Process Identification Using Discrete Wavelet Transforms: Design of Prefilters

Srinivas Palavajjhala, Rodolphe L. Motard, and Babu Joseph
Process Control and Systems Laboratory, Dept. of Chemical Engineering, Washington University, St. Louis, MO 63130

Most advanced control applications rely on good dynamic process models. The performance of the control system depends on the accuracy of the model used. Typically, such models are developed by conducting off-line identification experiments on the process. These identification experiments often result in input–output data with small output signal-to-noise ratio, and using these data results in inaccurate model parameter estimates. Prefilters are used to separate useful information from the noise in the input–output data and to improve parameter estimates. A systematic design procedure for selecting a prefilter using discrete wavelet transforms is presented. The design procedure provides explicit information on the compromises in prefilter design, interpreted in terms of parameter variance and bias. The prefilter design procedure is then applied to identify a second-order output error model.

Introduction

Linear time-invariant (LTI) dynamic models of chemical processes are widely used in the chemical process industry for process control and optimization. Some applications where such models are used include: (1) linear model predictive control, (2) inferential control, (3) proportional-integral-derivative (PID) controller tuning using internal model control (IMC) rules (Rivera et al., 1986), and (4) linear programming optimization for bias update (Forbes and Marlin, 1994; Harkins, 1991; Brosilow and Zhao, 1988). The stability and the performance of the control system designed using the LTI models depends to a large extent on the accuracy of the model parameters. This in turn dictates the design of identification experiments.

Designing an identification experiment involves deciding on the following: (1) the type of input signal, (2) plant frequencies excited by the input signal, (3) the sampling rate for the input and the output signal, (4) the amplitude of the input signal, (5) the test duration, and (6) the filters used on the input–output data. Some guidelines on making these decisions to yield an informative identification experiment are presented in this article. The relationships between pseudo-random binary sequence (PRBS) design parameters and the process knowledge are used to design identification experiments. In most process-control-related identification, the objective is to minimize the modeling error in the frequency band over which the controller will operate. Prefilters are used to improve parameter estimates in the frequency band of interest.

Rivera et al. (1992) introduce the notion of *control-related process identification* in which the information content of an identification experiment is defined in term of its impact on the control system design. They develop algorithms to design prefilters in terms of the closed-loop time constant, the assumed plant model structure and the desired controller performance characteristics. They also illustrate how a prefilter can be designed to preserve good model fit in the frequency range of interest from a control perspective. Rivera et al. assume that the output data generated using a white noise test signal has a high output signal-to-noise ratio (hereafter referred as output S/N ratio). This assumption makes it possible to relate the closed-loop properties of the controller and the prefilter design.

Identification experiments conducted in practice often result in data with small output S/N ratio, the reason for which can be attributed to:

1. Identification test is usually conducted while the plant is in normal operation. It is desirable that the external signal introduced for identification does not significantly change the

Correspondence concerning this article should be addressed to B. Joseph.
Present address of S. Palavajjhala: Dynamics Matrix Control Corporation, Houston, TX 77036.

controlled variable values from the normal operating conditions (for economic or safety reasons). This restricts the magnitude of the output signal.

2. Lack of process knowledge along with the lack of on-line validation tools.

3. In dynamic matrix identification (Cutler and Yocum, 1991), differencing is used to eliminate unmeasured disturbances. This may result in input–output data with small S/N ratio.

Identifiation experiments are expensive to conduct and are time-consuming. It is therefore desirable to extract the maximum possible process knowledge from the data collected during an identification experiment. Prefilters are used to achieve this objective. Prefilters are usually low-pass filters used to reduce the contribution of noise, while retaining the useful information, in the input–output data collected during an identification experiment. In this article, we develop some tools to design prefilters that yield better parameter estimates.

The output S/N ratio influences the prefilter design. If the output S/N ratio is high (> 5), then the contribution of the noise in the measurements is small and a prefilter does not significantly improve the parameter estimates. However, better parameter estimates are obtained using prefilters when the S/N ratio is small (< 2). This article provides a systematic procedure for selecting a prefilter that improves the output S/N ratio.

In this article we present a prefilter design methodology using discrete wavelet transforms. The input–output time-series data collected from the identification experiment is projected onto a family of scaling functions to yield a hierarchical decomposition called *multiresolution analysis*. In multiresolution analysis, the input–output data are decomposed into various components (based on frequency) using a *deck of sieves*. The fine-scale (high-frequency) information is retained on the top sieves, and the coarse-scale (low-frequency) information is decomposed as we move to the bottom of the deck. The sieved input–output data can then be used to obtain the cumulative sum (also referred to as the projections on the scaling function) starting from the bottom of the deck. The cumulative sum at the top of the deck is the raw signal. Eliminating fine-scale (high-frequency) information from the decomposition is equivalent to light filtering, while eliminating coarser scales yield heavier filtering. The sieved input–output data facilitate the selection of a prefilter that yields the maximum output S/N ratio. An efficient computing algorithm using orthonormal wavelets is used to decompose the data. The property that the wavelet transform of Gaussian white noise with zero mean and variance one is again white noise with variance no larger than 1, makes wavelets a suitable choice to study the properties of the estimator as compared to the traditional Butterworth low-pass filters.

This article is organized as follows: In the second section, statistical properties of estimates obtained using the prediction error method (Ljung, 1987) are summarized, and the methods to reduce the parameter variance and bias using prefilters are described. Properties of wavelet transform of PRBS, white noise, and output signal from an autoregressive moving average model (ARMA) process are important for understanding and interpreting the results obtained while designing prefilters. The third section presents a method to estimate the variance of the projections of the PRBS, white noise, and output signal from an ARMA process on the scaling functions. The estimates are compared with actual values using examples. The fourth section describes the prefilter design methodology. In the fifth section, the prefilter design is applied to identify a second-order output error model in the Tennessee-Eastman problem (Downs and Vogel, 1993). We conclude this article with some remarks on prefilter design based on the examples we studied.

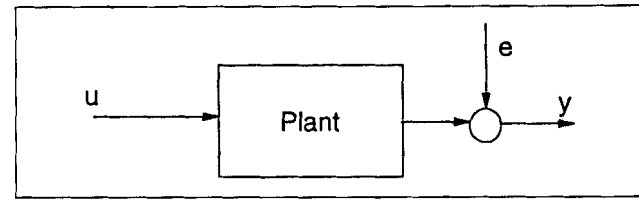

Figure 1. Black-box input–output modeling.

Prediction Error Problem

Consider the plant to be a black-box with input (manipulated variable) u, unmeasured disturbance e, and the output (measurement) y (see Figure 1). The objective in process identification is to determine a dynamic model of the process by performing an identification experiment. Here we assume that (1) the identification experiment is conducted off-line; (2) the plant is linear, time-invariant, causal, and stable; (3) the plant can be modeled using the output error (OE) model (though the design methodology developed is not limited to this assumption); and (4) the test signal used in the identification experiment is a PRBS.

PRBS test signal design

In this article we generate the PRBS signal using shift registers. The PRBS generator consists of n shift registers and one XOR gate in the feedback circuit (see Davies, 1970, for details). The inputs to the XOR gate are the output of registers k and n, and the output from the exclusive-OR (XOR) gate is fed back as input to register 1. The output of register n is the PRBS. Each register is connected to a shift pulse or a clock. The output from each shift register, which is either 0 or 1, is moved one stage to the right at each clock time.

The properties of PRBS generated using shift registers is a function of (1) the feedback location k, (2) the number of registers n, and (3) the clock time T_c. Therefore, the PRBS design parameters are:

n or T = number of registers or PRBS length $T = 2^n - 1$
a = amplitude of PRBS
T_c = clock time
k = register whose output is an input to the XOR gate
N = test duration

Rivera (1992) gives the criteria for selecting n and T_c in terms of the dominant closed-loop identification. We use a similar approach to design a PRBS signal for off-line identification.

Choice of T_c and T. At frequency $\omega = 0$, the PRBS frequency spectrum has a value equal to $a^2(T+1)/T$ (see Davies, 1970). The spectrum is reduced to half this value at frequency $\omega = 2.8/T_c$. Therefore, the frequency band considered useful for excitation is

$$\frac{2\pi}{TT_c} \le \omega_{\text{PRBS}} \le \frac{2.8}{T_c}. \tag{1}$$

We introduce two PRBS design parameters, α and $\beta > 0$, that define the minimum and the maximum frequency of interest in terms of the smallest corner-frequency of the closed-loop system. By introducing these parameters we can relate the PRBS design with the control system design specifications. If the *closed-loop dominant time constant* is denoted by τ_c, and parameters α, $\beta > 0$, then

$$\frac{2\pi}{\beta\tau_c} \le \omega \le \frac{2\pi\alpha}{\tau_c} \tag{2}$$

is the frequency band of interest. Comparing Eqs. 1 and 2:

$$T_c \le \frac{1.4\tau_c}{\pi\alpha} \tag{3}$$

$$T \ge \frac{\pi\alpha\beta}{1.4}. \tag{4}$$

The PRBS length $T = (2^n - 1)$ is chosen such that n is an integer and T satisfies Eq. 4. The number of registers n fixes the number of harmonics in the PRBS, and from Isermann (1980) 2^*n should be greater than the total number of model parameters. The criteria for selecting the clock time and the number of registers are thus obtained in terms of parameters α and β. Parameters α and β should be selected based on (1) how accurate the knowledge of the dominant time constant is, and (2) the frequency band of interest for control system design. For the examples we studied, $\beta \approx 3$ (from Goodwin and Payne, 1977) and $\alpha \approx \tau_p/\tau_c$ (or τ_c/τ_p, whichever is greater than 1) gave good results, where τ_p denotes the dominant process time constant and τ_c is the dominant closed-loop time constant. These values seem intuitively reasonable, because if β is very large and α is very small, then only steady-state gain information will be obtained. On the other hand, if β is very small and α is large, then the output signal will be lost in noise. If high accuracy of steady-state gain is required, then $\alpha \approx \tau_p/\tau_c$ (or τ_c/τ_p) and $\beta \approx 2$–4 should be used.

Choice of Amplitude, a. The amplitude of PRBS is selected based on the region over which the variable is expected to change and the knowledge about the process gain. For instance, if the controlled variable is expected to vary by 20% of the nominal steady-state value, and if the process gain is about 2, then the input signal is expected to vary by about 10% of the nominal value. Therefore, the amplitude of the PRBS is selected as 10% of the nominal value. Another criterion is to determine the change in the output variable that is necessary to obtain an output S/N > 5 and then using the process gain determine the PRBS amplitude. The amplitude of the PRBS may sometimes be restricted to a lower value because of process constraints, or constraints to maintain plant operation close to the nominal steady state.

Choice of k. The location of the kth register, the output of which is one of the inputs to the XOR gate, is selected to obtain a maximum-length binary sequence (MLBS) (Davies, 1970, for details). The length of a MLBS in one period is $(2^n - 1)T_c$. A table listing the location of the kth register for different values of n that generate a MLBS is given in Davies (1970).

Duration of the Test. The test duration should be fixed by considering the compromises between the cost of the identification test vs. the accuracy of the model required in the application. The longer the test duration, the better the quality of the parameter estimates. The minimum is one cycle of the PRBS sequence; the more we repeat the cycle the better the estimates.

Parameter estimation problem

The input–output data set generated by introducing a PRBS test signal at the plant input is used to estimate the model parameters using the *prediction error method*. Assume that the relationship between the plant input and the plant output is given by

$$y(k) = G(z^{-1})u(k - n_k) + e(k) \tag{5}$$

where $G(z^{-1})$ is an OE model,

$$G(z^{-1}) = \frac{B(z^{-1})}{A(z^{-1})} = \frac{b_1 z^{-1} + \cdots + b_{n_b} z^{-n_b}}{1 + a_1 z^{-1} + \cdots + a_{n_a} z^{-n_a}}, \tag{6}$$

$n_k \ge 1$ represents the time-delay, and $e(k)$ is Gaussian white noise with zero mean and unit variance. Let the plant be modeled as an OE model,

$$y_m(k) = \tilde{G}(z^{-1}, \theta)u(k - n_k) \tag{7}$$

where,

$$\tilde{G}(z^{-1}, \theta) = \frac{\tilde{B}(z^{-1})}{\tilde{A}(z^{-1})} = \frac{\tilde{b}_0 + \tilde{b}_1 z^{-1} + \cdots + \tilde{b}_{\tilde{n}_b} z^{-\tilde{n}_b}}{1 + \tilde{a}_1 z^{-1} + \cdots + \tilde{a}_{\tilde{n}_a} z^{-\tilde{n}_a}} \tag{8}$$

and $\theta = (\tilde{b}_0, \tilde{b}_1, \ldots, \tilde{b}_{\tilde{n}_b}, a_1, \ldots, \tilde{a}_{\tilde{n}_a})$. The order of the numerator and the denominator polynomials in Eq. 8 are assumed to be fixed. The order of the process and that of the model may be different, that is, $n_a \ne n_{\tilde{a}}$, $n_b \ne n_{\tilde{b}}$. Usually, a class of models with different values for $n_{\tilde{a}}$ and $n_{\tilde{b}}$ are used. An information criterion, such as Akaike's Information Theoretic Criteria, is then used to select the "best" model. The prediction error $\{\epsilon(k, \theta)\}$ attributed to $e(k)$ and the modeling error, is:

$$\epsilon(k, \theta) = y(k) - \tilde{G}(z^{-1})u(k - n_k) \tag{9}$$

$$\epsilon(k, \theta) = [G(z^{-1}) - \tilde{G}(z^{-1}, \theta)]u(k - n_k) + e(k). \tag{10}$$

The parameter θ is estimated by minimizing the l_2-norm of the prediction error. If

$$V(\theta) = \|\epsilon(k, \theta)\|_2^2,$$

then the parameter estimation problem can be stated as

$$\min_{\theta} V(\theta). \tag{11}$$

From Parseval's theorem, the frequency-domain equivalent of this optimization problem is

$$\hat{V}(\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi_{\epsilon}(\omega,\theta)\, d\omega \tag{12}$$

where Φ_{ϵ} is the power spectrum of the prediction error $\epsilon(k, \theta)$. Or, substituting $z = e^{-i\omega}$ and denoting the input power by Φ_u and the noise power by Φ_e, Eq. 12 can be rewritten as

$$\hat{V}(\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} |G(e^{i\omega}) - \tilde{G}(e^{i\omega},\theta)|^2 \Phi_u(\omega) + \Phi_e(\omega)\, d\omega. \tag{13}$$

The solution to the parameter estimation problem stated in Eq. 11 is

$$\theta = (\Phi^T\Phi)^{-1}\Phi^T Y_m, \tag{14}$$

where

$$\Phi = \begin{bmatrix} -y_m(1) & \cdots & -y_m(1-n_a) & u(1) & \cdots & u(1-n_b) \\ -y_m(2) & \cdots & -y_m(2-n_a) & u(2) & \cdots & u(2-n_b) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ -y_m(N-1) & \cdots & -y_m(N-1-n_a) & u(N-1) & \cdots & u(N-1-n_b) \end{bmatrix} \tag{15}$$

$$Y_m = [y_m(2) \quad y_m(3) \cdots y_m(N)]^T \tag{16}$$

$$\theta = \left[\tilde{a}_1\ \tilde{a}_2 \cdots \tilde{a}_{n_{\tilde{a}}}\ \tilde{b}_1\ \tilde{b}_2 \cdots \tilde{b}_{n_{\tilde{b}}}\right]^T. \tag{17}$$

The solution, θ, exists only when the covariance matrix $\Phi^T\Phi$ is nonsingular. This can be guaranteed by exciting the plant sufficiently at different frequencies. The condition of exciting the plant so that $\Phi^T\Phi$ is nonsingular is called *persistence of excitation.*

In the presence of measurement noise, $e(k)$, the model parameter estimates computed using a finite input–output data set has a bias and a variance given, respectively, by (Ljung, 1987):

$$\theta - \theta_0 = [\Phi^T\Phi]^{-1}\sum_{k=1}^{N-1}\varphi(k)e(k) \tag{18}$$

$$\sigma_{\theta}^2 = [\Phi^T\Phi]^{-1}\Phi^T R\Phi[\Phi^T\Phi]^{-1}, \tag{19}$$

where

$$\varphi(k) = [-y_m(k-1)\cdots -y_m(k-n_{\tilde{a}}-1)u(k-1) \cdots u(k-n_{\tilde{b}}-1)] \tag{20}$$

and R is the noise covariance matrix. If $e(k)$ can be characterized by white noise with zero mean and variance σ_n^2, Eq. 19 reduces to

$$\sigma_{\theta}^2 = \sigma_n^2[\Phi^T\Phi]^{-1}. \tag{21}$$

To determine the parameter covariance matrix σ_{θ}^2, the noise variance σ_n^2 should be known. If the modeling error $\epsilon(t, \theta)$ is attributed completely to noise, then for d model parameters ($d = n_a + n_b$) an unbiased estimate of the noise variance is (Ljung, 1987)

$$\hat{\sigma}_n^2 = V(\theta)/(N-d). \tag{22}$$

To minimize the parameter variance, the influence of the prefilter on $\hat{\sigma}_n^2$ and $\Phi^T\Phi$ will be used later in this article.

The parameter bias and variance (Eqs. 18 and 21) depend on the noise variance estimate. If the noise estimate is equal to the "true" noise, then there is no modeling error and the output S/N ratio is maximum. When the output S/N is maximum, the parameter bias and variance are minimum (Eqs. 18 and 21). We use this as a design criterion to select a prefilter.

Improving parameter estimates using prefilters

A prefilter (denoted by L) enhances the frequency band over which the input–output data have a large-output S/N ratio, and suppresses the frequencies over which the S/N ratio is small. By doing so, certain properties of the model can be enhanced or suppressed. When a prefilter is used in the identification of a linear and time-invariant model, the parameter estimation problem (Eq. 11) reduces to

$$\min_{\theta} V_F(\theta) \tag{23}$$

where

$$V_F(\theta) = \|\epsilon_F(k,\theta)\|_2^2$$

$$\epsilon_F(k,\theta) = L(z^{-1})\epsilon(k,\theta) = L(z^{-1})y(k) - \tilde{G}(z^{-1})L(z^{-1})u(k-n_k). \tag{24}$$

Equations 23 and 24 suggest that prefiltering noise is equivalent to filtering the input–output data using $L(z^{-1})$ followed by the standard least squares estimation. The problem stated in Eq. 23, in the frequency domain, is

$$\hat{V}_F(\theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[|G(e^{i\omega}) - \tilde{G}(e^{i\omega},\theta)|^2 \Phi_u(\omega) + \Phi_e(\omega) \right] \times L(e^{i\omega}) d\omega. \quad (25)$$

The bias in the parameter estimation therefore depends on the input power Φ_u, the error power spectrum Φ_e, the prefilter L, and the plant–model mismatch. After an identification test is conducted and the input–output data collected, the choice of the prefilter and the model order determine the parameter bias. In this article the focus is on the selection of the prefilter given that the model order is fixed.

The bias in parameter estimates can be distributed in the frequency domain by choosing a prefilter to suppress the undesired high-frequency noise contributions while enhancing the low-frequency information. In most control applications the low-frequency behavior of the plant is most important. By appropriate design of a prefilter, one can reduce the bias in this frequency band (Eq. 25). This can be achieved by

- Increasing $\Phi^T\Phi$ (Eq. 14) and
- Decreasing the cross-spectrum between $\varphi(k)$ and $e(k)$ (Eq. 18).

The parameter variance can be distributed by

- Increasing the input–output data set size (Eq. 22)
- Reducing the number of estimated parameters (Eq. 22) and
- Using prewhitening filters. This involves solving a weighted least-squares problem instead of the least-squares problem stated in Eq. 11, with the weighting matrix equal to the inverse of noise covariance matrix R^{-1}. The estimates obtained have minimum variance among the class of all linear unbiased estimators and this estimate is therefore called the *best linear unbiased estimate* (BLUE) (Goodwin and Payne, 1977).

While results on how the variance and bias can be distributed in the frequency domain are available in the literature (see Ljung, chap. 13, for example), the design of a prefilter has not been addressed directly. Some aspects of prefilter design were presented in Rivera et al. (1992). In the present article, we address some issues that were not explored by Rivera et al., namely,

- How should one select the prefilter $L(z^{-1})$ to obtain minimum bias and variance in the parameter estimates?
- How does the output S/N ratio of the data collected during an identification experiment influence this decision?

The knowledge of noise properties is limited by the ability to extract the contribution from the process in the noisy output measurements. If a low-pass filter with a small cutoff frequency is used, inaccurate parameter estimates are obtained as the plant data available at very low frequency is small. A low-pass filter with very high cutoff frequency yields a signal with large noise contribution; therefore the poor parameter estimates. The prefilter should be designed to have a cutoff frequency at some intermediate value while realizing its influence on the parameter bias and variance. Wavelets transforms provide a hierarchical decomposition of a signal with decreasing cutoff frequencies. This makes it possible to obtain a comprehensive picture of the compromises in the prefilter design. Thus, wavelet transforms are a natural choice in developing a prefilter design procedure.

Wavelet Transform

The reasons for using wavelet transforms as the tool for prefilter design are:

1. Wavelet transforms provide *multiresolution analysis* of the signal. Multiresolution analysis is equivalent to applying a family of low-pass filters with decreasing cutoff frequency on the signal.
2. Efficient wavelet transform algorithms are available for signal decomposition and reconstruction.
3. The projection of white noise on the scaling function is white noise, with variance inversely proportional to scale. This makes it possible to derive properties of output signal from an ARMA process driven by white noise, and to determine the scale at which the output S/N ratio will peak.

An alternative is to use a series of low-pass filters (e.g., Butterworth filters) with decreasing cutoff frequency. However, filtering white noise using Butterworth filters does not yield white noise. Further, evaluating the frequency at which the output S/N ratio will peak requires a number of iterations.

Discrete wavelet transforms and multiresolution analysis

The discrete wavelet transform (DWT) of a function f in the space of absolutely square integrable functions, $L^2(R)$, is defined as (Daubechies, 1992):

$$\langle f, \psi_{m,n} \rangle = |a_0|^{-m/2} \int_{-\infty}^{\infty} f(t) \psi(a_0^{-m} t - n b_0) dt \quad (26)$$

$$= W_d f(m,n) \quad (27)$$

Typically, $a_0 = 2$ and $b_0 = 1$ are used. The inner product in Eq. 26 for various values of parameters (m, n) are called the *discrete wavelet coefficients*. The family of functions $\{\psi_{m,n}\}$—called *wavelets*—are generated by dilation and translation of a single prototype function, $\psi(t)$, called the *mother wavelet*. The family of wavelets are generated using $\psi_{m,n}(t) = |2|^{-m/2} \psi(2^{-m}t - n)$; m and n are called the dilation and the translation parameters, respectively.

The function f is reconstructed from the discrete wavelets coefficients using

$$f(t) = \frac{2}{A+B} \sum_{m,n \in Z} \langle f, \psi_{m,n} \rangle \psi_{m,n}(t) + R, \quad (28)$$

where

$$A\|f\|^2 \le \sum_{m,n} |\langle f, \psi_{m,n} \rangle|^2 \le B\|f\|^2, \quad (29)$$

$A > 0$ and $B < \infty$ are called the frame bounds, and R is the error in reconstruction. For orthogonal wavelets $B = A = 1$, $R = 0$, and an exact reconstruction is obtained. The reconstruction of a function from the discrete wavelet coefficients is guaranteed if the wavelet $\psi(t)$ satisfies the zero mean admissibility condition:

$$\int_{-\infty}^{\infty} \psi(t)\, dt = 0. \quad (30)$$

In this article we use orthonormal wavelets and we restrict the discussion to this class of wavelets only. Daubechies orthonormal wavelets are used in this article. Filter coefficients for these wavelets with different regularity and support are available in Daubechies (1992).

The discrete wavelet coefficients of a signal are obtained by computing the inner product, defined in Eq. 26, for all values of dilation parameter $m \in Z$ and translation parameter $n \in Z$. Computing the wavelet coefficients using this approach, however, is very inefficient. Mallat (1989) proposed an efficient hierarchical decomposition algorithm for the computation of discrete wavelet coefficients. The algorithm, called *multiresolution analysis*, interprets Eq. 28 as

$$f(t) = \sum_{m \in [1,L]} \langle f, \psi_{m,n} \rangle \psi_{m,n}(t) + \langle f, \phi_{L,n} \rangle \phi_{L,n}(t). \quad (31)$$

The function f is represented as sum of the projections on wavelets between scales $m = 1$ to L, and the projection on the scaling function $\phi(t)$ at $m = L$. Wavelets have a frequency response like that of a band-pass filter. The frequency response peaks at a lower frequency as m increases. The scaling functions $\phi_{m,n}(t)$, generated by dilating and translation of the scaling function $\phi(t)$, has a frequency response like a low-pass filter. The cutoff frequency of this filter decreases as m increases. Equation 31 can therefore be interpreted as a set of signals resulting from smoothing of a function f from scale $m = 1$ to scale $m = L$. The information lost in obtaining the smooth version is captured by the projections on the wavelets, and are referred to as the *details*.

Let the operator $\boldsymbol{P}_m$ denote the projection of f on the scaling function at scale m, that is,

$$\boldsymbol{P}_m f(t) = \sum_n \langle f, \phi_{m,n} \rangle \phi_{m,n}(t) \quad (32)$$

and $\boldsymbol{Q}_m$ be the projection of f on the wavelet at scale m,

$$\boldsymbol{Q}_m f(t) = \sum_n \langle f, \psi_{m,n} \rangle \psi_{m,n}(t). \quad (33)$$

Since the space spanned by the wavelets and the scaling functions at a given scale m is orthogonal to the space spanned by the scaling function at scale $m-1$, we have the following recursive decomposition algorithm:

$$\boldsymbol{P}_{j-1} f = \boldsymbol{P}_j f + \boldsymbol{Q}_j f. \quad (34)$$

Implementation of the preceding decomposition can be done efficiently using the filter coefficients. Computational aspects of wavelet transforms and some applications in chemical engineering are given in Joseph and Motard (1994).

Before describing the prefilter design, it is necessary to derive some properties of the wavelet transform of (1) a PRBS signal, (2) the output signal from an ARMA model, and (3) white noise. The properties (like variance) of the wavelet transform of an ARMA model driven by PRBS yield nonlinear equations, the analytical solution for which is difficult to simplify. Since the PRBS signal has autocovariance like that of white noise, the relationship between variance of the output signal from an ARMA process driven by white noise and the scale is derived. Expressions to estimate the scale at which the input signal to output noise ratio (hereafter referred to as the input S/N ratio and the output S/N ratio) peak are derived. Using results from Fourier analysis and linear control theory, it can be shown that the scale at which the input S/N ratio and the output S/N ratio peak is the same if the plant is linear and time-invariant. However, the input and output S/N ratios may peak at different scales if the plant is quasilinear and/or time-variant, or if the noise estimate is inaccurate.

Wavelet transform of white noise

Let $e(t)$ denote the Gaussian white noise signal with zero mean and variance σ_e^2. Denote the DWT of $e(t)$ by $W_d e(m, n)$, where m denotes the scale and n is the translation parameter. Grossmann (1986) and later Mallat and Hwang (1992) show that the DWT of $e(t)$ is white noise with zero mean and variance inversely proportional to scale m,

$$E(|W_d e(m,t)|^2) = \frac{\|\psi\|^2 \sigma_e^2}{2^m} \quad (35)$$

It is also shown that the projection of white noise $e(t)$ on the scaling function, $\boldsymbol{P}_m e(t)$, has variance inversely proportional to scale m:

$$E(|P_m e(t)|^2) = \frac{\|\phi\|^2 \sigma_e^2}{2^m}. \quad (36)$$

Wavelet transform of ARMA process driven by white noise

Consider a single input, single output (SISO) process driven by white noise $u(k)$. The process output $y_1(k)$ is clouded with Gaussian white noise $e(k)$ (Figure 1). Thus,

$$y(k) = y_1(k) + e(k), \quad (37)$$

where $y_1(k)$ is the output from a process driven by white noise with zero mean and variance σ_u^2. Assume the process to be an ARMA model with a denominator polynomial of order $n_a = 1$, numerator polynomial order of $n_b = 1$, and a time-delay of $n_k = 1$ [denoted by ARMA(1,1,1)],

$$y_1(k) = \frac{b_1 z^{-1}}{1 + a_1 z^{-1}} u(k) \quad (38)$$

or,

$$y_1(k) = b_1 u(k-1) - a_1 y_1(k-1). \quad (39)$$

The variance of time series $\{y_1\}$ is

$$\sigma_{y_1}^2 = \frac{b_1^2}{1 - a_1^2} \sigma_u^2. \quad (40)$$

This equation relates the input and the output variance of the process at scale $m = 0$.

Next, we compute the variance of $\{y_1\}$ as a function of scale. Since $\{u\}$ is assumed to be white noise, the projections of $\{u\}$ on the scaling function at scale m (denoted as $\boldsymbol{P}_m u$) has variance decaying as given in Eq. 36. Due to the cross-correlation between the projections of $\{y_1\}$ on the scaling function at scale $m(\boldsymbol{P}_m y_1)$ at different scales, it *does not* have its variance decreasing as given in Eq. 40. An intuitive way to determine the variance of $\boldsymbol{P}_m y_1$ as a function of scale can be obtained using the Haar scaling functions and wavelets (see Daubechies, 1992). The Haar scaling function at $m = 0$ is a pulse of unit amplitude and width equal to the sample time. At scale m, the pulse has a width of 2^m and an amplitude of $1/2^m$. $\boldsymbol{P}_m y_1$ is equivalent to taking averages over nonoverlapping time windows of width 2^m. Therefore, the variance of $\boldsymbol{P}_m y_1$, denoted by $\sigma^2_{y1,m}$, is equivalent to computing the autocovariance function of $\{y_1\}$ at lag 2^m. We use the autocovariance function as an estimate of $\sigma^2_{y1,m}$, as it is not possible to derive an analytical expression for $\sigma^2_{y1,m}$ independent of the choice of scaling function and wavelet. Similar approximations are made while computing the autocovariance function of a PRBS. This interpretation can be extended to smoother scaling functions and wavelets as well (such as, Daubechies wavelets).

Multiplying Eq. 39 by $y_1(k+l)$, the autocovariance function at lag l is given by

$$\gamma_{y_1,l} = -a_1\gamma_{y_1,l-1} + b_1\gamma_{uy_1,l-1}, \tag{41}$$

where $\gamma_{y_1,l}$ is the autocovariance function of $\{y_1\}$ at lag l; $\gamma_{uy_1,l}$ is the cross-covariance function between $\{u\}$ and $\{y_1\}$ at lag l. Since $\{u\}$ and $\{y_1\}$ are uncorrelated for $l > 0$,

$$\begin{aligned} \gamma_{yu_1,l} &= 0 \qquad l > 0 \\ &\neq 0 \qquad l \le 0. \end{aligned} \tag{42}$$

Therefore, for $l \ge 1$

$$\gamma_{y_1,l} = -a_1\gamma_{y_1,l-1}. \tag{43}$$

For $l = 0$,

$$\gamma_{y_1,0} = -a_1\gamma_{y_1,1} + b_1\gamma_{uy_1,-1}. \tag{44}$$

Multiplying Eq. 39 by $u(k-1)$ and setting $E[y_1(k-1)\,u(k-1)] = 0$, we get

$$\gamma_{uy_1,-1} = b_1\sigma_u^2. \tag{45}$$

From Eqs. 43, 44 and 45,

$$\sigma^2_{y_1} = \gamma_{y_1,0} = \frac{b_1^2}{1-a_1^2}\sigma_u^2 \tag{46}$$

and

$$\gamma_{y_1,l} = (-a_1)^l\frac{b_1^2}{1-a_1^2}\sigma_u^2. \tag{47}$$

For, $l \ge 1$, setting $l = 2^m$ the autocovariance function for the output y is

$$\gamma_{y,l=2^m} = a_1^{2^m}\frac{b_1^2}{1-a_1^2}\sigma_u^2 + \frac{\sigma_e^2}{2^m}. \tag{48}$$

Thus, $\boldsymbol{P}_m y$ has variance decaying with scale and the plant dynamics (parameters a_1 and b_1 in Eq. 48). This result is used in the section on estimation of scale to determine the scale at which the output S/N ratio peaks.

Wavelet transform of a PRBS signal

The autocovariance function of a PRBS peaks at lag $\tau = 0$, is linear for $\tau \in [0, T_c]$, and is a constant for lag $T_c \le \tau \le (T-1)T_c$ (Davies, 1970),

$$\gamma_{u,\tau} = a^2\left[1 - \frac{\tau(T+1)}{TT_c}\right]. \tag{49}$$

For small values of clock time T_c, the autocovariance function is like that of white noise ($+1$ at lag 0, and 0 elsewhere). And for large values of T_c, the autocovariance function is like that of white noise filtered through a low-pass filter with a cutoff frequency ω_c. As $T_c \to \infty$, the cutoff frequency $\omega_c \to 0$.

The variance of $\boldsymbol{P}_m u$ can be approximated by the autocovariance function at lag 2^m:

$$\sigma^2_{u,2^m} = a^2\left[1 - \frac{2^m(T+1)}{TT_c}\right]. \tag{50}$$

Accurate estimates are obtained using this expression only for lags less than the clock time, that is, $l < T_c$. This is because the autocovariance function of a PRBS with $l > T_c$ yields an inaccurate approximation to $\boldsymbol{P}_m u$. Next we compare the estimates using Eq. 50 with those obtained numerically and illustrate the properties of wavelet transform of PRBS. Daubechies wavelet with eight filter coefficients is used in all computations.

Figure 2 graphs the PRBS generated using clock time $T_c =$ 50, 20, and 5. Figure 3 plots the variance of $\boldsymbol{P}_m u$ vs. the scale. Following are some observations that can be made from these figures:

- As the clock time $T_c \to 0$, the PRBS has properties similar to Gaussian white noise, and the variance is inversely proportional to 2^m (Eq. 36 and Figure 3).
- As T_c increases, the PRBS properties are like that of filtered Gaussian white noise with cutoff frequency of the filter $\omega_c \to 0$. $\boldsymbol{P}_m u$ has a larger variance when the clock time increases (Figure 3).
- The estimate for the variance of $\boldsymbol{P}_m u$ using the autocovariance function is a reasonable approximation to the true value (Figure 3).

The properties of wavelet transform of a PRBS and the estimate of the variance are used in the next section to estimate the scale at which input S/N ratio peaks.

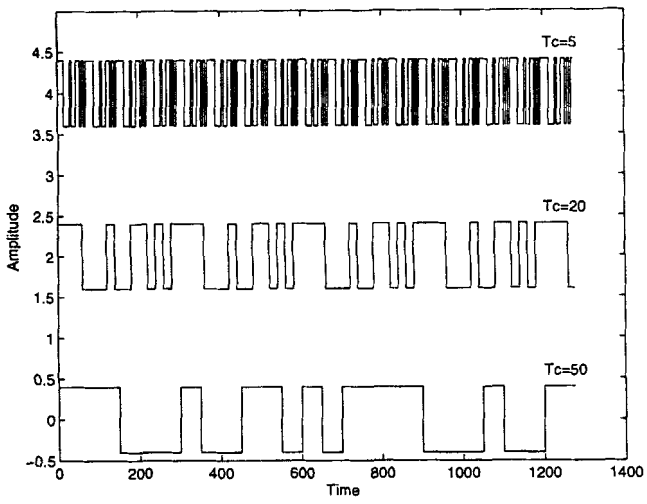

Figure 2. PRBS generated using $T_c = 50$, 20, 5, $a = 0.4$, $n = 4$, $T = 15$, $k = 3$, sample time = 5, and PRBS length = 256.

Estimation of scale at which input and output S/N ratio peak

The variance of white noise decays with scale at a rate faster than the variance of the output series $\{y_1\}$ and the PRBS signal $\{u\}$ (see Eqs. 47 and 50). The input and output S/N ratios may therefore peak at an intermediate scale. An intuitive explanation for the output S/N ratio to peak in terms of filtering is *a low-pass filter with a very small cutoff frequency yields a signal with small strength* (S/N ratio is small); *a very high cutoff frequency yields a signal with high noise strength* (S/N ratio is small); *while an intermediate cutoff value yields a stronger output signal strength than the noise strength* (S/N ratio large). The parameter estimates at the scale where the output S/N ratio peaks have a lower bias and variance than at scale $m = 0$ (unfiltered signal). The scale at which the input and the output S/N peak is derived next.

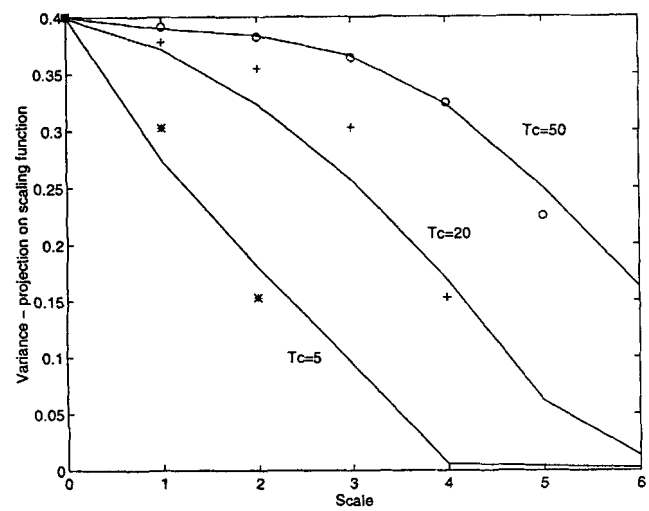

Figure 3. Influence of clock time T_c on the variance of the PRBS projections on the scaling function.

0, +, and * represent estimates of variance (using Eq. 50) for $T_c = 50$, 20 and 5, respectively.

Expression for Scale m_u^* at Which Input S/N Ratio Peaks. From Eqs. 50 and 36 the ratio of the PRBS variance to the noise variance is

$$r_{m,ue}^2 = \sigma_u^2 \left[1 - \frac{2^m(T+1)}{TT_c}\right] \frac{2^m}{\sigma_e^2} \quad 0 \le 2^m < T_c. \tag{51}$$

Differentiating Eq. 51 with respect to m and setting it to zero yields

$$m_u^* = \frac{\ln\left(\frac{TT_c}{T+1}\right)}{\ln 2} - 1. \tag{52}$$

The second derivative with respect to m at $m_u^* = \{\ln [TT_c/(T+1)]/\ln 2\} - 1$ is

$$1 - 2 \ln 2 < 0. \tag{53}$$

Thus, the input S/N ratio, peaks at scale m_u^*. Note that the preceding derivation assumes $0 \le 2^{m_u^*} \le T_c$.

Expression for Scale m_y^* at Which the Output S/N Ratio Peaks. From Eqs. 36 and 48, the output S/N ratio at scale m is

$$r_m = \frac{a_1^{2^m} b_1^2}{1 - a_1^2} \times \frac{1}{1/2^m}, \tag{54}$$

Differentiating Eq. 54 with respect to scale m and equating to zero, we get

$$m_y^* = [\ln (-1/\ln |a_1|)]/\ln 2, \tag{55}$$

where m_y^* is the scale at which the output S/N ratio r_m is an extremum. Differentiating Eq. 54 twice with respect to m, substituting for m from Eq. 55, and since $|a_1| < 1$ and $r_m > 0$, we get,

$$\frac{d^2 r_{m_y^*}}{dm^2} < 0. \tag{56}$$

The output S/N ratio is a maximum at scale m_y^* and depends only on the value of parameter a_1. Since the autocovariance function of a PRBS at small values of T_c is similar to that of white noise, the preceding derivation can be extended to PRBS signals as well.

Prefilter Design Methodology Using Wavelets

The *prefilter design* using wavelets consists of the following steps:

- Input–output data from an off-line identification experiment are collected.
- The input and the output data are projected on the scaling function at different scales.
- Model parameters are estimated using projections of the input and the output at the same scale.
- The unmodeled portion of the signal at each scale is assumed to be noise. The output S/N ratio is then computed.

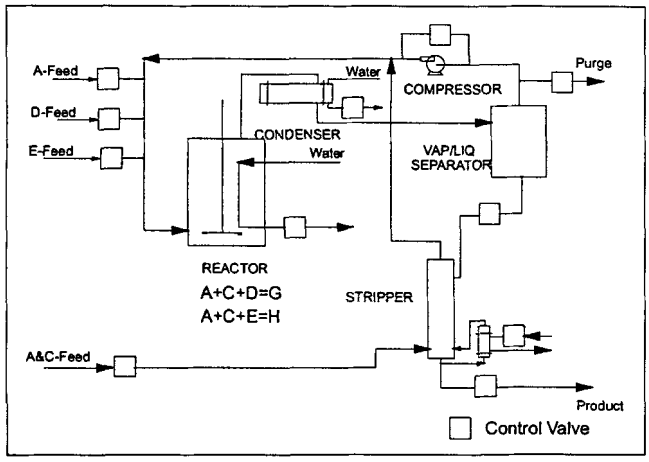

Figure 4. Tennessee Eastman Problem Process Flowsheet.

• Parameter estimates at scale where the output S/N ratio is maximum is recommended for use.

Parameter estimates up to the scale at which the energy of the projected signal is less than 20% of the raw signal are considered in the prefilter design. Parameters for higher scales are inaccurate, as the amplitude of excitation is small and the output S/N ratio is small. The preceding design steps are equivalent to using a number of low-pass filters with different cutoff frequencies (ranging from light filtering to heavy filtering) and then selecting the filter that yields the best parameter estimate.

Case Study: Tennessee-Eastman Challenge Problem

For a thorough evaluation of the prefilter design described in the preceding section, one must apply the technique to identify models for a real chemical plant. However, it is difficult in a university to conduct full-scale experiments for this purpose. For this reason we use the Tennessee-Eastman (TE) Industrial Problem put forth by Downs and Vogel (1993) to evaluate the prefilter design. The TE problem simulates an industrial chemical process consisting of a nonlinear, multicomponent, two-phase reactor, a condenser, a vapor–liquid separator, and a stripper (Figure 4).

Process description

Figure 4 shows the flowsheet of the TE process. The process produces two products (G, H) from four reactants (A, C, D, E). There is an inert (B) that enters with a feed stream and a byproduct (F) that is produced due to a side reaction. The reactions involved are

$$A_{(g)} + C_{(g)} + D_{(g)} \rightarrow G_{(l)} \quad \text{(R1)}$$

$$A_{(g)} + C_{(g)} + E_{(g)} \rightarrow H_{(l)} \quad \text{(R2)}$$

$$A_{(g)} + E_{(g)} \rightarrow F_{(l)} \quad \text{(R3)}$$

$$3D_{(g)} \rightarrow 2F_{(l)}. \quad \text{(R4)}$$

All reactions are irreversible and exothermic. The reaction rates are a function of temperature through an Arrhenius expression.

The gaseous reactants are fed to the reactor where they form volatile liquid products. The reactor product stream consists of unreacted noncondensable reactants and condensable products. This stream passes through a cooler that condenses the products partially. The stream then passes through a vapor–liquid separator. Noncondensable components are recycled through a compressor to the reactor feed, while the condensed components are fed to a product stripper. To avoid accumulation of inert B, part of the recycle stream is purged. Products G and H exit the system from the product stripper bottoms.

The inner loop controllers that reject fast and local disturbances are summarized in Table 1. The proportional gain for these controllers is computed from the base case values assuming a proportional band of 100%. Integral time for flow controllers is set equal to 0.1 min (Luyben, 1989). The tuning parameters for the temperature controllers are obtained using a trial-and-error procedure. We start with a large gain and integral time and half it till satisfactory performance is obtained. The inner loop controller sample time is 1 s.

Stabilizing controllers required to maintain material and energy balance at the nominal steady state (50G/50H and 14,078 kg/h production rate) are then designed. The level controllers for the reactor, separator, and stripper using E-feed, separator underflow, and stripper underflow, respectively, stabilize the process. This configuration is similar to that used by McAvoy and Ye (1994). However, the tuning parameters (Table 2) used here are different from those reported in McAvoy and Ye. A sample time of 30 s is used for level controllers.

The next step in the decentralized/multivariable controller design is to obtain the process model. The process model is typically used for steady-state/dynamic RGA analysis, controller tuning, and prediction of controlled variable trajectory.

Problem definition

The main reactions that form desired products G and H in the process are (R1) and (R2). The production rate is therefore most influenced by A and C feed, provided sufficient reactants D and E are available and enough cooling is available. In this case study:

• We design an off-line identification test to determine the LTI model relating A and C feed flow rate setpoint and the production rate.

• We design a prefilter using the method of sieves from the plant input–output data.

• We compare the accuracy of the process gain estimates and the prediction error using (1) a prefilter designed using the method described in the section on prefilter design, (2) an autoregressive (ARX) model, and an ARMA model.

Introducing a positive step change in A and C feed flow rate causes the reactor pressure to reach the shutdown limit. Therefore, step testing is considered impractical for identification in this problem. A PRBS signal that perturbs the plant inputs such that the output value is close to the nominal steady-state value is used instead. However, the plant data

Table 1. Inner Loop Controllers

Controlled Variable	Manipulated Variable	Proporational Gain	Integral Time (min)
A-feed xmeas(1)	A-feed flow xmv(3)	100 (%/kscmh)	0.1
D-feed xmeas(2)	D-feed flow xmv(1)	0.0172 (%/kg/h)	0.1
E-feed xmeas(3)	E-feed flow xmv(2)	0.008 (%/kg/h)	0.1
A & C-feed xmeas(4)	A & C-feed flow xmv(4)	6.5573 (%/kscmh)	0.1
Purge rate xmeas(10)	Purge valve xmv(5)	118 (%/kscmh)	0.1
Prod. sep. underflow xmeas(14)	Separator pot liquid flow xmv(6)	1.2 (%/m^3/h)	0.1
Stripper underflow xmeas(17)	Stripper liquid production flow xmv(7)	1.2 (%/m^3/h)	0.5
Stripper steam flow xmeas(19)	Stripper steam valve xmv(9)	0.206 (%/kg/h)	0.1
Reactor cooling water outlet temperature xmeas(21)	Reactor cooling water flow xmv(10)	−5 (%/°C)	0.3
Sep. cooling water outlet temperature xmeas(22)	Cond. cooling water flow xmv(11)	−3 (%/°C)	0.3

collected by performing a PRBS testing has a small output S/N ratio due to the constraint on the PRBS amplitude. A prefilter is therefore necessary to obtain reliable parameter estimates.

Test signal design

Preliminary tests suggest that the gain and the dominant time constant to the transfer function relating A and C feed flow setpoint and the product flow is about 4 m^3/h/% and 30 min, respectively. Further, the preliminary tests also suggest that the output response is like that of an underdamped second-order process. A closed-loop control system with time constant about three times the open-loop time constant is to be designed using the identified model. The control system designed is to be tested for a −15% step change in production rate. If A and C feed flow setpoint is used for production rate control, then A and C feed flow rate will decrease by about 4% when the production rate decreases by 15%. Therefore, the magnitude of PRBS is fixed at 4% of the nominal steady-state value. Since the dominant closed-loop time constant is 90 min, substituting $\alpha = 2.5$ and $\beta = \sqrt{3}$ in Eqs. 3 and 4, we get

$$T \geq 9.72,\ T_c \leq 16.04. \tag{57}$$

Therefore, $n = 4$ and $T_c = 15$ are used to generate the PRBS test signal. A sample time of 5 min is used for identification. The PRBS test signal is added to the A and C feed flow setpoint and the output data are collected.

Table 2. Stabilizing Controller Tuning Parameters

Controller	Prop. Gain	Integral Time
Reactor level	300 %/kg/h	0.32 h
Separator level	−0.301 %/m^3/h	0.8 h
Stripper level	−0.1789 %/m^3h	1 h

Parameter estimation

Wavelet transform of the input–output data using Daubechies wavelet with eight filter coefficients is graphed in Figure 5. The variance of the sieved output data at scale 5 is less than 20% of the output data variance. Hence, the sieved set of input–output data between scales 0 and 4 are used to estimate model parameters. A second-order OE model [denoted as OE(2,1,1)] of the form

$$y(k) = \frac{b_1 z^{-1}}{1 + a_1 z^{-1} + a_2 z^{-2}} u(k) + e(k) \tag{58}$$

is assumed. This model results in minimum Akaike's information theoretic criteria among first and second order OE models. Model parameter estimates and their variance at each

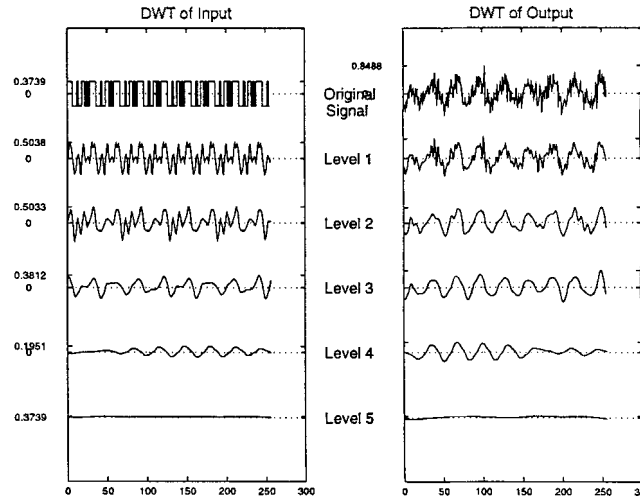

Figure 5. Discrete wavelet transform of input–output data collected from TE process.

Table 3. Parameter Estimates at Various Scales

Scale	b_1	a_1	a_2	σ_{b1}	σ_{a1}	σ_{a2}
4	−0.6927	0.4338	−0.3684	0.3856	0.8864	0.5351
3	0.0437	−1.8423	0.8728	0.0024	0.0074	0.0079
2	0.0407	−1.8457	0.8774	0.0030	0.0095	0.0101
1	0.3017	−0.0876	−0.8276	0.0511	0.2559	0.2530
0	0.2632	−0.1604	−0.7390	0.0478	0.2571	0.2545

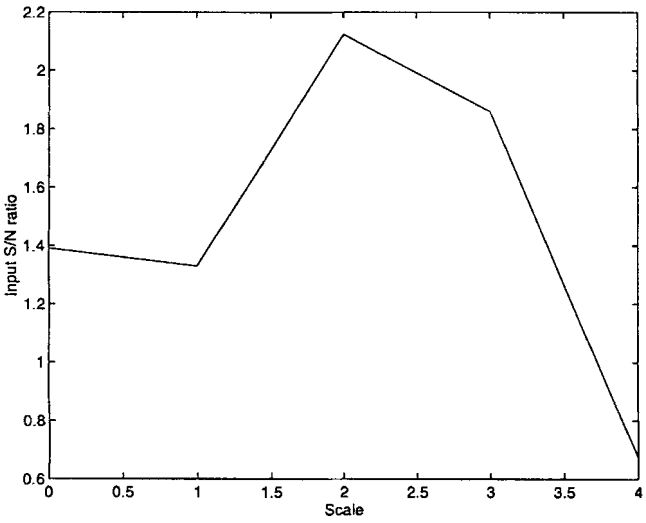

Figure 6. Input S/N ratio.

scale computed using the prediction error method are given in Table 3. Figures 6 and 7 plot the influence of scale on the estimated input S/N ratio and output S/N ratio, respectively. Figure 8 compares the predictions using the identified model with the plant data. Figure 9 plots the output residual or the prediction error. The plant data used in Figures 8 and 9 were taken from a test data set that was not used to fit the model parameters.

Following are some observations and corresponding deductions based on the results in Table 3 and Figures 6–9:

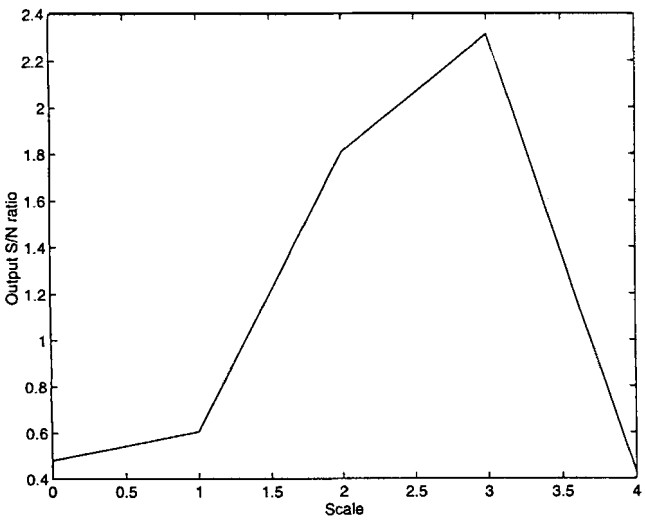

Figure 7. Output S/N ratio.

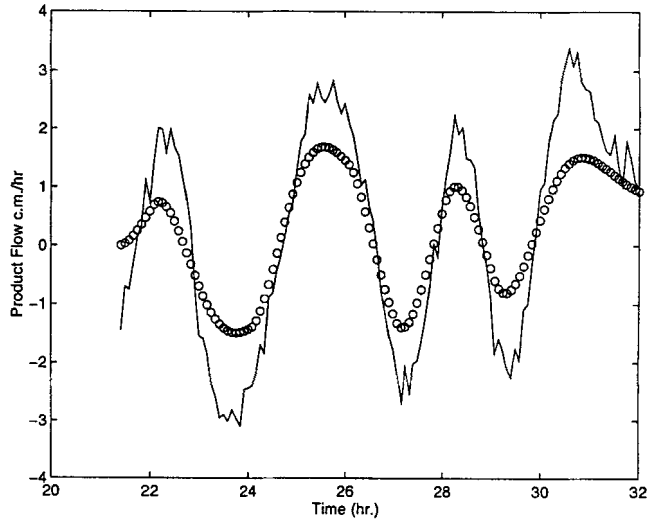

Figure 8. Prediction output value and the "true" plant measurements for the model obtained using the method of sieves.

o estimated value, — true value.

- The actual input S/N ratio and output S/N ratio peak at scales 2 and 3, respectively; see Figures 6 and 7. Using Eqs. 52 and 55, the input S/N ratio is expected to peak at scale $2.81 \approx 3$ and the output S/N ratio (assuming a process time constant of 30 min) is estimated to peak at scale $2.58 \approx 3$.
- As the output S/N ratio peaks at scale 3, the parameter estimates at scale 3 will have minimum variance and bias. The parameter estimates at this scale yield a process gain of 1.44.
- The presence of plant–model mismatch is indicated by the fact that the input S/N ratio and the output S/N ratio peak at different scales. However, Figures 8 and 9 suggest that the identified model is a reasonable approximation to the plant dynamics.

We next compare the process gain estimate and the prediction error obtained using the wavelet prefilter with the ARX

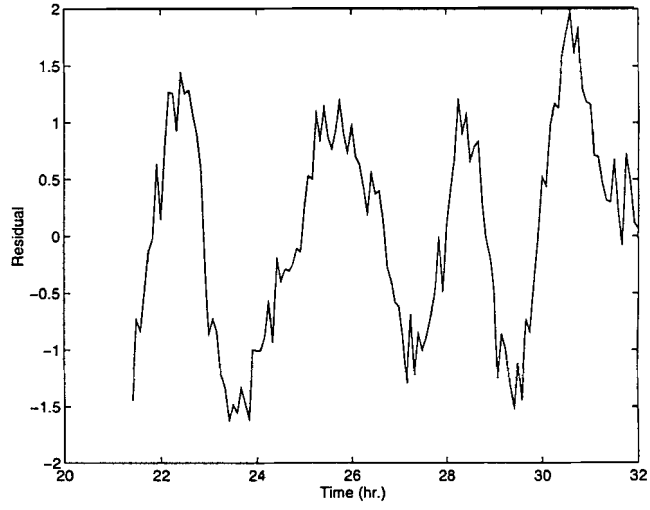

Figure 9. Output residual or prediction error for the model obtained using the method of sieves.

model and the ARMA model. The process gain of the "true" plant is obtained using step testing. The *average* process gain, obtained by introducing a step change of magnitude $\pm 4\%$ in A and C feed flow setpoint, estimated using least squares estimation, is 2.58 $m^3/h/\%$. Note that such a step test would be considered impractical as the reactor pressure reaches close to the shutdown limit. Since we could conduct such a test on the simulation of the plant, we can get parameter estimates from step tests.

For comparison we use an ARX(2,1,1) model,

$$y(k) = \frac{b_1 z^{-1}}{1 + a_1 z^{-1} + a_2 z^{-2}} u(k) + \frac{1}{1 + a_1 z^{-1} + a_2 z^{-2}} e(k), \quad (59)$$

a second-order ARMA(2,1,1,1) model,

$$y(k) = \frac{b_1 z^{-1}}{1 + a_1 z^{-1} + a_2 z^{-2}} u(k) + \frac{1 + c_1 z^{-1}}{1 + a_1 z^{-1} + a_2 z^{-2}} e(k), \quad (60)$$

and a second-order ARMA(2,1,2,1) model,

$$y(k) = \frac{b_1 z^{-1}}{1 + a_1 z^{-1} + a_2 z^{-2}} u(k) + \frac{\cdot 1 + c_1 z^{-1} + c_2 z^{-2}}{1 + a_1 z^{-1} + a_2 z^{-2}} e(k). \quad (61)$$

Parameter estimates for each of the preceding models using the same plant input–output data as in Figure 5, the process gain, and the prediction error are summarized in Table 4.

The following remarks can be made based on the results in Table 4:

- When no prefilter is used, the process gain is 2.6 (Table 3, scale 0) and the prediction error is 15.3.
- Prefiltering using wavelets gives better process gain estimates compared to ARX(2,1,1), ARMA(2,1,1,1), and ARMA(2,1,2,1). The gain estimates using the prefilter compared with ARX(2,1,1), ARMA(2,1,1,1), and ARMA(2,1,2,1) are better by 96%, 43%, and 22%, respectively.

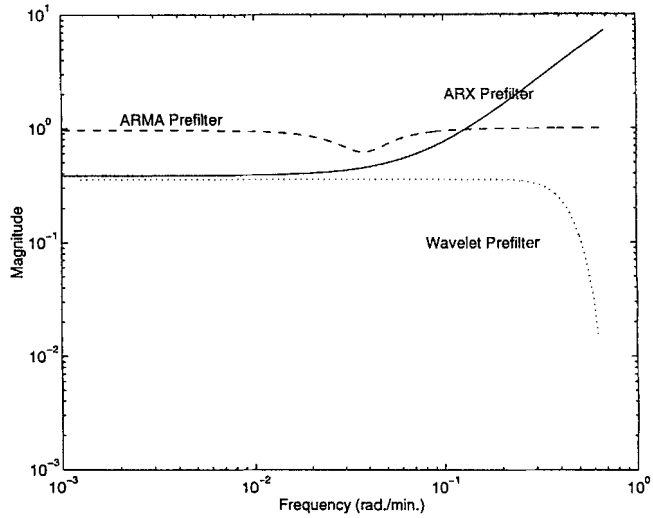

Figure 10. Frequency response of ARX, ARMA noise models and the wavelet prefilter.

- The gain obtained when no prefiltering is used, though better than the other methods, results in large prediction error.
- The prediction error using the wavelet prefilter is the least.

The prefilter used in ARX(2,1,1), ARMA(2,1,2,1), and using wavelets determine the accuracy of the parameter estimates. The prefilter using wavelets is more effective in eliminating the high-frequency noise. Figure 10 shows the frequency response of the prefilter used in these three methods. For the wavelet method, the prefilter is Daubechies scaling function with eight filter coefficients at scale 3. This prefilter has a frequency response like that of a low-pass filter, and the high-frequency information is filtered before the parameter estimation stage. The ARX model and the ARMA model use the following prefilters, respectively:

$$\frac{1 + a_1 z^{-1} + a_2 z^{-2}}{1} \quad (62)$$

$$\frac{1 + a_1 z^{-1} + a_2 z^{-2}}{1 + c_1 z^{-1} + c_2 z^{-2}}. \quad (63)$$

The parameters in Eqs. 62 and 63 estimated using the prediction error are given in Table 4.

Table 4. Summary of Results

Method	b_1	a_1	a_2	c_1	c_2	Gain	Pred. Error
No filtering OE(2,1,1)	0.2632	−0.1604	−0.7390	0	0	2.61	15.29
ARX(2,1,1)	0.0251	−0.3421	−0.2751	0	0	0.06	20.61
ARMA(2,1,1,1)	0.0829	−0.870	−0.0288	0.5796	0	0.82	18.09
ARMA(2,1,2,1)	0.0378	−1.8488	0.8825	−1.7755	0.8105	1.12	11.74
Filtering using wavelets OE(2,1,1)	0.0437	−1.8423	0.8728	0	0	1.44	10.03
Step test (1,1,1)	0.4234	−0.8358	0	0	0	2.58	14.19

When an ARMA model is used, it is necessary to specify the order of the denominator polynomial in Eq. 63. A wrong choice of the polynomial order gives poor parameter estimates (compare estimates for ARMA(2,1,1,1) and ARMA(2,1,2,1) in Table 4). Assuming the model order selected is correct and infinite input–output data are available, the theoretically optimal filter for this problem is no filtering. This is because for the TE problem it is known that uncorrelated noise is being added to the measurements and for this case the least squares estimate is unbiased. Therefore, by forcing a noise model using ARX(2,1,1) and ARMA(2,1,1,1) results in poor parameter estimates. Using ARMA(2,1,2,1) improves the parameter estimates significantly as the coefficients c_1, a_1, c_2, and a_2 are such that the noise model is nearly uncorrelated.

Figure 10 plots the prefilter frequency response obtained for ARX, ARMA, and wavelets. The frequency response obtained for the ARX prefilter suggests that a significant amount of high-frequency information is being retained. The ARMA prefilter suggests that the prefilter is unity, except for the hump. This results in poor parameter estimates using the ARX model and explains the improvement in estimates using ARMA(2,1,2,1). The wavelet prefilter maximizes the S/N ratio by separating the low-frequency information, the noise retained is uncorrelated (note that wavelet transform of white noise is again white noise), and this yields better parameter estimates than ARMA(2,1,2,1). The estimated S/N ratio for ARX(2,1,1), ARMA(2,1,2,1) and the wavelet prefilters are 0.047, 0.9529, and 2.3, respectively. Therefore the improved parameter estimates using the wavelet prefilter.

Conclusions

In this article we have given an explicit relationship between the PRBS design parameters and the process knowledge. We presented a systematic prefilter design using multiresolution analysis. Designing such prefilters yields parameter estimates with smaller variance and bias. We illustrate the application of the design procedure on the TE process.

Comparisons were also made between the method of sieves using wavelets with that using Butterworth low-pass filter (of order 5) with different cutoff frequencies. Results (not reported here) suggest that wavelet *does not* provide superior estimates than those obtained using Butterworth filters when the cutoff frequency is close to that obtained using the data up to the scale where the S/N ratio peaks. The advantage of using this design procedure is that it allows the engineer to choose the cutoff frequency based on a comprehensive and an efficient tool with explicit information about the compromises between various designs.

Acknowledgments

Financial support from National Science Foundation (Grant CTS-91-08649), is gratefully acknowledged. Srinivas Palavajjhala would like to acknowledge the use of the MATLAB PRBS script file written by Michael Altmann, Northern Telecom Limited, Canada.

Notation

G = a transfer function
P = PRBS period
R = low-pass filter
t = time
V = l_2-norm of prediction error
z = shift operator
$\langle \cdots, \cdots \rangle$ = inner product
$\| \cdots \|$ = L^2 norm (l^2 norm) of a continuous function (vector)

Greek letters

γ = autocovariance function
γ_u = autocovariance function of the input signal
γ_{uy} = cross-covariance function between the input and the output signal
γ_y = autocovariance function of the output signal
τ = lag
τ_c = closed-loop dominant time constant
τ_p = process time constant
$\in$ = belongs to
θ = model parameters
σ, σ_θ = standard deviation
$\phi_{m,n}$ = scaling function at scale m and translation n
$\psi_{m,n}$ = wavelet at scale m and translation n
ω = frequency
ω_b = corner frequency or break frequency
ω_c = cutoff frequency
Ω = frequency band

Literature Cited

Anderson, H. W., K. H. Rasmussen, and S. B. Jorgensen, "Advances in Process Identification," *Chemical Process Control—CPC IV*, Y. Arkun and W. H. Ray, eds., AIChE, p. 237 (1991).

Astrom, K. J., and P. Eykhoff, "System Identification—A Survey," *Automatica*, **7**(2), 123 (1971).

Brosilow, C., and G. Q. Zhao, "A Linear Programming Approach to Constrained Multivariable Process Control," *Control Dyn. Syst.*, **27**(3), 141 (1988).

Chui, C. K., "Wavelets—With Emphasis on Spline-Wavelets and Applications to Signal Analysis," *Approximation Theory, Spline Functions and Applications*, S. P. Singh, ed., Kluwer, Dordrecht, The Netherlands, p. 19 (1992).

Cutler, C. R., and F. H. Yocum, "Experience with the DMC Inverse for Identification," *Chemical Process Control—CPC IV*, Y. Arkun and W. H. Ray, eds., AIChE, p. 297 (1991).

Daubechies, I., "The Wavelet Transform, Time-Frequency Localization and Signal Analysis," *IEEE Trans. Inform. Theory*, **IT-36**, 961 (1990).

Daubechies, I., *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, Philadelphia (1992).

Davies, W. D. T., *System Identification for Self-Adaptive Control*, Wiley, London (1970).

Downs, J. J., and E. F. Vogel, "Plant-Wide Industrial Process Control Problem," *Comput. Chem. Eng.*, **17**, 245 (1993).

Eykhoff, P., *System Identification—Parameter and State Estimation*, Wiley, London (1974).

Forbes, J. F., and T. E. Marlin, "Model Accuracy for Economic Optimizing Controllers—The Bias Update Case," *Ind. Eng. Chem. Res.*, **33**(8), 1919 (1994).

Garcia, C. E., B. L. Ramaker, and J. F. Pollard, "Total Process Control—Beyond the Design of Model Predictive Controllers," *Chemical Process Control—CPC IV*, Y. Arkun and W. H. Ray, eds., AIChE (1991).

Gevers, M., and L. Ljung, "Optimal Experiment Designs with Respect to the Intended Model Application," *Automatica*, **22**(5), 543 (1986).

Godfrey, K. R., "Dynamic Analysis of an Oil-Refinery Unit under Normal Operating Conditions," *Proc. IEE*, **6**(5), 879 (1969).

Goodwin, G. C., and Payne, R. L., *Dynamic System Identification—Experiment Design and Data Analysis*, Academic Press, New York (1977).

Goodwin, G. C., J. C. Murdoch, and R. L. Payne, "Optimal Test Signal Design for Linear S.I.S.O. System Identification," *Int. J. Control*, **17**(1), 45 (1973).

Grenander, U., *Abstract Inference*, Wiley, New York (1981).

Grossmann, A., "Wavelet Transform and Edge Detection," *Stochastic Processes in Physics and Engineering*, M. Hazewinkel, ed., Reidel, Dordrecht, The Netherlands (1986).
Harkins, B. L., "The DMC Controller," *ISA Proc.*, Instr. Soc. of Amer., p. 853 (1991).
Hope, G. S., and O. P. Malik, "Identification: Pseudo-Random Signal Method," in *Systems and Control Encyclopedia: Theory, Technology, Applications*, M. G. Singh, ed., Pergamon Press, Oxford (1987).
Isermann, R., "Practical Aspects of Process Identification," *Automatica*, **16**, 575 (1980).
Joseph, B., and R. L. Motard, eds., *Wavelet Applications in Chemical Engineering*, Kluwer, Boston (1994).
Ljung, L., *System Identification—Theory for the User*, Prentice-Hall, Englewood Cliffs, N.J. (1987).
Ljung, L., *System Identification Toolbox—User's Guide*, The MathWorks, Natick, MA (1991).
Luyben, W. L., *Process Modeling, Simulation and Control for Chemical Engineers*, McGraw-Hill, New York (1989).
Mallat, S., "Zero Crossings of a Wavelet Transform," *IEEE Trans. Inf. Theory*, **37**, 1019 (1991).
Mallat, S. and W. L. Hwang, "Singularity Detection and Processing with Wavelets," *IEEE Trans. Inform. Theory*, **IT-38**(2), 617 (1991).
McAvoy, T. J., and N. Ye, "Base Control for the Tennessee Eastman Problem," *Comput. Chem. Eng.*, **18**(5), 383 (1994).
Palavajjhala, S., R. L. Motard, and B. Joseph, "Process Identification Using Discrete Wavelet Transforms," in *Proc. IFAC Symp.—ADCHEM'94*, Kyoto, Japan, p. 477 (1994).
Palavajjhala, S., R. L. Motard, and B. Joseph, "Plant-Wide Control of the Tennesee Eastman Problem," AIChE Meeting, St. Louis (1993).
Pati, Y. C., "Wavelets and Time-Frequency Methods in Linear Systems and Neural Networks," PhD Thesis, Univ. of Maryland, College Park (1992).
Richalet, J., S. Abu el At-Doss, and A. Coic, "Global Identification and Optimal Input Design," *Proc. IFAC Symp. on Identification and System Parameter Estimation*, Budapest, p. 1121 (1991).
Rivera, D. E., "Monitoring Tools for PRBS Testing in Closed-Loop System Identification," AIChE Meeting, Miami Beach, paper 131d (1992).
Rivera, D. E., J. F. Pollard, and C. E. Garcia, "Control-Relevant Prefiltering: A Systematic Design Approach and Case Study," *IEEE Trans. Automat. Contr.*, **AC-37**(7), 964 (1992).
Rivera, D. E., M. Morari, and S. Skogestad, "Internal Model Control, 4. PID Controller Design," *Ind. Eng. Chem., Process. Des. Dev.*, **25**, 252 (1986).
Roberts, P. D., and R. H. Davis, "Statistical Properties of Smoothed Maximal-Length Linear Binary Sequences," *Proc. IEE*, **113**(1), 190 (1966).

Manuscript received Oct. 31, 1994, and revision received Apr. 17, 1995.