Classification of process dynamics with Monte Carlo singular spectrum analysis

by C. Aldrich* and M. Barkhuizen*† †Paper written on project work carried out in partial fulfilment of M.Sc. (Chemical Engineering) degree

Synopsis

Identification of non-linear systems can be a daunting task and in the process industries the problem is complicated considerably by the presence of noise from various sources, non-stationarity of the data and intermittence, such as observed in particulate flows. Traditional methods, such as frequency analysis and linear modelling do not handle these systems well. Similar problems arise with the application of non-linear theory developed over the last few decades, where much of the analysis depends on embedding of the data in a phase space or pseudophase space, since these methods were not originally designed to deal with noisy systems.

In contrast, singular spectrum analysis is a method which can be used to identify structure and distinguish noise from important information in multivariate data, such as the trajectory matrices of time series measurements. It provides an orthogonal basis onto which the data can be transformed, so that components of the time series can be investigated individually. With Monte Carlo methods, the components can be compared with those from linear surrogate data in order to identify possibly non-linearities in the data. In this paper, this versatile technique is briefly explained and illustrated by means of case studies.

Introduction

The modelling of process systems has attracted considerable interest in industry and academia, owing to increased emphasis on the automation of plant operations, which are driven in turn by environmental constraints, safety legislation and competitive pressures. Since it is often not feasible to model complex systems from first principles, models are usually constructed from process data. Even so, identification of non-linear systems can be a daunting task and in the process industries the problem is complicated considerably by the presence of noise from various sources, non-stationarity of the data and intermittence, such as observed in particulate flows. Traditional methods, such as frequency analysis and linear modelling do not handle these systems well. Similar problems arise with the application of nolinear theory developed over the last few decades, where much of the analysis depends on embedding of the data in a phase space or pseudophase space, since these methods were not originally designed to deal with noisy systems.

In contrast, singular spectrum analysis does not suffer from these drawbacks. It allows a time series to be decomposed into different components, e.g. components representing the underlying signal itself, as well as various noise components, which can be used to classify the time series. For example, a basic distinction could be made between stochastic and deterministic time series, as this would have major implications with regard to the modelling and control of the system. Such classification is not a trivial process, but singular spectrum analysis is one of the most promising methods to have emerged during the last decade. The basic theory is briefly explained as follows.

Singular spectrum analysis

The spectral decomposition of matrices has only recently been applied to time series analysis (Vautard and Ghil, 1989; Vautard *et al.*, 1992; Różyński *et al.*, 2001; Pereira and Maciel, 2001) and has had its roots mostly in the application of chaos theory (Broomhead and King, 1986). In essence, the data are embedded in a high-dimensional reconstruction, followed by the introduction of a new coordinate system, where the origin is moved to the centroid of the reconstructed system states and the axes are represented by the dominant principal components of the states (points).

More formally, singular spectrum analysis can be seen as a four-step process, as indicated in Figure 1 (Golyandina *et al.*, 2001). In the first step (embedding), the one-dimensional time series is recast as an L-dimen-

* *Department of Chemical Engineering, University of Stellenbosch, Matieland, South Africa.*

 SA ISSN 0038–223X/3.00 + 0.00. Paper received Jan. 2003; revised paper received Feb. 2003.

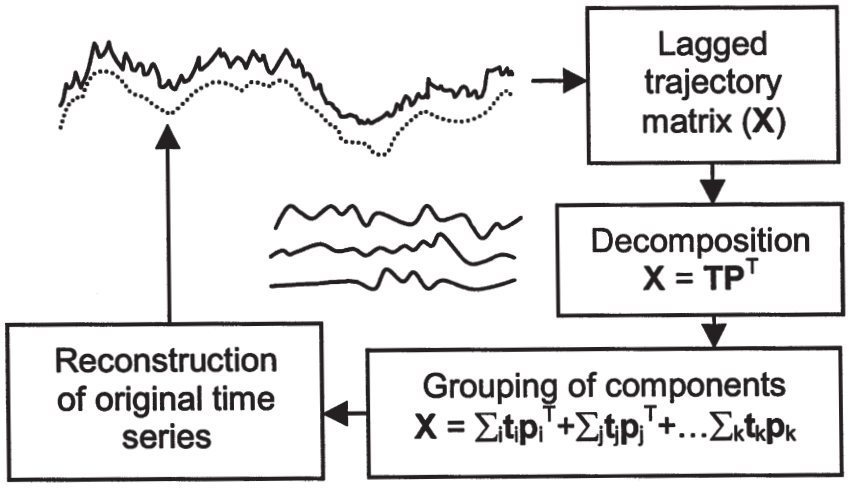

Figure 1—Decomposition and reconstruction of a time series by use of singular spectrum analysis

sional time series (trajectory matrix). In the second step (singular value decomposition), the trajectory matrix is decomposed into a sum of bi-orthogonal matrices of rank one. These two steps constitute the decomposition stage of singular spectrum analysis. In the third and fourth steps, which form the reconstruction stage of the process, the components are grouped and the time series associated with the groups are reconstructed. The process can be explained briefly as follows.

Decomposition stage

Embedding of time series

The embedding step is used to expand the original time series into what is referred to as the *trajectory matrix* of the system. The matrix is associated with a certain window length, i.e. the number of columns of the matrix into which the series is embedded. The series is expanded by giving it a lag of unity and creating a number of lagged vectors.

This can be illustrated by letting y_t be a selected time series.

$$y_t = [0;\ 2;\ 4;\ 6;\ 8;\ 10;\ 12;\ 14;\ 16;\ 18;\ 20]$$

If the time series were to be embedded in a matrix with a window length of four, the resulting trajectory matrix X would be:

$$X = \begin{bmatrix} 0 & 2 & 4 & 6 \\ 2 & 4 & 6 & 8 \\ 4 & 6 & 8 & 10 \\ 6 & 8 & 10 & 12 \\ 8 & 10 & 12 & 14 \\ 10 & 12 & 14 & 16 \\ 12 & 14 & 16 & 18 \\ 14 & 16 & 18 & 20 \end{bmatrix}$$

The optimal size of the embedding window (number of columns in the trajectory matrix) should be sufficient in order to capture the global behaviour of the system, as indicated by the minimum of either the point of linear decorrelation (where the correlation between the first and last column in matrix X is negligible), or the point where the correlation between the first and last column (lagged vectors) reaches its first minimum.

Decomposition of time series

Once the time series has been embedded into a trajectory matrix X, singular value decomposition (SVD) or equivalently, principal component analysis, is performed on the matrix lagged covariance matrix $A = X^T X/(n\text{-}1)$. With principal component analysis the trajectory matrix is decomposed into the product of a score (T) and a transposed loading (P) matrix which is equivalent to the sum of the outer products of pairs of vectors t_i and p_i, from which T and P are composed.

$$X = TP^T = t_1 p_1^T + t_2 p_2^T + ...t_k p_k^T + E \qquad [1]$$

The number of principal components that should be retained depends strongly on the aims of the analysis. If it is purely for visualization purposes, it could be adequate to explain as little as 60% of the variance and therefore only a few principal components need to be retained. However, for modelling purposes it may be desirable to explain a much larger amount of the variance, thereby ensuring that none of the essential information from the data is lost. To accomplish this it would be necessary to retain a larger number of the principal components.

Reconstruction stage

The aim of the grouping step is to separate the additive components of the time series. In this situation, it can be seen as separating the time series into two groups, viz. the original signal and the noise component. The criteria for this separation are once again a matter that has not been formalized completely and it depends on knowledge of the data.

Once the principal components to be retained have been identified, a new time series with a reduced number of principal components can be reconstructed. This is done in the fourth step, diagonal averaging. The calculations involve

taking the average of the matrix elements on each 'diagonal' of the matrix of the retained principal components. If properly reconstructed (smoothed), the time series should represent the underlying process dynamics more reliably, which should expedite process modelling, optimization and control.

Multivariate time series

Generally, in applications in process engineering, one is presented with series of measurements on a set of variables, rather than just a single variable. Multivariate or multichannel SSA as it is sometimes referred to, is a natural extension of the approach discussed above for a single time series and has been considered (Elsner and Tsonis, 1996, Golyandina *et al*, 2001). When considering n observations at equal time intervals on a set of Q variables or Q-channel time series $\{X_q(t):, q = 1, 2, \ldots Q\}$, the generalization of SSA to a multivariate time series can be approached by first computing a trajectory matrix for each variable X, i.e. for the k'th variable

$$X_k = \begin{bmatrix} X_{k,1} & X_{k,2} & . . & X_{k,m} \\ X_{k,2} & X_{k,3} & . . & X_{k,m+1} \\ . & . & . . & . \\ X_{k,n'-1} & . & . . & X_{k,n-1} \\ X_{k,n'} & X_{k,n'+1} & . . & X_{k,n} \end{bmatrix} \quad [2]$$

with $1 \le k \le Q$. These lagged trajectory matrices are subsequently used to form an augmented trajectory matrix, $D = (X_1, X_2, X_3, \ldots X_Q)$, from which a grand lag covariance matrix, C_X, can be constructed, analogous to that obtained with a single time series (Broomhead and King, 1986). Each block $C_{i,j}$ is a matrix containing estimates of the lag covariance between channels i and j.

$$C_X = \begin{bmatrix} C_{1,1} & C_{1,2} & . & . & C_{1,Q} \\ C_{2,1} & C_{2,2} & . & . & . \\ . & . & . & . & . \\ . & . & . & . & C_{Q\text{-}1,Q} \\ C_{Q,1} & . & . & C_{Q\text{-}1,Q} & C_{Q,Q} \end{bmatrix} \quad [3]$$

Monte Carlo SSA

Monte Carlo singular spectrum analysis (MC-SSA) is a methodology for discriminating between various components of the time series, particularly components containing meaningful information and other components containing mostly noise (Allen and Smith, 1996). This problem is especially important in process engineering applications, such as modelling, control, data validation and filtering. Although so-called white noise (additive measurement noise) is relatively easy to detect and remove, the situation becomes more complicated when the noise also drives the system, such as is the case in autoregressive moving average stochastic processes. These stochastic processes have frequency spectra that decrease monotonically with frequency and are often referred to as warm-coloured.

MC-SSA involves a null hypothesis against which the data are tested, as well as a discriminating statistic. The data are first assumed to belong to a specific class of dynamic processes, e.g. 1st order autoregressive processes or more broadly stationary linear Gaussian processes in general, perhaps distorted by some non-linear measurement system (sensor). Surrogate data are subsequently generated from this process and various statistics are calculated from both the surrogate and the original data (Theiler *et al*., 1992). If the calculated statistics of the surrogate and the original data are different, then the null hypothesis that the process that has generated the original data is of the same class as the system that has generated the surrogate data is rejected.

More specifically, let $x \in \Re^n$ be a time series consisting of n observations, ψ a specific hypothesis, $\Im\psi$ the set of process systems consistent with the hypothesis, and T: $\Re^N \rightarrow U$ be a statistic that will be used to evaluate the hypothesis ψ that x was generated by some process $\Im \in \Im_\psi$. In this paper, the statistic is the eigenspectrum of the time series from which it is possible to discriminate between the original data x and the surrogate data x_s consistent with the hypothesis given by the probability density of T, given $\Im$, i.e. $p_{T,\Im}(t)$.

Generation of surrogate data

Surrogate data

Surrogate data are artificially generated stochastic data that mimic certain features of an observed time series. For instance, surrogate data may have the same mean, variance, Fourier power spectrum or autocorrelation function as the measured time series. More formally, let $\{x_j\}$, $i = 1, 2, \ldots N$ be an observed time series. According to the null hypothesis, $x_j = h(s_j)$, where s_j is the realization of a (linear) Gaussian stochastic process, while h is a static measurement function, which could be non-linear. For a surrogate data set (z_j) to represent the null hypothesis, it has to have the same autocorrelation function (R), as well as empirical amplitude distribution (A) as the original time series at a sufficiently large time lag τ", i.e.

$$R_Z(\tau) = R_X(\tau), \text{ for } \tau = 1,2,\ldots\tau'' \quad [4]$$

and

$$A_Z(z) = A_X(x) \quad [5]$$

The general approach in surrogate data analysis is as follows.

(i) Generation of sets of surrogate data, each similar to the original time series, i.e. of the same length and statistically indistinguishable from the original time series with regard to certain specified characteristics, as specified by the user.

(ii) Calculation of a discriminating statistic for the measured time series and the sets of surrogate data. Any statistic quantifying some aspect of the time series can be used, such as forecasting error, correlation dimension, largest Lyapunov exponent, etc.

(iii) Setting up of a hypothesis that there is no difference between the discriminating statistic (eigenspectrum) of the original time series and the surrogates (null hypothesis).

(iv) Testing of the null hypothesis based on the values of the discriminating statistic and acceptance or rejection of the null hypothesis. Acceptance of the null hypothesis implies that the original data are stochastic (i.e. the same as the surrogate data). Since the probability distributions of the test statistics are usually unknown, Monte Carlo methods have to be used to estimate the acceptance or rejection criteria for the test.

Since a linear stochastic time series is completely characterized by its Fourier spectrum (or equivalently, the autocorrelation function), it should be a minimum requirement that the surrogate data and the original time series have the same power spectra (or autocorrelation functions). Assuming that the original observed time series is derived from a specific class of a linear random process, the probability distribution of the eigenspectrum is determined for an ensemble of surrogate data sets, which are just different realizations of the hypothesized linear stochastic process (Monte Carlo simulation).

In this investigation, all the surrogate data were generated by an iterative amplitude adjusted Fourier Transform algorithm, which consisted of the following steps

(i) Generation of a normally distributed (random) data set y, reordered to have the same rank distribution as the observed (original) data set, x
(ii) Phase-shuffling the Fourier transform of y, to create a data set y_s,
(iii) Rank ordering of y_s and replacing the amplitudes y_{sj} with that of x_i of corresponding rank.
(iv) Adjustment of the observations to ensure that the autocorrelation function of the surrogate data and the original data are identical. This may mean that the amplitude adjustments in the previous step have to be modified and steps (iii) and (iv) are therefore repeated until discrepancies between the amplitudes and autocorrelation are minimal.

Case studies

Flow in a series of tanks

To illustrate the use of singular spectrum analysis, consider the response of the flow of four tanks in series, the overall transfer functions of which can be described by the transfer function

$$G(s) = 1/\left[(0.38s+1)(2s+1)(2.62s+1)(0.1s+1)\right] \quad [6]$$

Figure 2 shows the actual (solid line) and simulated measured (+) response to a pulsed input (broken line). The measured response was simulated by adding zero mean Gaussian noise with a standard deviation of 0.1 to the actual response. The trajectory matrix derived from the time series data consisted of 23 columns, each copy delayed by a time step of one. This matrix formed the basis from which 23 principal components were extracted.

The toroidal shape of the attractor described by the scores of the first three principal components (Figure 3) reflects the roughly periodic behaviour of the system. The loose end portrayed at the bottom left of the Figure indicates the initial transient behaviour of the system.

Surrogate data were generated from the measured response time series and confidence limits for the eigenspectrum were estimated by means of Monte Carlo simulations. These are indicated in Figures 4 and 5 for the data, including the transient response and excluding it respectively. As can be seen from Figure 4, the eigenspectrum of the non-stationary system lies outside its estimated confidence limits, suggesting that the null hypothesis of a stationary, linear Gaussian system has to be rejected. Since the system is Gaussian and linear, this rejection can only be attributed to the non-stationarity of the system. This is confirmed by the data in Figure 5, showing that the null hypothesis should be accepted if the transient part of the time series (first 100 seconds) is omitted.

Autocatalysis in a continuous stirred tank reactor

The second case study concerns an autocatalytic process in a continuous stirred tank reactor originally considered by Gray and Scott (1983, 1984) and subsequently investigated by Lynch (1992). The system is capable of producing self-sustained oscillations based on cubic autocatalysis with

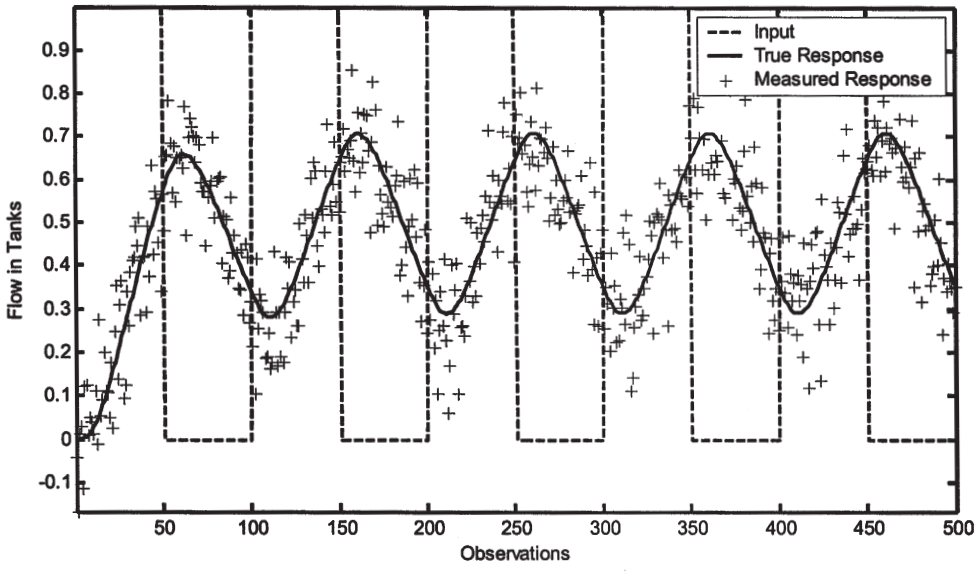

Figure 2— Actual response (solid line) of a 4th order system to a pulsed input (dashed line) and simulated observations ('+' markers)

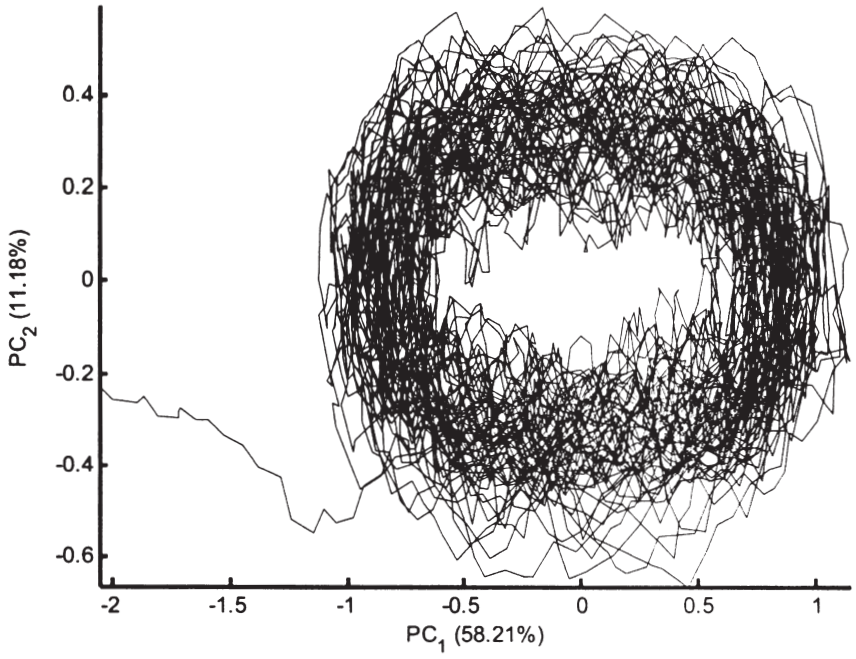

Figure 3—Attractor of the measured response shown in Figure 2, including transient behaviour from approximately 0-100 seconds

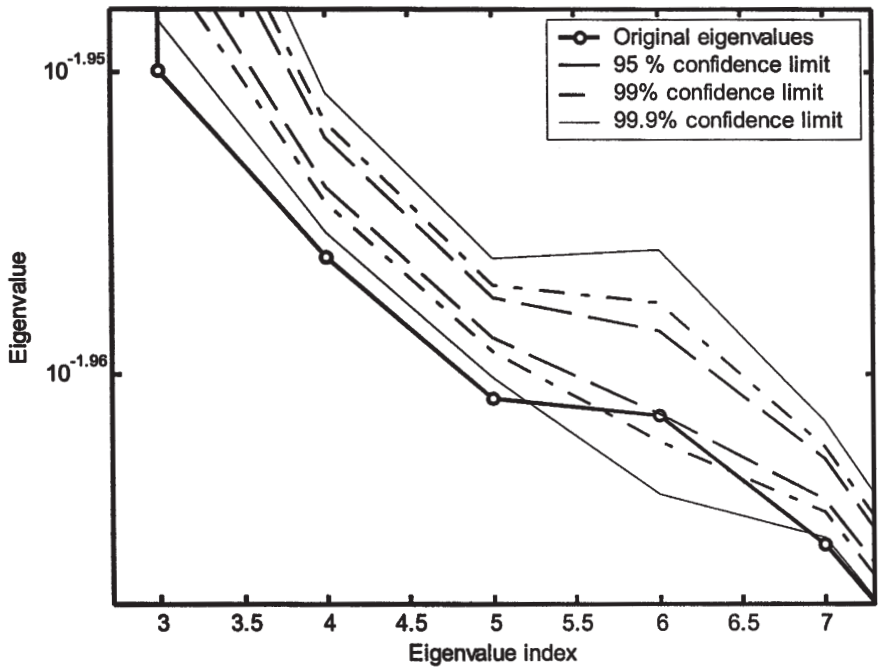

Figure 4—Confidence limits of the eigenspectrum for the response ('+' markers) shown in Figure 2, with non-stationary data

catalyst decay and proceeds mechanistically as follows.

$$\begin{aligned} A + 2B &\rightarrow 3B, \quad -r_A = k_1 c_A c_B^2 \\ B &\rightarrow C, \qquad r_C = k_2 c_B \\ D + 2B &\rightarrow 3B, \quad -r_D = k_3 c_D c_B^2 \end{aligned} \qquad [7]$$

where A, B, C and D are the participating chemical species and k_1, k_2 and k_3 the rate constants for the chemical reactions. This process is represented by the following set of ordinary differential equations.

$$\begin{aligned} \frac{dX}{dt} &= 1 - X - aXZ^2 \\ \frac{dY}{dt} &= 1 - Y - bYZ^2 \\ \frac{dZ}{dt} &= 1 - (1 + c)Z + daXZ^2 + ebYZ^2 \end{aligned} \qquad [8]$$

where X, Y, and Z denote the dimensionless concentrations of species A, B and D, while a, b and c denote the Damköhler

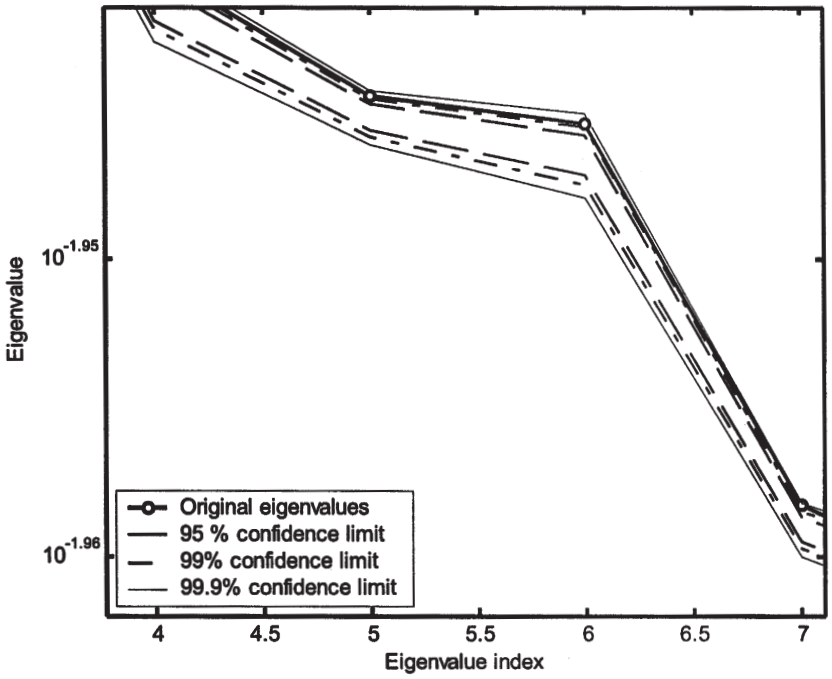

Figure 5—Confidence limits of the eigenspectrum for the response ('+' markers) shown in Figure 2, with stationary data

numbers for A, B and D respectively. The ratio of feed concentration of A to that of B is denoted by d and the same ratio of D to B by e.

The process is chaotic, with a well-defined attractor for specific ranges of the two parameters, d and e. For the settings $a = 18000$; $b = 400$; $c = 80$; $d = 1.5$; $e = 4.2$, and initial conditions $[0, 0, 0]^T$, the set of equations was solved by using a 5th order Runge Kutta numerical method over 100 simulated seconds. This gave approximately 10 000 observations, which were resampled with a constant sampling period of 0.01 s. The Y state was taken as the output variable. Figure 6 shows the attractor of the process reconstructed from the process states X, Y and Z.

In Figure 7, the eigenspectrum and estimated confidence limits of the simulated measurements from the autocatalytic system are shown. The eigenspectrum of the original observations falls outside these confidence limits by a wide margin, owing to the non-linearity of the data (which are otherwise known to be stationary and non-Gaussian (not a stochastic time series). In practice, systems are usually contaminated with measurement noise, and not as readily classified as the previous two simulated systems. Such a system is considered in the next case study.

Feed composition of a base metal flotation plant

The data in the third case study were obtained from a South African copper flotation plant. The plant consists of a crushing section and milling circuit, followed by a magnetic separation section. The purpose of the magnetic separation is to remove the high percentage of magnetic material in the ore and thereby reduce the load on the flotation circuit. The flotation circuit itself is designed to operate with feed grades of 0.6% Cu, 9.0% Pb, 2.4% Zn and 130 g/t silver. The circuit configuration consists of two conditioners, in which sulphurous acid, two copper collectors and a frother is added. From the two rougher banks the concentrate is circulated to the three cleaner banks, where zinc is depressed by acid in the first cleaner and lead depressed in the second and third cleaners by adding lime. The cleaner tails and the scavenger concentrate are returned to the copper feed of the flotation circuit, and these two streams make up the bulk of the feed.

The feed grades (%Cu, %Pb and %Zn) sampled every 12 minutes over a period of one month is shown in Figure 8. These data were subsequently analysed by means of singular spectrum analysis. Each of the three time series consisted of 1234 observations and these were embedded in trajectory matrices $X_{Cu} \in \Re^{1205 \times 29}$, $X_{Pb} \in \Re^{1034 \times 200}$ and $XZ_n \in \Re^{1185 \times 48}$, based on the estimated autocorrelation of each time series.

These trajectory matrices were subsequently decomposed into principal components, each with its own eigenspectrum, with upper and lower confidence limits computed from 15 surrogate data sets generated by the iterative amplitude adjusted Fourier transform algorithm. The results are shown in Figures 9–11.

As can be seen from Figure 9, the eigenspectrum of the copper data falls outside the 99% confidence limits for the eigenspectrum, which means that the null hypothesis of a stationary linear Gaussian system has to be rejected. Further analysis suggests that the data are roughly normal, as indicated by the top quantile-quantile plot in Figure 12.

The stationarity of the system can be assessed by comparing the mean of the first half of the observations with the mean of the second half in an ANOVA test. The results (with those for the other two metals) are summarized in Table I. In this Table the total sums of squares (SS) is partitioned into the sums of squares for the variables (Period 1 and Period 2), as well as the error terms. As can be seen from the results, the probability of the Pb and Zn being stationary is negligible, while that of the copper is approximately 12%. It is a crude test, but nonetheless an indication

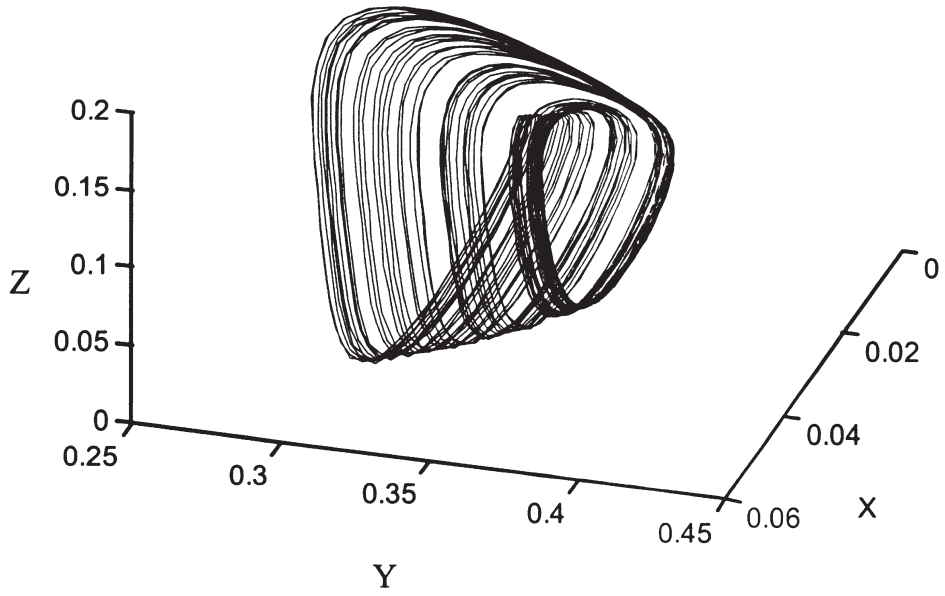

Figure 6—Attractor of autocatalytic process constructed from process states X, Y, Z

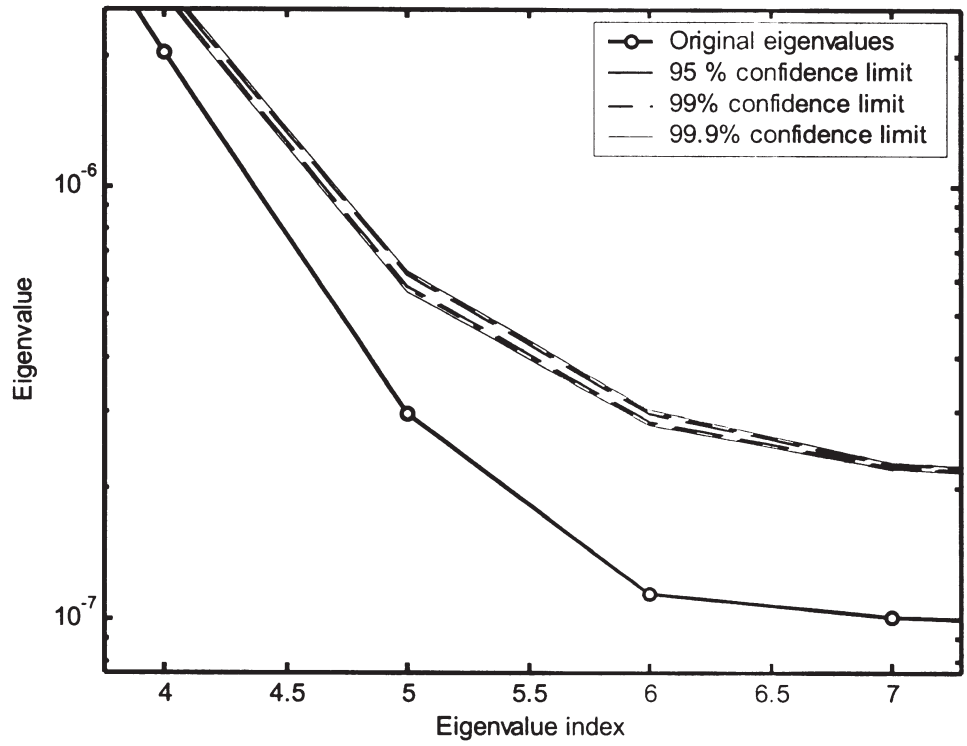

Figure 7—Confidence limits of the eigenspectrum for the autocatalytic system shown in Figure 6

of the stationarity of the data. On this basis alone, the eigenspectra of all three metals should fall outside the confidence limits generated by the Monte Carlo analyses.

Nonetheless, it can be seen from Figure 9, that the deviation of the copper eigenspectrum is larger than those of the other two metals, although it appears to be more stationary than the other two metals. This may imply that the copper data are more non-linear than the other two metals. However, these possibilities would have to be investigated further before firm conclusions would be possible.

The eigenspectrum of the %Pb falls mostly outside the confidence limits derived from the surrogate data. Although not shown, only eigenvalues 1, 3, 4, and 5 are within the limits. On this basis, these data are also not strictly stationary and Gaussian (as suggested by further analysis of the data). Characterization and analysis of the contribution of the components associated with the 1st, 2nd 4th and 5th eigenvalues of the time series is likely to shed more information on the character of the time series, but this is beyond the scope of the present investigation.

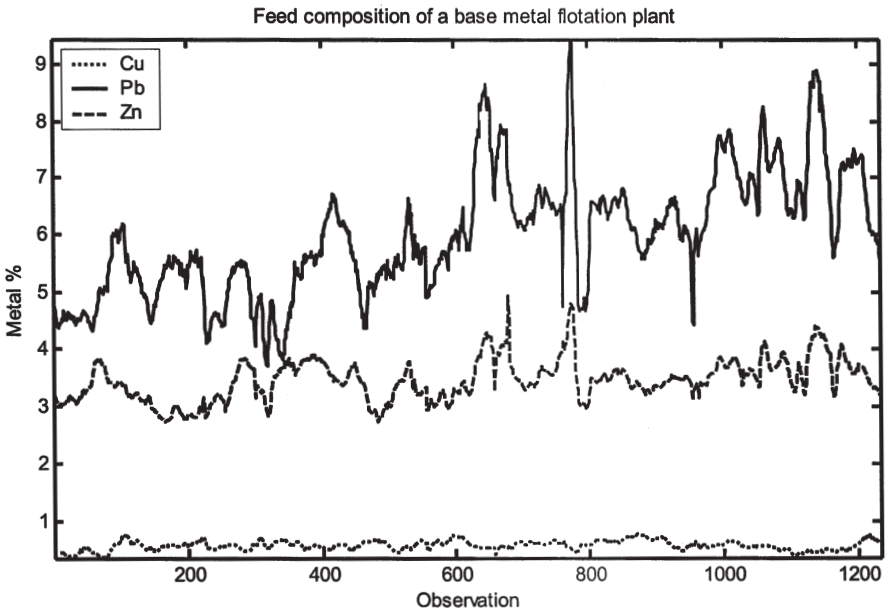

Figure 8—Base metal grades in the feed to base metal flotation plant

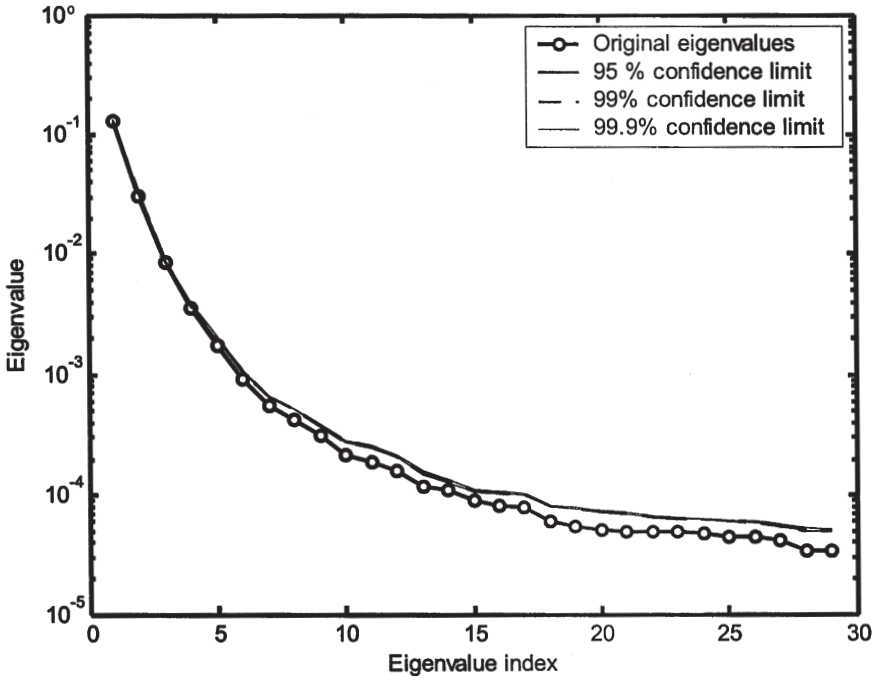

Figure 9—Confidence limits of the eigenspectrum for the %Cu in the feed

In the eigenspectrum of the %Zn, only eigenvalues 3, 6, 7, 8, 9, and 10 fall outside the confidence limits derived from the surrogate data. Again, strictly speaking, the data are therefore not entirely stochastic, or stationary and further tests are necessary. As with the Pb data, non-stationarity is clearly evident, the extent of which is shown by Figure 13, where most of the deviations between the first and second half of the time series can be attributed to the lead and zinc contents in the feed.

Discussion and conclusions

For a discrete time series of finite length, singular spectrum analysis makes use of the principal component decomposition of an estimate of the correlation matrix that is based on m lagged copies of the time series, which forms the trajectory matrix of the time series. The resultant eigenvectors form an optimal basis that is orthonormal at zero lag and permit the signal to be decomposed into its possibly oscillatory and

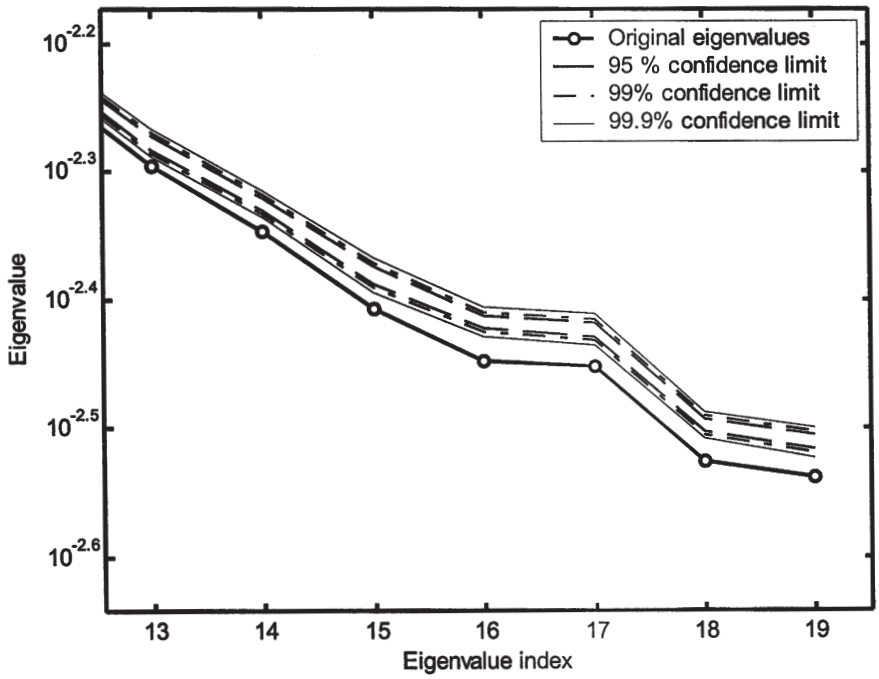

Figure 10—Confidence limits of the eigenspectrum for the %Pb in the feed for eigenvalues 13-19

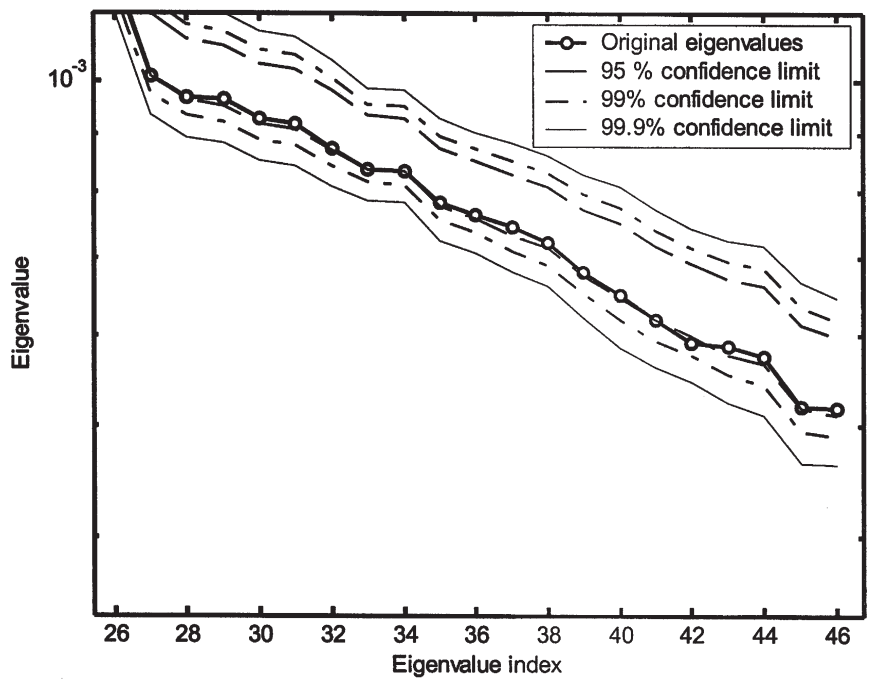

Figure 11—Confidence limits of the eigenspectrum for the %Zn in the feed, for eigenvalues 27–46

aperiodic components. The eigenspectrum associated with these eigenvectors are a generalized statistic that characterizes the nature of the time series and can be used to discriminate between time series or time series components on the basis of stochasticity/determinism, linearity/non-linearity, etc.

In this paper, the eigenspectrum as a whole was considered in the characterization of the time series. This is a very stringent approach, as the eigenspectrum consists of a ranked series of eigenvalues and more sophisticated analysis is possible by testing the individual eigenvalues. For example, the first few eigenvalues of the time series is usually associated with the trend or major variation in the time series, while eigenvalues with higher indices are associated with the fine structure (noise) of the time series. By requiring the null hypothesis to be valid for all

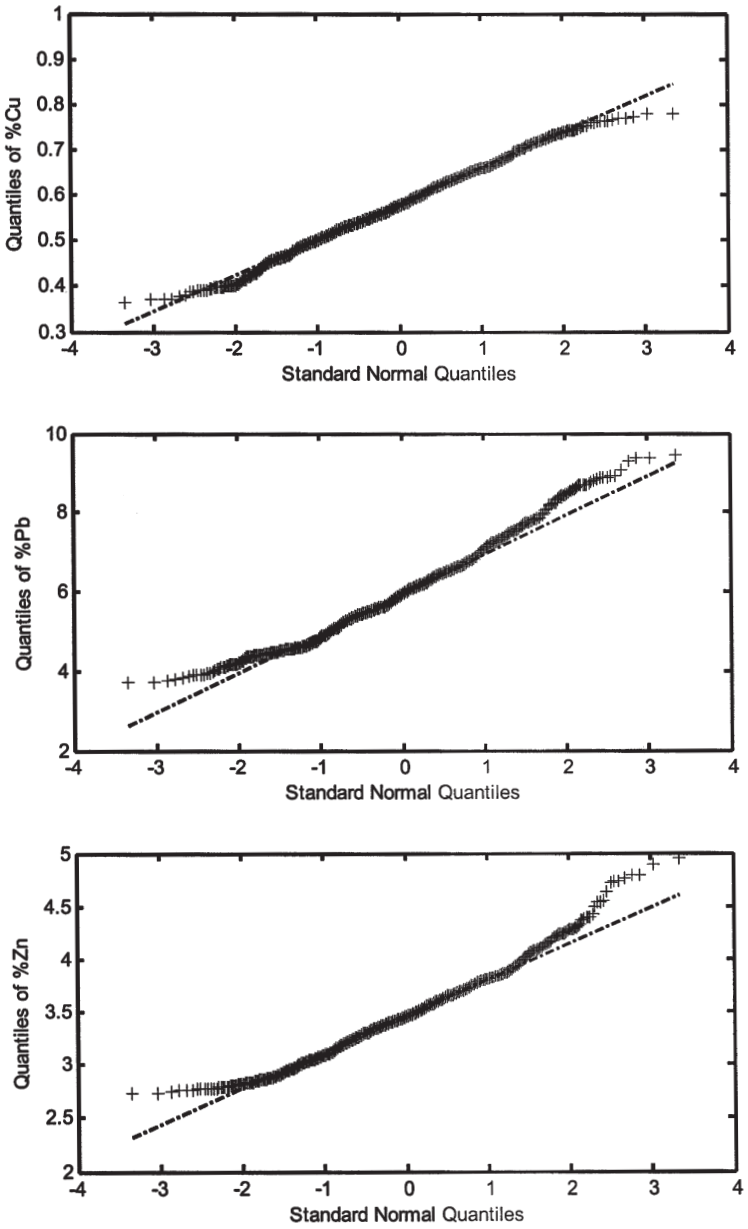

Figure 12—Q-Q plots of Cu, Pb and Zn feed data versus standard normal

components of the time series, it may mean that the time series is classified on the basis of the nature of its least significant components as well. Note however, that distinguishing between significant and insignificant components is not a trivial matter. Moreover, the interpretation of the statistical tests should be done very carefully, to account for possible biases in the tests and flawed surrogate data. Despite these caveats, the general approach outlines in this paper constitutes a promising route towards the classification and analysis of time series, and ultimately better system identification, process control and optimisation of plant operations through better data analysis.

Acknowledgements

Mintek's financial support for this project is gratefully acknowledged.

References

BROOMHEAD, D.S. and KING, G.P. Extracting qualitative dynamics from experimental data. *Physica D*, 1986, 20, pp. 217–236.

ELSNER, J.B. and TSONIS, A.A. *Singular Spectrum Analysis— A New Tool in Time Series Analysis,* 1996, Plenum Press, New York.

GOLYANDINA, N., NEKRUTKIN, V., and ZHIGLJAVSKY, A. *Analysis of Time Series Structure – SSA and Related Techniques,* 2001, Chapman and Hall/CRC, Boca Raton.

GRAY, P. and SCOTT, S. K. Autocatalytic reactions in the isothermal, continuous stirred tank reactor. Isolas and other forms of multi-stability, *Chemical Engineering Science,* 38, 1983. pp. 29–43.

GRAY, P. and SCOTT, S. K. Autocatalytic reactions in the isothermal, continuous stirred tank reactor. Oscillations and instabilities in the system A + 2B → 3B, B → *C. Chemical Engineering Science,* 39, 1984. pp. 1087–1097.

LYNCH, D.T. Chaotic behavior of reaction systems: parallel cubic autocatalators. *Chemical Engineering Science*, 47(2), 1992. pp. 347–355.

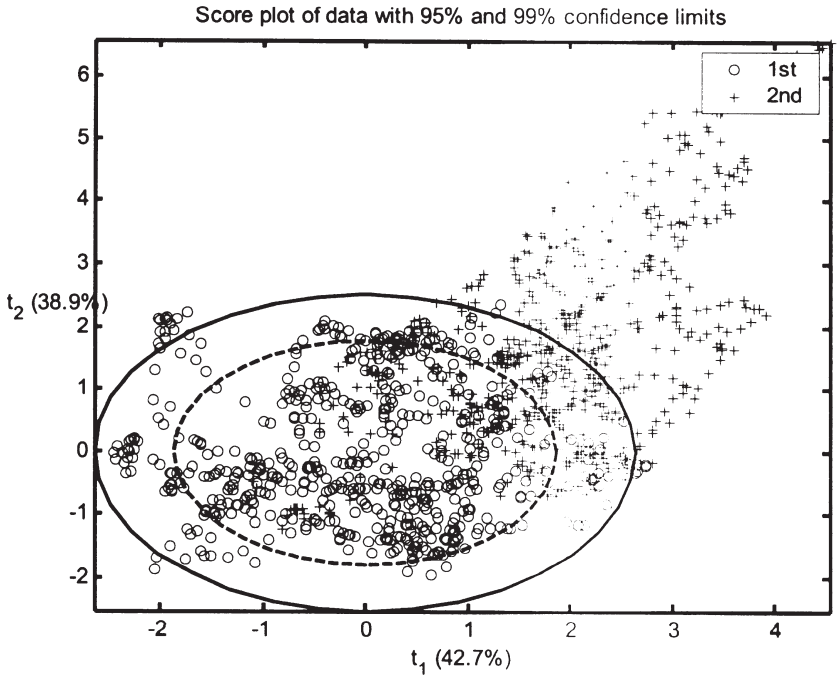

Figure 13—Non-stationarity of the feed to the flotation plant, showing the principal component scores of the first two weeks (o) and the last two weeks of the feed (+). The 95% (broken line) and 99% (solid line) confidence limits are indicated by the ellipses containing most of the (o) data

Table I

Analysis of variance for Cu, Pb and Zn

Cu	Source	SS	df	MS	F	Prob>F
	Periods	0.0403	1	0.0403	6.41	0.012
	Error	7.76	1232	0.0063		
	Total	7.80	1233			
Pb	Source	SS	df	MS	F	Prob>F
	Periods	700.47	1	700.47	1268.6	0
	Error	680.23	1232	0.5521		
	Total	1380.7	1233			
Zn	Source	SS	df	MS	F	Prob>F
	Periods	35.45	1	35.45	341.08	0
	Error	128.06	1232	0.104		
	Total	163.52	1233			

PEREIRA, W.C.D.A. and MACIEL, C.D. Performance of ultrasound echo decomposition using singular spectrum analysis. *Ultrasound in Medicine and Biology*, 2001, 27(9), pp. 1231–1238.

RÓŻYŃSKI, G., LARSON, M., and PRUSZAK, Z. Forced and self-organized shoreline response for a beach in the southern Baltic Sea determined through singular spectrum analysis. *Coastal Engineering*, 2001, 43, pp. 41–58.

THEILER, J., EUBANK, E., LONGTIN, A., and GALDRIKIAN, B. Testing for non-linearity in time series: The method of surrogate data. *Physica D*, 58, 1992. pp. 77–94.

VAUTARD, R. and GHIL, M. Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series. *Physica D*, 1989, 35, pp. 395–424.

VAUTARD, R., YIOU, P., and GHIL, M. Singular spectrum analysis: A toolkit for short, noisy, chaotic signals. *Physica D*, 1992, 58, pp. 95–126. ◆

Metallurgical coke conference set for October 7–9, 2003 in Toronto

February 21, 2003—The Seventh Annual Conference on Metallurgical Coke and Coal will be held in Toronto, Canada on October 7–9 this year. Chaired by Dr Neil Bristow (BHP Billiton) and Mr Patrick Cleary (CRU International), the conference is titled 'The Met Coke World Summit 2003—Meeting the Challenges of Booming Demand, Rising Costs, Chinese Supply Dominance and New Environmental Restriction', and organized by the Intertech Conference Group in Portland, USA.

'This conference will focus on current trends affecting the metallurgical coke industry—including supply and demand dynamics, cost and pricing trends, environmental issues, coking coal availability/quality/costs, and new technologies', said Dr Hugh Olmstead, Intertech's Coke and Coal Program Manager. The conference will be devoted to helping producers and users of coke to successfully respond to pressures associated with increased globalization, industry consolidation, rising energy and coal costs, and tightening environmental regulations. Olmstead noted that 'Attendees will get a comprehensive appraisal of coke production and capacities by region, coke trade balances, new coke-making technologies, the rate of closing/building/rebuilding coke batteries, by-products control and marketing, and the practical impact of the on-going rationalization of the world steel industry'.

Over 200 executives are expected to attend this event from the fields of coal, coke, iron and steel, plant design, project development, and international trade and transport. The conference will feature 23 speakers on October 8–9 and two panel discussions. Two 3-hour pre-conference seminars will be held on October 7 and a field trip is planned for October 10 to coke batteries in Hamilton, Ontario at the Dofasco and Stelco steel works. There will also be a number of sponsored networking functions, and an exhibit area for display of new products and processes.

2002 saw record steel and pig iron production, and also record coke consumption. Booming Chinese steel demand coupled to environmental closures of coke capacity in China, led to strongly rising coke prices and shortages. This situation has major implications for steel companies that face critical decisions pertaining to meeting their current and future coke needs.

This conference is timely because the entire international coke and coking coal sector faces four critical uncertainties today—the resolution of which will have a profound effect on international steel, foundry and ferroalloy practice over the next decade. These are:

- Can we make enough coke and coal to meet booming world demand, especially for developing regions of the world?
- Should we repair, up-grade or build new coke facilities—and with what designs if we go new—to meet the shortage of coke plants?
- Can we get enough high quality coking coal at reasonable prices? and
- How can we meet the increasingly stringent environmental regulations that affect our industry?

Clearly, coke is at a crossroads today. In 2003 over 350 million tons of coke will be consumed globally, with demand forecast to grow by at almost 10 million tons each year through 2007. But widespread rationalization and consolidation of the global steel industry, growing shortages of domestic and merchant coke, increasingly stringent coke quality demands from blast furnace ironmakers, enforced coke oven battery closures, the steep rise in energy costs, growing EAF steelmaking facilities, tight coking coal supplies, oven design and rebuild issues, and Chinese dominance of the export market all present major challenges. On the other hand, there are many opportunities that will enable coke to continue to play a vital and growing role—advanced value-in-use techniques, improved blast furnaces operating practices, new coke plant designs, new business models and partnerships, and new coke plant technology.

Conference attendees will get a comprehensive assessment in three days from experts in industry, transportation, trade, and research, as they explore all the business practices, market developments, and technological advances affecting the coke industry.

The major conference sessions are:

- Met coke market analysis and forecast
- New developments in coke making technology
- Deployment of coke, PCI, heat and by-products
- The role of China in the merchant coke market
- Met coal: A growing challenge for the coke industry
- Meeting the new environmental and regulatory challenges
- New strategies for making coke cost competitive.

Intertech, based in Portland, Maine, USA, is a consultancy specializing in iron and steel, coke, metals, ores and specialty chemicals. The group produces 25 industrial conferences each year in Western Europe, North America, and Asia. Additional information about current programmes may be found at www.intertechusa.com

Speaker recruitment for this conference is currently under way. To submit a topic idea, or to obtain more information, please contact Dr H.D. Olmstead, Conference Director, Intertech Conferences, 19 Northbrook Drive, Portland, Maine 04105 USA.
Telephone: +(207) 781-9800, Fax: +(207) 781-2150,
Email: olmstead@intertechusa.com ◆