Process Control and Quality, 4 (1992) 77–86
Elsevier Science Publishers B.V., Amsterdam

Identification of finite impulse response models by principal components regression: Frequency-response properties

B.M. Wise [1] and N.L. Ricker

Center for Process Analytical Chemistry and Department of Chemical Engineering BF-10, University of Washington, Seattle, WA 98195 (USA)

(Received October 24, 1991; accepted in revised form April 9, 1992)

Abstract

Principal components regression (PCR) is used to identify finite impulse response (FIR) dynamic models. It is demonstrated that the PCR method has a predictable influence on the frequency-response of the resulting models. If the process input is white noise passed through a first-order filter, its principal-component decomposition is a set of sinusoidal factors at specific frequencies. The first principal component (PC) is associated with the lowest-frequency, maximum-power sinusoid. Succeeding PCs are associated with increasingly higher frequencies having progressively less power. The decomposition of more complex input signals produces similar results. Models retaining a relatively small number of PCs fit the true process characteristics in a limited frequency range (where the input power is concentrated). Adding PCs to the regression broadens the model's bandwidth.

Keywords: Principal components regression; Identification; Dynamic model; Finite impulse response

INTRODUCTION

Finite impulse response (FIR) models, and the closely related step–response models, form the basis of many model-based control schemes [1]. There is, however, relatively little information in the literature on the identification of FIR models. The definitive work by Ljung [2] emphasizes identification of transfer-function and state–space models. Papers specifically concerned with FIR modeling include the work of Ricker [3], who used partial least squares (PLS) and a method based on the singular value decomposition (SVD). Rivera et al. [4,5] also studied the use

[1] Present address with Battelle Pacific Northwest Laboratories, P.O. Box 999, Richland, WA 99352.
Correspondence to: Prof. N.L. Ricker, Center for Process Analytical Chemistry and Department of Chemical Engineering BF-10, University of Washington, Seattle, WA 98195 (USA). Fax: (206) 543-3778, E-mail: ricker @ cheme.washington.edu.

of PLS. Cutler and Yocum [6] discuss identification of step–response models of complex chemical processes, but omitted details of the numerical methods.

It is possible to use multiple linear regression (MLR) to obtain FIR coefficients. The resulting models are usually unsatisfactory, however. Many parameters must be estimated, making the problem ill-conditioned. The parameter estimates are too easily corrupted by noise in the data. The use of a low-order transfer-function or state–space model avoids this problem (see, e.g., Ljung [2]). The disadvantage is that the results can be sensitive to the choice of model structure. Another possibility is the use of orthogonal polynomials, such as Laguerre networks [7,8]. This is conceptually similar to the principal components regression (PCR) approach studied here.

In the sequel, we show that the singular-value and principal-components decomposition filter the input signal, affecting its frequency content. This allows us to predict certain qualitative properties of PCR; we demonstrate that they occur in practice. This paper is an outgrowth of previous work on the use of PCR, PLS and related techniques (see for example Wise [9], Wise and Ricker [10–12] and Wise et al. [13]).

BACKGROUND: FIR MODELS AND PCR

Finite impulse–response models are easily extended to the multiple input single output (MISO) case, but we will restrict our attention to SISO systems. In a discrete, SISO, FIR model, the output of a dynamic system is a linear combination of n past values of the input:

$$y(t) = b_1 u(t-1) + b_2 u(t-2) + \cdots + b_n u(t-n) \quad (1)$$

Here, y is the output, u is the input, and b_1 to b_n are model parameters (constants). We assume that the process signals are sampled at a constant frequency; the notation $u(0)$, $u(1), \ldots, u(t)$ implies a sequence of such samples. The process should be asymptotically stable with a settling time approximately equal to n sampling periods.

The PCR method is outlined in several articles and texts such as Geladi and Kowalski [14], Lorder et al. [15] and Naes and Martens [16]. There are several points concerning PCR that must be emphasized here. The first concerns the relationship between principal components (PCs), the singular value decomposition (SVD) and eigenvectors. For a matrix $\mathbf{X}$, the PCs of $\mathbf{X}$, the right singular vectors of $\mathbf{X}$ and the eigenvectors of $\mathbf{X}'\mathbf{X}$ are all the same. This fact will be used in the development below.

The critical decision in PCR is the number of principal components to retain when building the regression model. This is analogous to the choice of the number of orthogonal polynomials in a Laguerre network representation [8]. Cross-validation usually allows one to eliminate some of the PCs. This, however, causes certain anomalies in the resulting model. The nature of these anomalies is a major focus of the paper.

THE INPUT AUTOCORRELATION MATRIX

When data are arranged for FIR modeling, the resulting matrices resemble those shown in eq. (2) below. Here, the first four samples are shown for identification of a five-coefficient SISO, FIR model (compare with eq. (1)). Note how the values of the measured input, $u(t)$, in the $\mathbf{X}$ matrix are repeated along diagonals.

$$\underset{\mathbf{X}}{\begin{bmatrix} u(5) & u(4) & u(3) & u(2) & u(1) \\ u(6) & u(5) & u(4) & u(3) & u(2) \\ u(7) & u(6) & u(5) & u(4) & u(3) \\ u(8) & u(7) & u(6) & u(5) & u(4) \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}} \times \underset{\mathbf{y}}{\begin{bmatrix} y(6) \\ y(7) \\ y(8) \\ y(9) \\ \vdots \end{bmatrix}} \tag{2}$$

As pointed out in Box and Jenkins ([17], p. 53), if $u(t)$ is a random sequence with zero mean and unit variance, the correlation matrix, $\mathbf{X}'\mathbf{X}/(n-1)$, approaches the autocorrelation matrix (ACM) as more data are collected. Each entry in the autocorrelation matrix, a_{ij}, is equal to the correlation coefficient between $u(t+j)$ and $u(t+i)$. Note, however, that the correlation between any $u(t+j)$ and $u(t+i)$ depends only on the difference between i and j. This can be compared to the autocovariance functin (ACF) (see, for example, Box and Jenkins [17]), which is a vector of the covariance between a (mean centered) signal u at time t and time $t-\tau$. Thus:

$$ACF(\tau) = E\{u(t)u(t-\tau)\} \tag{3}$$

where $E\{\cdot\}$ denotes the expectation operator. If u is scaled to unit variance, the ACF becomes the autocorrelation function. Because every value in the ACM depends only upon the difference of the indices, every entry a_{ij} in the ACM can be replaced by $ACF(i-j)$. This is shown in eq. (4), where each value in the ACM is replaced by the corresponding value in the autocorrelation function (ACF).

$$\underset{\mathbf{x'x}}{\begin{bmatrix} ACF(0) & ACF(1) & ACF(2) & ACF(3) & ACF(4) \\ ACF(-1) & ACF(0) & ACF(1) & ACF(2) & ACF(3) \\ ACF(-2) & ACF(-1) & ACF(0) & ACF(1) & ACF(2) \\ ACF(-3) & ACF(-2) & ACF(-1) & ACF(0) & ACF(1) \\ ACF(-4) & ACF(-3) & ACF(-2) & ACF(-1) & ACF(0) \end{bmatrix}} \tag{4}$$

The ACM is a symmetric Toeplitz matrix. Furthermore, if the characteristics of the input signal, u, are known, the expected value of the ACM can be calculated. For instance, if $u(k)$ is generated by passing a white noise signal, $e(k)$, through a first order filter

$$u(k) = \alpha u(k-1) + (1-\alpha)e(k) \tag{5}$$

then the expected values of the entries in the autocorrelation matrix are

$$ACM(i,j) = a_{ij} = (\alpha)^{|i-j|} \tag{6}$$

EIGENVECTOR DECOMPOSITION OF THE *ACM*

The basis of the PCR method is an eigenvector decomposition of the covariance matrix, $\mathbf{X}'\mathbf{X}/(n-1)$. For FIR models, this matrix is the autocovariance matrix (ACM) of the input signal, $u(t)$. Wise and Ricker [11] observed that eigenvectors (PCs) of such a matrix have coefficients that plot as sine and cosine functions. For example, Fig. 1 shows the coefficients of the first five eigenvectors of the theoretical ACM for a $u(k)$ sequence generated by eq. (5) with $\alpha = 0.8$. This ACM was calculated for $|i-j|$ up to ± 100 resulting in an ACM matrix that is 101 by 101.

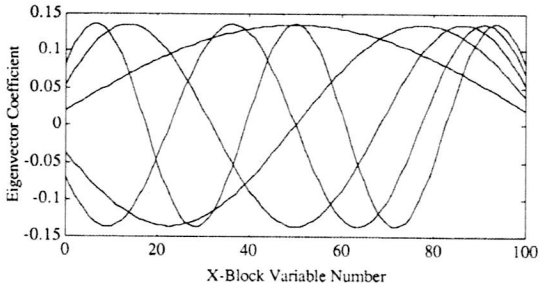

Fig. 1. Coefficients in first five eigenvectors of autocorrelation matrix of white noise process through first order filter.

There would be 101 coefficients in the FIR model corresponding to this *ACM*. As explained above, the loadings vectors from PCA are identical to these eigenvectors. For the corresponding continuous but finite case, the eigenvector coefficients are identically sine and cosine functions (see Appendix).

The theoretical results for the continuous case can be compared to the above calculations for the discrete case. To define the analogous continuous time problem, the integration limits must be specified. We require that

$$ACM(1, m) = \alpha^{m-1} = e^{-2a} \tag{7}$$

which when rearranged yields

$$a = \frac{1}{2}\log\left(\frac{1}{\alpha^{m-1}}\right) \tag{8}$$

For our example this gives $a = 5.268$. This makes the values in the continuous function map onto the same values in the discrete case. Equations (A5) and (A7) in the Appendix can now be used to determine the periods of the cosine and sine eigenvectors, respectively.

Figure 2 demonstrates the agreement between the continuous and discrete cases. The coefficients of the first three cosine eigenvectors of the discrete problem (eigenvectors 1, 3 and 5) are the solid lines, and the predicted values based on the continuous problem are the plus symbols. The coefficients are plotted against the position in the x-domain (see the Appendix). The coefficients of the continuous problem were scaled to yield unit vectors. The agreement is nearly perfect. Small discrepancies are confined to the higher frequency terms.

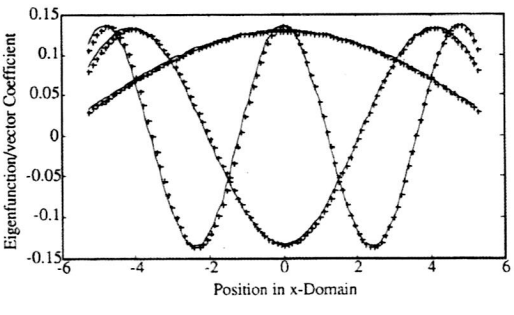

Fig. 2. Coefficients of the cosine eigenvectors of the discrete matrix (solid lines) shown with eigenfunction solutions to the continuous problem (+).

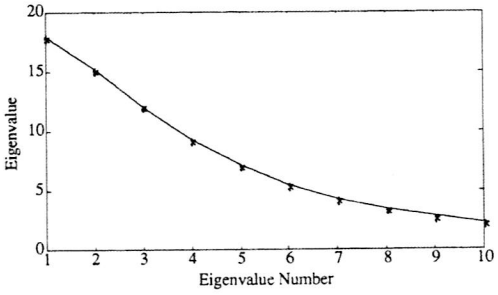

Fig. 3. Eigenvalues of the discrete case (solid line) shown with scaled eigenvalues of the continuous case (*).

It is possible to check the agreement of the corresponding eigenvalues. Because the scaling and size of the *ACM* affects the magnitude of the eigenvalues, it is not possible to compare the values directly. Figure 3 shows, however, that the distribution of the two sets of eigenvalues is the same. The solid line reprents the calculated eigenvalues of the discrete problem. The stars are the scaled eigenvalues of the continuous problem. The eigenvalues of the continuous problem were scaled to equate the first eigenvalues in the two distributions. The agreement is good; there are slight discrepancies for the smaller eigenvalues.

Numerical simulations show that the agreement between eigenvector decompositions of ideal discrete *ACM*s and the continuous analog is very good, even as the matrices become quite small. The agreement between cases suffers as the α parameter in the discrete case is decreased. This happens because for small α the *ACM* loses its smoothness, i.e., the differences between ad-

jacent entries in the matrix is large. A decrease in α is equivalent to an increase in the sampling frequency.

In practice, when identifying FIR models the input covariance matrix will not be ideal even when the input is generated by white noise through a first order filter, i.e., the matrix will not be perfectly Toeplitz due to the finite data record. However, for typical cases where the number of coefficients is > 10 and there are several hundred samples, the agreement between calculated principal components and pure sines and cosines is very good. Other *ACM* forms arrived at through higher order filters produce similar, though not identical, results. Certain cases, such as white noise through a second order under-damped filter, produce eigenvectors that still have periodic behavior, but are essentially combinations of frequencies and can be rather complex.

FREQUENCY DOMAIN EFFECTS OF PCR

The results of the previous section show that PCR decomposes the input signal into components with differing frequency content. This has a direct effect on the models obtained from the technique. Two identification experiments are used here to illustrate this effect. In the first case the true system is first order, while in the second case the true system is second order (under-damped). Both systems have unity gain. A pseudo-random binary sequence (PRBS) input signal of desired frequency content was used in both experiments. The PRBS was generated by filtering white noise through a second order Butterworth filter, then taking the sign of the result. Thus, the input signal was 1 when the output from the filter was positive and -1 when it was negative. The transfer function form of the process models and filter parameters are given in Table 1. The processes and filter are of the form

TABLE 1

Model parameters for case studies

	Numerator, $B(q^{-1})$	Denominator, $A(q^{-1})$
Case 1	$0.1426q^{-1}$	$1-0.8574q^{-1}$
Case 2	$0.1129q^{-1}+0.1038q^{-2}$	$1-1.5622q^{-1}+0.7788q^{-2}$
Filter	$0.0015+0.0029q^{-1}$ $+0.0015q^{-2}$	$1-1.8890q^{-1}+0.8949q^{-2}$

$$y(k)=a_1y(k-1)+\cdots+a_my(k-m) + b_1u(k-1)+\cdots+b_n(k-n) \quad (9)$$

This is conveniently expressed in terms of the backwards shift operator, q^{-1}, (see for example Åstrom and Wittenmark [18]) defined as

$$q^{-1}x(k)=x(k-1) \quad (10)$$

The individual terms in eq. 9 can be grouped to form two polynomials in q^{-1}:

$$A(q^{-1})y(k)=B(q^{-1})u(k) \quad (11)$$

where, e.g., $A=1-a_1q^{-1}-\cdots-a_mq^m$. Table 1 gives the values of the polynomial coefficidents used. Figure 4 shows the Bode magnitudes of the models, and the input signal power spectrum.

The calculated PRBS was used to generate a calibration set of 500 samples from both processes. A segment of the PRBS in-

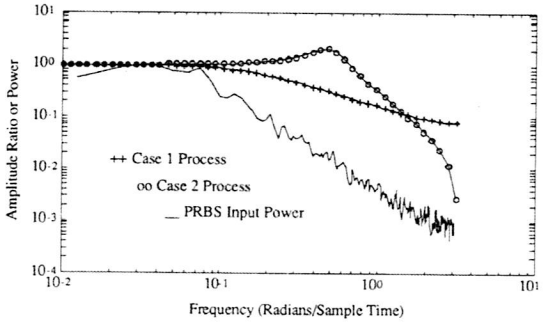

Fig. 4. Bode magnitude for test processes with input signal power spectrum.

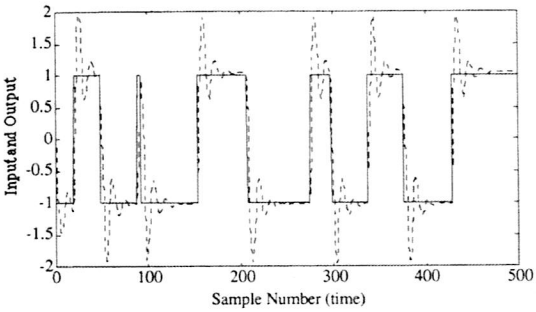

Fig. 5. Example PRBS input (solid) and Case 2 process output (dashed).

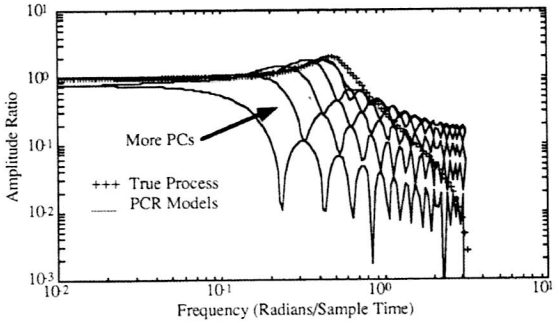

Fig. 7. Bode gain for second order process (+) and PCR models (solid lines).

put generated with these parameters is shown in Fig. 5, along with the output from the Case 2 process. Principal components regression was used to identify FIR models of the noise-free processes. In each case 30 FIR coefficients were used. The frequency-response of these models was calculated as a function of the number of PCs retained in the regression.

Figures 6 and 7 show the Bode magnitudes for each of the true systems. They also show the frequency responses of a series of FIR models (identified with PCR, for 1 to 6 principal components, PCs). The 1-PC models only match the process behavior at low frequency. As PCs are added, the model fit extends to progressively higher frequencies.

The dips in the gain for the PCR models are a consequence of the sinusoidal nature of the PCs. The FIR model identified using one PC has coefficients that are a cosine function of a particular frequency. Certain input frequencies are orthogonal to this frequency and are attenuated in the model output. Note, e.g., that the gain dips in Figures 6 and 7 occur at frequency intervals of 2π, as one would expect.

Figures 8 and 9 show the step response of the processes and the PCR models. As PCs are added, the shape and gain of the model step response improves. In no case is the high-frequency (short-time) response modeled well, however, confirming the results shown in Figs. 6 and 7. The primary reason for this is the limited input power at high frequencies (see Fig. 4).

The consequences of this behavior are important and deserve further clarification.

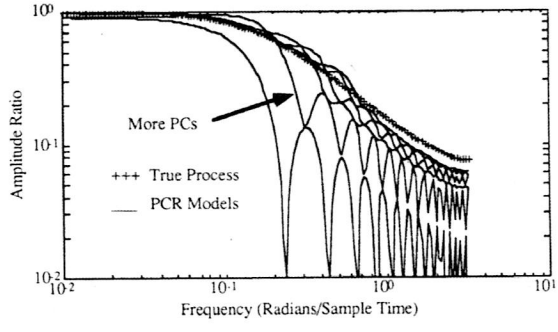

Fig. 6. Bode gain for first order process (+) and PCR models (solid lines).

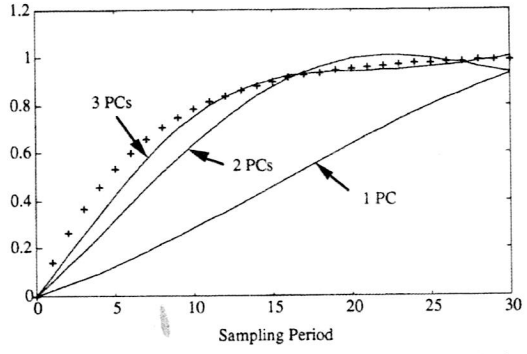

Fig. 8. Step response for first order process (+) and PCR models (solid lines).

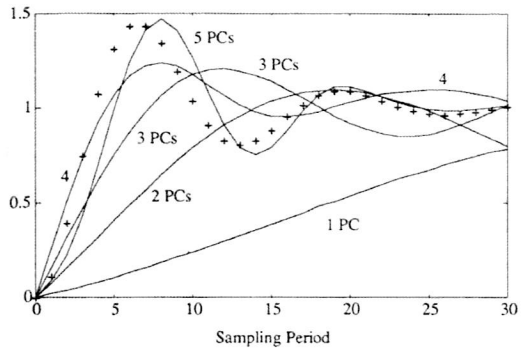

Fig. 9. Step response for second order process (+) and PCR models (solid lines).

Suppose that the first PC from the decomposition of the *ACM* can be represented as a continuous cosine function of period 4π (frequency = 0.5) over the interval from $-\pi$ to π. The function goes through one-half cycle in this interval. It is similar to the first PC from a typical ACM decomposition, as demonstrated previously (see Fig. 1). For a sinusoidal input, the process output $y(t)$ of the 1-PC (continuous) FIR model must then be

$$y_1(t) = c_1 \int_{-\pi}^{\pi} \cos(0.5\ x)\ \cos(m(x+t))\ \mathrm{d}x \tag{12}$$

where m is the frequency of the input signal, and c_1 is the constant determined from regression of $\mathbf{y}$ onto the scores vector $\mathbf{t}_1$. The solution to this integral is

$$\begin{aligned} y_1(t) = {} & \{\sin[-mt + (0.5-m)\pi] \\ & + \sin[mt + (0.5-m)\pi]\} \\ & \times \{2(0.5-m)\}^{-1} \\ & + \{\sin[mt + (0.5+m)\pi] \\ & + \sin[-mt + (0.5+m)\pi]\} \\ & \times \{2(0.5+m)\}^{-1} \end{aligned} \tag{13}$$

When m is an integer plus 0.5, the value of the numerator in both terms is identically zero for all t. Thus the frequencies 1.5, 2.5, 3.5, etc., will not pass through the 1-PC model.

Let us further assume that the second PC can be represented as a sine function of period 2π (frequency 1) over the same interval (again, compare with Fig. 1). The contribution of the second PC to the model will then be

$$y_2(t) = c_2 \int_{-\pi}^{\pi} \sin(x)\ \cos(m(x+t))\ \mathrm{d}x \tag{14}$$

When m is an integer greater than 1, the value of the integral in (12) is zero for all t. Thus, the second PC will make no contribution to the 2-PC model at frequencies of 2, 4, 6, etc. The response at these frequencies will equal that of the 1-PC model. There are no values of $m < 1$ for which the integral vanishes.

The behavior shown in Figs. 6 and 7 agrees with the above mathematical argument. The 1-PC model does not pass certain frequencies occurring at even intervals. The 2-PC model passes these frequencies but adds nothing to the centers of the intervals of the 1-PC model. This pattern repeats as we include more PCs in the model. Changes in the frequency content of the input signal will vary the location of the dips, but they will appear in any case.

Numerical results indicate that as PCs are added, the FIR models fit the true system behavior first in the frequency range where the power if the input was concentrated. For instance, in a band-pass filter is used to generate the input signal, the models will first emphasize the pass-band. As PCs are added, the fit in the neighboring regions will improve. This is consistent with the theory of eigenvector decomposition of the *ACM* developed here.

EFFECT OF INPUT FILTERING ON ACM DECOMPOSITION

The ACM for a process other than white noise through a first-order filter will have a more complex decomposition. Figures 5 and 6 show that the results are quite similar to the ideal *ACM* case, however.

The following example illustrates how different filtering options affect the *ACM* decomposition. An input signal was generated by filtering a 1000-sample white-noise sequence through a first-order process as in eq. (5) with $\alpha = 0.8$. Its 30-point autocorrelation matrix (including time shifts from -30 to $+30$ units) was computed. The original signal was filtered through a first-order process with α varied from 0.1 to 0.7. Figure 10 compares the eigenvalues of the *ACM* from the filtered signal with those of the original signal. Note how the filtering has a larger effect on the small eigenvalues (associated with higher frequencies) than on the large eigenvalues. This is expected because the filter is a low-pass type. Increasing the value of α lowers the cutoff frequency of the filter and begins to reduce the larger eigenvalues (associated with lower frequencies) while having an even greater impact on the smaller ones. The eigenvectors of this *ACM* change very little as α increases (not shown – see Wise [9]).

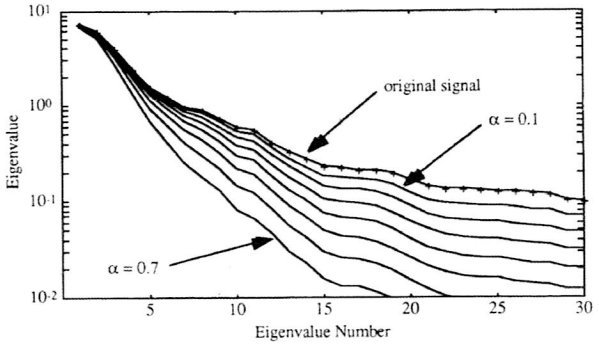

Fig. 10. Eigenvalues of autocovariance matrix for original signal (+) and filtered signals (solid lines) as cutoff frequency is lowered.

We note in passing that, for FIR model identification, low-pass pre-filtering of the input and output data is probably not a good idea. This is particularly true when the Multiple linear regression (MLR) method is to be used. Low pass filtering can make the autocovariance matrix extremely ill-conditioned. Multiple linear regression uses the inverse of this matrix, amplifying the noise in the data.

CONCLUSIONS

The frequency-domain properties of FIR models identified via Principal Components Regression have been explored. We have shown that PCR decomposes certain types of input signals by frequency. The decomposition is similar to a Fourier transform, but the frequencies are orthogonal over the "FIR window" rather than over an infinite domain.

An FIR model identified by PCR fits the true process behavior in a limited frequency range — that associated with the PCs retained in the model. This causes artifacts ("gain dips") in the frequency response of the model. We also showed that pre-filtering of the input signal has a significant effect on the eigenvalues of the resulting *ACM*, but not on the PCs. Thus, FIR models derived from complex input signals will also have frequency-domain artifacts. The effect of such artifacts on model-based controllers has yet to be reported.

Related work [12] has shown that the optimal choice of the number of PCs reduces the impact of noise. This, in combination with the above results, suggests that PCR can be viewed as a tunable filter. By choosing the number of PCs, one adjusts the filter cutoff to optimize the predictive ability of the re-

sulting model. The simplicity of this approach should be attractive to practitioners.

REFERENCES

1 C.E. Garcia, D.M. Prett and M. Morari, Model predictive control: Theory and practice — A survey, *Automatica*, 25 (1989) 335–348.

2 L. Ljung, *System Identification: Theory for the User*, Prentice-Hall, Englewood Cliffs, NJ, 1987.

3 N.L. Ricker, The use of biased least-squares estimators for parameters in discrete-time pulse–response models, *Ind. Eng. Chem. Res.*, 27(2) (1988) 343–350.

4 D.E. Rivera, J.F. Pollard and C.E. Garcia, Control-Relevant parameter estimation via prediction-error methods: Implications for Digital PID and QDMC control, Presented at AIChE Annual Meeting, Chicago IL, 1990.

5 D.E. Rivera, J.F. Pollard, L.E. Sterman and C.E. Garcia, An industrial perspective on control-relevant identification, Proc. 1990 American Control Conference, San Diego, CA, IEEE, 1990, pp. 2406–2411.

6 C.R. Cutler and F.H. Yocum, Experience with the DMC inverse for identification, in: Y. Ackun and W.H. Ray (Eds.), *Chemical Process Control–CPCIV*, AIChE, New York, 1991, pp. 297–317.

7 C. Zervos, P.R. Belanger and G.A. Dumont, On PID controller tuning using orthonormal series identification, *Automatica*, 24 (1988) 165–175.

8 W.R. Cluett and L. Wang, Modelling and robust controller design using step response data, *Chem. Eng. Sci.*, 46(8) (1991) 2065–2077.

9 B.M. Wise, Adapting Multivariate Analysis for Monitoring and Modeling Dynamic Systems, *Ph. D. Dissertation*, University of Washington, Seattle, WA, 1991.

10 B.M. Wise and N.L. Ricker, Feedback strategies in multiple sensor systems, *AIChE Symposium Series of Process Sensing*, 85 (267) (1989) 19–23.

11 B.M. Wise and N.L. Ricker, The effect of biased regression on the identification of FIR and ARX models, *AIChE 1990 Annual Meeting*, Nov. 1990.

12 B.M. Wise and N.L. Ricker, Identification of finite impulse response models with continuum regression. *J. Chemometrics* (submitted) 1992.

13 B.M. Wise, N.L. Ricker, D.J. Veltkamp and B.R. Kowalski, A theoretical basis for the use of principal component models for monitoring multivariate processes, *Process Control and Quality*, 1 (1990) 41–51.

14 P. Geladi and B.R. Kowalski, PLS tutorial, *Anal. Chim. Acta*, 185(1) (1986) 1–17.

15 A. Lorber, L.E. Wangen and B.R. Kowalski, A Theoretical Foundation for the PLS Algorithm, *J. Chemometrics*, 1 (1987) 19–31.

16 T. Naes and H. Martens, Principal component regression in NIR analysis: Viewpoints, background details and selection of components, *J. Chemometrics*, 2 (1988) 155–167.

17 G.E.P. Box and G.M. Jenkins, *Time Sereis Analysis: Forecasting and Control*, Holden-Day, Oakland CA, 1976.

18 K.J. Åstrom and B. Wittenmark, *Computer Controller Systems: Theory and Design*, Prentice-Hall, Englewood Cliffs, NJ, 1984.

APPENDIX

The continuous-time autocorrelation matrix is a continuous function of two variables, which we will call x and y. These vary over a finite domain. In the discrete-time case, it is possible to multiply the *ACM* (a function of two indices) by a vector (a function of one index) to obtain a vector (also function of one index). In an analogous manner, one can mutliply the continuous *ACM* by a function in one variable, say x, then integrate to obtain a function in one variable, y. Let us define the continuous *ACM* for x and y in the interval $[-a, a]$. The width of the interval is proportional to the FIR window width. For simplicity, assume that the continuous *ACM* can be defined as

$$ACM(x, y) = e^{-|x+y|} \tag{A1}$$

This does not affect the generality of the solution.

If a cosine function is an eigenfunction of the analogous continuous-time problem,

$$\int_{x=-a}^{x=a} e^{-|x+y|} \cos\left(\frac{x}{n}\right) dx = \lambda \cos\left(\frac{y}{n}\right) \tag{A2}$$

for some choices of n and λ. The integral in eq. (A2) can be evaluated as follows:

$$\int_{x=-a}^{x=a} e^{-|x+y|} \cos\left(\frac{x}{n}\right) dx = \begin{Bmatrix} \int_{-y}^{a} e^{-(x+y)} \cos\left(\frac{x}{n}\right) dx & \text{for } x+y>0 \\ \int_{-a}^{-y} e^{(x+y)} \cos\left(\frac{x}{n}\right) dx & \text{for } x+y<0 \end{Bmatrix} \tag{A3}$$

$$\int_{x=-a}^{x=a} e^{-|x+y|}\cos\left(\frac{x}{n}\right) dx$$
$$= (e^{-(\pi+y)} + e^{-\pi+y})$$
$$\times\left(\frac{-\cos\left(\frac{a}{n}\right)}{1-n^{-2}} + \frac{\sin\left(\frac{a}{n}\right)}{(1+n^{-2})n}\right)$$
$$+ \frac{2\cos\left(\frac{-y}{n}\right)}{1+n^{-2}} \qquad \text{(A4)}$$

For certain values of n, the first term on the right hand side of eq. (A4) is identically zero. This occurs when

$$n = \frac{\sin(a/n)}{\cos(a/n)} = \tan(a/n) \qquad \text{(A5)}$$

This gives the allowed periods of the cosine eigenfunctions. (The true period would be equal to $2\pi n$.) The eigenvalues associated with each period n, λ_n are given by

$$\lambda_n = 2/(1+n^{-2}) \qquad \text{(A6)}$$

In a similar manner, one can rewrite eq. (A2), with sine replacing cosine. The result is that the allowed periods of the sine eigenfunction are given by

$$n = \frac{-1}{\tan(a/n)} \qquad \text{(A7)}$$

The associated eigenvalues are are in eq. (A6).