

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

DANILO CUZZUOL PEDRINI

**PROPOSTA DE UM MÉTODO PARA APLICAÇÃO DE
GRÁFICOS DE CONTROLE DE REGRESSÃO NO
MONITORAMENTO DE PROCESSOS**

Porto Alegre
2009

DANILO CUZZUOL PEDRINI

**PROPOSTA DE UM MÉTODO PARA APLICAÇÃO DE
GRÁFICOS DE CONTROLE DE REGRESSÃO NO
MONITORAMENTO DE PROCESSOS**

Dissertação submetida ao Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal do Rio Grande do Sul como requisito parcial à obtenção do título de Mestre em Engenharia de Produção, modalidade Acadêmica, na área de concentração em Sistemas de Qualidade.

Orientadora: Carla Schwengber ten Caten, Dr.^a.

Porto Alegre
2009

DANILO CUZZUOL PEDRINI

**PROPOSTA DE UM MÉTODO PARA APLICAÇÃO DE GRÁFICOS DE
CONTROLE DE REGRESSÃO NO MONITORAMENTO DE
PROCESSOS**

Esta dissertação foi julgada adequada para a obtenção do título de Mestre em Engenharia de Produção na modalidade Acadêmica e aprovada em sua forma final pelo Orientador e pela Banca Examinadora designada pelo Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal do Rio Grande do Sul.

Prof^a. Carla Schwengber ten Caten, Dr^a.

PPGEP / UFRGS

Orientadora

Prof. Flávio Sanson Fogliatto, Ph.D.

Coordenador PPGEP / UFRGS

Banca Examinadora:

Eugênio Kahn Epprecht, Dr. (DEI / PUC-Rio)

Morgana Pizzolato, Dr^a. (FEENG / UFRGS)

Flávio Sanson Fogliatto, Ph.D. (PPGEP / UFRGS)

*“Essencialmente, todos os modelos são
incorretos, mesmo assim alguns são úteis”*

George P. E. Box

DEDICATÓRIA

Dedico este trabalho aos meus pais, Dilma e Henrique, e às minhas irmãs Gisele e Ligia.

AGRADECIMENTOS

Essa dissertação foi o resultado de muito esforço, dedicação e persistência. A conclusão deste trabalho somente foi possível com o auxílio financeiro, intelectual e emocional de outras pessoas. Assim, eu gostaria de agradecer a todos que me ajudaram, inspiraram e auxiliaram durante o mestrado na Universidade Federal do Rio Grande do Sul.

O primeiro agradecimento é para minha orientadora, Carla, por seu apoio moral, valiosas sugestões, competente orientação e o conhecimento repassado nas áreas de projetos de experimentos, controle estatístico de qualidade e modelagem estatística. Agradeço também à oportunidade disponibilizada de poder aplicar e testar meus conhecimentos através de projetos de consultoria.

Agradeço aos membros da banca examinadora – professores Eugênio, Morgana e Flávio – por suas valiosas contribuições e críticas ao trabalho. Suas opiniões, fundamentadas em suas experiências e conhecimentos, melhoraram significativamente esta dissertação.

Agradeço ao ‘primo’ Ricardo Pedrini por seu suporte inicial em Porto Alegre e por acolher um estranho em sua família.

Também agradeço aos amigos que fiz em Porto Alegre – em especial aos colegas de apartamento e à galera do PPGEP – pelo apoio, discussões e inúmeros momentos de descontração durante estes últimos dois anos. Aos amigos de Aracruz e de Ouro Preto, pelo suporte nos momentos em que pude estar presente.

Agradeço ao nobre povo gaúcho pelo acolhimento de mais um estrangeiro em terras sulistas.

Uma das mais importantes bases de minha vida é minha família, que me ajudou em inúmeros momentos que necessitei e compreendeu a necessidade de minha ausência para melhorar profissionalmente. Para eles vão os meus mais profundos agradecimentos.

Agradeço ao PPGEP e à UFRGS por toda a estrutura disponibilizada. Também agradeço à inesquecível Ouro Preto, juntamente com a Escola de Minas e a UFOP, por minha graduação em engenharia e por ser um dos pilares de sustentação.

Por último, mas não menos importante, gostaria de agradecer a Deus, que me capacitou e meu deus saúde e amor para acabar este trabalho.

RESUMO

O presente trabalho propõe um método para a aplicação do gráfico de controle de regressão para o monitoramento de processos industriais. O método proposto inclui uma modificação do gráfico de controle de regressão múltipla, permitindo o monitoramento direto da característica de qualidade do processo ao invés do monitoramento dos resíduos padronizados do modelo de regressão, facilitando a interpretação dos operadores do processo. O método é dividido em duas fases principais: (i) Fase I – análise retrospectiva e (ii) Fase II – monitoramento do processo. A Fase I é composta pela coleta das amostras iniciais, estimação do modelo de regressão e análise de estabilidade dos dados coletados e, a partir desta fase, define-se alguns parâmetros a serem utilizados na fase seguinte. A Fase II do método consiste na coleta periódica de amostras, verificação da extrapolação dos valores das variáveis de controle e monitoramento do processo propriamente dito. O método proposto foi validado através da aplicação em um processo produtivo e de uma comparação do número médio de amostras (NMA) do gráfico de controle de regressão proposto, gerado através de simulação de Monte Carlo, com outros procedimentos similares encontrados na literatura. Como principais resultados esta dissertação apresenta: (i) proposta de um método sistematizado para nortear a aplicação de gráficos de controle de regressão; (ii) adaptação do gráfico de controle de regressão, de forma a permitir o monitoramento direto da característica de qualidade; (iii) proposta de um procedimento gráfico para a verificação da extrapolação das variáveis de controle e (iv) obtenção do NMA do gráfico de controle de regressão proposto e de outros procedimentos encontrados na literatura. O método proposto foi aplicado em um processo produtivo de uma indústria de borrachas.

Palavras-chave: gráfico de controle de regressão, número médio de amostras, modelos de regressão

ABSTRACT

This work proposes a method for the application of regression control charts in the monitoring of industrial processes. In order to facilitate the interpretation by the process operators, a modification in the multiple regression control chart is proposed allowing the direct monitoring of the values of quality characteristic of the process, instead of monitoring the regression standardized residuals. The proposed method is divided into two Phases: (i) Phase I, called retrospective analysis, and Phase II, called process monitoring. Phase I is composed by sampling, estimation of linear regression model and verification of stability of these samples. This phase defines some parameters to be used in the following phase. Phase II consists in periodic sampling of the process, altogether with verification of the extrapolation of process control variables and the process monitoring itself. The proposed method was validated through practical application in an industrial process and compared with other procedures found in literature. This work has also achieved the average run length (ARL) of the proposed regression control chart, which was compared with the other procedures consulted. The main contributions of this work may be pointed: (i) the proposal of a method to guide the application of regression control chart; (ii) the adaptation of the multiple regression control chart, allowing the direct monitoring of the quality characteristic; (iii) the proposal of a control chart to monitor the extrapolation of the process control variable and (iv) the obtaining of the ARL of the proposed regression control chart and other similar procedures. The proposed method was applied in a process of a rubber manufactory.

Key words: regression control chart, average run length, regression models.

LISTA DE FIGURAS

Figura 1: Análise de variância para testar a significância da regressão	29
Figura 2: Exemplo de extrapolação escondida.....	33
Figura 3: Padrões de comportamento de gráficos de resíduos: (a) satisfatório, (b) funil, (c) arco duplo e (d) não linear	35
Figura 4: Gráfico de dispersão para um modelo de regressão linear simples ilustrando alguns casos de <i>outliers</i>	37
Figura 5: Exemplo de um típico gráfico de Controle	46
Figura 6: Limites de controle para os gráficos X- R e X- S	47
Figura 7: Limites de controle para os gráficos X-MR.....	47
Figura 8: Exemplo de um gráfico de controle CUSUM.....	50
Figura 9: Exemplo de um gráfico de controle EWMA	52
Figura 10: Exemplo genérico de um gráfico de controle de regressão	57
Figura 11: Gráfico de controle de regressão para o processamento de correspondências	61
Figura 12: Gráfico de controle de regressão múltipla para o monitoramento do tempo de manutenção de softwares	63
Figura 13: Gráfico de controle de regressão para o processo de galvanização	66
Figura 14: Gráfico de controle de medidas individuais aplicado aos resíduos	66
Figura 15: Regras de decisão para o gráfico de controle para seleção de causas.....	69
Figura 16: Gráfico de controle de regressão para o processamento de resíduos.....	72
Figura 17: Gráfico de controle de regressão para o monitoramento da produção de tijolos...	73
Figura 18: Gráfico de controle de regressão para o monitoramento de sucatas na produção de colheitadeiras	74
Figura 19: Sumário de trabalhos sobre gráfico de controle de regressão consultados.....	75
Figura 20: Fluxograma da Fase I do método proposto.....	79
Figura 21: Fluxograma da Fase II do método proposto	84
Figura 22: Esquema do processo de extrusão de bandas de rodagens	90
Figura 23: Distância de Cook para as amostras utilizadas para estimar o modelo.....	95
Figura 24: Gráfico dos resíduos <i>versus</i> valores estimados de y	96
Figura 25: Gráfico de probabilidade normal dos resíduos	96
Figura 26: Gráfico de controle de regressão da Fase I do método proposto	97
Figura 27: Gráfico de controle de extrapolação	99

Figura 28: Gráfico de controle de regressão da Fase II do método proposto.....	100
Figura 30: Gráfico de medidas individuais para o processo de extrusão	103
Figura 31: Gráfico EWMAREG para os resíduos padronizados do modelo	103
Figura 32: Gráfico CUSUMREG para os resíduos padronizados do modelo.....	104

LISTA DE TABELAS

Tabela 1: ANOVA para o modelo estimado.....	93
Tabela 2: Testes t individuais e valores FIV para as variáveis de controle e coeficiente de intercepto.....	93
Tabela 3: Valores NMA para alterações no coeficiente de intercepto	108
Tabela 4: Valores NMA para alterações no coeficiente de inclinação β_1	110
Tabela 5: Valores NMA para alterações no coeficiente de inclinação β_0 e β_1	111

SUMÁRIO

1	INTRODUÇÃO	13
1.1	COMENTÁRIOS INICIAIS.....	13
1.2	TEMA	15
1.3	OBJETIVOS	15
1.4	JUSTIFICATIVAS DO TEMA E OBJETIVOS	15
1.5	MÉTODOS	18
1.6	DELIMITAÇÕES DO TRABALHO.....	19
1.7	ESTRUTURA DO TRABALHO.....	20
2	REVISÃO BIBLIOGRÁFICA	22
2.1	MODELOS DE REGRESSÃO LINEAR.....	22
2.1.1	Modelos de Regressão Linear Simples	23
2.1.2	Modelos de Regressão Linear Múltipla	24
2.1.3	Estimação da Equação de Regressão Linear.....	25
2.1.4	Testes de Hipóteses e Intervalos de Confiança na Regressão Linear	27
2.1.5	Aplicação da Regressão Linear: a Estimação de Valores	30
2.1.6	Coeficientes de Determinação	33
2.1.7	Análise Residual	34
2.1.8	Observações Influentes	38
2.1.9	Seleção de Variáveis e de Modelos.....	40
2.1.10	Multicolinearidade	43
2.2	CONTROLE ESTATÍSTICO DE PROCESSOS	45
2.2.1	Gráficos de Controle para Variáveis.....	46
2.2.2	Gráficos de Controle CUSUM e EWMA	48
2.2.3	Análise do Comportamento dos Gráficos de Controle	52
2.2.4	Sensibilidade do Gráfico de Controle	53
2.3	GRÁFICO DE CONTROLE DE REGRESSÃO.....	55
2.3.1	Comparações entre os Gráficos de Controle de Regressão e de Shewhart.....	56
2.3.2	Limites de Controle para o Gráfico de Controle de Regressão.....	58
2.3.3	Construção do Gráfico de Controle de Regressão	60
2.3.4	Extensões do Gráfico de Controle de Regressão	62
2.3.5	Procedimentos Similares ao Gráfico de Controle de Regressão.....	68
2.3.6	Exemplos de Aplicações de Gráfico de Controle de Regressão	71
2.4	CONCLUSÃO DO CAPÍTULO.....	74
3	MÉTODO PROPOSTO	76
3.1	INTRODUÇÃO	76
3.1.1	Suposições.....	77
3.1.2	Aplicações.....	78
3.2	DETALHAMENTO DO MÉTODO PROPOSTO	78
3.2.1	Fase I – Análise Retrospectiva.....	78
3.2.2	Fase II – Monitoramento do Processo	83
3.3	CONCLUSÃO DO CAPÍTULO.....	87

4	APLICAÇÃO DO MÉTODO PROPOSTO	89
4.1	APLICAÇÃO DO MÉTODO EM UM SISTEMA PRODUTIVO	89
4.1.1	Descrição do Processo	89
4.1.2	Aplicação da Fase I.....	92
4.1.3	Aplicação da Fase II.....	97
4.2	COMPARAÇÃO COM OUTROS GRÁFICOS DE CONTROLE.....	100
4.2.1	Comparação Gráfica	101
4.2.2	Comparação de Sensibilidade	105
4.3	CONCLUSÃO DO CAPÍTULO.....	112
5	CONSIDERAÇÕES FINAIS	114
5.1	CONCLUSÕES	114
5.2	SUGESTÕES PARA TRABALHOS FUTUROS	116
	REFERÊNCIAS BIBLIOGRÁFICAS	117

1 INTRODUÇÃO

O presente capítulo tem como objetivo introduzir o leitor aos tópicos a serem abordados ao longo da dissertação. São apresentados os comentários iniciais acerca do tema a ser pesquisado, os objetivos do trabalho, as justificativas práticas e acadêmicas para o tema e objetivos, as delimitações e a estrutura desta dissertação.

1.1 COMENTÁRIOS INICIAIS

A qualidade tornou-se um dos mais importantes fatores de decisão dos consumidores na escolha de produtos e serviços. Dessa forma, as indústrias necessitam ter um melhor controle sobre seus processos, para tentar assegurar que as propriedades e características de seus produtos satisfaçam as especificações. Nesse contexto, o controle estatístico de processo (CEP) é extremamente útil, já que é um conjunto de ferramentas para a coleta, análise e interpretação de dados, com o objetivo de melhorar a qualidade através da monitoração de processos, permitindo a identificação e a eliminação de causas especiais de variação, podendo ser utilizado para a maioria dos processos (MONTGOMERY, 2004).

De acordo com Woodall e Montgomery (1999), os gráficos de controle de Shewhart destacam-se dentre as ferramentas do CEP, devido principalmente à sua simplicidade operacional e à sua efetividade na detecção de problemas no processo, sendo utilizados com sucesso no monitoramento do desempenho de muitos processos industriais. Estes gráficos de controle podem ser classificados em duas categorias: os gráficos de controle para variáveis, quando a característica de qualidade é avaliada através de medições em escala contínua e os gráficos de controle para atributos, para o caso em que a característica de qualidade é uma classificação, tal como, conforme ou não-conforme (MONTGOMERY, 2004).

Para a utilização dos gráficos de controle, deve-se assumir que os dados sejam independentes e identicamente distribuídos. Na situação em que o ajuste das variáveis de controle é freqüentemente alterado, pode ocorrer a alteração na distribuição dos dados, sendo necessário um gráfico de controle para cada ajuste, o que pode não ser possível devido ao baixo número de amostras a serem analisadas para cada ajuste.

Na abordagem tradicional do CEP, o processo é avaliado mediante a observação da característica de qualidade, sem serem consideradas as variáveis de controle do processo, essa prática pode ser inadequada e pouco efetiva quando há correlação entre a característica de

qualidade e as variáveis de controle. Nesses casos, a característica de qualidade de um produto ou processo é melhor representada pelo seu relacionamento com as variáveis de controle do processo podendo-se combinar a ferramenta de regressão linear com o CEP, tornando possível a avaliação da relação entre as variáveis de controle do processo ao longo do tempo, além de permitir prever o comportamento da característica de qualidade (MANDEL, 1969; JACOBI, 2001; MONTGOMERY, 2004).

Embora existam alguns trabalhos anteriores, Mandel (1969) é considerado como o trabalho inaugurador do gráfico de controle de regressão, já que foi o primeiro trabalho a propor uma metodologia clara para esta ferramenta. Mandel (1969) também foi o primeiro trabalho a definir claramente os elementos do gráfico de controle de regressão, além de destacar as semelhanças e diferenças com os gráficos de controle clássicos, aplicando esse novo gráfico de controle para a entrega de correspondências e o total de mão-de-obra utilizadas nas 74 maiores agências de correios dos Estados Unidos entre 1965 e 1967.

De acordo com Almeida (2003), o gráfico de controle de regressão tem como principal função controlar uma variação média da característica de qualidade, que ocorre devido à ação das variáveis de controle do processo, ao invés de controlar uma média constante do processo como é usualmente feito nos gráficos clássicos. Para Montgomery (2004) e Cai et al. (2002), esse gráfico de controle também pode ser aplicado para monitorar e analisar processos que apresentem tendências, que ocorrem devido à interferência de variáveis de controle do processo, como por exemplo, o desgaste de ferramentas. Almeida (2003) comenta que esse gráfico de controle é pouco utilizado em empresas, pois exige a existência de uma relação de causa e efeito entre as variáveis, sendo que essa relação deve ser corretamente estabelecida sob pena de levar a decisões erradas acerca do processo.

Mandel (1969), Olin (1998) e Jacobi et al. (2002) mostraram em seus estudos que o gráfico de controle de regressão pode ser construído a partir de dados históricos do processo, o que possibilita avaliar se as novas observações estão sob controle. O método comum encontrado na literatura é a utilização da regressão linear simples para modelar o comportamento da característica de qualidade com a fixação de limites de controle baseados em limites simples ou em intervalos de predição. Por sua vez, Spanos (1992) e Haworth (1996) aplicaram o gráfico de controle de regressão, utilizando um gráfico de controle para os resíduos da regressão. Shu et al. (2004) propuseram a utilização de gráficos de controle de médias móveis exponencialmente ponderadas (EWMA) para os resíduos da regressão, chamando-o de EWMAREG. Estes três últimos trabalhos apresentam o diferencial de preservar a ordem temporal dos dados, fato que não aconteceu no trabalho de Mandel (1969).

1.2 TEMA

O tema desta dissertação são os gráficos de controle de regressão. Este tema contempla as áreas de engenharia de qualidade, concentrando em gráficos de controle e modelos de regressão linear simples ou múltipla.

1.3 OBJETIVOS

O objetivo principal deste trabalho é propor um método para a aplicação dos gráficos de controle de regressão em processos industriais, considerando as fases de análise retrospectiva e de monitoramento do processo. Apresentam-se os seguintes objetivos específicos desta pesquisa:

- a) Propor uma modificação no gráfico de controle de regressão múltipla, com a adaptação dos limites de controle, de forma que seja possível monitorar os dados diretamente nas unidades originais, além de preservar a ordem temporal dos dados;
- b) Propor um gráfico de controle para a verificação da extrapolação das variáveis de controle;
- c) Comparar o gráfico de controle de regressão proposto com os gráficos encontrados na literatura: gráfico de controle de regressão linear múltipla, gráfico de controle de medidas individuais aplicado aos resíduos do modelo de regressão, gráfico EWMAREG e gráfico de somas cumulativas aplicado aos resíduos do modelo (CUSUMREG);
- d) Obter o número médio de amostras até o sinal (NMA) para o gráfico de controle de regressão proposto e comparar com os demais gráficos citados anteriormente.

1.4 JUSTIFICATIVAS DO TEMA E OBJETIVOS

Ao se trabalhar com gráficos de controle, supõe-se que as amostras provenientes do processo sejam independentes e identicamente distribuídas. Quando existir relação de causa e efeito entre a característica de qualidade monitorada e as variáveis de controle do processo e

estas últimas forem constantemente alteradas, o modelo não será válido para todas as amostras do processo. Neste caso, os valores da característica de qualidade irão variar, conforme os valores das variáveis de controle, aumentando desnecessariamente o número de alarmes falsos do processo. Se a função que relaciona a característica de qualidade às variáveis de controle for linear, é possível aplicar os gráficos de controle de regressão (MANDEL, 1969; SHU et al., 2004). No caso de regressão não linear pode-se aplicar o gráfico de regressão não linear proposto por Olin (1998).

Uma justificativa prática para o tema é a necessidade de monitoramento das características de qualidade de alguns processos, considerando conjuntamente as variáveis de controle do processo ao invés de se monitorar apenas o valor médio da característica de qualidade. Em alguns casos, alguma variável de controle pode estar ajustada em um nível extremo, fazendo com que as saídas do processo aparentem estar fora de controle, quando na verdade o processo não está. Nessa situação pode ser mais efetivo construir um gráfico com a característica de qualidade ajustada em relação às variáveis de controle do que plotar a saída do processo diretamente (JACOBI et al., 2002; SHU et al., 2004;2007).

A justificativa para o objetivo principal do presente trabalho é que não foram encontrados trabalhos que estruturassem o método proposto de forma a facilitar a adoção do gráfico de controle de regressão em sistemas produtivos. Embora seja uma linha de pesquisa consideravelmente antiga, foram encontrados poucos artigos que propuseram métodos estruturados para a aplicação de gráficos de controle de regressão em processos industriais. No Brasil, essa linha de pesquisa é pouco explorada, sendo que o primeiro trabalho encontrado foi Teixeira (1997). Além deste trabalho, foram encontrados na literatura nacional alguns outros trabalhos, como: Jacobi (2001), Jacobi et al. (2002), Almeida (2003), Carvalho Jr. (2006) e Casarin et al. (2007).

Outra necessidade relevante é a descrição de um método para a implantação de um gráfico de controle de regressão para processos industriais que contenham mais do que uma variável de controle, ou seja, utilizando modelos de regressão linear múltipla. Spanos (1992) e Haworth (1996) foram os únicos trabalhos localizados que utilizaram o gráfico de controle de regressão múltipla, embora não tenham apresentado métodos claros para sua aplicação. Através da proposta de um método simples, espera-se que o gráfico de controle de regressão também possa ser aplicado em processos produtivos que contenham mais do que uma variável de controle, aumentando a eficiência do CEP nesses processos, se as suposições estabelecidas no método proposto forem satisfeitas.

A modificação do gráfico de controle de regressão proposta no presente trabalho tem por objetivo monitorar diretamente a característica de qualidade do processo estudado, ao invés do monitoramento dos resíduos. Montgomery e Mastrangelo (1991), Montgomery (2004) e Moreira Jr. (2005) argumentam que o monitoramento dos valores reais da característica de qualidade permite um melhor entendimento e aplicação dos gráficos de controle por parte dos operários, já que estes não consideram os resíduos como uma representação direta do processo. Montgomery e Mastrangelo (1991) também salientam que esse tipo de adaptação é bastante benéfico, pois combina informações sobre a dinâmica do processo e do controle estatístico em um único gráfico.

Ao contrário da abordagem mais usual para os gráficos de controle de regressão, a modificação proposta preservará a ordem temporal dos dados, o que permitirá a utilização de diferentes testes para a verificação da estabilidade do processo.

O gráfico de controle de extrapolação visa facilitar a verificação de que o ajuste utilizado para as variáveis de controle do processo está extrapolando a região definida pelo conjunto de valores das variáveis de controle utilizadas para estimar o modelo de regressão. Caso esta extrapolação ocorra, Montgomery et al. (2001), Neter et al. (2005) e Weisberg (2005) afirmam que as previsões do modelo podem ser ruins.

A comparação dos resultados do método proposto com outros métodos encontrados na literatura visa provar a eficiência do gráfico de controle proposto em relação a outros procedimentos e, assim, validá-lo. Serão comparados quatro gráficos diferentes:

- a) Gráfico de controle de regressão múltipla, proposto por Haworth (1996);
- b) Gráfico de controle de medidas individuais, com limites de controle calculados com base na amplitude móvel dos resíduos, proposto por Rothshild e Roth (1986) e Loredó et al. (2002);
- c) Gráfico de controle EWMAREG, proposto por Teixeira (1997) e Shu et al. (2004);
- d) Gráfico de controle de somas cumulativas aplicado aos resíduos do modelo de regressão, definido neste trabalho como gráfico CUSUMREG.

Segundo Costa et al. (2005), o NMA é uma medida de desempenho do gráfico de controle, e seu cálculo visa determinar a taxa de alarmes falsos do gráfico de controle proposto e a sensibilidade deste gráfico às alterações no processo pré-definidas. Com os valores do NMA, espera-se validar o gráfico de controle proposto.

1.5 MÉTODOS

O método do trabalho deve ser apresentado segundo dois aspectos: (i) a caracterização do tipo de pesquisa e (ii) descrição das etapas, técnicas e ferramentas que serão utilizadas para atingir os objetivos da pesquisa.

Do ponto de vista da sua natureza, segundo Silva e Menezes (2000), um trabalho pode ser classificado como sendo uma pesquisa básica ou uma pesquisa aplicada. Dessa forma, o presente trabalho classifica-se como sendo uma pesquisa aplicada, já que possui foco na geração de conhecimentos dirigidos ao atendimento de necessidades e à solução de problemas específicos (OLIVEIRA, 2000; SILVA; MENEZES, 2000).

Quanto à abordagem, um trabalho científico pode ser classificado como uma pesquisa qualitativa ou uma pesquisa quantitativa. Seguindo esta classificação, o presente trabalho é uma pesquisa quantitativa, pois possui ênfase em análises numéricas, baseando-se no uso exaustivo de técnicas estatísticas (OLIVEIRA, 2000; SILVA; MENEZES, 2000).

Do ponto de vista dos objetivos, uma pesquisa pode ser exploratória, descritiva ou explicativa. O presente trabalho é classificado como sendo uma pesquisa exploratória, já que visa proporcionar um maior conhecimento sobre o problema estudado, visando torná-lo explícito (GIL, 1991).

O desenvolvimento do presente trabalho será realizado em 3 etapas: (i) pesquisa bibliográfica, (ii) proposta de método e (iii) validação dos resultados. A primeira etapa, pesquisa bibliográfica, contém o levantamento do estado da arte dos gráficos de controle de regressão encontrados na literatura nacional e internacional. Com a pesquisa bibliográfica, são identificados os principais procedimentos metodológicos da literatura, o que fornece as bases necessárias para a proposta de um novo método para a aplicação do gráfico de controle de regressão. Ressalta-se que pesquisa bibliográfica, de acordo com Oliveira (2000), é o levantamento de informações já existentes, em livros e artigos científicos, com a finalidade de dar fundamentação teórica ao trabalho.

A segunda etapa será a definição do método de implantação do gráfico de controle de regressão em um sistema produtivo, englobando adaptações dos métodos encontrados na literatura e a modificação proposta no presente trabalho. O método proposto engloba os procedimentos para estimação da equação de regressão múltipla, o cálculo dos limites de controle e a construção dos gráficos de controle de regressão. Serão apresentados os testes de estabilidade a serem aplicados para avaliar o estado do processo e as formas de identificação

de causas especiais que venham porventura a surgir. Nessa etapa também será elaborado o aplicativo em Excel, de forma a facilitar a aplicação do gráfico de controle de regressão.

A terceira etapa do trabalho será a validação do método proposto, realizada em duas partes: (i) aplicação do método proposto e (ii) comparação com os resultados de outros gráficos de controle encontrados na literatura. Estas partes, embora sejam inter-relacionados, apresentam objetivos diferentes, já que a primeira visa provar a aplicabilidade e a praticidade do método proposto e a segunda objetiva verificar a sua eficiência quando comparado a outros gráficos de controle encontrados na literatura. A comparação dos resultados será realizada em duas formas: aplicação dos outros gráficos de controle para dados utilizados na aplicação e simulação para obtenção do NMA de cada gráfico. Como resultado desta validação, espera-se provar que o método proposto pode ser utilizado em outros processos produtivos, se as suposições estabelecidas forem satisfeitas.

Para a aplicação do método proposto serão utilizados procedimentos de estudo de caso, já que este procedimento metodológico pode ser utilizado para entender, descrever e testar os resultados advindos da aplicação do método proposto, além de permitir fazer algumas generalizações dos resultados encontrados (EISENHARDT, 1989; MEREDITH, 1998). Para finalizar o trabalho, serão apresentadas as considerações finais e conclusões, com a apresentação dos principais resultados obtidos na implantação do método proposto, além de sugestões para trabalhos futuros.

1.6 DELIMITAÇÕES DO TRABALHO

Como citado anteriormente, a delimitação geral dessa pesquisa é a definição de um método para a aplicação do gráfico de controle de regressão em um sistema produtivo, englobando uma modificação do gráfico de controle de regressão que permita a monitoração direta dos valores originais da característica de qualidade estudada, ao invés do monitoramento dos resíduos. No presente trabalho será estimado um modelo de regressão linear múltipla para a modelagem de apenas uma característica de qualidade, portanto, não será contemplada a modelagem não linear, modelagem multivariada e modelos lineares generalizados (GLM).

Autores como Almeida (2003) e Carvalho Jr. (2006) estudaram o gráfico de controle de regressão considerando os modelos com erros nas variáveis, ou seja, modelos em que as

variáveis estão sujeitas a erros de medidas. No presente trabalho, serão consideradas como sem erros de medidas, não sendo necessário o uso de modelos com erros nas variáveis.

A utilização dos índices de capacidade, embora seja útil para avaliar a capacidade do processo em atender às especificações nominais do produto, não será escopo do presente trabalho.

O método a ser proposto será aplicado e validado em apenas um processo produtivo de uma única empresa, assim generalizações deverão ser realizadas com algumas ressalvas. O método proposto não incluirá procedimentos para a identificação das causas especiais que alteraram o comportamento do sistema, focando apenas em apontar quais amostras estão fora de controle estatístico.

1.7 ESTRUTURA DO TRABALHO

Esta pesquisa está organizada em cinco capítulos principais. No primeiro capítulo apresenta-se o tema, sendo justificada sua importância acadêmica e prática. Também são apresentados o objetivo geral e os objetivos específicos do trabalho, suas justificativas, o método de trabalho, as delimitações e a estrutura da dissertação.

O segundo capítulo apresenta uma revisão bibliográfica que contém: (i) modelos de regressão linear múltipla, (ii) revisão breve do CEP, aprofundando-se nos gráficos de controle EWMA e CUSUM e (iii) combinação de regressão e gráficos de controle, onde é apresentado o gráfico de controle de regressão.

No terceiro capítulo é apresentada a modificação do gráfico de controle de regressão proposta, também é apresentado o método proposto para a implementação do gráfico de controle de regressão em um sistema produtivo.

No quarto capítulo apresenta-se brevemente o sistema produtivo a ser estudado e o processo onde será aplicado o método. Também é apresentado o processo de implantação do gráfico de controle de regressão, suas particularidades, dificuldades de utilização e, assim, apresenta-se a validação do método proposto. Por fim, serão apresentados os resultados do gráfico de controle de regressão e as comparações com outros tipos de gráficos de controle encontrados na literatura, como o gráfico de controle de regressão múltipla, o gráfico de controle para medidas individuais para os resíduos do modelo de regressão e os gráficos EWMAREG e CUSUMREG.

No quinto capítulo serão explicitados os principais resultados do trabalho, as conclusões e algumas sugestões para trabalhos futuros.

2 REVISÃO BIBLIOGRÁFICA

O presente capítulo apresenta a fundamentação teórica para o desenvolvimento do restante do trabalho. Essa seção será subdividida em três partes: a primeira parte abordará os modelos de regressão linear, englobando o contexto histórico dessa ferramenta, a forma geral dos modelos de regressão, o método dos mínimos quadrados ordinários, suas suposições e pressupostos, os testes realizados, critérios de seleção de variáveis e de modelos, a análise residual e os métodos de estimação usando a equação de regressão estimada.

A segunda parte do capítulo abordará, de forma geral, o controle estatístico de processo e um breve histórico, descrevendo sucintamente os principais gráficos de controle e se aprofundando nos gráficos de controle para médias móveis exponencialmente ponderadas (EWMA) e de somas cumulativas (CUSUM).

A terceira parte do capítulo abordará os gráficos de controle de regressão, englobando sua evolução histórica, suas similaridades e diferenças em relação aos demais gráficos de controle, a estimação dos limites de controle, algumas técnicas similares, modificações propostas ao gráfico de controle de regressão por alguns autores e alguns exemplos de aplicação do gráfico de controle de regressão encontrados na literatura.

2.1 MODELOS DE REGRESSÃO LINEAR

O termo regressão foi introduzido por Francis Galton, em seu estudo sobre a altura de crianças. Nesse estudo, Galton constatou que a altura média dos filhos de pais de uma determinada altura tendia a “regredir” até a altura média da população. Esse fenômeno foi chamado de lei de regressão universal de Galton e foi confirmado posteriormente por Pearson (GUJARATI, 2000; NETER et al., 2005).

O significado atual da análise de regressão é diferente da abordagem inicial, sendo definida por Montgomery et al. (2001, p. 1), como “uma técnica estatística para modelar e investigar a relação entre variáveis”. Weisberg (2005) defende que a regressão, assim como a maioria das técnicas estatísticas, tem como meta sumarizar os dados observados da forma mais simples, útil e elegante possível. Para Gujarati (2000) e Neter et al. (2005) a análise de regressão estuda a dependência da característica de qualidade, em relação a uma ou mais variáveis, aqui denominadas de variáveis de controle, com o objetivo de estimar o valor

médio da característica de qualidade em função dos valores conhecidos das variáveis de controle.

A análise de regressão é uma das ferramentas mais utilizadas para a análise de dados, sendo aplicada em quase todas as áreas de conhecimento, incluindo: engenharia, física, química, economia, administração, ciências biológicas e da vida e ciências sociais. A simplicidade e utilidade dessa técnica são resultantes do conceito lógico de utilizar uma equação matemática para expressar a relação entre variáveis. A análise de regressão é também interessante devido à sua elegante base matemática e bem-desenvolvida teoria estatística (MONTGOMERY et al., 2001). Montgomery (2005) ainda aponta o uso da regressão para a análise de experimentos planejados que tiveram algumas falhas, como por exemplo, experimentos em que uma das rodadas não foi realizada corretamente (e essa não pode ser refeita) e casos em que, por algum descuido, os níveis de alguns fatores não puderam ser bem controlados com a precisão desejada para o experimento.

2.1.1 Modelos de Regressão Linear Simples

O modelo de regressão linear simples (MRLS) é representado pela seguinte equação (MONTGOMERY, et al., 2001):

$$y = \beta_0 + \beta_1 x + \varepsilon \quad (1)$$

Segundo Neter et al. (2005), a equação (1) é dita simples, pois representa a relação entre uma característica de qualidade e uma variável de controle, é linear quanto aos parâmetros, pois nenhum dos parâmetros aparece como expoente ou está sendo multiplicado ou dividido por outros parâmetros.

O termo ε é chamado de erro aleatório, sendo assumido como normal e independentemente distribuído, com média zero e variância constante e desconhecida σ^2 . Essa suposição é importante para a estimação da reta de regressão e deve ser verificada posteriormente para a validação do modelo estimado (MONTGOMERY et al., 2001; NETER et al., 2005).

Montgomery et al. (2001) afirmam que a variável x deve ser vista como sendo uma variável de controle e medida com erro desprezível, enquanto a variável y é vista como uma variável aleatória. Dessa forma, existe uma distribuição de probabilidade para y a cada valor de x . A média dessa distribuição é estimada por:

$$E(y|x) = \hat{\beta}_0 + \hat{\beta}_1 x \quad (2)$$

A variância é:

$$var(y|x) = var(\hat{\beta}_0 + \hat{\beta}_1 x + \varepsilon) = \sigma^2 \quad (3)$$

Para o uso de um modelo de regressão linear simples, é necessário utilizar um gráfico de dispersão, que, de acordo com Montgomery e Runger (2003), é uma representação de cada par (x_i, y_i) em um ponto plotado em um sistema bidimensional. Ainda segundo Montgomery e Runger (2003), a análise desse gráfico indica se os pontos repousam aleatoriamente dispersos em torno de uma linha reta.

2.1.2 Modelos de Regressão Linear Múltipla

De acordo com Montgomery et al. (2001), o modelo de regressão linear múltipla (MRLM) com k variáveis de controle é representado pela equação (4):

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon \quad (4)$$

Os coeficientes de regressão $\beta_0, \beta_1, \dots, \beta_k$ são interpretados por Gujarati (2000) e Montgomery et al. (2001) como:

- O coeficiente β_0 é chamado de coeficiente de intercepto, sendo interpretado como a média de y quando todas as variáveis de controle são iguais a zero;
- Os coeficientes β_1, \dots, β_k são chamados de coeficientes de regressão parcial. O coeficiente β_k pode ser interpretado como a derivada parcial de y em relação a x_k , ou seja, é variação de y causada por uma variação unitária em x_k dado que as outras variáveis de controle foram mantidas constantes.

É mais prático trabalhar com modelos de regressão linear múltipla quando a equação (4) é expressa em notação matricial, já que essa notação permite uma apresentação mais compacta do modelo, dados e resultados para qualquer número de variáveis (GUJARATI, 2000; MONTGOMERY et al., 2001).

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (5)$$

onde:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & \dots & x_{1k} \\ 1 & x_{21} & \dots & x_{2k} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n1} & \dots & x_{nk} \end{bmatrix} \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

2.1.3 Estimação da Equação de Regressão Linear

Na maioria das aplicações, os coeficientes de regressão reais são desconhecidos e devem ser estimados a partir de n vetores de dados amostrais. Esses dados amostrais podem ser obtidos através de projetos experimentais, dados históricos provenientes de outros estudos ou observações de um fenômeno.

Quando o número de observações (n) for maior que o número de variáveis de controle (k), o método mais usual para estimar a equação de regressão é o método dos mínimos quadrados ordinários, que visa estimar os parâmetros da regressão de forma a minimizar as somas quadráticas dos resíduos da regressão (MONTGOMERY et al., 2001; NETER et al., 2005; WEISBERG, 2005):

$$L = \sum_{i=1}^n \varepsilon_i^2 = \boldsymbol{\varepsilon}' \boldsymbol{\varepsilon} = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) = \mathbf{y}'\mathbf{y} - 2\boldsymbol{\beta}'\mathbf{X}'\mathbf{y} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \quad (6)$$

Derivando-se a equação (6) em relação aos parâmetros $\boldsymbol{\beta}$ e igualando-se essa derivada a zero:

$$\frac{\partial L}{\partial \boldsymbol{\beta}} = 2\mathbf{X}'\mathbf{y} + \mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = 0 \quad (7)$$

O estimador dos mínimos quadrados ordinários de $\boldsymbol{\beta}$ é obtido pela resolução da equação (7):

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y} \quad (8)$$

Segundo Montgomery et al. (2001) e Weisberg (2005), essa estimativa é válida se a matriz $(\mathbf{X}'\mathbf{X})$ for inversível. Dessa forma, a matriz $(\mathbf{X}'\mathbf{X})^{-1}$ sempre existirá se nenhuma coluna da matriz \mathbf{X} for uma combinação linear de outra coluna, ou seja, se as variáveis de controle forem linearmente independentes. Para utilizar o método dos mínimos quadrados ordinários é necessário supor que os resíduos sejam normalmente distribuídos com média zero e desvio padrão constante. Para Weisberg (2005) as suposições de normalidade dos resíduos na forma matricial podem ser representadas da como a equação (9).

$$\mathbf{e} \sim \text{NID}(0, \sigma^2 \mathbf{I}_n) \quad (9)$$

Montgomery et al. (2001) ressaltam que a matriz $(\mathbf{X}'\mathbf{X})$ é uma matriz simétrica, cuja diagonal é a soma dos quadrados dos elementos da coluna de \mathbf{X} e os demais elementos são a soma dos produtos cruzados dos elementos das colunas de \mathbf{X} . O vetor $\mathbf{X}'\mathbf{y}$ é um vetor coluna,

cujos elementos são as somas dos produtos cruzados das variáveis de controle e dos valores observados da característica de qualidade.

$$(\mathbf{X}'\mathbf{X}) = \begin{bmatrix} n & \sum_{i=1}^n x_{i1} & \sum_{i=1}^n x_{i2} & \dots & \sum_{i=1}^n x_{ik} \\ \sum_{i=1}^n x_{i1} & \sum_{i=1}^n x_{i1}^2 & \sum_{i=1}^n x_{i1}x_{i2} & \dots & \sum_{i=1}^n x_{i1}x_{ik} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n x_{ik} & \sum_{i=1}^n x_{i1}x_{ik} & \sum_{i=1}^n x_{i1}x_{ik} & \dots & \sum_{i=1}^n x_{ik}^2 \end{bmatrix} \quad \mathbf{X}'\mathbf{y} = \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_{i1}y_i \\ \vdots \\ \sum_{i=1}^n x_{ik}y_i \end{bmatrix} \quad (10)$$

Os valores estimados da característica de qualidade são representados por:

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \mathbf{H}\mathbf{y} \quad (11)$$

Montgomery et al. (2001) e Weisberg (2005) comentam que a matriz \mathbf{H} é chamada de matriz chapéu (*hat*), pois transforma os valores observados \mathbf{y} em valores estimados $\hat{\mathbf{y}}$. Essa matriz é de grande importância na regressão, principalmente na detecção de informações influentes.

A diferença entre os valores observados e os valores estimados são os resíduos de regressão \mathbf{e} , que podem ser escritos em notação matricial. O vetor \mathbf{e} , apresentado na equação (12), representa um estimador do vetor dos erros aleatórios $\boldsymbol{\varepsilon}$.

$$\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}} \quad (12)$$

2.1.3.1 Propriedades dos Estimadores de MQO e do Modelo Estimado

Segundo Montgomery et al. (2001), o estimador dos mínimos quadrados ordinários $\hat{\boldsymbol{\beta}}$ é um estimador não-viciado de $\boldsymbol{\beta}$, já que:

$$\mathbf{E}(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta} \quad (13)$$

A variância desses estimadores é dada pela equação (14).

$$\mathbf{var}(\hat{\boldsymbol{\beta}}_j) = \sigma^2 C_{jj} \quad (14)$$

onde: C_{jj} é o j -ésimo elemento da diagonal da matriz $(\mathbf{X}'\mathbf{X})^{-1}$

Segundo Gujarati (2000) e Montgomery et al. (2001), com a suposição de que os resíduos são independentes, com média zero e desvio padrão σ^2 desconhecido, o teorema de Gauss-Markov assegura que $\hat{\beta}$ é um estimador não-viciado de β , pois cada um de seus elementos é função linear da característica de qualidade y . O teorema também assegura que $\hat{\beta}$ é um estimador eficiente, isto é, apresenta menor variância que qualquer outro estimador.

2.1.3.2 *Estimação da Variância do Modelo*

Para os testes de hipóteses e construção de intervalos de confiança para as estimativas do modelo é necessário estimar a variância σ^2 . Uma estimativa não-viciada é obtida pelo quadrado médio dos erros (MONTGOMERY et al., 2001; WEISBERG, 2005), apresentado na equação (15):

$$\hat{\sigma}^2 = \frac{\mathbf{e}'\mathbf{e}}{n-p} = \frac{SQR}{n-p} = QMR \quad (15)$$

2.1.4 Testes de Hipóteses e Intervalos de Confiança na Regressão Linear

Neste item, apresentam-se os testes de hipóteses e intervalos de confiança comumente utilizados na análise de regressão.

2.1.4.1 *Teste para a Significância da Regressão*

A análise de variância (ANOVA) é utilizada para o teste da significância da regressão, ou seja, para verificar a existência de relação linear entre a característica de qualidade y e um subconjunto de variáveis de controle (MONTGOMERY et al., 2001; NETER et al., 2005). O teste de significância da regressão pode ser representado pelas seguintes hipóteses:

$$\begin{aligned} H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0 \\ H_1: \beta_j \neq 0 \text{ para no mínimo um } j \end{aligned} \quad (16)$$

A rejeição da hipótese nula $H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$, implica que pelo menos uma das variáveis de controle é estatisticamente significativa para o modelo. O teste da significância da regressão é baseado na decomposição dos quadrados totais (SQT), apresentados na equação (17).

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 - \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (17)$$

$$SQT = SQR + SQReg$$

onde: $SQT = S_{yy} = \mathbf{y}'\mathbf{y} - \frac{(\sum_{i=1}^n y_i)^2}{n}$ $SQReg = \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - \frac{(\sum_{i=1}^n y_i)^2}{n}$

De acordo com Montgomery et al. (2001) e Neter et al. (2005), os graus de liberdade para a análise de variância são distribuídos da seguinte forma:

- a) A soma dos quadrados totais (SQT) possui $n-1$ graus de liberdade, já que está sujeita à restrição apresentada na equação (18).

$$\sum_{i=1}^n (y_i - \bar{y}) = 0 \quad (18)$$

- b) A soma dos quadrados da regressão (SQReg) possui k graus de liberdade, pois é definido pelos k coeficientes de inclinação do modelo;
- c) A soma dos quadrados dos erros (SQR) possui $n - p$ graus de liberdade, pois a estimativa de cada um dos p parâmetros do modelo acarreta a perda de um grau de liberdade.

Estes graus de liberdade apresentam propriedade aditiva, conforme equação (19):

$$\begin{aligned} gl_T &= gl_{Reg} + gl_R \\ n - 1 &= k + (n - p) \end{aligned} \quad (19)$$

Assim, para testar as hipóteses de (16), ou seja, testar se existe relação linear entre a característica de qualidade e algumas variáveis de controle, define-se a seguinte estatística de teste (MONTGOMERY et al., 2001; NETER et al., 2005):

$$F_0 = \frac{SQReg/gl_{Reg}}{SQR/gl_R} = \frac{SQReg/k}{SQR/(n-p)} = \frac{QMReg}{QMR} \quad (20)$$

Esta estatística de teste apresenta uma distribuição F-Snedecor com k graus de liberdade no numerador e $n - p$ graus de liberdade no denominador se a hipótese H_0 for verdadeira. Dessa forma, rejeita-se a hipótese nula $H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$ se:

$$F_0 > F_{\alpha, k, n-p} \quad (21)$$

A análise de variância descrita anteriormente pode ser resumida na Figura 1:

Fonte de Variação	Soma dos Quadrados	Graus de Liberdade	Média Quadrática	F ₀
Regressão	<i>SQReg</i>	<i>k</i>	$QMReg = SQReg/k$	$QMReg/QMR$
Resíduos	<i>SQR</i>	$n - p$	$QMR = SQR/(n - p)$	
Total	<i>SQT</i>	$n - 1$		

Figura 1: Análise de variância para testar a significância da regressão
Fonte: Adaptado de Montgomery et al. (2001)

2.1.4.2 Testes Individuais para os Coeficientes Regressão

Em alguns casos, é necessária a formulação de hipóteses e construção de intervalos de confiança para os coeficientes da regressão. Por exemplo, diante da hipótese de que algum dos coeficientes de regressão é igual a uma constante, a seguinte estrutura de hipóteses pode ser montada (MONTGOMERY e RUNGER, 2003):

$$\begin{aligned} H_0: \beta_j &= \beta_{j,0} \\ H_1: \beta_j &\neq \beta_{j,0} \end{aligned} \quad (22)$$

onde: $\beta_{j,0}$ é uma constante qualquer

Para esse teste, utiliza-se a estatística t_0 , apresentada na equação (23):

$$t_0 = \frac{\hat{\beta}_j - \beta_{j,0}}{\sqrt{\hat{\sigma}^2 C_{jj}}} \quad (23)$$

De acordo com Montgomery et al. (2001), Neter et al. (2005) e Weisberg (2005), a estatística de teste da equação (23) segue uma distribuição t-student com $n - p$ graus de liberdade se a hipótese nula H_0 é verdadeira. Dessa forma, a hipótese nula é rejeitada se:

$$|t_0| > t_{\alpha/2, n-p} \quad (24)$$

Em particular, quando a constante $\beta_{j,0}$ é igual a zero, testa-se a significância do coeficiente de regressão β_j :

$$\begin{aligned} H_0: \beta_j &= 0 \\ H_1: \beta_j &\neq 0 \end{aligned} \tag{25}$$

Caso se falhe em rejeitar a hipótese nula $H_0: \beta_j = 0$, a variável de controle x_j não é estatisticamente significativa para modelar a característica de qualidade y e poderá ser retirada do modelo (MONTGOMERY et al., 2001). Uma forma alternativa para tomar decisões sobre as hipóteses é o uso do valor p , que é definido por Montgomery e Runger (2003, p. 50) como sendo “o menor nível de significância que conduz à rejeição da hipótese nula H_0 com os dados fornecidos”. Nesse caso, quando o valor p for menor que o nível de significância adotado existem evidências estatísticas suficientes para se rejeitar H_0 .

Como os coeficientes da reta de regressão estimados ($\hat{\beta}_j$) são estimativas pontuais de β_j , é bastante útil estimar intervalos de confiança para essas estimativas. Os intervalos bilaterais com $100(1 - \alpha)\%$ de confiança para os coeficientes da regressão são apresentados na equação (26) (MONTGOMERY et al., 2001; NETER et al., 2005; WEISBERG, 2005):

$$\hat{\beta}_j - t_{\alpha/2, n-p} \sqrt{\hat{\sigma}^2 C_{jj}} \leq \beta_j \leq \hat{\beta}_j + t_{\alpha/2, n-p} \sqrt{\hat{\sigma}^2 C_{jj}} \tag{26}$$

O intervalo de confiança bilateral para β_j da equação (26) também pode ser utilizado para o teste de hipóteses em (22), já que caso o intervalo de confiança contenha a constante $\beta_{j,0}$, então aceita-se a hipótese nula $H_0: \beta_j = \beta_{j,0}$.

2.1.5 Aplicação da Regressão Linear: a Estimação de Valores

A aplicação mais comum dos modelos de regressão é a estimativa da característica de qualidade y para um conjunto de valores das variáveis de controle do modelo. Esta estimativa pode ser utilizada de duas formas: (i) estimação da resposta média e (ii) predição de novas observações. A diferença entre estas duas aplicações recai no fato de que a estimativa da resposta média é calculada com base em um conjunto de valores das variáveis de controle que foi utilizado para estimar o modelo de regressão, enquanto a previsão de novos valores é estimada através de um novo conjunto de valores para as variáveis de controle.

2.1.5.1 Estimação da Resposta Média

Dado um conjunto de valores das variáveis de controle, representados pelo vetor \mathbf{x}_i , a estimativa da resposta média de y dado este vetor, é obtida através da equação (27).

$$\hat{\mu}_{y|x_i} = \mathbf{x}'_i \hat{\boldsymbol{\beta}} \quad (27)$$

onde: $\mathbf{x}'_i = [1 \quad x_{i1} \quad x_{i2} \quad \dots \quad x_{ik}]$

De acordo com Montgomery et al. (2001) e Neter et al. (2005), a estimativa da resposta média de y para o vetor \mathbf{x}_i da equação (27) é obtida pela substituição desse valor na equação (11). Como $\hat{\boldsymbol{\beta}}$ é um estimador não-viciado, $\hat{\mu}_{y|x_i}$ também é um estimador não-viciado de $\mu_{y|x_i}$. De acordo com Neter et al. (2005), esta estimativa é normalmente distribuída, com média estimada pela equação (27) e variância:

$$\text{var}(\mu_{y|x_i}) = \text{var}(\mathbf{x}'_i \hat{\boldsymbol{\beta}}) = \hat{\sigma}^2 \mathbf{x}'_i (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_i = \hat{\sigma}^2 h_{ii} \quad (28)$$

Com a média e a variância da resposta média já estimada e diante da definição da distribuição t , Neter et al. (2005) e Weisberg (2005) fornecem o seguinte intervalo com $100(1-\alpha)\%$ de confiança para a resposta média y :

$$\hat{\mu}_{y|x_i} - t_{\alpha/2, n-p} \sqrt{\hat{\sigma}^2 h_{ii}} \leq \mu_{y|x_i} \leq \hat{\mu}_{y|x_i} + t_{\alpha/2, n-p} \sqrt{\hat{\sigma}^2 h_{ii}} \quad (29)$$

Montgomery et al. (2001) e Neter et al. (2005) chamam atenção para a largura do intervalo de confiança ser uma função do vetor \mathbf{x}_i . Essa largura é mínima quando o valor de todas as variáveis de controle são iguais à média de cada uma, aumentando à medida que se afasta da média. De acordo com Montgomery et al. (2001) isso acontece pois espera-se melhores estimativas de y realizadas com valores de \mathbf{x}_j próximos ao centro dos dados e que essa precisão piora quando se utiliza valores mais extremos no espaço de dados de \mathbf{x} .

2.1.5.2 Predição de Novas Observações

Uma importante aplicação dos modelos de regressão é a previsão de novos valores de y correspondentes a certo conjunto de valores das variáveis de controle. Esse valor pode ser estimado pela equação (30).

$$\hat{y}_j = \mathbf{x}'_j \hat{\boldsymbol{\beta}} \quad (30)$$

Esta estimativa somente é validada se for assumido que o modelo de regressão estimado pela equação (11) seja válido para o novo conjunto de dados e que essa nova observação seja independente dos dados utilizados para a estimação do modelo de regressão. Embora estas estimativas apresentem o mesmo valor, o intervalo de confiança da equação (29) é inadequado para novas observações, pois a estimativa da equação (27) se refere à resposta média verdadeira e não às futuras observações. Assim, um intervalo de confiança para as novas observações deve ser estimado (MONTGOMERY; RUNGER, 2003; NETER et al., 2005). Para estimar esse novo intervalo de confiança, Montgomery et al. (2001) definem a seguinte variável aleatória:

$$\Psi = y_j - \hat{y}_j \quad (31)$$

Essa variável é normalmente distribuída, com média zero e variância:

$$Var(\Psi) = Var(y_j - \hat{y}_j) = \hat{\sigma}^2(1 + \mathbf{x}'_j(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_j) = \hat{\sigma}^2(1 + h_{jj}) \quad (32)$$

Como \hat{y}_j está sendo utilizado para prever y_j , a variância da variável aleatória ψ estimada na equação (31) é a estatística a ser utilizada para estimar a variância de y_0 . Assim, de acordo com Montgomery et al. (2001) e Neter et al. (2005), um intervalo de previsão com $100(1 - \alpha)\%$ de confiança para uma nova observação \mathbf{x}_j é apresentado na equação (33).

$$\hat{y}_j - t_{\alpha/2, n-p} \sqrt{\hat{\sigma}^2(1 + h_{jj})} \leq y_j \leq \hat{y}_j + t_{\alpha/2, n-p} \sqrt{\hat{\sigma}^2(1 + h_{jj})} \quad (33)$$

Gujarati (2000) e Montgomery et al. (2001) aconselham cautela ao se extrapolar estimativas para novas observações cujos valores de \mathbf{x}_j estejam muito afastados do espaço de dados inicial utilizado para estimar a reta de regressão, já que é bem possível que o modelo ajuste bem na região original de dados, mas que seja pobre para um vetor \mathbf{x}_j fora dessa região.

De acordo com Montgomery et al. (2001) e Weisberg (2005), quando se utiliza um modelo de regressão linear múltipla, é bastante comum a extrapolação dos valores das variáveis de controle, já que a região de dados é definida por todos os valores das variáveis de controle. Na Figura 2 ilustra-se um exemplo em que o ponto (x_{01}, x_{02}) está entre os intervalos de dados originais de ambas as variáveis de controle, mas está fora da região original de dados utilizados para estimar o modelo de regressão. Neste caso, a estimação de uma nova observação para este ponto seria uma extrapolação escondida do modelo original de regressão.

Os elementos h_{jj} da diagonal de \mathbf{H} são úteis para detectar a extrapolação escondida, já que dependem tanto da distância euclidiana do ponto em relação ao centro da nuvem de

dados quanto da densidade de pontos da mesma. Segundo Montgomery et al. (2001), o maior elemento h_{jj} do conjunto inicial de dados, definido como h_{\max} , estará localizado no limite do elipsóide que define a nuvem de dados. Dessa forma, para o caso em que os valores das variáveis de controle das novas observações apresentem um valor h_{jj} maior que h_{\max} , realiza-se uma extrapolação.

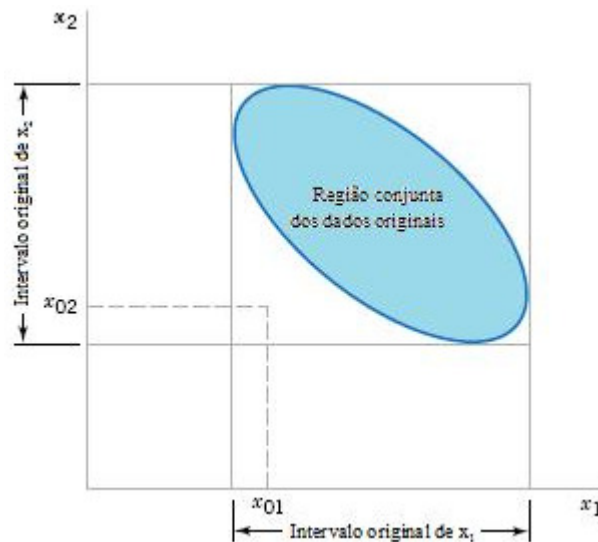


Figura 2: Exemplo de extrapolação escondida
Fonte: Montgomery et al. (2001)

2.1.6 Coeficientes de Determinação

O coeficiente de determinação é uma medida simples da variabilidade da característica de qualidade que pode ser explicada pelo modelo estimado contendo as variáveis de controle x_1, x_2, \dots, x_k (MONTGOMERY et al., 2001, WEISBERG, 2005):

$$R^2 = \frac{SQReg}{SQT} = 1 - \frac{SQR}{SQT} \quad (34)$$

onde: $0 \leq R^2 \leq 1$

De acordo com Montgomery e Runger (2003) e Weisberg (2005), a raiz quadrada positiva de R^2 é o coeficiente de regressão múltipla entre y e o conjunto de variáveis de controle usadas no modelo. Montgomery et al. (2001) comentam que a estatística R^2 deve ser utilizada com cuidado, pois é possível fazê-la maior simplesmente pela adição de variáveis de controle ao modelo, o que não indica que um modelo com mais variáveis seja necessariamente melhor que um modelo com menos variáveis de controle.

Assim, é necessário ter uma medida que leve em consideração tanto a variabilidade de y que é explicada pelo modelo quanto o número de variáveis de controle utilizado no mesmo. Para Montgomery et al. (2001) e Neter et al. (2005), o coeficiente de determinação ajustado resolve esse problema, pois penaliza a adição de novos termos que não são úteis:

$$R_{aj}^2 = 1 - \frac{n-1}{n-p}(1 - R^2) \quad (35)$$

onde: $0 \leq R_{aj}^2 \leq R^2 \leq 1$

Como é possível observar na equação (35), o acréscimo de uma variável ao modelo pode fazer com que R_{aj}^2 decresça, se o aumento de R^2 não for suficientemente grande para compensar a perda de um grau de liberdade em $n-p$ (MONTGOMERY; RUNGER, 2003).

2.1.7 Análise Residual

Para a estimação dos coeficientes de regressão pelo método dos mínimos quadrados ordinários supõe-se que os erros do modelo sejam normal e independentemente distribuídos, com média igual a zero e desvio-padrão constante. Para validar o modelo de regressão é necessário verificar se esses pressupostos acerca dos erros são atendidos e, para isso, utiliza-se a análise residual. Autores como Montgomery et al. (2001) e Neter et al. (2005) sugerem quatro técnicas para verificar essas suposições: (i) gráfico de probabilidade normal dos resíduos, (ii) gráficos dos resíduos em uma seqüência temporal, (iii) gráficos dos resíduos contra os valores preditos e (iv) gráficos dos resíduos contra as variáveis independentes.

O gráfico de probabilidade normal é a forma mais utilizada para a avaliação de que os resíduos seguem uma distribuição normal: se os pontos estiverem dispostos próximos a uma linha reta pode-se considerar que os resíduos são normalmente distribuídos (MONTGOMERY e RUNGER, 2003). Esse método é bastante subjetivo, podendo dar margens a interpretações errôneas, já que em alguns casos esse alinhamento não é evidente, o que torna a avaliação muito difícil. Para isso, existem procedimentos de testes para avaliar a normalidade como o teste de aderência qui-quadrado, estatísticas de Anderson-Darling, de Shapiro-Wilk e de Kolmogorov-Smirnov, entre outros.

Os gráficos dos resíduos contra os valores estimados da característica de qualidade ou contra os valores das variáveis de controle são utilizados para verificar se a variância dos resíduos pode ser considerada aproximadamente constante e também ajudam a verificar se a linearidade do modelo é válida (MONTGOMERY et al., 2001; NETER et al., 2005;

WEISBERG, 2005). Na Figura 3, são apresentados alguns exemplos de comportamentos típicos desses gráficos.

A Figura 3.a apresenta um exemplo de padrão aceitável, pois os resíduos aparentemente estão contidos dentro de uma faixa, o que indica que o modelo linear é apropriado e que os resíduos apresentam variância aproximadamente constante. As Figuras 3.b e 3.c indicam casos em que a variância dos resíduos não pode ser considerada como constante. Na Figura 3.d apresenta-se um caso em que o modelo linear proposto não é adequado (MONTGOMERY et al., 2001; NETER et al., 2005).

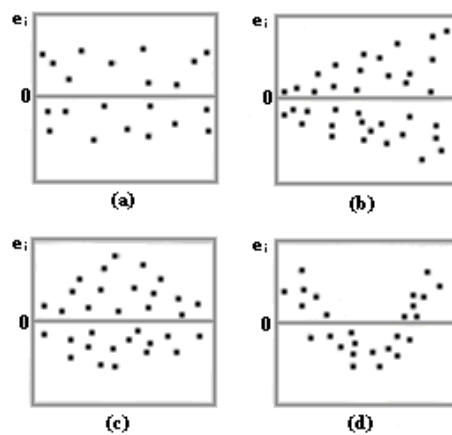


Figura 3: Padrões de comportamento de gráficos de resíduos: (a) satisfatório, (b) funil, (c) arco duplo e (d) não linear

Fonte: Montgomery e Runger (2003)

Para estabilizar a variância e , por conseqüência, eliminar os problemas presentes nas Figuras 3.b e 3.c, Montgomery e Runger (2003) sugerem a utilização de transformações lineares da característica de qualidade. As transformações mais usuais são: $\ln y$, \sqrt{y} ou $1/y$. Para resolver a não linearidade da Figura 3.d pode ser necessária a adição de outros termos de ordem superior ao modelo, como, por exemplo, um termo quadrático. Os autores também apontam transformações lineares nas variáveis do modelo ou a adição de novas variáveis como solução para a não-linearidade da Figura 3.d. Uma outra solução é a utilização dos mínimos quadrados ponderados, tendo o inverso da variância como pesos.

Caso seja conhecida a ordem de coleta dos dados, pode-se plotar um gráfico dos resíduos em ordem seqüencial, o que permite verificar se a variável tempo interfere na variância do modelo. Um padrão como a Figura 3.a indica que a variância do modelo não é afetada pela ordem de coleta; já padrões como Figura 3.b, Figura 3.c e Figura 3.d indicam interferência do tempo na variância do modelo (MONTGOMERY et al., 2001). Para testar a

autocorrelação dos resíduos do modelo pode ser necessária a utilização de testes, como o teste de Durbin-Watson, por exemplo. Segundo Montgomery et al. (2005), as hipóteses a serem testadas no teste de Durbin-Watson são:

$$\begin{aligned} H_0: \rho &= 0 \\ H_1: \rho &\neq 0 \end{aligned} \quad (36)$$

No entanto, o teste bilateral da equação (36) não pode ser realizado diretamente, devendo ser dividido em dois testes diferentes: (i) teste para a autocorrelação positiva e (ii) teste para a autocorrelação negativa. Dessa forma, definem-se as estatísticas d_p e d_n para os testes (i) e (ii), respectivamente:

$$d_p = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2} \quad (37)$$

$$d_n = 4 - d_p \quad (38)$$

As decisões de cada teste são feitas com base em duas constantes tabeladas, chamadas de d_l e d_u . No teste de Durbin-Watson podem haver três resultados diferentes:

- a) A autocorrelação dos resíduos é significativa – isto ocorre se a estatística de teste for menor que d_l ;
- b) Os resíduos não são autocorrelacionados: neste caso, a estatística de teste é maior que d_u ;
- c) O teste de Durbin-Watson é inconclusivo: isto ocorre se a estatística de teste ficar em um valor intermediário entre d_l e d_u .

2.1.7.1 *Análise de Outliers*

Os *outliers* são valores extremos das observações, ou seja, valores que estão, de alguma forma, separados do restante do conjunto de dados e que podem alterar significativamente as estimativas dos coeficientes de regressão. Os *outliers* podem indicar um comportamento atípico de alguma das variáveis do processo, uma falha de medição ou o efeito de alguma variável não incluída no modelo (MONTGOMERY et al., 2001; NETER et al., 2005). Na Figura 4, são apresentados alguns exemplos de valores extremos.

De acordo com Neter et al. (2001) os pontos 1-4 da Figura 4 são valores extremos. O ponto 1, embora esteja fora da nuvem de dados, provavelmente não é um *outlier*, já que seu valor de y não é muito superior à média da resposta dado o valor da variável de controle. O

ponto 2, mesmo estando fora do intervalo de valores de y e de x , não é um *outlier*, pois está localizado próximo à extensão da reta de regressão. Os pontos 3 e 4, por estarem muito distantes do conjunto de dados, podem ser considerados *outliers*.

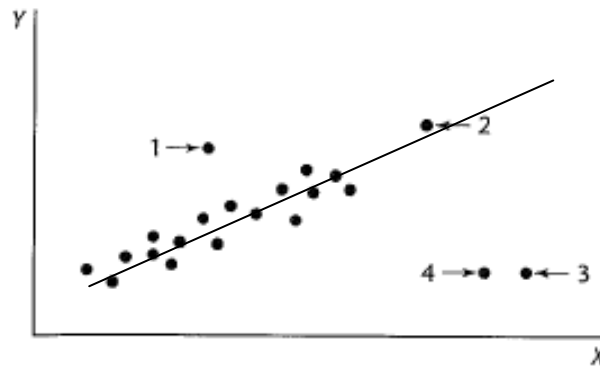


Figura 4: Gráfico de dispersão para um modelo de regressão linear simples ilustrando alguns casos de *outliers*
Fonte: Adaptado de Neter et al. (2005)

Weisberg (2005) afirma que as observações com grandes resíduos são candidatas a serem *outliers*, mas a avaliação das observações não pode ser feita de forma subjetiva, sendo necessária alguma forma de dimensionar os resíduos e assim avaliar quais resíduos são verdadeiramente *outliers*. Assim, de acordo com Montgomery et al. (2005) e Neter et al. (2005), o método mais simples e lógico é a padronização dos resíduos utilizando a variância do modelo, estimada pelo QMR, conforme equação (39).

$$d_i = \frac{e_i}{\sqrt{\hat{\sigma}_2}} = \frac{e_i}{\sqrt{QMR}} \quad (39)$$

Segundo Montgomery et al. (2005), os resíduos padronizados calculados pela equação (39) possuem média zero e desvio-padrão aproximadamente unitário. Os autores recomendam considerar como sendo *outliers* as observações que apresentam resíduos padronizados maiores que 3.

O uso do QMR para estimar a variância do i -ésimo resíduo e_i é apenas uma aproximação. É possível aprimorar o dimensionamento dos resíduos usando o desvio-padrão exato do i -ésimo resíduo. Com a utilização da matriz chapéu (\mathbf{H}), apresentada na equação (11), pode-se reescrever o vetor de resíduos (MONTGOMERY et al., 2001; NETER et al., 2005; WEISBERG, 2005):

$$\mathbf{e} = (\mathbf{I} - \mathbf{H})\mathbf{y} \quad (40)$$

A variância do i -ésimo resíduo é dada pela equação (41).

$$\text{var}(e_i) = \hat{\sigma}^2(1 - h_{ii}) \quad (41)$$

onde $0 \leq h_{ii} < 1$

Assim, um procedimento lógico é dimensionar o i -ésimo resíduo utilizando a estimativa da variância fornecida pela equação (41). Essa padronização, apresentada na equação (42), também pode ser chamada de padronização na forma *student*, possuindo média 0 e desvio-padrão constante igual a 1.

$$r_i = \frac{e_i}{\sqrt{\hat{\sigma}^2(1 - h_{ii})}} \quad (42)$$

2.1.8 Observações Influentes

Ao se utilizar a regressão múltipla, é bastante comum encontrar um ponto ou um subconjunto de dados que estão relativamente longe da vizinhança do restante dos dados. Esse subconjunto de informações pode influenciar fortemente a estimativa dos coeficientes de regressão (MONTGOMERY et al., 2001; WEISBERG, 2005).

Basicamente, existem dois tipos de pontos de alavanca: os pontos de alavanca e os pontos de influentes. Na Figura 4, o ponto 2 é um exemplo de um ponto de alavanca, pois está localizado na extensão da reta de regressão e não afeta as estimativas dos coeficientes de regressão, mas altera fortemente algumas estatísticas da regressão, como o R^2 e os desvios-padrão dos coeficientes de regressão. Os pontos 3 e 4 da Figura 4, apresentam valores anormais para x e y , sendo chamados de pontos influentes, pois têm grande impacto na estimação dos parâmetros da regressão (MONTGOMERY et al., 2001; NETER et al., 2005).

Assim, há grande interesse em identificar pontos de alavanca e pontos influentes dentre o conjunto inicial de dados. Montgomery et al. (2001) apresenta um método de identificação de pontos de alavanca, com o uso de um ponto de corte, apresentado na equação (43).

$$c = \frac{2p}{n} \quad (43)$$

Caso a i -ésima amostra apresente um h_{ii} maior que o ponto de corte c , é remoto o bastante para ser considerado um ponto de alavanca. Montgomery et al. (2001) consideram que pontos de alavanca que também apresentem resíduos grandes são pontos influentes.

Ressalta-se que esse método só é válido quando $2p/n \leq 1$. Para a identificação de pontos influentes, a técnica mais utilizada é a distância de Cook.

2.1.8.1 Distância de Cook

A distância de Cook, proposta por Cook (1977), é uma técnica para a medida de influência de cada um dos pontos de um conjunto de dados. De acordo com Montgomery et al. (2001) e Weisberg (2005), a idéia básica dessa técnica é comparar a estimativa dos mínimos quadrados dos coeficientes de regressão contendo todas as n observações com a estimativa destes coeficientes obtida através da remoção do i -ésimo ponto, representada por $\hat{\beta}_{(i)}$. A distância de Cook do i -ésimo ponto é apresentada na equação (44) (COOK, 1977).

$$D_i = \frac{(\hat{\beta}_{(i)} - \hat{\beta})' \mathbf{X}'\mathbf{X} (\hat{\beta}_{(i)} - \hat{\beta})}{p \hat{\sigma}^2} = \frac{r_i^2 h_{ii}}{p(1 - h_{ii})} \quad (44)$$

onde: $\hat{\beta}_{(i)}$ é a estimativa de β , desconsiderando-se o i -ésimo ponto

Cook (1977) comenta que se $D_i > F(0,5; p; n-p)$ há indícios de que o i -ésimo ponto exerce influência sobre a estimativa dos coeficientes de regressão. Como este valor aproxima-se de 1 à medida que aumentam os graus de liberdade da distribuição F, Montgomery et al. (2001) recomenda que pontos com $D_i > 1$ sejam considerados influentes.

2.1.8.2 DFFITS e DFBETAS

Belsley et al. (1980) apresentam formas úteis para a medida de influência de alguns pontos sobre a estimativa dos coeficientes de regressão e da característica de qualidade. A primeira medida é o DFBETAS, que mede, em desvios-padrão, o quanto muda a estimativa do j -ésimo coeficiente de regressão ao se omitir a i -ésima observação:

$$DFBETAS_{j,i} = \frac{\hat{\beta}_j - \hat{\beta}_{j(i)}}{\sqrt{s_{(i)}^2 C_{jj}}} \quad (45)$$

onde: $\hat{\beta}_{j(i)}$ é a estimativa de β_j desconsiderando-se a i -ésima observação

$s_{(i)}^2$ é a estimativa de variância do modelo desconsiderando a i -ésima observação

Ainda segundo os autores, a i -ésima observação é influente sobre a estimativa do j -ésimo coeficiente quando:

$$|DFBETAS_{j,i}| > \frac{2}{\sqrt{n}} \quad (46)$$

A outra medida de influência proposta por Belsley et al. (1980) é o DFFITS, que mede, em desvios-padrão, a estimativa da i -ésima observação quando esta não é considerada na estimativa do modelo de regressão:

$$DFFITS_i = \frac{\hat{y}_i - \hat{y}_{(i)}}{\sqrt{s_{(i)}^2 h_{ii}}} \quad (47)$$

onde: $\hat{y}_{(i)}$ é a estimativa da i -ésima observação, quando esta não é considerada na estimativa do modelo.

Neste caso, a i -ésima observação é influente sobre a estimativa quando:

$$|DFFITS_i| > \frac{2}{\sqrt{p/n}} \quad (48)$$

2.1.9 Seleção de Variáveis e de Modelos

Uma das maiores dificuldades ao se trabalhar com a análise de regressão é selecionar um conjunto de variáveis de controle para serem utilizadas no modelo. Em alguns casos, considerações teóricas e a experiência prévia podem ajudar tanto a selecionar as variáveis de controle quanto as transformações das variáveis e as interações que podem vir a integrar o modelo. No entanto, em alguns casos práticos o maior problema consiste em selecionar um conjunto de variáveis dentre todas as variáveis consideradas como sendo importantes, pois nem todas as variáveis candidatas são realmente necessárias para modelar a característica de qualidade (MONTGOMERY et al., 2001; MONTGOMERY; RUNGER, 2003).

Uma das formas de escolha de variáveis apontadas por Montgomery et al. (2001) é o uso de gráfico de dispersão entre a característica da qualidade e cada uma das variáveis de controle. O gráfico de dispersão também indica a forma de relação entre as variáveis, como por exemplo, uma relação linear entre variáveis é indicada por um repouso aleatório de pontos em torno de uma reta e um posicionamento aleatório de pontos em torno de uma parábola é uma indicação de uma relação quadrática.

Essa análise é bastante subjetiva e, de acordo com Montgomery et al. (2001), pode se tornar errônea à medida que aumentam os números de variáveis do modelo, principalmente

quando algumas variáveis de controle são correlacionadas entre si. Também existem métodos analíticos para a escolha das variáveis de controle significativas para o modelo, dentre os quais se destacam os métodos de todas as regressões possíveis e o método passo-a-passo.

De acordo com Montgomery et al. (2001), esses dois métodos de escolha do melhor conjunto de variáveis de controle que devem integrar o modelo lidam com dois objetivos conflitantes: (i) deseja-se escolher o maior número possível de variáveis para o modelo de forma a maximizar o coeficiente de determinação e (ii) deseja-se escolher o menor número de variáveis, pois a variância da previsão aumenta com o incremento de variáveis no modelo. A seguir, serão mostradas as duas técnicas mais utilizadas para a seleção de variáveis e alguns critérios adotados para isto.

2.1.9.1 *Todas as Regressões Possíveis*

Nessa abordagem, são estimados todos os modelos de regressão contendo as combinações de uma, duas e mais variáveis do conjunto de variáveis importantes. Assumindo que o coeficiente de intercepto está presente em todos os modelos e considerando k o número de variáveis candidatas, existirão 2^k modelos possíveis de serem estimados, o que faz com que o número de modelos aumente rapidamente com o aumento do número de variáveis de controle candidatas. Para a escolha dos melhores modelos existem vários critérios, que serão mostrados a seguir. (MONTGOMERY et al., 2001; NETER et al., 2005; WEISBERG, 2005).

O primeiro, e mais lógico critério, é a escolha do modelo que maximiza o coeficiente de determinação, estimado pela equação (34). Como afirmado anteriormente, esse critério deve ser utilizado com cuidado, pois a adição de novas variáveis ao modelo acarreta um aumento do R^2 , sem de fato melhorar o modelo. Um outro critério lógico é a escolha do modelo que apresente menor estimativa para os erros, ou seja, deve-se escolher o modelo que minimize QMR.

Também é possível utilizar o coeficiente de determinação ajustado, estimado pela equação (35). Essa estatística é uma modificação do coeficiente de determinação, que considera também o número de variáveis do modelo, penalizando modelos que possuam mais variáveis. O segundo critério seria a escolha do modelo que apresente esses dois coeficientes aproximadamente iguais. Outra alternativa é a utilização da estatística C_p , que é uma medida da média quadrática dos erros (MALLOWS, 1973; 1995):

$$C_p = \frac{SQR_p}{\hat{\sigma}^2} - n + 2p \quad (49)$$

Segundo Mallows (1995), se o modelo de regressão da equação (11) for correto e se a estimativa de σ^2 for não-tendenciosa, a estatística C_p tem a propriedade de ser igual ao número de parâmetros do modelo. Logo, deve-se adotar como critério de escolha a minimização do C_p para valores aproximadamente iguais a p . Para facilitar a escolha do modelo, também é útil a construção de gráficos C_p versus p .

É importante lembrar que o SQR_p diminui à medida que se aumenta o número de parâmetros p do modelo, mas esse aumento do número de parâmetros também aumenta a complexidade da equação a ser estimada, o que pode vir a dificultar o controle do fenômeno estudado e o ajuste de um ponto ótimo de operação. Dessa forma, os critérios para escolha de modelos devem balancear a minimização do SQR_p com o aumento da complexidade da equação.

2.1.9.2 Método Passo-a-passo

Quando existir um número muito grande de variáveis candidatas ao modelo ($n > 30$ ou 40), a análise do modelo de regressão pode tornar-se muito complexa por parte do usuário. Dessa forma, sugere-se a utilização de algum método que pré-seleciona as variáveis mais importantes para caracterizar a característica de qualidade. Segundo Montgomery et al. (2001) e Neter et al. (2005), existem três categorias diferentes do método passo-a-passo: (i) seleção progressiva, (ii) eliminação regressiva e (iii) regressão por etapas.

Antes da apresentação dessas três categorias, apresenta-se a estatística F parcial, que é o valor do teste F para a variável x_i dado que as variáveis x_j ($j \neq i$) já estão no modelo:

$$F_i = \frac{SQReg(\beta_i | \beta_0, \beta_j)}{QMR(x_j, x_i)} \quad (50)$$

Também é necessário definir as variáveis aleatórias f_{entra} e f_{sai} , que são os valores críticos da variável aleatória F para adicionar e retirar uma variável do modelo, respectivamente.

Resumidamente, a seleção progressiva consiste na adição gradual de variáveis de controle ao modelo. Esse método inicia-se com o modelo com apenas o parâmetro de intercepto e, a seguir, seleciona-se a variável de controle que apresenta maior coeficiente de correlação simples com a característica de qualidade. Caso o F parcial para essa variável seja

maior que f_{entra} , deve-se adicionar essa variável ao modelo. A cada etapa deste método, seleciona-se a variável que apresente maior correlação com a característica de qualidade, essa é adicionada desde que o valor do teste parcial F seja maior que f_{entra} . O procedimento finaliza assim que nenhuma variável puder ser adicionada ao modelo (MONTGOMERY et al., 2001; NETER et al., 2005).

A eliminação regressiva consiste na remoção gradual de variáveis de controle do modelo que contenha todas as variáveis candidatas. Como primeiro passo deve-se calcular o valor F parcial para todas as variáveis como se cada uma delas fosse a última a entrar no modelo. A seguir, seleciona-se a variável que possua o menor valor parcial F, se esse valor for menor que f_{sai} , deve-se retirar essa variável do modelo. A cada etapa desse método, seleciona-se a variável que apresentar menor valor parcial F, essas são removidas sempre que o valor do teste parcial F for menor que f_{sai} . Quando não existirem variáveis a serem excluídas do modelo, o método se finaliza (MONTGOMERY et al., 2001; NETER et al., 2005).

Segundo Montgomery et al. (2001) e Neter et al. (2005), a regressão por etapas é a combinação da seleção progressiva e da eliminação regressiva. Inicia-se com o modelo com nenhuma variável e a cada etapa escolhe-se a variável com maior correlação com a característica de qualidade, que deverá ser adicionada ao modelo se $F_i > f_{\text{entra}}$. Caso a variável x_i seja inserida no modelo, analisa-se a possibilidade de retirada das outras variáveis já selecionadas dado que x_i está presente. O procedimento termina assim que não existir nenhuma variável para ser incluída ou removida do modelo.

2.1.10 Multicolinearidade

Uma das suposições para o modelo de regressão é a independência entre as variáveis de controle do modelo de regressão, já que quando a correlação entre as variáveis é significativa, as inferências para o modelo de regressão são errôneas. Para o caso em que existem dependências aproximadamente lineares entre as variáveis de controle do modelo, diz-se que existe a presença de multicolinearidade (GUJARATI, 2000; MONTGOMERY et al., 2001; NETER et al., 2005).

Segundo Montgomery et al. (2001) e Neter et al. (2005), a multicolinearidade afeta principalmente as estimativas da variância e da covariância dos parâmetros de regressão, pois tende a inflacionar os termos da matriz $(\mathbf{X}'\mathbf{X})^{-1}$, o que pode acarretar a não-aceitação de significância de algumas variáveis de controle, mesmo quando a característica de qualidade

for linearmente relacionada a essas variáveis. Assim, a multicolinearidade pode prejudicar os resultados dos testes individuais para os coeficientes de regressão, apresentados anteriormente.

Montgomery e Runger (2003) e Neter et al. (2005) apontam quatro situações que indicam a presença de multicolinearidade em um modelo de regressão:

- a) Grandes alterações nas estimativas dos parâmetros do modelo quando uma variável é adicionada ou removida do modelo;
- b) O resultado do teste F de significância do modelo de regressão indica que a regressão é estatisticamente significativa, mas nenhum dos testes t indica significância dos coeficientes de inclinação;
- c) Os coeficientes de regressão apresentam sinais algébricos contrários ao esperados em experimentos anteriores ou conhecimentos teóricos;
- d) Variáveis teoricamente importantes para o modelo não são consideradas estatisticamente significantes no modelo estimado.

Gujarati (2000) e Montgomery et al. (2005) recomendam duas técnicas principais para o diagnóstico da multicolinearidade: (i) a análise da matriz de correlação e (ii) o fator de inflação da variância (FIV). A primeira técnica consiste na inspeção dos elementos r_{ij} que não pertencem à diagonal da matriz de correlação ($\mathbf{X}'\mathbf{X}$). Se as variáveis de controle x_i e x_j forem linearmente independentes, então todos os r_{ij} serão aproximadamente igual a 0.

Definindo R_j como sendo o coeficiente de determinação múltiplo de um modelo de regressão de x_j sobre as demais $k-1$ variáveis de controle, define-se, na equação (51), o FIV_j (GUJARATI, 2000; MONTGOMERY et al., 2001; NETER et al., 2005):

$$FIV_j = C_{jj} = \frac{1}{1 - R_j^2} \quad (51)$$

Na verdade, FIV_j é o elemento C_{jj} da diagonal da matriz $(\mathbf{X}'\mathbf{X})^{-1}$ e indica quantas vezes a estimativa da variância do parâmetro β_j está sendo inflacionada. Se o FIV_j for aproximadamente igual a 1, a j -ésima variável de controle é independente das demais variáveis de controle. Autores como Gujarati (2000), Montgomery et al. (2001) e Neter et al. (2005) recomendam adotar como limite máximo do FIV valores entre 5 e 10.

Para a solução da multicolinearidade, Gujarati (2000), Montgomery e Runger (2003) e Montgomery et al. (2001) propõem algumas soluções: coleta de novos dados, remoção de algumas variáveis de controle do modelo, combinações de algumas variáveis (divisões, multiplicações, somas, entre outras) e a utilização da regressão corrigida. Outra forma de

combater a multicolinearidade é o uso de projetos de experimentos, já que envolvem um planejamento de coleta de dados que garante a ortogonalidade das variáveis de controle.

2.2 CONTROLE ESTATÍSTICO DE PROCESSOS

O início formal do controle estatístico de processos ocorreu em 1924, quando Walter A. Shewhart desenvolveu e aplicou os gráficos de controle na Bell Telephones Laboratories. A base teórica do CEP somente foi apresentada sete anos depois, quando Shewhart publicou o livro *Economic Control of Quality of Manufactured Product* (MONTGOMERY, 2004; COSTA et al., 2005).

De acordo com Montgomery (2004) e Costa et al. (2005), o grande mérito de Shewhart foi perceber que qualquer processo produtivo, independentemente de quão bem planejado ou controlado este seja, possui uma componente de variabilidade inerente ou natural que está sempre presente. Segundo Woodall e Montgomery (1999), esta variabilidade é causada pelas variáveis não-controláveis do processo, sendo comumente chamadas de causas comuns ou aleatórias e que não podem ser facilmente removidas do processo.

O processo também pode estar submetido a outros tipos de causas, que aumentam significativamente a variabilidade do processo, podendo causar até o deslocamento da média do processo para valores inaceitáveis para o padrão normal de funcionamento do processo. Essas causas são chamadas de causas especiais ou atribuíveis e ocorrem devido a alguma falha de execução ou configuração das variáveis de controle do processo. As causas especiais sempre podem ser retiradas do processo e são oriundas de três fontes principais: matéria-prima, operador e máquinas (MONTGOMERY, 2004; COSTA et al., 2005).

Como a variabilidade do processo é, basicamente, a diferença existente entre as unidades produzidas, uma forma de monitorá-la é o acompanhamento dos valores das características de qualidade ao longo do tempo. Essa é a lógica por trás dos gráficos de controle, que cumprem, de maneira simples e efetiva, essa função e permitem a identificação do momento no qual o processo está sob ação de causas especiais, apontando se é necessário intervir no processo (MONTGOMERY, 2004; COSTA et al., 2005). Uma representação genérica de um gráfico de controle é apresentada Figura 5.

Montgomery (2004) define, de maneira simplificada, o gráfico de controle como sendo uma representação gráfica de uma característica de qualidade medida ou calculada *versus* o número da amostra ou o tempo. O gráfico de controle, apresentado na Figura 5,

possui uma linha central (LC), que representa o valor médio da característica de qualidade, e duas linhas horizontais paralelas à linha central, chamadas de limite superior de controle (LSC) e limite inferior de controle (LIC).

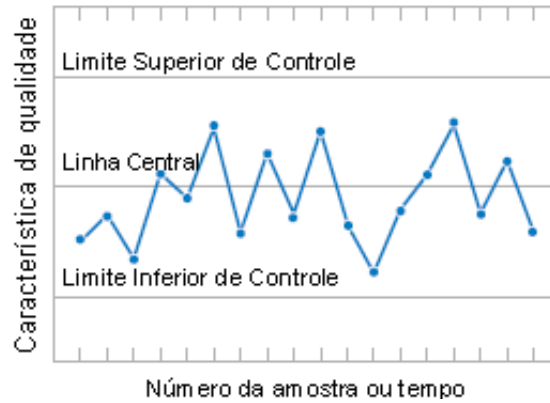


Figura 5: Exemplo de um típico gráfico de Controle
Fonte: Adaptado de Montgomery (2004)

Pontos localizados acima da LSC ou abaixo da LIC são chamados de pontos fora de controle e são a evidência mais clássica de presença de causas especiais no processo. Outros comportamentos que também evidenciam que o processo está fora de controle serão apresentados na próxima subseção. De acordo com Woodall e Montgomery (1999), quando o gráfico de controle aponta a possível ocorrência de uma causa especial, é necessária a execução de um procedimento para a investigação e ação sobre a causa especial que está perturbando o processo.

Segundo Montgomery (2004) e Costa et al. (2005), os gráficos de controle são classificados em dois tipos principais: os gráficos de controle para variáveis, quando a característica de qualidade é uma variável contínua resultante de uma medição e os gráficos de controle para atributos, quando a característica de qualidade é uma avaliação, como por exemplo, conforme ou não-conforme.

2.2.1 Gráficos de Controle para Variáveis

Uma característica de qualidade que é medida em uma escala numérica é chamada de variável, como por exemplo, tamanho, largura, diâmetro, temperatura e volume. Nesses casos,

segundo Montgomery (2004) e Costa et al. (2005), os gráficos de controle mais usuais são: gráficos \bar{X} , gráficos R e S e gráficos para medidas individuais.

O gráfico \bar{X} utiliza as médias amostrais para monitorar a média, que é uma medida de localização do processo. Os gráficos R e S utilizam, respectivamente, as amplitudes e os desvios-padrão amostrais para monitorar a variabilidade do processo. Montgomery (2004) desaconselha a utilização da amplitude como estimativa de variabilidade do processo para os casos em que o tamanho da amostra é maior que 10 ou quando trabalha-se com amostras de tamanhos variáveis. Na Figura 6 são apresentados os limites de controle para os gráficos \bar{X} , gráficos R e S:

Gráfico	Limite Inferior de Controle	Linha Central	Limite Superior de Controle
\bar{X}	$\bar{\bar{X}} - 3 \frac{\hat{\sigma}}{\sqrt{n}}$	$\bar{\bar{X}}$	$\bar{\bar{X}} + 3 \frac{\hat{\sigma}}{\sqrt{n}}$
R	$D_3 \bar{R}$	\bar{R}	$D_4 \bar{R}$
S	$B_3 \bar{S}$	\bar{S}	$B_4 \bar{S}$

Figura 6: Limites de controle para os gráficos \bar{X} - R e \bar{X} - S

Fonte: Figura elaborada pelo autor (2009)

Lembrando-se que as constantes utilizadas nas equações apresentadas na Figura 6 são dependentes apenas do tamanho da amostra e que para o gráfico das médias, o desvio padrão pode ser estimado por R/d_2 ou por s/c_4 , caso se use os gráficos R e S para monitorar a variabilidade, respectivamente. Para as situações em que a amostra consiste de uma única unidade, pode-se utilizar o gráfico para medidas individuais para monitorar a média do processo e a amplitude móvel para monitorar a variabilidade do processo. Esse gráfico é chamado de X-MR, cujos limites de controle são apresentados na Figura 7 (MONTGOMERY, 2004).

Gráfico	Limite Inferior de Controle	Linha Central	Limite Superior de Controle
X	$\bar{\bar{X}} - E_2 \bar{MR}$	$\bar{\bar{X}}$	$\bar{\bar{X}} + E_2 \bar{MR}$
MR	$D_3 \bar{MR}$	\bar{MR}	$D_4 \bar{MR}$

Figura 7: Limites de controle para os gráficos X-MR

Fonte: Figura elaborada pelo autor (2009)

2.2.2 Gráficos de Controle CUSUM e EWMA

Embora sejam ferramentas simples e efetivas para o controle da maioria dos processos, os gráficos de controle clássicos apresentados anteriormente são relativamente insensíveis para detectar pequenas mudanças no processo, da ordem de $1,5\sigma$. Duas alternativas são indicadas por Montgomery (2004) e Costa et al. (2005) para o monitoramento de processos sujeitos a pequenas perturbações: o gráfico de soma cumulativa (CUSUM) e o gráfico de médias móveis exponencialmente ponderadas (EWMA).

2.2.2.1 Gráfico para a Soma Cumulativa (CUSUM)

O gráfico CUSUM, proposto inicialmente por Page em 1954, plota as somas acumuladas dos desvios de cada amostra em relação ao valor-alvo da estatística do processo que está sendo monitorada. Essa característica de combinar as informações de todas as amostras anteriores faz com que o gráfico CUSUM seja mais eficaz que os gráficos de controle tradicionais na detecção de pequenas mudanças no processo (MONTGOMERY, 2004).

Existem gráficos CUSUM para diversas estatísticas amostrais, como amplitude, desvio-padrão, fração de não-conformes, número de defeitos e para a média, sendo que esta última é a aplicação mais comum para esse tipo de gráfico de controle. Considerando μ_0 como o valor alvo do processo, o gráfico é construído plotando-se a quantidade S_i , calculada pela equação (52), *versus* o número da amostra (MONTGOMEY, 2004; COSTA et al., 2005):

$$S_i = \sum_{i=1}^n (x_i - \mu_0) \quad (52)$$

Segundo estes autores, caso a média do processo esteja centrada no alvo, a estatística S_i oscilará aleatoriamente em torno de zero, pois os desvios positivos de S_i serão compensados pelos desvios negativos. Mas, caso a média do processo não seja realmente igual ao valor alvo μ_0 , a estatística S_i crescerá ou decrescerá, dependendo se a média do processo for, respectivamente, maior ou menor que o valor-alvo.

O gráfico CUSUM é usualmente utilizado para observações individuais, mas no caso de amostras de tamanhos maiores, basta-se substituir x_j pela média de cada amostra j . Existem dois procedimentos que auxiliam na interpretação de um gráfico CUSUM: o CUSUM tabular

ou algorítmico e a máscara V. Como o uso da máscara é desaconselhado por Montgomery (2004), este não será apresentado no presente trabalho.

O CUSUM tabular consiste no acúmulo dos desvios acima do alvo pela estatística S^+ , chamada de CUSUM unilateral superior, e dos desvios abaixo do alvo pela estatística S^- , chamada de CUSUM unilateral inferior (GAN, 1991; MONTGOMERY, 2004; COSTA et al., 2005). Estas estatísticas são apresentadas nas equações (53) e (54), respectivamente.

$$S_i^+ = \max [0, x_i - (\mu_0 + K) + S_{i-1}^+] \quad (53)$$

$$S_i^- = \max [0, (\mu_0 + K) - x_i + S_{i-1}^-] \quad (54)$$

Montgomery (2004) e Costa et al. (2005) comentam que, nas equações (53) e (54), K é chamado de valor de tolerância, sendo escolhido como a metade da magnitude de mudança que se pretende detectar com o gráfico CUSUM. Caso S_i^+ ou S_i^- excedam o valor de uma constante de decisão H , que é pré-escolhida, o processo é considerado fora de controle. No caso do CUSUM tabular, para determinar o instante em que o processo saiu fora de controle deve-se contar para trás a partir do primeiro valor que excede o valor H até o período em que o CUSUM deixou de ser nulo.

A escolha dos valores de K e H são de extrema importância para a efetividade do CUSUM e dependem da magnitude de mudança que se planeja detectar e do número de alarmes falsos teórico desejado ao gráfico. Definindo $H = h\sigma$ e $K = k\sigma$ e tendo como magnitude 1σ como a magnitude da mudança que se deseja detectar, Montgomery (2004) aconselha usar $h = 4$ ou 5 e $k = 0,5$. Para essa mesma magnitude de mudança, Gan (1991) aponta $k = 0,5$ e $h = 4,39$ como os melhores valores. Recomenda-se o trabalho de Vance (1986) e Gan (1991) para o planejamento de um gráfico de controle CUSUM, mediante a definição de outros valores para essas constantes.

Como exemplo, apresenta-se na Figura 8 um gráfico de controle CUSUM para um processo com alvo $\mu_0 = 0$ e $\sigma = 1$, onde utilizou-se: $k = 0,5$ e $h = 4$. Como é possível observar na Figura 8, a partir da amostra 28 os pontos do CUSUM superior excederam o valor H , assim o processo é considerado como estando fora de controle a partir da amostra 23, que é a primeira amostra com valor de ponto não nulo antes da amostra 28.

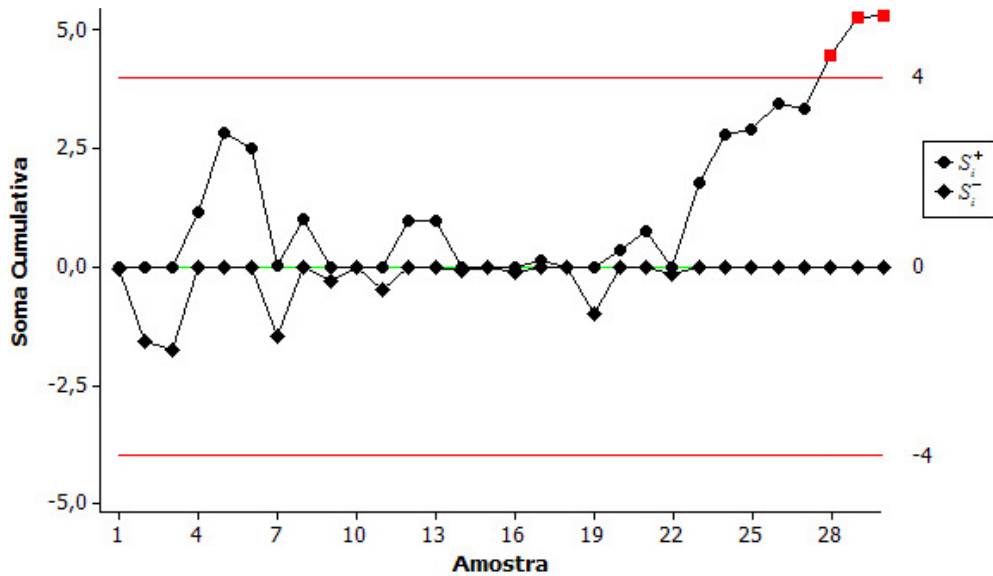


Figura 8: Exemplo de um gráfico de controle CUSUM
 Fonte: Adaptado de Montgomery (2004)

2.2.2.2 Gráfico para Médias Móveis Exponencialmente Ponderadas (EWMA)

O gráfico EWMA, proposto por Roberts em 1959, é outra alternativa quando o interesse é detectar pequenas mudanças no processo. Este gráfico apresenta desempenho bastante similar ao CUSUM e também é usualmente utilizado para observações individuais. O gráfico é baseado na seguinte estatística (MONTGOMERY, 2004; COSTA et al., 2005):

$$Z_i = \lambda Y_i + (1 - \lambda)Z_{i-1} \quad (55)$$

onde: $0 < \lambda \leq 1$

O valor inicial Z_0 para a equação (55) é usualmente escolhido como sendo 0. De acordo com Montgomery (2004), a constante de alisamento λ é o peso associado à informação mais recente, fazendo com que o gráfico EWMA seja uma ponderação entre todas as informações passadas e a atual. Fazendo $\lambda = 1$, o gráfico EWMA se reduz aos gráficos de Shewhart, pois todo o peso é dado somente à informação mais recente, tirando a capacidade de memória dos valores anteriores. Os limites de controle do gráfico EWMA são apresentados a seguir (LUCAS e SACUCCI, 1990; MONTGOMERY, 2004; COSTA et al., 2005):

$$LSC_i = \mu_0 + L\sigma \sqrt{\frac{\lambda}{2 - \lambda} [1 - (1 - \lambda)^{2i}]} \quad (56)$$

$$LC = \mu_0 \quad (57)$$

$$LIC_i = \mu_0 - L\sigma \sqrt{\frac{\lambda}{2-\lambda} [1 - (1-\lambda)^{2i}]} \quad (58)$$

À medida que i cresce, ou seja, o número das amostras aumenta, o termo $(1-\lambda)^{2i}$ se aproxima de 0 e os limites de controle das equações (56) e (58) podem ser aproximados de acordo com as equações (59) e (60):

$$LSC = \mu_0 + L\sigma \sqrt{\frac{\lambda}{2-\lambda}} \quad (59)$$

$$LIC = \mu_0 - L\sigma \sqrt{\frac{\lambda}{2-\lambda}} \quad (60)$$

A eficiência do gráfico EWMA está intimamente ligada aos valores escolhidos para as constantes λ e L . Alguns trabalhos, como Crowder (1987; 1989), Lucas e Saccucci (1990) e Jones (2002) buscam otimizar o desempenho do gráfico EWMA através das melhores combinações das constantes λ e L . De forma geral, os parâmetros são escolhidos com base na magnitude de mudança que se deseja detectar, no número médio máximo de falsos alarmes e no número de amostras necessárias para se detectar a mudança escolhida (NMA), chamado também de comprimento médio de seqüência (CMS).

De acordo com Montgomery (2004), os valores de λ mais utilizados estão dentro do intervalo $[0,05; 0,25]$. No geral, menores valores de λ são utilizados para se detectar menores mudanças no processo, já que estes valores atribuem maior peso às informações do passado. A constante L é o afastamento dos limites de controle em relação à linha central.

Na Figura 9 é apresentado um exemplo de um gráfico de controle EWMA plotado com os mesmos dados utilizados para o gráfico CUSUM da Figura 8. Para esse gráfico EWMA, adotou-se $\lambda = 0,1$ e $L = 2,7$. Como é possível observar na Figura 9, a partir da amostra 29 o processo está fora de controle. Fazendo uma simples comparação com o gráfico CUSUM da Figura 8, percebe-se que o EWMA demorou uma amostra a mais para sinalizar a mudança no processo.

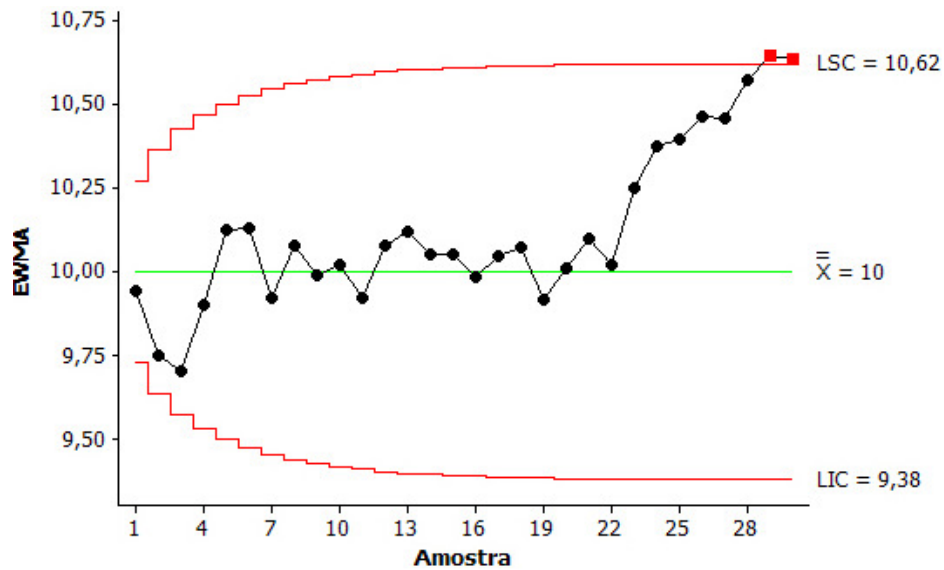


Figura 9: Exemplo de um gráfico de controle EWMA
 Fonte: Adaptado de Montgomery (2004)

2.2.3 Análise do Comportamento dos Gráficos de Controle

Além do teste padrão de um ponto afastado a mais de 3σ da LC, Nelson (1984), Montgomery (2004) e Costa et al. (2005) aconselham a adoção de outros diferentes testes que ajudam a melhorar a sensibilidade dos gráficos de controle às pequenas mudanças no processo (o teste 1 é o teste padrão):

- a) Teste 2: dois pontos em três consecutivos localizados a mais de 2σ da LC;
- b) Teste 3: quatro de cinco pontos consecutivos localizados a mais de 1σ da LC;
- c) Teste 4: seqüência de nove pontos localizados do mesmo lado da linha central;
- d) Teste 5: seis pontos em seqüência crescente ou decrescente;
- e) Teste 6: quinze pontos em seqüência a menos de 1σ da LC, tanto acima quanto abaixo;
- f) Teste 7: quatorze pontos em seqüência alternadamente para cima e para baixo;
- g) Teste 8: seqüência de oito pontos localizados a mais de 1σ em ambos os lados da LC.

Os autores ainda recomendam atenção a qualquer outro sinal de comportamento não-aleatório dos gráficos de controle. Costa et al. (2005) recomenda cautela no uso dos 8 testes descritos acima, já que estes podem implicar no aumento da incidência de alarmes falsos do processo, ou seja, o processo ser considerado como estando fora de controle quando na verdade está operando dentro dos padrões normais de funcionamento.

2.2.4 Sensibilidade do Gráfico de Controle

A simplicidade operacional e a eficiência na detecção de causas especiais do processo são características indispensáveis a um gráfico de controle, sendo apontadas por Montgomery (2004) como as principais razões pela ampla difusão e aplicação dos gráficos de controle nos mais diversos sistemas produtivos.

A efetividade na detecção de causas especiais está diretamente ligada à escolha dos limites de controle, do intervalo de amostragem, do tamanho da amostra e da escolha das regras sensibilizantes. Segundo Costa et al. (2005), uma medida de sensibilidade de um gráfico de controle é obtida através do número médio de amostras até o sinal (NMA), que é o número de amostras que devem ser coletadas até que o processo indique uma condição de fora de controle.

Quando o processo está sob controle estatístico, o NMA_0 indica o número médio de pontos necessários para a ocorrência do primeiro falso alarme no gráfico de controle. Quando o processo está fora de controle, o NMA_1 , chamado de NMA fora de controle indica, o número de amostras necessárias para a detecção da ocorrência da alteração na estatística monitorada (MONTGOMERY, 2004; COSTA et al., 2005).

Assim, quando se projeta um gráfico de controle, deseja-se que o NMA_0 seja o maior possível, pois este é um indício de que serão necessárias poucas interrupções no processo quando este está sob controle estatístico, e que o NMA_1 seja o menor valor possível, já que este valor está diretamente relacionado ao tempo necessário para a detecção de uma causa especial (BORROR et al., 1999; MONTGOMERY, 2004; COSTA et al., 2005). De forma geral, para um gráfico de controle de Shewhart, o NMA_0 pode ser escrito em função da probabilidade de erro tipo I, que é a probabilidade, em cada amostra, do gráfico sinalizar que o processo está fora de controle, quando o processo na verdade está sob controle estatístico (MONTGOMERY, 2004):

$$NMA_0 = \frac{1}{\alpha} \quad (61)$$

Por exemplo, para um gráfico de controle \bar{x} com limites de controle três sigma, a probabilidade de um ponto ser apontado como fora de controle é de 0,0027, logo NMA_0 é de 370,4.

O NMA_1 pode ser escrito em função da probabilidade de erro tipo II, que é a probabilidade do gráfico não detectar uma alteração na estatística monitorada quando o processo acabou de sair fora de controle estatístico (MONTGOMERY, 2004):

$$NMA_1 = \frac{1}{1 - \beta} \quad (62)$$

onde $\beta = P\{LIC \leq \bar{x} \leq LSC | \mu = \mu_1 = \mu_0 + k\sigma\}$

Quando a característica de qualidade monitorada apresenta distribuição conhecida e independente, os valores de σ e β podem ser obtidos pelo uso de simulação de Monte Carlo ou de Cadeias de Markov. Neste caso, para os gráficos de controle de Shewhart o NMA pode ser obtido diretamente, sem a necessidade de uso de simulação.

No caso de gráficos de controle CUSUM e EWMA, o NMA pode ser obtido através de Cadeias de Markov ou pela solução de equações integrais. O uso de equações integrais calcula o NMA de forma mais precisa que o uso de Cadeias de Markov, embora este último método seja mais versátil (LUCAS; CROSIER, 1982; SHU et al., 2004).

Assumindo que $L(u)$ é o NMA do gráfico EWMA, dado que o gráfico iniciou-se com valor inicial $Z_0 = u$, o NMA pode ser obtido através da solução da seguinte equação integral (CROWDER, 1987):

$$L(u) = 1 + \frac{1}{\lambda} \int_{LIC}^{LSC} L(y) f\left(\frac{y - (1 - \lambda)u}{\lambda}\right) dy \quad (63)$$

Nesta equação, assume-se que $f(\cdot)$ seja uma distribuição normal com média μ e desvio-padrão σ , ambos conhecidos. $L(\cdot)$ é uma integral Fredholm que pode ser resolvida através do método numérico de quadratura de Gauss-Legendre com r pontos (CROWDER, 1987; SHU et al., 2004). Para maiores informações sobre o método de quadratura de Gauss-Legendre, recomenda-se Press et al. (1997).

O gráfico CUSUM bilateral consiste na aplicação simultânea dos CUSUM unilaterais inferior e do superior. Segundo Lucas e Crosier (1982) e Gan (1991), o NMA de um CUSUM unilateral inferior, dado que o valor inicial de C_i^- é u , pode ser obtido através da equação (64).

$$L_I(u) = 1 + L_I(0)F(k + u) + \int_{-h}^0 L(y) f(y - k - u) dy \quad (64)$$

Similarmente, o NMA de um CUSUM unilateral superior pode ser calculado através da equação (65).

$$L_S(u) = 1 + L_S(0)F(k - u) + \int_0^{+h} L(y)f(y + k - u) dy \quad (65)$$

Nas equações (64) e (65), $f(\cdot)$ e $F(\cdot)$ são, respectivamente, a função de densidade de probabilidade e a função de distribuição acumulada de uma distribuição normal com média μ e desvio-padrão σ , ambos conhecidos. Estas equações podem ser resolvidas pelo método de quadratura de Gauss-Legendre com r pontos. O NMA do CUSUM bilateral pode ser calculado pela equação (66).

$$\frac{1}{L} = \frac{1}{L_I(u)} + \frac{1}{L_S(u)} \quad (66)$$

2.3 GRÁFICO DE CONTROLE DE REGRESSÃO

Embora existam alguns trabalhos anteriores, Mandel (1969) é considerado como o trabalho inaugural do gráfico de controle de regressão, pois foi esse autor quem apresentou a metodologia para esta ferramenta, mostrando também suas similaridades e diferenças em relação aos gráficos de controle de Shewhart. Mandel (1969) estudou o número de homens-hora necessários para processar o volume de correspondências processadas nas 74 maiores agências de correios dos Estados Unidos.

A primeira descrição para o gráfico de controle de regressão, utilizado para a monitoração de uma característica de qualidade em relação a uma variável de controle, foi apresentada por Lyle (1944). Neste trabalho, o autor descreve o controle do consumo de vapor em uma indústria de produção de açúcar em função da produção total da empresa.

DiPaola (1945) usou uma ferramenta que chamou de gráfico de controle para a linha de regressão para a monitoração e controle simultâneo de duas características da qualidade na produção de aviões. Weis (1957) também usa a mesma lógica utilizada por DiPaola para o controle simultâneo de duas características de qualidade de disjuntores para carro. Como os dois autores apenas utilizaram a correlação entre duas variáveis, sem apontar uma relação de causa e efeito entre estas, Mandel (1969) aconselha que a ferramenta utilizada por DiPaola e Weis poderia ser renomeada para gráfico de controle de correlação.

Wallis e Roberts (1956) apresentaram um exemplo do uso de limites de confiança para o controle de despesas totais de viagem em relação à duração das viagens dos engenheiros de uma empresa. A principal justificativa de Wallis e Roberts é o uso de duas fontes de variação para o monitoramento da característica de qualidade, ao contrário de

utilizar apenas a característica de qualidade monitorada. Mansfield e Wein (1958) aplicaram um gráfico de controle de regressão para o controle de custos de uma indústria automobilística.

Em alguns processos em que a característica de qualidade é fortemente dependente de algumas variáveis de controle, pode ocorrer uma situação em que alguma dessas variáveis de controle assumam um valor extremo e altere significativamente a média da característica de qualidade, podendo até causar tendência nos dados. Nessa situação, um gráfico de Shewhart apontaria que o processo estará fora de controle mesmo não havendo nada de errado com o processo. A melhor opção para essa situação seria ajustar a influência das variáveis de controle sobre a característica de qualidade do produto ou processo por meio de uma equação de regressão, para então plotar um gráfico com a característica de qualidade prevista por este modelo de regressão e o valor observado da característica de qualidade (ROTHSCHILD; ROTH, 1986; MONTGOMERY, 2004; SHU et al., 2004;2007).

Loredo et al. (2002) e Montgomery (2004) destacam o uso do ajuste de um modelo de regressão para processos autocorrelacionados, já que os resíduos de um modelo de regressão usualmente não apresentam autocorrelação, mesmo que a característica de qualidade ou as variáveis de controle sejam autocorrelacionadas. Uma breve revisão sobre o gráfico de controle de regressão é apresentada por Shu et al. (2007).

2.3.1 Comparações entre os Gráficos de Controle de Regressão e de Shewhart

O gráfico de controle de regressão é basicamente uma combinação dos gráficos de controle de Shewhart com a análise de regressão. De acordo com Mansfield e Wein (1958), Mandel (1969) e Almeida (2003), enquanto os gráficos de controle monitoram a média do processo, o gráfico de controle de regressão monitora uma variação média da característica de qualidade de acordo com os valores das variáveis de controle. Um exemplo genérico de um gráfico de controle de regressão é apresentado na Figura 10.

Comparando a representação genérica de um gráfico de controle de Shewhart, apresentado na Figura 5, com a representação genérica do gráfico de controle de regressão da Figura 10, observa-se que a linha central do gráfico de controle de regressão é inclinada em relação ao eixo horizontal, enquanto que para os gráficos de controle convencionais, essa linha é paralela a esse eixo. Segundo Mandel (1969), a linha central de um gráfico de controle

convencional é estimada pela média geral do processo, $\bar{\bar{X}}$, enquanto que a média do gráfico de controle de regressão é estimada pela linha de regressão apresentada na equação (2).

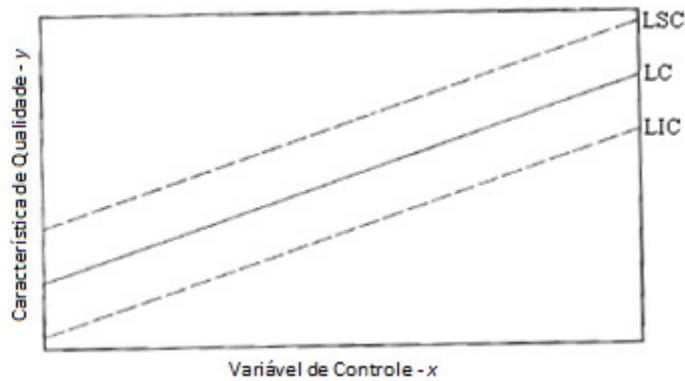


Figura 10: Exemplo genérico de um gráfico de controle de regressão
Fonte: Adaptado de Mandel (1969)

Os limites de controle dos gráficos de controle de Shewhart são paralelos ao eixo horizontal e à linha central, sendo calculados utilizando-se múltiplos do desvio-padrão do processo. Como é possível observar na Figura 10, os limites de controle do gráfico de controle de regressão, da forma como proposto por Mandel (1969) não são paralelos ao eixo horizontal, mas continuam sendo paralelos à linha central. Segundo Rothschild e Roth (1986), os limites de controle de um gráfico convencional são mais largos, pois esse gráfico trata a variação da característica de qualidade que ocorre devido às alterações como sendo variação não-explicada.

A maior diferença entre o gráfico de controle de regressão e os gráficos de Shewhart, segundo Mandel (1969) e Almeida (2003), está nas suposições em cada situação. Enquanto que para o uso dos gráficos de controle convencionais assume-se que as médias das amostras são normalmente distribuídas, com média constante estimada por $\bar{\bar{X}}$, para o gráfico de controle de regressão assume-se que a característica de qualidade é dependente das variáveis de controle e, para cada valor da variável de controle, as características de qualidade são normalmente distribuídas, com média estimada pela equação de regressão.

O gráfico de controle de regressão da forma como foi apresentado por Mandel (1969) apresenta outra diferença importante em relação aos gráficos de controle convencionais: o eixo horizontal do gráfico é a variável de controle e não uma unidade de tempo ou seqüência de amostragem, logo, as amostras não estão ordenadas temporalmente, o que dificulta a realização da maioria dos testes de estabilidade do processo. Uma possibilidade antecipada

por Mandel seria a construção de gráficos de controle para os resíduos de regressão, que teriam aparência igual aos gráficos de controle de Shewhart.

2.3.2 Limites de Controle para o Gráfico de Controle de Regressão

Como afirmado anteriormente, a linha central do gráfico de controle de regressão é estimada pela equação de regressão. De acordo com Mandel (1969), Olin (1998) e Almeida (2003), os limites de controle genéricos desse gráfico são fixados em relação ao desvio-padrão da variável ajustada:

$$LSC_i = \hat{y}_i + k\hat{\sigma}_e \quad (67)$$

$$LC_i = \hat{y}_i \quad (68)$$

$$LIC_i = \hat{y}_i - k\hat{\sigma}_e \quad (69)$$

O desvio-padrão para o gráfico de controle de regressão é estimado com base nos desvios dos valores observados da característica de qualidade em relação aos valores estimados pela equação de regressão. Mandel (1969) utilizou a estimativa da variância do modelo calculada pela equação (15) para calcular os limites de controle do gráfico de controle de regressão. Para o modelo de regressão linear simples, essa estimativa é:

$$\hat{\sigma}_e = s_e = \sqrt{QMR} = \sqrt{\frac{\sum_{i=1}^n e_i^2}{n-2}} \quad (70)$$

Dessa forma, os limites de controle simples serão:

$$LSC_i = \hat{y}_i + k\sqrt{QMR} \quad (71)$$

$$LIC_i = \hat{y}_i - k\sqrt{QMR} \quad (72)$$

De acordo com Almeida (2003), um dos objetivos do gráfico de controle de regressão é controlar uma variação média, dessa forma, pode-se utilizar a estimativa do desvio-padrão para a resposta média de um modelo de regressão linear simples:

$$\hat{\sigma}(\mu_{y|x_i}) = s_e \sqrt{\frac{1}{n} + \frac{(x_i - \bar{x})^2}{S_{xx}}} \quad (73)$$

Substituindo o desvio-padrão para a resposta média, estimado pela equação (73), nos limites de controle genéricos das equações (66) e (69), apresenta-se os seguintes limites de controle de predição (OLIN, 1998; JACOBI, 2001; ALMEIDA, 2003):

$$LSC_i = \hat{y}_i + ks_e \sqrt{\frac{1}{n} + \frac{(x_i - \bar{x})^2}{S_{xx}}} \quad (74)$$

$$LIC_i = \hat{y}_i - ks_e \sqrt{\frac{1}{n} + \frac{(x_i - \bar{x})^2}{S_{xx}}} \quad (75)$$

Os limites de controle apresentados nas equações (74) e (75) não são recomendados por Carvalho Jr. et al. (2007), pois produzem limites de controle muito estreitos podendo apresentar muitos alarmes falsos.

De acordo com Mandel (1969), Jacobi (2001) e Almeida (2003), a principal aplicação da regressão é a estimativa da característica de qualidade dado um conjunto de valores das variáveis de controle. Dessa forma, obtêm-se limites de controle mais exatos ao se utilizar a estimativa do desvio-padrão para a previsão de novos valores da característica de qualidade:

$$\hat{\sigma}(\hat{y}_{x_i}) = s_e \sqrt{1 + \frac{1}{n} + \frac{(x_i - \bar{x})^2}{S_{xx}}} \quad (76)$$

Usando a estimativa de variância para a predição de novas observações, obtêm-se os seguintes limites de controle para o gráfico de controle de regressão (JACOBI, 2001; ALMEIDA, 2003):

$$LSC_i = \hat{y}_i + ks_e \sqrt{1 + \frac{1}{n} + \frac{(x_i - \bar{x})^2}{S_{xx}}} \quad (77)$$

$$LIC_i = \hat{y}_i - ks_e \sqrt{1 + \frac{1}{n} + \frac{(x_i - \bar{x})^2}{S_{xx}}} \quad (78)$$

Segundo Ryan (1989), os limites de controle apresentados nas equações (77) e (78) diferem-se levemente de um intervalo de confiança para uma nova observação. Para limites baseados em intervalos de confiança, o k é substituído pelo valor da estatística t-student para o nível de confiança adotado e os graus de liberdade do termo de erro. Nos limites de controle do gráfico de controle de regressão, o k é fixado em um valor, geralmente 2 ou 3, independente dos graus de liberdade do erro. Lyle (1946) e Olin (1998) utilizaram intervalos de confiança com 95% como limites de controle.

Os limites de controle baseados em intervalos de confiança serão curvos, alargando-se à medida que os valores de x afastam-se da média. Mandel (1969) aponta duas justificativas para o uso de limites paralelos: o primeiro é que se a variação sobre a linha de

regressão não for grande, os limites de controle não serão tão curvos, o segundo é que esses limites paralelos controlam mais rigidamente os pontos mais afastados da média, embora isto possa aumentar o número de alarmes falsos. Assim, os limites paralelos não consideram os erros na estimativas dos coeficientes de regressão.

Olin (1998) comparou a estimação dos limites de controle utilizando os limites simples, apresentados nas equações (71) e (72), intervalos de confiança para a predição de novas observações, apresentados nas equações (77) e (78) e outro método, chamado de limites de tolerância. Segundo Olin (1998), os limites de tolerância apresentam uma menor sensibilidade que os limites de controle simples, sendo também mais difíceis de calcular e interpretar por parte dos operários.

Como os limites de controle baseados em intervalos de confiança apresentam um alto número de alarmes falsos, Olin (1998) defende o uso dos limites de controle simples, pois estes são mais fáceis de calcular, além de combinarem uma alta sensibilidade e um reduzido número de alarmes falsos. Esta conclusão está de acordo com Mandel (1969) e Jacobi et al. (2002), que defendem o uso de limites simples

2.3.3 Construção do Gráfico de Controle de Regressão

O método proposto por Mandel (1969) ainda é a base da maioria dos outros trabalhos consultados. Jacobi (2001) incluiu um passo preliminar no método de Mandel, que é uma análise descritiva dos dados, incluindo um *box-plot*, com a finalidade de identificar possíveis pontos que mereçam uma maior atenção e sejam passíveis de verificação.

O primeiro passo descrito por Mandel (1969) é a construção de um diagrama de dispersão, visando fazer uma análise preliminar para verificar a linearidade da relação entre as variáveis do processo, além de permitir a detecção de possíveis dados atípicos. Jacobi et al. (2002) aconselham a construção de um gráfico temporal dos dados, o que também ajuda a verificar a presença de dados atípicos.

Os dados que forem considerados atípicos nessas análises devem ser investigados, caso sejam detectadas causas especiais que tenham influenciado o comportamento desses pontos, pode-se desconsiderar os respectivos pares de dados dos cálculos. Os dados restantes são utilizados para estimar a equação de regressão e seus parâmetros.

De posse dos parâmetros do modelo, deve-se escolher qual das formas de cálculo do desvio-padrão será utilizada para estimar os limites de controle. Mandel (1969) utilizou os

limites de controle simples, apresentados nas equações (71) e (72). A constante k é escolhida conforme o número máximo de alarmes falsos e a sensibilidade desejada pra o gráfico de controle de regressão.

A abordagem tradicional do CEP é utilizar como limites de controle $k = 3$. Essa estratégia foi utilizada por Olin (1998) e Jacobi et al. (2002), mas Mandel (1969), Almeida (2003) e Carvalho Jr. (2006) utilizaram $k = 2$, alegando ser essa uma decisão gerencial baseada em critérios econômicos e de experiência prática, mesmo que essa escolha de k aumente o número de alarmes falsos. O gráfico de controle de regressão apresentado por Mandel (1969) para o controle de horas-homem necessários para entregar o volume mensal de cartas é apresentado na Figura 11.

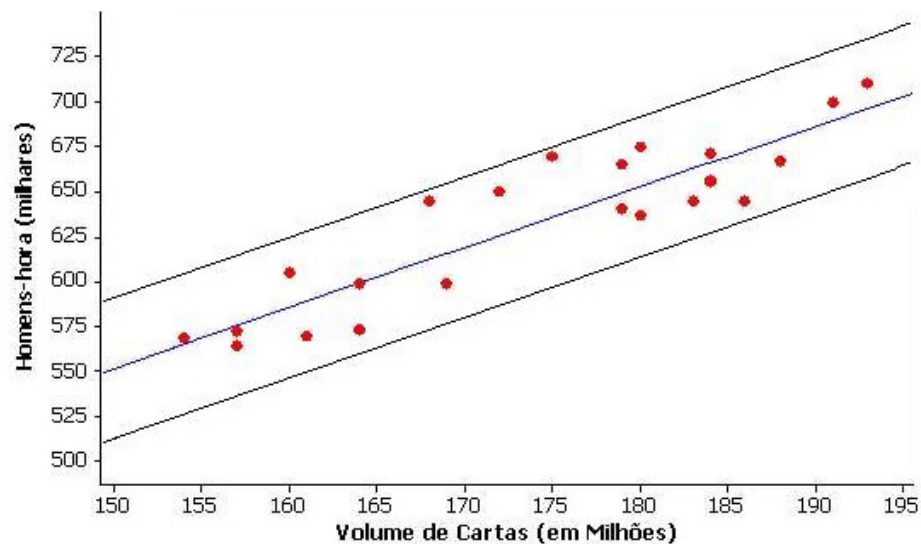


Figura 11: Gráfico de controle de regressão para o processamento de correspondências
Fonte: Adaptado de Mandel (1969)

Assim como nos gráficos de controle de Shewhart, analisa-se a presença de pontos fora de controle, sendo que o único dos testes de estabilidade que pode ser aplicado no gráfico de controle de regressão proposto por Mandel (1969) é o teste de pontos estarem acima do LSC ou abaixo do LIC. Outra análise nessa fase é o significado de cada ponto fora de controle, por exemplo, para a aplicação de Mandel (1969) um ponto localizado acima da LSC indica uma queda na produtividade do sistema, enquanto um ponto abaixo do LIC indica uma melhoria de produtividade.

2.3.4 Extensões do Gráfico de Controle de Regressão

Alguns autores estenderam o gráfico de controle de regressão proposto por Mandel (1969), aplicando esse gráfico para modelos de regressão linear múltipla, regressão não linear, modelos lineares generalizados, modelos com erros nas variáveis, entre outros. A seguir, serão apresentados o gráfico de controle de regressão múltipla, o gráfico EWMA para os resíduos padronizados (EWMAREG), o gráfico de medidas individuais aplicado ao modelo de regressão e uma breve descrição de outras extensões.

2.3.4.1 Gráfico de Controle de Regressão Múltipla

Conforme apresentado anteriormente, o gráfico de controle de regressão proposto por Mandel (1969) apresentava apenas uma variável de controle e uma característica de qualidade, utilizando o modelo de regressão linear simples para modelar a relação entre essas variáveis. Caso um processo apresente mais de uma variável de controle, Mandel (1969) sugere a utilização de gráficos de controle para os resíduos do modelo. Desta forma, Haworth (1996) utilizou um modelo de regressão linear múltipla para modelar o tempo de manutenção de softwares, gerando o gráfico de controle de regressão múltipla.

Para gerar o gráfico de controle de regressão múltipla, Haworth (1996) plotou um gráfico de controle para os resíduos padronizados na forma *student* do modelo de regressão linear múltipla ordenados temporalmente, elaborando um algoritmo em Visual BASIC para a construção do gráfico de controle. Haworth utilizou a estimativa do desvio-padrão para novas observações, apresentado na equação (32), para gerar os resíduos padronizados na forma *student*.

$$z_i = \frac{e_i}{\hat{\sigma} \sqrt{1 + x_i'(X'X)^{-1}x_i}} \quad (79)$$

Os limites de controle escolhidos por Haworth (1996) seriam o valor da estatística t, com o grau de significância desejado:

$$LSC = +t_{\alpha/2, n-p} \quad (80)$$

$$LC = 0 \quad (81)$$

$$LIC = -t_{\alpha/2, n-p} \quad (82)$$

Também é possível reescrever os limites de controle das equações (80) e (82) para uma forma genérica, adotando um k constante, geralmente escolhido como sendo 3.

$$LSC = +k \quad (83)$$

$$LIC = -k \quad (84)$$

Assim, no gráfico de controle de regressão múltipla proposto por Haworth (1996) são monitorados os resíduos padronizados. O gráfico de controle de regressão utilizado pelo autor para monitorar o tempo de manutenção de softwares em função de variáveis do processo é apresentado na Figura 12.

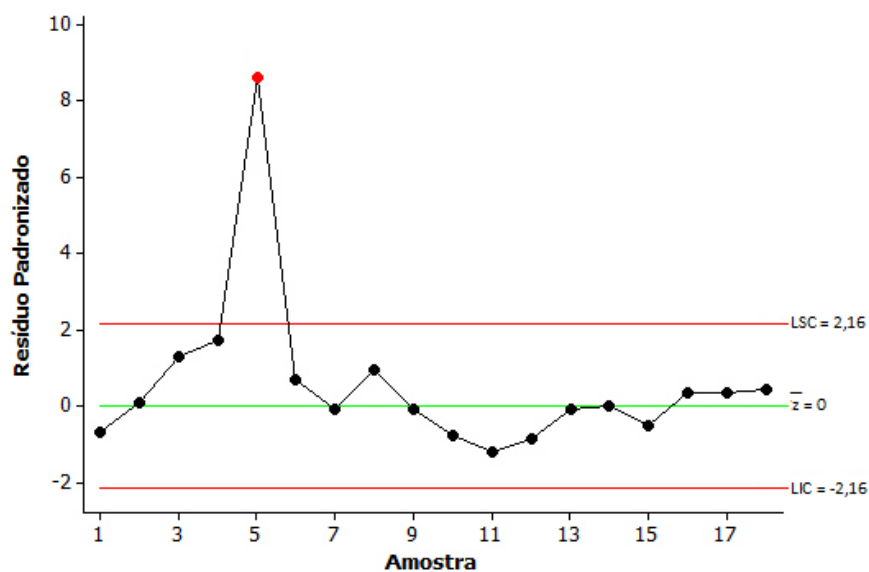


Figura 12: Gráfico de controle de regressão múltipla para o monitoramento do tempo de manutenção de softwares

Fonte: Adaptado de Haworth (1996)

O gráfico de controle de regressão, como proposto por Mandel (1969), permite apenas um teste de estabilidade: um ponto acima do LSC ou abaixo do LIC. A construção de gráfico para os resíduos da regressão, além de permitir a análise da influência de mais de uma variável de controle sobre a característica de qualidade, também permite a realização dos testes de estabilidade descritos na subseção 2.2.3.

A idéia do gráfico de controle de regressão múltipla é similar ao gráfico de controle de causas especiais, proposto por Alwan e Roberts (1988), já que ambos os gráficos monitoram resíduos de modelos. A diferença é que Alwan e Roberts (1988) utilizam modelos ARIMA para modelar os dados, enquanto Haworth (1996) utiliza modelos de regressão linear múltipla. Pedrini e Caten (2008) apresentam um exemplo em que o gráfico de controle de

regressão é capaz de detectar uma causa especial propositalmente inserida, enquanto o gráfico de causas especiais não detectou nenhuma causa especial.

2.3.4.2 Gráfico EWMA para os Resíduos da Regressão

Shu et al. (2004) chamaram de EWMAREG o gráfico de controle EWMA aplicado aos resíduos padronizados de um modelo de regressão, que pode ser considerado similar a um dos procedimentos de monitoramento de perfis lineares proposto por Kang e Albin (2000). Embora com pequenas diferenças, Teixeira (1997) e Shu et al. (2004) utilizaram o gráfico EWMAREG: enquanto Teixeira (1997) aplicou o gráfico EWMA diretamente aos resíduos de um modelo de regressão linear múltipla, Shu et al. (2004) aplicaram aos resíduos padronizados de um modelo de regressão linear simples. Por exemplo, o EWMAREG para os resíduos padronizados, seria:

$$U_i = \lambda \frac{e_i}{\hat{\sigma}} + (1 - \lambda)U_{i-1} \quad (85)$$

Os limites de controle para o EWMAREG são os mesmos limites de controles do gráfico de controle do gráfico EWMA, apresentados nas equações (56), (57) e (58), tomando-se o cuidado de adotar $\mu_0 = 0$.

$$LSC_i = +L\sigma \sqrt{\frac{\lambda}{2 - \lambda} [1 - (1 - \lambda)^{2i}]} \quad (86)$$

$$LC = 0 \quad (87)$$

$$LIC_i = -L\sigma \sqrt{\frac{\lambda}{2 - \lambda} [1 - (1 - \lambda)^{2i}]} \quad (88)$$

Quando $\lambda = 1$ e $L = t_{\alpha/2, n-p}$, o gráfico EWMAREG se reduz ao gráfico de controle de regressão, da forma como aplicada por Haworth (1996). Na maioria das aplicações, os parâmetros reais da regressão são desconhecidos e devem ser estimados. Devido a isso, Shu et al. (2004) estudaram a influência teórica da estimação de β_0 , β_1 e σ no NMA do gráfico de controle de regressão. Segundo os autores, quando os parâmetros da regressão são estimados, ocorre um aumento do número de alarmes falsos do EWMAREG. Se o σ for subestimado, sempre ocorrerá o aumento do número de alarmes falsos, mas no caso de ocorrer uma superestimação de σ , a alteração da sensibilidade depende do erro de estimação dos outros parâmetros da regressão.

Shu et al. (2004) obteve apenas o NMA deste gráfico para mudanças na média dos resíduos do modelo de regressão. Em função disto, na comparação do método proposto nesta dissertação com outros procedimentos, o NMA do EWMAREG somente será calculado apenas quando esta envolver alterações exclusivas na média.

2.3.4.3 Gráfico X-MR para os Resíduos da Regressão

Rothschild e Roth (1986) calcularam os limites de controle para o gráfico de controle de regressão utilizando a amplitude móvel dos resíduos da regressão como estimativa do desvio-padrão. Na verdade, esse método consiste em plotar gráficos de controle para os valores individuais da característica de qualidade, com o desvio-padrão sendo estimado através da amplitude móvel dos resíduos da regressão. Embora tenha utilizado as amplitudes móveis para estimar o desvio-padrão do processo, Rothschild e Roth (1986) não construíram o gráfico de controle para monitorar as amplitudes móveis. Os limites de controle são estimados por uma adaptação das equações apresentadas na Figura 7, em que se substitui a média aritmética pelo valores previstos pelo modelo de regressão:

$$LSC_i = \hat{y}_{x_i} + E_2 \overline{MR}_{e_i} \quad (89)$$

$$LC_i = \hat{y}_{x_i} \quad (90)$$

$$LIC_i = \hat{y}_{x_i} - E_2 \overline{MR}_{e_i} \quad (91)$$

Rothschild e Roth (1986) aplicaram o gráfico de controle de regressão para monitorar a concentração de íons de ouro em um processo de galvanização de superfícies. A principal justificativa apresentada pelos autores para aplicar o gráfico de controle de regressão é que esse processo apresenta uma média que é variável em relação à quantidade de energia utilizada. Segundo Rothschild e Roth (1986), o gráfico de controle de regressão permitiu um melhor entendimento do processo, possibilitando a estimação do momento exato para a reposição da solução metálica, diminuindo custos de produção e aumentando a qualidade, além de apontar falhas na preparação da solução.

Na Figura 13 apresenta-se o gráfico de controle de regressão, construído com base nos dados dos autores para o processo de galvanização a ouro. Analisando os dados utilizados por Rothschild e Roth (1986), chegou-se à conclusão de que o uso da amplitude móvel dos resíduos de regressão para estimar o desvio-padrão, neste caso resultou em limites de controle ligeiramente mais estreitos do que os limites de controle estimados pelas equações (71) e (72).

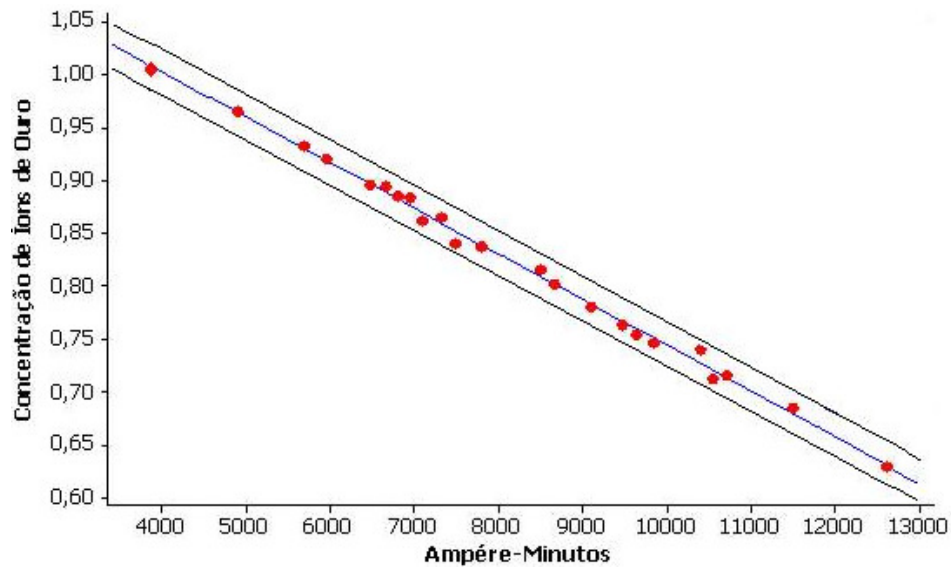


Figura 13: Gráfico de controle de regressão para o processo de galvanização
 Fonte: Adaptado de Rothschild e Roth (1986)

Loredo et al. (2002) propõem um método de monitoramento de processos autocorrelacionados utilizando gráficos de controle de regressão. Neste trabalho, os limites de controle foram definidos através da amplitude móvel dos resíduos do modelo de regressão, mas, neste caso, a característica monitorada foram os resíduos, ao invés da própria característica de qualidade.

O gráfico de controle proposto por Loredo et al. (2002) é apresentado na Figura 14. Através de simulação, os autores concluíram que quando a característica de qualidade é autocorrelacionada, o método proposto é mais eficiente que os gráficos de controle clássicos.

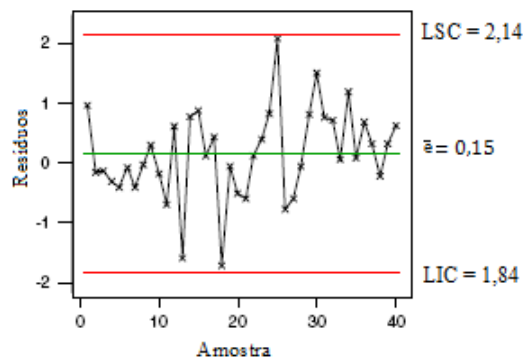


Figura 14: Gráfico de controle de medidas individuais aplicado aos resíduos
 Fonte: Adaptado de Loredo et al. (2002)

2.3.4.4 Outras Extensões do Gráfico de Controle de Regressão

Hawkins (1991) desenvolveu um procedimento alternativo para o controle de processos multivariados, que consiste basicamente do ajuste de modelos de regressão para cada variável separadamente em função das outras variáveis. As variáveis ajustadas são monitoradas por gráficos CUSUM e, segundo o autor, permitem a identificação da variável que está fora de controle, o que não é possível com o uso da estatística T^2 de Hotelling. Hawkins (1993) modificou esta abordagem para o monitoramento de processos em cascata, que são processos compostos por vários subprocessos.

Olin (1998) estendeu o método de Mandel (1969) para modelos de regressão não linear e para os modelos lineares generalizados, especialmente a regressão Poisson e binomial. Skinner et al. (2003) propuseram um método de monitoramento de dados de contagem, utilizando a regressão Poisson para a modelagem e gráficos de controle para os resíduos desta regressão.

Em um estudo similar ao de Skinner et al. (2003), Jearkpaporn et al. (2003; 2004) propuseram um gráfico de controle para os resíduos *deviances* de um modelo linear generalizado, utilizando a regressão gama como técnica de modelagem. Através de simulações, Jearkpaporn et al. (2003) e Skinner et al. (2003) concluíram que os gráficos de controle propostos apresentam desempenhos superiores ao de gráficos de controle clássicos. Jearkpaporn et al. (2007) estenderam o uso de gráficos de controle baseados em modelos lineares generalizados para o monitoramento de processos em cascata.

Skinner et al. (2004) aplicaram o método proposto por Skinner et al. (2003), comparando também os resultados do gráfico de controle proposto com o gráfico de controle de regressão proposto por Mandel (1969). Skinner et al. (2004) concluíram que os gráficos de controle baseados em modelos lineares generalizados apresentaram desempenho superior ao gráfico de controle de regressão. Uma justificativa para esta conclusão é o tipo de relação existente entre as variáveis de controle e a característica de qualidade, já que, segundo Hawkins (1991), o desempenho de gráficos de controle baseados em resíduos é aperfeiçoado conforme o modelo utilizado é melhorado.

Almeida (2003) e Carvalho Jr. (2006) estenderam o trabalho de Mandel (1969) para modelos com erros nas variáveis (MEV). Almeida (2003) utilizou o modelo funcional, em que a variável controlada é considerada como sendo fixa e medida com erro, gerando o gráfico de controle de regressão funcional. Por sua vez, Carvalho Jr. (2006) considerou a

variável de controle como sendo uma variável aleatória, dando origem ao gráfico de controle de regressão estrutural. Ambos aplicaram seus estudos no controle da densidade aparente de eletrodos de carbono para a produção de alumínio, em relação ao peso dos eletrodos. Uma comparação entre essas duas técnicas é apresentada por Carvalho Jr. et al. (2007).

2.3.5 Procedimentos Similares ao Gráfico de Controle de Regressão

Foram encontrados na literatura consultada, dois procedimentos similares ao gráfico de controle de regressão: (i) os gráficos de controle para seleção de causas e (ii) gráficos de controle para o monitoramento de perfis lineares.

2.3.5.1 *Gráficos de Controle para Seleção de Causas*

A maioria dos produtos manufaturados atualmente é resultado de várias etapas de processamento e, freqüentemente, adota-se um gráfico de controle de Shewhart para cada etapa individual de processamento. No entanto, a característica de qualidade de uma etapa do processo é significativamente afetada pelas características de qualidade das etapas anteriores, assim, o uso de gráficos de controle para cada etapa individualmente podem não ser tão eficientes (WADE; WOODALL, 1993; SHU et al., 2005).

O gráfico de controle para seleção de causas (GCSC), proposto por Zhang¹ *apud* Wade e Woodall (1993), é uma alternativa para esse caso, pois tem como idéia base a monitoração de uma variável somente após as observações terem sido ajustadas pelo efeito de outras variáveis aleatórias. A idéia geral do GCSC é bastante similar ao gráfico de controle de regressão, proposto por Mandel (1969) e aos gráficos de controle para variáveis ajustadas por regressão, propostos por Hawkins (1993). O GCSC consiste no ajuste de um modelo, linear ou não linear, relacionando a característica de qualidade do atual processo às características de qualidade dos processos anteriores. De acordo com Hawkins (1993), o uso do ajustamento de modelos de regressão para processos com várias etapas permite a identificação da etapa de processamento que não está sob controle estatístico.

¹ ZHANG, G. X. Cause-Selecting Control Charts – A New Type of Quality Control Charts. **The QR Journal**, Bangalore, v. 12, n. 4, p. 221-225, 1985.

Wade e Woodall (1993) e Yang (1998) destacam que o GCSC permite a distinção de duas causas especiais: (i) causas especiais de controle e (ii) causas especiais não-controláveis. Segundo estes autores, as causas especiais controláveis são as variáveis de controle do subprocesso atual e as causas especiais não-controláveis são as variáveis de controle dos subprocessos anteriores.

Considerando um processo composto por duas etapas, com y sendo a característica de qualidade do último processo e x sendo a característica de qualidade do primeiro processo, é possível ajustar um modelo de y em função de x . Definindo Z_i como sendo os valores para seleção de causas (WADE; WOODALL, 1993):

$$Z_i = y_i - \hat{y}_i \quad (92)$$

Os limites de controle para o GCSC podem ser estimados pelas amplitudes móveis de Z_i , ou seja, da mesma forma que os limites de controle para os gráficos de controle de regressão utilizando a amplitude móvel dos resíduos como estimativa do desvio-padrão, apresentados na equação (89, (90) e (91). Assim, quando se analisa um processo, utiliza-se gráficos de Shewhart para monitorar as características de qualidade do processo anterior e o GCSC para o processo atual. As regras de decisão apresentadas por Wade e Woodall (1993) são apresentadas na Figura 15.

Wade e Woodal (1993) propuseram o uso de intervalos de confiança para a previsão de novas observações ao invés do uso dos limites de controle utilizando a amplitude móvel. Dessa forma, os limites de controle para o GCSC são os mesmos limites apresentados para o gráfico de controle de regressão, apresentados anteriormente nas equações (80) e (82). Com o uso de simulação, Wade e Woodal (1993) mostram que esta modificação diminui o número de alarmes falsos do GCSC. Yang (1998) mostra o planejamento econômico do GCSC, adotando os limites de controle utilizados por Mandel (1969), apresentados nas equações (71) e (72).

Gráfico para X	Gráfico para Z	Interpretação
Fora de controle	Fora de controle	Ambos os processos estão fora de controle
Fora de controle	Sob Controle	Apenas o processo anterior está fora de controle
Sob Controle	Fora de controle	Apenas o processo atual está fora de controle
Sob Controle	Sob Controle	Ambos os processos estão sob de controle

Figura 15: Regras de decisão para o gráfico de controle para seleção de causas
Fonte: Adaptado de Wade e Woodal (1993)

2.3.5.2 *Monitoramento de Perfis Lineares*

Em muitas situações, a qualidade do produto ou processo é melhor representada por uma relação entre as características de qualidade do processo e outras variáveis. Um exemplo apresentado por Kang e Albin (2000) é a produção de adoçantes artificiais, onde a característica de qualidade do produto é representada por uma função desejável para a relação entre a quantidade de aspartame dissolvida em um litro de água e a temperatura desta.

De forma geral, o monitoramento das relações lineares é feito em duas fases, a primeira consistindo no ajustamento de um modelo de regressão linear para a característica de qualidade do processo e verificação inicial da estabilidade do processo. Com a estimação do modelo, tem-se início a Fase II, onde se estima um modelo de regressão para cada amostra de tamanho n e verifica-se da estabilidade dos parâmetros de regressão (KANG; ALBIN, 2000; KIM et al., 2003; ZOU et al., 2007).

Para o monitoramento do processo, é usual construir gráficos de controle X-MR, EWMA ou CUSUM para os parâmetros da regressão, tendo como padrão de comparação a equação de regressão obtida na Fase I. Kang e Albin (2000) propuseram dois procedimentos em seu trabalho, sendo que o primeiro consiste no monitoramento do coeficiente de intercepto e do coeficiente de inclinação através de um gráfico de controle T^2 de Hotelling, considerando que estes dois coeficientes de regressão são correlacionados. O segundo procedimento proposto por Kang e Albin (2000) consiste no monitoramento dos resíduos do modelo de regressão através de um gráfico EWMA e de um gráfico R, a lógica defendida pelos autores é de que valores altos para os resíduos indicam grandes alterações nos coeficientes do modelo de regressão. Com base na simulação realizada, os autores concluem que ambos os procedimentos propostos apresentam bons desempenhos para o monitoramento de perfis lineares.

Kim et al. (2003) propuseram um método para o monitoramento dos coeficientes de inclinação e intercepto e do desvio-padrão do modelo de regressão linear simples de um processo utilizando o gráfico EWMA. O método basicamente consiste na construção de gráficos EWMA para cada parâmetro de regressão estimado em cada amostra, com os limites de controle sendo calculados com base nos parâmetros do modelo estimado pela Fase I desse método.

2.3.6 Exemplos de Aplicações de Gráfico de Controle de Regressão

Um gráfico de controle, por mais sofisticado que seja, deve apresentar duas características: efetividade na detecção de alterações no processo e simplicidade de uso. O gráfico de controle de regressão possui essas duas características, pois não requer muitos recursos computacionais para sua construção e é efetivo para detectar causas especiais de variação, que não estejam associadas às variações planejadas da variável de controle. A seguir, serão apresentados os resultados de alguns estudos que aplicaram com sucesso o gráfico de controle de regressão.

Spanos (1992) revisou as ferramentas do CEP aplicadas na produção de semicondutores e dentre estas ferramentas, apontou a aplicação do gráfico de controle de regressão múltipla na monitoração da taxa de deposição de um vapor químico em um tubo foto-litográfico, em função da temperatura, fluxo de SiH_4 e pressão. Embora o autor não tenha apresentado como os limites de controle foram calculados, ressaltou que o gráfico de controle de regressão múltipla foi efetivo no controle do processo.

Johnson et al. (1995) mostram a aplicação do gráfico de controle de regressão para o monitoramento da produção de pesticidas nos Estados Unidos. Neste trabalho, a característica de qualidade monitorada foi o coeficiente de partição de carbono orgânico no solo, em função da solubilidade em água do pesticida. Os autores utilizaram limites de controle baseados em intervalos de predição de novas observações com 95% e 99% de confiança. Segundo os autores, o uso desses gráficos de controle permite a identificação das fontes de erro existentes na produção de pesticidas.

Koksalan *et al.* (1999) criaram um procedimento utilizando o gráfico de controle de regressão múltipla conjuntamente com variáveis *dummy* para detectar e incorporar variáveis que não foram incluídas em modelos de regressão, calculando limites de controle utilizando as equações (77) e (78), adotando $k = 1$. Os autores aplicaram o procedimento em modelos para as previsões de curto e médio prazo de demanda de cerveja na Turquia.

Em um estudo brasileiro, Jacobi (2001) aplicou o gráfico de controle de regressão para a monitoração da distância percorrida por caminhões na coleta de resíduos em Santa Maria (RS), em relação ao peso desses resíduos. Os autores aplicaram a mesma metodologia apresentada por Mandel (1969), mas utilizando os limites de controle baseados na previsão de novas observações, dados pelas equações (77) e (78), adotando $k = 3$. O gráfico de controle de

regressão para o monitoramento do processo de recolhimento de resíduos é apresentado na Figura 16.

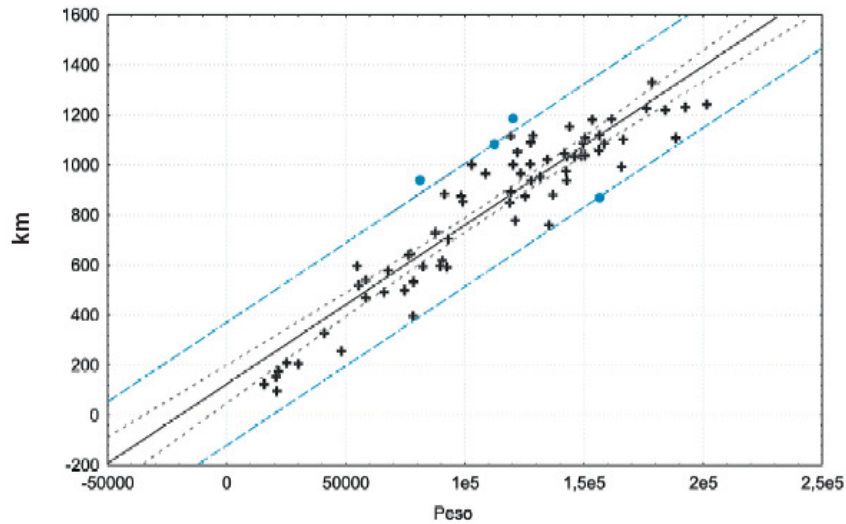


Figura 16: Gráfico de controle de regressão para o processamento de resíduos
Fonte: Jacobi et al. (2002)

Na análise de Jacobi (2001), um ponto acima do LSC indica uma queda na produtividade do processo e um ponto abaixo do LIC indica o contrário. Como é possível observar na Figura 16, foram encontrados pontos fora de controle que indicavam que o sistema de coleta de resíduos não estava sendo realizado de forma satisfatória. Essa ferramenta também permitiu uma melhoria dos processos da empresa, pois revelou a necessidade de transformações internas na empresa, como melhor organização do banco de dados e contratação de novos funcionários.

Omura e Steffe (2003) utilizaram o gráfico de controle de regressão com o objetivo de controlar a viscosidade de alimentos líquidos com grandes partículas em relação à taxa de rotação média do equipamento experimental de processamento. Como a relação entre essas variáveis é intrinsecamente linear, foi necessária uma transformação logarítmica para ajustar uma reta de regressão.

Omura e Steffe (2003) calcularam os limites de controle através das equações (71) e (72), com a escolha de $k = 3$, para diminuir o número de alarmes falsos. Os autores conseguiram atingir o objetivo do trabalho, construindo gráfico de controle de regressão para cada tipo de alimento a ser processado no equipamento, com um adicional de que os modelos reológicos obtidos podem ser usados para resolver problemas de engenharia do processo.

Lopes et al. (2006) aplicaram o gráfico de controle de regressão para o monitoramento da produção de tijolos em função da temperatura ambiente. Inicialmente, foram analisadas variáveis climáticas, como temperatura média, umidade relativa do ar, chuva acumulada e a insolação total. Após a análise de regressão, somente a temperatura ambiente foi considerada estatisticamente importante para explicar a variação da quantidade de tijolos produzidos. Os autores aplicaram um gráfico de controle de regressão utilizando os limites de controle das equações (71) e (72), adotando $k = 3$. O gráfico de controle para a quantidade de tijolos produzidos, em função da temperatura ambiente, é apresentado na Figura 17.

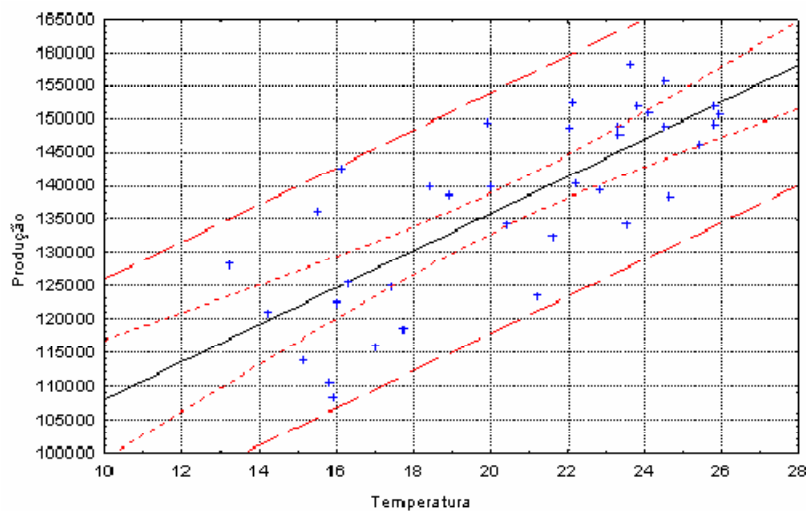


Figura 17: Gráfico de controle de regressão para o monitoramento da produção de tijolos
Fonte: Lopes et al. (2006)

De acordo com Lopes et al. (2006), a utilização do gráfico de controle de regressão permitiu avaliar a estabilidade do processo, além de ser utilizado como uma ferramenta de tomada de decisões, possibilitando a decisão dos níveis de produção com base na previsão da temperatura. Através da análise da Figura 17, conclui-se que o processo está sob controle estatístico e que maiores temperaturas acarretam uma produção maior de tijolos. É importante salientar que esta aplicação do gráfico de controle de regressão foi diferente da proposta inicial de Mandel (1969), já que a temperatura ambiente não é uma variável de controle.

Casarin et al. (2007) aplicaram o método de Jacobi et al. (2002) para o controle do peso de sucatas em relação ao número de colheitadeiras produzidas em uma indústria de máquinas agrícolas. Com a aplicação do gráfico de controle de regressão, os autores puderam concluir que o método de produção das máquinas estava produzindo uma elevada quantidade de sucatas em relação ao número de máquinas produzidas no período. O gráfico de controle

para a monitoração da sucata gerada na produção de colheitadeiras agrícolas é apresentado na Figura 18.

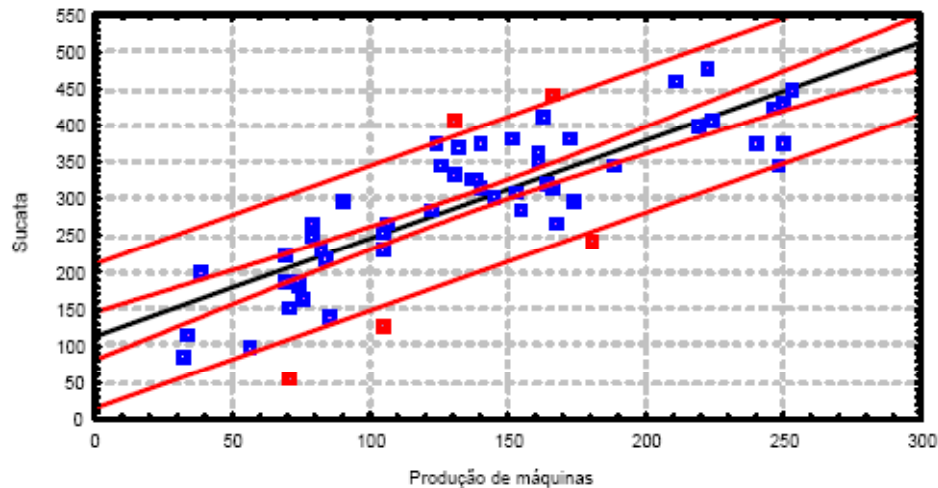


Figura 18: Gráfico de controle de regressão para o monitoramento de sucatas na produção de colheitadeiras
Fonte: Casarin et al. (2007)

Ainda, segundo Cai et al. (2002) e Montgomery (2004), se a taxa de desgaste de uma ferramenta for conhecida, ou pelo menos, puder ser estimada, o gráfico de controle de regressão pode ser aplicado para monitorar o desgaste de ferramentas, substituindo métodos apresentados por outros autores para esse problema. Em linhas gerais, quando a linha de tendência exceder o máximo admissível, a ferramenta deve ser substituída.

2.4 CONCLUSÃO DO CAPÍTULO

No capítulo 2, foi apresentada a revisão bibliográfica sobre modelos de regressão, controle estatístico de processo e gráficos de controle que fundamentam teoricamente o presente trabalho. Dentro da subseção relativa aos modelos de regressão linear, apresentou-se: o método dos mínimos quadrados para a estimação do modelo regressão e suas propriedades; os testes de hipóteses utilizados na análise de regressão; as aplicações do modelo de regressão para a previsão de novas observações; os coeficientes de determinação e sua interpretação; a análise residual para a validação do modelo; métodos para a detecção de observações influentes; métodos para a seleção de variáveis e modelos quando há um grande número de variáveis de controle e a multicolinearidade em um modelo de regressão. Na

subseção relativa ao controle estatístico de processo, apresentou-se um breve histórico dos gráficos de controle, uma descrição básica sobre os gráficos de controle para variáveis, os gráficos de controle EWMA e CUSUM, a análise do comportamento e sensibilidade dos gráficos de controle.

Dentro do gráfico de controle de regressão foi apresentado um breve histórico dessa ferramenta, as formas de calcular os limites de controle, algumas extensões e métodos similares ao gráfico de controle de regressão e, por fim, apresentou-se exemplos de aplicação do gráfico de controle de regressão. Um sumário dos principais trabalhos sobre o gráfico de controle de regressão encontrados na literatura é apresentado na Figura 19.

Tipo de Gráfico	Obras Consultadas
Gráfico de Controle de Regressão	Lyle (1944); Wallis e Roberts (1956); Mansfield e Wein (1958); Mandel (1969); Ryan (1989); Olin (1998); Jacobi (2001); Jacobi et al. (2002); Omura e Steffe (2003); Lopes et al. (2006); Casarin et al. (2007)
Gráfico de Controle de Regressão Múltipla	Spanos (1992); Haworth (1996); Koksalan (1999); Pedrini e Caten (2008)
Gráfico EWMA para os Resíduos da Regressão	Teixeira (1997); Shu et al. (2004)
Gráfico X-MR para os Resíduos da Regressão	Rothschild e Roth (1986); Loredó et al. (2002)
Gráfico de Controle de Regressão Funcional	Almeida (2003); Carvalho Jr. et al. (2007)
Gráfico de Controle de Regressão Estrutural	Carvalho Jr. (2006); Carvalho Jr. et al. (2007)
Gráfico de Controle para Regressão Não linear	Olin (1998)
Gráfico de Controle para Modelos Lineares Generalizados	Olin (1998); Jearkpaorn et al. (2003; 2004; 2007); Skinner et al. (2003; 2004)
Revisão Bibliográfica	Shu et al. (2007)

Figura 19: Sumário de trabalhos sobre gráfico de controle de regressão consultados
 Fonte: Figura elaborada pelo autor (2009)

3 MÉTODO PROPOSTO

Este capítulo apresenta o desenvolvimento do método proposto para a utilização do gráfico de controle de regressão múltipla em um processo produtivo, em que também estará inserida a modificação proposta no presente trabalho. É importante destacar que o método e modificação propostos visam simplificar a aplicação do gráfico de controle de regressão múltipla, mas também podem ser aplicados aos gráficos de controle de regressão simples. Também será apresentado o gráfico de controle de extrapolação para verificar se os valores das variáveis de controle extrapolaram a região original dos dados.

3.1 INTRODUÇÃO

De acordo com Woodall e Montgomery (1999), um método de implantação de gráficos de controle deve ser dividido em duas fases: (i) fase de análise retrospectiva, também chamada de Fase I; (ii) fase de monitoramento do processo, também chamada de Fase II. Segundo Faltin et al. (1997), a maioria dos trabalhos encontrados na literatura sobre gráficos de controle falha justamente em não separar essas duas fases em sua metodologia.

Dessa forma, o método proposto é subdividido nestas duas fases, sendo a Fase I composta pela coleta inicial de amostras, ajuste de um modelo de regressão, cálculo dos limites de controle e verificação da estabilidade do processo. A Fase II é a descrição do monitoramento do processo propriamente dito, onde se compara as amostras coletadas com os limites de controle calculados com base nas estatísticas calculadas na fase anterior. Segundo Faltin et al. (1997) para essa fase é necessário supor que os dados monitorados e os dados utilizados na estimação dos limites de controle sigam a mesma distribuição. O uso de modelos de regressão também permitirá a previsão dos valores da característica de qualidade em relação ao conjunto de valores das variáveis de controle do processo.

Ressalta-se que o método de monitoramento proposto incluirá uma modificação relativamente simples que permitirá o monitoramento direto das observações referentes a uma característica de qualidade dependente de uma ou mais variáveis de controle do processo, ao invés do monitoramento dos resíduos da regressão, conforme proposto por Haworth (1996). Montgomery e Mastrangelo (1991), Montgomery (2004) e Moreira Jr. (2005) ressaltam que alguns operários possuem dificuldades em interpretar os gráficos de controle baseados em resíduos, já que estes nem sempre podem ser considerados uma referência direta do processo.

Estes autores recomendam a utilização de procedimentos que monitorem as características de qualidade, em substituição ao monitoramento de resíduos. Esta modificação visa facilitar a aplicação do gráfico de controle de regressão múltipla, pois combina informações sobre a dinâmica do processo e o controle estatístico em um único gráfico.

3.1.1 Suposições

Antes da descrição do método proposto é necessário apresentar as quatro suposições necessárias à aplicação deste: (i) a característica de qualidade deve ter uma relação de causa-e-efeito com pelo menos uma variável de controle do processo, (ii) a característica de qualidade do processo deve ser uma variável contínua e sua relação com as variáveis de controle deve ser representável por um modelo de regressão linear simples ou múltipla, (iii) com o processo sob controle, os resíduos da regressão devem ser normalmente e independentemente distribuídos, com média zero e desvio-padrão constante e (iv) assume-se que os dados utilizados para o ajuste da regressão na Fase I do método e os dados monitorados na Fase II devem ser ter o mesmo comportamento, ou seja, apresentarem a mesma distribuição de probabilidade.

A primeira condição é relativa à relação de causa-e-efeito entre a característica de qualidade monitorada e as variáveis de controle do processo. A verificação dessa suposição pode ser realizada através de uma pré-investigação técnica do processo.

A segunda suposição, de que a relação entre essas variáveis deve ser ajustada por um modelo de regressão linear simples ou múltipla, é referente ao tipo específico de modelo que pode ajustar os dados. De maneira geral, a relação entre duas ou mais variáveis pode ser linear ou não linear quanto aos parâmetros, mas, o escopo do presente trabalho contempla apenas modelos lineares. Também se inclui algumas transformações lineares das variáveis analisadas, fato que não impede um ajuste através de um modelo de regressão linear múltipla. Ressalta-se que para a utilização de um modelo de regressão linear, é necessário que a característica de qualidade seja contínua, caso contrário, outros modelos devem ser utilizados.

A terceira suposição, referente à normalidade dos resíduos da regressão, é segundo Montgomery et al. (2001), Neter et al. (2005) e Weisberg (2005) uma condição necessária à utilização do método dos mínimos quadrados ordinários para estimar os parâmetros do modelo de regressão que expresse a relação entre as variáveis do problema. Esta condição deve ser verificada após a estimação do modelo de regressão, de forma a validar o modelo

estimado. A análise de que o desvio-padrão seja constante pode ser analisada por um gráfico de dispersão entre os valores dos resíduos e os valores ajustados. A suposição de normalidade pode ser verificada por um gráfico de probabilidade normal ou por testes de normalidade, como os testes de Anderson-Darling, Kolmogorov-Smirnov, Shapiro-Wilk, Ryan-Joiner, entre outros. A suposição de independência será feita com o uso do teste de Durbin-Watson.

A quarta suposição, ressaltada por Faltin et al. (1997), é necessária pois os limites de controle, obtidos através da Fase I do método serão válidos para o monitoramento do processo na Fase II se os dados utilizados em ambas as fases apresentem o mesmo comportamento. A existência de muitos pontos fora de controle, sem indícios de que o processo esteja realmente fora de controle, pode indicar uma alteração no modelo ajustado na Fase I.

3.1.2 Aplicações

O método proposto pode ser aplicado a qualquer processo que apresente uma característica de qualidade contínua linearmente dependente de variáveis de controle do processo, desde que as quatro suposições apresentadas anteriormente sejam satisfeitas.

3.2 DETALHAMENTO DO MÉTODO PROPOSTO

Nessa subseção são detalhadas todas as etapas das duas fases do método proposto. Ao decorrer desse detalhamento, também serão descritos a modificação do gráfico de controle de regressão proposta pelo presente trabalho e o gráfico de controle de extrapolação.

3.2.1 Fase I – Análise Retrospectiva

Como afirmado anteriormente, a Fase I do método proposto consiste na coleta inicial de dados do processo, da estimação de um modelo de regressão linear múltipla que relacione a característica de qualidade com as variáveis de controle do processo e da verificação da estabilidade do processo no período em que a amostra inicial foi coletada. As estimativas dos coeficientes de regressão e do desvio-padrão do modelo proveniente desta fase serão utilizados para a definição dos limites de controle da Fase II.

As etapas de execução desta fase são: (i) coleta de dados, (ii) ajuste do modelo de regressão linear simples ou múltipla e (iii) verificação da estabilidade do processo, que inclui a construção do gráfico de controle de regressão e identificação de causas especiais. Um roteiro esquemático da fase I do método é apresentado na Figura 20.

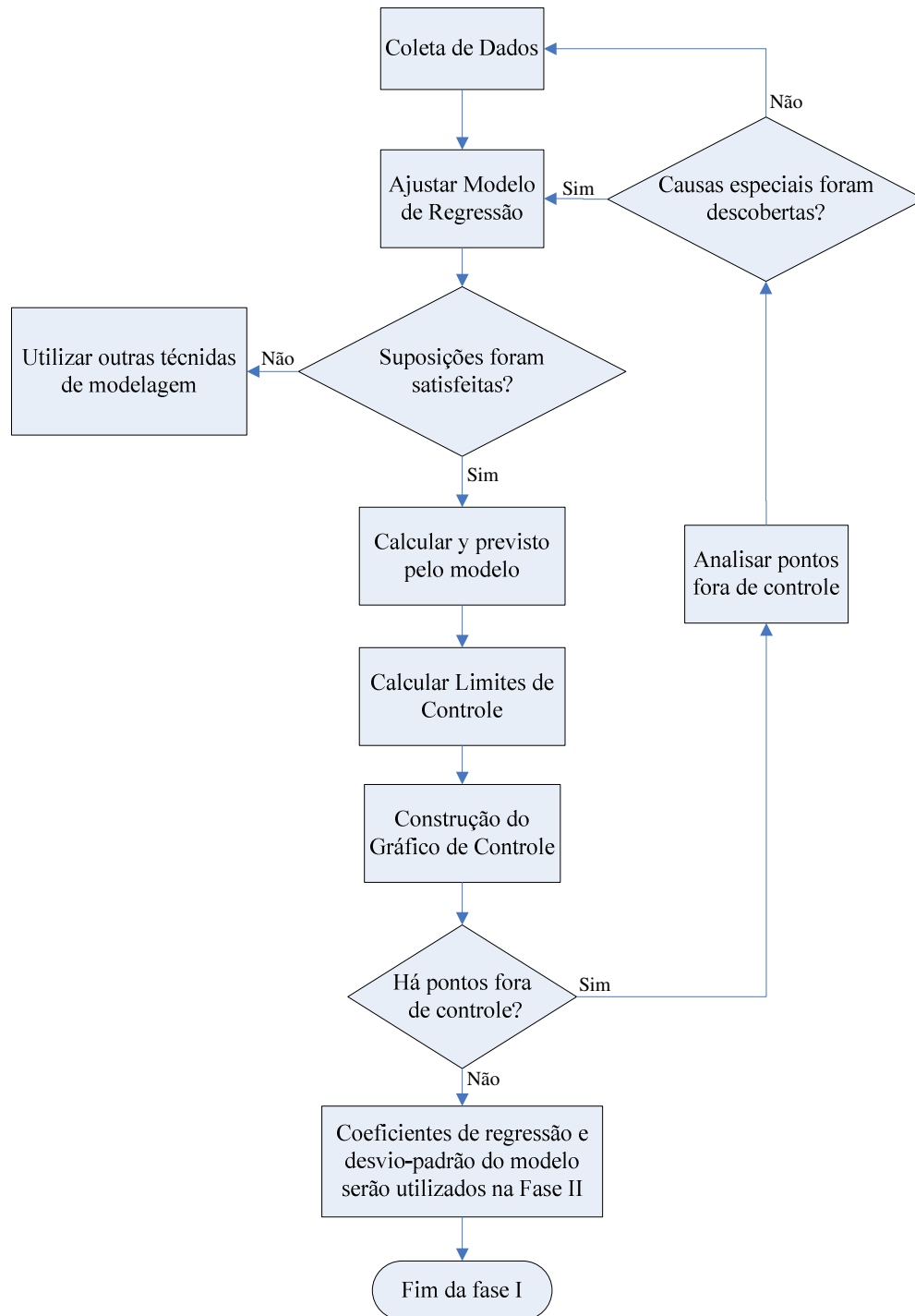


Figura 20: Fluxograma da Fase I do método proposto
 Fonte: Figura elaborada pelo autor (2009)

3.2.1.1 *Coleta de Dados*

Para a Fase I do método proposto é necessária a coleta de uma amostra suficientemente grande do processo, que permita estimar os coeficientes de regressão para todas as variáveis de controle, além de disponibilizar o maior número possível de graus de liberdade para o termo de erro do modelo, o que aumenta a confiabilidade dos resultados. Os dados coletados devem conter o valor da característica de qualidade monitorada e os valores das variáveis de controle do processo no momento de coleta. Outra alternativa para a coleta inicial de dados é a utilização da metodologia de projeto de experimentos, dessa forma, é possível garantir algumas características desejáveis a um modelo, como por exemplo a ortogonalidade das variáveis de controle.

Supõe-se que a característica de qualidade a ser monitorada seja y e as variáveis de controle do processo sejam $x_1, x_2, x_3, \dots, x_k$. Para este trabalho, pretende-se monitorar apenas valores individuais, logo devem ser coletadas m amostras de tamanho $n = 1$. Ressalta-se que para a utilização do método dos mínimos quadrados ordinários para estimar os coeficientes de regressão, o número de amostras deve ser maior que o número de variáveis de controle mais um ($m > k+1$).

Os dados devem ser ordenados na ordem temporal de coleta, o que permite um registro histórico dos dados do processo, sendo, também, indispensável para a execução da maioria dos testes de estabilidade do processo, apresentados no capítulo 2 desta dissertação.

3.2.1.2 *Ajuste do Modelo de Regressão Linear Múltipla*

Essa etapa é dividida em duas sub-etapas, a primeira referente à escolha das variáveis de controle a serem incluídas no modelo e a outra é a estimação e análise do modelo. Conforme apresentado anteriormente, os dois principais métodos para a escolha de variáveis e de modelos são o método passo a passo e o método de todas as regressões possíveis. Neste trabalho, o método passo a passo será utilizado como um filtro inicial de variáveis de controle. Dentre as abordagens do método passo a passo, será adotada a regressão por etapas, com um nível de significância de 15% tanto para a entrada quanto para a saída de variáveis do modelo.

Após a primeira sub-etapa de escolha do conjunto inicial de variáveis de controle a serem inseridas no modelo, inicia-se a segunda sub-etapa que utiliza o método dos mínimos

quadrados ordinários para estimar o modelo de regressão linear múltipla. Esta sub-etapa contempla cinco passos: (i) teste de significância do modelo de regressão; (ii) uso do fator de inflação da variância (FIV) para a verificação da presença de multicolinearidade no modelo; (iii) teste dos coeficientes individuais de regressão, de forma a verificar a significância da relação entre cada uma das variáveis de controle e a característica de qualidade; (iv) uso da distância de Cook para análise de pontos influentes na estimativa do modelo; (v) análise das suposições de que os resíduos são normalmente distribuídos com média zero e desvio-padrão aproximadamente constante.

No primeiro passo, testa-se a significância do modelo de regressão estimado, utilizando-se a estatística F proveniente da análise de variância da regressão. Neste teste, se o valor p for menor que o nível de significância adotado, a característica de qualidade é linearmente relacionada a pelo menos uma das variáveis de controle utilizadas para estimar o modelo, caso contrário, o modelo ajustado não é adequado para descrever a relação entre a característica de qualidade e as variáveis de controle.

O segundo passo é o uso do fator de inflação da variância (FIV) para a verificação da presença de multicolinearidade no modelo. Para este trabalho, considera-se um FIV_j maior que 5 como sendo um indício de multicolinearidade, conforme sugerido por Montgomery et al. (2005). Para correção do problema de multicolinearidade do modelo, poderão ser excluídas as variáveis não-significativas ou pode ser necessária a coleta de mais dados.

A seguir, testam-se os coeficientes individuais de regressão, de forma a verificar a significância da relação entre cada uma das variáveis de controle e a característica de qualidade. Neste passo, se o teste t para algum dos coeficientes de inclinação β_j apresentar um valor p maior que o nível de significância adotado, a variável de controle x_j não é estatisticamente significativa para o modelo e, desta forma, o modelo de regressão deve ser estimado novamente, com a retirada desta variável de controle.

O quarto passo é o uso da distância de Cook para análise de pontos influentes na estimativa do modelo. Neste caso, considera-se uma $F(0,5; p; n-p)$ como estatística de corte, assim caso a distância de Cook para o i -ésimo ponto seja maior que a estatística de corte, há um forte indício de que o ponto analisado é influente na estimativa dos coeficientes de regressão, conforme sugerido por Cook (1977). Caso um ponto seja influente na estimativa, é necessário investigar as causas para os valores encontrados neste ponto e, se necessário, estima-se novamente o modelo com a exclusão deste ponto.

Uma vez realizados os passos anteriores, deve-se analisar as suposições de que os resíduos são normalmente distribuídos com média zero e desvio-padrão aproximadamente

constante. Para a verificação de que os resíduos tenham desvio-padrão constante, é necessário construir um gráfico dos resíduos *versus* os valores estimados. Neste passo, deve-se realizar uma análise subjetiva, em que o padrão de comportamento dos resíduos deve ser próximo ao apresentado na Figura 3.a.

A verificação da normalidade dos resíduos pode ser realizada de maneira subjetiva através do gráfico de probabilidade normal. Para uma análise objetiva, recomenda-se o teste de normalidade de Kolmogorov-Smirnov, adotando um nível de significância de 5%. Para a verificação de que os resíduos não são auto-correlacionados, utiliza-se o teste de Durbin-Watson, apresentado nas equações (37) e (38), adotando-se um nível de significância de 5% para o teste bilateral. Ressalta-se que o teste de Durbin-Watson bilateral é composto por dois testes unilaterais, com um nível de 2,5% de significância cada. Caso a suposição dos resíduos não seja satisfeita, deve-se utilizar outros tipos de modelos.

3.2.1.3 Verificação de Estabilidade do Processo

Uma vez estimado o modelo de regressão, é possível obter as estimativas da resposta média da característica de qualidade e, assim, estimar os resíduos do modelo através da equação. O gráfico de controle proposto por Haworth (1996) consistia no monitoramento dos resíduos padronizados.

A modificação proposta pode ser obtida através da multiplicação dos resíduos padronizados, os limites de controle e a linha central, apresentados nas equações (82) e (83), pela estimativa do desvio-padrão do modelo, calculada pela raiz quadrada do QMR. Dessa forma, altera-se o gráfico de controle, de forma a serem monitorados os resíduos ao invés do monitoramento dos resíduos padronizados. Os limites de controle e linha central para este caso são dados por:

$$LSC_i = +3\sqrt{\hat{\sigma}^2} = +3\sqrt{QMR} \quad (93)$$

$$LC_i = 0 \quad (94)$$

$$LIC_i = -3\sqrt{\hat{\sigma}^2} = -3\sqrt{QMR} \quad (95)$$

A seguir, soma-se o valor previsto \hat{y}_i aos resíduos, à linha central e aos limites de controle apresentados nas equações (93), (94) e (95). Desta forma, é possível monitorar o

valor observado da característica de qualidade, ao invés do monitoramento dos resíduos padronizados. Os limites de controle e linha central para esta modificação são:

$$LSC_i = \hat{y}_i + 3\sqrt{\hat{\sigma}^2} = \hat{y}_i + 3\sqrt{QMR} \quad (96)$$

$$LC_i = \hat{y}_i \quad (97)$$

$$LIC_i = \hat{y}_i - 3\sqrt{\hat{\sigma}^2} = \hat{y}_i - 3\sqrt{QMR} \quad (98)$$

Nota-se que os limites de controle são os mesmos que os propostos por Mandel (1969), com a diferença de que no trabalho de Mandel (1969) plotava-se a característica de qualidade em função da variável de controle, perdendo-se a ordem temporal dos dados. Na modificação proposta plota-se a característica de qualidade em ordem temporal, com o eixo-x sendo a ordem das amostras.

Embora esta modificação proposta seja simples, há significativas alterações nos gráficos de controle, já que o valor da linha central e dos limites de controle não são linhas retas, como nos trabalhos anteriores. Neste caso, a linha central e os limites de controle irão variar, de acordo com o valor previsto pelo modelo de regressão. Serão utilizados os seguintes testes de estabilidade:

- a) Teste 1: um ponto acima do LSC ou abaixo do LIC;
- b) Teste 2: seqüência de oito pontos localizados do mesmo lado da linha central;
- c) Teste 3: seis pontos em seqüência crescente ou decrescente;

Assim, caso ocorram estes três tipos de causas especiais, o processo será considerado como estando fora de controle estatístico e, dessa forma, as causas especiais que alteraram o estado do processo deverão ser investigadas. Se as causas especiais forem detectadas, a equação de regressão e os limites de controle deverão ser calculados novamente, desconsiderando-se os pontos fora de controle. Caso não existam novos pontos fora de controle, os coeficientes de regressão e o desvio-padrão do modelo serão utilizados na Fase II do método proposto.

3.2.2 Fase II – Monitoramento do Processo

A Fase II é o monitoramento propriamente dito do processo. Para esta fase, é necessário assumir que o modelo de regressão estimado na Fase I do método proposto está correto e que os dados do processo a ser monitorado apresentem o mesmo comportamento dos

dados utilizados na Fase I, já que assim é possível utilizar a equação de regressão e o desvio-padrão do modelo para calcular o valor previsto da característica de qualidade e os limites de controle, referentes ao valor das variáveis de controle do processo.

No entanto, existem dois problemas para a utilização do método proposto durante o monitoramento do processo: (i) ao se utilizar uma equação de regressão calculada com base em um conjunto de dados anteriores, está-se trabalhando com a predição de novas observações; (ii) é possível que os novos dados extrapolem o região dos dados utilizados para a estimação da equação de regressão. A solução destes dois problemas será apresentada posteriormente, no decorrer da apresentação da Fase II do método proposto.

As etapas de execução da Fase II são: (i) coleta de dados e (ii) monitoramento do processo propriamente dito, que é representado por duas sub-etapas - monitoramento das variáveis de controle e monitoramento da característica de qualidade do processo. O fluxograma da Fase II do método proposto é apresentado na Figura 21.

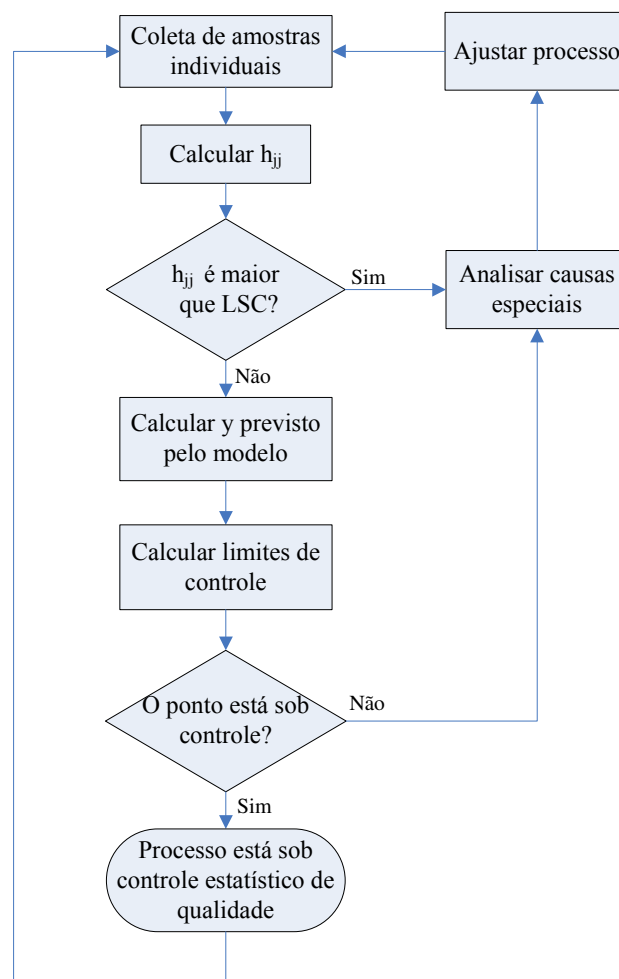


Figura 21: Fluxograma da Fase II do método proposto
Fonte: Figura elaborada pelo autor (2009)

3.2.2.1 *Coleta de Dados*

Para a Fase II do método proposto é necessária a coleta de amostras individuais periódicas do processo, coletadas em intervalos regulares de tempo. Ressalta-se que os dados coletados devem conter o valor da característica de qualidade monitorada e respectivos valores das variáveis de controle do processo.

3.2.2.2 *Monitoramento do Processo*

O monitoramento do processo é sub-dividido em monitoramento das variáveis de controle e monitoramento da característica de qualidade.

3.2.2.2.1 *Monitoramento da Extrapolação das Variáveis de Controle*

Antes da construção do gráfico de controle para monitorar a característica de qualidade do processo, é necessário verificar se os valores das variáveis de controle extrapolaram a região original de dados utilizados para estimar o modelo de regressão, já que nestes casos, o modelo de regressão estimado inicialmente pode não apresentar boas estimativas. Para verificar a extrapolação ou não da região original dos dados propõe-se o gráfico de controle de extrapolação, que é um procedimento gráfico que visa facilitar a aplicação do método apresentado por Montgomery et al. (2001) e Weisberg (2005) para a verificação da extrapolação da região original dos dados.

O gráfico de controle de extrapolação apresenta a particularidade de possuir apenas o LSC, que é igual h_{\max} , que é o maior valor de todos os elementos diagonais h_{ii} da matriz chapéu da Fase I do método proposto. A variável monitorada pelo gráfico de controle de extrapolação é h_{jj} , que é o valor da diagonal da matriz chapéu para o vetor x_j da j -ésima amostra monitorada:

$$h_{jj} = x_j'(X'X)^{-1}x_j \quad (99)$$

Como é possível observar na equação (99), h_{jj} é calculado com base nos dados utilizados para estimar o modelo de regressão da Fase I, representados pela inversa da matriz

$X'X$, e com os valores das variáveis de controle da amostra coletada. Ressalta-se que um h_{jj} igual a 0 corresponde a todas as variáveis de controle ajustadas exatamente no meio do intervalo dos dados utilizados para estimar o modelo de regressão e este valor aumenta à medida que as variáveis de controle se afastam da metade do intervalo. Para o caso de $h_{jj} = \text{LSC}$, o ponto definido pelo vetor x_j está localizado exatamente na fronteira da região original definida pelos dados utilizados para estimar o modelo de regressão, não caracterizando uma extrapolação.

Se o monitoramento for realizado de forma *on-line*, não se tem previamente o cálculo dos valores h_{jj} de todas as amostras. Conseqüentemente, para cada nova amostra coletada, deve-se inicialmente calcular o h_{jj} e comparar com o LSC do gráfico de controle de extrapolação. Caso $h_{jj} < \text{LSC}$, o valor coletado está contido nos intervalos de investigação das variáveis de controle utilizados na Fase I e pode ser monitorado pelo gráfico de controle de regressão. Na seqüência calcula-se os limites de controle para este valor e plota-se o valor no gráfico de controle de regressão para monitorar se a nova amostra é uma causa especial de variação do processo.

Se $h_{jj} > \text{LSC}$, o gráfico de controle de extrapolação aponta o processo como estando fora de controle, já que neste caso, o ponto está fora da região original da nuvem de dados e a amostra coletada não poderá ser monitorada pelo gráfico de controle de regressão proposto, devendo ser ignorada. Caso ocorram muitas situações de extrapolação da região original dos dados, recomenda-se voltar à Fase I do método proposto, alterando-se o intervalo de variação das variáveis de controle.

Especificamente para o gráfico de controle de extrapolação, um ponto fora de controle pode indicar:

- a) Erro de ajuste do processo por parte dos operadores;
- b) Erro de coleta de dados;
- c) Alteração nos parâmetros de ajuste do processo, indicando a necessidade de início da Fase I do método proposto.

3.2.2.2.2 Monitoramento da Característica de Qualidade

Uma vez estimado o modelo de regressão e verificado se a nova amostra do processo está dentro da região dos dados originais, é possível obter, através da equação de regressão ajustada, a previsão do valor da característica de qualidade. Para o presente trabalho, o valor

previsto pelo modelo é considerado como o valor ideal para o processo, dados os valores das variáveis de controle.

Por se tratar de uma nova amostra, independente da amostra utilizada para estimar o modelo de regressão, é necessário corrigir os limites de controle apresentados na Fase I, utilizando-se a para isso o valor h_{jj} de cada amostra (HAWORTH, 1996).

$$LSC_j = \hat{y}_j + 3\sqrt{QMR(1 + h_{jj})} \quad (100)$$

$$LC_j = \hat{y}_j \quad (101)$$

$$LIC_j = \hat{y}_j - 3\sqrt{QMR(1 + h_{jj})} \quad (102)$$

Para a verificação da estabilidade do processo, serão utilizados os mesmos testes de estabilidade da Fase I. Assim, caso o ponto não esteja sob controle, o processo será considerado como estando fora de controle estatístico e, dessa forma, as causas especiais que alteraram o estado do processo deverão ser investigadas. A principais causas de pontos fora de controle podem ser:

- a) Ação significativa de um fator não controlável no processo;
- b) Existência de uma variável de controle significativa que não foi incluída no modelo;
- c) Alteração no valor do coeficiente de uma variável de controle significativa em relação ao valor estimado na Fase I.

Para este trabalho, não serão descritas as investigações necessárias para a descoberta das causas especiais que agiram no processo. Caso o gráfico de controle de regressão múltipla apresente muitos pontos fora de controle, pode ser um sinal de uma alteração na relação entre a característica de qualidade e as variáveis de controle do processo, violando-se a suposição (iv) do método proposto. Neste caso, é necessário realizar a Fase I do método proposto novamente, a fim de revisar o modelo de regressão linear múltipla estimado.

3.3 CONCLUSÃO DO CAPÍTULO

No presente capítulo, apresentou-se um método para o monitoramento de um processo utilizando o gráfico de controle de regressão múltipla. Neste capítulo, também foram apresentadas as suposições necessárias à aplicação do método proposto. Seguindo

recomendações da literatura, o método proposto foi subdividido em 2 fases: Fase I – Análise Retrospectiva e Fase II – Monitoramento do Processo.

A Fase I do método proposto consistiu na coleta inicial de dados do processo, estimação de um modelo de regressão linear múltipla que relacione a característica de qualidade com as variáveis de controle do processo e a verificação da estabilidade do processo no período em que a amostra inicial foi coletada.

Assumindo que o modelo de regressão estimado da Fase I do método seja correto, prossegue-se para a Fase II, que consiste na coleta de amostras individuais do processo e no monitoramento do processo, representado pelo monitoramento das variáveis de controle, utilizando o gráfico de controle de extrapolação, e pelo monitoramento da característica de qualidade, utilizando o gráfico de controle de regressão.

4 APLICAÇÃO DO MÉTODO PROPOSTO

O presente capítulo apresenta os resultados da aplicação do método proposto, sendo dividido em duas partes: (i) aplicação do método proposto e (ii) comparação do gráfico de controle de regressão com outros gráficos de controle. A primeira seção deste capítulo será a apresentação dos resultados da aplicação do método proposto em um sistema produtivo, onde são apresentados detalhadamente as etapas de aplicação do método apresentado no capítulo 3.

Os resultados do gráfico de controle de regressão proposto serão comparados com outros gráficos de controle descritos na revisão bibliográfica desta dissertação. Esta comparação será realizada de duas formas: (i) comparação gráfica dos resultados e (ii) comparação de sensibilidade destes gráficos, obtida através do número médio de amostras (NMA) de cada gráfico.

4.1 APLICAÇÃO DO MÉTODO EM UM SISTEMA PRODUTIVO

O método proposto foi aplicado em um processo de Extrusão de Bandas de Rodagem de uma indústria de borrachas, cujo nome será omitido. Os valores observados da característica de qualidade foram alterados, com a soma de uma constante, de forma a manter o sigilo dos dados.

4.1.1 Descrição do Processo

O processo a ser monitorado é chamado de “Linha de Extrusão de Bandas de Rodagens HF10+10Y”, que é representado na Figura 22. Este processo consiste basicamente no bombeamento do composto através de uma fieira (matriz), acoplada à saída (cabeça) da extrusora, a qual dá a forma desejada ao perfilado. No caso deste processo, a fase móvel, responsável pelo transporte do material no interior da extrusora, é a hélice. A vazão da extrusora depende de vários fatores, sendo rotação da hélice (rpm) e volume de alimentação considerados como os mais importantes.

A extrusora utilizada é chamada de Extrusora HF10+10Y, pois é um equipamento alimentado a quente, composto por duas extrusoras acopladas em Y. O trabalho foi desenvolvido apenas na extrusora posterior, que é responsável pelo maior fluxo de material.

Após a extrusão, o perfil é transportado através de uma linha de resfriamento, no início da qual existem equipamentos de monitoramento *on-line* do dimensional do produto: telecâmara (largura) e balança dinâmica.

O composto de borracha inicia seu processamento no moinho quebrador (Q), onde as cadeias poliméricas são cindidas e é gerado calor, contribuindo para a redução de viscosidade do composto (plastificação). Em seguida, o material é transferido para o moinho pré-alimentador (P), onde ocorre a plastificação adicional. A etapa final de plastificação do composto é realizada no moinho alimentador (A), de onde uma fita de alimentação com dimensões bem definidas é transportada a uma velocidade constante até o funil de alimentação da extrusora.

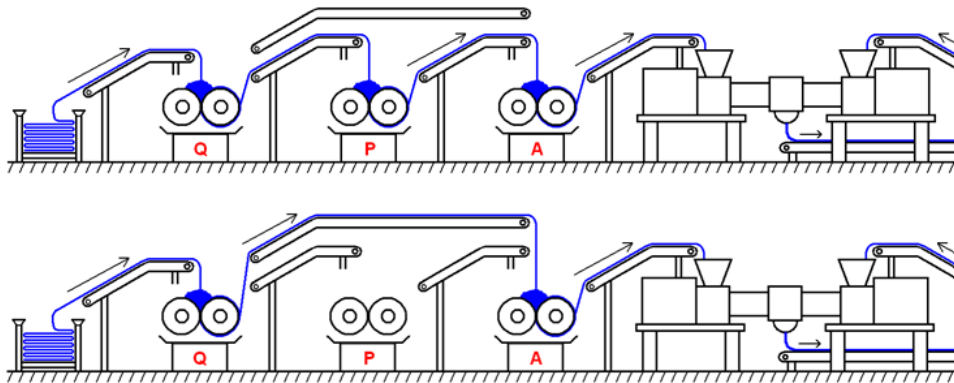


Figura 22: Esquema do processo de extrusão de bandas de rodagens
Fonte: Empresa estudada (2008)

A característica de qualidade a ser monitorada no processo estudado é o tempo de estabilização, que é medido a partir do momento em que o processo começa a operar após o ajuste de uma nova combinação de valores das variáveis de controle até o momento em que o processo se estabiliza. Esta característica de qualidade é uma variável contínua, que está diretamente relacionada à perda de matéria-prima e de tempo útil de processo, já que enquanto o processo está desestabilizado o produto produzido apresenta baixa qualidade, sendo posteriormente descartado. As variáveis de controle deste processo são:

- a) x_1 - temperatura na fieira;
- b) x_2 - nível de plastificação;
- c) x_3 - parâmetro de carregamento;
- d) x_4 - fator de enchimento;

- e) x_5 - temperatura da hélice;
- f) x_6 - temperatura Y;
- g) x_7 – tempo de setup;

Em síntese, as temperaturas das diversas zonas de uma extrusora (x_1 , x_5 e x_6) são especificadas de forma a aumentar ou diminuir o coeficiente de atrito entre a parede metálica e o composto. Baixas temperaturas da hélice garantem alto coeficiente de atrito e, conseqüentemente, elevada taxa de arraste de material. Por outro lado, nas zonas fixas (canhão, cabeça, fieira, etc.) é desejável que o coeficiente de atrito seja o menor possível para facilitar o fluxo do material, o que se consegue com temperaturas mais elevadas. No primeiro caso, a redução de temperatura da hélice tem como fator limitante a redução do nível de plastificação do composto, decorrente da elevação da viscosidade. Já no segundo caso, a pré-reticulação do material (início da transformação de termoplástico para termofixo) limita o emprego de temperaturas mais elevadas.

O nível de plastificação (x_2) pode ser definido como a capacidade do material em fluir quando sujeito a uma tensão cisalhante. Em função da dificuldade em se medir esta propriedade reológica, atribuiu-se o valor -1 para a plastificação com passagem por apenas dois moinhos (quebrador e alimentador) e, como valor +1, a passagem pelos três moinhos (quebrador, pré-alimentador e alimentador). Teoricamente, espera-se que quanto maior o trabalho mecânico sobre o material, maior seja o nível de plastificação.

O parâmetro de carregamento (x_3) é a diferença percentual da velocidade de rotação da hélice utilizada para o carregamento da extrusora em relação à velocidade de rotação utilizada em regime de produção. Por exemplo, um parâmetro de carregamento igual a +1 significa que a extrusora foi carregada com a mesma velocidade de rotação utilizada para a produção, enquanto que um parâmetro de -1 significa que a máquina foi carregada mais lentamente, possibilitando maior troca térmica com todos os componentes da extrusora.

O fator de enchimento (x_4) é definido como a relação entre a rotação da hélice em regime de produção e a largura da fita de alimentação (velocidade de alimentação e espessura constantes). Esta variável controlável é uma medida do volume de material no interior da extrusora, já que para uma mesma largura de fita, à medida que se aumenta a velocidade de rotação da hélice ocorre o esvaziamento da extrusora.

Durante as operações de *set up*, o tempo de residência do material na extrusora (x_7) também afeta o nível de plastificação do composto. Como a temperatura da fita de alimentação fica em torno de 90 °C e as temperaturas das zonas termorreguladas ficam no

intervalo de 15 a 75 °C, quanto maior o tempo de *set up* maior a viscosidade do composto, reduzindo vazão e estabilidade dimensional.

Dessa forma, a suposição da existência de uma relação de causa-e-efeito entre a característica de qualidade e as variáveis de controle é satisfeita, já que há evidências técnicas do processo que uma alteração nos valores das variáveis de controle acarreta uma alteração nos valores da característica de qualidade do processo de Extrusão de Bandas de Rodagem.

4.1.2 Aplicação da Fase I

Nesta subseção, são detalhadas as etapas da Fase I do método proposto.

4.1.2.1 *Coleta de Dados*

Os dados utilizados para a Fase I foram obtidos através de um experimento realizado na empresa. Neste experimento, utilizou-se um projeto fatorial fracionado em quatro partes com as variáveis de controle ajustadas em dois níveis (2^{7-2}) resultando em 32 amostras individuais. Devido à escolha deste experimento fatorial fracionado, as interações x_4*x_6 , x_5*x_6 e x_6*x_7 estarão confundidas com outras interações entre dois fatores e não poderão ser incluídas no modelo a ser estimado. Assim, há um total de 31 efeitos candidatos a entrarem no modelo de regressão linear múltipla, sendo 7 efeitos principais, 24 efeitos de interações entre as variáveis de controle.

4.1.2.2 *Ajuste do Modelo de Regressão Linear Múltipla*

Devido ao grande número de variáveis de controle e interações, aplica-se o procedimento passo-a-passo para a pré-seleção das variáveis a serem incluídas no modelo, conforme o método proposto anteriormente. Ao aplicar o método passo-a-passo do Minitab 14, as variáveis de controle x_1 , x_4 e x_7 e as interações x_1*x_3 , x_2*x_5 , x_2*x_6 e x_3*x_4 são escolhidas para compor o modelo. O modelo, estimado pelo método dos mínimos quadrados ordinários (MQO), é apresentado na equação (103).

$$\hat{y}_i = 227,1 + 14,8x_1 - 16,9x_4 + 12,9x_7 + 11,4x_1x_3 + 14,4x_2x_5 + 11,9x_2x_6 + 8,1x_3x_4 \quad (103)$$

Este modelo apresenta um coeficiente de determinação de aproximadamente 76,8% e um coeficiente de determinação ajustado de 76,1%. A estatística C_p de Mallows para este modelo foi igual a 7,9, que é bastante próxima do número de parâmetros do modelo. Para testar a significância do modelo de regressão estimado, utiliza-se a estatística F, calculada através da análise de variância (ANOVA) da regressão, que é apresentada na Tabela 1.

Tabela 1: ANOVA para o modelo estimado

Fonte	SQ	GL	QM	F	Valor p
Regressão	38913,7	7	5559,1	11,37	0,00
Erro	11732,7	24	488,9		
Total	50646,4	31			

Fonte: Tabela elaborada pelo autor (2009)

Como o valor p para o teste de significância do modelo de regressão, apresentado na Tabela 1, é menor que o nível de significância adotado (5%) rejeita-se a hipótese nula deste teste, logo, há fortes evidências estatísticas para considerar que a característica de qualidade seja linearmente dependente de pelo menos uma variável de controle.

Seguindo o método proposto, apresenta-se na Tabela 2 o Fator de Inflação de Variância (FIV) para verificar a presença de multicolinearidade no modelo. Na Tabela 2 também são apresentados os resultados dos testes individuais para a significância dos coeficientes do modelo de regressão, que utilizam a estatística t, calculada pela equação (23).

Tabela 2: Testes t individuais e valores FIV para as variáveis de controle e coeficiente de intercepto

Termo	Estimativa	Desvio-padrão	t	Valor p	FIV
Intercepto	227,076	3,91	19,72	0,000	-
x_1	14,805	3,91	3,79	0,001	1,0
x_4	-16,922	3,91	-4,33	0,000	1,0
x_7	12,924	3,91	3,31	0,003	1,0
$x_1 * x_3$	11,373	3,91	2,91	0,008	1,0
$x_2 * x_5$	14,399	3,91	3,68	0,001	1,0
$x_2 * x_6$	11,884	3,91	3,04	0,006	1,0
$x_3 * x_4$	8,098	3,91	2,07	0,049	1,0

Fonte: Tabela elaborada pelo autor (2009)

Como é possível observar na Tabela 2, os valores do FIV de todas variáveis incluídas no modelo foram menores que 5, logo, o modelo estimado na equação (103) não apresenta multicolinearidade e as variáveis de controle incluídas no modelo podem ser consideradas independentes. Ressalta-se que este é um resultado importante, já que a multicolinearidade inflaciona as estimativas da variância e covariância dos parâmetros do modelo de regressão, prejudicando a qualidade das estimativas feitas pelo modelo. Estes resultados eram esperados, devido ao fato do modelo ter sido estimado com base em um experimento planejado para ser ortogonal.

Como o modelo não apresenta multicolinearidade, é possível realizar os testes de hipóteses para avaliar a significância dos coeficientes de regressão. Analisando os resultados apresentados na Tabela 2, é possível observar que os testes t para o coeficiente de intercepto e para os coeficientes de inclinação apresentam valores p menores que 5%, que é o nível de significância adotado. Dessa forma, conclui-se que estes coeficientes são estatisticamente significativos e, portanto, devem ser mantidos no modelo. Isto também significa que as variáveis de controle e interações incluídas no modelo têm um efeito estatisticamente significativo sobre a característica de qualidade monitorada.

Assim, a segunda suposição estabelecida para a aplicação do método proposto é satisfeita, já que a característica de qualidade do processo é uma variável contínua e sua relação com as variáveis de controles foi ajustada por um modelo de regressão linear múltipla, apresentado na equação (103).

Na Figura 23, apresenta-se a distância de Cook (D_i) para as 32 amostras utilizadas para estimar o modelo de regressão. Neste caso, considera-se que valores D_i maiores que a estatística $F(0,5; 8; 24) = 0,944$ são um forte indício de que o ponto analisado é influente na estimativa dos coeficientes de regressão. Como nenhum valor D_i foi superior a 0,944, conclui-se que não há pontos influentes entre os dados utilizados para estimar o modelo de regressão da equação (106).

Para finalizar a estimação do modelo de regressão para o tempo de estabilização, deve-se verificar se a terceira suposição do método proposto pode ser aceita, ou seja, que os resíduos sejam normal e independentemente distribuídos com média zero e desvio-padrão aproximadamente constante.

Para verificar se o desvio-padrão pode ser considerado como aproximadamente constante, apresenta-se na Figura 24 o gráfico dos resíduos *versus* os valores estimados da característica de qualidade. Analisando a Figura 24, o desvio-padrão pode ser considerado

como aproximadamente constante, já que este gráfico apresenta um comportamento bastante parecido com o gráfico a da Figura 3, apresentada no capítulo 3.

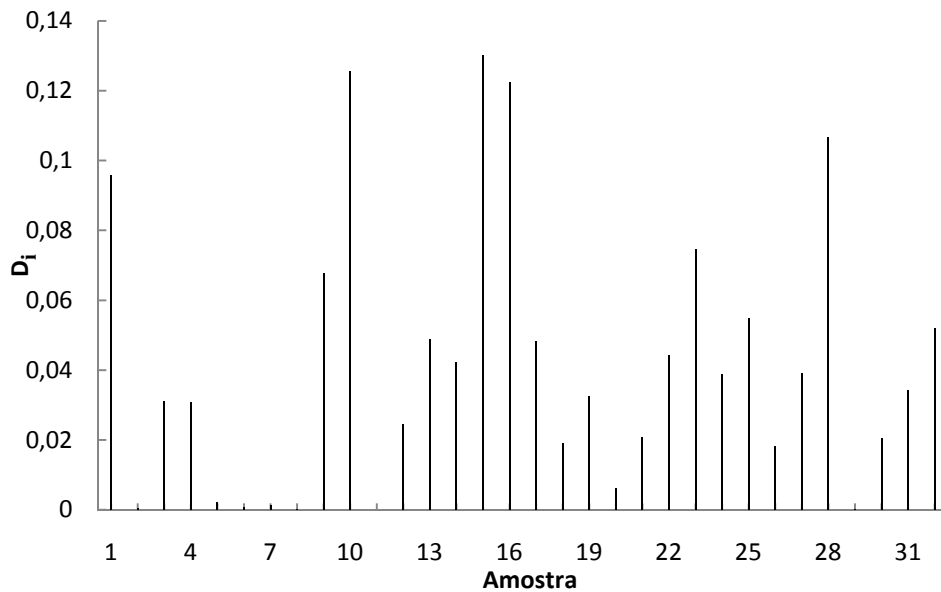


Figura 23: Distância de Cook para as amostras utilizadas para estimar o modelo
 Fonte: Figura elaborada pelo autor (2009)

Para analisar se os resíduos apresentam uma distribuição normal, apresenta-se o gráfico de probabilidade normal na Figura 25. Através da análise da Figura 25, pode-se considerar que os resíduos apresentam uma distribuição aproximadamente normal, já que os pontos estão muito próximos da reta que caracteriza a distribuição normal. Esta análise é confirmada pelo teste de Kolmogorov-Smirnov, que apresentou um valor p de 0,15, que é maior que o nível de significância adotado. Dessa forma, a quarta suposição é satisfeita e o modelo de regressão estimado na equação (103) é válido.

Para verificar se os resíduos são independentes, aplicou-se o teste de Durbin-Watson bilateral, com um nível de significância de 5%, o valor tabelado de d_l e d_u são de 0,886 e 1,904, respectivamente. Para os resíduos do modelo estimado da Fase I a estatística de teste d_p , para testar a autocorrelação positiva foi de 2,053, e a estatística de teste d_n , que testa a autocorrelação negativa, foi de 1,947. Como ambos os valores são superiores ao valor tabelado d_u conclui-se que os resíduos são independentes.

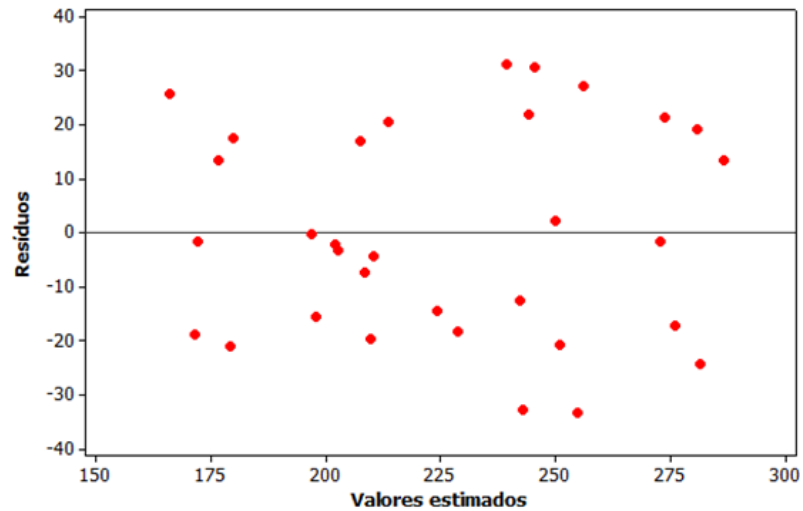


Figura 24: Gráfico dos resíduos *versus* valores estimados de y
 Fonte: Figura elaborada pelo autor (2009)

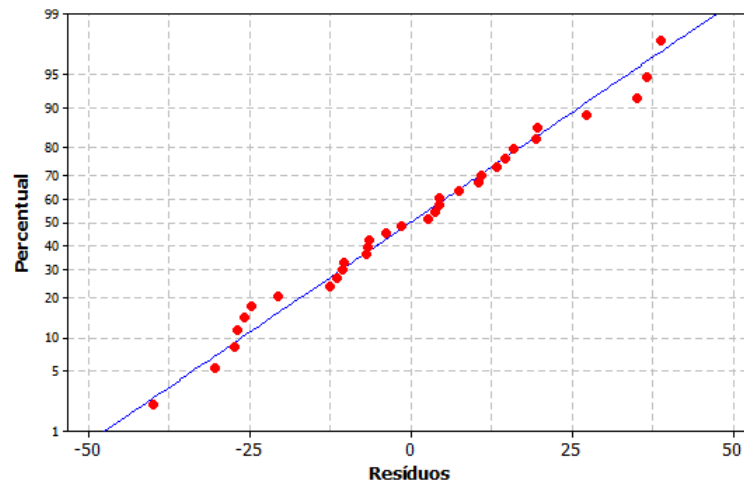


Figura 25: Gráfico de probabilidade normal dos resíduos
 Fonte: Figura elaborada pelo autor (2009)

4.1.2.3 Construção do Gráfico de Controle de Regressão da Fase I

Como o modelo estimado foi validado, é possível utilizá-lo para estimar a resposta média da característica de qualidade y , dados os valores das variáveis de controle do processo. O desvio-padrão do processo pode ser calculado através da raiz quadrada do quadrado médio dos resíduos, conforme a equação (70). Os resíduos do modelo estimado apresentam um desvio-padrão de aproximadamente 22,1 e, desta forma, os limites de controle para a Fase I do método proposto são dados por:

$$LSC_i = \hat{y}_i + 66,3 \quad (104)$$

$$LC_i = \hat{y}_i \quad (105)$$

$$LIC_j = \hat{y}_i - 66,3 \quad (106)$$

Como a característica de qualidade não pode ser menor que zero, caso o valor previsto de y_i seja menor que 66,3, o LIC_i será igual a zero. Dessa forma, o gráfico de controle de regressão múltipla para o processo de Extrusão de Bandas de Rodagem, elaborado no software Excel, é apresentado na Figura 26.

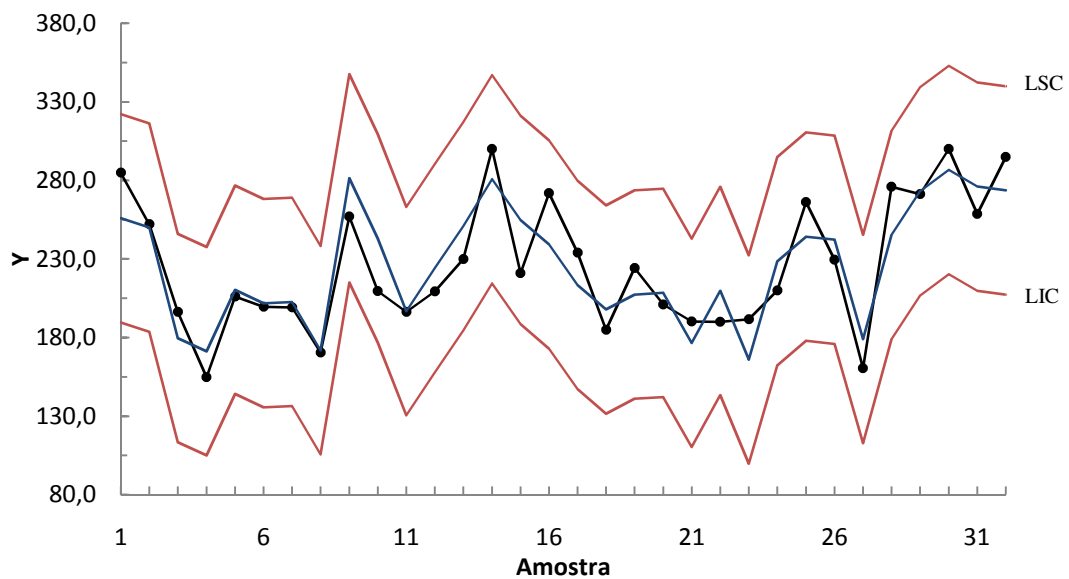


Figura 26: Gráfico de controle de regressão da Fase I do método proposto
Fonte: Figura elaborada pelo autor (2009)

Como o gráfico de controle de regressão, apresentado na Figura 26, não apresentou nenhum ponto fora de controle, o processo pode ser considerado como sob controle estatístico. Dessa forma, o modelo de regressão apresentado na equação (103) e o desvio-padrão do processo podem ser utilizados para estimar a característica de qualidade do processo e monitorar o processo de Extrusão de Bandas de Rodagem.

4.1.3 Aplicação da Fase II

Após a Fase I, inicia-se a Fase II do método proposto, que consiste basicamente na coleta de amostras individuais provenientes do processo e no monitoramento do processo, que é representado pela verificação da extrapolação das variáveis de controle e, por fim, pelo

monitoramento da característica de qualidade do processo. Os dados coletados devem conter: o valor da característica de qualidade e os valores das variáveis de controle, no momento de coleta dos dados. Para a aplicação da Fase II do método proposto, utilizaram-se 122 amostras provenientes do processo de Extrusão de Bandas de Rodagem, coletadas pelos operadores do processo.

4.1.3.1 *Monitoramento das Variáveis de Controle*

Após a coleta dos dados, deve-se verificar se os valores do conjunto de variáveis de controle extrapolam a região formada pelos dados utilizados para estimar o modelo de regressão, o que pode prejudicar a estimativa do modelo de regressão e o desempenho do gráfico de controle de regressão. Utiliza-se a equação (99) para calcular os elementos h_{jj} referentes à j -ésima amostra coletada, lembrando que para o cálculo desta variável é necessário utilizar os valores amostrados de todas variáveis de controle incluídas no modelo e a inversa da matriz $\mathbf{X}'\mathbf{X}$ dos dados utilizados para estimar o modelo de regressão na Fase I do método proposto.

Se o monitoramento for realizado de forma *on-line*, não se tem previamente o cálculo dos valores h_{jj} de todas as amostras. Conseqüentemente, para cada nova amostra coletada, deve-se inicialmente calcular o h_{jj} e comparar com o h_{\max} dos dados utilizados no ajuste do modelo de regressão na Fase I. Caso $h_{jj} < h_{\max}$, o valor coletado está contido nos intervalos de investigação das variáveis de controle utilizados na Fase I e pode ser monitorado pelo gráfico de controle de regressão. Na seqüência calcula-se os limites de controle para este valor e plota-se o valor no gráfico de controle de regressão para monitorar se a nova amostra é uma causa especial de variação do processo.

O LSC do gráfico de controle de extrapolação é igual a 0,25, já que este foi o maior valor dos h_{ji} dentre os vetores de variáveis de controle utilizados para estimar o modelo de regressão na Fase I. Dessa forma, se $h_{jj} > 0,25$ a amostra coletada está extrapolando a região original de dados, devendo ser descartada para a utilização no gráfico de controle de regressão proposto. Os valores h_{jj} calculados para as 122 amostras coletadas são apresentados na Figura 27.

Como é possível observar na Figura 27, a amostra 41 apresentou um valor de h_{jj} superior ao LSC de 0,25 e, portanto, é considerada como sendo uma amostra que extrapola a região formada pelos dados utilizados para estimar o modelo de regressão. Assim, conforme o

método proposto, esta amostra será desconsiderada do gráfico de controle de regressão da Fase II.

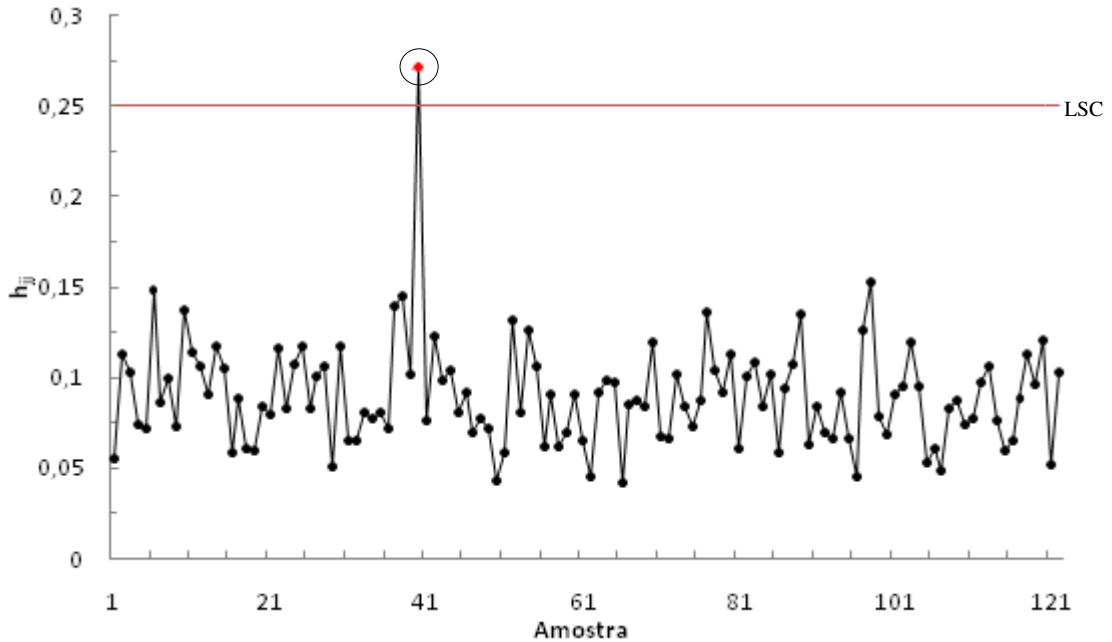


Figura 27: Gráfico de controle de extrapolação
Fonte: Figura elaborada pelo autor (2009)

4.1.3.2 Monitoramento da Característica de Qualidade

Após a exclusão da amostra 41, que extrapolou a região dos dados utilizados para estimar o modelo de regressão, restam 121 amostras que serão utilizadas no gráfico de controle de regressão proposto. Como os resíduos do modelo apresentaram um desvio-padrão de aproximadamente 22,1, os limites de controle para o monitoramento de novas observações do processo são apresentados nas equações (107) e (109):

$$LSC_j = \hat{y}_j + 66,3\sqrt{1 + h_{jj}} \quad (107)$$

$$LC_j = \hat{y}_j \quad (108)$$

$$LIC_j = \hat{y}_j - 66,3\sqrt{1 + h_{jj}} \quad (109)$$

O gráfico de controle de regressão para o monitoramento do processo de Linhas de Extrusão de Bandas de Rodagem é apresentado na Figura 28. Ressalta-se que as amostras fora de controle estão assinaladas no gráfico.

O processo foi considerado como estando fora de controle estatístico, já que o gráfico de controle de regressão proposto indicou que as amostras 2 e 30 estavam fora da região definida pelos limites de controle. Também se observa no gráfico de controle da Figura 28 que não ocorreram seqüências de oito pontos localizados do mesmo lado da linha central ou seqüências de seis pontos em ordem crescente ou decrescente.

O gráfico de controle de regressão proposto visa detectar se o valor da variável de qualidade monitorada apresentou uma grande diferença em relação ao valor estimado pelo modelo, que possa ser considerado como devido a alguma causa especial de variação do processo. Conforme afirmado anteriormente no método proposto, a etapa de investigação da causa especial não será descrita no presente trabalho.

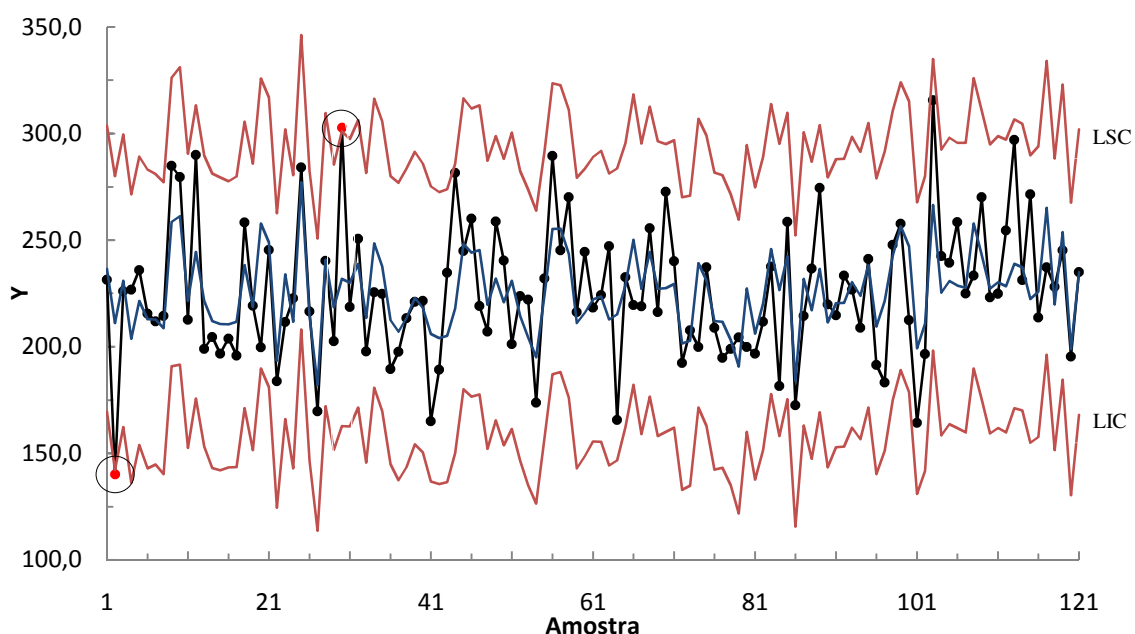


Figura 28: Gráfico de controle de regressão da Fase II do método proposto
Fonte: Figura elaborada pelo autor (2009)

4.2 COMPARAÇÃO COM OUTROS GRÁFICOS DE CONTROLE

Nesta subseção, o gráfico de controle de regressão será comparado com os gráficos de controle de regressão múltipla, com o gráfico de controle para os resíduos com desvio-padrão estimado pela amplitude móvel dos resíduos e com os gráficos EWMA e CUSUM para resíduos de um modelo de regressão, chamados de EWMAREG e CUSUMREG,

respectivamente. Esta comparação será realizada de duas formas: (i) comparação gráfica e (ii) obtenção do NMA e comparação de sensibilidade.

A comparação gráfica será a aplicação destes gráficos de controle aos dados utilizados na seção 4.1 deste trabalho e a análise dos resultados da aplicação de cada gráfico. A comparação de sensibilidade dos gráficos de controle será realizada através da obtenção do NMA de cada gráfico, através de simulação de Monte Carlo e solução de equações integrais no *software* R.

4.2.1 Comparação Gráfica

Conforme afirmado anteriormente, os gráficos de controle EWMAREG, CUSUMREG e gráfico de medidas individuais podem ser adaptados de forma que os resíduos ou os resíduos padronizados do modelo de regressão sejam as variáveis monitoradas por estes gráficos. O gráfico de controle de regressão múltipla, apresentado por Haworth (1996), também será apresentado, já que esse é similar ao gráfico de controle proposto. Para essas técnicas, será utilizado o modelo de regressão apresentado na equação (103) e somente será apresentado o respectivo gráfico de controle da Fase II de cada uma dessas técnicas.

O primeiro gráfico de controle a ser utilizado é o gráfico de controle de regressão múltipla, proposto por Haworth (1996). Conforme apresentado anteriormente, este gráfico possui os resíduos padronizados na forma *student* como característica de qualidade monitorada, com limites de controle definidos através de uma distribuição t-Student, com n-p graus de liberdade. Adotando um nível de confiança de 0,0027, e substituindo nas equações (80) e (82), os limites de controle para este gráfico são dados por:

$$LSC = 3,34 \quad (110)$$

$$LC = 0 \quad (111)$$

$$LIC = -3,34 \quad (112)$$

A linha central do gráfico de controle de regressão múltipla é igual a zero, que deve ser a média dos resíduos padronizados de qualquer modelo de regressão linear estimado corretamente. Neste caso, o gráfico de controle de regressão múltipla, apresentado na Figura 29, não apontou nenhuma amostra como estando fora de controle.

Uma alternativa ao procedimento proposto por Haworth (1996), é estreitar os limites de controle, utilizando limites de controle 3 sigma, o que é similar aos gráficos de controle

clássicos. Esta adaptação apontaria as amostras 2 e 30 como estando fora de controle, ou seja, apresenta os mesmos resultados do gráfico de controle proposto.

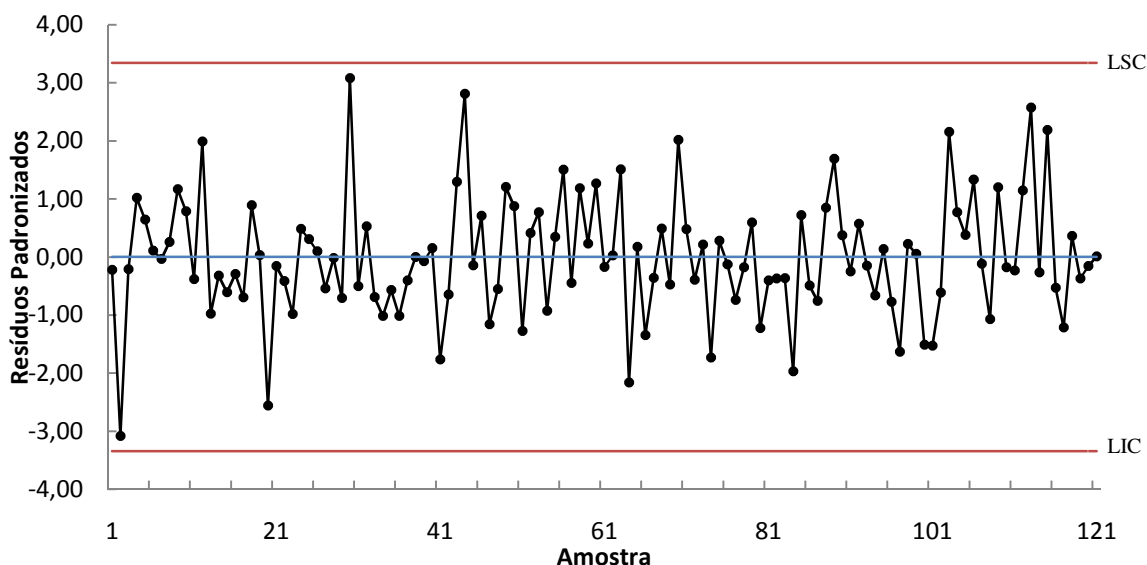


Figura 29: Gráfico de controle de regressão múltipla para os resíduos do modelo estudado
Fonte: Figura elaborada pelo autor (2009)

O segundo procedimento a ser comparado é o uso da amplitude móvel dos resíduos do modelo de regressão para estimar o desvio-padrão do processo, conforme utilizado por Rothschild e Roth (1986) e Loredó et al. (2002). Nos resultados obtidos, a amplitude móvel para os resíduos do modelo estimado pela equação (103) foi de 26,6. Desta forma, aplicando as equações apresentadas na Figura 7, os limites de controle para os gráficos de controle de medidas individuais para os resíduos do modelo de regressão são:

$$LSC = 71,0 \quad (113)$$

$$LC = 0 \quad (114)$$

$$LIC = -71,0 \quad (115)$$

Nota-se que o uso da amplitude móvel gerou limites de controle ligeiramente mais largos que o gráfico de controle proposto, já que esta técnica superestimou o desvio-padrão do processo. O gráfico de controle para os resíduos, com o desvio-padrão estimado pela amplitude móvel dos resíduos é apresentado na Figura 30. Como é possível observar, este gráfico indicou apenas a amostra 30 como estando fora de controle, enquanto a amostra 2 não foi indicada.

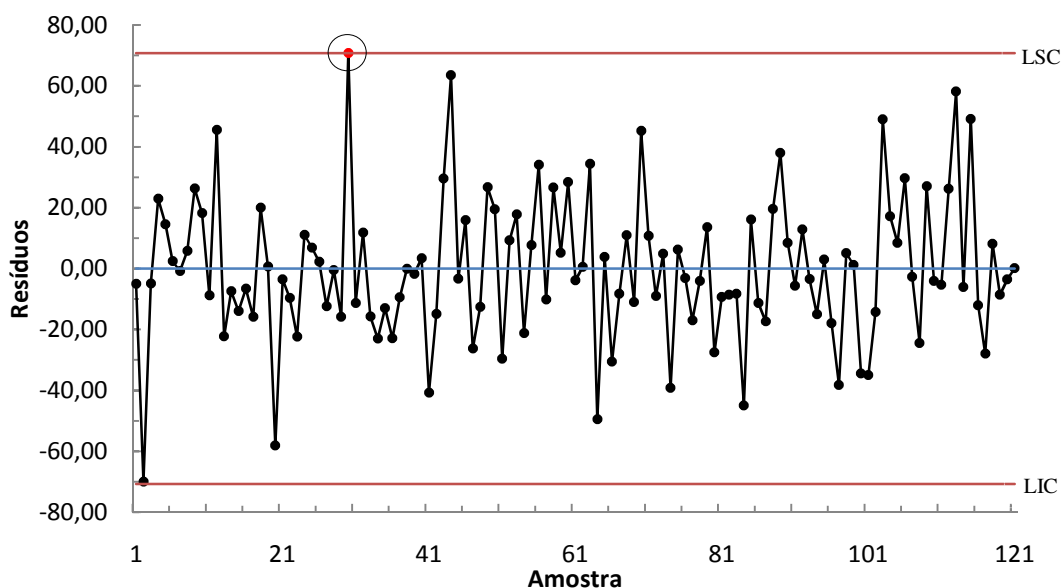


Figura 30: Gráfico de medidas individuais para o processo de extrusão
 Fonte: Figura elaborada pelo autor (2009)

Shu et al. (2004) propuseram a aplicação do gráfico de controle de EWMAREG, que é o gráfico EWMA aplicado aos resíduos padronizados do modelo de regressão. Na Figura 31 apresenta-se o gráfico EWMAREG, adotando-se $\lambda = 0,15$ e $L = 2,8$. Estes valores, segundo Crowder (1989) e Lucas e Sacucci (1990), conferem um NMA_0 de 370, quando o processo está sob controle, e um NMA_1 igual a 2,8, para mudanças de 3σ na média.

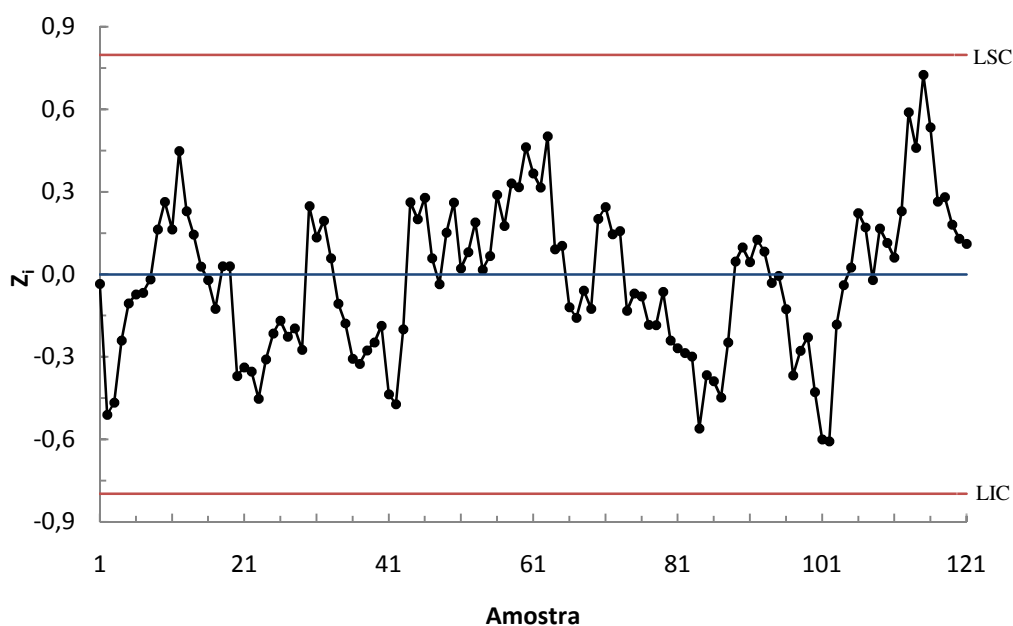


Figura 31: Gráfico EWMAREG para os resíduos padronizados do modelo
 Fonte: Figura elaborada pelo autor (2009)

Destaca-se que a característica monitorada neste gráfico são os resíduos padronizados exponencialmente ponderados, que podem ser obtidos através da aplicação direta da equação (85). O gráfico apresentado na Figura 31 não apontou nenhuma amostra como estando fora de controle.

Na Figura 32, apresenta-se o gráfico de controle CUSUM para os resíduos padronizados, adotando-se $k = 0,5$ e $h = 4,77$. De acordo Vance (1986) e Gan (1991) estes valores para as constantes fazem com que o NMA_0 sob controle do gráfico CUSUM seja de 370 e, para mudanças de 3σ na média, o NMA_1 seja aproximadamente igual a 2,48. Em analogia ao gráfico EWMAREG de Shu et al. (2004), este gráfico será chamado de CUSUMREG. Como é possível observar, este gráfico não indicou nenhuma amostra como estando fora de controle.

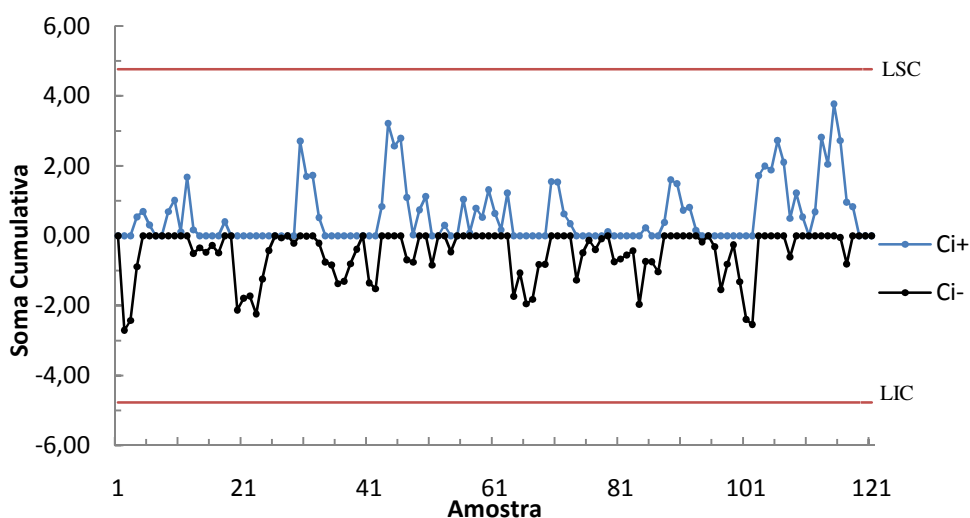


Figura 32: Gráfico CUSUMREG para os resíduos padronizados do modelo
Fonte: Figura elaborada pelo autor (2009)

É importante destacar que, de forma geral, os gráficos EWMA e CUSUM são mais sensíveis às mudanças na ordem de 1 a 2σ , mas respondem de forma mais lenta a mudanças isoladas como as detectadas, quando comparados aos gráficos de controle de Shewhart (MONTGOMERY, 2004). Desta forma, estes gráficos não foram capazes de apontar as amostras 2 e 30 como estando fora de controle, já que estas ocorreram de forma isolada.

Concluindo a comparação gráfica apresentada, o gráfico de controle proposto apontou dois pontos como estando fora de controle, enquanto o gráfico de controle de regressão de Haworth (1996) e os gráficos EWMAREG e CUSUMREG não apontaram nenhum ponto como estando fora de controle e o gráfico de medidas individuais aplicado aos

resíduos não apontou nenhum ponto. No entanto, é necessário comentar que, por não ter sido realizada uma investigação das causas especiais que poderiam ter causado essa alteração, não se sabe se os pontos apontados se referem ou a amostras em que o processo estava realmente fora de controle.

4.2.2 Comparação de Sensibilidade

Nesta subseção, apresentam-se o método adotado para a obtenção do NMA e a comparação com os resultados de outros procedimentos encontrados na literatura.

4.2.2.1 Descrição do Método Adotado

A comparação de sensibilidade do gráfico de controle proposto com os demais gráficos utilizados na subseção 4.2.1 será realizada através do cálculo do NMA de cada gráfico de controle, para determinadas alterações nos parâmetros do modelo de regressão que foi estabelecido na Fase I do método proposto. Devido a diferenças entre os gráficos de controle apresentados, serão utilizados dois métodos para o cálculo do NMA: (i) simulação de Monte Carlo e (ii) solução de equações integrais.

O método de simulação de Monte Carlo será utilizado para calcular o NMA para o gráfico de controle de regressão proposto, para o gráfico de controle de regressão múltipla proposto por Haworth (1996) e para o gráfico de controle com limites calculados através da amplitude móvel dos resíduos proposto por Rothshild e Roth (1986) e Loredó et al. (2002). Para maiores informações sobre o método de Monte Carlo, recomenda-se Bremaud (1999) e Casella e Robert (2004). Para os gráficos de controle EWMA e CUSUM, que monitoram os resíduos padronizados do modelo de regressão, solucionam-se as equações integrais apresentadas em (63), (64) e (65). Em todos os casos utiliza-se o *software* R para o cálculo do NMA de cada método.

O primeiro passo da simulação consiste na geração da amostra a ser monitorada pelos gráficos de controle que serão analisados. gerando-se vetores de dados até que um ponto fora de controle seja detectado. Como a variável de controle x_2 é uma variável discreta, foram sorteados valores -1 e 1, sem incluir valores intermediários. O restante das variáveis de

controle, por serem contínuas, foram modeladas segundo uma distribuição uniforme, cujos valores possuem a mesma probabilidade de ocorrência, com valor mínimo de -1 e valor máximo de +1. Para compor o erro aleatório, gera-se um valor obtido através de uma distribuição normal com média 0 e desvio-padrão 22,1.

Para a simulação de um processo sob controle, utiliza-se o modelo estimado na Fase I do método proposto, para gerar a previsão da característica de qualidade dado cada vetor de variáveis de controle. A seguir, soma-se o resíduo aleatório, de forma a constituir o valor simulado da característica de qualidade.

No caso da simulação de um processo fora de controle, altera-se a equação (103) conforme a modificação desejada e, ao se somar os resíduos aleatórios a este resultado, gera-se o vetor de valores simulados da característica de qualidade. Segundo Shu et al. (2004), a modificação dos coeficientes do modelo de regressão pode alterar a média e o desvio-padrão dos resíduos, conforme apresentado nas equações (116) e (117).

$$E(\mathbf{e}) = (\beta_0 - \hat{\beta}_0) + (\beta_1 - \hat{\beta}_1)\mu_{x_1} + (\beta_2 - \hat{\beta}_2)\mu_{x_2} + \dots + (\beta_k - \hat{\beta}_k)\mu_{x_k} \quad (116)$$

$$V(\mathbf{e}) = (\beta_1 - \hat{\beta}_1)^2 \sigma_{x_1}^2 + (\beta_2 - \hat{\beta}_2)^2 \sigma_{x_2}^2 + \dots + (\beta_k - \hat{\beta}_k)^2 \sigma_{x_k}^2 + \sigma^2 \quad (117)$$

onde: $\hat{\beta}_k$ é o valor estimado do k -ésimo coeficiente de regressão

β_k é o valor alterado do k -ésimo coeficiente de regressão

Em ambos os casos, a amostra gerada é comparada com os limites de controle obtidos através do método proposto no presente trabalho e dos métodos da amplitude móvel, proposto por Rothshild e Roth (1986) e Loredó et al. (2002) e pelo método proposto por Haworth (1996). Cada gráfico possui seu contador específico, que recebe o valor da amostra do primeiro ponto fora de controle. Não compara-se as novas amostras obtidas com os limites de controle do gráfico que apresentou uma amostra fora de controle e o processo continua até que todos os pontos tenham um ponto fora de controle.

Quando um ponto fora de controle é detectado por algum dos gráficos, contador $C(i)$, referente ao número de ocorrências de pontos fora de controle para a amostra detectada, recebe o valor anterior mais um. Repetindo-se todo o procedimento descrito anteriormente 5 mil vezes é possível calcular o NMA através da equação (118):

$$NMA = \frac{\sum_{i=1}^n iC(i)}{5000} \quad (118)$$

Para os gráficos de controle CUSUMREG e EWMAREG calcula-se, para cada alteração inserida nos coeficientes de regressão, a variação correspondente na média dos resíduos padronizados. Essa alteração é utilizada como entrada nas equações integrais das

equações (63), para o gráfico EWMAREG, e equações (64) e (65) para o gráfico CUSUMREG. Estas equações são resolvidas utilizando-se o método de quadratura de Gauss-Legendre com 40 pontos, através do pacote SPC do software R. Ressalta-se que o gráfico EWMAREG será aplicado apenas às alterações na média do processo, já que, segundo Kang e Albin (2000), o gráfico EWMA não apresenta bom desempenho no monitoramento de alterações na variabilidade dos processos. Como o gráfico CUSUMREG apresenta, em conceito, desempenho similar ao gráfico EWMAREG, também não será aplicado para o caso de alterações na variabilidade.

4.2.2.2 *Resultados da Simulação*

Além da situação de processo sob controle estatístico, foram simuladas três situações de processos fora de controle: (i) alterações no coeficiente de intercepto β_0 , com os coeficientes de inclinação constantes; (ii) alterações no coeficiente de inclinação β_1 , com os demais coeficientes de regressão mantidos constantes e (iii) alterações nos coeficientes β_0 e β_1 , com os demais coeficientes de inclinação mantidos constantes. As explicações dos motivos destas alterações serão informadas a seguir.

Analisando estas equações, uma alteração no coeficiente de intercepto β_0 visa alterar a média dos valores simulados da característica de qualidade e, conseqüentemente, a média dos resíduos do modelo, mas mantendo o desvio-padrão dos resíduos constante. Dessa forma, a primeira situação simulada visa analisar isoladamente a sensibilidade do gráfico de controle de regressão proposto às modificações na média dos valores observados em relação aos valores estimados pelo modelo de regressão estimado na Fase I do método proposto, já que apenas a média dos resíduos é alterada, enquanto a variância permanece constante.

De acordo com as equações (116) e (117), uma alteração em um coeficiente de inclinação β_k altera a variância dos resíduos do modelo de regressão e, se a média da variável de controle x_k não for igual a zero, altera também a média dos resíduos. Como as variáveis de controle simuladas apresentam média igual a zero, uma alteração em qualquer coeficiente de inclinação altera apenas o desvio-padrão, que é o objetivo da segunda situação de processo fora de controle.

Para a terceira situação fora de controle simulada, altera-se simultaneamente o coeficiente de intercepto β_0 e o coeficiente de inclinação β_1 , visando assim, verificar a sensibilidade do gráfico de controle proposto às alterações conjuntas da média e desvio-

padrão dos resíduos. Ressalta-se que em todas essas situações, as alterações nos coeficientes são planejadas em valores múltiplos do desvio-padrão dos resíduos estimado na aplicação da Fase I do método proposto.

Na Tabela 3 apresentam-se os valores do NMA, obtidos através da simulação, para o processo sob controle e para as alterações do coeficiente de intercepto β_0 do modelo de regressão, para o gráfico de controle proposto, para o gráfico de medidas individuais para os resíduos, para o gráfico de controle de regressão múltipla proposto por Haworth (1996) e para os gráficos EWMAREG e CUSUMREG.

Tabela 3: Valores NMA para alterações no coeficiente de intercepto

Alteração	Método Proposto	Medidas Individuais	Haworth (1996)	EWMAREG	CUSUMREG
Sob controle	576,75	736,63	2083,85	369,80	368,56
$\beta_0+0,5\sigma_e$	226,31	281,43	690,58	31,75	35,21
$\beta_0+1,0\sigma_e$	60,13	72,20	158,33	9,58	9,92
$\beta_0+1,5\sigma_e$	19,47	22,74	43,58	5,40	5,52
$\beta_0+2,0\sigma_e$	7,81	8,82	14,98	3,80	3,86
$\beta_0+2,5\sigma_e$	3,83	4,16	6,28	2,98	3,00
$\beta_0+3,0\sigma_e$	2,24	2,40	3,19	2,47	2,48
$\beta_0+3,5\sigma_e$	1,54	1,60	2,00	2,16	2,16
$\beta_0+4,0\sigma_e$	1,24	1,27	1,44	1,96	1,96

Fonte: Tabela elaborada pelo autor (2009)

Para a situação de processo sob controle, maiores valores do NMA_0 indicam menores taxas de alarmes falsos do gráfico de controle e, assim, menos intervenções desnecessárias no processo serão realizadas. Mas, altos valores do NMA_0 tem como consequência uma relativa ineficiência do gráfico de controle, já que os limites de controle podem estar muito largos.

Assim, analisando a Tabela 3, o gráfico de controle de regressão múltipla proposto por Haworth (1996) apresentou o maior NMA_0 e, por conseguinte, a menor taxa de alarmes falsos. Isto ocorreu porque este gráfico apresentou os limites de controle mais largos. De forma contrária, os gráficos EWMAREG e CUSUMREG apresentaram os menores NMA_0 entre os gráficos estudados, com valores NMA_0 ligeiramente inferiores a 370. O gráfico de controle para medidas individuais, com limites de controle estimados pela amplitude móvel dos resíduos, apresentou um NMA_0 de aproximadamente 730, já que o uso deste método forneceu limites de controle um pouco maiores que o gráfico de controle proposto.

O gráfico de controle de regressão proposto apresentou um NMA_0 de aproximadamente 572, ou seja, o gráfico de controle de regressão proposto apresentará, em média, um alarme falso a cada 572 amostras monitoradas. Este desempenho foi considerado satisfatório, embora seja superior ao NMA_0 de um gráfico de controle de Shewhart.

O principal motivo para este valor do NMA_0 é o uso da raiz de $1+h_{jj}$ como fator de correção nos limites de controle do método proposto. Este também é o fator de correção utilizado para o cálculo de intervalos de confiança para a previsão de novas observações da característica de qualidade. Caso este fator de correção não fosse utilizado, amostras com valores extremos para as variáveis de controle teriam limites de controle mais rígidos e o nível de confiança, implicitamente adotado quando se escolhe a constante k para os limites de controle, não seria constante para todos os valores monitorados.

Analisando as alterações no coeficiente de intercepto β_0 , é possível observar que o gráfico de controle de regressão múltipla de Haworth (1996) apresenta um desempenho insatisfatório se comparado com os demais gráficos de controle estudados. O NMA deste gráfico somente fica próximo dos demais para mudanças superiores a $3\sigma_e$ na média de β_0 .

Para mudanças na média de β_0 inferiores a $2,5\sigma_e$, os gráficos EWMAREG e CUSUMREG apresentaram desempenhos superiores aos dos demais gráficos, embora apresentem um NMA_0 um pouco menor, o que indica uma taxa de alarmes falsos superior aos dos demais gráficos. Quando se compara apenas estes dois gráficos, o gráfico EWMAREG possui um desempenho ligeiramente superior, já que apresenta um maior NMA_0 e um menor valor de NMA_1 para todas as mudanças em β_0 inferiores a $3,0\sigma_e$.

Para alterações na média de β_0 superiores a $2,5\sigma_e$, o gráfico de controle de regressão proposto apresenta o melhor desempenho entre todos os gráficos de controle estudados. O gráfico de controle para medidas individuais aplicado aos resíduos apresenta um desempenho muito próximo ao gráfico de controle proposto, embora apresente um NMA_1 maior. Isto indica uma proximidade entre o desempenho das duas técnicas de controle estatístico.

Na Tabela 4, apresentam-se os valores do NMA_1 para as alterações do coeficiente de inclinação β_1 dos gráficos de controle de regressão proposto, do gráfico de medidas individuais para os resíduos e do gráfico de controle de regressão múltipla proposto por Haworth (1996). Segundo Kang e Albin (2000), o gráfico EWMA não apresenta bom desempenho no monitoramento de alterações na variabilidade dos processos e, de acordo com Montgomery (2004) o gráfico CUSUM é semelhante ao gráfico EWMA. Dessa forma, os gráficos EWMAREG e CUSUMREG não serão estudados os desempenhos destes para alterações na variabilidade do processo.

Tabela 4: Valores NMA para alterações no coeficiente de inclinação β_1 .

Alteração	Método Proposto	Medidas Individuais	Haworth (1996)
Sob controle	576,75	736,63	2083,85
$\beta_1+0,5\sigma_e$	393,17	488,31	1311,10
$\beta_1+1,0\sigma_e$	172,59	204,02	502,31
$\beta_1+1,5\sigma_e$	71,90	82,10	180,61
$\beta_1+2,0\sigma_e$	32,03	35,69	70,00
$\beta_1+2,5\sigma_e$	16,04	17,52	30,71
$\beta_1+3,0\sigma_e$	9,10	9,79	15,46
$\beta_1+3,5\sigma_e$	5,82	6,18	8,90
$\beta_1+4,0\sigma_e$	4,14	4,35	5,78

Fonte: Tabela elaborada pelo autor (2009)

Analisando a Tabela 4, observa-se que o desempenho do gráfico de controle de regressão proposto por Haworth (1996) é novamente inferior aos dos demais gráficos estudados. Neste caso, o gráfico de controle proposto apresenta NMA_1 sempre menor que o gráfico de controle de medidas individuais, especialmente para alterações inferiores a $2,5\sigma_e$ na média de β_1 .

Os resultados da Tabela 4 mostram que estes três gráficos apresentados não são adequados para a detecção de mudanças no desvio-padrão dos resíduos do modelo. A principal razão para este fraco desempenho é que estes gráficos foram projetados especificamente para detectar alterações na média do valor observado em relação ao valor previsto pelo modelo e não para detectar mudanças no desvio-padrão. Isto indica a necessidade do desenvolvimento de procedimentos para detectar alterações no desvio-padrão do processo, quando este é monitorado por um gráfico baseado no gráfico de controle de regressão ou algum outro gráfico similar.

Na Tabela 5 apresentam-se os valores do NMA_1 para as alterações simultâneas do coeficiente de intercepto β_0 e do coeficiente de inclinação β_1 dos gráficos de controle de regressão proposto, do gráfico de medidas individuais para os resíduos e do gráfico de controle de regressão múltipla proposto por Haworth (1996). Nesta situação de processo fora de controle, não será apresentado o NMA dos gráficos EWMAREG e CUSUMREG. Ressalta-se que os gráficos EWMAREG e CUSUMREG serão aplicados apenas às alterações na média do processo, já que estes gráficos não apresentam bom desempenho no monitoramento de alterações na variabilidade dos processos.

Tabela 5: Valores NMA para alterações simultâneas nos coeficientes β_0 e β_1

Alterações	Método Proposto	Medidas Individuais	Haworth (1996)	
$\beta_0+0,5\sigma_e$	$\beta_1+0,5\sigma_e$	170,01	205,35	490,63
	$\beta_1+1,0\sigma_e$	89,99	104,08	227,02
	$\beta_1+1,5\sigma_e$	43,75	49,36	97,72
	$\beta_1+2,0\sigma_e$	22,54	24,89	44,34
	$\beta_1+2,5\sigma_e$	12,72	13,79	22,39
	$\beta_1+3,0\sigma_e$	7,92	8,47	12,61
$\beta_0+1,0\sigma_e$	$\beta_1+0,5\sigma_e$	49,40	57,98	120,81
	$\beta_1+1,0\sigma_e$	31,36	35,73	68,05
	$\beta_1+1,5\sigma_e$	18,82	20,87	36,07
	$\beta_1+2,0\sigma_e$	11,84	12,87	20,21
	$\beta_1+2,5\sigma_e$	8,00	8,56	12,34
	$\beta_1+3,0\sigma_e$	5,82	6,16	8,29
$\beta_0+1,5\sigma_e$	$\beta_1+0,5\sigma_e$	17,11	19,57	36,06
	$\beta_1+1,0\sigma_e$	12,63	14,12	23,92
	$\beta_1+1,5\sigma_e$	8,95	9,78	15,13
	$\beta_1+2,0\sigma_e$	6,80	7,06	10,46
	$\beta_1+2,5\sigma_e$	5,11	5,41	7,13
	$\beta_1+3,0\sigma_e$	4,20	4,40	5,48
$\beta_0+2,0\sigma_e$	$\beta_1+0,5\sigma_e$	7,20	8,03	13,14
	$\beta_1+1,0\sigma_e$	6,05	6,63	10,11
	$\beta_1+1,5\sigma_e$	4,93	5,31	7,49
	$\beta_1+2,0\sigma_e$	4,10	4,35	5,69
	$\beta_1+2,5\sigma_e$	3,54	3,71	4,58
	$\beta_1+3,0\sigma_e$	3,17	3,30	3,89
$\beta_0+2,5\sigma_e$	$\beta_1+0,5\sigma_e$	3,67	3,99	5,85
	$\beta_1+1,0\sigma_e$	3,39	3,65	5,06
	$\beta_1+1,5\sigma_e$	3,09	3,28	4,27
	$\beta_1+2,0\sigma_e$	2,83	2,98	3,66
	$\beta_1+2,5\sigma_e$	2,64	2,75	3,24
	$\beta_1+3,0\sigma_e$	2,51	2,60	3,96
$\beta_0+3,0\sigma_e$	$\beta_1+0,5\sigma_e$	2,22	2,36	3,13
	$\beta_1+1,0\sigma_e$	2,19	2,31	2,96
	$\beta_1+1,5\sigma_e$	2,16	2,26	2,76
	$\beta_1+2,0\sigma_e$	2,12	2,21	2,59
	$\beta_1+2,5\sigma_e$	2,10	2,17	2,46
	$\beta_1+3,0\sigma_e$	2,08	2,14	2,38

Fonte: Tabela elaborada pelo autor (2009)

Como é possível observar na Tabela 5, para alterações simultâneas em β_0 e β_1 , o gráfico de controle proposto possui melhor desempenho, isto é, um NMA_1 menor, que os

demais gráficos estudados em todas as situações simuladas. Quando comparado com o gráfico de controle para medidas individuais aplicado aos resíduos do modelo de regressão, o gráfico de controle proposto é notoriamente superior em mudanças inferiores a $1,5\sigma_e$ na média de β_0 e β_1 , sendo que a partir desta faixa de alteração, o desempenho destes gráficos se torna bastante similar. Para todas as faixas de alteração em β_0 , com alterações acima de $2,5\sigma_e$ em β_1 , o gráfico do método proposto e o de medidas individuais para os resíduos apresentam desempenhos muito próximos.

Analisando a Tabela 5, também é possível observar que o gráfico de controle de regressão múltipla de Haworth (1996) possui desempenho muito inferior a esses dois gráficos. Este gráfico só apresenta desempenho próximo aos dos demais gráficos em situações de grandes alterações nos coeficientes de regressão β_0 e β_1 .

Dessa forma, o gráfico de controle proposto apresentou um bom desempenho para a detecção de mudanças isoladas no coeficiente de intercepto β_0 e em mudanças simultâneas nos coeficientes de regressão β_0 e β_1 . Para mudanças simuladas no coeficiente de inclinação β_1 , que neste caso acarretam especificamente alterações multiplicativas no desvio-padrão dos resíduos, o gráfico de controle proposto teve um desempenho razoável, embora superior ao dos outros gráficos estudados.

4.3 CONCLUSÃO DO CAPÍTULO

No presente capítulo apresentou-se a validação do método proposto, dividida em duas partes: (i) aplicação deste método em um processo produtivo e (ii) comparação dos resultados do gráfico de controle de regressão proposto. Na primeira parte do capítulo, apresentou-se a descrição da aplicação do método proposto para o monitoramento do tempo de estabilização em um processo de Extrusão de Bandas de Rodagem de uma indústria de borrachas.

Na aplicação da Fase I do método proposto, apresentou-se a coleta de dados, os passos para a estimação do modelo de regressão linear múltipla, a análise de resíduos e a construção do gráfico de controle de regressão da Fase I. Na aplicação da Fase II do método proposto, apresentou-se o gráfico de controle de extrapolação para verificar se os valores das variáveis de controle estão extrapolando o intervalo de investigação original das variáveis de controle utilizado no modelo de regressão definido na Fase I. Ainda na aplicação da Fase II,

mostraram-se os resultados do uso do gráfico de controle de regressão proposto para o monitoramento do processo.

Na segunda etapa de validação do método proposto, comparou-se os resultados da aplicação do gráfico de controle de regressão proposto com outros quatro gráficos:

- a) Gráfico de controle de regressão múltipla, proposto por Haworth (1996);
- b) Gráfico de controle de medidas individuais para os resíduos, proposto por Rothschild e Roth (1986) e Loredó et al. (2002);
- c) Gráfico de controle EWMAREG, proposto por Shu et al. (2004);
- d) Gráfico de controle CUSUMREG.

Primeiramente, comparou-se graficamente o gráfico de controle de regressão proposto com estes quatro gráficos. A seguir, calculou-se o NMA destes gráficos de controle para a comparação da sensibilidade do gráfico de controle de regressão proposto. Nestas duas comparações, o gráfico de controle de regressão proposto apresentou um bom desempenho para a detecção de mudanças isoladas no coeficiente de intercepto β_0 e em mudanças simultâneas nos coeficientes de regressão β_0 e β_1 e, dessa forma, o método proposto foi validado.

5 CONSIDERAÇÕES FINAIS

Para finalizar este trabalho, apresentam-se as conclusões dos principais resultados encontrados e sugestões para trabalhos futuros.

5.1 CONCLUSÕES

Para o uso dos gráficos de controle tradicionais é necessário supor que as variáveis monitoradas sejam independentes e identicamente distribuídas. Quando ocorrem freqüentes ajustes nas variáveis de controle, a distribuição da característica de qualidade do processo pode ser alterada significativamente, fazendo com que os gráficos de controle clássicos sejam ineficientes. Nesta situação, se existir uma relação linear entre a característica de qualidade e as variáveis de controle, é possível utilizar o gráfico de controle de regressão.

O presente trabalho teve como objetivo principal propor um método para a aplicação dos gráficos de controle de regressão em um sistema produtivo, que foi elaborado a partir de outros métodos encontrados na literatura. Dessa forma, foi proposto um método dividido em duas fases inter-relacionadas: Fase I – análise retrospectiva – e Fase II – monitoramento do processo propriamente dito.

A Fase I do método proposto inclui as etapas de coletas de dados, estimação do modelo de regressão linear que relacione a característica de qualidade às variáveis de controle do processo e a verificação da estabilidade do processo no período em que a amostra foi coletada. Se o processo for considerado como estando sob controle estatístico e, assumindo que o modelo de regressão estimado na Fase I do processo esteja correto, prossegue-se à Fase II do método proposto, que consiste na coleta de amostras individuais do processo, verificação da extrapolação da região original dos dados e monitoramento da característica de qualidade.

O método proposto tem como principal justificativa facilitar a aplicação dos gráficos de controle de regressão e, seguindo este princípio, o primeiro objetivo específico desta pesquisa foi a proposta de uma modificação no gráfico de controle de regressão múltipla. Este objetivo específico foi atendido, já que foram adaptados os limites de controle deste gráfico, de forma a permitir o monitoramento direto da característica de qualidade do processo, facilitando a interpretação por parte dos operários do processo.

Durante a aplicação do método proposto, surgiu a necessidade de criar um procedimento que facilite a verificação da extrapolação do intervalo de investigação das

variáveis de controle. Neste caso, adaptou-se graficamente um procedimento encontrado na literatura para a verificação se o conjunto das variáveis de controle está extrapolando a região original de dados. Esta adaptação deu origem ao gráfico de controle de extrapolação, que pode ser considerado uma das principais contribuições deste trabalho.

O método proposto foi aplicado e validado em um processo de extrusão de bandas de rodagem de uma indústria de borrachas. Nesta aplicação explicitaram-se as etapas que compõem as fases do método proposto, de forma a ressaltar a facilidade que esta sistematização pode gerar. A seguir, comparou-se a aplicação do método proposto com outros métodos encontrados na literatura, sendo que o método proposto apresentou um desempenho satisfatório.

O quarto objetivo específico deste trabalho foi a obtenção do número médio de amostras até o sinal (NMA) para o gráfico de controle de regressão proposto, que foi calculado através do uso da simulação de Monte Carlo. Também se calculou o NMA de outros gráficos de controle baseados no gráfico de controle de regressão, como: gráfico de controle de regressão múltipla, gráfico de medidas individuais aplicado aos resíduos do modelo de regressão e gráficos EWMAREG e CUSUMREG.

Analisando os valores do NMA do gráfico de controle de regressão proposto para o processo simulado como estando sob controle, observa-se que este possui uma baixa ocorrência de alarmes falsos. Para a simulação de alterações isoladas no coeficiente de intercepto β_0 , o gráfico de controle de regressão proposto apresentou um bom desempenho, sobretudo para mudanças acima de $2,5\sigma_e$ em β_0 . Desta forma, o gráfico de controle proposto possui uma boa sensibilidade para a detecção de grandes mudanças na média do coeficiente de intercepto e, por consequência, uma boa sensibilidade às diferenças entre o valor observado e o valor previsto pelo modelo de regressão. Destaca-se, que para situações de alterações inferiores a $2,5\sigma_e$ em β_0 , os gráficos EWMAREG e CUSUMREG apresentaram os melhores desempenhos.

Para alterações isoladas no coeficiente de inclinação β_1 , o gráfico de controle de regressão proposto não apresentou um bom desempenho, embora melhor que os demais gráficos de controle. Lembrando que, como a variável de controle x_1 , possui média igual a zero, alterações em β_1 somente alteram o desvio-padrão dos resíduos, sem alterar a média destes. Assim, o gráfico de controle proposto não apresenta um bom desempenho para detectar mudanças no desvio-padrão do processo. Este resultado pode ser explicado pelo fato do gráfico de controle de regressão ter sido desenvolvido para detectar apenas mudanças na média.

No caso de mudanças simultâneas em um coeficiente de inclinação e no coeficiente de intercepto, o NMA do gráfico de controle de regressão possui um bom desempenho, sempre melhor que os demais gráficos estudados. Ressalta-se que estas situações simuladas acarretam na mudança da média e do desvio-padrão dos resíduos.

5.2 SUGESTÕES PARA TRABALHOS FUTUROS

Através deste trabalho, é possível identificar as sugestões para trabalhos futuros:

- a) Aplicação do método proposto neste trabalho em outros processos produtivos;
- b) Comparação entre as técnicas de monitoramento de perfis lineares e o gráfico de controle de regressão;
- c) Desenvolvimento de um gráfico de controle alternativo para monitorar a variabilidade de um processo monitorado pelo gráfico de controle de regressão e obtenção do NMA deste gráfico para mudanças na variabilidade;
- d) Utilização de gráficos de controle baseados em modelos para frações de não-conformes, utilizando-se, por exemplo, modelos lineares generalizados ou modelos de regressão Beta;
- e) Proposta de um método orientativo, baseado no método proposto neste trabalho, para a utilização do gráfico de controle de regressão não linear em um sistema produtivo;
- f) Proposta de um método para a aplicação do gráfico EWMAREG, incluindo uma alteração similar à proposta no presente trabalho;
- g) Planejamento econômico do gráfico de controle de regressão;
- h) Uso do método dos mínimos quadrados parciais para o monitoramento de processos multivariados, em abordagem semelhante ao gráfico de controle de regressão;
- i) Obtenção de índices de capacidade adaptados ao gráfico de controle de regressão.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALMEIDA, S. S. **Desenvolvimento de Gráficos de Controle Aplicados ao Modelo Funcional de Regressão**. 2003. 112 f. Tese (Doutorado em Engenharia de Produção) - Escola de Engenharia, Universidade Federal de Santa Catarina, Florianópolis, 2003.
- ALWAN, L. C.; ROBERTS, H. V. Time-Series Modeling for Statistical Process Control. **Journal of Business & Economic Statistics**, Washington, v. 6, n. 1, p. 87-95, 1988.
- BELSLEY, D. A.; KUH, E.; WELSCH, R. E. **Regression Diagnostics: identifying influential data and sources of collinearity**. New York: John Wiley & Sons, 1980. 292 p.
- BREMAUD, P. **Markov chains: Gibbs fields, Monte Carlo simulation, and queues**. New York: Springer, 1999, 444 p.
- CAI, D. Q.; XIE, M.; GOH, T. N.; TANG, X. Y. Economic Design of Control Chart for Trended Processes. **International Journal of Production Economics**, Amsterdam, v. 79, n. 1, p. 85-92, 2002.
- CARVALHO Jr., J. G. **Gráfico de Controle de Regressão Estrutural**. 2006. 86 f. Dissertação (Mestrado em Matemática e Estatística) - Centro de Ciências Exatas e Naturais, Universidade Federal do Pará, Belém, 2006.
- CARVALHO Jr., J. G.; ALMEIDA, S. S.; RAMOS, E. M. L. S. Gráfico de Controle de Regressão Estrutural. **Tendências de Matemática Aplicada e Computacional**, São José do Rio Preto, v. 8, n. 3, p. 1-10, 2007.
- CASARIN, V. A.; SOUZA, A. M.; BOHM, S. I. H.; JACOBI, L. F. Aplicação de Gráficos de Controle de Regressão no Monitoramento do Processo de Montagem de Colheitadeiras. In: ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO, 27, 2007, Foz do Iguaçu. **Anais...** Foz do Iguaçu: Associação Brasileira de Engenharia de Produção, 8 p., 2007.
- CASELLA, G.; ROBERT, C.P. **Monte Carlo Statistical Methods**. New York: Springer, 2. edition, 2004. 645 p.
- COOK, D. R. Detection of Influential Observation in Linear Regression. **Technometrics**, Washington, v. 19, n. 1, p. 15-18, 1977.
- COSTA, A. F. B.; EPPRECHT, E. K.; CARPINETTI, L. C. R. **Controle Estatístico de Qualidade**. São Paulo: Atlas, 2. ed., 2005. 334 p.
- CROWDER, S. V. A Simple Method for Studying Run-Length Distributions of Exponentially Weighted Moving Average Charts. **Technometrics**, Washington, v. 29, n. 4, p. 401-407, 1987.
- CROWDER, S. V. Design of Exponentially Weighted Moving Average Schemes. **Journal of Quality Technology**, Milwaukee, v. 21, n. 3, p. 155-162, 1989.
- DIPAOLA, P. P. Use of Correlation in Quality Control. **Industrial Quality Control**, Milwaukee, v. 2, n. 1, p. 10-14, 1945.

- EISENHART, K. M. Building Theories from Case Study Research. **Academy of Management Review**, Mississippi, v. 14, n. 4, p. 532-550, 1989.
- FALTIN, F. W.; MASTRANGELO, C. M.; RUNGER, G. C.; RYAN, T. P. Considerations in the Monitoring of the Autocorrelated and Independent Data. **Journal of Quality Technology**, Milwaukee, v. 29, n. 2, p. 131-133, 1997.
- GAN, J.J. An Optimal Design of CUSUM Quality Control Charts. **Journal of Quality Technology**, Milwaukee, v. 23, n. 4, p. 279-286, 1991.
- GIL, A. C. **Como Elaborar Projetos de Pesquisa**. 3. ed. São Paulo: Atlas, 1991, 159 p.
- GUJARATI, D. N. **Econometria Básica**. 3. ed. São Paulo: Makron Books, 2000. 846 p.
- HAWKINS, D. M. Multivariate Quality Control Based on Regression-Adjusted Variables. **Technometrics**, Washington, v. 33, n. 1, p. 61-75, 1991.
- HAWKINS, D. M. Regression Adjustment for Variables in Multivariate Quality Control. **Journal of Quality Technology**, Milwaukee, v. 25, n.3, p. 170-182, 1993.
- HAWORTH, D. A. Regression Control Charts to Manage Software Maintenance. **Journal of Software Maintenance**, Sussex, v. 8, n. 1, p 35-48, 1996.
- JACOBI, L. F. J. **Monitoração de Processos de Coleta de Resíduos em Santa Maria – Rs**, Usando Gráficos de Controle de Regressão. 2001. 85 f. Dissertação (Mestrado em Engenharia de Produção) – Centro de Tecnologia, Universidade Federal de Santa Maria, Santa Maria, 2001.
- JACOBI, L. F.; SOUZA, A. M.; PEREIRA, J. E. S. Gráfico de Controle de Regressão Aplicado na Monitoração de Processos. **Revista Produção**, São Paulo, v. 12, n. 1, p. 46-59, 2002.
- JEARKPAPORN, D.; MONTGOMERY, D. C.; RUNGER, G. C.; BORROR, C. M. Process Monitoring for Correlated Gamma-Distributed Data Using Generalized-Linear-Model-Based Control Charts. **Quality and Reliability Engineering International**, Chichester, v. 19, n. 6, p. 477-491.
- JEARKPAPORN, D.; MONTGOMERY, D. C.; RUNGER, G. C.; BORROR, C. M. Model-based Process Monitoring Using Robust Generalized Linear Models. **International Journal of Production research**, London, v. 43, n. 7, p. 1337-1354, 2005.
- JEARKPAPORN, D.; MONTGOMERY, D. C.; RUNGER, G. C.; BORROR, C. M. Process Monitoring for Mean Shifts for Multiple Stage Processes. **International Journal of Production Research**, London, v. 45, n. 23, p. 5547-5570, 2007.
- JONES, L. A. The Statistical Design of EWMA Control Charts with Estimated Parameters. **Journal of Quality Technology**, Milwaukee, v. 34, n. 3, p. 277-288, 2002.

JOHNSON, B.; JOHNSON, C.; SEIBER, J. The Use of Regression Equations for Quality Control in a Pesticide Physical Property Database. **Environmental Management**, New York, v. 19, n. 1, p. 127-134, 1995.

KANG, L.; ALBIN, S. On-line Monitoring When the Process Yields Linear Profiles. **Journal of Quality Technology**, Milwaukee, v. 32, n. 4, p. 418-426, 2000.

KIM, K.; MAHMOUD, M. A.; WOODALL, W. H. On the Monitoring of Linear profiles. **Journal of Quality Technology**, Milwaukee, v. 35, n. 3, p. 317-328, 2003.

KOKSALAN, M.; ERKIP, N.; MOSKOWITZ, H. Explaining Beer Demand: A Residual Modeling Regression Approach Using Statistical Process Control. **International Journal of Production Economics**, Amsterdam, v. 58, p. 265-276, 1999.

LYLE, P. **Regression Analysis of Production Costs and Factory Operations**, 1. edition. New York: Hafner Publishing Company, 1944. 204 p.

LOPES, L. F. D.; SEIDEL, E. J.; MOREIRA Jr., F. J. M. Análise de Regressão Associada ao Controle Estatístico de Processos Aplicado na Produção de Cerâmicas. In: SIMPÓSIO DE ENGENHARIA DE PRODUÇÃO, 13, 2006, Bauru. **Anais...** Bauru: UNESP, 7 p., 2006.

LOREDO, E. N.; JEARKPAPORN, D.; BORROR, C. M. Model-based Control Chart for Autoregressive and Correlated Data. **Quality and Reliability Engineering International**, Chichester, v. 18, n. 6, p. 489-496, 2002.

LUCAS, J. M.; CROSIER, R. B. Fast Initial Response for CUSUM Quality Control Schemes: Give Your CUSUM a Head Start. **Technometrics**, Washington, v. 24, n. 3, p. 199-205, 1982.

LUCAS, J. M.; SACCUCCI, M. S. Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements. **Technometrics**, Washington, v. 32, n.1, p. 1-12, 1990.

MALLOWS, C. L. Some Coments on C_p . **Technometrics**, Washington, v. 15, n. 4, p. 661-675, 1973.

MALLOWS, C. L. More Coments on C_p . **Technometrics**, Washington, v. 37, n. 4, p. 362-371, 1995.

MANDEL, B. J. The Regression Control Chart. **Journal of Quality Technology**, Milwaukee, v. 1, n. 1, p. 1-9, 1969.

MANSFIELD, E.; WEIN, H. H. A Regression Control Chart for Costs. **Journal of Royal Statistical Society Series C – Applied Statistics**, London, v. 7, n. 1, p. 48-57, 1958.

MEREDITH, J. Building Operations Management Theory Through Case and Field Research. **Journal of Operational Management**, v. 16, n. 4, p. 441-454, 1998.

MONTGOMERY, D. C. **Introdução ao Controle Estatístico da Qualidade**. 4. ed. Rio de Janeiro: Editora LTC, 2004. 513 p.

- MONTGOMERY, D. C. **Design and Analysis of Experiments**. 6. edition. New York: John Wiley & Sons, 2005, 643 p.
- MONTGOMERY, D. C.; MASTRANGELO, C. M. Some Statistical Process Control Methods for Autocorrelated Data. **Journal of Quality Technology**, Milwaukee, v. 23, n. 3, p. 179-193, 1991.
- MONTGOMERY, D. C.; RUNGER, G C. **Estatística Aplicada e Probabilidade para Engenheiros**. 2. ed., Rio de Janeiro: Editora LTC, 2003, 463 p.
- MONTGOMERY, D. C.; VINING, G. G.; PECK, E. A. **Introduction to Linear Regression Analysis**. 3. ed., New York: John Wiley & Sons, 2001, 641 p.
- MOREIRA Jr., F. J. **Proposta de um Método para o Controle Estatístico de Processo para Observações Autocorrelacionadas**. 2005. 137 f. Dissertação (Mestrado em Engenharia de Produção) - Escola de Engenharia, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2005.
- NELSON, L. S. The Shewhart Control Chart-Tests for Special Causes. **Journal of Quality Technology**, Milwaukee, v. 16, n. 4, p. 237-239, 1984.
- NETER, J.; KUTNER, M H.; NACHTSHEIM, C. J.; LI, W. **Applied Linear Statistical Models**. 5. ed. New York: Mc Graw-Hill/Irwin, 2005, 1396 p.
- OLIN, B. D. Regression Control Charts Revisited: Methodology and Cases Studies. In: ANNUAL FALL TECHNICAL CONFERENCE, 42, 1998, New York. **Proceedings** ..., New York: American Society for Quality, 17p., 1998.
- OLIVEIRA, S. L. **Tratado de Metodologia Científica**. 2. ed. São Paulo: Pioneira, 2000, 320 p.
- OMURA, A. P.; STEFFE, J. H. Mixer Viscometry to Characterize Fluid Foods with Large Particulates. **Journal of Food Process Engineering**, Westport, v. 26, n. 3, p. 435-445, 2003.
- PEDRINI, D. C.; CATEN, C. S. ten. Comparação entre Gráficos de Controle para Resíduos de Modelos. In: SIMPÓSIO DE ENGENHARIA DE PRODUÇÃO, 16, 2008, Bauru. **Anais...** Bauru: UNESP, 7 p., 2008.
- PRESS, W. H.; TEUKOLSKY, S. A.; VETTERLING, W. T.; FLANNERY, B. P. **Numerical Recipes: The Art of Scientific Computing**. 3. ed. Cambridge: Cambridge University Press, 1997, 1265 p.
- RYAN, T. P. **Statistical Methods for Quality Improvement**. New York: John Wiley & Sons, 1989, 446p.
- ROTHSCHILD, B. F.; ROTH, S. R. Statistical Process Control of Plating Solutions with Regression Control Charts. **The SAMPE Journal**, Covina, v. 22, n. 5, p. 37-41, 1986.

- SHU, L.; TSUNG, F.; TSUI, K. L. Run-length Performance of Regression Control Charts with Estimated Parameters. **Journal of Quality Technology**, Milwaukee, v. 36, n. 3, p. 280-292, 2004.
- SHU, L.; TSUI, K. L.; TSUNG, F. A Review of Regression Control Charts. In: Ruggeri, F.; Faltin, F.; Kenett, R. **Encyclopedia of Statistics in Quality and Reliability** New York: John Wiley & Sons, p. 1569-1573, 2007.
- SILVA, E. L.; MENEZES, E. M. **Metodologia da Pesquisa e Elaboração de Dissertação**. 4. ed. Florianópolis: UFSC, 2005. 138 p. Disponível em: <www.posarq.ufsc.br/download/metPesq.pdf>. Acesso em: 04 set. 2007.
- SKINNER, K. R.; MONTGOMERY, D. C.; RUNGER, G. C. Process Monitoring for Multiple Count Data Using Generalized Linear Model-Based Control Charts. **International Journal of Production Research**, London, v. 41, n. 6, p. 1167-1180, 2003.
- SKINNER, K. R.; MONTGOMERY, D. C.; RUNGER, G. C. Generalized Linear Model-based Control Charts for Discrete Semiconductor Process Data. **Quality and Reliability Engineering International**, Chichester, v. 20, n. 8, p. 777-786, 2004.
- SPANOS, C. J. Statistical Process Control in Semiconductor Manufacturing. **Proceedings of the IEEE**, New York, v. 80, n. 6, p. 819-830, 1992.
- TEIXEIRA, J. M. G. **Análise de Regressão Aplicada ao Controle Estatístico de Processos**. 1997. 89 f. Dissertação (Mestrado em Métodos Quantitativos) – Centro de Ciências Exatas e Naturais, Universidade Federal de Santa Maria, Santa Maria, 1997.
- VANCE, L. C. Average Run Lengths of Cumulative Control Charts for Controlling Normal Means. **Journal of Quality Technology**, Milwaukee, v. 18, n. 3, p. 189-193, 1986.
- WADE, M. R.; WOODALL, W. H. A Review and Analysis of Cause-Selecting Control Charts. **Journal of Quality Technology**, Milwaukee, v. 25, n. 3, p. 161-169, 1993.
- WALLIS, W. A.; ROBERTS, H. V. **Statistics: A New Approach**. Chicago: The Free Press, v. 2, p. 549-553, 1956.
- WEIS, P. E. An Application of Two-way X-Bar Chart. **Industrial Quality Control**, Milwaukee, v. 14, n. 6, p. 23-27, 1957.
- WEISBERG, S. **Applied Linear Regression**. 3. edition. New York: John Wiley & Sons, 2005, 310 p.
- WOODALL, W. H.; MONTGOMERY, D.C. Research Issues and Ideas in Statistical Process Control. **Journal of Quality Technology**, Milwaukee, v. 31, n. 4, p. 376-386, 1999.
- YANG, S. F. Optimal Process Control for Multiple Dependent Subprocesses. **Quality and Reliability Engineering International**, Chichester, v. 14, n. 4, p. 347-355, 1998.
- ZOU, C.; ZHOU, C.; WANG, Z.; TSUNG, F. A Self-Starting Control Chart for Linear Profiles. **Journal of Quality Technology**, Milwaukee, v. 39, n. 4, p. 364-375, 2007.