

ESTUDO COMPARATIVO DOS MÉTODOS DE DETECÇÃO DE MUDANÇAS EM SÉRIES NÃO LINEARES

Danielle Mayumi Campos Tamaki (UNIFEI)
danielletamaki@gmail.com

Danillo Lopes Nunes (UNIFEI)
dlopesn13@gmail.com

Luiz Carvalho Neto (UNIFEI)
luiz.carvalho.neto@hotmail.com

Pedro Paulo Balestrassi (UNIFEI)
ppbalestrassi@gmail.com



Este artigo apresenta o estudo comparativo dos métodos de detecção de mudanças em séries não lineares. Os métodos estudados são as cartas de controle de soma acumulada (CUSUM), carta de controle exponencialmente ponderada (EWMA) e os Tracking Signals. O objetivo deste estudo é verificar qual método possui um desempenho melhor na detecção de mudanças em séries não lineares. Neste presente estudo foram realizados testes nestes três métodos com a mistura de séries lineares e não lineares no ponto determinado. O resultado destes testes comparou a porcentagem de acerto e em qual ponto da série foi encontrada a mudança da série. A conclusão presente no artigo aponta que o Tracking Signal possui um desempenho superior aos demais métodos.

Palavras-chave: CUSUM, EWMA, Tracking signal, cartas de controle, controle de qualidade, séries não lineares

1. Introdução

Atualmente a qualidade já se tornou parte em todos os produtos e serviços e continua passando por etapas importantes para continuar a criar métodos para seu controle e sua melhoria. A cada dia mais, o processo pelo desenvolvimento da qualidade ganha maior importância. Para auxiliar neste processo de desenvolvimento da qualidade estão os conceitos estatísticos com suas ferramentas e métodos. Assim, umas das principais ferramentas estatísticas para melhoria e controle da qualidade são chamadas cartas de controle. Desenvolvidas por Walter A. Shewhart em 1924, é considerado o começo do controle estatístico de qualidade.

Para fazer um controle e monitoramento da qualidade em um processo de produção é necessário utilizar ferramentas estatísticas para a detecção, análise e identificação das causas que são responsáveis e que geram variações no comportamento do processo. Os gráficos de controle estatístico como as cartas de controle de Soma Acumulada (CUSUM) e as cartas de controle Exponencialmente Ponderada (EWMA), que são estudados neste artigo, são os meios que revelam quando um processo se altera inadequadamente e necessita de uma ação corretiva. Estes gráficos, também chamados de cartas de controle, são ferramentas eficientes que permitem a redução das variabilidades de qualidade dos produtos. Ou seja, estes gráficos contribuem para a melhoria da qualidade, produtividade, confiabilidade e, conseqüentemente, do custo do que se produz.

Segundo Montgomery (2000) a popularidade dos gráficos de controle se dá porque são técnicas comprovadas de melhorias da produtividade, são eficazes em prevenção de possíveis defeitos, ajudam a evitar ajustes desnecessários nos processos, e fornecem confiáveis informações sobre o desempenho e capacidade dos processos para eventuais diagnósticos.

Muitos são os trabalhos que envolvem as análises sobre estes gráficos de controle, principalmente sobre o comparativo das cartas de controle CUSUM e EWMA. Entretanto, outro método que vem ganhando destaque são os *Tracking Signals* (TS). A grande diferença entre as cartas de controle e os *Tracking Signals* é que estes não estão disponíveis em softwares e não utilizam gráficos para a visualização das anormalidades. Entretanto, Jardim (2013) realizou uma comparação entre estes

métodos e mostrou que para as situações apresentadas no estudo, o método *Tracking Signal* apresentava ser mais eficiente na detecção de viés que os demais métodos.

Para realizar este estudo os procedimentos adotados iniciaram com a simulação de séries normais geradas aleatoriamente e séries não lineares que foram geradas de acordo com certos modelos propostos. Foi realizada uma junção nestas séries em um ponto estratégico estabelecido para que fosse possível o estudo das mesmas com os métodos de controle. Os resultados obtidos, de cada um dos métodos constituem no objeto de estudo para comparar estas técnicas estatísticas, mostrando os pontos fortes e fracos de cada um dos métodos.

2. Carta de Controle CUSUM

Muito utilizadas e estudadas por diversos autores, as cartas de controle de soma acumulada (CUSUM), foram, no início, propostas por Page (1954) na Inglaterra. Este trabalho se baseia na definição dada por Montgomery (1996) e Montgomery, Jennings e Kulahci (2011). Em seu trabalho, Montgomery (1996) apresenta as cartas CUSUM aplicadas à média e à variabilidade do processo, além disso, apresenta a possibilidade de projetar procedimentos da carta para outras variáveis estatísticas, como desvio padrão e amplitude de subgrupos, variáveis de Poisson e binomial em modelos e processos contínuos.

Segundo Montgomery, Jennings e Kulahci (2011) a carta de controle de soma acumulada é muito eficaz na detecção de mudanças da variável que é monitorada. Seu funcionamento se dá através da acumulação de desvios dos erros de previsão que estão acima do valor-alvo T desejado ou abaixo. Este valor-alvo desejado possui estatísticas C+ e C- que são os chamados CUSUMs superior e inferior. Ainda segundo os autores, estes valores são calculados seguindo as equações (1) e (2).

$$C_t^+ = \max[0, e_t(1) - (T + K) + C_{t-1}^+] \quad (1)$$

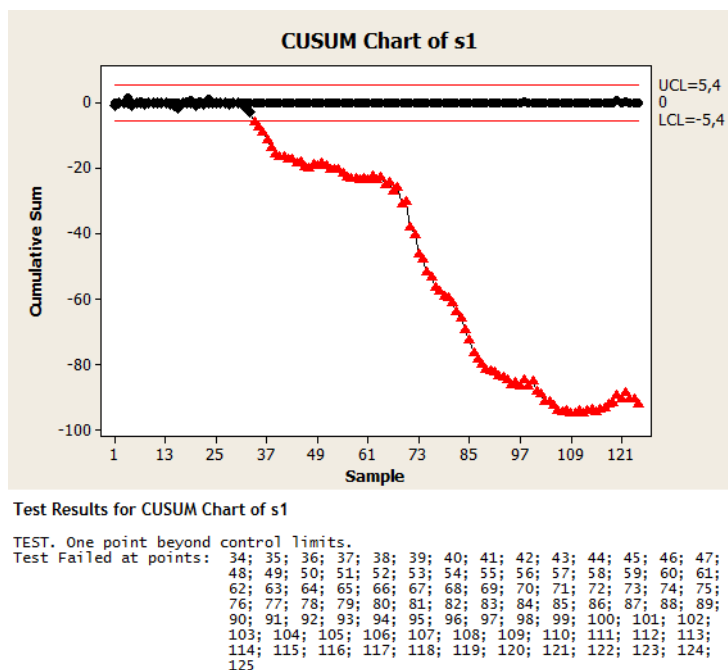
$$C_t^- = \min[0, e_t(1) - (T + K) + C_{t-1}^-] \quad (2)$$

De acordo com estas equações a constante K é chamado de valor de referência. Segundo os autores ela é normalmente escolhida como $k = 0.5\sigma_{e(1)}$ e $\sigma_{e(1)}$. É o desvio padrão dos erros.

A lógica segundo Montgomery, Jennings e Kulahci (2011) é que se os erros começam a cair de um lado do valor alvo, um dos CUSUMs aumenta em magnitude gerando um sinal de falta de controle. A regra de decisão é para sinalizar se a estatística C+ excede um intervalo de decisão $H = 5\sigma_{e(1)}$ ou se C- excede -H. Este sinal indica que a série possui anormalidades e não está satisfatória.

A Figura 1 traz um exemplo da análise CUSUM realizada em uma das séries temporais que serão estudadas nesta pesquisa.

Figura 1 – Resultado de uma análise CUSUM realizada para uma série temporal



De acordo com a Figura 1, o teste realizado mostra que a série possui anormalidades a partir do ponto 34. Todos pontos em vermelho revelam que estão fora do padrão da série. O valor alvo T é igual a zero e os valores UCL e LCL são os chamados CUSUMs superior e inferior.

Entre outros autores que estudaram as cartas de controle de soma acumulada, destacam-se Bissel (1969) que considerou o método e sua relevância em controle de qualidade propondo técnicas que facilitassem sua aplicação em situações mais práticas. Conforme Johnson e Bagshaw (1974), para observações não independentes o teste CUSUM não é apropriado. Hawkins e Olwell (1998) propuseram os gráficos como ferramentas que são essenciais para os profissionais de qualidade

para descobrir e diagnosticar as trocas persistentes em processos. Já Hawkins (1981) propõe uma nova técnica que emprega o mesmo procedimento do gráfico CUSUM para controlar a variabilidade.

3. Carta de controle EWMA

Tendo um desempenho aproximadamente equivalente a carta CUSUM, a carta de controle exponencialmente ponderada (EWMA) é usada, tipicamente, em observações individuais. Montgomery (1996) apresentou as cartas EWMA como alternativa para as cartas de controle de Shewhart, na detecção de pequenas mudanças. A média móvel ponderada é definida como na equação (3):

$$Z_i = \lambda x_i + (1 - \lambda)Z_{i-1} \quad (3)$$

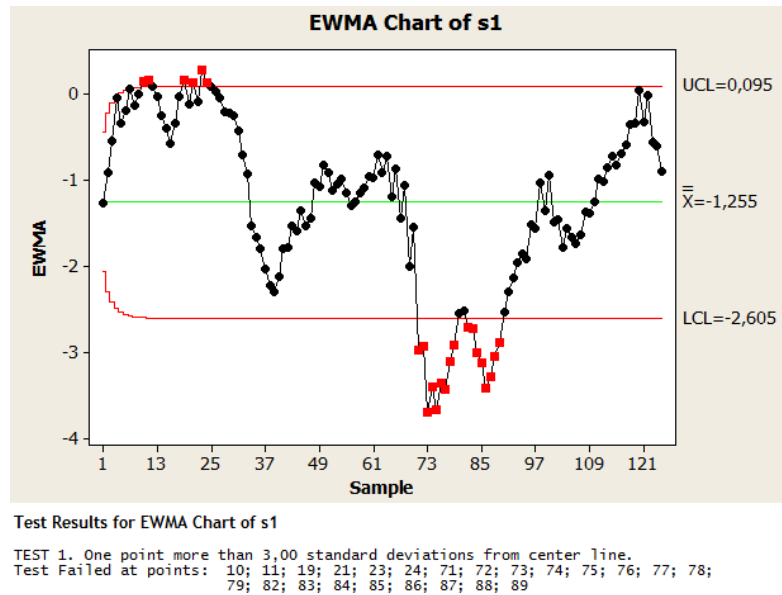
Onde $0 < \lambda \leq 1$ é uma constante e o valor inicial (para primeira amostra $i = 1$) é que se pretende para o processo, tal que $Z_0 = \mu_0$. Segundo Montgomery (1996), a carta EWMA é eficiente em situações de pequenas mudanças no processo.

Esta carta de controle foi introduzida inicialmente por Roberts (1959), porém, muitos outros autores apresentaram suas contribuições para esta carta. Hunter (1986) apresenta a diferença entre as três cartas de controle, Shewhart, CUSUM e EWMA. Segundo o autor a carta de Shewhart depende inteiramente do ponto demarcado por último, já a CUSUM atribui peso praticamente igual para toda a sequência de dados e a carta EWMA concede o maior peso para as mais atuais informações e menor para as mais longínquas.

Outro estudo comparativo das cartas de controle CUSUM e EWMA foi realizado por Lucas e Saccucci (1990) que concluíram que ambos os métodos possuíam resultados similares, porém a carta de EWMA possuíam ARL menor que a CUSUM algorítmica.

A Figura 2 mostra a análise EWMA realizada em uma das séries temporais que serão apresentadas nesta pesquisa.

Figura 2 – Resultado de uma análise EWMA realizada para uma série temporal



Conforme a Figura 2, o teste realizado mostra que a série possui anormalidades a partir do ponto 10. Todos os pontos em vermelho revelam que estão fora do padrão da série.

4. Tracking Signal

A primeira proposta da formulação denominada *Tracking Signal* foi realizada por Brown (1963). Ela foi definida como sendo o coeficiente entre a soma cumulativa do erro (EAC) e o desvio absoluto médio do erro (DAM), que é calculada a cada período, segundo a equação (4):

$$TS_t = \frac{EAC_t}{DAM_t} \quad (4)$$

Além disso, a soma cumulativa de erros é atualizada a cada período pela expressão (5)

$$EAC_t = e_t + EAC_{t-1} \quad (5)$$

Onde e é o erro do período t . O valor inicial de EAC é definido como zero. O cálculo do valor inicial de DAM, assim como Brown (1962) e Gardner (1983 e 1985) é realizada com o valor esperado em função do valor real do desvio padrão da componente aleatória da série temporal. Este cálculo não utiliza os dados históricos. Os limites de controle para TS são independentes do período de tempo. Segundo Corrêa *et al.* (1997) os valores devem se distribuir em torno de zero, não ultrapassando o intervalo de $\pm 3 e \pm 4$. Neste trabalho foi selecionado o limite de ± 4 .

Entre outros autores que estudaram os *Tracking Signals*, Gardner (1985) sugere substituir o DAM pela raiz quadrada dos erros quadráticos médio. Outro autor que sugere uma nova mudança é Trigg (1964) onde o *Tracking Signals* seria o coeficiente entre o erro de previsão amortecido exponencialmente e o desvio absoluto médio, calculado a cada período de tempo.

Segundo Montgomery, Jennings e Kulahci (2011), CUSUM e EWMA estão disponíveis em softwares de estatística, como o Minitab® e fornecem a plotagem e periodicidade para examinar visualmente os gráficos de erros, já os *Tracking Signals* não possuem estas facilidades.

5. Séries temporais lineares e não lineares

Segundo Morettin e Toloí (1985), uma série temporal é uma sequência de dados que são obtidos em intervalos regulares de tempo durante um período específico. Estes dados surgem de diversos campos do conhecimento como medicina, economia, epidemiologia, meteorologia, engenharia, entre outros. A linearidade, segundo Enders (1995), Morettin e Toloí (2006), se refere a média e variância que são constantes ao longo do tempo. Portanto, séries não lineares são o oposto de séries lineares, pois podem variar com relação à média ou com relação a variância. Por se tratar de um problema muito comum em séries lineares é que foram escolhidas as séries não lineares para o presente estudo. Segundo Enders (1995), desconsiderar fatores não lineares pode resultar em diversos erros graves, como erros probabilísticos falhos e estimativas tendenciosas.

Para o presente trabalho foram selecionadas séries não lineares já implementadas e estudadas por Balestrassi *et al.* (2009). Segundo os autores, os oito modelos de séries temporais foram escolhidos porque representam uma variedade de problemas que possuem diferentes características. A Tabela 5.1 apresenta os oito modelos não lineares e suas respectivas equações.

Tabela 5.1 - Modelos não lineares

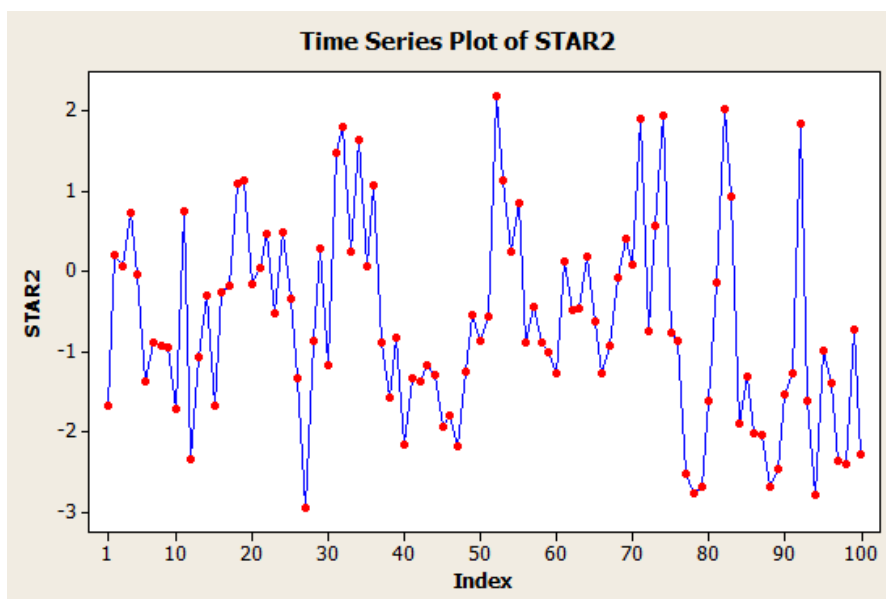
Modelo não linear	Equação
Sinal autor regressivo (SAR)	$y_t = \text{sinal}(y_{t-12}) + \varepsilon_t, \text{ onde } \text{sinal}(x) = 1, 0, -1 \text{ se } x > 0, x = 0, x < 0, \text{ respectivamente}$
Bilinear (BL1)	$y_t = 0.7y_{t-1}\varepsilon_{t-2} + \varepsilon_t$
Bilinear (BI2)	$y_t = 0.4y_{t-1} - 0.3y_{t-2} + 0.5y_{t-1}\varepsilon_{t-1} + \varepsilon_t$
Limiar autor	$y_t = 0.9y_{t-1} + \varepsilon_t \text{ para } y_{t-1} \leq 1 \text{ e}$

regressivo (TAR)	$y_t = -0.3y_{t-1} - \varepsilon_t \text{ para } y_{t-1} > 1$
Não linear autor regressivo (NAR)	$y_t = \frac{0.7 y_{t-1} }{ y_{t-1} + 2} + \varepsilon_t$
Não linear média móvel (NMA)	$y_t = \varepsilon_t - 0.3\varepsilon_{t-1} + 0.2\varepsilon_{t-2} + 0.4\varepsilon_{t-1}\varepsilon_{t-2} - 0.25\varepsilon_{t-2}^2$
Transição suave autor regressiva (STAR1)	$y_t = 0.8y_{t-1} - 0.8y_{t-1}[1 + \exp(-10y_{t-1})]^{-1} + \varepsilon_t$
Transição suave autor regressiva (STAR2)	$y_t = 0.3y_{t-1} + 0.6y_{t-2} + (0.1 - 0.9y_{t-1} + 0.8y_{t-2})[1 + \exp(-10y_{t-1})]^{-1} + \varepsilon_t$

Fonte: Adaptado de Balestrassi *et al.* (2009)

Um gráfico dos modelos mencionados é mostrado logo abaixo na Figura 3. Nesta imagem é plotado o gráfico do modelo STAR 2 da série gerada para a análise deste estudo. A série foi gerada de acordo com a equação apresentada na Tabela 5.1.

Figura 3 – Gráfico da série não linear STAR2



6. Metodologia

Para a realização deste estudo, foram realizadas as seguintes etapas:

1º etapa: Geração das séries

Geraram-se três amostras para realizar o estudo. Em cada uma das amostras foram geradas oito séries aleatórias de 25 valores com distribuição normal, com média 0 e desvio padrão igual a 1 através do software Minitab®. Além disso, foram geradas também as oito séries de 100 valores seguindo os modelos de séries não lineares apresentados anteriormente. Com isso, foram geradas um total de 24 séries aleatórias com distribuição normal e 24 séries não lineares.

Adicionou-se a cada uma das séries aleatórias as séries não lineares. Portanto, cada uma das três amostras contém apenas 8 séries e cada uma dessas séries é formada por uma parte linear (os 25 primeiros valores gerados pela distribuição normal) e uma parte não linear (100 últimos valores gerados pelo modelo não linear)

2º etapa: Análise das séries

Para realizar as devidas análises, foi utilizado o software Minitab® e o Excel®. Em cada uma das séries das três amostras, realizou-se o estudo através dos métodos CUSUM, EWMA e *Tracking Signal*. Para realizar os testes pelo método CUSUM foi utilizado em todas as séries o valor-alvo igual a zero e pelo método EWMA com peso igual a 0,2. Conforme mencionado anteriormente, o intervalo selecionado para o estudo do *Tracking Signal* foi de ± 4 .

7. Resultados e Conclusão

A proposta deste trabalho foi identificar dentre as cartas de controle apresentadas, CUSUM, EWMA e *Tracking Signal*, qual o melhor método para se detectar as mudanças nas séries com uma maior fidelidade e precisão. Os resultados obtidos com a simulação realizada estão apresentados nas Tabelas 7.1, 7.2 e 7.3.

Tabela 7.1 – Resultados da simulação de cada uma das cartas de controle para Amostra 1

AMOSTRA 1									
Série	EWMA			CUSUM			Tracking Signal		
	Qtd pontos detectados	Primeiro Ponto Detectado	%	Qtd pontos detectados	Primeiro Ponto Detectado	%	Qtd pontos detectados	Primeiro Ponto Detectado	%
STAR1	15	1	15,00%	67	37	67,00%	92	34	92,00%
STAR2	23	10	23,00%	92	34	92,00%	94	32	94,00%
NMA	0	-	0,00%	1	88	1,00%	95	27	95,00%
NAR	3	22	3,00%	25	44	25,00%	77	22	77,00%
TAR	0	-	0,00%	0	-	0,00%	88	18	88,00%
BI1	0	-	0,00%	1	70	1,00%	52	25	52,00%
BI2	8	18	8,00%	19	19	19,00%	81	17	81,00%
SAR	12	42	12,00%	34	43	34,00%	84	21	84,00%

Tabela 7.2 – Resultados da simulação de cada uma das cartas de controle para Amostra 2

AMOSTRA 2									
Série	EWMA			CUSUM			Tracking Signal		
	Qtd pontos detectados	Primeiro Ponto Detectado	%	Qtd pontos detectados	Primeiro Ponto Detectado	%	Qtd pontos detectados	Primeiro Ponto Detectado	%
STAR1	11	29	11,00%	80	45	80,00%	82	44	82,00%
STAR2	37	13	37,00%	80	32	80,00%	97	28	97,00%
NMA	3	81	3,00%	7	81	7,00%	99	23	99,00%
NAR	0	-	0,00%	5	29	5,00%	97	29	97,00%
TAR	0	-	0,00%	0	-	0,00%	58	27	58,00%
BI1	0	-	0,00%	0	-	0,00%	64	34	64,00%
BI2	9	20	9,00%	46	51	46,00%	88	20	88,00%
SAR	9	65	9,00%	47	43	47,00%	82	42	82,00%

Tabela 7.3 – Resultados da simulação de cada uma das cartas de controle para Amostra 3

AMOSTRA 3									
Série	EWMA			CUSUM			Tracking Signal		
	Qtd pontos detectados	Primeiro Ponto Detectado	%	Qtd pontos detectados	Primeiro Ponto Detectado	%	Qtd pontos detectados	Primeiro Ponto Detectado	%
STAR1	21	14	21,00%	89	32	89,00%	96	17	96,00%
STAR2	40	5	40,00%	92	34	92,00%	93	33	93,00%
NMA	0	-	0,00%	0	-	0,00%	62	23	62,00%
NAR	2	48	2,00%	42	44	42,00%	104	22	104,00%
TAR	0	-	0,00%	1	8	1,00%	51	29	51,00%
BI1	4	32	4,00%	27	31	27,00%	109	16	109,00%
BI2	5	28	5,00%	58	27	58,00%	110	13	110,00%
SAR	10	55	10,00%	47	30	47,00%	84	14	84,00%

Com os resultados obtidos, pode-se observar um grande desempenho da carta de controle *Tracking Signal* em encontrar as mudanças nas séries. Para as três amostras analisadas, o *Tracking Signal* retornou em média mais de 80% dos pontos aleatórios presentes nas amostras enquanto que as outras duas cartas de controle apresentaram um resultado bastante inferior em suas análises. Outra característica notada através desta pesquisa é a posição do primeiro ponto detectado pelas cartas de controle. Realizou-se a medição do erro médio entre o primeiro ponto detectado pelos três métodos em cada uma das amostras.

Tabela 7.4 – Medição do erro médio dos métodos para as três amostras.

AMOSTRA 1		
EWMA ABS (erro)	CUSUM ABS (erro)	Tracking Signal ABS (erro)
13,80	23,86	5,25
AMOSTRA 2		
EWMA ABS (erro)	CUSUM ABS (erro)	Tracking Signal ABS (erro)
23,20	20,83	7,13

AMOSTRA 3		
EWMA ABS (erro)	CUSUM ABS (erro)	Tracking Signal ABS (erro)
15,33	8,57	7,63

Os resultados, apresentados na Tabela 7.4, mostraram que o *Tracking Signal* apresenta um melhor desempenho também neste ponto, uma vez que detecta o primeiro ponto bem próximo do ponto esperado, superando desta forma as cartas de controle EWMA e CUSUM.

Apesar da facilidade do uso dos métodos EWMA e CUSUM por estarem presentes em *softwares* conhecidos e a sua visualização em gráficos, conclui-se por fim o melhor desempenho do *Tracking Signal* em comparação com as outras cartas de controle, obtendo uma resposta mais precisa e fiel ao esperado para as amostras analisadas. Sugere-se para trabalhos futuros a comparação do *Tracking Signal* com outros métodos e uma análise estatística para se aprofundar ainda mais nos resultados.

Agradecimentos

Os autores agradecem à CAPES, à FAPEMIG e ao CNPq pelo suporte à pesquisa.

8. Referências

- BALESTRASSI, P. P., POPOVA, E., PAIVA, A. D., e LIMA, J. M. (2009). **Design of experiments on neural network's training for nonlinear time series forecasting.** *Neurocomputing*, 72(4), 1160-1178.
- BISSEL, A.F. (1969) **Cusum Techniques for Quality Control.** *Applied Statistics*, v.18.
- BROW, R.G. **Smoothing, Forecasting and Prediction of Discrete Time Series.** Prentice-Hall, Englewood Cliffs, New Jersey, 1963
- CORRÊA, H.; GIANESI, I.; CAON, M. **Planejamento, Programação e Controle da Produção MRP II / ERP: Conceitos, Uso e Implementação.** São Paulo: Atlas 1997.
- ENDERS, W. **Applied econometric time series.** New York: Wiley, 1995. P 433
- GARDNER JR, E.S. **Automatic Monitoring of Forecasting Errors.** *International Journal of Forecasting*. Vol.2, 1-21, 1983.
- GARDNER JR, E.S. **CUSUM vs Smoothed-Error Forecast Monitoring Schemes: Some Simulation Results.** *Operational Research Society*, 36, 43-47, 1985
- HAWKINS, D.M. **A CUSUM for a Scale Parameter,** *Journal of Quality Technology*, 13(4), p.228-231, 1981.
- HAWKINS, D.M.; OLWELL, D.H; **Cumulative Sum Charts and Charting for Quality Improvement,** *Statistics for Engineering and Physical Science*, Springer,1998.
- HUNTER, J.S. (1986) **The Exponentially Weighted Moving Average.** *Journal of Quality Technology*, v. 18, n. 4.

JARDIM, F. S.; **Técnicas para Detecção de Viés em Previsão de Demanda–Comparação de Desempenhos**, Tese de Doutorado. PUC-Rio, 2013.

JOHNSON, R.A. e BAGSHAW M. **The Effect of Serial Correlation on the Performance of CUSUM Tests**. Technometrics, v.16, n. 1, 1974.

LUCAS, J.M.; SACCUCCI, M.S. **Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements**. Technometrics, v. 32, n. 1, 1990.

MONTGOMERY, D. C. **Introduction to Statistical Quality Control**. 4th Edition, New York: John Wiley, 2000.

MONTGOMERY, D. C., JENNINGS, C. L.; KULAHCI, M.; **Introduction to time series analysis and forecasting** (Vol. 526). John Wiley & Sons, 2011.

MONTGOMERY, D.C. **Introduction to Statistical Quality Control**. 3th. ed. New York: John Wiley, 1996.

MORETTIN, P.A.; TOLOI, C.M.C.; **Análise de séries temporais**. 2a. ed. São Paulo: Edgard Blucher, 2006.

MORETTIN, P.A.; TOLOI, C.M.C.; **Previsão de séries temporais**. 2a. ed. São Paulo:Atual Editora; 1985.

PAGE, E.S.; **Continous Inspection Schemes**, Biometrika, v. 41, 1954.

ROBERTS, S.W.; **Control Chart Tests Based on Geometric Moving Averages**, Technometrics, v. 1, 1959.

TRIGG, D. W. **Monitoring a Forecasting System**. Operational Research Quarterly. Vol. 15, 3, pp. 271-274, 1964.