Journal of Applied Statistics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/cjas20

A multivariate descriptor method for change-point detection in nonlinear time series

P. P. Balestrassi [a] , A. P. Paiva [a] , A. C. Zambroni de Souza [a] , J. B. Turrioni [a] & Elmira Popova [b]

[a] Instituto de Engenharia de Produção e Gestão, Federal University of Itajuba, MG, 37500, Brazil

[b] Operations Research and Industrial Engineering Department, The University of Texas at Austin, Austin, TX, 78712, USA

Available online: 04 Oct 2010

To cite this article: P. P. Balestrassi, A. P. Paiva, A. C. Zambroni de Souza, J. B. Turrioni & Elmira Popova (2011): A multivariate descriptor method for change-point detection in nonlinear time series, Journal of Applied Statistics, 38:2, 327-342

To link to this article: http://dx.doi.org/10.1080/02664760903406496

Journal of Applied Statistics
Vol. 38, No. 2, February 2011, 327–342

A multivariate descriptor method for change-point detection in nonlinear time series

P.P. Balestrassi[a], A.P. Paiva[a], A.C. Zambroni de Souza[a], J.B. Turrioni[a] and Elmira Popova[b]*

[a] *Instituto de Engenharia de Produção e Gestão, Federal University of Itajuba, MG 37500, Brazil;*
[b] *Operations Research and Industrial Engineering Department, The University of Texas at Austin, Austin, TX 78712, USA*

(*Received 11 April 2008; final version received 10 September 2009*)

The purpose of this paper is to present a novel method that is applied to detect dynamic changes in nonlinear time series. The method combines a multivariate control chart that monitors the variation of three normalized descriptors – Hjorth's descriptors of activity, mobility and complexity – and is applied to the change-point detection problem of nonlinear time series. The approach is estimated using six simulated nonlinear time series. In addition, a case study of six time series of short-term electricity load consumption was used to illustrate the power of the method.

Keywords: nonlinear time series; Hjorth's descriptors; Hotelling control chart; change point

1. Introduction

This paper proposes a novel method to detect dynamic changes in nonlinear time series. This detection is of fundamental importance for many applications that result in nonlinear time series such as electricity load, electroencephalogram (EEG) signals, econometrics, water consumption, machining and welding signals, etc. The detection of dynamic changes is crucial for system control and is related to the management of special causes of variation that often are indications of process problems.

The proposed methodology is based on combining the Hotelling T^2 control chart and Hjorth's descriptors. This way, a multivariate control chart that monitors special causes of variation on Hjorth's normalized descriptors of activity, mobility and complexity of a time series is applied to detect dynamic change-point problems. In this study, six artificially created nonlinear time series were used for comparison purpose. Also, a case study of six series of electricity load consumptions was used to illustrate the method.

*Corresponding author. Email: elmira@mail.utexas.edu

ISSN 0266-4763 print/ISSN 1360-0532 online

DOI: 10.1080/02664760903406496
http://www.informaworld.com

The following outlines the remaining sections of this paper. Section 2 describes recent literature review on the detection of dynamic changes in nonlinear time series. The multivariate approach considering the Hotelling T^2 control chart and Hjorth's descriptors is also presented in addition to a simplified example that shows the statistical calculations for the proposed method. Section 3 presents the simulation strategy and the results for the overall approach considering the average run length (ARL) measure. Section 4 presents a case study for short-term electricity load problem using the novel framework. Finally, Section 5 states the main conclusions and further research.

2. Background and literature review

2.1 *Detecting changes in nonlinear time series*

Temporal variations in process parameters are often both complex and non-stationary. While nonlinear reconstruction and other approaches can be used to attempt analyses of such data, it is sometimes necessary only to detect substantial changes in the underlying processes. Quick detection of change(s) enables users to control a system dynamically because it allows them to make necessary adjustments in a timely manner. For example, there is evidence that epileptic seizures are preceded by changes in EEG signals. Due to these changes, seizure prediction may be possible. Also, for example, in a sequential manufacturing process, a product unit proceeds through different manufacturing stages. At these different stages, sensors monitor the features of the unit to detect abrupt changes in process variables, as well as forecast their future behavior. The information thus received can be further utilized as preventive maintenance and quality control. These kinds of problems arise in many fields.

The problem of detecting break points for a broad class of non-stationary time series models has been investigated in the literature. It is a classical problem in signal processing that can be used, i.e. for event detection. Several methods involving sequential algorithms, wavelet decomposition, Markov chain Monte Carlo method, Bayesian approach, sequential probability ratio test, cumulative sum (CUSUM) control chart, neural network and many others can be found in the literature [4,5,7,15,17,19–21,25,27,31].

Many previous papers in this area employ a time domain approach to detect changes in the mean and/or variance structure of time series. For example, Inclan [13] and Inclan and Tiao [14] discuss time domain methods for detecting multiple changes in the variances of time series with application to stock returns. Basseville and Nikiforov [2] provide extensive notes on change-point detection methods for correlated data. Online methods for detecting changes in the second-order properties of the process receive less attention. For retrospective change-point detection, Kluppelberg and Mikosch [16] present a test based on the cumulative integrated periodogram. Ombao *et al.* [23] present a parametric approach to online change-point detection for time series based on fitting autoregressive (AR) models to blocks of series and compare them using likelihood-based methods. Although AR models can capture a wide range of process behavior, the results of the methodology are likely to be poor if the structure differs significantly from that of an AR process. Non-parametric methods, such as those based on spectral properties of the process, have been rarely examined.

Generally, the current change-point detection methods in literature tend to be quite sophisticated in nature. Additionally, *a priori* knowledge about the possibilities of the changes and their distributions is often required. This may make the implementation of these methods difficult as an automatic, online change-detection application. Some other drawbacks are: the requirement of several estimation parameters, the need of several monitoring descriptors and the great number of tuning variables. The concerns are increased when multiple time series are used simultaneously.

This work differs from previous methods by considering the idea that changes in the second-order structure of a stochastic process are reflected by changes in the Fourier or wavelet-based spectra, which are the decompositions of variance across frequency and scale, respectively.

By partitioning the series into short blocks (so that the stationarity can be safely assumed), this work develops a spectral-based method for comparing the process behavior across blocks to determine whether a change is occurred. As mentioned previously, an advantage of the spectral-based methods is that they do not assume a particular process structure (such as AR), and thus can be effective for series with non-AR and/or non-moving average (MA) spectra.

2.2 *The multivariate approach*

The main idea of the multivariate approach is illustrated in Figure 1 and the methods are described next. The whole process consists of extracting some features of the nonlinear time series in a certain window and submitting these features – Hjorth's descriptors – to a multivariate control chart. An out-of-control alarm coming from this chart is a signal that the nonlinear time series have changed and a change point was detected.

2.2.1 *Hjorth's descriptors*

Hjorth formulated three parameters capable of characterizing any signal and its derivatives in the frequency and time domains. The Hjorth parameters are called normalized slope descriptors because they can be defined as first and second derivatives. They are named *activity*, *mobility* and *complexity* and are calculated as follows [12]:

$$A = \text{Activity} = m_0 = \sigma_0^2 \tag{1}$$

$$M = \text{Mobility} = \left(\frac{m_2}{m_0}\right)^{1/2} = \frac{\sigma_1}{\sigma_0} \tag{2}$$

$$C = \text{Complexity} = \frac{(m_4/m_2)^{1/2}}{(m_2/m_0)^{1/2}} = \frac{\sigma_2/\sigma_1}{\sigma_1/\sigma_0} \tag{3}$$

Their computation in discrete time involves the variance (σ_0^2) of the nonlinear time series $f(t)$ on the segment to be analyzed, as well as the variance of the first and second derivatives of $f(t)$, σ_1^2 and σ_2^2, respectively. The nonlinear time series $f(t)$ can also be expressed by means of the

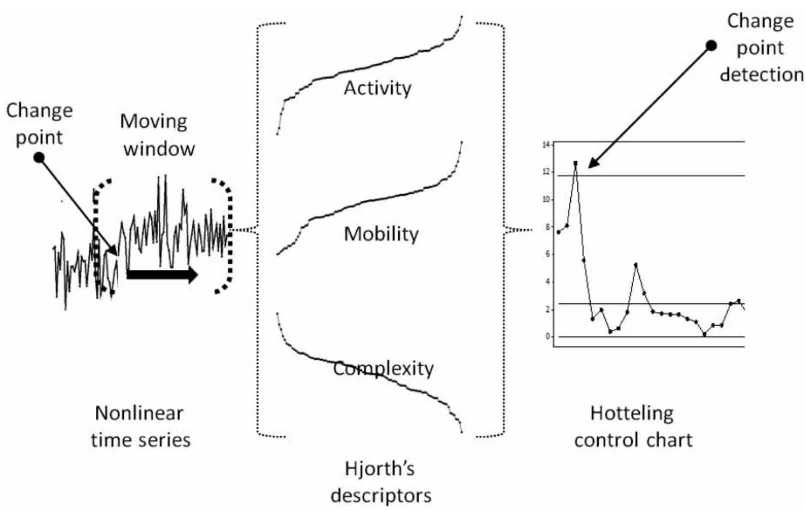

Figure 1. The multivariate approach combining Hjorth's descriptors and Hotelling control chart.

Fourier transform, translated into a function of frequency $F(\omega)$. The phase is excluded by the multiplication of $F(\omega)$ by its conjugate $F^*(\omega)$, giving the power spectrum $S(\omega)$ as a result.

$$F(\omega) \cdot F*(\omega) = S(\omega) \tag{4}$$

The complete frequency description as derived by means of the Fourier transform is always symmetrical with respect to zero frequency. A consequence of this symmetry is that, in a statistical approach to the shape of the frequency distribution, all odd moments will become zero. Thus, there will be no information in the linear average or in the skewness, since these qualities constitute the first and the third moments. The general definition of a moment m of order n is

$$m_n = \int_{-\infty}^{+\infty} \omega^n S(\omega)\,\mathrm{d}\omega \tag{5}$$

The transformation of the parameters between the frequency and time domains is based on the energy equality within the actual epoch, i.e. the total power in the frequency domain is identical to the mean power in the time domain [12]. Figure 2 shows the time domain properties of an arbitrary signal. The moments are calculated as follow:

$$m_0 = \frac{1}{T}\int_{t-T}^{t} f^2(t)\,\mathrm{d}t \tag{6}$$

$$m_2 = \frac{1}{T}\int_{t-T}^{t} \left(\frac{\mathrm{d}f(t)}{\mathrm{d}t}\right)^2 \mathrm{d}t \tag{7}$$

$$m_4 = \frac{1}{T}\int_{t-T}^{t} \left(\frac{\mathrm{d}^2 f(t)}{\mathrm{d}t^2}\right)^2 \mathrm{d}t \tag{8}$$

The first derivatives are obtained using central differentiation as shown in Figure 3. The small broken line may be thought of as giving a reasonable approximation to the required slope (shown by the longer broken line), if h is small enough.

So, we might approximate

$$f'(a) \approx \left(\frac{f(t-h) - f(t+h)}{2h}\right) \tag{9}$$

An approach that has been found to work well for second derivatives involves applying the notion of a central difference three times. It begins with

$$f''(a) \approx \left(\frac{f'(a+(1/2)h) - f'(a-(1/2)h)}{h}\right) \tag{10}$$

Next, it approximates the two derivatives in the numerator of this expression using central differences as follows:

$$f'\left(a+\frac{1}{2}h\right) \approx \frac{f(a+h)-f(a)}{h} \quad \text{and} \quad f'\left(a-\frac{1}{2}h\right) \approx \frac{f(a)-f(a-h)}{h} \tag{11}$$

Combining these three results gives

$$\begin{aligned} f''(a) &\approx \frac{f'(a+(1/2)h) - f'(a-(1/2)h)}{h} \\ &\approx \frac{1}{h}\left\{\left(\frac{f(a+h)-f(a)}{h}\right) - \left(\frac{f(a)-f(a-h)}{h}\right)\right\} \\ &\approx \frac{f(a+h)-2f(a)+f(a-h)}{h^2} \end{aligned} \tag{12}$$

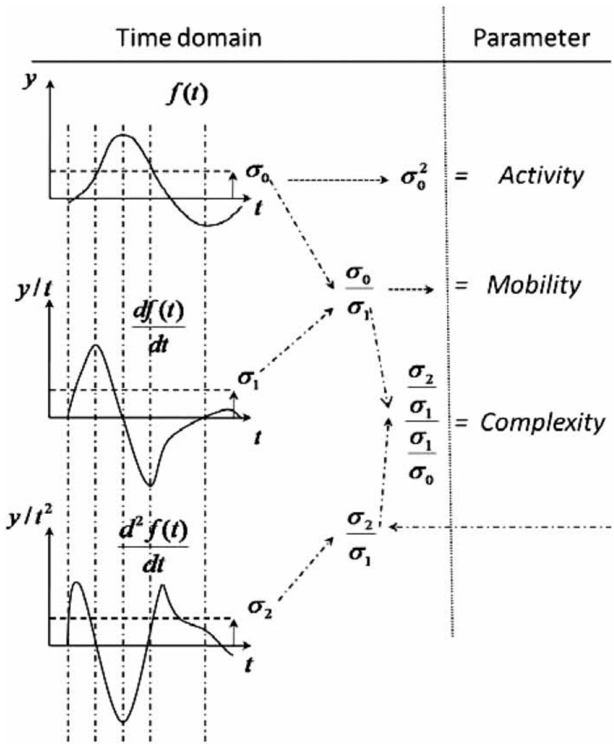

Figure 2. Time domain properties of an arbitrary signal.

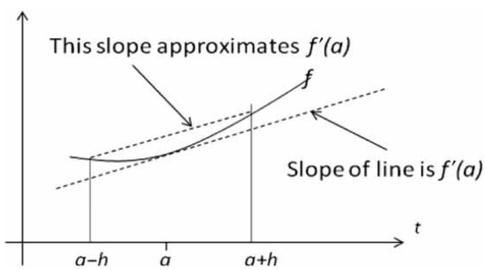

Figure 3. Numerical differentiation using a central difference approximation.

Hjorth's [12] descriptors have the following characteristics:

- Activity: The activity is quantified by means of the amplitude variance, which has the necessary additive property to allow integration of different observations during the epoch into one representative figure. In the frequency domain, the variance of the time function can be conceived as the surface of the power spectrum.
- Mobility: The mobility is defined as the square root of the ratio between the variances of the first derivative and the amplitude. Since these quantities are equally dependent on the mean amplitude, the ratio will be dependent on the curve shape only and in such a way that it measures the relative average slope.
- Complexity: This parameter is dimensionless and derived as the ratio between the mobility of the first derivative of the nonlinear time series and the mobility of the nonlinear time series.

The mobility of the first derivative is obtained in a way analogous to the mobility of the original signal using the variance of the second derivative. Complexity indicates the deviation of the slope and can be seen as a measure of change in frequency of the input signal.

Although Hjorth's descriptors are based on spectral moments, they are also calculated by time variances where the computational cost is more affordable than other methods. Due to that online methods considering this approach can be easily implemented for real applications.

In this study, Hjorth parameters are used to discriminate distinct stages, where some nonlinear properties (seasonality, mean shift, volatility, autocorrelation, etc.) changes over time.

2.2.2 *Multivariate control charts*

Multivariate control charts are designed to incorporate the correlation structure of several variables into a chart when determining the state of statistical control of a process. Hotelling T^2, multivariate CUSUM and multivariate exponentially weighted moving average (EWMA) are examples of such control charts designed to detect changes in the mean vector of several quality characteristics under a constant covariance matrix **S**. The interested reader is referred to Ref. [18], a valuable resource for multivariate control charts.

For notation purpose, arranging the observations as row vectors in an $m \times p$ data matrix, where m is the window size and p is the number of characteristics (in this case $p = 3$, the number of descriptors), gives

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1' \\ \mathbf{x}_2' \\ \vdots \\ \mathbf{x}_m' \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ \vdots & \vdots & \vdots \\ x_{m1} & x_{m2} & x_{m3} \end{bmatrix}$$

The ith individual observation of the three descriptors will be given by the vector

$$\mathbf{x}_i = \begin{bmatrix} X_{i1} \\ X_{i2} \\ X_{i3} \end{bmatrix}$$

The estimated mean vector, whose components are the means of each descriptor, is

$$\bar{\mathbf{x}} = \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \bar{X}_3 \end{bmatrix}$$

where

$$\bar{X}_j = \frac{1}{m} \sum_{i=1}^{m} x_{ij}, \quad j = 1, \ldots, 3$$

and the estimated covariance matrix is

$$\mathbf{S} = \frac{1}{m-1} \sum_{i=1}^{m} (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})' = \begin{bmatrix} S_1^2 & S_{12} & S_{13} \\ \ldots & S_2^2 & S_{23} \\ \ldots & \ldots & S_3^2 \end{bmatrix} \quad (13)$$

To construct a multivariate control chart based on Hotelling T^2-statistic, for observation $\mathbf{x}_i$, one uses the charting statistics

$$T_i^2 = (\mathbf{x}_i - \bar{\mathbf{x}})' \mathbf{S}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}}) \quad (14)$$

The distribution of the T^2-statistic is not widely known, and thus most multivariate control chart practitioners approximate it with a beta distribution to obtain the control chart limits. Specifically,

$$T^2 \approx \frac{(m-1)^2}{m} B\left(\frac{1-\alpha}{2}, \frac{p}{2}, \frac{m-p-1}{2}\right) \tag{15}$$

Taking into account this relationship, the control limit for the T^2 control chart is based on an upper-tail percentile from the central B distribution to obtain the desired in-control ARL and α is the significance level (usually at 1%).

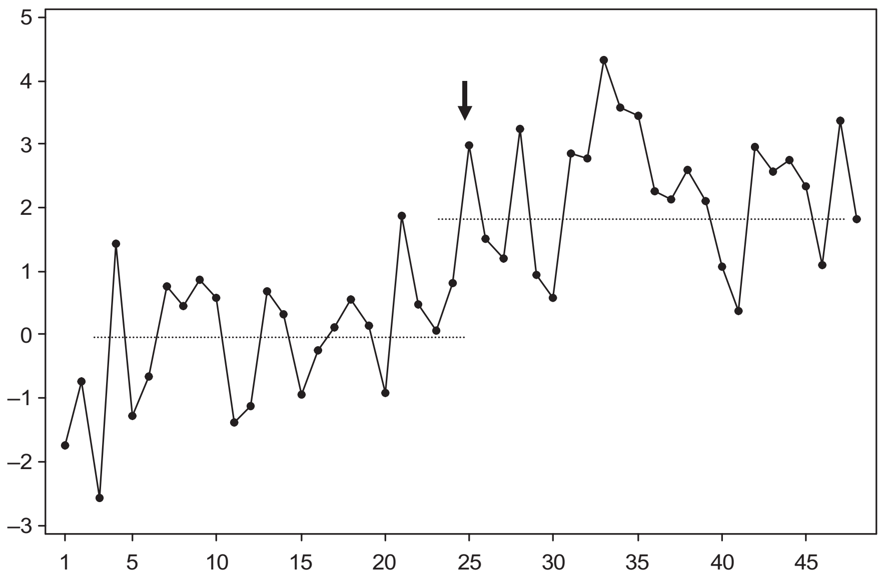

Figure 4. SAR model with two segments of data and a change point on $t = 25$.

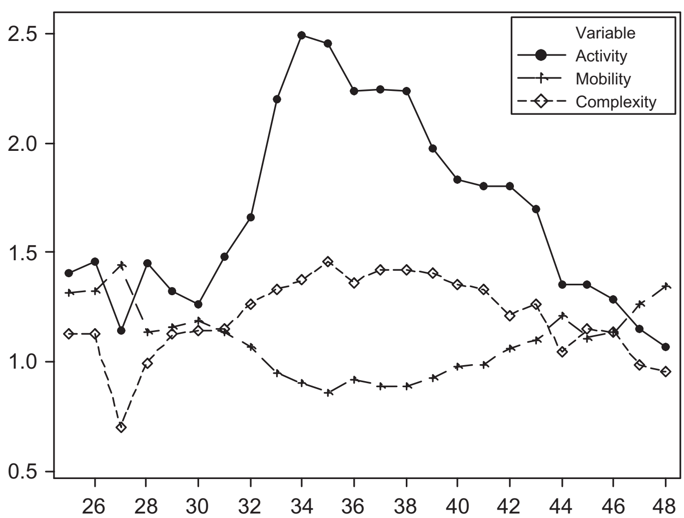

Figure 5. Hjorth's descriptors for the second segment of the SAR model in Figure 1.

The following procedure considering the above-mentioned formulas is used to run the Hotelling T^2 control chart:

(i) Calculate the means for each variable (X_1: activity, X_2: mobility and X_3: complexity).
(ii) Calculate the variances and covariances S_1^2, S_2^2, S_3^2, S_{12}, S_{13}, S_{23} of $\mathbf{X}$.
(iii) Calculate the T^2-statistic.
(iv) Calculate the upper control limit for the Hotelling T^2 control chart.

Table 1. Hjorth's descriptors and T^2 values.

	$f(t)$	$A = \sigma_0^2$	$f'(t)$	σ_1^2	$M = \frac{\sigma_1}{\sigma_0}$	$f''(t)$	σ_2^2	$C = \frac{\sigma_2/\sigma_1}{\sigma_1/\sigma_0}$	T^2
1	−1.74								
2	−0.74		1.01						
3	−2.58		−1.84			−2.85			
4	1.45		4.03			5.87			
5	−1.28		−2.73			−6.75			
6	−0.65		0.63			3.35			
7	0.76		1.41			0.79			
8	0.47		−0.29			−1.70			
9	0.87		0.41			0.70			
10	0.60		−0.28			−0.69			
11	−1.37		−1.97			−1.69			
12	−1.12		0.26			2.22			
13	0.69		1.81			1.55			
14	0.32		−0.37			−2.18			
15	−0.94		−1.26			−0.89			
16	−0.24		0.70			1.96			
17	0.13		0.37			−0.33			
18	0.56		0.43			0.05			
19	0.14		−0.42			−0.85			
20	−0.91		−1.05			−0.63			
21	1.88		2.80			3.85			
22	0.49		−1.40			−4.19			
23	0.07		−0.41			0.98			
24	0.83		0.76			1.17			
25	3.00	1.41	2.17	2.44	1.32	1.42	7.33	1.13	7.88
26	1.51	1.45	−1.50	2.52	1.32	−3.67	7.58	1.13	8.31
27	1.22	1.14	−0.29	2.36	1.44	1.21	6.05	0.69	12.87
28	3.25	1.45	2.03	1.86	1.13	2.32	4.21	0.99	5.56
29	0.94	1.32	−2.31	1.77	1.16	−4.35	4.59	1.12	1.26
30	0.58	1.26	−0.36	1.76	1.18	1.95	4.74	1.14	1.99
31	2.85	1.48	2.27	1.89	1.13	2.63	4.92	1.15	0.39
32	2.78	1.66	−0.07	1.89	1.07	−2.34	5.16	1.26	0.56
33	4.35	2.20	1.57	1.97	0.95	1.64	5.25	1.33	1.80
34	3.59	2.49	−0.76	2.00	0.90	−2.33	5.37	1.37	5.30
35	3.45	2.45	−0.14	1.81	0.86	0.62	5.17	1.46	3.18
36	2.26	2.24	−1.19	1.89	0.92	−1.05	5.09	1.36	1.86
37	2.14	2.24	−0.11	1.76	0.89	1.08	4.95	1.42	1.67
38	2.60	2.24	0.45	1.76	0.89	0.56	4.92	1.42	1.67
39	2.12	1.98	−0.48	1.69	0.93	−0.93	4.79	1.40	1.55
40	1.09	1.83	−1.03	1.73	0.97	−0.55	4.79	1.35	1.42
41	0.37	1.80	−0.72	1.75	0.99	0.31	4.80	1.33	1.10
42	2.97	1.80	2.60	2.03	1.06	3.32	5.25	1.21	0.23
43	2.59	1.69	−0.38	2.03	1.09	−2.99	5.66	1.26	0.65
44	2.76	1.35	0.17	1.97	1.21	0.56	4.99	1.04	0.79
45	2.36	1.35	−0.40	1.66	1.11	−0.58	4.22	1.15	2.29
46	1.12	1.28	−1.24	1.64	1.13	−0.84	4.20	1.13	2.68
47	3.38	1.15	2.27	1.84	1.26	3.51	4.70	0.98	1.71
48	1.83	1.07	−1.55	1.93	1.34	−3.81	5.24	0.95	2.27

(v) Plot the T^2-statistic on the control chart and compare it to the control limits to determine if individual points are out of control.

2.3 *An example*

The following simplified example illustrates the proposed methodology. The example considers a nonlinear sign autoregressive (SAR) time series described by the following model:

$$y_t = \text{sign}(y_{t-12}) + \varepsilon_t \tag{16}$$

where $\text{sign}(x) = 1, 0, -1$ if $x > 0$, $x = 0$ and $x < 0$, respectively. The error term ε_t follows a normal distribution $N(0,1)$.

This nonlinear time series is divided into two segments. The first segment with 24 points represents the pure SAR model. The second segment considers a change point where the mean is increased by two standard deviations from the mean. Figure 4 shows these segments.

Hjorth's descriptors A, M and C are computed using Equations (1)–(3) for the points 24–48. The idea of the sliding window is used to calculate all the point statistics. This means that all the statistics for $t = 24$ are obtained considering the points 1–24. The statistics for $t = 25$ considers points 2–25 and so on. The correlations between Hjorth's descriptors are considered significant: $r_{\text{AM}} = -0.901$, $r_{\text{AC}} = 0.868$ and $r_{\text{MC}} = -0.923$. Figure 5 shows Hjorth's descriptors for the second segment. One can reproduce the proposed methodology by computing the results in Table 1.

The Hotelling T^2 control chart is shown in Figure 6. The points are calculated using Equations (13)–(15). In this sample, there are 24 observations, so $m = 24$. The vector of sample mean is

$$\bar{\mathbf{x}} = \begin{bmatrix} 1.680 \\ 1.094 \\ 1.199 \end{bmatrix}$$

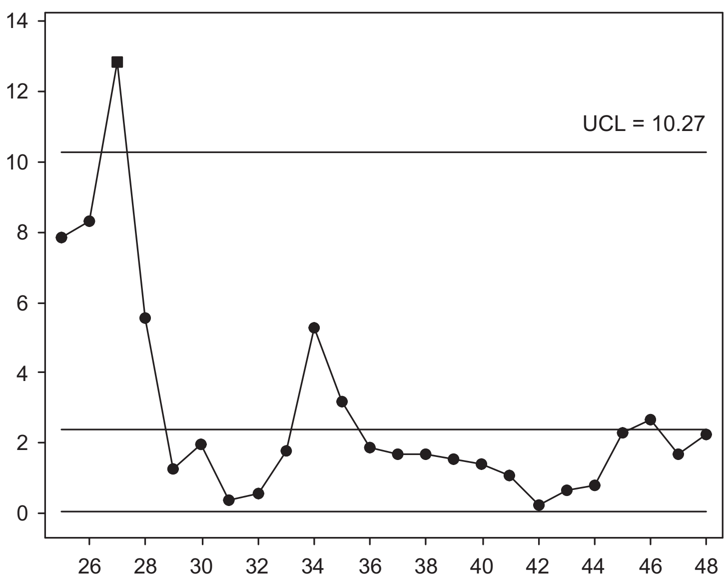

Figure 6. Hotelling T^2 control chart for Hjorth's descriptors.

and the sample covariance matrix of $\mathbf{X}$ is

$$\mathbf{S} = \begin{bmatrix} 0.193 & -0.065 & 0.071 \\ 0.065 & 0.027 & -0.028 \\ 0.071 & -0.028 & 0.035 \end{bmatrix}$$

The T^2 calculation for the point $\mathbf{x}_{27} = \begin{bmatrix} 1.14 \\ 1.44 \\ 0.69 \end{bmatrix}$, for example, gives $T^2 = 12.87$.

The upper control limit using beta distribution gives

$$T^2 \approx \frac{(24-1)^2}{24} B\left(1 - \frac{0.01}{2}, \frac{3}{2}, \frac{24-3-1}{2}\right) \approx \frac{23^2}{24} B(0.995, 1.5, 10) \approx 10.27$$

An out-of-control point is found on the control chart (point 27) based on the most popular criteria of T^2 falling over three standard deviations from the mean. In a simplified manner, this means that the change-point detection is obtained with only three points after the new regime.

3. Simulation results

3.1 *Nonlinear time series: the data set*

Linear time series methods have been widely used to represent processes for the past two decades. Recently, however, there has been increasing interest in extending the classical framework of Box and Jenkins [3] to incorporate non-standard properties, such as nonlinearity, non-Gaussianity and heterogeneity. A great number of nonlinear models have been developed, such as the bilinear model [9], the threshold autoregressive (TAR) model [29], the state-dependent model [26], the Markov switching model [11], the functional-coefficient AR model [6] and many others. Although the properties of these models tend to overlap somewhat, each is able to capture a wide variety of nonlinear behavior. This kind of time series is even more complex due to features like high frequency, daily and weekly seasonality, calendar effect on weekend and holidays, high volatility and presence of outliers.

Consider a univariate time series x_t observed at equally spaced time points. We denote the observations by $\{x_t | t = 1, \ldots, T\}$, where T is the sample size. A purely stochastic time series x_t is said to be linear if it can be written as

$$x_t = \mu + \sum_{i=0}^{\infty} \psi_i a_{t-i} \tag{17}$$

where μ is a constant, ψ_i are real numbers with $\psi_0 = 1$ and a_t is the sequence of independent and identically distributed (i.i.d.) random variables with a well-defined distribution function. We assume that the distribution of a_t is continuous and $E(a_t) = 0$. In many cases, we further assume that $\mathrm{Var}(a_t) = \sigma_a^2$ or, even stronger, that a_t is Gaussian. If $\sigma_a^2 \sum_{i=1}^{\infty} \psi_i^2 < \infty$, then x_t is weakly stationary (i.e. the first two moments of x_t are time-invariant). The autoregressive moving average (ARMA) process is linear because it has an MA representation in the above-mentioned equation. Any stochastic process that does not satisfy the condition of this equation is said to be nonlinear (see [30] for more details). For previous and more general surveys on nonlinear time series models, one is referred to Refs [8,10].

Table 2 shows the collection of nonlinear time series implemented and simulated for the present study. In each case, $\varepsilon_t : N(0, 1)$ is assumed to be i.i.d. These six time series models are chosen

Table 2. A collection of nonlinear time series models.

Model	Equation
SAR	$y_t = \text{sign}(y_{t-12}) + \varepsilon_t$, where $\text{sign}(x) = 1, 0, -1$, if $x > 0, x = 0, x < 0$, respectively
Bilinear (BL)	$y_t = 0.3y_{t-1} - 0.4y_{t-24} + 0.6y_{t-1}\varepsilon_{t-1} + \varepsilon_t$
TAR	$y_t = 0.7y_{t-1} + \varepsilon_t$ for $\lvert y_{t-1}\rvert \leq 1 = -0.4y_{t-12} - \varepsilon_t$ for $\lvert y_{t-1}\rvert > 1$
Nonlinear autoregressive (NAR)	$y_t = \dfrac{0.6y_{t-1}}{\lvert y_{t-24}\rvert + 3} + \varepsilon_t$
NMA	$y_t = \varepsilon_t - 0.5\varepsilon_{t-24} + 0.1\varepsilon_{t-2} + 0.3\varepsilon_{t-1}\varepsilon_{t-2} - 0.3\varepsilon_{t-2}^2$
Smooth transition autoregressive (STAR)	$y_t = 0.4y_{t-1} + 0.5y_{t-24} + (0.25 - 0.81y_{t-1} + 0.7y_{t-24})[1 + \exp(-10y_{t-1})]^{-1} + \varepsilon_t$

Source: Zhang *et al.* [32].
Note: The seasonality is represented by the number 12 (or 24) in the equations.

to represent a variety of problems that have different characteristics. For example, some of the series have pure AR or pure MA correlation structures, while others have mixed AR and MA components. Similar models (with different lags) were explored in Zhang *et al.* [32].

3.2 *Results*

The main results of the simulation study are presented in Table 3 where the ARL is computed for the six nonlinear time series. The ARL shows the number of samples needed to detect a process shift (a change point) in the time series. Two kinds of shifts are considered here: the mean shift and

Table 3. ARL for the nonlinear time series.

			Nonlinear time series					
Mean	Standard deviation	Lag	SAR	BL	TAR	NAR	NMA	STAR
μ	σ	12	360	362	367	385	195	356
μ	σ	24	370	413	370	378	180	409
μ	1.5σ	12	180	170	174	166	180	167
μ	1.5σ	24	190	180	186	206	140	192
μ	3σ	12	4	4	3	2	2	5
μ	3σ	24	6	9	6	4	6	6
$\mu + 1\sigma$	σ	12	33	43	35	40	34	28
$\mu + 1\sigma$	σ	24	43	53	36	40	37	47
$\mu + 1\sigma$	1.5σ	12	15	19	15	27	17	17
$\mu + 1\sigma$	1.5σ	24	25	23	16	28	29	18
$\mu + 1\sigma$	3σ	12	4	2	3	2	4	3
$\mu + 1\sigma$	3σ	24	6	2	3	4	5	3
$\mu + 2\sigma$	σ	12	4	3	4	1	4	4
$\mu + 2\sigma$	σ	24	6	4	7	5	10	7
$\mu + 2\sigma$	1.5σ	12	4	4	2	2	3	4
$\mu + 2\sigma$	1.5σ	24	6	7	4	6	4	6
$\mu + 2\sigma$	3σ	12	4	4	3	5	2	5
$\mu + 2\sigma$	3σ	24	6	5	6	6	6	6
$\mu + 3\sigma$	σ	12	2	2	2	2	2	3
$\mu + 3\sigma$	σ	24	4	4	5	3	3	3
$\mu + 3\sigma$	1.5σ	12	2	1	2	4	3	4
$\mu + 3\sigma$	1.5σ	24	3	2	2	2	2	3
$\mu + 3\sigma$	3σ	12	1	2	1	1	2	1
$\mu + 3\sigma$	3σ	24	2	2	2	2	2	2

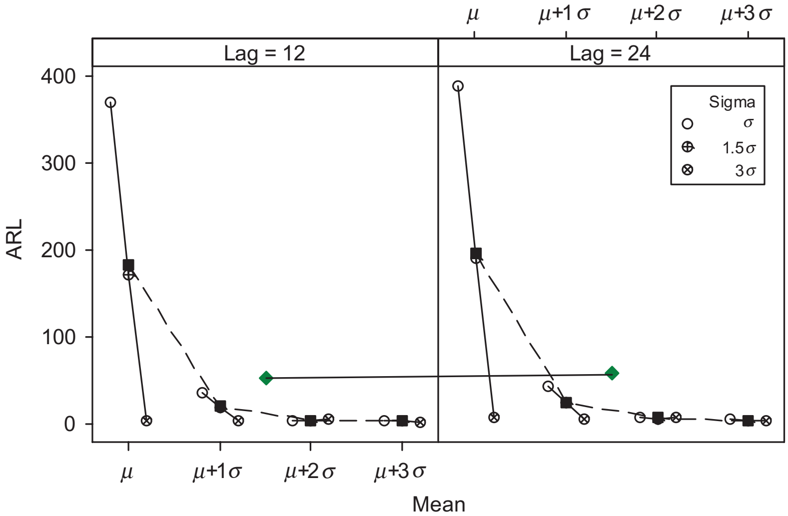

Figure 7. ARL as a function of the factors mean, sigma and lag in a multivariate chart.

the standard deviation shift. The control limits for the Hotelling T^2 control chart were defined by the in-control sample (the first 24 points). The results were obtained considering 100 simulations for each time series. A user-friendly Statistica [28] macro was developed to run the simulation. For each run, a sliding window of 24 points was processed. This way, the first control chart used the in-control points from t_1 to t_{24}. The second control chart added one point (t_{25}) under a new regime and removed the point at t_1. Keeping with this method, the second control chart processes the points from t_2 to t_{25}, the third chart processes the points from t_3 to t_{26} and so on. This process is repeated until an out-of-control point is obtained. When an outlier is found after introducing the new regime at t_{ARL}, the ARL value is obtained subtracting $t_{\mathrm{ARL}} - 24$. High values of ARL indicate when the time series is in control and low values of ARL when the time series change.

The obtained ARLs in Table 3 for all the time series are comparable to traditional results related to an Xbar control chart. ARL for large shifts (both in mean and/or in variance) are very low. This means that the change-point detection is obtained with just a few points after the introduction of a new regime. ARLs without shift are very large, meaning that false alarms are rarely obtained. A paired t-test comparing ARLs for the lags 12 and 24 results in a $p < 0.05$. This means that change point in nonlinear time series with small seasonality is detected faster than with large seasonality.

For the nonlinear moving average (NMA) time series, the correspondent ARLs of false alarms (without shift) are statistically different when compared with the other time series. When compared with traditional control charts the results are satisfactory. This time series, with only NMA realizations, represents the worst-case scenario for the simulation results because false alarms were obtained with fewer points. This is related to the high volatility of this time series.

Figure 7 shows a multivariate chart for the results in Table 3, considering averaged values for all the time series. It can be observed that the mean shift and sigma shift are significant factors for the ARL response. Lag is not statistically significant.

4. Case study: short-term electricity load

Companies that trade in deregulated electricity markets in the USA, Brazil, England and most other countries use time series to predict demand of electricity as part of the information necessary to set buying and selling contracts. If they are production or trade companies, their interests lie particularly in optimizing the energy load and prices implemented for their customers.

This case study deals with the modeling and forecasting of an important variable that affects achieving a good portfolio of electricity contracts: the electricity load (or consumption). Knowledge of its future value and its variation is essential to calculate the portfolio risk and return. Although it is possible to contract a certain amount of energy with no allowed variation, for a consumer it represents high risk. In the industrial process, there are a lot of unpredictable factors that can cause an increase or decrease of energy consumption. If an electricity producer can accept and manage this risk, the producer will have more opportunities in the market. Thus, it is important to manage these risks.

As considered by Angelus [1] and Mount [22], forecasting electricity prices and loads is an arduous task due to a variant of factors such as electricity supply and demand, number of generators, transmission and distribution constraints, etc. For deregulated electricity markets, a great volume of electricity is bought and sold in the spot market (as well as in the bilateral market) among agents at different geographical regions. A simple unpredictable climate change can add volatility in demand and, therefore, volatility in prices. Pai and Hong [24] also state that electrical load forecasting is complex to predict due to nonlinearity of its influenced factors and that the most important factor in regional or national power system strategy management is accurate load forecasting of future demand.

This case study uses the present framework to change the point detection of electricity load of six industrial consumers in Brazil using historical hourly time series. The change point here represents the time when the consumers go to the spot market and change the time series' properties. This is also related to the consumer's fidelity to a previously established contract of buying and selling electricity.

Figure 8 represents the time series of the hourly electricity load for each of the six industrial consumers. Procedures to improve data collection were enforced to treat missing data, outliers, typos, seasonality and errors. Week seasonality was observed in all time series of lag 168. The available data set comprises 3 years of hourly electricity load from Duke Energy, a production company that operates in Brazil.

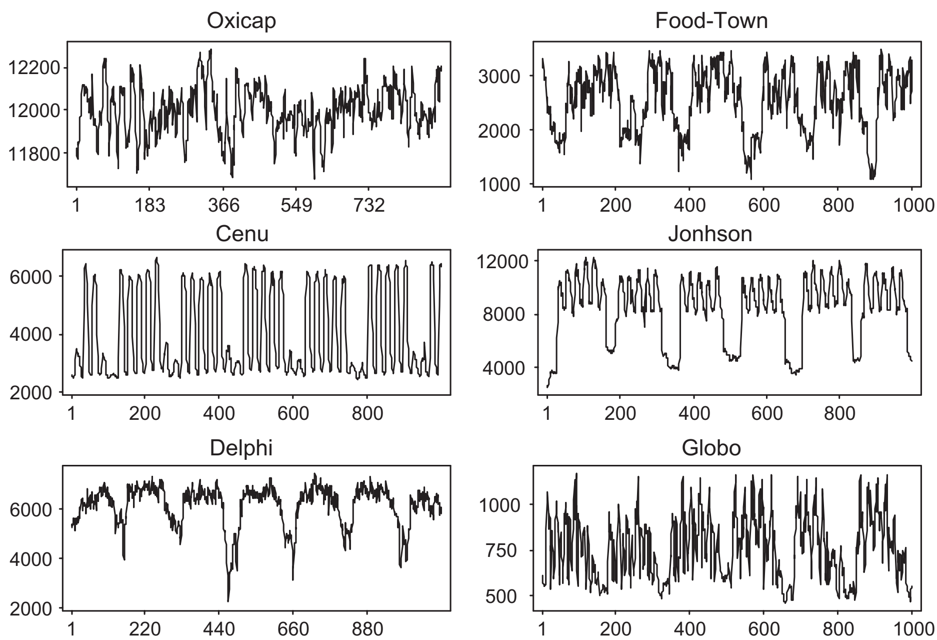

Figure 8. Hourly based time series (in Watts) of six industrial consumers.

Table 4. Change-point detection results for the six industrial consumers.

Series characteristics	Electricity load					
	Cenu	Food-Town	Oxicap	Delphi	Johnson	Globo
Seasonality	168	24	24	168	24	24
Mean	2554.3	11,999	3984.2	8402.6	6139	759.2
SE mean	18.1	3.69	46.3	78.2	24.9	5.41
Standard deviation	572.7	112	1462.1	2471.9	827.1	171.03
Skewness	−0.44	−0.16	0.57	−0.71	−1.42	0.23
Kurtosis	−0.7	−0.12	−1.46	−0.7	1.95	−0.8
Spot market trading	Yes	Yes	Yes	No	No	Yes
Number of trading	12	23	18			18
Results						
False alarm (ARL_0)	–	–	–	1847	–	–
Change-point detection	52	12	11	–	–	8
Hjorth's descriptors	*A*, *M*	*M*	*C*	*C*	–	*C*

One problem that comes up when dealing with multiple time series is their relation to one another. It is worth mentioning that for this entire set of electricity load, the behavior of the six time series cannot be considered as a multivariate process. The correlation between variables is not significant in any pair of the six time series. Also, some factors, such as missing data, different number of samples and different intervals, make this kind of approach less likely to occur. Accordingly, each series needs to be independently addressed.

Table 4 presents the results for the six industrial consumers. The time series characteristics of lag seasonality, mean, standard error of the mean, standard deviation, skewness and kurtosis are computed. Historical data and consumer's information about trading in the spot market is also provided.

It was previously known that the consumers Delphi and Johnson did not trade any electricity in the spot market. The detection of change point using the multivariate approach means only a false alarm because it does not present any real change point. Delphi presents only one change point after 1847 points mainly due to the *complexity* component of the Hjorth descriptor. This result is not significant considering that the nonlinear time series contains the highest volatility among the six time series. In a traditional control chart, this result would be around 370 points.

For the industrial consumers Cenu, Food-Town, Oxicap and Globo, the number of points to detect the change point is computed, and the results are inferior to the lag seasonality. The seasonality is computed using a fast Fourier transform algorithm. The components of Hjorth descriptors are obtained, but a pattern is not found for the six time series. The number of points related to the change-point detection is averaged once the aforementioned have several change points in the historical data (corresponding to the number trading in the spot market). The method is able to recognize the dynamic changes in less than half of the period of seasonality for the time series, which is a satisfactory interval for the hourly based time series.

5. Conclusions and further research

Throughout this study, a novel multivariate method based on Hjorth's descriptors of activity, mobility and complexity was applied to detect dynamic changes in nonlinear time series. A multivariate control chart monitoring special causes of variation on the normalized descriptors was applied to a change-point detection problem on six nonlinear time series. The simulated results showed that the method was able to detect the change point adequately and with a satisfactory ARL.

It is then correct to conclude that the proposed approach was quite appropriate for detecting the change point of the short-term electricity load for the six industrial consumers.

As for further research, the authors are developing a pattern recognition algorithm to an online control chart monitoring. In this new study, several run tests will be monitored in real time using neural networks. This mechanism is necessary for the change-point detection of several time series like EEG signals and machining processes. Further research will be the evaluation of the present approach in face of the Kolmogorov–Sinai entropy.

Acknowledgements

This work was supported by CAPES, a Brazilian governmental agency, under Grant 023/05.

References

[1] A. Angelus, *Electricity price forecasting in deregulated markets*, Electr. J. 4 (2001), pp. 32–41.
[2] M. Basseville and I. Nikiforov, *Detection of Abrupt Changes: Theory and Application*, Prentice-Hall, New Jersey, 2000.
[3] G.E.P. Box and G.M. Jenkins, *Time Series Analysis: Forecasting and Control*, Revised ed., Holden-Day, San Francisco, CA, 1976.
[4] A. Burrell and T. Papantoni, *Sequential algorithms for detecting changes in acting stochastic processes and on-line learning of their operational parameters*, Proceedings of the 15th International Conference on Pattern Recognition, Barcelona, Spain, Vol. 2, 2000, pp. 656–659.
[5] F. Castanie and F. Denjean, *Mean value jump detection using wavelet decomposition*, Proceedings of IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Victoria, Canada, 1992, pp. 181–184.
[6] R. Chen and R.S. Tsay, *Functional-coefficient autoregressive models*, J. Amer. Statist. Assoc. 88 (1993), pp. 298–308.
[7] H. Cho and P. Fryzlewicz, *Multiscale breakpoint detection in piecewise stationary AR models*, Proceedings of IASC 2008, Yokohama, Japan, 2008.
[8] J.G. de Gooijer, *On threshold moving-average models*, J. Time Ser. Anal. 19 (1998), pp. 1–18.
[9] C.W.J. Granger and A.P. Anderson, *An Introduction to Bilinear Time Series Models*, Vandenhoeck & Ruprecht, Göttingen, 1978.
[10] C.W.J. Granger and T. Terasvirta, *Modeling Nonlinear Economic Relationship*, Oxford University Press, New York, 1993.
[11] J.D. Hamilton, *A new approach to the economic analysis of nonstationary time series and the business cycle*, Econometrica 57 (1989), pp. 357–384.
[12] B. Hjorth, *EEG analysis based on time domain properties*, Electroencephalogr. Clin. Neurophysiol. 29 (1970). pp. 2–10.
[13] C. Inclan, *Detection of multiple changes of variance using posterior odds*, J. Bus. Econ. Stat. 11 (1993), pp. 289–300.
[14] C. Inclan and G. Tiao, *Use of cumulative sums of squares for retrospective detection of changes of variance*, J. Amer. Statist. Assoc. 89 (2004), pp. 913–923.
[15] V. Karathanassi, C. Iossifidis, and D. Rokos, *Application of machine vision techniques in the quality control of pharmaceutical solutions*, Comput. Ind. 32 (1996), pp. 169–179.
[16] C. Kluppelberg and T. Mikosch, *Gaussian limit fields for the integrated periodogram*, Ann. Appl. Probab. 6 (1996), pp. 969–991.
[17] M. Lavielle and E. Lebarbier, *An application of MCMC methods for the multiple change-points problem*, Signal Process. 81 (2001), pp. 39–53.
[18] C.A. Lowry and D.C. Montgomery, *A review of multivariate control charts*, IIE Trans. 27 (1998), pp. 800–810.
[19] S. MacDougall, A.K. Nandi, and R. Chapman, *Multiresolution and hybrid Bayesian algorithms for automatic detection of change points*, Proc. IEEE Vis. Image Signal Process. 145(4) (1998), pp. 280–286.
[20] D.P. Malladi and J.L. Speyer, *A generalized Shiryayev sequential probability ratio test for change detection and isolation*, IEEE Trans. Automat. Control 44 (1999), pp. 1522–1534.
[21] V.M. Morgenstern, B.R. Upadhyaya, and M. Benedetti, *Signal anomaly detection using modified Cusum method*, Proceedings of the 27th IEEE Conference on Decision and Control, Texas, USA, Vol. 3, 1988, pp. 2340–2341.
[22] T. Mount, *Market power and price volatility in restructured markets for electricity*, Decis. Support Syst. 30 (2001), pp. 311–325.
[23] H. Ombao, J. Heo, and D. Stoffer, *Statistical analysis of seismic signals: An almost real-time approach*, Time Series Analysis and Applications to Geophysical Systems: IMA Series, Vol. 139, Wiley, New York, 2004, pp. 53–72.
[24] P.F. Pai and W.C. Hong, *Support vector machines with simulated annealing algorithms in electricity load forecasting*, Energy Convers. Manage. 46 (2005), pp. 2669–2688.

[25] D.B. Percival and A.T. Walden, *Wavelet Methods for Time Series Analysis*, Cambridge University Press, Cambridge, 2000.

[26] M.B. Priestley, *State-dependent models: A general approach to nonlinear time series analysis*, J. Time Ser. Anal. 1 (1980), pp. 47–71.

[27] N.D. Ramirez-Beltran and J.A. Montes, Neural networks for on-line parameter change detections in time series models, Comput. Ind. Eng. 33 (1997), pp. 337–340.

[28] StatSoft, Inc., *Statistica (data analysis software system)*, version 7.1, 2005; software available at www.statsoft.com.

[29] H. Tong, On a threshold model, in *Pattern Recognition and Signal Processing*, C.H. Chen, ed., Sijhoff & Noordhoff, Amsterdam, 1978.

[30] R. Tsay, *Analysis of Financial Time Series*, 2nd ed., Wiley-Interscience, New Jersey, 2005.

[31] N. Ye, *Confidence assessment of quality prediction from process measurement in sequential manufacturing processes*, IEEE Trans. Electron. Packag. Manuf. 23 (2000), pp. 177–184.

[32] G.P. Zhang, B.E. Patuwo, and M.Y. Hu, *A simulation study of artificial neural networks for nonlinear time-series forecasting*, Comput. Oper. Res. 28 (2001), pp. 381–396.