

**UNIVERSIDADE FEDERAL DE SANTA MARIA
CENTRO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE
PRODUÇÃO**

**MODELOS DE PREVISÃO APLICADOS AO
CONTROLE DE QUALIDADE COM DADOS
AUTOCORRELACIONADOS**

DISSERTAÇÃO DE MESTRADO

Regiane Klidzio

**Santa Maria, RS, Brasil
2009**

MODELOS DE PREVISÃO APLICADOS AO CONTROLE DE QUALIDADE COM DADOS AUTOCORRELACIONADOS

por

Regiane Klidzio

Dissertação apresentada ao Curso de Mestrado do Programa de
Pós-Graduação em Engenharia de Produção, Área de Concentração em
Gerência de Produção, da Universidade Federal de Santa Maria (UFSM, RS),
como requisito parcial para obtenção do grau de
Mestre em Engenharia de Produção

Orientador: Prof. Dr. Adriano Mendonça Souza

Santa Maria, RS, Brasil

2009

**Universidade Federal de Santa Maria
Centro de Tecnologia
Pós-Graduação em Engenharia de Produção**

A Comissão Examinadora, abaixo assinada,
aprova a Dissertação de Mestrado

**MODELOS DE PREVISÃO APLICADOS AO CONTROLE DE
QUALIDADE COM DADOS AUTOCORRELACIONADOS**

elaborada por
Regiane Klidzio

como requisito parcial para obtenção do grau de
Mestre em Engenharia de Produção

COMISSÃO EXAMINADORA:

Adriano Mendonça Souza, Dr. (UFSM)
(Presidente/Orientador)

Leandro Cantorski da Rosa, Dr. (UFSM)

Wesley Vieira da Silva, Dr. (PUCPR)

Santa Maria, 4 de setembro de 2009

AGRADECIMENTOS

Agradeço a Deus, que me proporcionou a graça de viver, aprender e crescer, por ter me dado força para superar as longas jornadas de estudo e trabalho e por ter guiado meus passos.

A minha família, por depositarem confiança para a realização desta dissertação, em especial, aos meus pais, com quem aprendi a ter orgulho, pela confiança, paciência, apoio e oportunidade concebida e aos meus irmãos, pelo companheirismo e motivação.

Ao Prof. Dr. Adriano, pela oportunidade dada e pelos conhecimentos transmitidos em busca de novos caminhos, agradeço a grande e positiva influência que exerceu em minha formação acadêmica e pela valiosa contribuição neste trabalho.

Aos membros da banca examinadora, que contribuíram significativamente para a melhoria desta pesquisa.

Às empresas e a seus colaboradores da área industrial, que atenciosamente disponibilizaram a coleta de dados e informações sobre os processos industriais para a realização desta pesquisa.

Ao meu namorado, pelo suporte, incentivo, motivação e compreensão dos momentos dedicados aos meus estudos.

Aos amigos que fiz durante a realização do mestrado, pelas horas de estudo compartilhadas e pelo apoio mútuo, em especial, a minha amiga Cristina, pela amizade e colaboração na interação dos conhecimentos e pela parceria nas disciplinas que cursamos no decorrer deste curso.

À Universidade Regional Integrada do Alto Uruguai e das Missões – URI pela oportunidade concedida.

Enfim, a todos que de uma forma direta ou indireta contribuíram para a realização deste trabalho.

RESUMO

Dissertação de Mestrado
Programa de Pós-Graduação em Engenharia de Produção
Universidade Federal de Santa Maria

MODELOS DE PREVISÃO APLICADOS AO CONTROLE DE QUALIDADE COM DADOS AUTOCORRELACIONADOS

Autora: Regiane Klidzio
Orientador: Adriano Mendonça Souza, Dr.
Data e Local da Defesa: Santa Maria, 4 de Setembro de 2009.

A presente pesquisa tem como tema a abordagem de modelos de previsão, aplicados a processos produtivos industriais, com o objetivo de verificar a estabilidade do processo por meio de gráficos de controle, aplicado aos resíduos oriundos de modelagem linear e não-linear. Como as observações eram autocorrelacionadas, foi necessário buscar um modelo matemático pelo qual foram obtidos resíduos independentes e normalmente distribuídos. A investigação da estabilidade do processo passa pela verificação da influência da volatilidade na detecção de pontos amostrais que são potenciais para afetar o desempenho do processo produtivo. Esse fato comprova a existência da volatilidade em processos produtivos que, até o momento, é trabalhada apenas em variáveis econômicas. Os dados utilizados para análise pertencem a três empresas de segmentos distintos. O modelo matemático foi ajustado utilizando modelo de regressão dinâmica multivariada, modelo ARIMA e modelo ARIMA-ARCH. De acordo com os gráficos de controle, as técnicas estatísticas empregadas para eliminar a autocorrelação serial dos dados mostraram-se adequadas estatisticamente, se comparadas com o modelo clássico utilizado por cada empresa analisada. Além disso, verificou-se que, no período que ocorre volatilidade corresponde a um período fora de controle detectado nos gráficos de controle de Shewhart. Os modelos matemáticos encontrados foram capazes de representar os processos produtivos, possibilitando compreender o comportamento das variáveis e auxiliaram na realização das previsões e na monitoração do processo.

Palavras-chave: Séries temporais; Modelos lineares e não-lineares; Autocorrelação; Previsão; Gráficos de controle.

ABSTRACT

Master's Degree Thesis
Post-Graduate Program in Production Engineering
Federal University of Santa Maria

FORECAST MODEL APPLIED TO QUALITY CONTROL WITH AUTOCORRELATIONAL DATA

Author: Regiane Klidzio
Supervisor: Adriano Mendonça Souza, Dr.
Submission: Santa Maria, September 4th, 2009.

This research has a topic forecast models applied to industrial productive processes with the objective of verifying the stability of the process through control charts applied to the residues originated from linear and non-linear model. In the presence of autocorrelation data, it was necessary to look for a mathematical model which are produce independent and identically distributed residues. This investigation about the stability of the process goes by the verification of the volatility is influence in the detection of points that are capable to affect the productive process performance. This fact shows the existence of the volatility in productive processes, which it is just used until now in economic variables. The data used for analysis belong to three different industries in different segments. The mathematic models were used multivariate dynamic equation, ARIMA and ARIMA-ARCH model. According to the control charts the statistical techniques used to eliminate the serial autocorrelation was statistically adequate comparing to the classic model used by each industry analyzed. Besides, it was verified, in the period that the volatility occurs corresponds to the period the shows a lack of stability detected by Shewhart control charts. The mathematic models were able to represent the productive process, facilitating understands the behavior of the variables, and help to accomplish the forecast and monitoring the process.

Words key: Time series; Linear e non-linear models; Autocorrelation; Forecast; Control charts.

LISTA DE FIGURAS

Figura 1 – Gráfico de controle adaptado de Montgomery, 2004.....	26
Figura 2 – Representação gráfica para os testes de sequência	27
Figura 3 – Comportamento da FAC e FACP de um modelo AR(1)	33
Figura 4 – Comportamento da FAC e FACP de um modelo MA(1).....	33
Figura 5 – Roteiro metodológico	62
Figura 6 – Função de autocorrelação para as variáveis teor de umidade, permeabilidade e rcv	70
Figura 7 – Valores originais e estimados para as variáveis teor de umidade, permeabilidade e rcv	71
Figura 8 – Previsão dinâmica para o teor de umidade, permeabilidade e rcv.....	73
Figura 9 – Gráfico \bar{x} do teor de umidade, considerando os dados originais	75
Figura 10 – Gráfico \overline{MR} do teor de umidade, considerando os dados originais	76
Figura 11 – Gráfico \bar{x} da permeabilidade, considerando os dados originais.....	76
Figura 12 – Gráfico \overline{MR} da permeabilidade, considerando os dados originais	77
Figura 13 – Gráfico \bar{x} da rcv, considerando os dados originais.....	78
Figura 14 – Gráfico \overline{MR} da rcv, considerando os dados originais	78
Figura 15 – Gráfico \bar{x} do teor de umidade com a exclusão dos pontos amostrais fora de controle, considerando os dados originais.....	79
Figura 16 – Gráfico \overline{MR} do teor de umidade com a exclusão dos pontos amostrais fora de controle, considerando os dados originais.....	79
Figura 17 – Gráfico \bar{x} da permeabilidade com a exclusão dos pontos amostrais fora de controle, considerando os dados originais.....	80
Figura 18 – Gráfico \overline{MR} da permeabilidade com a exclusão dos pontos amostrais fora de controle, considerando os dados originais.....	80
Figura 19 – Gráfico \bar{x} da rcv com a exclusão dos pontos amostrais fora de controle, considerando os dados originais	80
Figura 20 – Gráfico \overline{MR} da rcv com a exclusão dos pontos amostrais fora de controle, considerando os dados originais	81
Figura 21 – Gráfico \bar{x} aplicado ao resíduo do teor de umidade.....	81
Figura 22 – Gráfico \overline{MR} aplicado ao resíduo do teor de umidade	82
Figura 23 – Gráfico \bar{x} aplicado ao resíduo da permeabilidade	83
Figura 24 – Gráfico \overline{MR} aplicado ao resíduo da permeabilidade.....	83
Figura 25 – Gráfico \bar{x} aplicado ao resíduo da rcv.....	84
Figura 26 – Gráfico \overline{MR} aplicado ao resíduo da rcv.....	84

Figura 27 – Série original da variável índice de acidez	90
Figura 28 – Correlograma, estatística de Ljung-Box - Q(K) e <i>p-valor</i> da série original para a variável índice de acidez	90
Figura 29 – Distribuição e estatísticas descritivas dos resíduos do modelo ARMA(2,1) para a variável índice de acidez	93
Figura 30 – Gráfico Q-Q plot dos resíduos do modelo ARMA(2,1)	94
Figura 31 – Série residual do modelo ARMA(2,1)	94
Figura 32 – Correlograma, estatística de Ljung-Box (Q(K)) e <i>p-valor</i> da série dos resíduos quadráticos do modelo ARMA(2,1) para a variável índice de acidez	95
Figura 33 – Gráfico Q-Q plot dos resíduos padronizados estimados pelo modelo ARMA(2,1)-ARCH(1) para a variável índice de acidez.....	97
Figura 34 – Valores originais, valores estimados e os resíduos do modelo ARMA(2,1)-ARCH(1)	98
Figura 35 – Volatilidade condicional do modelo misto ARMA(2,1)-ARCH(1).....	99
Figura 36 – Gráfico \bar{x} do índice de acidez, considerando os dados originais.....	100
Figura 37 – Gráfico \overline{MR} do índice de acidez, considerando os dados originais	100
Figura 38 – Gráfico \bar{x} do índice de acidez com a exclusão dos pontos amostrais fora de controle, considerando os dados originais.....	101
Figura 39 – Gráfico \overline{MR} do índice de acidez com a exclusão dos pontos amostrais fora de controle, considerando os dados originais.....	101
Figura 40 – Gráfico \bar{x} aplicado ao resíduo do modelo ARMA(2,1) para a variável índice de acidez	102
Figura 41 – Gráfico \overline{MR} aplicado ao resíduo do modelo ARMA(2,1) para a variável índice de acidez	103
Figura 42 – Gráfico \bar{x} aplicado ao resíduo do modelo ARMA(2,1) para a variável índice de acidez com a exclusão dos pontos amostrais fora de controle	104
Figura 43 – Gráfico \overline{MR} aplicado ao resíduo do modelo ARMA(2,1) para a variável índice de acidez com a exclusão dos pontos amostrais fora de controle	104
Figura 44 – Gráfico \bar{x} aplicado ao resíduo quadrático do modelo misto ARMA(2,1)-ARCH(1) para a variável índice de acidez.....	105
Figura 45 – Gráfico \overline{MR} aplicado ao resíduo quadrático do modelo misto ARMA(2,1)-ARCH(1) para a variável índice de acidez.....	105
Figura 46 – Gráfico \bar{x} aplicado ao resíduo quadrático do modelo misto ARMA(2,1)-ARCH(1) para a variável índice de acidez com a exclusão dos pontos amostrais fora de controle.....	106
Figura 47 – Gráfico \overline{MR} aplicado ao resíduo quadrático do modelo misto ARMA(2,1)-ARCH(1) para a variável índice de acidez com a exclusão dos pontos amostrais fora de controle.....	107
Figura 48 – Gráfico \bar{x} aplicado a volatilidade do modelo misto ARMA(2,1)-ARCH(1) para a variável índice de acidez	107
Figura 49 – Gráfico \overline{MR} aplicado a volatilidade do modelo misto ARMA(2,1)-ARCH(1) para a variável índice de acidez.....	108
Figura 50 – Gráfico \bar{x} aplicado a volatilidade do modelo misto ARMA(2,1)-ARCH(1) para a variável índice de acidez com a exclusão dos pontos amostrais fora de controle	109
Figura 51 – Gráfico \overline{MR} aplicado a volatilidade do modelo misto ARMA(2,1)-ARCH(1) para a variável índice de acidez com a exclusão dos pontos amostrais fora de controle	109
Figura 52 – Série original para as variáveis TATQ1 e TATQ2	113

Figura 53 – Correlograma, estatística de Ljung-Box - Q(K) e <i>p-valor</i> da série original para a variável TATQ1	114
Figura 54 – Correlograma, estatística de Ljung-Box - Q(K) e <i>p-valor</i> da série original para a variável TATQ2.....	114
Figura 55 – Distribuição e estatísticas descritivas dos resíduos do modelo ARMA(1,1) para a variável TATQ1	117
Figura 56 – Gráfico Q-Q plot dos resíduos do modelo ARMA(1,1).....	118
Figura 57 – Correlograma, estatística de Ljung-Box - Q(K) e <i>p-valor</i> da série dos resíduos quadráticos do modelo ARMA(1,1) para a variável TATQ1.....	119
Figura 58 – Distribuição e estatísticas descritivas dos resíduos do modelo AR(3) para a variável TATQ2.....	120
Figura 59 – Gráfico Q-Q plot dos resíduos do modelo ARMA(1,1).....	121
Figura 60 – Série residual do modelo AR(3)	121
Figura 61 – Correlograma, estatística de Ljung-Box - Q(K) e <i>p-valor</i> da série dos resíduos quadráticos do modelo AR(3) para a variável TATQ2.....	122
Figura 62 – Q-Q plot dos resíduos padronizados estimados pelo modelo AR(3)-ARCH(1) para a variável TATQ2	124
Figura 63 – Volatilidade condicional do modelo misto AR(3)-ARCH(1).....	124
Figura 64 – Valores originais, valores estimados e os resíduos do modelo ARMA(1,1).....	125
Figura 65 – Valores originais, valores estimados e os resíduos do modelo AR(3)-ARCH(1)	126
Figura 66 – Gráfico \bar{x} da TATQ1, considerando os dados originais	128
Figura 67 – Gráfico \overline{MR} da TATQ1, considerando os dados originais	128
Figura 68 – Gráfico \bar{x} da TATQ1 com a exclusão dos pontos amostrais fora de controle, considerando os dados originais	129
Figura 69 – Gráfico \overline{MR} da TATQ1 com a exclusão dos pontos amostrais fora de controle, considerando os dados originais	129
Figura 70 – Gráfico \bar{x} aplicado ao resíduo do modelo ARMA(1,1) para a variável TATQ1	130
Figura 71 – Gráfico \overline{MR} aplicado ao resíduo do modelo ARMA(1,1) para a variável TATQ1	130
Figura 72 – Gráfico \bar{x} aplicado ao resíduo do modelo ARMA(1,1) para a variável TATQ1 com a exclusão dos pontos amostrais fora de controle	131
Figura 73 – Gráfico \overline{MR} aplicado ao resíduo do modelo ARMA(1,1) para a variável TATQ1 com a exclusão dos pontos amostrais fora de controle	132
Figura 74 – Gráfico \bar{x} da TATQ2, considerando os dados originais	132
Figura 75 – Gráfico \overline{MR} da TATQ2, considerando os dados originais	133
Figura 76 – Gráfico \bar{x} da TATQ2 com a exclusão dos pontos amostrais fora de controle, considerando os dados originais	133
Figura 77 – Gráfico \overline{MR} da TATQ2 com a exclusão dos pontos amostrais fora de controle, considerando os dados originais	133
Figura 78 – Gráfico \bar{x} aplicado ao resíduo quadrático do modelo misto AR(3)-ARCH(1) para a variável TATQ2	134
Figura 79 – Gráfico \overline{MR} aplicado ao resíduo quadrático do modelo misto AR(3)-ARCH(1) para a variável TATQ2.....	135
Figura 80 – Gráfico \bar{x} aplicado ao resíduo quadrático do modelo misto AR(3)-ARCH(1) para a variável TATQ2 com a exclusão dos pontos amostrais fora de controle	136

Figura 81 – Gráfico \overline{MR} aplicado ao resíduo quadrático do modelo misto AR(3)-ARCH(1) para a variável TATQ2 com a exclusão dos pontos amostrais fora de controle	136
Figura 82 – Gráfico \bar{x} aplicado a volatilidade do modelo misto AR(3)-ARCH(1) para a variável TATQ2.....	137
Figura 83 – Gráfico \overline{MR} aplicado a volatilidade do modelo misto AR(3)-ARCH(1) para a variável TATQ2.....	138
Figura 84 – Gráfico \bar{x} aplicado a volatilidade do modelo misto AR(3)-ARCH(1) para a variável TATQ2 com a exclusão dos pontos amostrais fora de controle.....	138
Figura 85 – Gráfico \overline{MR} aplicado a volatilidade do modelo misto AR(3)-ARCH(1) para a variável TATQ2 com a exclusão dos pontos amostrais fora de controle.....	138

LISTA DE TABELAS

Tabela 1 – Estatísticas descritivas para as variáveis teor de umidade, permeabilidade e rcv da Empresa A no período de 29/04/2008 a 30/06/2009	68
Tabela 2 – Estimação dos coeficientes, erro padrão, t-calc e <i>p-valor</i> do modelo de regressão dinâmica multivariada para as variáveis teor de umidade, permeabilidade e rcv	69
Tabela 3 – Diagnóstico dos resíduos multivariados	72
Tabela 4 – Diagnóstico dos resíduos individuais	72
Tabela 5 – Período, valores originais, valores previstos, erro de previsão, MAPE e U de Theil para as variáveis teor de umidade, permeabilidade e rcv da Empresa A	73
Tabela 6 – Estatísticas descritivas da variável índice de acidez da Empresa B no período de 1º/04/2004 a 30/06/2009	89
Tabela 7 – Testes de raiz unitária PP e KPSS para a variável índice de acidez.....	91
Tabela 8 – Estimação dos parâmetros dos modelos ARIMA concorrentes, dos critérios AIC e SBC para a variável índice de acidez.....	92
Tabela 9 – Resultados das estatísticas <i>F</i> e ARCH-LM até a defasagem 13.....	95
Tabela 10 – Estimação dos coeficientes, erro-padrão, estatística <i>z</i> e <i>p-valor</i> do modelo ARMA-ARCH para a variável índice de acidez	96
Tabela 11 – Período, valores originais, valores previstos, erro de previsão, MAPE e U de Theil para a variável índice de acidez da Empresa B	98
Tabela 12 – Estatísticas descritivas das variáveis TATQ1 e TATQ2 da Empresa C no período de 29/12/2007 a 30/06/2009	113
Tabela 13 – Testes de raiz unitária PP e KPSS para as variáveis TATQ1 e TATQ2.....	115
Tabela 14 – Estimação dos parâmetros dos modelos ARIMA concorrentes, dos critérios AIC e SBC para as variáveis TATQ1 e TATQ2.....	116
Tabela 15 – Resultados das estatísticas <i>F</i> e ARCH-LM até a defasagem 13.....	122
Tabela 16 – Estimação dos coeficientes, erro-padrão, estatística <i>z</i> e <i>p-valor</i> do modelo AR-ARCH para a variável TATQ2	123
Tabela 17 – Período, valores originais, valores previstos, erro de previsão, MAPE e U de Theil para a variável TATQ1 da Empresa C	126
Tabela 18 – Período, valores originais, valores previstos, erro de previsão, MAPE e U de Theil para a variável TATQ2 da Empresa C	127

LISTA DE QUADROS

Quadro 1 – Comportamentos da FAC e da FACP de um processo ARIMA(p,d,q).....	35
---	----

LISTA DE ABREVIATURAS E SIGLAS

- ADF – Teste de raiz unitária de Dickey e Fuller Aumentado
- AIC – Critério de Informação Akaike
- AR – Autorregressivo
- ARCH – Heterocedasticidade Condicional Autorregressiva
- ARCH-LM – Teste Multiplicador de Lagrange
- ARFIMA – Autorregressivo Fracionalmente Integrado de Médias Móveis
- ARIMA – Autorregressivo Integrado de Média Móvel
- ARIMA-ARCH – Modelo ARIMA considerando Heterocedasticidade Autorregressiva
- ARMA – Autorregressivo de Média Móvel
- BM&F – Bolsa de Mercadorias & Futuros
- BOVESPA – Bolsa de Valores de São Paulo
- CEP – Controle Estatístico de Processo
- DF – Teste de raiz unitária Dickey e Fuller
- EGARCH – GARCH Exponencial
- Estatística F – Estatística F-Snedecor
- EWMA – Média Móvel Exponencialmente Ponderada
- FAC – Função de Autocorrelação
- FACP – Função de Autocorrelação Parcial
- FIGARCH – Modelo ARCH Fracionalmente Integrado
- FIML – Máxima Verossimilhança Informação Completa
- GARCH – Heterocedasticidade Condicional Autorregressiva Generalizada
- GED – Distribuição Erro Generalizado
- iid* – Independente e Identicamente Distribuída
- IC – Intervalo de Confiança

JB – Estatística Jarque-Bera
KOH – Hidróxido de Potássio
KPSS – Teste de raiz unitária proposto pelos autores Kwiatkowski, Phillips, Schmidt e Shin
LC – Linha Central
LIC – Limite Inferior de Controle
LSC – Limite Superior de Controle
MA – Média Móvel
MAE – Erro Médio Absoluto
MAPE – Erro Médio Percentual Absoluto
MQO – Mínimos Quadrados Ordinários
MQO3E – Mínimos Quadrados Ordinários em Três Estágios
MSE – Erro Médio Quadrático
MV – Máxima Verossimilhança
PET – Politereftalato de Etileno
PIB – Produto Interno Bruto
PP – Teste raiz unitária de Phillips-Perron
Q – Estatística Q de Box e Pierce
Q(K) – Estatística Q de Ljung-Box
RB – Ruído Branco
rcv – Resistência de Compressão a Verde
SBC – Critério Bayesiano de Schwarz
SQR – Soma dos Quadrados dos Resíduos
SUR – Método de Regressões Aparentemente Não-Correlacionadas
TARCH – Heterocedasticidade Condicional Autorregressiva com Threshold
TATQ1 – Temperatura da Água para o Tanque 1
TATQ2 – Temperatura da Água para o Tanque 2
U – Estatística U de Theil
VAR – Vetor Autorregressivo

LISTA DE ANEXOS

ANEXO A – Fatores para construção de gráficos de controle para variáveis.....	154
---	-----

ÍNDICE

1 INTRODUÇÃO.....	17
1.1 Tema da pesquisa	20
1.2 Justificativa e importância do trabalho	20
1.3 Problema da pesquisa	21
1.4 Objetivos.....	21
1.4.1 Objetivo geral.....	21
1.4.2 Objetivos específicos	22
1.5 Delimitação da pesquisa	22
1.6 Organização do trabalho	22
2 REVISÃO DE LITERATURA	24
2.1 Gráficos de controle	24
2.1.1 Gráficos de controle para medidas individuais e amplitude móvel.....	29
2.2 Modelos de previsão.....	30
2.2.1 Modelos ARIMA.....	31
2.2.2 Modelos ARCH.....	38
2.2.3 Modelo de regressão dinâmica multivariada	44
2.3 Testes de diagnósticos	48
2.3.1 Testes de raiz unitária	48
2.3.2 Testes de normalidade	51
2.3.3 Testes para autocorrelação	52
2.3.4 Testes para heterocedasticidade condicional	53
2.3.5 Critérios penalizadores	56
2.3.6 Critérios de avaliação de previsão	57
2.4 Comentários gerais do capítulo	59
3 METODOLOGIA	60
3.1 Etapas metodológicas	60
3.2 Comentários gerais do capítulo	63
4 ANÁLISE DOS RESULTADOS	64
4.1 Empresa A.....	64
4.1.1 Descrição do processo e das variáveis a serem analisadas	65
4.1.2 Análise descritiva	68
4.1.3 Estimção do modelo de regressão multivariado	69
4.1.4 Diagnóstico dos resíduos	71
4.1.5 Análise da estabilidade do processo	74
4.1.6 Comentários gerais da análise	85
4.2 Empresa B.....	86

4.2.1	Descrição do processo e das variáveis a serem analisadas	87
4.2.2	Análise descritiva	89
4.2.3	Estimação de modelos ARIMA concorrentes e escolha do melhor modelo.....	91
4.2.4	Verificação da presença de heterocedasticidade condicional	92
4.2.5	Estimação do melhor modelo misto ARMA-ARCH.....	96
4.2.6	Análise da estabilidade do processo	99
4.2.7	Comentários gerais da análise	109
4.3	Empresa C.....	110
4.3.1	Descrição do processo e das variáveis a serem analisadas	111
4.3.2	Análise descritiva	112
4.3.3	Estimação de modelos ARIMA concorrentes e escolha do melhor modelo.....	115
4.3.4	Diagnóstico dos resíduos	117
4.3.5	Estimação do melhor modelo misto AR-ARCH	123
4.3.6	Previsão dos modelos matemáticos estimados.....	125
4.3.7	Análise da estabilidade do processo	127
4.3.8	Comentários gerais da análise	139
4.4	Comentários gerais do capítulo	140
5	CONCLUSÕES E RECOMENDAÇÕES	141
5.1	Conclusões	141
5.2	Recomendações.....	144
6	REFERÊNCIAS BIBLIOGRÁFICAS	146

1 INTRODUÇÃO

A qualidade do produto final está presente na linguagem da vida organizacional, profissional e pessoal. Entende-se que o que prevalece é a idéia de que a preferência do consumidor pelo produto de uma empresa é determinante para a sua sobrevivência. “Certamente, quando as necessidades ou expectativas não são atendidas em relação a um produto ou serviço, o cliente buscará outra perspectiva de satisfação dessas necessidades ou expectativas” (MALUCHE, 2000).

Para que as empresas possam sobreviver no atual cenário competitivo, é preciso gerenciar mudanças que ensinem a antecipar os concorrentes e a identificar tendências com a finalidade de auxiliar nas tomadas de decisões. Para isso, as técnicas e ferramentas estatísticas são poderosas armas à disposição da indústria.

Para que um produto corresponda às exigências do cliente, este deve, em geral, ser produzido por um processo que seja estável. O controle estatístico de processo (CEP) é uma poderosa coleção de ferramentas para resolução de problemas na obtenção da estabilidade do processo e na melhoria da capacidade através da redução da variabilidade. Entre essas ferramentas, o gráfico de controle de Shewhart, desenvolvido por Walter A. Shewhart do Bell Telephone Laboratories por volta de 1924 é, provavelmente, a ferramenta mais simples.

Segundo Montgomery (2004), os gráficos de controle são utilizados, geralmente, para alcançar um estado de controle estatístico e para monitorar o processo produtivo em qualquer tipo de processo. Há, pelo menos, cinco razões para a popularidade e importância dos gráficos de controle:

- a) são uma técnica comprovada para a melhoria da produtividade;
- b) são eficazes à prevenção de defeitos;
- c) evitam o ajuste desnecessário do processo;
- d) fornecem informação de diagnóstico;
- e) fornecem informação sobre a capacidade do processo.

As suposições geralmente necessárias para a utilização dos gráficos de controle são que os dados gerados pelo processo, quando sob controle, são distribuídos normal e independentemente com média μ e desvio padrão σ .

De acordo com esse mesmo autor, a suposição mais importante é a independência das observações, porque os gráficos de controle convencionais não são eficazes se a característica da qualidade a ser monitorada apresenta altos níveis de correlação ao longo do tempo. Nesse caso, é necessário realizar um tratamento especial nos dados, antes de proceder com a metodologia de implantação de gráficos de controle.

Em alguns processos produtivos, a suposição de observações não-correlacionadas ou independentes nem sempre é satisfeita, daí a necessidade de se buscar novas metodologias que englobem uma classe maior de situações práticas. Uma das abordagens utilizadas para tratar dados autocorrelacionados consiste em modelar diretamente a estrutura correlacional utilizando um modelo matemático apropriado de série temporal para remover a autocorrelação serial dos dados e aplicar gráficos de controle aos resíduos desse modelo. Ressalta-se que os resíduos são aproximadamente normais e independentes com média zero e variância constante, o que satisfaz plenamente as suposições do uso adequado de gráficos de controle.

Alwan e Roberts (1988) sugerem usar a metodologia de séries temporais, que ajuda a detectar a existência de variações sistemáticas e a obter resultados mais concretos das causas especiais, através do gráfico de resíduos resultante do modelo matemático.

Muitos artigos têm sido publicados na literatura com o objetivo de avaliar e tratar os efeitos da autocorrelação serial nos gráficos de controle. Para maiores detalhes, ver, por exemplo, Mingoti e Yassukawa (2008), Silva, Fontanini e Del Corso (2007), Claro, Costa e Machado (2007), Moreira Junior e Ten Caten (2004) e Ramos e Ho (2003).

Tradicionalmente, na modelagem de séries temporais, a variância dos erros é assumida constante ao longo do tempo. Nesse caso, os modelos lineares da família ARIMA são adequados para modelar a média condicional de uma série. No entanto, muitas séries temporais, principalmente financeiras, exibem períodos de grande volatilidade, seguidos de períodos de relativa tranquilidade. Nesses casos, supor variância constante, chamada de homocedasticidade, pode não ser a maneira apropriada. Embora a variância incondicional dos erros ainda possa ser assumida constante, sua variância condicional pode estar mudando ao longo do tempo. Portanto, os modelos ARIMA não são adequados a essas restrições.

Existem situações em que é possível modelar e prever a variância condicional da série, assim como existem situações onde há interesse em modelar e prever a variância condicional

da série, além da série propriamente dita. Para uma série temporal que tenha esta característica, os modelos heterocedásticos condicionais da família ARCH são os mais adequados. É importante frisar que, em séries financeiras, a variância condicional é conhecida por volatilidade, mas também é chamada de heterocedasticidade condicional (MORETTIN, 2006).

No entanto, quando há um conjunto de séries temporais com mais de três variáveis envolvidas num determinado processo produtivo, a análise univariada nem sempre é apropriada. Se existir correlação entre os erros das equações analisadas individualmente, é aconselhável realizar uma estimação conjunta a partir de um modelo de regressão multivariado como o SUR¹. Esse método possibilita que cada equação tenha sua própria forma funcional, levando em consideração apenas a correlação existente entre os resíduos dessas equações. Esse método é discutido pelos autores Hamilton (1994), Hill, Griffiths e Judge (1999), Maddala (1992), Souza, Samohyl e Malavé (2004a).

Além disso, os modelos de previsão são importantes recursos que têm a finalidade de analisar e prever comportamentos e situações que podem ocorrer numa organização, no intuito de reconfigurar a ação das competências organizacionais. O monitoramento do êxito da gestão de uma empresa está diretamente relacionado aos instrumentos utilizados. Dentre eles, destaca-se a previsão, que pode ser realizada e aprimorada com o uso de metodologias estatísticas, entre elas, os modelos lineares – autorregressivos integrados e de médias móveis (ARIMA) e regressão dinâmica multivariada e os modelos não-lineares – heterocedásticos condicionais autorregressivos (ARCH).

Na gestão empresarial, os modelos de previsão são ferramentas importantes a partir dos quais os riscos gerados pela incerteza são praticamente eliminados, oferecendo maior credibilidade no planejamento e na tomada de decisões, uma vez que a eficácia de uma decisão depende obviamente dos fatores que a sucedem. Além da previsão, o CEP também tem sua relevância no setor industrial por ser uma forma de se fazer uma monitoração contínua, possibilitando uma melhoria no controle do sistema.

Este estudo possui caráter relevante no setor industrial por ser uma forma de se fazer uma monitoração preventiva, possibilitando uma melhoria no controle de qualidade. Se a distribuição do processo é mantida sobre o tempo, ele está sob controle, mas, se a distribuição mudou, ele encontra-se fora de controle. Nesse caso, é preciso tentar detectar a mudança

¹Que em inglês denomina-se *Seegmingly Unrelated Regression (SUR)*, e, traduzindo, o termo quer dizer método de estimação de equações de regressões aparentemente não-correlacionadas, proposto por Zellner (1962).

rapidamente e determinar a sua causa física, isso justifica a aplicação dos gráficos de controle. Dessa maneira, a equipe de produção será capaz de detectar onde um futuro problema poderá ocorrer como forma de diminuir os custos associados às falhas no processo.

Estudos revelam que o desempenho dos modelos de previsão é diferente para cada série temporal, dependendo do seu comportamento, do tamanho da amostra e das características dessa série como estacionariedade, não-normalidade, autocorrelação, heterocedasticidade condicional, *clusters* de volatilidade, *outliers*, entre outras.

Contudo, pode-se afirmar a importância dos modelos de previsão, cuja escolha adequada para representar uma série temporal não é uma tarefa trivial, mas um desafio para os pesquisadores.

1.1 Tema da pesquisa

A pesquisa aborda estudos ex-post-facto com dados autocorrelacionados, sob os quais é necessário estimar um modelo matemático para remover a autocorrelação presente nas variáveis em estudo. O tema central da presente pesquisa é a abordagem de modelos de previsão aplicados a processos produtivos industriais com o objetivo de manter a estabilidade do processo por meio da utilização de técnicas de controle de qualidade. Esta investigação da estabilidade do processo passa pela verificação da influência da volatilidade na detecção de pontos amostrais que são potenciais para afetar o desempenho do processo produtivo.

1.2 Justificativa e importância do trabalho

Os modelos da família ARCH-GARCH, que tratam da modelagem e previsão da volatilidade, são de grande interesse na macroeconomia e especialmente em finanças. Ao invés de considerar a heterocedasticidade dos resíduos um problema, os modelos ARCH e GARCH, consideram como uma característica a ser modelada, ou melhor, como uma variância a ser modelada (SÁFADI e ANDRADE FILHO, 2007).

Ao realizar uma revisão de literatura sobre o assunto, verificou-se que os modelos da família ARCH-GARCH são aplicados apenas nas áreas de Economia e Econometria. A inovação dessa pesquisa é verificar se existe volatilidade em séries de processos produtivos industriais, e, se ela pode ser modelada através desses modelos heterocedásticos.

Segundo Morettin (2006), a maioria das técnicas de análises de séries temporais é aplicável em diversas áreas. Contudo, a variância condicional ou a volatilidade está sendo estudada, atualmente, apenas em séries financeiras.

A proposta dessa pesquisa justifica-se para mostrar a existência da volatilidade em processos produtivos, trabalhada até o momento, em variáveis econômicas. Além disso, investiga-se se o período em que houve a presença de volatilidade corresponde a um período de falta de controle detectado pelos gráficos de controle de Shewhart.

1.3 Problema da pesquisa

Tendo em vista o que foi até aqui abordado, pode-se definir o problema de pesquisa com o seguinte questionamento: verificar a influência dos modelos lineares e não-lineares na estabilidade do processo produtivo via análise de resíduos.

1.4 Objetivos

Nesta seção, descrevem-se os objetivos a serem traçados ao longo da pesquisa.

1.4.1 Objetivo geral

Modelar processos produtivos industriais autocorrelacionados com o fim de analisar a estabilidade do processo, por meio de gráficos de controle aplicado aos resíduos, oriundo de modelagem linear e não-linear.

1.4.2 Objetivos específicos

Para atingir o objetivo geral, propõem-se executar os seguintes objetivos específicos:

- realizar a modelagem linear para tratar a autocorrelação serial;
- identificar a presença de volatilidade em processos produtivos e modelar essa volatilidade por meio de modelos não-lineares de séries temporais;
- comparar os efeitos da correlação, utilizando gráficos de controle nos dados originais e nos resíduos sem a presença de autocorrelação serial;
- monitorar a volatilidade estimada no processo produtivo.

1.5 Delimitação da pesquisa

A delimitação da pesquisa é aplicar a metodologia de séries temporais em séries de processos produtivos, que possuam características de autocorrelação serial.

A pesquisa engloba três estudos ex-post-facto. Não será realizado um estudo comparativo para determinar quais são as vantagens e desvantagens da metodologia empregada individualmente para cada aplicação, até por que os processos produtivos são diferentes entre si. Além disso, o estudo fica restrito apenas as três pesquisas ex-post-facto estudadas, não fazendo referência a outros tipos de empresas, nem sequer a outros segmentos.

É importante lembrar que nem sempre é possível encontrar processos produtivos em que a modelagem do seu processo produtivo apresente resíduos com a presença de volatilidade. O método utilizado para estimar essa característica temporal é um modelo misto ARIMA-ARCH. Não são utilizados outros modelos econométricos como modelos GARCH, EGARCH, TARCH, FIGARCH, nem de volatilidade estocástica.

1.6 Organização do trabalho

Neste tópico, apresenta-se a estrutura do trabalho, disposta da seguinte forma:

O primeiro capítulo enfoca o tema da pesquisa, justificativa e importância do trabalho, problema de pesquisa, objetivos e a delimitação da pesquisa.

O segundo capítulo expõe a revisão de literatura que irá dar suporte para o desenvolvimento desta pesquisa.

O terceiro capítulo identifica a metodologia que conduziu o desenvolvimento desta pesquisa, elencando as etapas metodológicas para alcançar os objetivos propostos.

O quarto capítulo apresenta uma descrição do processo produtivo, das variáveis a serem analisadas e apresenta uma análise dos dados com a aplicação das técnicas, bem como os resultados obtidos, procurando compreender os efeitos desses resultados para alcançar os objetivos deste trabalho.

O quinto capítulo aborda uma análise conclusiva da pesquisa através da verificação dos objetivos, demonstrando se foram atingidos e, em seguida, são feitas as recomendações.

No sexto capítulo, são especificadas as referências bibliográficas utilizadas no desenvolvimento do trabalho.

2 REVISÃO DE LITERATURA

As constantes mudanças no mercado consumidor mundial têm exigido das empresas um grau máximo de racionalização, competitividade e modernização, tudo em busca de qualidade. Produtos com altos níveis de qualidade possuem maior durabilidade, o que, conseqüentemente, assegura uma boa imagem frente aos clientes. Sendo assim, é importante analisar o processo produtivo da empresa.

A qualidade está presente na linguagem da vida organizacional, profissional e pessoal, sendo que os colaboradores estão envolvidos em alcançá-la. Há diversas definições de qualidade empregadas nas organizações. Juran (1992) caracteriza a qualidade como sendo a “adequação ao uso”, porque existem vários usos e usuários para um determinado produto ou serviço, e o termo cliente inclui todos aqueles que são afetados por nossos processos e produtos. Na verdade, a qualidade é avaliada pelo usuário ou cliente. O objetivo é satisfazer o cliente com a “quantidade certa”.

Este capítulo apresenta a pesquisa bibliográfica realizada para a elaboração deste estudo. O embasamento teórico primeiramente apresentará a metodologia dos gráficos de controle, na seqüência, abordam-se os modelos de previsão, os modelos ARIMA, os modelos de heterocedasticidade condicional e o modelo de regressão dinâmica multivariada, seguidos dos principais testes estatísticos utilizados.

2.1 Gráficos de controle

Os gráficos de controle auxiliam na melhoria contínua da qualidade do produto, podendo ser usados para reduzir não-conformidades em tempo real, mostrando, quando necessário, a necessidade da existência de ações no processo que está sendo monitorado.

O objetivo principal da aplicação dos gráficos de controle é indicar quando os processos de produção sofrem alterações e o quanto essas mudanças podem afetar a qualidade

do produto. Também permitem verificar se o processo em análise é estatisticamente estável e se apresenta tendência de permanecer estável, indicando quando é necessário atuar sobre ele, e, possibilitam o aprimoramento contínuo do processo, mediante a redução de sua variabilidade (RYAN, 1989; MONTGOMERY, 1997).

Em qualquer processo de produção, uma determinada quantidade de variabilidade inerente ou natural sempre existirá. As causas para ocorrência de variações em processos produtivos são classificadas em: causas comuns ou aleatórias, pequenos efeitos no processo e causas especiais ou assinaláveis, que produzem grandes efeitos comprometendo a qualidade do produto.

Os problemas em um processo produtivo têm suas causas associadas aos elementos básicos do processo. Qualquer um dos elementos básicos, matéria-prima, máquina, medida, mão-de-obra, método e meio ambiente, de um processo produtivo pode sofrer variações e, por consequência, provocar reflexos negativos na qualidade do produto que, dessa forma, requer cuidadoso controle.

A variação provocada por causas comuns, também conhecida como variabilidade natural do processo, é aquela inerente ao processo, estando presente mesmo que todas as operações sejam executadas empregando métodos padronizados. Quando apenas as causas comuns atuam sobre o processo, a quantidade de variabilidade se mantém em uma faixa estável, conhecida como faixa característica do processo. Assim, diz-se que o processo está sob controle estatístico, apresentando um comportamento estável e previsível.

De acordo com Costa, Epprecht e Carpinetti (2004), Montgomery e Runger (2003), Montgomery (2004), Zanini (2006), a representação de um gráfico de controle é realizada por três linhas horizontais: a linha central (LC) representa a média dos valores encontrados da característica investigada, e, as outras linhas representam simetricamente em relação à central os valores referentes a três desvios-padrão (3σ), sendo estes, o limite superior de controle (LSC) e o limite inferior de controle (LIC). Os pontos amostrais assinalados no gráfico representam as médias ou medidas de cada subamostra observada por determinado tempo.

A Figura 1 apresenta um gráfico de controle que mostra quando o processo está sob controle e quando está fora de controle. Essas conclusões são extraídas a partir da interpretação dos limites de controle e dos testes de sequência, também chamados de *runs tests*, de modo a detectar mais rapidamente a presença de causas especiais de variação.

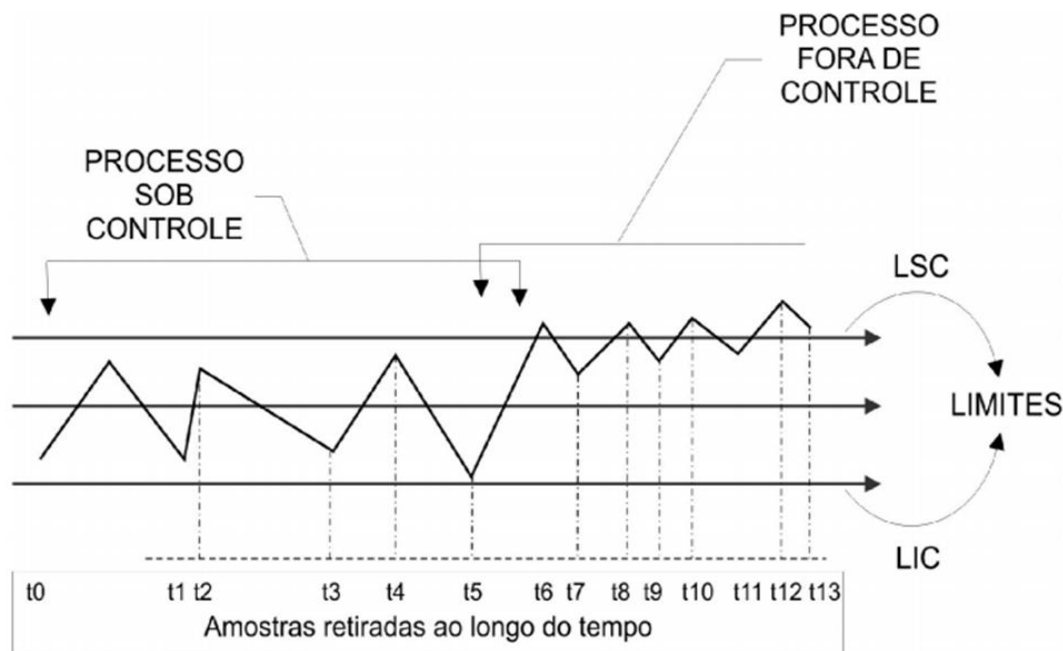


Figura 1 – Gráfico de controle adaptado de Montgomery, 2004

Fonte: Montgomery (2004)

Analisando a Figura 1, pode-se dizer que o processo está sob controle quando a maioria dos pontos amostrais grafados no gráfico está próximo da linha central; poucos pontos amostrais se espalham e se aproximam dos limites de controle, e nenhum dos pontos amostrais excede os limites de controle.

Conforme detalhado no capítulo 1, para que os resultados dos gráficos de controle sejam válidos, as seguintes suposições devem ser satisfeitas: as observações devem ser independentes e identicamente distribuídas (*iid*), seguindo uma distribuição normal com média zero e variância constante, ou seja, homocedásticos. Entretanto, os dados nem sempre são independentes e, conseqüentemente, os gráficos de controle não devem ser utilizados por conduzirem a alarmes falsos sobre o desempenho do processo. Maiores detalhes sobre a normalidade, a independência e a homocedasticidade serão apresentados posteriormente.

Nas situações em que não ocorra a independência entre as observações, ou seja, quando existe autocorrelação significativa nos dados do processo, Montgomery e Runger (2009) sugerem modelar a autocorrelação de modo que seja possível construir gráficos de controle mais confiáveis. Uma das soluções é estimar um modelo matemático do tipo ARIMA e, posteriormente, aplicar os gráficos de controle aos resíduos desse modelo. Para níveis de correlação serial baixos a moderados, um gráfico de controle de Shewhart aplicado às

observações será muito melhor para detectar uma mudança na média do processo do que um gráfico de Shewhart nos resíduos.

Contudo, se o processo é normalmente distribuído, as observações não são autocorrelacionadas, e a média é aproximadamente igual ao valor objetivo, espera-se que 99,7% das observações estejam dentro dos limites de controle. No entanto, se ocorrer mudanças na média do processo, presume-se que haverá observações fora dos limites de controle (MICHELSON, 1994).

Em geral, um gráfico de controle pode indicar uma condição fora de controle quando um ou mais pontos amostrais se localizam além dos limites de controle, ou quando exibem algum padrão de comportamento não-aleatório. Porém, segundo Montgomery (2004), um processo pode ser considerado fora de controle mesmo quando todos os pontos amostrais caem dentro dos limites de controle, ocorrendo um padrão de variação anormal no processo. A identificação dos pontos amostrais fora de controle é dada pelos testes de sequência também conhecidos por *runs tests* que podem ser aplicados aos gráficos de controle em busca de padrões não-aleatórios. Esses testes nos gráficos de controle de Shewhart são denominados de regras sensibilizantes

A probabilidade de ocorrência de um padrão anormal é, aproximadamente, igual à probabilidade de um ponto estar fora dos limites de $\pm 3\sigma$. A representação gráfica para a interpretação dos testes de sequência é apresentada na Figura 2.

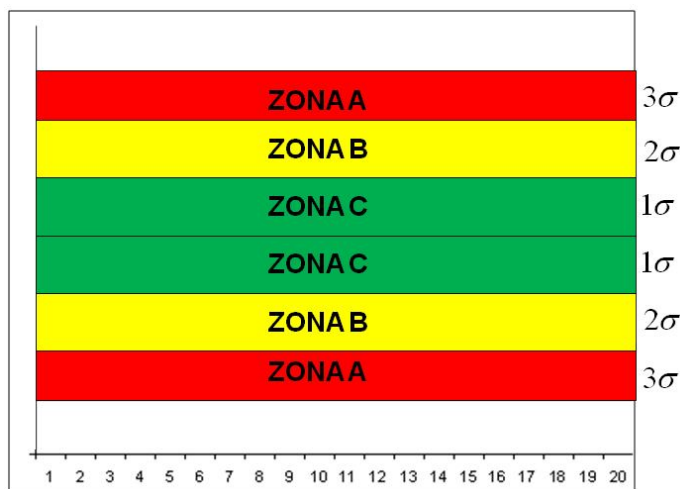


Figura 2 – Representação gráfica para os testes de sequência

Fonte: Montgomery e Runger (2003)

A Figura 2 mostra uma representação gráfica com limites de um-sigma, dois-sigma e três-sigma, chamados de limites de ação, isto é, quando um ponto se situa fora desse limite, procura-se uma causa especial e toma-se uma ação corretiva, se necessário. Esses limites dividem o gráfico de controle em três zonas: A, B e C, para cada lado da linha central, sob as quais se aplica um conjunto de regras de decisão com a finalidade de detectar padrões não-aleatórios, que podem indicar uma condição fora de controle estatístico.

Portanto, conclui-se que o processo está fora de controle se um ou mais dos critérios suplementares² apresentados a seguir forem verificados:

- um ou mais pontos amostrais fora dos limites de controle;
- dois ou três pontos amostrais consecutivos fora dos limites de alerta dois-sigma;
- quatro ou cinco pontos amostrais consecutivos além dos limites um-sigma;
- uma sequência de oito pontos amostrais consecutivos de um mesmo lado da linha central;
- seis pontos amostrais em uma sequência sempre crescente ou decrescente;
- quinze pontos amostrais em sequência na zona C (acima ou abaixo);
- quatorze pontos amostrais em sequência alternadamente para cima e para baixo;
- oito pontos amostrais em sequência de ambos os lados da linha central com nenhum na zona C;
- um padrão não-usual ou não-aleatório nos dados;
- um ou mais pontos amostrais perto dos limites de controle.

Segundo esse mesmo autor, as regras sensibilizantes podem ser úteis quando se aplica o gráfico de controle pela primeira vez e o foco é a estabilização de um processo fora de controle. Mas, uma vez que o processo esteja realmente estável, o uso rotineiro dessas regras sensibilizantes para detectar pequenas mudanças, ou para tentar reagir mais rapidamente às causas especiais, deve ser desencorajado. Os gráficos de controle da soma cumulativa e da média móvel exponencialmente ponderada (EWMA) devem ser usados quando são de interesse mudanças menores no processo.

Os gráficos de controle são classificados em duas categorias: gráficos de controle para variáveis e gráficos de controle para atributos, de acordo com a característica que se pretende

² Esses critérios, denominados como regras sensibilizantes para os gráficos de controle de Shewhart, foram, inicialmente, sugeridos por Western Electric Handbook (1956) para aumentar a sensibilidade dos gráficos de controle a uma pequena mudança no processo, de modo a responder mais rapidamente a uma causa especial, e estão citados em Montgomery (2004).

monitorar. Os gráficos de controle para variáveis são mais sensíveis do que os para atributos. Eles podem alertar para problemas antes que uma real rejeição ocorra (MONTGOMERY e RUNGER, 2009; COSTA, EPPRECHT e CARPINETTI, 2004).

Existem vários tipos de gráficos de controle para variáveis contínuas. Descreve-se a seguir, os gráficos de controle de Shewhart para medidas individuais e amplitude móvel, empregados nesta pesquisa.

2.1.1 Gráficos de controle para medidas individuais e amplitude móvel

Existem muitas situações nas quais o tamanho da amostra para monitoramento do processo é $n = 1$, isto é, a amostra consiste de uma única medida individual. Citam-se alguns exemplos de situações em que os gráficos de controle para medidas individuais são úteis:

- inspeção e medição automática em toda a unidade produzida;
- taxa de produção muito lenta, sendo inconveniente acumular amostras $n > 1$;
- várias medidas são tomadas em uma mesma unidade do produto;
- nos resíduos de modelos matemáticos ajustados aos dados que possuam autocorrelação.

Além de verificar a presença de autocorrelação, Montgomery (1997) afirma que é importante verificar a hipótese de normalidade e homocedasticidade antes de aplicar os gráficos de controle para medidas individuais. Como esses gráficos serão aplicados aos resíduos de modelos matemáticos estimados, as suposições necessárias para aplicá-los corretamente são plenamente satisfeitas.

Os gráficos de controle para medidas individuais e amplitude móvel também são conhecidos por gráficos de \bar{x} e \overline{MR} . Esse tipo de gráfico de controle possui habilidade para detectar grandes deslocamentos na média do processo. Um deslocamento na média do processo resultará em um ponto amostral (ou mais) fora dos limites de controle ou de um padrão, consistindo de uma sequência em um dos lados da linha central.

A média entre as amostras, a amplitude móvel e a média da amplitude móvel são definidas respectivamente por:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} \quad (1)$$

onde x_1, x_2, \dots, x_n é uma amostra de tamanho n e \bar{x} é a média dessa amostra.

$$MR_i = |x_i - x_{i-1}| \quad (2)$$

onde MR_i é a amplitude móvel de duas observações consecutivas como base para estimar a variabilidade do processo.

$$\overline{MR} = \frac{MR_1 + MR_2 + \dots + MR_n}{n} \quad (3)$$

onde \overline{MR} é a média das amplitudes móveis.

Para o gráfico de medidas individuais (\bar{x}), os limites situados a três desvios-padrão da média são dados por:

$$LSC = \bar{x} + 3 \frac{\overline{MR}}{d_2} \quad (4)$$

$$LC = \bar{x} \quad (5)$$

$$LIC = \bar{x} - 3 \frac{\overline{MR}}{d_2} \quad (6)$$

Para o gráfico de amplitude móvel (\overline{MR}), os limites de controle são dados pelas expressões:

$$LSC = D_4 \overline{MR} \quad (7)$$

$$LC = \overline{MR} \quad (8)$$

$$LIC = D_3 \overline{MR} \quad (9)$$

As constantes d_2 , D_4 e D_3 encontram-se tabuladas no Anexo A.

A próxima seção apresenta os modelos de previsão empregados nesta pesquisa.

2.2 Modelos de previsão

De forma geral, há vários modelos para realização de previsões, utilizando séries temporais: modelo de regressão simples, modelo de regressão com equações simultâneas, modelo autorregressivo integrado de média móvel (ARIMA), modelo vetor autorregressivo (VAR) e modelo autorregressivo de heterocedasticidade condicional (ARCH) (GUJARATI, 2000).

Neste trabalho, são empregados os modelos ARIMA para modelar a média condicional, os modelos mistos ARIMA-ARCH para modelar a média e a variância condicional, também chamada de volatilidade, e, o modelo de regressão dinâmica multivariada, que realiza uma estimação conjunta, considerando as inter-relações e o comportamento dinâmico entre as variáveis em dados de processos produtivos.

2.2.1 Modelos ARIMA

Uma série temporal é um conjunto de observações ordenadas em intervalos de tempo, comumente iguais. Essas observações apresentam dependência serial e constituem um dos objetivos do estudo de séries temporais, analisar e modelar essa dependência (EHLERS, 2007).

Em algumas situações, o objetivo de aplicar séries temporais é fazer previsões de valores futuros, enquanto, em outras, a estrutura da série ou sua relação com outras séries pode ser o interesse principal. De um modo geral, segundo Ehlers (2007), os principais objetivos da análise de séries temporais consiste em:

- compreender o mecanismo gerador da série: descrever efetivamente o comportamento da série; encontrar periodicidades na série; tentar obter razões para o comportamento da série (possivelmente através de variáveis auxiliares) e controlar a trajetória da série;
- prever o comportamento futuro da série, possibilitando: fazer planos a longo, médio e curto prazo e tomar decisões apropriadas.

Uma das técnicas usadas pela econometria contemporânea para a previsão do comportamento de variáveis se constitui no emprego de modelos univariados. A construção dos modelos de séries temporais univariados é fundamentada na teoria de que existe uma grande quantidade de informação presente na série de dados, sendo esses dados capazes de fornecer estimativas sobre o comportamento futuro da variável. Assim, será apenas o próprio comportamento da variável que responderá pela sua dinâmica futura. Este tipo de modelo é conhecido na literatura como o modelo autorregressivo integrado de médias móveis, ou definido simplesmente por modelo ARIMA.

A construção do modelo ARIMA parte da concepção de que as séries temporais envolvidas na análise são geradas por um processo estocástico estacionário, cuja natureza pode ser representada a partir de um modelo matemático. Segundo Box, Jenkins e Reinsel (1994), um processo é estacionário quando ele oscila em torno de uma média constante, com variância também constante, conhecida por homocedástica, e, é definido como estocástico por ser controlado por leis probabilísticas.

O ponto importante a destacar é que, para usarmos a metodologia proposta por Box e Jenkins em 1970, deve-se ter uma série temporal estacionária ou uma série temporal que se torna estacionária depois de uma ou mais diferenciações. Sendo assim, os modelos são classificados em estacionários e não-estacionários.

Apresentam-se a seguir os modelos estacionários em conjunto com o modelo não-estacionário, ARIMA. Maiores detalhes ver Box e Luceño (1997), Bueno (2008), Del Castillo (2002), Hill, Griffiths e Judge (1999) e Morettin (2008).

- Modelo AR(p)

Um modelo ou processo autorregressivo é caracterizado quando uma variável aleatória está relacionada com os próprios valores passados e com os erros aleatórios. A forma genérica de um modelo autorregressivo, sendo denotado por AR(p), onde p indica a ordem do modelo, é definido por:

$$x_t - \mu = \phi_1(x_{t-1} - \mu) + \phi_2(x_{t-2} - \mu) + \dots + \phi_p(x_{t-p} - \mu) + e_t \quad (10)$$

onde μ , ϕ_1, \dots, ϕ_p são os parâmetros reais e $e_t \sim RB(0, \sigma^2)$. Segue-se que $E(x_t) = \mu$ e escreve-se o processo na forma

$$x_t = \phi_0 + \phi_1 x_{t-1} + \dots + \phi_p x_{t-p} + e_t \quad (11)$$

$$\mu = E(x_t) = \frac{\phi_0}{1 - \phi_1 - \dots - \phi_p} \quad (12)$$

O modelo AR(1) é representado pela expressão (13)

$$x_t = \phi_1 x_{t-1} + e_t \quad (13)$$

A identificação do modelo será feita por meio das funções de autocorrelação (FAC) e autocorrelação parcial (FACP), que indicam a ordem do modelo, conforme a Figura 3.

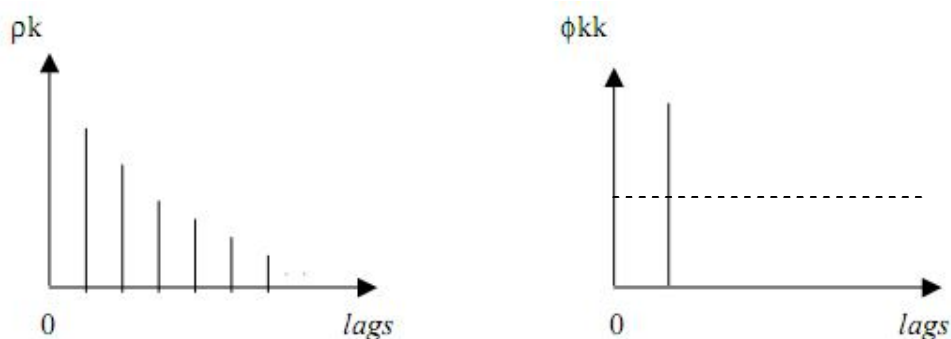


Figura 3 – Comportamento da FAC e FACP de um modelo AR(1)

- Modelo MA(q)

O nome do modelo vem do fato de que x_t é uma soma algébrica ponderada dos ε_t que se movem no tempo. Um modelo de médias móveis de ordem q , MA(q), é definido por:

$$x_t = \mu + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \quad (14)$$

onde $\mu, \theta_1, \dots, \theta_q$ são constantes reais e $e_t \sim RB(0, \sigma^2)$.

Segue-se que x_t é estacionário, de média μ e como e_t são não-correlacionados, a variância do processo pode ser obtida pela expressão (15)

$$\sigma_x^2 = \sigma^2(1 + \theta_1^2 + \dots + \theta_q^2) \quad (15)$$

Em particular, para um processo MA(1), tem-se a seguinte expressão

$$x_t = e_t - \theta_1 e_{t-1} \quad (16)$$

A identificação do modelo é feita através da FAC e FACP, sendo que a FAC fornece a ordem do modelo, conforme a Figura 4.

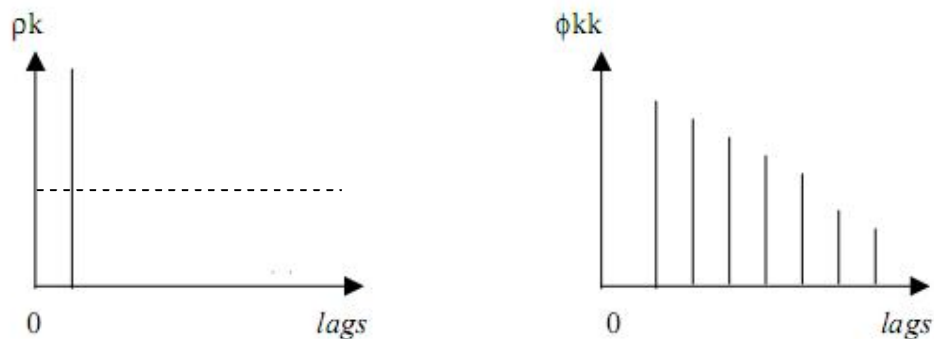


Figura 4 – Comportamento da FAC e FACP de um modelo MA(1)

- *Modelo ARMA(p,q)*

O processo autorregressivo de médias móveis é, simplesmente, a combinação dos processos vistos anteriormente, descrito por seus valores passados (parte AR) e pelos erros aleatórios correntes e passados (parte MA), denotado por ARMA(p,q). Esse processo está definido na expressão (17)

$$x_t - \mu = \phi_1(x_{t-1} - \mu) + \dots + \phi_p(x_{t-p} - \mu) + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \quad (17)$$

onde $e_t \sim RB(0, \sigma^2)$. Segue-se que a média do processo é μ . Usando os operadores autorregressivos e de médias móveis, o modelo pode ser escrito na forma:

$$\phi(B)\tilde{x}_t = \theta(B)e_t \quad (18)$$

onde $\tilde{x}_t = x_t - \mu$, sendo que $\mu = 0$. Quando $\mu \neq 0$, o modelo é escrito dessa forma:

$$x_t = \phi_0 + \phi_1 x_{t-1} + \dots + \phi_p x_{t-p} + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \quad (19)$$

onde $\theta_0 = \mu(1 - \phi_1 - \dots - \phi_p)$.

Um modelo frequentemente usado é o ARMA(1,1), ou seja,

$$x_t = \phi_1 x_{t-1} + e_t - \theta_1 e_{t-1} \quad (20)$$

- *ARIMA(p,d,q)*

Como grande parte das séries temporais são naturalmente não-estacionárias, é necessário aplicar um modelo que transforme o conjunto de dados em estacionários. Os modelos ARIMA são eficientes para atingir a estacionariedade na série, recorrendo-se ao processo de diferenciações, ao passo que as ordens p e q podem ser obtidas pela FAC e FACP, sendo auxiliada pelos critérios penalizadores que serão apresentados na seção 2.3.5.

Segundo Morettin (2006), os modelos ARIMA são modelos responsáveis por capturar e modelar processos de memória curta, pois a FAC decresce rapidamente para zero. Caso o processo não tenha esta característica, outros modelos devem ser empregados para estimar processos de memória longa.

Um processo $\{x_t, t \in Z\}$ segue um processo ARIMA(p,d,q) se $\Delta^d x_t$ seguir um modelo ARMA(p,q), ou seja,

$$\phi(B)\Delta^d x_t = \theta(B)e_t \quad (21)$$

onde B é operador retroativo, ou operador de defasagens, d é inteiro positivo e representa a diferença, conforme a ordem de integração; ϕ e θ são os parâmetros dos processos autorregressivos e de média móvel de ordem p e q , respectivamente, e $e_t \sim RB(0, \sigma^2)$.

A construção de modelos ARIMA segue um ciclo iterativo da metodologia de Box e Jenkins, que é composto por quatro etapas: a) identificação; b) estimação; c) verificação do diagnóstico e d) previsão (GUJARATI, 2000). Esses modelos também são discutidos pelos autores Box, Jenkins e Reinsel (1994), Makridakis, Wheelwright e Hyndman (1998), Morettin e Toloi (2004) e Morettin (2008).

No Quadro 1, apresentam-se os comportamentos da FAC e FACP para o modelo ARIMA(p,d,q).

Ordem	(1,d,0)	(0,d,1)
Comportamento de ρ_k	Decai exponencialmente	Somente $\rho_1 \neq 0$
Comportamento de ϕ_{kk}	Somente $\phi_{11} \neq 0$	Decaimento exponencial dominante
Estimativas iniciais	$\phi = \rho_1$	$\rho_1 = -\frac{\theta}{1+\theta^2}$
Região de admissibilidade	$-1 < \phi < 1$	$-1 < \theta < 1$
Ordem	(2,d,0)	(0,d,2)
Comportamento de ρ_k	Mistura de exponenciais ou ondas senóides amortecidas	Somente $\rho_1 \neq 0$ e $\rho_2 \neq 0$
Comportamento de ϕ_{kk}	Somente $\phi_{11} \neq 0$ e $\phi_{22} \neq 0$	Dominada por mistura de exponenciais ou senóides amortecidas
Estimativas iniciais	$\begin{cases} \phi_1 = \frac{\rho_1(1-\rho_2)}{1-\rho_1^2} \\ \phi_2 = \frac{\rho_2-\rho_1^2}{1-\rho_1^2} \end{cases}$	$\begin{cases} \rho_1 = \frac{-\phi_1(1-\theta_2)}{1+\theta_1^2+\theta_2^2} \\ \rho_2 = \frac{-\theta_2}{1+\theta_1^2+\theta_2^2} \end{cases}$
Região de admissibilidade	$\begin{cases} -1 < \phi_2 < 1 \\ \phi_2 - \phi_1 < 1 \\ \phi_2 + \phi_1 < 1 \end{cases}$	$\begin{cases} -1 < \theta_2 < 1 \\ \theta_2 - \theta_1 < 1 \\ \theta_2 + \theta_1 < 1 \end{cases}$
Ordem	(1,d,1)	
Comportamento de ρ_k	Decai exponencialmente após o lag 1	
Comportamento de ϕ_{kk}	Dominada por decaimento exponencial após o lag 1	
Estimativas iniciais	$\rho_1 = \frac{(1-\phi\theta)(\phi-\theta)}{1+\theta^2-2\phi\theta}$, $\rho_2 \neq \rho_1$	
Região de admissibilidade	$-1 < \phi < 1$, $-1 < \theta < 1$	

Quadro 1 – Comportamentos da FAC e da FACP de um processo ARIMA(p,d,q)

Fonte: Morettin e Toloi, 2004

A etapa de identificação do modelo é a etapa mais difícil da metodologia de Box e Jenkins. Muitas vezes, identificam-se vários modelos candidatos para representar o processo gerador da série em estudo. Isso porque, trabalhando com a FAC e a FACP amostrais, torna-se complexo, muitas vezes, decidir se elas são decrescentes ou se são truncadas.

Na tentativa de encontrar o melhor modelo que represente corretamente o processo de gerador da série, muitos pesquisadores preferem utilizar procedimentos complementares às funções de autocorrelação, como os critérios penalizadores AIC (Critério de Informação Akaike) e SBC (Critério Bayesiano de Schwarz). Esse procedimento faz uso de critérios de seleção de modelos construídos com base na variância estimada, no tamanho da amostra e nos valores de p e q .

A presença de p e q , nas fórmulas dos critérios penalizadores, AIC e SBC, tem por objetivo penalizar os modelos com muitos parâmetros, levando em consideração que modelos mais parcimoniosos devem ser privilegiados, por apresentarem menor número de parâmetros a serem estimados.

Na prática, quanto mais parâmetros são estimados no mesmo período da amostra, menor será o erro estimado, mas isso será penalizado na segunda parcela da estatística. Por isso, escolhe-se aquela especificação que apresentar o menor valor para AIC ou SBC.

O processo de estimação de um modelo ARIMA pode ser realizado pelo método dos mínimos quadrados ordinários (MQO), que é um método simples, ou através do método de máxima verossimilhança (MV). Diversos autores indicam que MV é superior ao MQO quando o tamanho da série é pequeno.

A etapa de verificação da metodologia de Box e Jenkins consiste em verificar se o modelo identificado e estimado é adequado. Se o modelo for adequado, pode-se adotá-lo para fazer a previsão. Caso contrário, outra especificação deve ser escolhida para modelar a série, o que implica refazer as etapas de identificação e estimação dos parâmetros.

A verificação é uma das etapas mais importantes, e, é baseada na análise dos resíduos obtidos do modelo. Nessa etapa, busca-se por resíduos do tipo ruído branco.

Definição 1 – Ruído branco: Diz-se que $\{e_t, t \in x\}$ é um ruído branco discreto se as variáveis aleatórias e_t são não-correlacionadas, isto é, $Cov\{e_t, e_s\} = 0, t \neq s$. Admitindo que, e_t seja um ruído branco, com média μ e variância σ^2 , denota-se $e_t \sim RB(\mu, \sigma^2)$.

Assim, pode-se admitir que um processo ruído branco caracteriza-se com média μ , variância σ^2 , não-autocorrelacionado e que segue uma distribuição normal. Obviamente, se

as variáveis aleatórias são independentes, elas também serão não-autocorrelacionadas. Para tanto, os coeficientes de autocorrelação devem ser, estatisticamente, iguais a zero.

Uma das formas mais simples de identificar se existe autocorrelação nos resíduos do modelo estimado é observar se os coeficientes de autocorrelação são iguais a zero através da FAC e da FACP, aplicadas à série de resíduos. Se esses coeficientes forem maiores que zero, significa que ainda existem informações, as quais não foram captadas pelo modelo, podendo gerar previsões pobres, pois o modelo não é adequado.

Um dos testes conjuntos empregados para verificar as autocorrelações dos resíduos estimados é a estatística de *Ljung-Box*, cujas hipóteses, nula e alternativa, são melhores inferidas, pois as autocorrelações são identificadas conjuntamente. Essa estatística tem distribuição Qui-quadrado com p graus de liberdade.

É importante destacar que pode ocorrer de os resíduos serem ruído branco, mas, se em cada instante de tempo o resíduo for elevado ao quadrado, tem-se uma série que representa a variância condicional ou a volatilidade dos resíduos, podendo apresentar estruturas de autocorrelação. Se isso acontecer, essa volatilidade pode ser modelada por meio de modelos não-lineares, denominados modelos autorregressivos de heterocedasticidade condicional (ARCH), apresentados na próxima seção.

Além de verificar a autocorrelação nos resíduos, necessita-se verificar a normalidade dos resíduos e sinais de heterocedasticidade condicional nos resíduos quadráticos, apresentados na seção 2.3.

A seleção do modelo matemático representativo para a série sob estudo ocorre de acordo com o modelo que apresentar o menor valor para AIC e SBC entre os modelos concorrentes estimados. Após a escolha do modelo matemático adequado à série, chega-se à última etapa da metodologia de Box e Jenkins, que consiste na realização de previsões da série em instantes de tempo posteriores a n .

Feitas as previsões e definidos os intervalos de confiança, algumas medidas de desempenho são utilizadas para avaliar as previsões. De acordo com Bueno (2008), as três medidas básicas são: erro médio quadrático (MSE), erro médio absoluto (MAE) e erro médio percentual absoluto (MAPE). Além das medidas mencionadas pelo autor, existem muitas outras, entre elas, a estatística U de Theil.

Para calcular as medidas de desempenho, deve-se deixar uma porção de observações fora da amostra, ou seja, em uma amostra com $T + H$ observações, deixam-se as últimas H observações fora da amostra, e, estima-se o modelo com as T observações restantes.

Os critérios, MAPE e U de Theil, utilizados para avaliação das previsões, serão apresentados na seção 2.3.6.

2.2.2 Modelos ARCH

O termo volatilidade, até pouco tempo atrás, era descrito como movimentos erráticos que não seguiam um padrão regular em uma série temporal. Ele é representado por uma série de resíduos aleatórios que resultariam de uma série temporal após a retirada dos componentes da série (LAMOUNIER, 2006).

No início da década de 1980, teve início o surgimento dos modelos autorregressivos de heterocedasticidade condicional, ARCH, propostos por Engle. Posteriormente, esses modelos foram estendidos por Bollerslev, e, denominados de GARCH.

Estudos como os de Engle (1982) e Bollerslev (1986) indicaram que uma série de resíduos poderia não apresentar um comportamento não-aleatório puro. Eles verificaram que, em grande parte das séries temporais, particularmente em financeiras, a variância não é constante. Isso significa que a série não apresenta a propriedade desejável de homocedasticidade. Sendo assim, essa variância condicional, também chamada de volatilidade, pode ser modelada utilizando-se modelos não-lineares ditos heterocedásticos.

Segundo Silva (2000), os estimadores obtidos com os modelos ARCH podem ser mais eficientes caso a variância seja modelada. Porém, a especificação desse modelo de regressão deve ser adotada com maior cuidado, pois a má escolha da forma funcional pode levar a problemas com heterocedasticidade. Infelizmente, ainda existe uma limitada quantidade de métodos que corrijam os problemas dela decorrentes.

De acordo com Bueno (2008), a principal razão do surgimento desses modelos é que, os modelos econométricos de séries temporais, fossem os financeiros ou macroeconômicos, enfatizavam apenas o primeiro momento condicional (média). As dependências temporais de ordem superior eram perturbações aleatórias, consideradas em seus momentos incondicionais. Essas dependências expressavam a existência de aglomerações na série e a alternância de períodos de baixa volatilidade com períodos de alta volatilidade, fenômeno denominado de *clusters* de volatilidade. Assim, desenvolveram-se técnicas que permitem a modelagem

temporal de variâncias e covariâncias. De fato, os modelos heterocedásticos fundamentam-se na estimação da variância condicional, em vez de considerá-la constante ao longo do tempo.

A contribuição principal desses modelos é a distinção de uso entre momentos de segunda ordem condicionais e não-condicionais. Enquanto a matriz de covariância não-condicional para as variáveis de interesse pode ser invariante no tempo, a matriz de covariância condicional depende de estados passados da natureza.

Assim, os modelos de heterocedasticidade condicional da família ARCH são modelos não-lineares utilizados na determinação da volatilidade de séries temporais, principalmente financeiras, onde modelos lineares se provaram incompletos em função da incapacidade de explicar algumas características dessas séries. Esse modelo é usado para dados com coletas de tempo bem pequenas, por exemplo, de uma em uma hora, de quinze em quinze minutos. Esse tipo de formulação precisa que as observações sejam equidistantes no tempo, o que impede seu uso para modelagem de dados de baixa frequência.

Diversas pesquisas tem sido realizadas por meio dos modelos ARCH, citam-se algumas:

Berti (2005) realizou uma avaliação empírica da previsão de volatilidade do índice BOVESPA, tanto de modelos tradicionais em base diária, como algumas adaptações para a inclusão da informação intradiária, com a intenção de obter melhores previsões no horizonte de um dia à frente. A conclusão do trabalho aponta a viabilidade do uso dos dados intradiários para realizar previsões de horizonte diário, e mais, há indícios de que podem melhorar a acurácia das previsões.

Souza (2005) apresentou uma metodologia de modelagem e previsão da volatilidade do café Arábica, com o intuito não só de explicar e prever sua heterocedasticidade condicional, como de obter o risco de mercado para seus investimentos. Os resultados revelaram o modelo GARCH ajustado aos resíduos de um modelo AR, dentre os modelos considerados, como o modelo de maior poder preditivo um passo à frente.

Goulart (2004) apresentou um estudo comparativo da capacidade preditiva dos modelos EWMA (adotado pela Clearing de Câmbio da BM&F), GARCH(1,1), EGARCH(1,1) e TARARCH(1,1), quando utilizados para prever a volatilidade das taxas de câmbio praticadas no mercado interbancário brasileiro. Os resultados demonstraram que o modelo TARARCH(1,1) apresenta o melhor desempenho preditivo, acompanhado de perto pelo modelo EGARCH(1,1), seguindo-se o modelo GARCH(1,1), e, por último, o modelo EWMA.

Além disso, constatou-se que todos os modelos revelaram uma propensão a superestimar a volatilidade futura.

Apesar de os modelos não-lineares ARCH serem aplicados inicialmente em séries financeiras, constata-se que muitas séries também podem apresentar alguma estrutura de autocorrelação nos resíduos quadráticos, de modo que modelos ARIMA podem ser inicialmente aplicados para remover essa autocorrelação antes de usar os modelos ARCH. Pode-se afirmar que, muitas vezes, os modelos ARIMA não conseguem remover completamente a autocorrelação devido à presença da volatilidade, justificando a necessidade de estimar um modelo da família ARCH, ou, um modelo misto, que estima a média condicional e a variância condicional conjuntamente.

O modelo ARCH, propriamente dito, assume que, em uma regressão, a variância do erro se correlaciona com a variável independente (erro da regressão). Desta forma, a variância é dita condicional (não constante) e seu valor depende da observação do próprio erro (ÉLIN DUXUS, 2009).

A idéia principal do modelo ARCH é o fato de que a variância de e_t , no período de tempo t , depende de e_{t-1}^2 . O termo de erro e_t , condicionado à informação disponível no período $(t-1)$ seria distribuído conforme a seguinte notação:

$$e_t \sim N[0, (\alpha_0 + \alpha_1 e_{t-1}^2)] \quad (22)$$

em que α_0 e α_1 são parâmetros explicativos da variância do termo de erro e_t (CAMPOS, 2007).

Um modelo ARCH(m), em que m denota a ordem do modelo, expressa a variância condicional do modelo anterior para a média condicional como uma função das inovações quadráticas passadas (SILVA, SÁFADI e CASTRO JÚNIOR, 2005).

Segundo Morettin (2008), o modelo ARCH(m) é definido por:

$$r_t = \sqrt{h_t} e_t \quad (23)$$

$$h_t = \alpha_0 + \alpha_1 r_{t-1}^2 + \dots + \alpha_m r_{t-m}^2 \quad (24)$$

em que e_t é *iid* com média zero, e, h_t é a variância condicional.

Para assegurar que a variância condicional seja positiva e fracamente estacionária, as seguintes restrições paramétricas são necessárias: $\alpha_0 > 0$, $\alpha_t \geq 0$ para todo $t = 2 \dots m$ e

$$\sum \alpha_t < 1.$$

Sob a condição de estacionariedade, a variância será dada por:

$$\text{Var}(r_t) = \frac{\alpha_0}{1 - \sum_{t=1}^m \alpha_t} \quad (25)$$

Pela própria definição, valores grandes de r_t são seguidos por outros valores grandes da série. Se admitirmos que r_t siga uma modelo ARCH, as caudas serão mais pesadas do que as da normal, o que é uma propriedade característica do modelo.

Na prática, seguindo o modelo original de Engle, assume-se que e_t tem distribuição normal ou *t-Student* padronizada (SILVA, SÁFADI e CASTRO JÚNIOR, 2005).

Campos (2007) afirma que, se a assimetria é diferente de zero e a curtose é maior que três, isso significa que a série pode exibir sinais de volatilidade e de *clusters* de volatilidade. A assimetria retrata a forma de distribuição dos dados, e, a curtose indica que a série apresenta uma distribuição leptocúrtica em relação à normal, isto é, os dados estão agrupados no centro juntamente com algumas observações nos extremos das caudas.

Apresentam-se, a seguir, as etapas pertencentes aos modelos ARCH. Maiores detalhes encontram-se disponíveis em Morettin (2006), Bueno (2008), Enders (1995) e Hamilton (1994).

Identificação

Inicialmente, antes de efetuar a construção de modelos ARCH, é preciso ajustar um modelo matemático para remover a autocorrelação na série, se esta existir. Nesse estudo será estimado um modelo ARIMA.

Para verificarmos se a série apresenta heterocedasticidade condicional, examina-se a série r_t^2 , utilizando testes específicos para esse propósito, tais como:

- Teste de *Ljung-Box*, ver seção 2.3.3;
- Teste ARCH-LM (multiplicador de Lagrange), descrito na seção 2.3.4;
- Estatística *F*, teste assintoticamente equivalente que pode ter propriedades melhores para amostras pequenas.

A identificação da ordem de um modelo para uma série é usualmente difícil, assim como, no processo de identificação dos modelos ARIMA. Portanto, além dos testes específicos apresentados anteriormente, aconselha-se definir o melhor modelo entre os modelos concorrentes de acordo com os critérios penalizadores.

Estimação

Para a estimação dos parâmetros, usa-se o método de máxima verossimilhança, cuja implementação pressupõe que se faça uma hipótese sobre a verdadeira distribuição. Escolhe-se, em geral, a distribuição normal (CAMPOS, 2007).

Entretanto, Hentschel (1995 apud BUENO, 2008) argumenta que a presença de excesso de curtose sugere o uso da distribuição *t-Student* ou distribuição erro generalizado (GED) para a estimação, ainda que esta venha a falhar se a série possuir elevado número de distribuições extremas. Não obstante, mesmo na falta de normalidade dos erros, pode-se utilizar a distribuição normal, entendendo isso como uma estimação de quase-máxima-verossimilhança.

Admitindo que a distribuição dos erros seja normal, a função de verossimilhança é a seguinte, segundo Morettin (2008):

$$L(r_1, \dots, r_T | \alpha) = \prod_{t=m+1}^T \frac{1}{(2\pi h_t)^{1/2}} \exp\left\{-\frac{r_t^2}{2h_t}\right\} f(r_1, \dots, r_m | \alpha) \quad (26)$$

em que:

$(r_1, \dots, r_T | \alpha)$ é o vetor de parâmetros a serem estimados da equação da média e da variância;

h_t são as variâncias iniciais do processo de otimização, necessárias ao algoritmo;

m é o número de variâncias para iniciar o processo de otimização.

A obtenção dos parâmetros iniciais é obtida por MQO ou por ajustamento de um modelo da família ARIMA, ou, ainda, pelos parâmetros obtidos com o teste ARCH-LM.

Verificação

Para um modelo ARCH(r) com e_t normal ou *t-Student*, os resíduos padronizados

$$\tilde{r}_t = \frac{r_t}{\sqrt{h_t}} \quad (27)$$

são variáveis aleatórias *iid* com distribuição normal padrão ou *t-Student*. Logo, uma maneira de verificar a adequabilidade do modelo é calcular a estatística $Q(K)$ de *Ljung-Box* para a sequência \tilde{r}_t . Além disso, podem-se calcular os coeficientes de assimetria e curtose estimados e fazer o gráfico Q-Q plot a fim de avaliar a suposição de normalidade.

Para a verificação da existência de heterocedasticidade condicional nos resíduos, pode-se aplicar o teste ARCH-LM para a série dos resíduos quadráticos padronizados (\tilde{r}_t^2).

Morettin e Tolo (2004) afirmam que, devem-se obter modelos de ordem baixa, e, escolher o melhor com base nos critérios penalizadores, AIC ou SBC, de acordo com os valores assumidos pela assimetria e curtose, valores da log-verossimilhança ou de alguma função perda.

Previsão

Nos modelos econométricos convencionais, com variância não-condicional, a incerteza do erro de predição é crescente com o horizonte de predição e não depende da origem da previsão. Na presença de modelos ARCH, a acuidade das previsões dependerá do conjunto de informações disponível de forma não trivial. Assim, a construção do intervalo de confiança do erro e os testes de estabilidade estrutural amostrais exigem a avaliação das variâncias condicionais do erro futuro.

O problema reside no fato de que a distribuição dos erros é leptocúrtica, alterando sensivelmente a maneira de se construir o intervalo de confiança. Na realidade, ainda estão sendo feitas pesquisas no sentido de encontrar as formulações corretas para a determinação do intervalo de confiança e da própria função de predição. Não obstante, a maneira das previsões segue os mesmos passos do modelo ARIMA. Prevê-se um passo à frente, com o resultado obtido, faz-se o passo seguinte, e assim sucessivamente.

As previsões para a volatilidade utilizando o modelo ARCH(m) são obtidas recursivamente. Assim,

$$\hat{h}_t(1) = \alpha_0 + \alpha_1 r_t^2 + \dots + \alpha_m r_{t-m+1}^2 \quad (28)$$

é a previsão um passo a frente (h_{t+1}), com origem fixada no instante t . As previsões ℓ passos a frente com origem em t , são dadas pela expressão (29)

$$\hat{h}_t(\ell) = \alpha_0 + \sum_{i=1}^m \alpha_i \hat{h}_t(\ell - i) \quad (29)$$

em que $\hat{h}_t(\ell - i) = r_{t+\ell-i}^2$, se $\ell - i \leq 0$.

Portanto, para aplicação dos modelos ARCH, é necessário que a série temporal analisada apresente volatilidade nos seus resíduos quadráticos oriundos do modelo estimado para a série de interesse.

No entanto, quando o processo gerador da série original é um modelo ARIMA e seu resíduo ao quadrado apresentar autocorrelação, sugerindo dependência não-linear entre eles, seguindo um modelo ARCH, ajusta-se aos resíduos quadráticos um modelo misto do tipo ARIMA-ARCH para a média e variância condicional, respectivamente. Com base nesse modelo, surgem outras extensões derivadas das combinações desses dois modelos, por exemplo, AR-ARCH, ARMA-ARCH e ARFIMA-ARCH (COSTA e BAIDYA, 2001).

Nos dados desta pesquisa, foram aplicados modelos mistos ARMA-ARCH e AR-ARCH, que apontam para uma estimação conjunta dos seus parâmetros, modelando, assim, a média e a variância condicional para a série sob estudo.

2.2.3 Modelo de regressão dinâmica multivariada

Os modelos de regressão dinâmica combinam a dinâmica de séries temporais e o efeito de variáveis explicativas. Atenta-se que o termo “regressão dinâmica” não indica que os parâmetros do modelo evoluem no tempo, que é o caso dos modelos de espaço de estados que usam o Filtro de Kalman. Nessa metodologia, a palavra “dinâmica” significa um modelo de regressão no qual se inclui a estrutura de dependência de uma série temporal (HAMILTON, 1994).

Modelos de regressão dinâmica podem ser usados quando existe uma estrutura de dependência entre a variável de interesse e possíveis variáveis exógenas e, ao mesmo tempo, quando a estrutura de correlação da série dependente indicar que não se pode supor a independência dos erros. Inicialmente, estimam-se equações separadas, se os resíduos de cada equação são correlacionados, então, a aplicação de modelos de regressão dinâmica é aceitável.

Nos modelos de regressão dinâmica, a variável dependente é explicada por seus valores defasados e pelos valores atuais e passados de variáveis exógenas (ZANINI, SOUZA e PEDREIRA, 2000). Entretanto, a estimação nesse tipo de modelo envolve um procedimento iterativo com vários estágios. Assim, a elaboração de um modelo de regressão dinâmica é, muitas vezes, um procedimento difícil, pois é preciso não apenas escolher as variáveis a serem incluídas no modelo, mas também as defasagens destas variáveis. Essa metodologia é conhecida por regressões aparentemente não-correlacionadas, também denominadas de vetores autorregressivos (VAR).

A metodologia de VAR possibilita uma estimação conjunta dos parâmetros, em que as inter-relações são consideradas, e o comportamento dinâmico dos dados é capturado, fornecendo uma estrutura entre as variáveis de entrada e de saída do sistema. Dessa maneira, torna-se possível aumentar a acurácia das estimativas do modelo, utilizando-se informações adicionais fornecidas pelas inter-relações.

Define-se VAR como um sistema de equações lineares dinâmicas, em que cada variável é escrita em função de um erro não-correlacionado, e todas as variáveis que pertencem ao sistema possuem o mesmo número de defasagens. Esses modelos de previsão multivariados são bastante discutidos pelos autores Lütkepohl (1991), Maddala (1992), Charemza e Deadman (1997), Reinsel (1993), Hamilton (1994) e Souza, Samohyl e Malavé (2004b).

Um vetor autorregressivo é um sistema de equações lineares dinâmicas, em que cada variável endógena é escrita como uma combinação linear de suas defasagens e também defasagens das variáveis endógenas de outras equações. Quando todas as variáveis, que pertencem ao sistema, possuem o mesmo número de defasagens, representadas por p , estas determinam a ordem do modelo, que, genericamente, é representado por VAR(p), conforme a equação (30).

$$Z_t = v + \varphi_1 Z_{t-1} + \dots + \varphi_p Z_{t-p} + \varepsilon_t \quad (30)$$

onde Z_t é um vetor aleatório, φ_i é a matriz dos coeficientes, v é o vetor dos interceptos, o qual permite que a média do processo seja diferente de zero e ε_t é o vetor ruído branco. Este processo ruído branco garante que os erros sejam independentes e identicamente distribuídos, indicando a ausência de qualquer correlação serial entre os erros, quer dizer, que os resíduos serão homoscedásticos.

Considerando-se o exemplo de um sistema composto por duas variáveis (x e y) com uma defasagem, torna-se, assim, mais compreensível o processo VAR(1), descrito a seguir:

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \quad (31)$$

Essa forma matricial pode ser reescrita da seguinte forma:

$$x_t = a_1 x_{t-1} + b_1 y_{t-1} + \varepsilon_t \quad (32)$$

$$y_t = c_1 x_{t-1} + d_1 y_{t-1} + \varepsilon_t \quad (33)$$

Observa-se que ambas as variáveis defasadas, x e y , aparecem, simultaneamente, em cada equação. Logo o modelo VAR(1) capta o efeito dinâmico nas inter-relações das variáveis, possuindo um caráter simultâneo.

Como pode ser observado até agora, na estimação das equações do modelo VAR, o sistema apresenta uma estrutura fixa, com as mesmas variáveis em todas as equações e com o mesmo número de defasagens, sendo conhecido como “VAR puro”.

Segundo Enders (1995) se algumas das equações possuírem regressores não incluídos nas outras, ou seja, variáveis diferentes do lado direito de cada equação, ou, quando as variáveis possuírem defasagens diferentes, o modelo é denominado de *quase VAR*, que vem do termo original em inglês “near VAR”.

Em modelos *quase VAR*, se há correlação serial nos erros contemporâneos das diferentes equações do sistema, é necessário a utilização do estimador de regressões aparentemente não correlacionadas, no qual chamaremos de estimador SUR, que vem do termo inglês *Seegmingly Unrelated Regression* (SUR).

O método de estimação SUR possibilita que cada equação tenha sua própria forma funcional, levando em consideração apenas a correlação existente entre os resíduos das equações. É um modelo de componentes comuns, bem difundido na literatura de métodos econométricos (SALLES, 2006).

Existem vários métodos para estimar um modelo através do SUR, descreve-se a seguir o método de Mínimos Quadrados Ordinários de Três Estágios (MQO3E), empregado nessa pesquisa.

- Método MQO3E

Segundo Kennedy (1985 apud SANVICENTE, 1998) pode-se considerar, inicialmente, cinco métodos de estimação de cada equação de um sistema separadamente: (1) mínimos quadrados ordinários, (2) mínimos quadrados indiretos, (3) variáveis instrumentais, (4) mínimos quadrados de dois estágios e (5) máxima verossimilhança, informação limitada. Como desvantagem, esses métodos consideram apenas as restrições existentes em cada equação; conseqüentemente, não fazem uso de toda informação disponível e, tendem, portanto, a ser estimadores menos eficientes.

Por outro lado, o método de mínimos quadrados ordinários de três estágios e o método de máxima verossimilhança, informação completa (FIML), são métodos que usam

explicitamente o fato de que se está lidando com um sistema, e, são considerados, assim, estimadores mais eficientes, pois realiza uma estimação conjunta por meio de relações simultâneas.

Na estimação do modelo dinâmico multivariado, considera-se uma estrutura de defasagens em todas as variáveis. O método MQO3E segue, também, essa metodologia do geral para o específico. As variáveis que, inicialmente constituíam o modelo são defasadas com uma estrutura de três defasagens, obtém-se não somente o nível de ajuste para a variável identificada, mas o nível de ajuste que deve ser realizado para o conjunto de variáveis. Aquelas defasagens que não foram estatisticamente significativas mediante teste de hipóteses são retiradas gradualmente, resultando em um modelo simples e bem especificado, com uma equação para cada variável. Esse método é bem discutido por autores, como Hill, Giffiths e Judge (1999), Gujarati (2000), Reinsel (1993), Kmenta (1988), Charemza e Deadman (1997) e Bayer e Souza (2007). Dessa forma, é possível a estimação dos parâmetros das equações simultâneas, possibilitando diferentes restrições na estrutura das equações e ganhando-se na eficiência das estimativas.

O procedimento de estimação do método MQO3E considera todas as equações em conjunto, e não separadamente. Os três estágios do procedimento são os seguintes:

- a) estimar as equações separadamente utilizando mínimos quadrados;
- b) utilizar os resíduos de mínimos quadrados, do passo anterior, para estimar as variâncias e covariância dos erros;
- c) utilizar as estimativas das variâncias dos erros para estimar as equações conjuntamente.

Segundo Hill, Griffiths e Judge (1999), o emprego do método de três estágios usa a matriz dos momentos dos erros para estimar todos os coeficientes envolvendo todo o sistema simultaneamente. Sendo assim, o método possui todas as informações, de modo que, se a estrutura de covariância dos erros não for zero, a estimação dos coeficientes de qualquer equação ganha em eficiência tão logo outras equações sejam identificadas. Além disso, o método leva em consideração as restrições impostas nas diferentes estruturas de equações.

Tendo estabelecido o modelo que melhor represente as variáveis em estudo e que satisfaça as hipóteses e a análise dos resíduos, ele será utilizado para realizar previsões em curto prazo, procurando-se sempre o menor valor para o MAPE, garantindo, assim, os melhores valores previstos para as variáveis. Essas previsões poderão ser utilizadas pela

equipe gerencial como suporte no planejamento das decisões e como medidas preventivas para o bom andamento do processo produtivo.

O sistema multivariado deve apresentar um processo ruído branco de forma que os resíduos sejam independentes e identicamente distribuídos, garantindo a ausência de autocorrelação nos resíduos, que os resíduos sejam homocedásticos, apresentando uma variância constante e seguindo a distribuição normal. Então, realizada a análise dos resíduos, todas estas hipóteses são satisfeitas a um bom nível de significância, e opta-se por utilizar o modelo encontrado para fazer as previsões (HILL, GIFFITHS e JUDGE, 1999; GUJARATI, 2000).

Na etapa de definir o modelo adequado, é necessário levar em conta não só a significância dos parâmetros, mas também certa estrutura lógica do modelo. Em síntese, na escolha de um modelo de regressão, não é necessário apenas encontrar um ajuste de parâmetros adequado, mas fundamentalmente verificar se os coeficientes estimados são coerentes.

2.3 Testes de diagnósticos

No processo de modelagem de séries temporais, têm-se diversos pressupostos, já mencionados, que devem ser satisfeitos e, portanto, testados.

Na seção 2.3.1, apresentam-se os testes de raiz unitária para analisar a estacionariedade; na seção 2.3.2, os testes de normalidade; na seção 2.3.3 encontra-se a estatística *Ljung-Box* para detectar autocorrelação; na seção 2.3.4, os testes para detectar a heterocedasticidade condicional residual; na seção 2.3.5, os critérios penalizadores para escolha do melhor modelo e, na seção 2.3.6, os critérios para avaliar a qualidade da previsão.

2.3.1 Testes de raiz unitária

Segundo Arêdes e Pereira (2008), o teste mais simples para analisar a estacionariedade em séries temporais é dado pela obtenção dos coeficientes de autocorrelação e autocorrelação

parcial, a partir dos quais são construídos os correlogramas: FAC e FACP. O correlograma traça um intervalo de confiança para os coeficientes no qual as estatísticas da FAC e da FACP devem variar. Esse intervalo é definido, frequentemente, por:

$$IC(95\%) = \pm 1,96\sqrt{n} \quad (34)$$

onde $IC(95\%)$ é o intervalo de confiança com 95% de significância estatística e n o tamanho da amostra.

Gurajati (2000) afirma que coeficientes de autocorrelação e autocorrelação parcial fora do intervalo de confiança, exceto para a defasagem um, são estatisticamente diferentes de zero, sugerindo não-estacionariedade na série.

Entretanto, os testes de raiz unitária aplicados antes da modelagem têm a finalidade de garantir estatisticamente os pressupostos sobre a estacionariedade dos modelos.

Nesta seção, apresentam-se os testes propostos por Phillips-Perron e KPSS.

- *Teste Phillips-Perron*

Segundo Morettin (2006), o teste Phillips-Perron (PP) detecta o problema de autocorrelação nos resíduos a partir de correções semi-paramétricas, considerando as variáveis defasadas dependentes e leva em conta a heterocedasticidade condicional.

O teste PP faz uma correção não-paramétrica ao teste de Dickey e Fuller (DF), permitindo que seja consistente mesmo que haja variáveis defasadas dependentes e autocorrelação nos resíduos, tornando-se desnecessária a especificação de um modelo com ordem autorregressiva suficiente para retirar a autocorrelação existente (BUENO, 2008).

A estatística de PP é dada pela expressão (35)

$$\hat{z}_{t,\mu} = \hat{\tau}_{\mu} \left(\frac{\hat{\sigma}}{\hat{\hat{\nu}}} \right) - \frac{1}{2} \left(\frac{\hat{\nu}^2 - \hat{\sigma}^2}{\hat{\nu} \sqrt{T^{-2} \sum_{t=1}^T y_{t-1}^2}} \right) \quad (35)$$

onde $\hat{\nu}^2$ representa a variância de longo prazo, em que estão incluídas todas as autocovariâncias do processo u_t ; $\hat{\tau}_{\mu}$ é a estatística DF e $\hat{\sigma}^2$ é a variância populacional da regressão.

As hipóteses para o teste são descritas a seguir:

H_0 : A série possui raiz unitária, então, é dita não-estacionária;

H_1 : A série não possui raiz unitária, concluindo-se que a série é estacionária.

Regra de decisão: se a estatística calculada é menor que os valores críticos, rejeita-se H_0 , conclui-se que a série não possui raiz unitária, portanto, ela é estacionária.

- *Teste Kwiatkowski, Phillips, Schmidt e Shin*

Um dos problemas do teste de raiz unitária, desenvolvido por Dickey e Fuller, é seu baixo poder. Por essa razão, outros testes vêm sendo desenvolvidos, entre eles, o teste KPSS dos autores Kwiatkowski, Phillips, Schmidt e Shin, idealizado em 1992. O argumento dos autores quanto ao uso desse teste é que se pode distinguir a raiz unitária de séries cujos dados não são suficientemente conclusivos. A idéia é testar a variância de passeio aleatório. Se essa variância for nula, então o processo é estacionário (BUENO, 2008).

A estatística para o teste KPSS é dada pela expressão (36)

$$KPSS = \sum_{t=1}^T \frac{S_t^2}{T^2 \hat{\sigma}^2} \quad (36)$$

onde $\hat{\sigma}^2$ representa a variância de longo prazo, definida como no teste de PP.

Nessa estrutura, as hipóteses testadas são:

H_0 : A série é estacionária ($\sigma^2 = 0$);

H_1 : A série é não-estacionária ($\sigma^2 > 0$).

Regra de decisão: se a estatística calculada é menor que os valores críticos, não se rejeita H_0 , conclui-se que a série é estacionária.

Analisando-se as hipóteses nulas dos testes PP e KPSS, verifica-se que as mesmas apresentam decisões contrárias. Desse modo, é interessante aplicar os testes PP e KPSS, conjuntamente, para confirmar estatisticamente a estacionariedade ou não da série, obtendo-se, assim, um resultado mais preciso (BAILLIE, CHUNG e TIESLAU, 1996).

Se a série de interesse for não-estacionária, os modelos estimados independentemente da metodologia aplicada serão explosivos. Para evitar esse risco, adota-se a especificação na forma de primeira diferença ou de segunda diferença da série para torná-la estacionária (BEM, 1998).

Salienta-se que os testes para identificar a estacionariedade são aplicados nos dados originais, como requisito para uma determinação mais precisa da ordem de integração da série quando a metodologia dos modelos ARIMA é empregada na série sob estudo.

2.3.2 Testes de normalidade

O pressuposto da normalidade precisa ser atendido com o objetivo de oferecer confiabilidade aos testes estatísticos, uma vez que, em geral, eles se baseiam na distribuição normal do termo residual. Sua violação afeta as qualidades de não-tendenciosidade e variância mínima do método de mínimos quadrados ordinários.

Vários testes estatísticos têm sido usados para avaliar a normalidade dos resíduos do modelo estimado, dentre os quais se destacam: Jarque-Bera e gráfico Q-Q plot.

- Jarque-Bera

É uma estatística que tem a vantagem de ser simples na sua operacionalização, em termos de compreensão intuitiva e depende de dois parâmetros, assimetria e curtose, que podem ser calculados dos dados históricos.

A estatística Jarque-Bera (JB) possui um poder de teste elevado, não erra na presença de normalidade. Há uma baixa probabilidade de indicar a normalidade na presença de não-normalidade das distribuições, ou seja, pode-se cometer o Erro Tipo II, aceitando-se a hipótese de normalidade quando a hipótese alternativa de não-normalidade é a verdadeira. É mensurada pela equação (37), a qual segue uma distribuição Qui-quadrado com dois graus de liberdade:

$$JB = N \left[\frac{A_3^2}{6} + \frac{(K_4 - 3)^2}{24} \right] \sim \chi_2^2(\alpha = 5\%) \geq 5,99 \quad (37)$$

onde A_3 é coeficiente de assimetria; K_4 é o coeficiente de curtose excessivo e N é o tamanho da amostra.

As hipóteses a serem testadas são:

H_0 : Assimetria = zero e curtose = três, então a série é dita normal;

H_1 : Assimetria \neq zero e curtose \neq três, então a série é dita não-normal.

Regra de decisão: o teste baseia-se nos coeficientes de assimetria e curtose. Se *p-valor* for maior que o nível de significância de 5% cujo valor crítico é 5,99, aceita-se a hipótese nula de normalidade.

- *Q-Q plot*

O Q-Q plot representa um dos métodos gráficos mais utilizados na verificação da normalidade de séries temporais. Apesar de ser simples, é uma ferramenta poderosa na comparação de duas distribuições e consiste na comparação gráfica dos quantis teóricos da distribuição normal com os quantis dos dados amostrais.

Regra de decisão: se as duas distribuições são aproximadamente iguais, ou seja, uma distribuição sobrepõe a outra, pode-se dizer que a série segue uma distribuição normal. Caso contrário, as duas distribuições diferem ao longo de alguma dimensão, e, portanto, a distribuição não é normal.

2.3.3 Testes para autocorrelação

A autocorrelação indica se existe uma dependência temporal entre os valores sucessivos dos resíduos, ou seja, se os resíduos são correlacionados entre si. A presença de autocorrelação significa que as estimativas através do método MQO dos parâmetros não são tão eficientes, ou seja, não apresentam variância mínima. Ela provoca um viés nos erros-padrão, conduzindo a testes e intervalos de confiança incorretos.

- *Estatística Ljung-Box*

Segundo Ehlers (2007), ao invés de verificar as autocorrelações residuais individualmente, pode-se testar se um grupo de autocorrelações é significativamente diferente de zero através das chamadas estatísticas Q .

Box e Pierce (1970) sugeriram um teste de hipótese para as autocorrelações dos resíduos estimados, apesar de não detectar quebras específicas no comportamento de ruído branco, pode indicar se esses valores são muito altos. Uma modificação desse teste foi proposta por Ljung e Box (1978), na qual a estatística de teste é denotada pela expressão (38)

$$Q(K) = n(n+2) \sum_{k=1}^h \frac{\hat{r}_k^2}{(n-k)} \quad (38)$$

em que: n representa o número de observações; \hat{r}_k^2 é a autocorrelação no *lag* k e h representa o número de *lags* que estão sendo testados, seguindo uma distribuição Qui-quadrado (χ^2) com p graus de liberdade. Essa estatística é reportada nos correlogramas pelos *softwares* econométricos.

As hipóteses a serem testadas são:

H_0 : Os dados são ruído branco (RB);

H_1 : Os dados não são ruído branco (RB).

Regra de decisão: se a estatística calculada excede o valor crítico da tabela de χ^2 para determinado nível de significância, podemos rejeitar a hipótese nula de que todos os (\hat{r}_k) são iguais a zero. Isso implica em aceitar a hipótese alternativa de que pelo menos uma autocorrelação (\hat{r}_k) é estatisticamente diferente de zero. Nesse caso, o resíduo não é ruído branco. No caso da aceitação da hipótese alternativa, rejeita-se o modelo, pois há uma estrutura de correlação serial significativa nos resíduos (GUJARATI, 2000; SÁFADI, 2004).

Na prática, utiliza-se entre as 15 e 20 primeiras autocorrelações (\hat{r}_k) para o cálculo da estatística $Q(K)$ de Ljung-Box. Em amostras grandes, tanto a estatística Q de Box e Pierce como a $Q(K)$ de Ljung-Box seguem a distribuição χ^2 , mas a estatística $Q(K)$ possui as melhores propriedades em amostras pequenas do que a estatística Q .

A estatística *Ljung-Box*, em conjunto com os correlogramas FAC e FACP, foi aplicada nas seguintes situações: inicialmente, nos dados originais para analisar, em primeira instância, a estacionariedade da série e a presença de autocorrelação; em seguida, nos resíduos do modelo ARIMA com o propósito de verificar se os resíduos do modelo estimado formam um ruído branco; e, posteriormente, nos resíduos quadráticos desse modelo para identificar se existe autocorrelação significativa. Se houver, é necessário aplicar testes para verificar se existe heterocedasticidade na série quadrática. Se a heterocedasticidade for confirmada, a série é modelada através dos modelos ARCH.

2.3.4 Testes para heterocedasticidade condicional

Apresentam-se a seguir os testes para verificar a presença de heterocedasticidade.

Conforme dito anteriormente, na seção 2.2.2, até pouco tempo atrás, a heterocedasticidade era considerada um problema muito comum em dados seccionais ou *cross section*³, embora também esteja presente em algumas séries temporais, mas estudos de Engle (1982) e por Bollerslev (1986) provaram que essa anomalia também pode ser modelada.

Na prática, é desejável que a variância dos resíduos do modelo estimado seja constante, ou seja, $Var(\mu_i) = \sigma^2$. Caso isso ocorra, diz-se que os resíduos do modelo estimado são homocedásticos. A violação desse pressuposto revela que a variância irá diferir ao longo do tempo, isto é, $Var(\mu_i) = \sigma_i^2$, e os resíduos serão heterocedásticos.

Na presença de heterocedasticidade, o método de mínimos quadrados ordinários não produz estimativas eficientes para seus parâmetros nem mesma uma variância mínima. A não minimização da variância conduz a erros-padrão viesados e a incorreção dos testes *t-student* e F-Snedecor, além dos intervalos de confiança. Isso justifica a necessidade de aplicar modelos não-lineares da classe ARCH em séries que apresentam heterocedasticidade (SILVA, 2000).

Existem diversos testes estatísticos para detectar a presença de heterocedasticidade nos resíduos quadráticos do modelo estimado. Apresentam-se a seguir, os testes de White aplicado no caso multivariado, o ARCH-LM empregado nos casos univariados e, posteriormente, a estatística *F*, presente em ambos os testes.

- *Teste de White*

Os testes de heterocedasticidade requerem, em geral, que se especifique de certa forma a possível natureza da heterocedasticidade. O teste de White, entretanto, pode ser conduzido sem que se faça nenhuma suposição sobre a natureza da heterocedasticidade (SOARES e CASTELAR, 2003).

Considerando o modelo, exposto a seguir:

$$Y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + e_i \quad (39)$$

o teste White é conduzido tomando nR^2 da regressão auxiliar de e_i^2 (resíduos quadráticos) como variável dependente em função de um termo constante, das variáveis explicativas x_2 e x_3 , de seus valores ao quadrado e dos produtos cruzados dos regressores. Assim,

$$e_i^2 = \alpha_1 + \alpha_2 x_{2i} + \alpha_3 x_{3i} + \alpha_4 x_{2i}^2 + \alpha_5 x_{3i}^2 + \alpha_6 x_{2i} x_{3i} \quad (40)$$

³ As variáveis são observações e medidas em um dado intervalo de tempo, sob um conjunto de objetos.

e, então, nR^2 dessa regressão segue uma distribuição Qui-quadrado com cinco graus de liberdade ($nR^2 \sim \chi^2_{5gl}$).

Acontece que, como na regressão auxiliar foram incluídos os quadrados das variáveis explicativas do modelo original e seus produtos cruzados, em casos de amostra pequena e/ou grande número de variáveis explicativas, perdem-se muitos graus de liberdade. Por esse motivo, o EViews possibilita realizar o teste de White, incluindo ou não os produtos cruzados na regressão auxiliar.

A estatística nR^2 é a estatística do teste de White e se distribui assintoticamente com uma distribuição χ^2 com $(p-1)$ graus de liberdade, em que p é o número de coeficientes da regressão auxiliar.

As hipóteses testadas são as seguintes:

H_0 : Homocedasticidade dos resíduos;

H_1 : Heterocedasticidade dos resíduos.

Regra de decisão: se *p-valor* for maior que o nível de significância adotado, a hipótese de homocedasticidade não será violada (BEM, 1998).

No caso específico desta pesquisa, o teste de White foi aplicado na presença e na ausência dos produtos cruzados nos resíduos do modelo estimado pelo método SUR.

- ARCH-LM

O teste ARCH-LM, do tipo multiplicador de Lagrange, proposto por Engle (1982), testa a hipótese de que os erros resultantes da estimação do modelo são dependentes dos erros $r_{t-1}^2, r_{t-2}^2, \dots, r_{t-m}^2$, e, baseia-se na estimação da regressão auxiliar (BEM, 1998).

Silva, Sáfiadi e Castro Júnior (2005) afirmam que, se os coeficientes de autocorrelações estimados revelam um padrão previsível para a média condicional da série analisada e se as autocorrelações estimadas para os resíduos quadráticos também exibem sinais de previsibilidade, há uma evidência de efeitos ARCH.

O teste ARCH-LM utiliza a seguinte expressão:

$$r_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \dots + \alpha_m r_{t-m}^2 + u_t \quad (41)$$

onde $t = m+1, \dots, T$. A estatística do teste $T \times R^2$ tem distribuição assintótica $\chi^2(m)$ sob H_0 .

Nesse caso, R^2 é o quadrado do coeficiente de correlação múltipla da regressão (41).

Para comprovar a presença de efeitos ARCH, testam-se as seguintes hipóteses:

H_0 : A série não possui efeitos ARCH;

H_1 : A série possui efeitos ARCH.

Regra de decisão: se p -valor é menor que o nível de significância adotado, rejeita-se a hipótese nula, portanto, existe heterocedasticidade condicional.

Assim, o teste ARCH-LM foi aplicado nos resíduos do modelo estimado em dois momentos. No primeiro momento, aplica-se o teste à série de resíduos quadráticos do modelo ARIMA. Se esta série possui heterocedasticidade condicional, necessita-se estimar um modelo por meio dos modelos não-lineares da classe ARCH. Em um segundo momento, aplica-se o teste na série de resíduos quadráticos do modelo ARCH e nos resíduos padronizados oriundos do modelo ARCH. Após a escolha do melhor modelo matemático por meio da modelagem ARCH, é necessário aceitar a hipótese nula do teste de hipótese ARCH-LM, indicando que a série residual do modelo estimado é puramente aleatória, sem dependência na média e na variância.

- Estatística F

De acordo com Morettin (2008), um teste assintoticamente equivalente, que pode ter propriedades melhores para amostras pequenas, consiste em utilizar a estatística

$$F = \frac{(SQR_0 - SQR_1) / m}{SQR_1 / (T - 2m - 1)} \sim F(m, T - 2m - 1) \quad (42)$$

na qual $SQR_0 = \sum_{t=m+1}^T (r_t^2 - \bar{r})^2$ e $SQR_1 = \sum_{t=m+1}^T \hat{u}_t^2$, com \bar{r} a média amostral dos r_t^2 e \hat{u}_t os resíduos de MQO da regressão (42). Se o valor de F for significativo, diz-se que há heterocedasticidade condicional na série.

2.3.5 Critérios penalizadores

Aplicam-se critérios penalizadores com o objetivo de escolher o melhor modelo matemático, levando em conta a parcimônia do modelo.

Em muitas aplicações de pesquisa, vários modelos podem ser julgados adequados em termos do comportamento dos resíduos. Uma forma de distinguir qual é o modelo mais adequado entre os modelos candidatos é utilizar os critérios penalizadores, que levam em consideração não apenas a qualidade do ajuste, mas também penalizam modelos que possuem um elevado número de parâmetros, ou seja, modelos parcimoniosos são os melhores (EHLERS, 2007).

Os critérios de informação AIC e SBC são calculados com base na variância estimada, e, são também denominados de critérios penalizadores, pois eles levam em consideração o número de parâmetros estimados. Algebricamente, estes critérios são definidos como:

$$\text{AIC (Akaike Info Criterion): } AIC = -2(l/T) + 2(k/T) \quad (43)$$

$$\text{SBC (Schwarz Criterion): } SC = -2(l/T) + k \log(T)/T \quad (44)$$

onde T é o tamanho da amostra; l é o valor da função de log verossimilhança e k é o número de parâmetros estimados.

Regra de decisão: escolher o modelo que minimiza os critérios penalizadores.

É importante considerar que os critérios AIC e SBC devem ser usados como procedimentos complementares e não como alternativos àqueles baseados na FAC e FACP.

Maiores informações sobre esse assunto são encontradas em Morettin (2006) e Priestley (1981).

2.3.6 Critérios de avaliação de previsão

Como o principal objetivo do modelo matemático é a previsão, conclui-se que as estimativas em curto prazo devem se aproximar da realidade. No entanto, para avaliar o desempenho da previsão dos modelos candidatos com a finalidade de obter valores previstos mais próximos possíveis aos observados, utilizam-se os indicadores: erro médio percentual absoluto (MAPE) e U de Theil.

- MAPE

O motivo da utilização desses critérios nesta pesquisa é avaliar a capacidade preditiva dos modelos perante os dados originais.

O MAPE indica o valor médio do erro percentual absoluto das previsões sobre um conjunto de dados. É calculado a partir das previsões um passo à frente, geradas pelo modelo matemático estimado (RUSSO, 2002). Algebricamente, é apresentado como:

$$MAPE(\%) = \frac{\sum \left(\frac{y_t - \bar{y}_t}{y_t} \right)}{n} \times 100 \quad (45)$$

onde: n é o número de observações; y_t é o valor observado no instante t e \bar{y}_t é o valor previsto no instante t .

- U de Theil

A estatística U de Theil mede o quanto o método de previsão é melhor se comparada à previsão ingênua, atuando como parâmetro de avaliação. Quanto mais próximo de zero, melhor o desempenho do modelo e, conseqüentemente, os erros de previsão são menores. Por outro lado, valores maiores que um é uma indicação de que, globalmente, os erros de previsão tendem a ser grandes em relação aos erros de um passeio aleatório (SOUZA, SAMOBYL e MIRANDA, 2008).

A estatística pode ser expressa tal como encontra-se em (46)

$$U = \sqrt{\frac{\sum_{t=1}^{n-1} \left(\frac{\bar{y}_{t+1} - y_{t+1}}{y_t} \right)^2}{\sum_{t=1}^{n-1} \left(\frac{y_{t+1} - y_t}{y_t} \right)^2}} \quad (46)$$

onde:

n é o número de observações;

y_t é o valor observado no instante t ;

y_{t+1} é o valor observado no instante $t+1$;

\bar{y}_t é o valor previsto para o instante t .

Essa estatística mede o quanto os resultados estão melhores que uma previsão ingênua ou trivial. Através desse coeficiente, pode-se analisar a qualidade de uma previsão da seguinte maneira:

- quando $U > 1$, o erro do modelo é maior que o erro ingênuo,
- quando $U < 1$, o erro do modelo é menor que o erro ingênuo (boa previsão).

Assim, um U de Theil menor que um já indica uma previsão melhor que a previsão trivial. Portanto, quanto mais próximo de zero for esse coeficiente melhor a previsão.

2.4 Comentários gerais do capítulo

Inicialmente, neste capítulo, foram apresentadas as técnicas de gráficos de controle, os quais serão as ferramentas utilizadas para classificar o processo quanto a sua estabilidade. Por meio dos gráficos \bar{x} e \overline{MR} , é mostrado se um ajuste deve ou não ser feito, e, em caso afirmativo, onde se deve interferir no processo.

Em seguida, apresentou-se uma descrição teórica sobre os modelos de previsão empregados. Os modelos ARIMA servem para modelar a média condicional, os modelos ARCH modelam a volatilidade, já os modelos mistos ARIMA-ARCH servem para modelar a média e a variância condicional. Esses modelos de previsão serão empregados em séries univariadas. E, por fim, descreve-se o modelo de regressão dinâmica, que será aplicado no caso multivariado, o qual realiza uma estimação conjunta considerando as inter-relações e o comportamento dinâmico entre as variáveis.

Vale lembrar que os modelos de previsão precisam atender certas pressuposições para não resultar em modelos ineficientes que possam comprometer as conclusões tanto nas previsões como nos gráficos de controle.

No próximo capítulo, descreve-se a metodologia empregada nas três pesquisas ex-post-facto analisadas.

3 METODOLOGIA

Este capítulo apresenta a metodologia utilizada na pesquisa, englobando: o método da pesquisa, a coleta de dados, as etapas metodológicas e os programas computacionais utilizados para desempenhar da melhor forma o seu desenvolvimento. As técnicas escolhidas para o desenvolvimento deste trabalho são de fundamental importância, pois servirão de base para alcançar os objetivos propostos.

O tipo de pesquisa caracterizado neste trabalho integra pesquisa bibliográfica e pesquisa ex-post-facto, necessárias para validar a metodologia proposta. Além disso, também é uma pesquisa quantitativa. A pesquisa quantitativa “procura quantificar os dados e aplica alguma forma de análise estatística” (MALHOTRA, p. 153, 2001).

O detalhamento da metodologia engloba os três estudos de caso, nos quais diferentes métodos de previsão são empregados em diferentes processos produtivos.

3.1 Etapas metodológicas

As etapas metodológicas para o cumprimento dos objetivos perfazem os seguintes passos:

- 1) Aplicação de estatística descritiva para verificar o comportamento das variáveis em estudo.
- 2) Encontrar um modelo matemático para testar o efeito da autocorrelação serial de modo a possibilitar a aplicação dos gráficos de controle:
 - 2a) Pode ser um modelo matemático de regressão univariada;
 - 2b) Pode ser um modelo matemático de regressão multivariada, a qual pode ser estática ou dinâmica.
- 3) Eliminada a autocorrelação das variáveis em estudo, passa-se ao exame dos resíduos para verificar a presença da volatilidade nos mesmos, por meio dos seguintes testes:
 - Teste de White aplicado no caso multivariado;

- Teste ARCH-LM empregado no caso univariado.
- 4) Se existir heterocedasticidade nos dados, passa-se à estimação dos modelos ARCH, utilizando os resíduos quadráticos que são oriundos do modelo matemático que foi estimado no passo 2a ou 2b.
- 5) Com a série livre de autocorrelação, tanto na série de resíduos como na série de resíduos quadráticos, passa-se para a elaboração dos gráficos de controle, utilizando os resíduos do modelo matemático estimado:
- 5a) Se os resíduos são ruído branco e não existe volatilidade, traçam-se os gráficos de controle \bar{x} e \overline{MR} para medidas individuais e amplitude móvel nos dados originais e nos resíduos, e, avalia-se se o processo está sob controle;
- 5b) Se os resíduos são ruído branco, mas a volatilidade está presente, encontra-se o modelo misto ARIMA-ARCH, e, traçam-se os seguintes gráficos de controle:
- \bar{x} e \overline{MR} nos dados originais e nos dados sem o efeito de autocorrelação (resíduos) para mostrar o que acontece no processo;
 - \bar{x} e \overline{MR} nos resíduos do modelo misto;
 - \bar{x} e \overline{MR} na volatilidade.

Assim, pretende-se mostrar que processos produtivos têm volatilidade, e, que essa volatilidade influencia para que o processo sofra alterações.

- 6) Na busca de encontrar o melhor modelo matemático que represente corretamente o processo gerador de cada série, estimaram-se diversos modelos concorrentes, e, escolhe-se o modelo que obteve o menor valor para AIC ou SBC.
- 7) Como forma de avaliar se as previsões geradas por meio dos modelos matemáticos estimados por meio do passo 2a ou 2b são satisfatórias, utilizam-se os critérios de avaliação: MAPE e U de Theil.
- 8) O nível de significância adotado para os testes de hipóteses é de 5%.
- 9) Os softwares utilizados são: Statistica versão 7.0, EViews versão 6.0 e planilha eletrônica.
- 10) Aplicação nas empresas A, B e C. Como não foi possível obter autorização para divulgar o nome verdadeiro da empresa por se tratar da qualidade de seus produtos, escolheu-se um nome fictício para nominar as três empresas: Empresa A, Empresa B e Empresa C.

Empresa A: atua no setor de fundição e usinagem,

Empresa B: atua no setor de alimentos,

Empresa C: atua no setor de bebidas.

Apresenta-se, a seguir, o roteiro metodológico empregado nas empresas sob estudo.

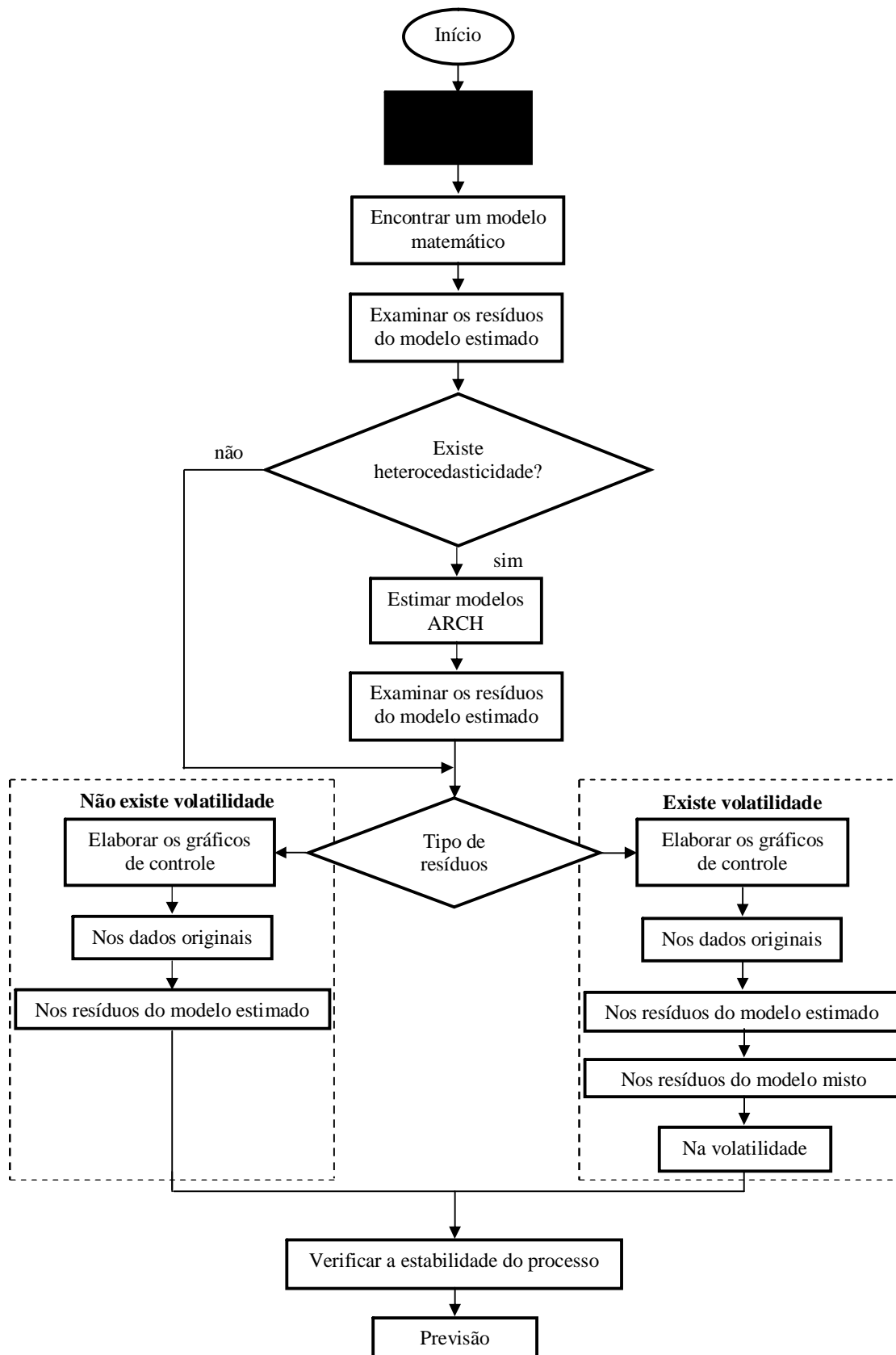


Figura 5 – Roteiro metodológico

A Figura 5 mostra, graficamente, as etapas metodológicas empregadas para analisar a estabilidade dos processos produtivos das respectivas empresas por meio dos gráficos de controle aplicado aos resíduos oriundos de modelagem linear e não-linear e fazer previsões utilizando o modelo matemático representativo para cada uma das variáveis analisadas.

Com os dados coletados, foi encontrado um modelo matemático através dos modelos de regressão univariado ou multivariado. Existindo heterocedasticidade nos resíduos desse modelo, estimou-se um novo modelo matemático por meio dos modelos ARCH. Com os resíduos obtidos, realizaram-se as previsões e, após, aplicou-se a técnica de gráficos de controle para medidas individuais e amplitude móvel nos dados originais, nos resíduos do modelo ARIMA, nos resíduos do modelo misto e na volatilidade.

3.2 Comentários gerais do capítulo

Para que os objetivos desta pesquisa pudessem ser alcançados, uma revisão da literatura foi realizada, com vistas a demonstrar a importância CEP. Além disso, foi necessário estudar as teorias de gráficos de controle, modelos ARIMA, modelos ARCH e modelo de regressão dinâmica multivariada.

Além de apresentar as etapas metodológicas de forma descritiva, elaborou-se um roteiro metodológico representado na forma de fluxograma, elencando os passos para a estimação dos modelos lineares, ARIMA e modelo de regressão dinâmica multivariada, e, dos modelos não-lineares, ARCH, para aplicação dos gráficos de controle e para a realização das previsões.

Após o estudo das teorias utilizadas neste trabalho, utilizou-se a metodologia proposta para os três estudos de caso relacionados a processos produtivos industriais.

4 ANÁLISE DOS RESULTADOS

Este capítulo apresenta a análise dos resultados mediante a aplicação da metodologia proposta em conjuntos de dados reais referentes a processos produtivos industriais de três empresas atuantes em segmentos diferentes.

De acordo com a metodologia detalhada no capítulo 3, serão utilizadas técnicas de séries temporais, incluindo modelos lineares e não-lineares com o fim de verificar a estabilidade do processo por meio de gráficos de controle aplicado aos resíduos do modelo matemático estimado para as variáveis em estudo.

Este capítulo está dividido em três seções: primeiramente, abordará a pesquisa ex-post-facto da empresa A, após, apresentará a pesquisa ex-post-facto da empresa B e, por fim, apresentará a pesquisa ex-post-facto da empresa C.

4.1 Empresa A

É uma empresa especializada na fundição de ferro fundido cinzento e nodular, e usinagem. Possui cerca de 800 colaboradores distribuídos entre os processos de produção de peças (fundição) e usinagem, operando em três turnos, sete dias da semana.

Atualmente, é uma das empresas mais consolidadas do segmento, contemplando uma capacidade produtiva de 2.200 toneladas/mês, fornecendo uma diversidade de peças, pequenas, médias e grandes, dentro das mais rigorosas especificações, com peso até 100 kg.

Os setores que a empresa abrange são os seguintes:

- linha automotiva, por exemplo: carro, ônibus, caminhão leve e pesado;
- linha agrícola, por exemplo: trator, plantadeira, colhedeira e implementos agrícolas;
- fora de estrada, por exemplo: patrola, esteira, retro escavadeira;
- outras linhas, por exemplo: elevadores, jardinagem e ginástica.

Possui as seguintes certificações: SA 8000:2001⁴, sendo a primeira empresa de fundição do mundo a ser certificada nessa norma, ISO 14001:2004⁵, OHSAS 18001:2007⁶ e ISO/TS 16949:2002⁷, um diferencial que a coloca dentro do mercado como uma empresa possuidora de Qualidade Assegurada.

Para suprir as necessidades do mercado, a empresa conta com oito máquinas semi-automáticas e uma automática. Com a recente crise econômica mundial, a produção de peças, bem como a solicitação de pedidos pelos clientes, diminuiu, ocasionando queda na produção, redução de funcionários, entre outras conseqüências na produção e administração da empresa.

Atualmente, a produção está em equilíbrio, sendo executada com a operação de uma máquina automática que substitui a capacidade produtiva de três máquinas semi-automáticas e com a operação de três máquinas semi-automáticas, atuando em dois turnos.

4.1.1 Descrição do processo e das variáveis a serem analisadas

Existem vários tipos de processos de fundição. O processo de fundição utilizado na empresa é a moldagem em areia verde, também conhecido como areia sintética, sob o qual será realizado esta pesquisa ex-post-facto.

Basicamente, o processo de fundição pode ser resumido nas seguintes operações: fabricação do molde, elaboração do metal e vazamento do metal. Um molde é construído para cada peça a ser fundida e, subsequentemente, ele é rompido para que a peça fundida possa ser removida.

Define-se moldagem de areia verde como um processo de confecção de moldes utilizando mistura preparada com areia de retorno, areia base (nova), argila e os aditivos necessários como pó de carvão, mogul (amido de milho), entre outros, devidamente balanceados para atender às aplicações específicas.

Segundo Agostinho, Vilela e Button (2004), a areia de fundição deve ser previamente preparada através de homogeneização. A gravidade é usada para vazamento do metal líquido. Para produção em larga escala, são adotados processos automáticos ou semi-automáticos,

⁴ SA 8000:2001 – Sistema de Gestão de Responsabilidade Social.

⁵ ISO 14001:2001 – Sistema de Gestão Ambiental.

⁶ OHSAS 18001:2007 – Sistema de Gestão de Saúde e Segurança Ocupacional.

⁷ ISO/TS 16949:2002 – Norma de Sistema de Gestão da Qualidade voltada para o segmento automotivo.

utilizando máquinas de moldagem. As perdas de material do molde são pequenas, já que a areia pode ser recuperada.

Para controlar o processo de fundição de moldagem em areia verde, é preciso realizar ensaios de laboratório num período determinado, de hora em hora, a fim de mensurar as características da areia verde. As características são as seguintes:

- a) **teor de umidade:** depende da quantidade de argila ativa, da granulometria da areia. O teor de umidade ideal aproximado é de 3,2%, e, depende, principalmente, da argila ativa;
- b) **permeabilidade:** é a capacidade que o molde tem de permitir a passagem de gases. Quanto menor a permeabilidade mais resistência haverá no modelo do molde.
- c) **resistência de compressão a verde (rcv):** capacidade da areia em copiar o modelo (referente ao molde) para que no processo de desmoldagem do modelo, o molde não sofra danos;
- d) **compactabilidade:** capacidade do molde em realizar a compactação, copiando as diversas formas do modelo com precisão;
- e) **plasticidade:** deve ser suficientemente elevada para evitar quebra de bolos e de cantos do molde, erosão, e, a consequente inclusão de areia na peça fundida.

As vantagens do processo de fundição de moldagem em areia verde são:

- menor custo;
- facilidade de elaboração;
- matérias-primas disponíveis nacionalmente;
- não é prejudicial à saúde humana;
- não é tão prejudicial ao meio-ambiente, comparado a outros tipos de moldagem.

Apesar das vantagens, existem também as desvantagens, pois podendo ocorrer defeitos. Os tipos de defeitos e suas possíveis causas estão, respectivamente, relacionados a seguir:

- **Escama:** água em excesso, baixa resistência do choque metal-molde, presença de teor de carvão alto (fora de especificação), teor de umidade alta;
- **Molde quebrado ou mordido:** baixa rcv, ferramental não estar bem plano, o funcionário esqueceu-se de colocar o produto para desmoldagem, teor de umidade baixo, modelo não está bem elaborado, desalinhamento na máquina;

- **Molde explodido:** areia muito úmida e permeabilidade baixa, caixa de moldar gasta;
- **Rechupe:** temperatura de vazamento, permeabilidade baixa;
- **Inchamento:** compactação baixa da areia, granulometria da areia.

Analisando os defeitos que podem ocorrer na peça fundida juntamente com um engenheiro mecânico, verificou-se que muitos deles são causados pelas variações nas especificações das características da areia verde. Por este motivo, justifica-se fazer uma estimação multivariada somente das três variáveis consideradas geradoras de maiores instabilidades, entre elas: teor de umidade (%), permeabilidade (AFS) e resistência de compressão a verde (L/POL²). Essas variáveis estão intimamente relacionadas, e, possuem as seguintes especificações determinadas pela Engenharia de Produção e Engenharia Técnica: o teor de umidade máximo é de 3,8%; a permeabilidade mínima é de 130 AFS e a rcv mínima é de 13 L/POL².

A coleta de dados foi feita mediante autorização do diretor industrial da empresa e executada diretamente na empresa. Cada uma das variáveis acima citadas formará uma série temporal diária composta por dias úteis de produção, abrangendo o período de 29 de abril de 2008 a 30 de junho de 2009, totalizando 319 observações.

Apesar de a empresa atuar em três turnos e as análises de laboratório acontecerem de hora em hora, selecionou-se apenas as amostras referentes à primeira hora de cada um dos turnos, pois esta é a forma mais adequada estatisticamente de se montar uma série temporal, levando em consideração a alta frequência dos dados.

Nesta pesquisa ex-post-facto, foi realizado um estudo de controle de qualidade utilizando-se um modelo de regressão dinâmica multivariada, tendo em vista que as variáveis sob estudo são inter-relacionadas. Ao encontrar um modelo matemático para cada uma das variáveis envolvidas, construíram-se os respectivos gráficos de controle utilizando-se os resíduos do modelo estimado com o objetivo de verificar a estabilidade do processo, identificando as possíveis variáveis causadoras dessa instabilidade.

Para avaliar as previsões do modelo estimado, deixou-se uma porção de observações fora da amostra. Assim, as seis últimas observações foram utilizadas para validar o modelo estimado, ficando fora da amostra, e, o modelo de regressão dinâmica multivariada foi estimado com as observações restantes. As previsões foram avaliadas pelos critérios MAPE e U de Theil.

4.1.2 Análise descritiva

Após a aquisição dos dados, é preciso conhecer o comportamento das variáveis no decorrer do período analisado. Na Tabela 1, apresentam-se algumas estatísticas descritivas do teor de umidade, permeabilidade e rcv.

Tabela 1 – Estatísticas descritivas para as variáveis teor de umidade, permeabilidade e rcv da Empresa A no período de 29/04/2008 a 30/06/2009

Estatísticas	teor de umidade	permeabilidade	rcv
Média	3,490	155,893	18,737
Mediana	3,500	160,000	19,000
Desvio Padrão	0,242	11,239	1,757
Coefficiente de variação (%)	6,930	7,210	9,378
Assimetria	-0,148	0,009	0,960
Curtose	-0,222	-0,159	2,554
Mínimo	2,900	130,000	15,000
Máximo	4,200	190,000	27,000

Pela análise descritiva das variáveis, verificou-se, pelo coeficiente de variação, que as médias das variáveis são representativas. Isso não significa que todos os seus pontos amostrais analisados estejam dentro dos limites de controle desejados, pois o coeficiente de variação é uma medida de dispersão utilizada quando se deseja comparar a variação de conjuntos de dados, mas indica que há uma pequena variabilidade no decorrer do período analisado. Percebeu-se também que as variáveis possuem uma distribuição assimétrica negativa, uma vez que suas médias são menores que a mediana.

Embora a análise descritiva revele o comportamento do processo, é necessário investigar sua estabilidade por meio dos gráficos de controle apresentados na seção 4.1.5, os quais irão caracterizá-lo em relação a sua estabilidade.

Inicialmente, estimaram-se as equações independentemente por meio do método de mínimos quadrados ordinários, analisando as variáveis individualmente. Com os resíduos oriundos das equações, calculou-se a matriz de correlação. A matriz de correlação revelou que existe correlação significativa ao nível de significância adotado, indicando ligação entre as variáveis por meio de seus resíduos.

Como existe correlação, é aconselhável fazer uma análise conjunta, a qual poderá fornecer informações relevantes sobre a relação existente a partir do comportamento multivariado, se comparada com a análise individual. Esse pressuposto comprova a estimação

de um modelo matemático utilizando equações aparentemente não-correlacionadas, que em inglês denomina-se *Seegmingly Unrelated Regression* (SUR), para estimar um modelo de regressão dinâmica multivariado.

4.1.3 Estimação do modelo de regressão multivariado

Para estimar o modelo matemático, foram construídas, a partir dos dados coletados, séries temporais correspondentes às variáveis sob estudo. Como os dados em análise apresentam-se inter-relacionados, não é possível a utilização direta destes valores nos gráficos de controle, pois os pressupostos de normalidade, homocedasticidade e independência são violados. Então, foi necessário estimar um modelo matemático que possibilitasse uma estimação conjunta, considerando as relações existentes entre as variáveis e as relações dinâmicas que ocorrem entre elas, obtendo-se após a estimação uma equação para cada variável, para, em seguida, aplicar os gráficos de controle nos resíduos desse modelo.

A metodologia empregada para esta pesquisa ex-post-facto foi utilizar o método de estimação SUR para encontrar um modelo de regressão dinâmica multivariada que representasse as variáveis. Com os resíduos desse modelo constroem-se os gráficos de controle para verificar a estabilidade das variáveis: teor de umidade, permeabilidade e rcv medidas e controladas no processo de fundição de moldagem em areia verde.

Na Tabela 2, apresentam-se as equações estimadas para o conjunto de variáveis, levando em consideração uma defasagem inicial de ordem três, as quais foram ajustadas por meio do método de mínimos quadrados ordinários de três estágios, que considera um nível de ajuste para todo o conjunto de variáveis.

Tabela 2 – Estimação dos coeficientes, erro padrão, t-calc e p-valor do modelo de regressão dinâmica multivariada para as variáveis teor de umidade, permeabilidade e rcv

Modelagem da variável - teor de umidade				
Variáveis	Coefficientes	Erro padrão	t-calc	p-valor
teor de umidade _{t-1}	0,329728	0,05489	6,010	0,000
teor de umidade _{t-2}	0,285004	0,05587	5,100	0,000
teor de umidade _{t-3}	0,178503	0,05457	3,270	0,001
permeabilidade _{t-1}	0,002515	0,00069	3,670	0,000
rcv _{t-2}	0,017489	0,00590	2,960	0,003

Modelagem da variável – permeabilidade				
Variáveis	Coefficientes	Erro padrão	t-calc	p-valor
permeabilidade _{t-1}	0,308247	0,05348	5,760	0,000
permeabilidade _{t-2}	0,337874	0,05314	6,360	0,000
permeabilidade _{t-3}	0,239161	0,05271	4,540	0,000
rcv _{t-3}	0,945208	0,27600	3,430	0,001

Modelagem da variável – rcv				
Variáveis	Coefficientes	Erro padrão	t-calc	p-valor
teor de umidade _{t-2}	0,158920	0,33490	3,460	0,001
permeabilidade _{t-3}	0,020881	0,00574	3,640	0,000
rcv _{t-1}	0,198212	0,05352	3,700	0,000
rcv _{t-2}	0,201790	0,05303	3,810	0,000
rcv _{t-3}	0,210247	0,05225	4,020	0,000

A Figura 6 exibe a função de autocorrelação para o teor de umidade, a permeabilidade e o rcv, e a Figura 7 apresenta, graficamente, os valores originais e estimados das respectivas variáveis.

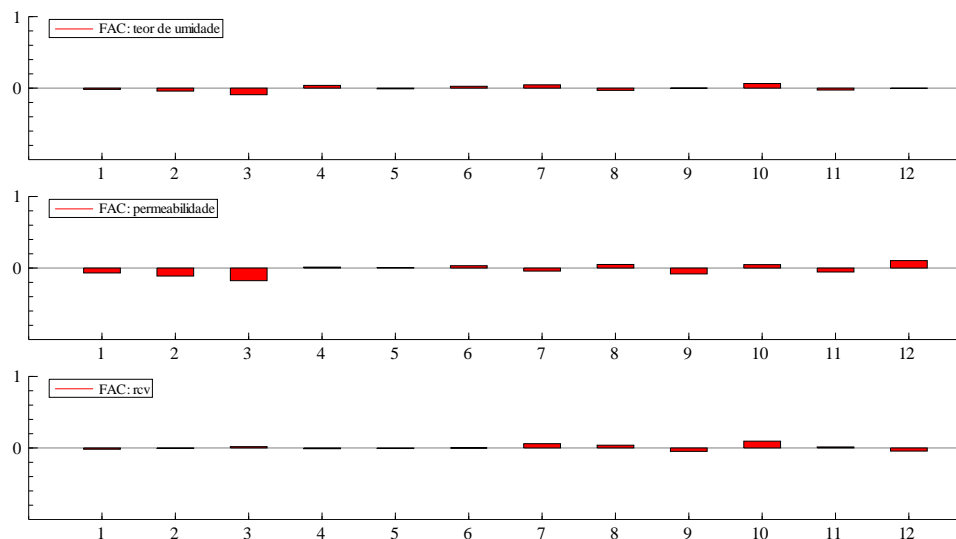


Figura 6 – Função de autocorrelação para as variáveis teor de umidade, permeabilidade e rcv

A FAC da Figura 6 revela que os resíduos das variáveis são independentes. Sendo assim, o processo é um ruído branco.

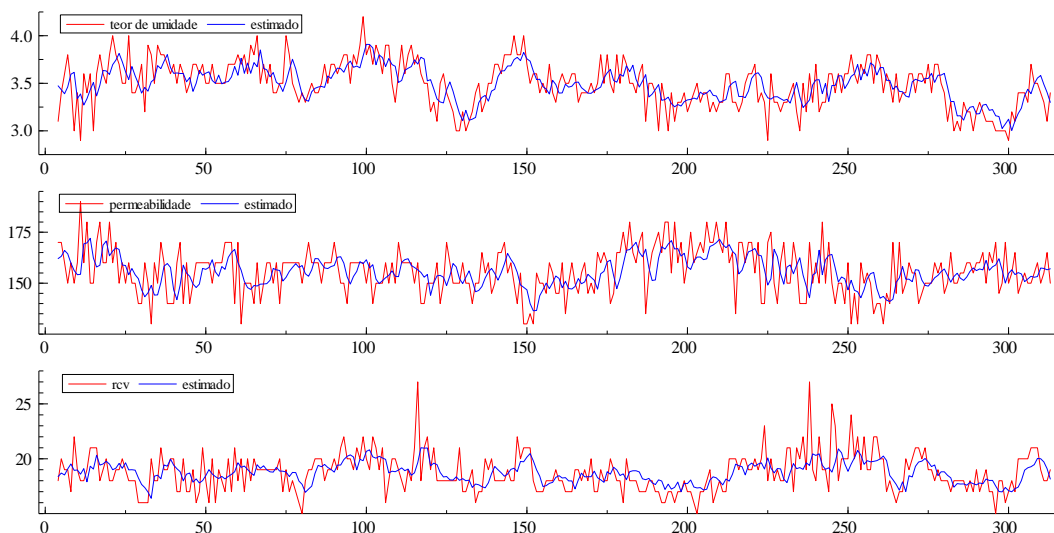


Figura 7 – Valores originais e estimados para as variáveis teor de umidade, permeabilidade e rcv

Pela análise gráfica do modelo estimado, conforme a Figura 7, percebe-se o bom ajuste do modelo de regressão dinâmica multivariada, sob o qual é realizada uma previsão dinâmica fora da amostra para seis períodos à frente, os quais serão comparados aos dados originais com a finalidade de avaliar a eficiência do modelo.

Para comprovar que o modelo matemático de equação dinâmica multivariada é adequado, mostra-se a seguir um diagnóstico dos resíduos do modelo estimado.

4.1.4 Diagnóstico dos resíduos

Além de o modelo matemático eliminar a autocorrelação das variáveis em estudo, é importante analisar a normalidade e a homocedasticidade dos resíduos com o intuito de avaliar se os pressupostos básicos dos gráficos de controle são atendidos.

A Tabela 3 apresenta os resultados dos testes de normalidade e dos testes de heterocedasticidade, considerando o conjunto de variáveis.

As hipóteses testadas são as seguintes:

H_0 : Homocedasticidade dos resíduos;

H_1 : Heterocedasticidade dos resíduos.

Para o teste de normalidade, utilizou-se a estatística χ^2 e, para os testes de heterocedasticidade, empregou-se a estatística F , que é a estatística calculada. Se o valor de F for significativo, então, há heterocedasticidade condicional na série.

Tabela 3 – Diagnóstico dos resíduos multivariados

Teste	Resultado	Interpretação
Teste de normalidade	Chi ² (6)= 1,814 [0,404]*	Aceita Ho
Teste de White sem os produtos cruzados	F(108,1623)= 0,772 [0,959]*	Aceita Ho
Teste de White com os produtos cruzados	F(324,1477)= 1,127 [0,079]*	Aceita Ho

* Os valores entre colchetes representam os *p-valores* associados a cada teste

Analisando o *p-valor* para cada teste, constatou-se que os resultados dos testes caem na região de aceitação da hipótese nula. Sendo assim, o teste de normalidade apontou que os resíduos do modelo matemático são normais, e, os testes de heterocedasticidade assinalaram que os mesmos são homocedásticos.

Os testes apresentados na Tabela 4, que levam em conta as variáveis individualmente, confirmaram os resultados obtidos na Tabela 3.

Para comprovar a presença de efeitos ARCH, testam-se as seguintes hipóteses:

$H_0 : \alpha_i = 0$, ou seja, não existem efeitos ARCH;

$H_1 : \alpha_i \neq 0$, ou seja, existem efeitos ARCH.

Tabela 4 – Diagnóstico dos resíduos individuais

Teste	Variáveis	Resultado	Interpretação
Teste ARCH-LM	Teor de umidade	F(1,304)= 0,108 [0,742]*	Aceita Ho
	Permeabilidade	F(1,304)= 0,173 [0,678]*	Aceita Ho
	Rcv	F(1,304)= 2,226 [0,137]*	Aceita Ho
Teste de White sem os produtos cruzados	Teor de umidade	F(18,287)= 0,480 [0,965]*	Aceita Ho
	Permeabilidade	F(18,287)= 1,279 [0,201]*	Aceita Ho
	Rcv	F(18,287)= 0,511 [0,952]*	Aceita Ho
Teste de White com os produtos cruzados	Teor de umidade	F(54,251)= 0,666 [0,963]*	Aceita Ho
	Permeabilidade	F(54,251)= 1,433 [0,052]*	Aceita Ho
	Rcv	F(54,251)= 1,059 [0,376]*	Aceita Ho

* Os valores entre colchetes representam os *p-valores* associados a cada teste

Observando-se os *p-valores* dos testes apresentados na Tabela 4, constatou-se que o modelo de regressão dinâmica multivariada não apresenta heterocedasticidade condicional em

nível de significância de 5%, ou seja, os resíduos apresentam variância constante, conhecida por homocedasticidade.

Mediante os resultados obtidos pelo diagnóstico dos resíduos constatou-se que os mesmos atendem os pressupostos básicos para aplicação dos gráficos de controle. Antes de verificar a estabilidade do processo através desses gráficos, apresenta-se, na Figura 8, a previsão dinâmica para as variáveis e, na Tabela 5, os valores originais e previstos.

Conforme mencionado anteriormente, o modelo estimado foi elaborado com 313 observações. As últimas seis ficaram de fora da amostra com o objetivo de avaliar o modelo matemático. Sendo assim, a previsão ocorre da observação 314 à observação 319.

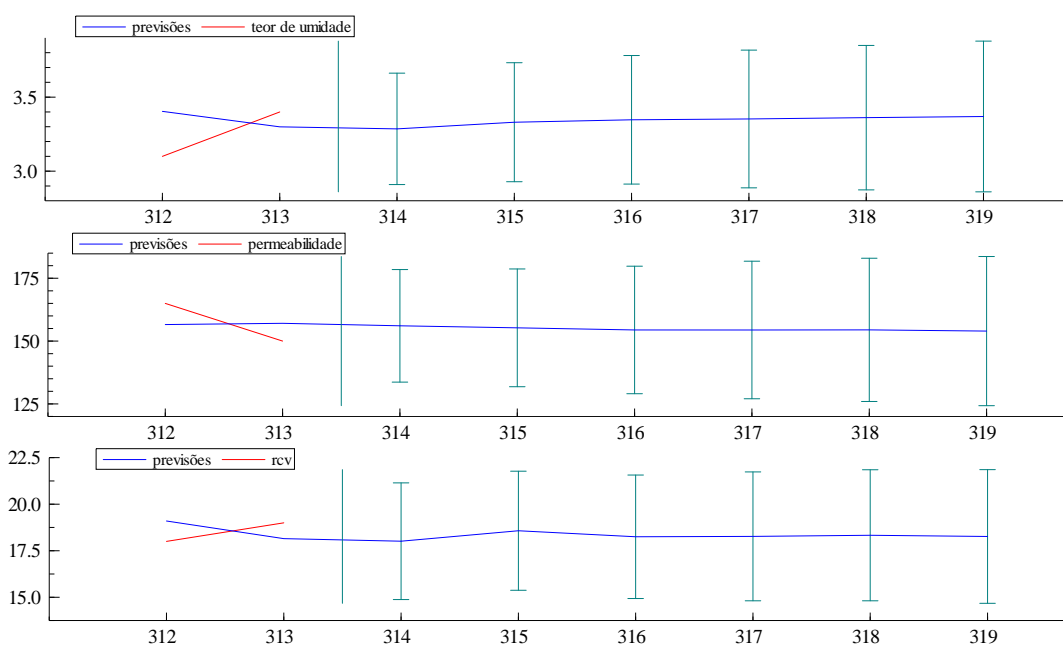


Figura 8 – Previsão dinâmica para o teor de umidade, permeabilidade e rcv

Tabela 5 – Período, valores originais, valores previstos, erro de previsão, MAPE e U de Theil para as variáveis teor de umidade, permeabilidade e rcv da Empresa A

Previsão para o teor de umidade			
Período	Originais	Previstos	Erro de previsão
24/6/2009	3,500	3,286	0,214
25/6/2009	3,400	3,331	0,069
26/6/2009	3,600	3,347	0,253
27/6/2009	3,500	3,353	0,147
28/6/2009	3,400	3,361	0,039
30/6/2009	3,400	3,369	0,031

MAPE(%)= 3,574 e U de Theil= 0,954

Previsão para a permeabilidade			
Período	Originais	Previstos	Erro de previsão
24/6/2009	140,000	156,070	-16,070
25/6/2009	155,000	155,265	-0,265
26/6/2009	165,000	154,425	10,575
27/6/2009	145,000	154,412	-9,412
28/6/2009	150,000	154,463	-4,463
30/6/2009	155,000	153,968	1,032

MAPE(%)= 4,698 e U de Theil= 0,522

Previsão para a rcv			
Período	Originais	Previstos	Erro de previsão
24/6/2009	19,000	18,012	0,988
25/6/2009	19,000	18,574	0,426
26/6/2009	18,000	18,251	-0,251
27/6/2009	20,000	18,271	1,729
28/6/2009	20,000	18,331	1,670
30/6/2009	17,000	18,267	-1,267

MAPE(%)= 5,547 e U de Theil= 0,745

Analisando o erro de previsão, é possível afirmar que, se os erros possuem valores negativos, é provável que esteja ocorrendo algum problema no processo relacionado às causas comuns ou especiais. Além disso, o MAPE e o U de Theil comprovaram estatisticamente que o modelo estimado apresenta uma previsão melhor que a previsão ingênua ou trivial. Vale também lembrar que a modelagem das séries analisadas é um dos objetivos principais, de modo a eliminar a autocorrelação serial e a correlação entre as variáveis, de modo que gráficos de controle pudessem ser aplicados.

4.1.5 Análise da estabilidade do processo

Nesta etapa, o objetivo é monitorar o processo de fundição de moldagem em areia verde, a partir do modelo de regressão multivariada, para verificar sua estabilidade e tentar identificar as variáveis responsáveis pela instabilidade utilizando os gráficos de controle para medidas individuais e amplitude móvel.

Para iniciar este estudo, apresentam-se os gráficos de controle \bar{x} e \overline{MR} nas Figuras 9, 10, 11, 12, 13 e 14, usando os dados originais para mostrar como está o comportamento das

variáveis: teor de umidade, permeabilidade e rcv com base nas amostras coletadas. Nos gráficos de controle, foram adicionados os *runs tests*⁸, os quais revelam a existência de algum tipo de padrão nos dados. O gráfico \bar{x} serve para avaliar se o processo é afetado por grandes variações na média do processo.

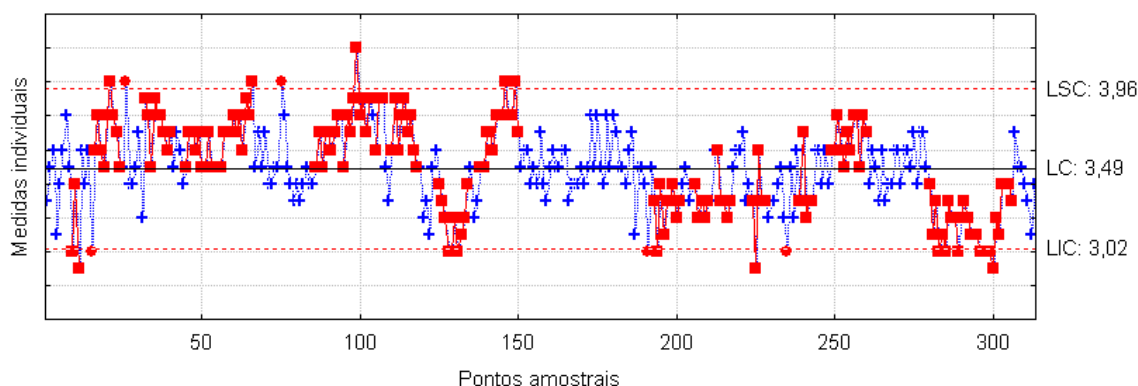


Figura 9 – Gráfico \bar{x} do teor de umidade, considerando os dados originais

De acordo com a Figura 9, os pontos vermelhos grafados no gráfico sugerem a instabilidade do processo de moldagem, usando os *runs tests*. Detectou-se que **9 pontos amostrais em sequência estão no mesmo lado da linha central**, representadas pelo conjunto de amostras iniciadas nas observações 16, 32, 45, 54, 86, 95, 110, 125, 138, 193, 249, 280, 289 e 298. Existem também um conjunto de 13 amostras em que, **a cada 3 pontos amostrais traçados, 2 se encontram localizados na zona A ou fora dela**, correspondendo às observações com início nas amostras 9, 64, 97, 101, 105, 127, 130, 194, 281, 284, 292, 295 e 298. Observa-se, também, um conjunto de 22 amostras, em que, **a cada 5 pontos amostrais traçados, 4 estão na zona B ou fora dela**, as quais são representadas pelo conjunto de amostras iniciadas nas observações 19, 32, 56, 61, 89, 94, 99, 111, 125, 130, 140, 145, 193, 206, 213, 224, 239, 256, 280, 285, 290 e 295. Além disso, observam-se **8 pontos amostrais em sequência situados fora da zona C**, correspondendo às observações com início em 96, 126, 143, 281 e 289.

Observando a Figura 10, constatou-se que **9 pontos amostrais em sequência estão no mesmo lado da linha central**, representadas pelo conjunto de amostras iniciadas nas observações 4, 78 e 291. Existem também um conjunto de 4 amostras em que, **a cada 3 pontos amostrais traçados, 2 se encontram localizados na zona A ou fora dela**,

⁸ Também chamados testes de sequência, ou regras sensibilizantes segundo Montgomery (2004).

correspondendo às observações com início nas amostras 9, 14, 25 e 30. Observa-se, também, um conjunto de 3 amostras, em que, **a cada 5 pontos amostrais traçados, 4 estão na zona B ou fora dela**, as quais são representadas pelo conjunto de amostras iniciadas nas observações 8, 54 e 294.

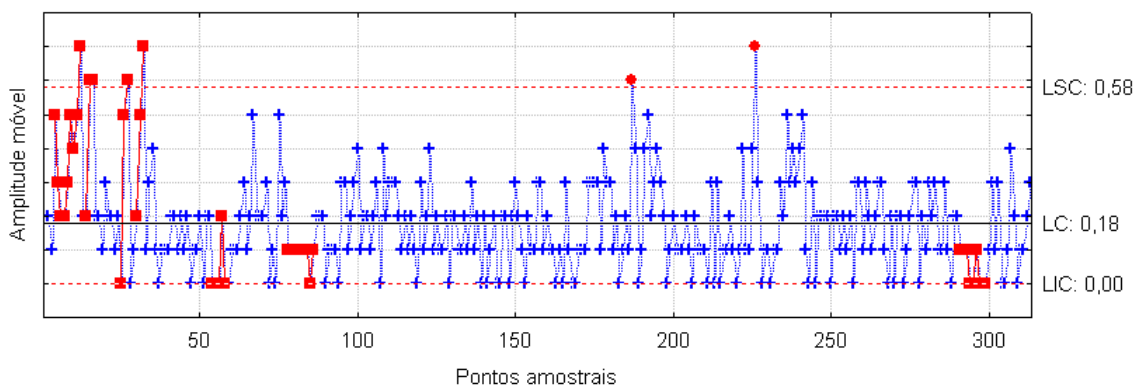


Figura 10 – Gráfico \overline{MR} do teor de umidade, considerando os dados originais

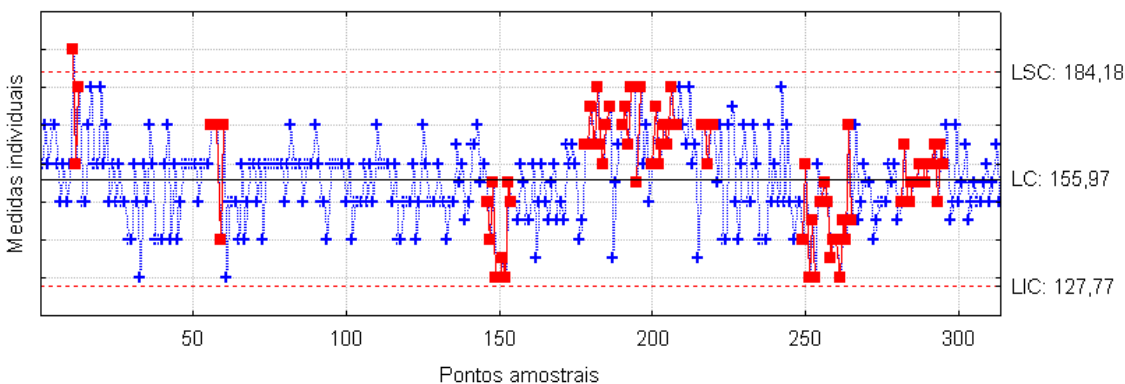


Figura 11 – Gráfico \bar{x} da permeabilidade, considerando os dados originais

Analisando-se a Figura 11, verificou-se que **9 pontos amostrais em sequência estão no mesmo lado da linha central**, representadas pelo conjunto de amostras iniciadas nas observações 146, 178, 200 e 255. Há um conjunto de 7 amostras em que, **a cada 3 pontos amostrais traçados, 2 se encontram localizados na zona A ou fora dela**, correspondendo às observações com início nas amostras 11, 148, 151, 180, 191, 194 e 251, e um conjunto de 7 amostras em que, **a cada 5 pontos amostrais traçados, 4 se encontram localizados na zona B ou fora dela**, correspondendo às observações com início nas amostras 56, 147, 190, 204, 216, 249 e 257. Além disso, observam-se **15 pontos amostrais em sequência situados na**

zona C, correspondendo à observação com início em 281 e **8 pontos amostrais em sequência situados fora da zona C**, correspondendo à observação com início em 258.

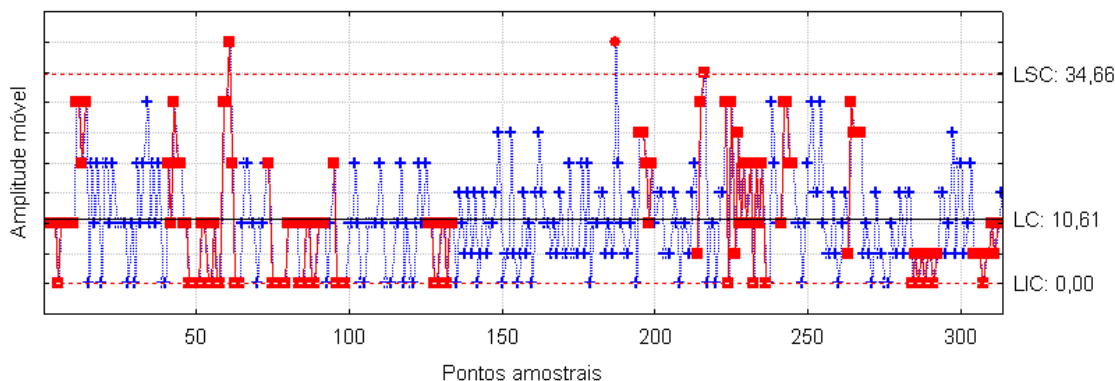


Figura 12 – Gráfico \overline{MR} da permeabilidade, considerando os dados originais

Pela Figura 12, detectou-se que **9 pontos amostrais em sequência estão no mesmo lado da linha central**, representadas pelo conjunto de amostras iniciadas nas observações 2, 46, 75, 84, 126, 284 e 304. Existem também um conjunto de 5 amostras em que, **a cada 3 pontos amostrais traçados, 2 se encontram localizados na zona A ou fora dela**, correspondendo às observações com início nas amostras 10, 58, 214, 223 e 241. Observa-se, também, um conjunto de 10 amostras, em que, **a cada 5 pontos amostrais traçados, 4 estão na zona B ou fora dela**, as quais são representadas pelo conjunto de amostras iniciadas nas observações 10, 41, 47, 54, 59, 74, 95, 195, 241 e 263. Além disso, observam-se **8 pontos amostrais em sequência situados fora da zona C**, correspondendo à observação com início em 54 e **14 pontos amostrais em sequência alternando-se para cima e para baixo**, correspondendo à observação 223.

Pela inspeção da Figura 13, constatou-se que **9 pontos amostrais em sequência estão no mesmo lado da linha central**, representadas pelo conjunto de amostras iniciadas nas observações 65, 88, 97, 182, 191, 213, 248, 268, 279 e 294. Há um conjunto de 10 amostras em que, **a cada 3 pontos amostrais traçados, 2 se encontram localizados na zona A ou fora dela**, correspondendo às observações com início nas amostras 28, 31, 51, 78, 191, 201, 236, 244, 251 e 257, e um conjunto de 19 amostras em que, **a cada 5 pontos amostrais traçados, 4 se encontram localizados na zona B ou fora dela**, correspondendo às observações com início nas amostras 28, 41, 76, 90, 95, 115, 146, 189, 194, 199, 208, 213, 231, 245, 250, 257, 262, 268 e 302. Além disso, observam-se **8 pontos amostrais em**

seqüência situados fora da zona C, correspondendo às observações com início em 190 e 302.

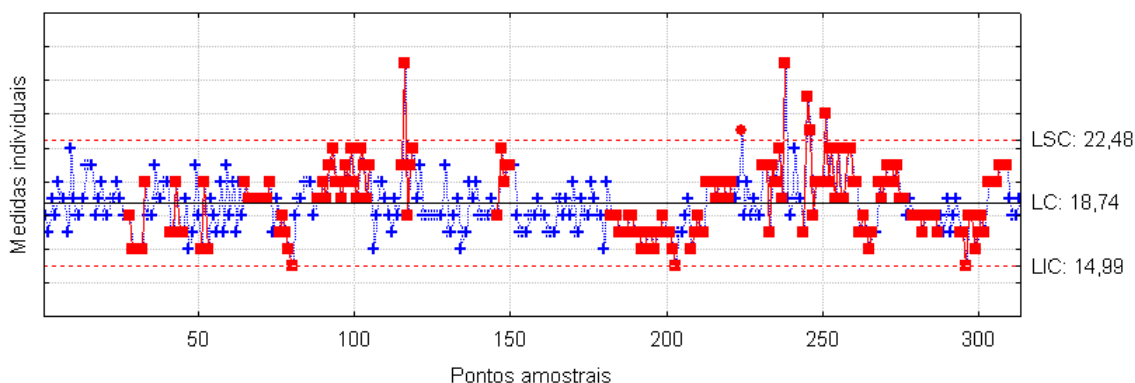


Figura 13 – Gráfico \bar{x} da rcv, considerando os dados originais

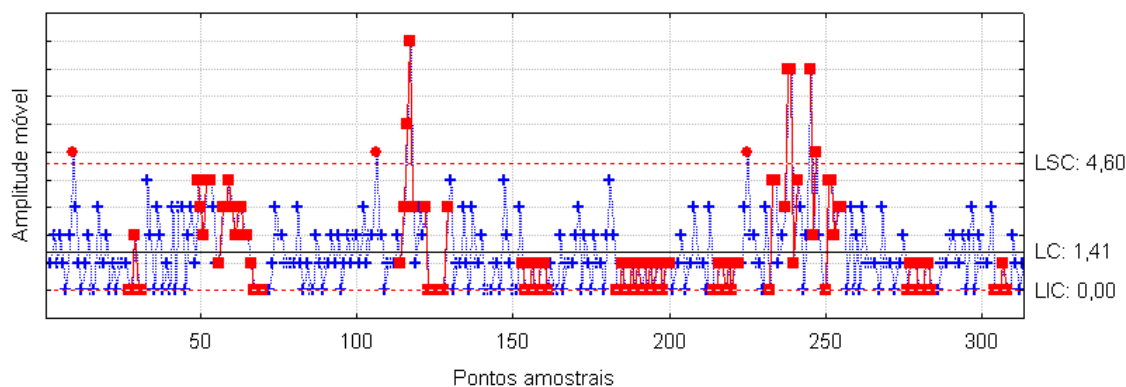


Figura 14 – Gráfico \overline{MR} da rcv, considerando os dados originais

Analisando a Figura 14, verificou-se que **9 pontos amostrais em seqüência estão no mesmo lado da linha central**, representadas pelo conjunto de amostras iniciadas nas observações 57, 153, 183, 192, 214 e 276. Existem também um conjunto de 6 amostras em que, **a cada 3 pontos amostrais traçados, 2 se encontram localizados na zona A ou fora dela**, correspondendo às observações com início nas amostras 51, 115, 232, 237, 245 e 250. Observa-se, também, um conjunto de 11 amostras, em que, **a cada 5 pontos amostrais traçados, 4 estão na zona B ou fora dela**, as quais são representadas pelo conjunto de amostras iniciadas nas observações 27, 49, 56, 66, 114, 122, 154, 183, 237, 251 e 304. Além disso, observam-se **8 pontos amostrais em seqüência situados fora da zona C**, correspondendo à observação com início em 122.

Ao realizar outra análise, simulando a exclusão dos pontos amostrais fora de controle, tem-se outra visão da instabilidade que está ocorrendo no processo. É importante salientar que, com essa exclusão dos pontos amostrais fora de controle, os limites de controle dos gráficos \bar{x} e \overline{MR} são recalculados. As Figuras 15, 16, 17, 18, 19 e 20 mostram essa simulação.

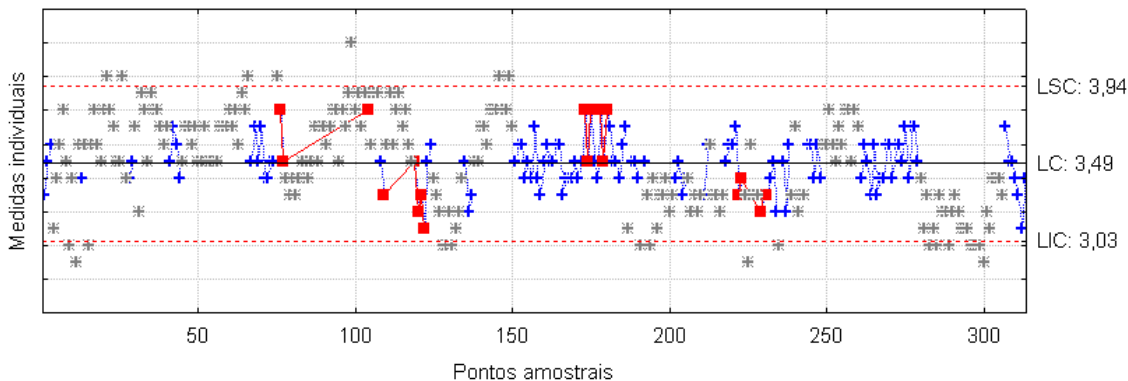


Figura 15 – Gráfico \bar{x} do teor de umidade com a exclusão dos pontos amostrais fora de controle, considerando os dados originais

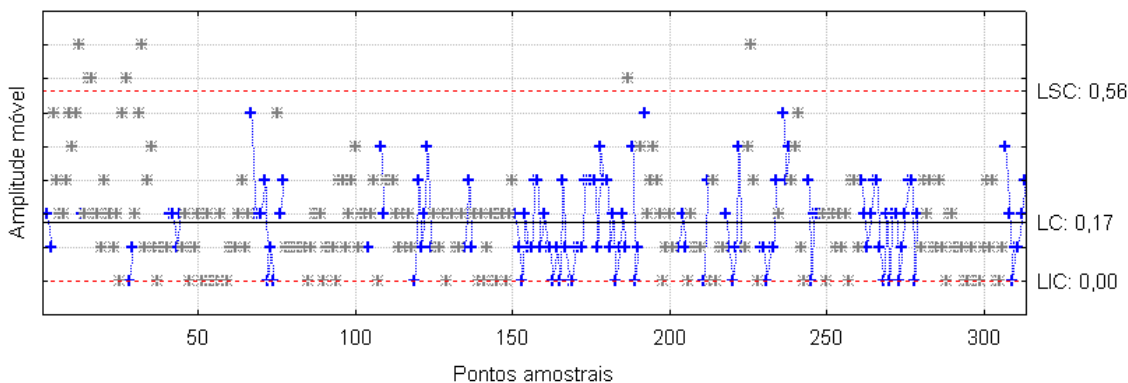


Figura 16 – Gráfico \overline{MR} do teor de umidade com a exclusão dos pontos amostrais fora de controle, considerando os dados originais

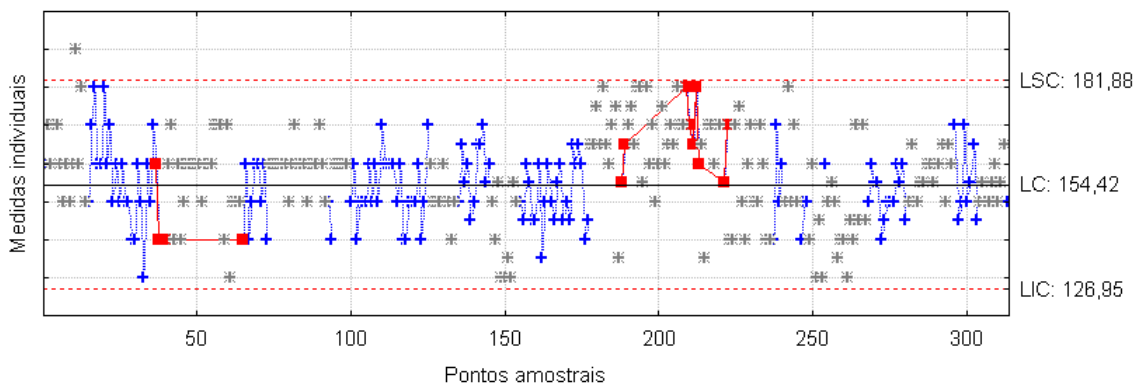


Figura 17 – Gráfico \bar{x} da permeabilidade com a exclusão dos pontos amostrais fora de controle, considerando os dados originais

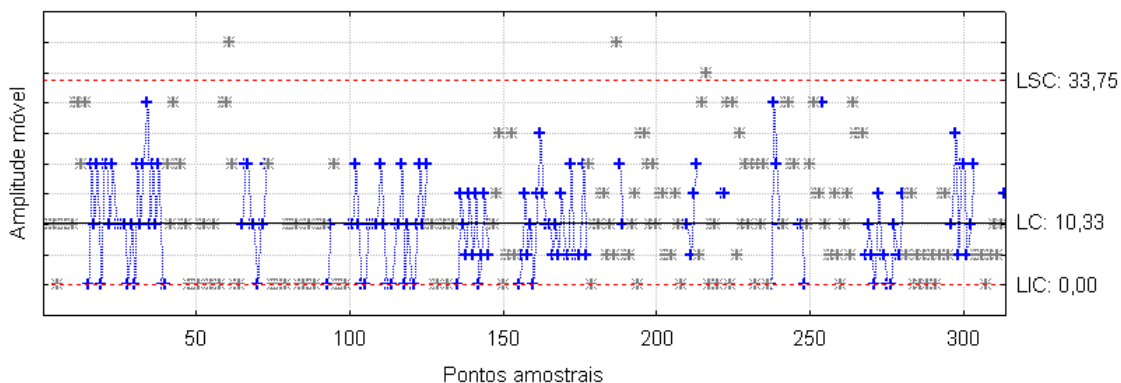


Figura 18 – Gráfico \overline{MR} da permeabilidade com a exclusão dos pontos amostrais fora de controle, considerando os dados originais

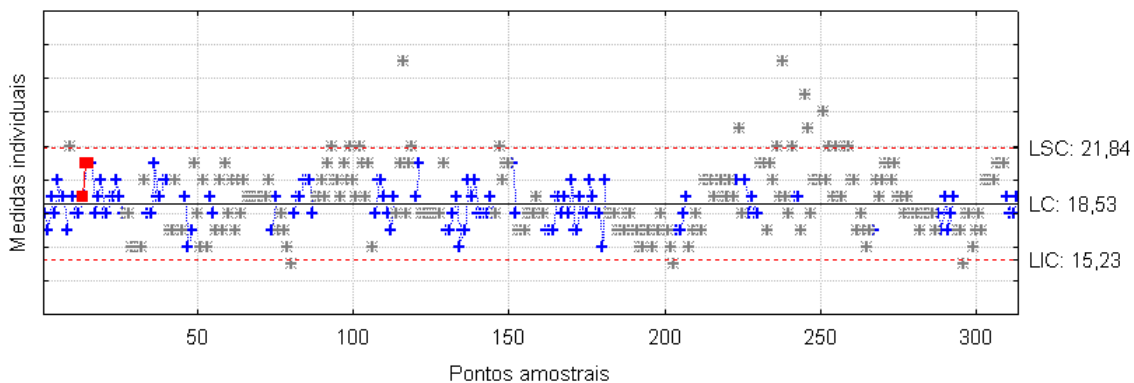


Figura 19 – Gráfico \bar{x} da rev com a exclusão dos pontos amostrais fora de controle, considerando os dados originais

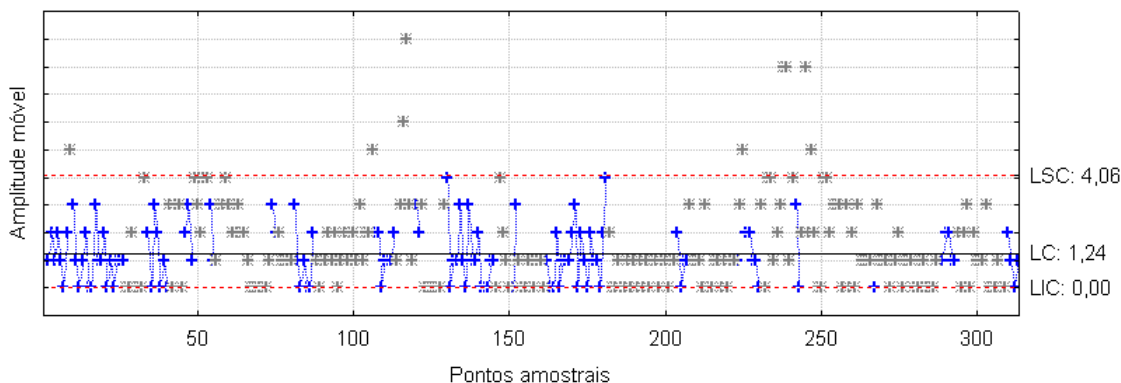


Figura 20 – Gráfico \overline{MR} da rcv com a exclusão dos pontos amostrais fora de controle, considerando os dados originais

Mediante os resultados obtidos, justifica-se que as técnicas convencionais do CEP não podem ser utilizadas diretamente nos dados originais, ou seja, é necessário um tratamento na variável de interesse para remover a autocorrelação serial, pois a mesma pode inserir pontos amostrais fora de controle, quando na verdade o processo pode estar sob controle, nesse caso, gerando alarmes falsos.

Para sanar o efeito de autocorrelação serial e a correlação entre as variáveis, estimou-se um modelo de regressão dinâmica multivariada, atendendo perfeitamente os pressupostos para a aplicação de gráficos de controle. Sendo assim, foram utilizados os resíduos do modelo estimado para traçar os gráficos \bar{x} e \overline{MR} , tendo como auxílio na decisão a aplicação dos *runs tests*.

Pelas Figuras 21, 22, 23, 24, 25 e 26 é possível identificar a ocorrência de padrões especiais, caracterizando o comportamento do processo de fundição de moldagem em areia verde quanto a sua estabilidade e/ou instabilidade de modo que seja feito ou não um ajuste.

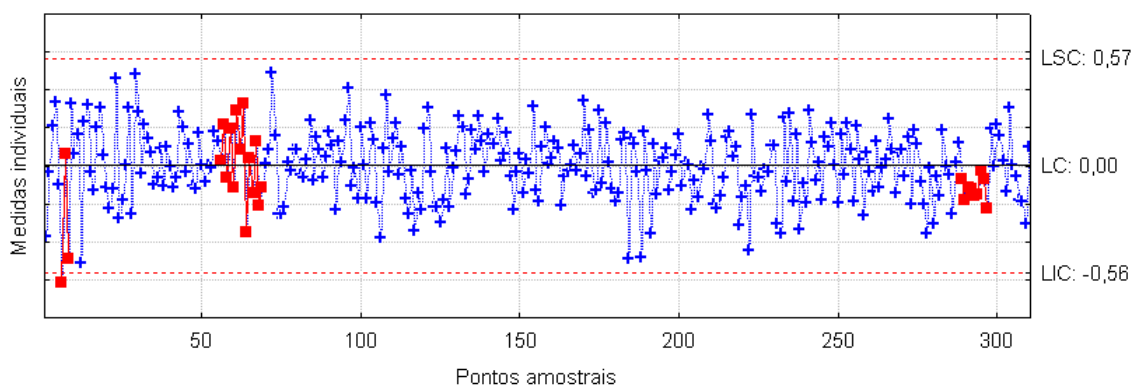


Figura 21 – Gráfico \bar{x} aplicado ao resíduo do teor de umidade

De acordo com os resíduos do modelo estimado, traçou-se o gráfico \bar{x} , mostrado na Figura 21, pelo qual se detecta que **9 pontos amostrais em sequência estão no mesmo lado da linha central**, representadas pela amostra com início na observação 289. Existem também **14 pontos amostrais em sequência alternando-se para cima e para baixo**, com início na observação 56. Observa-se ainda que, a **cada 3 pontos amostrais traçados, 2 se encontram localizados na zona A ou fora dela**, correspondendo à observação com início na amostra 6.

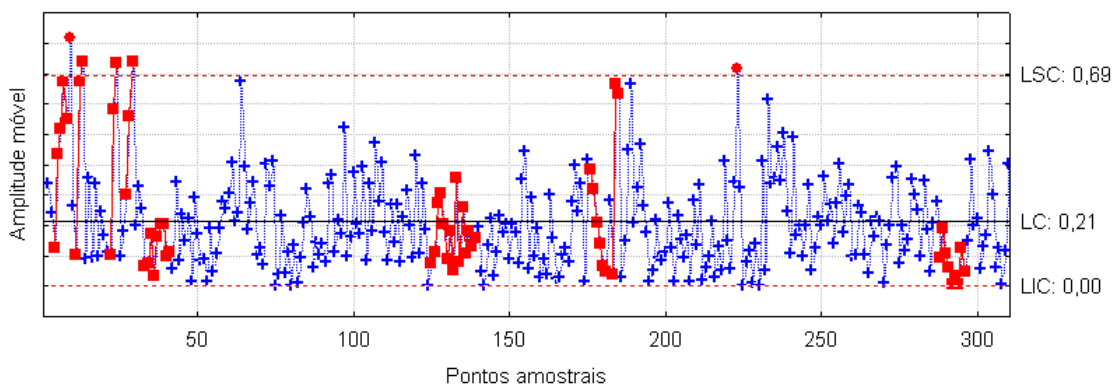


Figura 22 – Gráfico \overline{MR} aplicado ao resíduo do teor de umidade

Utilizando os resíduos do modelo estimado, traçou-se o gráfico \overline{MR} , mostrado na Figura 22, pelo qual se constata que **9 pontos amostrais em sequência estão no mesmo lado da linha central**, representadas pela amostra com início nas observações 33 e 288. Existem também **6 pontos amostrais em sequência com tendência crescente e/ou decrescente**, que corresponde à observação com início na amostra 176. Observa-se ainda que, a **cada 3 pontos amostrais traçados, 2 se encontram localizados na zona A ou fora dela**, correspondendo às observações com início nas amostras 6, 11, 22, 27 e 183. Verifica-se, também, um conjunto de 2 amostras, em que, **a cada 5 pontos amostrais traçados, 4 estão na zona B ou fora dela**, as quais são representadas pelo conjunto de amostras iniciadas nas observações 4 e 292. Além disso, observam-se **15 pontos amostrais em sequência situados na zona C**, correspondendo à observação com início em 125.

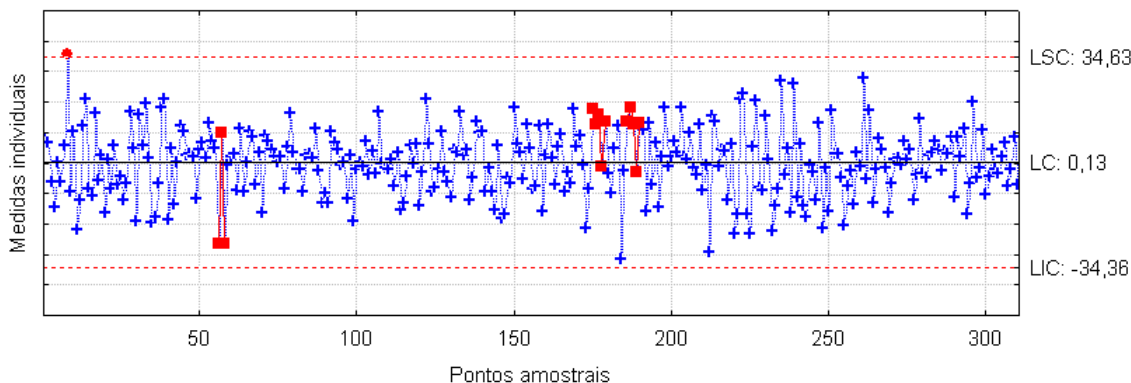


Figura 23 – Gráfico \bar{x} aplicado ao resíduo da permeabilidade

Analisando-se os resíduos do modelo estimado, traçou-se o gráfico \bar{x} , apresentado na Figura 23, pelo qual se verifica um conjunto de amostras em que, **a cada 3 pontos amostrais traçados, 2 se encontram localizados na zona A ou fora dela**, correspondendo à observação com início na amostra 56. Há também um conjunto de 2 amostras em que, **a cada 5 pontos amostrais traçados, 4 se encontram localizados na zona B ou fora dela**, correspondendo às observações com início nas amostras 175 e 186.

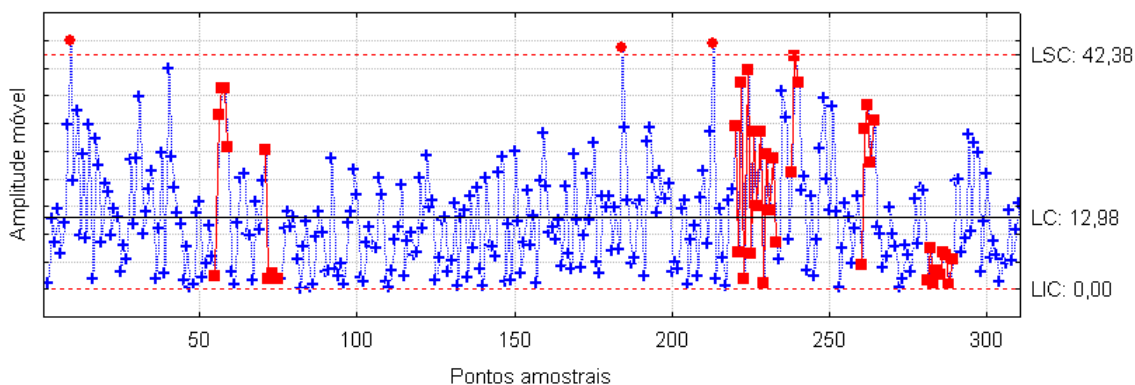


Figura 24 – Gráfico \overline{MR} aplicado ao resíduo da permeabilidade

De acordo com os resíduos do modelo estimado, traçou-se o gráfico \overline{MR} , mostrado na Figura 24, pelo qual se detecta que **9 pontos amostrais em sequência estão no mesmo lado da linha central**, representadas pela amostra com início na observação 281. Existem também **14 pontos amostrais em sequência alternando-se pra cima e para baixo**, que corresponde à observação 220. Observa-se ainda que, **a cada 3 pontos amostrais traçados, 2 se encontram localizados na zona A ou fora dela**, correspondendo às observações com início nas amostras 56, 222 e 238. Verifica-se, também, um conjunto de 3 amostras, em que, **a cada**

5 pontos amostrais traçados, 4 estão na zona B ou fora dela, as quais são representadas pelo conjunto de amostras iniciadas nas observações 55, 71 e 260.

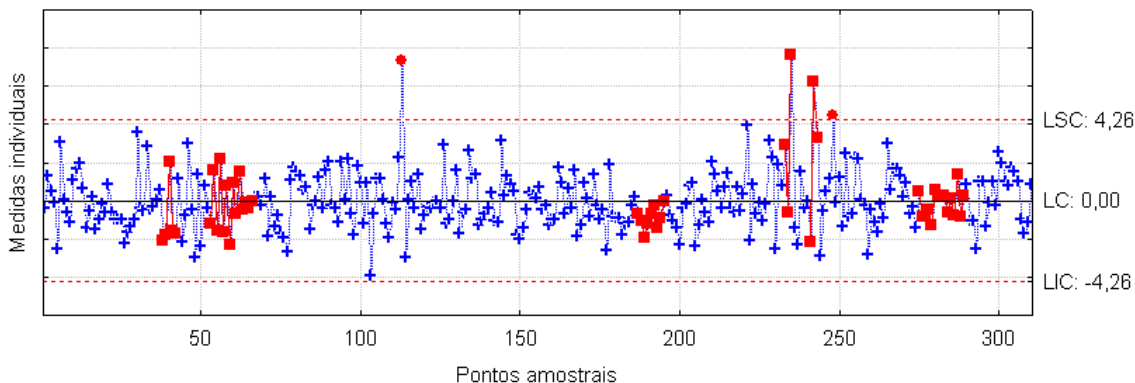


Figura 25 – Gráfico \bar{x} aplicado ao resíduo da rcv

Pela inspeção da Figura 25, que traça o gráfico \bar{x} usando os resíduos do modelo estimado, constata-se que **9 pontos amostrais em sequência estão no mesmo lado da linha central**, representadas pela amostra com início na observação 187; **14 pontos amostrais em sequência estão se alternando para cima e para baixo**, com início na observação 53. Existe também um conjunto de 2 amostras em que **cada 3 pontos amostrais traçados, 2 se encontram localizados na zona A ou fora dela**, correspondendo às observações com início nas amostras 233 e 241, e um conjunto de uma amostra em que, a **cada 5 pontos amostrais traçados, 4 se encontram localizados na zona B ou fora dela**, correspondendo à observação com início na amostra 38. Além disso, há **15 pontos amostrais em sequência situados na zona C**.

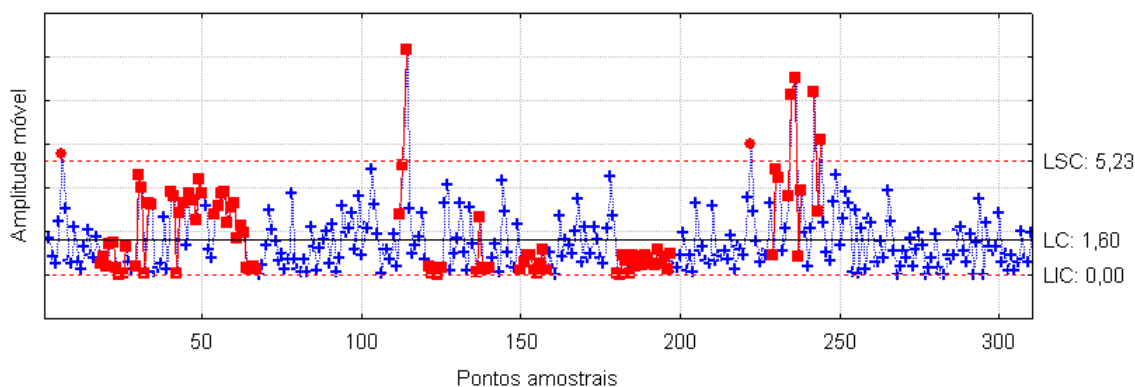


Figura 26 – Gráfico \overline{MR} aplicado ao resíduo da rcv

Com os resíduos do modelo estimado, traçou-se o gráfico \overline{MR} , mostrado na Figura 26, pelo qual se verifica que **9 pontos amostrais em sequência estão no mesmo lado da linha central**, representadas pela amostra com início nas observações 18, 54, 150, 180 e 189. Observa-se ainda que, a **cada 3 pontos amostrais traçados, 2 se encontram localizados na zona A ou fora dela**, correspondendo às observações com início nas amostras 29, 112, 229, 234 e 242. Verifica-se, também, um conjunto de 8 amostras, em que, **a cada 5 pontos amostrais traçados, 4 estão na zona B ou fora dela**, as quais são representadas pelo conjunto de amostras iniciadas nas observações 30, 40, 46, 55, 63, 121, 136 e 234. Além disso, observam-se **15 pontos amostrais em sequência situados na zona C**, correspondendo à observação com início em 125.

Analisando o gráfico \bar{x} das Figuras 21, 23 e 25, constatou-se a existência de poucos pontos amostrais fora de controle se comparado ao gráfico \bar{x} das séries originais. Isso indica que o modelo de regressão dinâmica multivariada foi capaz de captar as variações que havia nas séries: teor de umidade, permeabilidade e rcv, e também foi eficiente em modelar o processo da Empresa A, em que o efeito da autocorrelação serial foi tratado.

4.1.6 Comentários gerais da análise

Mediante a metodologia empregada, foi possível monitorar o comportamento do processo de fundição de moldagem em areia verde, utilizando o modelo de regressão dinâmica multivariada, que possibilita uma estimação conjunta das variáveis que compõem uma das etapas críticas de tal processo.

A modelagem realizada não teve apenas o objetivo de apresentar o valor futuro para cada uma das variáveis analisadas, mas também propiciou a detecção de pontos amostrais fora dos limites de controle.

Com relação aos resultados obtidos após a aplicação dos gráficos de controle \bar{x} , apresentam-se as seguintes considerações para cada uma das variáveis:

- Nos dados originais da variável teor de umidade havia vinte e seis pontos amostrais fora dos limites de controle; após o ajuste do modelo de regressão dinâmica multivariada, constataram-se apenas dois pontos amostrais fora dos limites de controle.

- Nos dados originais da variável permeabilidade havia catorze pontos amostrais fora dos limites de controle; após o ajuste do modelo de regressão dinâmica multivariada, constatou-se a existência de dois pontos amostrais fora dos limites de controle.
- Nos dados originais da variável rcv havia vinte pontos amostrais fora dos limites de controle; após o ajuste do modelo de regressão dinâmica multivariada, constatou-se a existência de quatro pontos amostrais fora dos limites de controle.

Investigações revelam que, após o ajuste dos dados por meio do modelo de regressão dinâmica multivariada, o efeito da autocorrelação foi tratado. Assim, o modelo evidenciou uma melhora significativa na detecção de pontos amostrais fora de controle.

Com os resultados obtidos, pode-se afirmar que a variável teor de umidade foi considerada como a variável mais instável em relação às outras, pois apresentou mais pontos amostrais fora dos limites de controle. Assim, pretende-se alertar a equipe de produção a fazer os ajustes necessários para obter um processo o mais próximo possível da especificação previamente estabelecida.

4.2 Empresa B

A empresa B, fundada em 1960, iniciou suas atividades com a industrialização de linhaça, matéria-prima na época, tradicional da região. Com o incremento da cultura da soja, montou-se uma indústria de óleos vegetais. A fábrica possui quatro plantas industriais, totalmente independentes, que podem trabalhar simultaneamente com produtos diferenciados, conferindo à empresa flexibilidade e disponibilidade perante o mercado:

- planta de extração de óleos vegetais com capacidade de processamento de 600 toneladas/dia;
- planta de extração de óleos vegetais com capacidade de processamento de 132 toneladas/dia;
- fábrica de farelo de soja integral, atualmente desativada, com capacidade de 400 toneladas/dia;
- planta de extração com prensas tipo *expeller*.

Para expandir-se num mercado competitivo, optou pela diversificação. Através de pesquisas de mercado, interação internacional, investimentos e, principalmente, inovação, a empresa passou a oferecer produtos diferenciados em oleaginosas, por exemplo: óleo de soja bruto degomado, óleo de canola em bruto, óleo de linhaça em bruto, óleo de girassol em bruto, proteína de soja, soja integral desativado, entre outros.

No local são industrializadas todas as matérias-primas recebidas na indústria. A empresa atua com a função de compra e armazenamento de matérias-primas, contando com moegas, armazéns, secadores, e, toda a infraestrutura montada para recebimento, armazenagem, vendas e expedição dos produtos.

Atualmente, as vendas estendem-se por todo o estado, e, também em nível nacional e internacional, sendo que o transporte é realizado por frota própria e terceirizado.

4.2.1 Descrição do processo e das variáveis a serem analisadas

Na atualidade, a soja domina o mercado mundial, tanto de proteína vegetal como de óleo comestível. O óleo de soja surgiu como um subproduto do processamento do farelo de soja e tornou-se um dos líderes mundiais no mercado de óleos. Esse óleo é considerado ímpar em suas propriedades, sendo indicado para um vasto número de aplicações, tais como preparação de assados, margarinas, óleo para salada, maionese, etc. (MORRETO e FETT, 1998).

Segundo a Resolução RDC nº 482, o óleo de soja segue a seguinte classificação:

- **óleo de soja:** óleo obtido pelos processos de extração e refino;
- **óleo de soja semi-refinado:** óleo obtido pelos processos de extração e neutralização;
- **óleo de soja degomado:** óleo obtido pelos processos de extração e degomagem;
- **óleo de soja bruto:** óleo obtido pelo processo de extração.

A industrialização das sementes oleaginosas divide-se em duas partes importantes: a produção de óleo bruto e de farelo residual e a refinação dos óleos brutos produzidos.

Na produção de óleo bruto e dos resíduos industrializados, caso específico da empresa analisada, três etapas do processo global devem ser concretizadas: armazenamento das sementes oleaginosas; preparação da matéria-prima e extração do óleo bruto. A empresa

trabalha no segmento de produção de óleo bruto, gerando o óleo de soja degomado pronto para refino, além de outros óleos provenientes de diferentes matérias-primas. Atualmente, o processo de extração do óleo vegetal da soja opera com capacidade de processamento entre 1.500 a 2.000 toneladas/dia.

A partir dessas atribuições sobre a importância da soja, esta pesquisa *ex-post-facto* refere-se apenas a matéria-prima soja, mais especificadamente ao óleo de soja bruto degomado, visto que este é o principal segmento da empresa. Da industrialização da soja, tem-se o óleo de soja degomado, destinado ao consumo humano e animal; o farelo de soja, hoje uma das principais fontes de proteína na alimentação animal, e, a casca de soja, também utilizada na alimentação animal.

No contato mantido com a equipe de qualidade da empresa, procurou-se, inicialmente, conhecer o processo de industrialização da soja a partir do qual é extraído o óleo de soja bruto degomado. Diariamente são realizados ensaios de laboratório, nos quais se medem várias características físico-químicas presentes na soja. Essas características formam as séries dos ensaios de laboratório que diagnosticam as propriedades existentes no grão de soja para dar seguimento ao processo de industrialização do produto.

Entre as propriedades químicas dos lipídios, encontram-se o índice de acidez, o índice de saponificação e o índice de iodo. Sendo assim, optou-se por fazer um estudo do índice de acidez (%), que é medido diariamente na fábrica pela importância que ele exerce na qualidade do produto final.

Segundo Morreto e Fett (1998), o índice de acidez é o número de miligramas de hidróxido de potássio (KOH) necessários para neutralizar os ácidos livres de um grama de gordura. Em síntese, o índice de acidez está intimamente relacionado à natureza e à qualidade da matéria-prima e, principalmente, às condições de conservação de gordura, revelando, assim, o estado de conservação do óleo.

A coleta de dados foi feita mediante autorização do diretor industrial da empresa e realizada diretamente na empresa. Os dados sobre o índice de acidez abrangem o período de 01 de abril de 2004 a 30 de junho de 2009, totalizando 960 observações, considerando os dias úteis de produção.

No processo de estimação do modelo matemático aplicou-se uma modelagem mista ARIMA-ARCH nos resíduos quadráticos resultantes do modelo ARIMA, pois seus resíduos apresentavam heterocedasticidade. O modelo foi estimado com 954 observações. As seis últimas observações foram usadas para avaliar as previsões por meio do MAPE e U de Theil.

4.2.2 Análise descritiva

Após a aquisição dos dados, é preciso conhecer o comportamento das variáveis no decorrer do período analisado. Na Tabela 6, apresentam-se algumas estatísticas descritivas do índice de acidez, cuja unidade de medida é a percentagem.

Tabela 6 – Estatísticas descritivas da variável índice de acidez da Empresa B no período de 1º/04/2004 a 30/06/2009

Estatísticas	Índice de acidez
Média	0,558
Mediana	0,510
Desvio padrão	0,239
Coefficiente de variação (%)	42,841
Assimetria	2,329
Curtose	14,126
Mínimo	0,130
Máximo	2,980

Com a análise descritiva das variáveis, verificou-se, pelo coeficiente de variação, que a média da variável é considerada pouco representativa, visto que esse valor expresso em percentual está muito próximo de 50%. Porém, o valor da amplitude 2,85 é relativamente alto, indicando que há uma grande dispersão do índice de acidez no decorrer do período. Percebe-se, também, que o índice de acidez possui uma distribuição assimétrica positiva, uma vez que a mediana mostra-se menor que a média.

Ao analisar o valor da assimetria e da curtose, observou-se o seguinte: como a assimetria é diferente de zero e a curtose é maior que três, é um indicativo de que a série exibe sinais de volatilidade e de *clusters* de volatilidade, conforme Campos (2007). Segundo Soares e Castelar (2003), se a curtose excede três, a distribuição é dita leptocúrtica.

A Figura 27 mostra o comportamento da série original do índice de acidez. Avaliando essa figura, observa-se que a série possui uma grande variabilidade, e, aparentemente não apresenta um comportamento estacionário em relação à média. Além disso, constata-se a presença de *outliers*, sendo que, o mais acentuado ocorre na observação 796 referente ao dia 14/08/08.

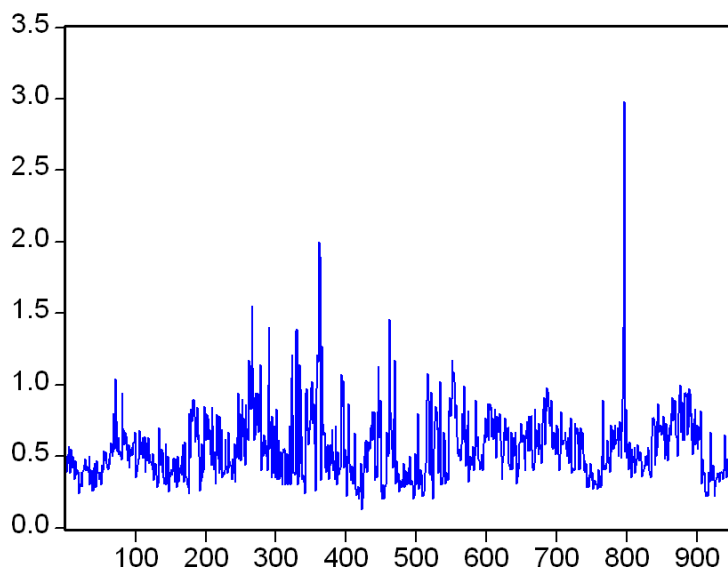


Figura 27 – Série original da variável índice de acidez

A verificação da estacionariedade foi realizada através da construção de correlogramas, apresentados na Figura 28, também conhecidos por função de autocorrelação (FAC) e função de autocorrelação parcial (FACP), em conjunto com as estatísticas *Ljung-Box* e, também pelos testes de raiz unitária.

O gráfico da Figura 28 apresenta as estatísticas *Ljung-Box*, que evidenciam a presença de autocorrelação para os primeiros 15 *lags*, pois os *p-valores* são menores que o nível de significância adotado. Sendo assim, pode-se afirmar que a série não é estacionária.

Autocorrelação	Autocorrelação Parcial	FAC	FACP	Q(K)	p-valor
0.565	0.565	0.565	0.565	305.75	0.000
0.405	0.126	0.405	0.126	462.81	0.000
0.331	0.089	0.331	0.089	567.81	0.000
0.277	0.050	0.277	0.050	641.24	0.000
0.247	0.052	0.247	0.052	699.79	0.000
0.249	0.078	0.249	0.078	759.58	0.000
0.204	-0.005	0.204	-0.005	799.70	0.000
0.188	0.030	0.188	0.030	833.89	0.000
0.161	0.000	0.161	0.000	858.84	0.000
0.151	0.022	0.151	0.022	880.75	0.000
0.171	0.060	0.171	0.060	908.89	0.000
0.170	0.027	0.170	0.027	937.01	0.000
0.154	0.009	0.154	0.009	960.02	0.000
0.145	0.012	0.145	0.012	980.38	0.000
0.131	0.005	0.131	0.005	996.94	0.000

Figura 28 – Correlograma, estatística de Ljung-Box - Q(K) e p-valor da série original para a variável índice de acidez

Para comprovar estatisticamente a estacionariedade ou não-estacionariedade da série índice de acidez, aplicaram-se os testes de raiz unitária, PP e KPSS, conjuntamente, conforme Bayer (2008). O resultados desses testes são mostrados na Tabela 7.

O teste PP tem as seguintes hipóteses: H_0 : **a série possui raiz unitária** e H_1 : **a série não possui raiz unitária**, e o teste KPSS tem as seguintes hipóteses: H_0 : **a série é estacionária** e H_1 : **a série não é estacionária**.

Tabela 7 – Testes de raiz unitária PP e KPSS para a variável índice de acidez

PP Estatística do teste	Valor crítico do teste PP*	KPSS Estatística do teste	Valor crítico do KPSS*
-18,999	-2,864	0,176	0,463

* O valor crítico considerado é de 5%

Conforme a Tabela 7, o teste PP prova que, se a estatística calculada é menor que o valor crítico, rejeita-se a hipótese nula, concluindo-se pela estacionariedade. Quanto ao teste KPSS, constata-se que, como a estatística calculada é menor que o valor crítico, não há rejeição da hipótese nula de estacionariedade. Assim, os testes, PP e KPSS, asseguraram que o índice de acidez é estacionário. Segundo Baillie, Chung e Tieslau (1996), esse resultado de conjugação de testes indica uma forte evidência de que o processo é estacionário.

4.2.3 Estimação de modelos ARIMA concorrentes e escolha do melhor modelo

Comprovou-se anteriormente que os dados originais do índice de acidez são autocorrelacionados, logo, deve-se buscar um modelo matemático capaz de eliminar a autocorrelação serial e, a partir de seus resíduos, traçam-se os gráficos de controle \bar{x} e \overline{MR} . Desse modo, uma das alternativas iniciais é a modelagem por meio dos modelos lineares ARIMA(p,d,q), que modelam a média condicional.

O correlograma da Figura 28 sugere que existe uma variedade de modelos ARIMA candidatos para ajustar a série índice de acidez. Esses modelos concorrentes foram testados e avaliados segundo a quantidade de parâmetros significativos e as estatísticas AIC e SBC. Salienta-se que todos os modelos ARIMA concorrentes, apresentados na Tabela 8, formam um ruído branco e são significativos no nível de significância adotado.

Tabela 8 – Estimação dos parâmetros dos modelos ARIMA concorrentes, dos critérios AIC e SBC para a variável índice de acidez

Modelos	Parâmetros	AIC	SBC
ARMA(1,1)	p(1)= 0,989316 q(1)= -0,578358	-0,353	-0,342
ARMA(1,2)	p(1)= 0,996624 q(1)= -0,524071 q(2)= -0,238767	-0,399	-0,384
ARMA(2,2)	p(1)= 0,688061 p(2)= -0,688118 q(1)= -0,214189 q(2)= 0,233632	-0,416	-0,396
ARMA(2,1)	p(1)= 0,410792 p(2)= -0,411677 q(1)= -0,907670	-0,415	-0,399
ARMA(1,3)	p(1)= 0,997771 q(1)= -0,515629 q(2)= -0,188766 q(3)= -0,103364	-0,408	-0,387
ARMA(2,3)	p(1)= 0,862792 p(2)= -0,862777 q(1)= -0,393415 q(2)= 0,269900 q(3)= 0,124744	-0,424	-0,399

Mediante os modelos concorrentes obtidos, o melhor modelo ajustado para o índice de acidez é o ARMA(2,1). Esse modelo foi determinado pelos critérios penalizadores AIC e SBC.

4.2.4 Verificação da presença de heterocedasticidade condicional

Tendo encontrado o modelo matemático que eliminou a autocorrelação serial da série, é necessário aplicar o teste ARCH-LM nos resíduos quadráticos desse modelo com a finalidade de verificar se apresentam heterocedasticidade condicional.

Antes de aplicar o teste, analisaram-se as principais estruturas estatísticas dos resíduos, como: estatísticas descritivas, normalidade, aglomerados de volatilidade e *outliers*.

A Figura 29 mostra a distribuição dos resíduos do modelo ARMA(2,1) e algumas de suas estatísticas descritivas.

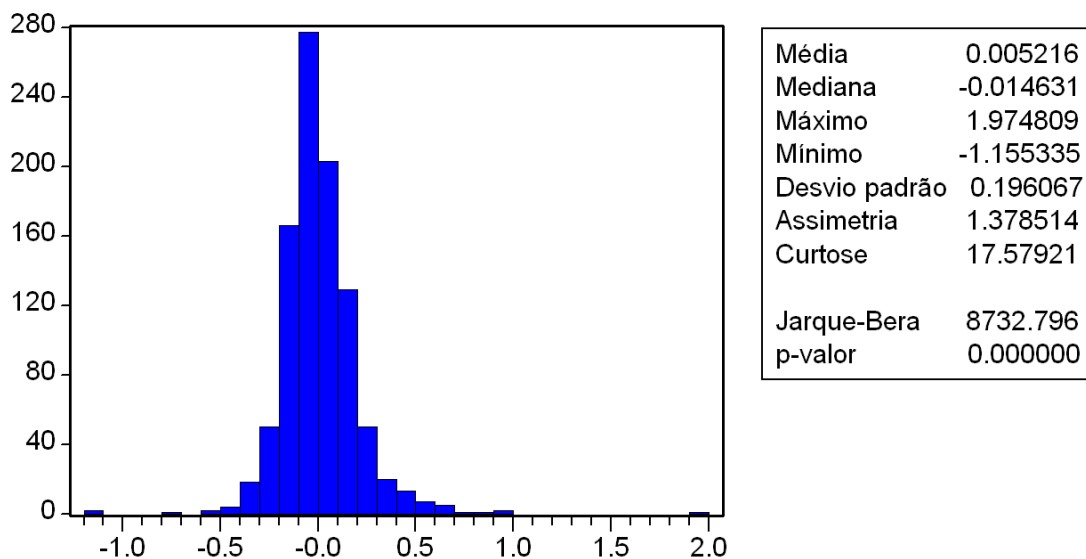


Figura 29 – Distribuição e estatísticas descritivas dos resíduos do modelo ARMA(2,1) para a variável índice de acidez

Pela inspeção visual da Figura 29, constata-se que há indícios de que a série residual não apresenta distribuição normal, visto que sua curtose é 17,57921, assumindo uma distribuição leptocúrtica. Para comprovar essa hipótese, foram feitos dois testes de normalidade: Jarque-Bera e Q-Q plot.

As hipóteses testadas pela estatística Jarque-Bera são:

H_0 : Assimetria = zero e curtose = a três (normalidade);

H_1 : Assimetria \neq zero e curtose \neq três (não-normalidade).

Observando o valor da estatística Jarque-Bera e seu respectivo *p-valor*, conclui-se que a estatística Jarque-Bera rejeita a hipótese nula de normalidade para a série residual do índice de acidez. Esse resultado condiz com o que foi indicado pela curtose da série.

O gráfico Q-Q plot exibido na Figura 30 compara os quantis teóricos da distribuição normal com os quantis dos resíduos do modelo. Observando este gráfico, verifica-se que existe uma relação não-linear entre estes quantis, indicando a presença de caudas mais pesadas.

Analisando os resultados do teste Jarque-Bera e do gráfico Q-Q plot, pode-se assegurar que a série residual rejeita fortemente a hipótese nula de normalidade, ou seja, os resíduos não são normais.

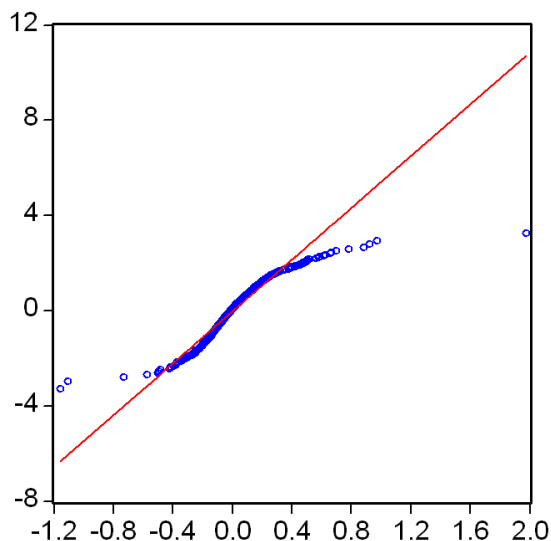


Figura 30 – Gráfico Q-Q plot dos resíduos do modelo ARMA(2,1)

A Figura 31 ilustra o comportamento temporal dos resíduos do modelo ARMA(2,1) para o índice de acidez.

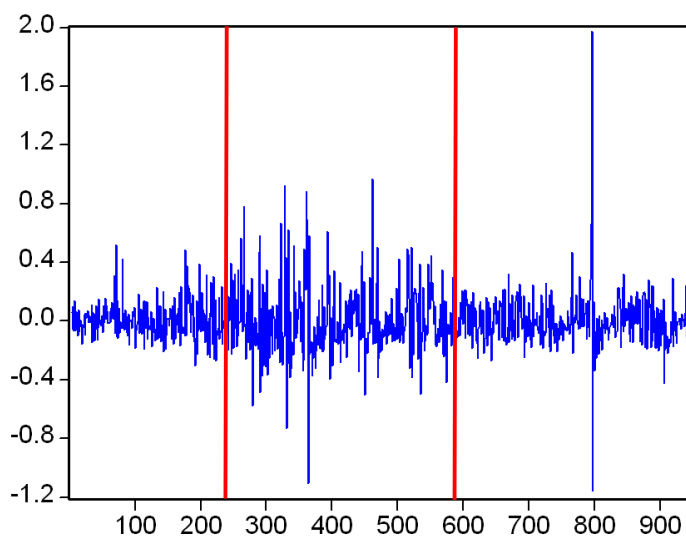


Figura 31 – Série residual do modelo ARMA(2,1)

A partir de uma simples inspeção do gráfico acima, determina-se que o processo é estacionário, mas apresenta *clusters* de volatilidade, localizados entre as linhas de cor vermelha. Em certos períodos, há altas volatilidades, enquanto que, nos demais constam, comparativamente, baixas volatilidades – períodos de maior equilíbrio. Esse fato comprova, claramente, a existência de dependência na variância condicional da série, conforme a Figura

32. Além disso, verifica-se a presença de *outliers*, um mais extremo em 08/11/2005 (observação 362), e outro, mais extremo ainda, em 14/08/2008 (observação 796).

Autocorrelação	Autocorrelação Parcial	FAC	FACP	Q(K)	p-valor	
		1	0.353	0.353	118.90	0.000
		2	0.084	-0.046	125.64	0.000
		3	0.090	0.087	133.43	0.000
		4	0.043	-0.017	135.18	0.000
		5	0.019	0.010	135.54	0.000
		6	0.052	0.044	138.17	0.000
		7	0.022	-0.016	138.63	0.000
		8	0.022	0.022	139.11	0.000
		9	0.016	-0.006	139.34	0.000
		10	0.004	-0.001	139.36	0.000
		11	0.009	0.007	139.43	0.000
		12	0.012	0.003	139.56	0.000
		13	0.003	-0.002	139.57	0.000
		14	0.001	-0.001	139.57	0.000
		15	-0.018	-0.023	139.90	0.000

Figura 32 – Correlograma, estatística de Ljung-Box (Q(K)) e p-valor da série dos resíduos quadráticos do modelo ARMA(2,1) para a variável índice de acidez

Para testar estatisticamente a presença de heterocedasticidade condicional, aplicou-se o teste de hipótese ARCH-LM nos resíduos quadráticos do modelo ARMA(2,1). Como foi constatada a presença desse fenômeno, justifica-se a utilização de um modelo não-linear da família ARCH para modelar a volatilidade. Encontra-se, na Tabela 9, o resultado desse teste.

Para comprovar a presença de efeitos ARCH, foram testadas as seguintes hipóteses:

H_0 : Não existem efeitos ARCH (homocedasticidade);

H_1 : Existem efeitos ARCH (heterocedasticidade).

Tabela 9 – Resultados das estatísticas F e ARCH-LM até a defasagem 13

Lags	Estatística F	p-valor (F)	Estatística ARCH-LM	p-valor (ARCH-LM)
Lag 1	135,023	0,000	118,454	0,000
Lag 3	48,303	0,000	126,175	0,000
Lag 5	28,932	0,000	126,184	0,000
Lag 7	20,917	0,000	127,714	0,000
Lag 9	16,258	0,000	127,838	0,000
Lag 11	13,244	0,000	127,562	0,000
Lag 13	11,155	0,000	127,260	0,000

A partir dos resultados encontrados para as diferentes defasagens, verifica-se que todas as estatísticas F e as estatísticas ARCH-LM foram significativas. Assim, rejeita-se a hipótese

nula, implicando que existe heterocedasticidade condicional nos resíduos quadráticos do modelo ARMA(2,1).

Mediante os resultados obtidos no teste de hipótese ARCH-LM, constata-se a necessidade de modelar a volatilidade presente no índice de acidez.

4.2.5 Estimação do melhor modelo misto ARMA-ARCH

Como foi estimado um modelo matemático por meio de um modelo ARMA, será empregado um modelo misto denominado de ARMA-ARCH, visto que os resíduos do modelo ARMA apresentaram heterocedasticidade. Assim, justifica-se a utilização de uma estimação mista dos parâmetros do modelo linear (ARMA) e não-linear (ARCH).

A Tabela 10 apresenta o modelo misto ARMA-ARCH para a variável índice de acidez, pelo qual foram estimadas equações para a média e variância condicionais.

Tabela 10 – Estimação dos coeficientes, erro-padrão, estatística z e p-valor do modelo ARMA-ARCH para a variável índice de acidez

Método: ML – ARCH (Marquardt) - Distribuição Normal				
Equação para a média condicional				
	Coefficiente	Erro-Padrão	Estatística z	p-valor
AR(1)	0,257	0,098	12,857	0,000
AR(2)	-0,283	0,093	-3,049	0,002
MA(1)	-0,608	0,065	-9,348	0,000
Equação para a variância condicional				
C	0,037	0,002	18,714	0,000
ARCH(1)	0,199	0,016	12,158	0,000

O modelo para a volatilidade é descrito como um ARMA(2,1)-ARCH(1) apresentando parâmetros estatisticamente significativos. Constata-se que a soma dos parâmetros é menor que um, assegurando que a variância condicional seja positiva e estacionária. Como o valor de alfa, parâmetro do ARCH(1), é razoavelmente baixo, igual a 0,199, significa que, após um curto período de tempo, a variância da série deverá convergir à média histórica, ou seja, o índice de acidez no período anterior não influencia fortemente no próximo período.

Para fazer o diagnóstico do modelo encontrado, analisam-se os resíduos padronizados. É importante verificar a distribuição dos resíduos padronizados para constatar se existe heterocedasticidade condicional nos resíduos remanescentes do modelo, apresentada na Figura 33.

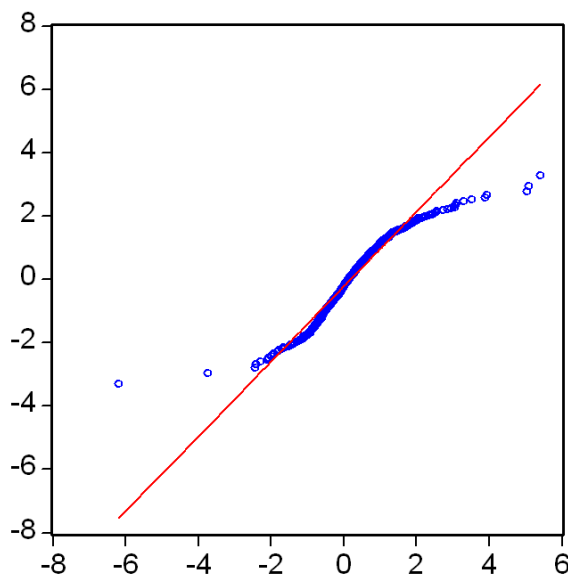


Figura 33 – Gráfico Q-Q plot dos resíduos padronizados estimados pelo modelo ARMA(2,1)-ARCH(1) para a variável índice de acidez

Pelo gráfico Q-Q plot, constata-se que os resíduos padronizados são aproximadamente normais, apresentando uma cauda mais pesada que a normal padronizada, o que não afeta a análise futura de estimação do modelo ARCH.

Para verificar a possível existência de alguma heterocedasticidade condicional nos resíduos do modelo misto, realizou-se, além do gráfico Q-Q plot, o teste ARCH-LM, o qual aceitou a hipótese nula de ausência de efeitos ARCH.

A Figura 34 mostra os valores originais, os valores estimados e os resíduos para a série índice de acidez, e, a Tabela 11 apresenta o período, os valores originais e previstos, o erro de previsão, e, os critérios, MAPE e U de Theil, considerando o modelo misto estimado.

Conforme mencionado anteriormente, tanto o modelo ARMA(2,1) como o modelo misto ARMA(2,1)-ARCH(1) foram estimados com 954 observações. As últimas seis ficaram de fora da amostra com o objetivo de avaliar o modelo matemático por meio do MAPE e do U de Theil, também expostos na Tabela 11. Sendo assim, a previsão ocorre da observação 955 à observação 960.

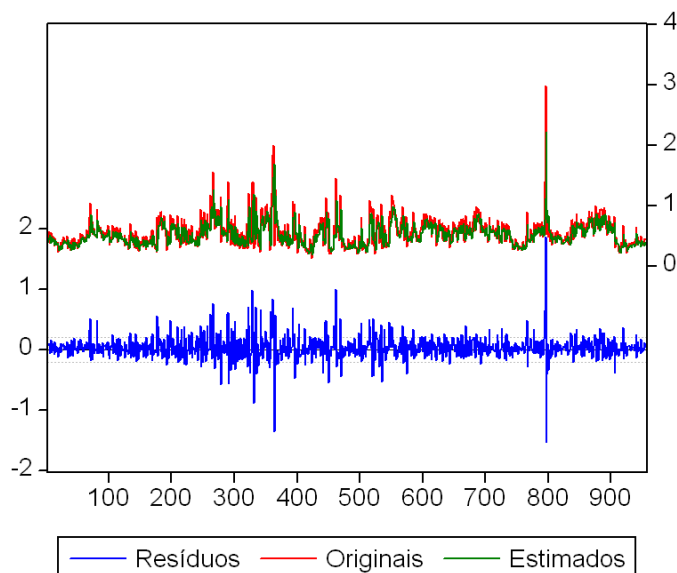


Figura 34 – Valores originais, valores estimados e os resíduos do modelo ARMA(2,1)-ARCH(1)

Tabela 11 – Período, valores originais, valores previstos, erro de previsão, MAPE e U de Theil para a variável índice de acidez da Empresa B

Período	Originais	Previstos	Erro de previsão
19/6/2009	0,530	0,431	0,099
22/6/2009	0,450	0,374	0,076
23/6/2009	0,370	0,370	0,000
24/6/2009	0,500	0,374	0,126
29/6/2009	0,610	0,341	0,269
30/6/2009	0,440	0,400	0,040
MAPE= 23,866 e U de Theil= 0,175			

Analisando o MAPE e o U de Theil, comprova-se estatisticamente que o modelo matemático estimado apresenta uma boa previsão, se comparada com a previsão ingênua. Como o coeficiente U de Theil está mais próximo de zero, a previsão realizada é melhor que a previsão ingênua ou trivial.

A partir do modelo misto ARMA(2,1)-ARCH(1), mostra-se na Figura 35 a sua volatilidade condicional.

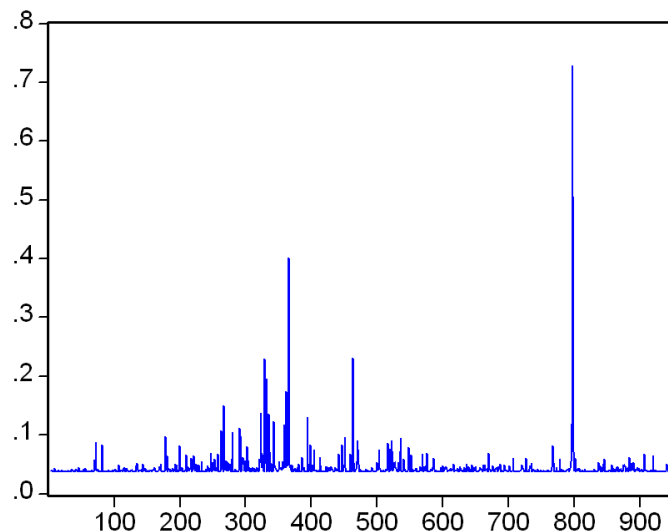


Figura 35 – Volatilidade condicional do modelo misto ARMA(2,1)-ARCH(1)

Analisando o gráfico da volatilidade condicional observa-se que o modelo matemático misto reproduziu bem os períodos marcantes dos *clusters* de volatilidade presentes na série residual do modelo ARMA(2,1), apresentado na Figura 31.

4.2.6 Análise da estabilidade do processo

Embora a análise descritiva revele algo sobre comportamento do processo, é necessário investigar sua estabilidade por meio dos gráficos de controle, os quais vão caracterizá-lo como estando sob controle ou fora de controle.

Para dar início a esta etapa, apresentam-se os gráficos de controle para medidas individuais e amplitude móvel, nas Figuras 36 e 37, usando os dados originais com base nas amostras coletadas. Nos gráficos de controle \bar{x} e \overline{MR} , foram adicionados os *runs tests*, os quais podem revelar a existência de algum tipo de padrão nos dados.

De acordo com a Figura 36, observa-se uma grande quantidade de pontos vermelhos grafados no gráfico, ocasionando uma elevada instabilidade da série índice de acidez. Essa instabilidade é explicada pelo fato de que os *runs tests* não funcionam em processos autocorrelacionados. Por isso, há uma grande quantidade de pontos vermelhos.

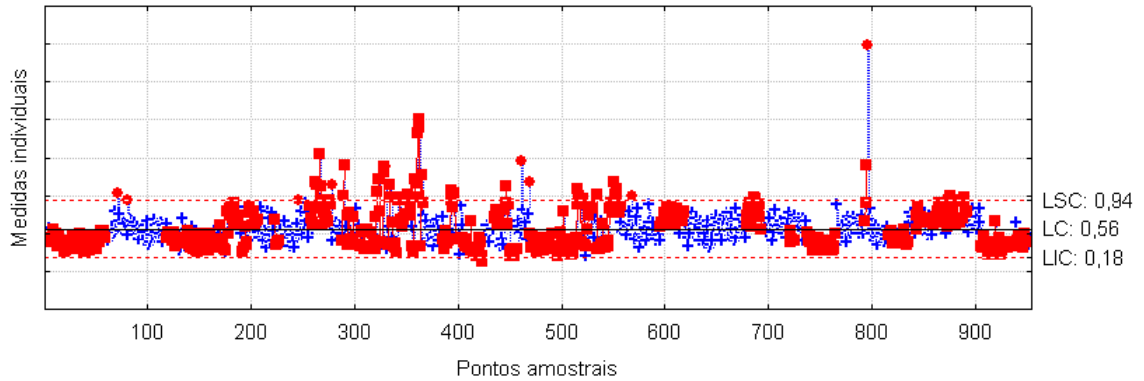


Figura 36 – Gráfico \bar{x} do índice de acidez, considerando os dados originais

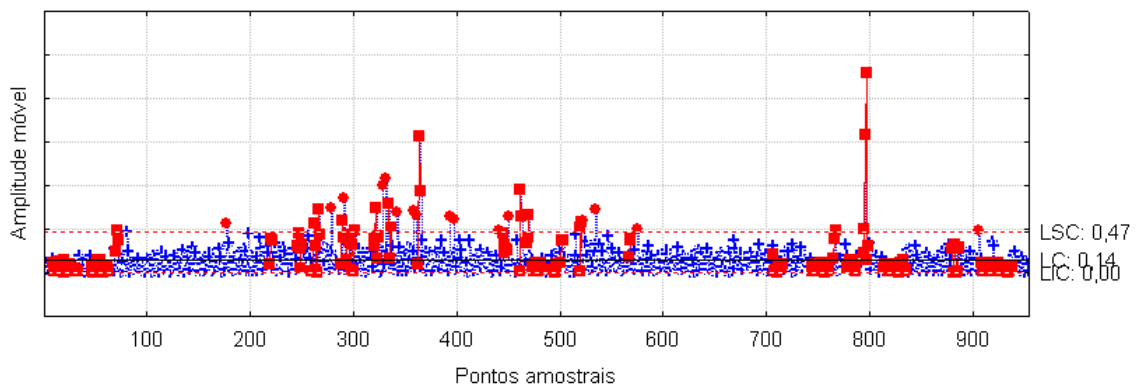


Figura 37 – Gráfico \overline{MR} do índice de acidez, considerando os dados originais

Analisando o gráfico \overline{MR} da Figura 37, também verifica-se uma grande quantidade de pontos vermelhos grafados no gráfico, ocasionando uma elevada instabilidade da variável índice de acidez.

Ao realizar outra análise simulando a exclusão dos pontos amostrais fora de controle, tem-se uma visão da instabilidade que está ocorrendo no processo. Com a exclusão dos pontos amostrais fora de controle, os limites de controle dos gráficos \bar{x} e \overline{MR} são recalculados. As Figuras 38 e 39 mostram essa simulação utilizando os *runs tests*.

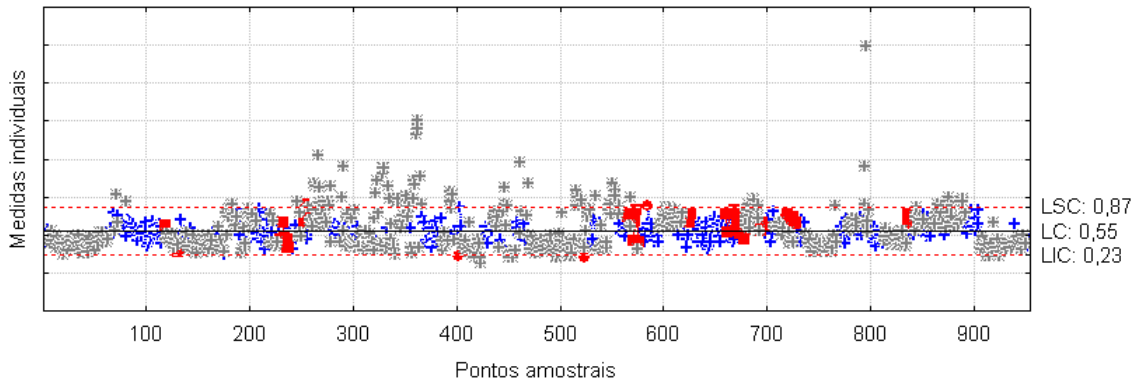


Figura 38 – Gráfico \bar{x} do índice de acidez com a exclusão dos pontos amostrais fora de controle, considerando os dados originais

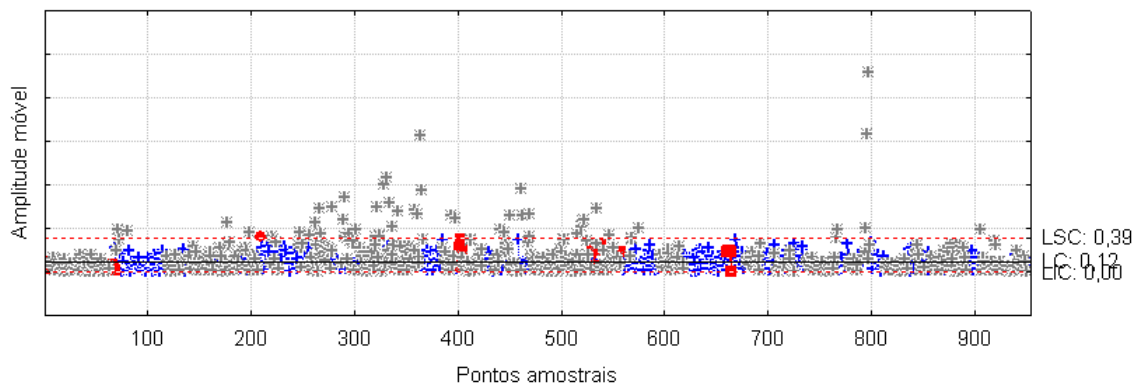


Figura 39 – Gráfico \overline{MR} do índice de acidez com a exclusão dos pontos amostrais fora de controle, considerando os dados originais

Mediante os resultados obtidos, justifica-se que as técnicas convencionais de controle estatístico de processo não podem ser utilizadas diretamente nos dados originais, ou seja, é necessário um tratamento na variável de interesse para remover a autocorrelação serial antes de monitorar a variável, que pode inserir pontos amostrais fora de controle, quando, na verdade o processo pode estar sob controle, nesse caso gerando alarmes falsos.

Para sanar o efeito da autocorrelação serial presente na série original do índice de acidez, estimou-se um modelo matemático utilizando a modelagem mista ARMA-ARCH, pois os resíduos do modelo linear ARMA(2,1) continham volatilidade. Como os resíduos são independentes e normalmente distribuídos, as técnicas do CEP podem ser utilizadas com a finalidade de identificar a ocorrência de padrões especiais, caracterizando o processo analisado quanto a sua estabilidade para que seja feito ou não um ajuste.

Nesta etapa, foram traçados os gráficos de controle \bar{x} e \overline{MR} para os resíduos do modelo ARMA(2,1), para os resíduos do modelo misto – ARMA(2,1)-ARCH(1) e, na sequência, para a volatilidade condicional do modelo misto.

- Análise dos gráficos de controle \bar{x} e \overline{MR} nos resíduos do modelo ARMA(2,1)

Os gráficos das Figuras 40 e 41 apresentam o monitoramento dos resíduos ARMA(2,1) para a média condicional, traçando os gráficos de controle \bar{x} e \overline{MR} , tendo como auxílio na decisão a aplicação dos *runs tests*.

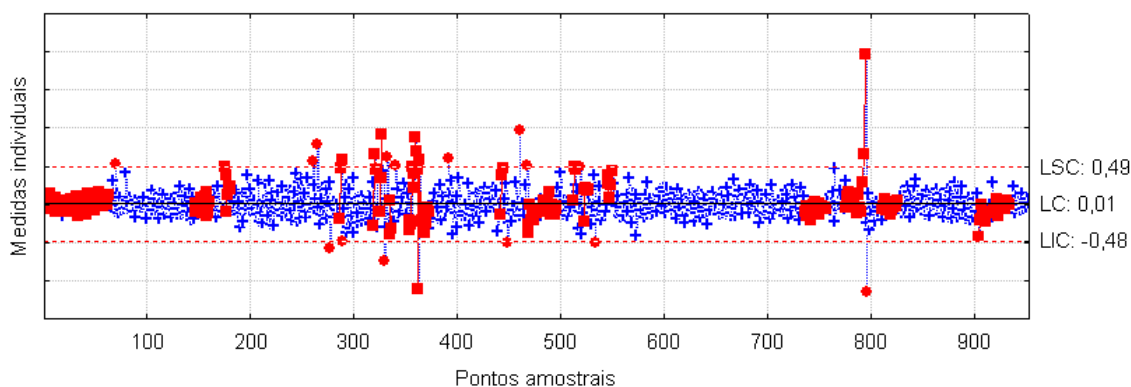


Figura 40 – Gráfico \bar{x} aplicado ao resíduo do modelo ARMA(2,1) para a variável índice de acidez

De acordo com os resíduos do modelo estimado, traçou-se o gráfico \bar{x} mostrado na Figura 40, pelo qual se detecta que **9 pontos amostrais em sequência estão no mesmo lado da linha central**, representadas pelo conjunto de amostras com início nas observações 364, 468, 736 e 904. Observa-se também que, a **cada 3 pontos amostrais traçados, 2 se encontram localizados na zona A ou fora dela**, correspondendo às observações com início nas amostras 175, 286, 319, 325, 334, 355, 358, 361, 442, 512 e 792, e um conjunto de 5 amostras em que, a **cada 5 pontos amostrais traçados, 4 se encontram localizados na zona B ou fora dela**, correspondendo às observações com início nas amostras 177, 355, 362, 522 e 545. Além disso, observam-se **15 pontos amostrais em sequência situados na zona C**, correspondendo às observações com início nas amostras 1, 19, 34, 49, 146, 482, 742, 776, 811 e 919, e, verificam-se **8 pontos amostrais em sequência situados fora da zona C**, correspondendo à observação com início na amostra 353.

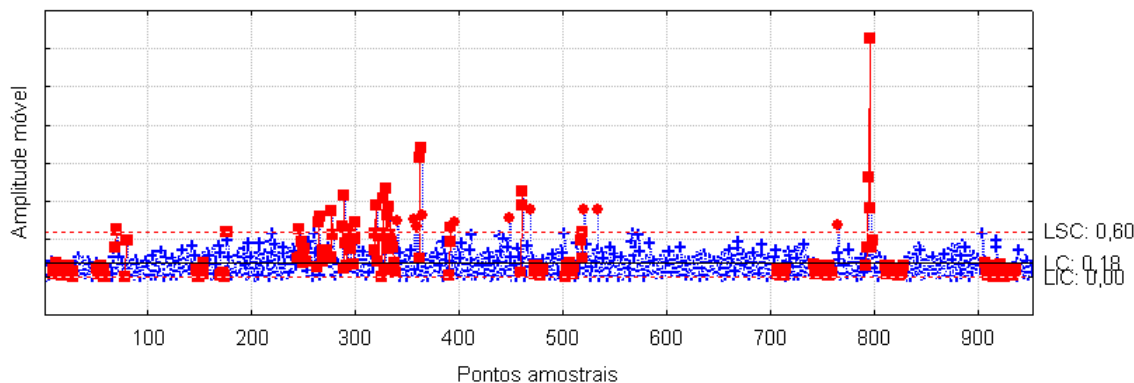


Figura 41 – Gráfico \overline{MR} aplicado ao resíduo do modelo ARMA(2,1) para a variável índice de acidez

Utilizando os resíduos do modelo estimado, traçou-se o gráfico \overline{MR} apresentado na Figura 41, pelo qual se constata que **9 pontos amostrais em seqüência estão no mesmo lado da linha central**, representadas pelo conjunto de amostras com início nas observações 8, 17, 51, 146, 244, 329, 472, 502, 706, 742, 751, 811, 820, 907, 919 e 928. Observa-se também que, a **cada 3 pontos amostrais traçados, 2 se encontram localizados na zona A ou fora dela**, correspondendo às observações com início nas amostras 68, 78, 174, 263, 276, 287, 292, 297, 318, 325, 328, 331, 361, 390, 459, 517, 793 e 796, e um conjunto de 12 amostras em que, a **cada 5 pontos amostrais traçados, 4 se encontram localizados na zona B ou fora dela**, correspondendo às observações com início nas amostras 24, 170, 268, 289, 296, 318, 326, 331, 751, 792, 819 e 924. Além disso, verificam-se **6 pontos amostrais em seqüência com tendência crescente e/ou decrescente**, começando na amostra 333 até a amostra 338, e observam-se **8 pontos amostrais em seqüência situados fora da zona C**, correspondendo à observação com início na amostra 329.

Mediante o levantamento das afirmações acima citadas, verifica-se que há poucos pontos amostrais fora de controle se comparado com o gráfico \bar{x} da série original. Isso indica que o modelo ARMA foi capaz de captar as variações que havia na série índice de acidez, pois o efeito da autocorrelação serial foi tratado.

Ao realizar outra análise simulando a exclusão dos pontos amostrais fora de controle, tem-se outra visão da instabilidade que está ocorrendo no processo. Com a exclusão dos pontos amostrais fora de controle, os limites de controle dos gráficos \bar{x} e \overline{MR} são recalculados. As Figuras 42 e 43 mostram essa simulação, aplicando os *runs tests*.

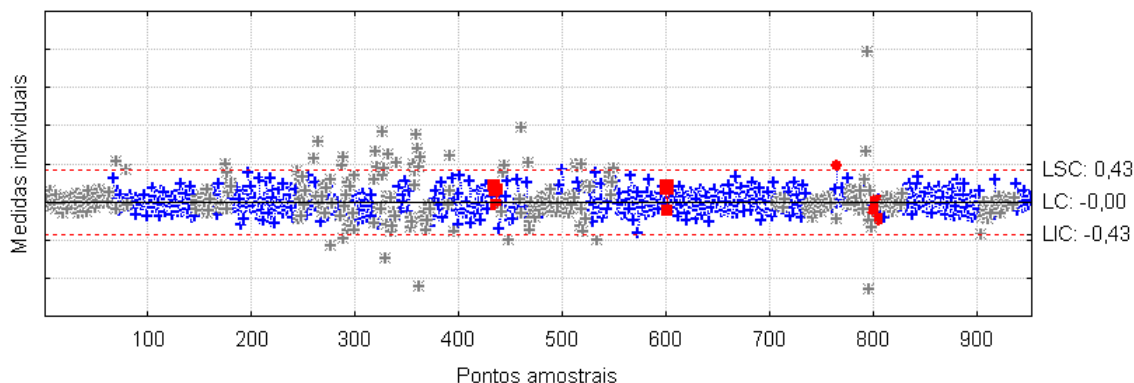


Figura 42 – Gráfico \bar{x} aplicado ao resíduo do modelo ARMA(2,1) para a variável índice de acidez com a exclusão dos pontos amostrais fora de controle

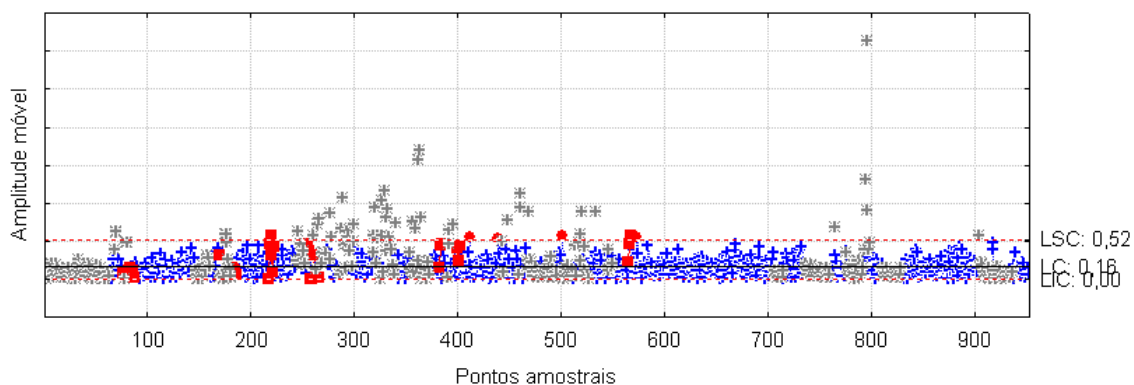


Figura 43 – Gráfico \overline{MR} aplicado ao resíduo do modelo ARMA(2,1) para a variável índice de acidez com a exclusão dos pontos amostrais fora de controle

- Análise dos gráficos de controle \bar{x} e \overline{MR} nos resíduos do modelo ARMA(2,1)-ARCH(1)

Os gráficos das Figuras 44 e 45 apresentam o monitoramento dos resíduos para o modelo misto ARMA(2,1)-ARCH(1), traçando-se os gráficos de controle \bar{x} e \overline{MR} , respectivamente, tendo como auxílio na decisão a aplicação dos *runs tests*.

De acordo com os resíduos do modelo misto estimado, traçou-se o gráfico \bar{x} mostrado na Figura 44, pelo qual se detecta que existem **6 pontos amostrais em sequência com tendência crescente e/ou decrescente**, começando na amostra 524 até a amostra 529, **14 pontos amostrais em sequência estão alternando-se para cima e para baixo**, começando na amostra 264 até a amostra 277. Observa-se, também, que a **cada 3 pontos amostrais traçados, 2 se encontram localizados na zona A ou fora dela**, correspondendo às observações com início nas amostras 286, 319, 355, 358, 442, 512, 792 e 795, e um conjunto de 2 amostras em que, a **cada 5 pontos amostrais traçados, 4 se encontram localizados na**

zona B ou fora dela, correspondendo às observações com início nas amostras 351 e 356. Além disso, verificam-se **15 pontos amostrais em sequência situados na zona C**, correspondendo às observações com início nas amostras 1, 16, 34, 49, 80, 113, 422, 472, 642, 742, 776, 802, 817 e 919.

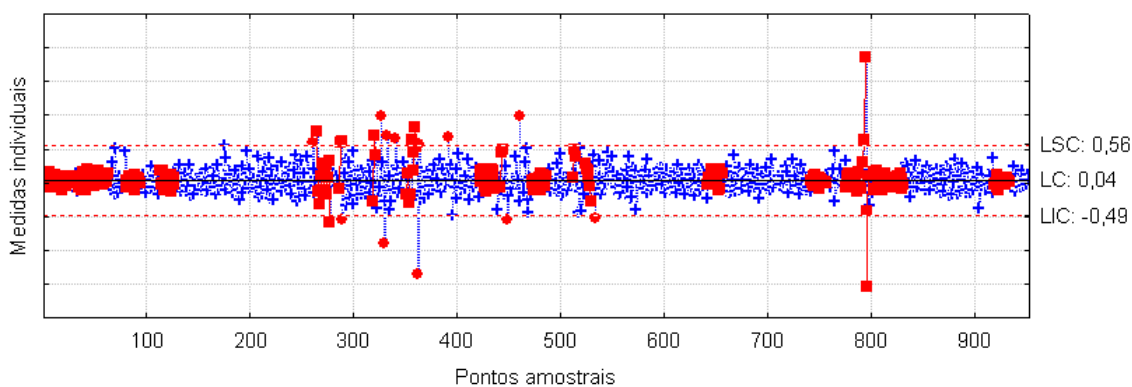


Figura 44 – Gráfico \bar{x} aplicado ao resíduo quadrático do modelo misto ARMA(2,1)-ARCH(1) para a variável índice de acidez

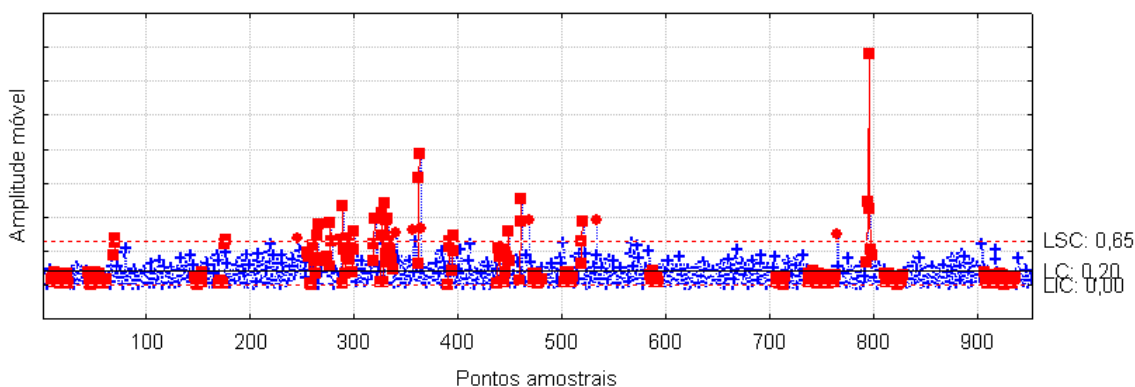


Figura 45 – Gráfico \overline{MR} aplicado ao resíduo quadrático do modelo misto ARMA(2,1)-ARCH(1) para a variável índice de acidez

Utilizando os resíduos do modelo estimado, traçou-se o gráfico \overline{MR} apresentado na Figura 45, pelo qual se constata que **9 pontos amostrais em sequência estão no mesmo lado da linha central**, representadas pelo conjunto de amostras com início nas observações 8, 17, 44, 53, 146, 329, 472, 502, 585, 706, 737, 746, 755, 811, 820, 907, 919 e 928. Observa-se também que, a **cada 3 pontos amostrais traçados, 2 se encontram localizados na zona A ou fora dela**, correspondendo às observações com início nas amostras 68, 174, 263, 276, 287, 292, 297, 318, 325, 328, 331, 361, 390, 394, 459, 518, 793 e 796, e um conjunto de 12

amostras em que, a **cada 5 pontos amostrais traçados, 4 se encontram localizados na zona B ou fora dela**, correspondendo às observações com início nas amostras 170, 260, 270, 289, 296, 326, 331, 445, 751, 793, 822 e 923. Além disso, verificam-se **8 pontos amostrais em sequência situados fora da zona C**, correspondendo às observações com início nas amostras 255 e 458.

Mediante o levantamento das afirmações acima citadas, verifica-se que há poucos pontos amostrais fora de controle se comparado ao gráfico \bar{x} da série original e ao gráfico do modelo ARMA(2,1), exposto anteriormente. Isso indica que o modelo misto ARMA(2,1)-ARCH(1) foi mais eficaz que o ARMA(2,1) em captar as variações que havia na série original.

Ao realizar outra análise simulando a exclusão dos pontos amostrais fora de controle do modelo misto ARMA(2,1)-ARCH(1), tem-se outra visão da instabilidade que está ocorrendo no processo. Com a exclusão dos pontos amostrais fora de controle, os limites de controle dos gráficos \bar{x} e \overline{MR} são recalculados. As Figuras 46 e 47 apresentam essa simulação, aplicando os *runs tests*.

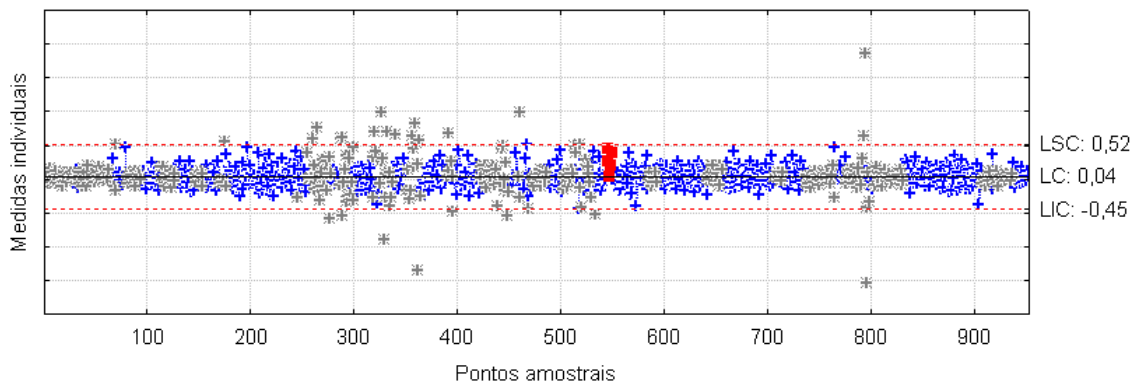


Figura 46 – Gráfico \bar{x} aplicado ao resíduo quadrático do modelo misto ARMA(2,1)-ARCH(1) para a variável índice de acidez com a exclusão dos pontos amostrais fora de controle

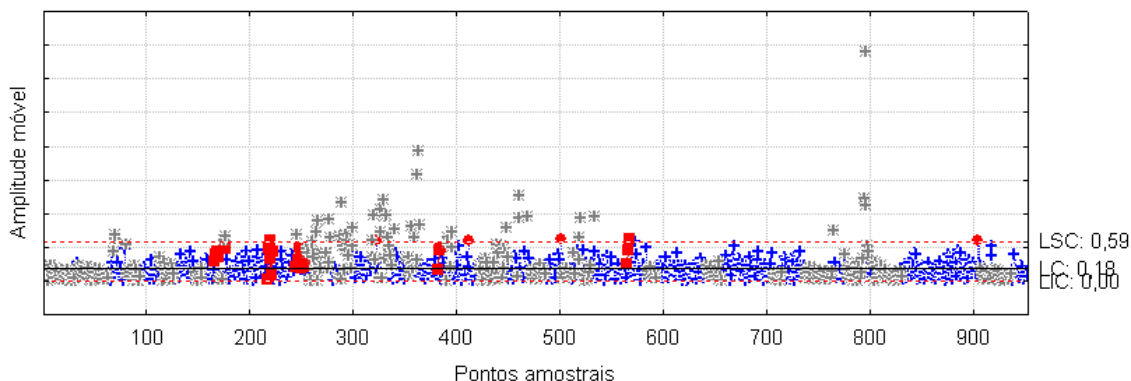


Figura 47 – Gráfico \overline{MR} aplicado ao resíduo quadrático do modelo misto ARMA(2,1)-ARCH(1) para a variável índice de acidez com a exclusão dos pontos amostrais fora de controle

- Análise dos gráficos de controle \bar{x} e \overline{MR} na volatilidade

No gráfico das Figuras 44 e 45 aplicado aos resíduos do modelo misto ARMA(2,1)-ARCH(1), consegue-se visualizar, conjuntamente, o comportamento da média e da variância condicional estimada. Portanto, traçou-se um gráfico \bar{x} e \overline{MR} para mostrar os *clusters* de volatilidade e justificar por que o gráfico do modelo misto é melhor, pois ele capta e elimina a volatilidade. Assim como, a modelagem ARIMA capta e elimina a autocorrelação serial dos resíduos.

Os gráficos das Figuras 48 e 49 apresentam o monitoramento da volatilidade procedente do modelo misto ARMA(2,1)-ARCH(1). Nessa análise, foram traçados os gráficos de controle \bar{x} e \overline{MR} , sem a aplicação dos *runs tests*.

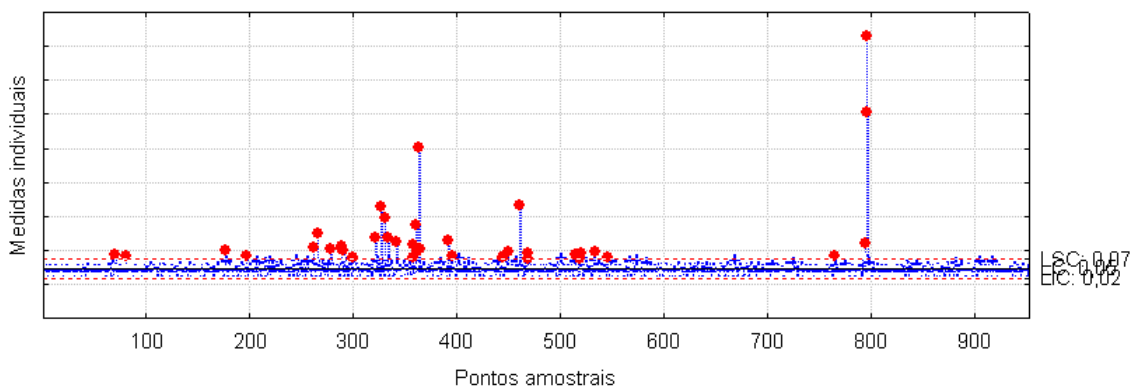


Figura 48 – Gráfico \bar{x} aplicado a volatilidade do modelo misto ARMA(2,1)-ARCH(1) para a variável índice de acidez

De acordo com a volatilidade do modelo misto ARMA(2,1)-ARCH(1), traçou-se o gráfico \bar{x} , mostrado na Figura 48, pelo qual se detecta que existem 39 pontos amostrais fora dos limites de controle. Ao comparar esse gráfico com o gráfico construído a partir dos resíduos do modelo misto, exposto na Figura 44, verifica-se que o gráfico \bar{x} para a volatilidade apresentou mais pontos amostrais fora de controle do que o gráfico para modelo misto. Além disso, observando o gráfico \bar{x} , percebeu-se que o mesmo representa bem os períodos marcantes dos *clusters* de volatilidade, representados na Figura 31.

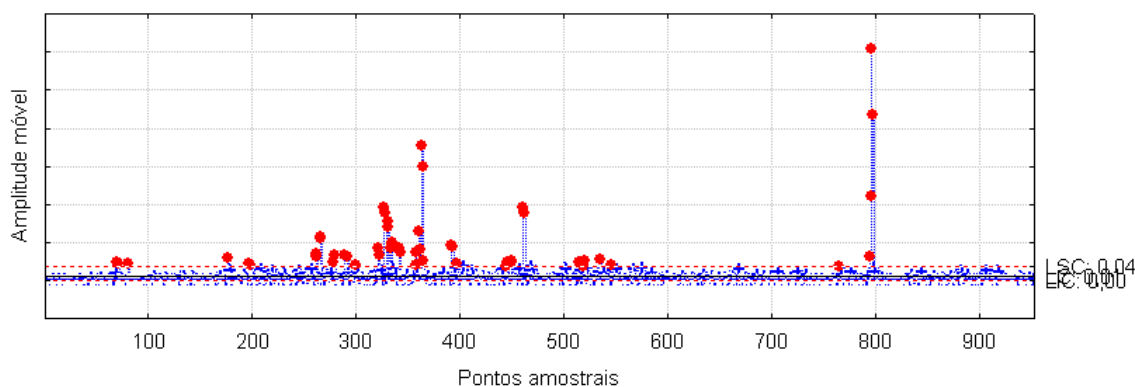


Figura 49 – Gráfico \overline{MR} aplicado a volatilidade do modelo misto ARMA(2,1)-ARCH(1) para a variável índice de acidez

Utilizando os resíduos do modelo estimado, traçou-se o gráfico \overline{MR} , apresentado na Figura 49, pelo qual se constata que existem 55 pontos amostrais fora dos limites de controle.

Mediante o levantamento das afirmações acima citadas, verifica-se que a volatilidade influencia na instabilidade tanto da média como da amplitude móvel do processo, pois os instantes em que foi detectada falta de controle também correspondem a um período em que há *clusters* de volatilidade. Isso indica que o modelo misto ARMA(2,1)-ARCH(1) foi eficaz no tratamento da autocorrelação do processo, pois o gráfico aplicado na volatilidade do modelo misto mostra que o processo apresenta uma variação que deveria ser levada em consideração no momento de analisar o processo.

Ao realizar outra análise simulando a exclusão dos pontos amostrais fora de controle referente a volatilidade resultante do modelo misto, tem-se outra visão da instabilidade que está ocorrendo no processo. Com a exclusão dos pontos amostrais fora de controle, os limites de controle dos gráficos \bar{x} e \overline{MR} são recalculados. As Figuras 50 e 51 apresentam essa simulação.

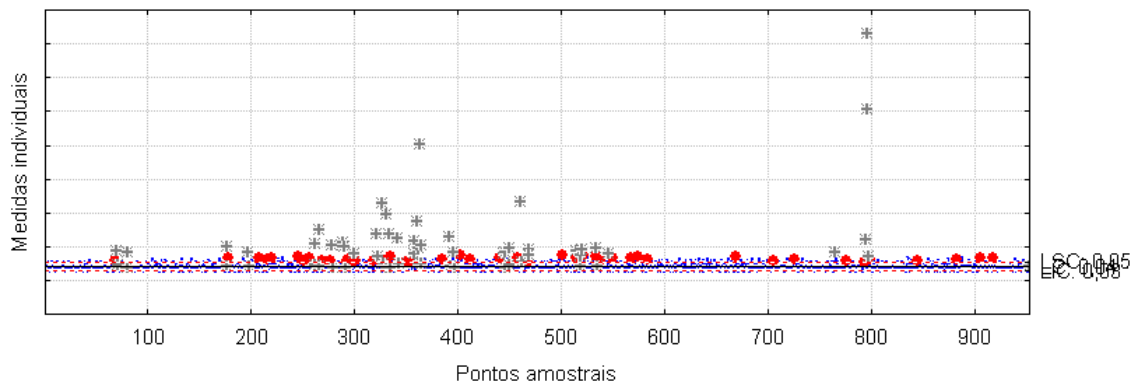


Figura 50 – Gráfico \bar{x} aplicado a volatilidade do modelo misto ARMA(2,1)-ARCH(1) para a variável índice de acidez com a exclusão dos pontos amostrais fora de controle

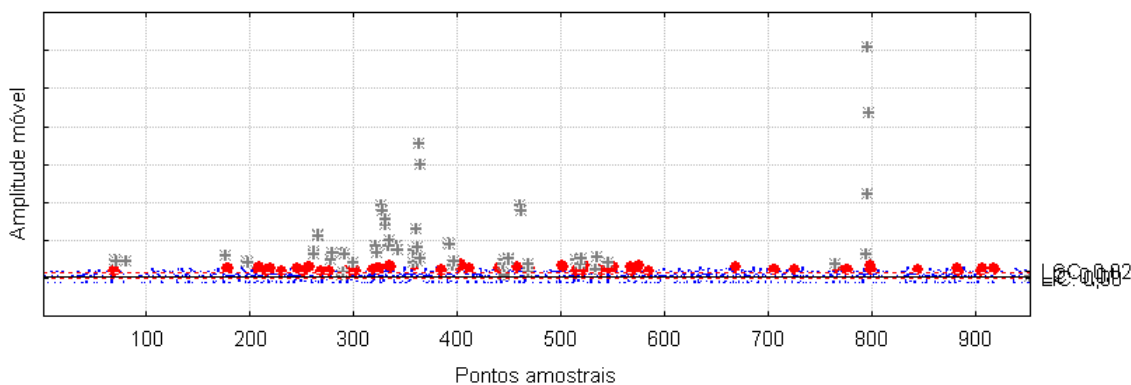


Figura 51 – Gráfico \overline{MR} aplicado a volatilidade do modelo misto ARMA(2,1)-ARCH(1) para a variável índice de acidez com a exclusão dos pontos amostrais fora de controle

4.2.7 Comentários gerais da análise

Com a aplicação da metodologia proposta, foi possível monitorar o comportamento da variável índice de acidez, que compõe uma das características críticas no processo de produção do óleo de soja bruto degomado, utilizando um modelo misto ARMA-ARCH, o qual propiciou uma estimação conjunta para a média e variância condicionais.

Com relação aos resultados obtidos após a aplicação dos gráficos de controle \bar{x} , apresentam-se as seguintes considerações sobre o índice de acidez:

- Nos dados originais da variável, havia cento e vinte e oito pontos amostrais fora dos limites de controle.

- Nos resíduos do modelo ARMA(2,1), constataram-se vinte e dois pontos amostrais fora dos limites de controle.
- Nos resíduos do modelo misto ARMA(2,1)-ARCH(1), constataram-se dezesseis pontos amostrais fora dos limites de controle.
- E, na volatilidade oriunda do modelo misto, verificaram-se trinta e nove pontos amostrais fora dos limites de controle. Isso mostra que o processo apresenta uma variação que deveria ser levada em consideração no momento de sua análise. Não foi possível detectar nesse estudo a quantificação da influência da volatilidade no processo.

As afirmações mencionadas acima revelaram que, após o ajuste dos dados por meio do modelo misto, apenas o efeito da autocorrelação foi eliminado, porém necessita-se quantificar a influência da volatilidade no processo. Dessa forma, justifica-se a acurácia do modelo misto, o qual evidenciou uma melhora significativa na detecção de pontos amostrais fora dos limites de controle em comparação com o modelo ARMA(2,1).

4.3 Empresa C

A empresa C atua no segmento de bebidas e possui franquia da Coca-Cola no Rio Grande do Sul e em Santa Catarina. Atualmente é a quarta maior fabricante do Sistema Coca-Cola no Brasil, com participação de 10% do volume total de vendas no país. No território gaúcho e catarinense é líder absoluta do mercado de refrigerantes com 54,9% de participação de mercado⁹.

A história do grupo da empresa começou no ramo de refrigerantes na década de 1940, com a produção da Laranjinha, seu primeiro refrigerante. Nos primeiros seis meses, o refrigerante era fabricado em Lajeado, mas, em 1949, foi construída uma fábrica em Porto Alegre que, então, passou a produzi-lo. Quatro anos mais tarde, a grande procura pelo produto levou à construção de uma fábrica em Santo Ângelo. Além da Laranjinha, essa unidade também distribuía cerveja e outras bebidas. Em 1956, a fábrica de Santo Ângelo iniciou a produção de Grapette e a distribuição de Coca-Cola na região.

⁹ Segundo dados do Instituto Nielsen divulgados em março de 2009.

A presente pesquisa ex-post-facto refere-se a uma das três fábricas da franquia. Além de produzir refrigerantes da linha Coca-Cola, também realiza a referida distribuição dos produtos.

Ressalta-se que, nessa empresa não foi possível obter autorização para divulgar informações mais concretas sobre a empresa por se tratar da qualidade de seus produtos.

4.3.1 Descrição do processo e das variáveis a serem analisadas

As garrafas de Politereftalato de Etileno (PET) possuem vantagens em relação às garrafas de vidro. Um dos grandes argumentos a favor da resina está no peso. Na hora da distribuição, o PET representa apenas 2% do peso da carga, o restante é produto, enquanto o vidro, por exemplo, representa 45% do peso da carga. Além de ser leve, é de fácil distribuição, dá maior durabilidade ao produto e barra o gás. O fato de ser descartável eliminou em grande parte o processo de lavagem de garrafas (cascos) para reuso, para o qual são necessários quase quatro litros de água por garrafa, mas, mesmo assim, a empresa vem investindo nas embalagens retornáveis (FRANÇA, 2009).

As garrafas são lavadas em máquinas próprias. Primeiramente, são imersas em uma solução detergente previamente aquecida para aumentar o efeito da lavagem e, assim, separar e emulsionar os detritos existentes. O processo de lavagem inicia colocando as garrafas em tanques de água quente, primeiro, no tanque um e, em seguida, elas passam para o tanque dois, cada qual com suas especificações quanto à temperatura da água. Depois de lavadas em tanques de água quente, as garrafas são enxaguadas em outros dois tanques com água à temperatura decrescente tendendo para fria. Para evitar o choque térmico, a amplitude de temperatura entre as duas zonas de lavagem não deve exceder a 35°C.

Quando as garrafas são reutilizadas, os cuidados com a limpeza devem ser redobrados. Portanto, a garrafa usada deve passar por uma limpeza física para eliminar todas as partículas estranhas visíveis no seu interior e, também no exterior. Quando a lavagem é feita adequadamente, a garrafa usada apresenta as mesmas condições de uso que a garrafa nova.

No contato com o diretor da empresa, optou-se por realizar um estudo no processo de lavagem de garrafas, que é composto por quatro tanques de água: dois tanques com água quente e dois tanques, subsequentes, com água fria.

O processo de lavagem é uma das etapas críticas do processo de engarrafamento do produto, pois a lavagem realizada inadequadamente compromete a estabilidade e a qualidade do produto. Além de ocasionar retrabalho, têm-se os custos com a segunda lavagem, podendo haver custos adicionais para aquisição de garrafas novas, entre outros. Assim, este estudo tem o objetivo de oferecer informações à equipe de produção sobre o comportamento futuro da temperatura da água, além de monitorar a estabilidade do referido processo.

Os dados em análise referem-se apenas aos tanques de água quente, nos quais tem-se início o processo de lavagem, tendo em vista que essas temperaturas precisam ser controladas, seguindo especificações determinadas pelo engenheiro. No tanque um, a temperatura da água deve estar acima de 50° C, e no tanque dois, a temperatura precisa ser superior a 60°C, não podendo passar de 80° C em ambos os casos. Caso não sejam atendidas, deve-se tomar uma medida corretiva para evitar que a garrafa seja mal higienizada e/ou esterilizada.

A variável de interesse para o tanque um foi nominada de TATQ1 e para o tanque dois, TATQ2. Os dados coletados abrangem o período de 29 de dezembro de 2007 a 30 de junho de 2009, totalizando 219 observações, considerando os dias úteis de produção e foram coletados diretamente na empresa.

Inicialmente, a metodologia proposta foi fazer uma modelagem de regressão dinâmica, tendo em vista que a temperatura da água do tanque um influencia na temperatura da água do tanque dois, pois são etapas sucessivas no processo de lavagem de garrafas. Porém, em sua execução, a metodologia não se adequou aos dados. Assim, para a série TATQ1, foi estimado um modelo matemático ARMA e, para a série TATQ2, estimou-se um modelo misto AR-ARCH.

No processo de estimação do modelo matemático, as seis últimas observações ficaram fora da amostra para avaliar as previsões por meio dos critérios de avaliação MAPE e U de Theil. E, os modelos matemáticos foram estimados com 213 observações.

4.3.2 Análise descritiva

Após a aquisição dos dados, é preciso conhecer o comportamento das variáveis no decorrer do período analisado. Na Tabela 12, apresentam-se algumas estatísticas descritivas da temperatura da água para o tanque um (TATQ1) e para o tanque dois (TATQ2).

Tabela 12 – Estatísticas descritivas das variáveis TATQ1 e TATQ2 da Empresa C no período de 29/12/2007 a 30/06/2009

Estatísticas	TATQ1	TATQ2
Média	56,041	68,201
Mediana	56,000	70,000
Desvio padrão	3,897	3,159
Coefficiente de variação (%)	6,954	4,632
Assimetria	-0,255	-1,042
Curtose	-1,002	0,791
Mínimo	50,000	60,000
Máximo	68,000	75,000

Com a análise descritiva das variáveis, verificou-se, pelo coeficiente de variação, que as médias são representativas, esse valor expresso em percentual deve estar abaixo de 50%, mostrando, assim, que a variabilidade é pequena. Percebe-se, também, que a TATQ1 exibe uma distribuição assimétrica positiva, uma vez que a mediana mostra-se menor que a média e a TATQ2 exibe uma distribuição assimétrica negativa, pois a média é menor que a mediana.

Ao analisar o valor da assimetria e da curtose, observa-se o seguinte: na variável TATQ1, a assimetria é próxima de zero e a curtose é inferior a três, e, na variável TATQ2, a assimetria é diferente de zero e a curtose é inferior a três. Por esses resultados, nenhuma das variáveis apresenta uma distribuição normal.

A Figura 52 mostra o comportamento dos dados originais para as variáveis TATQ1 e TATQ2.

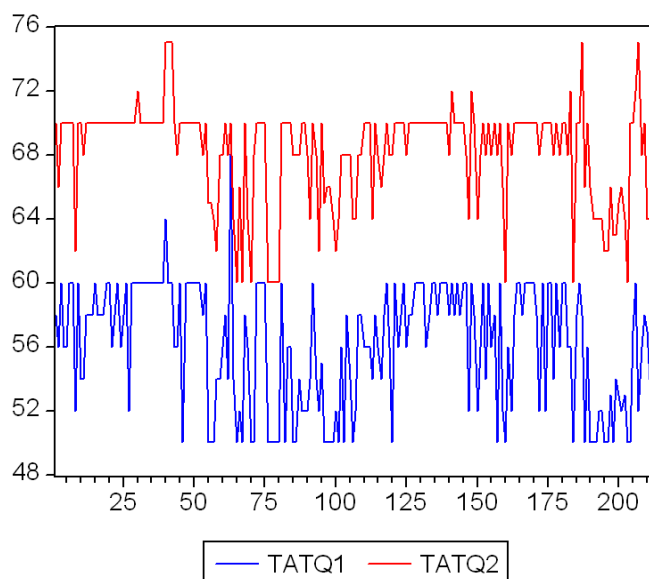


Figura 52 – Série original para as variáveis TATQ1 e TATQ2

Analisando a Figura 52, observa-se que as séries possuem uma grande variabilidade e, aparentemente, não apresentam um comportamento estacionário em relação à média. Além disso, constata-se a presença de *outliers*.

A verificação da estacionariedade pode ser feita, inicialmente, pela construção dos correlogramas, FAC e FACP, em conjunto com as estatísticas *Ljung-Box*, apresentados nas Figuras 53 e 54. Posteriormente, aplicaram-se os testes de raiz unitária, PP e KPSS, para testar estatisticamente a estacionariedade das séries.

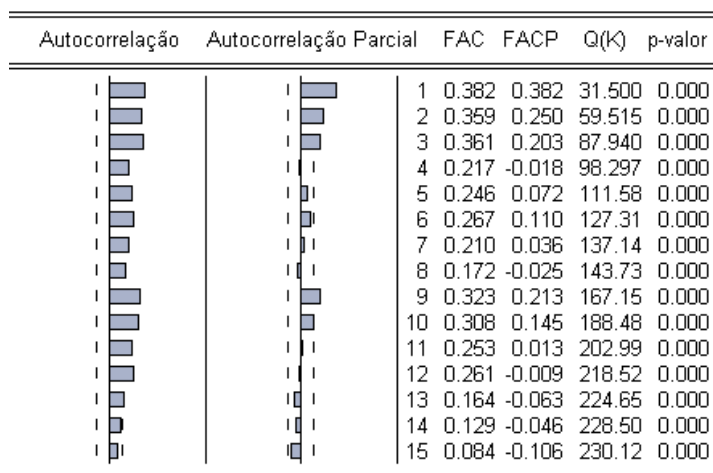


Figura 53 – Correlograma, estatística de Ljung-Box - Q(K) e p-valor da série original para a variável TATQ1

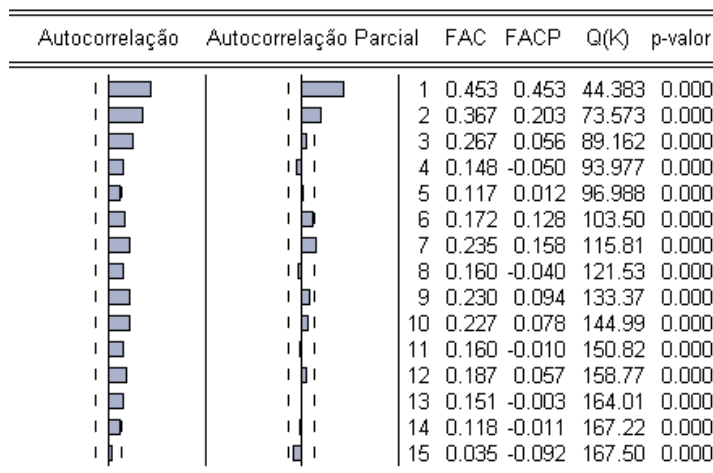


Figura 54 – Correlograma, estatística de Ljung-Box - Q(K) e p-valor da série original para a variável TATQ2

Os gráficos da Figura 53 e 54 apresentam as estatísticas *Ljung-Box*, que evidenciaram a presença de autocorrelação para os primeiros 15 *lags* das duas séries estudadas, pois os *p-valores* foram menores que o nível de significância adotado. Pela análise gráfica dos correlogramas, pode-se afirmar que ambas as séries não são estacionárias.

Para comprovar estatisticamente a presença ou ausência de estacionariedade das séries, aplicam-se os testes de raiz unitária, PP e KPSS, conjuntamente, conforme a pesquisa ex-post-facto da Empresa B. Estes testes são mostrados na Tabela 13.

O teste PP tem as seguintes hipóteses: H_0 : **a série possui raiz unitária** e H_1 : **a série não possui raiz unitária**, e o teste KPSS tem as seguintes hipóteses: H_0 : **a série é estacionária** e H_1 : **a série não é estacionária**.

Tabela 13 – Testes de raiz unitária PP e KPSS para as variáveis TATQ1 e TATQ2

Variáveis	PP	Valor crítico do	KPSS	Valor crítico
	Estatística do teste	teste PP*	Estatística do teste	do KPSS*
TATQ1	-10,750	-2,875	0,290	0,463
TATQ2	-9,156	-2,875	0,219	0,463

* O valor crítico considerado é de 5%

Conforme a Tabela 13, o teste PP prova que, se a estatística calculada é menor que valor crítico, rejeita-se a hipótese nula, concluindo-se pela estacionariedade. Quanto ao teste KPSS, constata-se que, como a estatística calculada é menor que o valor crítico, não há rejeição da hipótese nula de estacionariedade. Assim, os testes PP e KPSS asseguraram que as duas séries, TATQ1 e TATQ2, são estacionárias.

4.3.3 Estimação de modelos ARIMA concorrentes e escolha do melhor modelo

Comprovou-se anteriormente que os dados originais da TATQ1 e da TATQ2 são autocorrelacionados, logo, deve-se buscar um modelo matemático capaz de eliminar a autocorrelação serial e, a partir de seus resíduos, traçam-se os gráficos de controle \bar{x} e \overline{MR} . Desse modo, uma das alternativas iniciais é a modelagem por meio dos modelos lineares ARIMA(p,d,q), aplicado para as duas variáveis.

Os correlogramas apresentados, nas Figuras 53 e 54, sugerem que existe uma variedade de modelos ARIMA candidatos para ajustar as séries TATQ1 e TATQ2. Esses modelos concorrentes foram testados e avaliados segundo a quantidade de parâmetros significativos e as estatísticas AIC e SBC. Salienta-se que todos os modelos ARIMA concorrentes, expostos na Tabela 14, formaram um processo de ruído branco e são significativos.

Tabela 14 – Estimação dos parâmetros dos modelos ARIMA concorrentes, dos critérios AIC e SBC para as variáveis TATQ1 e TATQ2

Modelos	Parâmetros	AIC	SBC
TATQ1			
ARMA(1,1)	p(1)= 0,899728 q(1)= -0,809508	5,315	5,347
ARMA(2,2)	p(1)= 0,072575 p(2)= 0,926886 q(1)= 0,171487 q(2)= -0,807171	5,310	5,374
TATQ2			
AR(2)	p(1)= 0,588928 p(2)= 0,410668	5,052	5,084
AR(3)	p(1)= 0,508423 p(2)= 0,286499 p(3)= 0,204585	5,021	5,069
ARMA(1,1)	p(1)= 0,999905 q(1)= -0,721050	4,958	4,989
ARMA(1,2)	p(1)= 0,999890 q(1)= -0,639973 q(1)= -0,177238	4,935	4,982

A partir dos modelos concorrentes obtidos, o melhor modelo matemático representativo para a variável TATQ1 foi o ARMA(1,1) e, para a variável TATQ2, foi o modelo AR(3). Esses modelos foram selecionados com base nos critérios penalizadores, AIC e SBC.

Tendo encontrado o modelo matemático que eliminou a autocorrelação serial das séries, o próximo passo foi realizar o diagnóstico dos resíduos para os modelos estimados, e, em seguida, realizou-se a previsão.

Ressalta-se que, a partir da etapa de diagnóstico, a análise dos resultados para as duas séries temporais será realizada concomitantemente.

4.3.4 Diagnóstico dos resíduos

- Variável TATQ1

Antes de mostrar as previsões para a variável TATQ1, apresenta-se um diagnóstico do modelo estimado, analisaram-se as principais estruturas estatísticas dos resíduos, como: estatísticas descritivas, normalidade, estacionariedade, aglomerados de volatilidade e *outliers*.

A Figura 55 mostra a distribuição dos resíduos do modelo ARMA(1,1) e algumas de suas estatísticas descritivas.

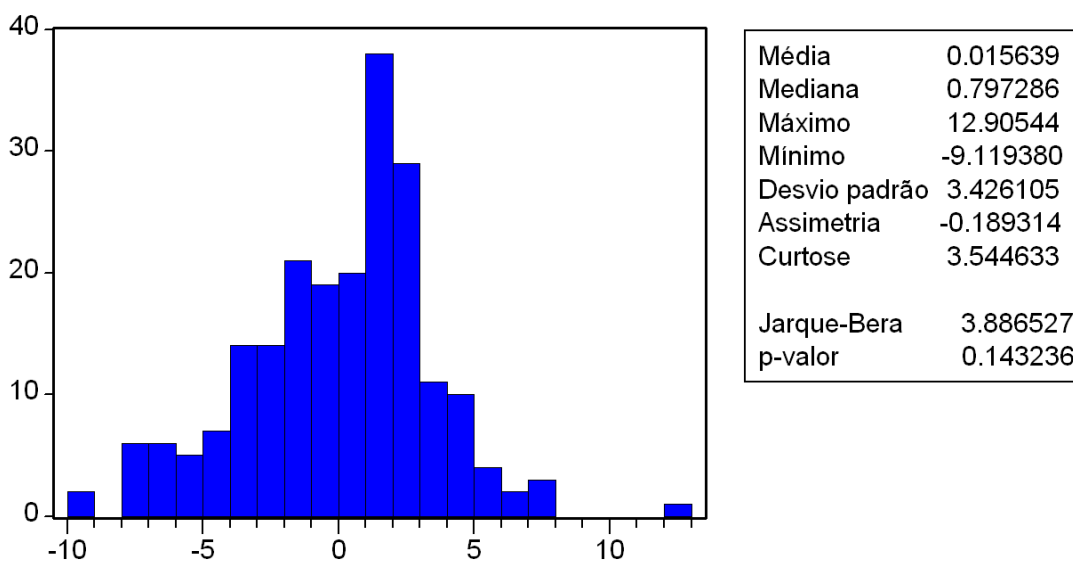


Figura 55 – Distribuição e estatísticas descritivas dos resíduos do modelo ARMA(1,1) para a variável TATQ1

Pela inspeção visual da Figura 55, constata-se que há indícios de que a série residual tem uma distribuição aproximadamente normal. Os critérios de uma distribuição normal assumem que a curtose é igual a três, e, a assimetria é igual a zero. Os resultados apontam que tanto o valor da curtose quanto da assimetria se igualam significativamente a três e zero, respectivamente. Para comprovar essa hipótese, foram feitos dois testes de normalidade: Jarque-Bera e Q-Q plot.

As hipóteses testadas na estatística Jarque-Bera são:

H_0 : Assimetria = zero e curtose = a três (normalidade);

H_1 : Assimetria \neq zero e curtose \neq três (não-normalidade).

Observando o valor da estatística Jarque-Bera e seu respectivo *p-valor*, concluiu-se que a estatística Jarque-Bera aceita a hipótese nula de normalidade para a série residual da TATQ1, considerando o nível de significância adotado.

Observando o gráfico Q-Q plot exibido na Figura 56, verifica-se que existe uma relação aproximadamente linear entre os quantis teóricos da distribuição normal e os quantis residuais do modelo estimado, indicando a sua normalidade.

Analisando os testes Jarque-Bera e o gráfico Q-Q plot, pode-se afirmar que a série residual aceita a hipótese nula de normalidade.

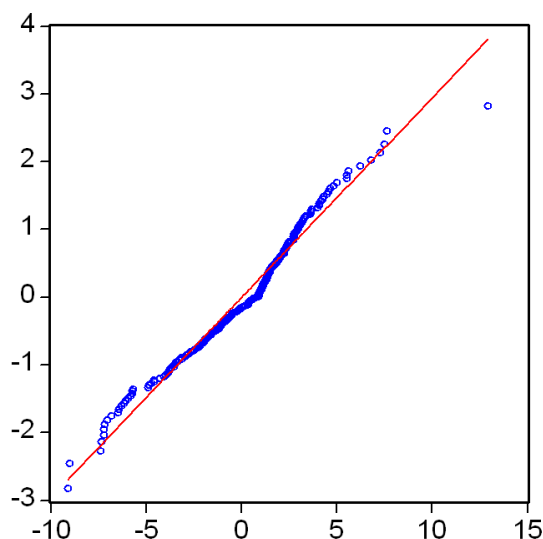


Figura 56 – Gráfico Q-Q plot dos resíduos do modelo ARMA(1,1)

Além da normalidade, é necessário averiguar se existe autocorrelação serial nos resíduos. O correlograma, exposto na Figura 57, apresenta a FAC e a FACP e as estatísticas *Ljung-Box*, usando a série dos resíduos quadráticos do modelo ARMA(1,1) estimado.

Pela análise gráfica do correlograma, pode-se afirmar que a série é estacionária, e, que a autocorrelação serial foi eliminada. As estatísticas *Ljung-Box* evidenciaram a ausência de dependência na variância condicional, pois os *p-valores* para os oito primeiros *lags* são maiores que o nível de significância.

Autocorrelação	Autocorrelação Parcial	FAC	FACP	Q(K)	p-valor	
		1	0.068	0.068	1.0012	0.317
		2	0.140	0.136	5.2562	0.072
		3	0.001	-0.017	5.2563	0.154
		4	0.035	0.017	5.5234	0.238
		5	-0.026	-0.028	5.6729	0.339
		6	0.034	0.031	5.9305	0.431
		7	0.131	0.138	9.7266	0.205
		8	0.152	0.132	14.895	0.061
		9	0.131	0.088	18.740	0.027
		10	0.140	0.100	23.173	0.010
		11	0.044	0.006	23.605	0.015
		12	0.037	0.012	23.921	0.021
		13	0.037	0.036	24.238	0.029
		14	-0.008	-0.035	24.253	0.043
		15	0.024	-0.012	24.384	0.059

Figura 57 – Correlograma, estatística de Ljung-Box - Q(K) e p-valor da série dos resíduos quadráticos do modelo ARMA(1,1) para a variável TATQ1

Além disso, para atender os pressupostos básicos da aplicação dos gráficos de controle, os resíduos do modelo estimado precisam ser homocedásticos. Sendo assim, aplicou-se o teste ARCH-LM, testando-se as seguintes hipóteses:

H_0 : Não existem efeitos ARCH (homocedasticidade);

H_1 : Existem efeitos ARCH (heterocedasticidade).

O teste revelou que os resíduos do modelo estimado para a variável TATQ1 são homocedásticos, ou seja, aceitou-se a hipótese nula de ausência de efeitos ARCH.

Como os resíduos do modelo ARMA(1,1) foram um processo de ruído branco, os mesmos serão usados para a construção dos gráficos de controle \bar{x} e \overline{MR} , apresentados na seção 4.3.7.

- Variável TATQ2

Antes de mostrar as previsões para a variável TATQ2, apresenta-se um diagnóstico do modelo matemático estimado para essa variável. Foram analisadas as principais estruturas estatísticas dos resíduos, como: estatísticas descritivas, normalidade, aglomerados de volatilidade e *outliers*.

A Figura 58 mostra a distribuição dos resíduos do modelo AR(3) e algumas de suas estatísticas descritivas.

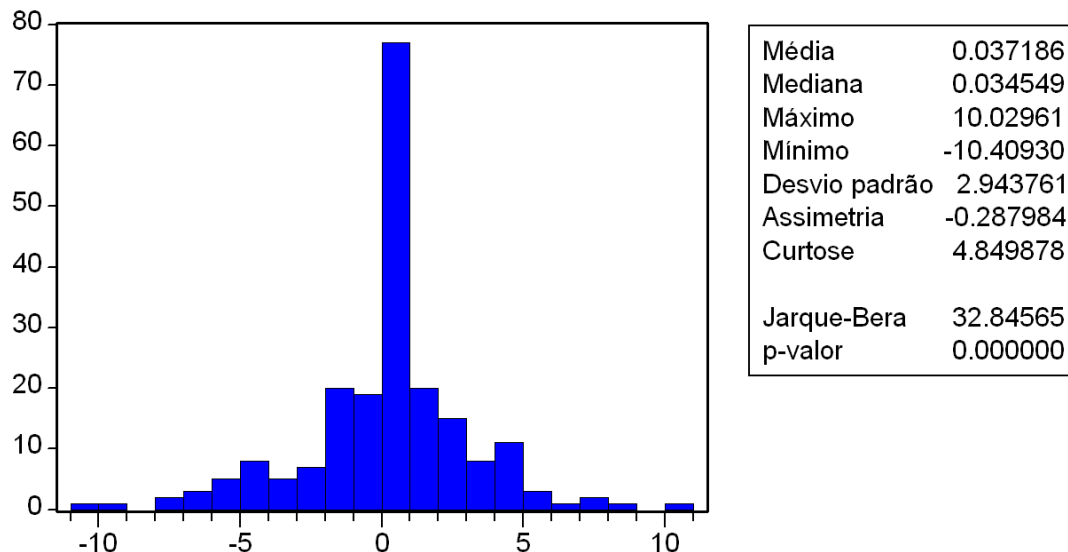


Figura 58 – Distribuição e estatísticas descritivas dos resíduos do modelo AR(3) para a variável TATQ2

Pela inspeção visual da Figura 58, constata-se que há indícios de que a série residual não apresenta uma distribuição normal. Os resultados apontam que o valor da curtose é superior a três, isso indica que a distribuição é leptocúrtica. Para comprovar essa hipótese, foram feitos dois testes de normalidade: Jarque-Bera e Q-Q plot.

As hipóteses testadas na estatística Jarque-Bera são:

H_0 : Assimetria = zero e curtose = a três (normalidade);

H_1 : Assimetria \neq zero e curtose \neq três (não-normalidade).

Observando o valor da estatística Jarque-Bera e seu respectivo *p-valor*, concluiu-se que a estatística Jarque-Bera aceita a hipótese alternativa de que a série residual não é normal. Esse resultado condiz com o que foi indicado pela curtose.

O gráfico Q-Q plot exibido na Figura 59 verifica-se que existe uma relação não-linear entre os quantis teóricos da distribuição normal e os quantis residuais do modelo estimado, indicando a presença de caudas mais pesadas, e conseqüentemente, ausência de normalidade.

Analisando os testes Jarque-Bera e o gráfico Q-Q plot, pode-se assegurar que a série residual rejeita a hipótese nula de normalidade.

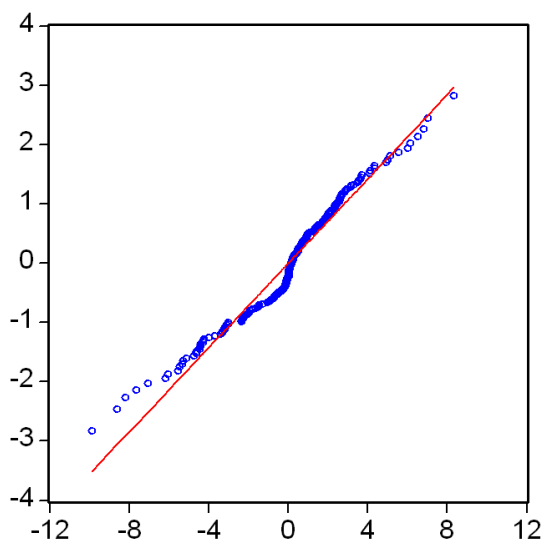


Figura 59 – Gráfico Q-Q plot dos resíduos do modelo ARMA(1,1)

A Figura 60 ilustra o comportamento temporal dos resíduos do modelo AR(3).

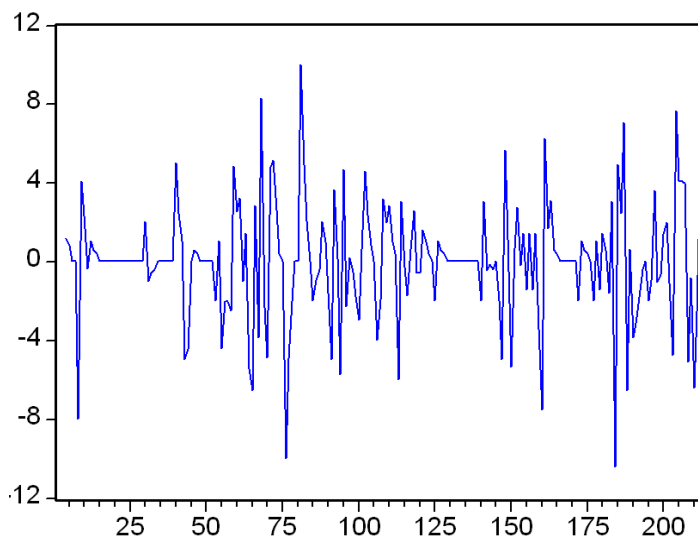


Figura 60 – Série residual do modelo AR(3)

A partir de uma simples inspeção do gráfico acima, determina-se que o processo apresenta *clusters* de volatilidade.

Para comprovar se a série residual apresenta *clusters* de volatilidade, verificou-se se existe autocorrelação serial nos resíduos quadráticos do modelo AR(3). O correlograma, exposto na Figura 61, apresenta a FAC e a FACP e as estatísticas *Ljung-Box*. Observando-se as estatísticas *Ljung-Box* evidencia-se, claramente, a existência de dependência na variância condicional.

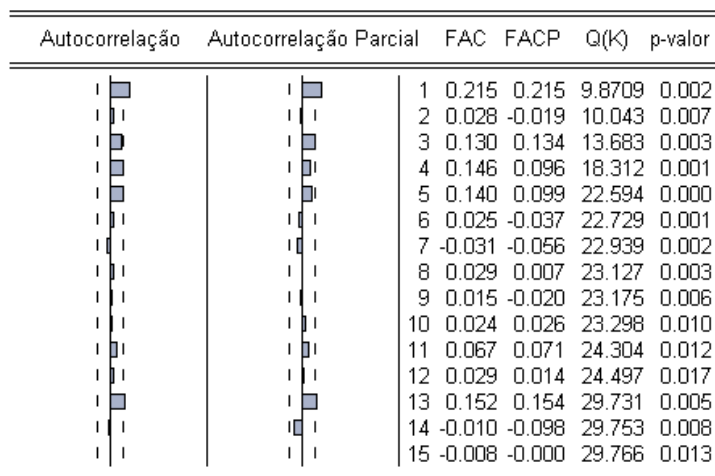


Figura 61 – Correlograma, estatística de Ljung-Box – Q(K) e p-valor da série dos resíduos quadráticos do modelo AR(3) para a variável TATQ2

Para comprovar estatisticamente a presença de heterocedasticidade condicional, aplicou-se o teste ARCH-LM nos resíduos quadráticos do modelo AR(3). Como foi constatada a presença desse fenômeno, justifica-se a utilização de um modelo não-linear da família ARCH para modelar a volatilidade. Encontra-se, na Tabela 15, o resultado desse teste.

Para comprovar a presença de efeitos ARCH, foram testadas as seguintes hipóteses:

H_0 : Não existem efeitos ARCH (homocedasticidade);

H_1 : Existem efeitos ARCH (heterocedasticidade).

A partir dos resultados encontrados para as diferentes defasagens, verifica-se que todas as estatísticas F e as estatísticas ARCH-LM calculadas foram significativas. Assim, rejeita-se a hipótese nula, implicando que existe heterocedasticidade condicional nos resíduos quadráticos do modelo AR(3).

Tabela 15 – Resultados das estatísticas F e ARCH-LM até a defasagem 13

Lags	Estatística F	p-valor (F)	Estatística ARCH-LM	p-valor (ARCH-LM)
Lag 1	10,074	0,002	9,670	0,002
Lag 3	4,442	0,005	12,751	0,005
Lag 5	3,950	0,002	18,506	0,002
Lag 7	2,860	0,008	18,900	0,008
Lag 9	2,307	0,018	19,705	0,020
Lag 11	1,977	0,033	20,729	0,036
Lag 13	2,114	0,015	25,723	0,018

4.3.5 Estimação do melhor modelo misto AR-ARCH

Como foi estimado um modelo matemático por meio de um modelo AR, será empregado um modelo misto denominado de AR-ARCH, visto que os resíduos do modelo AR apresentaram heterocedasticidade. Assim, justifica-se a utilização de uma estimação conjunta dos parâmetros do modelo linear (AR) e não-linear (ARCH).

A Tabela 16 apresenta o modelo misto AR-ARCH para a variável TATQ2, pelo qual foram estimadas equações para a média e variância condicionais.

Tabela 16 – Estimação dos coeficientes, erro-padrão, estatística *z* e *p*-valor do modelo AR-ARCH para a variável TATQ2

Método: ML – ARCH (Marquardt) - Distribuição Normal				
Equação para a média condicional				
	Coefficiente	Erro-Padrão	Estatística <i>z</i>	<i>p</i>-valor
AR(1)	0,525	0,084	6,275	0,000
AR(2)	0,308	0,068	4,528	0,000
AR(3)	0,164	0,054	3,072	0,002
Equação para a variância condicional				
C	0,773	0,417	13,829	0,000
ARCH(1)	0,381	0,121	3,152	0,002

O modelo para a volatilidade, exposto na Tabela 16, é descrito como um AR(3)-ARCH(1) apresentando parâmetros estatisticamente significativos. Constata-se que a soma dos parâmetros é menor que um, assegurando que a variância condicional seja positiva e estacionária. O valor de alfa, parâmetro do ARCH(1), é razoavelmente baixo, igual a 0,381, e, indica que, após um curto período de tempo, a variância da série deverá convergir à média histórica, ou seja, a TATQ2 no período anterior não influencia fortemente no próximo período.

Para fazer o diagnóstico do modelo encontrado, analisam-se os resíduos padronizados do modelo misto estimado. Observando a Figura 62, constata-se que os resíduos padronizados são aproximadamente normais, apresentando uma cauda mais pesada que a normal padronizada, o que não afeta a análise futura de estimação do modelo ARCH. O teste ARCH-LM comprovou a ausência de efeitos ARCH nesses resíduos.

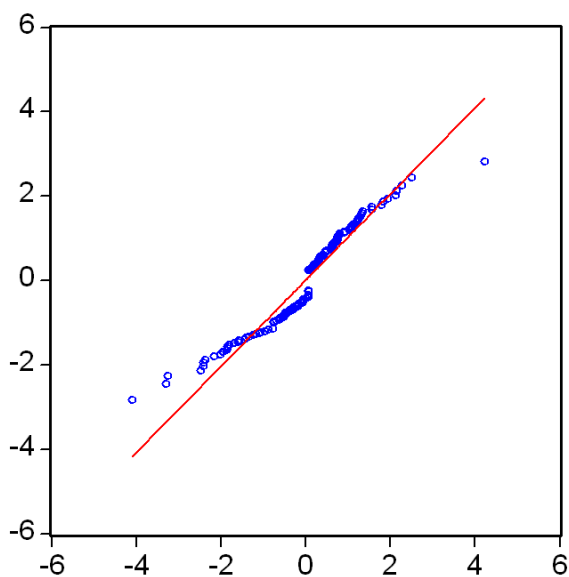


Figura 62 – Q-Q plot dos resíduos padronizados estimados pelo modelo AR(3)-ARCH(1) para a variável TATQ2

A partir do modelo misto, mostra-se na Figura 63 a sua volatilidade condicional.

Analisando o gráfico da volatilidade condicional observa-se que o modelo misto AR(3)-ARCH(1) reproduziu bem os períodos marcantes dos *clusters* de volatilidade presentes na série residual do modelo AR(3), apresentado na Figura 60.

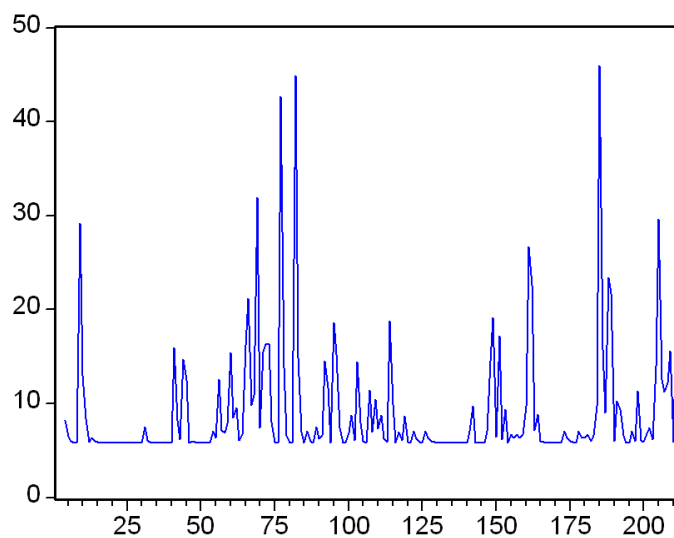


Figura 63 – Volatilidade condicional do modelo misto AR(3)-ARCH(1)

Antes de verificar a estabilidade do processo, apresenta-se a previsão.

4.3.6 Previsão dos modelos matemáticos estimados

Como foi mencionado anteriormente, tanto o modelo ARMA(1,1) como o modelo misto AR(3)-ARCH(1) foram estimados com 213 observações. As últimas seis ficaram de fora da amostra com o objetivo de avaliar o modelo matemático por meio do MAPE e do U de Theil. Sendo assim, a previsão ocorre da observação 214 à observação 219.

- Variável TATQ1

A Figura 64 mostra os valores originais, os valores estimados e os resíduos para a variável TATQ1.

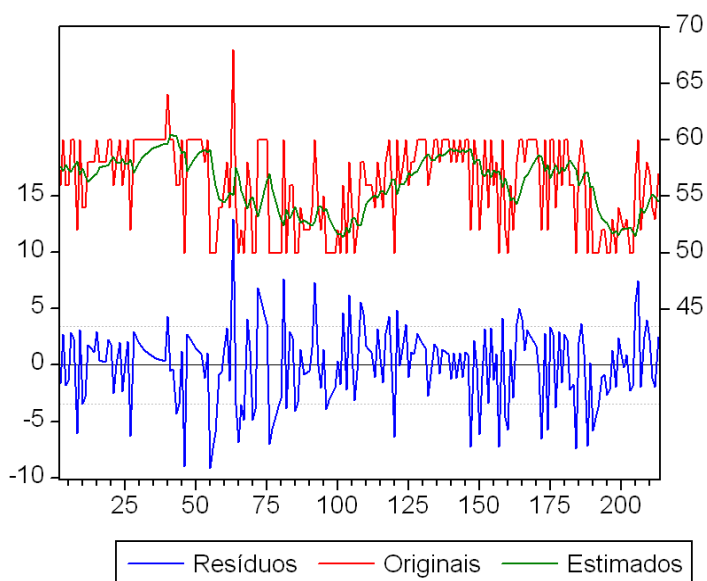


Figura 64 – Valores originais, valores estimados e os resíduos do modelo ARMA(1,1)

Na Tabela 17 apresenta-se a previsão, incluindo o período, os valores originais e previstos, o erro de previsão e os critérios MAPE e U de Theil.

Analisando o MAPE e o U de Theil, comprova-se estatisticamente que o modelo matemático estimado apresenta uma boa previsão, se comparada com a previsão ingênua. Como o coeficiente U de Theil está mais próximo de zero, a previsão realizada é melhor que a previsão ingênua ou trivial.

Tabela 17 – Período, valores originais, valores previstos, erro de previsão, MAPE e U de Theil para a variável TATQ1 da Empresa C

Período	Originais	Previstos	Erro de previsão
17/6/2009	53	54,830	-1,830
22/6/2009	51	54,815	-3,815
24/6/2009	52	54,800	-2,800
25/6/2009	53	54,785	-1,785
26/6/2009	50	54,770	-4,770
30/6/2009	50	54,755	-4,775

MAPE= 5,937 e U de Theil= 0,034

- Variável TATQ2

A Figura 65 mostra os valores originais, os valores estimados e os resíduos para a variável TATQ2.

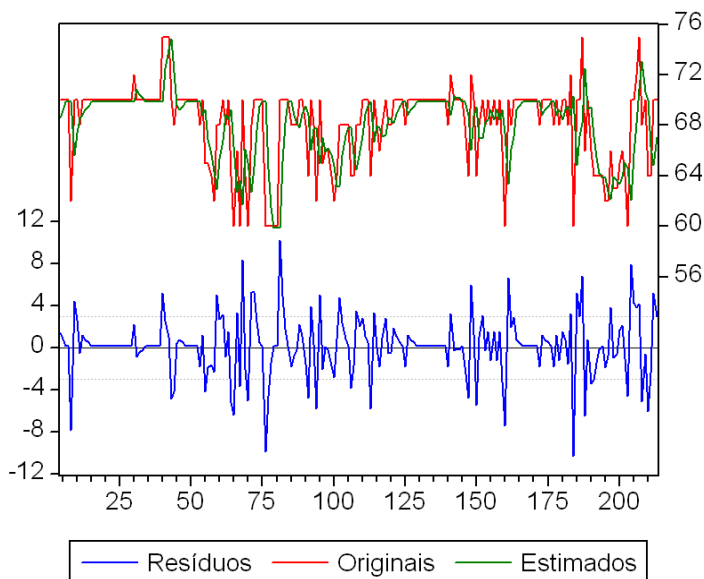


Figura 65 – Valores originais, valores estimados e os resíduos do modelo AR(3)-ARCH(1)

Na Tabela 18 mostra-se a previsão, incluindo o período, os valores originais e previstos, o erro de previsão e os critérios MAPE e U de Theil.

Tabela 18 – Período, valores originais, valores previstos, erro de previsão, MAPE e U de Theil para a variável TATQ2 da Empresa C

Período	Originais	Previstos	Erro de previsão
17/6/2009	70	51,902	18,098
22/6/2009	62	51,830	10,170
24/6/2009	68	51,759	16,241
25/6/2009	70	51,688	18,312
26/6/2009	65	51,617	13,383
30/6/2009	66	51,546	14,454
MAPE= 12,762 e U de Theil= 0,079			

Analisando o MAPE e o U de Theil, comprova-se estatisticamente que o modelo matemático estimado apresenta uma boa previsão, se comparada com a previsão ingênua. Como o coeficiente U de Theil está mais próximo de zero, a previsão realizada é melhor que a previsão ingênua ou trivial.

4.3.7 Análise da estabilidade do processo

Embora a análise descritiva revele o comportamento do processo, é necessário investigar sua estabilidade por meio dos gráficos de controle, os quais vão caracterizá-lo como estando sob controle ou fora de controle.

Apresentam-se, a seguir, os gráficos de controle \bar{x} e \overline{MR} , considerando os dados originais e, posteriormente, os resíduos do modelo estimado para TATQ1 e TATQ2.

- Variável TATQ1

Inicialmente, apresentam-se, nas Figuras 66 e 67, os gráficos de controle \bar{x} e \overline{MR} , usando os dados originais com base nas amostras coletadas para a variável TATQ1. Nesses gráficos de controle foram adicionados os *runs tests*, os quais podem revelar a existência de algum tipo de padrão nos dados.

De acordo com a Figura 66, observa-se uma grande quantidade de pontos vermelhos grafados no gráfico, os quais representam a instabilidade da variável TATQ1. Essa instabilidade é explicada pelo fato de que os *runs tests* não funcionam em processos autocorrelacionados.

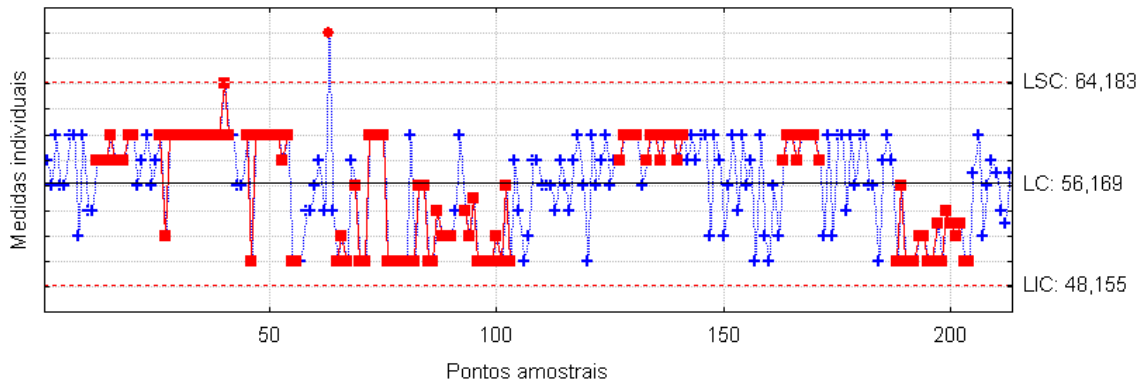


Figura 66 – Gráfico \bar{x} da TATQ1, considerando os dados originais

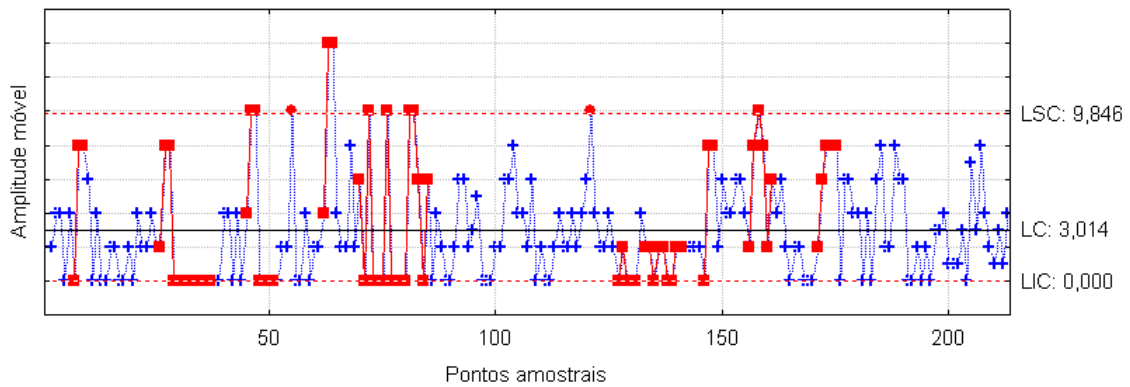


Figura 67 – Gráfico \overline{MR} da TATQ1, considerando os dados originais

Analisando o gráfico \overline{MR} da Figura 67, também verifica-se uma grande quantidade de pontos vermelhos grafados no gráfico, ocasionando uma elevada instabilidade da variável TATQ1.

Ao realizar outra análise simulando a exclusão dos pontos amostrais fora de controle, tem-se outra visão da instabilidade que está ocorrendo no processo. Com a exclusão dos pontos amostrais fora de controle, os limites de controle dos gráficos \bar{x} e \overline{MR} são recalculados. As Figuras 68 e 69 mostram essa simulação com os *runs tests*.

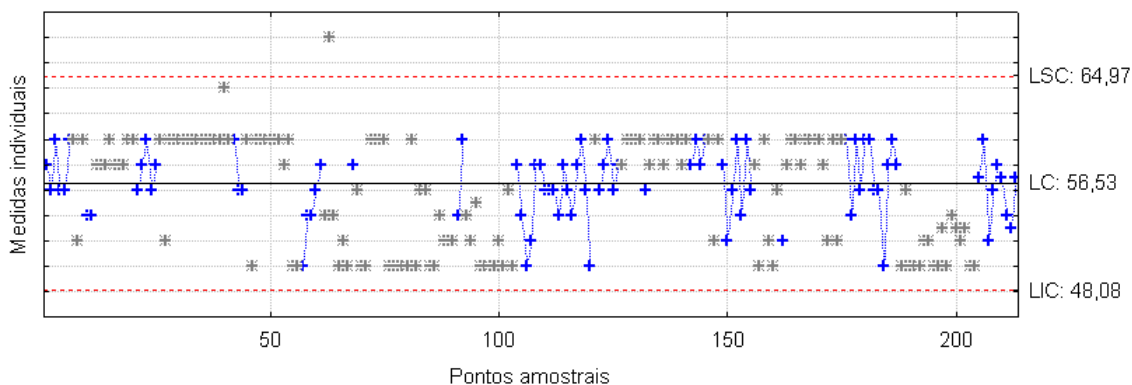


Figura 68 – Gráfico \bar{x} da TATQ1 com a exclusão dos pontos amostrais fora de controle, considerando os dados originais

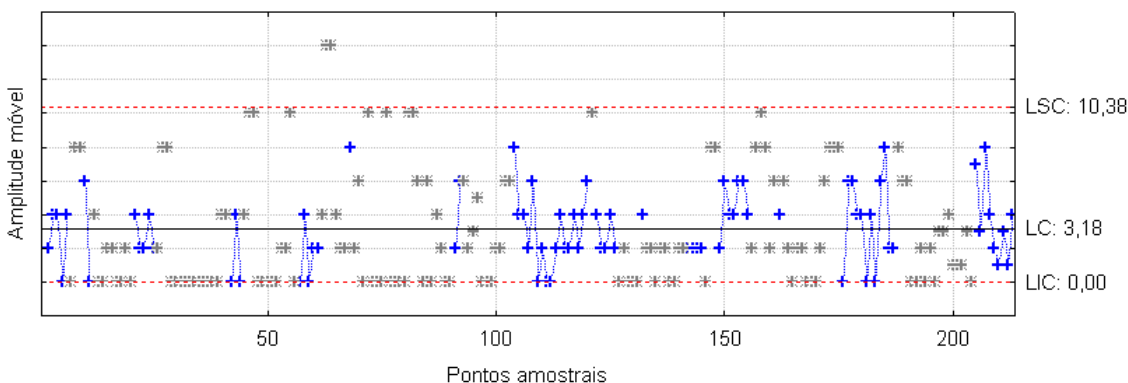


Figura 69 – Gráfico \overline{MR} da TATQ1 com a exclusão dos pontos amostrais fora de controle, considerando os dados originais

Pelos resultados obtidos, justifica-se que as técnicas convencionais de controle estatístico de processo não podem ser utilizadas diretamente nos dados originais, ou seja, é necessário um tratamento na variável de interesse para remover a autocorrelação serial antes de monitorar a variável, pois a mesma pode inserir pontos amostrais fora de controle, quando, na verdade, o processo pode estar sob controle, nesse caso, gerando alarmes falsos.

Para sanar o efeito da autocorrelação serial, estimou-se um modelo linear para a variável TATQ1, apresentado anteriormente. Como os resíduos do modelo estimado são independentes e normalmente distribuídos, as técnicas do CEP podem ser utilizadas com a finalidade de identificar a ocorrência de padrões especiais, caracterizando o processo analisado quanto a sua estabilidade para que seja feito ou não um ajuste.

Os gráficos das Figuras 70 e 71 apresentam o monitoramento dos resíduos ARMA(1,1) para a média condicional, traçando os gráficos de controle \bar{x} e \overline{MR} , tendo como auxílio na decisão a aplicação dos *runs tests*.

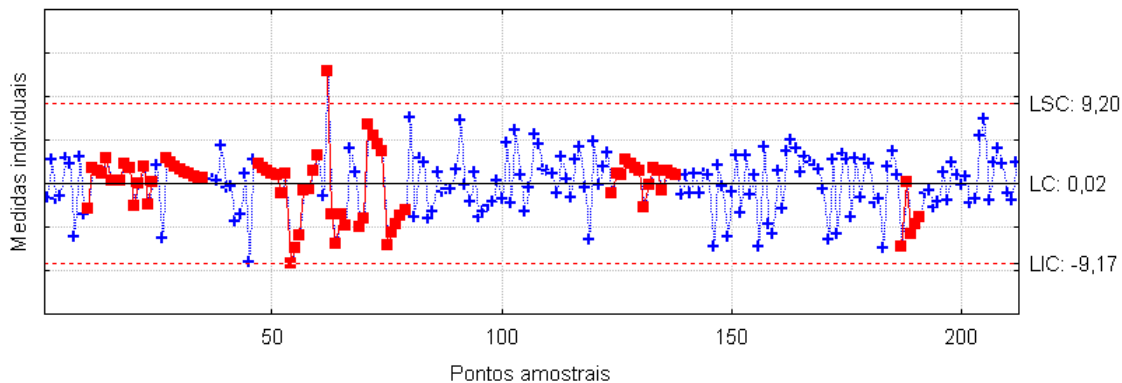


Figura 70 – Gráfico \bar{x} aplicado ao resíduo do modelo ARMA(1,1) para a variável TATQ1

De acordo com os resíduos do modelo estimado, traçou-se o gráfico \bar{x} mostrado na Figura 67, a partir do qual se detecta que **9 pontos amostrais em sequência estão no mesmo lado da linha central**, representadas pelo conjunto de amostras com início nas observações 11 e 27, e, que existem **6 pontos amostrais em sequência com tendência crescente e/ou decrescente**, correspondendo às observações com início nas amostras 28, 47 e 55. Observa-se também que, a **cada 3 pontos amostrais traçados, 2 se encontram localizados na zona A ou fora dela**, correspondendo à observação com início na amostra 53, e um conjunto de 4 amostras em que, a **cada 5 pontos amostrais traçados, 4 se encontram localizados na zona B ou fora dela**, correspondendo às observações com início nas amostras 62, 70, 75 e 187. Além disso, verificam-se **15 pontos amostrais em sequência situados na zona C**, correspondendo às observações com início nas amostras 10 e 124, e observam-se **8 pontos amostrais em sequência situados fora da zona C**, correspondendo à observação com início na amostra 69.

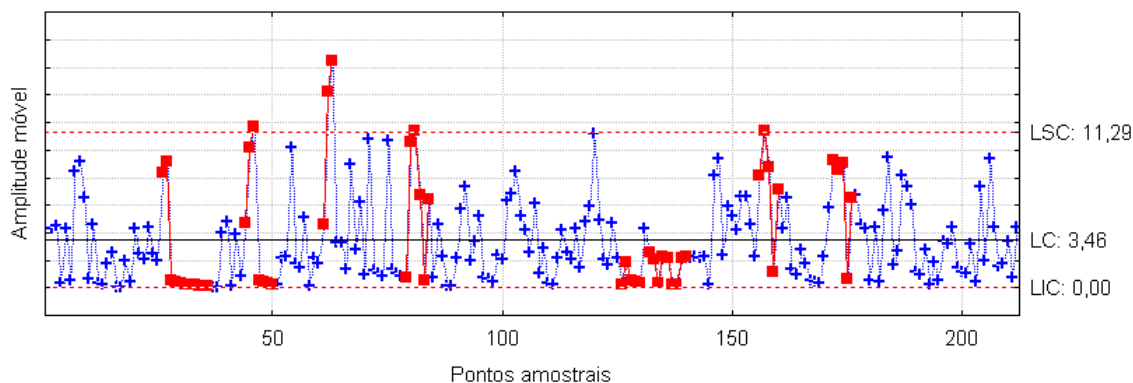


Figura 71 – Gráfico \overline{MR} aplicado ao resíduo do modelo ARMA(1,1) para a variável TATQ1

Utilizando os resíduos do modelo estimado, traçou-se o gráfico \overline{MR} apresentado na Figura 71, pelo qual se constata que **9 pontos amostrais em sequência estão no mesmo lado da linha central**, representadas pelo conjunto de amostras com início nas observações 28 e 132 e, que existem **6 pontos amostrais em sequência com tendência crescente e/ou decrescente**, correspondendo à observação com início na amostra 28. Observa-se também que, a **cada 3 pontos amostrais traçados, 2 se encontram localizados na zona A ou fora dela**, correspondendo às observações com início nas amostras 44, 61, 79, 156 e 172, e um conjunto de 7 amostras em que, a **cada 5 pontos amostrais traçados, 4 se encontram localizados na zona B ou fora dela**, correspondendo às observações com início nas amostras 27, 32, 46, 80, 126, 156 e 172. Além disso, verificam-se **8 pontos amostrais em sequência situados fora da zona C**, correspondendo à observação com início na amostra 26.

Mediante o levantamento das afirmações acima citadas, verifica-se que há poucos pontos amostrais fora de controle, se comparado ao gráfico \bar{x} da série original. Isso indica que o modelo ARMA foi capaz de captar as variações que havia na série TATQ1, pois o efeito de autocorrelação serial foi tratado.

Ao realizar outra análise simulando a exclusão dos pontos amostrais fora de controle, tem-se outra visão da instabilidade que está ocorrendo no processo. Com a exclusão dos pontos amostrais fora de controle, os limites de controle dos gráficos \bar{x} e \overline{MR} são recalculados. As Figuras 72 e 73 mostram essa simulação, aplicando os *runs tests*.

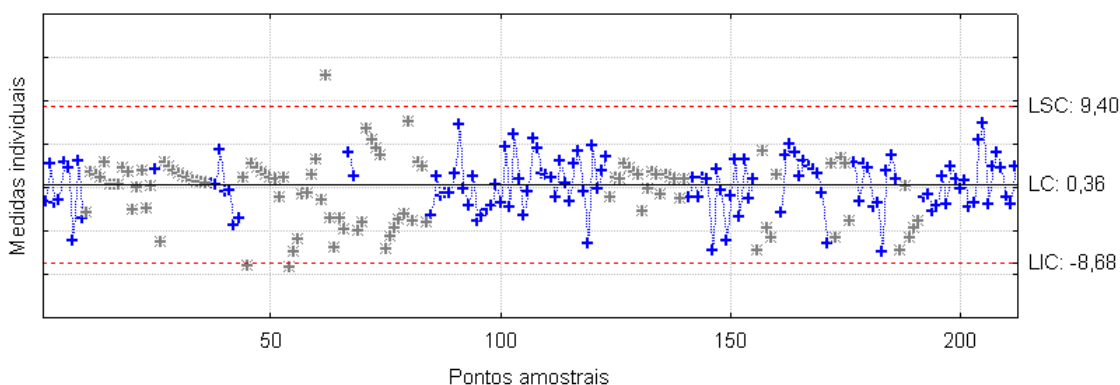


Figura 72 – Gráfico \bar{x} aplicado ao resíduo do modelo ARMA(1,1) para a variável TATQ1 com a exclusão dos pontos amostrais fora de controle

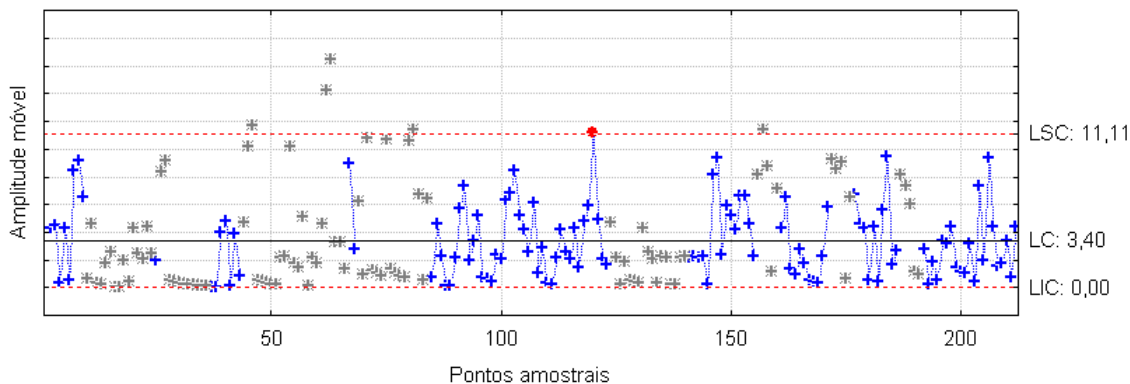


Figura 73 – Gráfico \overline{MR} aplicado ao resíduo do modelo ARMA(1,1) para a variável TATQ1 com a exclusão dos pontos amostrais fora de controle

- Variável TATQ2

Para dar início a esta etapa, apresentam-se os gráficos de controle \bar{x} e \overline{MR} , nas Figuras 74 e 75, usando os dados originais com base nas amostras coletadas para a variável TATQ2. Nesses gráficos foram adicionados os *runs tests*.

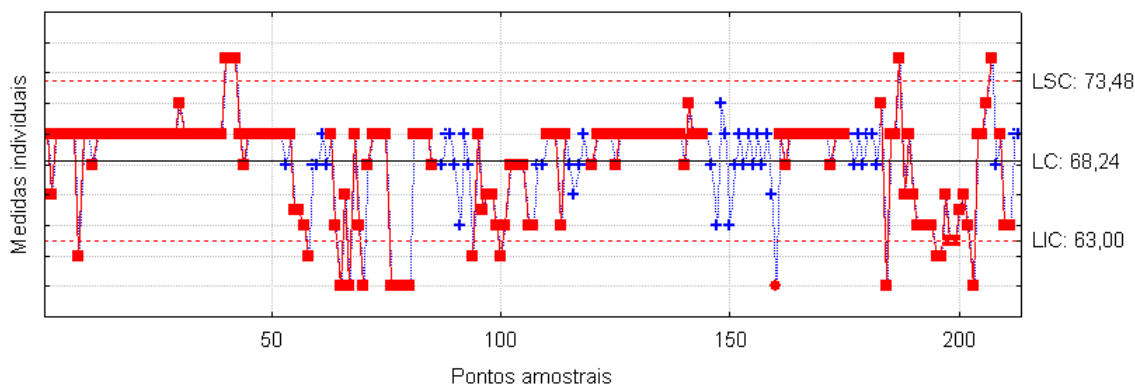


Figura 74 – Gráfico \bar{x} da TATQ2, considerando os dados originais

De acordo com a Figura 74, observa-se uma grande quantidade de pontos vermelhos grafados no gráfico, os quais revelam uma elevada instabilidade na variável TATQ2. Essa instabilidade é explicada pelo fato de que os *runs tests* não funcionam em processos autocorrelacionados. Por isso, há uma grande quantidade de pontos vermelhos.

Analisando o gráfico \overline{MR} da Figura 75, também verifica-se uma grande quantidade de pontos vermelhos grafados no gráfico, ocasionando uma elevada instabilidade da variável TATQ2.

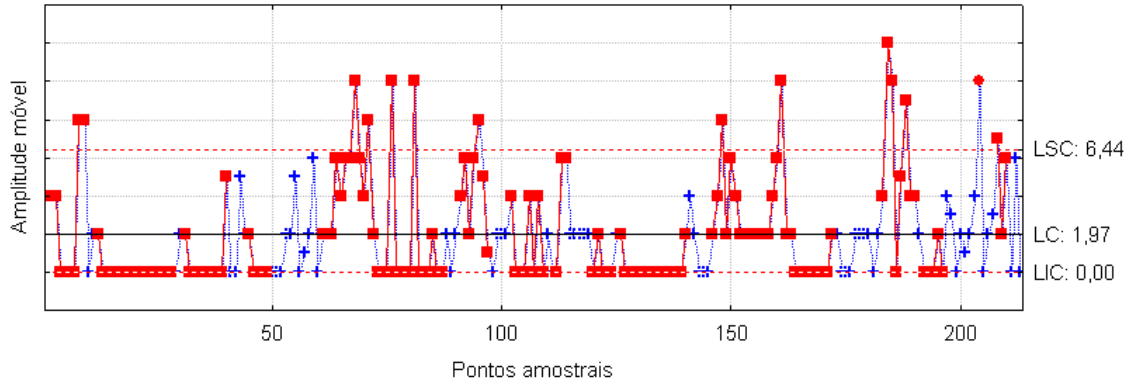


Figura 75 – Gráfico \overline{MR} da TATQ2, considerando os dados originais

Ao realizar outra análise simulando a exclusão dos pontos amostrais fora de controle, tem-se outra visão da instabilidade que está ocorrendo no processo. As Figuras 76 e 77 mostram essa simulação aplicando os *runs tests*.

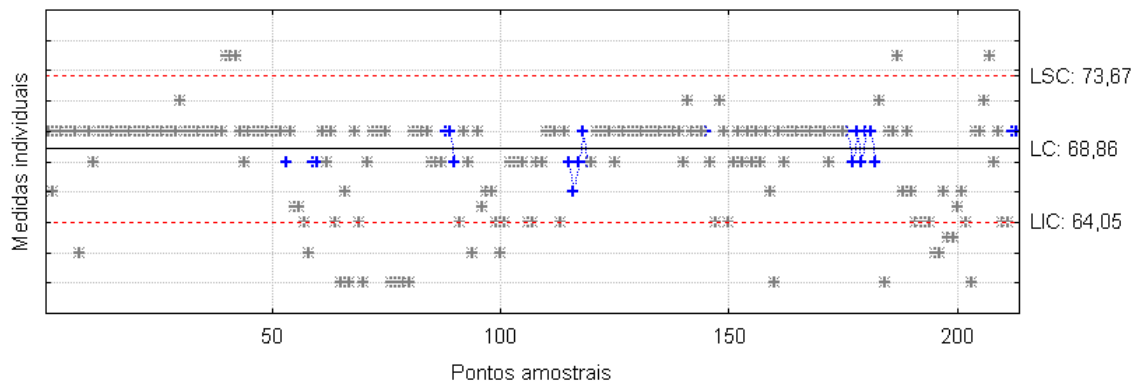


Figura 76 – Gráfico \bar{x} da TATQ2 com a exclusão dos pontos amostrais fora de controle, considerando os dados originais

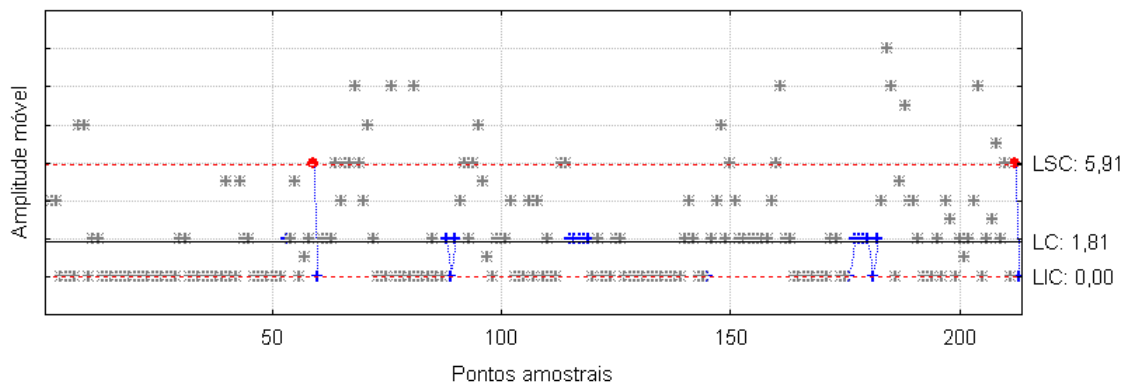


Figura 77 – Gráfico \overline{MR} da TATQ2 com a exclusão dos pontos amostrais fora de controle, considerando os dados originais

Pelos resultados obtidos, justifica-se que as técnicas convencionais de CEP não podem ser utilizadas diretamente nos dados originais, ou seja, é necessário um tratamento na variável de interesse para remover a autocorrelação serial antes de monitorar a variável, pois a mesma pode inserir pontos amostrais fora de controle, sendo capaz de gerar alarmes falsos.

Para sanar o efeito da autocorrelação serial, estimou-se um modelo misto, pois os resíduos do modelo linear AR(3) continham volatilidade, atendendo perfeitamente os pressupostos para a aplicação dos gráficos de controle. Como os resíduos do modelo estimado são independentes e normalmente distribuídos, as técnicas de CEP podem ser utilizadas com a finalidade de identificar o processo quanto a sua estabilidade.

Os gráficos das Figuras 78 e 79 apresentam o monitoramento dos resíduos para o modelo misto AR(3)-ARCH(1), traçando os gráficos de controle \bar{x} e \overline{MR} , tendo como auxílio na decisão a aplicação dos *runs tests*.

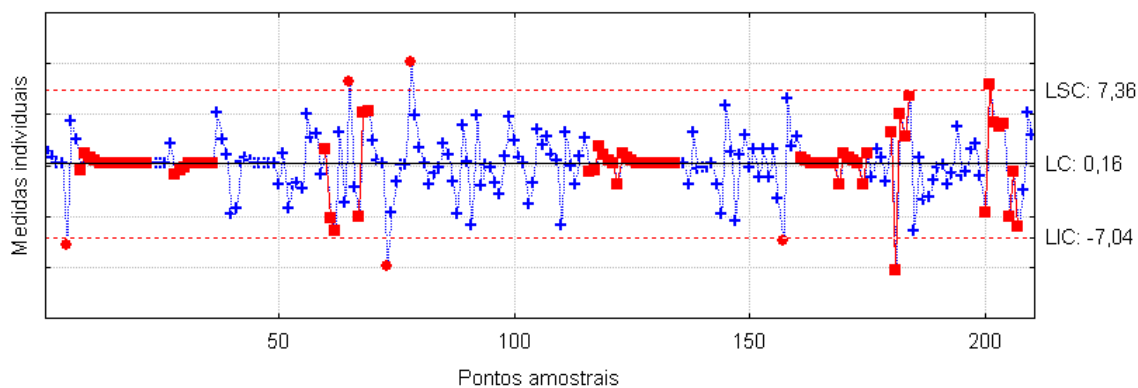


Figura 78 – Gráfico \bar{x} aplicado ao resíduo quadrático do modelo misto AR(3)-ARCH(1) para a variável TATQ2

De acordo com os resíduos do modelo estimado, traçou-se o gráfico \bar{x} mostrado na Figura 78, pelo qual se detecta que **9 pontos amostrais em sequência estão no mesmo lado da linha central**, representadas pelo conjunto de amostras com início nas observações 12, 28 e 126. Observa-se, também, que, a **cada 3 pontos amostrais traçados, 2 se encontram localizados na zona A ou fora dela**, correspondendo às observações com início nas amostras 60, 67, 182 e 205, e um conjunto de 2 amostras em que, a **cada 5 pontos amostrais traçados, 4 se encontram localizados na zona B ou fora dela**, correspondendo às observações com início nas amostras 180 e 200. Além disso, verificam-se **15 pontos amostrais em sequência situados na zona C**, correspondendo às observações com início nas amostras 8, 116 e 161.

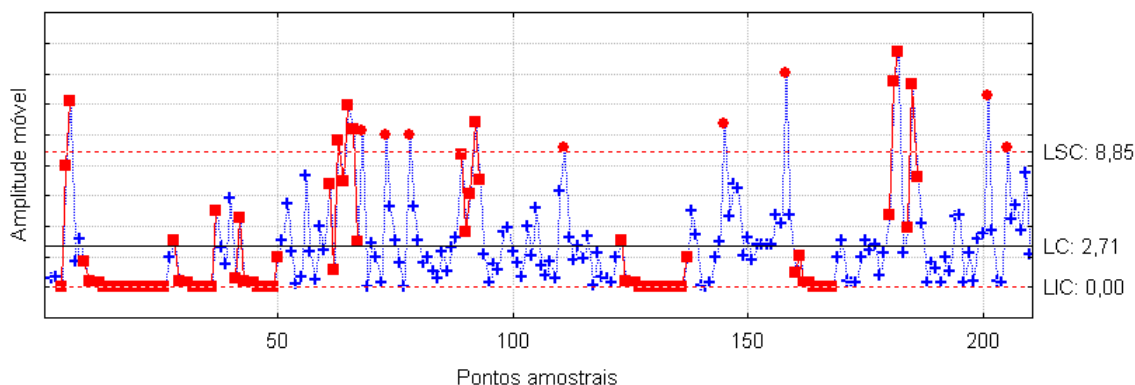


Figura 79 – Gráfico \overline{MR} aplicado ao resíduo quadrático do modelo misto AR(3)-ARCH(1) para a variável TATQ2

Utilizando os resíduos do modelo estimado, traçou-se o gráfico \overline{MR} apresentado na Figura 79, pelo qual se constata que **9 pontos amostrais em sequência estão no mesmo lado da linha central**, representadas pelo conjunto de amostras com início nas observações 9, 18, 124 e 160. Observa-se também que, a **cada 3 pontos amostrais traçados, 2 se encontram localizados na zona A ou fora dela**, correspondendo às observações com início nas amostras 4, 62, 65, 91, 180 e 184, e um conjunto de 13 amostras em que, a **cada 5 pontos amostrais traçados, 4 se encontram localizados na zona B ou fora dela**, correspondendo às observações com início nas amostras 9, 14, 19, 28, 33, 41, 46, 61, 89, 123, 128, 133 e 161. Além disso, verificam-se **8 pontos amostrais em sequência situados fora da zona C**, correspondendo às observações com início nas amostras 10, 18, 29 e 124.

Mediante as afirmações acima citadas, verifica-se que há poucos pontos amostrais fora de controle, se comparado ao gráfico \bar{x} da série original. Isso indica que o modelo estimado foi capaz de captar as variações que havia na série TATQ2 em que o efeito da autocorrelação foi tratado.

Ao realizar outra análise simulando a exclusão dos pontos amostrais fora de controle do modelo misto AR(3)-ARCH(1), tem-se outra visão da instabilidade que está ocorrendo no processo. Com a exclusão dos pontos amostrais fora de controle, os limites de controle dos gráficos \bar{x} e \overline{MR} são recalculados. As Figuras 80 e 81 apresentam essa simulação, aplicando os *runs tests*.

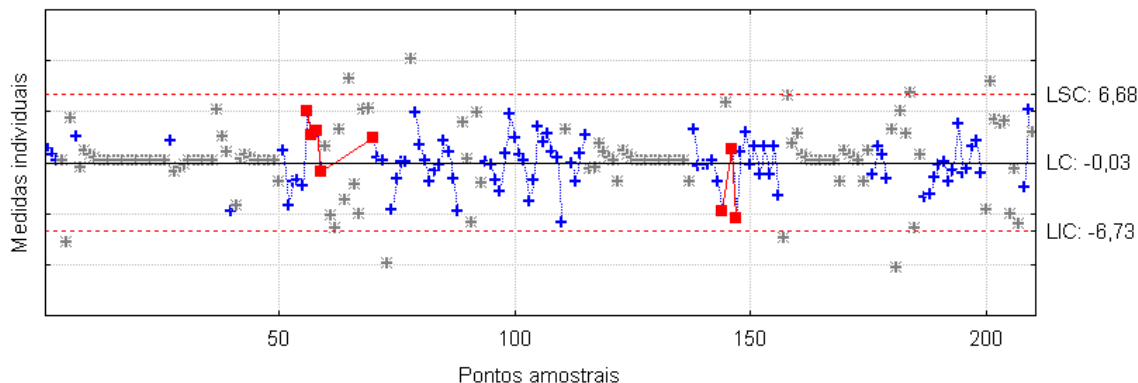


Figura 80 – Gráfico \bar{x} aplicado ao resíduo quadrático do modelo misto AR(3)-ARCH(1) para a variável TATQ2 com a exclusão dos pontos amostrais fora de controle

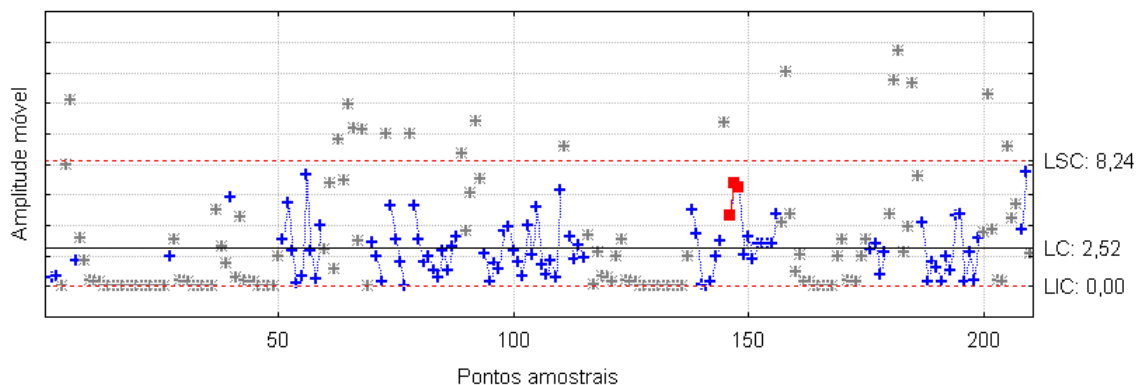


Figura 81 – Gráfico \overline{MR} aplicado ao resíduo quadrático do modelo misto AR(3)-ARCH(1) para a variável TATQ2 com a exclusão dos pontos amostrais fora de controle

Apresentam-se, na sequência, os gráficos de controle \bar{x} e \overline{MR} para realizar o monitoramento da volatilidade condicional oriunda do modelo misto AR(3)-ARCH(1).

No gráfico das Figuras 78 e 79 aplicados aos resíduos do modelo misto AR(3)-ARCH(1), consegue-se visualizar o comportamento da média e da variância condicional estimada. Portanto, foram traçados os gráficos \bar{x} e \overline{MR} para mostrar os *clusters* de volatilidade e justificar por que o gráfico do modelo misto é melhor, pois ele capta e elimina a volatilidade, assim como a modelagem ARIMA capta e elimina a autocorrelação serial dos resíduos.

Os gráficos das Figuras 82 e 83 apresentam o monitoramento da volatilidade procedente do modelo misto AR(3)-ARCH(1), traçando os gráficos de controle \bar{x} e \overline{MR} , sem a aplicação dos *runs tests*.

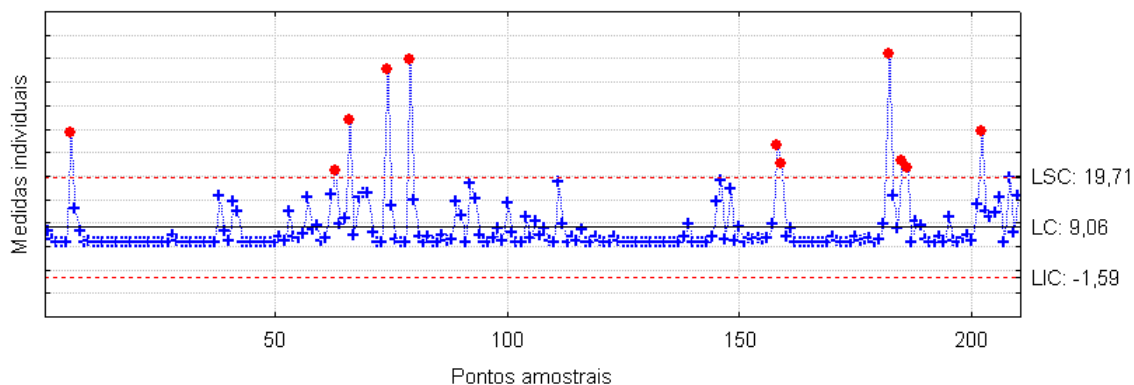


Figura 82 – Gráfico \bar{x} aplicado a volatilidade do modelo misto AR(3)-ARCH(1) para a variável TATQ2

De acordo com a volatilidade do modelo misto AR(3)-ARCH(1), traçou-se o gráfico \bar{x} , mostrado na Figura 82, pelo qual se detecta que existem 11 pontos amostrais fora dos limites de controle. Ao comparar esse gráfico com o gráfico construído a partir dos resíduos do modelo misto, exposto na Figura 78, verifica-se que o gráfico \bar{x} para a volatilidade apresentou mais pontos amostrais fora de controle do que o gráfico para modelo misto. Além disso, observando o gráfico \bar{x} , percebeu-se que o mesmo representa bem os períodos marcantes dos *clusters* de volatilidade, representados na Figura 60.

Utilizando os resíduos do modelo estimado, traçou-se o gráfico \overline{MR} , apresentado na Figura 83, pelo qual se constata que existem 17 pontos amostrais fora dos limites de controle.

Mediante as afirmações acima citadas, verifica-se que a volatilidade influencia na instabilidade da média e na amplitude móvel do processo. Isso indica que o modelo misto AR(3)-ARCH(1) foi eficaz no tratamento da autocorrelação do processo, pois o gráfico aplicado na volatilidade do modelo misto mostra que o processo apresenta uma variação que deveria ser levada em consideração no momento de analisá-lo.

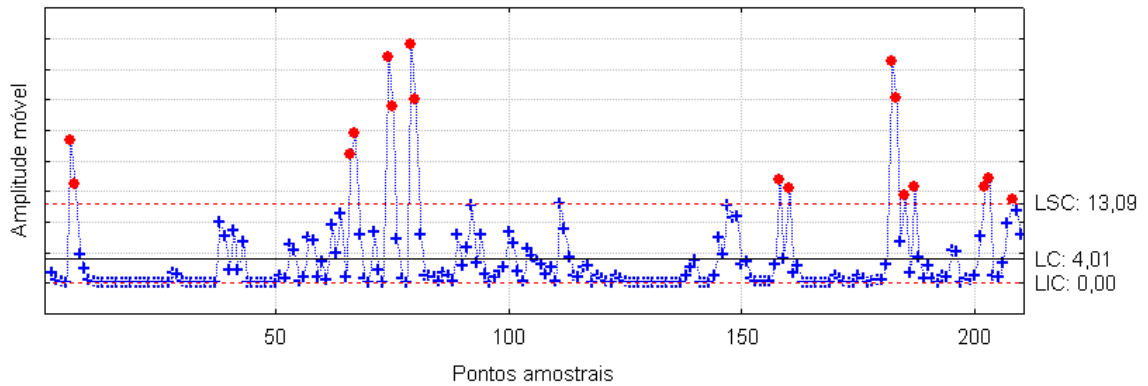


Figura 83 – Gráfico \overline{MR} aplicado a volatilidade do modelo misto AR(3)-ARCH(1) para a variável TATQ2

Ao simular a exclusão dos pontos amostrais fora de controle, tem-se outra visão da instabilidade. As Figuras 84 e 85 apresentam essa simulação, sem os *runs tests*.

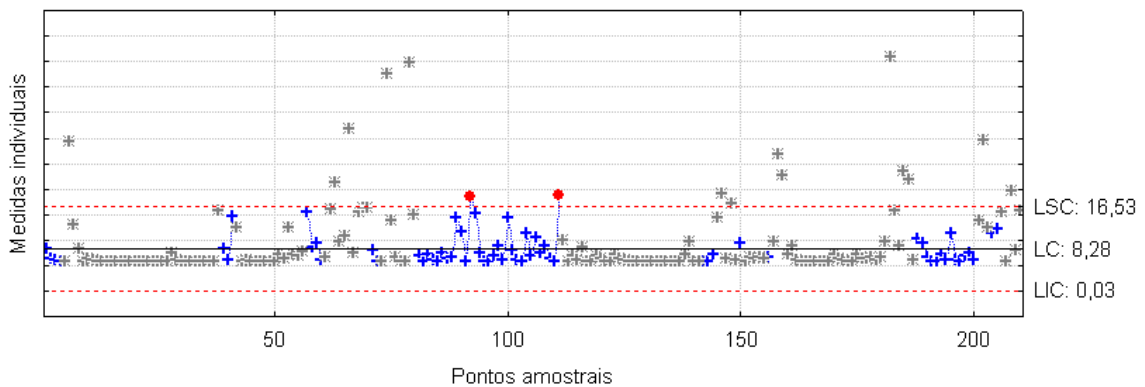


Figura 84 – Gráfico \bar{x} aplicado a volatilidade do modelo misto AR(3)-ARCH(1) para a variável TATQ2 com a exclusão dos pontos amostrais fora de controle

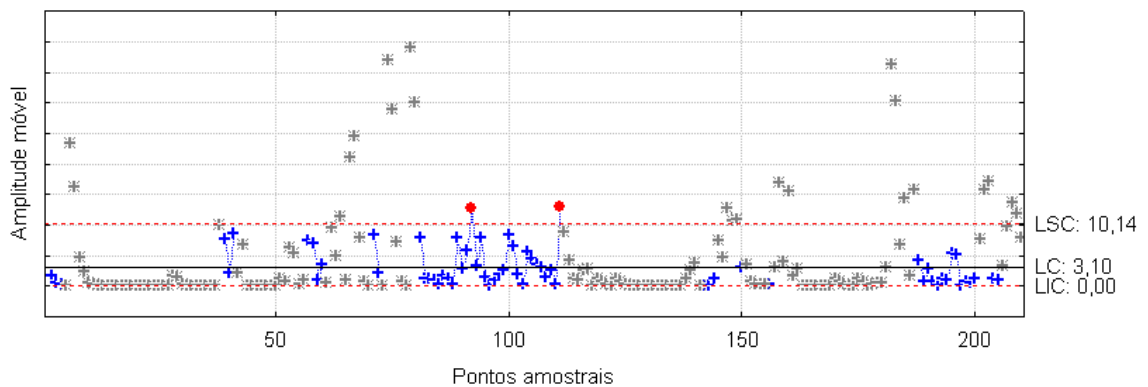


Figura 85 – Gráfico \overline{MR} aplicado a volatilidade do modelo misto AR(3)-ARCH(1) para a variável TATQ2 com a exclusão dos pontos amostrais fora de controle

4.3.8 Comentários gerais da análise

Com a aplicação da metodologia proposta, foi possível monitorar o comportamento das variáveis, TATQ1 e TATQ2, que compõem uma das características críticas controladas criteriosamente no processo de lavagem de garrafas. A modelagem realizada não teve apenas o objetivo de apresentar o valor futuro para cada uma das variáveis analisadas, mas também propiciou a detecção de pontos amostrais fora de controle.

Com relação aos resultados obtidos após a aplicação dos gráficos de controle \bar{x} , apresentam-se as seguintes considerações para a variável TATQ1:

- Nos dados originais da variável, havia vinte e seis pontos amostrais fora dos limites de controle.
- Nos resíduos do modelo ARMA(1,1), constataram-se apenas dois pontos amostrais fora dos limites de controle.

E, para a variável TATQ2 observam-se as seguintes considerações:

- Nos dados originais da variável, havia vinte e oito pontos amostrais fora dos limites de controle.
- Nos resíduos do modelo AR(3)-ARCH(1), constataram-se oito pontos amostrais fora dos limites de controle.
- E, na volatilidade obtida através do modelo misto, verificaram-se onze pontos amostrais fora dos limites de controle. Isso mostra que o processo apresenta uma variação que deveria ser levada em consideração no momento de sua análise. Não foi possível detectar nesse estudo a quantificação da influência da volatilidade no processo.

As afirmações citadas acima revelaram que, o modelo matemático estimado aos dados coletados para as variáveis, TATQ1 e TATQ2, eliminou o efeito da autocorrelação serial. Quanto à variável TATQ2, estimada por um modelo misto, apenas o efeito da autocorrelação foi eliminado, porém necessita-se quantificar a influência da volatilidade no processo. Assim, justifica-se a acurácia do modelo misto, o qual evidenciou uma melhora significativa na detecção de pontos amostrais fora dos limites de controle em comparação com o modelo AR(3).

4.4 Comentários gerais do capítulo

Neste capítulo realizou-se a parte prática da pesquisa, efetuando o ajuste de modelos matemáticos para os estudos de casos analisados. Com isto, exemplificou-se as etapas metodológicas descritas no capítulo 3.

Nos dados coletados pode-se observar uma grande variabilidade nos valores da amostra, sendo confirmada ao traçar os gráficos de controle \bar{x} e \overline{MR} . Assim, os processos produtivos apresentavam-se instáveis, principalmente pela estrutura de autocorrelação serial.

Contudo, após realizar a estimação dos modelos representativos às variáveis utilizando os modelos de previsão empregados para cada pesquisa ex-post-facto, encontraram-se modelos matemáticos capazes de modelar satisfatoriamente os dados e, assim, analisar a estabilidade do processo por meio de gráficos de controle.

Como todas as variáveis originais eram autocorrelacionadas, os modelos matemáticos estimados eliminaram essa autocorrelação, possibilitando uma análise mais concreta sobre os processos produtivos estudados.

5 CONCLUSÕES E RECOMENDAÇÕES

Este capítulo faz o encerramento das idéias e análises desenvolvidas durante o trabalho. Desta forma, inicialmente, serão apresentadas conclusões, buscando relacionar o referencial teórico do trabalho com as percepções levantadas durante a pesquisa. Em seguida, serão apresentadas recomendações relacionadas a aspectos de modelos de previsão e de possíveis desdobramentos para outras pesquisas.

5.1 Conclusões

A presente pesquisa possui caráter relevante no setor industrial por apresentar uma metodologia alternativa às técnicas tradicionais de Controle Estatístico de Qualidade. A inovação é verificar se existe volatilidade, até o momento trabalhada em variáveis econômicas, em séries de processos produtivos industriais. Essa nova abordagem, muito difundida internacionalmente, possui o diferencial de ser aplicada em séries de processos produtivos industriais, possibilitando uma melhoria no controle de qualidade.

Após a execução das etapas metodológicas conclui-se que, o procedimento proposto, de primeiro retirar a autocorrelação serial dos dados para depois ajustá-los, pois todos os processos produtivos avaliados apresentavam-se autocorrelacionados, mostrou-se adequado estatisticamente. Para tratar esse efeito presente nos dados originais, utilizou-se modelos de previsão, lineares e não-lineares.

Aplicando os modelos lineares, ARIMA e de regressão dinâmica multivariada, a autocorrelação serial foi eliminada, indicando que os resíduos do modelo matemático encontrado apresentam-se mais estáveis em relação aos dados originais. Além de encontrar um modelo matemático que eliminasse o efeito da autocorrelação serial, de modo a possibilitar a aplicação dos gráficos de controle, constatou-se que alguns resíduos oriundos desses modelos apresentavam volatilidade.

Ao identificar a presença de volatilidade em processos produtivos aplicaram-se modelos não-lineares em resíduos que continham esse fenômeno. Para modelar essa volatilidade estimou-se modelos matemáticos por meio da modelagem mista ARIMA-ARCH, utilizando os resíduos quadráticos oriundos modelos matemáticos lineares.

Com a série livre de autocorrelação, tanto nos resíduos como nos resíduos quadráticos, foram elaborados gráficos de controle nos dados originais, nos resíduos do modelo matemático e na volatilidade, resultante do modelo misto. Ao comparar esses gráficos, constatou-se que, os dados originais apresentaram mais pontos amostrais fora dos limites de controle em relação aos resíduos do modelo matemático. Porém, ao monitorar a volatilidade por meio dos gráficos de controle, observou-se mais pontos amostrais fora dos limites de controle em relação aos resíduos oriundos da modelagem linear e não-linear, evidenciando que o processo apresenta uma variação que deveria ser levada em consideração. Desta forma, foi possível mostrar que, processos produtivos possuem volatilidade, e, que essa volatilidade influencia para que o processo sofra alterações quanto a sua estabilidade, fazendo com que o mesmo esteja fora de controle.

Analisando os gráficos de controle pode-se constatar que, tanto os modelos lineares com os não-lineares foram eficazes no tratamento da autocorrelação serial, pois captaram as variações que haviam nas variáveis referentes aos processos produtivos analisados. Sendo assim, tem-se uma nova perspectiva de aprendizagem sobre o processo produtivo através das informações contidas na estrutura de autocorrelação serial, ignorada pelo modelo clássico de monitoramento.

Além disso, a investigação da estabilidade do processo passou também pela verificação da influência da volatilidade. Aplicando os gráficos de controle, detectou-se que o desempenho do processo produtivo foi influenciado pela volatilidade, ocasionando mais pontos amostrais fora de controle estatístico. O que não se pode detectar neste estudo foi a quantificação dessa influência no processo produtivo.

Os resultados apontam que o objetivo geral proposto de modelar processos produtivos industriais autocorrelacionados com o fim de analisar a estabilidade desses processos foi alcançado, sendo de fundamental importância para as pessoas envolvidas diretamente no processo como forma de fornecer informações gerenciais quanto à ação de medidas corretivas, possibilitando alcançar a melhoria contínua.

No caso específico desta pesquisa, optou-se por ajustar um modelo de regressão dinâmica multivariada, considerando um conjunto de variáveis do processo de fundição de

moldagem em areia verde para a empresa A; um modelo misto ARMA- ARCH a uma série temporal de um processo de óleo de soja bruto degomado para a empresa B, um modelo ARMA para uma das séries temporais da empresa C e um modelo misto AR-ARCH para a outra série temporal, relacionadas ao processo de lavagem de garrafas destinadas à produção de bebidas.

Na pesquisa ex-post-facto da Empresa A, os modelos matemáticos encontrados para as variáveis: teor de umidade, permeabilidade e rcv foram: teor de umidade = teor de umidade_{t-1} + teor de umidade_{t-2} + teor de umidade_{t-3} + permeabilidade_{t-1} + rcv_{t-2}, permeabilidade = permeabilidade_{t-1} + permeabilidade_{t-2} + permeabilidade_{t-3} + rcv_{t-3} e rcv = teor de umidade_{t-2} + permeabilidade_{t-3} + rcv_{t-1} + rcv_{t-2} + rcv_{t-3}, conforme a Tabela 2. Já, na pesquisa ex-post-facto da Empresa B, os modelos matemáticos para a variável, índice de acidez, foram: ARMA(2,1) para a média condicional e ARMA(2,1)-ARCH(1) estimado para a média e variância condicional, pois os resíduos quadráticos do modelo linear continham volatilidade. E, na pesquisa ex-post-facto da Empresa C, o modelo matemático para a variável, TATQ1, foi: ARMA(1,1) estimando a média condicional, e, os modelos matemáticos para a variável, TATQ2, foram: AR(3) para a média condicional e AR(3)-ARCH(1) para a média e variância condicional, pois os resíduos quadráticos do modelo linear também apresentaram volatilidade.

Neste trabalho, sustenta-se que os modelos de previsão lineares e não-lineares podem ser aplicados conjuntamente através de seus resíduos para estimar as variáveis em estudo da melhor forma possível, visando, com isso, a um crescimento de informações para a correta tomada de decisão. Pôde-se detectar, ainda, que houve um aumento de desempenho na monitoração dos processos produtivos comparado à monitoração tradicional.

Os modelos encontrados foram capazes de modelar satisfatoriamente os dados, possibilitando uma análise de seu comportamento, de mesmo modo que tornou-se possível fornecer previsões em curto prazo, sinalizando informações importantes no processo de tomada de decisões. Foram identificados vários pontos amostrais fora dos limites de controle nos dados originais, mas, após o ajuste dos dados por meio dos modelos de previsão, justificase a adequação dos modelos matemáticos estimados para cada pesquisa ex-post-facto.

Um dos fatores preponderantes para a utilização do CEP como ferramenta de análise foi a possibilidade de visualizar graficamente o comportamento da variável ao longo do tempo, permitindo, inclusive, detectar possíveis variações no processo. Além disso, a previsão é um recurso capaz de fornecer informações e subsídios para uma subsequente tomada de decisão. Assim, o emprego do CEP permitiu revelar o comportamento do processo,

mostrando que grandes variações são capazes de influenciar na qualidade do produto, sugerindo, então, a possibilidade de redução de custos através da aplicação de medidas corretivas.

Uma empresa que trabalhe com qualidade, utiliza racionalmente seus recursos e os custos operacionais são baixos. Dessa forma, pode-se dizer que o retorno do investimento em qualidade vem principalmente através da redução daquela parcela dos custos operacionais que decorrem da má qualidade.

Buscou-se, através do estudo proposto, evidenciar possibilidades de redução de custos associados aos produtos defeituosos relacionados a falhas internas antes de estes serem despachados para os clientes, como: refugo e retrabalho aos três processos produtivos estudados, principalmente quando a produção é bastante expressiva. Portanto, se as variáveis, em análise, estiverem fora de controle, ocasionarão um dispêndio ainda mais elevado de recursos financeiros, os quais deixarão de ser empregados em outros setores da empresa e em outras melhorias, seja de tecnologia ou de recursos humanos.

5.2 Recomendações

Visando a dar continuidade à proposta apresentada neste trabalho, algumas recomendações são feitas:

- avaliar a capacidade do processo produtivo para verificar o desempenho desse processo quando ele está operando sob controle estatístico;
- aplicar gráficos de controle EWMA para detectar se há pequenas variações no processo produtivo;
- aplicar o modelo GARCH, o qual descreve a variância condicional de uma série como dependendo de uma constante, da volatilidade passada e das previsões passadas da variância. Nessa equação de regressão, o segundo parâmetro, representado pelo ARCH, constitui o coeficiente de reação da volatilidade, e, o terceiro parâmetro, representado pelo GARCH, constitui o coeficiente de persistência da volatilidade;
- aplicar o modelo assimétrico EGARCH, que permite a absorção de impactos assimétricos com uma vantagem adicional: a possibilidade de que alguns

coeficientes sejam negativos. Nesse caso, os parâmetros do modelo podem ser negativos.

6 REFERÊNCIAS BIBLIOGRÁFICAS

AGOSTINHO, O. L.; VILLELA, R. C.; BUTTON, S. T. Processos de fabricação e planejamento de processos. In: Faculdade de Engenharia Mecânica. Campinas: UNICAMP, 2004. Disponível em: <www.fem.unicamp.br/~sergio1/graduacao/EM335/apostila.pdf>. Acesso em: 10 jun. 2009.

ALWAN, L. C.; ROBERTS, H. V. Times series modeling for statistical process control. **Journal of Business & Economics Statistics**, v. 6, n. 1, p. 87-95, 1988.

ARÊDES, A. F.; PEREIRA, M. W. G. Potencialidade da utilização de modelos de séries temporais na previsão do preço do trigo no estado do Paraná. **Revista de Economia Agrícola**, v. 55, n. 1, p. 63-76, jan-jun. 2008.

BAILLIE, R.; CHUNG, C. F.; TIESLAU, M. A. Analyzing inflation by the fractionally integrated ARFIMA-GARCH model. **Journal of Applied Econometrics**, v. 11, p. 23-40, 1996.

BAYER, F. M.; SOUZA, A. M. Modelagem e previsão da taxa de ocupação hospitalar por meio de equações simultâneas: uma ferramenta de gestão. **Revista GEPROS: Gestão da Produção, Operações e Sistemas**, ano 2, v. 3, p. 85-96, mai-jun. 2007.

BAYER, F. M. **Previsão do preço e da volatilidade de commodities agrícolas, por meio de modelos ARFIMA-GARCH**. 2008. 83f. Dissertação (Mestrado em Engenharia de Produção) – Universidade Federal de Santa Maria, Santa Maria, 2008.

BEM, A. B. **A utilização da regressão dinâmica para prever a demanda por matrícula no ensino fundamental em Santa Catarina**. 1998. 183f. Dissertação (Mestrado em Engenharia de Produção) – Universidade Federal de Santa Catarina, Florianópolis, 1998.

BERTI, A. F. **Estimação de volatilidade diária com dados de alta frequência: aplicações ao cálculo do valor em risco do IBOVESPA**. 2005. 74f. Dissertação (Mestrado em Estatística) – Universidade de São Paulo, São Paulo, 2005.

BOLLERSLEV, T. Generalized autoregressive conditional heteroskedasticity. **Journal of Econometrics**, v. 31, p. 307-327, 1986.

BOX, G. E. P.; JENKINS, G. M. **Time series analysis: forecasting and control**. San Francisco: Holden-Day, 1970.

BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C. **Time series analysis: forecasting and control**. 3.ed. San Francisco: Holden-Day, 1994.

BOX, G.; LUCEÑO, A. **Statistical control by monitoring and feedback adjustment**. Canadá: John Wiley & Sons, Inc., 1997.

BOX, G. E. P.; PIERCE, D. A. Distribution of autocorrelations in autoregressive moving average models. **Journal of The American Statistical Association**, v. 65, p. 1509-1526, 1970.

BUENO, R. L. S. **Econometria de séries temporais**. São Paulo: Cengage Learning, 2008.

CAMPOS, K. C. Análise da volatilidade de preços de produtos agropecuários no Brasil. **Revista de Economia de Agronegócio**, v. 5, n. 3, p. 303-328, 2007.

CHAREMZA, W. W.; DEADMAN, D. **New directions in econometric practice general to specific modelling, cointegration and vector autoregression**. 2.ed. Cheltenham: Edward Elgar Publisher Limited, 1997.

CLARO, F. A. E.; COSTA, A. F. B.; MACHADO, M. A. G. Gráficos de controle de EWMA e de X-barra para monitoramento de processos autocorrelacionados. **Revista Produção**, v. 17, n. 3, p. 536-546, set-dez. 2007.

COSTA, A. F. B.; EPPRECHT, E. K.; CARPINETTI, L. C. R. **Controle estatístico da qualidade**. São Paulo: Atlas, 2004.

COSTA, P. H. S.; BAIDYA, T. K. N. Propriedades estatísticas das séries de retornos das principais ações brasileiras. **Pesquisa Operacional**, v. 21, n. 1, p. 61-87, 2001.

DEL CASTILLO, E. **Statistical control adjustment for quality control**. Canadá: John Wiley & Sons, Inc., 2002.

EHLERS, R. S. Análise de séries temporais. In: Notas de aula de séries temporais. São Paulo: USP, 2007. Disponível em: <<http://www.icmc.usp.br/~ehlers/notas/stemp.pdf>>. Acesso em: 22 jan. 2009.

ÉLIN DUXUS Consultoria Ltda. In: Manual e ajuda para o add-in Metrixus: funções quantitativas para mercado de capitais. Disponível em: <<http://www.duxus.com.br/produtos/tecnologia/addin/metrixus/manual/ajuda.php>>. Acesso em: 16 fev. 2009.

ENDERS, W. **Applied econometric time series**. New York: John Wiley and Sons, Inc., 1995.

ENGLE, R. F. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. **Econometria**, v. 50, n. 4, p. 987-1008, 1982.

FRANÇA, A. L. Brasil recicla 53,4% das embalagens feitas de PET. **Gazeta Mercantil**, São Paulo, 28 mai. 2009. Disponível em: <<http://indexet.investimentosenoticias.com.br/arquivo/2009/05/28/464/Brasil-recicla-53-4-das-embalagens-feitas-de-PET.html>>. Acesso em: 18 jul. 2009.

GOULART, C. P. **Risco de Câmbio no Mercado Interbancário Brasileiro: Um Estudo Comparativo entre Modelos de Predição de Volatilidade**. 2004. 202f. Dissertação (Mestrado em Administração) – Universidade Federal de Minas Gerais, Belo Horizonte, 2004.

GUJARATI, D. N. **Econometria básica**. São Paulo: Makron Books, 2000.

HAMILTON, J. **Time series analysis**. New Jersey: Princeton University Press, 1994.

HILL, C.; GRIFFITHS, W.; JUDGE, G. **Econometria**. São Paulo: Saraiva, 1999.

JURAN, J. M. **A qualidade desde o projeto**. São Paulo: Pioneira, 1992.

KMENTA, J. **Elementos de Econometria: teoria econométrica básica**. 2.ed. São Paulo: Atlas, 1988.

LAMOUNIER, W. M. Análise da volatilidade dos preços no mercado de *spot* de cafés do Brasil. **Organizações Rurais & Agroindustriais**, v. 8, n. 2, p. 160-175, 2006.

LJUNG, G. M; BOX, G. E. P. On a measure of lack of fit in time series models. **Biometrika**, v. 65, n. 2, p. 297-303, 1978.

LÜTKEPOHL, H. **Introduction to multiple time series analysis**. 2.ed. Berlin: Springer-Verlag, 1991.

MADDALA, G. S. **Introduction to econometrics**. 2.ed. New Jersey: Prentice Hall, 1992.

MAKRIDAKIS, S. G.; WHEELWRIGHT, S. C.; HYNDMAN, R. J. **Forecasting: methods and applications**. 3ed. New York: John Willey & Sons, Inc., 1998.

MALHOTRA, N. K. **Pesquisa de Marketing: uma orientação aplicada**. 3.ed. Porto Alegre: Bookman, 2001.

MALUCHE, M. A. **Modelo de controle de gestão para a pequena empresa como garantia da qualidade**. 2000. 247f. Dissertação (Mestrado em Engenharia de Produção) – Universidade Federal de Santa Catarina, Florianópolis, 2000.

MICHELSON, D. K. **Statistical process control for correlated data**. 1994. 100f. Dissertation submitted to the Office of Graduate Studies (Doctor of Philosophy) – Texas A&M University, Bryan, 1994.

MINGOTI, S. A.; YASSUKAWA, F. R. S. Uma comparação de gráficos de controle para a média de processos autocorrelacionados. **Revista Eletrônica Sistemas & Gestão**, v.3, n. 1, p. 55-73, jan-abr. 2008.

MONTGOMERY, D. C. **Introdução ao controle estatístico da qualidade**. 4.ed. Rio de Janeiro: LTC, 2004.

_____. **Introduction to statistical quality control**. 3.ed. New York: John Wiley & Sons, Inc., 1997.

MONTGOMERY, D. C; RUNGER, G. C. **Estatística aplicada e probabilidade para engenheiros**. 2.ed. Rio de Janeiro: LTC, 2003.

_____. **Estatística aplicada e probabilidade para engenheiros**. 4.ed. Rio de Janeiro: LTC, 2009.

MOREIRA JUNIOR, F. J.; TEN CATEN, C. S. Estudo sobre o efeito da autocorrelação de modelos AR(1) no controle estatístico de processo. In: XXIV ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO, 2004, Florianópolis. **Anais...** Florianópolis: Associação Brasileira de Engenharia de Produção, 2004. p. 1705-1712.

MORETTIN, P. A. **Econometria financeira: um curso em séries temporais financeiras**. São Paulo: ABE, 2006.

_____. **Econometria financeira: um curso em séries temporais financeiras**. São Paulo: Blucher, 2008.

MORETTIN, P. A.; TOLOI, C. M. C. **Análise de séries temporais**. 2.ed. São Paulo: Edgard Blücher, 2004.

MORRETO, E.; FETT, R. **Tecnologia de óleos e gorduras vegetais na indústria de alimentos**. 1.ed. São Paulo: Varela, 1998.

PRIESTLEY, M. B. **Spectral Analysis and Time Series**. London: Academic Press, 1981.

RAMOS, A. W.; HO, L. L. Procedimentos inferenciais em índices de capacidade para dados autocorrelacionados via *bootstrap*. **Revista Produção**, v. 13, n. 3, p. 50-62, 2003.

REINSEL, G. C. **Elements of multivariate time series analysis**. New York: Springer-Verlag, 1993.

Resolução RDC nº 482, de 23 de setembro de 1999. Aprova o Regulamento Técnico para Fixação de Identidade e Qualidade de Óleos e Gorduras Vegetais. **Diário Oficial da União**, Poder Executivo, 13 out. 1999. Disponível em: <<http://e-legis.bvs.br/leisref/public/showAct.php?id=135>>. Acesso em: 12 jan. 2009.

RUSSO, S. L. **Gráficos de controle para variáveis não-conformes autocorrelacionadas**. 2002. 166f. Tese (Doutorado em Engenharia de Produção) – Universidade Federal de Santa Catarina, Florianópolis, 2002.

RYAN, T.P. **Statistical methods for quality improvement**. New York: John Wiley & Sons, Inc., 1989.

SÁFADI, T. Uso de séries temporais na análise de vazão de água na represa de Furnas. **Ciência Agrotécnica**, v. 28, n.1, p. 142-148, jan-fev. 2004.

SÁFADI, T.; ANDRADE FILHO, M. G. Abordagem bayesiana de modelos de séries temporais. In: 12ª ESCOLA DE SÉRIES TEMPORAIS E ECONOMETRIA, 2007, Gramado. **Minicurso...** Gramado: Associação Brasileira de Estatística, 2007.

SALLES, A. A. Análise multivariada do risco sistemático dos principais mercados de ações da América Latina: um enfoque bayesiano. In: XXVI ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO, 2006, Fortaleza. **Anais...** Fortaleza: Associação Brasileira de Engenharia de Produção, 2006. p. 1-9.

SANVICENTE, A. Z. Competição na indústria automobilística brasileira. In: Página pessoal. São Paulo: FEA/USP, 1998. Disponível em: <<http://www.ead.fea.usp.br/professores/sanvicente/pesq.htm>>. Acesso em: 10 ago. 2009.

SILVA, W. S.; SÁFADI, T.; CASTRO JÚNIOR, L. G. de. Uma análise empírica da volatilidade do retorno de commodities agrícolas utilizando modelos ARCH: os casos do café e da soja. **Revista de Economia e Sociologia Rural**, v. 43, n. 1, p. 120-134, 2005.

SILVA, W. V. da. **Curso de análise de regressão linear: teoria e aplicações**. Santa Maria: Universidade Federal de Santa Maria, 2000.

SILVA, W. V.; FONTANINI, C. A. C.; DEL CORSO, J. M. Garantia da qualidade do café solúvel com o uso do gráfico de controle de somas acumuladas. **Revista Produção on-line**, v. 7, n. 2, p. 43-63, 2007.

SOARES, I. G.; CASTELAR, I. **Econometria aplicada com o uso de EViews**. Fortaleza: UFC/CAEN, 2003.

SOUZA, A. M.; SAMOBYL, R. W.; MALAVÉ, C. O. Multivariate feedback control: an application in a productive process. **Computers & Industrial Engineering**, v. 46, p. 837-850, 2004a.

_____. Aplicação de um modelo paramétrico multivariado para o controle da temperatura de fornos de túnel. **Revista Produção**, v. 14, n. 2, p. 82-94, 2004b.

SOUZA, G. P.; SAMOBYL, R. W.; MIRANDA, R. G. **Métodos simplificados de previsão empresarial**. 1.ed. Rio de Janeiro: Ciência Moderna, 2008.

SOUZA, M. V. A. **Modelagem e previsão da volatilidade dos retornos do café arábica produzido no Brasil**. 2005. 70f. Dissertação (Mestrado em Biometria) – Universidade Federal Rural de Pernambuco, Recife, 2005.

ZANINI, A.; SOUZA, R. C.; PEDREIRA, C. E. Redes neurais e regressão dinâmica: um modelo híbrido para previsão de curto prazo da demanda de gasolina automotiva no brasil. In: XXXII SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, 2000, Viçosa. **Anais...** Viçosa: Sociedade Brasileira de Pesquisa Operacional, 2000. p. 1928-1946.

ZANINI, R. R. **A utilização dos gráficos de controle:** uma aplicação na área da saúde. 2006. 111p. Dissertação (Mestrado em Engenharia de Produção) – Universidade Federal de Santa Maria, Santa Maria, 2006.

ANEXOS

ANEXO A - Fatores para construção de gráficos de controle para variáveis

Observações na Amostra, <i>n</i>	Gráfico para Médias			Gráficos para Desvios Padrão						Gráficos para Amplitudes						
	Fatores para Limites de Controle			Fatores para Linha Central		Fatores para Limites de Controle				Fatores para Linha Central		Fatores para Limites de Controle				
	<i>A</i>	<i>A</i> ₂	<i>A</i> ₃	<i>c</i> ₄	1/ <i>c</i> ₄	<i>B</i> ₃	<i>B</i> ₄	<i>B</i> ₅	<i>B</i> ₆	<i>d</i> ₂	1/ <i>d</i> ₂	<i>d</i> ₃	<i>D</i> ₁	<i>D</i> ₂	<i>D</i> ₃	<i>D</i> ₄
2	2,121	1,880	2,659	0,7979	1,2533	0	3,267	0	2,606	1,128	0,8865	0,853	0	3,686	0	3,267
3	1,732	1,023	1,954	0,8862	1,1284	0	2,568	0	2,276	1,693	0,5907	0,888	0	4,358	0	2,575
4	1,500	0,729	1,628	0,9213	1,0854	0	2,266	0	2,088	2,059	0,4857	0,880	0	4,698	0	2,282
5	1,342	0,577	1,427	0,9400	1,0638	0	2,089	0	1,964	2,326	0,4299	0,864	0	4,918	0	2,115
6	1,225	0,483	1,287	0,9515	1,0510	0,030	1,970	0,029	1,874	2,534	0,3946	0,848	0	5,078	0	2,004
7	1,134	0,419	1,182	0,9594	1,0423	0,118	1,882	0,113	1,806	2,704	0,3698	0,833	0,204	5,204	0,076	1,924
8	1,061	0,373	1,099	0,9650	1,0363	0,185	1,815	0,179	1,751	2,847	0,3512	0,820	0,388	5,306	0,136	1,864
9	1,000	0,337	1,032	0,9693	1,0317	0,239	1,761	0,232	1,707	2,970	0,3367	0,808	0,547	5,393	0,184	1,816
10	0,949	0,308	0,975	0,9727	1,0281	0,284	1,716	0,276	1,669	3,078	0,3249	0,797	0,687	5,469	0,223	1,777
11	0,905	0,285	0,927	0,9754	1,0252	0,321	1,679	0,313	1,637	3,173	0,3152	0,787	0,811	5,535	0,256	1,744
12	0,866	0,266	0,886	0,9776	1,0229	0,354	1,646	0,346	1,610	3,258	0,3069	0,778	0,922	5,594	0,283	1,717
13	0,832	0,249	0,850	0,9794	1,0210	0,382	1,618	0,374	1,585	3,336	0,2998	0,770	1,025	5,647	0,307	1,693
14	0,802	0,235	0,817	0,9810	1,0194	0,406	1,594	0,399	1,563	3,407	0,2935	0,763	1,118	5,696	0,328	1,672
15	0,775	0,223	0,789	0,9823	1,0180	0,428	1,572	0,421	1,544	3,472	0,2880	0,756	1,203	5,741	0,347	1,653
16	0,750	0,212	0,763	0,9835	1,0168	0,448	1,552	0,440	1,526	3,532	0,2831	0,750	1,282	5,782	0,363	1,637
17	0,728	0,203	0,739	0,9845	1,0157	0,466	1,534	0,458	1,511	3,588	0,2787	0,744	1,356	5,820	0,378	1,622
18	0,707	0,194	0,718	0,9854	1,0148	0,482	1,518	0,475	1,496	3,640	0,2747	0,739	1,424	5,856	0,391	1,608
19	0,688	0,187	0,698	0,9862	1,0140	0,497	1,503	0,490	1,483	3,689	0,2711	0,734	1,487	5,891	0,403	1,597
20	0,671	0,180	0,680	0,9869	1,0133	0,510	1,490	0,504	1,470	3,735	0,2677	0,729	1,549	5,921	0,415	1,585
21	0,655	0,173	0,663	0,9876	1,0126	0,523	1,477	0,516	1,459	3,778	0,2647	0,724	1,605	5,951	0,425	1,575
22	0,640	0,167	0,647	0,9882	1,0119	0,534	1,466	0,528	1,448	3,819	0,2618	0,720	1,659	5,979	0,434	1,566
23	0,626	0,162	0,633	0,9887	1,0114	0,545	1,455	0,539	1,438	3,858	0,2592	0,716	1,710	6,006	0,443	1,557
24	0,612	0,157	0,619	0,9892	1,0109	0,555	1,445	0,549	1,429	3,895	0,2567	0,712	1,759	6,031	0,451	1,548
25	0,600	0,153	0,606	0,9896	1,0105	0,565	1,435	0,559	1,420	3,931	0,2544	0,708	1,806	6,056	0,459	1,541

Para *n* > 25,

$$A = \frac{3}{\sqrt{n}} \quad A_3 = \frac{3}{c_4 \sqrt{n}} \quad c_4 \approx \frac{4(n-1)}{4n-3}$$

$$B_3 = 1 - \frac{3}{c_4 \sqrt{2(n-1)}} \quad B_4 = 1 + \frac{3}{c_4 \sqrt{2(n-1)}}$$

$$B_5 = c_4 - \frac{3}{\sqrt{2(n-1)}} \quad B_6 = c_4 + \frac{3}{\sqrt{2(n-1)}}$$

Fonte: Montgomery, 2004