA for the nonlinear system

A semi-parametric CPA method for nonlinear systems was proposed to find the MVPLB for linear systems with valve stiction problems in [4, 5]. The approach was first to remove the nonlinearity in the observable time series y using smoothing B-splines where the degree of smoothing (tolerance, τ) is adjusted iteratively to just make the resultant series linear. It includes two parts: i) A non-parametric B-spline is used to fit the nonlinearity from the output and ii) a linear ARMA model is used to fit the residuals between the output and B-spline. Consequently the minimum variance performance bounds can be estimated given the regressed ARMA model to the now linear data.

The extent of smoothing to just remove the nonlinearity is established by checking the Gaussianity and linearity of the residuals using the nonparametric Hinich test, [10]. The application of this test on the diagnosis and detection of valve stiction was reported in [11].

3.3. ANOVA-based Performance Indices

The MVLPB-based performance indices such as described in section 3.2 exhibit significant theoretical difficulties when applied to general nonlinearities such as NARMAX models. Quite apart from the fact that the MVPLB may be difficult to estimate or even not exist, the idea of variance decomposition using the impulse response function and the concept of feedback invariance is not necessarily valid for nonlinear systems.

For the HW systems shown in Fig. 1 with a feedback controller, $u_t = g(y_t, \cdots, y_{n_t})$, the closed-loop outputs can be often approximated using a NARMA model form,

$$y_t = f(y_{t-1}, \cdots, y_{t-n_y}, a_t, \cdots, a_{t-n_t}) \tag{6}$$

The ANOVA-like variance decomposition method was used to provide the variance analysis for nonlinear systems with multi-disturbance sources [12]. However instead of the different disturbance sources, the ANOVA-like variance decomposition will be used with differing time horizons for the SISO nonlinear system with one disturbance as developed in [7].

We can take an analogous approach where we separate the disturbances entering the system in (6) after some nominal start time 0, say $[a_{t+b}, a_{t+b-1}, \cdots, a_1]$, into two groups. The first group includes all the disturbances entering the system *after* the current time t, namely $A_1 = [a_{t+b}, \cdots, a_{t+1}]$ and the remaining group which is all the disturbances entering the system from $t = 0$ to the current time t, namely $A_2 = [a_t, a_{t-1}, \cdots, a_1]$.

To decompose the variance of the nonlinear process using the ANOVA-like decomposition method, we use the following assumptions:

- Eqn. 6 can be computed subject to the initial condition, I_0, to give $y_{t+b-i}, i = 1, 2, \cdots, t+b-1$ for any choice of $a_{t+b}, a_{t+b-1}, \cdots, a_1$.
- 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 Eqn. 6 to variations of the two vector series A_1 and A_2. Since the future behavior of y_{t+b} is dependent on initial conditions due to the nonlinearity, we can use the well-known variance decomposition theorem [13] to deal with the initial conditions. The variance of y_{t+b} can be decomposed into two terms:

$$V[y_{t+b}] = E_{I_0}[V_A[y_{t+b}|I_0]] + V_{I_0}[E_A[Y_{t+b}|I_0]] \tag{7}$$

where $A \stackrel{\text{def}}{=} [A_1, A_2]$ denotes all 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 (7) is the sum of positive numbers, it follows that $V[y_{t+b}] \geq E_{I_0}[V_A[y_{t+b}|I_0]]$.

The first term in (7) is the fractional contribution to the variance of y from the disturbance signal and the interaction between the 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, $V_A[y_{t+b}|I_0]$, can be decomposed directly using the ANOVA-like decomposition method as:

$$V_A|I_0 = V_A[y_{t+b}|I_0] = V_1|I_0 + V_2|I_0 + V_{12}|I_0 \tag{8}$$

where

$$\begin{aligned} V_1|I_0 &= V_{A_1}\left[E_{A_2}\left[y_{t+b}|(A_1, I_0)\right]\right] \\ V_2|I_0 &= V_{A_2}\left[E_{A_1}\left[y_{t+b}|(A_2, I_0)\right]\right] \\ V_{12}|I_0 &= V_A|I_0 - V_1|I_0 - V_2|I_0 \end{aligned} \tag{9}$$

The variance decomposition with consideration of the initial condition can be obtained by simply calculating the expectation of the conditional variance decomposition in (8) with respect to the initial condition I_0. This procedure is not necessary if the initial condition has, (or 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 (8) based on the mean values of initial condition. Further information about this topic can be found in [14].

$E_{I_0}[V_1|I_0]$ denotes the main effect of A_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 in the main effects of A_1 and A_2. Consequently we propose a suitable performance index as

$$\eta_t = \frac{E_{I_0}[V_1|I_0]}{V[y_{t+b}]} \tag{10}$$

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 (7) dominates. However such cases are more pathological than common in industry.

Remarks:

- This performance index defined in (10) 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 [15].
- This performance index is strictly bounded in $[0, 1]$. If η reaches 1, it means that the variance of the outputs is contributed mostly by A_1, so the system controller is close to the minimum variance controller.

3.3.1. Computing the Performance Index. The practical computation of the performance index in (10) requires two steps. First one must estimate the closed loop model defined in (6), and then use a Monte-Carlo strategy to approximate the performance index.

For the first identification step, several strategies have been proposed such as orthogonal Least Squares (OLS) methods [16], Fast Orthogonal Search (FOS) methods [17], and approximations based on Artificial Neural Network (ANN) models [18].

Since the disturbance term in (6) 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), [16],

$$AIC(\lambda) = K \ln \hat{\sigma}_{\Xi}^2 + N\lambda \tag{11}$$

where N is the number of the model parameters, K is the number of outputs and $\hat{\sigma}_{\Xi}^2$ 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 [16].

Once the parameters of the closed loop are estimated, a more practical way is to use a Monte Carlo (MC) method to estimate the performance index. Although the computationally demanding due to the potentially high dimensionality, it can be offset somewhat by employing alternative strategies such as the Fourier Amplitude Sensitivity Test (FAST) [19] and Sobol's method [20].

For a model of the form given by (6), the MC estimates of the mean and variance of y_{t+b} given the initial condition I_0 can be calculated by,

$$\hat{\bar{y}}_{t+b}|I_0 \simeq \frac{1}{N}\sum_{k=1}^{N} f(A_1^{(k)}, A_2^{(k)})|I_0 \tag{12}$$

$$\hat{V}_A|I_0 \simeq \frac{1}{N}\sum_{k=1}^{N} (f(A^{(k)})|I_0)^2 - (\hat{\bar{y}}_{t+b}|I_0)^2 \tag{13}$$

where $A^{(k)} = (A_1^{(k)}, A_2^{(k)})$ is a set of N simulations of multidimensional inputs that have the requisite distribution. The partial variance $V_1|I_0$ can be estimated as [21],

$$\hat{V}_1|I_0 = \frac{1}{N}\sum_{k=1}^{N} f(A_1^{(k)}, A_2^{(k)}) f(A_1^{(k)}, \vec{A}_2^{(k)}) - (\hat{\bar{y}}_{t+b}|I_0)^2 \tag{14}$$

$\vec{A}_2^{(k)}$ is independent of the set of $A_2^{(k)}$. Hence (14) means that for computing $V_1|I_0$ we multiply values of f corresponding to a set $(A_1^{(k)}, A_2^{(k)})$ by values of f from a different set $(A_1^{(k)}, \vec{A}_2^{(k)})$ which includes the same partial set $(A_1^{(k)})$. 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 (10). 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.

4. Simulation comparisons

To highlight the different characteristics of the three CPA strategies, we will consider the four HW model structures given in Fig. 2. The models are similar except for the position of the noise input, d_t, and the order of the nonlinear elements N_1 and N_2.

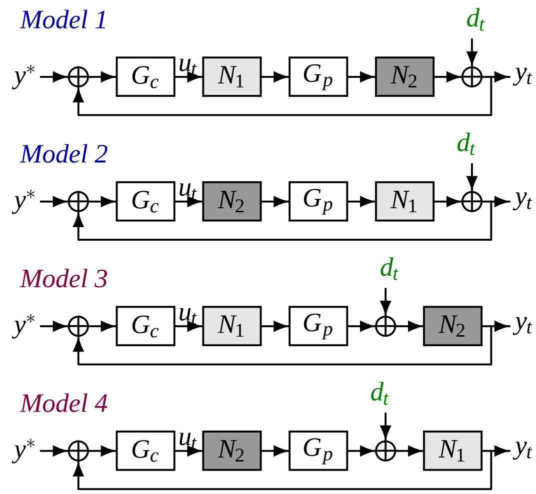

Figure 2. Four different HW model structures

In all cases, the linear plant is

$$G_p = \frac{q^{-b}B(q^{-1})}{A(q^{-1})} = \frac{q^{-3}(1-0.5q^{-1})}{1-1.5q^{-1}+0.7q^{-2}} \tag{15}$$

under feedback control with a PI controller,

$$G_c = \frac{0.05-0.02q^{-1}}{1-q^{-1}} \tag{16}$$

perhaps representing a standard process control type application, while the nonlinear element N_1, is a coulombic and viscous friction nonlinearity, $v_t(u_t) = \mathrm{sgn}(u_t)(|u_t| + 0.07)$, where 0.07 is the offset, and N_2, is a third order polynomial, $y_t(x_t) = x_t + 0.25x_t^2 + 0.125x_t^3$.

This plant is subjected to an additive disturbance of

$$d_t = \frac{a_t}{1-0.8q^{-1}} \tag{17}$$

where a_t is a sequence of independent and identically distributed Gaussian random variables with zero mean and nominal variance $\sigma_{a_t}^2 = 0.2$.

Models 1 and 3 where the upstream nonlinearity is non-differentiable represents the common situation where the manipulated variable such as a control valve suffers from stiction, while the trailing smooth nonlinearity represents the common case where the measuring transducer includes a calibration curve.

In models 3 and 4, the key point is that the disturbance d_t is placed *before* the output static nonlinearity block as opposed to the usual assumption where the disturbance is assumed additive to the output after the nonlinearity as in models 1 and 2. In some instances this early entrance of the disturbance signal is more realistic from a process operation point of view given that it is the transducer that often provides the nonlinear behaviour as discussed in [22].

Using the methods proposed in [15, 5], the MVPLB of the model 1 and 2 can be determined as 0.41. As the MVPLBs of the model 3 and 4 do not exist, an alternative performance lower bound is used as described in [23]. The estimates of the performance index using the three proposed strategies for the four different model structures can be compared in Fig. 3.

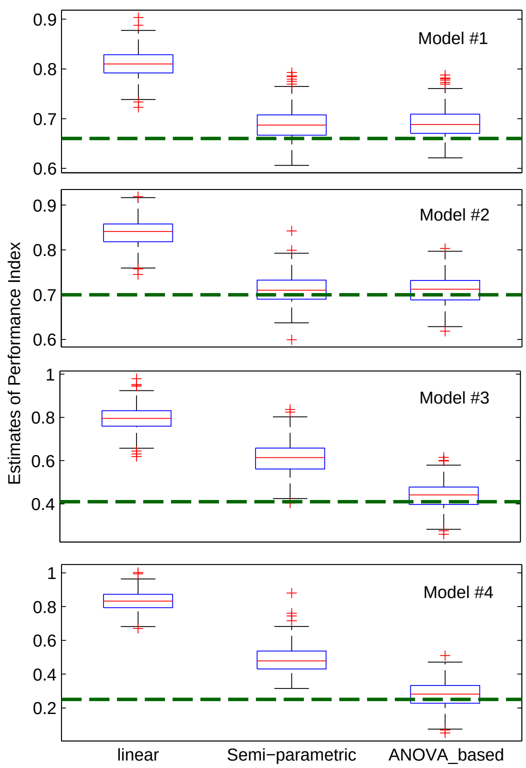

Figure 3. Comparative box plots of the quality estimates of the performance index for the four model structures given in Fig. 2.

From the simulation of these four models, we can observe that:

- The estimates of the performance index using linear CPA techniques (using a linear ARMA model to fit the process output) have the significant bias.
- The semi-parametric method provides the reliable estimates for models 1 and 2. However, the estimates for models 3 and 4 are not satisfactory.
- The ANOVA-like method provides reliable estimates of the performance index for all four models.

Although the results from the ANOVA-based method are preferable for control engineers, the drawback of this method is that it requires very intensive computations as indicated by the computation time given in the last line of Table 1. Consequently, the faster computation of the semi-parametric method may more than compensate for the small bias in the estimates of the performance index.

Table 1. Summarising the ability of the three strategies to estimate the performance index (PI) for HW models of different structures.

Model	*Techniques to estimate the PI*		
	Linear	Semi-parametric	ANOVA-based
1	×	✓	✓
2	×	✓	✓
3	×	×	✓
4	×	×	✓
time (min)	4	6	50

5. CONCLUSIONS

This paper compared three techniques for estimating the performance index for Hammerstein-Weiner nonlinear systems: a linear technique, a semi-parametric method and an ANOVA-based method. Not unexpectedly the linear CPA strategy delivered results with a significant bias leading to an overoptimistic assessment of the controllers performance. The semi-parametric method gave satisfactory results for the case when the disturbance was added to the output, but failed in the case where the disturbance entered the system upstream of the tailing nonlinearity. Only the ANOVA-based strategy could deliver reliable results in all cases, but that came at a cost of substantial computation compared to the alternatives.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support to this project from the Industrial Information and Control Centre (I^2C^2), Faculty of Engineering, The University of Auckland, New Zealand.

References

[1] T.J. Harris. A review of performance monitoring and assessment techniques for univariate and multivariate control systems. *J. Process Control*, 9(1):1–17, 1999.

[2] B. Huang and S.L. Shah. *Performance Assessment of Control Loops: Theory and Applications*. Springer, 1999.

[3] M. Jelali. An overview of control performance assessment technology and industrial applications. *Control Engineering Practice*, 14(5):441–466, 2006.

[4] Wei Yu, David I. Wilson, and Brent R. Young. Control performance assessment in the presence of valve stiction. In K. Grigoriadis, editor, *The Eleventh IASTED International Conference on Intelligent Systems and Control, ISC 2008*, pages 379–384, Orlando, Florida, USA, 16–18 November 2008.

[5] Wei Yu, David I. Wilson, and Brent R. Young. Nonlinear Control Performance Assessment in the Presence of Valve Stiction. *J. Process Control*, 20(6):754–761, July 2010.

[6] Wei Yu, David I. Wilson, and Brent R. Young. Control Performance Assessment for Block-oriented Nonlinear Systems. In *8th IEEE International Conference on Control & Automation ICCA'10*, pages 1151–1156, Xiamen, China, 9–11 June 2010.

[7] Wei Yu, David I. Wilson, and Brent R. Young. Control Performance Assessment for Nonlinear Systems. *J. Process Control*, February 2010. *Submitted*.

[8] T.J. Harris. Assessment of control loop performance. *Canadian Journal of Chemical Engineering*, 67:856–861, 1989.

[9] K. J. Åström. *Introduction to Stochastic Control Theory*. Academic Press, New York, 1970.

[10] M.J. Hinich. Testing for Gaussianity and linearity of a stationary time series. *Journal of Time Series Analysis*, 3(13):169–176, 1982.

[11] M.A.A.S. Choudhury, S.L. Shah, N.F. Thornhill, and D.S. Shook. Automatic detection and quantification of stiction in control valves. *Control Engineering Practice*, 14:1395–1412, 2006.

[12] T.J. Harris and W. Yu. Variance decompositions of nonlinear dynamic stochastic systems. *Journal of Process Control*, 20(2):195–205, 2010.

[13] E. Parzen. *Stochastic Processes*. Holden Day, San Francisco, 1962.

[14] W. Yu and T.J. Harris. Parameter uncrertainty effects on variance-based sensitivity analysis. *Reliability Engineering and System Safety*, 94(2):596–603, 2009.

[15] T.J. Harris and W. Yu. Controller assessment for a class of nonlinear systems. *J. Process Control*, 17:607–619, 2007.

[16] S. Chen, S. A. Billings, and W. Luo. Orthogonal least squares methods and their application to nonlinear system identification. *International Journal of Control*, 50(5):1873–1896, 1989.

[17] M. J. Korenberg. Identifying nonlinear difference equation and functional expansion representations: The fast orthogonal algorithm. *Annals of Biomedical Engineering*, 16:123–142, 1988.

[18] S. Chen and S. A. Billings. Neural networks for nonlinear dynamic system modeling and identification. *International Journal of Control*, 56(2):319–346, 1992.

[19] A. Saltelli, S. Tarantola, and K. Chan. A quantitative model-independent method of global sensitivity analysis of model output. *Technometrics*, 41(1):39–56, 1999.

[20] I. M. Sobol'. Sensitivity estimates for nonlinear mathematical models. Matematicheskoe Modelirovanie, 2, 112-118 (in Russian), translated in Mathematical Modeling and Computational Experiments, 1, 407-414, 1993.

[21] I. M. Sobol'. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computer in Simulation*, 55:271–280, 2001.

[22] A. Hagenblad, L. Ljung, and A. Wills. Maximum likelihood identification of wiener models. *Automatica*, 44:2697–2705, 2008.

[23] Wei Yu, David I. Wilson, Thomas Harris, and Brent R. Young. Control Performance Assessment for Block-Oriented Nonlinear Systems. In *9th International Symposium on Dynamics and Control of Process Systems (DYCOPS)*, pages 665–670, Leuven, Belgium, 5–7 July 2010.