Identification of nonlinearities in dynamic process systems

by J.P. Theron*† and C. Aldrich*

†Paper written on project work carried out in partial fulfilment of Phd. (Process Engineering) degree

Synopsis

Process modelling is an essential element in the development of advanced (model-based) process control systems, accounting for up to 80% of the cost of development. Often, models based on historic process data are the only viable option when dealing with processes that cannot be modeled cost-effectively from first principles. Despite major advances that have recently been made in the field of nonlinear process modeling, some basic idea of the dynamic behaviour of a process system is important prior to system identification. While many process systems seem unlikely to be linear, the possible nonlinear aspects of their dynamics are not necessarily supported by the data. This may have significant implications for subsequent system identification or interpretation of the data. For example, it is quite possible that an underlying nonlinear deterministic signal may be obscured by a high degree of measurement or process noise, to the extent that it may be critical in the choice of model to fit to the data. The method of surrogate data analysis has recently been proposed as a means to classify the underlying dynamics of complex systems, and in this paper the use of Fourier transform surrogates to classify process systems is explored.

Keywords: modelling, process control, simulation, surrogate data analysis, nonlinear systems.

Introduction

The last few decades have seen a steady increase in the use of computers for the advance control of process plants, owing to stiff competitive pressures and government regulations. Most of these advanced systems, of which model predictive control is the most popular, depend on accurate plant models. These predictive models are usually empirical and based on historic plant data, the availability of which has also increased considerably in recent years. Despite their ready availability, not all plant data are necessarily suitable for the development of process models. Critical variables may be unobservable or simply not measured, measurement intervals may be too long to capture essential plant dynamics, periods of observations may be too short, or where inferential measurements are used, the correlation between these and actual state variables may not be sufficiently precise. To complicate the problem, these shortcomings may be difficult to identify, as the plant engineer may have little, if any, other information related to the actual state variables of the system.

Although the quality of an accurate process model based on the data could ultimately be seen as a reflection of the high quality of the data themselves, the converse is not necessarily true. If a model fails, it does not automatically imply poor data. It may well imply a poor choice in the structure of the model, or a failure to parameterize the model properly. The cost of constructing models aside, this approach is clearly not satisfactory.

Before constructing any models, it may be a good idea to verify the assumptions being made about the data. These assumptions include notions about the (non)stationarity of the data, the stochastic nature of the data, and the (non) linearity of the data. In practice these classes will represent extremes and the question would rather be one of degree. Unfortunately, this is not a trivial task, since linear stochastic processes can create signals with a very complex appearance. As a consequence, not all structures in a time series can be attributed to nonlinear system dynamics. Irregular data can also result from random impulses to the system, or fluctuation of the process parameters. The method of surrogate data analysis provides some means of classifying the data and guiding the engineer towards the development of appropriate process models.

The method of surrogated data analysis

Surrogate data are stochastic data that are artificially generated to mimic certain features

* *Department of Process Engineering, University of Stellenbosch, Matieland, Stellenbosch*

 SA ISSN 0038–223X/3.00 + 0.00. Paper received Jan. 2004; revised paper received Mar. 2004.

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. The idea is to see whether the measured time series, which are similar to the surrogate data, have the same value of the selected measure, in which case it can be concluded that the original time series is the same as the surrogate data.

The general approach in surrogate data analysis (Theiler *et al.*, 1992; Takens, 1993; Kantz and Schmitz, 1997; Dolan and Spano, 2001) is as follows.

- 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
- 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, largest Lyapunov exponent, etc.
- Setting up a hypothesis that there is no difference between the discriminating statistics of the original time series and the surrogates (null hypothesis)
- Testing the null hypothesis based on the values of the discriminating statistics and acceptance or rejection of the null hypothesis.

Since a linear stochastic time series is completely characterized by its Fourier spectrum (or equivalently, 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). The discriminating statistic used to characterize the time series should be robust and not affected by the measurement of the system or other possible distortions. By using the correlation dimension of the time series as discriminating statistic (to be discussed later), a good indication of the topology of the underlying attractor of the time series can be obtained and it can indeed be considered to be a pivotal test statistic of the time series.

Assuming that the original observed time series is derived from a specific class of linear random process, the probability distribution of the correlation dimension is determined for an ensemble of surrogate data sets, which are just different realizations of the hypothesized linear stochastic process (Kantz and Schreiber). These issues are discussed in more detail in the following section.

Hypotheses

As an oversimplified example of the classification of a time series, the test could simply be to determine whether the data are random or not. In this case, the null hypothesis would be that the time series are simply realizations of a random (Gaussian) variable (white noise) and that there is no correlation between successive observations.

In this case, the surrogate data sets would consist of randomized versions of the original time series. By construction, the surrogates will have the same mean, standard deviations, amplitude distributions, etc., as the original time series. As a test statistic, the prediction error of some model could be used. If it is clear that the model can predict future observations of the original time series better than it can the future predictions of the randomized time series (surrogates), then the null hypothesis that the original time series is identical to the surrogates (i.e. random) has to be rejected. Otherwise, the null hypothesis is accepted.

Although the above is a simplistic example, the general idea with surrogate data analysis is the same, only more generalized assumptions are being evaluated. In fact, the most general assumption one can make about the original time series, is that the measurements were obtained from a stationary linear stochastic process, possible distorted by some nonlinear measurement function (i.e. the measurement of flow through a nonlinear valve). Consequently, this will serve as the null hypothesis in all further discussion.

Test statistic

The test statistic, T, (a single number), estimates the characteristics of the data and their variations in such a way as to be able to decide whether the time series is consistent with the null hypothesis (Kantz and Schmitz, 1997; Kugiumtzis, 2002). The probability distribution of the correlation dimension is the same for all processes, regardless of the source of the noise of the estimated model and it is calculated as follows.

Let $\{v_t\}$, $t = 1, 2, \ldots N$ be an embedding of a time series in $\Re^{dc}$. This means that the original time series has been transformed into a new set of coordinates in a phase space, as illustrated in Figure 1. In this figure a time series x = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10] is embedded with a lag of k = 2 and an embedding dimension of m = 3. The attractor of the time series can be visualized by plotting the embedding coordinates in three dimensions.

The correlation function, $C_N(\varepsilon)$, is defined by

$$C_N(\varepsilon) = \left[2/N(N-2)\right] \sum_{i=1}^{N-1}\sum_{j=i+1}^{N} I\left(\left\|v_i - v_j\right\| < \varepsilon\right) \quad [1]$$

where $I(\|\cdot\|)$ is a Heavyside function in $\Re^{dc}$ that returns 1, if the distance between point i and j is within ε and 0. Otherwise $\|\cdot\|$ denotes the Euclidean norm, but sometimes the supremum norm $\|(x_1, x_2, \ldots x_m)^T\| = sup_i|x_i| = max(|x_1|, |x_2|, \ldots |x_m|)$ is also used. The correlation dimension, d_c, is defined by

$$d_c = \lim_{\varepsilon\to 0}\lim_{N\to\infty}\log\left\{C_N(\varepsilon)/\log(\varepsilon)\right\} \quad [2]$$

The maximum and minimum value the correlation function can attain are $C(e) = 1$ and $C(\varepsilon) = 2/[N(N\text{-}1)]$

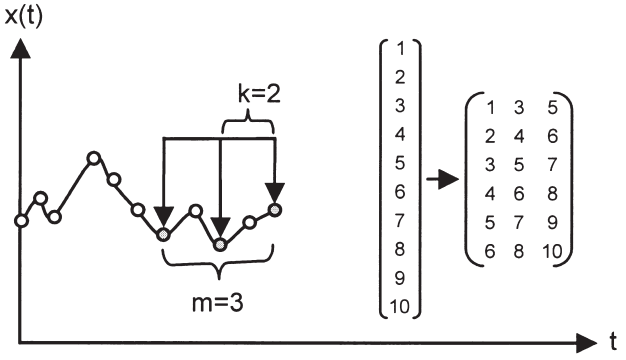

Figure 1— Embedding of a time series x = [1, 2, ... 10] with a lag of k =2 and an embedding dimension m = 3, giving the vectors Z = [(1, 3, 5), (2, 4, 6), ... (6, 8, 10)]

respectively. The idea is illustrated in Figure 2, which shows a hypersphere of radius e, centred on one of the points, v_j, defining the trajectory of the attractor. As the hypersphere is moved from one point to another on the attractor, the points that are enclosed in the given hypersphere are counted. This process is repeated for different values of ε and the correlation dimension is then estimated from a plot of the correlation function against the hypersphere size on logarithmic coordinates. At large values of ε, $d_c = 0$, i.e. where ε is large, than the attractor itself, and at very small ε-values, estimation of the correlation dimension is constrained by the statistics, as calculations have to be based on a minimum number of observations in each bin. Therefore, only the central region of the curve is considered and not the ends coinciding with very large and very small scales, which tend to be nonlinear (Grassberger and Procaccia, 1983a; Grassberger *et al.*, 1991).

With the Grassberger-Procaccia algorithm (Grassberger and Procaccia, 1983a; 1983b) used in this paper, a linear scaling region is required to calculate the correlation dimension reliably (Lai and Lerner, 1998). With this algorithm, a polynomial term is introduced to account for variations in the slope (scaling region) (Judd, 1992), which may be caused by noise in the data or less than optimal embedding of the time series. The correlation dimension, d_c, is thus a function of ε, rather than an invariant parameter for the time series (Judd, 1992). The polynomial $q(\varepsilon)$ is of the same order as the topological dimension of the embedding.

$$C_N(\varepsilon) \approx \varepsilon^{dc} q(\varepsilon) \quad [3]$$

The correlation dimension can be interpreted as the dimension of the attractor, or as a rough measure of the number of variables involved in the generating process of the time series (Diks, 1999; Cutler, 1993).

Generation of surrogate data

Amplitude adjusted Fourier transform (AAFT) surrogates (Small and Judd, 1998) and iterative amplitude adjusted Fourier transform (IAAFT) surrogates are generated as follows.

(i) Generation of a normally distributed data set x_r, reordered to have the same rank distribution as x_o, the observed (original) data set
(ii) Phase-shuffling of the Fourier transform of x_o to generate x_s
(iii) Rank ordering of x_s and replacing the amplitudes x_{sj} with those of x_{oi} of corresponding rank. This yields AAFT surrogates x_s. Step (iv) is necessary to generate IAAFT surrogates
(iv) Adjustment of the observations to ensure that the autocorrelation function of the surrogate data x_s and the original data x_o 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 (Schreiber and Schmitz, 2000; Kugiumtzis, 2002).

Note that when applied to multivariate systems, it is relatively straightforward to retain cross-spectral correlations between time series by simply ensuring that the randomization indices are the same for the relevant time series (Paluš, 1996).

Case study 1: stationary linear gaussian process

To illustrate the use of surrogate data analysis, consider the following linear Gaussian stochastic processes represented by Equations [4]–[6].

$$x_t = 0.7x_{t-1} + 0.6\varepsilon_{t-1} \quad [4]$$

$$x_t = 0.7x_{t-1} + 0.2x_{t-2} + 0.6\varepsilon_{t-1} \quad [5]$$

$$x_t = 0.4x_{t-1} + 0.2x_{t-2} + 0.1x_{t-3} + 0.3\varepsilon_{t-1} + 0.4\varepsilon_{t-2} \quad [6]$$

where x_t denotes a state variable, ε_t denotes a Gaussian process with zero mean and standard deviation of 0.15 and the subscript t refers to the time of the observation.

In each case, the state variables are assumed to be observed through a nonlinear measurement function of the form $y_t = (x_t+1)^{1/2}$, i.e.

$$y_1 = \left[0.7x_{t-1} + 0.6\varepsilon_{t-1} + 1\right]^{1/2} \quad [7]$$

$$y_2 = \left[0.7x_{t-1} + 0.2\varepsilon_{t-2} + 0.6\varepsilon_{t-1} + 1\right]^{1/2} \quad [8]$$

$$y_3 = \left[\begin{matrix}0.4x_{t-1} + 0.2x_{t-2} + 0.1x_{t-3} + \\ 0.3\varepsilon_{t-1} + 0.4\varepsilon_{t-2} + 1\end{matrix}\right]^{1/2} \quad [9]$$

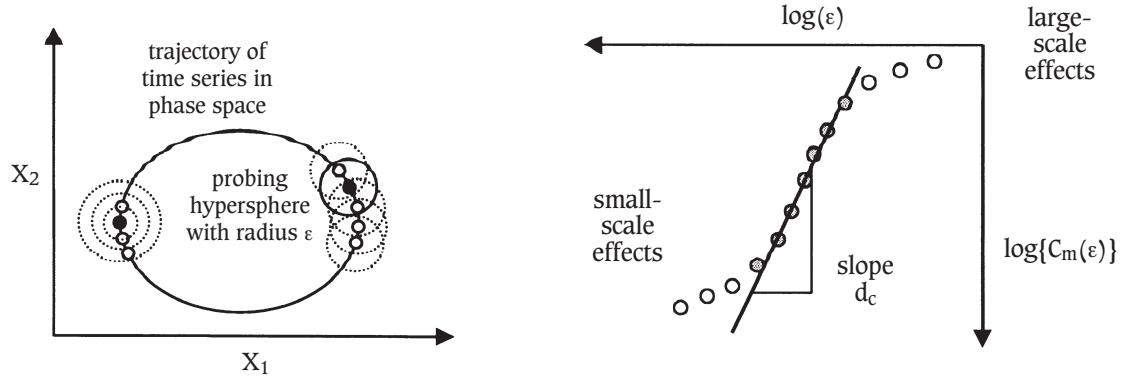

Figure 2—Calculation of the correlation dimension of a time series

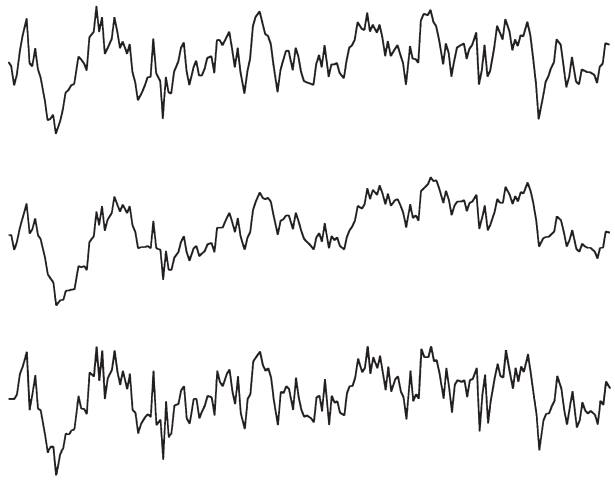

Figure 3—Samples of 1st, 2nd and 3rd order linear Gaussian processes observed through a nonlinear measurement function (Equations [7]–[9] in case study 1)

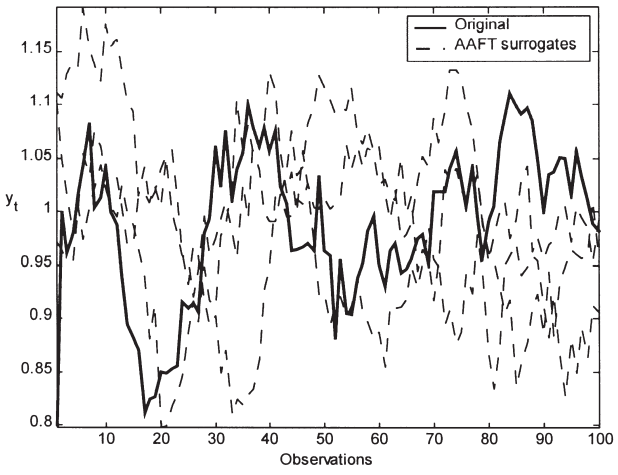

Figure 4—Samples of AAFT surrogates (broken lines) derived from the original time series (bold line) in case study 1 (Equations [7]–[9]

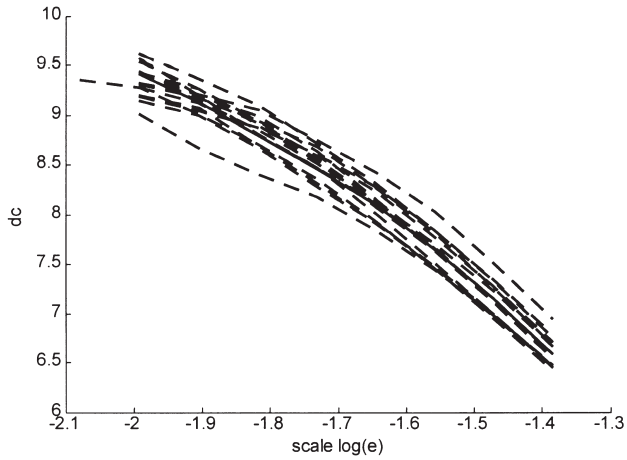

Figure 5—Correlation dimension (d_c) as a function of scale (e) for y_1 (solid line) and its AAFT surrogates (broken lines)

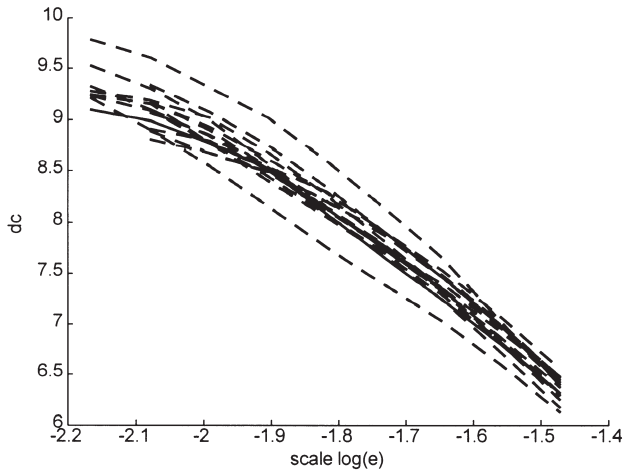

Figure 6—Correlation dimension (d_c) as a function of scale (e) for y_2 (solid line) and its AAFT surrogates (broken lines)

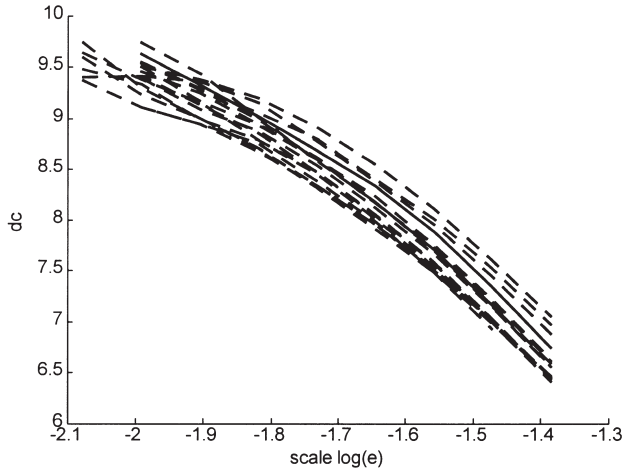

Figure 7— Correlation dimension (d_c) as a function of scale (e) for y_3 (solid line) and its AAFT surrogates (broken lines)

Ten thousand samples of each time series were generated, the first 200 observations of each of the time series y_1, y_2 and y_3 of which are shown in Figure 3. For each time series y_1, y_2 and y_3, a set consisting of 15 surrogate time series were generated by use of the amplitude adjusted Fourier transform method. As an example, Figure 4 shows three AAFT surrogate time series for the system represented by Equation 4. The correlation dimensions of each of the surrogate time series, as well as the original time series, were consequently calculated. The results are shown in Figures 5–7. Broken lines in these figures indicate surrogate data, and bold lines indicate the original data. For these figures, it is clear that the correlation dimensions of the original time series do not differ from the surrogate data over a large range of scales. It can therefore be concluded (acceptance of the null hypothesis) that the original time series are indeed linear Gaussian systems, with possible monotonic nonlinear transformation.

Case study 2: 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 catalyst decay and proceeds mechanistically as follows.

$$\begin{aligned} A + 2B &\rightarrow 3B, -r_A = k_1 c_A c_B^2 \\ B &\rightarrow C, r_C = k_2 c_B \\ D + 2B &\rightarrow 3B, -r_D = k_3 c_D c_B^2 \end{aligned} \quad [10]$$

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.

$$\begin{aligned} dX/dt &= 1 - X - aXZ^2 \\ dY/dt &= 1 - Y - bYZ^2 \\ dZ/dt &= 1 - (1+c)Z + daXZ^2 + ebYZ^2 \end{aligned} \quad [11]$$

where X, Y, and Z denote the dimensionless concentrations of species A, B and D, while a, b and c denote the Damköhler 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 8 shows the attractor of the process reconstructed from the process states X, Y and Z.

The smooth shape of the attractor in Figure 8a already hints at a nonlinear deterministic system, but in the presence of noise (Figure 8b) the character of the system is unclear. The correlation dimensions of the original time series, as well as their AAFT surrogates are shown in Figure 9a and 9b. Figure 9a shows that the system, (bold line) is a low order ($d_c < 2$ over the entire range of scales) nonlinear system, very distinct from its surrogate data (broken lines). In contrast, the presence of noise has a profound effect on the system (Figure 9b), as it is barely distinguishable from its surrogate data. While it can be shown that the system in Figure 8a can be modelled accurately, predictive modelling of the system in Figure 8b would be limited at best.

Case study 3: electrochemical noise generated by corrosion

Electrochemical noise is a general term to describe the apparently random fluctuations of current and potential in electrochemical processes. The measurement of electrochemical noise was studied in the 1970s and 1980s as a means to detect localized corrosion phenomena, such as pitting, crevice corrosion, and cavity attack. The corrosion current is related to the kinetics of the reaction, while the corrosion potential is related to the process thermodynamics. The idea is to identify specific corrosion phenomena, such as crevice corrosion or stress corrosion cracking, as well as the severity of these effects, from specific patterns in the electrochemical current and potential (Mansfeld and Xioa, 1993; Legat and Dolecek, 1995), which often exhibit low-dimensional chaotic behaviour (Lin *et al.*, 2001).

In this case study the aqueous corrosion of austenitic stainless steel is considered. The experimental set-up (see Figure 10) consisted of a voltammograph (CV-27) connected to a corrosion cell containing 500 ml NaOH and a Hewlett-Packard 34970A data acquisition unit connected to a computer. The electrochemical potential and current noise were measured simultaneously at 432 ms intervals with a three-sensor configuration. The current noise was measured between two of these sensor elements, i.e. two identical working electrodes consisting of austenitic stainless steel 304 strips, which were polished and degreased in hexane prior to use. The potential noise was measured between the third sensor or reference (an Ag/AgCl calomel) electrode and the other two coupled electrodes.

Samples of the data are shown in Figure 11, after removal of trends to ensure stationarity. As can be seen from the quantile-quantile plots in Figure 12, the corrosion current was distinctly non-Gaussian. In contrast, the corrosion potential measurements appear to have a normal distribution. From the localization index of the current measurements, $LI = s/RMS(I) = 0.996$, (Equation [12]), it appears as if localized (pitting) corrosion had taken place, although the experiment was not conducted over a sufficiently long period to validate the assumption.

$$LI = s / RMS(I) \quad [12]$$

where LI is the localization index of the current (LI close to 0 is associated with general non-local corrosion, while an LI

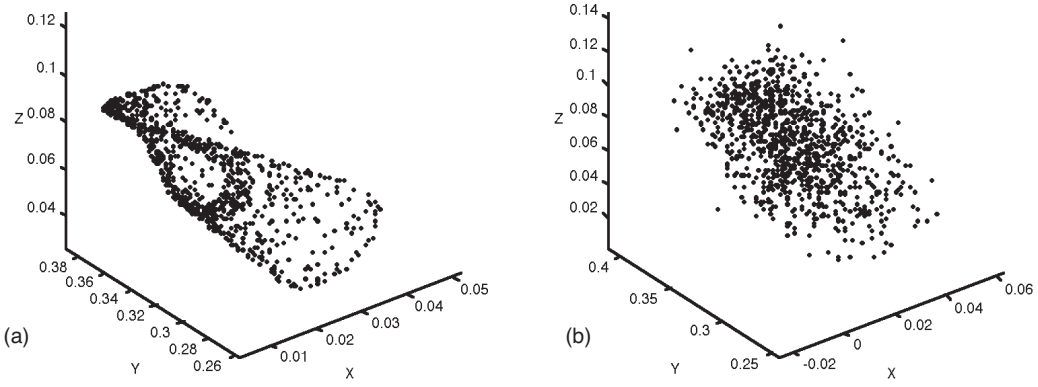

Figure 8—The attractor of the autocatalytic system (a) without noise, and (b) with Gaussian noise with zero mean and standard deviation of 0.015

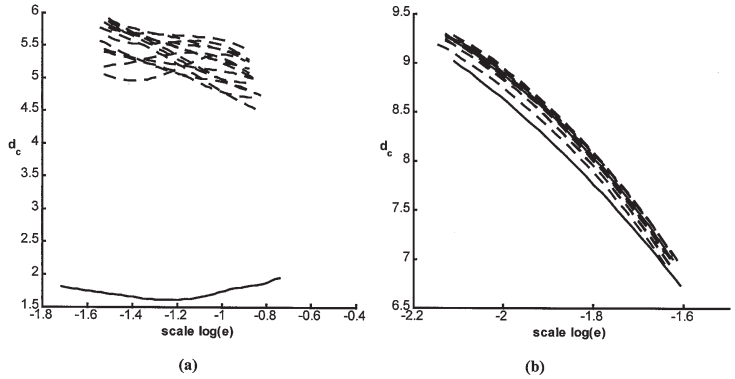

Figure 9—Correlation dimension curves for the data (solid lines) shown in Figure 5(a)-(c) and associated AAFT surrogate data (broken lines). In (c) the standard deviation was 0.015

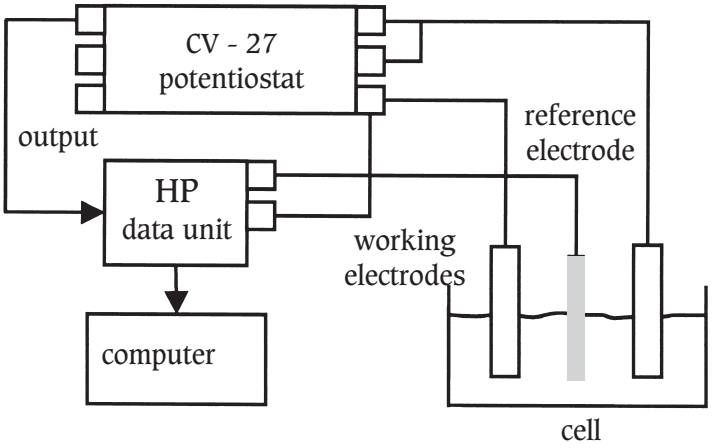

Figure 10—Experimental set-up for the measurement of electrochemical noise

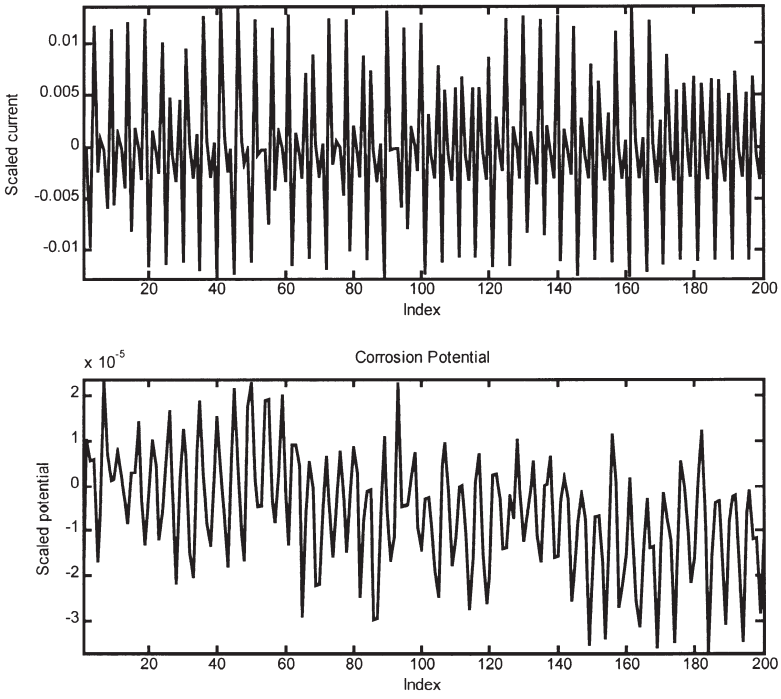

Figure 11—Residual values of detrended corrosion current and potential. (Only first 200 observations shown)

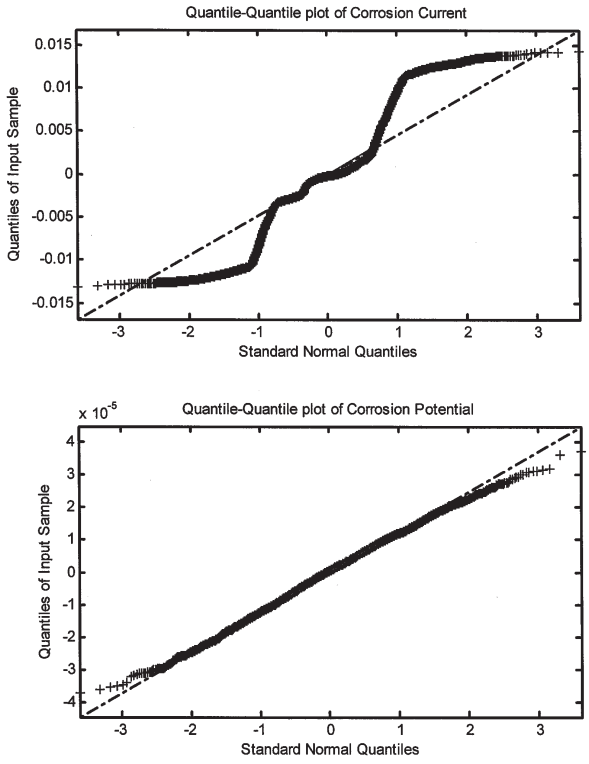

Figure 12—Quantile-quantile plots of the corrosion current (top) and potential (bottom)

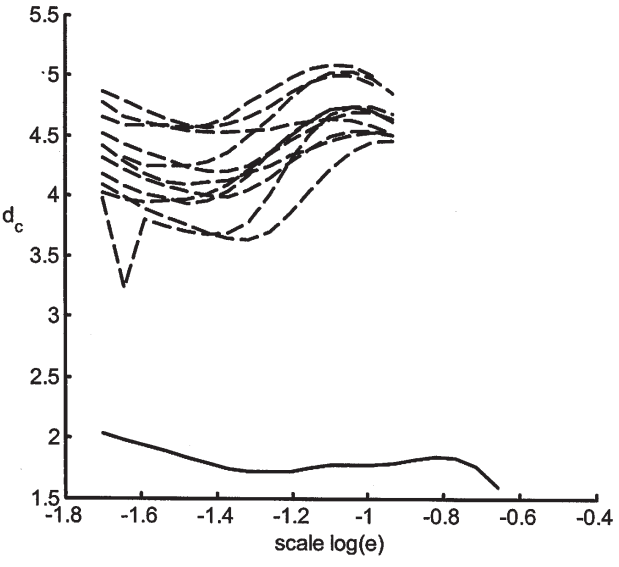

Figure 13—Correlation dimension (d_c) of the corrosion current measurements and their IAAFT surrogate data

close to 1 is associated with highly localized corrosion, such as pitting. S is the standard deviation of the current, and $RMS(I)$ is the root mean square value of the current, I, i.e. $RMS(I) = \{1/N\Sigma_{i=1}^{N}I^2\}^{1/2}$. The corrosion current and potential were embedded with lags of $k_C = 2$ and $k_P = 3$ and embedding dimensions, $m_C = 10$ and $m_P = 7$ respectively. The reconstructed attractors of the two time series are shown in

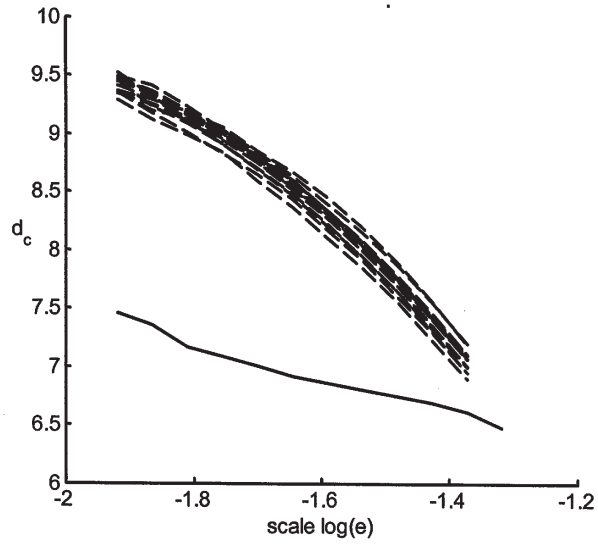

Figure 14—Correlation dimension (d_c) of the corrosion potential measurements and their IAAFT surrogate data

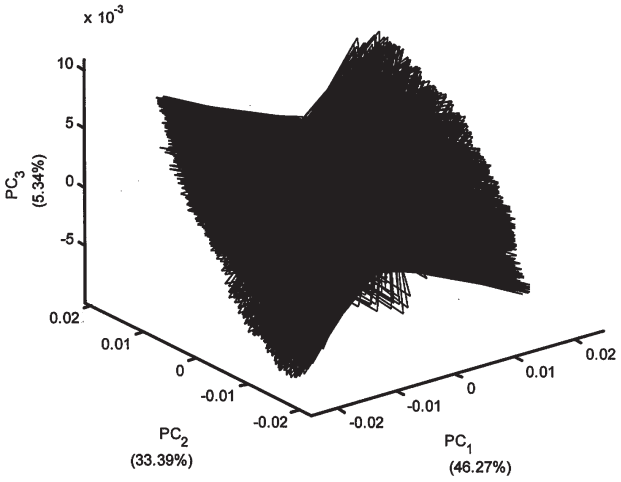

Figure 15—Reconstructed attractor of current data

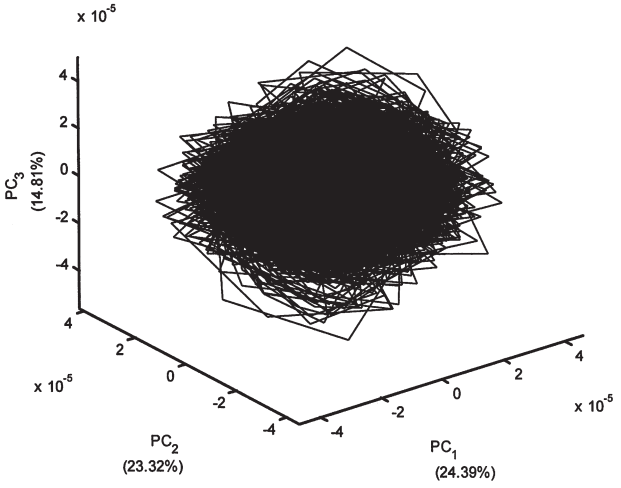

Figure 16—Reconstructed attractor of potential data

Table I

Embedding parameters of various mineral and metallurgical process systems

Process system	Lag	Embed dimension	No. of samples	Comments
Metal concentration in a leach plant	7	4	2282	Nonlinear
Rolling mill speed in a steel plant	14	8	3516	Nonlinear
Pulp density in concentrator plant	13	6	3412	Linear
Mill power on a con-centrator plant	22	9	4341	Nonlinear

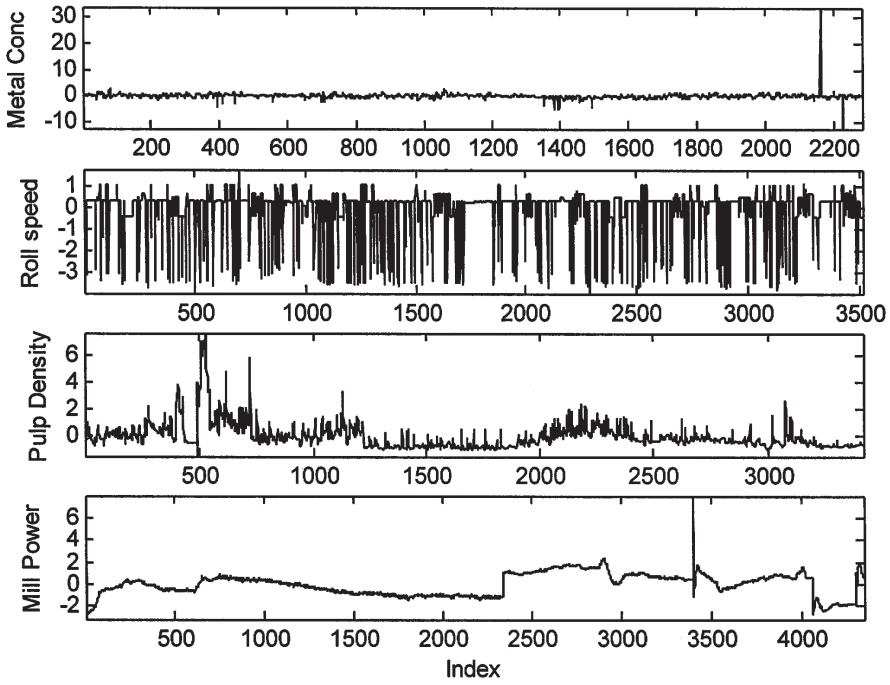

Figure 17—Plant data: metal concentrations in a leach plant, roll speed in a steel mill, pulp density in a base metal concentrator and mill power on a valuable metal concentrator plant (top to bottom respectively). All data have been normalized to zero mean and unit variance. The corresponding surrogate data analysis for each data set is shown in Figures 18–21 respectively

Figures 15–16, as projections in the principal component score space of the embedded data. The variance explained by each principal component is shown in parentheses on each axis.

Fifteen IAAFT surrogate data sets were generated for the time series, which consisted of 3 156 observations each. The correlation dimensions for these surrogates and the original time series are shown in Figures 13–14. Figure 13 shows that, unlike the corrosion potential, the corrosion current had a low dimensionality of approximately 1.5 to 2. The sharp distinction between the correlation dimension curves (d_c) of the original data and their associated surrogates in both cases validate their nonlinear character. Furthermore, the analysis indicates that it should be possible to construct accurate predictive models for the corrosion current, given its low dimensionality (< 2). The same cannot be said of the corrosion potential, which has a relatively high dimensionality. Current studies are under way to investigate the use of the correlation dimension of the time series measurements to distinguish between different corrosion phenomena and to serve as an early warning system for corrosion in equipment.

Case study 4: data from mineral and metallurgical process plants

In the final case study, some data from mineral and

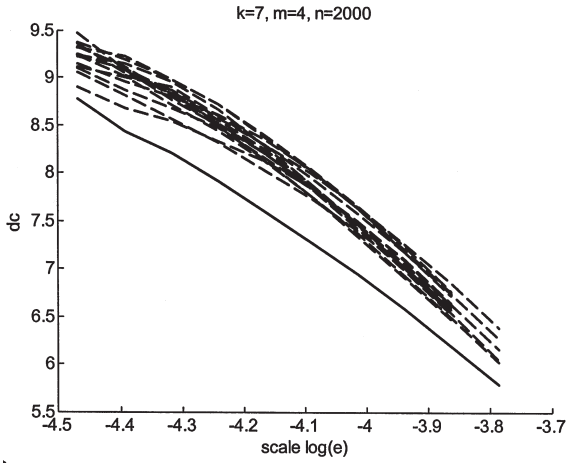

Figure 18—Correlation dimension (d_c) of the valuable metal concentrations on a leach plant and their IAAFT surrogate data

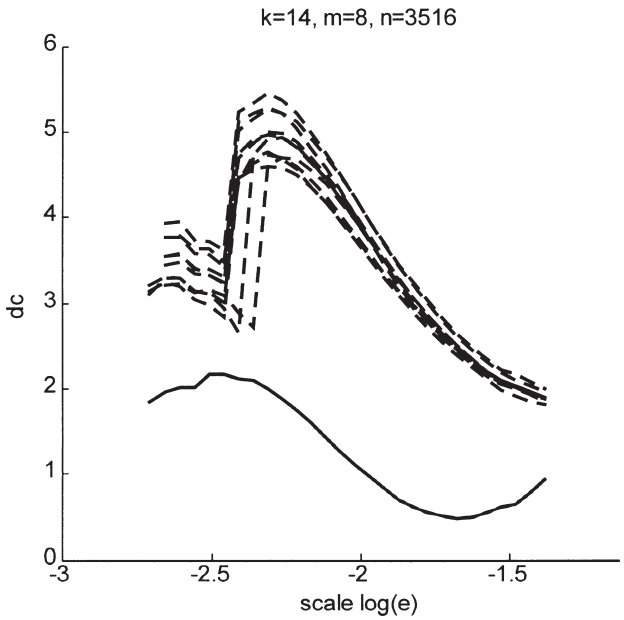

Figure 19—Correlation dimension (d_c) of roll speed observations in a steel mill and their IAAFT surrogate data

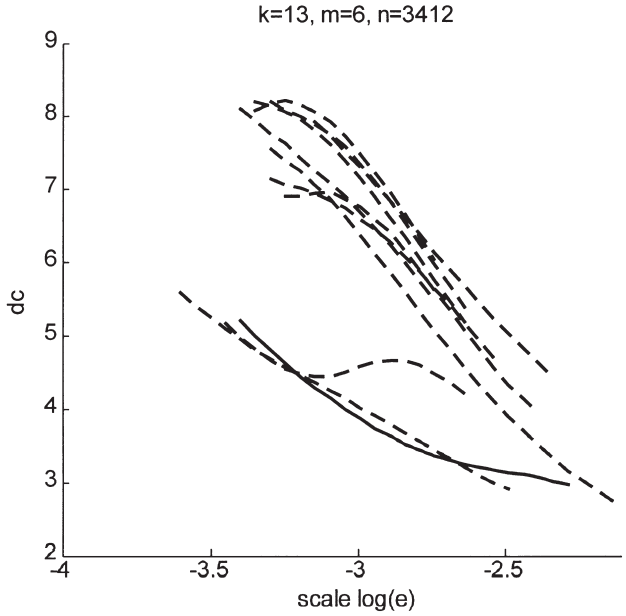

Figure 20—Correlation dimension (d_c) of the pulp densities on a base metal concentrator and their IAAFT surrogate data

metallurgical process plants are investigated. These data are shown in Figure 17 and the results of the surrogate data analyses, similar to the ones in the previous case studies, are shown in Figures 18–21. The embedding parameters of each system are summarized in Table I. In all cases, the lags were estimated from the data based on the average mutual information content of the data, and the embedding dimensions were estimated by means of the false nearest neighbour algorithm. The correlation dimensions in Figures 18–21 were calculated with the IAAFT algorithm mentioned previously. These results suggest that all the

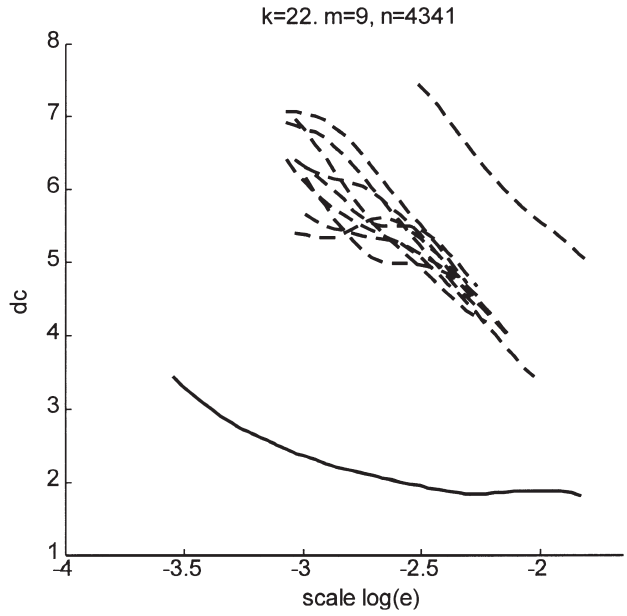

Figure 21—Correlation dimension (d_c) of mill power observations on a concentrator plant and their IAAFT surrogate data

process data are in fact nonlinear, with the exception of the pulp density measurements in the base metal concentrator, which can therefore be identified by a linear model. In all the other systems, the use of more complicated nonlinear models would be justified.

Discussion and conclusions

The classification of dynamic process systems from observed data is an important problem in process engineering, with potential applications ranging from system identification in automated control systems, process condition monitoring systems and the detection of anomalies in plant operations.

To summarize:

- Surrogate data are particularly useful for the screening of data prior to the construction of process models. It is not always easy to determine the degree of determinism or stochasticity of real-world data, and the technique allows the engineer to classify the data prior to building a model.
- With the composite null hypothesis used in this investigation, the results of the Monte Carlo tests should be interpreted with some care. In mineral processing, stationary process operation may be difficult to achieve, and some tests for stationarity should almost certainly be used when the null hypothesis is rejected, before the data can be classified as nonlinear. When iterated amplitude adjusted Fourier transform surrogates are used, the effect of non-Gaussianity in the data should not play a large role in the rejection of the null hypothesis
- Other issues that will influence the robustness of these tests are the embedding parameters (embedding dimension, m, and embedding lag, k). With short noisy time series characteristic of many process systems, it may be particularly difficult to optimize these

parameters, and it may be necessary to test a number of different embeddings before any conclusions about the process dynamics can be made

- Likewise, small data sets may also compromise the analysis. Although the correlation dimension is a pivotal test statistic, which can be used to discriminate between different time series, it requires large data sets for reliable calculation. As a rule of thumb, the Tsonis criterion considers the number of observations (N) required for reliable estimates of the correlation dimension (d_C) to be approximately exponentially related to the correlation dimension, that is $N = 10^{(2+0.4dc)}$. This means that for reliable estimates of high correlation dimension values ($d_C \geq 10$), in the order of 1 000 000 observations are required. Data sets of this size may not always be available, or may add a significant burden in terms of additional computational cost
- Finally, the use of surrogate data methods in conjunction with the correlation dimensions of time series attractors may not be suitable for online implementation in fast-changing process systems, as the high computational requirements make for relatively slow calculation.

References

DOLAN, K.T. and SPANO, M.L. Surrogate for Nonlinear Time Series Analysis. 2001, *Physical Review E,* 64, 046128. (Online journal).

CUTLER, C.D. A review of the theory and estimation of fractal dimension. *Dimension and Models* (ed.) H. Tong, Nonlinear Time Series and Chaos, 1, 1993, World Scientific, Singapore.

DIKS, C. *Nonlinear Time Series Analysis: Methods and Applications*, Nonlinear Time Series and Chaos, 4, 1999, World Scientific, Singapore.

GRASSBERGER, P. and PROCACCIA, I. Characterization of strange attractors. *Physical Review Letters*, 1983a, vol. 50, pp. 346–349.

GRASSBERGER, P. and PROCACCIA, I. Measuring the strangeness of strange attractors. *Physica D*, 1983b, vol. 9, pp. 189–208.

GRASSBERGER, P., SCHREIBER, T., and SCHAFFRATH, C. Nonlinear time sequence analysis. *International Journal of Bifurcation and Chaos*, 1991, vol. 1, no. 3, pp. 521–547.

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*, 1983, vol. 38, 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*, 1984, vol. 39, pp. 1087–1097.

JUDD, K. An improved estimator of dimension and some comments on providing confidence intervals. *Physica D*, 1992, vol. 56, pp. 216–228.

KANTZ, H. and SCHMITZ, A. *Non-Linear Time Series Analysis*. 1997, Cambridge University Press, Cambridge, England.

Kugiumtzis, D. *Surrogate Data Test on Time Series, in Modelling and Forecasting Financial Data, Techniques of Nonlinear Dynamics*, A. Soofi and L. Cao (eds.), 2002, Kluwer Academic Publishers, Chap 12, pp. 267–282.

LAI, Y-C. and LERNER, D. Effective scaling region for computing the correlation dimension from chaotic time series. *Physica D*, 1998, vol. 115, pp. 1–18.

LEGAT, A. and DOLECEK, V. Corrosion monitoring system based on measurement and analysis of electrochemical noise. *Corrosion*, 1995, vol. 51, no. 4, pp. 295–300.

LIN, Y.P., ST MARTIN, E.J., and FRANK, J.R. Monitoring and mitigation of sustained, localized pitting corrosion. *Proceedings of the 8th International Environmental Petroleum Conference*, Nov 6-9, 2001, Houston, TX, USA.

LYNCH, D. T. Chaotic behavior of reaction systems: parallel cubic autocatalators. *Chemical Engineering Science*, 1992, vol. 47, no. 2, pp. 347–355.

MANSFELD, F. and XIOA, H. Electrochemical noise analysis of iron exposed to NaCl solutions of different corrosivity. *Journal of the Electrochemical Society*, 1993, vol. 140, no. 8, pp. 2205–2209.

PALUŠ, M. Detecting nonlinearity in multivariate time series. *Physics Letters A*, 1996, A213, pp. 138–147.

SCHREIBER, T. and SCHMITZ, A. Improved surrogate data for nonlinearity tests. *Physics Review Letters*, 1996, vol. 77, p. 635.

SCHREIBER, T. and SCHMITZ, A. Surrogate time series. *Physica D*, 2000, vol. 142, pp. 346–382.

SMALL, M. and JUDD, K. Correlation dimension: A pivotal statistic for non-constrained realizations of composite hypotheses in surrogate data analysis. *Physica D*, 1998, vol. 129, pp. 386–400.

TAKENS, F. Detecting nonlinearities in stationary time series. *International Journal of Bifurcations and Chaos*, 1993, vol. 3, p. 241.

THEILER, J., EUBANK, E., LONGTIN, A., and GALDRIKIAN, B. Testing for nonlinearities in time series: The method of surrogate data. *Physica D*, 2000, vol. 58, pp. 77–94. ◆