



Paulo Henrique Coelho Maranhão

**Gráficos de Controle de Variáveis
Transformadas para o Monitoramento
de Processos Multivariados**

Tese de Doutorado

Tese apresentada ao Programa de Pós-graduação em Engenharia de Produção do Departamento de Engenharia Industrial da PUC-Rio como parte dos requisitos parciais para obtenção do título de Doutor em Engenharia de Produção.

Orientador: Prof. Eugenio Kahn Epprecht

Rio de Janeiro
Setembro de 2012



Paulo Henrique Coelho Maranhão

**Gráficos de Controle de Variáveis
Transformadas para o Monitoramento
de Processos Multivariados**

Tese apresentada como requisito parcial para
obtenção do grau de Doutor pelo programa de Pós-
Graduação em Engenharia de Produção da PUC-
Rio. Aprovada pela comissão examinadora abaixo
assinada

Prof. Eugenio Kahn Epprecht

Orientador

Departamento de Engenharia Industrial - PUC-Rio

Prof. Flávio Sanson Fogliatto

Departamento de Engenharia de Produção - UFRGS

Prof^a. Maysa Sacramento de Magalhães

Escola Nacional de Ciências Estatísticas - ENCE

Prof. Antonio Fernando de Castro Vieira

Departamento de Engenharia Industrial - PUC-Rio

Luiz Henrique Abreu Dal Bello

Centro Tecnológico do Exército - CTEx

Prof. José Eugenio Leal

Coordenador Setorial do Centro Técnico-Científico – PUC-Rio

Rio de Janeiro, 28 de setembro de 2012

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Paulo Henrique Coelho Maranhão

Graduou-se em Estatística pela UFC (Universidade Federal do Ceará) no final de 1994. Obteve o título de Mestre em Estatística (área de concentração: Estatística Aplicada) pela Universidade Federal de Pernambuco (UFPE) com a dissertação intitulada “Modelos para o Câncer Infantil em Pernambuco” em 1998. Atualmente é Professor do Instituto Militar de Engenharia (IME).

Ficha Catalográfica

Maranhão, Paulo Henrique Coelho

Gráficos de controle de variáveis transformadas para o monitoramento de processos multivariados / Paulo Henrique Coelho Maranhão ; orientador: Eugenio Kahn Epprecht. – 2012.

121 f. : il. (color.) ; 30 cm

Tese (doutorado)–Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Industrial, 2012.

Inclui bibliografia

1. Engenharia Industrial – Teses. 2. Gráficos de controle. 3. Processos multivariados. 4. Direções conhecidas. 5. Variáveis transformadas. 6. Componentes principais. I. Epprecht, Eugenio Kahn. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Industrial. III. Título.

CDD: 658.5

A minha esposa Flávia pelo o amor
e dedicação.

Aos meus Filhos Pedro e Diego,
fontes de luz no meu caminho.

Agradecimentos

Ao meu orientador e amigo Professor Doutor Eugenio Kanh Epprecht pela imensa paciência e pela excelente orientação, sem a qual esse trabalho jamais seria possível. Por acreditar que em mim, mesmo nos momentos de maiores adversidades.

A minha Esposa Flávia e aos meus filhos, Pedro e Diego, pelo o amor e compreensão, em especial, por todos os dias que deixei de dedicar meu tempo a eles por está envolvido em atividades relacionadas direta ou indiretamente ao doutorado.

Aos meus pais, Ana e Elias (In memorian), pelo o incentivo e por acreditarem que a educação era a maior riqueza que eles poderiam me dar.

A minha sogra, Iracy, por todo o suporte dispensados a minha família e ao meu amigo e sogro Dirceu (In memorian) por seu carinho e respeito.

Aos professores que participaram da Comissão Examinadora desta tese, do exame de qualificação e de proposta de tese pela valiosa contribuição a este trabalho.

Aos Professores do Departamento de Engenharia Industrial que contribuíram para a formação do conhecimento.

Aos funcionários do Departamento de Engenharia Industrial, pela presteza e cordialidade a qual sempre fui tratado.

Aos amigos, Professores João Maurício de Araújo Mota e Ronald Targino Norjosa da Universidade Federal do Ceará (UFC) por estarem sempre disponíveis e dispostos a ajudar.

A todos os amigos que contribuíram de alguma forma para a realização desta tese.

À PUC-Rio e ao Instituto Militar de Engenharia (IME) pelo o apoio de bolsa firmado em convênio.

Resumo

Maranhão, Paulo Henrique Coelho; Epprecht, Eugenio Kahn (Orientador). **Gráficos de Controle de Variáveis Transformadas para o Monitoramento de Processos Multivariados.** Rio de Janeiro, 2012. 121p. Tese de Doutorado – Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

A maioria dos trabalhos que propõem esquemas de Controle Estatístico de Processos Multivariados (CEPM) e que analisam o desempenho desses esquemas considera alterações nas variáveis observadas. Alguns autores mostraram que, quando alterações no vetor de médias de um processo multivariado ocorrem tipicamente em direções previsíveis, a estatística mais eficaz para o monitoramento do processo é o valor da projeção do vetor de observações (ou de sua média amostral) em cada uma dessas direções. Este trabalho propõe um método para o monitoramento de processos multivariados nos quais alterações nos parâmetros do processo são devidas a causas especiais que afetam variáveis não observáveis e ocorrem em direções conhecidas (ortogonais), e compara seu desempenho com o de gráficos de Shewhart nas variáveis observadas, nas componentes principais, e com o de gráficos de T^2 no vetor de variáveis observadas. Além disso, é proposto um esquema complementar de monitoramento para detectar alterações em novas direções fora do hiperplano formado pelas direções conhecidas. Resultados obtidos por simulação mostram que o esquema proposto, de gráficos de controle em variáveis transformadas (projeções do vetor de variáveis observadas nas direções conhecidas), tem melhor desempenho na maior parte dos casos analisados. A análise de desempenho é feita supondo deslocamentos da média nas direções conhecidas (já que estes é que são as alterações conhecidas, ocasionadas pelas causas especiais) e/ou aumentos da variância nessas mesmas direções. A comparação é baseada nas probabilidades de alarme falso e de alarme verdadeiro.

Palavras-chave

Controle estatístico de processos; direções conhecidas; variáveis transformadas; componentes principais; métodos de projeção; resíduos.

Abstract

Maranhão, Paulo Henrique Coelho; Epprecht, Eugenio Kahn (Advisor). **Control Charts on Transformed Variables for Monitoring Multivariate Process**. Rio de Janeiro, 2012. 121p. Doctorate Thesis – Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

Most of the works that propose schemes of Multivariate Statistical Process Control (MSPC) and that analyze the performance of these schemes consider changes in the observed variables. Previous authors have shown that when the shifts in the mean vector of a multivariate process typically occur in predictable directions, the most effective statistics for process monitoring are the values of the projections of the vector of observations (or of the sample average vector) in each of these directions. This paper proposes a method for the monitoring of multivariate processes in which changes in the process parameters are due to special causes that affect non-observable variables and occur in (orthogonal) known directions, and compares its performance with that of Shewhart's charts on the observed variables, on the principal components, and with that of T^2 charts on the vector of observed variables. In addition, it is proposed a supplementary scheme of monitoring to detect changes in new directions outside of the hyperplane formed by known directions. Results obtained by simulation show that the proposed scheme, consisting of control charts on the transformed variables (projections of the vector of observed variables on the known directions), has better performance in most of the cases analyzed. The analysis of performance is done assuming shifts in the mean of the known directions (since these are the known changes associated to special causes) and/or increases of the variance in these same directions. The comparisons are based on the in-control and out-of-control probabilities of signal.

Keywords

Statistical process control; known directions; transformed variables; principal components; projection methods; residuals.

Sumário

1. Introdução	17
1.1. Problema Tratado e Estado da Arte	17
1.2. Justificativa	22
1.3. Contribuição Original e Relevância	23
1.4. Objetivos	24
1.4.1. Objetivo Geral	24
1.4.2. Objetivos Específicos	24
1.5. Delimitação do Trabalho	25
1.6. Organização do Texto	25
2. O Uso de Variáveis Transformadas	27
2.1. Modelo do Processo	27
2.2. Definindo as Transformações	29
2.3. Utilizando as Componentes Principais para Determinar as Direções de Variação do Vetor de Variáveis Observadas	31
2.4. Utilizando o Resíduo para o Monitoramento do Processo	31
3. Análise de Desempenho	35
3.1. Modelo Particular para o Exemplo	35
3.2. Variáveis Transformadas	38
3.3. Análise de Componentes Principais	40
3.4. Vetor Resíduo	40
3.5. Metodologia para a Análise de Desempenho	46
4. Resultados	50
4.1. Casos de Alterações na Média	50
4.1.1. Uma Única Direção de Perturbação	50
4.1.2. Duas Direções de Perturbação com Mesma Variância	55
4.1.3. Três Direções de Perturbação com Mesma Variância	60
4.1.4. Duas Direções de Perturbação com Diferentes Variâncias	63
4.2. Casos de Aumentos na Variância	66
4.2.1. Uma Única Direção de Perturbação	67
4.2.2. Duas Direções de Perturbação com Mesma Variância	68
4.2.3. Três Direções de Perturbação com Mesma Variância	70
4.2.4. Duas Direções de Perturbação com Diferentes Variâncias	73
4.3. Casos de Alterações na Média e Aumentos na Variância	74
4.3.1. Uma Única Direção de Perturbação	74
4.3.2. Duas Direções de Perturbação com Mesma Variância	75
4.3.3. Três Direções de Perturbação com Mesma Variância	78
4.3.4. Duas Direções de Perturbação com Diferentes Variâncias	79

4.4. Monitorando o Resíduo	81
5. Conclusões e Sugestões para Pesquisa Futura	91
5.1. Conclusões	91
5.2. Sugestões para Pesquisa Futura	95
Referências Bibliográficas	96
Apêndice - Análise de Componentes Principais	99
Anexo - Gráficos de Controle para o Quadrado da Norma do Vetor Resíduo	101

Lista de figuras

Figura 1	Apresentação de quatro possíveis pontos de folgas do problema	35
Figura 2	Rotação de magnitude d no sentido anti-horário	36
Figura 3	Deslocamento para a direita de magnitude d	37
Figura 4	Contração de magnitude d	37
Figura 5	Representação geométrica do vetor resíduo (\mathbf{r})	41
Figura 6	Resultado da análise gráfica e do teste de Anderson-Darling para $\sigma_e = 0,1$	82
Figura 7	Resultado da análise gráfica e do teste de Anderson-Darling para $\sigma_e = 0,5$	83
Figura 8	Resultado da análise gráfica e do teste de Anderson-Darling para $\sigma_e = 0,75$	84
Figura 9	Resultado da análise gráfica e do teste de Anderson-Darling para $\sigma_e = 1,0$	85
Figura 10	Resultado da análise gráfica e do teste de Anderson-Darling para $\sigma_e = 1,5$	86

Lista de Quadros

Quadro 1	Resultado da Análise de Componentes Principais para o caso de alterações na média com uma única direção de perturbação	50
Quadro 2	Correlações entre as variáveis observadas para o processo em controle para o caso de alterações na média com uma única direção de perturbação	51
Quadro 3	ARL para o caso de alterações na média considerando apenas uma direção de perturbação	52
Quadro 4	ARL para o caso de alterações na média considerando apenas uma direção de perturbação e ARL_0 's aproximados	54
Quadro 5	Resultado da Análise de Componentes Principais para o caso de alterações na média considerando duas direções de perturbação com mesma variância	55
Quadro 6	ARL para o caso de alterações na média considerando duas direções de perturbação com mesma variância	56
Quadro 7	ARL para o caso de alterações na média considerando duas direções de perturbação com mesma variância e ARL_0 's aproximados	59
Quadro 8	Correlações entre as variáveis observadas para o processo em controle para o caso de alterações na média considerando duas direções de perturbação com mesma variância	59
Quadro 9	Resultado da Análise de Componentes Principais para o caso de alterações na média considerando três direções de perturbação com mesma variância	60
Quadro 10	Correlações entre as variáveis observadas para o processo em controle para o caso de alterações na média considerando três direções de perturbação com mesma variância	61
Quadro 11	ARL para o caso de alterações na média considerando três direções de perturbação com mesma variância	61
Quadro 12	ARL para o caso de alterações na média considerando três direções de perturbação com mesma variância e ARL_0 's aproximados	63
Quadro 13	Resultado da Análise de Componentes Principais para o caso de alterações na média com duas direções de perturbação com diferentes variâncias	64

Quadro 14	Correlações entre as variáveis observadas para o processo em controle para o caso de alterações na média considerando duas direções de perturbação com diferentes variâncias	64
Quadro 15	ARL para o caso de alterações na média considerando duas direções de perturbação com o desvio-padrão da Rotação maior que do Deslocamento	65
Quadro 16	ARL para o caso de alterações na média considerando duas direções conhecidas com diferentes variâncias e ARL_0 's aproximados	66
Quadro 17	ARL para o caso de aumentos na variância considerando uma única direção de perturbação	67
Quadro 18	ARL para o caso de aumentos na variância considerando uma única direção conhecida considerando e ARL_0 's aproximados	68
Quadro 19	ARL para o caso de aumentos na variância considerando duas direções de perturbação com mesma variância	69
Quadro 20	ARL para o caso de aumentos na variância considerando duas direções de perturbação com mesma variância e ARL_0 's aproximados	70
Quadro 21	ARL para o caso de aumentos na variância considerando três direções de perturbação com mesma variância	71
Quadro 22	ARL para o caso de aumentos na variância considerando três direções de perturbação com mesma variância e ARL_0 's aproximados	72
Quadro 23	ARL para o caso de aumentos na variância considerando duas direções de perturbação com o desvio-padrão da Rotação maior que do Deslocamento	73
Quadro 24	ARL para o caso de alterações na média e aumentos na variância considerando uma única direção de perturbação	75
Quadro 25	ARL para o caso de alterações na média e aumentos na variância considerando uma única direção de perturbação e ARL_0 's aproximados	75
Quadro 26	ARL para o caso de alterações na média e aumentos na variância considerando duas direções de perturbação com mesma variância	76
Quadro 27	ARL para o caso de alterações na média e aumentos na variância considerando duas direções de perturbações com mesma variância e ARL_0 's aproximados	77

Quadro 28	ARL para o caso de alterações na média e aumentos na variância considerando três direções de perturbação com mesma variância	78
Quadro 29	ARL para o caso de alterações na média e aumentos na variância considerando três direções de perturbação e ARL_0 's aproximados	79
Quadro 30	ARL para o caso de alterações na média e aumentos na variância considerando duas direções de perturbação com diferentes variâncias	80
Quadro 31	ARL para o quadrado da norma do vetor resíduo segundo a forma distribucional com probabilidade de alarme falso $\alpha = 0,0027$	87
Quadro 32	ARL para o quadrado da norma do vetor resíduo segundo a forma distribucional com probabilidade de alarme falso $\alpha = 0,005$	89
Quadro 33	ARL para o quadrado da norma do vetor resíduo segundo a forma distribucional com probabilidade de alarme falso $\alpha = 0,01$	90
Quadro 34	Resumo dos resultados da análise de desempenho dos esquemas propostos segundo o caso simulado	94

Lista de Abreviaturas e Siglas

CEPM	Controle Estatístico de Processos Multivariados
CEP	Controle Estatístico de Processos
MSSD	Diferença Sucessiva da Quadrado da Média (do inglês, <i>Mean Square Successive Difference</i>)
CUSUM	Controle das Somas Acumuladas
EWMA	Média Móvel Exponencialmente Ponderada
ACP	Análise de Componentes Principais
PEL	Projeção de Estruturas Latentes
CP	Componentes Principais
VL	Variáveis Latentes
ACPM	Análise de Componentes Principais Multidirecionais
PLS	Mínimos Quadrados Parciais (do inglês, <i>Partial Least Square</i>)
ACI	Análise de Componentes Independentes
ARL	Número médio de amostras até a ocorrência de um alarme (do inglês, <i>Average Run Length</i>)
ARL ₀	Número médio de amostras até a ocorrência de um alarme falso
AD	Teste de Anderson-Darling
NMA	Número médio de amostras até a ocorrência de um alarme
qcc	Pacote do programa R para construção de gráficos de controle univariados
mqcc	Pacote do programa R para construção de gráficos de controle multivariados
rnorm	Função do programa computacional R para geração de variáveis aleatórias normais com os parâmetros média e desvio-padrão especificados.
Prcomp	Função do programa computacional R para obtenção das componentes principais

Lista de Símbolos

T^2	Estatística multivariada de Hotelling
χ^2	Estatística Qui-Quadrado
\bar{X}	Média amostral da variável x
m	Número de observações individuais
\mathbf{x}	Vetor aleatório contendo as variáveis de qualidade de interesse observadas
\mathbf{d}	Vetor aleatório que representa as k distintas perturbações nas direções conhecidas.
\mathbf{C}	Matriz de constantes com colunas linearmente independentes
\mathbf{e}	Vetor de erros ou ruídos aleatórios de medição
\mathbf{c}_j	Vetor coluna unitário em uma direção conhecida
Σ_d	Matriz de variância-covariância referente às direções conhecidas
Σ_e	Matriz de variância-covariância referente aos ruídos aleatórios
$\text{Cov}(\mathbf{e}, \mathbf{d})$	Covariância entre os erros aleatórios e as direções conhecidas
$\mathbf{d}(\mathbf{x})$	Vetor aleatório contendo as variáveis transformadas
\mathbf{I}	Matriz identidade
\mathbf{r}	Vetor resíduo
$\ \mathbf{r}\ ^2$	Quadrado da norma do vetor resíduo
σ_e^2	Variância do erro aleatório
σ_e	Desvio-padrão do erro aleatório
T	Direção conhecida Rotação
L	Direção conhecida Deslocamento
C	Direção conhecida Contração/Dilatação
$T(x)$	Variável transformada na direção conhecida Rotação ou o seu respectivo gráfico
$L(x)$	Variável transformada na direção conhecida Deslocamento ou o seu respectivo gráfico
$C(x)$	Variável transformada na direção conhecida Contração/Dilatação ou o seu respectivo gráfico
$\mathbf{t}(\mathbf{x})$	Projeção do vetor \mathbf{x} na direção $T(x)$
$\mathbf{l}(\mathbf{x})$	Projeção do vetor \mathbf{x} na direção $L(x)$
$G(r, \alpha)$	Distribuição gama com parâmetros r e α

$\cup X_i$	Quantidade de alarmes em pelo menos um dos gráficos nas variáveis observadas x_1, x_2, x_3 e x_4 ;
CP_i	i-ésima Componente Principal ou seu respectivo gráfico
α	Probabilidade de alarme falso
$\cup X_{i\alpha-aj}$	ARL conjunto dos gráficos de controle nas variáveis observadas, considerando um risco de alarme falso (α) ajustado
$T(x)_{\alpha-aj}$	Gráfico de controle para a variável transformada $T(x)$, considerando um risco de alarme falso (α) ajustado
$PC_{1\alpha-aj}$	$PC_{1\alpha-aj}$ = gráfico de controle para a primeira componente principal considerando um risco de alarme falso (α) ajustado
$\cup d_i(x)$	ARL conjunto dos gráficos nas variáveis transformadas
$\cup PC_i$	ARL conjunto dos gráficos nas componentes principais
R2	Gráfico de $\ r\ ^2$ com limite de controle no quantil de $(1 - \alpha)$ da distribuição exponencial
$R2_{aj}$	Gráfico de $\ r\ ^2$ com probabilidade de alarme falso ajustado para produzir um limite de controle no quantil de $(1 - \alpha)$ da distribuição normal
R2T	Gráfico na transformada de Box-Cox de $\ r\ ^2$ com limite de controle ao quantil de $(1 - \alpha)$ da distribuição normal

1 Introdução

1.1 Problema Tratado e Estudo da Arte

O estudo do Controle Estatístico de Processos Multivariados (CEPM) tem ganho bastante atenção nos últimos anos, especialmente na busca de métodos que consigam realizar o monitoramento eficaz de processos produtivos. Atualmente, parcela significativa dos problemas de Controle Estatístico de Processos (CEP) se encontra dentro do contexto multivariado, especificamente aqueles onde as variáveis de interesse apresentam-se correlacionadas.

A ideia de utilizar métodos estatísticos multivariados em CEP não é recente. Hotelling (1947) propôs um método de controle para monitorar mais de uma variável em um único gráfico onde os valores dispostos são baseados na estatística T^2 , mais conhecida como T^2 de Hotelling, considerando que as variáveis são correlacionadas e seguem uma distribuição normal multivariada.

No entanto, o maior desenvolvimento do CEPM ocorreu a partir dos anos 80, particularmente pelo avanço computacional que facilitou a implementação das técnicas existentes. Desde então, muito do que foi feito até o momento foi no sentido de aperfeiçoar o método de Hotelling ou desenvolver novas metodologias visando o controle de processos multivariados.

Nesse sentido, alguns autores voltaram a atenção para o aperfeiçoamento das técnicas de interpretação do sinal no gráfico T^2 . Dentre esses, merecem destaque: Murphy (1987), que sugeriu a partição das p variáveis envolvidas no problema em dois subgrupos, um dos quais representando as variáveis suspeitas de causar o descontrole no processo, e propôs o controle da estatística definida como a diferença da estatística T^2 dos respectivos grupos; Mason, Tracy e Young (1995), Holmes e Mergen (1995), Nedumaran e Pignatiello (1998), propuseram implementar o gráfico T^2 de modo que, após a emissão do sinal de descontrole, fossem identificadas as variáveis que sofreram desvio. Mais especificamente, Mason *et al.* (1995) mostraram que a interpretação do sinal dos gráficos de controle multivariados, usando a estatística T^2 , pode ser aperfeiçoada se a estatística for decomposta em partes independentes. Nesse caso, a partição informa qual característica está contribuindo para o sinal de fora de controle.

Baseando-se na estimativa da MSSD (*Mean Square Successive Difference*), Holmes e Mergen (1995) propuseram outro método para ajudar na identificação das variáveis-problema, que contribuem para os sinais dados no gráfico de controle T^2 . Nedumaran e Pignatiello (1998) implementaram um método em que intervalos do tipo de Scheffé são usados para identificar as variáveis que causam os sinais de fora de controle nos gráficos T^2 e χ^2 . Este método faz uso de p gráficos de diagnóstico univariados para as p variáveis e p gráficos de diagnóstico univariados para as componentes principais, mantendo a probabilidade de alarme falso em um nível preestabelecido para todos os gráficos. Ainda, Mason *et al.* (1997) apresentaram uma estratégia para a decomposição de T^2 baseada na ordenação *a priori* das médias das p variáveis, chamado de decomposição *step-down*.

Com relação ao desenvolvimento de outras técnicas estatísticas multivariadas aplicadas aos problemas de controle de processos são referências importantes dentro do contexto multivariado, por suas contribuições, Woodall e Ncube (1985), Crosier (1988) e Pignatiello e Runger (1990), que trabalharam no desenvolvimento dos gráficos de Controle das Somas Acumuladas (CUSUM) Multivariado, e Lowry et al. (1992), que estudaram e propuseram a versão multivariada dos gráficos da Média Móvel Exponencialmente Ponderada (EWMA). O objetivo destas propostas é melhorar a eficácia (rapidez) na detecção de alterações pequenas e moderadas no vetor de médias do processo. Também possuem relevância, pelo grande número de artigos e aplicações em CEP, os procedimentos multivariados baseados em modelos de regressão, dos quais são destaques os trabalhos de Zhang (1985), Hawkins (1991, 1993) e Wade e Woodall (1993). Destaca-se ainda, numa outra frente de pesquisa, o uso de métodos de projeção, tais como Análise de Componentes Principais (ACP) e Projeção de Estruturas Latentes (PEL). Esses métodos visam essencialmente reduzir a dimensionalidade do problema de controle preservando a estrutura de variância-covariância, eliminando possíveis redundâncias causadas pela colinearidade entre as variáveis originais, e extrair a informação sobre a direção da variabilidade dos dados projetando-os em um espaço com menor dimensão definidas como Componentes Principais (CP) ou Variáveis Latentes (VL).

Jackson (1980, 1981), referência importante da utilização da ACP no monitoramento de processos multivariados, detalha a teoria da ACP e utiliza o gráfico

de T^2 nas componentes principais (CP's) obtidas das variáveis originais, com o objetivo de identificar causas especiais. Um pouco antes, Jackson e Mudholkar (1979) já haviam estabelecido um procedimento de controle dos resíduos utilizando a ACP. A ideia era, após obtenção das componentes principais, além de utilizar os resíduos do modelo para verificar a qualidade do ajuste e a presença de *outliers*, usá-lo no monitoramento de novos eventos que afetassem os parâmetros do processo. Posteriormente, Schall e Chandra (1987) propuseram um método que consiste na utilização da ACP em conjunto com a análise de regressão multivariada no intuito de prever possíveis causas de variação. Nesse caso, assume-se que o processo é composto por variáveis de entrada que contribuem para a formulação das variáveis de qualidade final do produto, chamadas de variáveis de saída. Dessa forma, a técnica consiste basicamente em estabelecer um modelo de regressão múltipla entre as variáveis de entrada do processo e os valores observados das componentes principais obtidas através das variáveis de saída, cujo objetivo é prever a variável de entrada responsável por problemas na qualidade final do processo.

Contudo, a maior difusão do uso da ACP dentro do contexto de CEPM ocorreu nos anos 90, fato que pode ser observado pelo aumento de publicações versando sobre o referido tema. Nesse sentido, são referências importantes os trabalhos de Nomikos e MacGregor (1994, 1995), que desenvolveram procedimentos para o monitoramento *on-line* de processos em batelada usando ACP, denotada neste caso como Análise de Componentes Principais Multidirecionais (ACPM); MacGregor e Kourti (1995) e Kourti e MacGregor (1996) que apresentaram uma visão geral da teoria dos métodos de projeção, mais especificamente, Análise de Componentes Principais (ACP) e Mínimos Quadrados Parciais ou *Partial Least Squares - PLS*, mostrando diversas aplicações dessas técnicas no CEPM. Destaca-se também Runger (1996), que definiu o gráfico U^2 baseando-se nos métodos de projeção, onde U é o vetor das k primeiras componentes principais e U^2 é o comprimento de uma projeção ortogonal, que pode ser escrito como uma forma quadrática baseada em uma matriz idempotente. Essa projeção é usada para simplificar a construção e o planejamento dos esquemas de controle multivariado. Souza e Rigão (2005) estudaram por meio de um exemplo o funcionamento de duas técnicas de identificação de variáveis fora de controle em processos produtivos multivariados, considerando dois casos: variáveis fracamente e fortemente relacionadas. Machado e Costa (2008) compararam o desempenho do gráfico de T^2 baseado em

componentes principais com gráficos de controle univariados usados simultaneamente, tanto com base nas variáveis originais, como com base nas componentes principais.

González e Sanchez (2008a) propuseram um método baseado na transformação das variáveis originais em fatores independentes usando ACP e considerando que as variáveis estão correlacionadas, o qual foi denotado de alarmes principais (*principal alarms*) cujo objetivo é observar alterações na média em processos multivariados normais, usando os dados em controle disponíveis. Em seguida, González e Sanchez (2008b) estenderam o método em questão para o caso de processos multivariados não-normais utilizando a Análise de Componentes Independentes (ACI). ACI é uma técnica recente para decompor dados multivariados não-normais em componentes estatisticamente independentes. Mais recentemente, González e Sanchez (2010) estabeleceram uma metodologia para selecionar um conjunto de variáveis, reduzindo com isso o número de variáveis a serem medidas no processo, com intuito de monitorar um processo multivariado. O método proposto pode ser resumido em dois estágios, a saber: seleção das variáveis e a avaliação da informação contida nas variáveis usando dois critérios. O primeiro baseia-se na quantidade de informação residual das variáveis não-selecionadas e o segundo, na análise de “alarmes principais”, conceito proposto pelos mesmos autores. Este último critério refere-se à capacidade de um subconjunto de variáveis, quando monitoradas, serem eficazes em sinalizar os desvios mais frequentes nos parâmetros do processo. Por fim, Aparisi *et al.* (2010) desenvolveram um programa que otimiza, usando algoritmos genéticos, um conjunto de gráficos de \bar{X} ou um conjunto de gráficos nas componentes principais do processo e compararam o seu desempenho com o de um gráfico de T^2 baseado respectivamente nas variáveis originais ou nas suas componentes principais.

Ainda baseados nos métodos de projeção de dados, Barton e Gonzalez-Barreto (1996) apresentaram uma forma alternativa de representar um processo multivariado, considerando uma nova base para expressar a variável de qualidade de interesse, partindo da suposição que é possível identificar padrões de erro ou problemas relacionados às prováveis causas potenciais de desequilíbrio no processo. Dessa forma, esses padrões de erro formarão os elementos ou coeficientes da nova base e o controle é feito por meio do gráfico da média para cada coeficiente dessa nova base. Em seguida, o trabalho de Runger e Montgomery (1997) chamou especial atenção por mostrar que, no caso de perturbações do processo ocorrerem tipicamente em direções pré-conhecidas, a forma de controle mais eficaz é usar um gráfico de controle tendo como estatística a

projeção do vetor de variáveis em cada uma dessas direções. Ainda, Apley e Shi (2001) propuseram um método para diagnosticar as fontes da variabilidade do processo supondo que falhas não correlacionadas afetam a medição da variável de interesse. Assim, a metodologia consiste em estabelecer um modelo, desenvolvido com base no modelo fatorial ortogonal (ver Johnson e Wichern, 2002), cujo intuito é descrever o relacionamento entre as possíveis falhas que podem ocorrer no processo e as causas de variabilidade. O objetivo é determinar quais são as falhas responsáveis pela variabilidade do processo.

Dessa forma, inspirando-se nos três últimos trabalhos supracitados, surge a ideia de analisar o desempenho de esquemas que consideram processos em que há direções de variação mais frequentes, que correspondam a fontes de variação ou de variações em variáveis do processo não diretamente observáveis, mas que se traduzam em variações do vetor de variáveis observáveis em direções específicas, que serão denotadas neste trabalho por *direções conhecidas* ou *direções de variação*. Assim, pretende-se nesta pesquisa estudar o desempenho de gráficos de controle das projeções das variáveis observáveis nas direções conhecidas, comparando seu desempenho com o de gráficos de Shewhart nas variáveis originais, com gráficos de Shewhart nas componentes principais e com gráficos de T^2 nas variáveis originais. Cabe ressaltar que este trabalho difere da proposta de González e Sanchez (2008a), visto que esses autores propõem simular a ocorrência de causas especiais que alteram a média do vetor de variáveis nas direções principais, obtidas por meio da ACP, para analisar o desempenho de esquemas de controle pré-existentes, enquanto que aqui é proposto e analisado um esquema específico para detecção dessas causas especiais. A análise de desempenho será feita supondo deslocamentos da média nas direções conhecidas e/ou aumentos da variância nessas mesmas direções. Ainda, a presente pesquisa obedece às seguintes suposições:

- a. As direções conhecidas estão relacionadas às causas primárias de perturbações aleatórias no processo. Quando o processo está em controle, as variações do vetor de variáveis se dão - a menos de pequenos ruídos aleatórios - nessas direções.
- b. Com o processo em controle, a média das alterações nessas direções é nula.

- c. Causas especiais atuando nas fontes primárias provocarão deslocamentos da média do vetor de variáveis observadas em alguma direção conhecida ou em mais de uma delas; ou, ainda, poderão provocar aumentos na variância das perturbações em uma ou mais das direções, assim, cada direção corresponderá a uma das fontes primárias de variação.
- d. Não é impossível, porém, que outras causas especiais possam provocar alterações (deslocamento da média ou aumento da variância do processo) em direções não contidas no hiperplano varrido pelas direções conhecidas.

Para detectar alterações do tipo mencionado no item (d) acima, propõe-se monitorar o quadrado da norma de um vetor resíduo, ortogonal ao hiperplano varrido pelas direções conhecidas.

Pretende-se, ainda, verificar se a ACP é capaz de identificar correta e univocamente as direções de variação conhecidas. Posteriormente, nos casos em que a ACP não identifica de forma eficaz essas direções, cabe analisar se o desempenho dos gráficos nas CP's é muito deteriorado em relação ao dos gráficos nas direções principais.

Devido à complexidade analítica do problema, as análises de desempenho serão conduzidas por simulação. A medida de desempenho que será utilizada é o Comprimento Médio da Sequência, mas conhecido como ARL (do inglês, *Average Run Length*).

1.2 Justificativa

Muitos dos trabalhos em CEPM, embora reconheçam que as variáveis são correlacionadas para a análise de desempenho dos métodos, consideram casos arbitrários de alterações em uma variável, e se preocupam em identificar a variável responsável pelo sinal quando talvez não seja esse o modelo mais adequado para o problema. Em outras palavras: a causa especial pode não estar atuando direta e exclusivamente sobre uma variável observada, isoladamente, mas sim sobre uma variável latente, não diretamente observável, que afeta as variáveis observáveis em conjunto, alterando sua média e/ou variância segundo direções “principais” ou “conhecidas”. Dessa forma, a principal motivação desse trabalho é estudar métodos de

CEPM para o caso em que os parâmetros do processo se deslocam em direções conhecidas.

Nesse sentido, a ideia é monitorar as projeções do conjunto de variáveis nas direções conhecidas, e não as variáveis isoladas (observáveis). Assim, espera-se que os esquemas propostos neste estudo, além de terem melhor desempenho (maior poder de detecção de descontroles), já identificariam automaticamente em qual variável latente ocorreu a alteração, e, com isso, qual a causa especial atuando sob o processo.

1.3 Contribuição Original e Relevância

A grande maioria dos trabalhos que propõem esquemas de CEPM e que analisam o desempenho desses esquemas considera alterações nas variáveis originais. Nos casos em que as alterações se dão em direções previamente conhecidas (direções conhecidas), a análise deve ser feita considerando variações nessas direções e, além disso, os sinais devem ser interpretados em termos dessas direções. Tal interpretação é automática no caso de gráficos nas direções conhecidas, como proposto nesta pesquisa. Portanto, a originalidade e relevância da presente pesquisa podem ser observadas nos seguintes pontos:

- i. Proposta de um método para o controle estatístico de processos multivariados nos quais as perturbações ocorrem em direções conhecidas, que além de visar a detecção de causas especiais, tem o benefício adicional de identificar diretamente a causa especial (associada a uma direção conhecida) responsável pelo sinal.
- ii. Análise do desempenho do método proposto em detectar deslocamentos de diversas magnitudes da sua média e/ou aumentos da sua variância nas direções conhecidas;
- iii. Identificação das situações em que a ACP de dados históricos do processo em controle falha em identificar as direções conhecidas;
- iv. Análise comparativa do desempenho do esquema proposto e dos demais esquemas “clássicos” (gráficos de T^2 e vários gráficos univariados nas variáveis diretamente observáveis).

1.4 Objetivos

1.4.1 Objetivo Geral

Propor um método de CEP destinado ao controle de processos produtivos multivariados em que as alterações na sua média e/ou na sua variância ocorram em algumas direções *conhecidas*; e analisar seu desempenho para diferentes magnitudes de alterações no vetor de médias ou na variância nas direções conhecidas, comparando-o com alternativas clássicas como o gráfico T^2 e gráficos univariados conjuntos nas variáveis observadas diretamente.

1.4.2 Objetivos Específicos

São objetivos específicos deste trabalho:

Para a Fase I

- Supondo que as direções conhecidas sejam as mesmas direções das variações aleatórias do processo quando em controle, verificar se a ACP aplicada a dados históricos com o processo em controle identifica corretamente essas direções, ou ainda, em que condições isso ocorre.

Para a Fase II

- Supondo alterações na média e no desvio-padrão do processo nas direções conhecidas, analisar a eficiência de um esquema que consiste de gráficos nessas direções;
- Proposta de um gráfico adicional para detecção de perturbações da média e do desvio-padrão do processo em direções fora do hiperplano varrido pelas direções conhecidas de variação e analisar o seu desempenho.

1.5 Delimitações do Trabalho

Considerou-se neste estudo apenas o caso de observações individuais (em contraste com subgrupos) do vetor de variáveis e os gráficos propostos são destinados à sinalização de alterações na média do vetor de variáveis nas direções conhecidas. Embora também seja analisado o desempenho do esquema proposto no caso de aumentos na variância do vetor de variáveis nas direções conhecidas, não está no escopo deste trabalho propor e analisar esquemas especificamente destinados ao controle estatístico de tais aumentos na variância, que fica como indicação de prosseguimento.

Por limitação de tempo de uma pesquisa de doutorado, os esquemas propostos e analisados são do tipo “de Shewhart”; aprimoramentos tais como versões EWMA ou CUSUM dos gráficos propostos também são deixados para pesquisa futura.

1.6 Organização do texto

A presente tese está organizada em cinco capítulos. Neste primeiro capítulo foi apresentado um breve histórico dos principais artigos que serviram de referência para o desenvolvimento deste trabalho, a justificativa, a contribuição e a relevância, os objetivos, bem como as delimitações do estudo.

O capítulo 2 trata da fundamentação teórica do uso de variáveis transformadas no monitoramento de processos multivariados. Além disso, mostra na teoria como a ACP pode ser útil para determinar direções de variação e como usar o vetor resíduo no monitoramento do processo para captar perturbações fora do hiperplano varrido pelas direções principais.

No capítulo 3 é apresentado um estudo de simulação baseado em um problema real onde as metodologias propostas no capítulo anterior são desenvolvidas. Este capítulo contém ainda a descrição dos diversos casos a serem simulados com o objetivo de analisar o desempenho dos gráficos usando as variáveis transformadas. A análise de desempenho será feita supondo deslocamentos da média nas direções conhecidas e aumentos da variância nessas mesmas direções.

O capítulo 4 dedica-se a consolidação dos resultados obtidos por meio das simulações descritas no capítulo 3 com o intuito de demonstrar a eficácia do gráfico de controle proposto em comparação com os esquemas tradicionais de controle.

As conclusões e as sugestões para trabalhos futuros estão contempladas no capítulo 5.

No Apêndice é apresentado um resumo teórico sobre a Análise de Componentes Principais e no Anexo encontram-se os gráficos de controle referentes ao quadrado da norma do vetor resíduo apenas para dar uma ideia de como os gráficos de controle foram gerados e utilizados na determinação dos indicadores de desempenho (ARL's)

2 O Uso de Variáveis Transformadas

2.1 Modelo do Processo

A suposição básica deste trabalho é que perturbações ou variações aleatórias ocorridas nas características de qualidade de um processo, que se deseja controlar estatisticamente, são efeitos de outras perturbações que são previamente conhecidas e independentes. Assim, é possível pensar em um modelo em que essas variações, não observadas ou medidas diretamente, provocam um conjunto de alterações no vetor de variáveis observadas do processo em direções específicas. Supõe-se ainda, que as direções de variação conhecidas sejam ortogonais, porém, como não necessariamente coincidem com as direções dos eixos coordenados que correspondem às variáveis observadas (x_1, x_2, \dots, x_p), a variação (deslocamentos da média e/ou aumento da variância do vetor \mathbf{x}) em cada uma dessas direções provoca alterações correlacionadas nas variáveis observáveis. Dessa maneira, as variações nessas direções é que são conhecidas e correspondem cada uma, a uma fonte de perturbação. Além disso, quando ocorre um problema no processo, este aumenta a variabilidade das perturbações nas direções conhecidas ou desloca o vetor de médias do processo em uma dessas direções.

Dessa forma, é necessário estabelecer um modelo que represente a relação linear entre as variáveis observáveis e as variações nas direções conhecidas. Logo, com base em estrutura semelhante ao modelo fatorial ortogonal (Ver Johnson e Wichern, 2002), pode-se adotar a seguinte equação:

$$\mathbf{x} = \mathbf{C}\mathbf{d} + \mathbf{e} \quad (1)$$

onde $\mathbf{x} = [x_1, x_2, \dots, x_p]'$ é um vetor aleatório contendo as p variáveis de qualidade de interesse observadas, $\mathbf{d} = [d_1, d_2, \dots, d_k]'$ é um vetor aleatório que representa as k distintas perturbações nas direções conhecidas, \mathbf{C} é uma matriz $p \times k$ ($k \leq p$) de constantes, com colunas linearmente independentes, cada coluna representando o efeito linear da variação em uma direção conhecida sobre as diversas variáveis medidas (elementos de \mathbf{x}), e $\mathbf{e} = [e_1, e_2, \dots, e_p]'$ é um vetor de erros ou ruídos aleatórios de medição, ou ainda,

de outras variações em direções não consideradas no modelo. O modelo em (1) pode ser reescrito como:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1k} \\ c_{21} & c_{22} & \cdots & c_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ c_{p1} & c_{p2} & \cdots & c_{pk} \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_k \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_p \end{bmatrix} \quad (2)$$

ou ainda,

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_k \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_k \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_p \end{bmatrix} \quad (3)$$

onde cada $\mathbf{c}_j = [c_{1j} \ c_{2j} \ \cdots \ c_{pj}]'$, $j = 1, 2, \dots, k$, é um vetor coluna unitário em uma direção conhecida. Assim, o modelo proposto representa o efeito linear das variações nas k direções conhecidas sobre as medidas do vetor aleatório \mathbf{x} , a qual se soma ainda um erro aleatório. Além disso, assume-se que o modelo obedeça às seguintes suposições:

- As variações nas direções conhecidas formam um vetor aleatório com distribuição Normal k -variada com vetor de médias nulo e matriz de variância-covariância diagonal Σ_d , isto é, $\mathbf{d} \sim N_k(\mathbf{0}, \Sigma_d)$.
- Os ruídos aleatórios também são independentes e identicamente distribuídos com distribuição Normal p -variada com vetor de médias nulo e matriz de variância-covariância diagonal Σ_e , isto é, $\mathbf{e} \sim N_p(\mathbf{0}, \Sigma_e)$.
- As variações nas direções conhecidas e os ruídos aleatórios são independentes entre si, ou seja, $\text{Cov}(\mathbf{e}, \mathbf{d}) = E(\mathbf{e}\mathbf{d}') = \mathbf{0}$.

Cabe registrar que outros autores também utilizaram estruturas lineares baseadas no modelo fatorial ortogonal em seus trabalhos, entre os quais, Apley e Shi (2001), que utilizaram o referido modelo no diagnóstico e na interpretação das causas de variabilidade em processos de manufatura, e Apley e Lee (2003), na identificação de padrões de variação espaciais.

2.2 Definindo as Transformações

Supondo que se conhece bem o processo produtivo que se deseja controlar, e as direções conhecidas inerentes a esse processo (o caso em que não se conhece essas direções será tratado em 2.3), podem-se definir transformações nas variáveis observadas que sejam estimativas das variações do vetor de variáveis nas direções conhecidas, para controlar estatisticamente essas variáveis transformadas em vez das variáveis medidas diretamente. Assim, será definido, respectivamente ao vetor \mathbf{d} (definido na seção 2.1), o vetor $\mathbf{d}(\mathbf{x})$ como segue:

$$\mathbf{d}(\mathbf{x}) = \mathbf{C}'\mathbf{x} \quad (4)$$

onde \mathbf{C}' é a transposta da matriz \mathbf{C} apresentada na seção 2.1.

Dessa forma, as variáveis transformadas são as projeções do vetor de variáveis observadas nessas direções. Assume-se que as linhas da matriz \mathbf{C}' são vetores linearmente independentes de norma um, formando assim uma base ortonormal, o que é conveniente para a simplificação da conversão de construções geométricas em cálculos algébricos. A hipótese de que as linhas da matriz \mathbf{C}' sejam linearmente independentes é uma hipótese básica deste trabalho – direções conhecidas ortogonais – e não está no seu escopo considerar o caso em que essa hipótese não valha; quanto aos vetores serem de norma um, isso pode ser forçado em qualquer caso, pela adequada definição dos seus elementos.

Substituindo (1) na expressão (4) para $\mathbf{d}(\mathbf{x})$, pode-se chegar ao seguinte resultado:

$$\mathbf{d}(\mathbf{x}) = \mathbf{C}'(\mathbf{C}\mathbf{d} + \mathbf{e}) = \mathbf{C}'\mathbf{C}\mathbf{d} + \mathbf{C}'\mathbf{e} \quad (5)$$

Assim, observa-se que $\mathbf{d}(\mathbf{x})$ é uma medida ruidosa de \mathbf{d} , vetor de variações nas direções conhecidas, e, portanto, pode-se dizer que $\mathbf{d}(\mathbf{x})$ é um estimador de \mathbf{d} . Ainda, com base na equação (5), os valores da esperança (média) e da variância para $\mathbf{d}(\mathbf{x})$ podem ser determinados. Logo,

$$E[\mathbf{d}(\mathbf{x})] = \mathbf{C}'\mathbf{C}E(\mathbf{d})$$

mas, note que $\mathbf{C}'\mathbf{C} = \mathbf{I}_{k \times k}$, pois, por hipótese, as colunas de \mathbf{C} (linhas de \mathbf{C}') são linearmente independentes, com norma unitária. Então,

$$E[\mathbf{d}(\mathbf{x})] = E(\mathbf{d}) \quad (6)$$

e

$$\text{Cov}[\mathbf{d}(\mathbf{x})] = E[\mathbf{d}(\mathbf{x})\mathbf{d}'(\mathbf{x})] = E[(\mathbf{C}'\mathbf{C}\mathbf{d} + \mathbf{C}'\mathbf{e})(\mathbf{d}'\mathbf{C}'\mathbf{C} + \mathbf{e}'\mathbf{C})] \Rightarrow$$

$$\Rightarrow \text{Cov}[\mathbf{d}(\mathbf{x})] = \mathbf{C}'\mathbf{C}E(\mathbf{d}\mathbf{d}')\mathbf{C}'\mathbf{C} + \mathbf{C}'E(\mathbf{e}\mathbf{e}')\mathbf{C} = \mathbf{C}'\mathbf{C}\Sigma_{\mathbf{d}}\mathbf{C}'\mathbf{C} + \mathbf{C}'\Sigma_{\mathbf{e}}\mathbf{C} \Rightarrow$$

$$\Rightarrow \text{Cov}[\mathbf{d}(\mathbf{x})] = \Sigma_{\mathbf{d}} + \mathbf{C}'\Sigma_{\mathbf{e}}\mathbf{C} = \Sigma_{\mathbf{d}} + \sigma_e^2\mathbf{I} \quad (7)$$

Portanto, se o processo estiver sob controle, tem-se que $E(\mathbf{d}) = \mathbf{0}$, e, consequentemente, $E[\mathbf{d}(\mathbf{x})] = \mathbf{0}$. Tem-se, ainda, que $\text{Cov}[\mathbf{d}(\mathbf{x})]$ é uma matriz diagonal. Assim, os elementos de $\mathbf{d}(\mathbf{x})$ são variáveis aleatórias descorrelacionadas, com média nula quando o processo está sob controle. Embora as parcelas $\Sigma_{\mathbf{d}}$ e $\sigma_e^2\mathbf{I}$ da matriz (diagonal) de variáveis-covariáveis de $\mathbf{d}(\mathbf{x})$ não possam ser diretamente estimadas, seus elementos, que são as variâncias de cada elemento de $\mathbf{d}(\mathbf{x})$, podem ser estimados a partir dos dados históricos. Esses resultados são importantes, pois permitem estabelecer um gráfico de controle para cada elemento de $\mathbf{d}(\mathbf{x})$, isto é, $d_1(x), d_2(x), \dots, d_k(x)$.

2.3 Utilizando a Análise de Componentes Principais para Determinar as Direções de Variação do Vetor de Variáveis Observadas

Quando as direções conhecidas do processo não são conhecidas, é importante ter como identificá-las ou estimá-las com precisão. A Análise de Componentes Principais (ACP) é uma das técnicas estatísticas que pode ser utilizada com essa finalidade. Assim, um dos objetivos deste estudo é investigar se a ACP é eficaz nessa tarefa. Para isso, deve-se considerar a definição de paralelismo entre dois vetores. Dois vetores, digamos $\mathbf{u} = (u_1, u_2, \dots, u_k)$ e $\mathbf{v} = (v_1, v_2, \dots, v_k)$, são paralelos, e portanto, têm mesma direção, se existe um número λ tal que $\mathbf{u} = \lambda \mathbf{v}$, ou de outra maneira, dois vetores \mathbf{u} e \mathbf{v} são paralelos quando seus elementos são proporcionais, isto é, $\frac{u_1}{v_1} = \frac{u_2}{v_2} = \dots = \frac{u_k}{v_k} = \lambda$.

Na prática, visto que as CP's devem ter norma igual a um e que cada coluna da matriz \mathbf{C} das magnitudes das direções de variação definida em (1) também possui norma igual a um, basta examinar se os coeficientes das CP's são aproximadamente iguais às colunas dessa matriz ou a vetores simétricos a eles. Entretanto, observa-se que a coincidência entre valores dos coeficientes jamais será perfeita, por causa do erro experimental; trata-se de examinar se os coeficientes coincidem a menos de uma pequena diferença.

Um resumo da técnica de ACP pode ser encontrado no Apêndice desta tese.

2.4. Utilizando o Vetor Resíduo para o Monitoramento do Processo

Pode haver casos em que, além das perturbações do processo produtivo nas direções conhecidas, que ocorram de forma mais frequente, eventualmente venham a ocorrer, devido a causas especiais até então não observadas, perturbações do vetor de médias em novas direções, não identificadas na história prévia do processo. Portanto, é importante complementar o esquema de monitoramento proposto com um esquema capaz detectar alterações em direções fora do hiperplano formado pelas direções identificadas como as mais frequentes. Com essa finalidade, é proposto aqui um gráfico de controle na norma, ou no quadrado da norma de um vetor resíduo (\mathbf{r}), definido como a diferença entre o vetor de observações e a sua projeção sobre o hiperplano varrido

pela base ortonormal formada pelos vetores nas direções conhecidas. Esta projeção é o vetor soma das projeções do vetor das variáveis observadas nas direções conhecidas.

Portanto, considerando que as direções de variação que afetam o processo mais frequentemente são direções conhecidas, pode-se definir o vetor resíduo na base x_1, x_2, \dots, x_p de acordo com a seguinte expressão:

$$\mathbf{r} = \mathbf{x} - \mathbf{d}(\mathbf{x}) \quad (8)$$

onde \mathbf{x} é o vetor de observações na base x_1, x_2, \dots, x_p e $\mathbf{d}(\mathbf{x})$ é a projeção de \mathbf{x} no subespaço das direções conhecidas.

Com o processo em controle ou com perturbações apenas nas direções conhecidas, espera-se que o vetor resíduo resulte apenas das flutuações aleatórias e_1, e_2, \dots, e_p . Porém, eventuais perturbações que surjam, devidas as novas causas especiais até então desconhecidas e em direções não contidas no subespaço das direções conhecidas, resultarão em um aumento do módulo de \mathbf{r} . Isso faz com que este módulo possa ser usado para sinalizar alterações nos parâmetros do processo em novas direções, não contidas no subespaço de direções conhecidas. Note que, pelo Teorema de Pitágoras:

$$\|\mathbf{r}\|^2 = \|\mathbf{x}\|^2 - \|\mathbf{d}(\mathbf{x})\|^2 \quad (9)$$

Desenvolvendo cada parte do lado direito da eq. (9) separadamente, tem-se:

$$\|\mathbf{x}\|^2 = x_1^2 + x_2^2 + \dots + x_p^2$$

mas,

$$\|\mathbf{x}\|^2 = \mathbf{x}' \mathbf{x} = (\mathbf{C}\mathbf{d} + \mathbf{e})' (\mathbf{C}\mathbf{d} + \mathbf{e}) = (\mathbf{d}' \mathbf{C}' + \mathbf{e}') (\mathbf{C}\mathbf{d} + \mathbf{e}) = \mathbf{d}' \mathbf{C}' \mathbf{C} \mathbf{d} + 2\mathbf{d}' \mathbf{C}' \mathbf{e} + \mathbf{e}' \mathbf{e}$$

como $\mathbf{C}' \mathbf{C} = \mathbf{I}_{k \times k}$, então

$$\|\mathbf{x}\|^2 = \mathbf{d}'\mathbf{d} + 2\mathbf{d}'\mathbf{C}'\mathbf{e} + \mathbf{e}'\mathbf{e} \quad (10)$$

Trabalhando agora com o termo $\|\mathbf{d}(\mathbf{x})\|^2$, tem-se que $\|\mathbf{d}(\mathbf{x})\|^2 = \mathbf{d}'(\mathbf{x})\mathbf{d}(\mathbf{x})$, mas foi visto em (4) que $\mathbf{d}(\mathbf{x}) = \mathbf{C}'\mathbf{x}$. Dessa forma,

$$\begin{aligned} \|\mathbf{d}(\mathbf{x})\|^2 &= \mathbf{d}'\mathbf{C}'\mathbf{C}\mathbf{C}'\mathbf{C}\mathbf{d} + 2\mathbf{d}'\mathbf{C}'\mathbf{C}\mathbf{C}'\mathbf{e} + \mathbf{e}'\mathbf{C}\mathbf{C}'\mathbf{e} \Rightarrow \\ \Rightarrow \|\mathbf{d}(\mathbf{x})\|^2 &= \mathbf{d}'\mathbf{d} + 2\mathbf{d}'\mathbf{C}'\mathbf{e} + \mathbf{e}'\mathbf{C}\mathbf{C}'\mathbf{e} \end{aligned} \quad (11)$$

logo,

$$\begin{aligned} \|\mathbf{r}\|^2 &= \|\mathbf{x}\|^2 - \|\mathbf{d}_x\|^2 \Rightarrow \\ \Rightarrow \|\mathbf{r}\|^2 &= (\mathbf{d}'\mathbf{d} + 2\mathbf{d}'\mathbf{C}'\mathbf{e} + \mathbf{e}'\mathbf{e}) - (\mathbf{d}'\mathbf{d} + 2\mathbf{d}'\mathbf{C}'\mathbf{e} + \mathbf{e}'\mathbf{C}\mathbf{C}'\mathbf{e}) \Rightarrow \\ \Rightarrow \|\mathbf{r}\|^2 &= \mathbf{e}'\mathbf{e} - \mathbf{e}'\mathbf{C}\mathbf{C}'\mathbf{e} \end{aligned} \quad (12)$$

Observa-se que, conforme esperado, a expressão (12) depende apenas das flutuações aleatórias e_1, e_2, \dots, e_p, e , portanto, o cálculo da esperança e da variância de $\|\mathbf{r}\|^2$ dependerá desses parâmetros.

Porém, note-se que não é fácil determinar uma forma fechada para a distribuição de probabilidade do quadrado da norma do vetor resíduo, visto que a expressão (12) contém somas de produtos de variáveis normais, de acordo com as suposições desta tese. Na verdade, os resultados que se tem a esse respeito são apenas aproximações do que seria a distribuição da soma de produto de variáveis aleatórias normais. Vários autores dedicaram-se a esse estudo. Entre eles, Craig (1936) foi o primeiro a determinar uma forma algébrica para a função geradora de momentos do produto de normais, e Aroian *et al.* (1978) provaram que, sob certas condições, o produto de duas variáveis normalmente distribuídas aproxima-se da distribuição padronizada de Pearson do tipo

III. Mais recentemente Ware e Lad (2003) compararam três métodos de aproximação para a distribuição do produto de duas variáveis normais.

Esse desenvolvimento tem interesse teórico, para melhorar a compreensão do problema, mas, como os σ_e^2 são desconhecidos e não diretamente mensuráveis, os limites de controle a serem usados para o gráfico de $\|\mathbf{r}\|^2$ deverão ser estabelecidos com base nas estimativas obtidas na Fase I, a partir dos dados históricos dos parâmetros da distribuição de $\|\mathbf{r}\|^2$.

3 Análise de Desempenho

Para exemplificar o uso do modelo e dos esquemas de controle tratados nos itens anteriores, bem como para realizar uma análise de desempenho desses esquemas, será considerado o processo descrito em Gaio e Epprecht (2006), referente à montagem da porta traseira em um modelo de automóvel, em que a peça (porta) pode apresentar problemas no ajuste com a carroceria. Dessa forma, a característica de interesse é o ajuste geométrico entre as peças externas, isto é, as variáveis observadas são as folgas entre a porta e a carroceria em diversos pontos ao redor da referida porta, que constituem um vetor de variáveis aleatórias correlacionadas entre si. Para efeito de simplificação, serão considerados apenas quatro pontos, distribuídos conforme a Figura 1, apresentada a seguir.

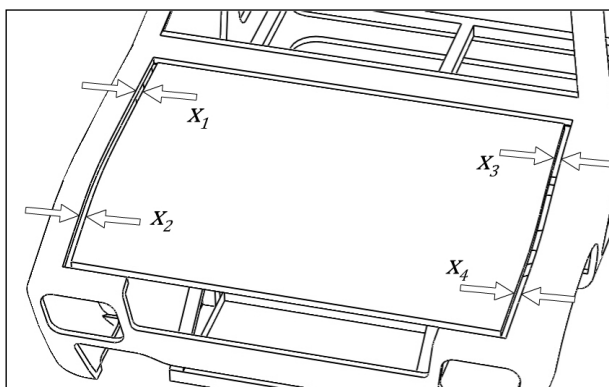


Figura 1: representação de quatro possíveis pontos de folgas do problema.

Nessa figura, os x_i 's, $i = 1, 2, 3, 4$, são as variáveis observadas (medidas) e representam as folgas em cada ponto específico. Por conveniência, trabalhar-se-á com as folgas centradas, isto é, com as diferenças entre o valor da folga e seu valor médio (supostamente igual ao valor-alvo se o processo estiver em controle). Em consequência, com o processo em controle, $E(x_i) = 0$, $i = 1, 2, 3, 4$. A seguir, serão aplicados ao problema em questão os procedimentos descritos no Capítulo 2.

3.1 Modelo Particular para o Exemplo

Suponha que os possíveis problemas no ajuste da peça possam ser explicados por rotações e/ou deslocamentos da peça, ou ainda, por uma contração ou dilatação após

o estágio de pintura (que se dá a altas temperaturas, em uma estufa), e que poderiam ser entendidas como as próprias variações essenciais – não observadas ou medidas diretamente, mas que provocam um conjunto de alterações correlacionadas nas variáveis (folgas) medidas. Cada uma dessas variações (rotação, deslocamento, contração/dilatação) corresponde a uma direção específica do vetor \mathbf{x} , das folgas, ou seja, uma direção conhecida de variação, pois cada uma corresponde a uma causa especial. Dessa maneira, essas direções de variação diretamente designadas por Rotação (T), Deslocamento (L) ou Contração/Dilatação (C).

Como dito, uma Rotação, um Deslocamento ou mesmo uma Contração/Dilatação alteram as variáveis observadas de maneira correlacionada, isto é, em certas direções específicas. Por exemplo, uma Rotação no sentido anti-horário em torno do centro de gravidade da peça provoca alterações nas médias das variáveis x_i , especificamente, uma redução de uma mesma magnitude, digamos d unidades, nas médias de x_1 e x_4 e um aumento de d unidades nas médias de x_2 e x_3 conforme Figura 2. Portanto, a média do vetor \mathbf{x} desloca-se de $2d$ unidades na direção do vetor $(-0,5 \ 0,5 \ 0,5 \ -0,5)'$. Já um Deslocamento de magnitude d para direita provoca um aumento de d unidades nas médias de x_1 e x_2 e uma redução de d unidades nas médias de x_3 e x_4 , ou seja, a média de \mathbf{x} desloca-se de $2d$ unidades na direção $(0,5 \ 0,5 \ -0,5 \ -0,5)'$, como pode ser observado na Figura 3 e uma Contração de magnitude d produz um aumento de d unidades nas médias de x_1 , x_2 , x_3 e x_4 , o que implica que a média desloca-se de $2d$ unidades na direção $(0,5 \ 0,5 \ 0,5 \ 0,5)'$ de acordo com a Figura 4.

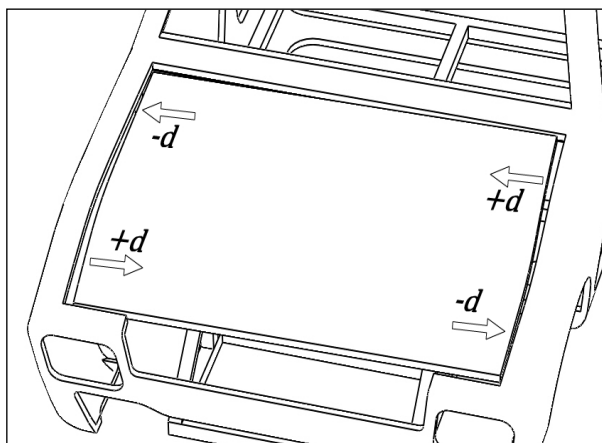


Figura 2: Rotação de magnitude d no sentido anti-horário.

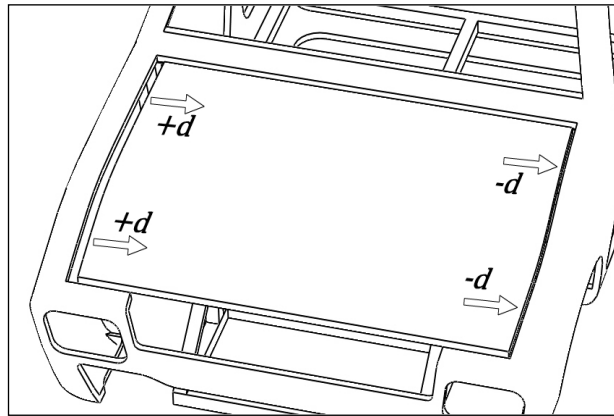


Figura 3: Deslocamento para direita de magnitude d .

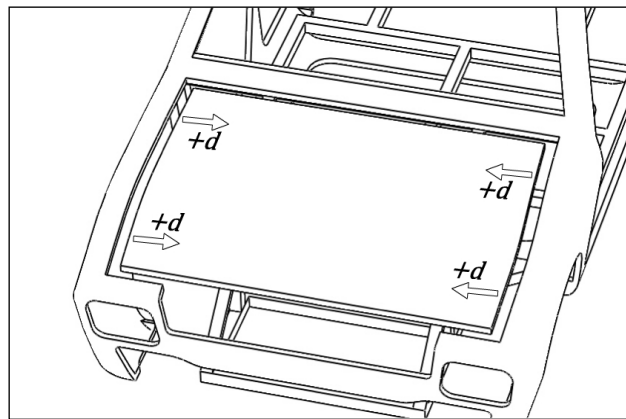


Figura 4: Contração de magnitude d .

Dessa forma, de acordo com as hipóteses assumidas inicialmente, considera-se que, se o processo estiver sob controle, as variações em cada direção conhecida seguem uma distribuição Normal com média zero e variância constante. Sendo assim, as magnitudes das variações nas direções conhecidas podem ser definidas da seguinte maneira:

- Deslocamento: $L \sim N(0, \sigma_l^2)$
- Rotação: $T \sim N(0, \sigma_t^2)$
- Contração/Dilatação: $C \sim N(0, \sigma_c^2)$

Note que nesse caso, $C > 0$ representa em uma contração e $C < 0$ representa em uma dilatação da porta traseira. Assim, considerando inicialmente, com toda

generalidade, que houve um Deslocamento, uma Rotação e uma Contração/Dilatação nas folgas, o seguinte modelo pode ser adotado para representar as variáveis observadas (folgas centradas):

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1/2 & -1/2 & 1/2 \\ 1/2 & 1/2 & 1/2 \\ -1/2 & 1/2 & 1/2 \\ -1/2 & -1/2 & 1/2 \end{bmatrix} \begin{bmatrix} L \\ T \\ C \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{bmatrix} \quad (13)$$

ou ainda,

$$\begin{aligned} x_1 &= 0,5(L - T + C) + e_1 \\ x_2 &= 0,5(L + T + C) + e_2 \\ x_3 &= 0,5(- L + T + C) + e_3 \\ x_4 &= 0,5(- L - T + C) + e_4 \end{aligned} \quad (14)$$

onde e_1 , e_2 , e_3 e e_4 são ruídos aleatórios, que foram supostos independentes e identicamente distribuídos com distribuição Normal com média zero e variância σ_e^2 .

A equação (13) corresponde ao modelo na eq. (1), com $p = 4$, $k = 3$, $\mathbf{d} = [L \ T \ C]'$, e a matriz \mathbf{C} tem seu elementos de módulo 0,5, pois suas colunas devem ter norma um. Ademais, conforme suposição básica, as variáveis L , T , C e os ruídos e_i são independentes entre si. Note que L , T e C terão cada um, sinal positivo ou negativo, em função do seu sentido.

3.2 Variáveis Transformadas

Conforme embasamento teórico contido na seção 2.2, pode-se definir, respectivamente a T , L e C , as variáveis $T(x)$, $L(x)$ e $C(x)$ como segue:

$$\begin{bmatrix} T(x) \\ L(x) \\ C(x) \end{bmatrix} = \begin{bmatrix} -1/2 & 1/2 & 1/2 & -1/2 \\ 1/2 & 1/2 & -1/2 & -1/2 \\ 1/2 & 1/2 & 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \quad (15)$$

ou ainda,

$$\begin{aligned} T(x) &= (-x_1 + x_2 + x_3 - x_4)/2 \\ L(x) &= (+x_1 + x_2 - x_3 - x_4)/2 \\ C(x) &= (+x_1 + x_2 + x_3 + x_4)/2 \end{aligned} \quad (16)$$

em notação matricial,

$$\begin{aligned} T(x) &= \mathbf{c}_1' \mathbf{x} \\ L(x) &= \mathbf{c}_2' \mathbf{x} \\ C(x) &= \mathbf{c}_3' \mathbf{x} \end{aligned} \quad (17)$$

onde $\mathbf{c}_1' = [-1/2 \ 1/2 \ 1/2 \ -1/2]$, $\mathbf{c}_2' = [1/2 \ 1/2 \ -1/2 \ -1/2]$ e $\mathbf{c}_3' = [1/2 \ 1/2 \ 1/2 \ 1/2]$ são vetores unitários nas direções de variação conhecidas (Rotação, Deslocamento e Contração/Dilatação). Dessa forma, as variáveis transformadas são as projeções do vetor de variáveis observadas nessas direções. Veja que os vetores unitários, \mathbf{c}_1' , \mathbf{c}_2' e \mathbf{c}_3' , formam uma base ortonormal. Além disso, substituindo (14) na expressão (16) para $T(x)$, pode-se chegar ao seguinte resultado:

$$T(x) = \frac{1}{2}(2T + \sum_{i=1}^4 e_i) = T + \frac{1}{2}(-e_1 + e_2 + e_3 - e_4) \quad (18)$$

Assim, conforme previsto, $T(x)$ é uma medida ruidosa de T . Ainda, com base nas equações (6) e (7) os valores da esperança (média) e da variância para $T(x)$ podem ser determinados. Assim,

$$E[T(x)] = \begin{cases} 0, & \text{se } E(T) = 0 \\ \delta, & \text{se } E(T) = \delta \end{cases} \quad \begin{matrix} \text{(se o processo estiver em controle)} \\ \text{(se houver uma alteração de } \delta \text{ na média de } T) \end{matrix} \quad (19)$$

e

$$V[T(x)] = \sigma_T^2 + \sigma_e^2 \Rightarrow DP[T(x)] = \sqrt{\sigma_T^2 + \sigma_e^2} \quad (20)$$

onde $DP[T(x)]$ denota o desvio-padrão de $T(x)$.

Raciocínio análogo mostra que $L(x)$ e $C(x)$ são medidas ruidosas de L e C , com média e variância dadas por expressões análogas a (19) e (20).

3.3 Análise de Componentes Principais

Para verificar se a ACP identifica eficazmente as direções conhecidas de variação, geramos aleatoriamente conjuntos de valores de L , T , C e e_i , combinando-os em seguida segundo o modelo (16) para simular dados históricos de medidas de folgas. A esses dados, então, aplicamos a ACP, para verificar se os vetores de coeficientes das primeiras componentes são (dentro de uma boa aproximação) paralelos aos vetores de coeficientes \mathbf{c}_1' , \mathbf{c}_2' e \mathbf{c}_3' , definidos de acordo com a notação matricial da expressão (17).

3.4 Vetor Resíduo

Para a análise de desempenho, será suposto que as direções de variação que afetam o processo mais frequentemente são Rotação (T) e Deslocamento (L), deixando a Contração/Dilatação (C) como causa especial “nova”, para observar a eficácia do monitoramento do resíduo em detectar a ocorrência dessa nova causa. Dessa maneira, de acordo com a expressão (12), pode-se definir o vetor resíduo na base x_1, x_2, x_3 e x_4 da seguinte forma:

$$\mathbf{r} = \mathbf{x} - [\mathbf{t}(\mathbf{x}) + \mathbf{l}(\mathbf{x})] \quad (21)$$

onde \mathbf{x} é o vetor de observações na base x_1, x_2, x_3 e x_4 , $\mathbf{t}(\mathbf{x})$ e $\mathbf{l}(\mathbf{x})$ são as projeções do vetor \mathbf{x} nas direções $T(x)$ e $L(x)$, respectivamente.

A representação geométrica do vetor resíduo para o caso considerado pode ser vista na figura a seguir.

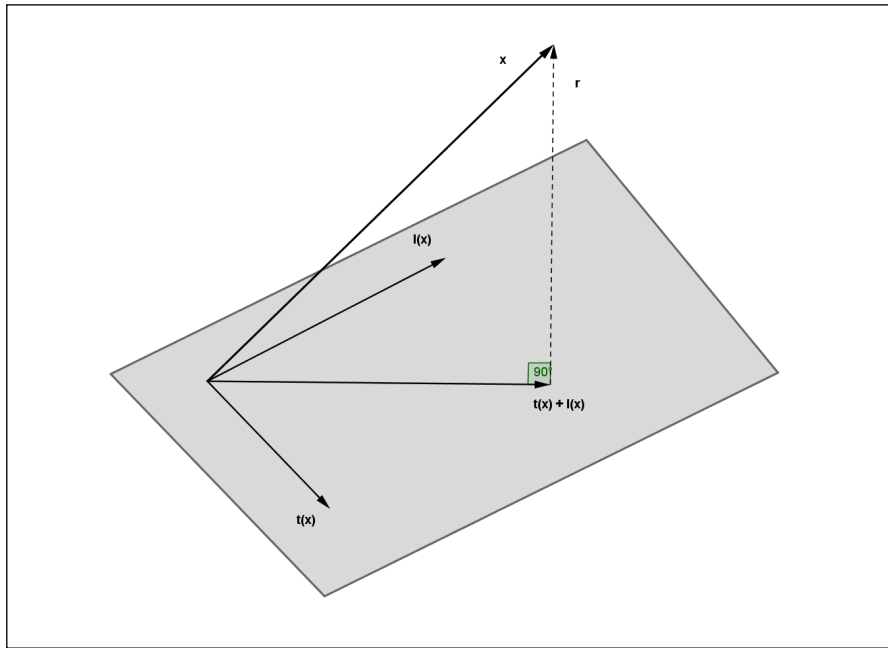


Figura 5: Representação geométrica do vetor resíduo (\mathbf{r})

Assim, pelo Teorema de Pitágoras, no caso proposto, tem-se:

$$\|\mathbf{r}\|^2 = \|\mathbf{x}\|^2 - \|\mathbf{t}(\mathbf{x}) + \mathbf{l}(\mathbf{x})\|^2 \quad (22)$$

mas, note que $\mathbf{t}(\mathbf{x})$ e $\mathbf{l}(\mathbf{x})$ são vetores ortogonais, e além disso, o vetor \mathbf{x} também é perpendicular, isto é, ortogonal ao vetor soma $\mathbf{t}(\mathbf{x}) + \mathbf{l}(\mathbf{x})$. Portanto,

$$\|\mathbf{r}\|^2 = \|\mathbf{x}\|^2 - \|\mathbf{t}(\mathbf{x})\|^2 - \|\mathbf{l}(\mathbf{x})\|^2 \quad (23)$$

Dessa forma, desenvolvendo cada parte do lado direito da equação (23) separadamente, temos:

$$\|\mathbf{x}\|^2 = x_1^2 + x_2^2 + x_3^2 + x_4^2$$

mas,

$$x_1 = -\frac{T}{2} + \frac{L}{2} + e_1 \Rightarrow x_1^2 = \left(-\frac{T}{2} + \frac{L}{2}\right)^2 + 2\left(-\frac{T}{2} + \frac{L}{2}\right)e_1 + e_1^2 \Rightarrow$$

$$\Rightarrow x_1 = \frac{T^2}{4} + \frac{L^2}{4} - \frac{TL}{2} - Te_1 + Le_1 + e_1^2$$

$$x_2 = \frac{T}{2} + \frac{L}{2} + e_2 \Rightarrow x_2^2 = \left(\frac{T}{2} + \frac{L}{2}\right)^2 + 2\left(\frac{T}{2} + \frac{L}{2}\right)e_2 + e_2^2 \Rightarrow$$

$$\Rightarrow x_2 = \frac{T^2}{2} + \frac{L^2}{2} + \frac{TL}{2} + Te_2 + Le_2 + e_2^2$$

$$x_3 = \frac{T}{2} - \frac{L}{2} + e_3 \Rightarrow x_3^2 = \left(\frac{T}{2} - \frac{L}{2}\right)^2 + 2\left(\frac{T}{2} - \frac{L}{2}\right)e_3 + e_3^2 \Rightarrow$$

$$\Rightarrow x_3 = \frac{T^2}{2} + \frac{L^2}{2} - \frac{TL}{2} + Te_3 - Le_3 + e_3^2$$

e

$$x_4 = -\frac{T}{2} - \frac{L}{2} + e_4 \Rightarrow x_4^2 = \left(-\frac{T}{2} - \frac{L}{2}\right)^2 + 2\left(-\frac{T}{2} - \frac{L}{2}\right)e_4 + e_4^2 \Rightarrow$$

$$\Rightarrow x_4 = \frac{T^2}{2} + \frac{L^2}{2} + \frac{TL}{2} - Te_4 - Le_4 + e_4^2$$

daí,

$$\begin{aligned} \|\mathbf{x}\|^2 &= x_1^2 + x_2^2 + x_3^2 + x_4^2 = T^2 + L^2 + (e_1^2 + e_2^2 + e_3^2 + e_4^2) + \\ &\quad + T(-e_1 + e_2 + e_3 - e_4) + L(e_1 + e_2 - e_3 - e_4) \end{aligned} \quad (24)$$

Analisando agora com os termos $\|\mathbf{t}(\mathbf{x})\|^2$ e $\|\mathbf{l}(\mathbf{x})\|^2$. Tem-se que $\|\mathbf{t}(\mathbf{x})\|^2 = T^2(x)$ e $\|\mathbf{l}(\mathbf{x})\|^2 = L^2(x)$, mas foi visto em (18) que $T(x) = T + \frac{1}{2}(-e_1 + e_2 + e_3 - e_4)$ e por analogia, $L(x) = L + \frac{1}{2}(e_1 + e_2 - e_3 - e_4)$. Dessa forma,

$$\begin{aligned} T^2(x) &= T^2 + T(-e_1 + e_2 + e_3 - e_4) + \frac{1}{4}(e_1^2 + e_2^2 + e_3^2 + e_4^2) + \\ &\quad + \frac{1}{4}(-2e_1e_2 - 2e_1e_3 + 2e_1e_4 + 2e_2e_3 - 2e_2e_4 - 2e_3e_4) \end{aligned} \quad (25)$$

e

$$L^2(\mathbf{x}) = L^2 + L(\mathbf{e}_1 + \mathbf{e}_2 - \mathbf{e}_3 - \mathbf{e}_4) + \frac{1}{4}(\mathbf{e}_1^2 + \mathbf{e}_2^2 + \mathbf{e}_3^2 + \mathbf{e}_4^2) + \frac{1}{4}(2\mathbf{e}_1\mathbf{e}_2 - 2\mathbf{e}_1\mathbf{e}_3 - 2\mathbf{e}_1\mathbf{e}_4 - 2\mathbf{e}_2\mathbf{e}_3 - 2\mathbf{e}_2\mathbf{e}_4 + 2\mathbf{e}_3\mathbf{e}_4) \quad (26)$$

assim,

$$T^2(\mathbf{x}) + L^2(\mathbf{x}) = T^2 + L^2 + T(-\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3 - \mathbf{e}_4) + L(\mathbf{e}_1 + \mathbf{e}_2 - \mathbf{e}_3 - \mathbf{e}_4) + \frac{1}{2}(\mathbf{e}_1^2 + \mathbf{e}_2^2 + \mathbf{e}_3^2 + \mathbf{e}_4^2) - \mathbf{e}_1\mathbf{e}_3 - \mathbf{e}_2\mathbf{e}_4 \quad (27)$$

então,

$$\|\mathbf{r}\|^2 = \|\mathbf{x}\|^2 - \|\mathbf{t}(\mathbf{x})\|^2 - \|\mathbf{l}(\mathbf{x})\|^2 = \frac{1}{2}(\mathbf{e}_1^2 + \mathbf{e}_2^2 + \mathbf{e}_3^2 + \mathbf{e}_4^2) + \mathbf{e}_1\mathbf{e}_3 + \mathbf{e}_2\mathbf{e}_4 \quad (28)$$

Dessa maneira, pode-se determinar a esperança e a variância do quadrado da norma do vetor resíduo conforme o que se segue

$$\begin{aligned} E\|\mathbf{r}\|^2 &= E\|\mathbf{x}\|^2 - E\|\mathbf{t}(\mathbf{x})\|^2 - E\|\mathbf{l}(\mathbf{x})\|^2 \Rightarrow \\ \Rightarrow E\|\mathbf{r}\|^2 &= \frac{1}{2}E(\mathbf{e}_1^2 + \mathbf{e}_2^2 + \mathbf{e}_3^2 + \mathbf{e}_4^2) + E(\mathbf{e}_1\mathbf{e}_3) + E(\mathbf{e}_2\mathbf{e}_4) \end{aligned} \quad (29)$$

Lembrando que os erros aleatórios são supostos independentes e identicamente distribuídos com $\mathbf{e}_i \sim N(0, \sigma_e^2)$, é fácil verificar que $E(\mathbf{e}_i^2) = V(\mathbf{e}_i) = \sigma_e^2$, pois $E(\mathbf{e}_i) = 0$. Além disso, pode-se concluir que $E(\mathbf{e}_i\mathbf{e}_j) = 0 \forall i \neq j$. Assim,

$$E\|\mathbf{r}\|^2 = 2\sigma_e^2 \quad (30)$$

e

$$V\|\mathbf{r}\|^2 = \frac{1}{4}V(\mathbf{e}_1^2 + \mathbf{e}_2^2 + \mathbf{e}_3^2 + \mathbf{e}_4^2) + V(\mathbf{e}_1\mathbf{e}_3) + V(\mathbf{e}_2\mathbf{e}_4) \quad (31)$$

Note que, embora o cálculo da variância não seja tão trivial como o da esperança, pois não se conhece *a priori* a distribuição do quadrado dos erros, isso não chega a ser um obstáculo, visto que essa distribuição pode ser determinada. Assim, seja

$$Y = e_i^2 \quad \text{onde} \quad e_i \sim N(0, \sigma_e^2) \Rightarrow f(e_i) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-e_i^2}{2\sigma_e^2}\right) I(e_i)_{(-\infty, \infty)}.$$

Função de Distribuição de Y dada por:

$$F_Y(y) = P(Y \leq y) = P(e_i^2 \leq y) = P(-\sqrt{y} \leq e_i \leq \sqrt{y}) = F_{e_i}(\sqrt{y}) - F_{e_i}(-\sqrt{y})$$

mas, $f_Y(y) = \frac{dF_Y(y)}{dy}$, pois Y é uma variável aleatória contínua. Então,

$$f_Y(y) = f_{e_i}(\sqrt{y}) \cdot \frac{1}{2} y^{-1/2} + f_{e_i}(-\sqrt{y}) \cdot \frac{1}{2} y^{-1/2} = \frac{1}{\sqrt{2\pi}\sigma} y^{-1/2} \exp\left(\frac{-y}{2\sigma_e^2}\right) I(y)_{[0, \infty)}$$

Observe que essa distribuição obtida coincide com uma distribuição Gama ($G(r, \alpha)$) com parâmetros $r = \frac{1}{2}$ e $\alpha = \frac{1}{2\sigma_e^2}$. De fato, se uma variável aleatória X segue uma distribuição Gama com parâmetros $r > 0$ e $\alpha > 0$, então a sua função densidade de probabilidade tem a seguinte expressão:

$$f_X(x) = \frac{\alpha}{\Gamma(r)} (\alpha x)^{r-1} \exp(-\alpha x)$$

Dessa forma, pode-se dizer que a distribuição dos quadrados dos erros aleatórios segue uma distribuição Gama com os referidos parâmetros e com média e variância dados por:

$$E(Y) = E(e_i^2) = \frac{r}{\alpha} = \sigma_e^2 \quad \text{e} \quad V(Y) = V(e_i^2) = \frac{r}{\alpha^2} = 2(\sigma_e^2)^2$$

$$\text{assim, } \frac{1}{4} V(e_1^2 + e_2^2 + e_3^2 + e_4^2) = 2(\sigma_e^2)^2$$

Além disso, para o cálculo da variância de $\|\mathbf{r}\|^2$ é necessário determinar $V(e_1e_3)$ e $V(e_2e_4)$. Desse modo, de maneira geral pode-se mostrar que:

$$V(e_i e_j) = V(e_i)V(e_j) + (E(e_i))^2 V(e_j) + (E(e_j))^2 V(e_i)$$

Mas, sabe-se que $E(e_i) = E(e_j) = 0$

então,

$$V(e_i e_j) = V(e_i)V(e_j)$$

logo, $V(e_1e_3) = V(e_2e_4) = (\sigma_e^2)^2$

Dessa forma, $V\|\mathbf{r}\|^2$ pode finalmente ser determinada. Assim,

$$V\|\mathbf{r}\|^2 = 4(\sigma_e^2)^2 \quad (32)$$

$$DP(\|\mathbf{r}\|^2) = 2\sigma_e^2 \quad (33)$$

Conforme visto no desenvolvimento do caso geral, não é trivial determinar a distribuição de probabilidade para o quadrado da norma do vetor resíduo. Contudo, para esse caso específico, verifica-se que a média e o desvio-padrão de $\|\mathbf{r}\|^2$ são iguais, o que pode ser um indício da forma da distribuição de probabilidade do quadrado da norma desse vetor. Será utilizada simulação para verificação da forma distribucional do vetor resíduo.

3.5 Metodologia para a Análise de Desempenho

Para todas as análises, foi utilizada a simulação, a qual foi implementada no programa computacional R, um *software* livre de código aberto. As análises realizadas foram:

- Avaliação da eficácia da ACP em identificar as direções conhecidas;
- Análise de desempenho dos gráficos nas variáveis transformadas, quer no caso ideal em que as direções de variação conhecidas são conhecidas, bem como no caso em que são estimadas pela ACP;
- Análise de desempenho dos demais gráficos concorrentes (T^2 , gráficos de Shewhart nas variáveis observadas) nos mesmos casos que os propostos, para comparação;
- Análise da forma da distribuição de probabilidade do quadrado da norma do vetor resíduo (com o processo em controle);
- Análise de desempenho do gráfico no quadrado da norma do vetor resíduo.

Para a realização das referidas análises, considerou-se inicialmente o modelo dado pela equação em (1). Em consequência, para a obtenção das variáveis observadas (x_1, x_2, x_3 e x_4) foram gerados valores das variações nas direções conhecidas T, L e C, e das flutuações aleatórias (e_1, e_2, e_3 e e_4) segundo distribuições normais, sendo essas geradas pela função **rnorm**, a qual reproduz uma quantidade determinada de valores pseudo-aleatórios seguindo uma distribuição normal com os parâmetros média e desvio-padrão especificados.

Em seguida, foram simuladas as variáveis transformadas e as componentes principais de acordo com as equações (4) e (8), respectivamente. Para obtenção das componentes principais foi usada a função **prcomp** e, para determinar quais componentes devem ser consideradas no modelo, foi utilizado como critério o

percentual de variância explicada das componentes principais, calculado pela função **summary(prcomp)**.

Com relação à análise de desempenho, utilizou-se como indicador o ARL, sigla em inglês de “*Average Run Length*”, chamado em português de comprimento médio da sequência (vide Montgomery, 2004), ou número médio de amostras até a ocorrência de um sinal (NMA) (vide Costa *et al.*, 2005). Para obtenção dos alarmes, tanto para a Fase I (de estimação dos parâmetros de controle) quanto para a Fase II (de monitoramento do processo), foram utilizados os gráficos de Shewhart em cada variável observada, nas variáveis transformadas e nas componentes principais, e o gráfico de T^2 de Hotelling no conjunto das variáveis observadas. Para tanto, utilizou-se os pacotes **qcc** (*quality control chart*) e **mqcc** (*multivariate quality control chart*) encontrados no próprio programa.

Ainda, no sentido de obter mais subsídios para comparação entre os desempenhos dos esquemas de controle assinalados, considerou-se também a contagem de alarmes em pelo menos um dos gráficos nas variáveis observadas, nas variáveis transformadas e nas CP's. Além disso, para que as comparações fossem realizadas em igualdade de condições, foram determinados os ARL's levando em consideração ARL₀'s aproximados ao obtido pelo gráfico de T^2 .

Para a estimação dos limites de controle (Fase I), considera-se que o processo esteja sob controle. Sendo assim, os valores das variações conhecidas (T, L e C) foram gerados segundo distribuições normais com média nula e variância constante.

Além disso, para verificar se na Fase I do CEP, a ACP identifica eficazmente as direções conhecidas de variação, foi aplicado o procedimento apresentado na seção 2.3, sendo os valores de T, L e C gerados segundo a distribuição em controle dessas variáveis (com média nula e variância constante).

Para analisar o desempenho dos gráficos de controle propostos na Fase II, de monitoramento do processo, foram usadas como limites de controle as estimativas dos parâmetros obtidas na Fase I. Assim, foram consideradas as seguintes situações:

- Para o monitoramento da média, foram consideradas e simuladas alterações na média das direções conhecidas (T, L e C) de um e dois desvios-padrão em relação ao valor alvo (média nula);

- Para o controle do desvio-padrão, foram considerados e simulados aumentos de 50% e 100% no verdadeiro desvio-padrão das direções conhecidas. Assim, considerando que as direções conhecidas sejam normalmente distribuídas com média zero e variância 1, os aumentos simulados foram de 1,5 e 2 desvios-padrão, respectivamente;
- No caso de alterações simultâneas na média e no desvio padrão, considerou-se à combinação dos valores usados na avaliação individual para alterações na média e aumentos na variância conforme descrito nos itens anteriores.

Foram considerados na simulação para o monitoramento da média e/ou da variância os seguintes casos:

- Uma única direção de perturbação: Rotação;
- Duas direções de perturbação: Rotação e Deslocamento, com variâncias iguais;
- Duas direções de variação: Rotação e Deslocamento, com a variância da Rotação superior à do Deslocamento;
- Três direções de variação: Rotação, Deslocamento e Contração/Dilatação, com variâncias iguais.

Em todos os casos, foram gerados os valores das variáveis observadas considerando um comprimento m de 40.000 observações individuais do vetor \mathbf{x} com três magnitudes do desvio-padrão do erro aleatório (σ_e): 0.1; 0.5 e 1.0. Ressalta-se ainda que, conforme o caso, algumas componentes de variação conhecidas não foram consideradas, isto é, em alguns casos adotou-se $C=0$, e em outro, além disso, $L=0$.

No caso da análise de desempenho do gráfico usado no monitoramento do vetor resíduo, o primeiro passo foi encontrar a forma da distribuição de probabilidade do $\|\mathbf{r}\|^2$, dado que esta pode ser bastante distinta de uma distribuição normal, o que pode resultar em probabilidades de alarme falso bastante elevadas, caso se adotem, sem maior análise, limites “de três sigmas”. Portanto, devido à dificuldade de determinar a distribuição exata de $\|\mathbf{r}\|^2$, optou-se pela utilização da análise gráfica juntamente com o teste de adequação de Anderson-Darling (AD) a fim de encontrar, com base nos dados

históricos obtidos (simulados) de $\|\mathbf{r}\|^2$ a distribuição que mais se ajusta a esses dados. O teste de Anderson-Darling (AD) verifica a qualidade do ajuste de uma variável a uma determinada distribuição de probabilidade, sendo considerado superior aos testes qui-quadrado e Kolmogorov-Smirnov, principalmente no ajuste dos dados às caudas da distribuição, o que é mais relevante para o estabelecimento de limites de controle.

Após ser identificada a forma distribucional mais adequada aos dados históricos, foi iniciada a simulação da Fase I, com novos dados, para estimação dos parâmetros e cálculo dos limites de controle.

Para a estimação dos limites de controle (Fase I), simularam-se como direções conhecidas mais frequentes a Rotação (T) e o Deslocamento (L), e os parâmetros foram estimados de acordo com a distribuição de probabilidade do quadrado da norma dos vetores resíduos testados considerando três valores para a probabilidade de alarme falso (α): 0,0027; 0,005 e 0,01.

No que diz respeito ao desempenho dos gráficos de controle para a Fase II, foram usados como limites de controle as estimativas dos parâmetros obtidos na Fase I. Foram adotados limites de acordo com a distribuição que mais se ajustou aos dados históricos. Ainda, para a análise de desempenho do gráfico do quadrado da norma do vetor resíduo foi simulada a presença de outra direção de variação, mais precisamente, uma Contração (C).

Para todas as análises referentes ao vetor resíduo utilizou-se para ambas as fases um comprimento de corrida $m = 40.000$ observações individuais, e cinco magnitudes do desvio-padrão do erro aleatório (σ_e): 0.1; 0.5; 0,75; 1.0 e 1,5. A razão de usar cinco magnitudes é explicada pela necessidade de verificar se a forma da distribuição de probabilidade da norma ao quadrado do vetor resíduo é independente da magnitude do desvio-padrão do erro aleatório.

4 Resultados

A seguir serão apresentados os resultados obtidos por meio das simulações realizadas dos casos mencionados no capítulo anterior.

4.1 Casos de alterações na média

4.1.1 Uma única direção de perturbação

Considerando inicialmente que o processo está sob controle, o Quadro 1 mostra o resultado da aplicação da técnica de análise de componentes principais às variáveis observadas para $m = 40.000$ observações individuais do vetor \mathbf{x} . Observe que quando só há uma direção conhecida, no caso, a rotação, o modelo em (1) reduz-se a:

$$x_1 = -0,5T + e_1$$

$$x_2 = 0,5T + e_2$$

$$x_3 = 0,5T + e_3$$

$$x_4 = -0,5T + e_4$$

Quadro 1: Resultado da Análise de Componentes Principais para o caso de alterações na média com uma única direção de perturbação.

Variáveis Observadas	$\sigma_e = 0.1$				$\sigma_e = 0.5$				$\sigma_e = 1.0$			
	CP ₁	CP ₂	CP ₃	CP ₄	CP ₁	CP ₂	CP ₃	CP ₄	CP ₁	CP ₂	CP ₃	CP ₄
x_1	-0.50	0.26	-0.58	0.60	-0.50	0.57	-0.50	0.42	0.56	-0.22	-0.26	-0.75
x_2	0.50	-0.43	0.14	0.74	0.50	-0.43	-0.60	0.46	-0.49	0.05	0.64	-0.59
x_3	0.51	0.85	0.09	0.14	0.49	0.46	0.50	0.55	-0.47	0.53	-0.65	-0.28
x_4	-0.50	0.17	0.80	0.28	-0.51	-0.54	0.38	0.56	0.48	0.82	0.31	0.01
Desvio-padrão	2.03	0.103	0.095	0.091	2.05	0.521	0.491	0.476	2.19	1.09	1.00	0.94
Prop. da Variância Explicada	0.99	0.003	0.002	0.002	0.85	0.06	0.05	0.05	0.61	0.15	0.13	0.11

Nota-se que a primeira componente é responsável pela maior parte da variância total explicada (99%), o que é esperado, visto que nesse caso específico, apenas uma direção conhecida foi considerada.

Ainda, observando os coeficientes da primeira componente principal para cada desvio-padrão do erro aleatório, pode-se dizer que para uma única direção de

perturbação, a técnica da ACP mostrou-se eficaz em identificar a direção de variação, ou seja, a direção do vetor \mathbf{x} obtida pela ACP coincide com a direção conhecida “Rotação”. Além disso, como seria de esperar, a percentagem da variação explicada cai com o aumento da variância do erro aleatório em x_1 , x_2 , x_3 e x_4 (a rigor com o aumento da razão entre a variância do erro aleatório e a variância do deslocamento na direção conhecida), e o módulo dos coeficientes da 1ª CP desviam-se um pouco dos valores dos elementos do vetor unitário na direção conhecida, ou seja, a precisão da identificação da direção conhecida pela 1ª CP diminui com o aumento do desvio-padrão do erro aleatório nas variáveis observadas.

Outra informação interessante, que ajudará a compreender o desempenho relativo dos diversos gráficos, é o grau de interdependência entre as variáveis observadas, que, como poderá ser visto no Quadro a seguir, apresenta correlações significativas, especialmente para σ_e pequeno.

Quadro 2: Correlações entre as variáveis observadas para o processo em controle para o caso de alterações na média com uma única direção de perturbação.

	$\sigma_e = 0.1$				$\sigma_e = 0.5$				$\sigma_e = 1.0$			
	x_1	x_2	x_3	x_4	x_1	x_2	x_3	x_4	x_1	x_2	x_3	x_4
x_1	1.00				1.00				1.00			
x_2	-0.99	1.00			-0.81	1.00			-0.55	1.00		
x_3	-0.99	0.99	1.00		-0.80	0.79	1.00		-0.52	0.46	1.00	
x_4	0.99	-0.99	-0.99	1.00	0.80	-0.80	-0.81	1.00	0.48	-0.45	-0.40	1.00

Para a análise de desempenho dos esquemas propostos foi utilizado o ARL, conforme mencionado na metodologia. Assim, o Quadro 3, apresenta o ARL para o gráfico de Shewhart em cada variável observada individualmente, na primeira CP e na variável transformada Rotação ($T(x)$); o ARL conjunto referente aos gráficos nas variáveis observadas e o ARL para o gráfico de T^2 de Hotelling no conjunto das variáveis observadas.

Quadro 3: ARL para o caso de alterações na média considerando apenas uma direção de perturbação.

DIREÇÃO: ROTAÇÃO									
Fase I		ARL ₀							
Casos	σ_e	X_1	X_2	X_3	X_4	$\cup X_i$	T^2	$T(x)$	CP_1
$T \sim N(0,1)$	0,1	434,78	400,00	396,04	384,62	261,44	87,91	396,04	396,04
	0,5	400,00	377,36	373,83	333,33	142,86	96,85	388,35	388,35
	1,0	430,11	377,36	392,16	412,37	101,52	91,12	392,16	392,16
Fase II		ARL							
$T \sim N(1,1)$	0,1	45,15	45,25	43,76	45,10	36,43	31,30	45,05	44,89
	0,5	60,06	66,45	61,82	61,44	24,38	33,67	52,29	52,77
	1,0	95,92	92,81	100,00	90,09	27,91	36,80	58,65	58,48
$T \sim N(2,1)$	0,1	6,39	6,31	6,30	6,32	5,52	6,98	6,29	6,25
	0,5	9,29	9,56	9,30	9,24	4,88	7,40	7,29	7,32
	1,0	18,29	17,25	18,31	19,28	6,51	9,36	9,06	9,05

onde:

X_i = gráfico de controle de Shewhart para cada variável observada x_i .

$\cup X_i$ = ARL conjunto dos gráficos nas variáveis observadas x_1, x_2, x_3 e x_4 (número médio de amostras até um sinal em pelo menos um dos gráficos).

T^2 = gráfico de controle T^2 de Hotelling.

$T(x)$ = gráfico de controle de Shewhart para a variável transformada rotação $T(x)$.

CP_1 = gráfico de controle de Shewhart para a primeira componente principal.

Dessa forma, analisando o referido quadro, observa-se que:

- Para a Fase I, todos os resultados mostram-se satisfatórios, pois o número médio de amostras até a ocorrência de um alarme falso encontra-se próximo do esperado, isto é, aproximadamente um alarme falso a cada 370 pontos observados para os gráficos de Shewhart, e um alarme falso a cada 93 pontos para o gráfico T^2 de Hotelling, considerando quatro variáveis observadas. Além disso, pode-se verificar que os ARL_0 de $T(x)$ e da CP_1 são iguais. Isso acontece porque, conforme visto no Quadro 1, a primeira componente coincide com a direção conhecida (Rotação);

- ii. Na Fase II, verificou-se que, a exemplo do que ocorreu na fase de estimação dos parâmetros (Fase I), os dois gráficos ($T(x)$ e CP_1) têm ARL muito próximos, dado que a variável transformada $T(x)$ e a primeira componente principal CP_1 praticamente coincidem.
- iii. Para esse caso específico, qualquer um dos gráficos nas variáveis observadas tende a emitir o mesmo número médio de alarmes que o gráfico $T(x)$ ou na 1ª CP. Isso porque, como as correlações entre as variáveis observadas são todas altas (ver Quadro 2), essas tendem a se comportar como uma só variável. Note-se, porém, que quando o desvio-padrão do erro aleatório cresce, o ARL fora de controle nos gráficos das variáveis observadas é maior, pois estas são mais afetadas pelo erro aleatório. Nestes casos, os gráficos de $T(x)$ e na 1ª CP desempenham melhor (possuem menor ARL) quando comparados ao ARL individual de cada variável observada, pois, combinando as informações das quatro variáveis (ao fazerem a projeção na direção $T(x)$ ou da 1ª CP), filtram melhor o ruído, sendo assim mais sensíveis ao “sinal”. Além disso, os gráficos na variável observada continuam com a desvantagem de utilizar quatro gráficos contra apenas um de $T(x)$ ou CP_1 .
- iv. Aparentemente os gráficos $T(x)$ e CP_1 têm menor eficácia de sinalização quando comparados ao gráfico de T^2 e ao número de alarmes conjunto nas quatro variáveis observadas. No entanto, verifica-se que esses gráficos ($T(x)$ e CP_1), não estão sendo comparados com os demais em condições de igualdade, pois tanto o gráfico de T^2 , como o desempenho conjunto dos gráficos nas variáveis observadas ($\bigcup X_i$) possuem menores ARL_0 .

Para que se possam fazer comparações entre os diversos esquemas propostos é necessário considerar um mesmo ARL_0 . Assim, foram ajustados todos os valores do risco de alarme falso (α) dos esquemas considerados no estudo para produzir ARL_0 's iguais a 200. Com os valores dos ARL_0 's ajustados para cada esquema, é possível recalcular os ARL's para a Fase II do controle. O resultado desse ajuste é apresentado no Quadro 4.

Quadro 4: ARL para o caso de alterações na média considerando apenas uma direção de perturbação e ARL_0 's ajustados.

DIREÇÃO: ROTAÇÃO					
Fase I		ARL_0			
Casos	σ_e	$\cup X_{i\alpha-aj}$	T^2	$T(X)_{\alpha-aj}$	$CP_{1\alpha-aj}$
$T \sim N(0,1)$	0,1	200			
	0,5				
	1,0				
Fase II		ARL			
$T \sim N(1,0)$	0,1	29,72	56,26	29,20	29,20
	0,5	30,94	59,00	28,43	28,43
	1,0	49,08	70,05	37,59	37,59
$T \sim N(2,0)$	0,1	5,10	10,90	5,03	5,03
	0,5	5,56	11,53	5,08	5,08
	1,0	9,47	14,20	6,62	6,62

onde:

$\cup X_{i\alpha-aj}$ = ARL conjunto dos gráficos de controle nas variáveis observadas, considerando um risco de alarme falso (α) ajustado para produzir um ARL_0 igual a 200;

$T(x)_{\alpha-aj}$ = gráfico de controle para a variável transformada $T(x)$, considerando um risco de alarme falso (α) ajustado para produzir um ARL_0 igual a 200;

$CP_{1\alpha-aj}$ = gráfico de controle para a primeira componente principal considerando um risco de alarme falso (α) ajustado para produzir um ARL_0 para o gráfico CP_1 próximo ao ARL_0 do gráfico de T^2 .

Baseado nos novos cálculos do ARL, considerando a igualdade dos respectivos ARL_0 's, constata-se por meio do Quadro 4, que os gráficos na variável transformada $T(x)$ e na primeira componente principal PC_1 obtiveram melhor desempenho em todos os casos analisados, demonstrando assim maior eficácia quando comparados em igualdade de condições.

4.1.2 Duas direções de perturbação com mesma variância

Seguindo procedimento de análise semelhante ao caso anterior, os Quadros 5 e 6 apresentam respectivamente, o resultado da ACP e do ARL considerando que duas direções de deslocamento atuem no problema. Nesse caso, o modelo adotado é dado por:

$$x_1 = 0,5L - 0,5T + e_1$$

$$x_2 = 0,5L + 0,5T + e_2$$

$$x_3 = -0,5L + 0,5T + e_3$$

$$x_4 = -0,5L - 0,5T + e_4$$

Com relação aos gráficos nas CP's, utilizou-se gráficos apenas nas componentes mais significativas, segundo a proporção de variância explicada. Dessa forma, apenas as duas primeiras CP's foram consideradas, visto que essas componentes são responsáveis por quase toda variabilidade explicada.

Quadro 5: Resultado da Análise de Componentes Principais para o caso de alterações na média considerando duas direções de perturbação com mesma variância.

Variáveis Observadas	$\sigma_e = 0.1$		$\sigma_e = 0.5$		$\sigma_e = 1.0$	
	CP ₁	CP ₂	CP ₁	CP ₂	CP ₁	CP ₂
x_1	-0.07	0.71	0.62	-0.34	0.26	-0.68
x_2	-0.71	-0.07	-0.34	-0.63	-0.62	-0.30
x_3	0.07	-0.70	-0.62	0.35	-0.24	0.64
x_4	0.70	0.07	0.35	0.61	0.70	0.19
Desvio-padrão	2.07	1.82	2.24	2.06	2.33	2.14
Proporção da Variância Explicada	0.56	0.436	0.51	0.435	0.45	0.38

Dessa forma, pode-se observar no Quadro 5 que, diferentemente do que ocorreu quando apenas uma direção conhecida atuou no problema, no caso em que há duas direções conhecidas, com a mesma variância, e estando o processo sob controle, as CP's não conseguiram identificar as direções de variação. Para ver isso, basta verificar que os coeficientes das componentes não são proporcionais aos coeficientes das direções

conhecidas consideradas, o que implica que as CP's e as direções conhecidas não seguem a mesma direção. A provável explicação para isso é que, se as variâncias das duas perturbações em direções conhecidas (T e L) são iguais, os eixos da elipse que representa os contornos de isoprobabilidade no plano definido por T e L possuem igual comprimento: assim, a elipse degenera em um círculo, o que tem como consequência uma indeterminação das direções dos eixos principais. Embora as duas primeiras componentes principais identificadas pela ACP continuem contidas no mesmo plano, as suas direções nesse plano tornam-se aleatórias: função dos valores particulares dos dados, dominadas pela aleatoriedade experimental.

Note que, conforme foi feito quando considerado apenas uma direção conhecida, no Quadro 6 a seguir também foi incluído o ARL conjunto dos gráficos nas variáveis observadas, o ARL conjunto dos gráficos nas CP's e o ARL conjunto dos gráficos nas variáveis transformadas.

Quadro 6: ARL para alterações na média considerando duas direções de perturbação com mesma variância.

DIREÇÕES: ROTAÇÃO E DESLOCAMENTO													
Fase I		ARL ₀											
Casos	σ_e	X ₁	X ₂	X ₃	X ₄	$\cup X_i$	T ²	T(x)	L(x)	$\cup d_i(x)$	CP ₁	CP ₂	UCP ₁
T ~ N(0,1) L ~ N(0,1)	0,1	384,62	330,58	363,64	341,88	157,48	92,17	392,16	366,97	191,39	377,36	366,97	186,05
	0,5	310,08	392,16	322,58	357,14	111,73	95,24	430,11	350,88	193,24	317,46	347,83	165,29
	1,0	412,37	373,83	377,36	421,05	106,10	94,12	439,56	430,11	218,58	400,00	412,37	203,05
Fase II		ARL											
T ~ N(1,1) L ~ N(0,1)	0,1	98,04	97,09	99,01	99,26	44,49	31,72	44,10	373,83	39,53	138,89	156,25	73,94
	0,5	101,27	114,29	107,24	108,40	36,07	32,47	46,84	400,00	41,93	90,50	113,96	50,89
	1,0	132,45	124,22	123,08	132,45	36,40	35,84	58,14	353,98	50,00	59,88	303,03	50,13
T ~ N(2,1) L ~ N(0,1)	0,1	18,16	17,86	18,16	17,77	8,55	6,83	6,30	380,95	6,21	6,53	285,71	6,40
	0,5	19,96	22,31	20,74	21,63	7,86	7,31	7,01	392,16	6,90	17,31	21,88	9,91
	1,0	30,65	30,72	31,27	30,21	9,49	9,27	9,04	408,16	8,87	9,46	251,57	9,15
T ~ N(1,1) L ~ N(1,1)	0,1	366,97	18,26	353,98	18,14	15,99	16,49	41,93	46,89	22,43	65,90	33,81	22,68
	0,5	384,62	20,68	373,83	20,25	13,90	16,74	46,57	46,46	23,46	366,97	18,82	17,91
	1,0	404,04	30,58	360,36	31,03	16,83	21,56	59,79	58,74	29,85	41,62	87,15	28,35
T ~ N(2,1) L ~ N(1,1)	0,1	94,56	5,40	92,38	5,39	4,87	5,13	6,29	46,89	5,65	8,36	23,65	6,50
	0,5	101,27	6,47	98,28	6,16	4,63	5,37	6,86	47,06	6,10	79,84	5,93	5,59
	1,0	129,87	9,70	141,34	9,68	5,81	7,02	8,91	52,84	7,76	7,26	120,12	6,90
T ~ N(2,1) L ~ N(2,1)	0,1	470,59	2,32	481,93	2,33	2,24	2,68	6,13	6,61	3,45	10,30	4,38	3,29
	0,5	327,87	2,77	373,83	2,73	2,28	2,83	6,97	7,15	3,81	350,88	2,53	2,52
	1,0	434,78	4,10	350,88	4,04	2,82	3,58	9,01	8,90	4,74	5,71	16,14	4,43

onde:

$L(x)$ = gráfico de controle para a variável transformada deslocamento $L(x)$;

CP_2 = gráfico de controle para a segunda componente principal.

$\bigcup d_i(\mathbf{x})$ = ARL conjunto dos gráficos nas variáveis transformadas, que nesse caso específico, equivalem a $T(x)$ e $L(x)$;

$\bigcup CP_i$ = ARL conjunto dos gráficos nas componentes principais, que nesse caso particular são: CP_1 e CP_2 .

Assim, da análise do quadro 6, chega-se às seguintes constatações:

- i. Considerando o processo em controle (Fase I), mais uma vez, observa-se que o ARL_0 está dentro do esperado (em torno de um alarme a cada 370 pontos para os gráficos de Shewhart e um alarme para cada 93 pontos para o gráfico T^2) para todas as representações, isto é, X_i 's, CP 's, variáveis transformadas $T(x)$ e $L(x)$, e para o gráfico de T^2 nas variáveis observadas;
- ii. Quando simulados alterações na média do processo em apenas uma das direções conhecidas T ou L (Fase II), verificou-se que para deslocamento de um desvio-padrão, o gráfico que teve maior rapidez de sinalização foi o de T^2 e para alterações de dois desvios-padrão o gráfico com melhor desempenho foi o de $T(x)$. Este último fato pode ser observado também pelo ARL conjunto nas variáveis transformadas, dado por $\bigcup d_i(\mathbf{x})$. No caso da ocorrência de alterações na média em ambas as direções (ou seja, rotação e deslocamento), verifica-se que o ARL conjunto dos gráficos nas variáveis observadas tem melhor desempenho;
- iii. Analisando os gráficos nas CP 's, verifica-se que, embora algumas vezes sinalizem com maior rapidez do que os gráficos nas variáveis transformadas, por não corresponderem às direções conhecidas (conforme observado na análise das componentes principais, registradas no quadro 5), na maior parte dos casos simulados, não indicam em que direção ocorreu a alteração, como pode ser verificado nas células destacadas em vermelho.

- iv. Examinando, para cada alteração simulada, as três linhas correspondentes a $\sigma_e = 0,1, 0,5$ e $1,0$, observa-se que o gráfico na CP que emite mais sinais varia: ora um, ora outro, ora ambos, *para uma mesma perturbação na média de T e L*. Isso porque, como já visto, as direções das CP's obtidas pela ACP neste caso (em que as variâncias das perturbações aleatórias nas direções conhecidas são iguais) são aleatórias, determinadas pelas flutuações aleatórias (erro experimental) dos dados usados na Fase I.
- v. Em alguns casos, os gráficos nos X_i 's têm menor ARL do que os gráficos nas CP's e os gráficos em $T(x)$ e $L(x)$. Contudo, a variável X_i que sinaliza varia conforme o caso considerado, de modo que para garantir proteção contra qualquer alteração seria necessário manter gráficos para as quatro variáveis, X_1, X_2, X_3 e X_4 . Isso implicaria maior sobrecarga operacional (comparativamente a manter apenas gráficos nas direções conhecidas ou nas CP's), visto que seria preciso uma maior quantidade de gráficos do que se fossem utilizados as CP's ou $T(x)$ e $L(x)$.

O Quadro 7 apresenta o número médio de amostras até a ocorrência de um sinal, considerando que os esquemas comparativos tenham aproximadamente o mesmo ARL_0 (200). Dessa forma, foram ajustados os ARL_0 's da $\bigcup X_i$, dos gráficos nas variáveis transformadas e nas CP's, com o ARL_0 do gráfico de T^2 de Hotelling.

Portanto, de acordo com os resultados, na comparação entre o desempenho do gráfico de T^2 e o desempenho conjunto dos gráficos nas variáveis observadas, nas variáveis transformadas e nas CP's com os respectivos ARL_0 's ajustados, observa-se que, quando ocorreu alteração na média em apenas uma direção conhecida, os gráficos nas variáveis transformadas foram mais eficazes e quando simuladas alterações simultâneas em duas direções de variação, os gráficos nas CP's tiveram melhor desempenho para o desvio-padrão do erro aleatório (σ_e) igual a 0,5 e 1,0, enquanto para $\sigma_e = 0,1$, o desempenho conjunto nas variáveis observadas obteve maior rapidez de detecção.

Quadro 7: ARL para alterações na média considerando duas direções de perturbação com mesma variância e ARL_0 's ajustados.

DIREÇÕES: ROTAÇÃO E DESLOCAMENTO					
Fase I		ARL_0			
Casos	σ_e	$\cup X_{i\alpha-aj}$	T^2	$\cup d_i(\mathbf{x})_{\alpha-aj}$	$\cup CP_{i\alpha-aj}$
T ~ N(0,1) L ~ N(0,1)	0,1	200			
	0,5				
	1,0				
Fase II		ARL			
T ~ N(1,1) L ~ N(0,1)	0,1	45,71	59,26	37,84	38,24
	0,5	54,79	61,63	48,90	50,19
	1,0	268,46	123,08	99,50	116,62
T ~ N(2,1) L ~ N(0,1)	0,1	9,48	10,52	6,20	6,25
	0,5	11,11	11,56	7,27	10,57
	1,0	39,53	16,28	11,71	19,19
T ~ N(1,1) L ~ N(1,1)	0,1	17,38	29,41	21,51	20,09
	0,5	19,98	29,78	24,04	19,36
	1,0	80,48	47,96	51,35	39,10
T ~ N(2,1) L ~ N(1,1)	0,1	5,09	7,49	5,52	6,16
	0,5	5,91	8,00	6,35	5,76
	1,0	17,18	11,00	10,98	9,73
T ~ N(2,1) L ~ N(2,1)	0,1	2,34	3,55	3,36	3,20
	0,5	2,67	3,83	3,82	2,57
	1,0	5,46	4,67	6,26	3,62

Por fim, no Quadro 8 são mostradas as correlações entre as variáveis observadas para o caso em que duas direções conhecidas atuam no problema. Dessa maneira, verifica-se que, nesse caso particular, as variáveis observadas deixaram de ser todas correlacionadas entre si, como ocorreu quando se tinha apenas uma direção de perturbação, porém, ainda há fortes correlações (no caso, negativas) entre x_1 e x_3 e entre x_2 e x_4 .

Quadro 8: Correlações entre as variáveis observadas para o processo em controle para o caso de alterações na média considerando duas direções de perturbação com mesma variância.

	$\sigma_e = 0.1$				$\sigma_e = 0.5$				$\sigma_e = 1.0$			
	x_1	x_2	x_3	x_4	x_1	x_2	x_3	x_4	x_1	x_2	x_3	x_4
x_1	1.00				1.00				1.00			
x_2	-0.001	1.00			0.001	1.00			-0.002	1.00		
x_3	-0.995	0.001	1.00		-0.889	-0.001	1.00		-0.664	0.003	1.00	
x_4	0.001	-0.995	0.001	1.00	-0.004	-0.888	0.002	1.00	0.003	-0.665	-0.001	1.00

4.1.3 Três direções de variação com iguais variâncias

Avaliou-se também o caso de três direções de perturbação com mesma variância, sob a ótica das CP's e das novas representações (Tx, Lx e Cx), conforme o seguinte modelo:

$$x_1 = 0,5(L - T + C) + e_1$$

$$x_2 = 0,5(L + T + C) + e_2$$

$$x_3 = 0,5(-L + T + C) + e_3$$

$$x_4 = 0,5(-L - T + C) + e_4$$

Quadro 9: Resultado da Análise de Componentes Principais para o caso de alterações na média considerando três direções de perturbação com mesma variância.

Variáveis Observadas	$\sigma_e = 0.1$			$\sigma_e = 0.5$			$\sigma_e = 1.0$		
	CP ₁	CP ₂	CP ₃	CP ₁	CP ₂	CP ₃	CP ₁	CP ₂	CP ₃
x_1	0.40	-0.63	0.45	-0.79	0.33	-0.13	-0.19	0.12	-0.86
x_2	-0.51	-0.70	0.04	-0.21	-0.10	-0.85	-0.24	-0.71	-0.35
x_3	-0.16	-0.29	-0.80	-0.01	-0.84	-0.18	-0.75	-0.29	0.37
x_4	0.75	-0.21	-0.39	-0.57	-0.41	0.49	-0.58	0.63	-0.06
Desvio-padrão	2.22	2.03	1.92	2.25	2.19	2.05	2.23	2.16	2.04
Proporção da variância Explicada	0.39	0.32	0.29	0.35	0.34	0.29	0.34	0.32	0.28

Novamente, considerando que o processo está sob controle e que as variâncias das direções consideradas são iguais, verifica-se pelo exame dos coeficientes das componentes que o método da ACP não conseguiu identificar corretamente as direções de variação. A interpretação desse fato é a mesma que com duas direções de variação; no caso, o elipsóide de isoprobabilidade degenera em uma esfera.

Analisando agora as correlações entre as variáveis observadas (Quadro 10), nota-se claramente uma redução no grau de correlação entre elas, especialmente quando o desvio-padrão do erro aleatório aumenta. No entanto, mesmo com essa redução, as correlações ainda permanecem significativas, o que pode ser facilmente comprovado por meio de um teste estatístico para o coeficiente de correlação.

Quadro 10: Correlações entre as variáveis observadas para o processo em controle para o caso de alterações na média considerando três direções de perturbação com mesma variância.

	$\sigma_e = 0.1$				$\sigma_e = 0.5$				$\sigma_e = 1.0$			
	X ₁	X ₂	X ₃	X ₄	X ₁	X ₂	X ₃	X ₄	X ₁	X ₂	X ₃	X ₄
X ₁	1.00				1.00				1.00			
X ₂	0.34	1.00			0.31	1.00			0.25	1.00		
X ₃	-0.33	0.33	1.00		-0.32	0.29	1.00		-0.25	0.25	1.00	
X ₄	0.33	-0.33	0.33	1.00	0.30	-0.31	0.31	1.00	0.25	-0.25	0.25	1.00

Quadro 11: ARL para o caso de alterações na média considerando três direções de perturbação com mesma variância.

DIREÇÕES: ROTAÇÃO, DESLOCAMENTO E CONTRAÇÃO															
Fase I		ARL ₀													
Casos	σ_e	X ₁	X ₂	X ₃	X ₄	$\cup X_i$	T ²	T(x)	L(x)	C(x)	$\cup d_i(x)$	CP ₁	CP ₂	CP ₃	UCP ₁
T ~ N(0,1) L ~ N(0,1) C ~ N(0,1)	0,1	388,35	380,95	444,44	360,36	101,78	98,77	392,16	377,36	408,16	131,58	366,97	357,14	459,77	129,87
	0,5	388,35	416,67	444,44	314,96	100,00	91,32	404,04	377,36	454,55	136,52	408,16	408,16	449,44	140,35
	1,0	430,11	370,37	416,67	380,95	101,52	95,01	344,83	350,88	388,35	120,48	357,14	366,97	439,56	128,62
Fase II		ARL													
T ~ N(1,1) L ~ N(0,1) C ~ N(0,1)	0,1	144,93	123,46	131,15	120,48	33,93	33,50	44,30	357,14	373,83	35,71	298,51	49,32	285,71	37,17
	0,5	133,33	148,15	153,26	125,79	36,63	34,01	49,02	344,83	333,33	36,83	91,12	136,05	216,22	43,72
	1,0	153,85	153,26	175,44	148,15	40,49	38,91	57,64	307,69	347,83	42,74	80,65	268,46	193,24	47,17
T ~ N(2,1) L ~ N(0,1) C ~ N(0,1)	0,1	30,53	30,03	30,40	28,63	8,51	7,03	6,22	347,83	357,14	6,04	162,60	7,56	145,99	6,98
	0,5	37,07	36,50	35,59	34,42	10,07	7,65	7,25	449,44	388,35	7,04	16,35	32,34	160,00	10,48
	1,0	43,76	45,40	43,91	42,42	11,76	9,35	8,81	430,11	307,69	8,44	14,31	141,34	77,97	11,31
T ~ N(1,1) L ~ N(1,1) C ~ N(0,1)	0,1	388,35	29,99	404,04	32,10	15,58	17,38	44,59	44,49	400,00	21,38	88,89	31,70	194,17	21,04
	0,5	408,16	33,90	392,16	32,98	16,56	17,44	49,57	48,25	404,04	23,41	353,98	24,42	136,05	19,71
	1,0	430,11	41,80	500,00	43,48	20,28	20,94	55,87	58,74	347,83	26,68	29,20	156,86	336,13	23,09
T ~ N(2,1) L ~ N(1,1) C ~ N(0,1)	0,1	133,33	9,88	133,33	9,82	5,33	5,22	6,36	44,79	408,16	5,62	186,05	5,54	105,26	5,19
	0,5	164,61	11,14	152,67	11,00	5,88	5,61	7,01	48,78	454,55	6,18	65,36	8,18	73,53	6,78
	1,0	149,81	14,78	158,10	14,60	7,36	6,93	8,70	59,17	350,88	7,54	7,31	273,97	141,34	6,85
T ~ N(2,1) L ~ N(2,1) C ~ N(2,1)	0,1	31,10	1,48	31,42	30,70	1,46	1,65	6,42	6,46	6,34	2,51	12,76	2,25	36,46	1,99
	0,5	36,13	1,60	34,78	34,75	1,57	1,77	7,07	7,08	6,88	2,70	373,83	1,64	103,36	1,63
	1,0	45,40	2,00	47,00	41,49	1,93	2,14	8,97	8,81	8,69	3,29	2,35	264,90	24,66	2,22

onde:

$C(x)$ = gráfico de controle para a variável transformada contração.

CP_3 = gráfico de controle para a terceira componente principal.

$\cup d_i(x)$ = ARL conjunto dos gráficos nas variáveis transformadas, que no caso, equivalem a T(x), L(x) e C(x);

$\cup PC_i$ = ARL conjunto dos gráficos nas componentes principais, as quais nesse caso específico são: CP_1 , CP_2 e CP_3 .

Analisando o Quadro 11, é possível obter as seguintes conclusões:

- i. Com relação à Fase I, os esquemas em análise têm ARL dentro do esperado.
- ii. Na Fase II, verificou-se que as CP's embora funcionem bem na detecção de pontos fora de controle, como evidenciado nos resultados dos ARL's para UPC_i , em que vários casos apresentam desempenham superior a $\bigcup d_i(\mathbf{x})$, não mostram a mesma eficácia na identificação das direções conhecidas, especialmente nos casos em que o desvio-padrão do erro aleatório cresce. Além disso, quando alterações ocorrem em duas ou três direções de variação, os alarmes tendem a concentrar-se em apenas uma componente, conforme pode ser observado nas últimas três linhas do quadro apresentado (células em vermelho).
- iii. Com relação ao desempenho dos esquemas considerados nos casos simulados, observou-se que, os gráficos nas variáveis transformadas tiveram maior eficácia, quando apenas uma direção conhecida sofreu alteração na média na ordem de dois desvios-padrão do valor alvo e, quando a alteração foi de apenas um desvio-padrão, o gráfico de T^2 desempenhou melhor. Nas simulações restantes, o desempenho conjunto em pelo menos um dos gráficos nas variáveis observadas foi mais eficaz (ver células em cinza) na maior parte dos casos. No entanto, para que possam ser realizadas comparações entre os esquemas propostos é interessante que os mesmos tenham aproximadamente o mesmo ARL_0 .

O Quadro 12 apresenta o resultado do desempenho dos esquemas supracitados, considerando aproximadamente o mesmo ARL_0 . Dessa forma, baseado nos resultados, pode-se constatar que mesmo ajustando os ARL_0 's conjuntos dos gráficos nas variáveis transformadas e dos gráficos nas CP's, o cenário obtido no Quadro 11 (sem ajuste do ARL_0) sofreu poucas mudanças, entre essas, destaque para os gráficos nas variáveis transformadas, que passaram a ser mais eficazes nos dois casos em que apenas uma direção conhecida sofreu alteração na média. Ressalta-se que esse mesmo fato ocorreu também no caso de duas direções de perturbação com mesma variância.

Quadro 12: ARL para o caso de alterações na média considerando três direções de perturbação com mesma variância e ARL_0 's ajustados.

DIREÇÕES: ROTAÇÃO, DESLOCAMENTO E CONTRAÇÃO					
Fase I		ARL_0			
Casos	σ_e	$\bigcup X_i$	T^2	$\bigcup d_i(\mathbf{x})_{\alpha-aj}$	$\bigcup CP_{i\alpha-aj}$
$T \sim N(0,1)$ $L \sim N(0,1)$ $C \sim N(0,1)$	0,1	200			
	0,5				
	1,0				
Fase II		ARL			
$T \sim N(1,1)$ $L \sim N(0,1)$ $C \sim N(0,1)$	0,1	59,17	53,19	50,70	55,40
	0,5	42,78	67,57	40,86	41,03
	1,0	84,21	75,61	66,67	74,63
$T \sim N(2,1)$ $L \sim N(0,1)$ $C \sim N(0,1)$	0,1	14,23	9,90	7,88	9,51
	0,5	9,89	8,48	7,45	10,40
	1,0	21,70	14,41	11,27	15,62
$T \sim N(1,1)$ $L \sim N(1,1)$ $C \sim N(0,1)$	0,1	26,54	28,63	31,10	31,55
	0,5	20,31	32,28	24,41	24,08
	1,0	39,92	38,02	40,24	36,70
$T \sim N(2,1)$ $L \sim N(1,1)$ $C \sim N(0,1)$	0,1	7,71	7,09	7,11	6,65
	0,5	7,35	8,46	6,85	8,07
	1,0	12,21	10,09	9,97	8,97
$T \sim N(2,1)$ $L \sim N(2,1)$ $C \sim N(2,1)$	0,1	1,66	1,92	3,02	2,32
	0,5	1,73	2,17	2,92	1,79
	1,0	2,43	2,64	4,18	2,61

4.1.4 Duas direções de perturbação, com diferentes variâncias

A seguir serão discutidos os resultados da simulação obtidos para duas direções de perturbação com diferentes variâncias, sumarizados nos quadros 13 a 16. O modelo é o mesmo considerado para duas direções conhecidas com variâncias iguais.

Dessa forma, considerando que o processo está sob controle, mas que as variâncias das direções conhecidas são diferentes, pode-se observar no Quadro 13, que nesse caso, as CP's são eficazes em identificar essas direções. O fato de que a direção com maior variância corresponde ao maior eixo do elipsóide de isoprobabilidade faz com que a ACP consiga separar essas direções, associando a primeira CP à direção de maior variância (no caso, T). Em consequência, pode-se pensar que quanto maior a diferença entre os desvios das direções, mais precisamente as CP's deverão coincidir com Tx e Lx.

Quadro 13: Resultado da Análise de Componentes Principais para o caso de alterações na média com duas direções de perturbação com diferentes variâncias.

Variáveis Observadas	$\sigma_e = 0.1$		$\sigma_e = 0.5$		$\sigma_e = 1.0$	
	CP ₁	CP ₂	CP ₁	CP ₂	CP ₁	CP ₂
x ₁	-0.51	-0.49	-0.50	-0.52	0.51	-0.48
x ₂	0.49	-0.51	0.51	-0.50	-0.50	-0.51
x ₃	0.51	0.49	0.49	0.49	-0.51	0.50
x ₄	-0.49	0.51	-0.50	0.49	0.48	0.51
Desvio-padrão	5.35	2.06	6.55	1.96	6.23	2.11
Proporção da Variância Explicada	0.87	0.129	0.91	0.08	0.86	0.098

Observando as correlações, apresentadas no quadro 14 a seguir, verifica-se que diferentemente do que ocorreu quando se considerou duas direções conhecidas com mesma variância, que apontou correlação apenas entre as variáveis x₁ e x₃, e as variáveis x₂ e x₄, aqui, nesse caso, todas as variáveis mostraram-se altamente correlacionadas.

Quadro 14: Correlações entre as variáveis observadas para o processo em controle para o caso de alterações na média considerando duas direções de perturbação com diferentes variâncias.

	$\sigma_e = 0.1$				$\sigma_e = 0.5$				$\sigma_e = 1.0$			
	X ₁	X ₂	X ₃	X ₄	X ₁	X ₂	X ₃	X ₄	X ₁	X ₂	X ₃	X ₄
X ₁	1.00				1.00				1.00			
X ₂	-0.799	1.00			-0.781	1.00			-0.723	1.00		
X ₃	-0.999	0.800	1.00		-0.976	0.780	1.00		-0.910	0.727	1.00	
X ₄	0.800	-0.999	-0.800	1.00	0.781	-0.976	-0.781	1.00	0.728	-0.909	-0.727	1.00

Avaliando os resultados no Quadro 15, tem-se que:

- i. Todos os resultados obtidos para a Fase I estão dentro do esperado, pelos mesmos motivos já mencionados na análise dos casos anteriores.
- ii. Com relação ao desempenho na detecção dos pontos fora de controle (Fase II) entre o gráfico nas CP's e o gráfico nas variáveis transformadas, os resultados são semelhantes. A coincidência de desempenho dos dois esquemas de controle

pode ser constatada simplesmente comparando o ARL dos gráficos nas CP's e dos gráficos nas variáveis transformadas. Entretanto, esse resultado já era esperado, dada a coincidência entre as CP's e as variáveis transformadas. Contudo, percebem-se algumas pequenas diferenças, por exemplo, quando o desvio-padrão do erro-aleatório é muito pequeno ($\sigma_e = 0,1$), as CP's são mais eficazes em sinalizar. Entretanto, à medida que o desvio-padrão do erro-aleatório aumenta o desempenho do gráfico nas variáveis transformadas torna-se superior ou igual ao gráfico nas CP's.

- iii. Para esse caso, o gráfico de T^2 foi o esquema mais eficaz, com desempenho superior aos demais em praticamente todos os casos simulados, conforme pode ser observado nas células em cinza. No entanto, para verificar a eficácia do referido gráfico serão obtidos novos ARL's para os outros esquemas propostos considerando um ARL_0 igual a 200 (ver Quadro 16).

Quadro 15: ARL para o caso de alterações na média considerando duas direções de perturbação com o desvio-padrão da Rotação maior que do Deslocamento.

DIREÇÕES: ROTAÇÃO E DESLOCAMENTO													
Fase I		ARL ₀											
Casos	σ_e	X_1	X_2	X_3	X_4	$\bigcup X_i$	T^2	$T(x)$	$L(x)$	$\bigcup d_i(x)$	CP_1	CP_2	UCP_i
$T \sim N(0,3)$ $L \sim N(0,1)$	0,1	396,04	408,16	412,37	396,04	209,42	94,12	412,37	370,37	195,12	412,37	370,37	195,12
	0,5	392,16	380,95	353,98	363,64	168,78	90,09	370,37	353,98	181,82	370,37	353,98	181,82
	1,0	373,83	350,88	373,83	350,88	134,23	93,46	353,98	347,83	175,44	353,98	347,83	175,44
Fase II		ARL											
$T \sim N(1,3)$ $L \sim N(0,1)$	0,1	275,86	246,91	272,11	245,40	147,06	82,30	256,41	434,78	161,29	168,07	425,53	120,85
	0,5	266,67	242,42	270,27	246,91	120,48	85,29	240,96	350,88	142,86	240,96	353,98	143,37
	1,0	253,16	243,90	259,74	229,89	94,34	82,82	231,21	357,14	140,35	232,56	357,14	140,85
$T \sim N(2,3)$ $L \sim N(0,1)$	0,1	103,63	105,26	108,70	106,95	62,50	52,63	98,52	430,11	80,16	70,80	430,11	60,79
	0,5	123,46	140,85	122,70	120,48	61,16	60,51	116,28	357,14	87,72	116,28	357,14	87,72
	1,0	108,99	119,05	104,44	112,68	47,17	56,66	93,24	307,69	72,20	93,90	307,69	72,60
$T \sim N(1,3)$ $L \sim N(1,1)$	0,1	300,75	100,25	314,96	101,78	80,16	29,87	198,02	46,62	37,88	136,99	46,40	34,90
	0,5	444,44	111,42	425,53	107,82	77,22	32,13	242,42	46,57	39,14	242,42	46,89	39,37
	1,0	320,00	124,22	333,33	125,00	65,68	33,87	231,21	59,35	47,34	231,21	59,35	47,34
$T \sim N(2,3)$ $L \sim N(1,1)$	0,1	232,56	50,51	232,56	51,28	44,84	24,48	104,99	47,23	32,95	70,55	46,95	28,57
	0,5	287,77	54,05	281,69	53,62	40,36	24,43	115,27	46,95	33,70	115,27	47,06	33,76
	1,0	263,16	55,56	239,52	57,31	35,97	27,66	98,77	53,84	35,12	99,50	53,84	35,21
$T \sim N(2,3)$ $L \sim N(2,1)$	0,1	396,04	23,23	412,37	23,07	21,63	6,07	93,02	6,41	6,05	65,25	6,39	5,89
	0,5	404,04	26,67	459,77	26,32	21,62	6,48	103,63	6,95	6,58	103,63	6,99	6,61
	1,0	317,46	27,78	350,88	27,10	18,94	7,91	97,80	8,96	8,29	98,04	8,96	8,29

Ajustados os ARL_0 's aproximados para os gráficos propostos, verifica-se que, quando simuladas alterações na média em somente uma das direções conhecidas, o ARL conjunto dos gráficos nas variáveis observadas obteve melhor desempenho. Contudo, no caso em que a alteração na média ocorreu em apenas uma direção de variação, o número médio de amostras até a sinalização ficou acima de 90 amostras para todos os esquemas, desempenho considerado fraco. Ainda, quando alterações na média ocorreram nas duas direções consideradas na simulação, os gráficos nas variáveis transformadas e os gráficos nas CP's sinalizaram com maior rapidez.

Quadro 16: ARL para o caso de alterações na média considerando duas direções conhecidas com diferentes variâncias e ARL_0 's ajustados.

DIREÇÕES: ROTAÇÃO E DESLOCAMENTO					
Fase I		ARL_0			
Casos	σ_e	$\bigcup X_{i\alpha-aj}$	T^2	$\bigcup d_i(x)_{\alpha-aj}$	$UCP_{i\alpha-aj}$
$T \sim N(0,3)$ $L \sim N(0,1)$	0,1	200			
	0,5				
	1,0				
Fase II		ARL			
$T \sim N(1,3)$ $L \sim N(0,1)$	0,1	132,01	161,94	148,15	148,15
	0,5	90,09	198,02	125,79	125,79
	1,0	139,37	153,85	161,94	161,94
$T \sim N(2,3)$ $L \sim N(0,1)$	0,1	67,57	112,68	87,15	87,15
	0,5	43,48	129,45	71,43	71,43
	1,0	68,38	95,69	95,69	95,24
$T \sim N(1,3)$ $L \sim N(1,1)$	0,1	87,91	54,05	40,49	40,49
	0,5	66,33	75,47	33,50	33,50
	1,0	95,69	58,91	50,76	50,38
$T \sim N(2,3)$ $L \sim N(1,1)$	0,1	43,15	40,53	32,08	32,08
	0,5	36,46	57,97	28,45	28,45
	1,0	50,96	45,66	42,37	42,33
$T \sim N(2,3)$ $L \sim N(2,1)$	0,1	22,26	9,25	6,30	6,30
	0,5	20,78	15,20	6,37	6,37
	1,0	22,98	17,09	7,38	7,38

4.2 Casos de aumentos na variância

Para as análises do desempenho dos gráficos: nas variáveis observadas; nas variáveis transformadas; nas CP's e de T^2 , no caso de aumentos na variância das direções conhecidas, serão considerados os mesmos casos explorados para alterações na média. Seguindo procedimento semelhante ao que foi realizado para alteração na média, será analisada a eficácia dos gráficos sugeridos nessa tese, inicialmente considerando

que o processo está sob controle e em seguida considerando que o processo sofreu aumentos na variância.

4.2.1 Uma única direção de perturbação

Quadro 17: ARL para o caso de aumentos na variância considerando uma única direção de perturbação.

DIREÇÃO: ROTAÇÃO									
Fase I		ARL ₀							
Casos	σ_e	X_1	X_2	X_3	X_4	$\cup X_i$	T^2	$T(x)$	$PC1$
$T \sim N(0,1)$	0,1	434,78	400,00	396,04	384,62	261,44	87,91	396,04	396,04
	0,5	400,00	377,36	373,83	333,33	142,86	96,85	388,35	388,35
	1,0	430,11	377,36	392,16	412,37	101,52	91,12	392,16	392,16
Fase II		ARL							
$T \sim N(0,1.5)$	0,1	21,80	21,59	21,74	21,33	19,11	20,88	21,27	21,18
	0,5	32,00	32,39	30,82	30,77	14,71	23,15	25,79	25,94
	1,0	56,02	53,48	58,91	60,88	18,86	26,32	31,01	30,98
$T \sim N(0,2)$	0,1	7,49	7,41	7,42	7,46	6,86	7,94	7,40	7,39
	0,5	9,95	10,06	9,96	9,89	6,09	8,59	8,21	8,26
	1,0	17,62	16,81	18,11	18,16	7,49	10,14	9,73	9,72

Do quadro 17, pode-se constatar que:

- Considerando que o processo está sob controle, os resultados para a Fase I do processo estão de acordo com o esperado. É interessante observar também, que, os valores do ARL₀ foram repetidos do Quadro 3, para o caso de alterações na média, pois, dado que o processo encontra-se em controle, é possível utilizar os seus limites de controle para o monitoramento, tanto da média, como da variância.
- Na Fase II, observou-se que os resultados são semelhantes aos obtidos para o monitoramento de alterações na média, ou seja, os gráficos na variável transformada $T(x)$ e na CP_1 sinalizam praticamente os mesmos pontos fora de controle, isso porque, como explicado no item 4.1, a CP_1 praticamente coincide com $T(x)$. Além disso, como neste primeiro caso analisado, as correlações entre as variáveis observadas são todas altas (ver Quadro 2), os gráficos nas variáveis observadas tendem a sinalizar tão bem quanto os gráficos na CP e em $T(x)$, como observado através do ARL conjunto dos gráficos nas variáveis observadas, dado por $\cup X_i$. Porém, para que se possa fazer alguma comparação entre os

esquemas de controle em estudo é necessário que os gráficos tenham aproximadamente o mesmo ARL_0 .

Novamente, para efeito de comparação, o Quadro 18 apresenta os novos ARL 's conjuntos dos gráficos nas variáveis observadas ($\bigcup X_i$), dos gráficos nas variáveis transformadas ($\bigcup d_i(\mathbf{x})$) e dos gráficos nas CP 's ($\bigcup CP_i$), considerando que os referidos gráficos tiveram o risco de alarme falso ajustado para produzir aproximadamente o mesmo ARL_0 .

Quadro 18: ARL para o caso de aumentos na variância considerando uma única direção conhecida considerando e ARL_0 's ajustados.

DIREÇÃO: ROTAÇÃO					
Fase I		ARL_0			
Casos	σ_e	$\bigcup X_{i\alpha-aj}$	T^2	$T(x)_{\alpha-aj}$	$CP_{1\alpha-aj}$
$T \sim N(0,1)$	0,1	200			
	0,5				
	1,0				
Fase II		ARL			
$T \sim N(0,1.5)$	0,1	17,52	31,65	17,17	17,17
	0,5	18,64	34,60	17,01	16,98
	1,0	31,70	42,74	22,60	22,61
$T \sim N(0,2)$	0,1	6,44	10,89	6,38	6,38
	0,5	6,93	11,18	6,47	6,46
	1,0	10,27	13,29	7,89	7,90

Neste quadro, observou-se que, quando considerado aproximadamente o mesmo ARL_0 , o gráficos nas variáveis transformadas e o gráfico nas CP 's foram mais eficazes nos dois casos simulados, resultado similar ao obtido para o caso de alterações apenas na média.

4.2.2 Duas direções de perturbação com mesma variância

Analisando o aumento da variância para o caso em que duas direções conhecidas (Rotação e Deslocamento) atuem no problema, verificou-se por meio do quadro 19 que:

- Com relação ao desempenho dos gráficos em sinalizar pontos fora de controle (Fase II), observa-se que houve uma alternância na sinalização dos alarmes entre os esquemas considerados neste trabalho. No entanto, será realizada uma análise

mais apurada considerando que ambos os esquemas tenham ARL_0 's aproximados (ver Quadro 20), para que se possa verificar a real eficácia dos gráficos.

- ii. Diferentemente do que ocorreu quando se analisou alterações na média, as CP's só não identificaram as direções conhecidas quando a magnitude do aumento na variância em uma direção foi diferente do aumento da variância na outra direção e particularmente, no caso em que $\sigma_e = 0,5$ (células do quadro em vermelho). Além disso, quando a magnitude do aumento na variância foi a mesma para ambas as direções de variação consideradas, observou-se que as CP's conseguiram identificar as direções conhecidas T e L para todos os desvios-padrão do erro aleatório.

Quadro 19: ARL para o caso de aumentos na variância considerando duas direções de perturbação com mesma variância.

DIREÇÕES: ROTAÇÃO E DESLOCAMENTO													
Fase I		ARL ₀											
Casos	σ_e	X_1	X_2	X_3	X_4	$\cup X_i$	T^2	$T(X)$	$L(X)$	$\cup d_i(x)$	CP_1	CP_2	$\cup CP_i$
T ~ N(0,1) L ~ N(0,1)	0,1	384,6	330,6	363,6	341,9	157,5	92,17	392,2	367,0	191,4	377,4	367,0	186,0
	0,5	310,1	392,2	322,6	357,1	111,7	95,24	430,1	350,9	193,2	317,5	347,8	165,3
	1,0	412,4	373,8	377,4	421,1	106,1	94,12	439,6	430,1	218,6	400,0	412,4	203,0
Fase II		ARL											
T ~ N(0,1.5) L ~ N(0,1)	0,1	53,3	52,9	53,8	52,7	25,4	19,84	21,6	381,0	20,5	22,1	330,6	20,7
	0,5	60,0	60,6	61,9	57,8	22,7	21,30	24,5	341,9	22,9	53,9	61,9	30,0
	1,0	85,3	80,8	82,8	83,7	25,4	25,66	28,1	373,8	26,2	30,1	330,6	27,6
T ~ N(0,2) L ~ N(0,1)	0,1	17,0	16,9	17,0	16,9	9,3	7,97	7,3	408,2	7,2	7,5	254,8	7,3
	0,5	19,6	20,6	19,5	20,0	9,0	8,19	7,9	367,0	7,7	16,8	20,2	10,5
	1,0	29,7	29,5	29,6	28,9	10,6	9,98	9,5	425,5	9,3	9,9	315,0	9,7
T ~ N(0,1.5) L ~ N(0,1.5)	0,1	22,7	22,8	22,7	23,4	11,0	10,17	22,2	23,2	11,7	22,3	23,1	11,6
	0,5	24,8	26,4	25,1	25,3	9,9	11,12	23,8	23,6	12,1	23,6	23,4	12,0
	1,0	40,3	37,0	37,0	36,8	12,2	13,55	29,2	30,9	15,3	30,1	31,3	15,6
T ~ N(0,2) L ~ N(0,1.5)	0,1	11,4	11,3	11,3	11,4	5,8	5,60	7,4	23,1	5,7	7,4	22,4	5,8
	0,5	12,5	13,1	12,7	12,8	5,6	5,93	7,9	24,8	6,2	11,4	12,5	6,4
	1,0	19,1	19,2	18,6	19,4	6,9	7,24	9,8	28,3	7,5	10,0	27,2	7,5

O Quadro 20 apresenta os novos ARL's, considerando que na Fase I do processo, os gráficos em estudo tenham aproximadamente o mesmo número médio de amostras até a ocorrência de um alarme falso (ARL_0).

Da análise do presente quadro, verificou-se que na maior parte dos casos analisados os gráficos nas variáveis transformadas obtiveram um ARL um pouco menor do que os outros esquemas, e, portanto, maior eficácia na detecção de pontos fora de controle.

Quadro 20: ARL para o caso de aumentos na variância considerando duas direções de perturbação com mesma variância e ARL_0 's ajustados.

DIREÇÕES: ROTAÇÃO E DESLOCAMENTO					
Fase I		ARL_0			
Casos	σ_e	$\bigcup X_{i\alpha-aj}$	T^2	$\bigcup d_i(\mathbf{x})_{\alpha-aj}$	$\bigcup CP_{i\alpha-aj}$
T ~ N(0,1) L ~ N(0,1)	0,1	200			
	0,5				
	1,0				
Fase II		ARL			
T ~ N(0,1,5) L ~ N(0,1)	0,1	27,87	32,39	20,83	21,20
	0,5	33,96	35,18	24,17	32,31
	1,0	121,58	52,02	40,53	60,51
T ~ N(0,2) L ~ N(0,1)	0,1	9,95	10,40	7,38	7,40
	0,5	11,78	11,24	8,33	11,43
	1,0	27,62	13,68	11,00	17,04
T ~ N(0,1,5) L ~ N(0,1,5)	0,1	11,35	15,36	10,96	10,94
	0,5	13,19	16,49	12,72	12,63
	1,0	36,53	22,27	20,83	21,39
T ~ N(0,2) L ~ N(0,1,5)	0,1	7,49	9,43	7,23	7,22
	0,5	6,88	7,86	6,26	6,63
	1,0	14,13	9,47	8,89	9,49

4.2.3 Três direções de perturbação

O Quadro 21 a seguir apresenta os ARL 's obtidos para aumentos no desvio-padrão em uma, duas e três direções de variação, para o caso em que o processo possui (mesmo sob controle) três direções de perturbação (Rotação, Deslocamento e Contração/Dilatação).

Da análise desse quadro, é possível observar:

- Com relação à Fase II, visto que o ARL_0 entre os gráficos nas variáveis transformadas e os gráficos nas CP's são praticamente iguais, pode-se verificar que em quase todos os casos simulados o esquema que utiliza as variáveis transformadas é mais eficaz do o esquema utilizando as CP's. Este fato é evidenciado mais uma vez observando o ARL individual de cada variável transformada e de cada CP, como também por meio do ARL conjunto ($\bigcup d_i(\mathbf{x})$, $\bigcup CP_i$)

- ii. Analogamente ao que ocorreu no caso anterior (de duas direções de perturbação), as CP's só não identificaram as direções conhecidas quando o aumento na variância ocorreu em apenas uma direção e quando $\sigma_e = 0,5$ (células do quadro em vermelho). Além disso, quando a magnitude do aumento na variância foi a mesma para ambas as direções de variação consideradas, observou-se que as CP's conseguiram identificar as direções conhecidas T, L e C, sendo o desempenho de seu gráfico tão bom quanto os gráficos nas variáveis transformadas (ver resultado do ARL no último caso simulado), para todos os desvios-padrão do erro aleatório considerados.
- iii. O gráfico T^2 apresentou bom desempenho quando os aumentos no desvio-padrão ocorreram simultaneamente em duas ou três direções de variação.

Quadro 21: ARL para o caso de aumentos na variância considerando três direções de perturbação com mesma variância.

DIREÇÕES: ROTAÇÃO, DESLOCAMENTO E CONTRAÇÃO															
Fase I		ARL ₀													
Casos	σ_e	X ₁	X ₂	X ₃	X ₄	$\bigcup X_i$	T ²	T(x)	L(x)	C(x)	$\bigcup d_i(x)$	CP ₁	CP ₂	CP ₃	UCP ₁
T ~ N(0,1)	0,1	388,3	381,0	444,4	360,4	101,8	98,77	392,2	377,4	408,2	131,6	367,0	357,1	459,8	129,9
L ~ N(0,1)	0,5	388,3	416,7	444,4	315,0	100,0	91,32	404,0	377,4	454,5	136,5	408,2	408,2	449,4	140,4
C ~ N(0,1)	1,0	430,1	370,4	416,7	381,0	101,5	95,01	344,8	350,9	388,3	120,5	357,1	367,0	439,6	128,6
Fase II		ARL													
T ~ (0,1.5)	0,1	82,6	83,3	82,0	84,7	23,4	19,74	21,0	373,8	449,4	19,1	248,4	25,0	277,8	21,2
L ~ N(0,1)	0,5	97,8	94,3	96,4	94,1	26,6	22,00	24,4	512,8	350,9	22,0	51,3	87,1	275,9	29,8
C ~ N(0,1)	1,0	111,1	110,8	123,1	103,4	29,5	26,37	31,1	330,6	330,6	26,2	45,9	261,4	170,9	32,3
T ~ N(0,2)	0,1	31,2	29,8	29,6	29,0	10,1	7,98	7,5	363,6	384,6	7,2	204,1	8,6	196,1	8,1
L ~ N(0,1)	0,5	34,4	31,3	33,7	31,9	10,7	8,48	7,7	315,0	339,0	7,4	15,3	27,8	163,9	10,4
C ~ N(0,1)	1,0	43,9	43,3	46,8	40,7	13,3	10,14	9,7	388,3	404,0	9,2	14,2	173,9	80,3	11,9
T ~ N(0,2)	0,1	19,1	19,6	19,4	18,8	6,3	5,76	7,6	22,4	336,1	5,8	20,4	8,7	144,9	6,2
L ~ (0,1.5)	0,5	21,3	20,7	21,9	20,4	6,8	5,88	8,0	24,4	370,4	6,1	11,3	16,1	126,6	6,8
C ~ N(0,1)	1,0	26,7	26,0	27,6	27,0	8,2	6,98	9,7	28,6	339,0	7,3	12,6	31,8	65,8	8,4
T ~ (0,1.5)	0,1	39,3	37,1	38,7	38,4	11,4	10,36	22,2	22,4	439,6	11,2	23,6	24,6	258,1	11,8
L ~ (0,1.5)	0,5	41,6	41,7	42,5	42,1	12,2	11,11	24,3	24,4	439,6	12,2	24,6	30,4	172,4	12,9
C ~ N(0,1)	1,0	53,8	56,3	59,3	48,9	15,1	13,36	29,7	28,3	325,2	14,1	33,6	36,5	134,2	15,9
T ~ N(0,2)	0,1	7,7	7,5	7,6	7,5	2,6	2,60	7,5	7,7	7,6	2,9	7,7	7,5	7,6	2,9
L ~ N(0,2)	0,5	8,2	8,4	8,3	8,2	2,8	2,66	8,3	8,1	8,1	3,1	8,2	8,1	8,2	3,1
C ~ N(0,2)	1,0	10,2	10,4	10,4	9,9	3,2	3,16	9,3	9,8	9,4	3,5	9,5	9,7	9,4	3,5

Considerando que o gráfico T^2 registrou melhor desempenho na maior parte dos casos simulados, mas que esse gráfico possui um ARL₀ menor do que os outros

esquemas sugeridos neste trabalho, houve necessidade de ajustar os valores do risco de alarme falso para esses esquemas.

Quadro 22: ARL para o caso de aumentos na variância considerando três direções de perturbação com mesma variância e ARL_0 's ajustados.

DIREÇÕES: ROTAÇÃO E DESLOCAMENTO					
Fase I		ARL_0			
Casos	σ_e	$\cup X_{i\alpha-aj}$	T^2	$\cup d_i(x)_{\alpha-aj}$	$\cup CP_{i\alpha-aj}$
$T \sim N(0,1)$ $L \sim N(0,1)$ $C \sim N(0,1)$	0,1	200			
	0,5				
	1,0				
Fase II		ARL			
$T \sim N(0,1.5)$ $L \sim N(0,1)$ $C \sim N(0,1)$	0,1	39,64	29,05	25,00	27,97
	0,5	23,42	38,20	22,08	27,99
	1,0	55,33	40,69	35,03	45,20
$T \sim N(0,2)$ $L \sim N(0,1)$ $C \sim N(0,1)$	0,1	15,00	10,21	8,54	9,97
	0,5	11,14	11,86	8,42	11,42
	1,0	20,79	13,53	11,01	14,89
$T \sim N(0,2)$ $L \sim N(0,1.5)$ $C \sim N(0,1)$	0,1	8,39	6,95	6,62	7,19
	0,5	7,31	8,22	6,77	7,38
	1,0	12,39	9,44	8,65	10,53
$T \sim N(0,1.5)$ $L \sim N(0,1.5)$ $C \sim N(0,1)$	0,1	16,79	14,63	13,95	15,01
	0,5	14,81	17,96	13,34	14,03
	1,0	26,02	20,50	18,96	22,74
$T \sim N(0,2)$ $L \sim N(0,2)$ $C \sim N(0,2)$	0,1	3,11	2,99	3,24	3,29
	0,5	3,04	3,29	3,32	3,35
	1,0	4,30	3,90	4,20	4,23

O Quadro 22 traz os novos valores do ARL, considerando que os esquemas propostos tenham os mesmos ARL_0 's. Dessa forma, constatou-se que o esquema que apresenta maior eficácia na detecção de pontos fora de controle, é o que considera as variáveis transformadas, o qual obteve menor ARL conjunto em praticamente todos os casos, com exceção do último caso, que simula aumentos de 100% do desvio-padrão nas três direções conhecidas de variação, nesta situação específica, os esquemas que desempenharam melhor foram o de T^2 de Hotelling e o desempenho conjunto dos gráficos nas variáveis observadas.

4.2.4 Duas direções de perturbação, com diferentes variâncias

Por fim, será analisada a última situação que leva em consideração aumentos na variância.

Quadro 23: ARL para o caso de aumentos na variância considerando duas direções de perturbação com o desvio-padrão da Rotação maior que do Deslocamento.

DIREÇÕES: ROTAÇÃO E DESLOCAMENTO													
Fase I		ARL ₀											
Casos	σ_e	X_1	X_2	X_3	X_4	$\cup X_i$	T^2	$T(x)$	$L(x)$	$\cup d_i(x)$	CP_1	CP_2	$\cup CP_i$
$T \sim N(0,3)$ $L \sim N(0,1)$	0,1	396,04	408,16	412,37	396,04	209,42	94,12	412,37	370,37	195,12	412,37	370,37	195,12
	0,5	392,16	380,95	353,98	363,64	168,78	90,09	370,37	353,98	181,82	370,37	353,98	181,82
	1,0	373,83	350,88	373,83	350,88	134,23	93,46	353,98	347,83	175,44	353,98	347,83	175,44
Fase II		ARL											
$T \sim N(0,3)$ $L \sim (0,1.5)$	0,1	95,9	114,9	98,8	114,9	54,6	20,4	191,4	16,5	15,3	191,4	16,5	15,3
	0,5	116,6	130,7	111,1	125,0	53,0	21,4	170,2	15,7	14,5	170,2	15,8	14,5
	1,0	145,5	161,3	152,1	149,8	55,4	26,7	191,4	20,4	18,5	191,4	20,3	18,4
$T \sim N(0,3)$ $L \sim N(0,2)$	0,1	59,6	68,6	58,1	69,1	31,7	8,1	195,1	6,3	6,2	195,1	6,3	6,1
	0,5	74,2	71,0	73,3	72,7	31,2	8,5	173,9	6,2	6,1	173,9	6,3	6,1
	1,0	102,0	100,0	95,2	99,8	36,3	10,2	216,2	7,8	7,5	216,2	7,7	7,5
$T \sim N(0,4.5)$ $L \sim N(0,1.5)$	0,1	14,5	15,3	14,4	15,3	9,4	10,1	16,2	16,5	8,4	16,2	16,5	8,4
	0,5	16,5	16,5	16,4	16,2	9,5	10,9	15,3	15,9	8,1	15,3	16,0	8,1
	1,0	21,4	21,1	20,7	20,9	10,5	12,1	17,4	20,2	9,6	17,4	20,1	9,6
$T \sim N(0,4.5)$ $L \sim N(0,2)$	0,1	12,4	13,3	12,5	13,3	7,6	5,5	16,4	6,3	4,8	16,4	6,3	4,7
	0,5	14,1	13,9	14,0	13,8	7,5	5,8	15,4	6,4	4,7	15,4	6,4	4,7
	1,0	17,9	17,6	17,8	17,4	8,4	6,7	16,7	7,8	5,6	16,7	7,8	5,6

Dessa forma, com base nos resultados mostrados no Quadro 23, chegam-se as seguintes constatações:

- Para a Fase I do processo, todos os gráficos considerados neste caso estão de acordo com o que prescreve a teoria, isto é, ARL de aproximadamente 370 amostras para o gráfico de Shewhart e 93 amostras para o gráfico de T^2 de Hotelling.
- Conforme visto anteriormente, para o caso em que os desvios-padrão das direções conhecidas são diferentes, as CP 's tendem a coincidir com as variáveis transformadas (ver Quadro 13). Portanto, é esperado que os gráficos tenham aproximadamente o mesmo ARL.

- iii. Diferentemente do que ocorreu quando analisado alterações na média, no caso de aumentos na variância, os gráficos nas CP's e nas variáveis transformadas são mais eficazes em sinalizar em todos os casos simulados, independente se o aumento ocorreu em uma ou nas duas direções conhecidas. Observe que, para este caso específico, os gráficos na CP's e nas variáveis transformadas desempenham melhor (menor ARL), mesmo tendo em quase todos os casos considerados maior ARL_0 , ou seja, maior número de amostras até a ocorrência de um alarme falso, tornando desnecessário o ajuste do ARL_0 para comparações entre os esquemas.

4.3 Casos de alterações na média e aumentos na variância

A seguir será analisado o desempenho dos gráficos propostos nesse trabalho, para o caso que ocorram simultaneamente aumentos na média e na variância.

4.3.1 Uma única direção de perturbação

Com base nos resultados do ARL apresentado no Quadro 24, observa-se que tanto para o processo em controle, como para o processo em descontrole, os resultados são semelhantes ao verificado nas análises obtidas para os casos separados de apenas alterações na média e de somente aumentos na variância, ou seja, os ARL's conjuntos dos gráficos nas variáveis originais aparentemente sinalizam com maior rapidez. Mais uma vez, para que se possa concluir, qual o esquema tem melhor desempenho, é necessário compará-los sob aproximadamente os mesmos ARL_0 's.

Dos resultados do Quadro 25 a seguir, observa-se que os gráficos nas variáveis transformadas e os gráficos nas CP's têm maior rapidez na detecção de alarmes.

Quadro 24: ARL para o caso de alterações na média e aumentos na variância considerando uma única direção de perturbação.

DIREÇÃO: ROTAÇÃO									
Fase I		ARL ₀							
Casos	σ_e	X_1	X_2	X_3	X_4	$\cup X_i$	T^2	$T(x)$	PC1
$T \sim N(0,1)$	0,1	434,78	400,00	396,04	384,62	261,44	87,91	396,04	396,04
	0,5	400,00	377,36	373,83	333,33	142,86	96,85	388,35	388,35
	1,0	430,11	377,36	392,16	412,37	101,52	91,12	392,16	392,16
Fase II		ARL							
$T \sim N(1,1.5)$	0,1	11,00	10,91	10,83	10,86	6,70	10,79	10,84	10,79
	0,5	14,93	15,04	14,67	14,63	6,88	12,25	11,94	11,97
	1,0	25,99	25,11	26,25	26,13	10,14	14,30	13,40	13,40
$T \sim N(1,2)$	0,1	5,56	5,50	5,52	5,52	4,04	5,91	5,49	5,48
	0,5	7,19	7,23	7,18	7,10	6,22	6,30	6,06	6,07
	1,0	12,44	11,81	12,25	12,71	6,66	7,42	7,00	6,99
$T \sim N(2,1.5)$	0,1	4,07	4,01	4,02	4,03	2,89	4,29	4,03	3,99
	0,5	5,12	5,18	5,16	5,11	3,06	4,71	4,29	4,30
	1,0	9,51	9,25	9,57	9,66	4,21	5,66	5,18	5,18
$T \sim N(2,2)$	0,1	3,20	3,18	3,19	3,19	2,53	3,43	3,17	3,17
	0,5	3,89	3,90	3,88	3,84	2,65	3,62	3,38	3,38
	1,0	6,42	6,31	6,47	6,48	3,43	4,21	3,92	3,92

Quadro 25: ARL para o caso de alterações na média e aumentos na variância considerando uma única direção de perturbação e ARL₀'s ajustados.

DIREÇÃO: ROTAÇÃO					
Fase I		ARL ₀			
Casos	σ_e	$\cup X_{i\alpha-aj}$	T^2	$T(x)_{\alpha-aj}$	$CP_{1\alpha-aj}$
$T \sim N(0,1)$	0,1	200			
	0,5				
	1,0				
Fase II		ARL			
$T \sim N(1,1.5)$	0,1	8,75	16,31	8,60	8,60
	0,5	9,64	16,92	8,89	8,87
	1,0	15,00	20,36	10,98	11,01
$T \sim N(1,2)$	0,1	4,87	7,67	4,82	4,82
	0,5	5,27	8,01	4,92	4,91
	1,0	7,58	9,55	5,95	5,95
$T \sim N(2,1.5)$	0,1	3,56	5,61	3,53	3,53
	0,5	3,84	6,00	3,58	3,57
	1,0	5,61	7,33	4,31	4,32
$T \sim N(2,2)$	0,1	2,95	4,12	2,93	2,93
	0,5	3,10	4,28	2,95	2,95
	1,0	4,18	5,03	3,45	3,44

4.3.2 Duas direções de perturbação com mesma variância

No caso em que duas direções de perturbação atuam no processo, o quadro abaixo apresenta o desempenho dos esquemas propostos neste trabalho.

Da análise do Quadro 26, constata-se que:

- i. Considerando o processo em controle (Fase I), tal como foi observado nos itens anteriormente analisados, constata-se que os resultados estão dentro do esperado;
- ii. Com relação à fase de monitoramento do processo (Fase II), no caso em que apenas uma direção de variação sofreu a alteração simultânea na média e na variância, os gráficos nas variáveis transformadas e o gráfico de T^2 foram mais eficazes. Porém, quando alterações na média e aumentos na variância ocorreram em duas direções conhecidas, o desempenho conjunto nas variáveis observadas teve maior rapidez de detecção;
- iii. Ainda, no caso em que ocorreu alterações na média e na variância em apenas uma das direções conhecidas, especificamente, quando $\sigma_e = 0,5$, a ACP confunde as direções de variação (ver célula destacada em vermelho);

Quadro 26: ARL para o caso de alterações na média e aumentos na variância considerando duas direções de perturbação com mesma variância.

DIREÇÕES: ROTAÇÃO E DESLOCAMENTO													
Fase I		ARL											
Casos	σ_e	X_1	X_2	X_3	X_4	$\bigcup X_i$	T^2	$T(x)$	$L(x)$	$\bigcup d_i(x)$	CP_1	CP_2	UCP_i
$T \sim N(0,1)$ $L \sim N(0,1)$	0,1	384,62	330,58	363,64	341,88	157,48	92,17	392,16	366,97	191,39	377,36	366,97	186,05
	0,5	310,08	392,16	322,58	357,14	111,73	95,24	430,11	350,88	193,24	317,46	347,83	165,29
	1,0	412,37	373,83	377,36	421,05	106,10	94,12	439,56	430,11	218,58	400,00	412,37	203,05
Fase II		ARL											
$T \sim N(1,1.5)$ $L \sim N(0,1)$	0,1	28,3	27,3	28,2	27,8	14,0	10,88	10,5	421,1	10,3	10,8	327,9	10,5
	0,5	29,9	30,4	30,1	30,3	12,1	11,38	11,3	347,8	10,9	25,9	30,4	15,0
	1,0	78,6	82,5	78,9	83,5	25,1	13,94	28,0	327,9	25,9	29,5	307,7	27,1
$T \sim N(1,1.5)$ $L \sim N(1,1)$	0,1	55,1	9,7	56,1	9,7	8,1	6,74	10,6	48,4	8,8	12,8	31,8	9,5
	0,5	65,4	11,0	65,8	10,9	7,5	7,19	11,3	50,2	9,4	56,7	10,3	9,0
	1,0	82,1	16,7	87,9	16,8	9,2	7,92	13,9	57,0	11,3	12,0	79,8	10,5
$T \sim N(0,1.5)$ $L \sim N(1,1)$	0,1	27,4	26,8	27,2	26,3	12,8	18,91	22,5	46,0	15,3	22,6	44,4	15,2
	0,5	29,0	32,4	29,4	32,2	11,7	20,13	25,1	46,7	16,5	26,5	30,0	14,3
	1,0	44,9	44,7	43,0	46,1	14,1	24,55	28,2	57,4	19,2	28,8	54,8	19,1
$T \sim N(0,2)$ $L \sim N(2,1)$	0,1	6,4	6,4	6,4	6,4	3,1	7,05	7,6	6,5	3,8	7,5	6,4	3,7
	0,5	7,0	7,3	7,1	7,2	3,1	7,46	8,1	6,9	4,0	6,7	6,8	3,4
	1,0	10,4	10,3	10,0	10,4	3,7	8,73	9,8	8,8	4,9	9,6	8,5	4,8
$T \sim N(1,2)$ $L \sim N(2,1)$	0,1	12,0	3,5	12,0	3,5	2,7	5,43	5,4	6,5	3,2	6,5	5,2	3,2
	0,5	13,6	4,0	13,9	3,9	2,7	5,58	5,9	6,9	3,5	12,7	3,7	3,0
	1,0	21,1	5,4	20,8	5,4	3,2	6,72	7,0	9,1	4,2	5,8	11,6	4,0

Analisando o Quadro 27, observa-se que, quando apenas uma direção de variação sofre alterações na média e na variância, o gráfico de controle nas variáveis transformadas tem maior probabilidade de detecção. No entanto, quando ocorrem alterações simultâneas na média e/ou na variância nas duas direções conhecidas, os possíveis resultados podem ser sintetizados da seguinte forma: quando $\sigma_e = 0,1$, para quaisquer alterações na média e quaisquer aumentos na variância, o desempenho conjunto dos gráficos nas variáveis observadas é mais eficaz; quando o σ_e é igual a 0,5 ou 1,0, os gráficos nas CP's desempenham melhor.

Quadro 27: ARL para o caso de alterações na média e aumentos na variância considerando duas direções de perturbações com mesma variância e ARL_0 's ajustados.

DIREÇÕES: ROTAÇÃO E DESLOCAMENTO					
Fase I		ARL_0			
Casos	σ_e	$UX_{i\alpha-aj}$	T^2	$Ud_i(x)_{\alpha-aj}$	$UCP_{i\alpha-aj}$
T ~ N(0,1) L ~ N(0,1)	0,1	200			
	0,5				
	1,0				
Fase II		ARL			
T ~ N(1,1,5) L ~ N(0,1)	0,1	14,52	15,70	10,22	10,27
	0,5	16,72	16,56	11,31	15,74
	1,0	48,84	21,92	16,93	27,40
T ~ N(1,1,5) L ~ N(1,1)	0,1	8,03	10,92	8,29	8,84
	0,5	9,60	12,21	9,58	9,43
	1,0	13,19	16,49	12,72	12,63
T ~ N(0,1,5) L ~ N(1,1)	0,1	13,42	19,52	14,14	14,03
	0,5	15,93	21,38	16,76	15,11
	1,0	6,88	7,86	6,86	6,63
T ~ N(0,2) L ~ N(2,1)	0,1	3,19	4,27	3,62	3,53
	0,5	3,58	4,58	3,99	3,45
	1,0	6,93	5,57	6,13	4,85
T ~ N(1,2) L ~ N(2,1)	0,1	2,78	3,61	3,16	3,14
	0,5	3,11	3,91	3,51	3,01
	1,0	5,45	4,50	4,98	4,04

4.3.3 Três direções de perturbação com mesma variância

Para o caso de três direções com mesma variância atuando no processo, o Quadro 28 resume os resultados das simulações considerando que essas direções sofreram alterações na média e aumentos na variância.

Quadro 28: ARL para o caso de alterações na média e aumentos na variância considerando três direções de perturbação com mesma variância.

DIREÇÕES: ROTAÇÃO, DESLOCAMENTO E CONTRAÇÃO															
Fase I		ARL													
Casos	σ_e	X_1	X_2	X_3	X_4	$\cup X_i$	T^2	$T(x)$	$L(x)$	$C(x)$	$\cup d_i(x)$	CP_1	CP_2	CP_3	$\cup CP_i$
T ~ N(0,1) L ~ N(0,1) C ~ N(0,1)	0,1	388,3	381,0	444,4	360,4	101,8	98,77	392,2	377,4	408,2	131,6	367,0	357,1	459,8	129,9
	0,5	388,3	416,7	444,4	315,0	100,0	91,32	404,0	377,4	454,5	136,5	408,2	408,2	449,4	140,4
	1,0	430,1	370,4	416,7	381,0	101,5	95,01	344,8	350,9	388,3	120,5	357,1	367,0	439,6	128,6
Fase II		ARL													
T ~ N(0,1.5) L ~ N(1,1) C ~ N(0,1)	0,1	47,8	44,4	47,3	45,5	13,2	12,7	21,9	48,4	400,0	14,8	43,3	24,6	272,1	15,2
	0,5	51,5	50,6	50,4	47,9	13,9	13,6	23,5	45,6	470,6	15,3	28,1	39,6	186,0	15,3
	1,0	62,8	64,3	63,6	59,0	16,8	16,1	29,4	58,4	325,2	18,6	34,3	62,3	127,8	19,3
T ~ N(0,1.5) L ~ N(1,1) C ~ N(2,1)	0,1	7,0	7,0	43,8	44,4	3,5	3,8	21,3	44,1	6,3	4,6	27,1	15,2	8,8	4,8
	0,5	7,8	7,8	50,4	48,0	3,8	4,0	23,5	47,2	7,0	5,1	27,4	10,2	13,8	5,1
	1,0	10,4	10,8	67,1	63,4	5,0	5,0	29,3	56,0	8,9	6,3	18,1	242,4	7,6	5,3
T ~ N(1,1.5) L ~ N(0,2) C ~ N(2,1)	0,1	14,7	4,8	4,9	14,6	2,3	2,4	10,4	7,6	6,5	3,0	8,1	6,3	9,9	2,9
	0,5	15,9	5,3	5,4	15,6	2,5	2,5	11,7	8,2	7,1	3,2	10,2	6,0	9,7	3,0
	1,0	21,5	6,9	6,8	20,4	3,0	3,0	14,1	9,8	9,1	3,9	8,1	8,0	17,8	3,5

De acordo com os resultados do quadro acima, é possível verificar que:

- Os resultados do ARL_0 encontram-se dentro dos valores esperados para todos os esquemas considerados neste estudo;
- Da mesma forma que ocorreu quando foram consideradas alterações apenas na média ou somente aumentos na variância, a ACP em alguns casos ainda não conseguiu identificar perfeitamente as direções de variação (ver campos destacados em vermelho). Note que, nesse caso específico, a direção conhecida Contração não sofreu quaisquer tipos de alterações na média e/ou na variância, e, portanto, seria esperado um ARL com valores em torno de 370 alarmes. Contudo, verificou-se que o ARL para o primeiro caso simulado diminui à medida que o σ_e cresce, chegando ao valor de aproximadamente 128 alarmes para $\sigma_e = 1,0$. Ainda, no segundo caso simulado, quando $\sigma_e = 1,0$, percebe-se

claramente que a ACP mais uma vez não identificou corretamente as direções de variação.

- iii. Sem considerar a igualdade dos ARL_0 's dos esquemas propostos, o desempenho conjunto dos gráficos nas variáveis observadas mostrou melhor sinalização na maior parte dos casos simulados.

Quadro 29: ARL para o caso de alterações na média e aumentos na variância considerando três direções de perturbação e ARL_0 's ajustados.

DIREÇÃO: ROTAÇÃO					
Fase I		ARL_0			
Casos	σ_e	$\bigcup X_{i\alpha-aj}$	T^2	$T(x)_{\alpha-aj}$	$CP_{1\alpha-aj}$
T ~ N(0,1) L ~ N(0,1) C ~ N(0,1)	0,1	200			
	0,5				
	1,0				
Fase II		ARL			
T ~ N(0,1.5) L ~ N(1,1) C ~ N(0,1)	0,1	40,94	30,56	26,02	30,10
	0,5	14,49	22,00	16,67	16,27
	1,0	7,25	26,14	8,16	6,72
T ~ N(0,1.5) L ~ N(1,1) C ~ N(2,1)	0,1	7,08	6,17	6,07	6,29
	0,5	4,88	5,87	4,91	6,09
	1,0	7,44	7,13	8,29	6,84
T ~ N(1,1.5) L ~ N(0,2) C ~ N(2,1)	0,1	2,73	2,83	3,41	3,38
	0,5	2,74	3,24	3,36	3,25
	1,0	3,97	4,61	4,62	4,13

O Quadro 29 mostra que o desempenho conjunto dos gráficos de controle nas variáveis observadas foi mais eficaz quando alterações simultâneas na média e na variância ocorreram em pelo menos duas direções de perturbação. No caso em que uma das direções sofreu alteração na média e a outra direção sofreu aumento na variância, ou ainda, quando apenas uma direção sofreu aumento na variância e as outras duas direções sofreram alterações na média, houve uma alternância na eficácia entre os esquemas considerados.

4.3.4 Duas direções de perturbação com diferentes variâncias

A seguir, a última situação abordada nesta tese considera que o processo tenha duas direções conhecidas com variâncias diferentes e que essas direções sejam afetadas com alterações na média e aumentos na variância.

Portanto, avaliando o resultado do Quadro 30, foi possível verificar que:

i. Os gráficos nas variáveis transformadas e nas CP's tem aproximadamente o mesmo ARL, isto é, a mesma probabilidade de sinalização. Isto ocorre, porque conforme análise feita para o Quadro 13, as CP's coincidem com as variáveis transformadas.

ii. Da mesma forma que ocorreu no caso de aumentos apenas na variância, quando simulados alterações na média e aumentos da variância, os gráficos nas variáveis transformadas e nas CP's foram mais eficazes em detectar pontos fora de controle em todos os caso simulados, mesmo tendo maior ARL_0 . Sendo assim, para este caso, não foi necessário recalculer os ARL's considerando um ARL_0 aproximado ao do gráfico de T^2 .

Quadro 30: ARL para o caso de alterações na média e aumentos na variância considerando duas direções de perturbação com diferentes variâncias.

DIREÇÕES: ROTAÇÃO E DESLOCAMENTO													
Fase I		ARL ₀											
Casos	σ_e	X_1	X_2	X_3	X_4	UX_1	T^2	$T(x)$	$L(x)$	$Ud_1(x)$	CP_1	CP_2	UCP_1
$T \sim N(0,3)$ $L \sim N(0,1)$	0,1	396,04	408,16	412,37	396,04	209,42	94,12	412,37	370,37	195,12	412,37	370,37	195,12
	0,5	392,16	380,95	353,98	363,64	168,78	90,09	370,37	353,98	181,82	370,37	353,98	181,82
	1,0	373,83	350,88	373,83	350,88	134,23	93,46	353,98	347,83	175,44	353,98	347,83	175,44
Fase II		ARL											
$T \sim N(0,3)$ $L \sim N(1,1.5)$	0,1	73,1	78,1	71,8	77,7	37,6	10,9	193,2	8,2	8,0	193,2	8,2	7,9
	0,5	87,1	97,1	86,4	94,8	39,0	11,4	174,7	8,4	8,0	175,4	8,4	8,1
	1,0	125,8	122,7	126,2	119,0	44,9	13,9	187,8	10,9	10,3	187,8	10,8	10,3
$T \sim N(1,3)$ $L \sim N(0,1.5)$	0,1	81,0	79,7	80,3	82,3	42,1	18,9	149,3	17,1	15,4	149,3	17,1	15,4
	0,5	96,6	91,1	91,7	89,7	40,5	20,1	124,6	15,4	13,8	124,6	15,4	13,8
	1,0	104,4	120,8	111,7	119,4	42,6	24,6	127,4	19,8	17,2	127,4	19,8	17,2
$T \sim N(0,4.5)$ $L \sim N(1,1.5)$	0,1	13,7	14,3	13,7	14,5	8,4	6,7	16,9	8,3	5,8	16,9	8,3	5,8
	0,5	15,0	15,3	15,3	15,1	8,3	7,2	15,8	8,5	5,7	15,8	8,5	5,8
	1,0	19,2	19,4	19,4	19,3	9,3	7,9	17,1	10,7	6,8	17,1	10,6	6,8
$T \sim N(2,3)$ $L \sim N(0,2)$	0,1	31,4	31,2	31,3	31,4	16,2	7,0	64,6	6,2	5,7	64,6	6,1	5,7
	0,5	35,1	34,1	34,6	33,6	15,5	7,5	55,6	6,2	5,7	55,7	6,2	5,7
	1,0	44,0	43,5	44,0	43,7	17,3	8,7	61,3	7,7	6,9	61,3	7,7	6,9

4.4 Monitorando o Resíduo

Nesta seção serão apresentados os resultados obtidos da análise de desempenho do gráfico de controle no quadrado da norma do vetor resíduo, cujo objetivo é detectar outras causas de perturbação do processo, diferentes das causas consideradas mais frequentes, que já estão sendo controladas individualmente pelos gráficos nas variáveis transformadas. Conforme mencionado na seção de metodologia (item 3.5), o primeiro passo para realização desse monitoramento, é determinar a forma distribucional do quadrado da norma do vetor resíduo, de maneira a estabelecer um limite de controle que corresponda a sua probabilidade de alarme falso especificada. Entretanto, foi visto por meio da equação (30) na seção 3.4, que o quadrado da norma do referido vetor depende da soma de produtos de variáveis aleatórias normalmente distribuídas, a qual sabe-se não ter uma distribuição exata de probabilidade, apenas aproximações, ou seja, não há um modelo teórico pré-existente da distribuição do quadrado da norma do vetor resíduo. Dessa forma, foi utilizado o método gráfico associado a um teste de adequação de ajuste, especificamente, o teste de Anderson-Darling, para verificar qual a distribuição de probabilidade aproximada do quadrado da norma do vetor resíduo.

A adequação da distribuição dos dados à distribuição teórica é considerada tanto melhor quanto maior for o p-valor dado pela estatística de Anderson-Darling. Assim, como poderá ser visto no conjunto de gráficos a seguir, embora várias distribuições de probabilidade tenham sido testadas, duas distribuições apresentaram melhor qualidade no ajuste; a saber, a distribuição exponencial e a distribuição normal com a aplicação da transformada de Box-Cox. Com relação à análise do ajuste, um detalhe interessante a ser observado é que para os desvios-padrão dos erros aleatórios testados menores que 0,75, a distribuição normal com a transformação de Box-Cox mostra melhor qualidade de ajuste (maior p-valor), e para desvios maiores que 0,75, a distribuição exponencial tem maior p-valor. Além disso, quando o desvio é exatamente igual a 0,75 as duas distribuições têm praticamente a mesma significância. A seguir, o resultado gráfico do ajuste e do teste de Anderson-Darling das principais distribuições testadas.

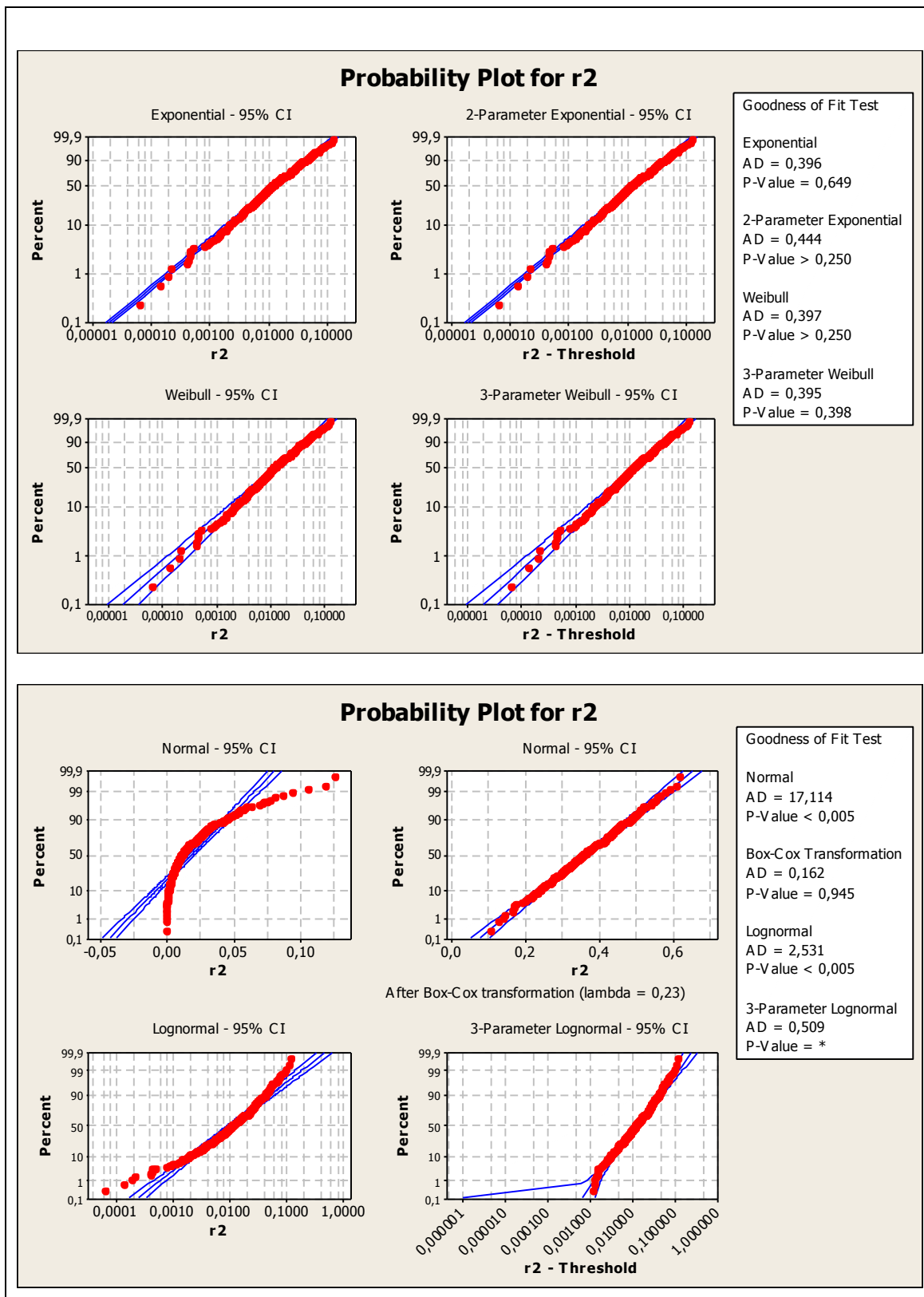


Figura 6: Resultado da análise gráfica e do teste de Anderson-Darling para $\sigma_e = 0,1$.

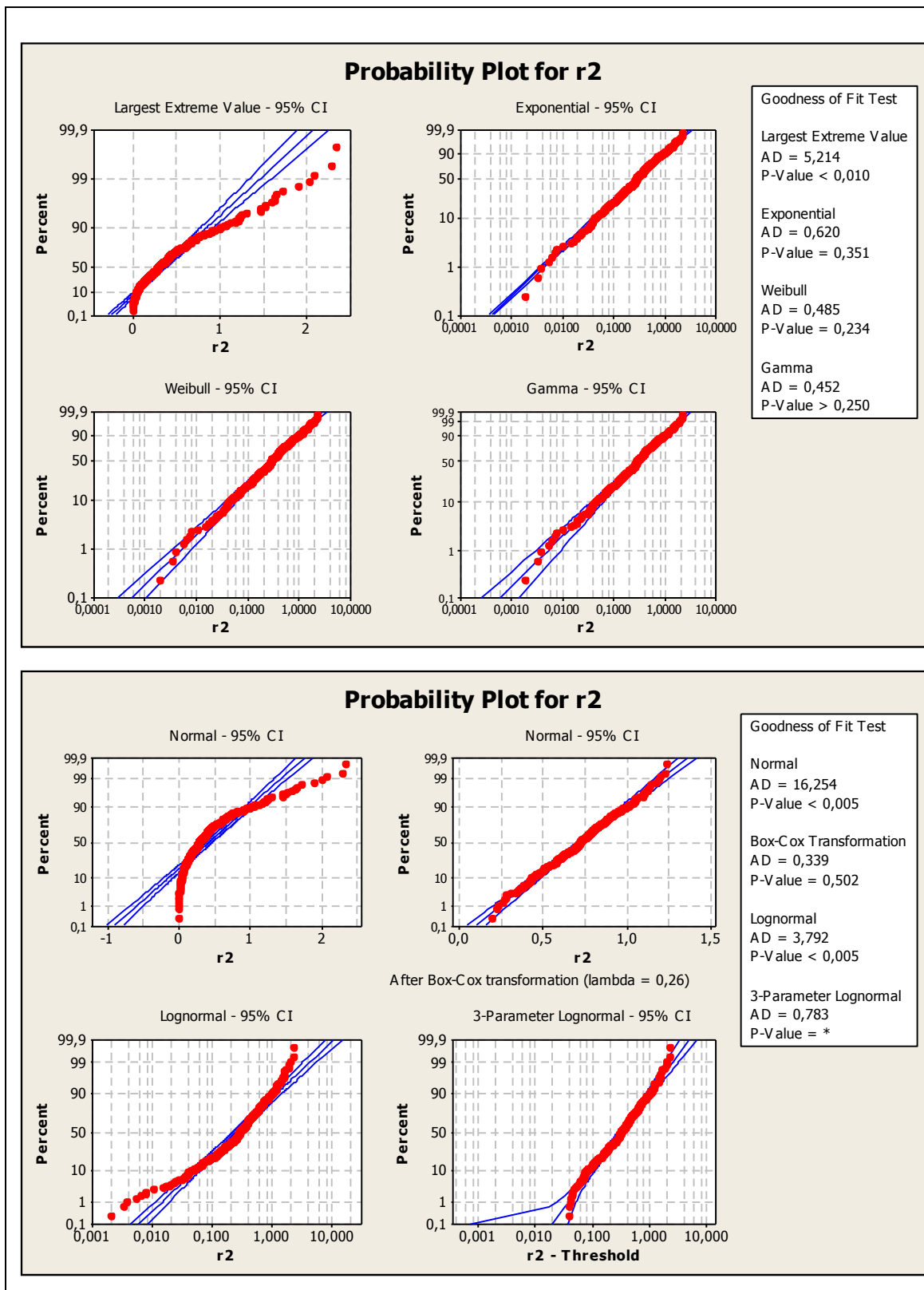


Figura 7: Resultado da análise gráfica e do teste de Anderson-Darling para $\sigma_e = 0,5$.

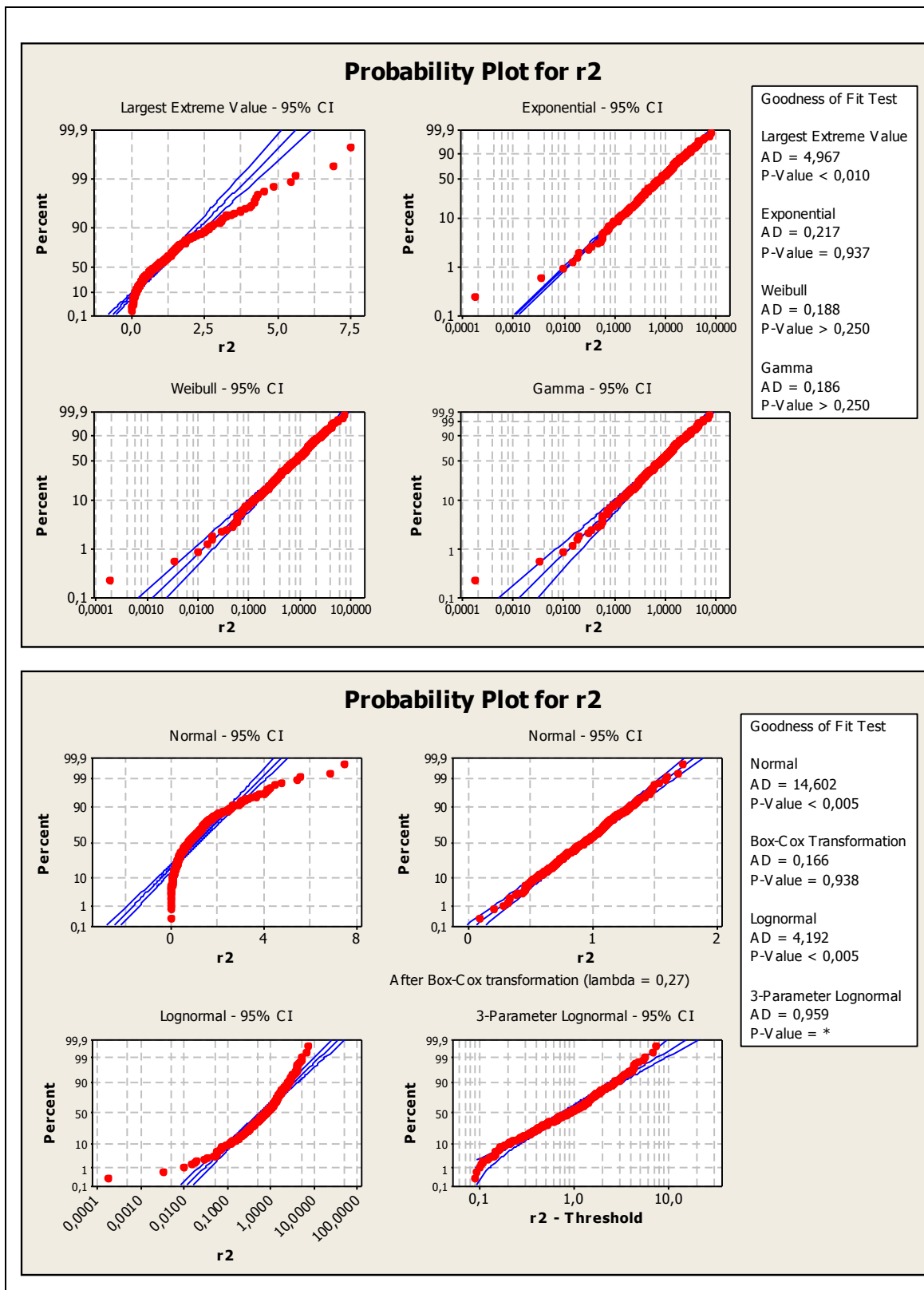


Figura 8: Resultado da análise gráfica e do teste de Anderson-Darling para $\sigma_e = 0,75$.

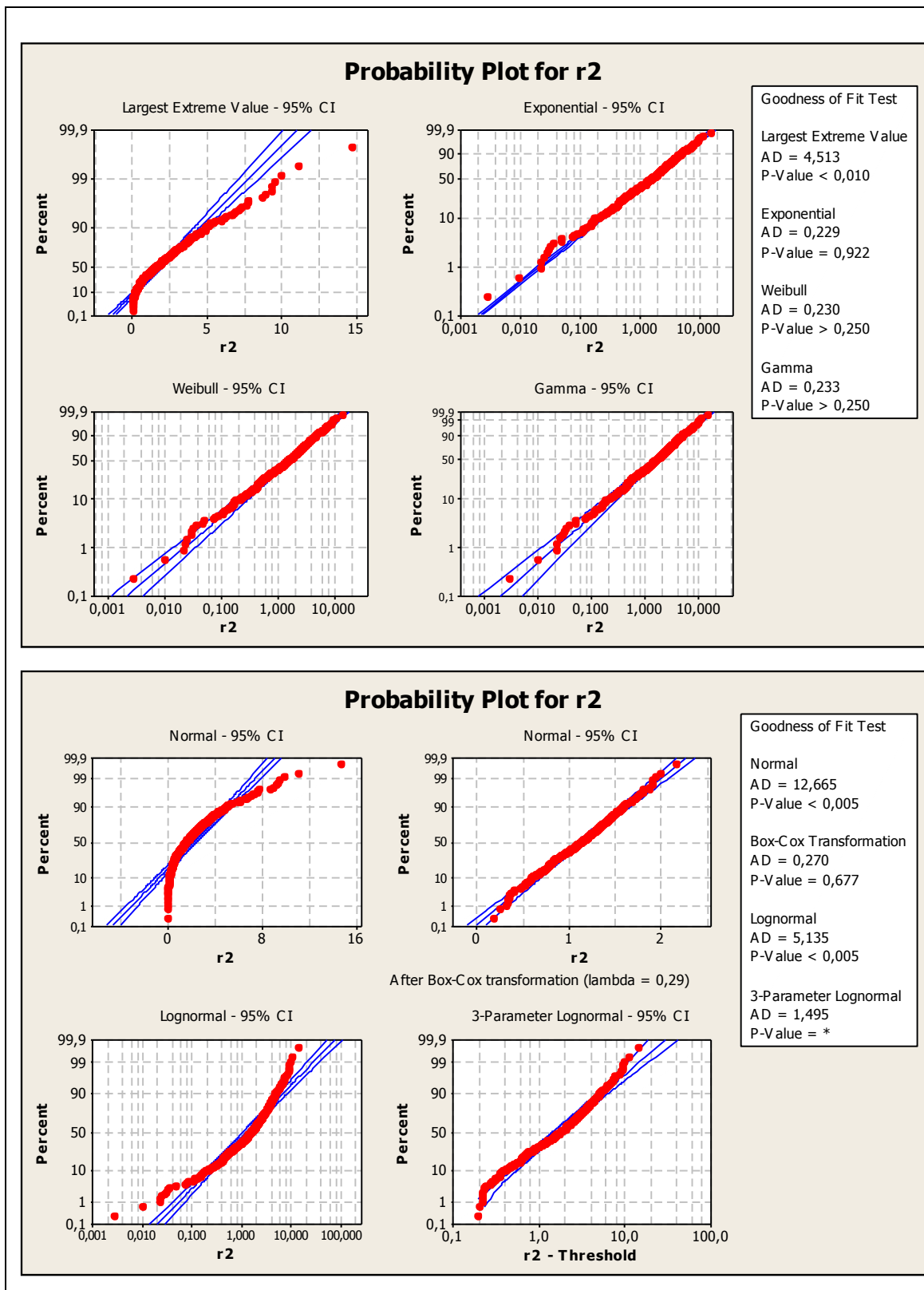


Figura 9: Resultado da análise gráfica e do teste de Anderson-Darling para $\sigma_e = 1,0$.

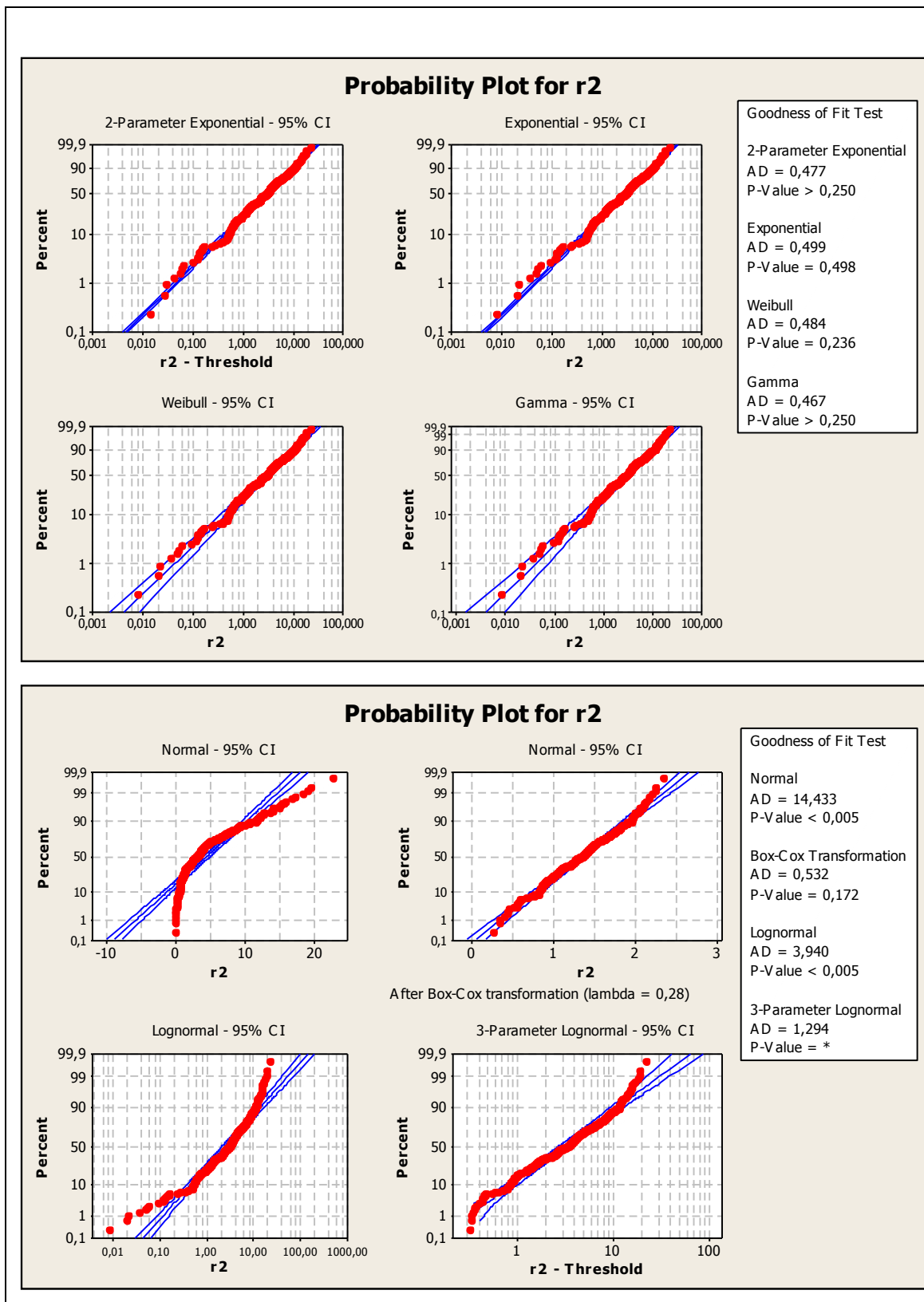


Figura 10: Resultado da análise gráfica e do teste de Anderson-Darling para $\sigma_e = 1,5$.

Conhecida as distribuições de probabilidade que mais se ajustaram ao quadrado da norma do vetor resíduo, será realizada agora a análise de desempenho considerando os limites de controle associados a essas distribuições. Dessa forma, foi analisado o desempenho dos gráficos: R2 para o caso em que $\|\mathbf{r}\|^2$ se ajustou a distribuição exponencial; R2T para o caso em que foi aplicada a transformada de Box-Cox em $\|\mathbf{r}\|^2$. Ainda, para efeito de comparação entre os dois gráficos (R2 e R2T) foi estabelecido outro gráfico (R2_{aj}), considerando uma probabilidade de alarme falso ajustada para o gráfico de R2, a fim de produzir um limite de controle no quantil de $(1 - \alpha)$ da distribuição normal. A seguir os resultados dos desempenhos dos referidos gráficos:

Quadro 31: ARL para o quadrado da norma do vetor resíduo segundo a forma distribucional com probabilidade de alarme falso $\alpha = 0,0027$.

RESÍDUO				
Fase I		ARL ₀		
Casos	σ_e	R2	R2 _{aj}	R2T
T ~ N(0,1) L ~ N(0,1)	0,1	53,91	370,40	1428,57
	0,5	53,69	368,99	634,92
	0,75	53,91	370,40	869,57
	1,0	54,13	371,83	547,95
	1,5	53,84	369,93	615,38
Fase II		ARL		
T ~ N(0,1) L ~ N(0,1) C ~ N(0,1)	0,1	1,12	7,04	1,17
	0,5	1,92	12,16	2,54
	0,75	2,81	18,01	4,39
	1,0	4,12	26,49	7,47
	1,5	8,20	53,02	22,90
T ~ N(0,1) L ~ N(0,1) C ~ N(1,1)	0,1	1,07	6,70	1,10
	0,5	1,46	9,28	1,72
	0,75	1,87	11,83	2,45
	1,0	2,45	15,59	3,60
	1,5	4,29	27,57	9,09

onde:

R2 – gráfico de $\|\mathbf{r}\|^2$ com limite de controle no quantil de $(1 - \alpha)$ da distribuição exponencial;

R2_{aj} – gráfico de $\|\mathbf{r}\|^2$ com probabilidade de alarme falso ajustado para produzir um limite de controle no quantil de $(1 - \alpha)$ da distribuição normal;

R2T – gráfico na transformada de Box-Cox de $\|\mathbf{r}\|^2$ com limite de controle ao quantil de $(1 - \alpha)$ da distribuição normal;

Da análise dos resultados apresentados no Quadro 31, pode-se destacar que:

- i. Para a Fase I, os resultados dos ARL_0 's estão dentro do esperado. Note-se que, para o gráfico de R2, como a distribuição de probabilidade ajustada para a variável em estudo é uma exponencial, é normal que ocorra um alarme falso a cada 55 observações. No caso, em que foi ajustado o valor da probabilidade de alarme falso (α) considerando a distribuição normal ($R2_{aj}$), foi obtido um alarme falso a cada 370 pontos. Entretanto, analisando o gráfico na transformada de Box-Cox de $\|\mathbf{r}\|^2$ (R2T), verificou-se um ARL_0 bem acima do esperado, o que não significa necessariamente um resultado anormal, pois, conforme observado por meio da análise gráfica, o quadrado da norma do vetor resíduo não é normalmente distribuído, sendo a transformada de Box-Cox apenas uma aproximação da distribuição normal.
- ii. Com relação à Fase II, verificou-se que ambos os esquemas simulados (R2, $R2_{aj}$ e R2T) conseguiram detectar a presença da componente Contração ($C(x)$), considerando a mesma com ou sem alteração na média, e especialmente quando o σ_e é pequeno. Contudo, observa-se que o gráfico na transformada de Box-Cox de $\|\mathbf{r}\|^2$ obteve melhor desempenho em sinalizar em todos os casos analisados, mesmo tendo um maior ARL_0 .

Considerando agora uma probabilidade de alarme falso de 0,005, os resultados podem ser observados no Quadro 32. Assim, da análise do referido quadro é possível constatar que:

- i. No que tange a fase de estimação dos parâmetros do processo (Fase I), mais uma vez, não é possível afirmar que os resultados do ARL_0 estão fora da normalidade, pois, conforme observado na análise do Quadro 31, o gráfico de R2 considera que $\|\mathbf{r}\|^2$ tem distribuição aproximadamente exponencial e o

gráfico de R2T é apenas uma aproximação da distribuição normal obtida através da transformação de Box-Cox. Nesse caso, um ARL_0 seria considerado dentro do esperado se ocorresse um alarme falso a cada 200 pontos observados, visto que foi usado um $\alpha = 0,005$.

- ii. Para a Fase II, observa-se novamente que os esquemas testados tiveram um bom desempenho em detectar alterações fora do plano $T(x)L(x)$, mesmo para um $\sigma_e = 1,5$. Nesse caso, mais uma vez, o gráfico na transformada de Box-Cox de $\|\mathbf{r}\|^2$ teve maior probabilidade de sinalização em todos os casos, especialmente quando comparado a $R2_{aj}$.

Quadro 32: ARL para o quadrado da norma do vetor resíduo segundo a forma distribucional com probabilidade de alarme falso $\alpha = 0,005$.

RESÍDUO				
Fase I		ARL ₀		
Casos	σ_e	R2	R2 _{aj}	R2T
T ~ N(0,1) L ~ N(0,1)	0,1	44,49	199,00	555,56
	0,5	45,25	204,12	333,33
	0,75	44,40	198,57	327,87
	1,0	44,59	199,86	272,11
	1,5	45,98	209,85	259,74
Fase II		ARL		
T ~ N(0,1) L ~ N(0,1) C ~ N(0,1)	0,1	1,12	2,25	1,15
	0,5	1,87	3,15	2,33
	0,75	2,69	5,52	3,83
	1,0	3,90	7,34	6,24
	1,5	7,54	19,03	16,60
T ~ N(0,1) L ~ N(0,1) C ~ N(1,1)	0,1	1,07	1,96	1,09
	0,5	1,45	2,39	1,64
	0,75	1,82	3,84	2,25
	1,0	2,36	5,35	3,21
	1,5	4,02	16,13	7,17

Por fim, foi analisado o Quadro 33, considerando um $\alpha = 0,01$. Dessa análise observou-se que:

- i. Os valores do ARL_0 do gráfico na transformada de Box-Cox ainda são maiores do que o esperado, mas, à medida que o desvio-padrão do erro aleatório

aumenta, esses valores tendem a se aproximar do ARL_0 esperado para $\alpha = 0,01$, que implica aproximadamente um alarme falso a cada 100 observações.

- ii. Considerando a ocorrência de alterações fora do hiperplano varrido pelas direções de variação mais frequentes, isto é, T e L, verificou-se, mais uma vez, que ambos os gráficos são eficazes na sinalização dessa alteração, sendo que, novamente, o gráfico na transformada de Box-Cox de $\|\mathbf{r}\|^2$ tem maior poder de detecção, mesmo ainda tendo um ARL_0 maior.

Quadro 33: ARL para o quadrado da norma do vetor resíduo segundo a forma distribucional com probabilidade de alarme falso $\alpha = 0,01$.

RESÍDUO				
Fase I		ARL ₀		
Casos	σ_e	R2	R2 _{aj}	R2T
T ~ N(0,1) L ~ N(0,1)	0,1	36,00	104,02	195,12
	0,5	36,23	105,51	145,99
	0,75	36,17	105,08	133,33
	1,0	35,27	109,28	112,99
	1,5	36,93	110,07	125,00
Fase II		ARL		
T ~ N(0,1) L ~ N(0,1) C ~ N(0,1)	0,1	1,11	1,19	1,14
	0,5	1,82	2,81	2,14
	0,75	2,57	3,69	3,28
	1,0	3,65	5,12	4,97
	1,5	6,76	13,96	11,53
T ~ N(0,1) L ~ N(0,1) C ~ N(1,1)	0,1	1,06	1,12	1,08
	0,5	1,42	1,92	1,56
	0,75	1,77	2,52	2,05
	1,0	2,24	2,85	2,78
	1,5	3,72	6,20	5,48

5 Conclusões e Sugestões para Pesquisa Futura

5.1 Conclusões

Nesta tese foi proposto um método destinado ao controle de processos multivariados em que as alterações na média e/ou na variância ocorrem em direções específicas, que foram denotadas nesse estudo como *direções conhecidas*. Assim, estudou-se nesta pesquisa o desempenho de gráficos de controle das projeções das variáveis observáveis nas direções conhecidas, chamados aqui de gráficos de controle nas variáveis transformadas, comparando seu desempenho com o de gráficos de Shewhart nas variáveis originais, com gráficos de Shewhart nas componentes principais e com gráficos de T^2 nas variáveis originais. Avaliou-se ainda, o desempenho conjunto dos gráficos nas variáveis observadas, dos gráficos nas variáveis transformadas e dos gráficos nas componentes principais.

A análise de desempenho foi feita supondo deslocamentos da média nas direções conhecidas e/ou aumentos da variância nessas mesmas direções. Não foi encontrado na literatura nenhum outro trabalho que avaliasse o desempenho de esquemas no qual os parâmetros do processo se deslocam em direções conhecidas.

Desta forma, para atingir os objetivos traçados nesta tese, inicialmente foi feita uma revisão das principais referências no contexto dos métodos de CEP multivariado, levando em consideração as pesquisas mais recentes nesta área. Em seguida, foi realizada a descrição do modelo geral e o desenvolvimento da teoria para o uso de variáveis transformadas no monitoramento de processos multivariados. Foi proposto também um gráfico adicional para detecção de alterações na média e/ou na variância fora do hiperplano varrido pelas direções conhecidas que já estão sendo controladas individualmente pelos gráficos nas variáveis transformadas. A análise de desempenho foi realizada considerando um exemplo particular. Neste caso, toda a teoria proposta foi aplicada ao exemplo. Ainda, fez-se uma descrição pormenorizada da metodologia para a análise de desempenho.

Por fim, a partir dos resultados obtidos por intermédio das simulações realizadas, levando-se em consideração os objetivos descritos neste estudo, pode-se concluir que:

a) Com relação à identificação das direções conhecidas pela ACP:

Constatou-se que a ACP só mostrou eficácia quando uma única direção de variação atuou no processo ou quando foram consideradas duas direções de perturbação (rotação e deslocamento), mas as variâncias dessas direções eram diferentes. Nos demais casos, a ACP não identificou adequadamente as direções conhecidas, com exceção de alguns casos específicos.

Contudo, no que se refere à detecção de causas especiais, isto é, Fase II do CEP, verificou-se que os gráficos de controle nas CP's, embora não identifiquem qual a causa especial, são *eficientes* na sinalização dessas causas na maior parte dos casos simulados, mesmo nos casos em que a ACP não conseguiu identificar as direções de variação.

b) Análise de desempenho do esquema proposto

Para o caso em que uma única direção de perturbação atua no processo e o processo sofre alterações na média e/ou aumentos na variância, verificou-se que tanto o gráfico na variável transformada quanto o gráfico na componente principal tiveram desempenho superior aos gráficos concorrentes em todos os casos simulados.

Considerando que duas direções de perturbação com mesma variância afetaram o processo, observou-se que para o caso de alterações somente na média ou aumentos simultâneos na média e na variância os gráficos nas variáveis transformadas só tiveram maior rapidez de sinalização quando apenas uma direção sofreu alteração. Nesse caso, os esquemas que desempenharam melhor foram os gráficos na CP's e os gráficos nas variáveis observadas (desempenho conjunto). Quando houve aumentos somente na variância, o esquema que considera as variáveis transformadas foi mais eficaz em praticamente todos os casos simulados. Além disso, considerando esse caso específico (duas direções de perturbação com mesma variância), foi observado que a ACP não identifica corretamente as direções de variação, e, por conseguinte, os gráficos nas componentes principais confundem as direções conhecidas em grande parte dos casos em que há apenas alterações da média ou somente aumentos da variância. Ainda, seriam necessários quatro gráficos nas variáveis observadas para realizar o controle estatístico do processo, implicando em maior sobrecarga operacional. Portanto, caso haja conhecimento prévio das direções conhecidas de variação, devido a conhecimento mais profundo da natureza do processo, que permita definir diretamente as transformações

das variáveis, os gráficos nas variáveis transformadas surgem como uma ótima opção para o monitoramento do processo.

No caso em que três direções de perturbação com mesma variância atuaram no processo, verificaram-se resultados semelhantes aos obtidos para o caso de duas direções com mesma variância. A diferença ficou por conta do gráfico de T^2 , que aparece com o melhor desempenho nos casos em que todas as direções sofreram aumentos de 100% apenas na variância. Para os casos de alterações simultâneas na média e na variância, os gráficos nas variáveis transformadas apresentaram melhor desempenho quando apenas duas direções conhecidas sofreram alteração e no caso em que as três direções tiveram alterações, o desempenho conjunto dos gráficos nas variáveis observadas foi mais eficaz. Dessa forma, considerando os mesmos motivos mencionados quando duas direções conhecidas atuaram no processo, no caso em que ocorrem três direções de perturbação, os gráficos nas variáveis transformadas ainda continuam sendo uma boa opção para o monitoramento do processo.

Por fim, avaliou-se o caso em que duas direções com diferentes variâncias afetaram o processo. Assim, no caso em que ocorreram alterações apenas da média, o desempenho conjunto dos gráficos nas variáveis observadas foi mais eficaz quando apenas uma das direções sofreu a alteração, e quando as duas direções sofreram alterações, os gráficos nas variáveis transformadas e os gráficos nas CP's foram mais eficazes na sinalização. Para os casos simulados de aumentos somente na variância e aumentos simultâneos na média e na variância, novamente os gráficos nas variáveis transformadas e os gráficos nas CP's desempenharam melhor. Contudo, argumenta-se mais uma vez que os gráficos nas variáveis transformadas têm a facilidade de identificar automaticamente a causa especial que produziu a alteração, enquanto que a incapacidade dessa identificação é uma das desvantagens no uso de gráficos nas CP's.

Portanto, com base no que foi exposto, pode-se concluir que o gráfico de controle baseado nas variáveis transformadas é preferível aos gráficos concorrentes, pois esse esquema, em geral, é mais eficaz em sinalizar descontroles no processo, inclusive identificando o tipo de variação e facilitando o diagnóstico. No entanto, cabe ressaltar que, com exceção de alguns casos simulados, todos os esquemas considerados nesse estudo foram eficazes no que se refere ao número médio de sinalizações. Um resumo dos resultados de desempenho podem ser observados no quadro a seguir.

Quadro 34 – Resumo dos resultados da análise de desempenho dos esquemas propostos segundo o caso simulado.

Casos	Alterações		
	Média	Variância	Média e Variância
Uma única direção de perturbação com mesma variância	VT CP	VT CP	VT CP
Duas direções de perturbação com mesma variância	VO CP VT (somente quando uma direção conhecida sofreu alteração)	VT VC	VO CP VT (somente quando uma direção conhecida sofreu alteração)
Três direções de perturbação com mesma variância	VO CP VT (somente quando uma direção conhecida sofreu alteração)	VT VO (quando as três direções sofreram alterações de mesma magnitude)	VO VT , CP (somente quando uma ou duas direções conhecidas sofreram alteração)
Duas direções de perturbação com diferentes variâncias	VT CP	VT CP	VT CP

onde: VT - Esquema que utiliza as Variáveis Transformadas;

VC - Esquema que utiliza as Componentes Principais;

VO - Esquema que utiliza a união das Variáveis Observadas.

c) Análise de desempenho do gráfico no quadrado da norma do vetor resíduo

Para a análise de desempenho do quadrado da norma do vetor resíduo ($\|\mathbf{r}\|^2$) foram analisados os seguintes gráficos: R2 para o caso em que $\|\mathbf{r}\|^2$ se ajustou à distribuição exponencial e R2T para o caso em que foi aplicada a transformada de Box-Cox em $\|\mathbf{r}\|^2$. Dessa forma, verificou-se que, embora ambos os esquemas sejam eficientes na sinalização de perturbações da média e da variância do processo em direções fora do hiperplano varrido pelas direções conhecidas, o gráfico na transformada de Box-Cox de $\|\mathbf{r}\|^2$ foi o esquema mais eficaz em todos os casos simulados.

5.2 Sugestões para Pesquisas Futuras

Considerando que o trabalho proposto obteve bons resultados, são enumeradas algumas sugestões para pesquisas futuras, a seguir:

- Análise do desempenho de esquemas T^2 aplicados às CP's, às variáveis transformadas e comparação do desempenho destes esquemas: entre eles, com o desempenho do gráfico T^2 aplicado às variáveis observadas, e com o desempenho dos três esquemas analisados neste trabalho.
- Extensão dos gráficos propostos para as versões EWMA e/ou CUSUM.
- Estudo da precisão das estimativas dos parâmetros (média e variância) e determinação de limites de controle para os métodos analisados, em função do número de observações iniciais utilizadas para a estimação.
- Estender o trabalho ao caso de processos em dois estágios e estudar o desempenho de métodos de CEP aplicados a cada estágio, métodos estes definidos em função de estudo de transmissão de variação, por exemplo; via de regressão das CP's (ou direções de variação independentes) de um estágio nas CP's (ou direções de variação independentes) do estágio anterior.

Referências Bibliográficas

APARISI, A.; LUNA, M.A.; EPPRECHT, E. (2010). Optimisation of a set of \bar{X} or principal components control charts using genetic algorithms. **International Journal of Production Research**, 48(18), 5345–5361.

APLEY, D. W.; SHI, J. (2001). A factor-analysis method for diagnosing variability in multivariate manufacturing processes. **Technometrics**, 43(1), 84-95.

APLEY, D. W.; LEE, Y. (2003). Identifying spatial variation patterns in multivariate manufacturing processes: a blind separation approach. **Technometrics**, 45(3), 220-234.

AROIAN, L. A.; TANEJA, V. S.; CORNWELL, L. W. (1978). Mathematical forms of the distribution of the product of two normal variables. **Communications in Statistics - Theory and Methods**, 7(2), 164-172.

BARTON, R. R.; GONZALEZ-BARRETO, D. R. (1996). Process-oriented basis representations for multivariate process diagnostics. **Quality Engineering**, 9(1), 107-118.

CATELL, R. B. (1966). The screen test for the number of factors. **Multivariate Behavioral Research**, 1, 140-161.

COSTA, A.F.B.; EPPRECHT E.K.; CARPINETTI, L.C.R. (2005). **Controle Estatístico de Qualidade**. 2. ed. São Paulo: Editora Atlas, 334 p.

CRAIG, C. C. (1936). On the frequency function of xy . **Annals of Mathematical Statistics**, 7, 1-15.

CROSIER, R. B. (1988). Multivariate generalizations of cumulative sum quality control schemes, **Technometrics**, 30, 291-303.

GAIO, A. P. C.; EPPRECHT, E. K. (2006). Controle estatístico de qualidade de processos em estágios, **Anais do XXVI ENEGEP**, Fortaleza-CE.

GONZÁLEZ, I.; SÁNCHEZ, I. (2008a). Principal alarms in multivariate statistical process control. **Journal of Quality Technology**, 40(1), 19-30.

GONZÁLEZ, I.; SÁNCHEZ, I. (2008b). Principal alarms in multivariate process control using independent component analysis. **International Journal of Production Research**, 46(22), 6345-6366.

GONZÁLEZ, I.; SÁNCHEZ, I. (2010). Variable selection for multivariate statistical process control. **Journal of Quality Technology**, 42(3), 242-259.

HAWKINS, D. M. (1991). Multivariate quality control based on regression-adjusted variables. **Journal of Quality Technology**, 33(1), 61-75.

HAWKINS, D. M. (1993). Regression Adjustment for variables in multivariate quality control. **Journal of Quality Technology**, 25(3), 170-182.

HOLMES, D.S. & MERGEN, A.E. (1995). Identifying the sources for out-of-control signals when the control charts is used. **Quality Engineering**, 8(1), 137-143.

HOTELLING, H. (1947). Multivariate quality control, illustrated by the air testing of sample bombsights. **Techniques of Statistical Analysis**, 111-184. New York, MacGraw Hill.

JACKSON, J. E.; MUDHOLKAR, G. S. (1979). Control procedures for Residuals Associated With Principal Components Analysis. **Technometrics**, 21(3), 341-349.

JACKSON, J. E. (1980). Principal components and factor analysis: part I – principal components. **Journal of Quality Technology**, 12(4), 201-213.

JACKSON, J. E. (1981). Principal components and factor analysis: part II – principal components. **Journal of Quality Technology**, 13(1), 46-58.

JOHNSON, R. A.; WICHERN, D. W. (2002). **Applied multivariate statistical analysis**. Prentice-Hall, Inc., Upper Saddle River, New Jersey.

KOURTI, T.; KOURTI, T. (1996). Multivariate SPC methods for process and product monitoring. **Journal of Quality Technology**. 28(4), 409-428.

LOWRY, C. A.; WOODALL W. H.; CHAMP C. W.; RIGDON S. E. (1992). A multivariate exponentially weighted moving average control chart. **Technometrics**, v. 34(1), 46-53.

MACGREGOR, J. F.; KOURTI, T. (1995). Statistical process control of multivariate processes. **Control Engineering Practice**, 3(3), 403-414.

MACHADO, M. A. G.; COSTA, A. F. B. (2008). The use of principal component and univariate charts to control multivariate processes. **Pesquisa Operacional**, 28, 173-196.

MASON, R.L.; TRACY, N.D.; YOUNG, J.C. (1995). Decomposition of T^2 for multivariate control chart interpretation. **Journal of Quality Technology**, 27(2), 99-108.

MASON, R.L.; CHAMP, C. W.; TRACY, N.D.; WIERDA, S. J.; YOUNG, J.C. (1997). Assessment of Multivariate Process Control Techniques. **Journal of Quality Technology**, 29(2), 140-143.

MINGOTI, S.A. (2005). **Análise de Dados Através de Métodos de Estatística Multivariada: Uma Abordagem Aplicada**. Editora UFMG, Belo Horizonte.

MONTGOMERY, D. C. (2004). **Introduction to statistical quality control**. John Wiley & Sons, Inc., New York, New York.

- MURPHY, B. J. (1987). Selecting out of control variables with the T2 multivariate quality control. **The Statistician**, 36, 571-583
- NEDUMARAN, G.; PIGNATIELLO, J.J. (1998). Diagnosing signals from T^2 and χ^2 multivariate control charts. **Quality Engineering**, 10(4), 657-667.
- NOMIKOS, P.; MACGREGOR, J. F. (1994). Monitoring of batch process using multiway principal components analysis. **AIChE Journal**, 40, 1361-1375.
- NOMIKOS, P.; MACGREGOR, J. F. (1995). Multivariate SPC charts for monitoring batch Processes. **Technometrics**, 37(1), 41-59.
- PIGNATIELLO, J.J.; RUNGER G. C. (1990). Comparisons of multivariate CUSUM charts. **Journal of Quality Technology**, 22(3), 173-186.
- RUNGER, G. C.; MONTGOMERY, D. C. (1997). Multivariate and univariate process control: geometry and shift directions. **Quality and Reliability Engineering International**, 13(3), 153-158.
- RUNGER, G. C. (1996). Projections and the U^2 multivariate control chart. **Journal of Quality Technology**, 28(3), 313-319.
- SCHALL, S.; CHANDRA, J. (1987). Multivariate quality control using principal components. **International Journal of Production Research**, 25(4), 571-588.
- SOUZA, A. M.; RIGÃO, M. H. (2005). Identificação de variáveis fora de controle em processos produtivos multivariados. **Revista Produção**, 15(1), 74-86.
- WADE, M. R.; WOODALL, W. H. (1993). A review and analysis cause-selecting control chart. **Journal of Quality Technology**, 25(3), 161-169.
- WARE, R.; LAD, F. (2003). **Approximating the distribution for sums of products of normal variables**. Report. The Mathematics and Statistics Department at Canterbury University.
- WOODALL, W. H.; NCUBE M. M. (1985). Multivariate CUSUM quality-control procedures. **Technometrics**, 27(3), 285-292.
- ZHANG, G. (1985). Cause-selecting control chart: A new type of quality control chart. **The QR Journal**, 12, 221-225.

Apêndice

Análise de Componentes Principais (ACP)

De acordo com Johnson e Wichern (2002) e Mingoti (2005) a ACP tem como objetivo a explicação da estrutura de variância e covariância de um vetor aleatório por meio de combinações lineares das variáveis envolvidas no processo. Estas combinações são chamadas de Componentes Principais (CP) e são não correlacionadas. Geometricamente, representam um novo sistema de coordenadas obtido pela rotação do sistema formado pelas variáveis observadas. Sendo assim, esse novo sistema representa as direções com as maiores variabilidades, produzindo uma descrição mais simples da estrutura de covariância.

Seja Σ a matriz de covariância de ordem $p \times p$ associada ao vetor aleatório $\mathbf{X}' = [X_1, X_2, \dots, X_p]$. Ainda, Σ tem os pares de autovalor-autovetor dado por $(\lambda_1, \mathbf{e}_1), (\lambda_2, \mathbf{e}_2), \dots, (\lambda_p, \mathbf{e}_p)$, onde $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$. Dessa forma, a i -ésima componente principal Y_i da matriz Σ é definida como:

$$Y_i = \mathbf{e}_i' \mathbf{X} = e_{i1} X_1 + e_{i2} X_2 + \dots + e_{ip} X_p \quad (A1)$$

onde \mathbf{e}_i , $i=1, 2, \dots, p$, são os autovetores de Σ . Além disso, se Σ é uma matriz simétrica, seus autovetores associados aos seus respectivos autovalores são ortogonais. Assim, as componentes principais $[Y_1, Y_2, \dots, Y_p]$ também serão ortogonais. Dessa forma, cada componente principal em (A1) tem variância dada pelo autovalor correspondente, isto é,

$$\text{Variância}(Y_i) = \mathbf{e}_i' \Sigma \mathbf{e}_i = \lambda_i \quad (A2)$$

Dessa forma, pode-se dizer que a ACP transforma a informação contida em p variáveis correlacionadas em p componentes principais independentes.

Ainda, a variância total da população é dada pelo traço de Σ , que pode ser obtido facilmente, simplesmente decompondo a matriz de covariância em função dos autovalores e autovetores. Então,

$$\text{Variância total da população} = \sigma_{11}^2 + \sigma_{22}^2 + \dots + \sigma_{pp}^2 = \lambda_1 + \lambda_2 + \dots + \lambda_p \quad (\text{A3})$$

Consequentemente, a proporção da variância total devido a i -ésima CP é dada por:

$$\frac{\lambda_i}{\lambda_1 + \lambda_2 + \dots + \lambda_p}, \quad i = 1, 2, \dots, p, \quad (\text{A4})$$

Assim, se a maior parte de (A3) é atribuída às k primeiras componentes, então essas k componentes podem substituir as p variáveis originais sem muita perda de informação.

Entretanto, é importante estabelecer critérios para escolher quais componentes devem permanecer no modelo, e para isso necessita-se um “ponto de corte”, pois o conceito de maior parte da variância explicada é subjetivo. Embora Johnson e Wichern (2002) afirmem que não há uma resposta totalmente correta para se fazer essa escolha, eles mesmos sugerem alguns indicadores que podem ser utilizados como critérios. Entre esses, um dos que tem mais apelo prático é o percentual de variância explicada pelas k componentes principais. Outro indicador bastante usado é o *scree-plot* desenvolvido por Cattell (1966). Esse gráfico traça os valores dos autovalores $\hat{\lambda}_i$ na ordenada contra a respectiva ordem i na abscissa. Dessa forma, para estabelecer o número de componentes, deve-se verificar o ponto em que os valores de $\hat{\lambda}_i$ tendem a se estabilizar, que coincide com os pontos onde os autovalores são pequenos e aproximadamente do mesmo tamanho. O número de pontos que se encontrarem acima desse ponto de estabilização é que será exatamente o número de componentes principais do modelo.

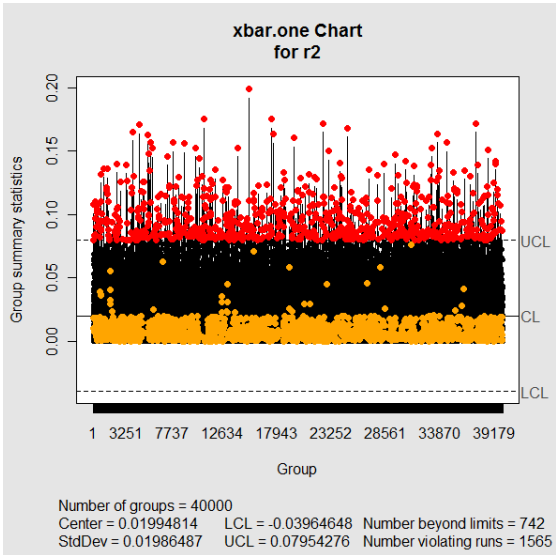
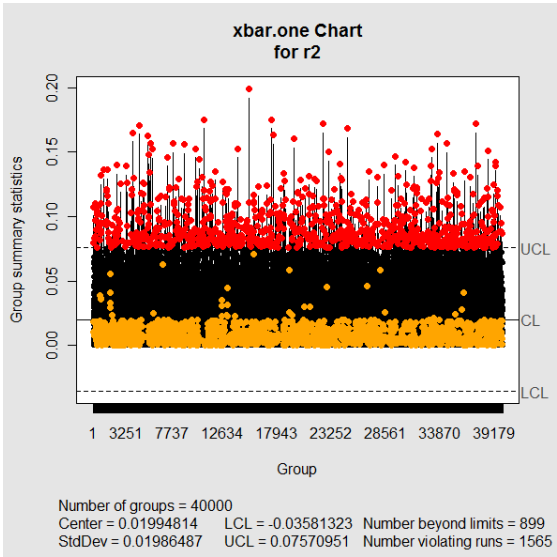
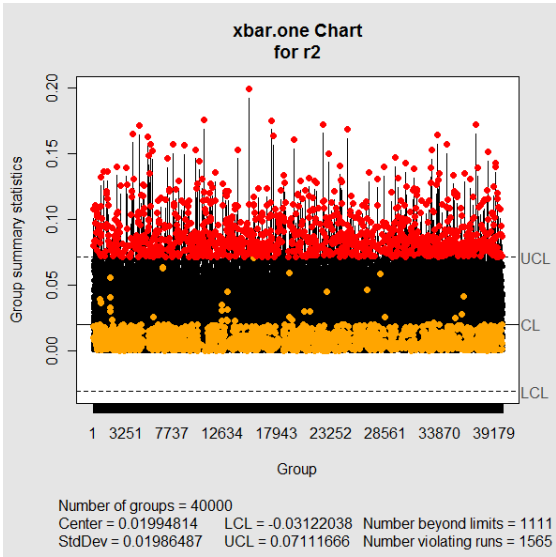
Anexo

Gráficos de Controle para o Quadrado da Norma do Vetor Resíduo

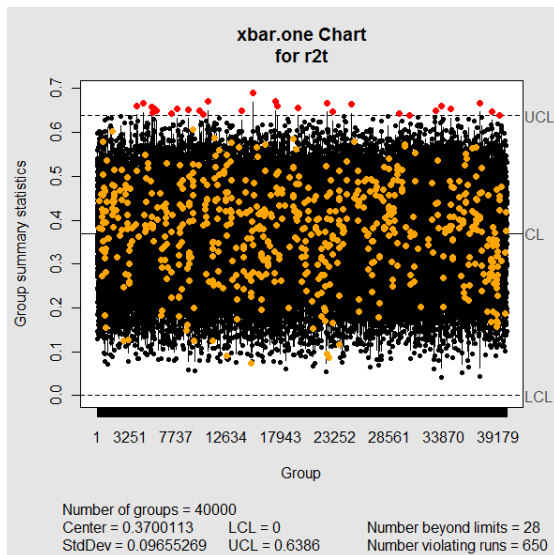
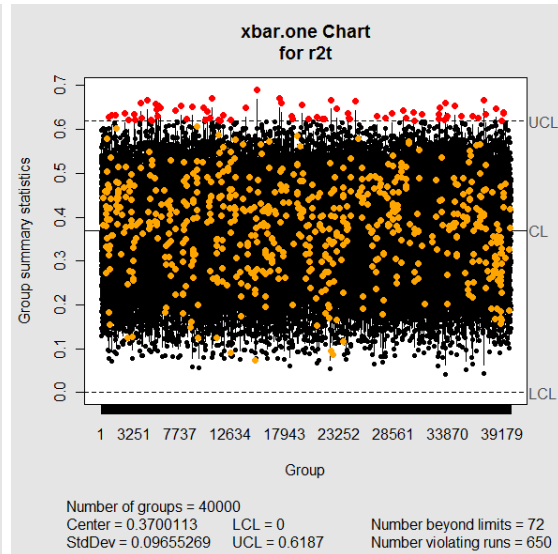
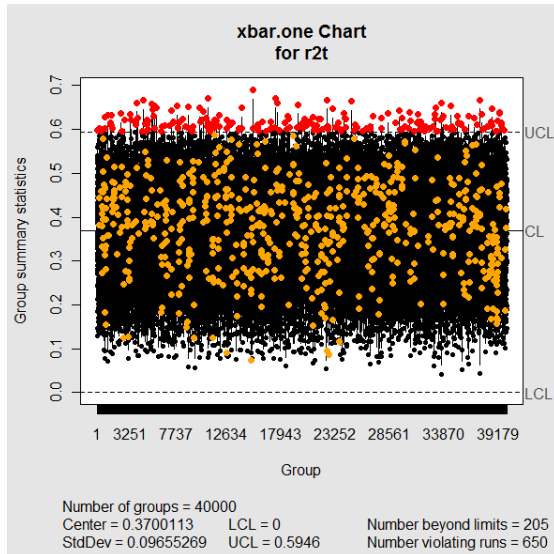
Resultados para a Fase I considerando valores para α (0,01; 0.005; 0.0027) e que apenas as direções conhecidas $T \sim N(0,1)$ e $L \sim N(0,1)$ atuem no processo.

• $\sigma_e = 0,1$

R2

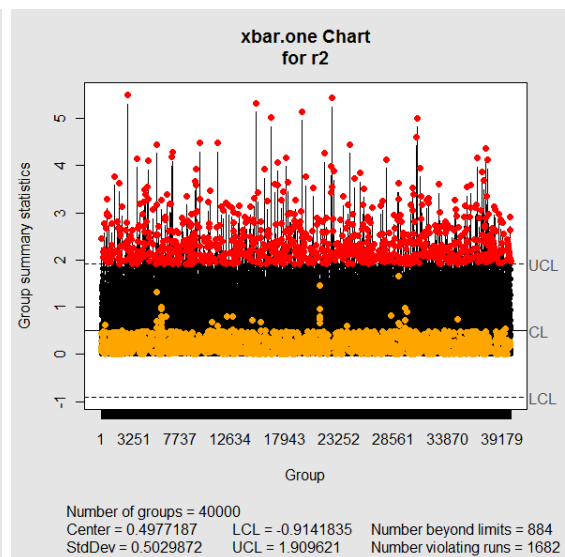
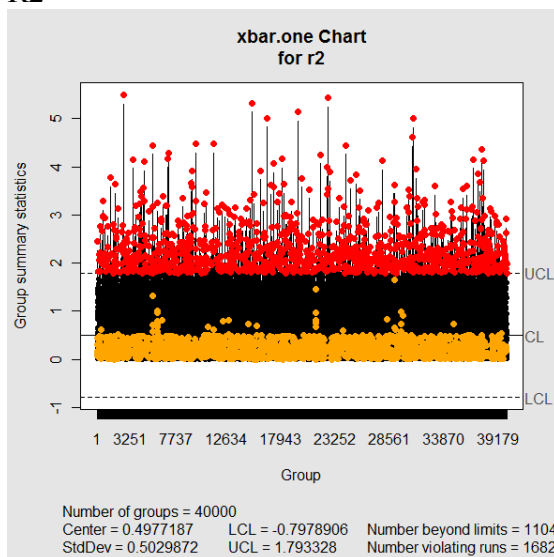


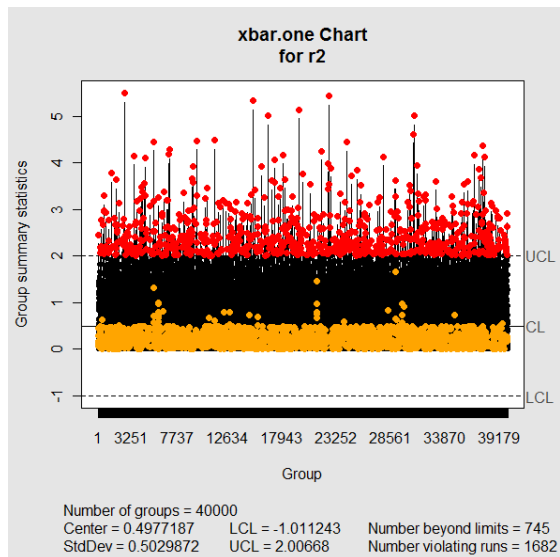
R2T



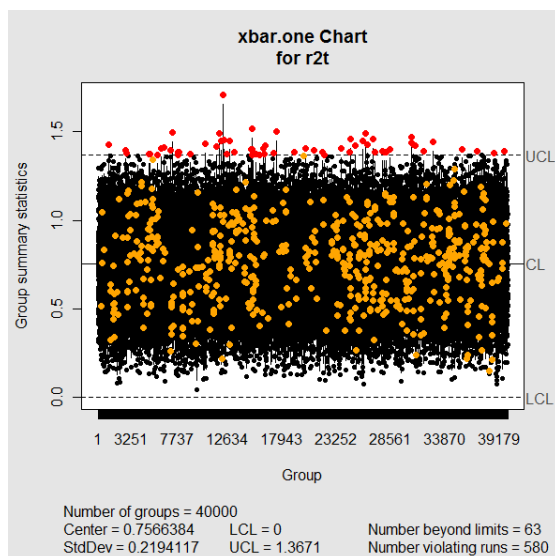
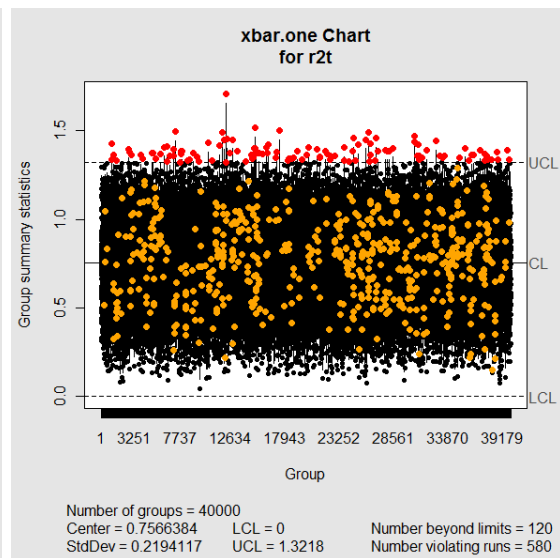
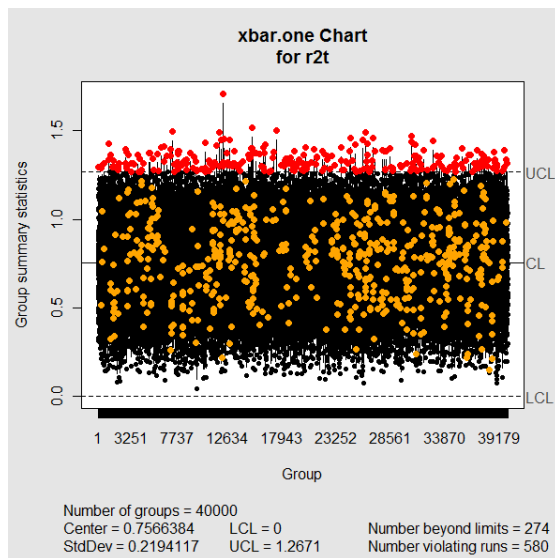
- $\sigma_e = 0,5$

R2



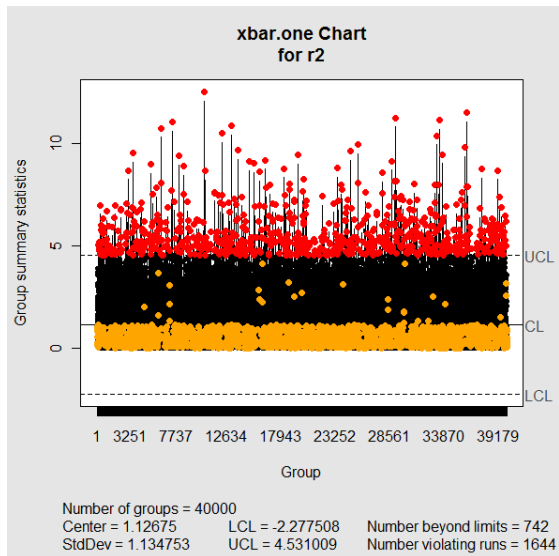
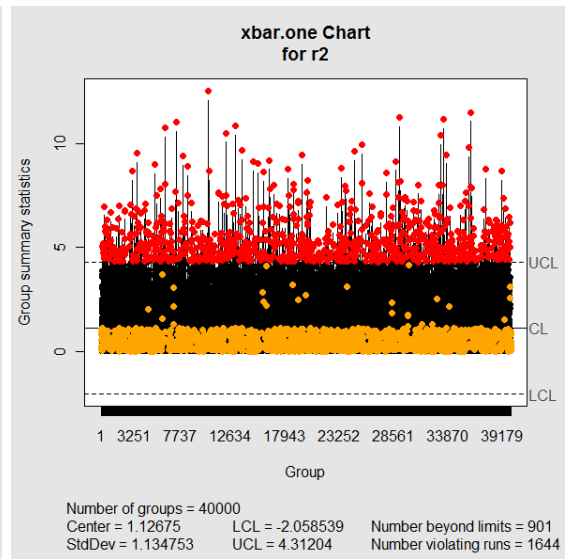
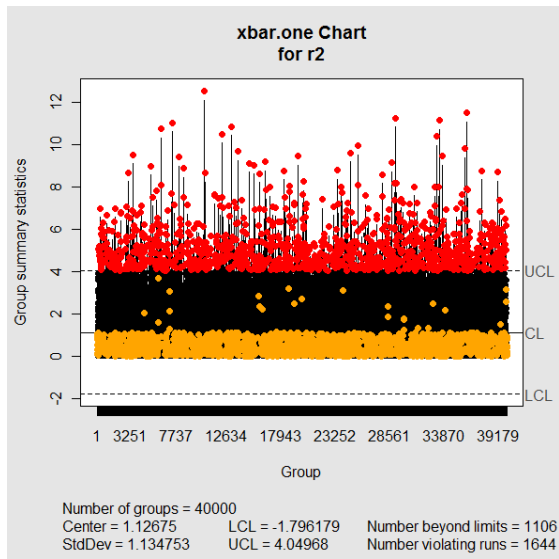


R2T

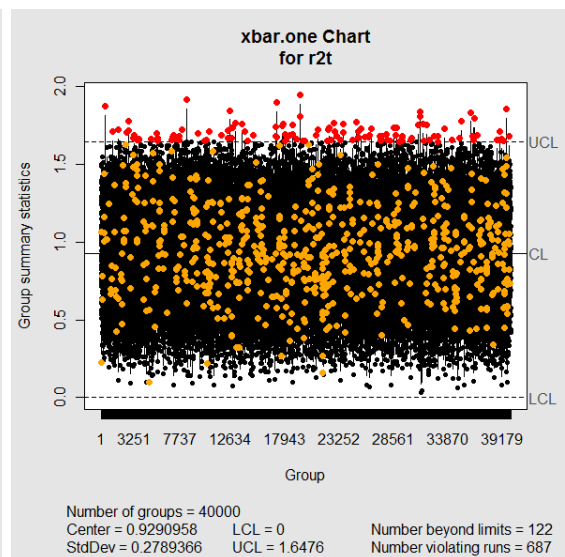
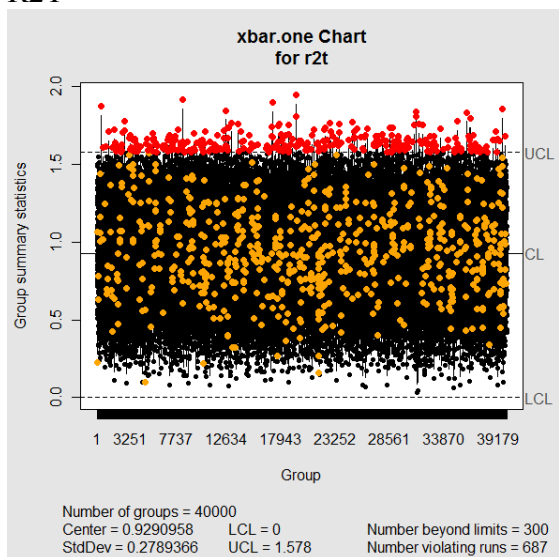


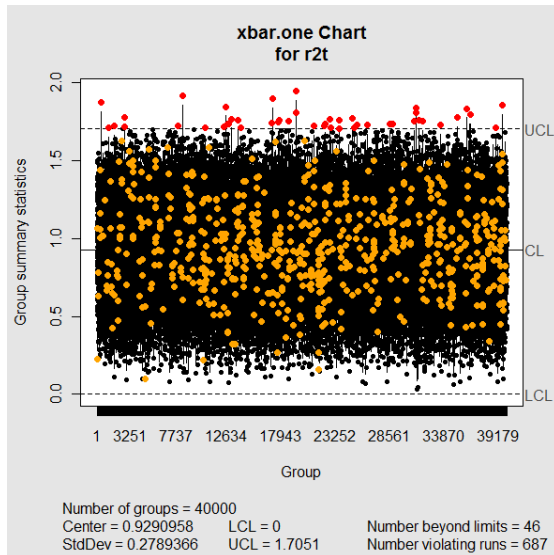
- $\sigma_e = 0,75$

R2



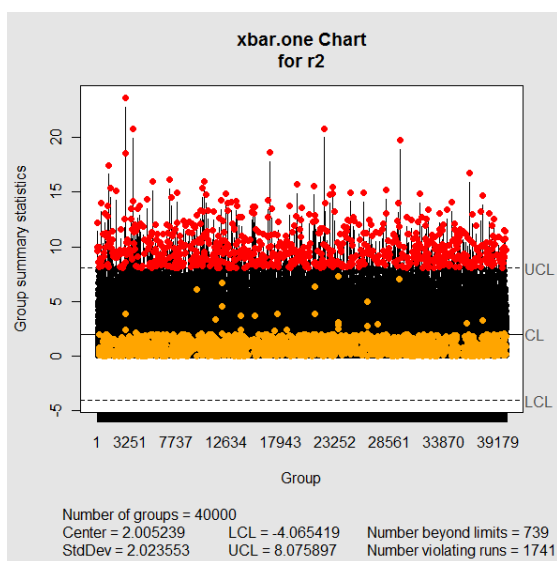
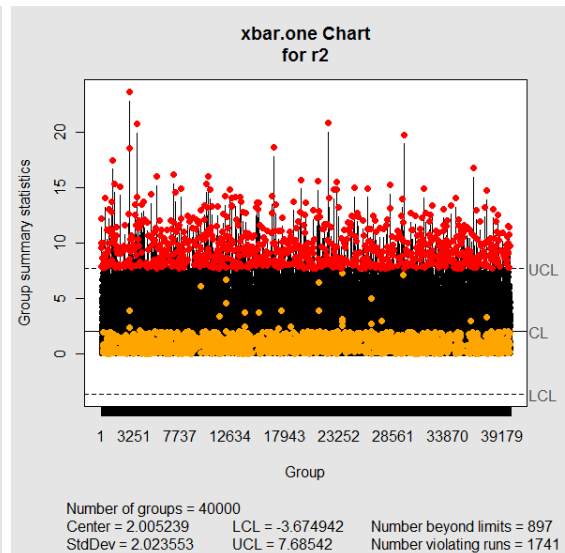
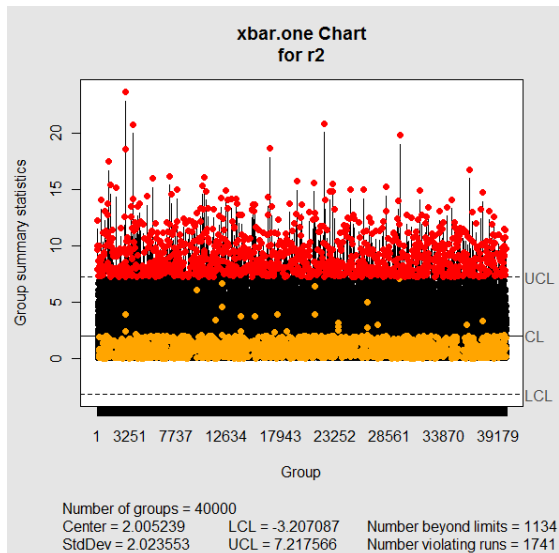
R2T



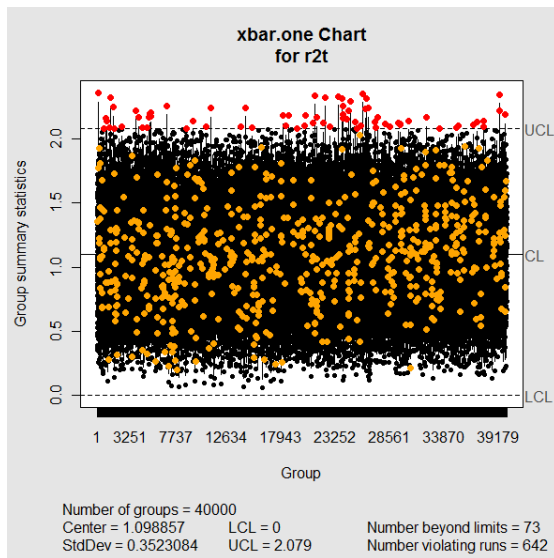
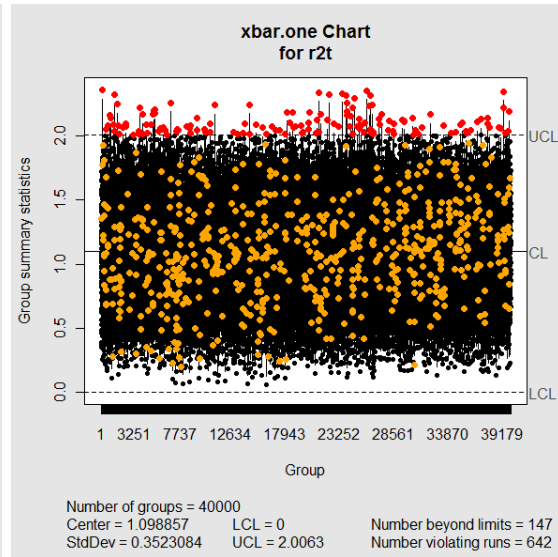
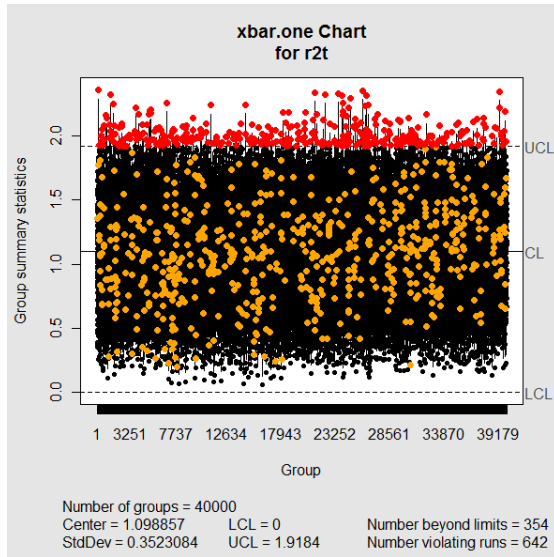


- $\sigma_e = 1,0$

R2

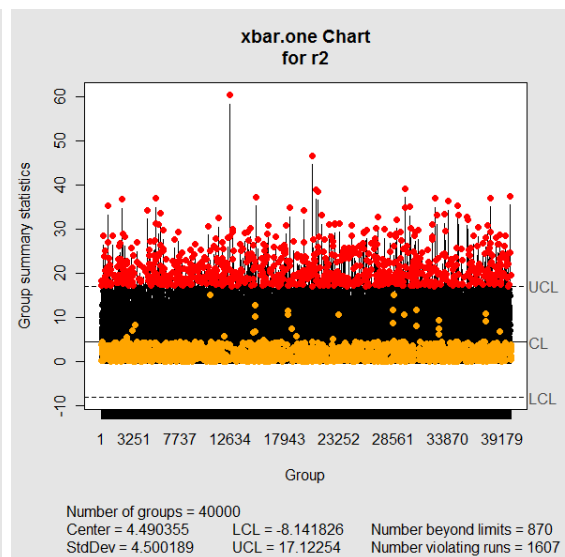
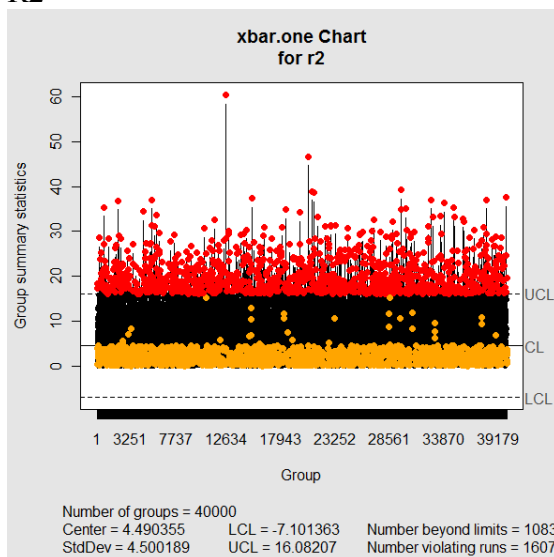


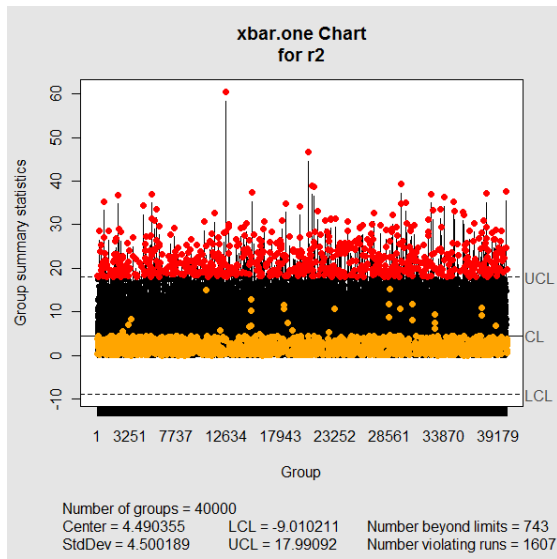
R2T



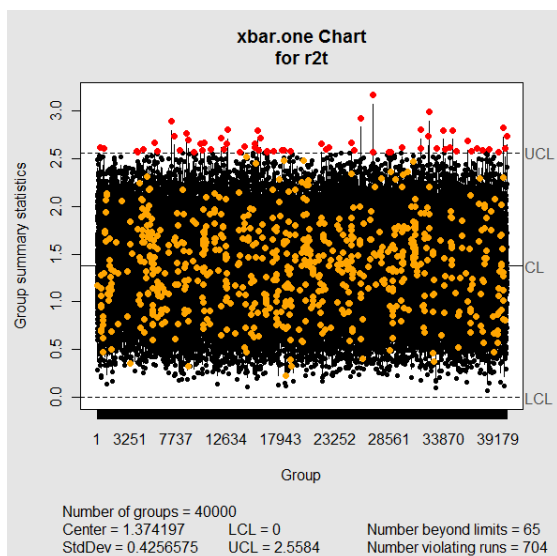
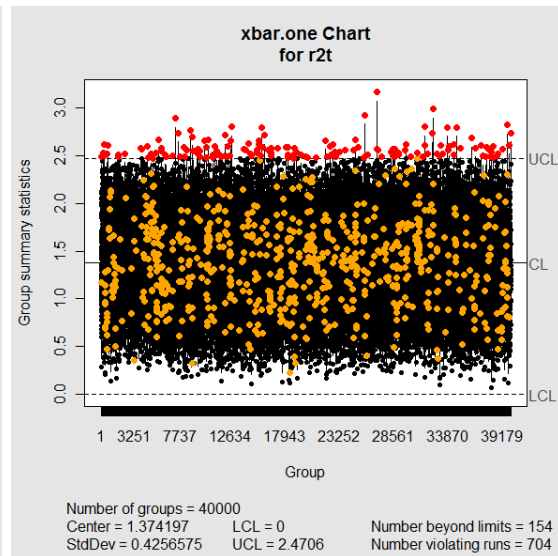
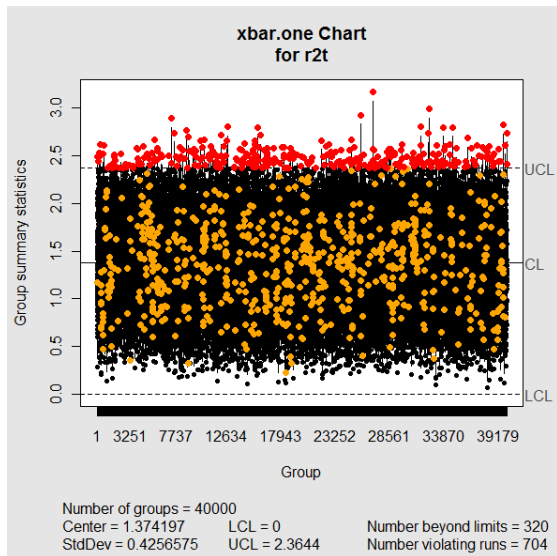
- $\sigma_e = 1,5$

R2





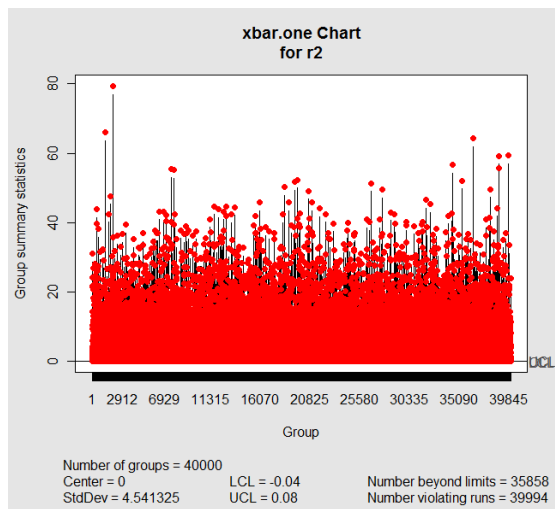
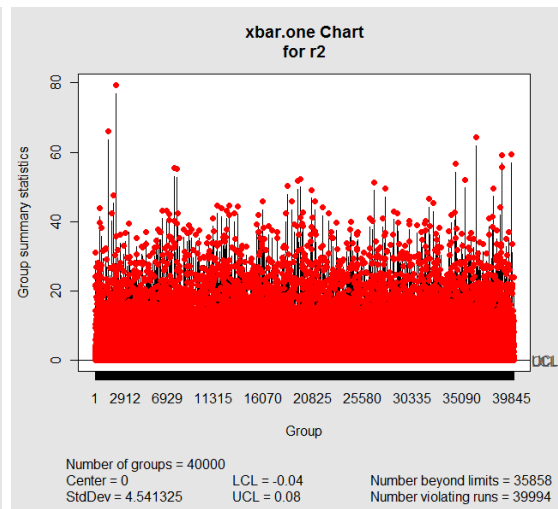
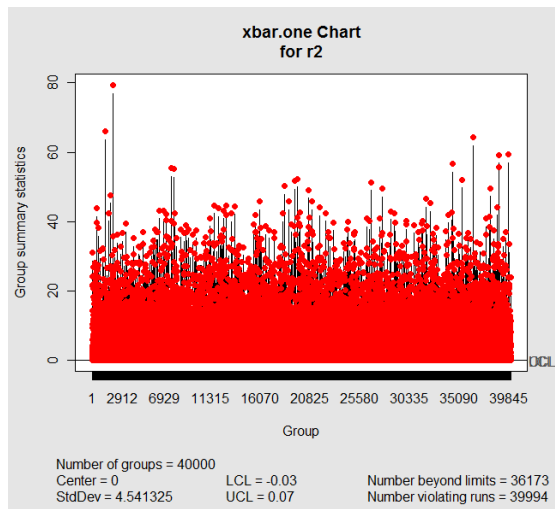
R2T



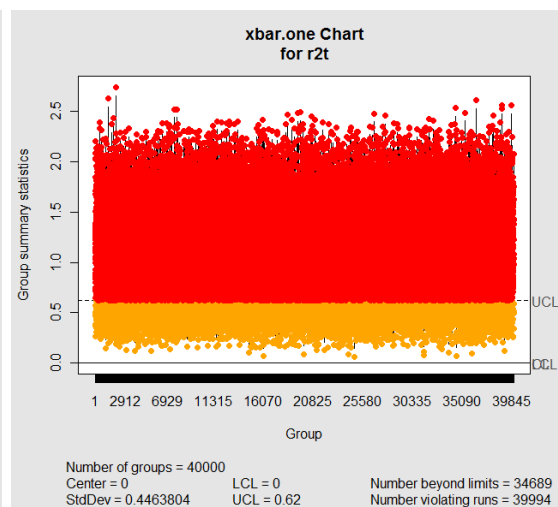
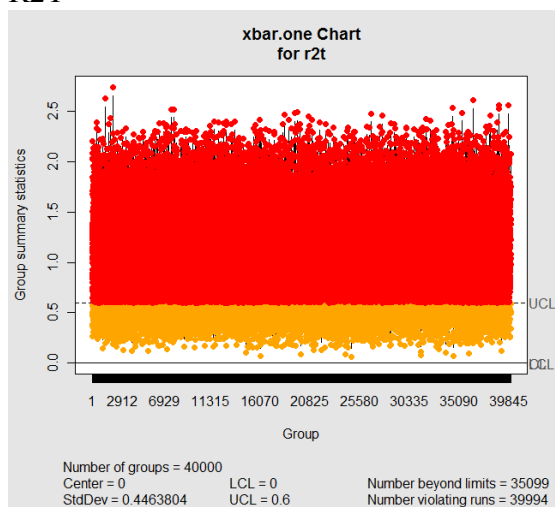
Fase II – $\alpha = 0,01$ e considerando a direção de variação $C \sim N(0,1)$

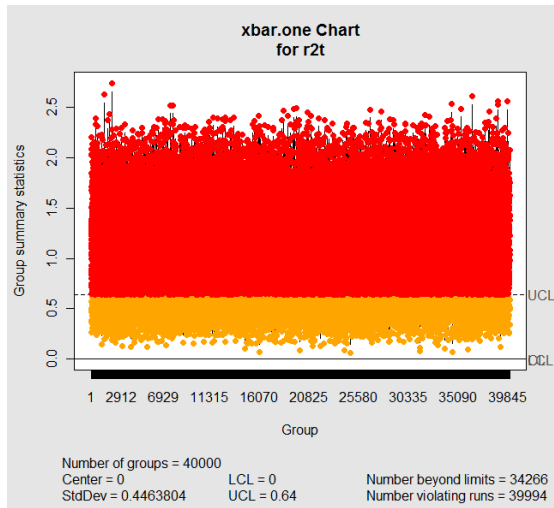
- $\sigma_e = 0,1$

R2



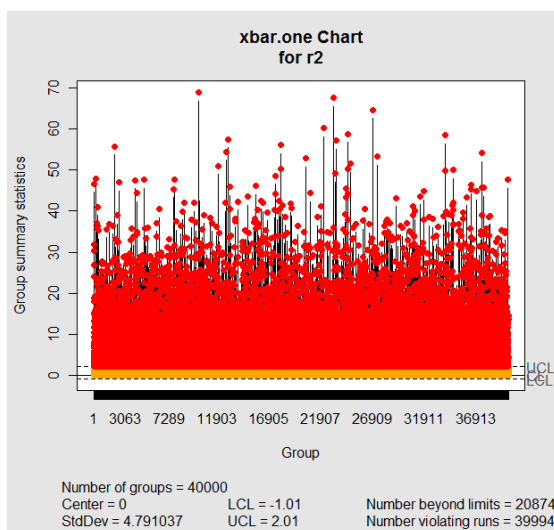
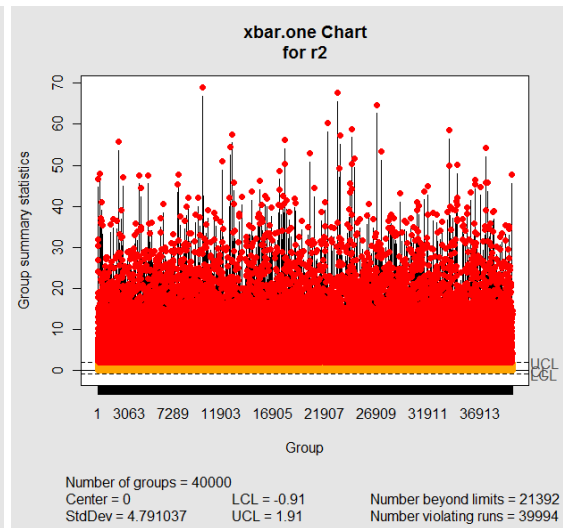
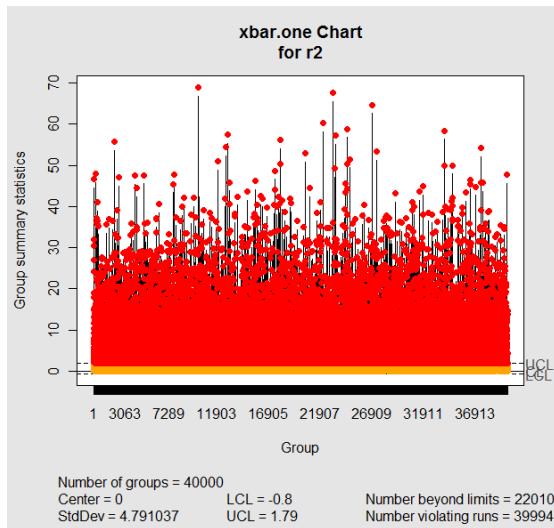
R2T



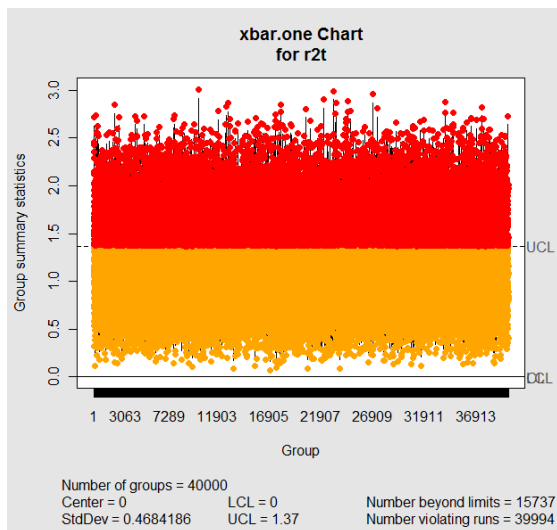
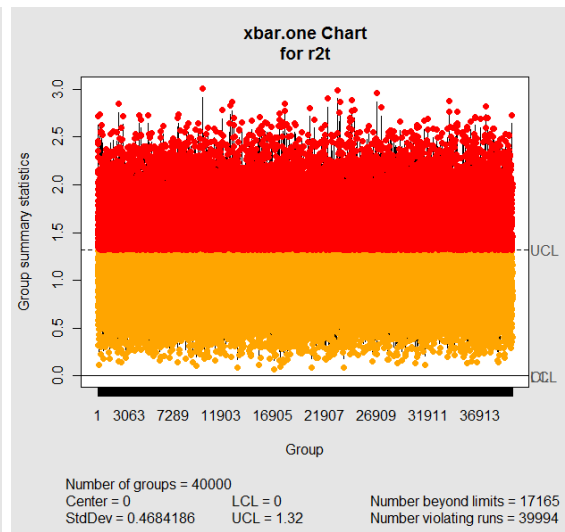
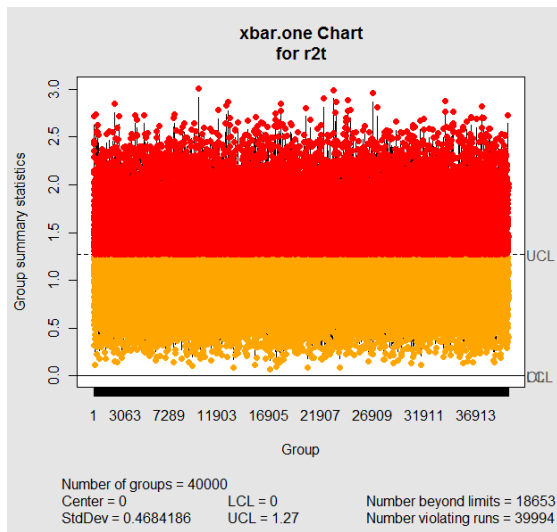


- $\sigma_e = 0,5$

R2

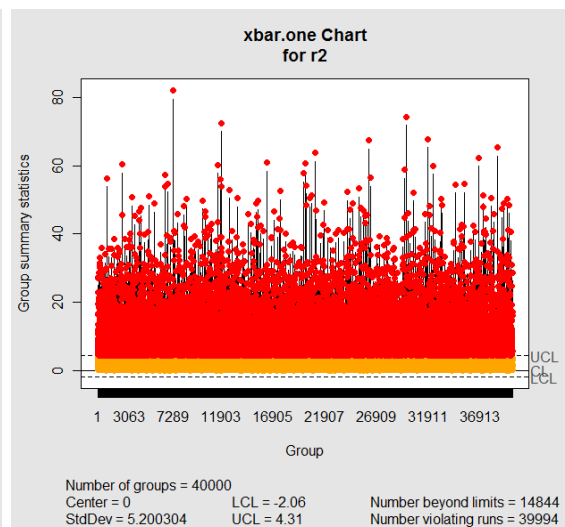
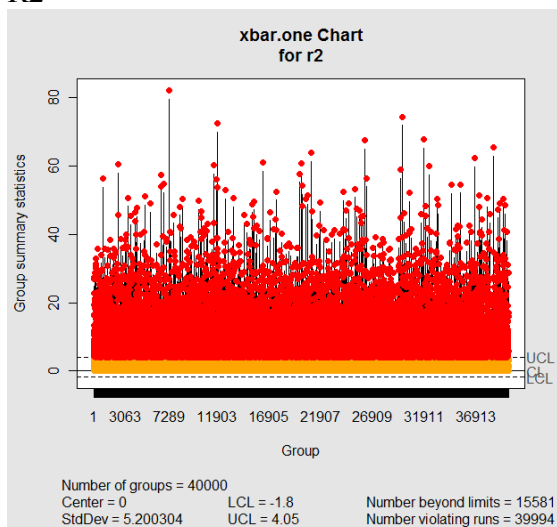


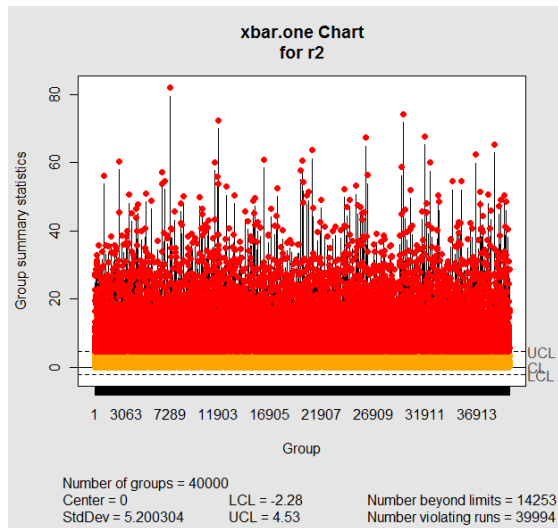
R2T



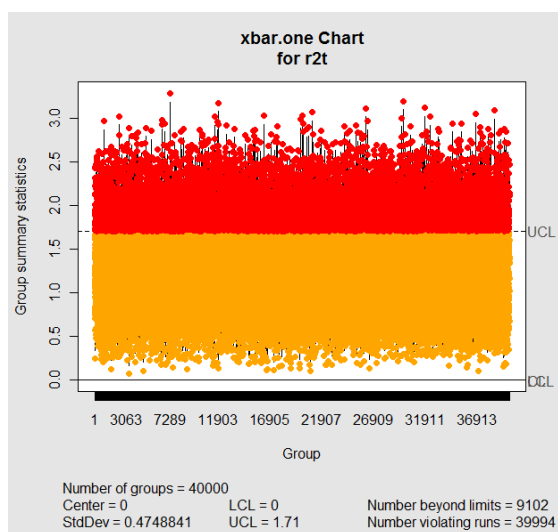
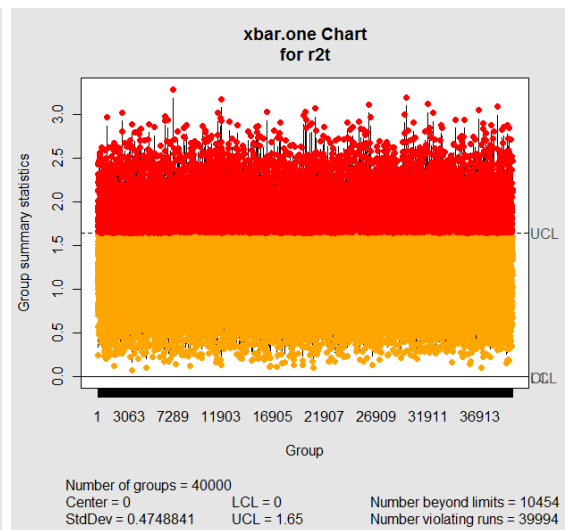
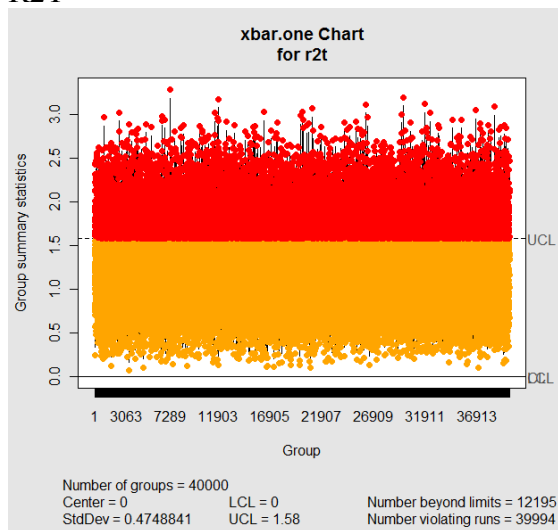
- $\sigma_e = 0,75$

R2



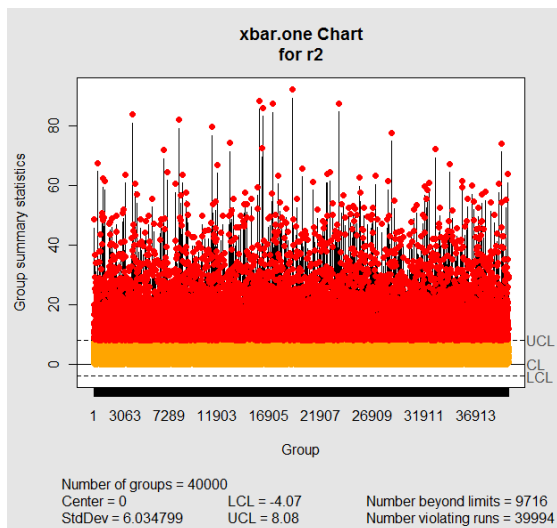
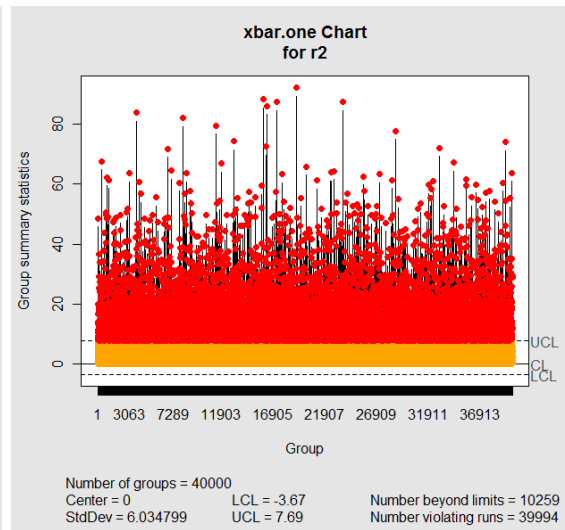
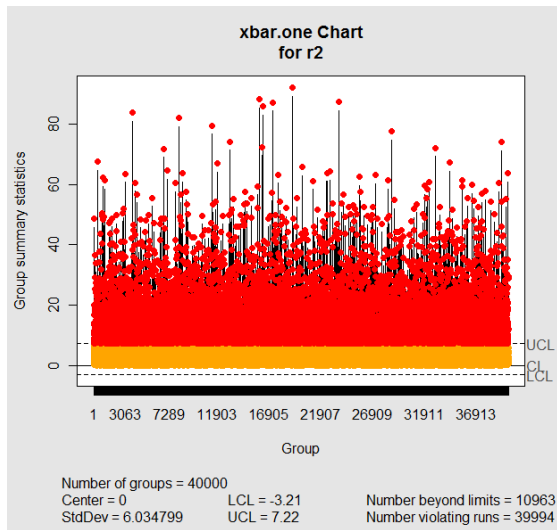


R2T

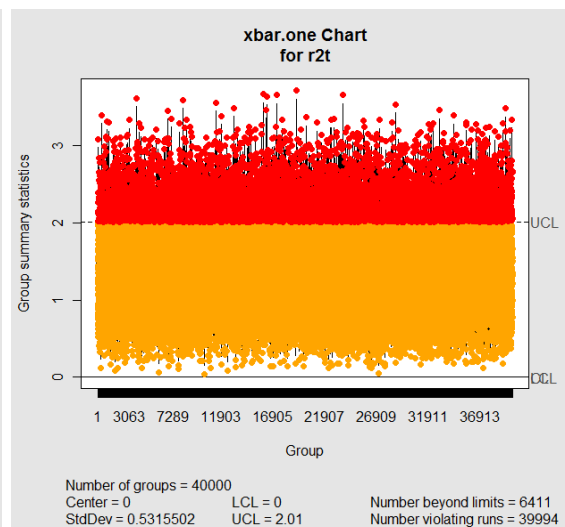
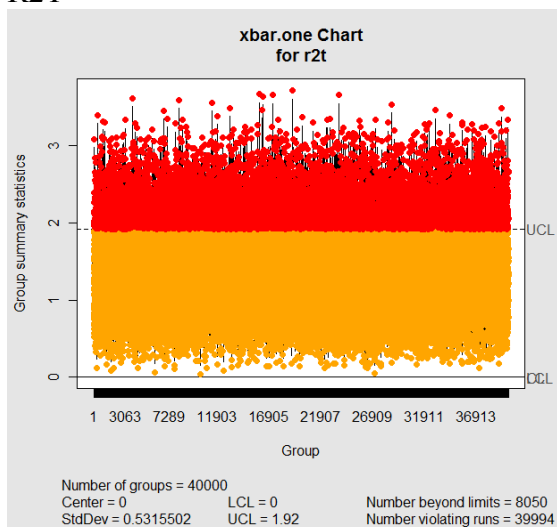


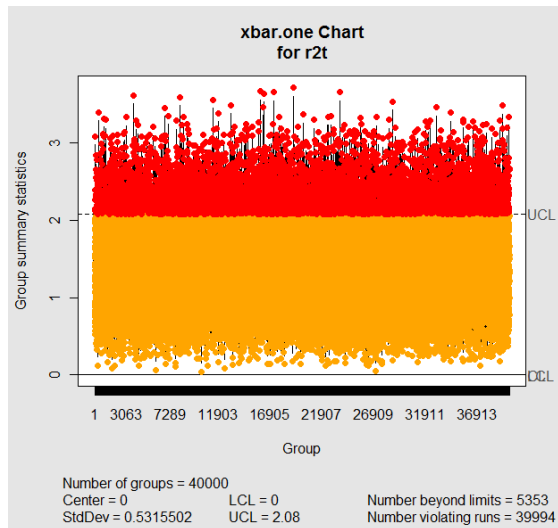
- $\sigma_e = 1,0$

R2



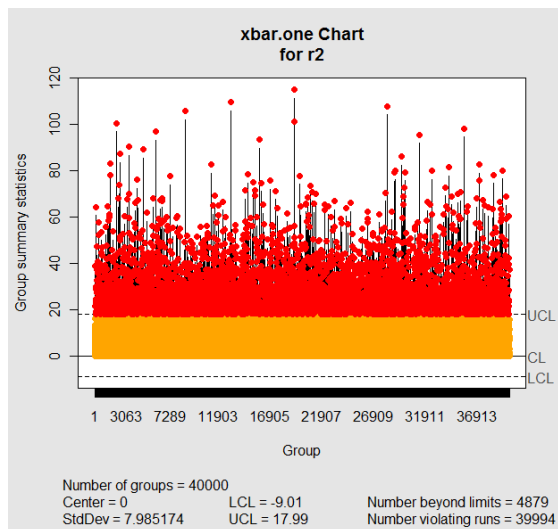
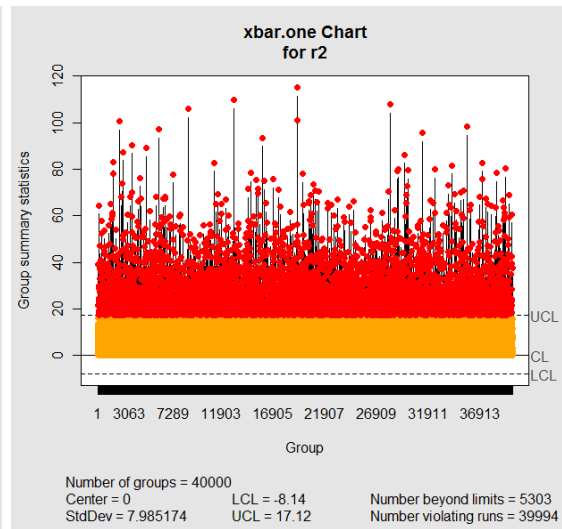
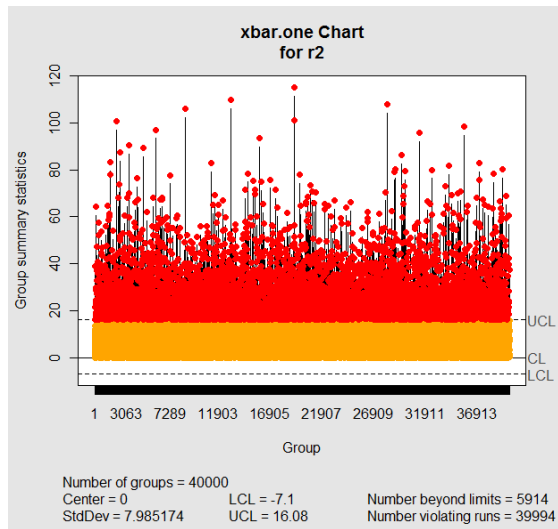
R2T



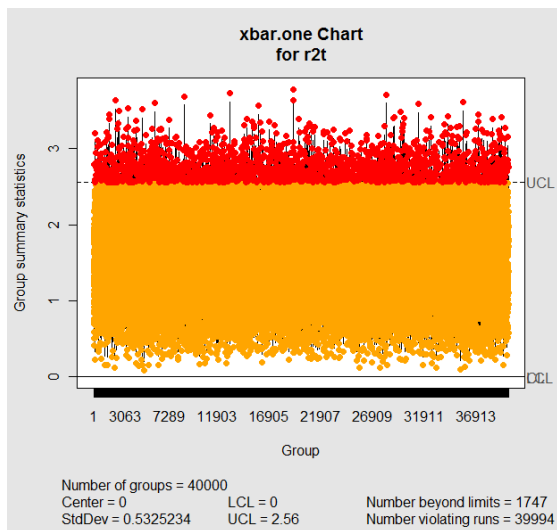
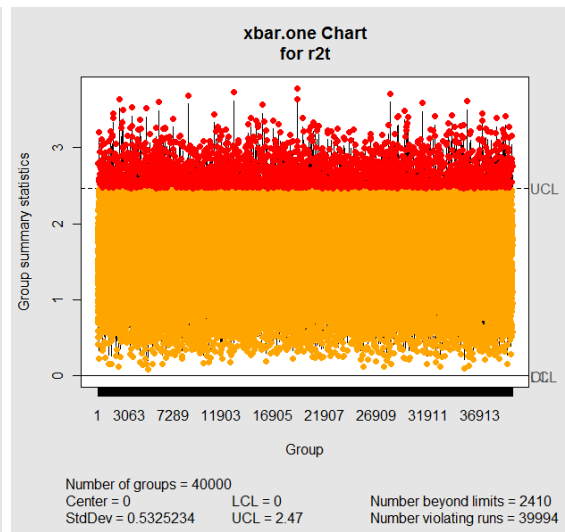
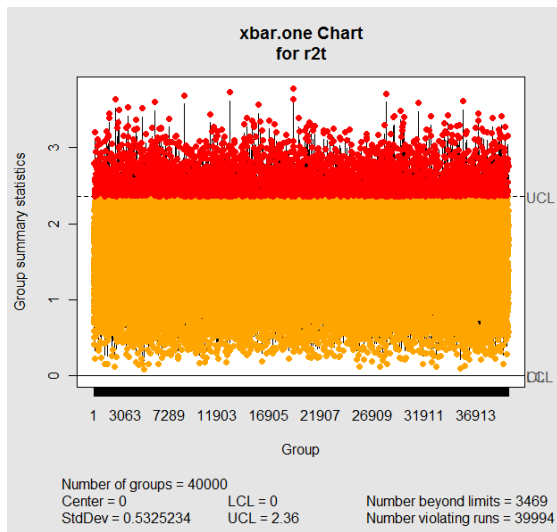


- $\sigma_e = 1,5$

R2



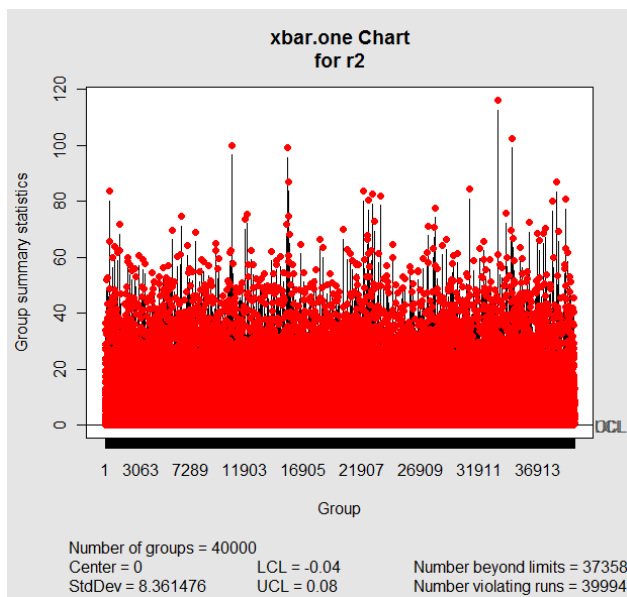
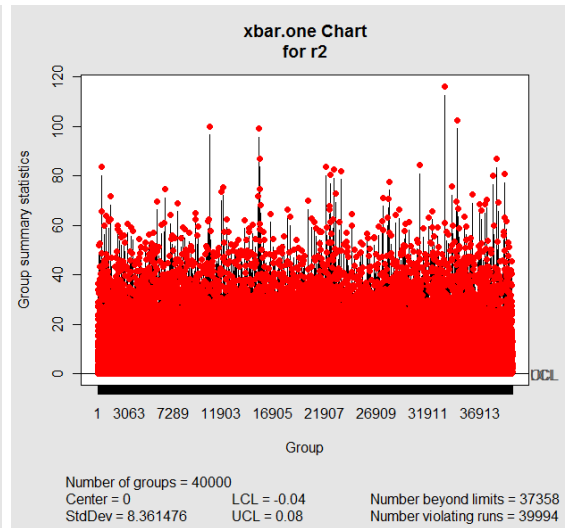
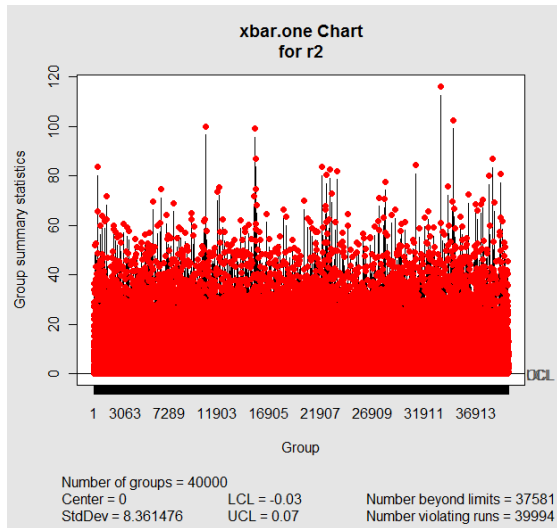
R2T



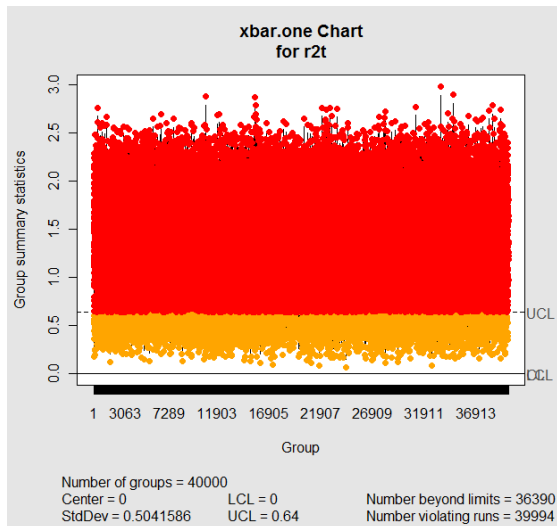
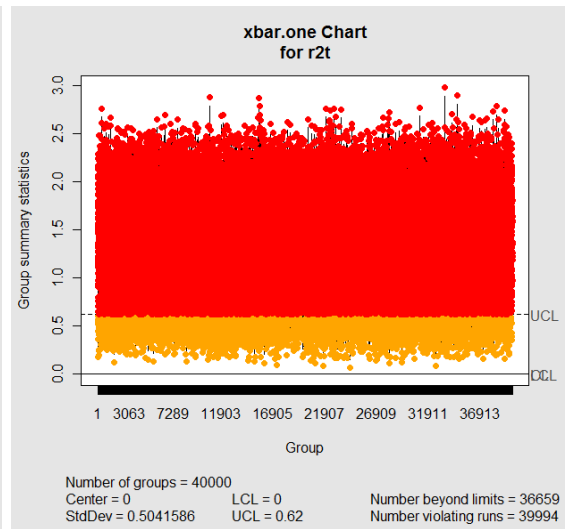
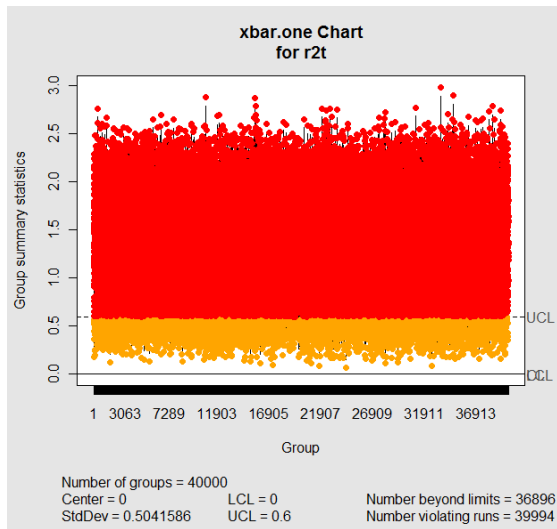
Fase II – $\alpha = 0,01$ e considerando a direção $C \sim N(1,1)$ com alteração na média.

- $\sigma_e = 0,1$

R2

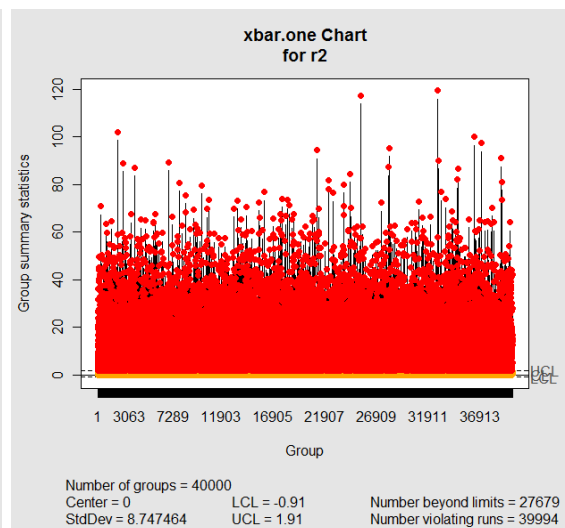
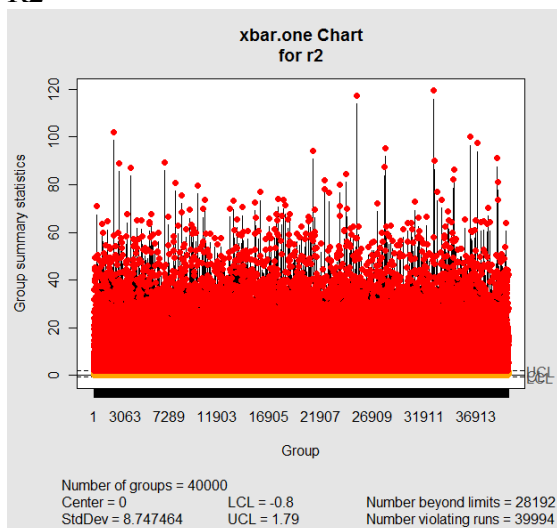


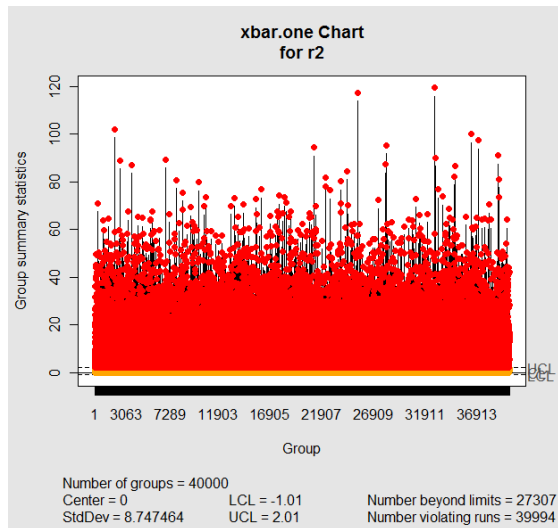
R2T



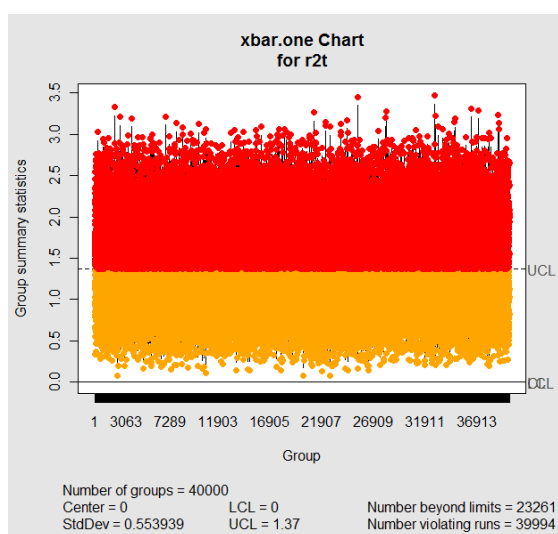
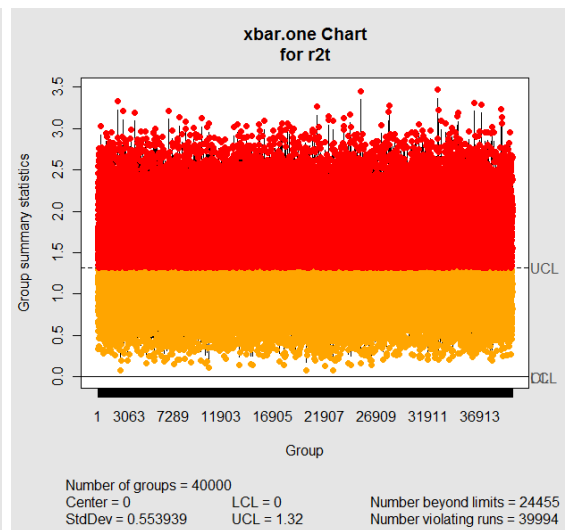
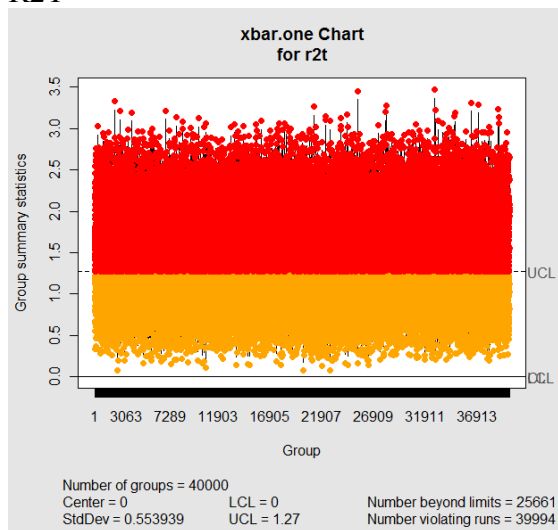
- $\sigma_e = 0,5$

R2



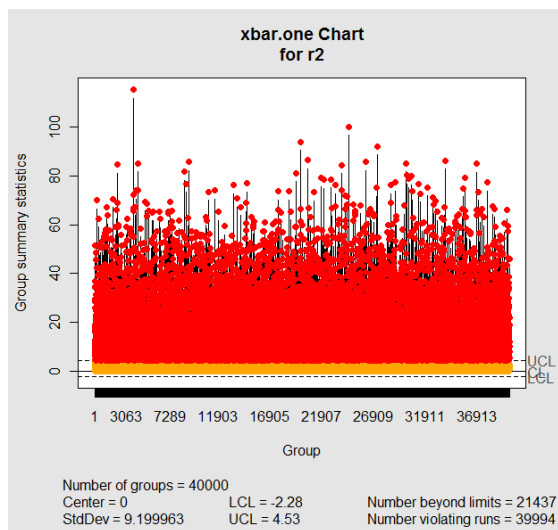
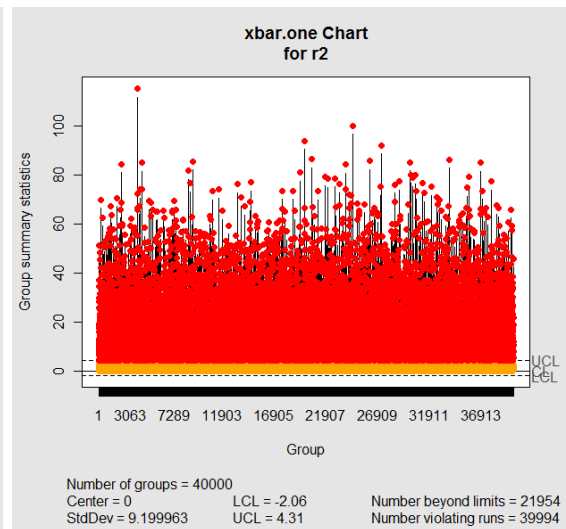
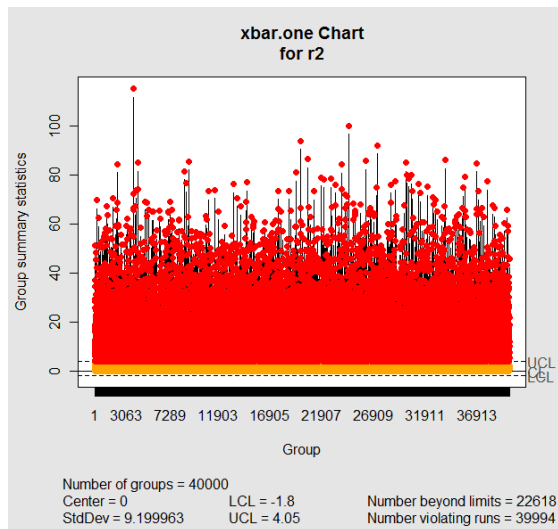


R2T

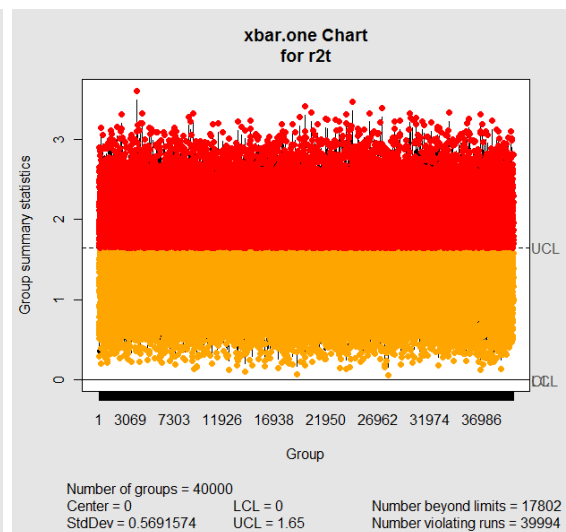
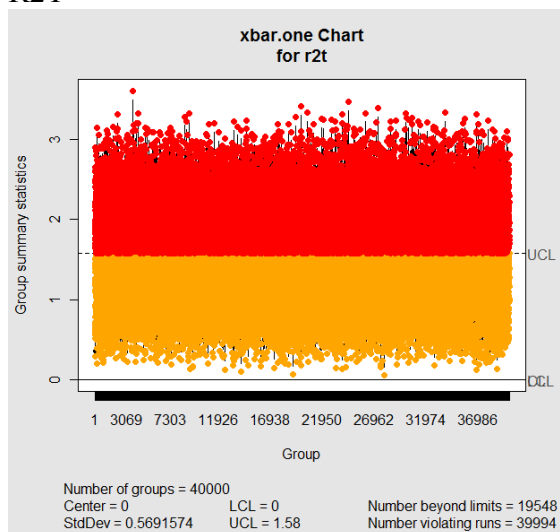


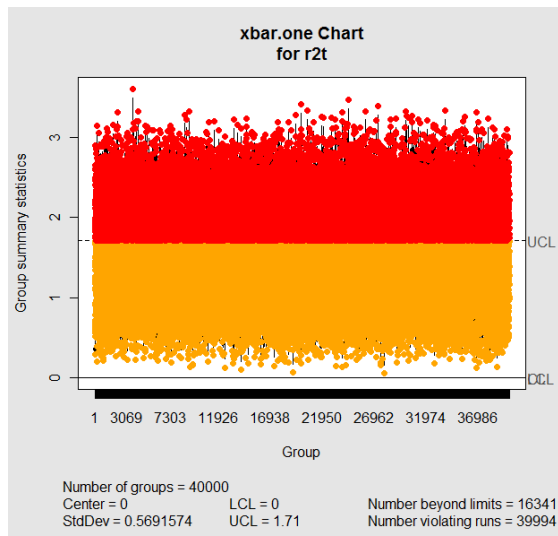
- $\sigma_e = 0,75$

R2



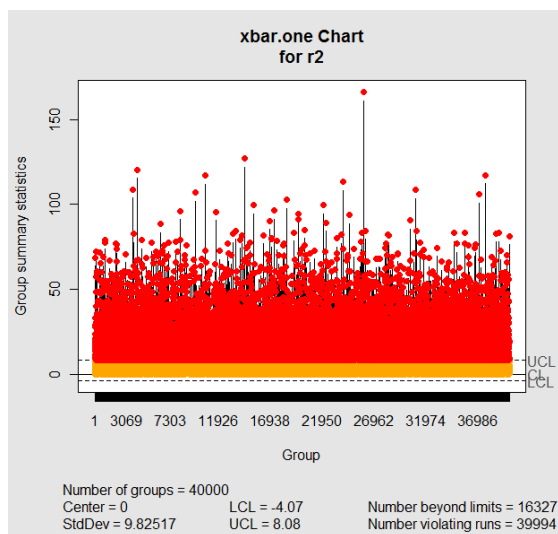
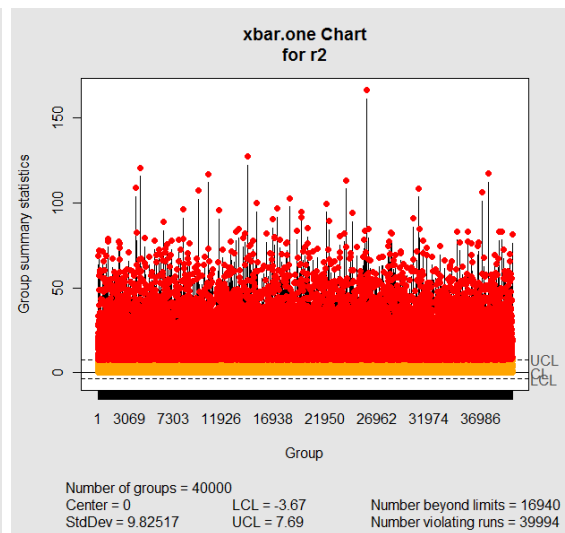
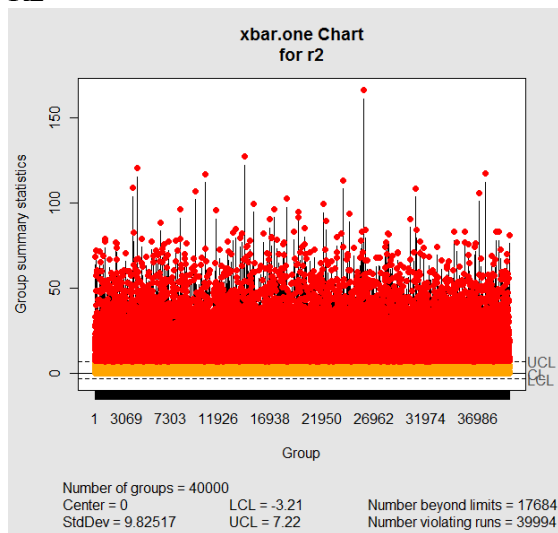
R2T



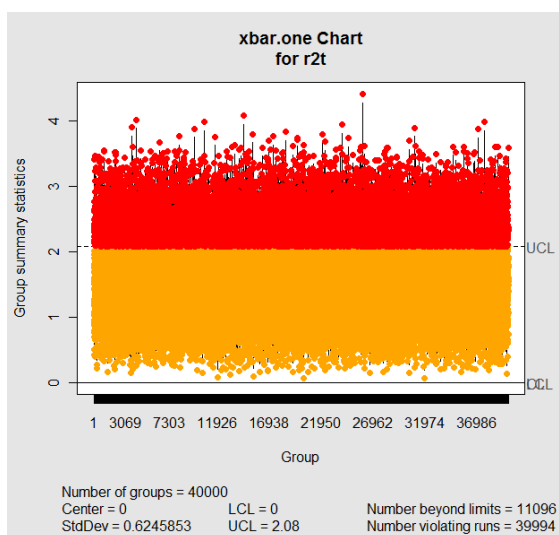
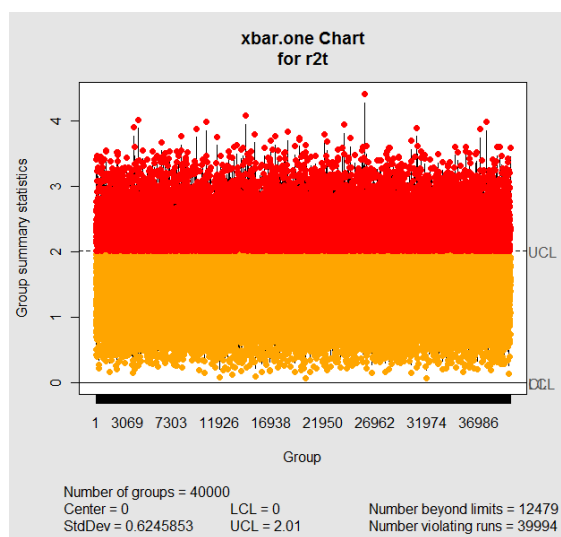
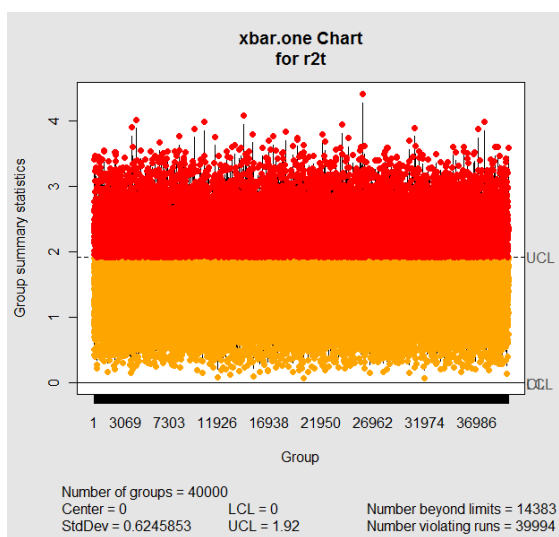


- $\sigma_e = 1,0$

R2

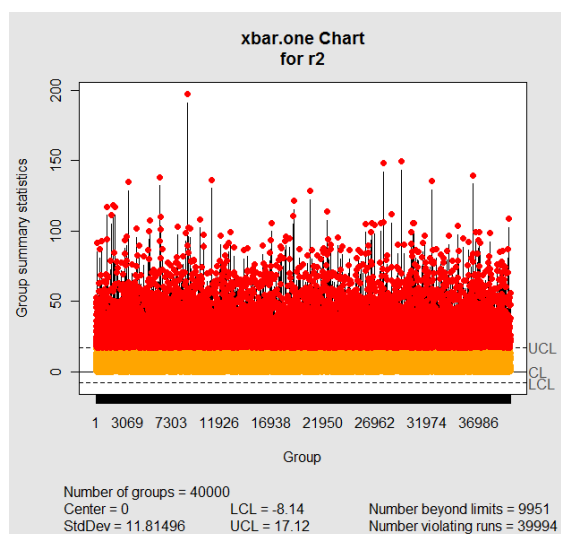
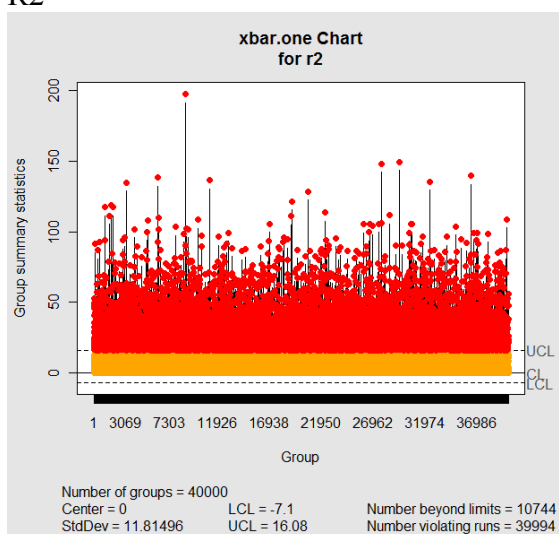


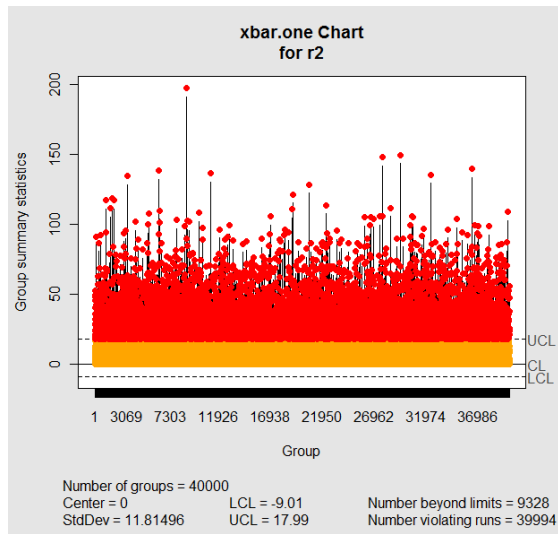
R2T



- $\sigma_e = 1,5$

R2





R2T

