Simple Denoising Algorithm Using Wavelet Transform

Manojit Roy, V. Ravi Kumar, and B. D. Kulkarni
Chemical Engineering Div., National Chemical Laboratory, Pune 411 008, India

John Sanderson and Martin Rhodes
Dept. of Chemical Engineering, Monash University, Clayton, Victoria, 3168, Australia

Michel vander Stappen
Unilever Research, Vlaardingen, AC Vlaardingen, The Netherlands

Application of wavelets and multiresolution analysis to reaction engineering systems from the point of view of process monitoring, fault detection, systems analysis and so on, is an important topic and of current research interest (see, Bakshi and Stephanopoulos, 1994; Safavi et al., 1997; Luo et al., 1998; Carrier and Stephanopoulos, 1998). In the present article, we focus on one such important application, where we propose a new and simple algorithm for the reduction of noise from scalar time-series data. The presence of noise in a time-varying signal restricts one's ability to obtain meaningful information from the signal. Measurement of the correlation dimension can be affected by a noise level as small as 1% of signal, making estimation of invariant properties of a dynamical system, such as the dimension of the attractor and Lyapunov exponents, almost impossible (Kostelich and Yorke, 1988). Noise in experimental data can also cause misleading conclusions (Grassberger et al., 1991). Numerous articles on various techniques for noise reduction exist in the literature (Kostelich and Yorke, 1988; Härdle, 1990; Farmer and Sidorowich, 1991; Sauer, 1992; Cawley and Hsu, 1992; Cohen, 1995; Donoho and Johnstone, 1995; Kantz and Schreiber, 1997). For instance, the fast Fourier transform (FFT) reduces noise effectively in those cases where the frequency distribution of noise is known (Kostelich and Yorke, 1988; Cohen, 1995; Kantz and Schreiber, 1997); and singular-value analysis methods (Cawley and Hsu, 1992) project the original time series onto an optimal subspace, whereby noise components are left behind in the remaining orthogonal directions, and so forth. In the existing wavelet-based denoising methods (Donoho and Johnstone, 1995), two types of denoising are introduced: linear denoising and nonlinear denoising. In linear denoising, noise is assumed to be concentrated only on the fine scales and that all the wavelet coefficients below these scales are cut off. Nonlinear denoising, on the other hand, treats noise reduction by either cutting off all coefficients below a certain threshold (so-called ''hard-thresholding''), or reducing all coefficients by this threshold (so-called ''soft-thresholding''). The threshold values are obtained by statistical calculations and have been seen to depend on the standard deviation of the noise (Nason, 1994).

The noise-reduction algorithm that we propose here makes use of the wavelet transform (WT), which in many ways complements the well-known Fourier-transform (FT) procedure. We apply our method first to three model flow systems, viz., the Lorenz, Autocatalator, and Rössler systems, all of which exhibit chaotic dynamics. The reasons for choosing these systems are the following: first, all of them are simplified models of well-studied experimental systems. For instance, Lorenz is a simple realization of convective systems (Lorenz, 1963), while the Autocatalator and Rössler systems have more complicated analogs in chemical multicomponent reactions (Rössler, 1976; Lynch, 1992). Second, chaotic dynamics is extremely nonlinear, highly sensitive, possesses only short-time correlations, and is associated with a broad range of frequencies (Guckenheimer and Holmes, 1983; Strogatz, 1994). Because of these properties it is well known that FT methods are not applicable in a straightforward way to chaotic dynamical systems (Abarbanel, 1993). On the other hand, WT methods are particularly suited to handle not only nonlinear but also nonstationary signals (Strang and Nguyen, 1996). This is because the properties of the data are studied at varying scales with superior time localization analysis when compared to the FT technique. Our noise-reduction algorithm is advantageous because, as will be shown, the threshold level for noise is identified automatically. In this study, we have used the discrete analog of the wavelet transform (DWT), which involves transforming a given signal with orthogonal wavelet basis functions by dilating and translating in discrete steps (Daubechies, 1990; Holschneider, 1995), although, other basis functions may be judiciously chosen. For study purposes we corrupt one variable $x(t)$ for each of these systems with noise of zero mean, and then apply our algorithm for denoising. We analyze the performance of this method in all three

Correspondence concerning this article should be addressed to V. Ravi Kumar.

systems for a wide range of noise strengths, and show its effectiveness. Note that we then validate the applicability of the method to experimental data obtained from two chemical systems. In one system, the time series data was obtained from pressure-fluctuation measurements of the hydrodynamics in a fluidized bed. In the other, the conductivity measurements in a liquid surfactant manufacturing experiment were analyzed.

Methodology

The noise-reduction algorithm based on DWT consists of the following five steps.

Step 1. In the first step, we differentiate the noisy signal $x(t)$ to obtain the data $x_d(t)$, using the central finite differences method with fourth-order correction to minimize error (Constantinides, 1987), that is,

$$x_d(t) = \frac{dx(t)}{dt}. \tag{1}$$

Step 2. We then take the DWT of the data $x_d(t)$ and obtain wavelet coefficients $W_{j,k}$ at various *dyadic scales* j and displacements k. A dyadic scale is the scale whose numerical magnitude is equal to 2 raised to an integer exponent, and is labeled by the exponent. Thus, the dyadic scale j refers to a scale of size 2^j. In other words, it indicates a resolution of 2^j data points. Thus a low value of j implies a finer resolution, while high j analyzes the signal at a larger resolution. This transform is the discrete analog of continuous WT (Holschneider, 1995), and is given by the formula

$$W_{j,k} = \int_{-\infty}^{+\infty} x_d(t)\,\psi_{j,k}(t)\,dt, \tag{2}$$

with

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k),$$

where j, k are integers. As for the wavelet function $\psi(t)$, we have chosen Daubechies' compactly supported orthogonal function with four filter coefficients (Daubechies, 1990; Press et al., 1996).

Step 3. In this step we estimate the *power* P_j contained in different dyadic scales j, via

$$P_j(x) = \sum_{k=-\infty}^{+\infty} |W_{j,k}|^2 \qquad (j = 1, 2, \ldots). \tag{3}$$

By plotting the variation of P_j with j, we see that it is possible to identify a scale j_m at which the power due to noise falls off rapidly. This is important because, as we shall see from the studies on the case examples, it provides an automated means of detecting the threshold. Identification of the scale j_m at which the power due to noise shows the first minimum allows us to reset all $W_{j,k}$ up to scale index j_m to zero, that is, $W_{j,k} = 0$, for $j = 1, 2, \ldots, j_m$.

Step 4. In the fourth step, we reconstruct the denoised data $\hat{x}_d(t)$ by taking the inverse transform of the coefficients $W_{j,k}$:

$$\hat{x}_d(t) = c_\psi \sum_{j=0}^{\infty} \sum_{k=-\infty}^{\infty} W_{j,k}\,\psi_{j,k}(t), \tag{4}$$

where c_ψ is normalization constant given by

$$c_\psi = 1 \Big/ \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|^2}{\omega}\,d\omega < \infty,$$

with $\hat{\psi}(\omega)$ as the FT of the wavelet function $\psi(t)$.

Step 5. In the fifth and final step $\hat{x}_d(t)$ is integrated to yield the cleansed signal $\hat{x}(t)$:

$$\hat{x}(t) = \int \hat{x}_d(t)\,dt. \tag{5}$$

A commutativity property exists between the operation of differentiation/integration and wavelet transform. Therefore, first differentiating the signal and then taking DWT is equivalent to carrying out the two operations in reverse order. This implies that the same result can be obtained by switching the order between the first and second steps, and then between the fourth and fifth steps.

The effectiveness of this method lies in the following observations. Upon differentiation, contribution due to white noise moves toward the finer scales because the process of differentiation converts the uncorrelated stochastic process to a first-order moving average process and thereby distributes more energy to the finer scales. That the differentiation of white noise brings about this behavior is known in the Fourier spectrum (Box et al., 1994). It should be noted that the nature and effectiveness of separation depend on the wavelet basis function chosen and also on the properties of the derivative of the WT, which is in itself a highly interesting and not fully understood subject (Strang and Ngyuyen, 1996). For the signals studied in this article—model as well as experimental—the wanted signal features lie in the larger wavelet scales, while the unwanted signal features after differentiation lie in the finer resolution wavelet scales. This is because the size of the data set handled determines the total number of scales available, and a suitable choice can bring out the noisy signal WT features lying in the larger scales. This justifies the assumption that fine-scale features can be removed by setting the corresponding wavelet coefficients to zero, and large-scale features retained after differentiation. For this reason we also see a clear separation in the scales attributed to noise and those for the signal. A threshold scale for noise removal is thus identified, and this leads to automation of noise removal.

Results and Discussion

First we take up the three model systems and discuss the observations. For test purposes the pure signal obtained from these systems was corrupted with noise of a certain strength. For the systems chosen for study, namely, Lorenz, Autocatalator, and Rössler, Table 1 summarizes the details, that is, the equations governing their dynamics, the values chosen for

Table 1. Model Systems Studied: Parameter Values and Nature of Dynamics

Systems	Lorenz	Autocatalator	Rössler
Dynamical equations	$dx/dt = -\sigma x + \sigma y$ $dy/dt = Rx - y - xz$ $dz/dt = -bz + xy$	$dx/dt = 1 - x - Da_1 xz^2$ $dy/dt = \beta - y - Da_2 yz^2$ $dz/dt = 1 - (1 + Da_3)z + \alpha(Da_1 x + Da_2 y)z^2$	$dx/dt = -y - z$ $dy/dt = x + ay$ $dz/dt = b + z(x - c)$
Parameters chosen	$\sigma = 10,\ R = 28$ $b = 8/3$	$Da_1 = 18000,\ Da_2 = 400$ $Da_3 = 80,\ \beta = 2.93$ $\alpha = 1.5$	$a = 0.398,\ b = 2$ $c = 4$
Dynamics	Chaotic	Chaotic	Chaotic

the parameters, and the nature of the evolution of these systems, of these sets of parameter values. These values were chosen appropriately so that the dynamics is chaotic. In our initial studies, purely for testing purposes, we studied situations where we ensure that all scales are affected by noise. In the wavelet domain this may be conveniently carried out by perturbing the wavelet coefficients in the following way. The differential equations are first numerically integrated to obtain a pure signal $x^0(t_i)$ at equidistant time steps t_i. We then take the DWT of the signal, and add white noise η of zero mean and certain strength, that is, $W_{j,k} = W^0_{j,k} + \eta$ where $W^0_{j,k}$, $W_{j,k}$ are the wavelet coefficients of the pure and noisy signal, respectively. We take the strength of noise as the relative percentage of the difference between the maximum and minimum of the signal value. Since each coefficient $W'_{j,k}$ is individually affected by the noise, this procedure ensures equal weight for the presence of noise at all scales. Reconstructing the time series signal with this perturbed set of wavelet coefficients gave us the noisy signal to be cleaned. Our studies following this routine did show that noise and signal separation was achieved. For the subsequent studies we followed the usual way of corrupting the signal by additive noise, that is,

$$x(t_i) = x^0(t_i) + \eta(t_i), \tag{6}$$

where $\eta(t_i) \in [-0.5, 0.5]$ is the noise with zero mean and uniform distribution, and J the number of available dyadic scales. We have taken a data size of 16,384 $(= 2^{14})$ points for all three of these systems, and so $J = 14$.

We plot our observations for the Lorenz system in Figure 1. Figure 1a shows the power at different scales in the pure signal $x^0(t_i)$. In Figure 1b we plot the scalewise power distribution after numerically differentiating the pure signal. We see that almost the entire power of the differentiated data is accumulated within the dyadic scales 4 and 9 (the signal power between scales 10 and 14 has disappeared by the process of differentiation). Figure 1c plots the scalewise power in the noisy signal $x(t_i)$ when the pure signal is infected with noise $\eta(t_i)$ of a typical strength of 5% of the signal (that is, 5% of the difference in the maximum and minimum values of $x^0(t_i)$). Because of the relatively larger contribution of the pure signal at all scales, compared to that from noise, it is impossible to distinguish between the two components, and the figure looks qualitatively very similar to Figure 1a. However, when we plot the scalewise power distribution to the differentiated noisy data in Figure 1d, the signal contribution can easily be identified and also compared with the plot in Figure 1b. It is to be noted that the difference in the values of the two peaks in Figure 1b and 1d is due to the power being normalized by the respective total signal power. The contribution due to noise shows up in the finer scales. A clear minimum with close to zero value separates out two distinct regions. Figure 1e exhibits a small segment of the signal after the noise has been successfully removed following the procedure just outlined. All the three signals—pure, noisy, and cleaned—are overlaid for the sake of comparison.

In Figure 2 we show the results for the Autocatalator and the Rössler reacting systems, Figures 2a and 2c plot, respectively, the scalewise power distribution for noise-infected signals obtained from the two systems. As in the case of the Lorenz system, it also is evident that here one cannot distinguish the noise and signal components. Figure 2b and 2d exhibit a scalewise power profile for the differentiated data of the two signals, respectively. The clear separation is again obvious.

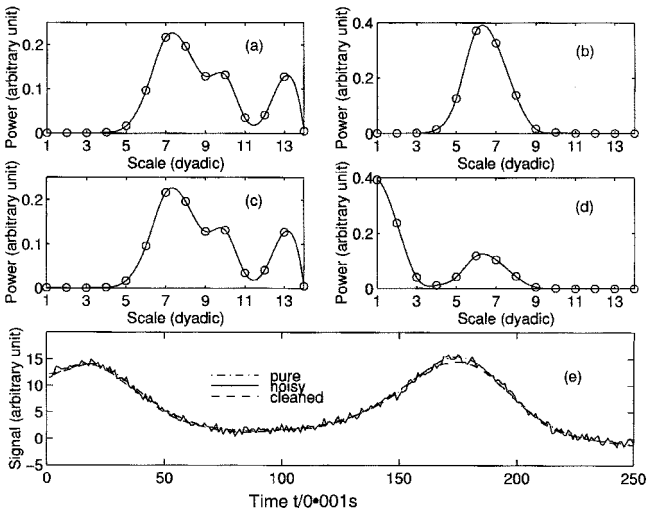

Figure 1. Lorenz system with parameter values as stated in Table 1; 16,384 $(= 2^{14})$ data points are considered.

Scalewise power distribution is plotted against the dyadic scale: (a) for the pure signal (the interpolated line through the data points is drawn for visualization); (b) using the data obtained after differentiating the signal; (c) for the signal corrupted by noise; and (d) using differentiated data of noisy signal. A segment of cleansed signal is shown in part (e), along with the pure and noisy signals overlaid for comparison.

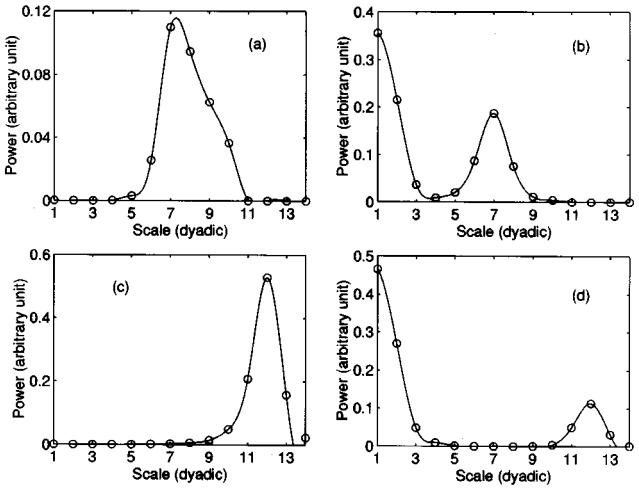

Figure 2. Autocatalator and Rössler systems for parameter values as in Table 1.

Scalewise power profile plotted: (a) for the noisy autocatalytic signal; (b) using data after differentiating the signal; (c) for the noisy Rössler signal; and (d) using the differentiated data of noisy signal.

In order to quantitatively estimate the efficiency of our denoising method, we have made the following error estimation (Kostelich and Schreiber, 1993) for the preceding three model systems. Since the pure signal is known in all three cases, a measure of the amount of error present in the cleaned data is obtained by taking the rms deviation of the cleaned signal $\hat{x}(t_i)$ from the pure signal $x^0(t_i)$ as follows:

$$\hat{E} = \left\{ \frac{1}{N} \sum_{i=1}^{N} \left[\hat{x}(t_i) - x^0(t_i) \right]^2 \right\}^{1/2}, \qquad (7)$$

where N is the length of the time series. A similar quantity E for the noisy data $x(t_i)$ is also computed. The condition $\hat{E}/E < 1$ guarantees that noise has been successfully reduced. The error estimator $\hat{E}/E$ is a natural measure for noise reduction when the original dynamics is known (Kostelich and Schreiber, 1993). In Figure 3 we plot $\hat{E}/E$ against noise strength for all three model systems. We see that for the entire range of noise values, and even with a noise level as high as 10% of the signal exhibiting chaotic dynamics, $\hat{E}/E$ remains appreciably below unity. Thus the plot demonstrates the efficiency of the approach. Different wavelet basis functions may change the nature and also improve the efficiency further.

We now discuss our method when applied to raw data obtained from two real chemical systems. In the first system, the time series data were obtained from the measurements of the pressure fluctuations in a fluidized bed, which consists of a vertical chamber inside of which a bed of solid particles is supported by an upwardly moving gas. Our system used a bed of silica sand particles (with a mean diameter of 200 microns) with a settled height of 500 mm, fluidized by ambient air in a transparent vessel 430 mm across and 15 mm wide. Beyond a critical inlet gas velocity, namely, the minimum bubbling velocity, the gas passes through the bed in the form of bubbles, thereby churning the solid and gas mixture in a turbulent manner. The time-series data have been taken by using a pressure transducer attached to a probe inserted into the fluid bed to measure the pressure fluctuations inside this mixture, relative to atmospheric pressure. The bed was operated at an inlet gas velocity of 0.85 m/s, and the pressure fluctuations were recorded at a sampling rate of 333 Hz (333 data points per second). As a standard procedure, we normalize the data by subtracting the mean and dividing by the standard deviation (Constantinides, 1987; Bai et al., 1997). In Figure 4 we show the results obtained after the data have been subjected to denoising. Figure 4a shows the power distribution at different scales in the original experimental signal, while in Figure 4b we plot the scalewise power profile of the differentiated data. Again, one clearly sees the two distinct contributions of the noise and signal components. Figure 4c shows short segments of the denoised signal and the original signal, which is overlaid for comparison. The cleaned signal is seen to be smooth, indicating that the noise has been removed.

In the second chemical system, the time-series data were obtained by sampling a measure related to the conductivity in a 3-L liquid surfactant manufacturing experiment, at a sampling rate of 500 Hz. The time series is highly nonstationary, since at various stages the operational parameters are altered (increasing the temperature for a certain period, then adding actives to the liquid, etc.). We studied unfiltered noisy data sets from the experiments, in order to check whether our method can filter the noise out and also bring out some intrinsic features of the system. We used our denoising algorithm to treat this data set in a slightly different way. The aim was to remove the finer scales from the differentiated data one by one, starting from the lowest (dyadic scale 1) and gradually going up, so that at each stage (after integrating the data) the observable frequencies in the filtered signal can be

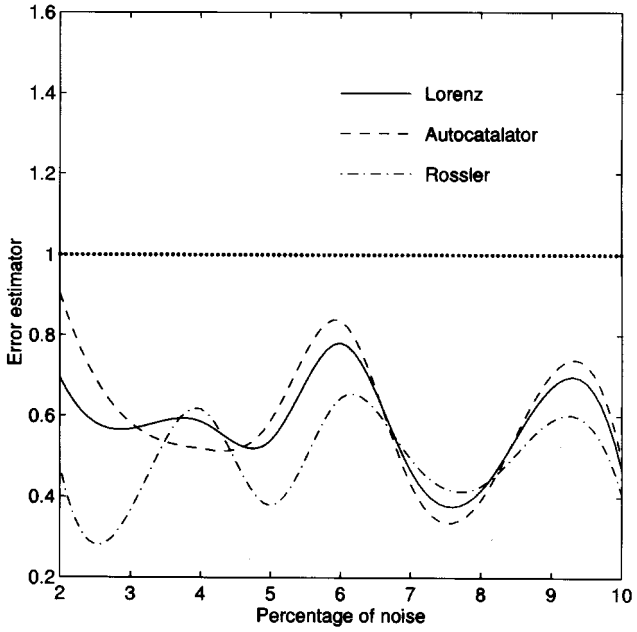

Figure 3. Error estimator $\hat{E}/E$ plotted against the noise strength for all three systems.

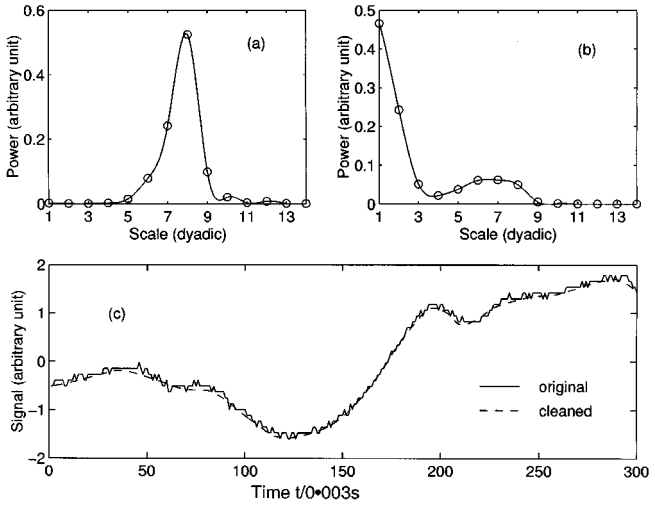

Figure 4. Fluidized-bed experiment.

(a) The scalewise power profile is shown for (a) the experimental signal, and (b) the data after the signal has been numerically differentiated. A small segment of the cleansed signal is shown in (c), along with the original signal for comparison.

related to identifiable physical sources. Figure 5a shows a small segment (1 s long) of the noisy data. In Figure 5b we plot, on the same scale as in the earlier figure, the filtered data, using our method to remove the lowest dyadic scale 1. One can now clearly identify a 50-Hz component, due to the signal from the electrical power supply (the ''net frequency''). By removing scale 2 along with scale 1, the net frequency goes away, and the filtered data exhibit a 13-Hz component superimposed with occasional spikes. This 13-Hz signal shows up clearly in the filtered data with scale 3 also removed. Figure 5b also shows these data overlaid on the data with scale 2 removed. The 13 Hz may have been caused by the stirring device, which has two blades and revolves with 260 rpm, which

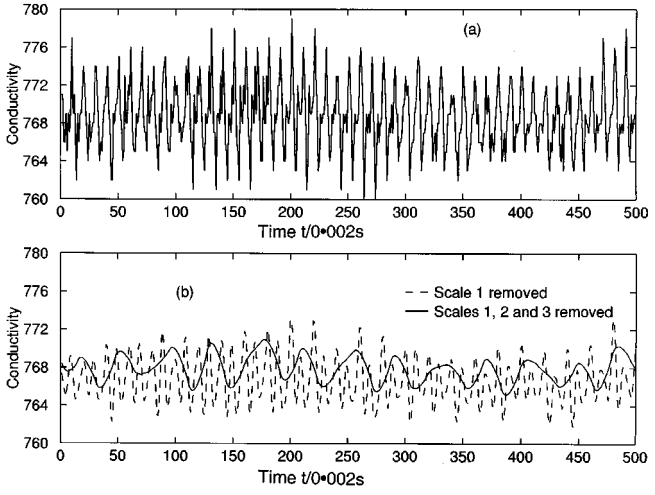

Figure 5. Liquid-surfactant experiment.

(a) A segment of the original noisy data, 1 s long; (b) the filtered data, with scale 2 removed, and with scales 2, 3, and 4 removed (on the same axes—scales as part (a)).

corresponds to approximately 10 Hz. The electronic signal had an antialiasing feature of no more than 250 Hz, and therefore aliasing (beating) can be ruled out. It may also be mentioned here that the Fourier power spectrum of the denoised signal shows a spike at 50 Hz frequency, whose removal resulted in a residual spectrum consisting mainly of a background continuum without any appreciable peak around 13 Hz. With the present example, this study suggests that the wavelet transform methodology offers considerable benefits in the recovery of intrinsic signal components.

Summary

We have presented a new and alternative algorithm for noise reduction using a discrete wavelet transform. We believe that our algorithm will be beneficial in various noise-reduction applications, and that it shows promise in developing techniques that can resolve an observed signal into its various intrinsic components. In our method the threshold for reducing noise comes out automatically. The algorithm has been applied to three model flow systems—the Lorenz, Autocatalator, and Rössler systems—all evolving chaotically. The method is seen to work quite well for a wide range of noise strengths, even as large as 10% of the signal level. We have also applied the method successfully to noisy time series data obtained from the measurement of pressure fluctuations in a fluidized bed, and also to that obtained by conductivity measurement in a liquid-surfactant experiment. In all the illustrations we have been able to observe that there is a clean separation in the frequencies covered by the differentiated signal and white noise. However, if the noise is colored, a certain degree of overlap between the signal and noise may exist, even after differentiation. The method needs to be improved upon for this complex situation.

Acknowledgment

The authors acknowledge Unilever Research, Port Sunlight, for financial and other assistance. Part of the work has been carried out under the aegis of Indo-Australian S&T program DST/INT/AUS/I-94/97.

Literature Cited

Abarbanel, H. D. I., ''The Observance of Chaotic Data in Physical Systems,'' *Rev. Mod. Phys.*, **65**, 1340 (1993).

Bai, D., T. Bi, and J. R. Grace, ''Chaotic Behavior of Fluidized Beds Based on Pressure and Voidage Fluctuations,'' *AIChE J.*, **43**, 1357 (1997).

Bakshi, B., and G. Stephanopoulos, ''Representation of Process Trends. Part IV. Induction of Real-Time Patterns from Operating Data for Diagnoses and Supervisory Control,'' *Comput. Chem. Eng.*, **18**, 303 (1994).

Box, G. E. P., G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis Forecasting and Control*, Prentice Hall, Englewood Cliffs, NJ (1994).

Carrier, J. F., and G. Stephanopoulos, ''Wavelet-Based Modulation in Control-Relevant Process Identification,'' *AIChE J.*, **44**, 341 (1998).

Cawley, R., and G.-H. Hsu, ''Local-Geometric-Projection Method for Noise Reduction in Chaotic Maps and Flows,'' *Phys. Rev. A*, **46**, 3057 (1992).

Cohen, L., *Time Frequency Analysis*, Prentice Hall, Englewood Cliffs, NJ (1995).

Constantinides, A., *Applied Numerical Methods with Personal Computers*, McGraw-Hill, New York (1987).

Daubechies, I., *Ten Lectures on Wavelets*, SIAM, Philadelphia (1990).
Donoho, D. L., and I. M. Johnstone, ''Adapting to Unknown Smoothness via Wavelet Shrinkage,'' *J. Am. Stat. Assoc.*, **90**, 1200 (1995).
Farmer, J. D., and J. J. Sidorowich, ''Optimal Shadowing and Noise Reduction,'' *Phys. D.*, **47**, 373 (1991).
Grassberger, P., T. Schreiber, and C. Schaffrath, ''Non-Linear Time Sequence Analysis,'' *Int. J. Bifurcation Chaos*, **1**, 521 (1991).
Guckenheimer, E., and P. Holmes, *Nonlinear Oscillation, Dynamical Systems and Bifurcations of Vector Fields*, Springer-Verlag, Berlin (1983).
Härdle, W., ''Applied Nonparametric Regression,'' Econometric Society Monographs, Cambridge Univ. Press, Cambridge (1990).
Holschneider, M., *Wavelets: An Analysis Tool*, Clarendon Press, Oxford (1995).
Kantz, H., and T. Schreiber, *Nonlinear Time Series Analysis*, Cambridge Univ. Press, Cambridge (1997).
Kostelich, E. J., and J. A. Yorke, ''Noise Reduction in Dynamical Systems,'' *Phys. Rev. A.*, **38**, 1649 (1988).
Lorenz, E., ''Deterministic Nonperiodic Flow,'' *J. Atmos. Sci.*, **20**, 130 (1963).
Luo, R., M. Misra, S. J. Qin, R. Barton, and D. M. Himmelblau, ''Sensor Fault Detection via Multiscale Analysis and Nonparametric Statistical Inference,'' *Ind. Eng. Chem. Res.*, **37**, 1024 (1998).
Lynch, D. T., ''Chaotic Behavior of Reaction Systems: Mixed Cubic and Quadratic Autocatalators,'' *Chem. Eng. Sci.*, **47**, 4435 (1992).
Nason, G. P., ''Wavelet Regression by Cross-Validation,'' Dept. of Mathematics, Univ. of Bristol, Bristol, England (1994).
Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes*, Cambridge Univ. Press, Cambridge (1987).
Rössler, O. E., ''Chaotic Behavior in Simple Reaction Systems,'' *Z. Naturforsch*, **31a**, 259 (1976).
Safavi, A. A., J. Chen, and J. A. Romagnoli, ''Wavelet-Based Density Estimation and Application to Process Monitoring,'' *AIChE J.*, **43**, 1227 (1997).
Sauer, T., ''A Noise Reduction Method for Signals from Nonlinear Systems,'' *Phys. D.*,. **58**, 193 (1992).
Strang, G., and T. Nguyen, *Wavelets and Filter Banks*, Wellesley–Cambridge Press, Wellesley, MA (1996).
Strogatz, S. H., *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*, Addison-Wesley, Reading, MA (1994).

Manuscript received Dec. 4, 1998, and revision received Apr. 5, 1999.