

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE MATEMÁTICA
DEPARTAMENTO DE ESTATÍSTICA**

**CONTROLE ESTATÍSTICO DE PROCESSO PARA DADOS
AUTOCORRELACIONADOS**

Fernando de Jesus Moreira Junior

Porto Alegre, 2003

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE MATEMÁTICA
DEPARTAMENTO DE ESTATÍSTICA**

**CONTROLE ESTATÍSTICO DE PROCESSO PARA DADOS
AUTOCORRELACIONADOS**

Fernando de Jesus Moreira Junior

Orientador: Prof. Márcia Elisa Echeveste

**Trabalho de Conclusão para a obtenção do
título de Bacharel em Estatística do
Departamento de Estatística da
Universidade Federal do Rio Grande do Sul.**

Porto Alegre, 2003

*Feliz o homem que acha sabedoria,
e o homem que adquire conhecimento;
porque melhor é o lucro que ela dá do que o da prata,
e melhor a sua renda do que o ouro mais fino.*

Pv. 3:13-14

AGRADECIMENTOS

Gostaria de agradecer à minha orientadora, Márcia E. Echeveste, e aos professores do Departamento de Estatística da UFRGS, em especial aos professores Liane Werner, Carlos A. Crusius, e Flávio A. Ziegelmann.

Agradeço aos meus colegas e amigos da UFRGS que me acompanharam e estiveram presentes durante o tempo em que estive no curso.

Agradeço às equipes de trabalho pelo companheirismo e aprendizado que me foi dados nos lugares onde eu estagiei.

Agradeço aos meus pais, Fernando J. Moreira e Wasila Moreira, que sempre me incentivaram a estudar durante toda a minha vida.

Acima de tudo, agradeço a Deus que me sustentou e me deu forças para cursar e concluir esse curso.

ÍNDICE

1	COMENTÁRIOS INICIAIS	12
1.1	INTRODUÇÃO.....	12
1.2	TEMA E OBJETIVOS	14
1.3	JUSTIFICATIVA DO TEMA.....	14
1.4	MÉTODO DE TRABALHO	15
1.5	ESTRUTURA DE TRABALHO	15
1.6	LIMITAÇÕES	16
2	REVISÃO BIBLIOGRÁFICA.....	17
2.1	INTRODUÇÃO.....	17
2.2	DETECTANDO A AUTOCORRELAÇÃO	21
2.2.1	<i>Gráfico de correlograma</i>	21
2.2.2	<i>Estatísticas Q de Box-Pierce</i>	22
2.2.3	<i>Estatística de Ljung-Box</i>	23
2.2.4	<i>Teste da raiz unitária</i>	23
2.2.5	<i>Teste da Correlação Cruzada</i>	24
2.2.6	<i>Teste do Periodograma Acumulado</i>	25
2.3	MÉTODOS PARA RETIRAR A AUTOCORRELAÇÃO	26
2.3.1	<i>Ajustando modelos ARIMA</i>	27
2.3.1.1	Modelo Auto-regressivo de primeira ordem: AR(1).....	27
2.3.1.2	Modelo Auto-regressivo de segunda ordem: AR(2).....	28
2.3.1.3	Modelo Auto-regressivo de ordem p: AR(p)	29
2.3.1.4	Modelo de Média Móvel de primeira ordem: MA(1).....	30
2.3.1.5	Modelo de Média Móvel de segunda ordem: MA(2).....	30
2.3.1.6	Modelo de Média Móvel de ordem q: MA(q).....	31
2.3.1.7	Modelo Auto-regressivo de Média Móvel.....	32
2.3.1.8	Modelo Auto-regressivo Integrado e de Média Móvel.....	32
2.3.2	<i>Aplicações de gráficos de controle após a estimação de um modelo ARIMA</i>	38
2.3.2.1	Gráfico de causa especial ou gráfico de resíduos (special-cause chart, SCC).....	38
2.3.2.2	Gráficos CUSUM	40
2.3.2.3	Gráficos EWMA	46
2.3.2.4	Gráficos de Controle $\lambda_{LS,max}$ e $\bar{\lambda}_{LS}$	50
2.3.2.5	Gráfico para a média com Intervalo Amostral Variável (VSI)	50
2.3.2.6	Gráfico de Controle ARMA	51
2.3.2.7	Gráfico de Médias de Grupos.....	51
2.3.2.8	Método da Estimação Conjunta (JE)	53
2.3.3	<i>O uso de gráficos EWMA com linha central móvel</i>	55

2.3.3.1	Estimação e monitoramento de σ	57
2.3.3.2	A sensibilidade do gráfico de controle de resíduos.....	57
2.3.3.3	Alguns outros aspectos do EWMA	58
2.3.4	<i>O uso de EPC (Controle da Engenharia do Processo)</i>	59
2.3.4.1	Monitoramento do processo e regulagem do processo	59
2.3.4.2	Combinando CEP e EPC.....	63
2.3.4.3	Comparando CEP e EPC.....	65
2.3.5	<i>Uso de redes neurais (RN)</i>	66
2.4	DISCUSSÃO DO CAPÍTULO.....	68
3	ESTUDO DE CASO.....	70
3.1	DESCRIÇÃO DO EXPERIMENTO	70
3.2	ANÁLISE PRELIMINAR DOS DADOS.....	71
3.3	DETECÇÃO DA AUTOCORRELAÇÃO	74
3.4	AJUSTE POR UM MODELO ARIMA.....	77
3.5	USO DO GRÁFICO DE CONTROLE DE RESÍDUOS	81
3.6	DISCUSSÃO DOS RESULTADOS	82
4	CONCLUSÕES E RECOMENDAÇÕES PARA FUTUROS TRABALHOS.....	83
4.1	CONCLUSÕES.....	83
4.2	RECOMENDAÇÕES.....	84
5	REFERÊNCIAS BIBLIOGRÁFICAS	86

LISTA DE FIGURAS

FIGURA 1. AS TRÊS PERSPECTIVAS DE MÉTODOS DE GRÁFICOS DE CONTROLE, SEGUNDO PALM, RODRIGUES, SPIRING, WHEELER, 1997.....	18
FIGURA 2. EXEMPLO DE GRÁFICO DE CONTROLE DE SHEWHART	18
FIGURA 3. EXEMPLO DE UM PROCESSO COM OBSERVAÇÕES AUTOCORRELACIONADAS.....	19
FIGURA 4. REGRAS PARA A SELEÇÃO DE GRÁFICOS DE CONTROLE SUGERIDAS POR MONTGOMERY, 1997.....	26
FIGURA 5. PADRÕES DE CORRELOGRAMAS DA FAC E DA FACP, RESPECTIVAMENTE, EM UM PROCESSO AR(1).....	28
FIGURA 6. PADRÕES DE CORRELOGRAMAS DA FAC E DA FACP, RESPECTIVAMENTE, EM UM PROCESSO AR(2).....	29
FIGURA 7. PADRÕES DE CORRELOGRAMAS DA FAC E DA FACP, RESPECTIVAMENTE, EM UM PROCESSO MA(1).....	30
FIGURA 8. PADRÕES DE CORRELOGRAMAS DA FAC E DA FACP, RESPECTIVAMENTE, EM UM PROCESSO MA(2).....	31
FIGURA 9. PADRÕES DE CORRELOGRAMAS DA FAC E DA FACP, RESPECTIVAMENTE, EM UM PROCESSO ARMA(1,1).....	32
FIGURA 10. REPRESENTAÇÃO ESQUEMÁTICA DA METODOLOGIA BOX-JENKINS PARA MODELOS ARIMA.....	35
FIGURA 11. EXEMPLO DE GRÁFICO CUSUM PARA DETECTAR MUDANÇAS PEQUENAS NA MÉDIA DE UM PROCESSO.....	40
FIGURA 12. EXEMPLO DE GRÁFICO CUSUM TABULAR.....	41
FIGURA 13. ESTRUTURA DA MÁSCARA V.....	42
FIGURA 14. EXEMPLO DE GRÁFICO CUSUM MÁSCARA V.....	43
FIGURA 15. EXEMPLO DE RELACIONAMENTO ENTRE A SOMA DOS QUADRADOS DOS ERROS DE PREVISÃO UM PASSO À DIANTE (EIXO DAS ORDENADAS) E O VALOR DE LAMBDA (EIXO DAS ABSCISSAS).....	46
FIGURA 16. EXEMPLO DE GRÁFICO DE CONTROLE EWMA.....	47

FIGURA 17.	EXEMPLO DE GRÁFICO DE CONTROLE EWMA COM LINHA CENTRAL MÓVEL.	56
FIGURA 18.	EXEMPLO DE GRÁFICO DE CONTROLE EWMA COM LINHA CENTRAL MÓVEL COM VALOR DO PARÂMETRO LAMBDA NÃO OTIMIZADO.....	56
FIGURA 19.	CONTROLE DA ENGENHARIA DO PROCESSO E MONITORAMENTO ESTATÍSTICO DO PROCESSO	65
FIGURA 20.	ESQUEMA PROPOSTO PARA O TRATAMENTO DOS DADOS AUTOCORRELACIONADOS.	69
FIGURA 21.	SÉRIE TEMPORAL DA VARIÁVEL UR/HORA.	71
FIGURA 22.	HISTOGRAMA DA VARIÁVEL UR/HORA.	72
FIGURA 23.	GRÁFICO DE BOX PLOT PARA A VARIÁVEL UR/HORA.....	72
FIGURA 24.	GRÁFICO DA PLOTAGEM DE PROBABILIDADE NORMAL	73
FIGURA 25.	GRÁFICO DE CONTROLE PARA AS OBSERVAÇÕES ORIGINAIS.	74
FIGURA 26.	CORRELOGRAMA DA SÉRIE UR/HORA	75
FIGURA 27.	CORRELOGRAMA DA FAC DOS DADOS ORIGINAIS.	75
FIGURA 28.	CORRELOGRAMA DA FAC DA SÉRIE DIFERENCIADA UM VEZ.....	76
FIGURA 29.	CORRELOGRAMA DA FACP DA SÉRIE DIFERENCIADA UMA VEZ.....	76
FIGURA 30.	VALORES DA FAC E ESTATÍSTICAS BOX-LJUNG DO MODELO ESTIMADO.....	77
FIGURA 31.	VALORES DA FACP DO MODELO ESTIMADO	78
FIGURA 32.	CORRELOGRAMA DA FAC DOS RESÍDUOS DE MODELO ARIMA(0,1,1).	78
FIGURA 33.	CORRELOGRAMA DA FACP DOS RESÍDUOS DO MODELO ARIMA(0,1,1).....	79
FIGURA 34.	HISTOGRAMA DOS RESÍDUOS DO MODELO ARIMA(0,1,1) AJUSTADO.....	79
FIGURA 35.	GRÁFICO DE PROBABILIDADE NORMAL DOS RESÍDUOS DO MODELO ARIMA(0,1,1).	80
FIGURA 36.	GRÁFICO DE CONTROLE PARA O MODELO ARIMA(0,1,1)	81
FIGURA 37.	GRÁFICO DE CONTROLE PARA OS RESÍDUOS DO MODELO ARIMA(0,1,1).	81

LISTA DE TABELAS

TABELA 1.	COMPARATIVO DE COMPORTAMENTO DAS FAC E FACP DE UM PROCESSO ARIMA(P,D,Q), SEGUNDO MORETIN E TOLOI (1985).	34
TABELA 2.	TESTE DE NORMALIDADE (SAÍDA DO SOFTWARE SPSS).....	73
TABELA 3.	COEFICIENTE DO TERMO MA DO MODELO ESTIMADO.....	77
TABELA 4.	TESTE DE NORMALIDADE NOS RESÍDUOS DO MODELO ARIMA(0,1,1) AJUSTADO. 80	
TABELA 5.	COMPARATIVO ENTRE DIVERSOS MODELOS AJUSTADOS EM RELAÇÃO ÀS ESTATÍSTICAS AIC E BIC.	82

RESUMO

No uso do Controle Estatístico de Processo tradicional (CEP), pressupõe-se que as observações sejam independentes e normalmente distribuídas. Entretanto, em algumas situações práticas a suposição de independência é violada, gerando autocorrelação entre as observações. Essa autocorrelação prejudica o desempenho dos gráficos de controle, causando um aumento na frequência de alarmes falsos. Essa monografia apresenta métodos que são aplicados em CEP quando os dados estão autocorrelacionados. Para ilustrar o problema, também é apresentado um estudo de caso em que os dados estão autocorrelacionados, onde é aplicado um método para retirar a autocorrelação.

ABSTRACT

In traditional Statistical Process Control (SPC) it is assumed that the observations are independent and normally distributed. However, sometimes this assumption of independence is violated causing autocorrelation in observations. This autocorrelation deteriorate the performance of control charts, causing an increase in the frequency of false alarms. This monograph presents methods for applying in SPC when the data are autocorrelated. To illustrate this problem it is presented a study of case that the data are autocorrelated and a method is applied to remove the autocorrelation.

1 COMENTÁRIOS INICIAIS

1.1 Introdução

Qualidade constitui-se hoje num fator fundamental na decisão do consumidor para um número explosivamente crescente de produtos e serviços. Qualidade tornou-se simplesmente o fator mais significativo, conduzido nas empresas nos mercados nacional e internacional ao êxito organizacional e ao crescimento. O retorno sobre o investimento, obtido por meios rigorosos e eficazes programas de qualidade, está gerando excelente rentabilidade nas empresas quando acompanhado de estratégias eficientes para a qualidade (Feigenbaum, 1994).

A melhora da qualidade transfere o desperdício para a fabricação de um bom produto e uma melhor prestação de serviços. Isso resulta em uma reação em cadeia – custos mais baixos devido a menos re-trabalhos e erros, melhor posição competitiva com melhor qualidade e preços menores, manutenção dos negócios, ampliação do mercado de trabalho com a geração de empregos (Deming, 1990).

Segundo Feigenbaum (1994), ao logo dos anos, as técnicas e metodologias estatísticas tornaram-se cada vez mais utilizadas e, atualmente, desempenha papel fundamental nos programas modernos de Controle da Qualidade Total. Uma das técnicas estatísticas mais utilizadas neste contexto é o controle estatístico de processo, o CEP ou SPC (*statistical process control*).

O CEP é uma técnica estatística que permite um monitoramento do desempenho de processos através de um sistema de inspeção por amostragem ao longo do tempo. No CEP, a ferramenta estatística utilizada é o gráfico de controle (ou carta de controle), desenvolvido pelo Dr. Walter A. Shewhart na década de 20, e utilizado mundialmente nos dias atuais.

Para Deming (1990), a maneira mais fácil de examinar os dados é colocar numa carta de controle os pontos na ordem em que são produzidos, para ver se pode ser tirado qualquer proveito da distribuição dos dados ao logo da carta; ou seja, questionar o estado de controle estatístico que produziu esses dados.

De acordo com a definição de Feigenbaum (1994), gráfico de controle é um método gráfico para avaliar se um processo se encontra ou não sob “controle estatístico”, isto é, ele compara a variação real na produção de itens com os limites de controle estipulados para esses mesmos itens.

Para a utilização das cartas convencionais de controle (cartas de Shewhart), os dados devem atender a suposição fundamental de serem *iid* (independentes e identicamente distribuídos) com distribuição normal. Entretanto, algumas vezes os dados que representam o processo estão autocorrelacionados, e assim, a utilização dos gráficos de controle convencionais poderia possivelmente resultar em “falsos alarmes”, isto é, conclusões errôneas quanto a estabilidade do processo, como, por exemplo, poderia se confundir causas comuns (aquelas que atuam de forma aleatória) do processo com causas especiais (aquelas que não seguem um padrão aleatório, interferem no processo com alterações bruscas).

Nestes casos, não há lógica em utilizar os gráficos de controle de Shewhart, pelo fato de conduzir à possíveis conclusões errôneas e possibilitar uma falta de segurança em verificar se o processo se encontra ou não sob controle estatístico. Entretanto, segundo Schissatti (1998), o erro mais comum é a adoção de um modelo de carta de controle sem que se faça a análise de correlação dos dados que representam o processo. Kume (1993) alerta para o fato que as ferramentas estatísticas são apenas ferramentas, ou seja, elas podem não funcionar, caso sejam aplicadas inadequadamente.

Para corrigir o problema da autocorrelação, o procedimento a seguir seria primeiro tratar os dados e depois controlá-los estatisticamente. Desta forma, algumas técnicas estatísticas são necessárias para abordar os dados, dentre as quais se destacam, segundo Montgomery (1997): estimação de modelos ARIMA (Auto-regressivos integrados e de média móvel), uso dos gráficos de controle EWMA (média móvel ponderada exponencialmente) com linha central móvel, e uso de Controle da Engenharia do Processo, o EPC (*engineering process control*).

Os processos que possuem dados autocorrelacionados caracterizam-se pelo fato de que a observação seguinte está correlacionada com a observação anterior. Isso pode ocorrer, por exemplo, em indústrias de produtos químicos, em processos realizados por determinados equipamentos, em linhas de montagens, em procedimentos de treinamento de funcionários, em monitoramentos de desempenho de equipes.

1.2 Tema e Objetivos

Esse trabalho tem como tema principal o estudo do uso do controle estatístico de processo (CEP) para dados autocorrelacionados. Esse estudo abrange um levantamento de uma quantidade razoável de documentos teóricos sobre o assunto mencionado, assim como um estudo de caso.

Esse trabalho pretende sugerir procedimentos adequados para a implantação de gráficos de controle em situações onde os dados estão correlacionados. O tema desenvolvido nesse trabalho tem a sua relevância pelo fato desse assunto ser pouco desenvolvido na literatura, entretanto muito útil e necessário na prática da implantação do controle estatístico de processo.

O objetivo principal desse trabalho é realizar uma revisão bibliográfica que fundamente o tratamento de observações autocorrelacionadas através da aplicação de modelos de série temporais a cartas de controle estatístico e de outras técnicas encontradas durante a pesquisa.

Como objetivo secundário desse trabalho, tem-se a aplicação de técnicas abordadas nessa monografia para a resolução do problema de gráficos de controle estatístico para dados autocorrelacionados em um estudo de caso de desempenho de funcionários de uma empresa de companhia elétrica.

1.3 Justificativa do tema

A maioria dos livros que tratam sobre o controle estatístico de processo, pressupõe que as observações são independentes (dados não correlacionados) e identicamente distribuídas com distribuição normal, caracterizando-se por uma média estável. Entretanto, muitas vezes,

na prática, essa suposição não é confirmada, e conseqüentemente, os gráficos de controle convencionais não podem ser utilizados por conduzirem conclusões errôneas. Alguns casos típicos, como desempenho de funcionários, desgaste de ferramentas, resistência de algum material, etc, caracterizam-se por possuírem observações autocorrelacionadas. Para esses casos, é necessário um tratamento especial dos dados com métodos estatísticos adequados antes de proceder com a metodologia de implantação de gráficos de controle. Por essa razão, esse trabalho justifica-se por sua constituição acadêmica e ampla possibilidade de aplicação na prática.

1.4 Método de trabalho

Essa seção descreve como será a metodologia desenvolvida neste trabalho.

Primeiramente será realizada uma revisão bibliográfica ampla sobre o assunto, com o objetivo de verificar e analisar as técnicas descritas por diversos autores, divulgadas nos mais variados periódicos do ramo.

A seguir, será realizado um estudo de caso onde serão aplicadas as técnicas adequadas vistas anteriormente na revisão bibliográfica, a fim de verificar a eficiência das ferramentas sugeridas pelos autores.

Para finalizar, serão feitas as considerações finais para concluir o trabalho e também serão disponibilizados e sugeridos temas para possíveis pesquisas ou estudos futuros.

1.5 Estrutura de Trabalho

O trabalho está dividido em quatro capítulos principais.

O capítulo 1 trata dos comentários iniciais, subdivididos em seis seções que compreendem os tópicos: introdução, tema e objetivos, justificativa, método, e estrutura do trabalho, e limitações.

O capítulo 2 consiste em uma sólida revisão bibliográfica, contendo um resumo das principais técnicas estatísticas utilizadas para a resolução do problema em questão e dos

resultados obtidos de estudos desenvolvidos por diversos autores. Essa parte é base teórica do assunto que está sendo estudado nessa monografia. O item 2.1 apresenta uma introdução ao assunto do uso de gráficos de controle com observações autocorrelacionadas. O item 2.2 apresenta técnicas para detectar a autocorrelação. Na seção 2.3 são apresentadas as técnicas para retirar a autocorrelação, e subdivide-se em estimação de modelos ARIMA e aplicação dos gráficos de controle para resíduos, utilização do gráfico EWMA com linha central móvel, uso de EPC, e uso de redes neurais.

No capítulo 3, é composto por um estudo de caso que consiste em analisar o desempenho de funcionários de uma companhia elétrica através do uso de gráfico de controles. Nesta parte são aplicadas as técnicas apropriadas revisadas no capítulo 2 deste trabalho.

O capítulo 4 engloba as considerações finais, conclusões, indicações, sugestões e recomendações para trabalhos futuros.

1.6 Limitações

Esse trabalho limita-se a estudar o uso de gráficos de controle para os casos em que os dados estão autocorrelacionados, através de técnicas tais como o ajuste de gráficos de séries temporais de modelos ARIMA, os gráficos de controle EWMA com linha central móvel, e o uso de EPC.

Não faz parte do escopo deste trabalho aprofundar os conhecimentos em técnicas mais recentes e complexas capazes de solucionar o problema em questão, como por exemplo o uso de redes neurais e alguns tipos de gráficos de controle.

2 REVISÃO BIBLIOGRÁFICA

2.1 *Introdução*

Os gráficos de controle ou cartas de controle são conhecidos pela ampla utilização na monitoração de diversos processos industriais. Entretanto, ultimamente, os gráficos de controle também têm sido utilizados em outros tipos de controles, tais como monitorar o desempenho de atividades de equipes, monitorar vendas, satisfação de clientes ou procedimentos rotineiros em departamentos dentro de uma empresa, etc. Segundo estudos de Palm, Rodrigues, Spiring, Wheeler (1997), a diversidade de aplicação do uso dos gráficos de controle pode ser explicitada em 3 grupos com perspectivas distintas nos métodos de aplicação de gráficos de controle. O grupo 1 está interessado no geral, em aplicações eficazes dos métodos dos gráficos de controle padrão em diferentes campos, indústrias, e departamentos dentro de uma empresa. Esse grupo é motivado pelo fluxo competitivo e mudanças nos padrões industriais, e é orientado pelos programas de melhoria de qualidade em larga escala, e pelo “pensamento estatístico” de Deming. O grupo 2 tem o seu interesse voltado para o desenvolvimento e aplicações de gráficos de controle avançados para processos industriais onde dados “em linha” são superabundantes e onde uma resposta rápida para condições fora de controle é crítica. Esse grupo é composto de estatísticos que são bem treinados em CEP e trabalham conjuntamente com engenheiros que possuem conhecimento do processo. Há um uso ativo de tecnologia computacional, e é particularmente bem representado em processos industriais químicos e de semicondutores. O grupo 3 está interessado academicamente nos métodos dos gráficos de controle. Esse grupo consiste de pesquisadores nos cursos acadêmicos de estatística, engenharia industrial, e departamentos de faculdades de negócios que desenvolvem variações dos gráficos de controle que freqüentemente otimizam ou

melhoram a eficiência utilizando métodos estatísticos. Esse grupo freqüentemente contribui para publicações dos periódicos. A figura 1 resume essas três perspectivas.

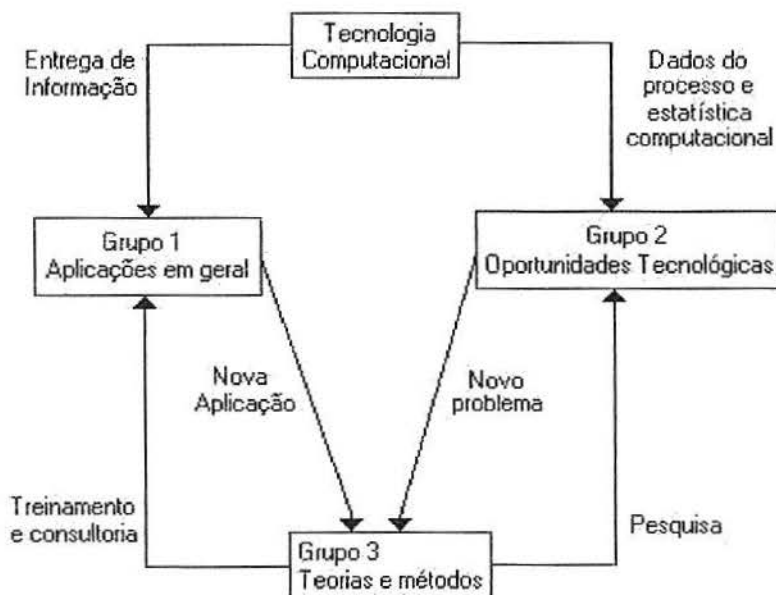


Figura 1. As três perspectivas de métodos de gráficos de controle, segundo Palm, Rodrigues, Spiring, Wheeler, 1997

Os gráficos de controle servem para monitorar um processo estatisticamente, isto é, verificar se um processo está sob controle o qual é caracterizado por oscilar aleatoriamente em torno de uma média e dentro de limites toleráveis estabelecidos ou calculados, chamados de limites de controle. A figura 2 mostra um típico gráfico de controle.

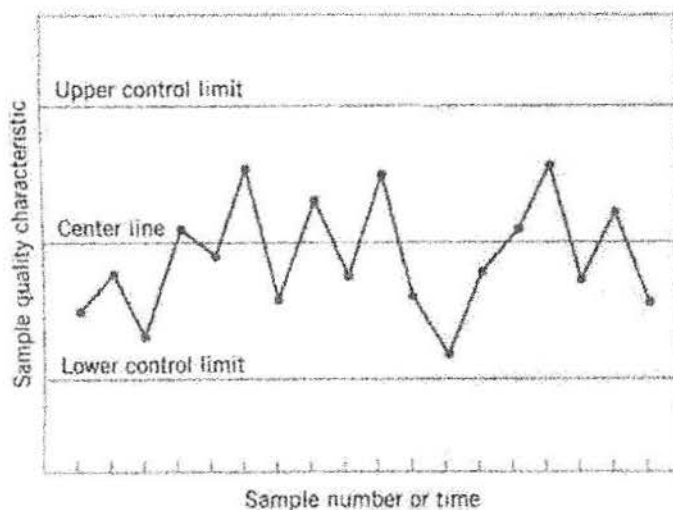


Figura 2. Exemplo de gráfico de controle de Shewhart

A utilização dos usuais gráficos de controle de Shewhart pressupõe que os dados sejam independentes e identicamente distribuídos com distribuição normal. Uma condição de “fora de controle” seria uma mudança na média ou no desvio padrão (ou em ambos) para um valor diferente. Entretanto, quando o processo está sob controle, a variável x_t em um tempo t , é representada pelo modelo

$$x_t = \mu + \varepsilon_t \quad (2.1)$$

onde ε_t é independente e normalmente distribuído com média zero e desvio padrão σ . Isso é freqüentemente chamado de “Modelo Shewhart do processo”.

Quando essas suposições estão satisfeitas, os gráficos de controle convencionais podem ser aplicados, e conclusões sobre o processo podem ser obtidas. Em algumas situações, quando a suposição de normalidade é violada leve ou moderadamente, esses gráficos de controle funcionam razoavelmente bem. Entretanto, quando as observações estão autocorrelacionadas (figura 3), isto é, não são independentes, o uso de gráficos de controle não possui um bom desempenho, podendo surgir falsos alarmes indicando uma situação “fora de controle” quando na verdade eles são devidos a causas naturais de um processo autocorrelacionado.

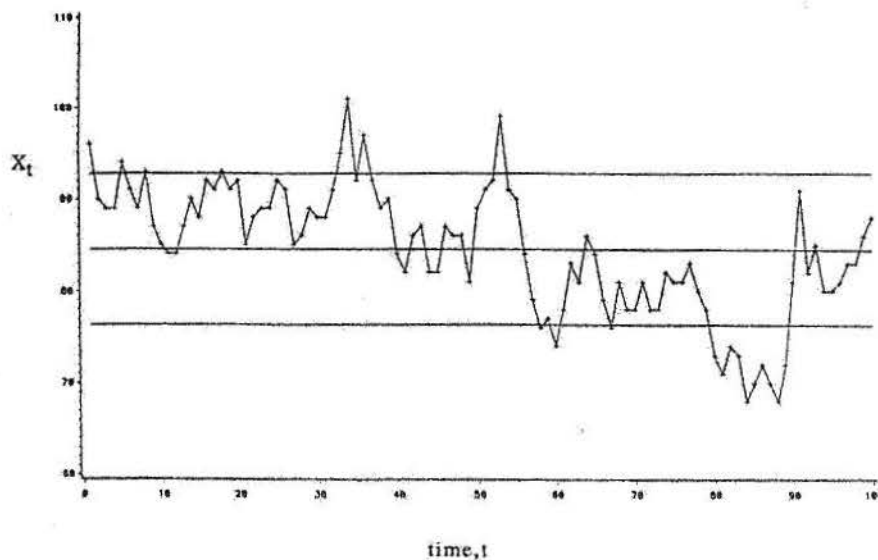


Figura 3. Exemplo de um processo com observações autocorrelacionadas

Montgomery (1997) faz uma demonstração analítica deste fenômeno (dependência ou autocorrelação). Consideremos um sistema simples que consiste em um tanque de volume V , com entrada e saída de um fluxo de material com uma taxa de fluxo f . Seja w_t a concentração de um certo material no fluxo de entrada no tempo t , e x_t a correspondente concentração do fluxo de saída no tempo t . Supondo homogeneidade dentro do tanque, o relacionamento entre x_t e w_t é

$$x_t = w_t - T \frac{dx_t}{dt} \quad (2.2)$$

onde $T = V/f$ é chamado de tempo constante do sistema. Se o fluxo de entrada experimenta uma troca de passo de w_0 num tempo $t=0$ (por exemplo), então a concentração de saída no tempo t é

$$x_t = w_0 (1 - e^{-t/T}) \quad (2.3)$$

Agora, na prática, não se observa x_t continuamente, mas em intervalos de tempo Δt igualmente espaçados. Nesse caso,

$$x_t = x_{t-1} + (w_e - x_{t-1})(1 - e^{-\Delta t/T}) = \alpha w_t + (1 - \alpha)x_{t-1} \quad (2.4)$$

onde $\alpha = 1 - e^{-\Delta t/T}$.

As propriedades da concentração do fluxo de saída x_t dependem da concentração do fluxo de entrada w_0 e do intervalo de tempo de amostragem Δt . Se é assumido que os valores w_t são variáveis aleatórias não correlacionadas, então a correlação entre sucessivos valores de x_t é dada por

$$\rho = 1 - \alpha = e^{-\Delta t/T} \quad (2.5)$$

Note que se Δt for muito maior que T , $\rho \cong 0$. Isto é, se o intervalo entre os Δt no fluxo de saída for muito maior que o tempo constante T , as observações na concentração de saída serão não autocorrelacionadas. Entretanto, se $\Delta t \leq T$, isso não acontece.

2.2 Detectando a autocorrelação

Segundo Gujarati (2000), o termo autocorrelação pode ser definido como “correlação entre membros de séries de observações ordenadas no tempo (como em uma série temporal) ou no espaço (dados de corte)”.

A autocorrelação caracteriza uma série temporal não-estacionária, isto é, não possui média ou variância constantes ao longo do tempo e a covariância entre dois períodos de tempo não depende apenas da distância entre os dois períodos.

Podemos detectar a presença de autocorrelação através dos seguintes meios: (i) o gráfico de correlograma, com base no cálculo da função de autocorrelação (FAC ou ACF – *autocorrelation function*) e da função de autocorrelação parcial (FACP ou PACF – *partial autocorrelation function*), (ii) as estatísticas Q de Box-Pierce, (iii) e de Ljung-Box, (iv) o teste da raiz unitária, (v) o teste da correlação cruzada, ou (vi) o teste do periodograma acumulado. Essas técnicas serão detalhadas na sequência.

2.2.1 Gráfico de correlograma

A FAC, indicada por ρ_k , é definida por:

$$\rho_k = \frac{\gamma_k}{\gamma_0}, \quad (2.6)$$

onde:

γ_k é a covariância na defasagem (*lag*) k , e

γ_0 é a variância.

Por ser um coeficiente de correlação, ρ_k varia entre -1 e 1 . O estimador de ρ_k é dado por:

$$r_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = \frac{\frac{\sum_{t=1}^{n-k} (x_t - \bar{x})(x_{t-k} - \bar{x})}{n}}{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}} = \frac{\sum_{t=1}^{n-k} (x_t - \bar{x})(x_{t-k} - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad (2.7)$$

onde:

x_i é o i -ésimo valor da variável aleatória X , e

\bar{x} é a média estimada da variável X .

A representação gráfica de r_k contra k é conhecida como correlograma amostral. Observado o correlograma, se os valores de autocorrelação não diminuem rapidamente a medida que o valor de k aumenta, há, então, um indicio de não estacionariedade no processo. Por outro lado, em um processo estacionário, os valores de sua autocorrelação a qualquer defasagem maior que zero é zero. Para testarmos a hipótese de que o verdadeiro valor de ρ_k é zero, calculamos um intervalo com $(1 - \alpha)\%$ de confiança para ρ_k , dado por:

$$0 \pm z_{(1-\alpha)/2} \sqrt{1/n}, \quad (2.8)$$

onde:

n é o tamanho da amostra, e

$z_{(1-\alpha)/2}$ é o valor da distribuição Normal (0; 1) para uma dada significância α .

Assim, se um ρ_k estimado se situar dentro do intervalo, não rejeitamos a hipótese de que o verdadeiro ρ_k seja significativamente igual a zero. Caso contrário, rejeitamos a hipótese. Usualmente, utiliza-se um intervalo de confiança de 95%.

A FACP mede a correlação entre as observações que sejam k períodos afastados, depois de remover o efeito dos X s intermediários, ou seja, é a correlação entre X_t e X_{t-k} depois de remover a influência de $X_{t-1}, X_{t-2}, \dots, X_{t-k+1}$. A FACP também é representada graficamente através de um correlograma, e indicada por ϕ_{kk} .

2.2.2 Estatísticas Q de Box-Pierce

Box e Pierce desenvolveram a estatística Q , para testar a hipótese conjunta de que todos os coeficientes de autocorrelação ρ_k são simultaneamente iguais a zero, e definida como:

$$Q = n \sum_{k=1}^m r_k^2 \quad (2.9)$$

onde:

n é o tamanho da amostra, e

m é a duração da defasagem, isto é, a quantidade de defasagens.

A estatística Q possui, aproximadamente, distribuição qui-quadrado com m graus de liberdade. Se o valor de Q calculado exceder o valor crítico de Q tabelado, ao nível de significância escolhido, rejeitamos a hipótese nula de que todos os ρ_k são iguais a zero, isto é, pelo menos um deles é significativamente diferente de zero, o que indica autocorrelação.

2.2.3 Estatística de Ljung-Box

Outra estatística variante da estatística Q de Box-Pierce é a estatística de Ljung-Box (LB), que também possui, distribuição aproximada qui-quadrado com m graus de liberdade definida como:

$$LB = n(n+2) \sum_{k=1}^m \left(\frac{r_k^2}{n-k} \right) \quad (2.10)$$

A estatística LB tem mostrado melhores propriedades (no sentido estatístico) para amostras pequenas do que a estatística Q (Gujarati, 2000).

2.2.4 Teste da raiz unitária

Outro teste para detectar a estacionariedade é o teste da raiz unitária. Considera-se o modelo abaixo:

$$x_t = \phi x_{t-1} + \varepsilon_t, \quad (2.11)$$

que é uma regressão de primeira ordem, isto é, um AR(1). Se o coeficiente ϕ for igual a 1, temos o problema da raiz unitária, ou seja, uma situação de não-estacionariedade. Essa equação pode ser alternativamente expressa como:

$$\Delta x_t = (\phi - 1)x_{t-1} + \varepsilon_t = \delta x_{t-1} + \varepsilon_t, \quad (2.12)$$

onde

$$\Delta \text{ é chamado de operador de primeira diferença, pois } \Delta x_t = (x_t - x_{t-1}) = \varepsilon_t.$$

Para verificar a existência ou não de estacionariedade, testa-se a hipótese de que $\delta = 0$ ou, equivalentemente, se $\phi = 1$. O valor t obtido é conhecido como estatística τ , cujos valores foram tabulados por Dickey e Fuller com base em simulações de Monte Carlo. Porém, essas tabelas não são totalmente adequadas, e elas vêm sendo ampliadas por MacKinnon. Para calcular a estatística τ de Dickey Fuller (DF), dividimos o coeficiente ρ estimado pelo seu erro padrão e comparamos com os valores tabelados de DF. Se o valor absoluto calculado da estatística τ exceder os valores críticos absolutos de DF ou MacKinnon-DF, então rejeita-se a hipótese de que $\delta = 0$ ou $\phi = 1$, isto é, rejeita-se a hipótese de que a série é não-estacionária.

Para expandir o teste para modelos AR(p), onde $p > 1$, foi criado o teste aumentado de Dickey-Fuller (ADF), que possui a mesma distribuição da estatística DF, podendo ser utilizados os mesmos valores críticos. Os valores das estatísticas DF e ADF podem ser encontrados em Dickey e Fuller (1979), em Fuller (1976) e em MacKinnon (1991), citados em Gujarati (2000).

2.2.5 Teste da Correlação Cruzada

Esse teste consiste na investigação da Função de Correlação Cruzada (FCC) dada por:

$$s_k = \frac{\sum \varepsilon_t (x_{t-k} - \bar{x})}{\left[\sum \varepsilon_t^2 \sum (x_t - \bar{x})^2 \right]^{1/2}} \quad (2.13)$$

O estimador de s_k é dado por

$$\hat{s}_k = \frac{\sum_{t=k+1}^n e_t (x_{t-k} - \bar{x})}{\left[\sum_{t=1}^n e_t^2 \sum_{t=1}^n (x_t - \bar{x})^2 \right]^{1/2}} \quad (2.14)$$

onde e_t são os resíduos dados por

$$e_t = x_t - \hat{x}_t \quad (2.15)$$

Se para um dado k_0 , s_{k_0} tiver valor alto, há evidência de que o modelo seja inadequado. Com um AR(p) de ordem baixa, pode-se incluir novos termos AR analisando a FCC. Quando \hat{s}_k não se apresentar mais significativo, a FAC pode indicar termos de MA (média móvel) a serem incluídos. Se termos de MA são incluídos num estágio anterior de identificação, a interpretação dos valores altos de $|s_k|$ não é tão simples.

2.2.6 Teste do Periodograma Acumulado

Um dos primeiros estimadores a serem propostos da Análise Espectral de séries temporais, com a finalidade de detectar periodicidade nos dados é dado por:

$$I_e(f_i) = \frac{2}{n} \left[\left(\sum_{t=1}^n e_t \cos \frac{2\pi i}{n} t \right)^2 + \left(\sum_{t=1}^n e_t \sin \frac{2\pi i}{n} t \right)^2 \right], \quad (2.16)$$

chamado periodograma,

onde $0 < f_i < \frac{1}{2}$

Um pico na frequência f_i indica uma periodicidade no período $1/f_i$. Se e_t é ruído branco, então seu espectro é constante e igual a $2\sigma_e^2$ no intervalo $[0, 1/2]$, isto é,

$$P_e(f) = 2\sigma_e^2, \quad 0 \leq f \leq \frac{1}{2}. \quad (2.17)$$

Conseqüentemente

$$P_e(f) = \int_0^f P_e(g) dg = \begin{cases} 0, & f < 0 \\ 2\sigma_e^2 f, & 0 \leq f \leq \frac{1}{2}, \\ \sigma_e^2, & f \geq \frac{1}{2} \end{cases}, \quad (2.18)$$

onde $P_e(f)$ é o “espectro acumulado” ou função de distribuição espectral. Uma estimativa de

$P_e(f_j)/\sigma_e^2$ é dada por

$$C(f_j) = \frac{\sum_{i=1}^j I_e(f_i)}{n\hat{\sigma}_e^2} \quad (2.19)$$

que é o periodograma acumulado (normalizado). Em um processo de ruído branco, o gráfico de $C(f_j) \times f_j$ estaria em torno da reta que passa pelos pontos (0;0) e (0,5;1). Desvios sistemáticos dessa reta sugerem que o modelo não é adequado. O teste de Kolmogorov-Smirnov pode ser usado para verificar a normalidade desses dados.

2.3 Métodos para retirar a autocorrelação

Para tratar do problema de autocorrelação dos dados, Montgomery (1997) propõe três alternativas: ajuste de um modelo ARIMA (auto-regressivo integrado e de média móvel) e aplicação de um gráfico de controle padrão para resíduos, os gráficos de controle EWMA (média móvel ponderada exponencialmente) com linha central móvel, e uso de EPC (Controle da Engenharia do Processo). Além dessas alternativas, será revisado sucintamente também o uso de redes neurais, que tem sido muito utilizado ultimamente. A figura 4 apresenta um roteiro para a seleção dos gráficos de controle.

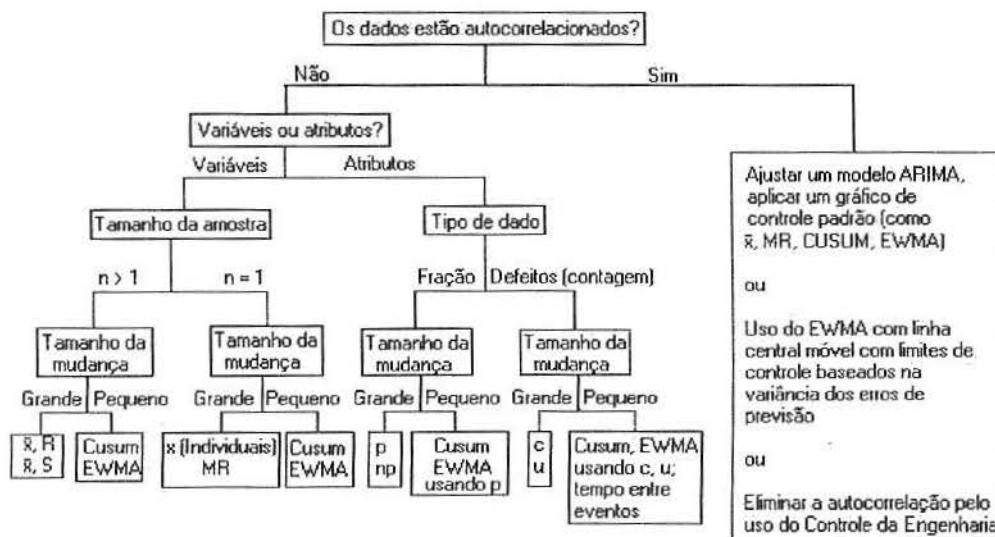


Figura 4. Regras para a seleção de gráficos de controle sugeridas por Montgomery, 1997

2.3.1 Ajustando modelos ARIMA

Um dos tratamentos sugeridos e um dos mais utilizados nos processos com dados autocorrelacionados e que tem sido útil em muitas aplicações é o ajuste de um modelo apropriado de série temporal ARIMA, que remova a autocorrelação dos dados. Uma vez removida a autocorrelação a partir desse modelo estimado, pode-se aplicar os gráficos de controle para os resíduos, tais como gráfico para a média, para os valores individuais, EWMA ou CUSUM (soma acumulada).

2.3.1.1 Modelo Auto-regressivo de primeira ordem: AR(1)

Suponha-se que a variável característica x_t possa ser ajustada pelo modelo

$$x_t = \xi + \phi x_{t-1} + \varepsilon_t \quad (2.20)$$

onde ξ e ϕ ($-1 < \phi < 1$) são constantes desconhecidas, e ε_t é independente e normalmente distribuído com média zero e desvio padrão σ . Esse modelo é conhecido como modelo auto-regressivo de primeira ordem, AR(1), onde as observações x_t tem média $\xi/(1-\phi)$, desvio padrão $\sigma/(1-\phi^2)^{1/2}$, e as observações que estão k períodos separados têm coeficiente de correlação ϕ^k . Isto é, a função autocorrelação (correlograma) deverá cair exponencialmente (alternando ou não o sinal). Por outro lado, a função de autocorrelação parcial deverá apresentar 1 pico significativo, pois, nesse caso, $p=1$. A figura 5 mostra aspectos de correlogramas da FAC e da FACP, respectivamente, que sugerem que o processo seja um AR(1).

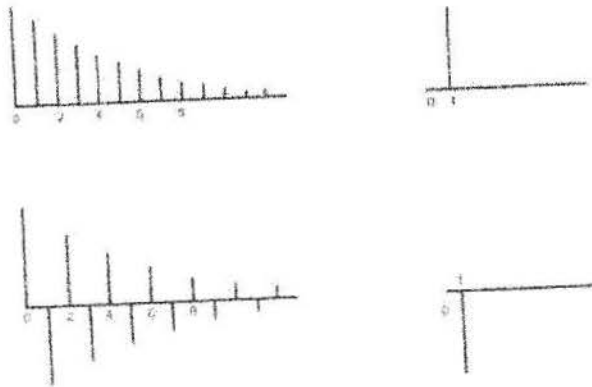


Figura 5. Padrões de correlogramas da FAC e da FACP, respectivamente, em um processo $AR(1)$

Suponha-se que $\hat{\phi}$ seja um estimador de ϕ , obtido da análise dos dados amostrais do processo, e que \hat{x}_t seja o valor estimado de x_t . Então, os resíduos e_t são independentes e aproximadamente normal com média zero e variância constante. Assim, os gráficos de controle convencionais poderão ser aplicados à seqüência dos resíduos. Pontos fora de controle ou modelos não usuais nos gráficos poderão indicar que o parâmetro ϕ foi alterado, sugerindo que a variável original x_t estava fora de controle. Os parâmetros do modelo auto-regressivo poderão ser estimados pelo método dos mínimos quadrados, onde a escolha dos valores de ξ e ϕ minimiza a soma do quadrado dos erros ε_t .

2.3.1.2 Modelo Auto-regressivo de segunda ordem: $AR(2)$

Uma extensão óbvia da equação (2.19) pode ser dada por

$$x_t = \xi + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \varepsilon_t \quad (2.21)$$

que é um modelo auto-regressivo de segunda ordem, um $AR(2)$. Nesse modelo, as observações x_t tem média $\xi/(1 - \phi_1 - \phi_2)$, desvio padrão $\phi_1^2 \gamma_0 + \phi_2^2 \gamma_0 + 2\phi_1 \phi_2 \gamma_1 + \sigma^2$, e as observações que estão k períodos separados têm coeficiente de correlação $\phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}$. Assim, o correlograma da FAC deverá cair exponencialmente (alternando ou não o sinal) ou com um padrão de onda senoidal amortecida, ou ambos; e a FACP deverá apresentar 2 picos significativos. A figura 6 mostra aspectos de correlogramas da FAC e da FACP,

respectivamente, que sugerem que o processo seja um AR(2).

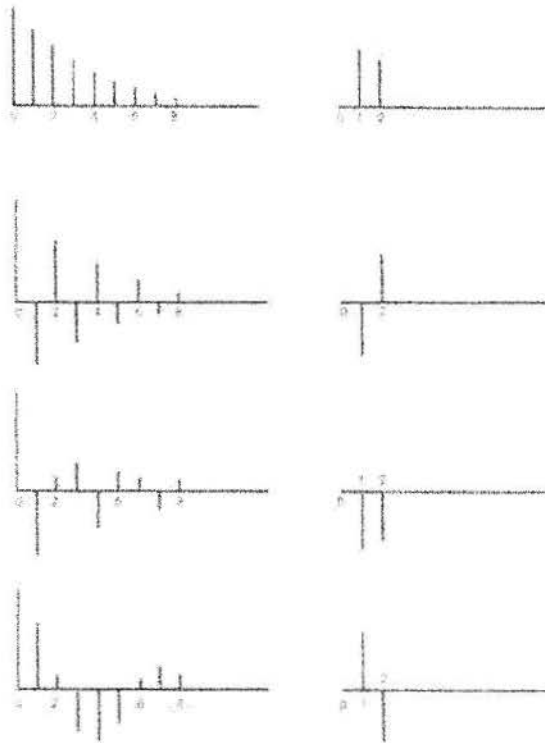


Figura 6. Padrões de correlogramas da FAC e da FACP, respectivamente, em um processo AR(2)

2.3.1.3 Modelo Auto-regressivo de ordem p : AR(p)

Generalizando o modelo da equação (2.19) para p ordens, obtemos o modelo

$$x_t = \xi + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + \varepsilon_t \quad (2.22)$$

o qual é um modelo auto-regressivo de ordem p , um AR(p). Como pode-se perceber, um modelo auto-regressivo depende de p observações anteriores, ou seja, a observação x_t é gerada pela média ponderada das p primeiras observações anteriores da variável. Se o processo é estacionário, a média é $\xi / (1 - \phi_1 - \phi_2 - \dots - \phi_p)$, onde $\phi_1 + \phi_2 + \dots + \phi_p < 1$. O correlograma da FAC deverá cair exponencialmente ou com um padrão de onda senoidal amortecida, ou ambos; e o correlograma da FACP deverá apresentar picos significativos através das defasagens p , que indicará a ordem do modelo AR.

2.3.1.4 Modelo de Média Móvel de primeira ordem: MA(1)

Suponha-se que a modelo que observação x_t dependa não dos seu valores anteriores, mas da componente aleatória ε_t , como na equação abaixo.

$$x_t = \mu + \varepsilon_t - \theta\varepsilon_{t-1}, \quad (2.23)$$

onde μ e θ são constantes desconhecidas.

Este modelo é chamado de modelo de média móvel de primeira ordem, um MA(1). Nesse modelo, a correlação entre x_t e x_{t-1} é $\rho_1 = -\theta/(1+\theta^2)$, e zero para todas as outras defasagens. O correlograma da FAC apresentará um pico significativo, pois $q = 1$ entretanto, o correlograma da FACP declinará exponencialmente (alternando ou não o sinal). A figura 7 mostra aspectos de correlogramas da FAC e da FACP, respectivamente, de um processo MA(1).

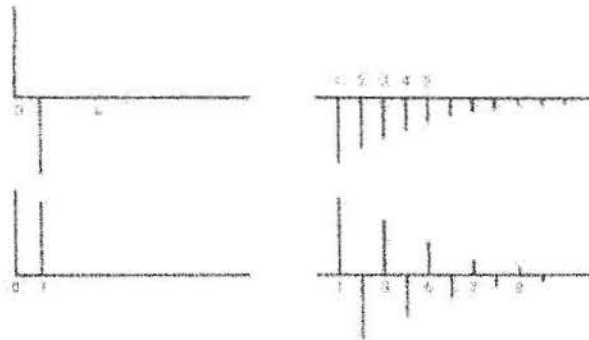


Figura 7. Padrões de correlogramas da FAC e da FACP, respectivamente, em um processo MA(1)

2.3.1.5 Modelo de Média Móvel de segunda ordem: MA(2)

Uma extensão do modelo da equação (2.23) pode ser dado por

$$x_t = \mu + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} \quad (2.24)$$

que é um modelo de média móvel de segunda ordem, um MA(2). Nesse modelo, os coeficientes de autocorrelação são dados por $\rho_1 = -\theta_1(1-\theta_2)/(1+\theta_1^2+\theta_2^2)$ e por

$\rho_2 = -\theta_2 / (1 + \theta_1^2 + \theta_2^2)$. O correlograma da FAC apresentará 2 picos significativos, pois $q = 2$ entretanto, o correlograma da FACP declinará exponencialmente (alternando ou não o sinal) ou apresentará o formato de uma onda senoidal amortecida. A figura 8 mostra aspectos de correlogramas da FAC e da FACP, respectivamente, de um processo MA(2).

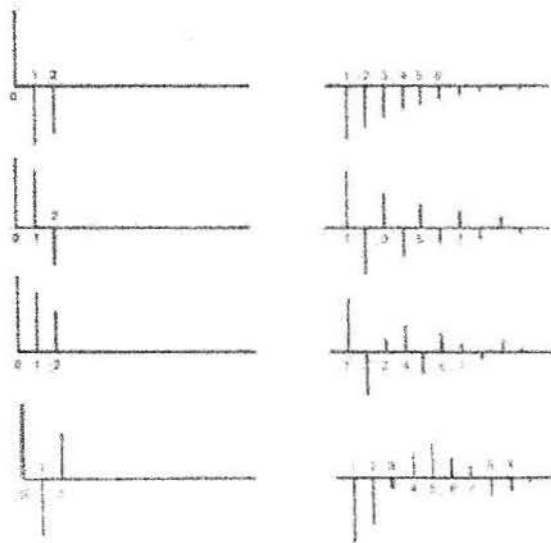


Figura 8. Padrões de correlogramas da FAC e da FACP, respectivamente, em um processo MA(2)

2.3.1.6 Modelo de Média Móvel de ordem q : MA(q)

Generalizando o modelo da equação (2.23) para q ordens, obtemos o modelo

$$x_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (2.25)$$

que é um modelo de média móvel de ordem p . Pode-se perceber que cada observação é gerada por uma média ponderada do valor presente e de q valores passados de um processo de ruído branco. O modelo de média móvel é sempre estacionário e possui média μ . Assim, o termo do erro ε_t , que possui média μ e variância σ^2 e tem valores que são não autocorrelacionados, é chamado de ruído branco. Comparando com o modelo AR, os correlogramas se invertem: o correlograma da FAC deverá apresentar picos significativos através das defasagens q , que indicará a ordem do modelo MA, e o correlograma da FACP deverá cair exponencialmente ou com um padrão de onda senoidal amortecida.

2.3.1.7 Modelo Auto-regressivo de Média Móvel

Uma combinação dos modelos AR(p) e MA(q) resulta em um modelo auto-regressivo e de média móvel, ou seja, um ARMA(p,q) . Um modelo auto-regressivo de média móvel de primeira ordem, um ARMA(1,1), é representado pela equação

$$x_t = \xi + \phi x_{t-1} + \varepsilon_t - \theta \varepsilon_{t-1}. \quad (2.26)$$

Um processo ARMA(p,q) é estacionário, ou seja, possui média constante ao longo do tempo. Os correlogramas desse processo, tanto da FAC como da FACP, decaem exponencialmente (positivamente ou negativamente, ou alternado entre os dois). A figura 9 apresenta os aspectos de correlogramas da FAC e da FACP, respectivamente, de um processo ARMA(1,1).

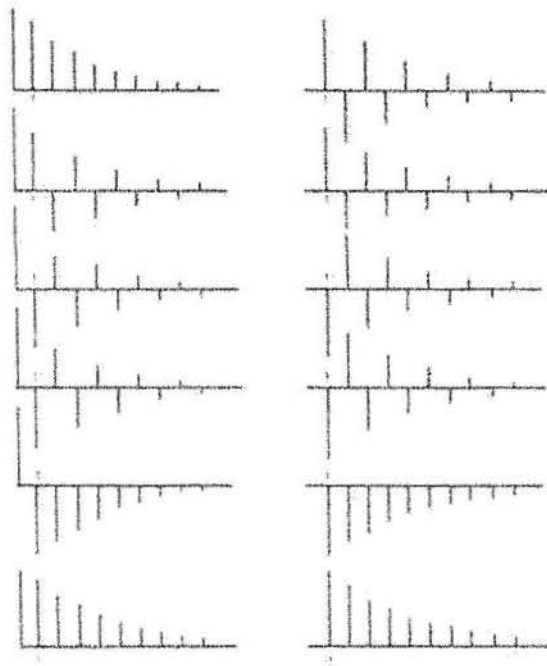


Figura 9. Padrões de correlogramas da FAC e da FACP, respectivamente, em um processo ARMA(1,1)

2.3.1.8 Modelo Auto-regressivo Integrado e de Média Móvel

Quando o processo é não estacionário, a combinação entre os modelos auto-regressivos e de média móvel, resulta em um modelo ARIMA(p,d,q), onde d é o número de

diferenças necessárias para tornar a série estacionária. O modelo abaixo é um ARIMA(1,1,1).

$$x_t = x_{t-1} + \varepsilon_t - \theta\varepsilon_{t-1} \quad (2.27)$$

Esse modelo descreve um procedimento não-estacionário (a variável x_t segue uma tendência, como se não existisse um valor fixo para a média). Isso freqüentemente ocorre em processos industriais e químicos quando x_t é um processo de saída não controlável, isto é, quando não há ações que possam ser feitas para o controle da variável a fim de atingir o valor objetivo.

Segundo Fischer (1982), na prática, a maioria das séries não estacionárias pode ser modelada por um $d = 1$ ou um $d = 2$, os quais caracterizam a série quanto ao nível e/ou quanto à inclinação. Necessita-se de $d = 1$ quando a série é não estacionária quanto ao nível do processo, isto é, quando o comportamento da série oscila em torno de uma média durante um certo período de tempo e depois ocorre uma mudança na média da série que passa a oscilar em torno dessa nova média, sem que haja uma troca significativa na direção da série. Nesse caso, diz-se que a série é não estacionária de grau um ou que apresenta uma tendência estocástica em relação ao nível da série. Por outro lado, necessita-se de $d = 2$ quando a série é não estacionária também quanto à inclinação, ou seja, quando o comportamento da série oscila em uma direção por um certo período, e depois muda para outra direção.

Um resumo comparativo do comportamento das FAC e FACP em um processo ARIMA(p,d,q) pode ser visto na tabela 1.

Ordem	(1,d,0)	(0,d,1)	(2,d,0)	(0,d,2)	(1,d,1)
comportamento do correlograma da FAC	decai exponencialmente	somente $\rho_1 \neq 0$	mistura de exponenciais ou ondas senóides amortecidas	somente $\rho_1 \neq 0$ e $\rho_2 \neq 0$	decai exponencialmente após o <i>lag</i> 1
comportamento do correlograma da FACP	somente $\phi_{11} \neq 0$	decai exponencialmente	somente $\phi_{11} \neq 0$ e $\phi_{22} \neq 0$	mistura de exponenciais ou ondas senóides amortecidas	decai exponencialmente após o <i>lag</i> 1
estimativas iniciais	$\phi = \rho_1$	$\rho_1 = \frac{-\theta}{1-\theta^2}$	$\begin{cases} \phi_1 = \frac{\rho_1(1-\rho_2)}{1-\rho_1^2} \\ \phi_2 = \frac{\rho_2-\rho_1^2}{1-\rho_1^2} \end{cases}$	$\begin{cases} \rho_1 = \frac{-\theta_1(1-\theta_2)}{1+\theta_1^2+\theta_2^2} \\ \rho_2 = \frac{-\theta_2}{1+\theta_1^2+\theta_2^2} \end{cases}$	$\begin{aligned} \rho_1 &= \frac{(1-\phi\theta)(\phi-\theta)}{1+\theta^2-2\phi\theta}, \\ \rho_2 &= \rho_1\phi \end{aligned}$
região admissibilidade	de $-1 < \phi < 1$	$-1 < \theta < 1$	$\begin{cases} -1 < \phi_2 < 1 \\ \phi_2 - \phi_1 < 1 \\ \phi_2 + \phi_1 < 1 \end{cases}$	$\begin{cases} -1 < \theta_2 < 1 \\ \theta_2 - \theta_1 < 1 \\ \theta_2 + \theta_1 < 1 \end{cases}$	$-1 < \phi < 1, -1 < \theta < 1$

Tabela 1. Comparativo de comportamento das FAC e FACP de um processo ARIMA(p,d,q), segundo Moretin e Toloí (1985).

A estimação de um modelo ARIMA pode ser feita através de uma metodologia desenvolvida por Box e Jenkins (1970), chamada de metodologia Box-Jenkins para modelos ARIMA. Essa metodologia consiste em 3 fases: Identificação (onde se preparam os dados, e se selecionam os modelos), estimação e teste (onde se estimam os modelos e são realizados testes de diagnósticos para selecionar o melhor modelo: caso os modelos não sejam adequados, retorna-se à fase da identificação), e a aplicação (onde se faz as previsões). A figura 10 resume as etapas dessa metodologia, em uma representação esquemática sugerida por Makridaskis (1998).

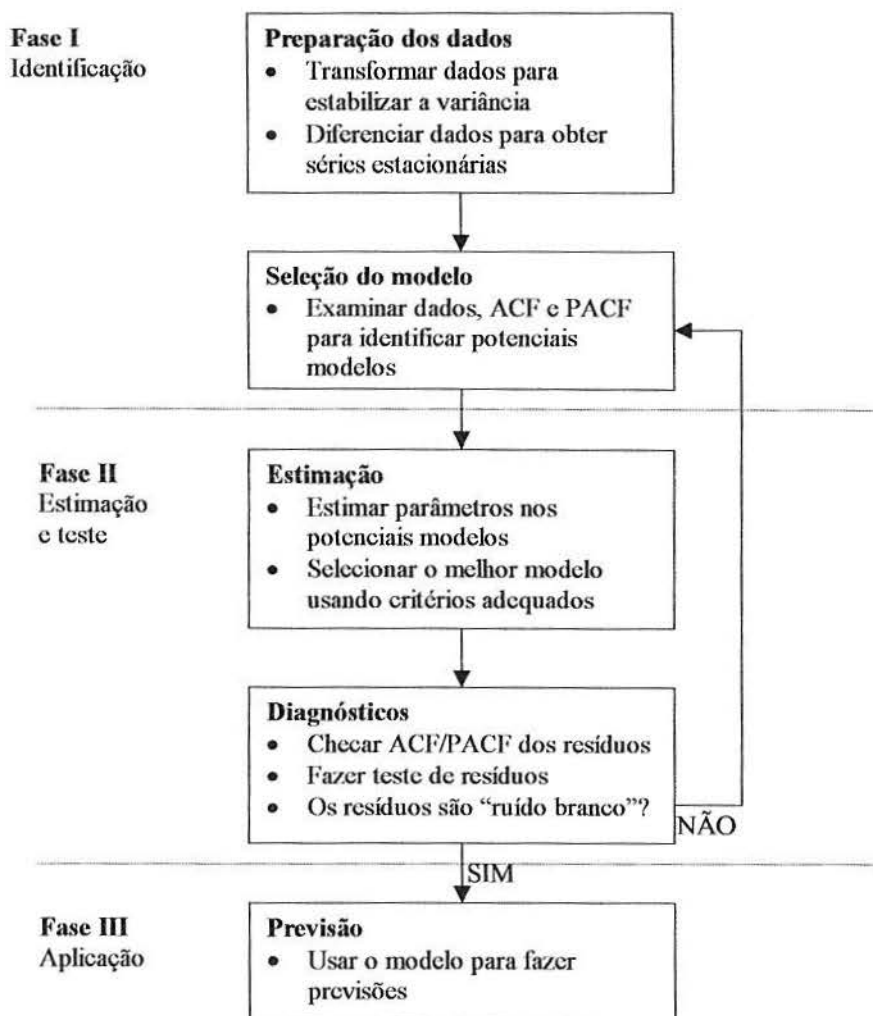


Figura 10. Representação esquemática da metodologia Box-Jenkins para modelos ARIMA

Para Moretin e Toloí (1985), a fase crítica dessa metodologia é a identificação, pois é possível que vários pesquisadores identifiquem modelos diferentes para uma mesma série temporal. Entretanto, um procedimento que muitas vezes é utilizado, é identificar não só um único modelo, mas alguns modelos para serem estimados e verificados. Um critério de decisão

recomendado, se o propósito da estimação for previsão, é escolher o modelo ajustado que fornecer o menor erro quadrático médio de previsão.

Existem outros critérios alternativos para identificar as ordens p e q de um modelo ARIMA, dentre os quais destacam-se os critérios: (i) AIC, (ii) BIC, (iii) $\phi(p, q)$ (iv) FPE, (v) CAT, (vi) critério de Anderson (1963), (vii) critério de Gray, Kelley, e McIntire (1978), (viii) critério de Nerlove, Grether e Carvalho (1979), e (ix) FACI. A seguir, encontra-se uma síntese de cada um deles, citados em Moretin e Toloí (1985).

Akaike (1973, 1974) sugere escolher o modelo tal que p e q minimizem o critério AIC (*Akaike Information Criterium*) dado por

$$AIC(p, d, q) = n \log \hat{\sigma}_e^2 + \frac{n}{n-d} 2(p + q + 1 + \delta_{d0}) + n \log 2\pi + n \quad (2.28)$$

$$\text{onde } \delta_{d0} = \begin{cases} 1, & d = 0 \\ 0, & d \neq 0 \end{cases}$$

Alternativamente, existe o critério BIC, sugerido por Akaike (1977), Rissanem (1978) e Schwarz (1978), dado por

$$BIC(p, q) = \log \hat{\sigma}_e^2 + (p + q) \log n/n \quad (2.29)$$

e o critério $\phi(p, q)$, sugerido por Hannan e Quinn (1979), dado por

$$\phi(p, q) = \log \hat{\sigma}_e^2 + (p + q)c \log \log n/n, \quad c > 2 \quad (2.30)$$

Para o caso de modelos AR(p), Akaike (1969) sugere o critério FPE (*final prediction error*), onde a ordem p é aquela que minimiza

$$FPE(k) = \left(1 + \frac{k}{n}\right) \left(1 - \frac{k}{n}\right)^{-1} S^2(k), \quad k \leq K, \quad (2.31)$$

$$\text{onde } S^2(k) = c_0 - \sum_{j=1}^k \hat{\phi}_j c_j.$$

Parzen (1976) também sugeriu um critério para o caso de AR(p), o CAT (*criterion autoregressive transfer function*), no qual a ordem p minimiza

$$CAT(p) = \frac{1}{n} \sum_{j=1}^p \frac{n-j}{n\hat{\sigma}_j^2} - \frac{n-p}{n\hat{\sigma}_p^2} \quad (2.32)$$

Um procedimento de decisão múltipla no caso de AR(p) foi desenvolvido por Anderson (1963). O procedimento consiste em testar seqüencialmente os coeficientes do modelo onde é estabelecido a priori que a verdadeira ordem p^* do modelo satisfaz $m \leq p^* \leq M$. Detalhes sobre esses testes, que são baseados na estatística t de Student, podem ser encontrados em Anderson (1976) e Mentz et al. (1980).

Um procedimento de identificação de modelos ARMA, através de inspeção de matrizes apropriadas, calculadas em função das autocorrelações, é apresentado por Gray, Kelley, e McIntire (1978).

Um procedimento alternativo para identificar modelos ARIMA a partir somente das autocorrelações, desenvolvido por Nerlove, Grether e Carvalho (1979), consiste em tentar separar as partes AR e MA do processo, transformando a série original de tal forma que ela possa ser aproximada por um processo MA puro.

Cleveland (1972) introduziu a função de autocorrelação inversa (FACI), que é definida em termos do recíproco do espectro da série. Chatfield (1979) apresenta um tratamento mais intuitivo, e McClave (1978) e McLeod et al. (1977) apresentam aplicações.

Após selecionar o modelo, deve-se verificar se ele representa ou não corretamente os dados. Se houver a suspeita de que um modelo mais elaborado (com mais parâmetros) é necessário, pode-se utilizar uma técnica chamada de superajustamento, que estima o modelo com parâmetros extras e verifica se eles são significativos e se a variância residual diminui significativamente.

A verificação do modelo ajustado pode ser feita pela análise dos resíduos (erros): se o modelo for adequado, os resíduos constituirão um ruído branco. Essa e outras técnicas utilizadas para a verificação, estão descritas na seção 2.2 dessa monografia.

2.3.2 Aplicações de gráficos de controle após a estimação de um modelo ARIMA

Depois de aplicar um modelo ARIMA adequado nos dados do processo, com o objetivo de remover a autocorrelação e obter os resíduos conhecidos como “ruído branco”, é necessário aplicar um tipo de gráfico de controle para monitorar o processo. Uma das mais importantes propriedades estudadas pelos pesquisadores na escolha de um gráfico para controle dos resíduos é a distribuição do comprimento de corrida (RLD – *run-length distribution*). O RL é o número de amostras ou observações necessárias para o gráfico sinalizar um estado fora de controle ou emitir um alarme falso (no caso de mudança na média). O conhecimento da distribuição RL permite estimar a ARL – quantidade média de amostras ou observações para que o gráfico sinalize um estado fora de controle ou emita um alarme falso – e o SRL ou SDRL – o desvio padrão de RL. Normalmente deseja-se que a ARL seja grande quando não há causas especiais, e pequena quando há (para que se possa identificá-la mais rapidamente). O conhecimento de ARL para uma particular causa especial (usualmente uma mudança na média do processo) permite desenvolver um gráfico de controle mais eficaz.

Entre os gráficos de controle, destacam-se: (i) gráfico de resíduos, (ii) CUSUM, (iii) EWMA, (iv) gráficos de controle $\lambda_{LS,max}$ e $\bar{\lambda}_{LS}$, (v) gráfico com intervalo variável, (vi) gráfico ARMA, e (vii) o método da estimação conjunta.

2.3.2.1 Gráfico de causa especial ou gráfico de resíduos (*special-cause chart, SCC*)

O gráfico de causa especial (SCC) é um gráfico tradicional de Shewhart para resíduos. Enquanto que o gráfico tradicional de Shewhart monitora as observações, o SCC monitora os resíduos ou erros, que são a diferença entre os valores observados e os valores estimados pelo modelo ajustado.

Wardell, Moskowitz, e Plante (1994) deduziram uma distribuição RL de gráfico de causa especial (SCC) proposto por Alwan e Roberts (1988), dado que a causa especial a ser detectada é uma mudança na média do processo. Ambas soluções recursivas e de forma fechada são deduzidas para a distribuição RL, o ARL e o SRL para qualquer processo AR(p), e soluções aproximadas são deduzidas para o processo mais geral ARMA(p,q). As expressões deduzidas não dependem do tipo de mudança da média do processo (por exemplo, se é uma mudança pequena, ou uma mudança brusca, etc). Os resultados mostram que ARL e SRL dos

gráficos de controle de resíduos são relativamente menores quando o processo é mais negativamente do que positivamente autocorrelacionado. Desconsiderando o sinal da autocorrelação, a forma da função massa de probabilidade do RL revela que a probabilidade de detectar mudanças antecipadamente é substancialmente superior para o SCC do que para os mais tradicionais gráficos de controle. Uma detecção antecipada faz com que a causa do indício seja identificada facilmente, resultando numa maior rapidez na implantação contínua da qualidade. Uma das conclusões obtidas foi que o SCC não é necessariamente o melhor gráfico para o uso em todo tipo de processo autocorrelacionado. Em alguns casos, gráficos convencionais de controle, como Shewhart ou EWMA, possuem bons resultados quando o processo é autocorrelacionado positivamente e são freqüentemente melhor em termos de ARL (porque são fáceis de implantar).

Programas em linguagem FORTRAN foram desenvolvidos com base na metodologia desenvolvida por Wardell, Moskowitz, e Plante (1994) para o cálculo da distribuição RL, da ARL e da SDRL para gráficos de controle de resíduos. RLD, ARL e SDRL são calculados para processos que possam ser modelados por AR(1), MA(1) e ARMA(p,1), dado que a causa especial a ser detectada é uma pequena mudança na média do processo.

Runger (2002) realizou um estudo recente sobre qual o tipo de gráfico de controle utilizar (para observações ou resíduos) nos casos de processos com causas especiais e autocorrelações. Um gráfico de controle baseado nas observações pode ser eficiente para o CEP, mas seu sucesso depende do modelo de séries temporais. Essa caracterização produz introspecção na pesquisa que compara gráficos baseados nos resíduos com gráficos baseados nos dados puros. A contribuição do artigo está no uso de um modelo de perturbação fisicamente realístico que difere daquele na maioria das pesquisas de autocorrelação em CEP. Além disso, é lógico que um gráfico de controle baseado nos resíduos é preferido em geral para esse modelo. Entretanto, há casos nos quais um gráfico de controle baseado nos dados originais pode ser tão eficaz quanto um baseado nos resíduos. O mesmo raciocínio simples é então usado para ilustrar as desvantagens dos métodos de aproximação que usam um preditor EWMA e as vantagens de um método de distribuição livre (*model free*) baseado em médias de grupos.

2.3.2.2 Gráficos CUSUM

Os gráficos de controle para média e para os valores individuais não são eficazes para detectar uma mudança pequena na média de um processo, isto é, mudanças menores que $1,5\sigma$ na média. Nessas situações, é eficaz utilizar o gráfico de soma acumulada (CUSUM – *Cumulative Sum*), apesar de não ser um gráfico de controle por não ter limites de controle. A estimativa da soma acumulada da i -ésima observação é dada por

$$C_i = \sum_{j=1}^i (\bar{x}_j - \mu_0), \text{ para } n \geq 1 \quad (2.33)$$

onde \bar{x}_j é a média da j -ésima amostra, no caso de um gráfico CUSUM para monitorar a média, e é a j -ésima observação no caso de um gráfico CUSUM para monitorar as observações individuais.

Nota-se que o gráfico CUSUM incorpora a seqüência da informação dos valores amostrais por plotar a soma acumulada dos desvios em torno da média. Montgomery (1997) fez um estudo simulado de 30 observações normalmente distribuídas, sendo as 20 primeiras com média 10 e desvio padrão 1, e as 10 últimas com média 11 e desvio padrão 1, ou seja, uma mudança da média do processo em 1 desvio padrão, a qual um gráfico tradicional de controle não detectaria. A figura 11 apresenta o comportamento do gráfico CUSUM para esse estudo simulado. Verifica-se que o gráfico consegue facilmente detectar a mudança na média.

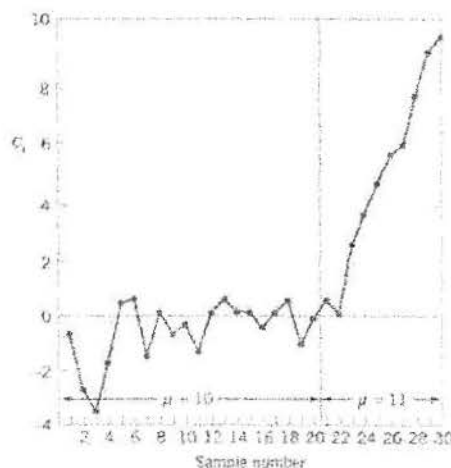


Figura 11. Exemplo de gráfico CUSUM para detectar mudanças pequenas na média de um processo.

Existem ainda mais duas maneiras de representar um gráfico CUSUM: o CUSUM tabular (ou algorítmico), e a máscara V (*V-mask*) de um CUSUM. Entretanto, Montgomery (1997) mostra que a maneira “tabular” é a mais eficaz dessas duas.

O gráfico CUSUM tabular pode ser feito tanto para a média quanto para os valores individuais. A construção desse gráfico é dada através do cálculo de desvios da média que estão acima do alvo pela estatística C_i^+ e abaixo do alvo pela estatística C_i^- . As estatísticas C_i^+ e C_i^- são chamadas de laterais superior e inferior do CUSUM, e são, respectivamente, dadas por

$$C_i^+ = \max\left[0, x_i - (\mu_0 + K) + C_{i-1}^+\right], \text{ e} \quad (2.34)$$

$$C_i^- = \max\left[0, (\mu_0 - K) - x_i + C_{i-1}^-\right], \quad (2.35)$$

onde os valores iniciais para $i=1$ são zero, e K é chamado de valor de referência.

Se a mudança da média é expressa em unidades de desvio padrão, como $\mu_1 = \mu_0 + \delta\sigma$, então

$$K = \frac{\delta}{2}\sigma = \frac{|\mu_1 - \mu_0|}{2}. \quad (2.36)$$

Se C_i^+ ou C_i^- excede o intervalo de decisão H , o processo está fora de controle. Um valor razoável para H é 5σ . A figura 12 ilustra um processo com o gráfico CUSUM tabular para o estudo simulado de Montgomery (1997).

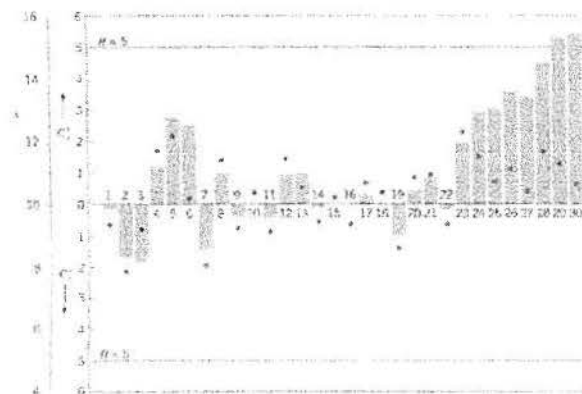


Figura 12. Exemplo de gráfico CUSUM tabular.

Para identificar quando ocorreu a mudança da média, olha-se o primeiro ponto fora de controle e diminui-se do valor de N^+ ou de N^- . O valor de N^+ ou de N^- é dado pela quantidade de observações consecutivas positivas ou negativas. No exemplo do gráfico, a mudança da média provavelmente ocorreu na observação $29 - 7 = 22$.

Montgomery (1997) apresenta um estudo de ARL para esse gráfico e argumenta que as escolhas de K e H devem ser feitas para proporcionar um bom desempenho do ARL.

O outro gráfico proposto é gráfico obtido através do método da máscara V, que é aplicado a valores sucessivos da estatística CUSUM

$$C_i = \sum_{j=1}^i \frac{(x_j - \mu_0)}{\sigma} \quad (2.37)$$

A estrutura da máscara V é apresentada na figura 13.

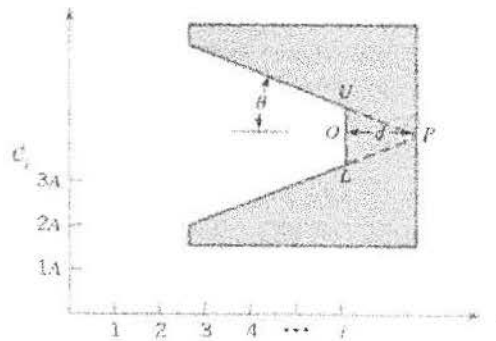


Figura 13. Estrutura da máscara V

O procedimento do gráfico de controle CUSUM Máscara V consiste em colocar o ponto O no último valor de C_i , e traçar uma linha OP paralela ao eixo horizontal. O desempenho do gráfico é determinado pela distância d e pelo ângulo θ . Montgomery (1997) sugere que esses parâmetros sejam calculados, respectivamente, por

$$d = \left(\frac{2}{\delta^2} \right) \ln \left(\frac{1-\beta}{\alpha} \right), \text{ e por} \quad (2.38)$$

$$\theta = \tan^{-1} \left(\frac{\delta}{2A} \right). \quad (2.39)$$

onde A é a distância horizontal na Máscara V entre os pontos sucessivos, em termos da unidade de distância da escala vertical; 2α é a maior probabilidade permissível de um sinal quando a média do processo está no alvo (um alarme falso); e β é a probabilidade de não detectar uma mudança em δ . Quando β é pequeno, d pode ser calculado por

$$d \cong -\frac{\ln(\alpha)}{\delta} \quad (2.40)$$

A figura 14 mostra o desempenho desse gráfico com os dados do estudo simulado de Montgomery (1997).

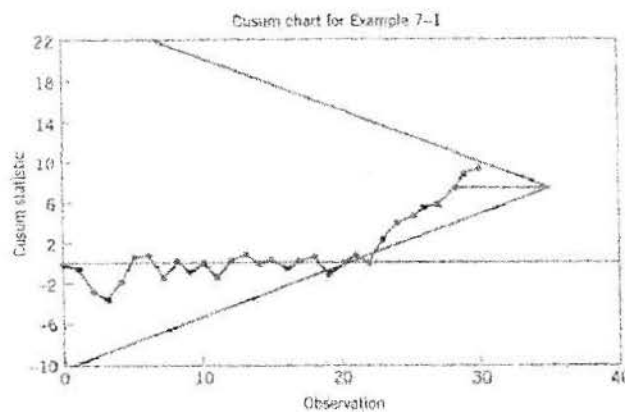


Figura 14. Exemplo de gráfico CUSUM Máscara V.

Algumas desvantagens são encontradas nesse tipo de gráfico: a máscara V é bilateral, não sendo muito útil para processos unilaterais; a técnica chamada de método do começo pelo topo (*headstart feature*) não pode ser utilizado; dificuldades na interpretação; e ambigüidade associada com os parâmetros α e β .

Yashchin (1993) examinou um método que possibilita estimar a distribuição RL de um método de controle CUSUM quando o dado subjacente apresenta correlação serial moderada. É demonstrado que problemas envolvendo monitoramento de dados correlacionados serialmente exigem a aplicação de um método. É mostrado que pode-se substituir a seqüência de observações por uma seqüência *iid* para os quais a distribuição RL é aproximadamente a mesma. O método é simples de aplicar, produz resultados exatos quando não há correlação serial presente, e conduz a aproximações de boa qualidade quando a magnitude da correlação serial não é tão grande, especialmente para processos Gaussianos (distribuição Normal). Os

exemplos incluídos ilustram como o método proposto pode ser usado em um gráfico CUSUM para a média e para o desvio padrão. Analogamente, podemos usá-lo para controlar outros tipos de gráficos de controle CUSUM. Quando a proporção da variância na sequência de controle $\{X_i\}$ explicada pela presença de correlação serial é muito alta (acima de 25%), não se deve tentar controlar a sua média pela aplicação do procedimento de Page (1954) diretamente, que é o método CUSUM. Em vez disso, deve-se primeiro transformar $\{X_i\}$ para reduzir a magnitude da correlação serial e então aplicar esse procedimento para a sequência transformada. O método proposto é então usado para encontrar os parâmetros apropriados para esse procedimento. Em muitas aplicações industriais o propósito de um conjunto de dados é conduzir o processo para o seu respectivo alvo com variabilidade pequena. Em situações semelhantes, a presença de uma correlação serial estável, às vezes, produz uma oportunidade para introduzir uma sistemática *feedforward-feedback* de ação corretiva que reduzirá o desvio do processo em torno da média. Quando relevante, aplicações de métodos descritos na literatura EPC poderão ser usadas.

Quando observações do processo estão autocorrelacionadas, o desempenho do gráfico CUSUM não é eficiente. Para eliminar o efeito da autocorrelação na monitoração da mudança na média, o gráfico CUSUM é tipicamente aplicado nos resíduos do modelo da série temporal para o processo. Como uma alternativa, Atienza, Tang, e Ang (2002) desenvolveram um método CUSUM para detectar uma mudança pequena na média do processo que utiliza observações de um processo autocorrelacionado. As análises mostram que o desempenho do método proposto é muito competitivo quando comparado com os gráficos Shewhart e CUSUM baseados nos resíduos. O gráfico de controle proposto por Shewhart (1932) é conhecido por ser o melhor para detectar mudanças grandes na média; por outro lado, os gráficos CUSUM, de Page (1953) e EWMA, de Roberts (1959) são melhores para detectar mudanças pequenas na média do processo. Na análise do desempenho desse método CUSUM proposto, o foco era a detecção de mudanças na média de um processo AR(1) com $\phi \geq 1$. Também foram analisados o comportamento do gráfico V e o desempenho da ARL. Dentre as conclusões, destaca-se que o esquema CUSUM proposto é menos sensível para especificações de ϕ quando o processo é fracamente autocorrelacionado positivamente do que quando o processo é fortemente autocorrelacionado positivamente.

Lu e Reynolds (2001) investigaram gráficos CUSUM baseados em observações e

gráficos CUSUM baseados nos resíduos, e mostraram que, para níveis moderados de autocorrelação, ambos os tipos de gráfico CUSUM necessitam quase a mesma quantidade de tempo para detectar mudanças na média dos processos. Para níveis altos de autocorrelação, os dois tipos dos gráficos CUSUM possuem desempenho similarmente para mudança pequena na média, mas o CUSUM dos resíduos é um pouco melhor que o CUSUM de observações para mudanças grandes na média. Os gráficos CUSUM foram comparados com os gráficos EWMA baseados nas observações e nos resíduos. O desempenho dos gráficos EWMA na detecção de mudanças na média do processo é similar ao desempenho dos gráficos CUSUM. Assim, para aplicações práticas, a escolha entre os gráficos CUSUM e EWMA podem ser baseadas em outros fatores tal como facilidade de interpretação. Os resultados obtidos por ambos gráficos CUSUM e EWMA mostram que é satisfatório usar um gráfico baseado nas observações originais, preferivelmente que o de resíduos, a não ser que o grau da autocorrelação seja relativamente alto. A maioria dos profissionais provavelmente preferirá usar um gráfico baseado nas observações originais, pois esse gráfico é mais fácil de interpretar do que um baseado nos resíduos.

Lu e Reynolds (2001) mostraram que a autocorrelação tem um efeito muito significativo em gráficos de controle baseados nas observações originais. Assim, para um gráfico baseado nas observações originais, é necessário calcular para a autocorrelação de ambos na estimação do desvio padrão do processo e na determinação dos limites de controle. Além do que, para níveis elevados de autocorrelação, isso pode ser útil para aumentar o $\delta/2$ no gráfico de observações CUSUM e λ no gráfico de observações EWMA para obter um bom desempenho na detecção de mudanças no processo. Um efeito importante da autocorrelação é que torna as mudanças pequenas na média do processo mais difícil de detectar do que no caso de observações independentes. Essa é a razão pela qual pode ser recomendável aumentar o $\delta/2$ no gráfico de observações CUSUM e λ no gráfico de observações EWMA. Quando há autocorrelação significativa, mudanças pequenas na média do processo não podem ser detectadas em um tempo razoável usando algum valor de $\delta/2$ e de λ . Assim, ajustar o gráfico para detectar mudanças pequenas piorará o desempenho para mudanças grandes, sem oferecer uma capacidade razoável de detecção para mudanças pequenas.

2.3.2.3 Gráficos EWMA

Montgomery e Mastrangelo (1991) sugeriram um procedimento de aproximação baseado nos gráficos de controle EWMA (média móvel ponderada exponencialmente). Eles o utilizaram pelo fato de que o EWMA pode ser utilizado em certas situações onde os dados estão autocorrelacionados. Suponha-se que o processo possa ser modelado por um modelo do tipo ARIMA. Pode ser facilmente mostrado que o EWMA com $\lambda = 1 - \theta$ é o otimizador da previsão um passo à diante para esse processo. Isto é, se $\hat{x}_{i+1}(i)$ é o valor previsto para a observação no período $i + 1$ feito até o final do período i , então

$$\hat{x}_{i+1}(i) = z_i \quad (2.41)$$

onde $z_i = \lambda x_i + (1 - \lambda)z_{i-1}$ é o EWMA. A seqüência dos erros de previsão um passo à diante

$$e_i = x_i - \hat{x}_i(i-1) \quad (2.42)$$

é independente e identicamente distribuída com média zero. Assim, os gráficos de controle podem ser aplicados a essa seqüência de erros de previsão um passo à diante. O parâmetro λ (ou equivalentemente, θ) deverá ser encontrado pela minimização da soma dos quadrados dos erros e_i , como pode-se ver na figura 15.

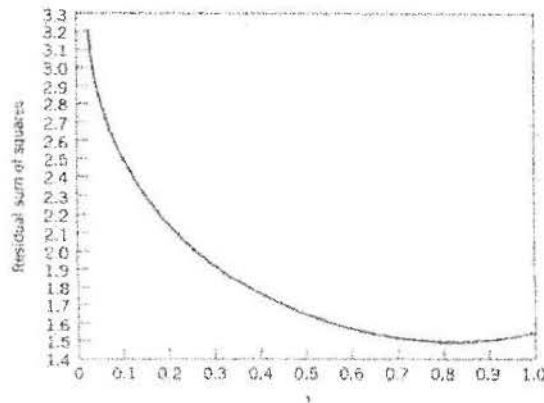


Figura 15. Exemplo de relacionamento entre a soma dos quadrados dos erros de previsão um passo à diante (eixo das ordenadas) e o valor de lambda (eixo das abscissas)

Os limites superior e inferior de controle para o gráfico EWMA são dados, respectivamente, por

$$UCL = \mu_0 + L\sigma\sqrt{\frac{\lambda}{2-\lambda}\left[1 - (1-\lambda)^{2i}\right]} \quad (2.43)$$

$$LCL = \mu_0 - L\sigma\sqrt{\frac{\lambda}{2-\lambda}\left[1 - (1-\lambda)^{2i}\right]} \quad (2.44)$$

onde L é a largura dos limites de controle.

A figura 16 apresenta o gráfico de controle EWMA para os dados do estudo simulado de Montgomery (1997). Percebe-se uma mudança brusca no comportamento do gráfico a partir da observação 23.

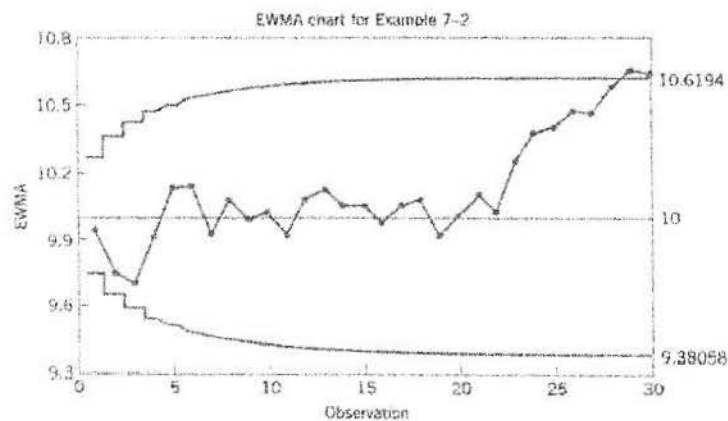


Figura 16. Exemplo de Gráfico de Controle EWMA

Lu e Reynolds (1999) estudaram o problema do monitoramento da média em gráficos de Shewhart e EWMA com observações autocorrelacionadas. O objetivo concentrava-se em avaliar o desempenho estatística desses gráficos e sugerir recomendações que possam ser utilizadas em aplicações práticas. Para obtenção de resultados numéricos foi utilizado um modelo simples de séries temporais, um AR(1), com um erro aleatório adicional. Uma conclusão geral obtida nesse estudo é que monitorar um processo desse tipo é consideravelmente mais difícil que monitorar um processo no qual as observações podem ser modeladas como variáveis aleatórias normais independentes. Uma razão para essa dificuldade é que a autocorrelação parece mascarar mudanças pequenas da média. Para um nível de autocorrelação razoavelmente alto, o tempo necessário para detectar uma mudança na média, no qual é uma dada fração do desvio padrão do processo, é muito maior do que para a mesma mudança no caso de independência. Outra razão é que o planejamento de um gráfico para

alcançar as propriedades desejadas requer um processo mais sofisticado do modelo e dos parâmetros estimados e um conjunto de dados iniciais maior do que no caso de independência. Adaptar um gráfico de controle para monitorar um processo usualmente requer parâmetros estimados usando um conjunto de dados iniciais, calculando os limites de controle, verificando se o processo estava estável quando os dados foram obtidos, e então usando os limites de controle para monitoramentos futuros. Para o caso de independência e normalidade nos dados, há no máximo dois parâmetros para estimar, e o processo de ajuste do gráfico é explicado em vários livros textos. Para situações onde a autocorrelação é significativa, não parece haver nenhuma alternativa para tentar modelar essa autocorrelação em algum modo. Apesar de que softwares para ajustar um modelo de séries temporais sejam amplamente disponíveis, o processo de seleção do modelo e de ajuste requer mais sofisticação por parte do usuário. No caso de observações independentes, uma prática comum é eliminar os pontos do grupo de dados iniciais que estão fora dos limites de controle e para que as causas possam ser identificadas. Então, os parâmetros do processo são re-estimados. No caso de dados autocorrelacionados, o processo de identificação de dados que não seguem o modelo de controle assumido pode ser muito mais difícil que no caso de dados independentes.

O tamanho do grupo de dados iniciais utilizados para estimar os parâmetros do processo é uma informação importante assim como no caso de observações independentes. Os resultados apresentados por Lu e Reynolds (1999) mostram que essa informação é mais crítica quando ajustada por um modelo de séries temporais. Em particular, um grupo de dados iniciais muito grande é requerido para que as propriedades dos gráficos de controle com os parâmetros estimados do processo sejam similares àquelas calculados sob a suposição de que os parâmetros sejam conhecidos. O tamanho do grupo de dados necessários é muito maior que o usado tradicionalmente para a construção dos gráficos de controle. Em muitas aplicações práticas, grupos de dados muito grandes não estão disponíveis; desse modo, há uma necessidade para mais pesquisa e desenvolvimento de trabalho na monitoração de procedimentos os quais não são altamente sensíveis para modelos com falhas de especificação e para erros nos parâmetros estimados.

Lu e Reynolds (1999) também estudaram simultaneamente o monitoramento da média e da variância de processos autocorrelacionados considerando o ajuste por um modelo AR(1) mais um erro aleatório e compararam o desempenho dos gráficos de Shewhart e EWMA.

Nesse caso, o problema da monitoração é mais complicado porque o modelo utilizado envolve vários parâmetros: a média global, dois parâmetros de variância e o parâmetro auto-regressivo. Em algumas aplicações, uma mudança no processo pode afetar um desses parâmetros ou uma combinação desses parâmetros. Por exemplo, em um processo de grupos, uma causa especial pode produzir um aumento na variabilidade dentro do grupo, na variabilidade entre grupos ou ambos. Em muitas aplicações práticas é importante saber os parâmetros que foram mudados, pois assim seria mais fácil de identificar a causa.

Muitos pesquisadores da área têm estudado o monitoramento da média na presença de autocorrelação. Entretanto, o estudo de Lu e Reynolds (1999) considera o monitoramento da variância e também simultaneamente a variância e a média. O desempenho relativo dos diferentes gráficos de controle depende dos valores dos parâmetros do modelo do processo e do tipo e da magnitude das possíveis mudanças do processo. Em geral, nenhum gráfico de controle simples terá um desempenho ótimo a uma ampla variedade de situações. Entretanto, um gráfico de controle pode ser escolhido por ter um bom desempenho a um tipo e a uma magnitude particulares da mudança do processo. Muitos métodos têm sido identificados por funcionar razoavelmente bem em uma variedade de situações. Lu e Reynolds (1999) concluíram, baseados em propriedades estatísticas e facilidade de interpretação, que o método que utiliza simultaneamente um gráfico de resíduos de Shewhart e um gráfico EWMA de observações pode ser recomendado para o caso de correlação baixa ou moderada.

Crowder (1987) desenvolveu um método simples para o estudo de distribuições RL em gráficos EWMA em processos independentes e identicamente distribuídos com distribuição normal, através do uso de equações integrais para a tabulação de momentos de RLD. Segundo ele, a aproximação por essas equações dá uma expressão exata para momentos de um RL associado com um esquema EWMA. O procedimento pode ser estendido para muitos casos de distribuição não-normal para versões unilaterais do gráfico EWMA.

Wieringa (1999) estudou o desempenho do ARL em gráficos CUSUM para a média e comparou com as curvas dos gráficos ARL dos gráficos de Shewhart e EWMA. Ele também comparou as curvas dos gráficos ARL em processos AR(1) com o caso de observações independentes e identicamente distribuídas em gráficos EWMA.

Harris e Ross (1991) estudaram o impacto das observações autocorrelacionadas no

desempenho dos gráficos CUSUM e EWMA. Eles utilizaram bases de medidas ponderadas para ilustrar o uso do método de séries temporais.

2.3.2.4 Gráficos de Controle $\lambda_{LS,max}$ e $\bar{\lambda}_{LS}$

Atienza, Tang e Ang (1998) analisaram o caso em que a aplicação das técnicas de CEP com a escolha de um modelo ARMA não é uma boa solução quando o processo está altamente correlacionado positivamente. Eles exploraram o uso de estatísticas no procedimento de séries temporais para detectar *outliers* (observações fora dos limites de controle) e mudanças na média do processo em processos autocorrelacionados com particular ênfase no modelo AR(1). Os resultados do estudo mostram que um método para monitorar mudanças em níveis de processos autocorrelacionados pode ser baseado em $\lambda_{LS,max}$ ou em $\bar{\lambda}_{LS}$. O gráfico de controle $\lambda_{LS,max}$ possui as propriedades desejáveis de ambos gráficos de Shewhart e de CUSUM, e oferece um desempenho ARL superior comparado com os procedimentos existentes para detectar níveis de mudanças em processos autocorrelacionados. Comparado com os gráficos SCC e $\lambda_{LS,max}$, $\bar{\lambda}_{LS}$ mostra-se mais sensível em detectar mudanças pequenas, mas menos sensível para mudanças grandes. O método de gráfico de controle proposto pode ser facilmente estendido para detectar a presença de *outliers* aditivos (AO) e *outliers* inovadores (IO). Um AO afeta o nível da série temporal observada apenas no tempo $t = d$, por outro lado, um IO afeta o nível da série temporal observada no tempo $t = d$, e depois de $t = d$ esse efeito se enfraquece exponencialmente. Uma identificação mais refinada do tipo de intervenção nos permitirá descobrir a origem da situação fora de controle, o que é um passo importante na eliminação de causas especiais de variação. O procedimento proposto também pode ser aplicado com um modelo geral ARIMA.

2.3.2.5 Gráfico para a média com Intervalo Amostral Variável (VSI)

Um estudo de monitoramento de gráficos de controle para a média com dados autocorrelacionados, utilizando um intervalo amostral variável (*variable sampling interval, VSI*) foi desenvolvido por Reynolds, Arnold e Baik (1996). Os resultados, baseados em um modelo AR(1), mostraram que a correlação entre as observações de um processo pode ter efeito significativo nas propriedades dos gráficos de intervalo amostral fixo (*fixed sampling*

interval, FSI) e VSI. Quando se usa o método VSI, a melhor escolha para a função de intervalo amostral para as observações correlacionadas não é necessariamente a mesma para as observações independentes. Quando o nível de correlação positiva aumenta, a vantagem o gráfico VSI para a média em relação ao FSI é reduzida, e o uso do gráfico da média no modo tradicional, com ou sem o método VSI, torna-se desejável.

2.3.2.6 *Gráfico de Controle ARMA*

Jiang, Tsui e Woodall (2000) propuseram um novo método de monitoramento chamado de gráfico ARMA (Auto-regressivo e de média móvel), baseado no monitoramento de uma estatística ARMA. Segundo eles, esse novo método, visto como uma estratégia informal, providencia uma escolha mais flexível dos parâmetros para descrever a estrutura de autocorrelação para o desempenho do gráfico e inclui os gráficos SCC e EWMAST (Schimid 1997; Zhang 1998; Adams e Tseng 1998; Lu e Reynolds 1999a,b) para processos autocorrelacionados, os quais são casos especiais de gráficos ARMA.

Os autores deduziram duas razões chamadas de razão transitória e razão estabilizada, que são criteriosas para a escolha apropriada dos parâmetros para gráficos ARMA em processos autocorrelacionados. A razão transitória mede a capacidade do gráfico de detectar as mudanças nas primeiras poucas observações e é mais apropriado para detectar mudanças grandes. Se o gráfico falhar em sinalizar a mudança nas primeiras observações, a razão estabilizada torna-se importante por detectar eficientemente a mudança nas observações posteriores. Se a razão transitória puder se tornar um valor alto suficiente (por exemplo, 4 ou 5), pela escolha apropriada dos parâmetros do gráfico ARMA, o gráfico correspondente será capaz de detectar mudanças rapidamente. Por outro lado, se essa razão for menor que 3, a mudança provavelmente será equivocada para o estado transitório e precisará ser detectada nas observações posteriores.

2.3.2.7 *Gráfico de Médias de Grupos*

Runger e Willemain (1995) investigaram o método do gráfico do grupo de médias ponderadas (WBM – *weighted batch means*) desenvolvido por Bischak, Kelton e Pollock (1993) e compararam com o gráfico do grupo de médias não ponderadas (UBM – *unweighted batch means*) proposto por Runger e Willemain (1995). Foi utilizado o critério ARL em um

modelo AR(1) para resolver o problema da autocorrelação, e comparados os resultados com ambos os gráficos de controle de Shewhart e CUSUM. Três gráficos de Shewhart em diferentes transformações dos dados do processo foram comparados, todos propostos para eliminar a autocorrelação. Em um gráfico de resíduos para valores individuais de um modelo de séries temporais; se o modelo estiver correto e seus parâmetros bem estimados, os resíduos terão autocorrelação desprezível. O WBM usa o mesmo modelo de séries temporais para determinar os pesos que retribuem o grupo de médias não correlacionadas. O gráfico UBM não precisa de modelo de séries temporais, desde que ele simplesmente assinale no gráfico as médias aritméticas dos valores dos dados consecutivos; neste caso, a autocorrelação é diminuída pela escolha de um tamanho do grupo grande o suficiente.

O gráfico do grupo de médias pode ter melhor desempenho ARL do que o gráfico para resíduos. A limitação do gráfico de resíduos origina-se de uma rápida reação do modelo AR(1) para uma mudança na média do processo. Essa reação minimiza a chance de detectar uma mudança na média após o acontecimento da primeira mudança. O gráfico WBM funciona melhor porque efetivamente calcula a média dos resíduos, permitindo uma detecção mais confiável para a presença de mudanças discretas nos resíduos segundo e subsequente depois da alteração. O gráfico UBM apresenta mais vantagens. A variância do UBM não é aumentada pelos grandes pesos dados para os pontos finais no WBM.

Os resultados para os dois gráficos do grupo de médias assumem que as causas especiais ocorrem antes do começo do gráfico de controle, ou seja, o processo já está fora de controle antes de ser monitorado. Por outro lado, uma causa especial quase sempre ocorre primeiramente dentro dos grupos do que entre os grupos. Uma meta do subgrupo racional no tradicional CEP (dados não correlacionados) é reduzir a probabilidade deste evento. Se assumirmos que um grupo que contém uma causa especial nunca excede o limite de controle, então os ARLs para os gráficos do grupo de médias deveriam ser acrescentados pelo número médio de observações do período da causa especial até o grupo ser completado. Essa suposição é realista para mudanças pequenas em processos, mas conservadora para mudanças grandes.

Um modo para interpretar o resultado é como uma comparação de tentativas de aproximação de modelos baseados em distribuições (*model-based*) e modelo de distribuição livre (*model-free*) para CEP com dados autocorrelacionados. Assume-se que a inferência de

modelos baseados em distribuições seria a mais poderosa alternativa, já que a inferência paramétrica é mais poderosa do que a não paramétrica. Entretanto, foi mostrado que gráficos de controle *model-free* (UBM) podem ter desempenho ARL superior.

Outra vantagem do gráfico UBM apontada por Runger e Willemain (1995) é sua similaridade para os métodos CEP hoje extensamente em uso. Gráficos de controle baseados em grupos de médias podem ser construídos e interpretados de acordo com as normas tradicionais para dados não correlacionados. O procedimento da média de observações sucessivas para gerar um ponto no gráfico é simples e consistente em aproximações convencionais tanto conceitualmente como mecanicamente. Os aspectos produzidos nos softwares comerciais de CEP podem ser usados para construir esses gráficos de controle.

Em um contexto similar, notou-se que implementações de sistemas digitais de controle tipicamente incorporam grupos de médias para filtrar ruídos, e que práticas de engenharia têm tido vantagens nesse aspecto em desenvolver gráficos por muitos anos. Finalmente, nota-se que os gráficos UBM não são os únicos meios disponíveis para tratar com dados autocorrelacionados sem um modelo de séries temporais. Willemain e Runger (1995) estabeleceram o valor dos gráficos de controle baseados em estatísticas *level-crossing*. A aproximação *level-crossing* pode ser útil especialmente com dados de distribuições assimétricas ou com cauda pesada.

2.3.2.8 Método da Estimação Conjunta (JE)

Wright, Booth, e Hu (2001) investigaram o uso do método de detecção de *outliers* de estimação conjunta (*joint estimation*, JE) como um método CEP para dados autocorrelacionados em corridas curtas (*short-run*). Já que JE é capaz de diferenciar entre quatro tipos diferentes de *outliers* (observações fora de controle), o seu desempenho é relatado com respeito a sua habilidade para encontrar a observação fora de controle e identificar o tipo de associação. Esse é um interesse particular para os profissionais da área porque os quatro tipos diferentes de *outliers* podem indicar diferentes problemas em um processo. Os resultados encontrados no estudo mostram que JE possui um desempenho melhor para modelos AR(1) quando a observação fora de controle é a última observação do que para modelos MA(1). Entretanto, JE é mais eficaz em distinguir entre os quatro tipos de *outliers* para modelos MA(1) do que para modelos AR(1).

Segundo Liu et al. (1992), os quatro tipos de *outliers* são: o *outlier* aditivo (AO), que é um evento em um período da série, mas que não afeta a série temporal, a não ser no período em que ocorre; a mudança de nível (LS), que ocorre em um certo período, mas que afeta a série a partir do período em que ocorre, isto é, há uma mudança na média do processo; *outlier* inovador (IO), o qual afeta a série temporal após a ocorrência do *outlier* no período t de acordo com o processo ARIMA; e mudança temporária (TC), que é um evento com um impacto inicial no período em que ocorre e que decai exponencialmente.

Wright, Booth, e Hu (2001) resumem o método JE, o qual pode ser encontrado na íntegra em Chen e Liu (1993a, 1993b). O primeiro estágio do método consiste em estimar os parâmetros para um modelo ARIMA ajustado e a detecção dos *outliers*. Se não houver *outliers*, o procedimento termina. No segundo estágio, o procedimento estima conjuntamente o efeito de *outliers*, através de regressão múltipla, e os parâmetros do modelo. No terceiro estágio, o procedimento procura detectar *outliers* baseado nos parâmetros finais estimados.

Algumas recomendações para o uso de JE são salientadas: determinar o mais apropriado modelo ARIMA para a série; utilizar as tabelas desenvolvidas no estudo para verificar os níveis aceitáveis de detecção de locais de *outliers*, identificação do tipo, e identificação de alarmes falsos para as necessidades no ambiente CEP; uso de JE para localizar observações fora de controle na série temporal; determinar como reagir frente à informação resultante do método JE.

Esse estudo é útil para especialistas que procuram métodos para dados CEP de corridas curtas que estão autocorrelacionados, onde a investigação de um alarme falso ou um tipo de erro é cara. Os autores desenvolveram tabelas com valores críticos, número de observações, localização fora de controle para séries AR(1) e MA(1); identificação do tipo quando a localização é detectada corretamente, para séries AR(1) e MA(1); e resultados para JE e EWMA. Pode acontecer que, poucas observações, como 9 ou 11 sejam suficientes para detectar observações fora de controle quando ela for a última observação. Geralmente, como o número de observações aumenta, é mais difícil para JE detectar observações fora de controle. Igualmente, 10 observações subsequentes à observação fora de controle não são sempre suficientes para identificar o tipo de *outlier*. JE é fácil de implantar e está disponível pela *Scientific Computing Associates*.

2.3.3 O uso de gráficos EWMA com linha central móvel

Suponha-se que um certo processo não é modelado exatamente como um modelo ARIMA. Em geral, se as observações do processo são correlacionadas positivamente e a média do processo não apresenta uma tendência muito rapidamente, o EWMA com valor apropriado para λ pode gerar um excelente preditor um passo à diante. Conseqüentemente, pode-se contar com muitos processos que obedecem a modelos dinâmicos de primeira ordem (que seguem uma “tendência” lenta) para ser bem representados pelo EWMA.

O procedimento consiste em assinalar no gráfico os erros de previsão um passo à frente do EWMA (ou modelo de resíduos) em um gráfico de controle. Esse gráfico deverá ser acompanhado de um gráfico das observações originais no qual a previsão EWMA é sobreposta. A experiência indica que ambos os gráficos são usualmente necessários. O gráfico das observações originais permite que a dinâmica do processo seja visualizada.

Montgomery e Mastrangelo (1991) mostram que é possível combinar informações sobre o estado do controle estatístico e processo dinâmico num gráfico de controle simples. Considere-se $z_t = \lambda x_t + (1 - \lambda)z_{t-1}$ sendo o EWMA. Suponha-se que os erros de previsão e_t um passo à frente do EWMA são normalmente distribuídos. Então os limites de controle 3 sigma do gráfico de controle desses erros satisfazem a seguinte condição de probabilidade:

$$P[-3\sigma \leq e_t \leq 3\sigma] = 0,9973 \quad (2.45)$$

onde σ é o desvio padrão dos erros ou resíduos e_t . Pode-se reescrever isso como

$$P[-3\sigma \leq x_t - \hat{x}_t(t-1) \leq 3\sigma] = 0,9973, \text{ ou} \quad (2.46)$$

$$P[\hat{x}_t(t-1) - 3\sigma \leq x_t \leq \hat{x}_t(t-1) + 3\sigma] = 0,9973 \quad (2.47)$$

Essa equação sugere que se o EWMA é um previsor um passo à frente adequado, então pode-se usar z_t como a linha central de um gráfico de controle para período $t + 1$ com limites de controle superior e inferior

$$UCL_{t+1} = z_t + 3\sigma, \text{ e} \quad (2.48)$$

$$LCL_{t+1} = z_t - 3\sigma \quad (2.49)$$

e a observação x_{t+1} poderá ser comparada com esses limites para testar o controle estatístico. Assim, tem-se o gráfico de controle EWMA com linha central móvel (figura 17).

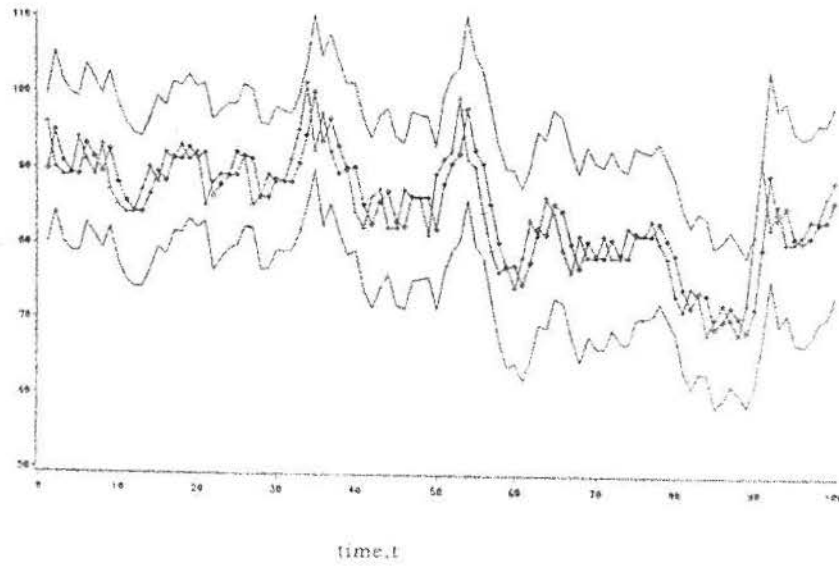


Figura 17. Exemplo de gráfico de controle EWMA com linha central móvel

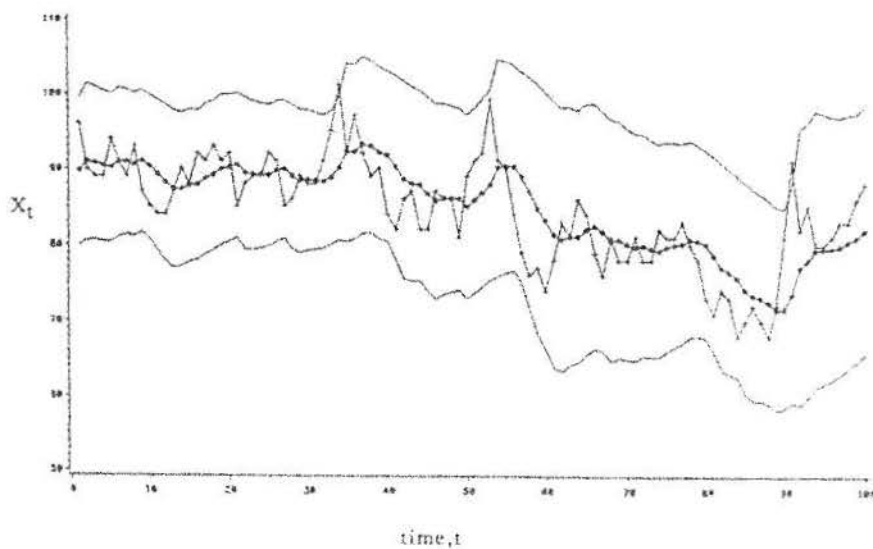


Figura 18. Exemplo de gráfico de controle EWMA com linha central móvel com valor do parâmetro lambda não otimizado

2.3.3.1 *Estimação e monitoramento de σ*

O desvio padrão dos erros de previsão um passo à frente modelo dos resíduos σ pode ser estimado de muitas formas. Se λ for escolhido, como sugerido anteriormente, em um registro de n observações, então dividindo a soma dos quadrados dos erros previstos para o otimizador λ por n produzirá um estimador de σ^2 . Esse é o método usado em muitos programas de computador de análise de séries temporais.

Outra aproximação é dada pelo cálculo do valor estimado de σ tipicamente feito em sistemas de previsão. O desvio absoluto da média (MAD) pode ser usado para isso. O MAD é calculado pela aplicação de um EWMA para o valor absoluto do erro de previsão

$$\Delta_t = \alpha |e_t| + (1 - \alpha)\Delta_{t-1}, \quad (2.50)$$

onde $0 < \alpha \leq 1$.

Já que o MAD de uma distribuição normal está relacionado com o desvio padrão por $\sigma \cong 1,25\Delta$ (ver Montgomery, Johnson, e Gardiner, 1990) podemos estimar o desvio padrão do erro de previsão no tempo t por

$$\hat{\sigma}_t \cong 1,25\Delta_t \quad (2.51)$$

Outra aproximação serve para calcular uma variância suavizada

$$\hat{\sigma}_t^2 = \alpha e_t^2 + (1 - \alpha)\hat{\sigma}_{t-1}^2, \quad (2.52)$$

onde $0 < \alpha \leq 1$.

MacGregor e Harris (1993) discutem o uso da variância móvel exponencialmente ponderada estimada na monitoração da variabilidade do processo. Eles mostram como encontrar os limites de controle para essas quantidades para ambos os processos correlacionados e não-correlacionados.

2.3.3.2 *A sensibilidade do gráfico de controle de resíduos*

Muitos autores têm defendido que os gráficos de controle para resíduos não são sensíveis para mudanças pequenas na média do processo (ver Wandell, Moskowitz, e Plante,

1994). Para melhorar a sensibilidade, Montgomery e Mastrangelo (1991) recomendam o uso de gráficos de controle CUSUM (soma cumulativa) ou EWMA para resíduos em vez dos gráficos de Shewhart, e sugerem o uso de procedimentos complementares chamados de “pista de sinais combinados” com o gráfico de controle para resíduos. Há evidências que nestes procedimentos suplementares aumentem consideravelmente o desempenho dos gráficos de controle de resíduos. Além disso, Mastrangelo e Montgomery (1997) mostram que se um método apropriado de delineamento de “pista de sinais” for combinado com o procedimento base de EWMA como foi descrito, pode ser alcançada um bom desempenho para detectar mudanças no controle.

2.3.3.3 *Alguns outros aspectos do EWMA*

O gráfico de controle EWMA com as modificações precedentes pode ser aplicado em uma grande variedade de situações quando os dados estão autocorrelacionados. Em geral, se as observações estão correlacionadas positivamente até abaixo do número da defasagem e se o processo possui uma tendência moderadamente lenta (ao contrário de oscilações rápidas como o que acontece freqüentemente com os dados autocorrelacionados negativamente), então o método da linha central do EWMA funciona razoavelmente bem. Esse procedimento de média “tendenciosa” ocorre em processos não estacionários, embora alguns processos estacionários possam exibir um procedimento similar. Por exemplo, um AR(1) com $\phi \geq 0,8$ exibirá valores distantes acima e abaixo da média. O procedimento da linha central do EWMA poderia ser visto como uma aproximação de um procedimento baseado em um modelo ARIMA exato.

Montgomery e Mastrangelo (1991) argumentam que em processos de variáveis não controláveis (por exemplo, em processos químicos) o gráfico de linha central do EWMA é útil para o controle de entrada de variáveis controladas, em oposição ao resultado do processo.

O estudo de Montgomery e Mastrangelo (1991) promoveu alguns debates entre personalidades especializadas sobre o assunto. Faltin e Woodall (1991) gostariam de explorar dois pontos: estimação de parâmetros e representação gráfica. Os exemplos usados no estudo são para usos retrospectivos das suas aproximações por EWMA. Na prática, é desejável continuar aplicando o procedimento prospectivamente como têm sido feito com as observações. Quanto à representação gráfica, Faltin e Woodall (1991) discordam do gráfico de controle recomendado com os resíduos e o EWMA exibidos juntamente; primeiro porque

consideram que a plotagem dos dados torna-se complicada; segundo, porque vêem algum mérito na utilização dos gráficos separadamente. Montgomery e Mastrangelo concordam com a argumentação de seus questionadores para situações com suspeitas de *outliers*, e argumentam que um ciclo iterativo de ajustamento de modelos, com identificação e remoção de *outliers*, seguido por um reajustamento é apropriado e similar ao gráfico de controle padrão na prática. Quanto ao gráfico, tem considerável aplicação em certos casos, e é freqüentemente confundido com muitos gráficos de controle.

MacGregor (1991) comenta que os gráficos de controle baseados nos resíduos não são eficazes em todos os tipos de mudanças de processo. Montgomery e Mastrangelo (1991) destacam dois pontos: primeiro, que essa aproximação assume que há um relacionamento de identificação fácil entre a variável resposta (a de interesse) e a variável manipulada; segundo, o uso de EPC que reduz a variabilidade na variável resposta por transferência para a variável manipulada em alguns casos pode aumentar a discrepância de outras características do sistema.

Ryan (1991) criticou o procedimento para a escolha de λ do EWMA, argumentando que pode ter boas propriedades ARL casualmente, já que falta ser demonstrado. Montgomery e Mastrangelo (1991) argumentaram que o procedimento é adequado para a escolha de λ , uma vez que o critério de seleção deve estar baseado na previsão “um passo à frente”.

2.3.4 O uso de EPC (Controle da Engenharia do Processo)

2.3.4.1 Monitoramento do processo e regulagem do processo

Para Montgomery (1997), a redução da variabilidade é uma parte importante no desempenho da implantação de um processo em qualquer indústria. O CEP é uma ferramenta eficaz para a redução da variabilidade pela habilidade dos gráficos de controle em detectar causas especiais. Quando as causas especiais são removidas, a variabilidade do processo é reduzida e o desempenho do processo melhora.

O CEP tem tido uma longa história de sucesso no uso em processos discretos de manufatura. Em processos contínuos, como aqueles encontrados em indústrias de processos e de químicas, outra aproximação é muitas vezes usada para reduzir a variabilidade. Essa aproximação é baseada em compensação e regulagem de processo, o qual alguma variável

manipulável do processo é ajustada com o objetivo de manter o resultado do processo no “alvo” (ou equivalentemente, minimizar a variabilidade do resultado do processo em torno do alvo). Esses processos de compensação ou formas de regulagem são amplamente conhecidos como *controle da engenharia do processo* (EPC), controle estocástico, ou controle *feedback* ou *feedforward*, dependendo da natureza dos ajustes.

O CEP é sempre aplicado em uma situação onde se assume que é possível conduzir o processo a um estado de controle estatístico. Por “controle estatístico”, entende-se que é somente a observação da variação aleatória estável em torno do alvo, isto é, da média do processo. Agora, ao eliminarem-se as causas especiais como diferenças adequadas para operadores e as variações na matéria prima, muitas vezes é possível obter esse estado de controle. Entretanto, o processo pode apresentar uma leve tendência num determinado sentido ou oscilar muito distante do valor alvo. Isso pode ocorrer por causa de fenômenos como a variação contínua do mecanismo de entrada de materiais, efeitos de temperatura, etc, ou pode ser devido inteiramente a forças desconhecidas que interferem no processo. A regulação do processo EPC pressupõe que exista outra variável que pode ser ajustada para compensar essa tendência no resultado do processo, e que uma série de ajustes regulares para essa variável manipulável direcione o resultado do processo para o objetivo desejável.

É de considerável interesse combinar ou integrar essas estratégias em um esforço para produzir um procedimento melhorado, isto é, uma intensificação para que o EPC seja capaz de detectar os tipos de perturbações das causas especiais. Um conhecimento mais aprofundado pode ser encontrado em Box e Kramer (1992); MacGregor (1987); Vander Weil, Tucker, Faltin, e Doganaksoy (1992); MacGregor e Harris (1990); e Montgomery, Keats, Runger, e Messina (1994).

É natural questionar a necessidade para a integração entre o EPC e o CEP. Historicamente, essas técnicas têm se desenvolvido em diferentes ambientes. CEP é muitas vezes parte de estratégias “empurradas” de organizações para melhorar a qualidade, e isso é usualmente conduzido por um gerenciamento *top-down*, gerenciamento com ênfase nas pessoas, nos métodos e nos procedimentos. EPC, por outro lado, é mais tático por natureza, com suas origens na organização da engenharia de processo, e seu foco primário é no processo. A estrutura estatística do CEP é similar à do teste de hipótese, enquanto que a estrutura estatística do EPC é a estimação do parâmetro, isto é, a estimação da quantidade de

perturbação que há no sistema do processo fora do alvo, e então fazer um ajuste para cancelar esse efeito. O que esses dois procedimentos compartilham é um objetivo comum: reduzir a variabilidade. EPC pressupõe que há um modelo dinâmico específico que liga a entrada e a saída do processo. Se esse modelo for correto, então o processo EPC ajustado regularmente minimizará a variação em torno do objetivo. Entretanto, quando ocorrem certos tipos de perturbações externas ou de causas especiais que estão fora da estrutura desse modelo dinâmico, então as medidas de compensação não serão completamente assimiladas por causa dessa queda. Como um resultado, variabilidade será acrescentada. Por aplicar-se CEP em um modo específico, essas causas especiais podem ser detectadas e o procedimento EPC/CEP combinado será mais eficaz do que o uso isolado do EPC.

Para motivar essa discussão, Montgomery (1997) apresenta um exemplo simples de um processo onde um ajuste compensatório direto do EPC é apropriado. O sistema desse exemplo também foi discutido por Box (1991-1992). Suponhamos que o resultado do processo no tempo t é y_t e que isso seja necessário para y_t produza um alvo T . Esse processo possui uma variável manipulável x , e uma mudança em x produzirá todo o seu efeito em y em um período, isto é

$$y_{t+1} - T = gx_t \quad (2.53)$$

onde g é uma constante usualmente chamada de “ganho do processo”. O ganho é como um coeficiente de regressão, no qual ele relata a magnitude de uma mudança em x_t para uma mudança em y_t . Se não for feito nenhum ajuste, o processo tende a sair do alvo, de acordo com

$$y_{t+1} - T = N_{t+1} \quad (2.54)$$

onde N_{t+1} é uma perturbação. A perturbação nessa equação é usualmente representada por um modelo de séries temporais, geralmente um modelo ARIMA. Suponha-se que a perturbação possa ser estimada adequadamente usando um EWMA:

$$\hat{N}_{t+1} = \hat{N}_t + \lambda(N_t - \hat{N}_t) = \hat{N}_t + \lambda e_t \quad (2.55)$$

onde $e_t = N_t - \hat{N}_t$ é o erro de previsão no período de tempo t e $0 < \lambda \leq 1$ é o fator de ponderação para o EWMA. Essa suposição é equivalente a assumir que o processo não correlacionado está tendendo de acordo com o ARIMA da equação 2.27 com parâmetro

$\theta = 1 - \lambda$. No tempo t , o processo ajustado é

$$y_{t+1} - T = N_{t+1} + gx_t. \quad (2.56)$$

A equação 2.56 diz que no tempo $t+1$ o desvio do resultado do alvo dependerá da perturbação no período t , ou apenas do ponto no período t . Obviamente pode-se acertar x_t de modo a tirar fora a perturbação. Entretanto, não se pode fazer isso, porque N_{t+1} é desconhecido no período t . Por outro lado, pode-se estimar N_{t+1} por \hat{N}_{t+1} usando a equação (do N_{t+1}). Então, obtém-se

$$y_{t+1} - T = e_{t+1} + \hat{N}_{t+1} + gx_t \quad (2.57)$$

desde que $e_{t+1} = \hat{N}_{t+1} - N_{t+1}$.

Da equação 2.57, fica claro que se for feito $gx_t = -\hat{N}_{t+1}$ ou o ponto $x_t = -(1/g)\hat{N}_{t+1}$ então o ajuste cancelaria a perturbação, e no período $t+1$ desvio do resultado do alvo seria $y_{t+1} - T = e_{t+1}$ onde e_{t+1} é o erro previsto no período t , que é $e_{t+1} = N_{t+1} + \hat{N}_{t+1}$. O ajuste atual para a variável manipulável feita no tempo t é

$$x_t - x_{t-1} = -\frac{1}{g}(\hat{N}_{t+1} - \hat{N}_t). \quad (2.58)$$

Já que $\hat{N}_{t+1} + \hat{N}_t = \lambda(N_t - \hat{N}_t) = \lambda e_t$ e $e_t = y_t - T$, então o ajuste na equação 2.58 pode ser escrito como

$$x_t - x_{t-1} = -\frac{\lambda}{g}(y_t - T) = -\frac{\lambda}{g}e_t. \quad (2.59)$$

O ponto atual para a variável manipulada no final do período t é simplesmente a soma de todos os ajustes no tempo t , ou

$$x_t = \sum_{j=1}^t (x_j - x_{j-1}) = -\frac{\lambda}{g} \sum_{j=1}^t e_j \quad (2.60)$$

Esse tipo de método de regulagem do processo é chamado de *controle integral*. É um método puro de controle *feedback* que acerta o nível da variável manipulável igual à soma

ponderada de todos os processos anteriores e atual desviados do alvo. Pode ser mostrado que se uma parte determinada do modelo do processo (2.55) for correta, e se a perturbação N_t for perfeitamente prevista, fora o erro aleatório do EWMA, então ela é um controle ótimo no sentido que minimiza o erro do quadrado da média dos resultados dos processos desviados do alvo T .

2.3.4.2 *Combinando CEP e EPC*

Há uma confusão considerável sobre ajustes ou regulagens de processos, e o papel importante que isso tem na redução da variabilidade. Por exemplo, o gráfico de controle não é sempre o melhor método para a redução da variabilidade em torno da média. Em processos industriais e químicos, técnicas como o papel do “controle integral simples” ilustrado anteriormente tem sido efetivamente usado para esse propósito. Em geral, a teoria do controle da engenharia é baseada na idéia de que se for possível (1) prever a próxima observação no processo, (2) ter algumas outras variáveis que se possa manipular em ordem para influenciar a saída do processo e (3) conhecer o efeito dessa variável manipulada, assim pode-se determinar o quanto aplicar dessa ação de controle, então pode-se fazer o ajuste na variável manipulada no tempo t que é o mais provável para produzir um valor no alvo do resultado do processo no período $t+1$. Claramente, isso requer um bom conhecimento do relacionamento entre o resultado ou a variável controlada e a variável manipulada assim como um entendimento de “processos dinâmicos”. Deve-se também haver capacidade em facilmente mudar a variável manipulada. De fato, se o custo de usar a “ação do controle” for omitido, então a variável na saída do processo é minimizada pelo uso da “ação do controle” em todo o período. Note que isso é um contraste severo com o CEP, onde a “ação do controle” ou um ajuste de processo é usado apenas quando há uma evidência estatística de que o processo está fora de controle. Essa evidência estatística é usualmente um ponto fora dos limites de um gráfico de controle.

Há muitos processos onde alguns tipos de métodos de controle *feedback* poderão ser preferíveis a um gráfico de controle. Por exemplo, considere-se o processo dirigir um carro, com o objetivo de mantê-lo no centro de uma travessa de “mão direita” (ou equivalentemente, minimizar a variação em torno do centro de uma travessa de “mão direita”). O motorista pode facilmente ver a estrada a frente, e os ajustes do processo (correções da “posição da direção”) podem ser feitos em algum tempo em um custo omitido. Conseqüentemente, se o motorista

conhece o relacionamento entre a variável resposta (posição do carro) e a variável manipulada (ajuste da direção), ele possivelmente preferiria em usar um método de controle *feedback* para controlar a posição do carro, melhor do que um gráfico de controle estatístico. “Dirigir um carro com um gráfico de controle Shewhart pode ser uma idéia interessante, mas não se quer estar no carro durante o experimento”.

Por outro lado, o EPC não se esforça em identificar uma causa especial que pode “causar impacto” no processo. A eliminação de causas especiais pode resultar em uma melhoria significativa no processo. O que todos os esquemas EPC fazem é reagir a uma perturbação no processo, eles não fazem nenhum esforço para remover as causas especiais. Conseqüentemente, em processos onde o controle *feedback* é usado, pode haver uma melhoria substancial nos gráficos de controle que são usados também para o monitoramento do processo estatístico (em oposição à “controle”, as “ações de controle” são baseadas em esquemas de engenharia). Alguns autores referem-se à sistemas onde ambos sistemas EPC e CEP para o monitoramento do processo tem sido implementados como algoritmos de CEP. Veja Vander Weil et al. (1992).

O gráfico de controle deve ser aplicado para cada controle do erro (a diferença entre a variável controlada e o alvo) ou para a seqüência de ajustes para a variável manipulada. Pontos que ficam caem fora dos limites de controle nesses gráficos identificam períodos onde o controle dos erros é grande ou onde grandes mudanças para a variável manipulada estão sendo feitas. Esses períodos provavelmente seriam boas oportunidades para procurar causas especiais. Montgomery et al. (1994) têm demonstrado a eficácia de um sistema assim. A figura 19 ilustra como uma combinação de EPC e CEP pode ser empregada.

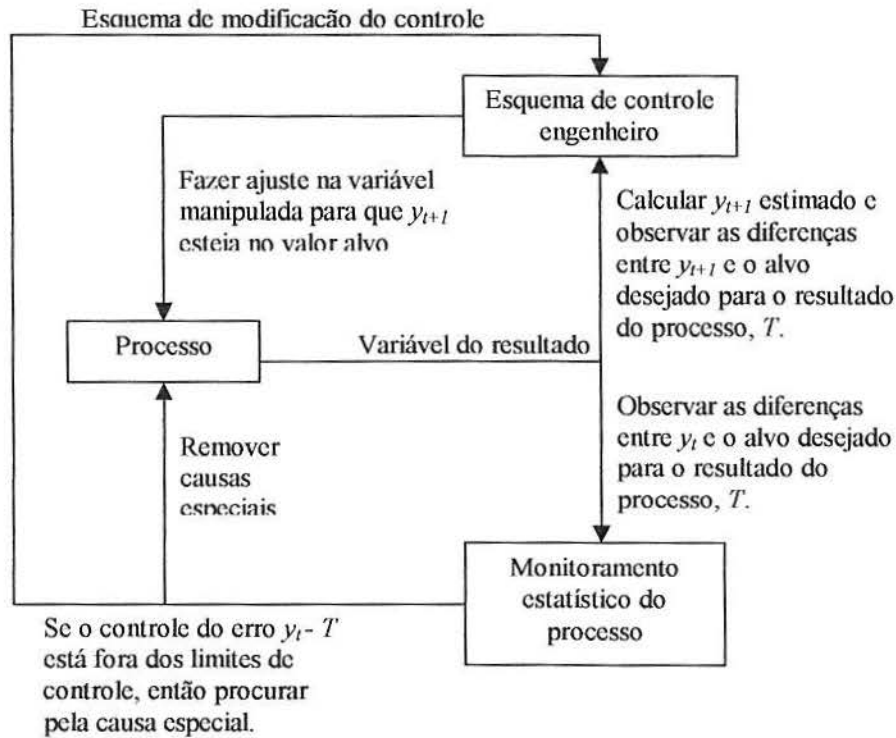


Figura 19. Controle da engenharia do processo e monitoramento estatístico do processo

2.3.4.3 Comparando CEP e EPC

Conforme o estudo de Box, Coleman, Baxley (1997), durante uma implementação bem sucedida de gráfico de controle, um especialista tomará três importantes decisões: (1) se o gráfico de controle é a ferramenta apropriada para essa aplicação; (2) qual o tipo de gráfico de controle que deve ser usado; (3) onde os limites de controle deverão ser colocados?

A primeira questão depende se existe ou não uma distribuição de referência, ou colocar outro caminho se há ou não períodos estáveis na média ou variância do processo. De acordo com Alwan e Roberts (1995), “um estudo de 235 aplicações de controle de qualidade sugere que as violações das suposições são a ordem (85% observados) melhor do que a exceção na prática ... conduz a: (a) uma garantia falsa de que o processo está estável, (b) uma falsa procura pelas causas especiais, (c) uma deficiência em procurar por causas especiais, as quais podem ser vistas com análises melhores, (d) deficiência em observar e agir na variação sistemática, como tendência, periodicidade, e variação auto-regressiva, e (e) gráficos de controle sendo ignorados.” Uma aproximação eficaz para esses casos é fazer um gráfico para

as observações individuais, subgrupos de médias, ou variâncias, e tornar visível para o responsável pelo processo. Se há uma variância estável, mas a média do processo exibir um comportamento tendencioso, então a estratégia EPC poderá ser considerada como uma medida prática para a redução da variabilidade.

De acordo com Calcut (1995), as questões 2 e 3 dependem primeiro em como os gráficos serão usados: (1) tempo real de monitoramento do processo, (2) resolução de problemas (ex post facto), ou (3) taxaço da estabilidade do processo.

Falhas para distinguir entre teste de hipóteses e estimação são freqüentemente problemas causados em estatísticas. Do mesmo modo isso é necessário para distinguir claramente entre a detecção de sinais fora de controle por monitoramento do processo e a estimação do nível de perturbações correntes necessário para compensação por controle *feedback*. Ambos devem ser usados mas não confundidos.

Filosofia e metodologia apropriadas para essa situação e instrução de ajuste como monitoração no contexto CEP não tem tido opinião suficiente. Deveria ser mais extensamente conhecido na comunidade CEP que as idéias obtidas do EPC não precisam ser misteriosas ou requerer matemática avançada. Em particular, os esquemas de controle PI (*proportional integral*) podem ser facilmente compreendidos e aplicados de gráficos simples apropriados. Como o controle é extremamente rigoroso, conseqüentemente a estimação exata do parâmetro é raramente necessária, e esse procedimento simples pode repor mais adiante um procedimento “ótimo forçado” mais complicado.

Quando há ajustes e custo amostral, um ajuste *bounded feedback* é necessário. Apesar de que os métodos *bounded* têm a aparência de gráficos de monitoramento EWMA, sua base lógica é completamente diferente. Em particular, a posição das linhas de limites é baseada em custos e não em probabilidades do excedente.

2.3.5 *Uso de redes neurais (RN)*

Balestrassi (2000) desenvolveu sistema híbrido de reconhecimento de padrões baseado em modelagem de séries temporais e no uso de redes neurais artificiais para os dados de uma carta de CEP sob o regime de tempo real. Algumas características das redes neurais, tais como

as citadas abaixo, ajudam na tarefa de reconhecimento de padrão de CEP:

- Podem aprender de exemplos e adaptar-se a novas situações;
- São robustas (podem abstrair informações de dados com ruído, incompletos ou não precisos);
- Podem generalizar a partir de exemplos (podem prover soluções corretas para dados similares, não vistos no treinamento);
- Podem construir soluções rapidamente sem muito domínio do conhecimento do processo;
- Podem aproximar qualquer função multivariada complexa (mal definida ou desconhecida) e formar uma decisão de classificação a partir do reconhecimento de padrões discriminantes;
- São computacionalmente eficientes (tem a velocidade para operar em tempo real).

Para se entender melhor o uso de redes neurais para monitorar processos autocorrelacionados é preciso se ter um bom entendimento sobre redes neurais. Uma estrutura comum em basicamente todas as RNs é o neurônio.

Foram também observadas considerações com respeito ao aprendizado, a topologia e a classificação de padrões das redes neurais. Existem essencialmente três fatores que foram considerados para a seleção da rede neural. O primeiro é como os neurônios processam as informações, o qual enfoca a determinação do somatório e da função de ativação. O segundo é como os pesos das conexões são modificados, que é determinado pelo tipo de paradigma de treinamento escolhido. O terceiro fator consiste na determinação da maneira com que os neurônios da rede são conectados, definindo a arquitetura ou topologia da rede. Em muitas circunstâncias esses fatores são influenciados pela definição do problema em questão.

Chiu, e Yang (2001) estudaram métodos de identificação de mudanças em processos correlacionados usando rede neural. No processo de construção da topologia da rede neural um procedimento de seleção de todas as regressões possíveis foi aplicado para encontrar o número apropriado de comandos de entrada para usar em um modelo de rede neural. Uma

melhoria significativa feita pela aproximação por redes neurais dos métodos atuais na capacidade disponível para identificar mudanças em processos de manufatura é apresentada. Os modelos atuais (gráficos de controle de séries temporais) podem classificar corretamente apenas o dado mudado no 9º e 7º percentuais nos casos de mudança de um ou dois desvios padrões com coeficiente de correlação 0,9. Modelos de redes neurais podem identificar, com sucesso, mudanças de 1,5 e 2 sigmas para coeficientes de autocorrelação significativos, entretanto as redes não podem identificar a mudança de 1 sigma para coeficiente de correlação 0,9. Melhorias futuras são necessárias para permitir as redes a identificarem mudanças de valores pequenos.

Chou, Chang e Chen (2002) desenvolveram um método baseado no modelo de custo apresentado por Duncan (1956) que determina a escolha de três parâmetros: o tamanho da amostra, o intervalo amostral entre os sucessivos subgrupos, e os limites de controle para o gráfico; aprimorado posteriormente por Alexander, Dillman, Usher e Damodaran (1995) com a combinação entre o modelo de custo de Duncan e a função perda de Taguchi para a escolha desses parâmetros. O método apresentado por Chou, Chang e Chen (2002) adapta esse modelo de custo para a situação onde existe autocorrelação nos dados.

2.4 *Discussão do Capítulo*

Esse capítulo mostrou algumas técnicas que podem ser utilizadas no tratamento de dados autocorrelacionados. A ênfase maior foi dada à aplicação de um modelo ARIMA seguido de um ajuste por um gráfico de controle, devido à grande quantidade de material encontrado.

O autor da monografia, com o intuito de elucidar o uso dos gráficos de CEP em dados autocorrelacionados, organizou e resumiu os métodos e gráficos discutidos nesse trabalho e propôs um modelo referencial de análise representado na figura 20.

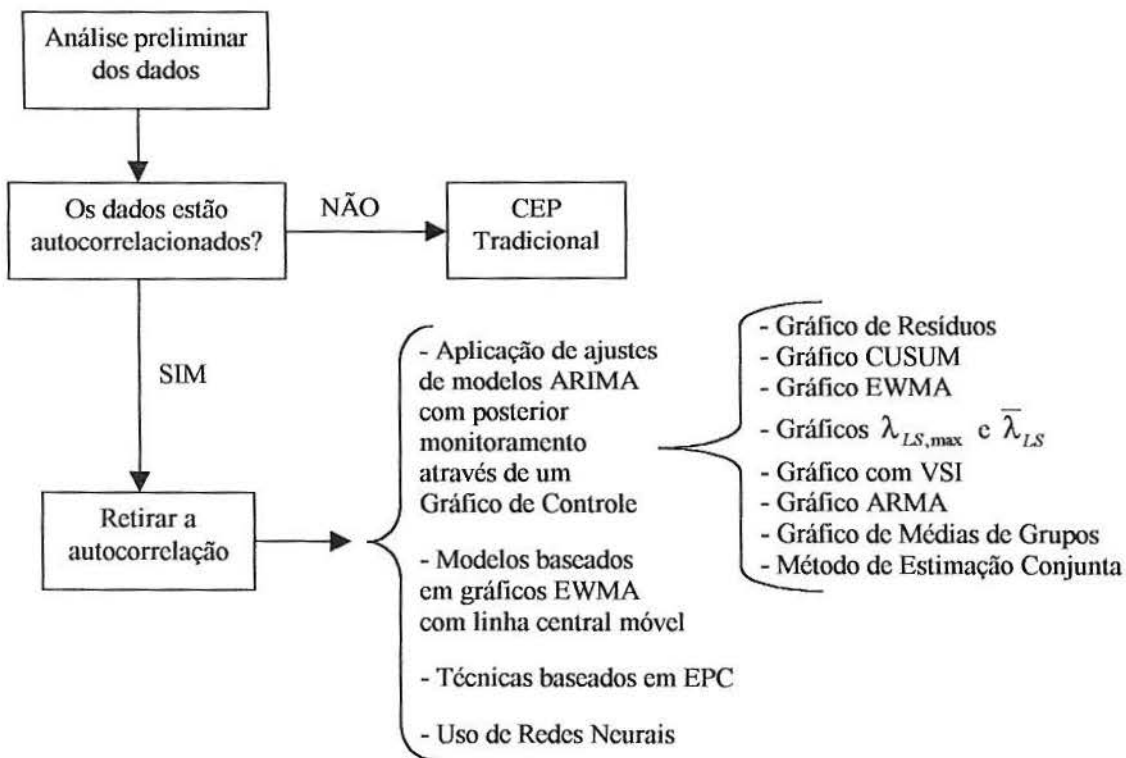


Figura 20. Esquema proposto para o tratamento dos dados autocorrelacionados.

3 ESTUDO DE CASO

3.1 Descrição do experimento

O experimento utilizado para ilustrar esse trabalho consiste em analisar o desempenho de uma equipe de funcionários de uma empresa de energia elétrica. Esse desempenho é medido por uma unidade de desempenho chamada de Unidade de Referência, UR. A UR é uma quantidade que depende da atividade realizada pela equipe, quanto ao tempo (duração) e quanto a complexidade da atividade. Quanto mais complexa ou demorada for a atividade, maior o valor da UR.

O objetivo da análise é acompanhar, através de um gráfico de controle, o desempenho de uma equipe nova no início do processo após três meses de treinamento, ao longo do tempo com base nos valores da UR. Considerando que o tempo de trabalho não é o mesmo para todos os dias, optou-se por monitorar a taxa UR/hora. Essa equipe, para controle do experimento, é composta de empregados novos, e serão observados seus respectivos desempenhos ao longo de 4 meses.

Os dados foram obtidos nos dias em que houve atividades para as equipes. Cada observação corresponde a um dia de atividade, ordenada, portanto, em uma série temporal. O primeiro dado foi coletado no dia 5 de agosto de 2002, uma segunda-feira; e o último dado no dia 21 de dezembro de 2002, um sábado; totalizando 85 observações. Entretanto, não foram coletados dados todos os dias.

Os dados foram analisados nos softwares *excel* e *SPSS*, e os resultados são apresentados nos próximos itens desse capítulo.

3.2 Análise Preliminar dos Dados

Para começar a análise, primeiramente foi feito um gráfico de série temporal dos dados para observar qual o comportamento da variável UR/hora ao longo do tempo. A figura 21 mostra esse comportamento.

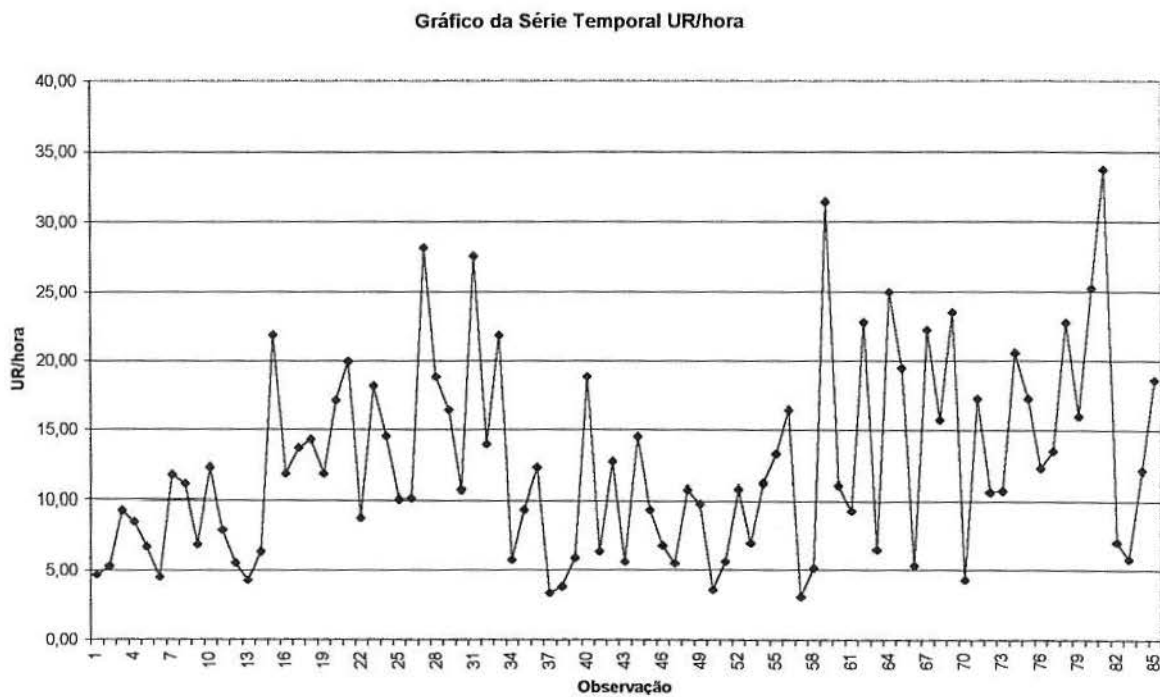


Figura 21. Série Temporal da variável UR/hora.

Em uma análise visual preliminar, percebe-se uma tendência na média das primeiras observações, uma mudança na média nas observações intermediárias, e uma variabilidade maior nos dados finais. A série ainda apresenta alguns picos isolados.

Na seqüência foi construído um histograma dos dados para verificar qual o comportamento da variável quanto a média, variabilidade e simetria. O histograma, apresentado na figura 22, apresenta também uma curva normal ajustada com base na média e no desvio padrão dos dados.

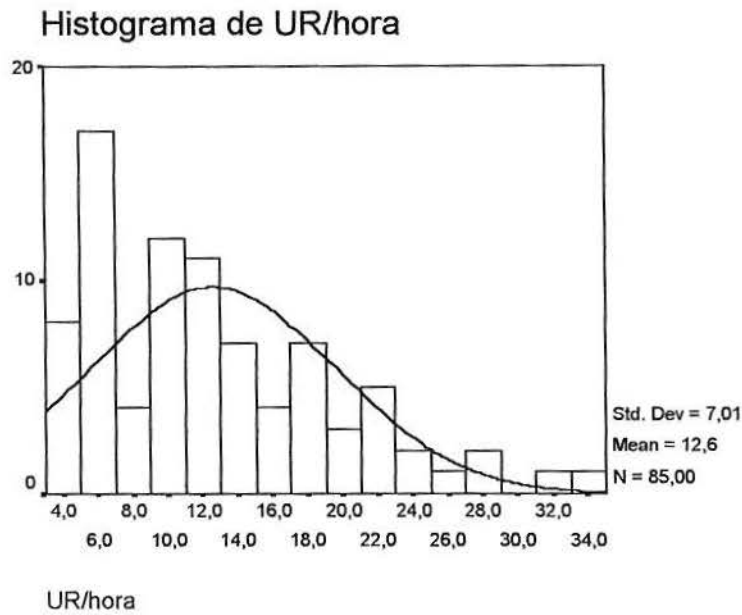


Figura 22. Histograma da variável UR/hora.

O histograma sugere que a variável não tem um comportamento simétrico, portanto não parece ter distribuição normal. O Gráfico de Box-Plot da figura 23 mostra que a distribuição é assimétrica à esquerda com uma concentração dos valores próximos do limite inferior.

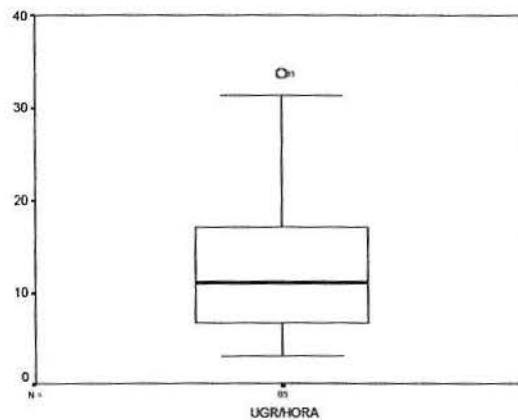


Figura 23. Gráfico de Box Plot para a variável UR/hora

A seguir foi realizado um teste de normalidade nos dados, chamado de teste de Kolmogorov-Smirnov. O resultado do teste está na tabela 2.

Tests of Normality			
Kolmogorov-Smirnov ^a			
	Statistic	df	Sig.
UR/HORA	,117	85	,006

a. Lilliefors Significance Correction

Tabela 2. Teste de Normalidade (saida do Software SPSS).

O teste revela que os dados não seguem uma distribuição normal, pois o valor da significância do teste é de 0,006, bem menor que 0,15, que é o valor de significância mínimo para que os dados possam ser considerados normais nesse teste. O gráfico de plotagem de probabilidade normal, apresentado na figura 24, evidencia graficamente a não-normalidade.

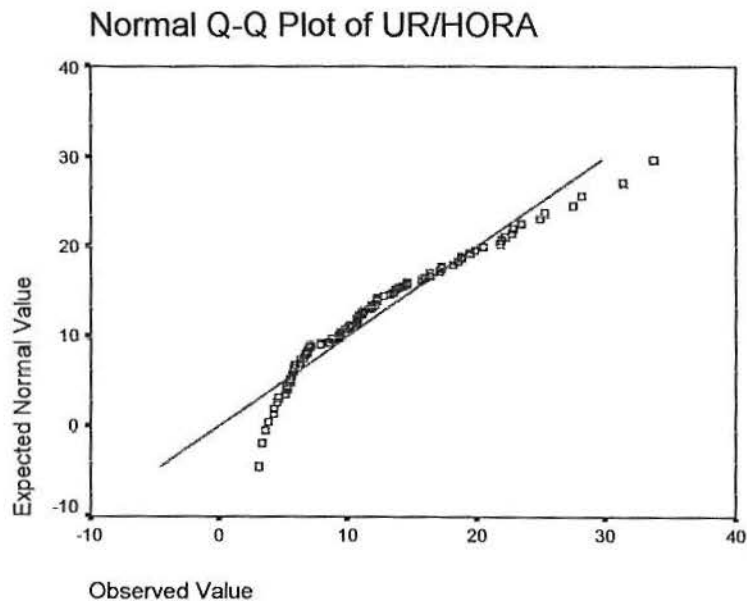


Figura 24. Gráfico da Plotagem de Probabilidade Normal

O objetivo desse estudo, como mencionado anteriormente, é monitorar a variável UR/hora através de um gráfico de controle. Na seqüência desse estudo será adotada uma combinação entre a metodologia de Box-Jenkins esboçada na figura 10 e esquema proposto pelo autor apresentado na figura 20 do item 2.4.

3.3 Detecção da Autocorrelação

Inicialmente, organizam-se os dados. A figura 25 apresenta um gráfico de controle tradicional de Shewhart para as observações da variável UR/hora. Como suspeitado anteriormente, os dados parecem ter uma média não constante ao longo do tempo, o que sugere que o processo não é estacionário.

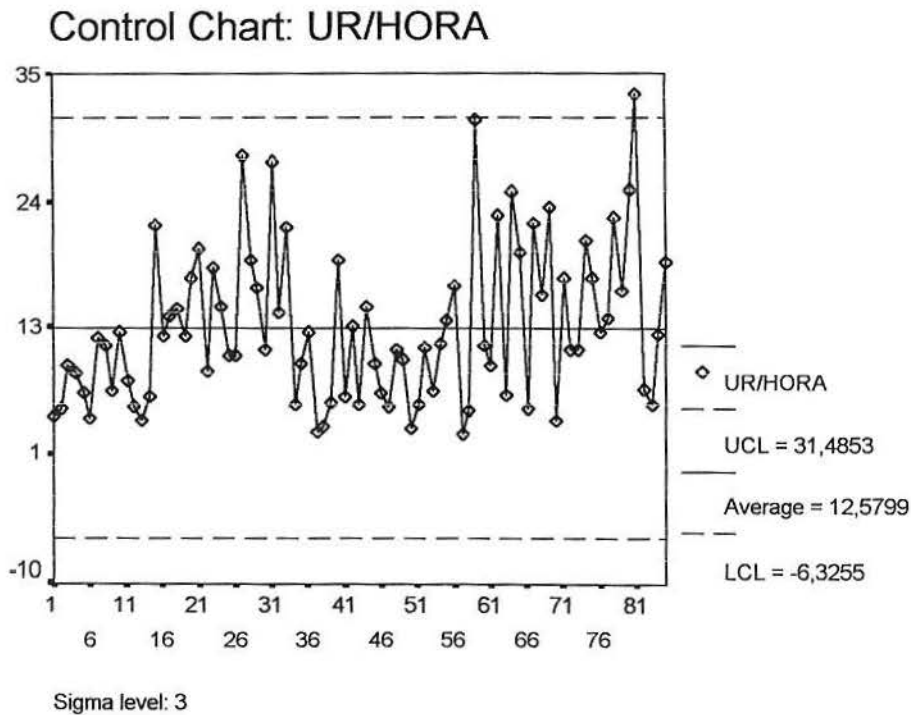


Figura 25. Gráfico de controle para as observações originais.

Para verificar a questão da estacionariedade, foi construído o gráfico de correlograma da FAC (figura 27). A seguir (figura 26) são apresentados os valores das autocorrelações em cada defasagem k (*lag*). Conforme visto nesse trabalho, o valor do coeficiente de autocorrelação para qualquer defasagem maior que zero é zero, em um processo estacionário. Como tem-se valores que excedem o intervalo de confiança, o processo não é estacionário. As estatísticas de Box-Ljung evidenciam esse argumento.

Autocorrelations: URH UR/HORA

Lag	Auto-Corr.	Stand. Err.	-1	-.75	-.5	-.25	0	.25	.5	.75	1	Box-Ljung	Prob.
1	,103	,107					I**					,940	,332
2	,107	,106					I**					1,966	,374
3	,228	,105					I***.*					6,639	,084
4	,199	,105					I****					10,238	,037
5	,165	,104					I****					12,740	,026
6	,097	,103					I**					13,619	,034
7	,095	,103					I**					14,473	,043
8	,017	,102					*					14,499	,070
9	,058	,101					I*					14,822	,096
10	,060	,101					I*					15,177	,126
11	-,037	,100					*I					15,314	,169
12	,122	,099					I**					16,819	,157
13	,093	,099					I**					17,716	,169
14	-,141	,098					***I					19,784	,137
15	-,104	,097					**I					20,933	,139
16	,032	,097					I*					21,041	,177
17	-,100	,096					**I					22,118	,180
18	-,229	,095					*.***I					27,916	,063
19	-,067	,094					*I					28,420	,076
20	-,133	,094					***I					30,432	,063
21	-,119	,093					**I					32,067	,058
22	-,121	,092					**I					33,784	,052
23	-,160	,092					***I					36,825	,034
24	-,167	,091					***I					40,225	,020
25	-,043	,090					*I					40,450	,026
26	-,128	,089					***I					42,516	,022
27	-,148	,089					***I					45,304	,015
28	-,096	,088					**I					46,503	,015
29	-,022	,087					*					46,568	,021
30	-,232	,086					**.***I					53,775	,005
31	-,024	,085					*					53,852	,007
32	,054	,085					I*					54,255	,008

Plot Symbols: Autocorrelations * Two Standard Error Limits .

Total cases: 85 Computable first lags: 84

Figura 26. Correlograma da Série UR/hora

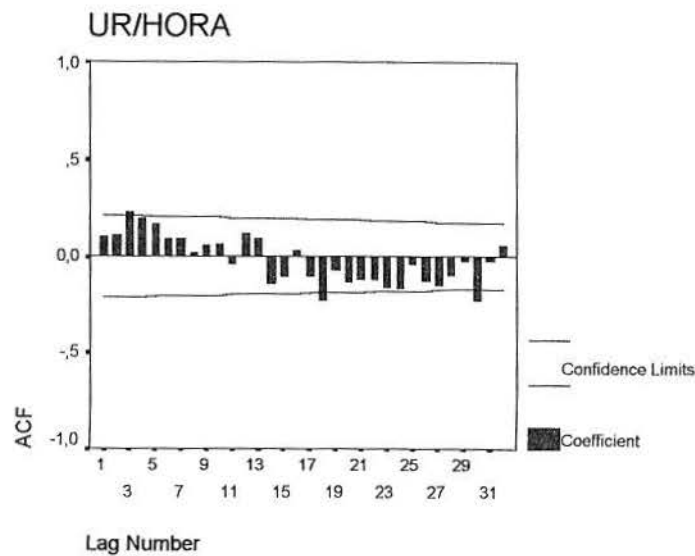


Figura 27. Correlograma da FAC dos dados originais.

Com a evidência de não estacionariedade, a série foi diferenciada uma vez. Os correlogramas das FAC e FACP das observações diferenciadas uma vez são apresentados, respectivamente, nas figuras 28 e 29.

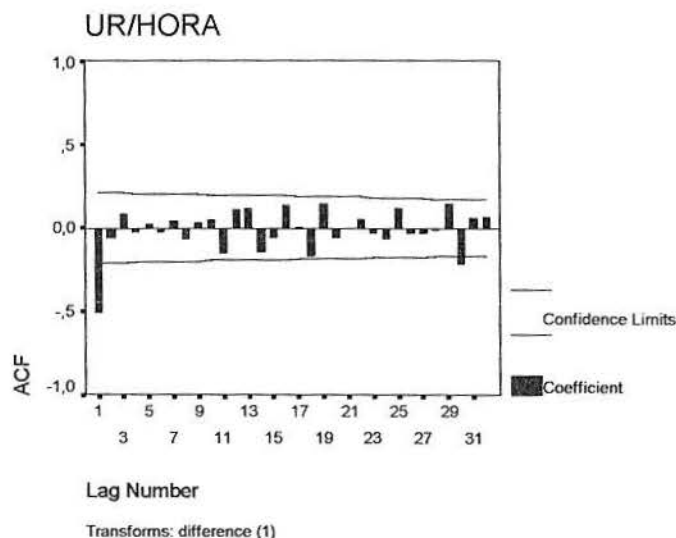


Figura 28. Correlograma da FAC da série diferenciada um vez.

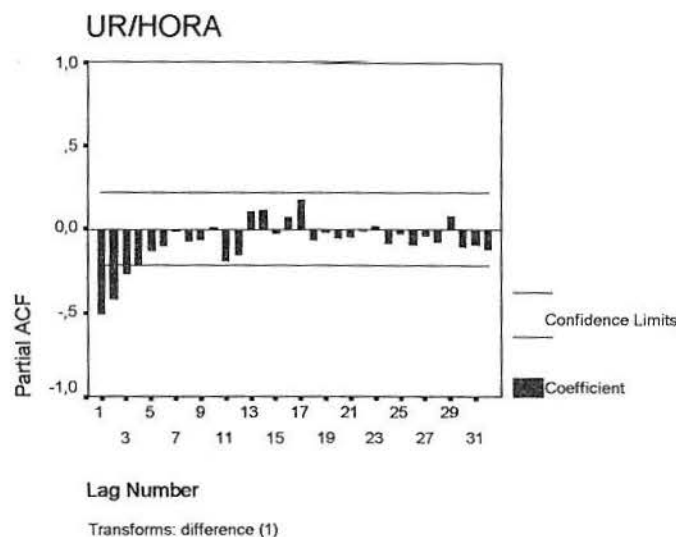


Figura 29. Correlograma da FACP da série diferenciada uma vez.

Os aspectos desses correlogramas se identificam com uma série MA(1), conforme foi visto na figura 7 do capítulo 2. Assim, a série diferenciada parece ser uma série MA(1), e, conseqüentemente, a série das observações originais deve ser uma série ARIMA(0,1,1), por causa da defasagem.

3.4 Ajuste por um Modelo ARIMA

Para verificar o modelo sugerido, foi, então, estimado um modelo ARIMA(0,1,1) nos dados originais. O coeficiente encontrado para o termo MA foi 0,86, conforme pode-se ver na tabela 3. O modelo não possui termo constante, por ele não ter sido considerado estatisticamente significativo.

	B	SEB	T-RATIO	APPROX. PROB.
MA1	,84504198	,06051691	13,963733	,0000000

Tabela 3. Coeficiente do termo MA do modelo estimado

Os valores das FAC e FACP dos resíduos mostrados nas figuras 30 e 31 respectivamente, são bem menores que os valores das observações originais, caracterizando uma estacionariedade na série.

Variable: ERR_1		Missing cases: 1		Valid cases: 84									
Autocorrelations:		ERR_1		Error for URH from ARIMA									
Auto- Stand.		ERR_9		Error for URH from ARIMA, MOD_6 NOCON									
Lag	Corr.	Err.	-1	-.75	-.5	-.25	0	.25	.5	.75	1	Box-Ljung	Prob.
1	-,097	,107					. **I					,826	,363
2	-,075	,107					. *I					1,319	,517
3	,075	,106					. I**					1,823	,610
4	,044	,105					. I*					1,995	,737
5	,044	,105					. I*					2,170	,825
6	,000	,104					. *					2,170	,903
7	,011	,103					. *					2,181	,949
8	-,064	,103					. *I					2,575	,958
9	,006	,102					. *					2,579	,979
10	,014	,101					. *					2,597	,989
11	-,085	,101					. **I					3,314	,986
12	,148	,100					. I***					5,509	,939
13	,141	,099					. I***					7,544	,872
14	-,122	,098					. **I					9,086	,826
15	-,081	,098					. **I					9,779	,833
16	,091	,097					. I**					10,669	,829
17	-,048	,096					. *I					10,921	,861
18	-,199	,096					. ****I					15,257	,644
19	,025	,095					. *					15,325	,702
20	-,068	,094					. *I					15,841	,726
21	-,055	,093					. *I					16,189	,759
22	-,030	,093					. *I					16,295	,801
23	-,096	,092					. **I					17,389	,790
24	-,094	,091					. **I					18,451	,781
25	,065	,090					. I*					18,961	,799
26	-,035	,090					. *I					19,114	,832
27	-,075	,089					. *I					19,823	,838
28	-,033	,088					. *I					19,966	,866
29	,035	,087					. I*					20,129	,889
30	-,218	,086					. **I					26,475	,651
31	-,004	,086					. *					26,477	,698
32	,078	,085					. I**					27,327	,702

Plot Symbols: Autocorrelations * Two Standard Error Limits .
 Total cases: 85 Computable first lags: 83

Figura 30. Valores da FAC e estatísticas Box-Ljung do modelo estimado

Partial Autocorrelations: ERR_1 Error for URH from ARIMA			Pr-Aut- Stand.								
Lag	Corr.	Err.	-1	-.75	-.5	-.25	0	.25	.5	.75	1
1	-,097	,109					. **I				
2	-,085	,109					. **I				
3	,060	,109					. I*				
4	,052	,109					. I*				
5	,065	,109					. I*				
6	,014	,109					. *				
7	,013	,109					. *				
8	-,073	,109					. *I				
9	-,014	,109					. *				
10	-,002	,109					. *				
11	-,078	,109					. **I				
12	,144	,109					. I***				
13	,177	,109					. I*****				
14	-,059	,109					. *I				
15	-,103	,109					. **I				
16	,025	,109					. *				
17	-,077	,109					. **I				
18	-,224	,109					. ****I				
19	-,028	,109					. *I				
20	-,057	,109					. *I				
21	-,019	,109					. *				
22	-,030	,109					. *I				
23	-,070	,109					. *I				
24	-,119	,109					. **I				
25	-,024	,109					. *				
26	-,073	,109					. *I				
27	-,015	,109					. *				
28	-,049	,109					. *I				
29	-,005	,109					. *				
30	-,171	,109					. ****I				
31	-,007	,109					. *				
32	,009	,109					. *				

Plot Symbols: Autocorrelations * Two Standard Error Limits .
Total cases: 85 Computable first lags: 83

Figura 31. Valores da FACP do modelo estimado

Os correlogramas das FAC e FACP dos resíduos (ou erros de previsão) do modelo ARIMA(0,1,1) ajustado são apresentados nas figuras 32 e 33, respectivamente.

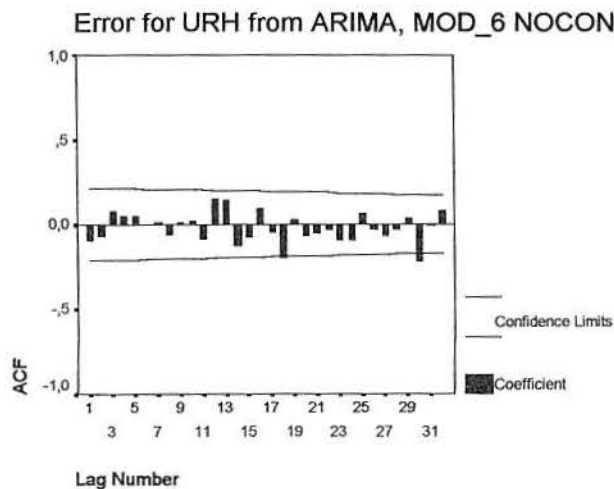


Figura 32. Correlograma da FAC dos resíduos de modelo ARIMA(0,1,1).

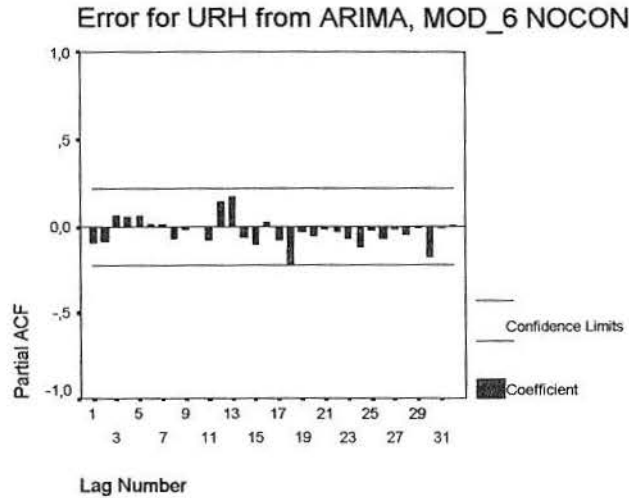


Figura 33. Correlograma da FACP dos resíduos do modelo ARIMA(0,1,1).

Os correlogramas parecem não apresentarem autocorrelação, exceto em algumas observações isoladas (18 e 30 na FAC, e 18 na FACP), provavelmente devido a pouca quantidade de observações, pois dificilmente haveria uma correlação entre as observações X_t e X_{t+18} ; e X_t e X_{t+30} nesse caso. Para verificar o comportamento dos resíduos, foram analisados o histograma (figura 34), o gráfico de probabilidade normal (figura 35) e a normalidade dos dados através do teste de Komogorov-Smirnov (tabela 4).

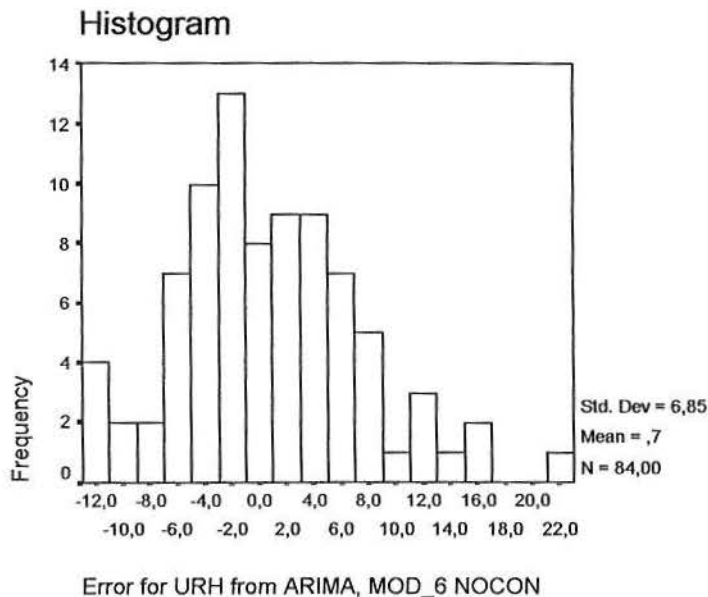


Figura 34. Histograma dos resíduos do modelo ARIMA(0,1,1) ajustado.

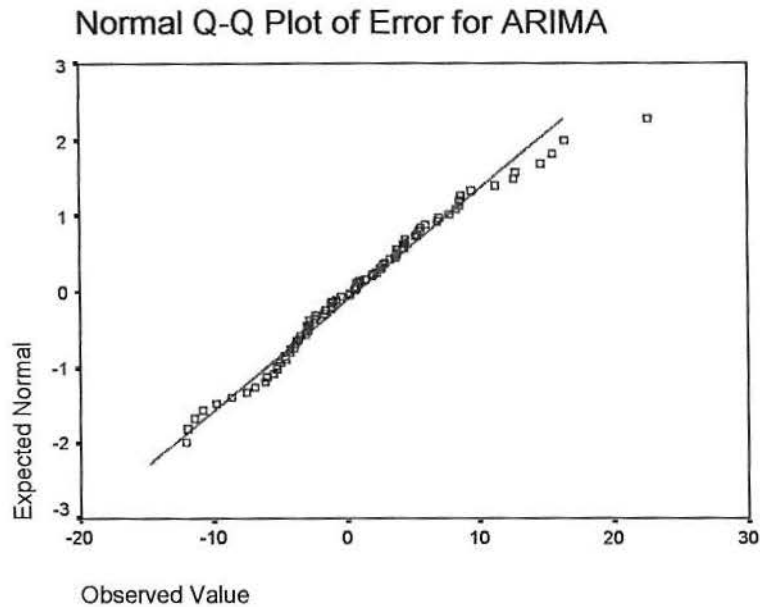


Figura 35. Gráfico de Probabilidade Normal dos resíduos do modelo ARIMA(0,1,1).

Tests of Normality			
	Kolmogorov-Smirnov ^a		
	Statistic	df	Sig.
Error for URH from ARIMA, MOD_6 NOCON	,058	84	,200*

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Tabela 4. Teste de normalidade nos resíduos do modelo ARIMA(0,1,1) ajustado.

O teste de normalidade evidencia que os resíduos são normalmente distribuídos, pois o valor do *p-value* do teste é de 0,2, maior que 0,15, que é a significância para esse teste. Os resíduos possuem média 0,7 e desvio padrão 6,85, como pode-se ver na figura 34. As figuras 34 e 35 também evidenciam que os resíduos são normalmente distribuídos.

3.5 Uso do Gráfico de Controle de Resíduos

O modelo estimado ARIMA(0,1,1) está representado no gráfico da figura 36.

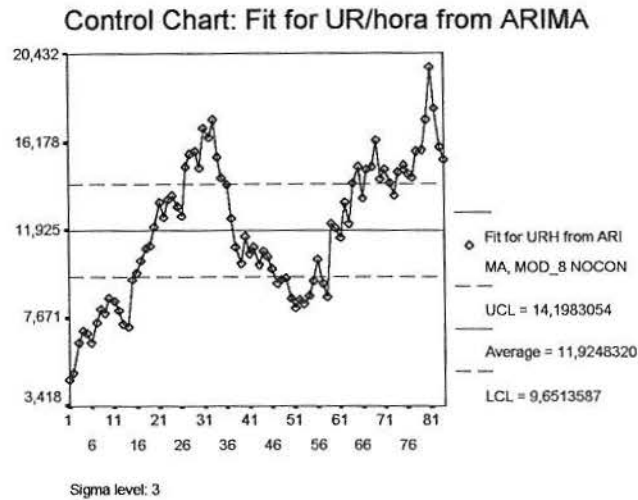


Figura 36. Gráfico de Controle para o modelo ARIMA(0,1,1)

Após o tratamento dos dados, que consistiu em um processo de retirada da autocorrelação e estimação de um modelo de séries temporais ARIMA para estabilizar a média, construiu-se um gráfico de controle para monitorar os resíduos do processo, o qual é representado na figura 37.

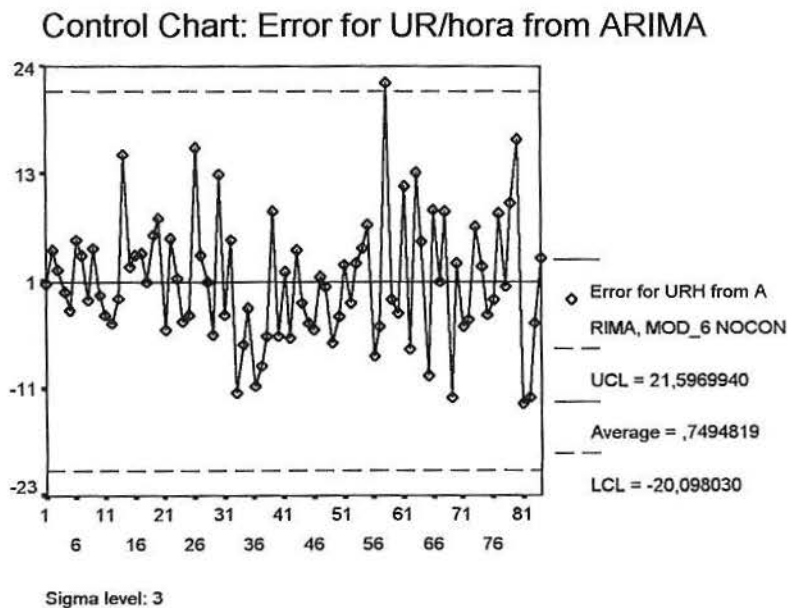


Figura 37. Gráfico de Controle para os resíduos do modelo ARIMA(0,1,1).

3.6 Discussão dos Resultados

Os resultados apresentados nesse capítulo são referentes ao modelo ARIMA(0,1,1) ajustado. As evidências desse estudo indicavam que esse era o melhor modelo para os dados originais. Entretanto, outros modelos também foram estimados: ARIMA(1,1,1), AR(1), ARIMA(1,1,0), MA(1), ARMA(1,1), mas nenhum deles obteve um bom desempenho na análise dos correlogramas de FAC e FACP e nas estatísticas Box-Ljung, AIC e BIC. A tabela 5 mostra um comparativo das estatísticas AIC e BIC entre vários modelos estimados. Como visto no capítulo 2, os melhores valores para p e q de um modelo ARIMA minimizam essas estatísticas.

MODELO	AIC	BIC
AR(1) com constante	574,32	579,21
MA(1) com constane	574,46	579,35
ARMA(1,1) com constante	571,40	578,73
ARIMA(1,1,0) sem constante	591,06	593,49
ARIMA(0,1,1) sem constante	564,64	567,07
ARIMA(1,1,1) sem constante	565,90	570,76

Tabela 5. Comparativo entre diversos modelos ajustados em relação às estatísticas AIC e BIC.

Quanto ao experimento, verificou-se uma certa instabilidade nos dados e uma variabilidade grande, principalmente nas últimas observações. Assim, pode-se sugerir analisar separadamente o tipo das atividades desenvolvidas pela equipe em gráficos de controle separados, e frisar que o aumento do desempenho é bom, mas o aumento da variabilidade do desempenho não é desejável. Assim, deve-se buscar melhorar o desempenho sem aumentar significativamente a sua variabilidade.

4 CONCLUSÕES E RECOMENDAÇÕES PARA FUTUROS TRABALHOS

4.1 Conclusões

Os gráficos de controle de Shewhart têm sido amplamente utilizados em diversos processos de monitoração. Entretanto, quando as observações estão autocorrelacionadas, eles não possuem um bom desempenho por violar as suposições de normalidade e independência, podendo muitas vezes assinalar uma causa especial quando ela não existe. Nesse caso, é necessário transformar os dados para retirar a autocorrelação antes de aplicar os gráficos de controle.

Essa monografia abordou uma ampla variedade de métodos para detectar e retirar a autocorrelação dos dados, dando ênfase para a estimação de um modelo ARIMA para retirar a autocorrelação, e posteriormente a aplicação de um gráfico de controle. Esse método recebeu maior destaque por ser eficiente na maioria dos casos, por ser mais conhecido e utilizado que os demais, e pela maior quantidade de estudos publicados. Além da aplicação dos modelos ARIMA, foram abordados também o uso do gráfico EWMA com linha central móvel; o uso do Controle da Engenharia do Processo, EPC; e o uso de Redes Neurais, que se caracteriza por ser um método recente para o tratamento de dados autocorrelacionados, complexo, e ainda com poucos estudos divulgados.

Praticamente em todos os artigos publicados sobre esse assunto, os pesquisadores estudam a distribuição *run length* (RL), que é a quantidade de amostras ou observações (no caso de gráfico para dados individuais) necessárias para detectar o primeiro ponto fora de controle, ou uma mudança na média do processo, quanto ao nível ou à inclinação. A

distribuição RL caracteriza a qualidade do método quanto a capacidade e o tempo para detectar *outliers*.

Apesar desse assunto já ser estudado há algum tempo, deve-se salientar que muitos métodos novos têm surgido nos últimos anos e publicados em periódicos do ramo. Por outro lado, os livros sobre CEP geralmente não tratam disso (entre os mais conhecidos, somente o Montgomery, 1997, trata sobre esse assunto). Também não há muitas referências em português, o que se constitui numa primeira contribuição desse trabalho, tornando acessível um assunto pouco discutido na língua portuguesa.

No final da revisão bibliográfica, propôs-se um diagrama denominado “modelo referencial de análise” que orienta como tratar os dados quando os resíduos estão autocorrelacionados, como outra contribuição dessa monografia.

Para ilustrar a pesquisa, a monografia apresentou um estudo de caso que avalia o desempenho de uma equipe formada por funcionários de uma empresa. O modelo encontrado para ajustar o processo foi um ARIMA(0,1,1), com os dados de resíduos posteriormente monitorados através de um gráfico de controle de resíduos.

Muitas aplicações de gráficos de controle, na prática, não têm se levado em conta verificar se os dados são normalmente distribuídos ou independentes. Isso implica que esses monitoramentos não terão um bom desempenho, sinalizando causa especial quando não há, e possivelmente deixando de assinalar essas causas quando elas ocorrerem. Essa monografia é um referencial bibliográfico para diversas situações de dados autocorrelacionados, destacando-se pelo fato de abordar um assunto muito importante na área de Controle Estatístico de Processo, porém pouco considerado na prática.

4.2 Recomendações

O assunto abordado nessa monografia dispõe de uma ampla quantidade de técnicas, como pode-se perceber no capítulo 2. Entre essas técnicas, muitas são recentes e surgiram nos últimos anos.

Uma grande quantidade de estudos pode ser realizada, principalmente na literatura em português, onde os tratamentos com dados autocorrelacionados são raros. Como sugestões para trabalhos futuros, destacam-se:

- o estudo do desempenho das técnicas mais recentes, como , por exemplo, o uso de redes neurais;
- um estudo teórico sobre o comportamento das distribuições RL dos diversos métodos existentes;
- o desenvolvimento de novas técnicas e procedimentos ou aprofundamento e extensões das já existentes;
- um comparativo mais aprofundado entre as diversas técnicas abordadas nessa monografia, etc.

5 REFERÊNCIAS BIBLIOGRÁFICAS

ATIENZA, O. O.; TANG, L. C.; ANG, B. W., A SPC Procedure for Detecting Level Shifts of Autocorrelated Processes. **Journal of Quality Technology**, v. 30, n. 4, p. 340-351, 1998.

ATIENZA, O. O.; TANG, L. C.; ANG, B. W., A CUSUM Scheme for Autocorrelated Observations. **Journal of Quality Technology**, v. 34, n. 2, p. 187-199, 2002.

BALESTRASSI, Pedro. P. **Identificação de Padrões em Gráficos de Controle Estatístico de Processos, em tempo real, utilizando Séries Temporais e Redes Neurais Artificiais**. Florianópolis, 2000. Tese de Doutorado em Engenharia de Produção – Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina.

BOX, G. E. P.; COLEMAN D. E.; BAXLEY, R. V. JR. A Comparison of Statistical Process Control and Engineering Process Control. **Journal of Quality Technology**, v. 29, n. 2, p. 128-130, 1997.

BOX, G. E. P.; JENKINS, G. M. **Times Series Analysis: Forecasting and Control**. 1ª Edição. San Francisco: Holden-Day, 1970.

CHIU, C.-C.; YANG, M.-H. Shifts Recognition in Correlated Manufacturing Process Parameters using Neural Network. **Journal of Chinese Institute of Industrial Engineers**, v. 18, n. 3, p. 86-94, 2001.

CHOU, C.-Y.; CHANG, C.-L.; CHEN, C.-H. Minimum-loss Design of X-Bar Control Charts for non-Normally Correlated Data. **Journal of Chinese Institute of Industrial Engineers**, v. 19, n. 1, p. 16-24, 2002.

CROWDER, S. V. A Simple Method for Studying Run-Length Distributions of Exponentially

- Weighted Moving Average Charts. **Technometrics**, v. 29, n. 4, p. 401-407, 1987.
- DEMING, W. Edwards. **Qualidade: A revolução da Administração**. 1ª Edição. Rio de Janeiro – RJ: Editora Marques Saraiva, 1990.
- FALTIN, F. W.; WOODALL, W. H. Discussion of “Some Statistical Process Control Methods for Autocorrelated Data”. **Journal of Quality Technology**, v. 23, n. 3, p. 194-197, 1991.
- FEIGENBAUM, Armand V. **Controle da Qualidade Total – Volume 3: Métodos Estatísticos Aplicados à Qualidade**. 3ª Edição. São Paulo – SP: Makron Books, 1994.
- FISCHER, S. **Séries Univariantes de Tempo – metodologia de Box & Jenkins**. Porto Alegre, FEE, 1982.
- GUJARATI, D. N. **Econometria Básica**. 3ª Edição. São Paulo – SP: Makron Books, 2000.
- HARRIS, T. J.; ROSS, W. H. Statistical Process Control Procedures for Correlated Observations. **The Canadian Journal of Chemical Engineering**, v. 69, p. 48-57, 1991.
- JIANG, W.; TSUI, K-L.; WOODALL, W. H. A New SPC Monitoring Method: The ARMA Chart. **Technometrics**, v. 42, n. 4, p. 399-410, 2000.
- KUME, H. **Métodos Estatísticos para Melhoria da Qualidade**. São Paulo: Editora Gente, 1993.
- LU, C-W.; REYNOLDS, M. R. JR., Cusum Charts for Monitoring an Autocorrelated Process. **Journal of Quality Technology**, v. 33, n. 3, p. 316-334, 2001.
- LU, C-W.; REYNOLDS, M. R. JR., EWMA Control Charts for Monitoring the Mean of Autocorrelated Processes. **Journal of Quality Technology**, v. 31, n. 1, p. 166-188, 1999.
- LU, C-W.; REYNOLDS, M. R. JR., Control Charts for Monitoring the Mean and Variance of Autocorrelated Processes. **Journal of Quality Technology**, v. 31, n. 3, p. 259-274, 1999.
- MACGREGOR, J. F. Discussion of “Some Statistical Process Control Methods for Autocorrelated Data”. **Journal of Quality Technology**, v. 23, n. 3, p. 198-199, 1991.
- MAKRIDASKIS, S. et al. **Forecasting – Method and Applications**. 3ª Edição. New York: John Wiley & Sons, 1998.

MONTGOMERY, D. C. **Introduction to Statistical Quality Control**. 3ª Edição. New York: John Wiley & Sons, 1997.

MONTGOMERY, D. C.; MASTRAGELO, C. M. Some Statistical Process Control Methods for Autocorrelated Data. **Journal of Quality Technology**, v. 23, n. 3, p. 179-193, 1991.

MONTGOMERY, D. C.; MASTRAGELO, C. M. Response of Discussion of “Some Statistical Process Control Methods for Autocorrelated Data”. **Journal of Quality Technology**, v. 23, n. 3, p. 203-204, 1991.

PALM, A. C.; RODRIGUEZ, R. N.; SPIRING, F. A.; WHEELER, D. J. Some Perspectives and Challenges for Control Chart Methods. **Journal of Quality Technology**, v. 29, n. 2, p. 122-127, 1997.

REYNOLDS, M. R.; ARNOLD, J. C.; BAIK, J. W. Variable Sampling Interval \bar{X} Charts in the Presence of Correlation. **Journal of Quality Technology**, v. 28, n. 1, p. 12-30, 1996.

RUNGER, G. C., Assignable Causes and Autocorrelation: Control Charts for Observations or Residuals? **Journal of Quality Technology**, v. 34, n. 2, p. 165-170, 2002.

RUNGER, G. C.; WILLEMAIN, T. R. Batch means Charts for Autocorrelated Data. **IIE Transactions**, 1995.

RUNGER, G. C.; WILLEMAIN, T. R. Model-Based and Model-Free Control of Autocorrelation Processes. **Journal of Quality Technology**, v. 27, n. 4, p. 283-192, 1995.

RYAN, T. P. Discussion of “Some Statistical Process Control Methods for Autocorrelated Data”. **Journal of Quality Technology**, v. 23, n. 3, p. 200-202, 1991.

SCHISSATTI, M. L. **Uma Metodologia de Implantação de Cartas de Shewhart para o Controle de Processos**. Florianópolis, 1998. Dissertação de Mestrado em Engenharia de Produção – Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina.

TOLOI, C. M. C.; MORETTIN, P. A. **Previsão de Séries Temporais**. São Paulo: Atual Editora, 1985.

WARDELL, D. G.; MOSKOWITZ, H.; PLANTE, R. D. Run-Length Distributions of Special-

Cause Control Charts for Correlated Processes. **Technometrics**, v. 36, n. 1, p. 3-17, 1994.

WARDELL, D. G.; MOSKOWITZ, H.; PLANTE, R. D. Run Length Distributions of Residual Control Charts for Autocorrelated Processes. **Journal of Quality Technology**, v. 26, n. 4, p. 308-317, 1994.

WIERINGA, J. E. **Statistical Process Control for Serially Correlated Data**. Zoutkamp, 1998. Tese de Doutorado em Economia, Rijksuniversiteit Groningen.

WRIGHT, C. M.; BOOTH, D. E.; HU, M. Y., Joint Estimation: SPC Method for Short-Run Autocorrelated Data. **Journal of Quality Technology**, v. 33, n. 3, p. 365-378, 2001.

YASHCHIN, E. Performance of CUSUM Control Schemes for Serially Correlated Observations. **Technometrics**, v. 35, n. 1, p. 37-52, 1993.