

GRÁFICOS DE CONTROLE PARA MONITORAR PROCESSOS AUTOCORRELACIONADOS

CONTROL CHARTS TO MONITOR AUTOCORRELATED PROCESSES

Cátia Michele Tondolo¹; Fernanda Maria Müller²; Leandro Cantorski da Rosa³

¹Programa de Pós-Graduação em Engenharia de Produção- PPGE
Universidade Federal de Santa Maria – UFSM – Santa Maria/RS – Brasil
cacamt@gmail.com

²Programa de Pós-Graduação em Engenharia de Produção- PPGE
Universidade Federal de Santa Maria – UFSM – Santa Maria/RS – Brasil
nandamuller90@gmail.com

³Programa de Pós-Graduação em Engenharia de Produção- PPGE
Universidade Federal de Santa Maria – UFSM – Santa Maria/RS – Brasil
leski78@hotmail.com

Resumo

Este documento tem como objetivo empregar gráficos de controle como ferramenta de controle estatístico, para monitorar processos autocorrelacionados. Existem estudos que mostram o efeito negativo da autocorrelação em gráficos de controle usuais denotando a importância do uso de ferramentas adequadas a este tipo de dados. Este estudo, portanto, apresenta resultados de eventos simulados e uma aplicação de gráficos de controle para monitorar o índice médio da inflação brasileira, representado pelo Índice Nacional de Preços ao Consumidor Amplo (IPCA).

Palavras-chave: gráficos de controle, processos autocorrelacionados, IPCA.

Abstract

This document aims to use control charts, as statistical control tool to monitor autocorrelated processes. There are studies that show the negative effect of autocorrelation in usual control charts showing the importance of using suitable tools such data. This study, therefore, presents the results of simulated events and an application to real data, an application of control charts to monitor the average rate of inflation in Brazil, represented by Índice Nacional de Preços ao Consumidor Amplo (IPCA).

Key-words: control chart, autocorrelated process, IPCA.

1. Introdução

O Controle Estatístico do Processo (CEP) é um conjunto de técnicas e ferramentas que permitem monitorar as características da qualidade, garantindo a manutenção do processo dentro de limites preestabelecidos (HENNING, et al. 2014). Dentre as técnicas mais simples e eficientes do CEP têm-se os gráficos de controle (GC). Os gráficos de controle são usados para estabelecer e manter o controle estatístico de um processo (FRANCO, et al. 2014a). Sua análise permite determinar as causas de variação e diferenciar as causas especiais de variação da variabilidade natural do processo (FLOTT, 2012).

As inferências e propriedades estatísticas dos gráficos de controle são baseadas no pressuposto de que os dados seguem distribuição normal e são independente e identicamente distribuídos (i.i.d.), $X \sim N(\mu, \sigma^2)$, sendo X uma característica da qualidade a ser avaliada (PSARAKIS; PAPALEONIDA, 2007). Em situações em que a suposição de normalidade é violada os gráficos de controle tradicionais ainda mostram-se eficientes (MONTGOMERY, 2009). Todavia, isso não é observado quando a suposição de independência é violada. Nesses casos, os gráficos de controle apresentam grande distorção de seu desempenho, observando um aumento na média da taxa de alarmes falsos e redução na capacidade de detectar alterações no processo (MONTGOMERY, 2009; FRANCO, et al. 2014a; PSARAKIS; PAPALEONIDA, 2007).

Em processos industriais a suposição de que as observações são independentes geralmente não é satisfeita, como destacado por Montgomery (2009). Uma das explicações, segundo o autor, é que em sistemas produtivos, para monitorar o processo, as observações são coletadas em intervalos de tempo muito pequenos, o que contribui para que sejam autocorrelacionadas ao longo do tempo.

Uma das abordagens mais usuais para lidar com dados autocorrelacionados é a utilização de estratégias de amostragem, que permitem reduzir os efeitos sobre a sensibilidade do gráfico de controle oriunda da correlação serial entre as observações (FRANCO, et al. 2014b; FRANCO, et al. 2014a). Outra abordagem utilizada é modelar os dados por meio de um modelo de séries temporais que removerá a autocorrelação dos dados e após aplicar gráficos de controle tradicionais ou modificados nos resíduos do modelo (MONTGOMERY, 2009). A justificativa para utilizar gráficos de controle residuais é que modelos corretamente especificados possuem resíduos que se comportam como variáveis aleatórias independente e identicamente distribuídas (PSARAKIS; PAPALEONIDA, 2007). Assim, os pressupostos exigidos para construção dos GC serão satisfeitos. Nos trabalhos de Henning et al. (2014), Montgomery e Mastrangelo (1991), entre outros trabalhos, são encontradas algumas aplicações de modelos de séries temporais para eliminar a dependência dos dados.

Com a finalidade de compreender as características dos gráficos de controle CUSUM (soma acumulada) e EWMA (média móvel exponencialmente ponderada) no âmbito de situações em que os dados apresentam autocorrelação foi realizado um estudo de simulações de Monte Carlo. Foram geradas amostras pseudo-aleatórias com diferentes níveis de autocorrelação. A autocorrelação das séries foi modelada por meio de um modelo ARIMA (modelo autoregressivo integrado de médias móveis) (BOX; JENKIS, 1970). Para analisar o desempenho dos gráficos de controle foi computado o Número Médio de Amostras (NMA) (SANT'ANNA; ten CATEN,2012) até a detecção de uma causa especial. Por último, foi realizada a aplicação do gráfico de controle que demonstrou melhor desempenho, segundo o estudo de simulação realizado, para monitorar o Índice Nacional de Preços ao Consumidor Amplo (IPCA).

Desse modo, com a realização desse trabalho pretende-se: i) Apresentar as alternativas para modelar dados autocorrelacionados; ii) Demonstrar entre os gráficos de controle CUSUM e EWMA qual apresenta melhor desempenho para modelar os resíduos do modelo ARIMA; iii) Monitorar a variabilidade média da inflação brasileira (IPCA) por meio de gráficos de controle.

2 Gráficos de controle

2.1 Modelo ARIMA

Antes de iniciar a construção dos gráficos de controle é necessário investigar se a série de dados atende os pressupostos exigidos para sua construção. Um dos detalhes que deve ser investigado é analisar se existe alguma estrutura de dependência na série de dados. Para detectar a presença de autocorrelação na série podem-se computar os gráficos da função de autocorrelação parcial (FACP) e da função de autocorrelação (FAC) (TSAY, 2010). Outra alternativa é utilizar testes estatísticos como o teste Ljung-Box nos resíduos, onde a hipótese nula é de que o resíduos são i.i.d., contra a hipótese alternativa de que os resíduos não são i.i.d. (LJUNG; BOX, 1978).

Caso os dados demonstrem alguma dependência o desempenho do gráfico de controle pode ser comprometido (HENNING, et al. 2014). Uma das alternativas para eliminar a dependência dos dados é utilizar um modelo estatístico para modelar a dependência e fazer uso dos resíduos deste modelo para monitorar a característica do processo. Nesse trabalho será monitorada somente a média do processo. Uma alteração na média da característica da qualidade dos dados observados se reflete diretamente nos resíduos do modelo ARIMA (BOX; LUCENO, 1997). Isso justifica a utilização de gráficos de controle de resíduos para monitorar o processo de interesse.

Um dos tratamentos sugeridos para remoção da dependência entre a série é o ajuste de um modelo ARIMA. Para a implementação dessa metodologia é necessário que sejam cumpridas as

seguintes etapas: identificação, estimação, verificação e previsão (para maiores detalhes sobre as etapas sugere-se consultar as obras de Brockwell e Davis (1991), Morettin e Toloí (2004), Box, Jenkins e Reinsel (2008) e Tsay (2010)).

Uma das suposições necessárias para realizar um modelo de séries temporais é que a série de dados seja estacionária, isto é, possua média e variância constante ao longo do tempo. Quando essa suposição é atendida, para modelar a autocorrelação dos dados, utiliza-se um modelo ARMA (autoregressivo e de médias móveis). O modelo ARMA (p,q) é definido como (BOX; JENKINS; REINSEL, 2008):

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \dots + \phi_p \tilde{z}_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}, \quad t = 0, 1, 2, \dots, \quad (1)$$

ou em sua forma compactada por:

$$\phi(B)\tilde{z}_t = \theta(B)a_t \quad (2)$$

em que, \tilde{z}_t é a série de dados que está sendo modelada no período t , ϕ parâmetro autorregressivo, θ parâmetros de médias móveis do modelo e a é o termo aleatório com média zero e variância constante.

Quando a série não atende o pressuposto de estacionariedade é necessário realizar uma transformação nos dados para torná-la estacionária. Para estacionarizar a série geralmente se utiliza a diferenciação dos dados. Assim, o modelo passa a ser denominado de ARIMA (p,d,q), em que d corresponde ao número de diferenças necessárias para que a série torne-se estacionária. Morettin e Toloí (2004) argumentam que o grau de diferenciação dos modelos ARIMA é um ou dois.

O modelo ARIMA é definido por (BOX; JENKINS; REINSEL, 2008):

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \dots + \phi_p \tilde{z}_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}, \quad (3)$$

ou em sua forma compactada por ser definido por:

$$\phi(B)\nabla^d \tilde{z}_t = \theta(B)a_t \quad (4)$$

onde ∇^d representa a diferenciação da série e a_t é o termo aleatório com média zero e variância constante, $a_t \sim RB(0, \sigma^2)$. Desse modo, o modelo ARIMA, incorpora termos autorregressivos (AR), filtro de integração (I) e termos de Médias Móveis (MA).

2.2 Gráfico de controle de CUSUM

O gráfico de controle CUSUM pode ser utilizado para monitorar a média do processo. Esse GC incorpora diretamente toda a informação na sequência de valores da amostra, plotando as somas cumulativas dos desvios dos valores da amostra de um valor alvo. O CUSUM tabular trabalha acumulando desvios de μ_0 que estão acima do valor alvo, com uma estatística C^+ , e acumulando desvios de μ_0 que estão abaixo do valor alvo, com outra estatística C^- (MONTGOMERY, 2009). As

estatísticas C^+ e C^- são chamadas CUSUM's unilaterais superior e inferior, respectivamente. Calculam-se através das Equações 5 e 6:

$$C_i^+ = \text{máx}[0, x_i - (\mu_0 + K) + C_{i-1}^+], \quad (5)$$

$$C_i^- = \text{máx}[0, (\mu_0 + K) - x_i + C_{i-1}^-], \quad (6)$$

onde os valores iniciais são $C_0^+ = C_0^- = 0$, K é o valor de tolerância ou folga e μ_0 é o valor alvo. K é dado por:

$$K = \frac{|\mu_1 - \mu_0|}{2}, \quad (7)$$

em que, μ_1 é o valor da média fora de controle.

Quanto menor o valor obtido de K , menor será a faixa de variação que o gráfico será capaz de detectar e maior será a sensibilidade do gráfico.

2.3 Gráfico de controle EWMA

O gráfico de controle EWMA foi introduzido por Roberts (1959). Esse método é utilizado para previsão e aplicação em séries temporais. Os gráficos de controle EWMA foram apresentados principalmente com o intuito de incorporar as limitações dos gráficos Shewhart. No EWMA o peso das observações decai à medida que ficam mais antigas. λ é o fator utilizado para realizar essa ponderação. O GC EWMA é definido por (MONTGOMERY, 2009):

$$z_i = \lambda x_i + (1 + \lambda)z_{i-1}, \quad i = 1, \dots, t, \quad (8)$$

em que, $0 < \lambda \leq 1$ é uma constante de amortecimento, x é a variável. Para o valor alvo utiliza-se $z_0 = \mu_0$.

Os limites de controle considerando o período $i=1$, são definidos pelas Equações 9, 10 e 11:

$$LSC = \mu_0 + L\sigma \sqrt{\frac{\lambda}{2-\lambda} [1 - (1-\lambda)^{2i}]}, \quad (9)$$

$$LC = \mu_0, \quad (10)$$

$$LIC = \mu_0 - L\sigma \sqrt{\frac{\lambda}{2-\lambda} [1 - (1-\lambda)^{2i}]}, \quad (11)$$

onde, σ corresponde ao desvio padrão do processo quando em controle e L é o coeficiente de abertura dos limites.

2.4 Análise de sensibilidade

Quando se avalia a eficiência de um gráfico de controle, espera-se que ele detecte a presença de uma causa especial logo que ela tenha ocorrido. O desempenho de um gráfico de controle costuma ser medido em termos do número de amostras coletadas até a emissão de um sinal

(NMA), ou comprimento médio de sequência (CMS), geralmente detectando uma alteração na média do processo.

Segundo Montgomery (2009) o NMA está relacionado aos erros tipo I e tipo II. O erro tipo I é o número médio de amostras coletadas até que um falso alarme seja disparado quando o processo está sob controle, sendo descrito por NMA_0 . O NMA_0 é dado por:

$$NMA_0 = \frac{1}{\alpha}, \quad (12)$$

onde, α é a probabilidade do erro tipo I, ou seja, a probabilidade do alarme falso. O número médio de amostras coletadas até que seja sinalizada uma situação fora de controle (NMA_1) é definido como:

$$NMA_1 = \frac{1}{1-\beta}, \quad (13)$$

em que, β é a probabilidade do erro do tipo II.

3 Procedimentos metodológicos

O estudo computacional foi realizado por meio de simulações de Monte Carlo. A rotina de programação foi implementada no *software* livre *R* (R Core Team, 2014). Para realizar os gráficos de controle utilizaram-se as funções disponíveis no pacote *qcc* (*Quality Control Charts*) (SCRUCCA, 2004). Para as simulações de Monte Carlo foram consideradas 5000 réplicas e os tamanhos amostrais utilizados foram $n = 50, 100, 200$ e 1000 . Para computar a autocorrelação da série de dados utilizaram-se diferentes valores para o parâmetro autoregressivo. Os valores do parâmetro AR foram iguais a $\phi = -0.8, -0.5, -0.2, 0.2, 0.5$ e 0.8 .

Para cada réplica de Monte Carlo simulou-se uma série temporal com determinado tamanho (n) e determinado grau de correlação (ϕ). Após os dados serem gerados ajustou-se um modelo ARIMA e obtiveram-se os resíduos do modelo ajustado. Com base nos resíduos do modelo ARIMA ajustaram-se os gráficos de controle EWMA e CUSUM. Conforme recomendado por Kandananond (2014) os dados foram obtidos por meio de um processo não estacionário ARIMA (1,1,0) e um processo estacionário ARIMA (1,0,0). No entanto, processos estacionários são comumente encontrados no ambiente de produção (YU; LIU, 2011). Por fim, para analisar o desempenho dos gráficos de controle foram computados os valores de probabilidade de erro do tipo I e do NMA_0 por meio da Equação 12. Nesse trabalho não foi analisado o NMA_1 .

Por último foi realizada uma aplicação do gráfico de controle que apresentou o melhor desempenho no estudo de simulação, na série de dados do Índice Nacional de Preços ao Consumidor Amplo, para o período de janeiro de 1995 a abril de 2015, totalizando 244 observações mensais. Essas observações foram coletadas no site Ipeadata (<http://www.ipeadata.gov.br/>). Para analisar se a série do índice IPCA atende ao pressuposto de estacionariedade realizou-se o teste

ADF aumentado de Said e Dickey (1984) e o teste KPSS de Kwiatkowski, Phillips, Schmidt e Shin (1992). A hipótese nula do teste ADF aumentado indica presença de raiz unitária (série não estacionária) e o teste KPSS tem como hipótese nula estacionariedade de uma série temporal. Para seleção das ordens do modelo ARIMA utilizou-se a FAC e a FACP. Os gráficos da FAC e da FACP serão omitidos por questão de brevidade. Para auxiliar na seleção do modelo também foi analisado o critério AIC (AKAIKE, 1974).

Definida a ordem do modelo, ajustaram-se os dados e computaram-se os resíduos. A fim de realizar a análise de adequação do modelo foi realizada a análise dos resíduos. Para isso investigou-se a normalidade dos resíduos por meio do teste de Doornik-Hansen (DOORNIK; HANSEN, 1994), a independência dos resíduos por meio do teste Ljung-Box e ausência de heterocedasticidade por meio do teste Multiplicador de Lagrange (ENGLE, 1982).

4. Resultados numéricos

Nesta seção são apresentados os resultados numéricos do estudo de simulação de Monte Carlo realizado. Para demonstrar os problemas acarretados ao realizar um gráfico de controle em que a série apresenta indícios de correlação serial, foi construído um gráfico de controle EWMA de uma série autocorrelacionada. O cenário correspondente foi computado com $\phi = 0.8$. Henning, et al. (2014) observa esse nível de autocorrelação na turbidez da água, ao construir gráficos de controle com os resíduos do modelo ARIMA. Desse modo, é um cenário plausível de ser encontrado em dados em que são aplicados GC para monitorar o processo. A Figura 1 apresenta o gráfico de controle EWMA da série temporal simulada sem ajuste do modelo ARIMA. Para construir o GC foi utilizada uma série de tamanho amostral igual a $n = 50$. Conforme observado, para o nível de autocorrelação utilizado $\phi=0.8$, a maior parte das observações está fora de controle no gráfico elaborado. Dessa forma, nota-se que ao realizar GC's com observações que apresentam correlação serial são observados altas taxas de alarmes no processo, como apontado por Psarakis e Papaleonida (2007). Isso inviabiliza a análise e a construção de gráficos para monitorar o processo. Uma alternativa sugerida na literatura para séries que apresentam níveis de autocorrelação significativos é modelar o gráfico com resíduos de um modelo de série temporal ou ainda utilizar GC para séries com esse problema.

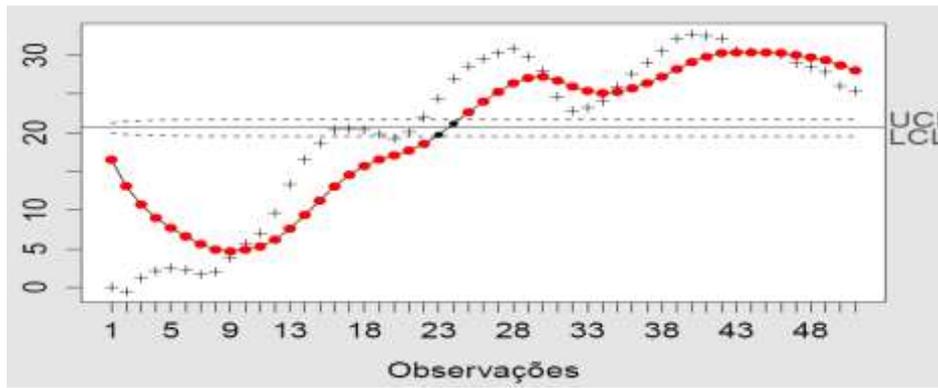


Figura 1 - Gráfico de controle EWMA da série simulada para $\phi=0.8$ e $n=50$. No gráfico UCL corresponde ao limite superior de controle, LCL é o limite inferior de controle, os pontos em preto correspondem às observações sob controle e os pontos em vermelho são as observações fora de controle.

Nas Figuras 2 e 3 são apresentados os gráficos de controle EWMA e CUSUM ajustados por meio dos resíduos do modelo ARIMA. As primeiras 25 observações dos resíduos foram usadas para calibrar os limites. O restante dos resíduos foram plotados nos limites construídos com as primeiras observações. Para realizar uma análise gráfica da performance dos gráficos na observação correspondente a 60% da amostra foi inserido um ponto aberrante (*outlier*). Objetiva-se com isso identificar visualmente se os gráficos conseguem detectar esse ponto fora de controle.

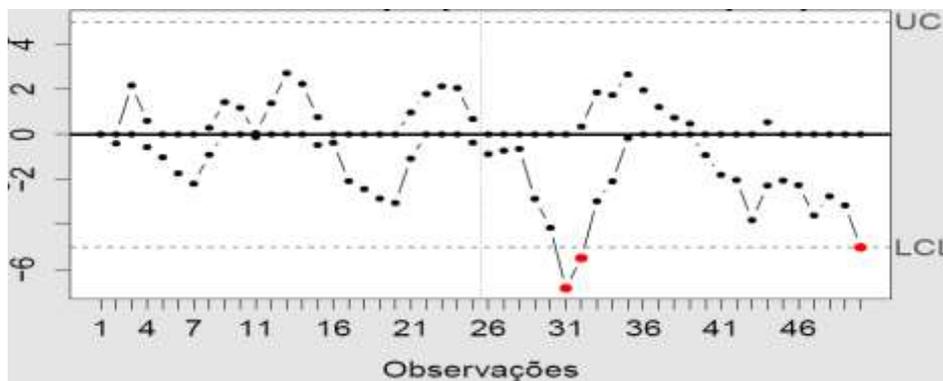


Figura 2 - Gráfico de controle CUSUM para os resíduos do modelo ajustado para a série simulada com $\phi=0.8$ e $n=50$. No gráfico UCL corresponde ao limite superior de controle, LCL é o limite inferior de controle, os pontos em preto correspondem às observações sob controle e os pontos em vermelho são as observações fora de controle.

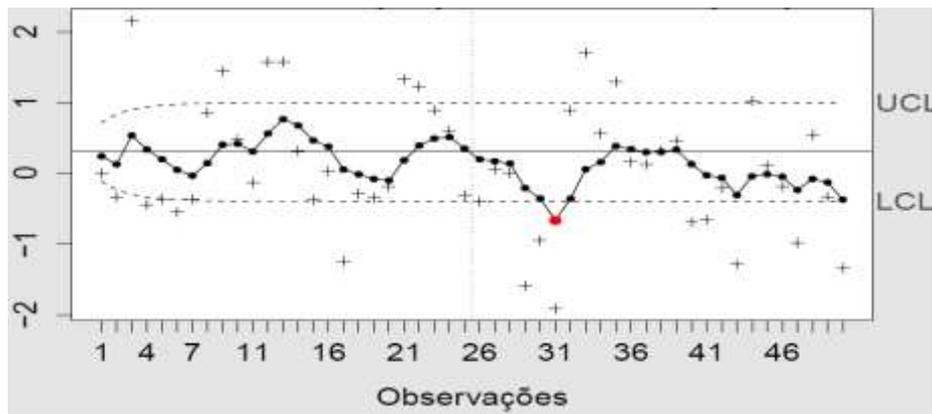


Figura 3 – Gráfico de controle EWMA para resíduos do modelo ajustado para a série simulada com $\phi=0.8$ e $n=50$. No gráfico UCL corresponde ao limite superior de controle, LCL é o limite inferior de controle, os pontos em preto correspondem as observações sob controle e o ponto em vermelho é a observação fora de controle.

Ao analisar as Figuras 02 e 03 percebe-se que tanto o gráfico CUSUM, quanto o EWMA identificaram o ponto que foi colocado fora de controle. Um interessante resultado verificado é que o gráfico CUSUM identifica uma maior quantidade de pontos fora de controle. Ao passo que o GC EWMA identificou apenas o ponto que foi inserido. Desse modo, pela análise gráfica pode-se concluir que o gráfico EWMA demonstrou ser mais eficiente para detectar alterações no processo.

Dando seguimento, a Tabela 1 apresenta os resultados da análise de sensibilidade dos GC's ajustados pelos resíduos, para os diferentes valores de ϕ e tamanhos amostrais (n). Foi analisado o número médio de amostras até detecção de um falso alarme para o gráfico de controle EWMA (NMA_0) e para o gráfico de controle CUSUM (NMA_{0b}). Além do erro tipo I. Para um GC apresentar melhor desempenho que seu concorrente, sendo gerado sob mesmas condições, este irá apresentar um maior valor para o NMA_0 , ou seja, menor será a probabilidade de alarmes falsos quando o processo está sob controle (MONTGOMERY, 2009). Ao analisar a Tabela 1, percebe-se que em todos os casos o NMA_0 do gráfico EWMA foi maior, indicando que esse gráfico apresenta melhor desempenho. Nota-se que os melhores valores, na maioria dos casos, são observados para baixos níveis de autocorrelação ($\phi = -0,2$ e $0,2$). Além disso, a performance dos gráficos melhora a medida que o tamanho amostral aumenta. Para o processo não estacionário gerado (ARIMA(1,1,0)) observa-se melhores valores de NMA , principalmente para o gráfico EWMA, comparado ao processo estacionário gerado (ARIMA(1,0,0)) para os tamanhos amostrais 50 e 100. Por exemplo, para $\phi=0,5$ e $n=50$, percebe-se que o proceso gerador estacionário resultou em um $NMA_0=179,212$ e o processo gerador não estacionário em um $NMA_0=233,863$. Outro detalhe observado, é que a performance dos gráficos de controle, na maioria dos casos, é melhor para os processos gerados com coeficientes de correlação positivo.

Tabela 1 – Valores de NMA dos gráficos de controle EWMA e CUSUM gerados por meio dos resíduos do modelo ARIMA e suas probabilidades de erro do tipo I, para diferentes valores de ϕ e de n

N		ARIMA (1,1,0)				ARIMA (1,0,0)			
		NMA_0	NMA_{0b}	α_0	α_{0b}	NMA_0	NMA_{0b}	α_0	α_{0b}
50	-0,8	140,528	57,287	0,007	0,017	133,547	54,478	0,007	0,018
	-0,5	146,113	61,020	0,007	0,016	154,512	63,613	0,006	0,016
	-0,2	177,054	67,385	0,006	0,015	181,291	71,388	0,006	0,014
	0,2	218,723	87,230	0,005	0,011	187,688	72,275	0,005	0,014
	0,5	233,863	94,697	0,004	0,011	179,212	73,228	0,006	0,014
	0,8	165,673	66,507	0,006	0,015	129,534	51,706	0,008	0,019
100	-0,8	212,947	73,303	0,005	0,014	203,749	65,686	0,005	0,015
	-0,5	238,095	76,864	0,004	0,013	233,754	81,222	0,004	0,012
	-0,2	253,678	84,904	0,004	0,012	259,471	91,025	0,004	0,011
	0,2	316,056	106,724	0,003	0,009	283,930	97,276	0,004	0,010
	0,5	325,945	122,369	0,003	0,008	293,427	105,619	0,003	0,009
	0,8	293,083	120,453	0,003	0,008	228,311	87,474	0,004	0,011
200	-0,8	283,930	97,286	0,004	0,010	298,686	97,172	0,003	0,010
	-0,5	307,503	104,976	0,003	0,010	324,570	105,664	0,003	0,009
	-0,2	337,041	113,045	0,003	0,009	319,387	107,945	0,003	0,009
	0,2	353,232	118,078	0,003	0,008	323,206	115,009	0,003	0,009
	0,5	376,081	138,389	0,003	0,007	344,353	124,860	0,003	0,008
	0,8	347,947	137,533	0,003	0,007	309,406	120,934	0,003	0,008
1000	-0,8	351,815	124,273	0,003	0,008	350,877	126,451	0,003	0,008
	-0,5	357,424	126,804	0,003	0,008	360,698	128,783	0,003	0,008
	-0,2	373,580	137,351	0,003	0,007	368,759	133,422	0,003	0,007
	0,2	374,869	136,983	0,003	0,007	371,333	135,336	0,003	0,007
	0,5	381,127	139,322	0,003	0,007	368,189	134,644	0,003	0,007
	0,8	372,940	137,859	0,003	0,007	359,609	134,459	0,003	0,007

Nota: α_0 e α_{0b} são a probabilidade de erro do tipo I para o EWMA e CUSUM.

Com base no estudo de simulação realizado, nota-se que o gráfico de controle EWMA residual possui um melhor desempenho em relação ao número médio de amostras coletadas até que um falso alarme seja disparado, para os diferentes tamanhos amostrais e valores de ϕ . Esse resultado vai ao encontro das evidências apresentadas por Kandananond (2014) e Santos (2012), que também encontram indícios de superioridade do gráfico EWMA. Desse modo, é sugerido que para processos produtivos que normalmente apresentam autocorrelação, como a fabricação de alimentos, produtos químicos, papel e outros produtos de madeira (PSARAKIS; PAPALEONIDA, 2007), a utilização do gráfico EWMA construído por meio dos resíduos do ARIMA. Em vista de o gráfico EWMA ter demonstrado maior acurácia, na próxima seção, foi realizada uma aplicação do gráfico para monitorar a média da inflação brasileira.

5. Aplicação da metodologia a dados do IPCA

Nesta seção, foi monitorado o índice médio do IPCA por meio do gráfico de controle EWMA. O IPCA é o índice oficial da inflação utilizado pelo Banco Central do Brasil para acompanhar o sistema de metas da inflação. Os dados utilizados nesse estudo correspondem ao período de janeiro de 1995 a abril de 2015, totalizando 244 observações mensais.

Para atender os princípios de estacionariedade realizou-se a primeira diferença do índice IPCA, dado por $VIPCA_t = \frac{IPCA_t}{IPCA_{t-1}}$, em que $IPCA_t$ é o índice do IPCA em t e $t-1$. Após transformação dos dados, realizou-se o teste KPSS e ADF aumentado com a finalidade de verificar se a primeira diferença do índice é estacionária. O valor do teste KPSS foi igual a 0,033 (p -valor = > 0.1) e o valor do teste ADF aumentado foi -8,982 (p -valor = $< 0,01$). Com base em ambos os testes verificou-se que a série transformada é estacionária. Ressalta-se que a série em nível não apresentou média e variância constante ao longo do tempo (estacionária) de acordo com os testes, por isso foi realizada a transformação do índice.

Antes de construir os gráficos de controle foi inspecionado se a série apresentava autocorrelação. Inicialmente calculou-se o teste Ljung-Box da série transformada para os 10 primeiros *lags*. O valor do teste de 34,632 (p -valor = $< 0,001$) indica indícios de presença de autocorrelação. Para eliminar o problema de dependência da série foi ajustado um modelo ARIMA e com os resíduos construído o gráfico de controle. Para seleção da ordem do modelo ARIMA foi utilizado a FAC e a FACP. De maneira auxiliar para seleção da ordem do modelo ARIMA fez-se uso do critério de seleção AIC, onde o modelo selecionado é aquele que possui um menor AIC, bem como a significância dos estimadores. O modelo selecionado está apresentado na Tabela 2.

Tabela 2 - Modelo ARIMA (1,1,2) para o IPCA que compreende o período de janeiro de 1995 a abril de 2015

	Intercepto	ϕ_1 (AR)	θ_1 (MA)	θ_2 (MA)
Estimador	-0,004	0,449	-0,660	-0,207
p-valor	0,488	$<0,001$	$<0,001$	0,048
AIC				159,346

Com a finalidade de investigar se o modelo foi corretamente ajustado realizou-se a análise dos resíduos. O teste Ljung-Box, com estatística 15,465 (p -valor 0,116), não rejeita a hipótese nula de que os resíduos sejam distribuídos de maneira independente para os 10 primeiros *lags* da série. O teste de normalidade (teste 86,124 e p -valor = $<0,001$) rejeita a hipótese nula de que os resíduos tem distribuição normal. No entanto, como apontado por Montgomery (2009) quando a suposição de

normalidade é violada os gráficos de controle tradicionais ainda mostram-se eficientes. Por último, o teste multiplicador de Lagrange utilizado para analisar a presença de heterocedasticidade nos resíduos, não rejeitou a hipótese nula de ausência de efeito ARCH (estatística de teste 14,777 e p -valor = 0,253). Desse modo, em virtude da análise dos resíduos demonstrar que o modelo foi corretamente ajustado foi construído o gráfico EWMA (média), por meio dos resíduos obtidos pelo modelo ARIMA (1,1,2).

Para construir o gráfico EWMA foram utilizadas as primeiras 157 observações dos resíduos, que correspondem ao período de janeiro de 1995 a fevereiro de 2008, para ajustar os limites do gráfico. O restante das observações foram plotados nos limites calculados. Selecionou-se o período correspondente a março de 2008 a abril de 2015 para plotar nos limites estabelecidos devido o Brasil apresentar nos últimos anos períodos com inflação crescente, principalmente nos anos de 2014 e 2015. Pretende-se com isso investigar se a inflação média permaneceu sobre controle estatístico nesse período. Manter a inflação sobre controle é primordial para o governo, principalmente em períodos com turbulência econômica e financeira, como foi marcado os últimos anos e agravos econômicos apresentados pela economia brasileira. Quando a inflação está fora de controle pode gerar sérios problemas para economia do país, como: desvalorização da moeda nacional, clima econômico desfavorável, aumento da especulação financeira, elevação das taxas de juros, aumento do desemprego e redução de investimentos estrangeiros (BANCO CENTRAL DO BRASIL, 2011).

Ao estimar os limites observou-se que os resíduos 94, 95 e 96 apresentam uma condição não usual em relação aos demais. Para construir os limites do gráfico de controle é necessária uma situação de controle das variáveis. Desse modo, para estabelecer os limites foram retiradas essas observações para construir os limites de controle, que foram utilizados para monitorar o processo médio da inflação. Após estabelecido os limites do GC foram plotadas o restante de observações nos limites ajustados. Na Figura 4 é apresentado o gráfico de controle EWMA (para a média) dos resíduos do modelo ARIMA (1,1,2), estimado com os dados do IPCA.

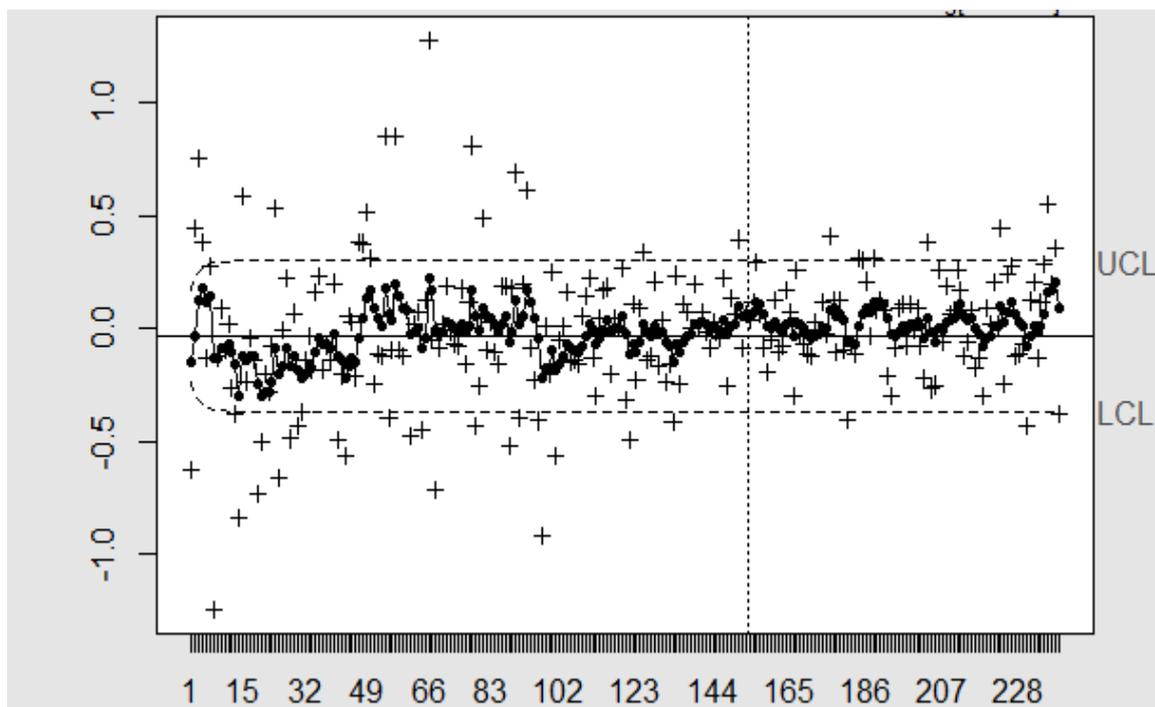


Figura 4 - Gráfico EWMA para os resíduos médios do modelo ARIMA (1,1,2) para o IPCA.

Conforme observado na Figura 4, ao plotar o restante dos resíduos nos limites calibrados não foi observado pontos fora de controle, indicando que nesses meses a inflação esteve sobre controle estatístico. O controle da inflação é de extrema importância, pois pode contribuir para o crescimento dos rendimentos reais da população e para o desenvolvimento do país (BANCO CENTRAL DO BRASIL, 2011). As avaliações sobre as metas da inflação, realizadas pelo Banco Central do Brasil, devem ser contantes para identificar tendências e ciclos de inflação. Para auxiliar nas decisões políticas que almejam controlar períodos com aumento desenfreado dos preços. Sugere-se nesse artigo, o gráfico de controle como uma ferramenta adicional para identificar tendências e períodos de alta inflação na economia brasileira. Com a utilização do GC seria possível acompanhar o percurso da inflação e assim fornecer informações adicionais para a tomada de decisão.

6. Considerações Finais

O presente trabalho teve a finalidade de analisar a performance do gráfico de controle CUSUM e EWMA (média) ajustados com os resíduos do modelo ARIMA, para dados com autocorrelação. Com intuito de alcançar o objetivo estabelecido realizou-se um estudo de simulação de Monte Carlo. Para comparar o desempenho entre os dois gráficos foi analisado o número médio de alarmes falsos e o erro tipo I para os diferentes níveis de autocorrelação e diferentes tamanhos

amostrais. Também foi considerado uma análise gráfica para analisar a performance dos gráficos de controle CUSUM e EWMA na presença de pontos fora de controle estatístico.

Os resultados demonstraram que o gráfico EWMA ajustado com os resíduos do modelo ARIMA apresenta melhor desempenho comparado ao CUSUM. Além disso, percebe-se que os melhores desempenhos são identificados para menores níveis de autocorrelação e em tamanhos amostrais maiores. Por último foi realizada uma aplicação do gráfico EWMA para monitorar a inflação brasileira (índice IPCA). Com a análise foi observado que os gráficos de controle podem ser uma ferramenta adicional para monitorar tendências e ciclos de alta inflação. Para pesquisas futuras sugere-se computar o NMA_1 (erro tipo II) e analisar a performance dos dois gráficos para modelar a dispersão de variáveis autocorrelacionadas.

Referências

- AKAIKE, H. A new look at the statistical model identification. **IEEE Transactions on Automatic Control.**, v.19, n.6, p.716-723,1974.
- BANCO CENTRAL DO BRASIL. Dez anos de metas para a inflação no Brasil 1999-2009. 2011. Disponível em: <http://www.bcb.gov.br/Pec/Metas/10_anos_metas_inflacao_completo.pdf>. Acesso em: 03 fev. 2015.
- BOX, G. E. P.; JENKINS, G. M. **Time series analysis forecasting and control**. Holden-Day, San Francisco, 1970.
- BOX, G. E. P.; JENKINS, G. M.; REINSEL. **Time series analysis forecasting and control**. 4ed. Wiley, 2008.
- BOX, G. P.; LUCENO, A. **Statistical control by monitoring and feedback adjustment**. New York: John Wiley, 1997.
- BROCKWELL, P. J.; DAVIS, R. A. **Time Series: Theory and Methods**. New York, Springer, 1991.
- DOORNIK, J. A.; HANSEN, H. **An omnibus test for univariate and multivariate normality**. Working paper, Nuffield College, Oxford, 1994.
- ENGLE, R.F. Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. **Econométrica**, v. 50, p. 987-1008, 1982.
- FLOTT, L. W. Introduction to control charts. **Metal finishing**, v. 110, n. 6, p. 36–38, 2012.
- FRANCO, B. C.; CASTAGLIOLA, P.; CELANO, G.; COSTA, A. F. B. A new sampling strategy to reduce the effect of autocorrelation on a control chart. **Journal of Applied Statistics**, v. 41, n. 7, p. 1408–1421, 2014a.
- FRANCO, B. C.; CELANO, G.; CASTAGLIOLA, P.; COSTA, A. F. B. Economic design of shewhart control charts for monitoring autocorrelated data with skip sampling strategies. **Int J Production Economics**, v. 151, p. 121–130, 2014b.
- HENNING, E.; WALTER, O. M. C. F.; DE SOUZA, N. S.; SAMOBYL, R. W. Um estudo para a aplicação de gráficos de controle estatístico de processo em indicadores de qualidade da água potável. **Revista Eletrônica Sistemas & Gestão**, v. 9, n. 1, p. 2–13, 2014.

- KANDANANOND, K. Guidelines for applying statistical quality control method to monitor autocorrelated processes. **Procedia Engineering**, v. 69, p. 1449- 1458, 2014.
- KWIATKOWSKI, D.; PHILLIPS, P. C. B.; SCHMIDT, P.; SHIN, Y. Testing the null hypothesis of stationarity against the alternative of a unit root. **Journal of Econometrics**, v.54, p.159–178, 1992.
- LJUNG, G.M.; BOX, G.E.P. On a Measure of a Lack of Fit in Time Series Models. **Biometrika**, v. 65, p. 297-303, 1978.
- MONTGOMERY, D. C. **Introdução ao Controle Estatístico da Qualidade**. Editora LTC, 2009.
- MONTGOMERY, D. C.; MASTRANGELO, C. M. Some statistical process control methods for autocorrelated data. **Journal of Quality Technology**, v. 23, p. 179-193, 1991.
- MORETTIN, P. A.; TOLOI, C. M. **Análise de Séries Temporais**. São Paulo: E. Blücher, 2004.
- PSARAKIS, S.; PAPALEONIDA, G. E. A. SPC procedures for monitoring autocorrelated processes. **Quality Technology & Quantitative Management**, v. 4, n. 4, p. 501–540, 2007.
- R CORE TEAM. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria, 2014.
- ROBERTS, S. W. Control Chart Tests Based on Geometric Moving Averages. **Technometrics**, v. 1. , 1959.
- SAID, S. E.; DICKEY, D. A. Testing for unit roots in autoregressive-moving average models of unknown order. **Biometrika** 71, 599–607, 1984.
- SANT’ANNA, A. M. O.; ten CATEN, C. S. Beta control charts for monitoring fraction data. **Expert Systems with Applications**, 2012, 39, 10236-10243.
- SANTOS, L. H. P. O uso de modelos autoregressivos e gráficos de controle para monitorar volatilidade de ativos financeiros. Trabalho de Graduação – Universidade de São Paulo: São Paulo, 2012.
- SCRUCCA, L. qcc: an R package for quality control charting and statistical process control. **R News**, p. 11-17, 2004.
- TSAY, R. S. **Analysis of financial time series**. A John Wiley & Sons, Inc., 2^a ed., 2010.
- YU, J.; LIU, J. LRProb control chart based on logistic regression for monitoring mean shifts of auto-correlated manufacturing processes. **International Journal of Production Research**, v. 49, n. 8, p. 2301–2326, 2011.

Recebido: 18/02/2015

Aprovado: 10/06/2015