

Detecção de mudança de nível em séries temporais não lineares usando Descritores de Hjorth

Gabriela da Fonseca de Amorim^{a*}, Pedro Paulo Balestrassi^a,
 Anderson Paulo de Paiva^a, Mariângela de Oliveira-Abans^a

^aUniversidade Federal de Itajubá, Itajubá, MG, Brasil, amorimgabi@yahoo.com.br

Resumo

O propósito deste artigo é apresentar um método de detecção de mudança de nível na dinâmica de séries temporais não lineares que consiste no uso de um gráfico de Controle Multivariado T^2 de Hotelling monitorando a variação de três descritores normalizados: os Descritores de Hjorth de atividade, mobilidade e complexidade. Esta abordagem foi aplicada em diferentes séries temporais não lineares criadas artificialmente e é ilustrada neste artigo por um exemplo detalhado. Também foi feito um estudo de caso com seis séries reais do consumo de energia elétrica no meio industrial, confirmando a eficácia do método.

Palavras-chave

Séries temporais não lineares. Descritores de Hjorth. Gráficos de Controle Multivariado T^2 de Hotelling. Mudança de nível.

1. Introdução

Este artigo propõe um método para detectar mudanças dinâmicas em séries temporais não lineares tais como carga elétrica, sinais de eletroencefalograma (EEG), econometria, consumo de água, sinais de usinagem e soldagem, abalos sísmicos etc. Uma série temporal representa uma classe de fenômenos cujo processo observacional, e consequente quantificação numérica, gera uma sequência de dados distribuídos no tempo (Souza, 1981). Quando a equação característica ou o modelo de regressão de uma série temporal pode ser escrito em função de parâmetros lineares (mesmo após transformação), essa série é dita linear. Caso não existam transformações que linearizem os parâmetros, essa série é denominada não linear e, para esse tipo de série, os métodos de controle e monitoramento tradicionais são normalmente complexos.

Uma série temporal que possua a média e a variância constantes ao longo do tempo é dita estacionária ou convergente, caso contrário é denominada não estacionária ou divergente. A detecção dinâmica de mudanças em séries temporais não lineares, sendo elas estacionárias ou não, é crucial para o controle em tempo real de sistemas que estão

sendo monitorados e está diretamente relacionada à identificação de causas especiais de variação, que são normalmente indicações de problemas. As causas especiais de variação ocorrem devido a uma situação particular (não aleatória) que faz com que o processo se comporte de um modo diferente do usual e devem ser, de modo geral, localizadas e eliminadas. Além disso, certas medidas devem ser tomadas para evitar a sua reincidência.

A metodologia proposta é baseada na combinação de Gráficos de Controle Multivariado T^2 de Hotelling e Descritores de Hjorth. A detecção dinâmica dos pontos de mudança é realizada a partir da aplicação de um gráfico de controle multivariado nos Descritores de Hjorth normalizados de atividade, mobilidade e complexidade, a fim de monitorar possíveis variações na série temporal. Esse estudo foi aplicado em diferentes tipos de séries temporais não lineares criadas artificialmente e, além disso, foi feito um estudo de caso em seis séries temporais reais do consumo de energia elétrica por indústrias brasileiras.

Este artigo está dividido em cinco seções, a iniciar pela introdução ao tema. A seção 2 descreve

uma revisão da literatura sobre detecção dinâmica de mudanças em séries temporais não lineares, além de apresentar os Descritores de Hjorth e o Gráfico de Controle T^2 de Hotelling, a partir de uma abordagem multivariada e de um exemplo simplificado aplicado em uma série criada artificialmente.

A seção 3 apresenta a estratégia de simulação, os tipos de séries com diferentes características nas quais o método foi aplicado com sucesso e seus resultados para uma abordagem global considerando a medida ARL (*Average Run Length*). Já a seção 4 mostra um estudo de caso da nova metodologia para problemas reais de consumo de carga elétrica e, finalmente, na seção 5 são reveladas as principais conclusões e sugeridas pesquisas futuras envolvendo a metodologia em questão.

2. Revisão da literatura

2.1. Detecção de mudanças em séries temporais não lineares

No caso de uma série temporal não linear, a análise de dados com fins de monitoramento e/ou detecção de mudanças pode ser feita por meio de reconstrução não linear ou outras abordagens excessivamente longas, *off-line* ou interativas difundidas na literatura, mas às vezes é necessário detectar rapidamente as mudanças nos processos para permitir ao usuário o controle dinâmico do sistema, possibilitando a realização de ajustes em tempo hábil.

Problemas como esses surgem nas mais diversas situações. Por exemplo, Cabrerizo et al. (2012) destacam o fato de que crises epiléticas são precedidas por alterações nos sinais de eletroencefalograma (EEG), isso significa que poderiam ser previstas. Existe também o processo de fabricação sequencial no qual um produto passa por vários estágios de produção e sensores monitoram seus recursos para identificar mudanças nas variáveis do processo e prever seu comportamento futuro; as informações dos sensores poderiam ser usadas na manutenção preventiva e no controle de qualidade. Outro exemplo, ainda, é abordado por Mariani et al. (2011) e consiste na caracterização de fases durante a alternância de padrões no sono.

A detecção de pontos de mudança em modelos de séries temporais não lineares e não estacionárias é um tipo clássico de problema relacionado ao processamento de sinais para a detecção de eventos que vem sendo estudado na literatura. Com essa finalidade, existem vários métodos envolvendo Algoritmos Sequenciais, Decomposição Wavelet, Monte Carlo, Abordagem Bayesiana, Teste de Taxa de Probabilidade

Sequencial, Cartas de Controle CUSUM, Redes Neurais, Cadeia de Markov e outros (Burrell & Papantoni, 2000; Castanie & Denjean, 1992; Ramirez-Beltran & Montes, 1997; Lavielle & Lebarbier, 2001; Malladi & Speyer, 1999; Morgenstern et al., 1988; Ye, 2000; Macdougall et al., 1998; Karathanassi et al., 1996; Percival & Walden, 2000).

Muitos trabalhos nessa área utilizam abordagens no domínio do tempo para detectar mudanças na média e/ou na variância das séries temporais. Por exemplo, Inclan (1993) e Inclan & Tiao (1994) discutem métodos no domínio do tempo para detectar alterações nos desvios das séries temporais dos retornos de aplicações em ações. Basseville & Nikiforov (2000) oferecem notas sobre os métodos de detecção de pontos de mudança em dados correlacionados. Quanto à detecção retrospectiva de mudanças, Klöppelberg & Mikosch (1996) apresentam um teste baseado no Periodograma Cumulativo Integrado, no entanto, métodos *on-line* para a detecção de mudanças nos processos recebem menos atenção.

Ombao et al. (2004) apresentam uma abordagem paramétrica para detecção *on-line* de pontos de mudança em séries temporais baseada na montagem de modelos autorregressivos (AR) para conjuntos de séries e no uso de métodos baseados em risco para comparação. Embora os modelos AR possam capturar boa parte do comportamento do processo, é provável que seus resultados sejam pobres no caso da estrutura do modelo se diferenciar significativamente do processo AR. No entanto, métodos não paramétricos, tais como os baseados nas propriedades espectrais das séries, têm sido pouco examinados.

Os atuais métodos de detecção de pontos de mudança existentes na literatura tendem a ser de natureza bastante sofisticada. Além disso, faz-se necessário em muitos casos um conhecimento prévio sobre possíveis variações no padrão e sua distribuição, o que pode dificultar a aplicação automática desses métodos, ou seja, a aplicação *on-line* para detecção de mudança. Outras desvantagens comuns são a exigência de estimativas de vários parâmetros, a necessidade de indicadores de acompanhamento e o grande número de variáveis de ajuste. As preocupações são ainda maiores quando há a necessidade de utilização de várias séries temporais simultaneamente.

Este estudo difere dos métodos citados anteriormente por considerar que as mudanças na estrutura de segunda ordem de um processo estocástico são refletidas por mudanças baseadas em espectros de Fourier e Wavelet, ou seja, em decomposições da variância em frequência e escala, respectivamente. A partir da repartição da série em blocos menores, para assumir-se seguramente que sejam estacionários é desenvolvido um método baseado em espectro para

comparar o comportamento do processo em blocos e determinar se uma mudança ocorreu.

Uma das vantagens do método baseado em espectro é não assumir um modelo particular de estrutura, como acontece com o AR, portanto essa abordagem pode também ser eficaz para séries com espectros não autorregressivos (não AR) e/ou sem média móvel (não MA).

2.2. A abordagem multivariada

A ideia principal da abordagem multivariada é ilustrada na Figura 1 e os métodos usados neste estudo são descritos nos itens 2.2.1 e 2.2.2. O processo consiste basicamente em extrair determinadas características das janelas de valores da série temporal não linear e apresentá-las em uma carta de controle multivariada. Neste trabalho, os Descritores de Hjorth de atividade, mobilidade e complexidade são extraídos do intervalo e a carta de controle utilizada para monitorar essas três características é Carta de Controle T^2 de Hotelling. Um alarme, proveniente dessa carta, indicando falta de controle é um sinal de que a série temporal não linear teve um ponto de mudança detectado.

2.2.1. Descritores de Hjorth

Hjorth formulou três parâmetros capazes de caracterizar qualquer sinal e seus equivalentes nos domínios da frequência e do tempo, são eles: atividade, mobilidade e complexidade. Esses parâmetros são chamados de descritores normalizados e podem ser definidos como medidas obtidas a partir da primeira

e da segunda derivadas de uma série temporal não linear $f(t)$.

A transformação dos parâmetros entre os domínios de frequência e tempo é baseada na igualdade de energia no momento atual, ou seja, a potência total no domínio da frequência é idêntica à potência média no domínio do tempo (Hjorth, 1970).

Por meio da Transformada de Fourier, as séries temporais não lineares $f(t)$ podem ser traduzidas em função da frequência e expressas por $F(\omega)$. A fase do sinal é excluída com a multiplicação de $F(\omega)$ pelo seu conjugado $F^*(\omega)$, resultando no espectro de potência $S(\omega)$.

$$S(\omega) = F(\omega) \cdot F^*(\omega) \quad (1)$$

Tem-se também que a frequência do descritor, obtida através da Transformada de Fourier, é simétrica com relação à frequência zero. A consequência dessa simetria é que todos os momentos ímpares da frequência do descritor serão iguais a zero por causa da abordagem estatística para a distribuição de frequência. Assim, não haverá dados para a média linear ou a assimetria, qualidades que constituem o primeiro e o terceiro momentos, respectivamente.

O cálculo dos Descritores de Hjorth de atividade, mobilidade e complexidade pode ser dado, conforme as Equações 2, 3 e 4, em função dos momentos pares da série temporal não linear $f(t)$ em questão:

$$A = m_0 \quad (2)$$

$$M = \left(\frac{m_2}{m_0} \right)^{\frac{1}{2}} \quad (3)$$

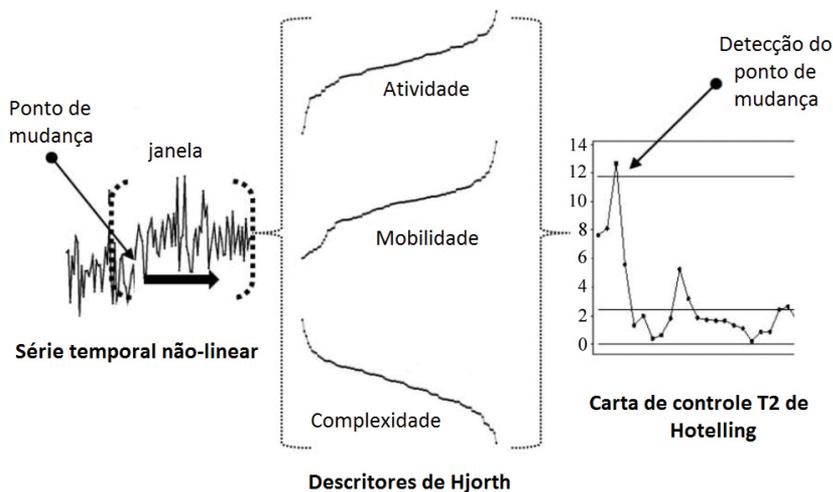


Figura 1. A abordagem multivariada combinando Descritores de Hjorth e Cartas de Controle T^2 de Hotelling.

$$C = \frac{\left(\frac{m_4}{m_2}\right)^{\frac{1}{2}}}{\left(\frac{m_2}{m_0}\right)^{\frac{1}{2}}} \quad (4)$$

É difundido na literatura que a definição geral de um momento m de ordem n é:

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

Os momentos m_0 , m_2 e m_4 são calculados conforme as Equações 6, 7 e 8, ou seja, a partir da função $f(t)$ e suas derivadas de primeira e segunda ordem.

$$m_0 = \frac{1}{T} \int_{-T}^T f^2(t) dt \quad (6)$$

$$m_2 = \frac{1}{T} \int_{-T}^T \left(\frac{df(t)}{dt}\right)^2 dt \quad (7)$$

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

O cálculo das derivadas pode ser feito usando-se a técnica de diferenciação central, como mostrado na Figura 2. A derivada de uma função em um ponto é dada pela inclinação da reta tangente a ele indicada pela linha pontilhada maior e essa inclinação pode ser aproximada para a linha pontilhada menor desde que h seja pequeno o bastante.

Uma aproximação usual para a derivada no ponto a é dada pela tangente do ângulo θ :

$$f'(a) \approx \left(\frac{f(a+h)-f(a-h)}{(a+h)-(a-h)}\right) \Rightarrow f'(a) \approx \left(\frac{f(a+h)-f(a-h)}{2h}\right) \quad (9)$$

O cálculo de derivadas de segunda ordem tem sido feito a partir de uma abordagem funcional muito utilizada, que envolve a aplicação da diferenciação central três vezes. Inicialmente,

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

As duas derivadas no numerador dessa expressão são calculadas como na Equação 9:

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

$$f'\left(a-\frac{1}{2}h\right) \approx \left(\frac{f(a)-f(a-h)}{h}\right) \quad (12)$$

Combinando-se as três equações anteriores, tem-se a derivada de segunda ordem no ponto a , calculada a partir da técnica de diferenciação central:

$$\begin{aligned} f''(a) &\approx \frac{f'\left(a+\frac{1}{2}h\right) - f'\left(a-\frac{1}{2}h\right)}{h} \\ &\approx \frac{1}{h} \left\{ \left(\frac{f(a+h)-f(a)}{h}\right) - \left(\frac{f(a)-f(a-h)}{h}\right) \right\} \\ \Rightarrow f''(a) &\approx \left(\frac{f(a+h) - 2f(a) + f(a-h)}{h^2}\right) \end{aligned} \quad (13)$$

Neste estudo, os parâmetros de Hjorth são usados para discriminar as fases distintas nas quais algumas propriedades não lineares como sazonalidade, volatilidade, autocorreção etc. mudam de acordo com o tempo. Outra maneira de calcular os Descritores de Hjorth de uma série temporal não linear $f(t)$ é a partir das características descritas por Hjorth (1970) para cada um dos três tipos:

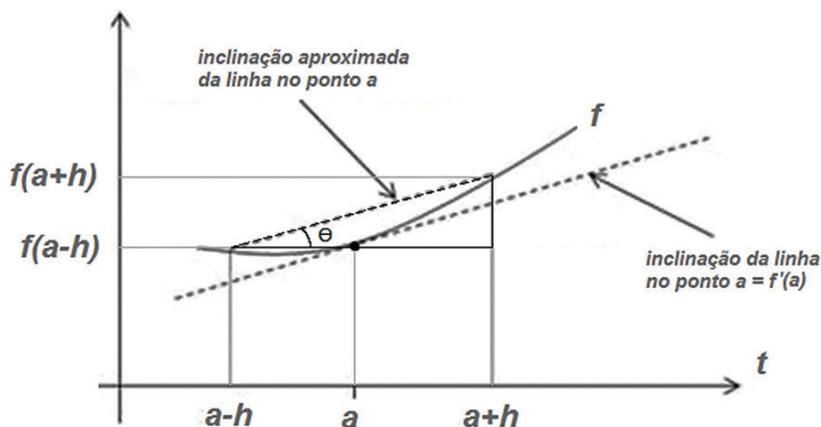


Figura 2. Derivada utilizando aproximação por Diferenciação Central.

- **Atividade:** É quantificada por meio da variância de amplitude (σ_0^2), por ter as propriedades aditivas necessárias para permitir a integração de diferentes observações. No domínio da frequência, a variância da função do tempo pode ser concebida como a superfície do espectro de potência.

$$A = m_0 = \sigma_0^2 \quad (14)$$

- **Mobilidade:** É definida como a raiz quadrada da razão entre a variância da derivada (σ_1^2) e a variância da amplitude (σ_0^2). Como essas quantidades são igualmente dependentes da amplitude média, a mobilidade é dependente apenas da forma da curva, ou seja, da inclinação em relação à média.

$$M = \left(\frac{m_2}{m_0} \right)^{\frac{1}{2}} = \left(\frac{\sigma_1^2}{\sigma_0^2} \right)^{\frac{1}{2}} = \frac{\sigma_1}{\sigma_0} \quad (15)$$

- **Complexidade:** É dada de acordo com a relação entre a mobilidade da derivada da série temporal não linear $(\sigma_2/\sigma_1)^2$ e a mobilidade da própria série $(\sigma_1/\sigma_0)^2$. A mobilidade da primeira derivada é obtida de forma análoga à mobilidade do sinal original, mas com a variância da derivada de segunda ordem. Complexidade indica o desvio da inclinação e pode ser vista como uma medida de mudança na frequência do sinal de entrada.

$$C = \frac{\left(\frac{m_4}{m_2} \right)^{\frac{1}{2}}}{\left(\frac{m_2}{m_0} \right)^{\frac{1}{2}}} = \frac{\left(\frac{\sigma_2^2}{\sigma_1^2} \right)^{\frac{1}{2}}}{\left(\frac{\sigma_1^2}{\sigma_0^2} \right)^{\frac{1}{2}}} = \frac{\sigma_2}{\sigma_1} \quad (16)$$

Embora os Descritores de Hjorth sejam baseados em momentos espectrais, esses parâmetros também podem ser calculados por variações no tempo para o custo computacional ser mais acessível. Dessa forma, o cálculo dos parâmetros em tempo discreto é simplificado, pois envolve a variância (σ_0^2) das séries temporais não lineares $f(t)$ na janela a ser analisada e a variância da primeira e segunda derivadas dadas por (σ_1^2) e (σ_2^2) , respectivamente. Essa abordagem pode, portanto, ser implementada como um método *on-line* para aplicações reais.

2.2.2. Cartas de controle para processos multivariados

Cartas multivariadas de controle são projetadas para incorporar a estrutura de correlação de diversas variáveis em um gráfico a fim de determinar o estado de controle estatístico de um processo. T^2 de Hotelling, CUSUM e EWMA são exemplos comuns de cartas de controle projetadas para detectar mudanças no vetor médio de diversas características de qualidade sob

uma matriz S de covariância constante. Os trabalhos de Lowry & Montgomery (1995), bem como de Bersimis et al. (2007), são recursos valiosos para o aprofundamento do estudo a esse respeito.

Para fins de notação, as observações deste estudo foram organizadas em uma matriz de dados com m linhas e p colunas, na qual m é o tamanho da janela e p é o número de características observadas, neste caso, os três Descritores de Hjorth. Assim:

$$X = \begin{bmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_m \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ \vdots & \vdots & \vdots \\ x_{m1} & x_{m2} & x_{m3} \end{bmatrix} \quad (17)$$

A observação individual de cada um dos três descritores (atividade, mobilidade e complexidade) pode ser denominada i e é dada pelo vetor x_i :

$$x_i = \begin{bmatrix} X_{i1} \\ X_{i2} \\ X_{i3} \end{bmatrix} \quad (18)$$

O vetor médio estimado cujos componentes são as médias de cada descritor é

$$\bar{x} = \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \bar{X}_3 \end{bmatrix} \text{ no qual } \bar{X}_j = \frac{1}{m} \sum_{i=1}^m x_{ij}, j = 1, \dots, 3 \quad (19)$$

e a matriz de covariância estimada é:

$$S = \frac{1}{m-1} \sum_{i=1}^m (x_i - \bar{x})(x_i - \bar{x})' = \begin{bmatrix} S_{11}^2 & S_{12} & S_{13} \\ \dots & S_{22}^2 & S_{23} \\ \dots & \dots & S_{33}^2 \end{bmatrix} \quad (20)$$

A construção de uma carta de controle para processos multivariados baseada na Carta T^2 de Hotelling para uma observação x_i é ilustrada na Equação 21 e parte dos cálculos estatísticos mostrados nas equações anteriores.

$$T_i^2 = (x_i - \bar{x})' S^{-1} (x_i - \bar{x}) \quad (21)$$

A distribuição da estatística T^2 não é amplamente conhecida, portanto a maioria dos usuários de cartas de controle multivariadas faz a aproximação com uma distribuição beta para obter os limites do gráfico de controle. Especificamente,

$$T^2 \approx \frac{(m-1)^2}{m} B \left(\frac{1-\alpha}{2}, \frac{p}{2}, \frac{m-p-1}{2} \right) \quad (22)$$

Levando em conta essa relação, o limite para a Carta de Controle T^2 de Hotelling é baseado em um percentual superior da distribuição beta central para obter o controle desejado (em ARL) e α é o nível de significância (geralmente 1%).

O seguinte procedimento é utilizado para executar a Carta de Controle T^2 de Hotelling a partir das fórmulas acima mencionadas:

- Calcular a média para cada descritor (X_1 Atividade, X_2 Mobilidade e X_3 Complexidade).
- Calcular a matriz de covariância S .
- Calcular a estatística T^2 .
- Calcular o limite de controle superior para a Carta de Controle T^2 de Hotelling.
- Plotar as estatísticas T^2 na Carta de Controle e comparar com os limites para determinar se os pontos individuais estão fora de controle.

2.3. Um exemplo

O exemplo simplificado a seguir ilustra a metodologia proposta aplicada a uma série temporal não linear SAR (*Smoothed Autoregressive*) $f(t)$ descrita pelo modelo da Equação 23, no qual $sign(x) = 1, 0, -1$ se $x > 0, x = 0$ e $x < 0$, respectivamente. O erro ε_t segue uma distribuição normal $N(0,1)$.

$$y_t = sign(y_{t-12}) + \varepsilon_t \tag{23}$$

A série linear e estacionária $f(t)$ mostrada na Tabela 1 é dividida em dois segmentos; um modelo SAR puro nos primeiros 24 pontos; no 25º valor há um ponto de mudança e tem início o segundo segmento, cuja média é acrescida em dois desvios padrão com relação à média do primeiro. A Figura 3 mostra essa divisão.

Os Descritores de Hjorth de atividade, mobilidade e complexidade são calculados para os pontos 25 a 48 partindo-se das Equações 2, 3 e 4, respectivamente. Uma janela deslizante com 24 pontos é utilizada para calcular os valores estatísticos, ou seja, todas as estatísticas para $t = 25$ consideram os pontos 2 a 25 da mesma maneira que as estatísticas para $t = 26$ consideram os pontos 3 a 26 e assim sucessivamente. A Figura 4 mostra o comportamento dos três Descritores de Hjorth para o segundo segmento da série $f(t)$.

Os pontos da Carta de Controle T^2 de Hotelling mostrada na Figura 5 são calculados pelas Equações 20, 21 e 22. Como nesse caso o primeiro segmento tem 24 observações, então $m = 24$ e o vetor médio dos descritores é dado por

$$\bar{x} = \begin{bmatrix} 1,680 \\ 1,092 \\ 1,544 \end{bmatrix} \tag{24}$$

e a matriz de covariância de x para a amostra é

$$S = \begin{bmatrix} 0,194 & -0,066 & 0,092 \\ -0,066 & 0,027 & -0,036 \\ 0,092 & -0,036 & 0,051 \end{bmatrix} \tag{25}$$

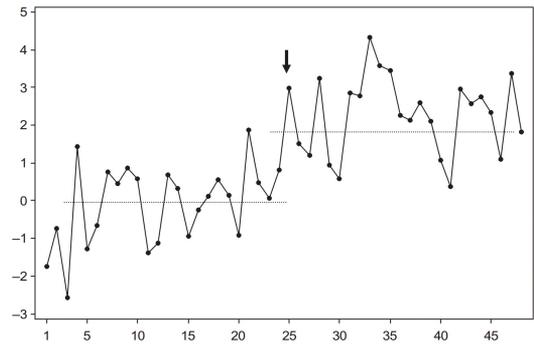


Figura 3. Modelo SAR com dois segmentos de dados e um ponto de mudança em $t = 25$.

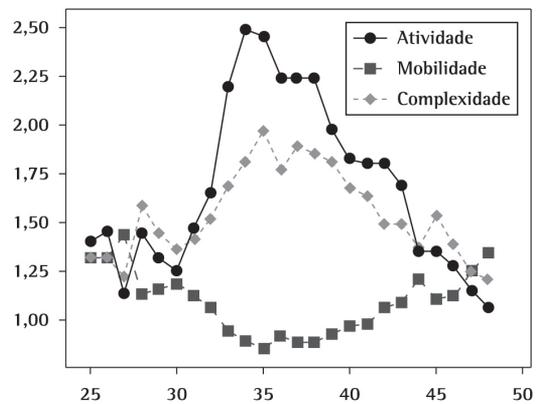


Figura 4. Descritores de Hjorth para o segundo segmento do modelo SAR em questão.

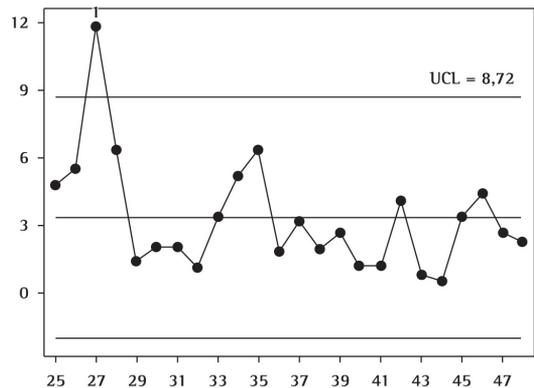


Figura 5. Carta de Controle T^2 de Hotelling para Descritores de Hjorth.

O limite superior de controle usando a distribuição beta é dado por:

$$T^2 \approx \frac{(24-1)^2}{24} B\left(1 - \frac{0,01}{2}, \frac{3}{2}, \frac{24-3-1}{2}\right) \tag{26}$$

$$\approx \frac{23^2}{24} B(0,995; 1,5; 10) \approx 8,72$$

De maneira simplificada, é possível afirmar que a mudança ocorrida entre os pontos 24 e 25 foi detectada no ponto 27, ou seja, apenas três pontos

após o início do novo regime. A Tabela 1 ilustra a reprodução da metodologia proposta através da computação dos resultados.

Tabela 1. Descritores de Hjorth e valores T^2 .

	$f(t)$	$X_1 = A = \sigma_0^2$	$f'(t)$	σ_1^2	$X_2 = M = \frac{\sigma_1}{\sigma_0}$	$f''(t)$	σ_2^2	$X_3 = C = \frac{\sigma_2}{\sigma_1 / \sigma_2}$	T^2
1	-1,74	-	-	-	-	-	-	-	-
2	-0,74	-	1,00	-	-	-	-	-	-
3	-2,58	-	-1,84	-	-	-2,84	-	-	-
4	1,45	-	4,03	-	-	5,87	-	-	-
5	-1,28	-	-2,73	-	-	-6,76	-	-	-
6	-0,65	-	0,63	-	-	3,36	-	-	-
7	0,76	-	1,41	-	-	0,78	-	-	-
8	0,47	-	-0,29	-	-	-1,70	-	-	-
9	0,87	-	0,40	-	-	0,69	-	-	-
10	0,60	-	-0,27	-	-	-0,67	-	-	-
11	-1,37	-	-1,97	-	-	-1,70	-	-	-
12	-1,12	-	0,25	-	-	2,22	-	-	-
13	0,69	-	1,81	-	-	1,56	-	-	-
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	-	-	0,06	-	-	-
19	0,14	-	-0,42	-	-	-0,85	-	-	-
20	-0,91	-	-1,05	-	-	-0,63	-	-	-
21	1,88	-	2,79	-	-	3,84	-	-	-
22	0,49	-	-1,39	-	-	-4,18	-	-	-
23	0,07	-	-0,42	-	-	0,97	-	-	-
24	0,83	-	0,76	-	-	1,18	-	-	-
25	3,00	1,41	2,17	2,44	1,32	1,41	7,33	1,32	4,81
26	1,51	1,45	-1,49	2,52	1,32	-3,66	7,59	1,32	5,57
27	1,22	1,14	-0,29	2,36	1,44	1,20	7,30	1,22	11,82
28	3,25	1,44	2,03	1,85	1,13	2,32	6,04	1,59	6,38
29	0,94	1,32	-2,31	1,76	1,15	-4,34	4,88	1,44	1,48
30	0,58	1,26	-0,36	1,75	1,18	1,95	4,55	1,36	2,09
31	2,85	1,47	2,27	1,89	1,13	2,63	4,82	1,41	2,06
32	2,78	1,66	-0,07	1,88	1,07	-2,34	4,94	1,52	1,16
33	4,35	2,20	1,57	1,97	0,95	1,64	5,03	1,69	3,45
34	3,59	2,49	-0,76	2,00	0,90	-2,33	5,25	1,81	5,26
35	3,45	2,46	-0,14	1,80	0,86	0,62	5,14	1,97	6,36
36	2,26	2,24	-1,19	1,88	0,92	-1,05	4,97	1,77	1,85
37	2,14	2,24	-0,12	1,76	0,89	1,07	4,91	1,89	3,21
38	2,60	2,24	0,46	1,76	0,89	0,58	4,73	1,85	1,97
39	2,12	1,98	-0,48	1,69	0,92	-0,94	4,73	1,81	2,72
40	1,09	1,83	-1,03	1,73	0,97	-0,55	4,57	1,67	1,23
41	0,37	1,80	-0,72	1,75	0,99	0,31	4,58	1,64	1,30
42	2,97	1,80	2,60	2,02	1,06	3,32	5,05	1,49	4,17
43	2,59	1,69	-0,38	2,02	1,09	-2,98	5,41	1,50	0,89
44	2,76	1,35	0,17	1,96	1,20	0,55	5,40	1,38	0,58
45	2,36	1,35	-0,40	1,65	1,10	-0,57	4,76	1,54	3,47
46	1,12	1,28	-1,24	1,64	1,13	-0,84	4,05	1,39	4,50
47	3,38	1,15	2,26	1,83	1,26	3,50	4,53	1,25	2,75
48	1,83	1,07	-1,55	1,93	1,34	-3,81	5,10	1,21	2,35

O cálculo de T^2 para o ponto 27, sendo $x_{27} = \begin{bmatrix} 1,14 \\ 1,44 \\ 1,22 \end{bmatrix}$,

resulta em $T^2 = 11,82$. Como pode ser visualizado na Figura 5, esse é um ponto que está fora do padrão.

3. Resultados da simulação

3.1. Séries temporais não lineares: conjuntos de dados

Diversos métodos para a análise de séries temporais têm sido utilizados para representar vários tipos de processos nas últimas duas décadas. Recentemente, entretanto, vem crescendo o interesse em alargar o quadro clássico de Box & Jenkins (1976) para incorporar propriedades tais como a não linearidade, a não normalidade e a heterocedasticidade. Um grande número de modelos não lineares tem sido desenvolvido, como, por exemplo, o modelo Bilinear de Granger & Anderson (1978), o modelo de TAR (*Threshold Autoregressive*) de Tong (1978), o modelo de estado dependente de Priestley (1980), o modelo de mudança de Markov, abordado por Hamilton (1989), o modelo de Coeficiente Funcional Autorregressivo de Chen & Tsay (1993), entre outros.

As séries temporais tornam-se ainda mais complexas por causa de características como alta frequência, sazonalidade, efeitos de calendário como finais de semana e feriados, alta volatilidade e presença de *outliers*. Cada uma dessas características é responsável por uma grande variedade de comportamentos não lineares que se sobrepõem no conjunto de dados.

Considerando uma série temporal univariada x_t observada em momentos espaçados, denotamos as observações por $(x_t | t = 1, \dots, T)$, onde T é o tamanho da amostra. Uma série temporal puramente estocástica x_t é dita linear se puder ser escrita como:

$$x_t = \mu + \sum_{i=0}^{\infty} \psi_i \cdot a_{t-i} \tag{27}$$

sendo μ uma constante, ψ_i números reais a começar por $\psi_0 = 1$ e a_t uma sequência independente e identicamente distribuída (IID) de variáveis aleatórias com uma função de distribuição bem definida.

Assume-se que a distribuição a_t é contínua e $\varepsilon(a_t) = 0$. Em muitos casos, assume-se também que $\text{var}(a_t) = \sigma_a^2$ ou, ainda, que a_t é Gaussiana. Se $\sigma_a^2 \cdot \sum_{i=1}^{\infty} \psi_i^2 < \infty$, então x_t é fracamente estacionário, isto é, os dois primeiros momentos de x_t são invariantes. O processo autorregressivo de média móvel ARMA, por exemplo, é linear porque pode ser representado como na Equação 27. Qualquer processo estocástico que não satisfaça a condição dessa equação é dito não linear. Mais detalhes a esse respeito podem ser encontrados em Tsay (2005). Para pesquisas a respeito de séries temporais não lineares, a leitura de Gooijer (1998) e Granger & Terasvirta (1993) é recomendada.

A Tabela 2 mostra o conjunto de diferentes séries temporais não lineares escolhidas, implementadas e simuladas para este estudo. Em cada caso, $\varepsilon_t \sim N(0,1)$ é assumido como sendo IID. Estes seis modelos de séries temporais representam uma variedade de problemas com características diversas. Assim sendo, algumas das séries possuem estrutura de correlação puramente autorregressiva (AR) ou puramente de média móvel (MA), enquanto outros apresentam uma mistura desses componentes. Modelos similares, mas com diferentes sazonalidades (Lag), foram explorados por Zhang et al. (2001).

3.2. Resultados

Os principais resultados do estudo de simulação são apresentados na Tabela 3, na qual o ARL é calculado para as seis séries não lineares criadas artificialmente. O ARL é o valor esperado do número de pontos necessários para que o gráfico sinalize o processo de mudança na série temporal. Dois tipos de mudanças são considerados: na média e/ou no desvio padrão da amostra.

Tabela 2. Modelos de séries temporais não lineares.

Modelo	Equação
SAR	$y_t = \text{sign}(y_{t-12}) + \varepsilon_t$ sendo $\text{sign}(x) = 1, 0, -1$ se $x > 0, x = 0, x < 0$, respectivamente
Bilinear (BL)	$y_t = 0,3y_{t-1} - 0,4y_{t-24} + 0,6y_{t-1} \cdot \varepsilon_{t-1} + \varepsilon_t$
TAR	$y_t = 0,7y_{t-1} + \varepsilon_t$ para $ y_{t-1} \leq 1$ $y_t = -0,4y_{t-12} + \varepsilon_t$ para $ y_{t-1} > 1$
Autorregressiva não linear (NAR)	$y_t = \frac{0,6y_{t-1}}{ y_{t-24} + 3} + \varepsilon_t$
NMA	$y_t = \varepsilon_t - 0,5\varepsilon_{t-24} + 0,1\varepsilon_{t-2} + 0,3\varepsilon_{t-1} \cdot \varepsilon_{t-2} - 0,3\varepsilon_{t-2}^2$
Autorregressiva com transição suave (STAR)	$y_t = 0,4y_{t-1} + 0,5y_{t-24} + (0,25 - 0,81y_{t-1} + 0,7y_{t-24})[1 + \exp(-10y_{t-1})]^{-1} + \varepsilon_t$

Fonte: Zhang et al. (2001). Nota: A sazonalidade é representada pelo número 12 ou 24 nas equações.

Tabela 3. ARL para séries temporais não lineares.

Média	Desvio padrão	Sazonalidade	Séries temporais não lineares					
			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	$1,5\sigma$	12	180	170	174	166	180	167
μ	$1,5\sigma$	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
$\mu+1\sigma$	$1,5\sigma$	12	15	19	15	27	17	17
$\mu+1\sigma$	$1,5\sigma$	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
$\mu+2\sigma$	$1,5\sigma$	12	4	4	2	2	3	4
$\mu+2\sigma$	$1,5\sigma$	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
$\mu+3\sigma$	$1,5\sigma$	12	2	1	2	4	3	4
$\mu+3\sigma$	$1,5\sigma$	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

Os limites para a Carta de Controle T^2 de Hotelling foram calculados com base nos 24 primeiros pontos da amostra e os resultados consideram uma simulação com t variando de 1 a 100 para cada série temporal. Uma interface no *software* Statistica (StatSoft, Inc., 2005) foi desenvolvida para executar a simulação.

Em cada passagem, uma janela deslizante de 24 pontos foi processada. Dessa forma, a primeira janela para a Carta de Controle utilizou os pontos t_1 a t_{24} . A segunda adicionou o ponto t_2 e removeu o ponto t_1 , analisando os pontos entre t_2 a t_{25} . A terceira, os pontos t_3 a t_{26} , e assim por diante. Esse processo foi repetido até a identificação de um ponto fora de controle. Quando esse *outlier* é encontrado, o valor RL pode ser obtido subtraindo 24 do t_{RL} . Altos valores de ARL indicam que a série temporal está sob controle e baixos valores de ARL, que a série temporal varia.

Os ARLs obtidos na Tabela 3 para todas as séries temporais são comparáveis aos resultados tradicionais relacionados com uma Carta de Controle X-barra. O ARL para grandes mudanças na média e/ou na variância é muito baixo, ou seja, a detecção de pontos de mudança ocorre após alguns poucos pontos a partir da introdução do novo regime. Quando não há mudança na série, os ARLs são muito grandes e por isso alarmes falsos são raramente obtidos. Um teste

paired t comparando os ARLs para as sazonalidades de 12 e 24 resulta em um $p_{value} < 0,05$; o que significa que pontos de mudança em séries temporais não lineares com pequenos ciclos sazonais são detectados mais rapidamente do que naquelas com grandes ciclos sazonais.

Para as séries temporais sem componente MA, os ARLs correspondentes a alarmes falsos são estatisticamente diferentes dos de outras séries. Uma série temporal com mudanças não lineares na média é o pior cenário para os resultados da simulação, pois os alarmes falsos foram obtidos com menos pontos. Isso está relacionado com a alta volatilidade da série temporal.

O procedimento de determinar o limite de controle para a Carta T^2 de Hotelling a partir de um percentual superior da distribuição da sua estatística é válido quando os sucessivos valores de T^2 são independentes; porém a janela deslizante introduz correlação serial entre os sucessivos valores dos descritores e, em consequência, entre os sucessivos valores de T^2 . Assim, a rigor, para garantir o ARL desejado seria necessário calibrar os limites de controle por simulação ou por um modelo analítico mais complexo que $ARL = 1/P$.

A Figura 6 mostra um gráfico multivariado explorando as possíveis interações para os resultados

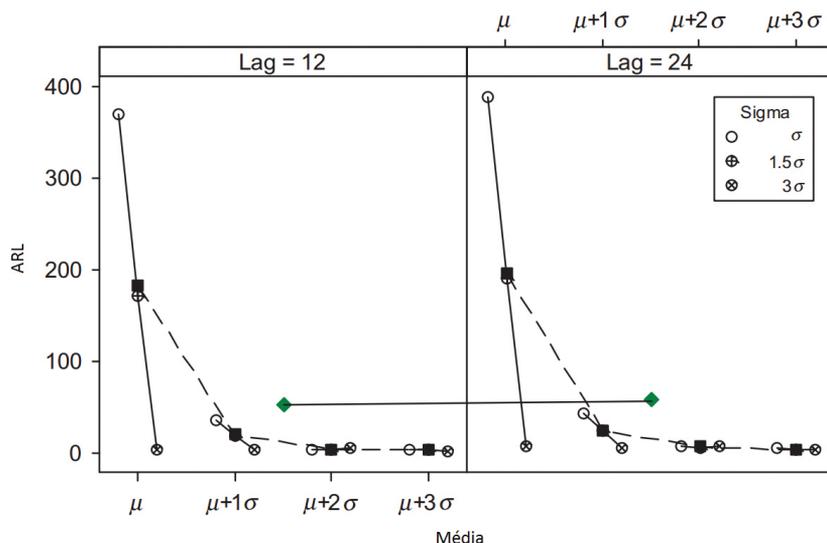


Figura 6. ARL como função dos fatores média (μ), desvio padrão (σ) e sazonalidade (Lag).

da Tabela 3: o ARL em função de variações na média (μ), de variações no desvio padrão (σ) e de variações na sazonalidade (Lag). É possível observar-se que os gráficos são bem similares para as duas sazonalidades (Lag = 12 e Lag = 24), indicando que não existe interação entre média, desvio padrão e sazonalidade no ARL. Ainda nesse gráfico percebe-se que a média e o desvio padrão são fatores estatisticamente significativos para a resposta ARL, já a sazonalidade não é estatisticamente significativa.

4. Aplicação: séries temporais do consumo de energia elétrica

Companhias que trabalham em mercados com pouca regularidade quanto ao consumo de energia elétrica, como os Estados Unidos, a Inglaterra, o Brasil e a maioria dos países, usam séries temporais para prever a demanda de energia elétrica como parte das informações necessárias para definir contratos de compra e venda. Quanto às empresas de produção ou comércio, seus interesses estão ligados particularmente à otimização da quantidade de energia comprada ou produzida e dos preços repassados a seus consumidores.

Este estudo de caso é baseado na modelagem e previsão de uma variável importante, cujos efeitos influenciam muitos contratos: o consumo de energia elétrica. Possuir um conhecimento sobre seus valores futuros ou variações é essencial para calcular o risco e o retorno de possíveis transações e investimentos. No processo industrial existem inúmeros fatores que ocasionam aumento ou redução do consumo de energia e que não podem ser previstos; por

isso, embora também seja possível contratar certa quantidade de energia sem variação admissível, para certos consumidores isso representa um risco muito alto. Assim, um produtor de eletricidade capaz de aceitar, assumir e gerenciar esse risco no lugar do consumidor terá mais oportunidades no mercado.

Como considerado por Angelus (2001) e Mount (2001), a dificuldade em prever preços e consumo de energia elétrica está nas muitas possibilidades de variação dos fatores relacionados, tais como fornecimento e demanda de eletricidade, limitação dos geradores, transmissores e distribuidores, entre outros. Até uma simples mudança de clima pode alterar a demanda e, consequentemente, o preço da energia. Pai & Hong (2005) também constataram a complexidade de prever o consumo de energia elétrica por causa da não linearidade de seus fatores de influência; além disso, deram grande importância às estratégias regionais ou nacionais de gerenciamento de sistemas de potência, que preveem com precisão a carga necessária em futuras demandas.

Em um mercado inconstante, um grande volume de energia elétrica é comprado e vendido por terceiros no mercado paralelo ou no mercado *spot* entre agentes de diferentes regiões geográficas. Este estudo de caso usa a metodologia proposta para detectar pontos de mudança no consumo de energia elétrica de seis consumidores industriais brasileiros cujos dados históricos são dispostos de hora em hora nas séries temporais. O ponto de mudança aqui representa a hora em que o consumidor vai para o mercado paralelo ou para o mercado *spot*, alterando as propriedades da série temporal. Isso está relacionado à fidelidade

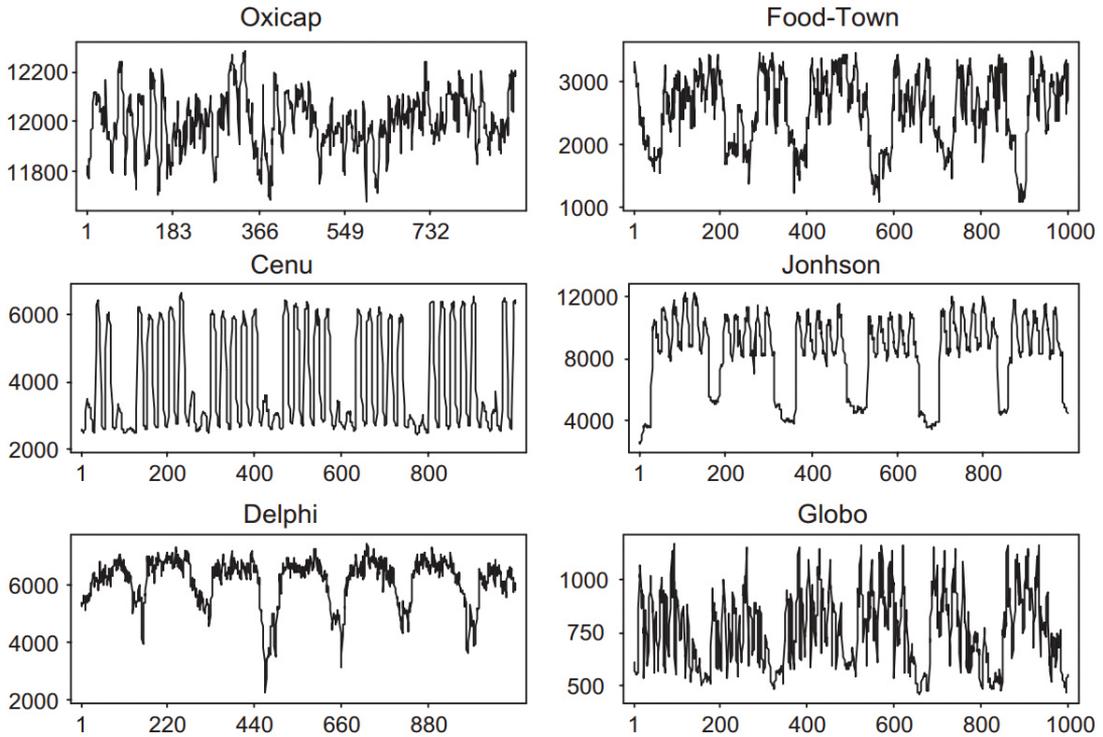


Figura 7. Séries temporais do consumo energético por hora de seis indústrias.

do consumidor ao contrato de compra e venda de energia elétrica previamente estabelecido.

A Figura 7 mostra as séries temporais do consumo de energia elétrica para cada um dos seis consumidores industriais estudados. Alguns procedimentos de melhoria na coleta de dados foram executados para tratar *outliers*, erros de digitação, sazonalidade, falta de dados etc. O conjunto de dados disponível abrange três anos com dados sobre a potência consumida a cada hora pelas empresas, e foi disponibilizado pela Duke Energy, uma empresa de produção de energia elétrica que opera no Brasil. Uma sazonalidade semanal foi observada em todas as séries temporais, resultando em 168 intervalos semelhantes.

Um problema típico ao se lidar com múltiplas séries temporais é a existência de correlação entre elas. Por isso é válido mencionar que, para o conjunto de dados utilizados neste trabalho, o comportamento das séries temporais não pode ser considerado um processo multivariado, uma vez que a correlação entre as variáveis não é significativa para nenhum par entre as seis séries. O fato de haver dados ausentes, ou seja, quantidades diferentes de dados nas amostras, torna a possibilidade de existência de correlação ainda menos provável. Como consequência, cada série deve ser trabalhada individualmente.

A Tabela 4 apresenta os resultados para os seis consumidores industriais. Foram computadas características das séries temporais como sazonalidade, média, erro, desvio padrão, assimetria e achatamento. Dados históricos e informações dos consumidores sobre negócios no mercado *spot* também foram providenciados.

Foi informado previamente que os consumidores Delphi e Johnson não negociaram eletricidade em mercados paralelos, portanto a detecção de um ponto de mudança para essas empresas, usando-se o método em questão, seria um falso alarme, pois não corresponderia à realidade. Para os dados referentes à empresa Delphi, houve a detecção de um ponto de mudança após 1.847 medidas, devido ao Descritor de Hjorth de complexidade, mas esse resultado não é significativo, pois essa é a série temporal com maior volatilidade dentre as seis estudadas. Em um gráfico de controle tradicional, o ponto de mudança seria detectado com aproximadamente 370 pontos, em média.

A quantidade de valores necessários para detectar um ponto de mudança também foi calculada para as empresas Cenu, Food-Town, Oxicap e Globo e o resultado foi inferior ao comprimento do ciclo sazonal calculado usando-se um algoritmo de Transformada

Tabela 4. Resultados da detecção de pontos de mudança para os seis consumidores industriais.

Características das séries temporais	Consumo de energia elétrica					
	Cenu	Food-Town	Oxicap	Delphi	Johnson	Globo
Sazonalidade	168	24	24	168	24	24
Média	2554,3	11,999	3984,3	8402,6	6139	759,2
Erro da média	18,1	3,69	46,3	78,2	24,9	5,41
Desvio padrão	572,7	112	1462,1	2471,9	827,1	171,03
Assimetria	-0,44	-0,16	0,57	-0,71	-1,42	0,23
Achatamento	-0,7	-0,12	-1,46	-0,7	1,95	-0,8
Negócios no mercado <i>spot</i>	Yes	Yes	Yes	No	No	Yes
Quantidade	12	23	18	-	-	18
Resultados						
Alarme falso (ARL ₀)	-	-	-	1847	-	-
Detecção de ponto de mudança	52	12	11	-	-	8
Descriptor de Hjorth	A, M	M	C	C	-	C

Rápida de Fourier. O método mostrou-se, portanto, capaz de reconhecer mudanças na dinâmica da série em um período menor do que o comprimento do ciclo sazonal, o que é um resultado satisfatório para séries temporais com dados de hora em hora.

Existem vários dados relacionados à detecção de pontos de mudança ao longo das séries temporais, pois correspondem aos momentos em que o comportamento da série foi alterado devido aos negócios no mercado *spot*.

5. Conclusões

Ao longo deste estudo, o novo método multivariado proposto foi aplicado para detectar mudanças dinâmicas de séries temporais não lineares. Esse método é baseado em uma Carta de Controle T² de Hotelling monitorando os Descritores de Hjorth de atividade, mobilidade e complexidade. A fim de ilustrar a detecção de pontos de mudança e o monitoramento de causas especiais de variação nos descritores normalizados, aplicou-se com sucesso esse novo método em seis séries temporais artificiais não lineares.

Os resultados simulados mostraram que o método foi capaz de detectar pontos de mudança de forma adequada e em um tempo (ARL) satisfatório. Além disso, a metodologia proposta foi bem-sucedida na detecção de pontos de mudança em seis séries reais relativas ao consumo de energia elétrica de seis empresas brasileiras.

Um exemplo da aplicabilidade deste trabalho é a empresa Neurotec (www.neurotec.com.br), que tem utilizado os recursos dele oriundos em tomada de decisões altamente críticas relacionadas à detecção de padrões epiléticos em sinais de eletroencefalograma. Tal empresa foi agraciada com os maiores prêmios de qualidade do país em Biomedicina.

Quanto às pesquisas futuras, estão sendo desenvolvidos algoritmos de reconhecimento de padrões para acompanhamento em tempo real de Cartas de Controle. Nesse novo estudo, vários ensaios serão monitorados em tempo real utilizando-se redes neurais. Esse mecanismo é necessário para a detecção de mudança de ponto em vários tipos de séries temporais como, por exemplo, sinais de EEG e processos de usinagem. Outras pesquisas adicionais serão a avaliação da presente abordagem em face da entropia de Kolmogorov-Sinai e a aplicação da metodologia abordada neste artigo a dados correlacionados.

Referências

- Angelus, A. (2001). Electricity price forecasting in deregulated markets. *Electricity Journal*, 14(3), 32-41. [http://dx.doi.org/10.1016/S1040-6190\(01\)00184-1](http://dx.doi.org/10.1016/S1040-6190(01)00184-1)
- Basseville, M., & Nikiforov, I. (2000). *Detection of abrupt changes: theory and application*. New Jersey: Prentice-Hall.
- Bersimis, S., Psarakis, S., & Panaretos, J. (2007). Multivariate statistical process control charts: an overview. *Quality and Reliability Engineering International*, 23(5), 517-543. <http://dx.doi.org/10.1002/qre.829>
- Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis: forecasting and control*. San Francisco: Holden-Day.
- Burrell, A., & Papantoni, T. (2000). Sequential algorithms for detecting changes in acting stochastic processes and on-line learning of their operational parameters. In *Proceedings of the 15th International Conference on Pattern Recognition*, Barcelona, Spain. <http://dx.doi.org/10.1109/ICPR.2000.906160>
- Cabrerizo, M., Ayala, M., Goryawala, M., Jayakar, P., & Adjouadi, M. (2012). A new parametric feature descriptor for the classification of epileptic and control EEG records in pediatric population. *International Journal of Neural Systems*, 22(2). PMID:23627587. <http://dx.doi.org/10.1142/S0129065712500013>
- Castanie, F., & Denjean, F. (1992). Mean value jump detection using wavelet decomposition. In *Proceedings of the*

- IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, Victoria, Canada. <http://dx.doi.org/10.1109/TFTSA.1992.274207>
- Chen, R., & Tsay, R. S. (1993). Functional-coefficient autoregressive models. *Journal of the American Statistical Association*, 88(421), 298-308.
- Gooijer, J. G. (1998). On threshold moving-average models. *Journal of Time Series Analysis*, 19(1), 1-18. <http://dx.doi.org/10.1111/1467-9892.00074>
- Granger, C. W. J., & Anderson, A. P. (1978). *An introduction to bilinear time series models*. Göttingen: Vandenhoeck & Ruprecht.
- Granger, C. W. J., & Terasvirta, T. (1993). *Modeling nonlinear economic relationship*. New York: Oxford University Press.
- Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2), 357-384. <http://dx.doi.org/10.2307/1912559>
- Hjorth, B. (1970). EEG analysis based on time domain properties. *Electroencephalography and Clinical Neurophysiology*, 29(3), 306-310. [http://dx.doi.org/10.1016/0013-4694\(70\)90143-4](http://dx.doi.org/10.1016/0013-4694(70)90143-4)
- Inclan, C. (1993). Detection of multiple changes of variance using posterior odds. *Journal of Business & Economic Statistics*, 11(3), 289-300.
- Inclan, C., & Tiao, G. (1994). Use of cumulative sums of squares for retrospective detection of changes of variance. *Journal of the American Statistical Association*, 89(427), 913-923.
- Karathanassi, V., Iossifidis, C., & Rokos, D. (1996). Application of machine vision techniques in the quality control of pharmaceutical solutions. *Computers in Industry*, 32(2), 169-179. [http://dx.doi.org/10.1016/S0166-3615\(96\)00063-2](http://dx.doi.org/10.1016/S0166-3615(96)00063-2)
- Klöppelberg, C., & Mikosch, T. (1996). Gaussian limit fields for the integrated periodogram. *The Annals of Applied Probability*, 6(3), 969-991. <http://dx.doi.org/10.1214/aop/1034968236>
- Lavielle, M., & Lebarbier, E. (2001). An application of MCMC methods for the multiple change-points problem. *Signal Process*, 81(1), 39-53. [http://dx.doi.org/10.1016/S0165-1684\(00\)00189-4](http://dx.doi.org/10.1016/S0165-1684(00)00189-4)
- Lowry, C. A., & Montgomery, D. C. (1995). A review of multivariate control charts. *IIE Transactions*, 27(5), 800-810. <http://dx.doi.org/10.1080/07408179508936797>
- Macdougall, S., Nandi, A. K., & Chapman, R. (1998). Multiresolution and hybrid bayesian algorithms for automatic detection of change points. *IEE Proceedings - Vision, Image and Signal*, 145(4), 280-286. <http://dx.doi.org/10.1049/ip-vis:19982150>
- Malladi, D. P., & Speyer, J. L. (1999). A generalized shiryayev sequential probability ratio test for change detection and isolation. *IEEE Transactions on Automatic Control*, 44(8), 1522-1534. <http://dx.doi.org/10.1109/9.780416>
- Mariani, S., Manfredini, E., Rosso, V., Mendez, M. O., Bianchi, A. M., Matteucci, M., Terzano, M. G., Cerutti, S., & Parrino, L. (2011). Characterization of a phases during the cyclic alternating pattern of sleep. *Clinical Neurophysiology*, 122(10), 2016-2024. PMID:21439902. <http://dx.doi.org/10.1016/j.clinph.2011.02.031>
- Morgenstern, V. M., Upadhyaya, B. R., & Benedetti, M. (1988). Signal anomaly detection using modified cusum method. In *Proceedings of the 27th IEEE Conference on Decision and Control*, Knoxville, Texas. <http://dx.doi.org/10.1109/CDC.1988.194756>
- Mount, T. (2001). Market power and price volatility in restructured markets for electricity. *Decision Support Systems*, 30, 311-325. [http://dx.doi.org/10.1016/S0167-9236\(00\)00108-1](http://dx.doi.org/10.1016/S0167-9236(00)00108-1)
- Ombao, H., Heo, J., & Stoffer, D. (2004). *Statistical analysis of seismic signals: an almost real-time approach, time series analysis and applications to geophysical systems* (IMA Series 139). New York: Wiley.
- Pai, PF., & Hong, W. C. (2005). Support vector machines with simulated annealing algorithms in electricity load forecasting. *Energy Conversion and Management*, 46, 2669-2688. <http://dx.doi.org/10.1016/j.enconman.2005.02.004>
- Percival, D. B., & Walden, A. T. (2000). *Wavelet methods for time series analysis*. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511841040>
- Priestley, M. B. (1980). State-dependent models: a general approach to nonlinear time series analysis. *Journal of Time Series Analysis*, 1, 47-71. <http://dx.doi.org/10.1111/j.1467-9892.1980.tb00300.x>
- Ramirez-Beltran, N. D., & Montes, J. A. (1997). Neural networks for on-line parameter change detections in time series models. *Computers & Industrial Engineering*, 33, 337-340. [http://dx.doi.org/10.1016/S0360-8352\(97\)00106-X](http://dx.doi.org/10.1016/S0360-8352(97)00106-X)
- Souza, R. C. (1981). Metodologias para a análise e previsão de séries temporais univariadas e multivariadas. *Revista de Econometria*, 77, 105.
- Statsoft, Inc. (2005). *Statistica: version 7.1*. Recuperado em 6 de fevereiro de 2013, de www.statsoft.com
- Tong, H. (1978). On a threshold model. In C. H. Chen (Ed.), *Pattern recognition and signal processing* (pp. 575-586). Amsterdam: Sijhoff & Noordhoff. http://dx.doi.org/10.1007/978-94-009-9941-1_24
- Tsay, R. S. (2005). *Analysis of financial time series*. New Jersey: Wiley. <http://dx.doi.org/10.1002/0471746193>
- Ye, N. (2000). Confidence assessment of quality prediction from process measurement in sequential manufacturing processes. *IEEE Transactions on Electronics Packaging Manufacturing*, 23(3), 177-184. <http://dx.doi.org/10.1109/6104.873245>
- Zhang, G. P., Patuwo, B. E., & Hu, M. Y. (2001). A simulation study of artificial neural networks for nonlinear time-series forecasting. *Computers & Operations Research*, 28, 381-396. [http://dx.doi.org/10.1016/S0305-0548\(99\)00123-9](http://dx.doi.org/10.1016/S0305-0548(99)00123-9)

Agradecimentos

Nossos agradecimentos à Capes, à Fapemig e ao CNPq pelo suporte à pesquisa.

Detecting a level change in a nonlinear time series using Hjorth's Descriptors

Abstract

The purpose of this paper is to present a method for detecting the dynamic changes in a nonlinear time series that uses the Hotelling T^2 multivariate control chart to monitor the variation in three normalized descriptors: Hjorth's descriptors of activity, mobility and complexity. This approach was applied to different simulated nonlinear time series and is illustrated in this paper with a detailed example. A case study with six time series of short-term electricity load consumption was also used to confirm the method's effectiveness.

Keywords

Nonlinear time series. Hjorth's Descriptors. Hotelling T^2 control chart. Level change.
