

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

DAMARIS SERIGATTO VICENTIN

**PROPOSTA DE USO DE DISTRIBUIÇÕES FINITAS DE
PROBABILIDADE E DISTRIBUIÇÃO T^2 -HOTELLING PARA
O MONITORAMENTO DE PROCESSOS DE MÚLTIPLOS
FLUXOS**

TESE DE DOUTORADO

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

**PROPOSTA DE USO DE DISTRIBUIÇÕES FINITAS DE
PROBABILIDADE E DISTRIBUIÇÃO T^2 -HOTELLING PARA
O MONITORAMENTO DE PROCESSOS DE MÚLTIPLOS
FLUXOS**

DAMARIS SERIGATTO VICENTIN

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal de São Carlos, como parte dos requisitos para o exame de defesa.

Área de concentração: Gestão de Qualidade

ORIENTADOR: PROF. DR. PEDRO CARLOS OPRIME

São Carlos
2021



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Exatas e de Tecnologia
Programa de Pós-Graduação em Engenharia de Produção

Folha de Aprovação

Defesa de Tese de Doutorado da candidata Damaris Serigatto Vicentin, realizada em 09/04/2021.

Comissão Julgadora:

Prof. Dr. Pedro Carlos Oprime (UFSCar)

Profa. Dra. Fabiane Letícia Lizarelli (UFSCar)

Prof. Dr. Ricardo Coser Mergulhão (UFSCar)

Prof. Dr. Mario Orestes Aguirre González (UFRN)

Prof. Dr. Gladys Dorotea Cacsire Barriga (UNESP)

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa de Pós-Graduação em Engenharia de Produção.

*Dedico este trabalho a minha filha ISABELE
SERIGATTO VICENTIN POSSATO*

AGRADECIMENTOS

Ao Professor Dr. Pedro Carlos Oprime, pelo cuidado, paciência, ensinamentos, orientação e competência.

Aos Professores Dra. Gladys Dorotea Cacsire Barriga, Dr. Mario Orestes Aguirre González, Dr. Ricardo Coser Mergulhão, Dra. Fabiane Letícia Lizarelli, Dr. Manoel Henrique Salgado, Dr. Andrei Aparecido de Albuquerque, Dr. Herick Morales, Dra. Geeta Dulpati pela participação e colaboração com minha formação acadêmica.

Aos funcionários do Departamento de Pós Graduação e do Departamento de Engenharia de Produção pelo auxílio. Sou grata também a todos os Professores do DEP que contribuíram de diferentes formas para minha formação.

Aos amigos de curso pelas contribuições, bons momentos e conhecimentos compartilhados durante o estudo.

A minha família, em especial minha filha Isabele, minha avó Therezinha, meu irmão Daniel e minha mãe Rosimeire, meu companheiro Daniel pela paciência, torcida, carinho e injeções de ânimo.

Ao CAPES pelo financiamento desta pesquisa.

RESUMO

Um tema ainda pouco pesquisado é o monitoramento e controle de processos com múltiplos fluxos (*Multiple Stream Process – MSP*). São casos nos quais os equipamentos possuem vários módulos e cada módulo produz as mesmas peças, componentes e materiais com diferentes parâmetros de médias e desvios padrão, constituindo-se em subpopulações. Os métodos existentes para o monitoramento com MSP estão baseados principalmente na detecção de mudanças em um único fluxo ou um deslocamento simultâneo em todos os fluxos. Dentro desse escopo, o objetivo desta tese foi propor um modelo de gráfico de controle de mistura finita para o monitoramento de MSP, ainda pouco encontrado na literatura e compará-lo com o gráfico de controle T^2 -Hotelling. Para atingir o objetivo pretendido, foi utilizada a abordagem quantitativa e a qualitativa, por meio da aplicação em um caso ilustrativo com dados reais, visto que os dados foram coletados durante a produção de massa de pastel de uma indústria alimentícia. Foi realizada uma análise comparativa quanto aos aspectos estatísticos e operacionais entre os gráficos de controle de mistura finita e T^2 -Hotelling. Observou-se que as abordagens se diferenciam em termos dos seus pressupostos, como na sua complexidade matemática e operacional. Em termos de desempenho, os gráficos de T^2 - Hotelling foram melhores na detecção de problemas (causas especiais do processo) e maior facilidade de aplicação por sua menor complexidade matemática. Já o gráfico de mistura finita é inovador, é aplicável, considera todos os fluxos e as variabilidades entre eles em um único gráfico de controle, possui maior facilidade quanto ao procedimento de amostragem, porém traz maior complexidade matemática para sua aplicação.

Palavras-chave: Mistura de distribuições, Gráficos de controle, Processos de múltiplos fluxos; Gráfico de T^2 - Hotelling, Setor Alimentício.

ABSTRACT

A topic still little researched is the monitoring and control of multiple stream processes (Multiple Stream Process - MSP). These are cases in which the equipment has several modules and each module produces the same parts, components and materials with different parameters of means and standard deviations, constituting subpopulations. Within this scope, the objective of this thesis was to propose a finite mixture control chart model for monitoring MSP, which is still rarely found in the literature, and to compare it with the control chart T^2 -Hotelling. To achieve the intended objective, was used the quantitative and qualitative approach based on modelling, through the application in an illustrative case with real data collected during the production of pastry dough from a food industry. A comparative analysis was performed regarding the statistical and operational aspects between the finite mixture control charts and T^2 -Hotelling. It was observed that the approaches differ in terms of their assumptions, as in their mathematical and operational complexity. In terms of performance, the T^2 -Hotelling charts were better at detecting problems (special causes of the process) and easier to apply due to their lower mathematical complexity. On the other hand, the finite mixture graph is innovative, it is applicable, it considers all the flows and the variability between them in a single control graph, it is easier in terms of the sampling procedure, but it brings greater mathematical complexity to its application.

Keywords: Mixture distributions, Control chart, Multiple-Stream Processes, T^2 - Hotelling chart, Alimentary sector.

LISTA DE FIGURAS

Figura 1.1 - Estrutura da tese.....	24
Figura 3.1 - Etapas do desenvolvimento da pesquisa.....	61
Figura 3.2 - Etapas do desenvolvimento do gráfico de controle de mistura finita.....	65
Figura 3.3 - Etapas do desenvolvimento do gráfico de controle de T^2 -Hotelling.....	66
Figura 4.1 - Caracterização de métodos para coleta de dados de processo de múltiplos fluxos para construção de gráficos de controle. Fonte: Grupo de Ação da Indústria Automotiva (AIAG, 1991).....	68
Figura 4.2 - Plano de produção do gráfico de múltiplos fluxos proposto.....	69
Figura 4.3 – Fluxograma dos modelos propostos para MSP.....	78
Figura 5.1 - Esquema da produção do processo de linha de massa folhada.....	79
Figura 5.2 - Procedimento amostral aplicado para a abordagem de T^2 -Hotelling.....	80
Figura 5.3 - Etapas para o uso do Gráfico de Mistura Finita.....	82
Figura 5.4 - Caracterização da distribuição de peso em quilogramas.....	82
Figura 5.5 - Plotando o gráfico normal das distribuições de probabilidade para o banco de dados da Figura 5.4.....	83
Figura 5.6 - Desvios esperados dos escores na distribuição normal padrão para o banco de dados da Figura 5.4.....	83
Figura 5.7 - Sentido horário da esquerda para a direita, plotando os escores normais da posição dos quartis de 1 a 4, considerando quatro subpopulações.....	84
Figura 5.8 - Etapas para o uso do Gráfico de Controle de T^2 - Hotelling.....	90
Figura 5.9 - Gráfico de Scree-plot da amostra mestre.....	92
Figura 5.10 - Gráfico de Controle de T^2 - Hotelling da amostra mestre.....	93
Figura 5.11 - Gráfico de Scree-plot amostra expandida.....	95
Figura 5.12 - Gráfico de controle T^2 -Hotelling da amostra expandida.....	96
Figura 6.1 - Análise comparativa entre gráfico de controle de mistura para ajuste específico [-0,5, 0,5, 0,0] ($K = 3,59$) e parâmetro K tradicional de Shewhart \bar{X}	106
Figura 6.2 - Fluxograma da construção do gráfico de controle de distribuição de probabilidade mistura finita.....	107
Figura 6.3 - Fluxograma da construção do gráfico de controle T^2 - Hotelling.....	107
Figura 6.4 - Análise comparativa entre gráfico de controle T^2 - Hotelling para ajuste específico $p=[0,00135$ a $0,09]$ e de mistura para $K=3,59$	109

LISTA DE TABELAS

Tabela 5.1- Média e desvio padrão para cada posição e as correlações obtidas entre elas.....	80
Tabela 5.2 - Reprodução 1000 amostras aplicando o mecanismo de interdependência entre os fluxos mostrado na Figura 5.2.....	81
Tabela 5.3 - Parâmetros estimados das subpopulações.....	85
Tabela 5.4 - Parâmetros estimados de π e δ	85
Tabela 5.5 - Estimativas dos parâmetros gerados pelo Maple.....	87
Tabela 5.6 - Valores de AARL para diferentes valores de t (quando $t \neq 0$, fora de controle).....	88
Tabela 5.7 - Estatística Descritiva e matriz de correlação da amostra mestre.....	90
Tabela 5.8 - Variância explicada da amostra mestre.....	91
Tabela 5.9 - Componentes Principais e suas Cargas Fatoriais em relação as posições no cilindro da amostra mestre.....	92
Tabela 5.10 - Estatística Descritiva da Fase II.....	94
Tabela 5.11 - Variância explicada da amostra expandida.....	94
Tabela 5.12 - Componentes Principais e suas Cargas Fatoriais em relação as posições no cilindro da amostra mestre.....	95
Tabela 5.13 - Medias finais das posições do cilindro em relação aos t da amostra expandida.....	96
Tabela 5.14 - Comparação dos números de desvios-padrão, t , para análise do desempenho dos gráficos de controle T^2 em termos do AARL, com amostra expandida para $\alpha_{Nominal} = 0,135\%$	97
Tabela 5.15 - Comparação dos números de desvios-padrão, t , para análise do desempenho dos gráficos de controle T^2 em termos do AARL, com amostra expandida para $\alpha_{Nominal} = 0,27\%$	98
Tabela 5.16 - Comparação dos números de desvios-padrão, t , para análise do desempenho dos gráficos de controle T^2 em termos do AARL, com amostra expandida para $\alpha_{Nominal} = 2,00\%$	99
Tabela 5.17 - Comparação dos números de desvios-padrão, t , para análise do desempenho dos gráficos de controle T^2 em termos do AARL, com amostra expandida para $\alpha_{Nominal} = 2,25\%$	99

Tabela 5.18 - Comparação dos números de desvios-padrão, t , para análise do desempenho dos gráficos de controle T^2 em termos do AARL, com amostra expandida para $\alpha_{Nominal} = 2,50\%$	100
Tabela 5.19 - Comparação dos números de desvios-padrão, t , para análise do desempenho dos gráficos de controle T^2 em termos do AARL, com amostra expandida para $\alpha_{Nominal} = 8,00\%$	101
Tabela 5.20 - Comparação dos números de desvios-padrão, t , para análise do desempenho dos gráficos de controle T^2 em termos do AARL, com amostra expandida para $\alpha_{Nominal} = 9,00\%$	102
Tabela 5.21 - Valores de AARL para diferentes valores de t e p ($\alpha_{Nominal}$).....	102
Tabela 6.1 - Desempenho do AARL para diferentes conjuntos de t para $K = 3,00$ e $K = 3,59$	105
Tabela 6.2 - Desempenho do AARL para diferentes conjuntos de t e p , sendo $K = 3,59$ para gráfico de controle de mistura finita, e $p=[0,00135$ a $0,09]$ para gráfico de controle T^2 -Hotelling.....	108

LISTA DE QUADROS

Quadro 3.1 - Classificação da pesquisa.....	60
---	----

LISTA DE ABREVIATURAS

ACP	Análise de Componentes Principais
AARL	<i>Average of Average Run Length</i> (Média do Comprimento de Corrida Média)
ARL	<i>Average Run Length</i> (Comprimento de Corrida Média)
CEP	Controle Estatístico de Processo
CUSUM	<i>Cumulative Sum Control Chart</i>
EWMA	<i>Exponentially Weighted Moving Average</i>
ISO	<i>International Standard Organization</i>
LC	Limite Central
LIC	Limite Inferior de Controle
LIE	Limite Inferior de Especificação
LM	Limite Médio ou Linha Média
LSC	Limite Superior de Controle
LSE	Limite Superior de Especificação
MEP	Monitoramento Estatístico de Processo
MMV	Métodos Multivariados
MSP	<i>Multiple Stream Process</i>
PCA	<i>Principal Component Analysis</i>
p.d.f.	Função de Distribuição de Probabilidade
PMF	Processo de Múltiplos Fluxos
SDARL	<i>Standard Deviation of the Average Run Length</i> (Desvio Padrão do Comprimento Médio de Corrida)
SPM	<i>Statistical Process Monitoring</i>

SUMÁRIO

1 INTRODUÇÃO.....	15
1.1 PROBLEMA E QUESTÃO DE PESQUISA.....	18
1.2 OBJETIVO.....	20
1.2.1 Objetivo Geral.....	20
1.2.2 Objetivos Específicos.....	21
1.3 JUSTIFICATIVA.....	21
1.4 ESTRUTURA DA TESE.....	23
2 REVISÃO BIBLIOGRÁFICA.....	25
2.1 FUNDAMENTOS DO CONTROLE E MONITORAMENTO ESTATÍSTICO DO PROCESSO.....	25
2.1.1 Característica da Qualidade e Valor Alvo.....	29
2.1.2 Causas Comuns e Especiais.....	31
2.1.3 Pressupostos para o Monitoramento Estatístico do Processo.....	33
2.1.4 Diferenças entre Controle Estatístico de Processo e Monitoramento Estatístico de Processo.....	35
2.2 RECENTES PESQUISAS SOBRE MONITORAMENTO ESTATÍSTICO DE PROCESSO.....	36
2.3 MONITORAMENTO DE PROCESSO DE MÚLTIPLOS FLUXOS.....	41
2.4 PROPOSTAS PARA MONITORAMENTO DE MSP.....	42
2.4.1 Distribuição de Probabilidade de Mistura Finita.....	43
2.4.2 Análise de Componentes Principais e T^2 -Hotelling.....	45
2.4.2.1 Análise de Componentes Principais (ACP).....	47
2.4.2.2 Gráfico de controle T^2 -Hotelling.....	51
3 MÉTODO DE PESQUISA.....	55
3.1 ABORDAGEM METODOLÓGICA.....	55
3.2 MÉTODOS DE PESQUISA.....	57
3.3 PROCEDIMENTO PARA O DESENVOLVIMENTO DA PESQUISA.....	58
3.4 ETAPAS DO DESENVOLVIMENTO DA PESQUISA.....	60
4 MODELAGEM DE GRÁFICOS DE CONTROLE COM DISTRIBUIÇÕES DE PROBABILIDADE DE MISTURA FINITA.....	68
4.1 CASO COM PARÂMETROS ESTATÍSTICOS, μ, CONHECIDO.....	70
4.2 CASO COM μ DESCONHECIDO.....	74
4.3 SÍNTESE DAS DUAS PROPOSTAS.....	76

5 APLICAÇÃO DOS MODELOS PROPOSTOS EM UM CASO ILUSTRATIVO COM DADOS REAIS	79
5.1 MODELO DE DISTRIBUIÇÃO DE PROBABILIDADE DE MISTURA FINITA	81
5.1.1 Determinação dos Limites de Controle	85
5.2 PROCEDIMENTO PARA O CASO DE ACP E T²-HOTELLING	89
5.2.1 Fase I – Amostra Mestre.....	90
5.2.2 Fase II – Dados Expandidos.....	93
5.2.3 Desempenho do Gráfico de T ² Hotelling Proposto.....	95
6 ANÁLISE COMPARATIVA DAS PROPOSTAS DE GRÁFICOS DE PARA PROCESSO DE MÚLTIPLOS FLUXO.....	103
6.1 ANÁLISE COMPARATIVA DO GRÁFICO DE CONTROLE DE MISTURAS COM O PARÂMETRO K DO GRÁFICO DE CONTROLE TRADICIONAL	103
6.2 ANÁLISE COMPARATIVA DO GRÁFICO DE CONTROLE DE MISTURAS COM O GRÁFICO DE CONTROLE T²-HOTELLING.....	106
6.3 ANÁLISE GERAL	110
7 CONCLUSÕES	113
REFERÊNCIAS.....	116
APÊNDICES	131
APÊNDICE 1 – REVISÃO SISTEMÁTICA DE LITERATURA PARA MÚLTIPLOS FLUXOS...131	
APÊNDICE 2 – TABELA DE LIMITE SUPERIOR DE CONTROLE PARA OS GRÁFICOS DE CONTROLE T²-HOTELLING.....149	
APÊNDICE 3 – ARTIGO COM PARTE DOS RESULTADOS DA TESE PUBLICADO	150

1. INTRODUÇÃO

O controle estatístico de processo (CEP) é um método amplamente utilizado em muitos setores industriais com o propósito de monitorar, diagnosticar e promover melhorias (WOODALL, 1985; PHALADIGANON *et al.*, 2013; VARIYATH; VATTATHOOR, 2014; AHANGAR; CHIMKA, 2016). Dentre as técnicas e ferramentas do CEP, a mais popular é o gráfico de controle utilizado para o monitoramento de características críticas de produtos e processos (CASTILLO; MONTGOMERY, 1994; JENSEN *et al.*, 2006; ASLAM *et al.*, 2014; WOODALL; MONTGOMERY, 2014). Segundo Woodall e Montgomery (2014), o objetivo principal do CEP, também denominado pelos autores como monitoramento estatístico de processos (*Statistical Process Monitoring – SPM*), é entender sua variação, avaliar seu estado atual, e agir para mantê-los sob controle estatístico.

Há vários tipos de gráficos de controle desenvolvidos. O primeiro a surgir foi o gráfico de Shewhart, em 1928. Posteriormente, surgiram os gráficos: de CUSUM (*Cumulative Sum Control Chart* - Gráfico de controle da soma cumulativa), o da média móvel ponderada exponencialmente (*Exponentially Weighted Moving Average – EWMA*) e os gráficos de atributos, p e c (MONTGOMERY, 2013). Todos esses gráficos de controle apresentam a mesma construção básica em que uma estatística de plotagem, que equivale a uma estatística de teste, calculada a partir de observações coletadas do processo monitorado, é comparada com limites de controle. Se a magnitude da estatística de plotagem for maior que a magnitude do (s) limite (s) de controle, sugere que o processo pode estar fora de controle.

Segundo os autores Jirasetpong e Rojanarowan (2011), Montgomery (2013), Woodall e Montgomery (2014) e Epprecht (2015) uma aplicação especial do gráfico de controle, ainda pouco estudada, é o monitoramento de processos de múltiplos fluxos. São casos nos quais os equipamentos possuem vários módulos e cada módulo produz os mesmos itens,

porém com diferentes parâmetros de médias e desvios padrão, constituindo-se em subpopulações. Espera-se que nesses casos os múltiplos fluxos, como os que foram estudados por Epretch *et al.* (2011), Bersimis *et al.* (2007), Liu *et al.* (2008) e Brown e Shaffer (2020), sejam presumivelmente idênticos, constituindo-se em uma única população, com valores-alvo idênticos. A literatura denomina esses processos de "*Multiple Stream Process*" (MSP).

Outra definição de MSP é dada por Eprecht *et al.* (2011), que o define como um processo que gera vários fluxos de saída, pois a variável de qualidade e suas especificações são as mesmas em todos os fluxos. Segundo Woodall e Montgomery (2014), existem métodos para monitorar MSP que se baseiam na detecção de mudanças em um único fluxo ou em um deslocamento simultâneo em todos os fluxos.

Eprecht (2015) fez um mapeamento das pesquisas sobre monitoramento de múltiplos fluxos e foi concluído que a literatura sobre o tema é escassa, com poucos avanços desde 1950 e mesmo ressurgindo a partir de meados da década de 1990. O autor também apresentou as diferenças essenciais nos modelos subjacentes de trabalhos antes e depois de 1995 e apontou questões para pesquisas futuras. Nesse mapeamento foi destacada a primeira publicação sobre MSP que aparece no ano de 1950, com o artigo de Boyd (1950). Em que é proposto um gráfico de controle de Shewhart para monitorar a média e a variação de um MSP, calculando a média da amostra para cada fluxo, considerando o valor mínimo e máximo nas amostras como estatística de plotagem. Também foi destacada a pesquisa de Mortell e Runger (1995) em que é apontada a ineficiência que pode surgir no gráfico de Boyd e propõem dois esquemas de gráficos. O primeiro para monitorar a variabilidade entre os fluxos, que usa uma variante do gráfico de Shewhart. O segundo para monitorar a variabilidade de um fluxo em relação aos outros, utilizando as estruturas dos gráficos de Shewhart, e também a dos gráficos de EWMA e CUSUM, para monitorar a amplitude máxima de todas as médias das amostras dos fluxos.

Outra proposta inicialmente evidenciada por Runger e Montgomery (1996), propõe gráficos de monitoramento de MSP multivariados usando a estatística de T^2 de Hotelling para detectar mudanças em um subconjunto do total de fluxos. O gráfico de controle T^2 de Hotelling é um método estatístico utilizado para o monitoramento de processo com mais de uma característica de qualidade ou para múltiplos fluxos interdependentes. Para tanto, os autores utilizaram os gráficos de controle de CUSUM e EWMA. Montgomery (2013) acrescenta que se todos os fluxos estiverem altamente correlacionados, apenas um fluxo precisará ser monitorado. Assim, para o monitoramento de MSP, os gráficos de controle multivariado enfatizaram vantagens em termos de desempenho de detectar se o processo está fora de controle, especificamente em processo que há muitas variáveis e com alta interdependência entre os fluxos (BROWN; SHAFFER, 2020).

Trabalhos sobre o assunto foram publicados por Lanning *et al.* (2002), Liu *et al.* (2008) e Meneces *et al.* (2008) desenvolvendo novas alternativas para detectar alterações na média do processo de múltiplos fluxos. Epprecht *et al.* (2011) contribuíram com a área ao expandir a ideia sobre o gráfico de amplitude máxima de Mortell e Runger (1995) e propuseram o uso das derivações máxima e mínima de uma média geral a ser usada como estatística de plotagem, semelhante a Boyd (1950). Bersimis *et al.* (2007), por sua vez, propuseram um gráfico multivariado de CUSUM e EWMA para MSP. Meneces *et al.* (2008) sugerem a utilização de um gráfico de Shewhart para cada fluxo, pois essa técnica, para os autores, é mais eficaz para obtenção de informação sobre o processo.

Diferentemente das abordagens acima, Vicentin *et al.* (2018) desenvolveram pela primeira vez na literatura um gráfico de controle que utiliza de uma distribuição de probabilidade de mistura finita para MSP. Assim o gráfico de distribuição de probabilidade de mistura finita proposto pelos autores pode ser utilizado quando o processo a ser monitorado

possui múltiplos fluxos provenientes de várias distribuições normais de probabilidade. A abordagem proposta vem da mesma família paramétrica, mas considera parâmetros estatísticos diferentes em cada fluxo para monitoramento em somente um gráfico de controle. Os autores salientaram um esquema alternativo aos gráficos de Shewhart, porém com múltiplas distribuições de probabilidade que pode ser monitorado utilizando apenas um gráfico de controle para todos os fluxos do processo.

Recentemente, Brown e Shaffer (2020) desenvolveram um gráfico de controle não paramétrico para soma acumulada (CUSUM) usando dados de mediana provenientes de distribuição normal e não normal. Para os autores, na prática, assumir que as medianas sejam conhecidas pode ser uma suposição irracional.

1.1 PROBLEMA DE PESQUISA E QUESTÃO DE PESQUISA

Apesar dos trabalhos publicados como Boyd (1950), Mortell e Runger (1995), Lanning *et al.* (2002), Zhang e Weng (2005), Bersimis *et al.* (2007), Liu *et al.* (2008) e Meneces *et al.* (2008), Epprecht *et al.* (2011), Epprecht (2015), Vicentin *et al.* (2018) e Brown e Shaffer (2020) observa-se na literatura que ainda há poucos artigos que adotam uma abordagem empírica sobre MSP, o que chama a atenção, pois o tema é relevante; e mais, com o desenvolvimento dos sistemas de produção, incorporando novas tecnologias digitais, a coleta de dados não é mais uma barreira, o que pode facilitar o monitoramento de múltiplos fluxos.

Dos trabalhos publicados, duas abordagens são destacadas por sua aderência e relevância para monitoramento de MSP: a referente ao uso de distribuições de probabilidade de mistura finita, proposta por Vicentin *et al.* (2018) e a segunda o uso de gráficos de controle

multivariado, o T^2 -Hotelling, proposto por Bersimis *et al.* (2007) e Zhang e Weng (2005). As propostas têm suas limitações e pressupostos que ainda não foram abordados pela literatura.

O uso de distribuições de probabilidade de mistura finita pondera que não existam correlações entre os múltiplos fluxos, bem como possuam a mesma distribuição de probabilidade. A proposta de Vicentin *et al.* (2018) não aborda o problema da correlação entre os múltiplos fluxos, assume que as amostras de cada fluxo sejam independentes, seguindo a distribuição normal. A proposta do uso da T^2 -Hotelling, ao contrário, pressupõe a existência da correlação entre os fluxos, com distribuição normal multivariada. Entretanto, Bersimis *et al.* (2007) utilizam gráficos de controle CUSUM e EWMA, diferentemente da proposta pelo uso da Análise de Componentes Principais e T^2 -Hotelling. Já Zhang e Weng (2005) examinam as correlações no processo utilizando a Análise de Componentes Principais (ACP) em conjunto com a abordagem de distribuição de probabilidade de mistura, propondo gráficos de T^2 -Hotelling, porém os autores modelam a ACP com base na distribuição de probabilidade de mistura não contextualizando casos de múltiplos fluxos de processo e não propõem um gráfico de controle específico de distribuição de probabilidade de mistura.

Nesses casos, pouco são mencionados os aspectos operacionais das suas aplicações e as suposições necessárias para suas aplicações em situações reais. Subentende-se que, para os gráficos de controle acerca das distribuições com mistura finita, supõem-se não haver interdependência entre os fluxos, ou seja, não haver correlação entre fluxos. Em razão disso, o procedimento da amostra é diferente da análise multivariada, sendo amostras que extraídas ao final da linha não ponderando as posições de fabricação, constituindo-se de uma população única, e os limites de controle são estabelecidos tendo em vista a mistura, ou distribuição multimodal dos dados. Já para os gráficos de controle multivariado, o procedimento expansão da amostra é feito para cada fluxo; e o ideal é, nesses casos, que haja interdependência entre os

fluxos, ou seja, a correlação entre eles, de modo que um método multivariado possa ser aplicado. Observa-se que na literatura não são encontrados trabalhos que analisem essas duas abordagens para MSP, pois a abordagem de mistura finita para MSP foi publicada somente por Vicentin *et al.* (2018) e as técnicas multivariadas como T^2 -Hotelling e APC são consagradas na literatura em diversas áreas acadêmicas, não existindo nenhum trabalho em termos estatísticos e operacionais. Com isso, buscando as lacunas de pesquisa será realizada essas análises comparativas na referente tese.

Tendo em vista esses aspectos teóricos, uma questão de pesquisa é qual dessas duas abordagens, mistura finita e multivariada, têm melhor desempenho em termos dos erros tipo I e II (gerar menor falso alarme) e qual delas é mais fácil de ser implementada. Assim, para responder a essa pergunta, pretende-se comparar a proposta dos tipos de gráfico de controle, de mistura finita e multivariada, e analisar os aspectos positivos e negativos de cada um deles.

1.2 OBJETIVO

Apesar da aplicação de métodos estatísticos ser reconhecida como uma importante ferramenta que possibilita monitorar e controlar processos, existem ainda situações em que esses métodos precisam ser aprimorados, é o que ocorre com os processos de múltiplos fluxos. Assim, o objetivo da tese pode ser dividido em objetivo geral e objetivos específicos, que serão salientados a seguir.

1.2.1 Objetivo Geral

Levando em conta os aspectos teóricos delineados na introdução desse trabalho, como também a falta de trabalhos que utilizem distribuição de probabilidade de mistura finita

nos gráficos de controle, apenas Vicentin et al. (2018), o objetivo geral deste trabalho é desenvolver um único gráfico de controle considerando distribuição de probabilidade de mistura finita para o monitoramento de MSP e comparar empiricamente o seu desempenho, em termos do AARL, com o gráfico de controle de T^2 - Hotelling.

1.2.2 Objetivos Específicos

A partir do objetivo geral, os objetivos específicos são seguintes:

1. Desenvolver o gráfico de controle para processo de múltiplos fluxos com característica de qualidade que possua distribuições de probabilidade de mistura finita baseada no artigo Vicentin *et al.* (2018) (apresentadas no Capítulo 4);
2. Aplicar os tipos de gráficos de controle de mistura finita e T^2 - Hotelling em um caso ilustrativo com dados reais (apresentadas no Capítulo 5);
3. Analisar o desempenho, em termos de AARL, os gráficos de mistura finita e T^2 - Hotelling. Adicionalmente, comparar o desempenho do gráfico de mistura finita com o parâmetro K do gráfico de controle de Shewhart (apresentadas no Capítulo 6).

1.3 JUSTIFICATIVA

Muitos setores industriais usam os mais variados recursos para o controle e melhoria da qualidade de produtos e processos. Para os casos de múltiplos fluxos em que a população pode ter como característica ser distribuição normal e multimodal, a aplicação de métodos estatísticos é complexa por ser ainda pouco explorada. Com isso, os gráficos tradicionais (univariados) figuram como inadequados por apresentarem uma distribuição de probabilidade de mistura finita. Sendo assim é importante explorar o tema MSP para que o

setor industrial encontre recursos na literatura para monitorar esses tipos de processos e assim melhorar a qualidade de seus produtos e satisfação dos clientes. Entretanto para realizar o monitoramento por gráfico de controle tradicional para MSP constrói-se um gráfico para cada fluxo, dificultando assim o monitoramento de um processo em que um equipamento produz todos os fluxos em que não é possível monitorar cada fluxo independentemente, conseqüentemente resultando em monitoramento não adequado, sendo possível aceitar um processo sob controle quando na verdade não está ou vice versa (EPPRECHT, 2015). O monitoramento adequado de um processo de múltiplos fluxos também contribui para a redução do retrabalho, desperdício de materiais, produtos com defeito e produtos de baixa qualidade que serão descartados rapidamente, ou seja, auxiliam a reduzir o lixo e os custos de produção (NOOR *et al*, 2020).

Outra limitação observada é suposição de distribuição normal para dados provenientes de processo de múltiplos fluxos, que não é necessariamente verdade (AHANGAR; CHIMKA, 2015). Em geral, existe uma tendência de que os dados sejam provenientes de uma distribuição multimodal para esse tipo de processo, o que fere a suposição sobre a aderência dos dados à distribuição normal padrão, quando se observam os dados dos fluxos em conjunto, sendo assim importante analisar casos em que ocorre essas situações para que seja possível averiguar essa questão.

Para muitos casos reais de processo com múltiplos fluxos, nenhuma solução anterior na literatura é aplicável. Sendo assim, é necessário desenvolver métodos e procedimentos operacionais adequados à realidade (EPPRECHT, 2015).

Utilizar técnicas multivariadas como ACP/ T^2 -Hotelling é algo recorrente em outras áreas de pesquisa, porém ainda são pouco utilizadas para monitorar MSP. Outro fato observado é que há pouca aplicação e testes empíricos dos modelos para múltiplos fluxos

(EPPRECHT, 2015). Nesse contexto, desenvolver um gráfico de controle utilizando a abordagem de distribuições de probabilidade de mistura finita, pode trazer contribuições teóricas e práticas para o monitoramento estatístico de múltiplos fluxos.

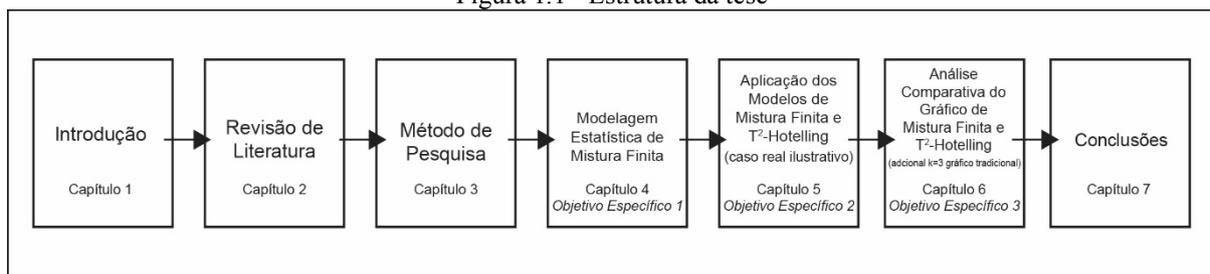
Portanto, esta tese trata do Monitoramento Estatístico de Processos (MEP) para múltiplos fluxos, assunto pouco tratado na literatura especializada. Observa-se como lacuna de pesquisa que as técnicas encontradas na literatura pressupõem condições de validade nem sempre verificadas na prática, tais como assumir que os fluxos são identicamente distribuídos e independentes (EPRECHT 2015). A contribuição dessa tese é auxiliar com um único gráfico de controle o monitoramento adequado de MSP e que este para isso considere a variância total e as variâncias entre os fluxos pela utilização da abordagem de distribuição de probabilidade de mistura finita no gráfico de controle, sendo essa abordagem estudada desta forma somente em Vicentin *et al.*, (2018), trazendo uma análise comparativa com gráfico de controle de T^2 -Hotelling com suas implicações e benefícios. Visto que para Zwetsloot e Woodall (2019) as observações dessas variâncias são pouco estudadas, porém importantes para não mascarar os efeitos de causas especiais e não perder informações sobre o processo, resultando em um monitoramento adequado do processo e melhoria da qualidade dos produtos e processos.

1.4 ESTRUTURA DA TESE

O trabalho terá a estrutura especificada a seguir (como na Figura 1.1). No capítulo 1, a introdução apresenta o tema de pesquisa, o estado da arte com objetivos, justificativa, questão de pesquisa da tese e estrutura da tese. No capítulo 2, a revisão bibliográfica traz um histórico do desenvolvimento do Monitoramento Estatístico de Processo seus desdobramentos, definições e conceitos, apresenta pesquisas recentes na área, como também a abordagem de processo de múltiplos fluxos e a literatura dos métodos utilizados na tese são apresentados,

sendo eles a distribuições de probabilidade de mistura finita e gráficos de controle de T^2 -Hotelling. No capítulo 3, método de pesquisa, os métodos usados na pesquisa são evidenciados. No capítulo 4, modela-se o gráfico de controle de mistura finita para MSP. No capítulo 5, são aplicados os modelos propostos de gráficos de controle para mistura finita e T^2 -Hotelling, retratado por um caso ilustrativo com dados coletados em indústria do setor alimentício, com o objetivo de realizar uma análise comparativa dos modelos. No capítulo 6, o estudo empírico é comparado com dois métodos, mistura finita proposto por Vicentin *et al.* (2018) e o gráfico de mistura finita com o gráfico de controle T^2 -Hotelling e, adicionalmente, traz uma comparação do gráfico de mistura finita com o parâmetro K do gráfico tradicional de Shewhart. No capítulo 7, são dissertadas as conclusões e apontamentos de pesquisas futuras.

Figura 1.1 - Estrutura da tese



Fonte: figura elaborada pela autora.

2 REVISÃO BIBLIOGRÁFICA

Nessa seção, serão apresentadas as abordagens nas quais a pesquisa proposta está baseada como: fundamentos do controle e monitoramento estatístico de processo, a características de qualidade e o valor-alvo do processo para realizar esse tipo de monitoramento, as causas comuns e especiais provenientes do processo a ser monitorado, os pressupostos para o uso desse monitoramento estatístico de processo, diferenças entre controle e monitoramento estatístico de processo e pesquisas recentes sobre esse tema.

Será discutido também o monitoramento de processo dos múltiplos fluxos, bem como a Revisão Sistemática de Literatura apresentada no Apêndice 1, as teorias utilizadas para propor e comprar gráficos de controle para processo de múltiplos fluxos, como a abordagem de distribuições de probabilidade de mistura finita e gráfico de controle multivariados que usa as técnicas de Análise de Componentes Principais e T^2 - Hotelling.

Posteriormente será mostrado uma síntese das duas abordagens usadas para desenvolver os gráficos de controle para processo de múltiplos fluxos.

2.1 FUNDAMENTOS DO CONTROLE E MONITORAMENTO ESTATÍSTICO DE PROCESSO

Walter A. Shewhart desenvolveu, em 1924, a teoria dos gráficos de controle estatístico, que é considerada, em geral, como o começo formal do monitoramento estatístico da qualidade. De acordo com Apsemidis *et al.* (2020), há muitas situações em que é necessário o monitoramento de uma ou mais características de processos e produtos. Para tanto, os gráficos de controle desempenham papel fundamental em garantir o monitoramento e controle de processos de características do produto e críticas da qualidade.

O monitoramento estatístico de processo (MEP) tem sido amplamente utilizado para detectar anomalias monitorando mudanças nos valores de parâmetros específicos de interesse (MONTGOMERY *et al.*, 2013). Com a implementação do MEP, obtêm-se melhores processos de produção, com menos variabilidade, proporcionando níveis adequados de qualidade nos resultados da produção. Quando se fala em melhores processos, isso significa não somente qualidade melhor, mas também custos menores (TASIAS; NENES, 2016).

Segundo Shang *et al.* (2013), o Monitoramento Estatístico de Processo é uma ferramenta essencial para o monitoramento do processo de fabricação, ou para monitoramento da qualidade de bens ou serviços.

Para Woodall (2016), há uma significativa distância entre a teoria e prática do MEP, sendo esse um dos desafios dos pesquisadores e os gestores. Conforme o mesmo autor, o grande desafio das indústrias é atender cada vez melhor às exigências dos consumidores, para tal o monitoramento estatístico desempenha um papel central. Outro aspecto apontado como relevante por esse autor, é que todo embasamento teórico mais recentemente desenvolvido dá pouca atenção à fase I da elaboração dos gráficos de controle. Portanto, dois problemas de pesquisa se destacam: aderência entre teoria e prática e maior foco na fase I da implantação do MEP.

A aplicação mais utilizada do MEP ocorre com o estudo dos gráficos de controle que corresponde aos gráficos de Shewhart, porém outros métodos mais recentes, tais como CUSUM e EWMA, também são empregados (MONTGOMERY *et al.*, 2013). Os gráficos de Shewhart têm sido amplamente utilizados na indústria de fabricação, em razão da simplicidade de sua implementação e a facilidade de sua interpretação, principalmente quanto ao fato de o processo estar ou não sob controle. Contudo, o rápido desenvolvimento da tecnologia levou a processos melhorados para tal finalidade, pois muitos gráficos de controle tradicionais podem

mostrar problemas de desempenho ou de implementação prática (ZWETSLOOT; WOODALL, 2019).

Berk e Berk (1997) afirmam que, para a implementação do MEP, é preciso compreender que essa técnica não é apropriada para todos os processos. Sendo assim, cabe ao líder identificar e conhecer o processo, como também os fatores dominantes relacionados à motivação e ao comprometimento da equipe com a empresa, pois eles podem influenciar provendo a alta qualidade (JURAN, 1997). Para o autor, esses fatores envolvem os processos gerenciais de planejamento, monitoramento e melhoramento da qualidade. Para o gerenciamento do monitoramento, o autor destaca a comparação do desempenho real com as metas e definição de ações em função de possíveis diferenças.

Para Woodall (2016), em função de sua própria experiência no tema, os processos repetitivos, no sentido de produção de itens semelhantes, são os mais indicados para a indústria, pois, em geral, salientam taxas de rejeição maiores que as desejadas e possuem características que permitem medições diretas. No entanto, em vista de sua finalidade, os gráficos de controle também sofrem críticas.

O pressuposto dos gráficos de Shewhart é de que a distribuição de probabilidade das características de controle se adere à distribuição normal, proveniente de uma única população ou único fluxo de produção. Para Epprech *et al.* (2015), há casos em que o produto é manufaturado simultaneamente em diferentes máquinas e fluxos, resultando em subpopulações com diferentes parâmetros e, eventualmente, também com diferentes distribuições de probabilidade e/ou diferentes fluxos como MSP.

Outro ponto relevante sobre monitoramento estatístico de processo e construção dos gráficos de controle tem a ver com os parâmetros estatísticos das características a serem monitoradas em termos de média e desvio padrão (μ e σ). Existem também parâmetros das

características que são desconhecidos em termos de média e desvio padrão (μ e σ) que influenciam na estimativa e cálculo dos limites de controle, afetando a eficiência do uso dos gráficos de controle. Também há duas fases do desenvolvimento desses gráficos, a fase I, em que se estima os limites de controle estatísticos utilizados, e a fase II, que se realiza o monitoramento do processo (OPRIME *et al.*, 2016). No entanto, quando os parâmetros são desconhecidos, estima-se na fase I os limites de controle a partir de k amostras de tamanho n . Já na fase II, são coletadas amostras de tamanho n do processo para monitorar o estado de controle do processo. Assim, a partir da amostra é calculado o valor da estimativa do parâmetro, caso não se encontre dentro dos limites de controle. O processo é considerado fora de controle por uma possível causa especial que deverá ser identificada para poder se tomar ações corretivas, retomando o estado de controle do processo (MONTGOMERY, 2013). Tradicionalmente, os gráficos têm Parâmetros Fixos (FP, do inglês *Fixed Parameters*), tais como o intervalo entre amostras, as constantes K para cálculo dos limites de controle e o tamanho das amostras extraídas do processo durante a coleta dos dados. Como no artigo de Jardim *et al.* (2019), que explora o tamanho das amostras e seus cálculos de limites para avaliar o desempenho do gráfico de controle \bar{X} na Fase II do monitoramento.

Trabalhos publicados nas últimas duas décadas, tais como os de Costa (1998), Magalhães *et al.* (2001) e Nenes *et al.* (2014) propõem variar esses parâmetros (VP, do inglês *Variavel Parameters*) ao invés de mantê-los fixos, tendo como uma das vantagens obtidas a redução dos custos do processo de controle. Esses trabalhos têm como premissa que os parâmetros conhecidos e que a variável de controle seja proveniente de distribuição normal de probabilidade.

Outras correntes de pesquisa com impacto significativo nos *designs* dos gráficos de controle são as análises dos efeitos da estimativa dos parâmetros no desempenho dos

gráficos \bar{X} . Jensen *et al.* (2006) fizeram uma revisão geral das pesquisas sobre os efeitos das estimativas de parâmetros no desempenho dos gráficos de controle e propuseram futuras pesquisas sobre o assunto, dentre elas, o uso de estimadores robustos dos parâmetros. Recentemente, Schoonhoven e Does (2012) analisaram esse assunto em um estudo sobre gráficos de controle \bar{X} robusto. Saleh *et al.* (2015) e Epprech *et al.* (2015) analisaram a dificuldade em se construir gráficos de controle \bar{X} quando os parâmetros são estimados. Goedhard *et al.* (2017), Jardim *et al.* (2017), Jardim *et al.* (2019) modelam ajustes no gráfico de controle \bar{X} com parâmetros estimados. Sobue *et al.* (2020) compara parâmetros do gráfico de controle \bar{X} .

2.1.1 Característica da Qualidade e Valor-alvo

A definição das características de qualidade relevantes em um produto é fundamental para a implantação de qualquer sistema de qualidade. Antes de se aplicar qualquer técnica estatística, é importante definir os pontos-chave do processo, para que os esforços sejam gastos no que de fato é importante. Ouvir os consumidores é uma forma para se conhecer quais características apresentam maior relevância. Geralmente, suas reclamações ou sugestões são bons indicativos de fatores relevantes e são considerados como principais. No entanto, existem outras maneiras de se identificar essas características na definição da qualidade de um produto. Elas podem estar relacionadas com o grau de aceitação ou à rejeição deste, à ocorrência de defeitos e à confiabilidade entre outros (MONTGOMERY, 2013).

Para Abbasi e Adegoke (2020), as características da qualidade provenientes de processo de manufatura ou de serviços geralmente estão associadas às variáveis aleatórias numéricas, podem ser mensuradas e expressas como valores em uma escala contínua de medidas. Entretanto, quando os itens de qualidade são classificados apenas como conforme e

não conforme, a característica pode ser resultado de algum tipo de contagem consistindo, então, em uma variável aleatória discreta. Em seu estudo, os autores destacam a atenção quanto às características da qualidade no MEP para propor seu gráfico de controle *S* na Fase I.

Realçando o conceito de característica de qualidade, Exler e Lima (2012) afirmam que as estimativas dos parâmetros encontrados descrevem o que pode ser considerado como qualidade, daí serem conhecidos como características da qualidade.

Em termos práticos, essas características são associadas a dois tipos de variáveis aleatórias: Contínua e Discreta. No caso contínua, são denominadas como variáveis. É comum encontrar nos textos a expressão “gráfico para variáveis”. As “variáveis” estão relacionadas com as medições com valores quantificáveis. Os processos nos quais os dados referem-se a algum tipo de classificação (conforme - não conforme), seguidos de algum tipo de contagem são referenciados pelo termo “atributos” (ZHOU *et al.*, 2019). Os autores propõem em estudo de um design ótimo para gráfico de controle para atributos buscando monitorar a média dos processos autocorrelacionados.

Montgomery (2013), explica que para monitorar e controlar um processo é preciso que a empresa, em vista de projeto do produto, especifique um valor para cada característica da qualidade em análise, bem como um intervalo de tolerância. Esse valor norteia as técnicas estatísticas empregadas e ajuda a interpretação dos resultados, sinalizando a possível interferência. Assim, cada produto deve conter descrito em seu projeto o seu valor especificado que representa o valor-alvo da variável aleatória. Associado ao valor-alvo, os limites de tolerância também devem ser definidos. Em resumo, o que se pretende é controlar o processo, minimizando sua variabilidade. Evidente que este processo sofre influências dos insumos, dos colaboradores, variáveis externas e equipamentos. Também é importante considerar que existe

a possibilidade de variabilidade em virtude do sistema de medidas, como aquelas associadas à repetitividade e à reprodutibilidade.

2.1.2 Causas Comuns e Especiais

No estudo de MEP, Montgomery (2013) relata que, em um processo, não existem duas unidades iguais, sempre haverá alguma variação em toda atividade produtiva e sua ocorrência é inevitável, isto para instrumentos de medidas perfeitos, o que não existe na prática. Lei e MacKenzie (2019) levam em conta três tipos de causas de variação; uma causa assinalável ou especial em geral única, porém suficientemente grande que ocorre, ocasionalmente, podendo ser eliminada ou ainda minimizada; outra causa ele define como estrutural e a diferencia da especial pelo fato de ser periódica, sendo que, quando o período é muito grande, ela confunde-se com a especial e a causa aleatória ou comum que inerente ao processo e não prejudica o processo, mantendo-o sob controle estatístico. Na realidade, o autor subdivide as causas especiais, definidas pela maioria dos autores, em outras duas: assinaláveis e estruturais.

As Causas Comuns ou Aleatórias fazem parte da variabilidade natural do processo, que é inerente a ele e ocorrerá mesmo que o setor de controle utilize métodos padronizados. Neste caso, a dimensão da variabilidade se mantém em uma faixa estável, em vista da característica do processo, indicando que está sob controle estatístico. As Causas Especiais ou Assinaláveis são as variações esporádicas que ocorrem em uma situação específica e causam anomalias no funcionamento normal do processo, resultando em redução no nível de qualidade. Esse comportamento indica que o processo pode estar fora do controle estatístico, ou que apresenta um alarme falso. Para solucioná-lo, é necessário localizar a possível causa, analisar a situação e decidir ou não pela intervenção no processo. Caso seja realizada alguma intervenção, eliminar a causa de variação e definir medidas que evitem sua persistência

(GOEDHART *et al.*, 2019). Ainda os autores em seu estudo destacam a importância da identificação das causas especiais e comuns no monitoramento de gráficos de controle multivariado.

Firka (2011) evidencia que não se deve esperar ter um processo perfeitamente sob controle estatístico, porém é preciso acompanhar de forma cuidadosa as corridas experimentais tentando prevenir as ocorrências de causas especiais. Pequenas causas levam à variabilidade natural do processo, são inevitáveis e geram uma variabilidade com a qual é preciso conviver. Quando ocorre um problema ou uma operação anormal, no processo, com maiores perturbações, as causas são denominadas especiais. As causas especiais são sempre possíveis de eliminar, originam-se de lotes com defeitos, quebras ou desajustes das máquinas, mas, certos casos, demandam correções significativas no processo. O processo ainda pode produzir itens com defeitos ou defeituosos que podem ser classificados como: defeitos crônicos, que são inerentes ao processo (sempre presentes no processo - causas comuns) e defeitos esporádicos, que representam desvios ao que o processo é capaz de fazer (mais fáceis de identificar - causas assinaláveis).

Como o conceito de causas comuns é inerente a todo processo produtivo, é fundamental que se identifique e analise as causas da variabilidade quando as causas não são naturais ou causas assinaláveis, visando estabelecer mecanismos de controle, a fim de atender às principais características de qualidade do produto. Não sendo naturais, podem ser provenientes de fatores como mão de obra, materiais, métodos, máquinas, medições e meio ambiente. Posteriormente ações corretivas são providenciadas para melhoramento da qualidade dos produtos (MONTGOMERY, 2013).

Em um processo de fabricação, é muito comum observar grandes alterações de parâmetros específicos, em um estágio inicial do processo, devido aos elementos que ainda não

tenham sido estabilizados (LEE *et al.*, 2012). No entanto, quando no processo encontram-se variações na característica de qualidade do produto, é necessário identificá-las quanto ao tipo: Causas Comuns ou Aleatórias ou Causas Especiais ou Assinaláveis (GOEDHART *et al.*, 2019).

2.1.3 Pressupostos para o Monitoramento Estatístico do Processo

Para a utilização de muitas técnicas estatísticas, como o MEP e a análise capacidade de processo, é necessário que os dados tenham uma distribuição normal. Segundo Levinson (2011) e Montgomery (2013), é importante enfatizar que, em alguns casos, a transformação da distribuição não-normal para uma distribuição normal irá simplificar significativamente a análise para os operadores, mas é importante salientar que tais transformações podem perder precisão nas regiões de interesse para o índice de desempenho do processo. A distribuição normal é expressa pela função de densidade de probabilidade $f(x)$ dada por:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-1}{2\sigma^2}(x - \mu)^2\right) \quad (2.1)$$

Onde, μ é a média do processo e σ^2 é a variância. A notação típica para tal distribuição é $N(\mu, \sigma^2)$. Mesmo para série de dados não normais, para amostras suficientemente grandes, as análises em função de médias, são justificáveis em função da existência do Teorema do Limite Central, que garante que, para amostras maiores do que 30, a distribuição amostral das médias, no mínimo, se aproxima da distribuição normal. A questão da aproximação envolve o questionamento do “suficientemente grande”, pois de acordo com Piovesana e Senior (2016), se uma amostra retirada de uma população com distribuição aproximadamente simétrica, então n pode até ser um valor relativamente pequeno, porém, se a distribuição for fortemente assimétrica, será necessário um tamanho de amostra relativamente grande. Segundo Ralaivaosaona e Wagner (2019), o teorema do limite central considera que

sendo $X_1, X_2, X_3, \dots, X_n$ observações independentes de uma população com média μ e variância σ^2 , então \bar{X} , a média amostral, tem uma distribuição aproximadamente normal com média μ e variância σ^2/n . Em seu trabalho os autores utilizam o teorema de limite central para aplicar a log-normal e analisar tamanhos de amostras. Para facilitar a utilização da distribuição normal em meios práticos é usual, conforme Huang *et al.* (2020), lidar com a variável padronizada Z , a qual tem média 0 e variância 1, com notação $N(0,1)$. A transformação de qualquer variável X , normal, em um padrão Z , é dada pela relação:

$$Z = \frac{x - \mu}{\sigma} \quad (2.2)$$

Onde, Z indica o número de desvios padrão da variável X a contar da média, x é o valor arbitrário, μ é a média populacional e σ é o desvio padrão da variável X . Para visualizar melhor a distribuição dos dados, utiliza-se a construção de um histograma que também tem a função de mostrar, graficamente, a normalidade dos dados (MONTGOMERY, 2013). Para Koh e Ahad (2020), a observação visual não deve predominar sobre os testes estatísticos para detecção da não normalidade. Sobre essa temática, os autores em seu trabalho estudam as distribuições não normais e normal em termos de quantidade amostral.

Além da observação da normalidade dos dados também é identificada as modas da distribuição analisada, podendo encontrar mais de uma moda, conhecidas por bimodais e multimodais. A interpretação mais comum de multimodalidade é a de uma mistura de distribuições contendo várias subpopulações com médias diferentes são multimodais, ou como uma indicação de agrupamentos (MINNOTTE, 1997; STRICKER; REDMAN, 2003; MCLACHLAN *et al.*, 2019). Outro caso ocorre quando uma distribuição possua diferentes picos, então, essa distribuição pode ser tratada como multimodal (DIXON, 2012; KAYID; IZADKHAH, 2015). Em distribuição de probabilidade de mistura finita é possível notar mais de moda, como no caso multimodal estudado por Vicentin *et al.* (2018). Ainda dentro desse

assunto de distribuição de probabilidade de mistura finita, Kayid e Izadkhah (2015) propõem um novo modelo de distribuição de probabilidade de mistura finita estendida para tempo de vida dos equipamentos. Já MClachlan *et al.* (2019), fornecem um relato atualizado da teoria e desenvolvimentos metodológicos subjacentes às aplicações de modelos de mistura finita.

2.1.4 Diferenças entre Controle Estatístico de Processo e Monitoramento Estatístico de Processo

As nomenclaturas Controle Estatístico de Processo (CEP) e Monitoramento Estatístico de Processo (MEP) aparecem em diversos estudos na área de conhecimento, porém o CEP foi o primeiro termo a ser utilizado nos trabalhos científicos. No entanto, autores como Antony e Preece (2002) defendem que o termo mais correto seria monitoramento, pois pela definição de monitoramento, os indicadores estatísticos são acompanhados, sendo que, para usar o termo controle, teria que ser algo mais relacionado à mudança em si para normalizar o processo.

A outra possível razão de utilizar o termo monitoramento é que o mesmo aparece mais em estudos com processos multivariados (ou que usam ACP), porque é possível monitorar (mesmo que não for controlar) algumas características não-controláveis e que afetam o processo (LUI *et al.*, 2013; VON ROSING *et al.*, 2015; RATO; REIS, 2017; PETER HE; WANG, 2018; STUBBS *et al.*, 2018).

Sendo assim, as diferenças entre esses termos seria apenas uma nova questão de definição comum ou uso comum para a aplicação do estudo no processo industrial buscando destacar o que pretende ser compreendido nos estudos de monitoramento, tanto para trabalhar com as técnicas estatísticas multivariadas ou para delimitar que o estudo monitora os indicadores estatísticos do processo (BAKER; BROBST, 1996; GRAVES *et al.*, 1999;

JENSEN *et al.*, 2006; ELG *et al.*, 2008; YANG *et al.*, 2012; YU; LIU, 2012; WOODAL, 2016; OPRIME *et al.*, 2019; WOODALL; FALTIN, 2019). No trabalho proposto, foi usado o termo monitoramento com o intuito de representar a ótica observada na literatura.

2.2 RECENTES PESQUISAS SOBRE MONITORAMENTO ESTATÍSTICO DE PROCESSO

Os gráficos de controle de Shewhart foram extremamente importantes para a engenharia, com a estatística facilitando a relação entre técnicos e operários, em um ambiente de produção, apesar de algumas vezes não serem utilizadas de forma correta (MONTGOMETY, 2013). Observam-se em títulos de artigos científicos o termo MEP associado às diversas formas de estudos práticos, cujos conteúdos podem não refletir o que se espera quando comparado com o que a literatura, em sua maioria, consagra. Nesse sentido, pode se perceber uma grande distância entre a realidade e a literatura. Woodall (2000) afirma que “Há um papel para a teoria na aplicação de gráficos de controle, mas a teoria não é o principal ingrediente para aplicações mais bem-sucedidas”. Tem-se constatado, na literatura sobre programas de gerenciamento de processos, relatos controversos a respeito dos melhores caminhos para o monitoramento estatístico da variabilidade em processos produtivos (BERK; BERK, 1997; WOODALL, 2000; BHOTE, 2008).

Saleh *et al.* (2015) relata que, nas aplicações práticas, a suposição dos parâmetros de um monitoramento de processo é tipicamente desconhecida e deve ser estimada. Os autores descrevem que, em alguns casos, seria usual estimar os valores dos parâmetros do processo sob controle na fase I. Também retratam a dificuldade dos pesquisadores na fase II das cartas de controle, na qual seu desempenho é deteriorado, contendo pontos fora de controle, pelo efeito

duplo de erro de estimação, sendo erro nominal maior do que literatura indica (0,0027). Montgomery (2013) ainda relata que, na prática, o processo pode sofrer alterações em termos de média e o valor dessa probabilidade tende a aumentar.

Com isso, evidencia-se a complexidade das decisões e, conseqüentemente, a necessidade de utilizar técnicas estatísticas aliadas à experiência e ao bom senso dos gestores aumentando as chances de intervir no processo sem gerar ainda mais descontroles (FIRKA, 2011). Montgomery e Runger (2003) comentam sobre essa questão, alertando que um resultado estatisticamente significativo pode ter pouca ou nenhuma significância prática.

Montgomery (2013) afirma também que a suposição de normalidade, quando não atendida, pode levar a uma análise incorreta em relação à capacidade do processo. Por isso, pode-se avaliar um processo como capaz quando, na realidade, encontra-se incapaz e outras vezes pode classificá-lo como incapaz sendo ele capaz. Ou mesmo em termos dos limites de controle, ele pode ser classificado como fora de controle quando, na realidade, está sob controle. Como pode ser observado no estudo de Hadian e Rahimifard (2019), a suposição de normalidade para capacidade do processo foi analisada e aplicou-se o gráfico de controle multivariado de T^2 - Hotelling em um projeto para otimizar os custos e tempo.

Esse problema ocorre porque existem várias outras distribuições não-normais como: *gamma*, *lognormal* e *Weibull* que não são identificadas e transformadas antes da construção dos gráficos de controle (ALEVIZAKOS; KOUKOUVINOS; 2019).

Todavia, Chakraborti e Sparks (2020) evidenciam que, para validar as interpretações dos gráficos de controle, é necessário que os valores da variável-resposta Y sejam independentes e se comportem de acordo com o modelo de distribuição normal.

Chakraborti e Graham (2019) relatam que os gráficos de controle não paramétricos podem ser o caminho a percorrer na prática, quando não é razoável supor uma distribuição

paramétrica específica, tal como o modelo normal. Isto pode acontecer quando a suposição de distribuição normal é insustentável por causa da teoria, da falta de conhecimento ou informação disponível sobre os dados.

Jensen *et al.* (2006) trazem uma revisão de literatura sobre os efeitos dos parâmetros nas propriedades dos gráficos de controle, apontando como pesquisas futuras mais trabalhos científicos sobre os gráficos de controle para dados de atributos que precisam ser investigados minuciosamente, sendo observada a distribuição marginal para fazer recomendações úteis sobre o tamanho da amostra e desenvolver métodos para ajustar os limites de controle para levar em consideração a estimativa de parâmetros e obter o desempenho desejado no controle. Ainda ressaltam a necessidade de mais estudo sobre o efeito do uso de estimadores robustos ou de outras alternativas na Fase I e na Fase II. Também sugerem pesquisas que determinem as propriedades dos gráficos de controle de Shewhart com parâmetros estimados que incorporam várias regras de execução, sendo pelos autores ainda apontado que provavelmente um gráfico usando várias regras de execução simultaneamente seja mais afetado pela estimativa de parâmetros do que um gráfico usando uma única regra. Os autores comentam que muito pouco trabalho foi realizado em gráficos multivariados e o existente concentrou-se na detecção de alterações no vetor médio e o desempenho deles sob alterações nos parâmetros da matriz de covariância não foi investigado, o que seria análogo às alterações na variação dos gráficos de controle univariados. Ainda relatam que o uso de outros estimadores, como o estimador de diferença sucessivas ou estimador robusto, não foi estudado e pouca atenção foi dada a outros gráficos multivariados, como o EWMA multivariado e o CUSUM multivariado.

Os autores também comentam que, em muitas aplicações, os limites de controle são atualizados à medida que mais dados se tornam disponíveis e que a maioria das pesquisas se baseia no pressuposto de que existe uma única amostra de dados da Fase I usada para calcular

os limites da Fase II, sendo necessária a investigação dessas questões. Ainda sobre essas questões, sugerem que o efeito nas propriedades da carta de controle, quando os limites de controle são atualizados em algum momento futuro que não seja necessariamente durante um período de inicialização (Fase I), não estando claro como o desempenho do gráfico de controle é afetado, apontando que o uso de dados anteriores da Fase II levaria aos melhores gráficos de controle e a atualização dos limites de controle se encaixaria naturalmente em um esquema de gráfico de controle bayesiano. Sugerem mais trabalhos sobre gráficos para monitorar mudanças na média do que gráficos para monitorar a variação, mesmo que o efeito de estimativa pareça ser mais severo para gráficos que monitoram mudanças na dispersão do que para aqueles que monitoram mudanças na média, como os gráficos EWMA ou CUSUM. Ainda mostram que o efeito da estimativa de parâmetros não foi incluído em muitos outros tipos de gráficos de controle como gráficos simultâneos (gráficos EWMA com limites de Shewhart) e gráficos usando tamanhos de amostra variáveis e/ou intervalos de amostragem, sendo necessário, para monitorar perfis, tamanhos de amostra maiores da Fase I, devido à natureza multivariada dos dados. Esses, segundo os autores, serão mais severamente afetados pela estimativa de parâmetros, pois são mais sensíveis a pequenas mudanças no processo, sendo necessário pesquisas para obter boas recomendações sobre como melhor projetar os gráficos com parâmetros estimados.

Os autores também recomendam correções nos limites de controle quando não há dados suficientes da Fase I, sendo que ainda não houve uma comparação desses métodos de correção para determinar quais são melhores, nem estudo de correções para gráficos que não sejam gráficos de Shewhart. Certificam ainda que, em alguns casos, a especificação incorreta do modelo é uma questão importante a ser considerada ao projetar e configurar gráficos de controle, como nos gráficos de controle para dados autocorrelacionados que dependam fortemente da correção do modelo de série temporal assumido. Sugerem que o efeito

combinado da especificação incorreta do modelo e estimativa de parâmetros não foi estudado. Verificam que há pouca orientação na literatura sobre como fazer uma análise da Fase I, que é fundamental para o uso bem-sucedido do gráfico na Fase II. Ainda recomendam que seria muito útil ter bons métodos gráficos para avaliar melhor o impacto da estimativa de parâmetros em vez de tabelas de valores, como normalmente é feito, especialmente em situações em que o impacto depende de determinadas variáveis. Identificam que os estimadores padrão frequentemente usados nas cartas de controle geralmente são justos e consistentes em relação a alguns critérios, sendo razoável que a restrição de considerar aqueles que sejam consistentes, tenham menor variabilidade, maior robustez e sejam mais exatos para que possam ser estimadores que produzem melhor desempenho no gráfico de controle. Posteriormente, ressaltam que quaisquer avaliações devem incluir várias medidas de desempenho em suas avaliações das propriedades do gráfico de controle. Os autores ainda demonstram preocupação com a ênfase na ARL e insuficiente no SDRL e nos percentuais da distribuição da RL, sendo que a consideração do efeito da estimativa de parâmetros, à medida que novos métodos estão sendo desenvolvidos, aumentará a probabilidade de novos métodos serem utilizados e produzirá aplicações mais bem-sucedidas dos esquemas de monitoramento de cartas de controle.

Já no trabalho de Zwetsloot e Woodall (2019), foi realizada uma revisão de literatura no tema de monitoramento de processo com apontamento para pesquisas futuras. Os autores recomendam que mais estudos sejam feitos relativos ao efeito de agregação temporal na Fase I, que são usados para avaliar a estabilidade histórica de um processo, visto que, na maior parte dos artigos, essa análise é realizada na Fase II. Sugerem investigar em processo proveniente de distribuição de Poisson o efeito da agregação na capacidade dos métodos de monitoramento para detectar diminuições na taxa e estudar o efeito da agregação de dados no monitoramento dos processos Poisson não homogêneos. Para gráfico de controle multivariado,

os autores relatam que o efeito da agregação de dados no desempenho desses tipos de gráficos ainda não foi estudado, visto que as taxas de amostragem para as variáveis podem não ser as mesmas, dificultando a análise. Recomendamos que mais trabalhos sejam feitos sob ótica do efeito possivelmente substancial da agregação de dados na detecção de mudanças transitórias. Em relação às redes sociais, apontamos muitas oportunidades para estudar o efeito da agregação de dados em seu monitoramento, destacando a importância da investigação do potencial de usar os dados do registro de data e hora diretamente sem agregação temporal. Apontamos também pesquisas realizadas na área de monitoramento de imagens para reduzir o ruído. Com relação ao erro de medição, recomendamos estudar como ele pode afetar as decisões sobre o nível de agregação e ainda sugerimos que as medições médias podem reduzir a variação do erro de medição. Apontamos também que estudos futuros podem ser feitos relacionados com a plotagem das estatísticas da carta de controle em uma frequência e a análise da carta que pode levar às novas métricas de desempenho, com base no número de exames (ou visualizações) da carta até que um sinal seja emitido. Ainda mostramos a importância de estudar o efeito de agregar os desvios individuais de um modelo ajustado usado com gráficos de seleção de causa para formar amostras que poderiam ser usadas com métodos de monitoramento de perfil. Incentivamos ainda pesquisas que justifiquem o nível de agregação selecionado usado em qualquer estudo de caso.

2.3 MONITORAMENTO DE PROCESSO DE MÚLTIPLO FLUXOS

Vários fluxos ocorrem em um processo que gera dados provenientes de diversas fontes ou canais ou está sob monitoramento do processo estatístico, mas a variável de qualidade e suas especificações são as mesmas em todos os fluxos (WANG, 2015; TIAN, 2016; ABDULAZIZ *et. al.*, 2019). Assim, é importante entender as fontes de variação entre e dentro dos subgrupos para analisar e interpretar o gráfico de controle e como eles afetam o processo

(EPPRECHT, 2015; BROWN; SCHAFFER, 2020). Presume-se que cada fluxo tenha a Função de Distribuição de Probabilidade normal (f.d.p).

Conforme Abdulaziz *et al.* (2019), para o ponto de vista estatístico, as diferenças entre os fluxos e dentro de cada fluxo são inevitáveis. Essas diferenças contribuem com a variabilidade total nas unidades de saída que determinam a qualidade do produto. Esses vários fluxos de processos, tradicionalmente, utilizam um conjunto diferente de gráficos de controle para monitorar o desempenho de cada fluxo ao longo do tempo.

Ainda nesse tema foi desenvolvida uma Revisão Sistemática de Literatura apresentada no Apêndice 1.

2.4 PROPOSTAS PARA MONITORAMENTO DE MSP

Nessa seção, são apresentadas os dois modelos para o monitoramento de processos de múltiplos fluxos, em termos de suas construções teóricas, sendo elas gráfico de controle utilizando distribuição de probabilidade de mistura finita e gráfico de controle T^2 - Hotelling.

A distribuição da probabilidade de mistura finita é o resultado da tese que foi publicado na revista *International Journal of Quality & Reliability Management*, no ano de 2018, após ser selecionado entre os melhores artigos no congresso *2nd International Conference on Quality Engineering and Management*, em 2016. Essa proposta de modelo/método é detalhada na seção 4.1 e o artigo publicado está no Apêndice 3.

2.4.1 Distribuição de Probabilidade de Mistura Finita

A abordagem da primeira proposta para gráfico de controle e processo de múltiplos fluxos foi usada primeiro no artigo de Vicentin *et al.* (2018) em que se observou pela pesquisa de literatura para Monitoramento Estatístico de Processo que ainda não havia sido explorada na área em questão. Nessa pesquisa, foi utilizada a literatura de Titterington *et al.* (1985) para construir o gráfico de controle proposto. Podemos ver a seguir a proposta construída a partir da literatura sobre Distribuição de Probabilidade de Mistura Finita.

Sendo assim, seja X uma variável aleatória, assumindo os valores contidos em um espaço de amostral, Ω , sua distribuição pode ser representada pelo modelo de função densidade de probabilidade para variáveis contínuas ou função de massa para o caso discreto, utilizando o mesmo modelo para MSP como (TITTERINGTON *et al.*, 1985):

$$p(x) = \pi_1 f_1(x) + \pi_2 f_2(x) + \dots + \pi_l f_l(x) \quad (2.3)$$

Onde $\pi_j > 0$, $\sum_j^l \pi_j = 1$; $j = 1, \dots, l$ e $\int f_j(x) dx = 1$. Os parâmetros π_j são pesos da mistura das distribuições ou subpopulações. Como as funções de densidade de probabilidade têm parâmetros, uma distribuição de probabilidade de mistura finita pode ser representada geralmente por $\pi_1 f_1(x|\theta_1) + \pi_2 f_2(x|\theta_2) + \dots + \pi_l f_l(x|\theta_l)$, onde θ representa os parâmetros $f_j(x)$.

Por exemplo, se houver uma mistura de duas distribuições normais com o mesmo desvio padrão σ , diferentes médias μ_1 e μ_2 , e a função densidade total dos parâmetros é $\omega = (\pi, \mu_1, \mu_2, \sigma)$, a probabilidade será representada por:

$$p(x|\omega) = \pi \phi(x|\mu_1, \sigma) + (1 - \pi) \phi(x|\mu_2, \sigma) \quad (2.4)$$

Pode ser escrito uma função genérica da função densidade de probabilidade de uma mistura finita com subpopulações l como:

$$p(x|\omega) = \sum_{j=1}^l \pi_j f(x|\theta_j) \quad (2.5)$$

A Equação 2.5 das funções de densidade de probabilidade da mistura indica que existem categorias (subpopulações ou fontes) nas quais a unidade experimental obtém o valor x da variável X a partir de uma dessas categorias. Assim $f_j(x)$, mostra a distribuição de probabilidade de X obtida a partir da categoria j and π_j que indica a probabilidade de a variável estar em uma dessas categorias (TITTERINGTON *et al.*, 1985).

Algumas propriedades importantes das distribuições de misturas são conhecidas e exploradas para identificar distribuições. Claramente, se as distribuições de referência tiverem variações comparáveis e forem suficientemente diferentes nas medições de posição central, uma mistura de duas distribuições será bimodal e misturas de três ou mais serão distribuições multimodais (DIXON, 2012; KAYID E IZADKHAH, 2015). Portanto, identificar a probabilidade da mistura não parece uma tarefa trivial. A existência de bi ou multimodal é uma forma de verificar a mistura de subpopulação. Os testes de normalidade também podem identificar se existe uma distribuição multimodal. No entanto, existe a possibilidade de erros: a distribuição unimodal pode esconder a existência de duas subpopulações e a identificação de mais de um modo por um histograma não necessariamente conclui que há mais de uma subpopulação em dados de massa. Em geral, considera-se que um componente puro é unimodal, mas isso nem sempre é verdade. Existem gráficos e métodos multivariados, como análise de cluster e função discriminante, que podem ser usados para identificar uma distribuição de probabilidade de mistura finita para posteriormente monitorar o processo (TITTERINGTON *et al.*, 1985). Como no trabalho de Thissen *et al.* (2005), em que os autores utilizam um algoritmo de maximização de expectativa e método multivariado como cluster para distribuições de probabilidade de mistura finita em gráfico de controle. Já no trabalho de Zhang e Albin (2007), os autores propõem utilizar métodos multivariados para identificar o

número correto de cluster nos dados coletados com distribuições de probabilidade de mistura finita para construção de gráficos de controle multivariado. No trabalho de Grasso *et al.* (2014), os autores investigam o uso de métodos multivariados e rede neural de *fuzzy* em gráfico de controle com distribuições de probabilidade de mistura finita. Shao *et al.* (2017) desenvolvem um esquema baseado em rede neural de dois estágios para aumentar a precisão da taxa de identificação de distribuições de probabilidade de mistura finita para gráfico de controle.

Neste trabalho, serão estudados nas seções 4.1 e 4.2 um indicador de mistura e será explorado o histograma com gráficos de testes de normalidade e bitangencialidade para identificar distribuições de probabilidade de mistura finita na seção 5.1.

2.4.2 Análise de Componentes Principais e T^2 -Hotelling

O Monitoramento Estatístico de Processo (MEP) torna-se mais complexo quando mais variáveis de monitoramento são adicionadas, para controlar variações menores no processo, podendo levar à sobrecarga de informação para os operadores. Aplicação de cartas de controle univariadas a cada variável individual pode não ser eficaz e levar a conclusões errôneas sobre o MEP. A diferença entre os gráficos univariados e multivariados é a interrelação entre os dados. Assim, ao terem reciprocidade e se comportem de forma dependente, as características da qualidade do processo necessitam ser analisadas em conjunto (DOĞU, 2015; MARQUEZ; VIVAS, 2020). Além disso, possíveis interrelações que as variáveis possuem podem não ser observadas, dificultando a interpretação do real comportamento do processo. Já os gráficos de controle multivariado mostram, em um só gráfico, todas as variáveis estudadas no controle do processo, indicando os pontos de controle de forma agregada (MONTGOMERY, 2013; JARDIM *et al.*, 2019).

O Monitoramento Estatístico de Processo Multivariado (Monitoring Statistical of Process Multivariate - MSPM) proposto por Wise e Gallagher (1996) e Mac Gregor *et al.* (1994), pode superar esse problema. Gráficos de controle multivariados funcionam bem quando o número de variáveis não é muito grande (sendo menor ou igual a 10 variáveis). No entanto, Marquez e Vivas (2020) em seu estudo relatam que os gráficos de controle multivariados perdem sua eficiência em detectar mudança à medida em que o número de variáveis cresce.

Dentro da abordagem de MSPM, as técnicas de análise multivariada de dados são usadas para reduzir o número de variáveis em alguns componentes, diminuindo a dimensão do problema e apresentar as observações em algumas parcelas com regiões de confiança, como a Análise de Componentes Principais (ACP). Isso proporciona a possibilidade de incluir um grande número de variáveis no sistema de supervisão e ainda manter uma visão global de conjuntos de dados de processos complexos (SIVASAMY; SUNDAN, 2015; AHSAN *et al.*, 2018).

Essa abordagem para monitoramento de processo multivariado tem a característica de avaliar cada variável no tempo, analisar o conjunto de variáveis e suas correlações existentes e verificar cada lote para determinar se um processo está fora de controle ou não (HUANG *et al.*, 2020). Para o monitoramento da estabilidade de um processo multivariado, segundo Mason e Young (2001), é necessário que as médias, variâncias e covariâncias das variáveis do processo estão estáveis nas observações realizadas.

No monitoramento, o controle do vetor de médias para um processo multivariado é similar ao gráfico de Shewhart, porém também compreende diferentes tipos de demonstrações para dados provenientes de subgrupos e observações individuais (ABBAS *et al.*, 2020).

Sendo assim, a proposta para essa pesquisa foi utilizar métodos multivariados como Análise de Componentes Principais (ACP) e T^2 -Hotelling para verificar o número adequado de gráficos de controle para processo de múltiplos fluxos e construí-lo, tendo em vista os fluxos como as variáveis empregadas nos métodos em questão. Os itens a seguir evidenciam, em termos teóricos, ambos métodos multivariados.

2.4.2.1 Análise de Componentes Principais (ACP)

Esse método consiste em resumir e identificar as características dos produtos e variáveis que mais influenciam o processo produtivo com o intuito de melhorar a qualidade. Esse método estatístico pode garantir melhor estabilidade para um processo com mais de uma variável correlacionada. O uso de tais técnicas é recomendado, pois os testes univariados não são adequados para determinar a plausibilidade das variáveis por conta do aumento na probabilidade de erro tipo I ou na taxa de alarme falso (NEZHAD; NIAK, 2013; SALO *et al.*, 2019).

Para Jonhson e Wichern (2007), o principal objetivo dessa técnica estatística multivariada é explicar, por meio da construção de combinações lineares das variáveis originais, a estrutura de variância e covariância de um vetor aleatório. Seu modelo pode ser representado da seguinte forma:

$$\begin{aligned}
 Y_1 &= a_{11}x_1 + a_{21}x_2 + \dots + a_{p1}x_p \\
 Y_2 &= a_{12}x_1 + a_{22}x_2 + \dots + a_{p2}x_p \\
 &\quad \vdots \quad \quad \quad \vdots \\
 Y_p &= a_{1p}x_1 + a_{2p}x_2 + \dots + a_{pp}x_p
 \end{aligned}
 \tag{2.6}$$

Este método consiste transformar um conjunto de variáveis originais X_1, X_2, \dots, X_p em um novo conjunto de variáveis Y_1 (CP₁), Y_2 (CP₂), ..., Y_p (CP_p). Onde essas p variáveis não são correlacionadas entre si e estão arranjadas em ordem decrescente de variâncias. O principal enfoque nesse procedimento é trabalhar para que poucas dentre as primeiras componentes principais incorporem a maior variabilidade dos dados originais, podendo racionalmente descartar as demais componentes que, conseqüentemente, irão reduzir o número de variáveis. A redução das variáveis é realizada com o intuito de observar quais são as que afetam mais o processo produtivo (JONHSON; WICHERN, 2007; SALAH *et al.*, 2015; VANHATALO *et al.*, 2017).

Os critérios para utilizar a ACP são: dados quantitativos, variáveis originais correlacionadas, o novo conjunto de variáveis deve ser menor ao número de variáveis originais, novos componentes ou novos conjuntos de variáveis devem ser completamente não correlacionados (MASTRANGELO *et al.*, 1996; AHSAN *et al.*, 2018).

Os resultantes da ACP são as Componentes Principais (CP) que contém as informações sobre as variáveis que influenciam no processo e são apresentados em ordem decrescente por essa influência. Sendo assim, a quantidade de informação aglutinada nos primeiros eixos da CP é bastante significativa. As variáveis das componentes principais CP 1 e CP 2 são os eixos do novo sistema de coordenadas obtido pela rotação do sistema original e os novos eixos representam as direções de variabilidade máxima. Sendo a primeira componente principal (CP 1) responsável pela maior parte da variabilidade das variáveis originais (MONTGOMERY, 2013; GU *et al.*, 2016).

Para escolher as Componentes Principais, pode ser utilizado o critério sobre as porcentagens ou o critério de maior peso. O critério das porcentagens define, segundo Hair *et al.* (2019), quais os componentes principais k podem substituir as p variáveis originais. Para

isso é comum obter em apenas duas ou três das primeiras componentes principais mais que 90% das informações necessárias sem que haja prejuízo para o estudo. Sendo assim, existe quase tanta informação em k componentes principais que nas p variáveis originais.

Outros preferem estimar o número de CPs com a utilização dos critérios de Kaiser que são os critérios mais utilizados. Esses critérios indicam que os autovalores das CPs devem ser maiores que 1 para representar o conjunto original, ainda que a soma das variâncias das CPs seja maior que 80%, utilizando uma matriz de correlação ou a matriz de covariância. Na matriz original, cada amostra apresenta um valor para cada variável e, na nova matriz da componente, cada amostra tem um valor para cada componente que são obtidos em conjunto e conhecidos como valores das componentes (JOHNSON; WICHERN, 2007; ALMEIDA *et al.*, 2019; de ALMEIDA *et al.*, 2020).

O núcleo de qualquer ACP é composto por autovetores e autovalores. Os autovetores são conjuntos de valores que representam o peso de cada variável original sobre cada componente. São escalados como coeficientes de correlação e variam de +1,0 a -1,0. Para cada componente, todas as variáveis têm um conjunto de autovetores correspondentes e quanto mais próximo de +1,0 ou -1,0 está o autovetor, mais importante é a variável para o componente. Os autovalores representam a contribuição relativa de cada componente na explicação da variação total dos dados, existindo então um autovalor para cada componente, e o tamanho do autovalor para o componente é uma indicação direta da importância do componente na explicação da variação total dentro do conjunto de dados. Já o vetor é um uma classe de equipolência de um segmento orientado no espaço que possui propriedades de comprimento e direção. Ele varia de -1,0 a +1,0 e cada variável na análise pode ser examinado um vetor. O comprimento de um vetor é representando por uma variável que está diretamente relacionada à sua variância, mas como o primeiro passo na PCA geralmente é padronizar as variáveis à variância 1,0, os vetores representando todas as variáveis na análise que possuirão o mesmo

comprimento de 1,0 (JOLLIFFE, 2002; MONTGOMERY, 2013; JOLLIFFE, I. T.; CADIMA, 2016).

Para aplicar essa técnica, conforme Manly e Alberto (2017), é preciso seguir os seguintes passos:

- 1) Codificar as variáveis X_1, X_2, \dots, X_P para terem médias zero e variâncias unitárias;
- 2) Calcular a matriz de covariâncias $S_{P \times P}$. Essa é uma matriz de correlação se o primeiro passo já foi executado;
- 3) Encontrar os autovalores $\hat{\lambda}_1, \hat{\lambda}_2, \dots, \hat{\lambda}_p$ e os correspondentes autovetores $\hat{e}_1, \hat{e}_2, \dots, \hat{e}_p$.

Os coeficientes da j -ésima componente principal são então os elementos de \hat{e}_j , enquanto que $\hat{\lambda}_j$ é a sua variância.

4) Descartar quaisquer componentes que explicam apenas uma pequena proporção da variação dos dados.

Pode-se também encontrar os *outlier* a partir de seus *x-scores*, com o uso do método das Componentes Principais que transformam os valores originais em outra escala. Usa-se, então, a distribuição amostral multivariada T^2 de Hotelling. Pode-se ainda avaliar a trajetória dos pontos ao longo do tempo e identificar a importância de cada variável para um componente com o método de Análise de Componentes Principais (MONTGOMERY, 2013; AHSAN *et al.*, 2018).

Jolliffe e Cadima (2016) trazem uma revisão de literatura sobre Análise de Componentes Principais (ACP) na qual analisam a importância de várias e diferentes formas de empregar esse método em diversas áreas científicas. Zhou *et al.* (2016) analisaram duas modificações nos métodos de monitoramento convencionais baseados em ACP para trabalhar com dados provenientes de distribuição não normal e a natureza variável no tempo dos dados

do processo. Os autores utilizaram dados reais de um processo de fabricação de ferro para verificar a eficácia da detecção precoce de anormalidades no processo. Vanhatalo *et al.* (2017) desenvolvem uma estrutura dinâmica de ACP para monitoramento estatístico de processo. Ahsan *et al.* (2018) utilizam ACP e T^2 -Hotelling para propor um método que supera o desempenho do gráfico de controle T^2 -Hotelling convencional em termos de precisão de detecção e tempo de execução. Val *et al.* (2019) desenvolveram uma estratégia utilizando ACP para monitoramento estatístico de processo com intuito de melhorar a tomada de decisão na avaliação da característica de qualidade. Salo *et al.* (2019) utilizaram métodos análise de componente principal (ACP) e algoritmos de aprendizagem baseados em instância para desenvolver um método híbrido que tem melhor desempenho sobre as abordagens de gráficos de controle individuais, alcançando alta precisão e baixas taxas de falsos alarmes.

2.4.2.2 Gráfico de controle T^2 -Hotelling

Harold Hotelling, em 1947, foi quem desenvolveu esse método estatístico, o que deu a ele o título de pioneiro na pesquisa sobre gráficos de controle multivariado. Esse gráfico é visto como o mecanismo mais comum no monitoramento desse tipo de processo para dados constituídos de subpopulações (MONTGOMERY, 2013; SIVASAMY; SUNDAN, 2015).

O método baseia-se na generalização da estatística t de *Student*, a qual se estende para uma aplicação multivariada que possui a forma da estatística T^2 , tendo em vista estimativas amostrais da matriz de covariância (HAIR, 2019; DOĞU; KIM, 2020). A aplicação desse método é feita por meio do cálculo da medida da distância entre os valores das variáveis e o vetor de suas médias, que é conhecida ou estimada por meio de uma amostra de base (AHSAN *et al.*, 2018).

Montgomery (2013) e Shams *et al.* (2015), na equação da estatística T^2 , afirmam ser possível estimar \bar{X} (vetor estimado das médias de cada p variável), de dimensão $p \times 1$ e S de dimensão $p \times p$, quando μ (média populacional) e σ (desvio padrão populacional) não são conhecidos. Sendo que:

$$\bar{X} = \frac{1}{m} \sum_{k=1}^m \bar{X}_k \quad (2.7)$$

$$\bar{S} = \frac{1}{m} \sum_{k=1}^m S_k \quad (2.8)$$

$$S = \frac{1}{m-1} \sum_{k=1}^m (X_i - \bar{X}_k)(X_i - \bar{X}_k)' \quad (2.9)$$

Onde, m é o número de amostras preliminares que foram retiradas do processo de tamanho n e X_i representa a i -ésima observação da k -ésima amostra da p -variável.

Lowry e Montgomery (1995) e Ahsan *et al.* (2018) relatam a existência de 2 fases para a construção de gráficos multivariados. A Fase I consiste em testar, após a retirada de m subgrupos de amostras, se os dados multivariados estão sob controle. Os limites de controle são relacionados com a escolha de um nível de significância α , tal que a probabilidade P dos dados testados estarem dentro dos limites estabelecidos é dada por: $P(T^2 \leq LIC)$. Sendo assim, para Montgomery (2013) e Shams *et al.* (2015) os limites de controle para o gráfico são calculados por:

$$LSC = \frac{p(m-1)(n-1)}{mn-m-p+1} F_{\alpha,p,mn-m-p+1} \quad (2.10)$$

$$LIC = 0$$

Onde, p é a quantidade de características que estão sendo analisadas simultaneamente, n é o tamanho do subgrupo, m é o total de subgrupos, $F_{\alpha,p,mn-m-p+1}$ é um percentil superior da distribuição F com p e $mn - m - p + 1$ graus de liberdade, sendo $n > 1$ (CHEN; HSIEH, 2007; SHAMS *et al.*, 2015).

Na Fase II, caso algum ponto da estatística esteja fora dos limites de controle, deve-se averiguar se isso ocorre por conta das causas especiais ou à variabilidade própria do processo. Se um ponto estiver fora dos limites de controle por causa especial, a amostra correspondente deve ser retirada dos subgrupos de amostras e os limites de controle devem ser recalculados, para verificar se, com os novos limites, os pontos permanecem sob controle. Após a Fase I, tomando como base os limites de controle estabelecidos na mesma, realiza-se o monitoramento de processo na Fase II. Na Fase II, novos subgrupos de amostras são extraídos para se juntarem as amostras sob controle obtidas na Fase I e os limites são recalculados. Esses limites recalculados têm a função de confirmar se o processo permanece sob controle na Fase II. A desvantagem desse método está relacionada com a dificuldade de encontrar o conjunto de variáveis ou a variável que está colaborando para causar o ponto fora de controle do processo (MONTGOMERY, 2013; KEEFE *et al.*, 2015).

Para Alt (1985), Mason, Tracy e Young (1997), Abbas *et al.* (2020), um ponto fora de controle representa um sinal do que pode ser causado por várias situações como: devido ao relacionamento entre as variáveis contraditórias estabelecidas por dados históricos; uma das observações de uma das p variáveis pode estar fora dos limites de controle; a combinação de algumas variáveis fora do controle com outras interrelacionadas.

Para verificar a variável, ou as variáveis, que contribui para que um ponto ultrapasse os limites de controle, deve-se decompor a estatística T^2 . Com isso, é possível obter as componentes e observar a contribuição de cada variável individualmente no processo (PENG *et al.*, 2020). Mason *et al.* (1995) desenvolveram um método de decomposição que é um dos mais usados e decompõe a estatística T^2 em componentes independentes que mostram a contribuição de uma variável individual. A decomposição geral da estatística de T^2 de Hotelling para variáveis é dada pela seguinte equação:

$$T^2 = T_1^2 + T_{2.1}^2 + T_{3.1,2}^2 + T_{4.1,2,3}^2 + \dots + T_{p.1,\dots,p-1}^2 \quad (2.11)$$

Onde, os termos do lado direito da equação representam as contribuições parciais das p variáveis estudadas na estatística T^2 .

Sivasamy e Sundan, (2015) desenvolveram um gráfico de controle T^2 de Hotelling com outras técnicas multivariadas como *Support Vector Machine* (SVM) e *Triangle Area-based Nearest Neighbors* (TANN) para monitoramento estatístico de processo. Shams *et al.* (2015) propõem um design econômico para o gráfico multivariado T^2 de Hotelling aplicado a um sistema de grande escala. Ahsan *et al.* (2018) utilizam ACP e T^2 -Hotelling para propor um método que supera o desempenho do gráfico de controle T^2 -Hotelling convencional em termos de precisão de detecção e tempo de execução. Ahsan *et al.* (2019) desenvolvem um gráfico de controle T^2 de Hotelling baseado no algoritmo de *Minimum Covariance Determinant* (MCD) e estimativa de densidade de Kernel. Doğu e Kim (2020) propõem gráficos T^2 de Hotelling para parâmetros não conhecidos. Huang *et al.* (2020) aplicam o gráfico de controle T^2 de Hotelling para identificação de paletes não conformes em fábricas de montagem automotiva.

No caso do estudo dessa tese, foi feita uma adaptação do método T^2 de Hotelling e ACP com relação as diferentes variáveis existentes na sua literatura para as diferentes posições dos produtos no cilindro, estudando, assim, o processo de múltiplos fluxos em uma indústria do setor alimentício, analisadas no caso ilustrativo.

No próximo item, será exposto como se realizou a construção das propostas e a suas aplicações, destacando suas diferenças. Para tanto, foi utilizado um caso ilustrativo com dados reais de uma indústria do setor alimentício.

3 MÉTODO DE PESQUISA

Este capítulo está dividido em três partes, sendo que as primeiras tratam da classificação da pesquisa e o último item deste capítulo apresenta os passos para sua execução.

Para Lakatos e Marconi (2017), primeiro é necessário destacar a atividade de pesquisa. Esta fase, por sua vez, representa um procedimento formal referente a um pensamento reflexivo que necessita de tratamento científico. Sendo assim, a atividade de pesquisa requer um método científico para ser relevante em termos acadêmicos. O método científico apresenta-se por meio de ideias, regras, técnicas e abordagens utilizadas na área de conhecimento (NEUMAN, 2006). Gay e Diehl (1992) evidenciam que, para trabalhar com o método científico de forma ordenada, é oportuno o reconhecimento e a definição do problema, coleta de dados, análise dos dados e conclusão, rejeitando ou aceitando a hipótese formulada na pesquisa.

Portanto, para este trabalho serão descritas seções com intuito de apresentar com mais detalhes o método de pesquisa, contendo seus componentes e as respectivas justificativas.

3.1 ABORDAGEM METODOLÓGICA

Existem três tipos de abordagem de pesquisa, qualitativa, quantitativa e o uso dos dois tipos de abordagens em conjunto. A abordagem quantitativa compreende o uso de medições e procedimentos estatísticos (MARTINS, 2010). Esse tipo de abordagem é iniciado com a identificação do problema para, posteriormente, desenvolver uma teoria, justamente com as variáveis utilizadas para tal. Para formular uma teoria, é preciso utilizá-la de forma generalizada ou não, em conformidade com seus resultados a partir de medições e métodos

estatísticos empregados. No uso da abordagem quantitativa, deve-se mensurar as variáveis com as suas características. Entretanto, não é exclusividade dessa abordagem, pois, mesmo com a linguagem matemática e uso de variáveis, pode não ser um bom critério para diferenciar as abordagens de pesquisa científica (CRESWELL; CRESWELL, 2017).

Bryman (1989) destaca que a abordagem quantitativa tem características como a mensurabilidade, causalidade, generalização e replicação. Sendo assim, a mensurabilidade consiste definir um conjunto de variáveis plausíveis de serem medidas, a causalidade compreende a prova da existência dos relacionamentos com as variáveis definidas, a generalização entende-se por existir possibilidade ampliar os resultados obtidos para além da pesquisa e a replicação que fornece a repetição da pesquisa por outros pesquisadores podendo chegar ao mesmo resultado.

A abordagem qualitativa tem característica como a ênfase na interpretação subjetiva do objeto de pesquisa, o delineando-o. Essa abordagem utiliza diversas fontes de evidências com necessidade de percepção da realidade na indústria e a proximidade com o fenômeno estudado. Ainda tem o interesse de investigar a forma de como se os resultados foram obtidos, sem o foco nos resultados em si.

Na junção das duas abordagens aqui discorridas, ainda pode se obter o terceiro tipo de abordagem que, tal como enfatizado por Creswell e Creswell (2017) e Creswell e Clark (2006), pode ser complementar possibilitando uma ampliação da visão da pesquisa e melhor entendimento do problema de pesquisa. Ela pode ser utilizada quando uma das abordagens não responde totalmente à questão de pesquisa, tornando-se necessário uma segunda fonte de dados para respondê-la.

No estudo proposto, foi utilizada a abordagem quantitativa que compreende o uso de métodos de modelagem e simulação salientados nos itens a seguir e a abordagem qualitativa e quantitativa que compreende o uso da Revisão Sistemática de Literatura.

3.2 MÉTODOS DE PESQUISA

Há diferentes métodos de pesquisa científicos utilizados dependendo da área de conhecimento. Na Engenharia de Produção, conforme Miguel *et al.* (2010), os principais métodos usados são o levantamento tipo *Survey*, estudo de caso, pesquisa-ação, desenvolvimento teórico-conceitual, revisão de literatura, pesquisas experimentais e modelagem e simulação. No trabalho em evidência, foram utilizados os métodos de modelagem e simulação, além de retratar um passo inicial de pesquisa com a revisão sistemática de literatura (método teórico-conceitual) exposta no Apêndice 1.

Fonseca (2002) profere que a Revisão Sistemática de Literatura (RSL) faz referência à busca de teorias publicadas objetivando obter informações sobre um problema específico que se deseja estudar. A métrica adotada para realizar a RSL consiste em análises e discussões a partir de perguntas sobre literatura do tema que se pretende aprofundar (PETERSEN *et al.*, 2008; KITCHENHAM *et al.*, 2009). A revisão sistemática de literatura identifica artigos que abordem temas relacionado com o processo de múltiplos fluxos no contexto de monitoramento estatístico de processos.

Para Meredith *et al.* (1989), a modelagem e simulação são métodos que geram conhecimento racional, sendo assim dedutiva e axiomática, pois tem como objetivo explicar o comportamento dos processos reais ou parte dele. A pesquisa axiomática quantitativa pode ser prescritiva ou descritiva. Sendo a prescritiva uma pesquisa que estrutura estratégias e ações,

com o objetivo de melhorar os resultados observados na literatura e encontrar uma solução ótima tendo um problema novo ou comparar o desempenho de estratégias que trabalham o mesmo problema. A descritiva analisa os modelos quantitativos para explicar as características do processo que o modela ou compreende-o, apoiando-se em modelos que descrevem o comportamento do sistema modelado. Ainda é possível, por meio desse método de modelagem e simulação, capturar problemas de tomada de decisões de gestores nos processos produtivos reais. Nesse método de modelagem, há também a classificação de modelo estático e dinâmico. O modelo estático analisa um sistema complexo visto por um instante de tempo particular. Já o modelo dinâmico é visto à medida que o tempo evolui como nas operações dos sistemas reais (MORABITO; PUREZA, 2010).

Portanto, a investigação proposta tem como método escolhido a modelagem e simulação como pesquisa axiomática descritiva, estudada pelo viés do modelo dinâmico.

3.3 PROCEDIMENTOS PARA DESENVOLVIMENTO DA PESQUISA

Para desenvolver os gráficos de controle de mistura finita, de T^2 -Hotelling e adicionalmente o parâmetro K do gráfico de controle de Shewhart, bem como as suas análises comparativas foram utilizados os *softwares Maple 13, Statistica 13.0 e Microsoft Office Excel*.

Sendo o Maple uma ferramenta utilizada como um sistema de matemática computacional comercial que permite o usuário programar suas próprias modelagens por meio de linguagem de programação procedural com sintaxes similares ao FORTRAN, PASCAL ou C, que permite inserir dados na forma de distribuição de probabilidade (MAPLESOFT, 2020). Neste trabalho, foi utilizada a distribuição de probabilidade de mistura finita para inserir os

dados dos produtos coletados no final da linha produtiva para modelar e simular o gráfico de controle em múltiplos fluxos.

Já o *software* Statistica, é um programa computacional estatístico comercial de análise e visualização de dados. Os recursos de análise de dados desse software compreendem diversos algoritmos, funções, testes e métodos, trabalhando com tabelas de detalhamento simples a modelagem não-linear avançada, modelos lineares generalizados, métodos de séries temporais e técnicas multivariadas (SÁ, 2007). Na investigação desenvolvida, foram utilizadas técnicas multivariadas como ACP e T^2 - Hotelling com dados coletados a partir das posições de produção dos itens estudados.

O *software Microsoft Office Excel* é um recurso computacional que contribui para a interpretação e as inter-relações possíveis a respeito do fenômeno pesquisado, minimizando o tempo dos cálculos e traçado de tabelas e gráficos (BIANCHINI, 2015). Sendo assim, essa pesquisa utiliza esse *software* para construir a amostra expandida utilizando a técnica *Bootstrap* e tabelas para organizar os dados coletados, analisados e comparados.

Tendo em vista os objetivos propostos, a metodologia adotada para o projeto é a pesquisa de campo com características de pesquisa exploratória e propositiva: proposta de dois gráficos de controle para monitoramento de processos de múltiplos fluxos utilizando distribuições de probabilidade de mistura finita e outras técnicas de estatística multivariada ACP e T^2 -Hotelling.

Foram adotadas diferentes estratégias: método de revisão sistemática de literatura e método de modelagem e simulação.

A revisão sistemática de literatura permite identificar artigos que abordem temas relacionados com o processo de múltiplos fluxos no contexto de monitoramento estatístico de

processos. O objetivo é sistematizar práticas, barreiras e dificuldades no uso em conjunto, previamente identificadas, podendo subsidiar a realização do presente estudo.

A modelagem e simulação evidenciam abordagem quantitativa, já que sua realização busca, primeiro, analisar a adoção de um novo método estatístico para monitoramento de processo diferente do tradicional e a identificação de problemas na adoção dos métodos propostos para distribuição mista utilizando um gráfico de controle com amostras aleatórias e outro de T^2 -Hotelling com amostras referentes às posições de fabricação.

Os objetivos propostos envolvem a análise quantitativa e qualitativa, pois, na modelagem e simulação, foram observados dados de variáveis contínuas e, na pesquisa de Revisão Sistemática de Literatura, localizada no Apêndice 1, foram tratados aspectos relativos ao seu conteúdo e valores numéricos referentes ao número de artigos publicados por ano.

Sendo assim, o Quadro 3.1 resume a classificação da pesquisa realizada de acordo com os fundamentos metodológicos expostos nesse item.

Quadro 3.1 - Classificação da pesquisa

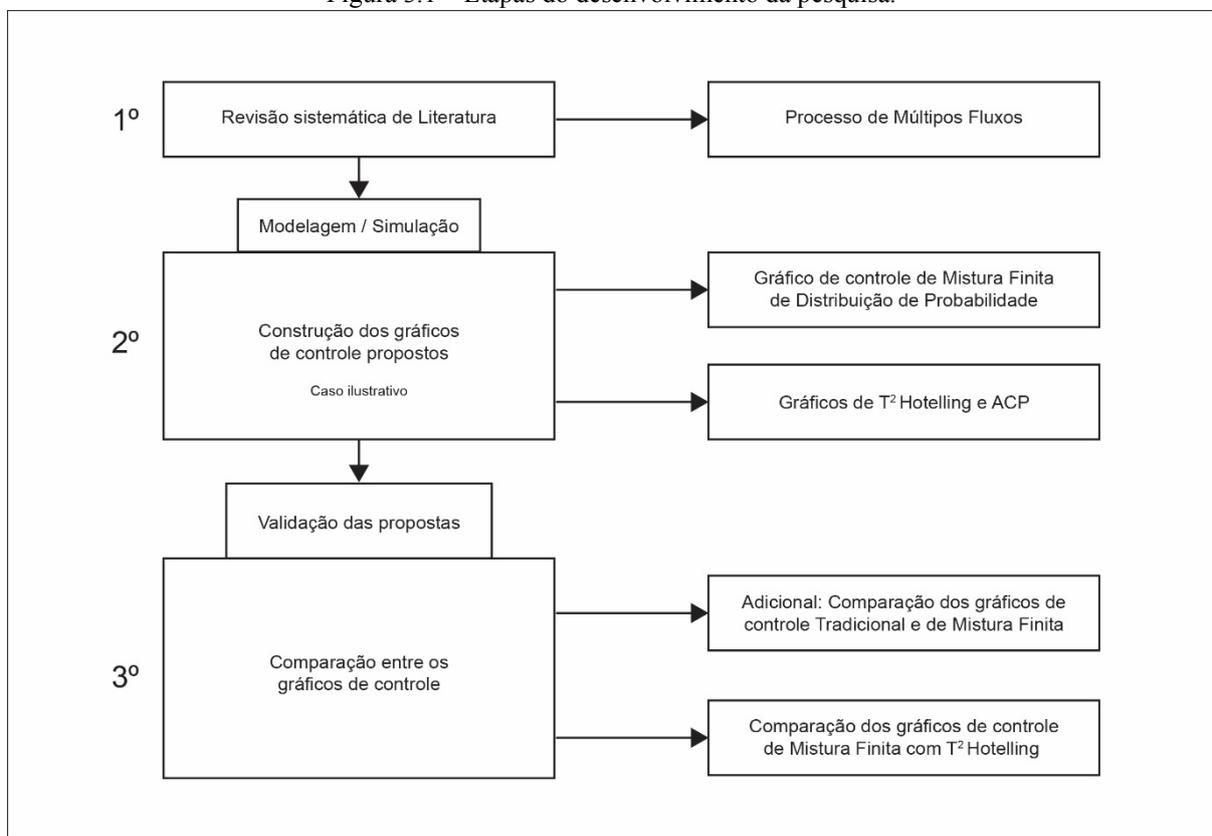
Abordagem de Pesquisa	Qualitativa e Quantitativa
Método de Pesquisa	Teórico-Conceitual (Revisão Sistemática de Literatura) e Modelagem e Simulação
Característica de Pesquisa	Axiomática Descritiva
Técnicas de Pesquisa	Revisão Sistema de Literatura e Modelagem e Simulação
Ferramenta de Pesquisa	<i>Softwares: Maple 13, Statistica 13.0 e Excel</i>

3.4 ETAPAS DO DESENVOLVIMENTO DA PESQUISA

A pesquisa foi desenvolvida em três etapas: i) Revisão Sistemática de Literatura sobre Processos de Múltiplos Fluxos, ii) Modelagem e Simulação com a construção dos

gráficos de controle usando as abordagens de distribuição de probabilidade de mistura finita e análise multivariada como Análise de Componentes Principais (ACP) e T^2 -Hotelling, com dados reais de uma indústria alimentícia como caso ilustrativo, e iii) Comparações entre o parâmetro K do gráfico de controle tradicional de Shewhart adicionalmente e entre os dois gráficos de controle de mistura finita e de T^2 -Hotelling. A Figura 3.1 mostra a disposição dessas etapas para o desenvolvimento da pesquisa.

Figura 3.1 – Etapas do desenvolvimento da pesquisa.



Fonte: figura elaborada pela autora.

Na primeira etapa, foi realizada Revisão Sistemática de Literatura (RSL) sobre o tema de processo de múltiplos fluxos e especificamente sobre Monitoramento Estatístico de Processo (MEP), comentando opiniões de diversos autores sobre a aplicação dessa abordagem. Fonseca (2002) menciona que a pesquisa faz referência à busca de teorias publicadas com o objetivo de obter informações sobre um problema específico que se deseja estudar. Na busca

inicial, com refinamento feito por meio das *strings*, foram encontrados 672 artigos. Neste trabalho, optou-se pela análise de 25 artigos que se encaixam no tema, considerando que esses artigos estavam de acordo com os critérios verificados por meio da introdução, metodologia e conclusão.

Os artigos foram buscados em bases de dados como Scopus, SciELO (Scientific Electronic Library Online) e Google Acadêmico publicados no período de 10 anos (janeiro de 2010 a janeiro de 2020). Para a busca dos artigos, foram observadas as palavras-chave “*Multiple Stream Process*” sendo escolhidos aqueles que, de fato, abordavam o MEP. A métrica adotada na RSL é obter discussões sobre a literatura do tema que se pretender aprofundar (PETERSEN *et al.*, 2008; KITCHENHAM *et al.*, 2009).

No estudo desses artigos selecionados para a RSL sobre MSP, foram analisadas as atividades realizadas e instrumentos dispostos, a abordagem, método e metodologia (dos tipos de gráficos de controle usufruídos), os resultados dos artigos, análises, discussões e as lições aprendidas pelos autores nos artigos ou apontamentos futuros.

Na segunda parte da pesquisa, foram realizadas a modelagem e simulação com os dados de uma empresa do setor alimentício no estado de São Paulo. O item em destaque usado para o estudo é uma massa para pastel, cujo produto final é uma embalagem composta por 17 discos, medindo 14cm de diâmetro cada um deles, separados por lâminas plásticas, sendo que, para o cliente final, cada disco gera “um pastel”. Esse produto tem peso líquido na embalagem dado por 500g. Atualmente, a empresa utiliza o controle que consiste na pesagem dos produtos de 30 em 30 minutos. No estudo desenvolvido, a característica de interesse monitorada é a variável peso (g) cuja especificação aponta o intervalo de tolerância [511; 526], já contando o peso das 17 lâminas plásticas (11g). Essa especificação foi elaborada em função da ocorrência de produtos com peso líquido inferior a 500g, gerando possibilidade de conflitos com órgãos

fiscalizadores. O intervalo para os pesos exigidos pelo INMETRO, de acordo com a empresa, é de 496g a 525g.

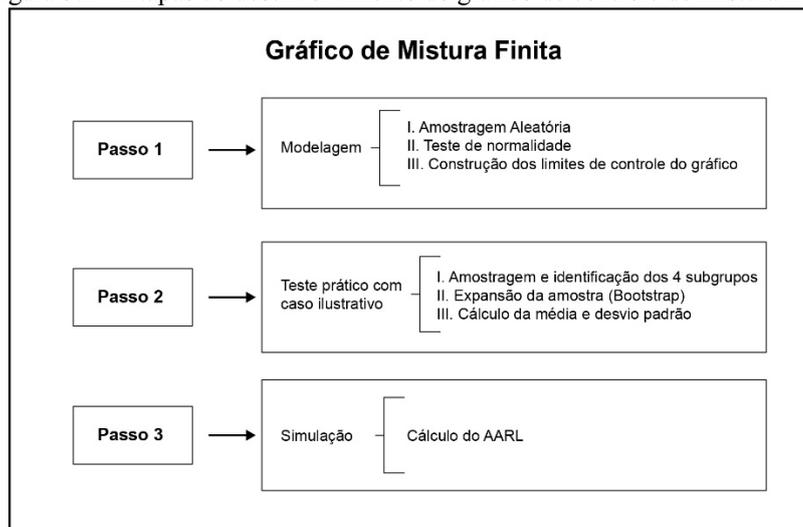
Sendo assim, os dados foram coletados ponderando os fluxos do processo, ou seja, as posições de sua fabricação para a construção dos gráficos de controle: i) uso de distribuição de mistura finita para amostras independentes, tomando-se as amostras aleatórias no final do processo, descondensando as posições de fabricação, apenas observando o raqueamento dos quartis; ii) uso do método T^2 de Hotelling para gráficos de controle mantendo o modo sistemático de extração das amostras, analisando os diferentes fluxos de produção oriundos das posições dos discos de massa de trigo ocupadas no cilindro.

Os dados dos 40 itens, visto que cada item representa um pacote com 17 discos de massa de trigo e os plásticos que os separam, foram coletados de cada fluxo, ou de cada posição dos discos no cilindro, em um intervalo de 30 minutos, durante dois dias, para formar a amostra e após serem expandidos utilizando a técnica de *Bootstrap*. Para a primeira proposta, foi considerada que a distribuição de probabilidade geral resultante de subconjuntos do processo de fabricação com uma função de distribuição de probabilidade (p.d.f) é conhecida. Um modelo teórico foi desenvolvido e avaliado examinando dois cenários: 1) assumindo que os parâmetros de p.d.f. são conhecidos e 2) no caso em que a média é desconhecida. A partir da amostra de 40 itens coletada na indústria, foi utilizada a técnica *Bootstrap* para sua expansão para 4000 itens, desconsiderando as posições de fabricação, cuja característica de qualidade é exatamente o peso de cada unidade processada (etapas detalhadas do processo são evidenciadas na Figura 4.1, que está na seção 4.1). Foi expandida a amostra mestre pontuando as quatro distribuições normais encontradas na observação gráfica do teste de normalidade, sendo 1000 itens para cada distribuição normal em que se buscou estabelecer uma relação com as posições dos quartis. A amostra também foi ordenada pelas suas médias. A distribuição de probabilidade de mistura finita decorrente da característica desse processo é proveniente do erro da forma do

cilindro que gera espessura não uniforme e, portanto, pesos diferentes para processo de múltiplos fluxos.

A proposta do estudo é desenvolver um gráfico de controle de mistura finita para monitoramento de MSP e comparar empiricamente, em termos de AARL, com o gráfico de controle T^2 - Hotelling. Assim, primeiramente constrói-se um gráfico de controle único, cuja amostra expandida não dependem da origem de fabricação da amostra mestre extraída e, portanto, tenha melhor desempenho que os gráficos de controle tradicionais. A comparação entre os gráficos foi feita por meio da análise do desempenho do AARL. A Figura 3.2 resume e ilustra essa etapa de desenvolvimento do gráfico de controle de mistura finita em que no passo 1 é realizada uma modelagem assumindo uma amostragem aleatória, sem se importar com os fluxos, e dados provenientes de uma distribuição de probabilidade de mistura finita. A partir dessas considerações construiu-se os limites de controle do gráfico de mistura finita. No passo 2 é feito um teste prático para implementação do gráfico de controle de mistura finita com dados reais, como caso ilustrativo, em que retirou-se a amostra aleatória, sem se preocupar com os fluxos e posteriormente fez-se a identificação dos 4 fluxos por meio de teste de normalidade, após isso expandiu-se a mostra de 40 elementos para 4000 elementos por meio da técnica *Bootstrap* e calculou-se a média e o desvio padrão da amostra expandida. Posteriormente no passo 3 foi realizada uma simulação do AARL, ou seja, de várias estimativas de limites de controle para AARL.

Figura 3.2 – Etapas do desenvolvimento do gráfico de controle de mistura finita.

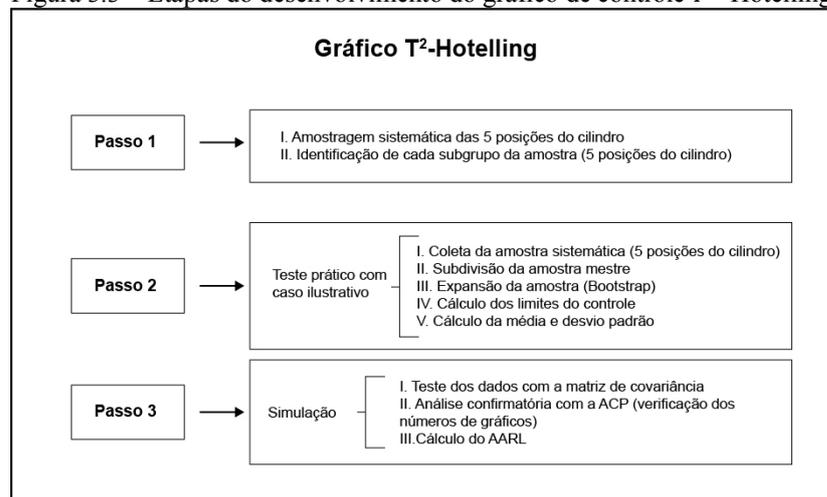


Posteriormente, ainda na segunda parte, desenvolveu-se um gráfico de controle multivariado utilizando o método T^2 -Hotelling para analisar as interdependências entre os múltiplos fluxos.

Primeiramente antes de usar o T^2 de Hotelling multivariado em conjunto com o ACP é checar a correlação entre as amostras dentro de cada fluxo, assim espera-se que haja correlação entre as medidas para cada uma das posições dos cinco fluxos de processo. Para tal, conforme a Figura 3.3 o primeiro passo foi realizar uma amostragem sistemática e identificar os fluxos do processo ou subgrupos. O segundo passo foi testar o gráfico de T^2 -Hotelling em um caso ilustrativo com dados reais. Para isso foi extraída amostra sistemática de 8 itens dos 5 fluxos ou 5 posições de fabricação no cilindro, portanto identificou-se que cada fluxo foi composto de 8 unidades, sendo eles subdivididos. A correlação foi observada para justificar o uso de técnicas multivariadas, como o ACP e o T^2 -Hotelling, assim o gestor poderá construir modelos gráficos para monitorar o processo sem a necessidade de construir um gráfico para cada posição. Entretanto, é necessário expandir a amostra original para avaliar a proposta por meio da técnica *Bootstrap*. Para tal, foi preciso construir o mecanismo de correlação linear entre os fluxos. O termo “linear” usado aqui se deve ao fato do pressuposto de que fatores

externos, ou seja, causas identificáveis ou não externas estudadas no modelo terão efeitos lineares, ou comuns, em todas as posições. Ou seja, um desgaste de ferramenta, um setup diferente ou mudanças nos inputs de matéria prima, afetará igualmente todas as variáveis na sua média. Sem a modelagem do mecanismo de correlação entre as posições, não será possível compreender e prever o comportamento da proposta de design multivariado de gráfico de controle. Foi realizada, então, uma simulação de 1000 amostras para cada posição no cilindro, totalizando 5000 amostras, feita a partir da expansão de sua amostra mestre. Essa expansão foi oportuna para analisar o desempenho do gráfico de controle de T^2 de Hotelling. Após isso foram calculados os limites de controle, bem como a média e o desvio padrão. Já no passo 3 foi feita uma simulação por meio de teste dos dados com a matriz de covariância, de uma análise confirmatória com a APC para verificar o número de gráfico de controle a ser construído, e do cálculo do AARL.

Figura 3.3 – Etapas do desenvolvimento do gráfico de controle T^2 -Hotelling.



Fonte: figura elaborada pela autora.

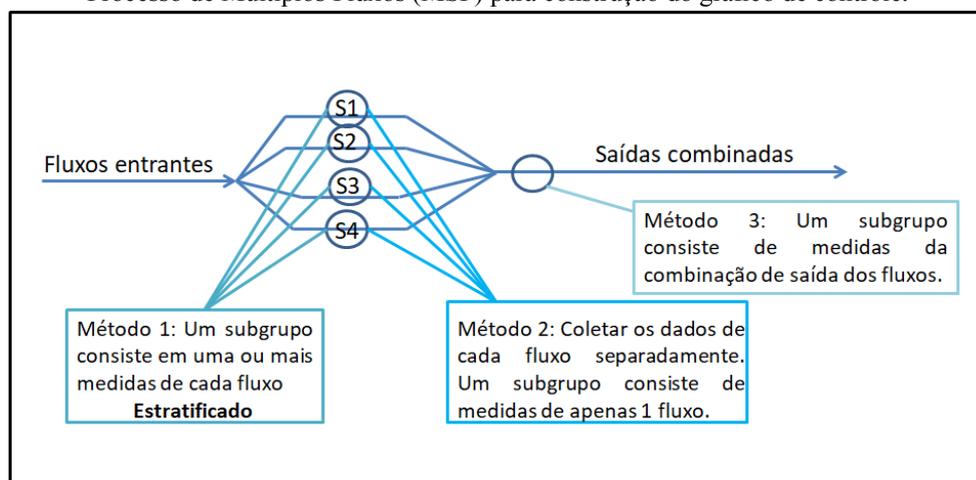
Posteriormente, foram desenvolvidos os procedimentos para a Fase I e a Fase II de monitoramento do processo com o gráfico de T^2 -Hotelling com discussões e análises sobre a aplicação desse método.

Na terceira parte, compreende a fase de validação das propostas para gráficos de controle feitos nessa tese que são o gráfico de controle de distribuições de probabilidade de mistura finita e o gráfico de controle multivariado de T^2 - Hotelling. Com isso, primeiro, será comparado o gráfico de controle de distribuições de probabilidade de mistura finita com o parâmetro K do gráfico tradicional de Shewhart e posteriormente serão comparados os gráficos de controle propostos, sendo eles, o gráfico de controle de distribuições de probabilidade de mistura finita e o gráfico de controle multivariado de T^2 - Hotelling. A distribuição de probabilidade de mistura finita para diferentes subpopulações do processo (ou seja, MSP) impede o uso de gráficos de controle tradicionais. Uma alternativa é monitorar esses fluxos (neste caso, cada posição do cilindro) com esse tipo gráfico de controle de T^2 -Hotelling que monitora as diferentes posições do cilindro.

4. MODELAGEM DE GRÁFICOS DE CONTROLE COM DISTRIBUIÇÕES DE PROBABILIDADE DE MISTURA FINITA

O desempenho dos gráficos de controle é avaliado pela distribuição do valor esperado na *Average Run Length* (ARL) — ou Comprimento de Corrida Média — quando os parâmetros são estimados para a determinação dos Limites de Controle (LSC e LIC). Para avaliar o desempenho dos gráficos de controle em caso de mistura finita, utiliza-se a ARL para o desvio padrão, conhecido igualmente para cada categoria, e para a média desconhecida. Assume-se, então, que as categorias com médias diferentes seriam ranqueadas a partir de quartis e, em seguida, identificadas e separadas, como mostra a Figura 4.1. Foi analisado, primeiramente, o cenário em que os parâmetros e as distribuições de probabilidade são conhecidos. Nesse caso, presume-se que cada um tenha a Função de Distribuição de Probabilidade (FDP) normal.

Figura 4.1 – Caracterização de métodos para coleta de dados em um Processo de Múltiplos Fluxos (MSP) para construção do gráfico de controle.



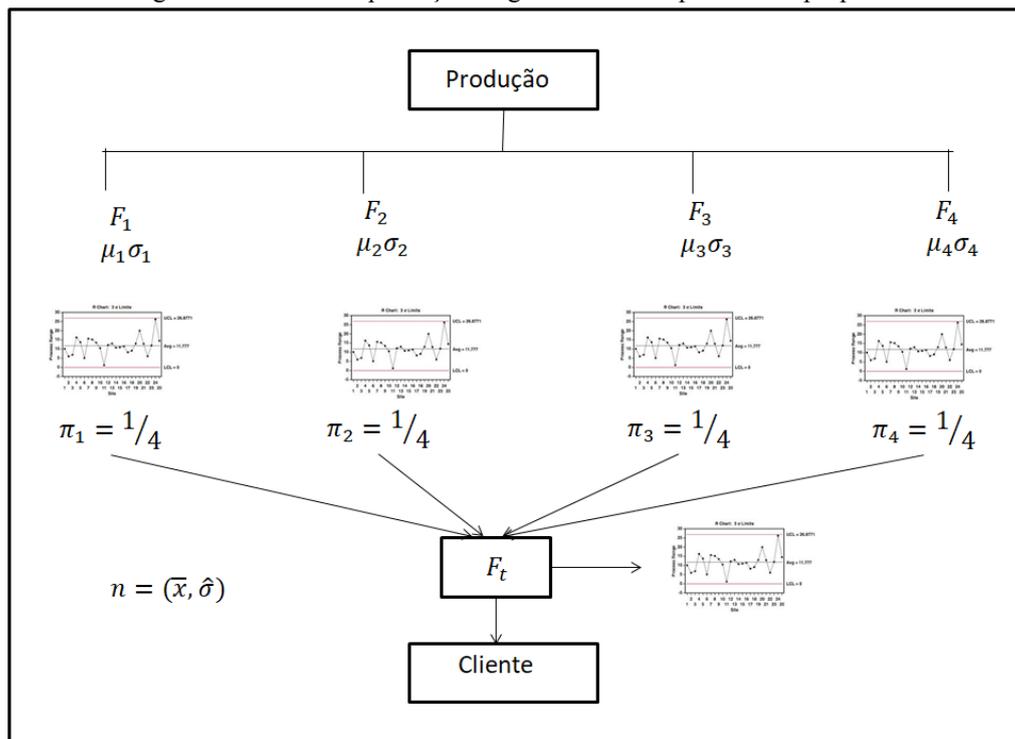
Fonte: Grupo de Ação da Indústria Automotiva (AIAG), 1991.

Segundo o Grupo de Ação da Indústria Automotiva (AIAG), as amostras podem ser coletadas considerando cada fluxo e examinadas em conjunto. Assim, quando o processo

de produção é composto por múltiplos fluxos, um gráfico de controle pode ser construído para todo o conjunto de dados coletados.

A Figura 4.2 mostra o plano de produção teórico e ilustrativo com quatro fluxos (F_1 , F_2 , F_3 e F_4), observados em termos de quartis, e a união deles em apenas um gráfico de controle — em detrimento da criação de um gráfico para cada fluxo. Essa separação se deu pela observação gráfica do teste de normalidade, o qual indicou um quadro de distribuições normais. As médias da amostra foram dispostas em ordem crescente. Este esquema de plano de produção pode ser muito útil para obter informações necessárias no final da linha de produção, para serem entregues ao cliente.

Figura 4.2 – Plano de produção do gráfico de múltiplos fluxos proposto.



Fonte: figura elaborada pela autora.

4.1 CASO COM PARÂMETROS ESTATÍSTICOS E μ CONHECIDO

Segundo Vicentin et al. (2018), sendo x uma variável aleatória proveniente de uma população, por exemplo, com $l=3$ categorias ou subpopulações, normalmente distribuídas com parâmetros $\theta = \{\mu_1, \mu_2, \mu_3, \sigma_1, \sigma_2, \sigma_3\}$, temos a seguinte função de densidade de probabilidade:

$$p(x|\omega) = \pi_1\phi(x|\mu_1, \sigma_1) + \pi_2\phi(x|\mu_2, \sigma_2) + \pi_3\phi(x|\mu_3, \sigma_3) \quad (4.1)$$

Por simplicidade, considerou-se $\sigma_1 = \sigma_2 = \sigma_3 = \sigma$ e adotou-se que a média geral, nesse caso, é obtida por $\mu_0 = \frac{\mu_1 + \mu_2 + \mu_3}{3}$. Outra suposição é que cada n amostra extraída tem probabilidade π_l de ser da subpopulação l . O estimador amostral é a média dada por $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ e os Limites de Controle Superior e Inferior (LSC e LIC, respectivamente) são calculados para parâmetros conhecidos por:

$$LSC = \mu_0 + \frac{K\sigma_t}{\sqrt{n}}; \quad LIC = \mu_0 - \frac{K\sigma_t}{\sqrt{n}}. \quad (4.2)$$

Onde μ_0 é a média global e σ_t é o desvio padrão considerando todos os fluxos. Sabendo que a linha central do gráfico de Limites de Controle (LCs) se encontra em μ_0 , a probabilidade de um ponto estar dentro dos Limites de Controle é:

$$P(\bar{x} \in [LIC, LSC]) = \pi_1 P(LIC \leq \bar{x} \leq LSC | \mu_1, \sigma) + \pi_2 P(LIC \leq \bar{x} \leq LSC | \mu_2, \sigma) + \pi_3 P(LIC \leq \bar{x} \leq LSC | \mu_3, \sigma) \quad (4.3)$$

Expandindo esta formulação, temos:

$$\begin{aligned} & \pi_1 P\left(\mu_0 - \frac{K\sigma_t}{\sqrt{n}} \leq \bar{x} \leq \mu_0 + \frac{K\sigma_t}{\sqrt{n}} \middle| \mu_1, \sigma\right) + \pi_2 P\left(\mu_0 - \frac{K\sigma_t}{\sqrt{n}} \leq \bar{x} \leq \mu_0 + \frac{K\sigma_t}{\sqrt{n}} \middle| \mu_2, \sigma\right) + \\ & \pi_3 P\left(\mu_0 - \frac{K\sigma_t}{\sqrt{n}} \leq \bar{x} \leq \mu_0 + \frac{K\sigma_t}{\sqrt{n}} \middle| \mu_3, \sigma\right) \end{aligned} \quad (4.4)$$

No estado de controle, cada subpopulação tem médias: μ_1, μ_2 e μ_3 , sendo $\mu_1 < \mu_0$ e $\mu_3 > \mu_0$. Dado que existe, de fato, relação matemática entre as distâncias médias de cada subpopulação e a média geral, quando assumimos $\mu_0 = \mu_2$, temos as seguintes relações entre a média geral e média de cada subpopulação: $\mu_0 = \mu_1 + \delta_1\sigma$; $\mu_0 = \mu_3 - \delta_3\sigma$, onde δ é o desvio padrão da diferença entre as subpopulações. Sendo assim expressa por:

$$\delta = \frac{\mu_0 - \mu_1}{\sigma} = \frac{\mu_1 - \mu_0}{\sigma} \quad (4.5)$$

Neste caso específico, para $l = 3$ e para as médias das subpopulações equidistantes, $\delta_3 = \delta_1 = \delta$ e $\delta_2 = 0$.

Após a expansão proposta e considerando, em termos do número de desvios padrão, a possibilidade de o processo estar fora de controle, esse estado de processo é representado da seguinte forma:

$$\begin{aligned} \mu_1 &= \mu_{10} + t\sigma_t \\ \mu_2 &= \mu_{20} + t\sigma_t \\ \mu_3 &= \mu_{30} + t\sigma_t \end{aligned} \quad (4.6)$$

Quando $t \neq 0$, assume-se que o processo está fora de controle e, caso contrário, tem-se que o processo está em estado de controle.

Considerando estas relações e sabendo que a distribuição de probabilidade de subpopulações é representada por uma distribuição normal, com parâmetros conhecidos e cujas amostras são de subpopulação com π_l de probabilidade (Equação 4.4), é possível determinar a probabilidade de uma amostra de tamanho n estar na região dentro dos Limites de Controle. Para isso, subtrai-se μ_l das desigualdades de cada subpopulação e obtém-se o seguinte resultado:

$$\begin{aligned}
P(\bar{x} \in [LIC, LSC]) &= \pi_1 P\left(\mu_0 - \frac{K\sigma_t}{\sqrt{n}} - \mu_1 \leq \bar{x} - \mu_1 \leq \mu_0 + \frac{K\sigma_t}{\sqrt{n}} - \mu_1 \mid \mu_1, \sigma\right) + \\
&\pi_2 P\left(\mu_0 - \frac{K\sigma_t}{\sqrt{n}} - \mu_2 \leq \bar{x} - \mu_2 \leq \mu_0 + \frac{K\sigma_t}{\sqrt{n}} - \mu_2 \mid \mu_2, \sigma\right) + \\
&\pi_3 P\left(\mu_0 - \frac{K\sigma_t}{\sqrt{n}} - \mu_3 \leq \bar{x} - \mu_3 \leq \mu_0 + \frac{K\sigma_t}{\sqrt{n}} - \mu_3 \mid \mu_3, \sigma\right)
\end{aligned} \tag{4.7}$$

Sendo $\mu_0 = \mu_l + \delta_l \sigma$ e substituindo μ_0 pelas relações mostradas acima, para δ e t , temos:

$$\begin{aligned}
&\pi_1 P\left(\mu_1 + \delta_1 \sigma - \frac{K\sigma_t}{\sqrt{n}} - \mu_1 - t\sigma \leq \bar{x} - \mu_1 \leq \mu_1 + \delta_1 \sigma + \frac{K\sigma_t}{\sqrt{n}} - \mu_1 - t\sigma \mid \mu_1, \sigma\right) + \\
&\pi_2 P\left(\mu_2 - \frac{K\sigma_t}{\sqrt{n}} - \mu_2 - t\sigma \leq \bar{x} - \mu_2 \leq \mu_2 + \frac{K\sigma_t}{\sqrt{n}} - \mu_2 - t\sigma \mid \mu_2, \sigma\right) + \\
&\pi_3 P\left(\mu_3 + \delta_3 \sigma - \frac{K\sigma_t}{\sqrt{n}} - \mu_3 - t\sigma \leq \bar{x} - \mu_3 \leq \mu_3 + \delta_3 \sigma + \frac{K\sigma_t}{\sqrt{n}} - \mu_3 - t\sigma \mid \mu_3, \sigma\right)
\end{aligned} \tag{4.8}$$

Multiplicando cada desigualdade por $\frac{\sqrt{n}}{\sigma}$, obtemos:

$$\begin{aligned}
&\pi_1 P\left(\delta_1 \sqrt{n} - \frac{K\sigma_t}{\sigma} - t\sqrt{n} \leq z \leq \delta_1 \sqrt{n} + \frac{K\sigma_t}{\sigma} - t\sqrt{n} \mid \mu_1, \sigma\right) + \\
&\pi_2 P\left(-\frac{K\sigma_t}{\sigma} - t\sigma \leq z \leq \frac{K\sigma_t}{\sigma} - t\sigma \mid \mu_2, \sigma\right) + \\
&\pi_3 P\left(\delta_3 \sqrt{n} - \frac{K\sigma_t}{\sigma} - t\sigma \leq z \leq \delta_3 \sqrt{n} + \frac{K\sigma_t}{\sigma} - t\sigma \mid \mu_3, \sigma\right)
\end{aligned} \tag{4.9}$$

Considerando a abordagem de um modelo aditivo da Análise de Variância (ANOVA), onde a variância total é calculada pela variância interna de cada subpopulação, σ_i^2 , e a variabilidade é determinada pelas diferenças nas médias entre subpopulações (SPIEGEL *et al.*, 2013), pode-se substituir a relação $\frac{\sigma_t}{\sigma} = \sqrt{1 + \frac{1}{l} \sum_l \delta_l^2}$ na Equação 4.9. Assim no contexto de distribuição de probabilidade de mistura finita, temos:

$$\begin{aligned}
P(\bar{x} \in [LIC, LSC]) = & \\
\pi_1 P\left(\delta_1\sqrt{n} - K\sqrt{1 + \frac{1}{l}\sum_l \delta_l^2} - t\sqrt{n} \leq z \leq \delta_1\sqrt{n} + K\sqrt{1 + \frac{1}{l}\sum_l \delta_l^2} - t\sqrt{n} \middle| \mu_1, \sigma\right) + & \\
\pi_2 P\left(-K\sqrt{1 + \frac{1}{l}\sum_l \delta_l^2} - t\sigma \leq z \leq K\sqrt{1 + \frac{1}{l}\sum_l \delta_l^2} - t\sigma \middle| \mu_2, \sigma\right) + & \\
\pi_3 P\left(\delta_3\sqrt{n} - K\sqrt{1 + \frac{1}{l}\sum_l \delta_l^2} - t\sigma \leq z \leq \delta_3\sqrt{n} + K\sqrt{1 + \frac{1}{l}\sum_l \delta_l^2} - t\sigma \middle| \mu_3, \sigma\right) & \quad (4.10)
\end{aligned}$$

Em termos de distribuição de probabilidade normal, temos a seguinte equação:

$$\begin{aligned}
P(\bar{x} \in [LIC, LSC]) = \pi_1 \left[\Phi\left(\delta_1\sqrt{n} + K\sqrt{1 + \frac{1}{l}\sum_l \delta_l^2} - t\sqrt{n}\right) - \Phi\left(\delta_1\sqrt{n} - \right. \right. & \\
\left. \left. K\sqrt{1 + \frac{1}{l}\sum_l \delta_l^2} - t\sqrt{n}\right) \right] + \pi_2 \left[\Phi\left(K\sqrt{1 + \frac{1}{l}\sum_l \delta_l^2} - t\sigma\right) - \Phi\left(-K\sqrt{1 + \frac{1}{l}\sum_l \delta_l^2} - t\sigma\right) \right] + & \\
\pi_3 \left[\Phi\left(\delta_3\sqrt{n} + K\sqrt{n}\sqrt{1 + \frac{1}{l}\sum_l \delta_l^2} - t\sigma\right) - \Phi\left(\delta_3\sqrt{n}\sqrt{1 + \frac{1}{l}\sum_l \delta_l^2} - K\sqrt{n}\sqrt{1 + \frac{1}{l}\sum_l \delta_l^2} - \right. \right. & \\
\left. \left. t\sigma\right) \right] & \quad (4.11)
\end{aligned}$$

Para este caso, os parâmetros para $P(\bar{x} \in [LIC, LSC])$ são $\omega = \{\pi_1, \pi_2, \pi_3, k, \delta, t, n\}$ e o *Average Run Length* (ARL), usado para detectar um ponto fora dos Limites de Controle, é dado por:

$$ARL = \frac{1}{1 - P(\bar{x} \in [LIC, LSC])} \quad (4.12)$$

4.2 CASO COM μ DESCONHECIDO

Segundo Vicentin et al. (2018), para, μ_1, μ_2 e μ_3 desconhecidos, $\bar{x} \in \left(\bar{\bar{X}} - \frac{K\sigma_t\sqrt{n}}{\sigma}; \bar{\bar{X}} + \frac{K\sigma_t\sqrt{n}}{\sigma} \right) = (\widehat{LIC}; \widehat{LSC})$, os $(\widehat{LIC}; \widehat{LSC})$ são suposições de Limites de Controle. Assumindo $(X_{i,1}, \dots, X_{i,n})$, $i = 1, 2, \dots, m$ são amostras independentes de variáveis aleatórias de subpopulações de mistura finita. A estimativa da média é dada por:

$$\bar{\bar{x}} = \frac{1}{mn} \sum_{j=1}^m \sum_{i=1}^n x_{ij} \quad (4.13)$$

Na fase II, o tamanho da amostra n é tirado sem distinção de categoria ou subpopulação e a média da amostra é representada por \bar{X}_i a partir de $\{X_{i,1}, X_{i,2}, \dots, X_{i,n}\}$. Desenvolvendo, temos a probabilidade de a média amostral cair dentro dos Limites de Controle, como segue:

$$\begin{aligned} P(\widehat{LIC}_{\bar{x}} \leq \bar{X} \leq \widehat{LSC}) = & \pi_1 \left[\Phi \left(\frac{w}{\sqrt{m}} + \delta_1 \sqrt{n} + K \sqrt{1 + \frac{1}{l} \sum_l \delta_l^2} - t \sqrt{n} \right) - \Phi \left(\frac{w}{\sqrt{m}} + \delta_1 \sqrt{n} - \right. \right. \\ & \left. \left. K \sqrt{1 + \frac{1}{l} \sum_l \delta_l^2} - t \sqrt{n} \right) \right] + \pi_2 \left[\Phi \left(\frac{w}{\sqrt{m}} + K \sqrt{1 + \frac{1}{l} \sum_l \delta_l^2} - t \sigma \right) - \Phi \left(\frac{w}{\sqrt{m}} - \right. \right. \\ & \left. \left. K \sqrt{1 + \frac{1}{l} \sum_l \delta_l^2} - t \sigma \right) \right] + \pi_3 \left[\Phi \left(\frac{w}{\sqrt{m}} + \delta_3 \sqrt{n} + K \sqrt{1 + \frac{1}{l} \sum_l \delta_l^2} - t \sigma \right) - \Phi \left(\frac{w}{\sqrt{m}} + \delta_3 \sqrt{n} - \right. \right. \\ & \left. \left. K \sqrt{1 + \frac{1}{l} \sum_l \delta_l^2} - t \sigma \right) \right] \end{aligned} \quad (2.14)$$

Os parâmetros envolvidos, neste caso, quando as médias são desconhecidas, são $\omega = \{\pi_1, \pi_2, \pi_3, K, \delta, t, m, n\}$, onde π_i são os pesos conhecidos e ω representa o Z da distribuição normal, e a esperança de ARL ou (E(ARL)), conhecida como *Average of Average*

Run Length (AARL) –E(ARL) a Média de ARL — (OPRIME *et al.*, 2016) até a ocorrência de um ponto fora dos Limites de Controle é dada por:

$$AARL(\pi_1, \pi_2, \pi_3, K, \delta_1, \delta_2, \delta_3, t, m, n) = \int_{-\infty}^{\infty} \frac{1}{[1-P(LIC_{\bar{X}} \leq \bar{X} \leq LSC)]} \varphi(w) dw \quad (4.15)$$

Onde φ é a função de densidade de probabilidade da distribuição normal padrão.

Além disso, a expansão para $l > 3$ pode ser feita da mesma maneira, conforme Vicentin *et al.* (2018).

Em termos de função de probabilidade de mistura, o tipo de erro I depende dos Limites de Controle. Portanto, segundo Vicentin *et al.* (2018) e Titterington *et al.* (1987) podemos reescrever a Equação 4.15 da seguinte forma:

AARL =

$$= \frac{1}{\pi_1 \int_{\frac{\bar{X} + K\sigma_t\sqrt{n}}{\sigma}}^{\infty} f(\bar{x}_1) dx_1 + \pi_2 \int_{\frac{\bar{X} + K\sigma_t\sqrt{n}}{\sigma}}^{\infty} f(\bar{x}_2) dx_2 + \pi_3 \int_{\frac{\bar{X} + K\sigma_t\sqrt{n}}{\sigma}}^{\infty} f(\bar{x}_3) dx_3 + \pi_1 \int_{-\infty}^{\frac{\bar{X} - K\sigma_t\sqrt{n}}{\sigma}} f(\bar{x}_1) dx_1 + \pi_2 \int_{-\infty}^{\frac{\bar{X} - K\sigma_t\sqrt{n}}{\sigma}} f(\bar{x}_2) dx_2 + \pi_3 \int_{-\infty}^{\frac{\bar{X} - K\sigma_t\sqrt{n}}{\sigma}} f(\bar{x}_3) dx_3} \quad (4.16)$$

Nesta equação, o problema é encontrar o valor tradicional da constante K para a probabilidade do erro tipo I nominal ou teórico, usualmente definido na distribuição normal, com 3 desvios padrão, que estabelece que $\alpha = 0,0027$. Foi proposto, neste caso, que os Limites de Controle Superior e Inferior (LSC e LIC, respectivamente) sejam determinados pelo problema de mistura geral para o erro α especificado, para isso conforme VICENTIN *et al.* (2018); TITTERINGTON *et al.* (1987) segue a seguinte equação:

$$\alpha = \pi_1 \int_{LSC}^{\infty} f(\bar{x}_1) dx_1 + \pi_2 \int_{LSC}^{\infty} f(\bar{x}_2) dx_2 + \dots + \pi_p \int_{LSC}^{\infty} f(\bar{x}_p) dx_p + \pi_1 \int_{-\infty}^{LIC} f(\bar{x}_1) dx_1 + \pi_2 \int_{-\infty}^{LIC} f(\bar{x}_2) dx_2 + \dots + \pi_p \int_{-\infty}^{LIC} f(\bar{x}_p) dx_p \quad (4.17)$$

Onde p é um número de funções de probabilidade e os parâmetros são os pesos da mistura das distribuições. O objetivo é encontrado para o LSC e LIC para $\alpha = 0,0027$ usando métodos numérico-teóricos para 3 desvios padrão.

4.3 SÍNTESE DAS DUAS PROPOSTAS

Como apresentado na introdução do presente trabalho, existe um consenso sobre o uso de gráficos de controle tradicional para o Processo de Múltiplos Fluxos (MSP). No entanto, a construção de apenas um gráfico de controle para o monitoramento estatístico desse tipo de processo ainda é discutida e nota-se, na literatura, a ausência de estudos que comparam as abordagens de mistura finita e multivariada como ACP e gráfico de controle de T^2 -Hotelling. Para desenvolver um gráfico de controle que atenda a essa situação, foi proposta a seguinte questão de pesquisa: qual dessas duas abordagens, a de mistura finita e a multivariada, tem melhor desempenho e é mais fácil de ser implementada?

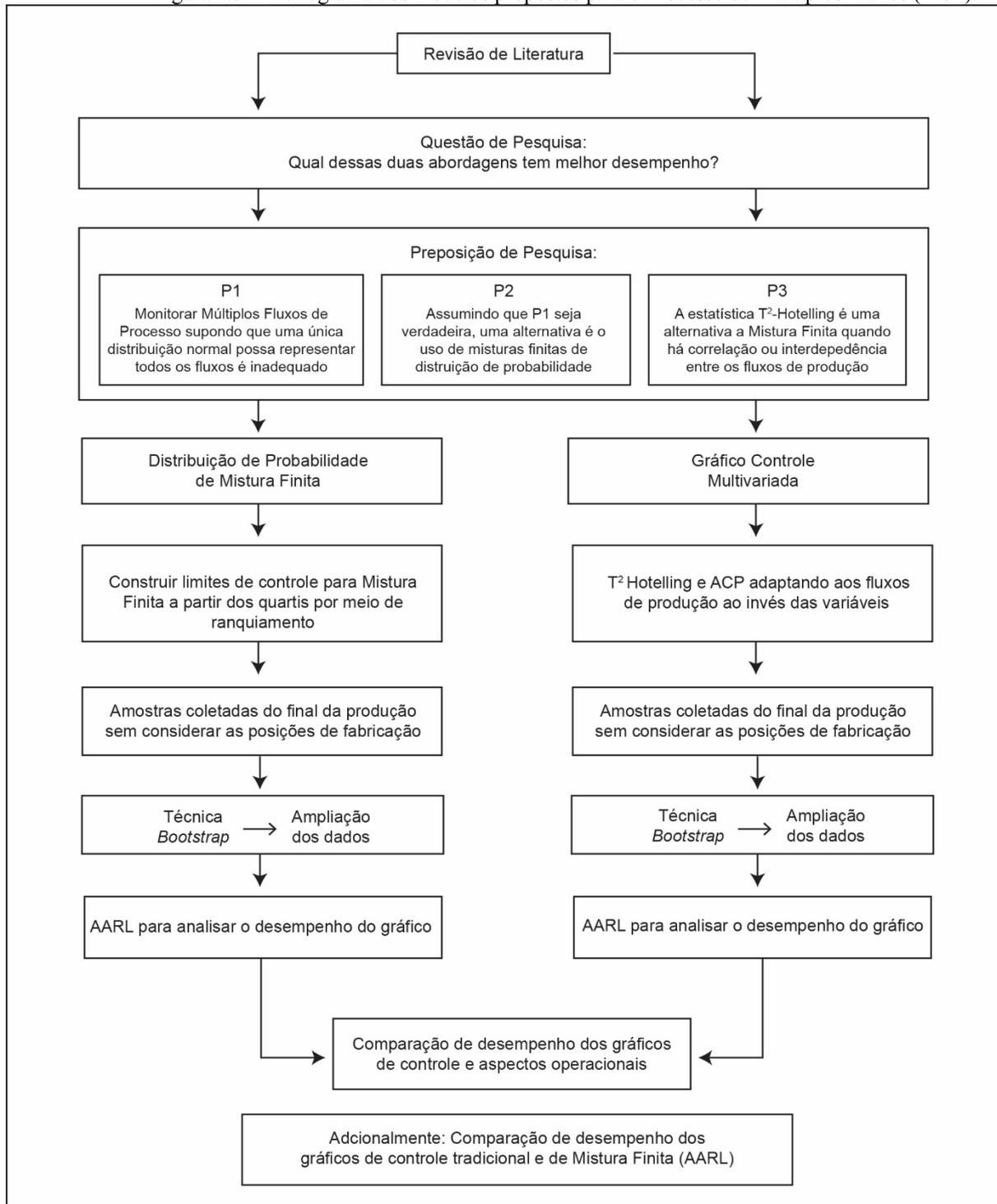
Ao observar a literatura e com a questão de pesquisa em mente, foram analisadas três proposições de pesquisa (P1, P2 e P3). A primeira (P1) consiste em monitorar os múltiplos fluxos do processo, supondo que representar todos os fluxos por meio de uma única distribuição normal é inadequado. A segunda (P2) parte do pressuposto que P1 é verdadeira e coloca que uma alternativa é o uso de misturas finitas de distribuição de probabilidade. Por fim, a última proposição (P3) apresenta que a carta de controle T^2 de Hotelling é uma alternativa à mistura finita quando há correlação ou interdependência entre os fluxos de produção.

Para respondê-las, primeiramente, foi proposto o gráfico de controle de distribuições de probabilidade de mistura finita, como ilustrado na Figura 4.3. A amostra foi

coletada sistematicamente considerando 8 itens dos 5 fluxos ou 5 posições de fabricação no cilindro, portanto identificou-se que cada fluxo foi composto de 8 unidades, sendo eles subdivididos. Entretanto, essas posições foram desconsideradas, de modo que utilizou-se a amostra de quarenta elementos colocados os em ordem crescente. Para o desenvolvimento do gráfico de controle, foram determinados os Limites de Controle por meio da análise dos quartis e do ranqueamento dos dados. Então, os quartis foram separados em quatro posições. É importante considerar que a amostra foi retirada no final da linha de produção, tendo em vista as quatro distribuições normais presentes nessa mistura finita. Para sua construção, foi necessário o uso do *software Maple 13*. A partir dele foi possível realizar a simulação da proposta no monitoramento estatístico do processo. Utilizou-se a técnica *Bootstrap* para expandir a amostra mestre o caso ilustrativo de uma indústria do setor alimentício.

Posteriormente, como mostra a Figura 4.3, o gráfico de controle multivariado foi desenvolvido utilizando a Análise de Componentes Principais (ACP) e a carta de controle T^2 de Hotelling. Para tal, as variáveis de seu modelo tradicional para os múltiplos fluxos, foram compostas pelas cinco posições do cilindro na saída final da linha de produção, já que, como observado na literatura (Apêndice 1), não se encontra o emprego desse método empírico para a situação apresentada. O *software Statistica 13.0* foi usado para simular a proposta, possibilitando a construção do gráfico. Novamente, fez-se uso da técnica *Bootstrap* para expandir a amostra mestre. Contudo, diferentemente do método anterior, optou-se por considerar as cinco posições do cilindro no final da linha de produção, junto a suas correlações, a fim de ilustrar o monitoramento estatístico de processo proposto para essa situação. Sendo assim, por meio da AARL, ele foi testado e comparado com o gráfico de controle de distribuições de probabilidade de mistura finita. Adicionalmente, o gráfico de mistura finita foi também comparado com o parâmetro K do gráfico de controle tradicional de Shewhart, em termos de ARRL.

Figura 4.3 – Fluxograma dos modelos propostos para o Processo de Múltiplos Fluxos (MSP).

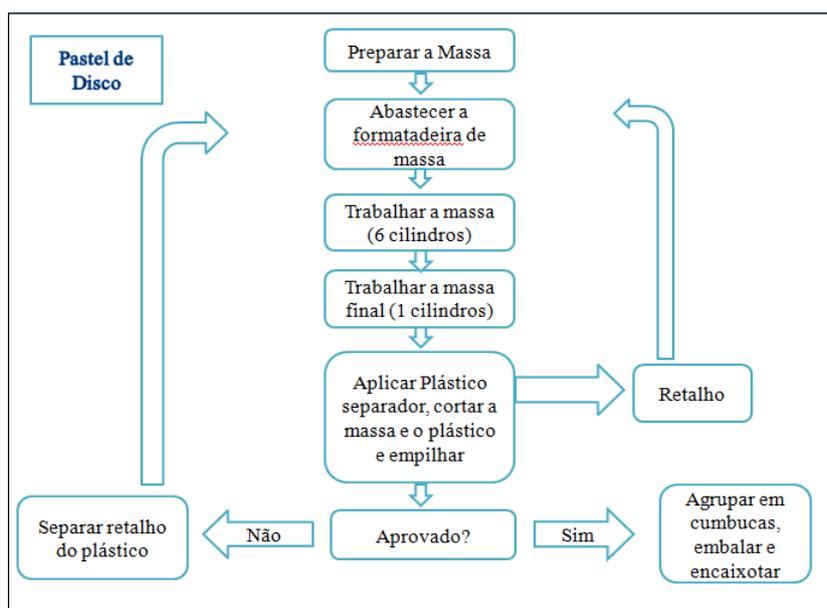


Fonte: figura elaborada pela autora.

5 APLICAÇÃO DOS MODELOS PROPOSTOS EM UM CASO ILUSTRATIVO COM DADOS REAIS

Para aplicação da melhoria proposta, foi selecionado um processo de uma empresa no setor de alimentos que fabrica massa à base de trigo para uma gama de produtos. O exemplo mostrado é semelhante aos processos das indústrias metalúrgicas, no qual o produto pode passar por diferentes linhas de produção, com parâmetros e distribuições de probabilidade distintos. A característica geral do processo é ilustrada na Figura 5.1. A manufatura começa com a preparação da massa, colocando o trigo na base da calha para homogeneização. Em seguida, a massa é removida em blocos, colocada em molde de cilindro e processada para obter uma espessura uniforme. Depois, a massa é dobrada, colocada de volta no cilindro e cortada em discos, que saem em fileiras horizontais dispostas em cinco posições.

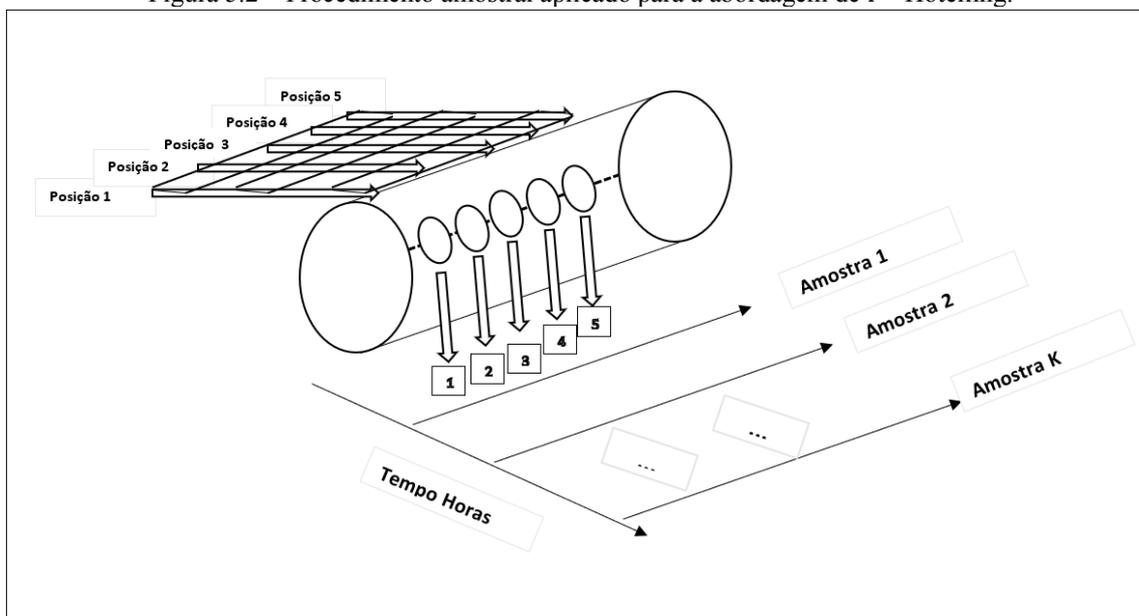
Figura 5.1 – Esquema da produção do processo de linha de massa folhada.



Fonte: figura elaborada pela autora.

A Figura 5.2 ilustra a estratégia de amostragem utilizada para o caso do uso do T^2 de Hotelling, considerando as posições de fabricação de 1 a 5 e amostras coletadas no final da produção de 1 a k.

Figura 5.2 – Procedimento amostral aplicado para a abordagem de T^2 -Hotelling.



Fonte: figura elaborada pela autora.

O peso foi medido em quilogramas e o tempo em horas. Ademais, propõe-se a aplicação de dois modelos: (i) mistura finita e (ii) gráfico T^2 -Hotelling para um Processo de Múltiplos Fluxos (MSP). Para construir o gráfico de controle multivariado foram extraídos oito itens dos cinco fluxos — sendo cada fluxo uma posição. Os resultados são mostrados na Tabela 5.1 e indicam correlação entre as posições dos referentes e os cinco fluxos de produção ilustrados na Figura 4.2.

Tabela 5.1 – Média e desvio padrão para cada posição e as correlações obtidas entre elas.

	Média	Desv.Pad	P1	P2	P3	P4	P5
P1	542,62	35,85	1,000000	0,635709	-0,6253	0,887908	0,882711
P2	556,12	24,37	0,635709	1,000000	-0,6480	0,859625	0,659989
P3	554,75	15,21	-0,62530	-0,64801	1,000000	-0,67724	-0,4201
P4	557,00	17,11	0,887908	0,859625	-0,6772	1,000000	0,801669
P5	519,25	21,76	0,882711	0,659989	-0,42011	0,801669	1,000000

Fonte: tabela elaborada pela autora.

Havendo correlações com valores expressivos entre os fluxos de produção, justifica-se o uso da ACP e do T^2 de Hotelling. Assim, o gestor poderá construir modelos gráficos para monitorar o processo sem a necessidade de construir um gráfico para cada uma das 5 posições do cilindro. Observou-se, então, que as relações que determinam a covariância (correlação) entre os fluxos foram estabelecidas, mantendo originalmente a mesma média e mesmo desvio padrão — como mostra a Tabela 5.2 em uma simulação de 1000 amostras para cada posição, totalizando 5000 amostras.

Tabela 5.2 – Reprodução de 1000 amostras aplicando o mecanismo de interdependência entre os fluxos mostrado na Figura 5.2.

Posição	N	Média	Mediana	Moda	Desv. Padrão
P5	1000	518,65	461,33	586,38	21,41
P1	1000	543,27	423,08	668,87	42,00
P3	1000	554,41	495,41	611,21	19,19
P4	1000	556,04	482,37	621,55	20,46
P2	1000	556,84	450,87	655,51	31,74

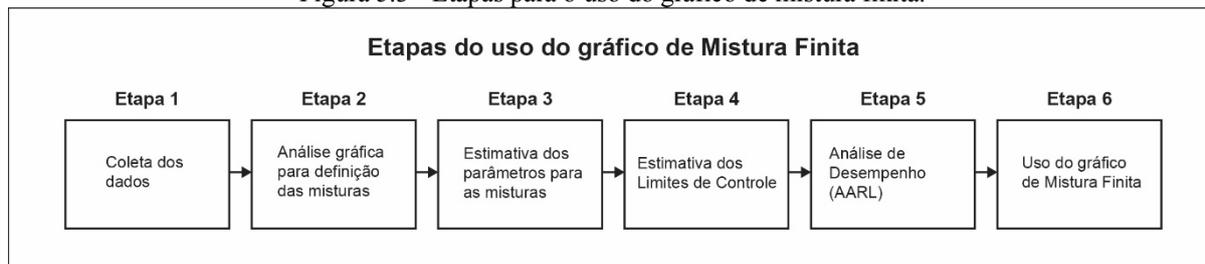
Fonte: tabela elaborada pela autora.

A seguir, são apresentados os roteiros para elaboração de cada modelo proposto.

5.1 MODELO DE DISTRIBUIÇÃO DE PROBABILIDADE DE MISTURA FINITA

O modelo de mistura finita pode ser desenvolvido conforme as etapas da Figura 5.3. Primeiramente, de acordo com a Figura 5.2, foi feita a coleta dos dados. Em seguida, foi realizada uma análise gráfica para definição das k misturas e, depois, estimou-se os parâmetros para as k misturas. Por fim, fez-se a estimativa dos Limites de Controle, seguida da análise de desempenho do modelo de mistura finita e do uso do modelo proposto.

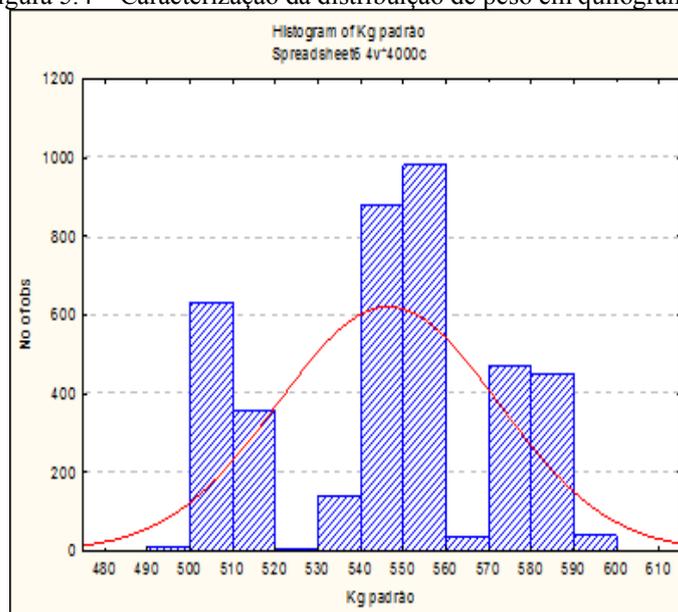
Figura 5.3 - Etapas para o uso do gráfico de mistura finita.



Fonte: figura elaborada pela autora.

O primeiro passo, que integra o processo mostrado na Figura 5.4, é identificar a mistura de subpopulação. Para isso, foram coletados 40 itens de cada fluxo em um intervalo de 30 minutos, de modo a analisar o peso enquanto característica de qualidade. Posteriormente, os itens foram expandidos utilizando a técnica *Bootstrap* para 4000 unidades. A posição de fabricação dos itens foi desprezada, buscando verificar uma mistura de distribuições de probabilidade referente à totalidade dos dados. Portanto, é possível identificar a quantidade dessas misturas e estimar os parâmetros para cada uma das k funções de probabilidade.

Figura 5.4 – Caracterização da distribuição de peso em quilogramas.

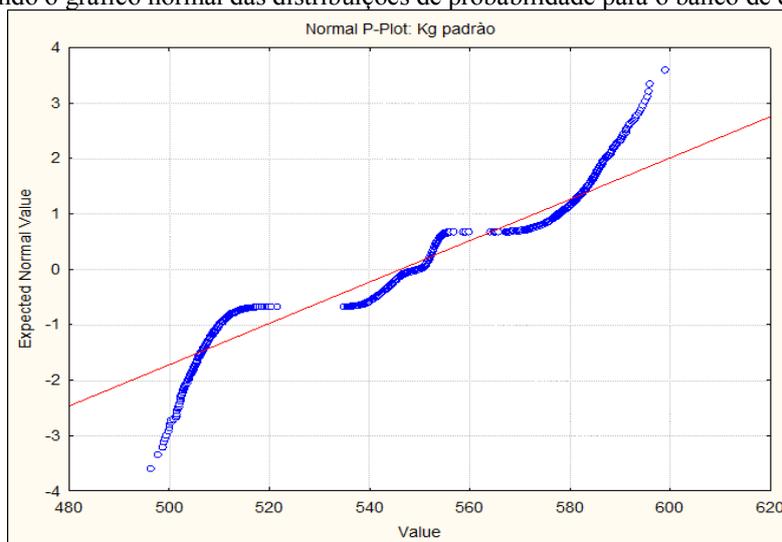


Fonte: figura elaborada pela autora.

De acordo com Vicentin et al. (2018), uma das técnicas para atingir esses objetivos é por meio de gráficos baseados na densidade da função de probabilidade. A Figura 5.4 mostra

pelo menos três distribuições de probabilidade e a Figura 5.5 mostra o gráfico normal das distribuições de probabilidade para o banco de dados da Figura 5.4. A Figura 5.5 indica dois pontos de inflexão claros nas extremidades da curva padrão e duas ou mais distribuições de probabilidade no centro da interseção.

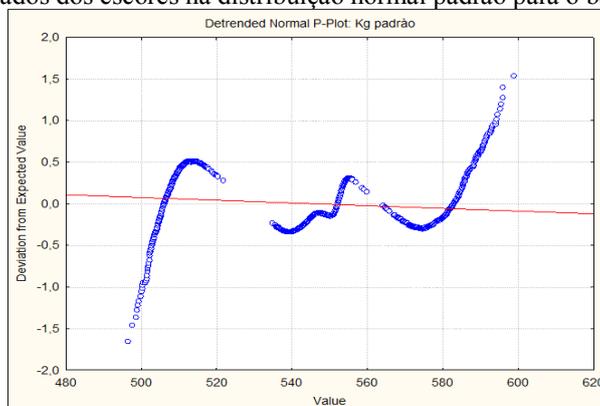
Figura 5.5 – Plotando o gráfico normal das distribuições de probabilidade para o banco de dados da Figura 5.4.



Fonte: figura elaborada pela autora.

A Figura 5.6 mostra a mesma informação, mas há desvios na pontuação das distribuições normais para cada valor na escala real. Se houver uma mistura pura sem subpopulações, então pode ser esperado que os padrões normais de valores flutuem em torno de zero. O desafio é descobrir quantas subpopulações estão no banco de dados.

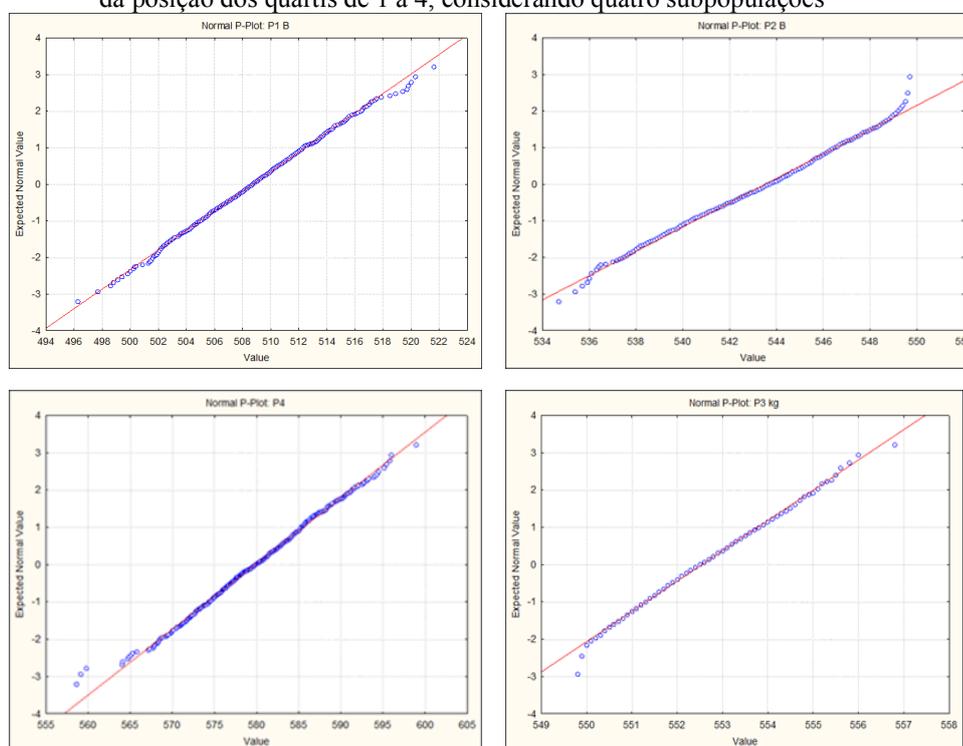
Figura 5.6 – Desvios esperados dos escores na distribuição normal padrão para o banco de dados da Figura 5.4.



Fonte: figura elaborada pela autora.

Como mostrado na Figura 5.7, os dados de massa foram separados em quatro subpopulações realizadas a partir do ranqueamento baseado nos seus quartis com funções de densidade de probabilidade gaussiana e desconsiderando as posições do cilindro, visto que o teste Kolmogorov–Smirnov e o teste gráfico para verificar a normalidade não indicou evidência estatística para rejeitar a hipótese de não-normalidade para cada uma das distribuições de probabilidade.

Figura 5.7 – Sentido horário da esquerda para a direita, plotando os escores normais da posição dos quartis de 1 a 4, considerando quatro subpopulações



Fonte: figura elaborada pela autora.

Com as distribuições normais de probabilidade identificadas para todas as subpopulações, é possível observar, na Tabela 5.3, as estatísticas para estimativas dos parâmetros de interesse.

Tabela 5.3 – Parâmetros estimados das subpopulações.

Sub populações	N	Média	Mediana	Moda	Mínimo	Máximo	Des.Pad	Assimetria	Curtoses
1	1000	508,7	508,7	508,1	496,3	521,6	3,7319	0,0824	0,0826
2	1002	543,5	543,7	544,6	534,7	549,7	2,9900	-0,1900	-0,4133
3	998	552,5	552,5	Múltiplo	549,8	556,8	1,2245	0,1583	-0,2983
4	1001	579,9	579,9	Múltiplo	558,6	598,9	5,6569	-0,0395	0,2906
Populações	4000	546,2	549,7	Múltiplo	496,3	598,9	25,6965	-0,1815	-0,9390

Fonte: tabela elaborada pela autora.

Tabela 5.4 – Parâmetros estimados de π e δ .

Subpopulações	$\pi_j(p)$ *	δ	$\sqrt{1 + \frac{1}{l} \sum_t \delta_t^2}$
1	0,2500	-9,9771	
2	0,2505	-0,6535	6,84919
3	0,2495	1,6908	
4	0,2503	8,9898	

* π_j representa também o peso de cada subpopulação na população geral

Fonte: tabela elaborada pela autora.

Na Tabela 5.4, π_j representa a probabilidade de pertencer à j -ésima subpopulações ($j = 1, \dots, p$), onde p é a quantidade de subpopulações. Já δ indica o número de deslocamento do desvio padrão da diferença entre as subpopulações. Observe na Tabela 5.4 que $\frac{\sigma_t}{\sigma} \cong 6,84919$ e que, na Tabela 5.3, $\sigma_t = 25,6965$. Isso resulta em $\sigma = 3,7559$, sendo σ o desvio padrão da subpopulação (supondo que $\sigma = \sigma_1 = \sigma_2 = \sigma_3 = \sigma_4$).

5.1.1 Determinação dos Limites de Controle (LCs)

Na função matemática da AARL, pode-se escrever que os parâmetros envolvidos são $\omega = \{\pi_1, \pi_2, \pi_3, \pi_4, K, \delta_1, \delta_2, \delta_3, \delta_4, t, m, n\}$. Entretanto, como optou-se por considerar todas Funções de Densidade de Probabilidade (FDPs) com mesmo peso, tem-se que

$\pi \cong \pi_1 \cong \pi_2, \cong \pi_4$. Assim, o número médio esperado de amostras, até a ocorrência de um ponto fora dos Limites de Controle, é dado pela expressão abaixo. É importante lembrar que, quando $t = 0$, o processo está sob controle.

$$E(ARL) = AARL(\pi, K, \delta_1, \delta_2, \delta_3, \delta_4, t, m, n) = \int_{-\infty}^{\infty} \frac{1}{[1 - P(\widehat{LIC}_{\bar{X}} \leq \bar{X} \leq \widehat{LSC})]} \varphi(w) dw. \quad (5.1)$$

Onde,

$$\begin{aligned} P(\widehat{LIC}_{\bar{X}} \leq \bar{X} \leq \widehat{LSC}) = & \pi_1 \left[\Phi \left(\frac{w}{\sqrt{m}} + \delta_1 \sqrt{n} + K \sqrt{1 + \frac{1}{l} \sum_l \delta_l^2} - t \sqrt{n} \right) - \Phi \left(\frac{w}{\sqrt{m}} + \delta_1 \sqrt{n} - K \sqrt{1 + \frac{1}{l} \sum_l \delta_l^2} - t \sqrt{n} \right) \right] + \\ & \pi_2 \left[\Phi \left(\frac{w}{\sqrt{m}} + \delta_2 \sqrt{n} + K \sqrt{1 + \frac{1}{l} \sum_l \delta_l^2} - t \sigma \right) - \Phi \left(\frac{w}{\sqrt{m}} + \delta_2 \sqrt{n} - K \sqrt{1 + \frac{1}{l} \sum_l \delta_l^2} - t \sigma \right) \right] + \pi_3 \left[\Phi \left(\frac{w}{\sqrt{m}} + \delta_3 \sqrt{n} + \right. \right. \\ & \left. \left. K \sqrt{1 + \frac{1}{l} \sum_l \delta_l^2} - t \sigma \right) - \Phi \left(\frac{w}{\sqrt{m}} + \sqrt{n} - K \sqrt{1 + \frac{1}{l} \sum_l \delta_l^2} - t \sigma \right) \right] + \pi_4 \left[\Phi \left(\frac{w}{\sqrt{m}} + \delta_4 \sqrt{n} + K \sqrt{1 + \frac{1}{l} \sum_l \delta_l^2} - t \sigma \right) - \right. \\ & \left. \Phi \left(\frac{w}{\sqrt{m}} + \delta_4 \sqrt{n} - K \sqrt{1 + \frac{1}{l} \sum_l \delta_l^2} - t \sigma \right) \right] \end{aligned} \quad (5.2)$$

De maneira a auxiliar na definição do gráfico de controle, em termos de tamanho e quantidade de amostras, fornece-se instruções para os seguintes casos: quando a produção possui vários fluxos e quando os dados pertencem às distribuições de probabilidade de mistura finita. Nestes cenários, $\mu_1 \neq \mu_2 \neq \dots \neq \mu_l$, $l = \{1, 2, \dots, i\}$. Logo, os Limites de Controle podem ser obtidos por:

$$\begin{aligned} \widehat{LIC} &= \bar{\bar{X}} - K \frac{\sigma_T}{\sqrt{n}} \\ \widehat{LSC} &= \bar{\bar{X}} + K \frac{\sigma_T}{\sqrt{n}} \end{aligned} \quad (5.3)$$

Sendo assim, essa é uma relação entre o desvio padrão global e o desvio padrão interno de cada subpopulação, $\frac{\sigma_t}{\sigma} = \tau$. Então, foi proposto que os LCs fossem calculados da seguinte forma:

$$\begin{aligned}\widehat{LIC} &= \bar{X} - K \frac{\tau\sigma}{\sqrt{n}} \\ \widehat{LSC} &= \bar{X} + K \frac{\tau\sigma}{\sqrt{n}}\end{aligned}\quad (5.4)$$

O procedimento indicado possui etapas similares ao clássico gráfico de controle de Shewhart \bar{X} . Ou seja, na Fase I temos m amostras de tamanho n de subpopulação especificada, além de estimativas de desvio padrão interno e de \bar{X} . Já na fase II temos uma amostra do tamanho n de uma subpopulação especificada e o valor de \bar{X} calculado.

O modelo desenvolvido nas Seções 4.1 e 4.2 foi usado para obter os resultados mostrados na Tabela 5.5, na qual pode-se observar os valores da AARL e o desempenho do gráfico de controle de mistura em detectar um alarme falso, usando o programa *Maple*. Alguns resultados são mostrados a seguir para valores diferentes de K .

Tabela 5.5 – Estimativas dos parâmetros gerados pelo *Maple*.

K	$AARL_0$	K	$AARL_0$
3,505	99,13	3,550	203,2
3,510	106,9	3,555	221,3
3,515	115,4	3,560	241,2
3,520	124,7	3,565	263,2
3,525	134,9	3,570	287,5
3,530	146,1	3,575	314,4
3,535	158,4	3,580	344,2
3,540	172,0	3,585	377,2
3,545	186,8	3,590	413,9

Fonte: tabela elaborada pela autora.

Para o procedimento mostrado, assumindo probabilidade de erro tipo I próximo de 0,0027 ($\alpha = 1/370,4$), foi proposto o uso de $K = 3,585$. Note, na Tabela 5.5, o valor de 377,2 para a AARL. Portanto, para a análise do desempenho do gráfico usando a AARL, que representa a média da ARL, foi alterado o valor de t que representa deslocamentos em torno da média. Então, tem-se as variações de sigma para $t = [0; 0,25; -0,25; 0,5; -0,5; 1,0; -1,0; 2,0; -2,0; 3,0; -3,0]$ com o mesmo valor de K mostrado na Tabela 5.5. Este resultado é mostrado na Tabela 5.6.

Tabela 5.6 – Valores da AARL para diferentes valores de t (quando $t \neq 0$, fora de controle).

K	t								
	0	0,25	-0,25	0,5	-0,5	1,0	-1,0	2,0	-2,0
3,505	99,13	33,69	349,1	14,87	686,7	5,759	93,26	4,013	5,679
3,510	106,9	35,73	383,4	15,52	762,1	5,862	100,50	4,014	5,777
3,515	115,4	37,92	421,6	16,22	846,7	5,970	108,40	4,016	5,880
3,520	124,7	40,28	464,1	16,96	941,7	6,083	117,00	4,018	5,989
3,525	134,9	42,84	511,4	17,75	1048,0	6,203	126,50	4,019	6,103
3,530	146,1	45,61	564,2	18,60	1169,0	6,329	136,90	4,021	6,224
3,535	158,4	48,60	623,0	19,50	1304,0	6,462	148,30	4,024	6,351
3,540	172,0	51,84	688,8	20,47	1457,0	6,602	160,80	4,026	6,486
3,545	186,8	55,36	762,3	21,51	1629,0	6,750	174,50	4,029	6,627
3,550	203,2	59,17	844,6	22,62	1823,0	6,906	189,60	4,031	6,776
3,555	221,3	63,32	936,8	23,81	2043,0	7,071	206,30	4,034	6,934
3,560	241,2	67,82	1040,0	25,09	2292,0	7,245	224,70	4,038	7,100
3,565	263,2	72,72	1156,0	26,46	2574,0	7,428	244,90	4,041	7,276
3,570	287,5	78,06	1287,0	27,94	2894,0	7,622	267,30	4,045	7,461
3,575	314,4	83,88	1434,0	29,53	3258,0	7,827	292,10	4,050	7,656
3,580	344,2	90,22	1599,0	31,23	3671,0	8,043	319,40	4,054	7,863
3,585	377,2	97,15	1785,0	33,07	4140,0	8,271	349,80	4,059	8,081
3,590	413,9	104,70	1995,0	35,05	4676,0	8,513	383,40	4,064	8,312

Fonte: tabela elaborada pela autora.

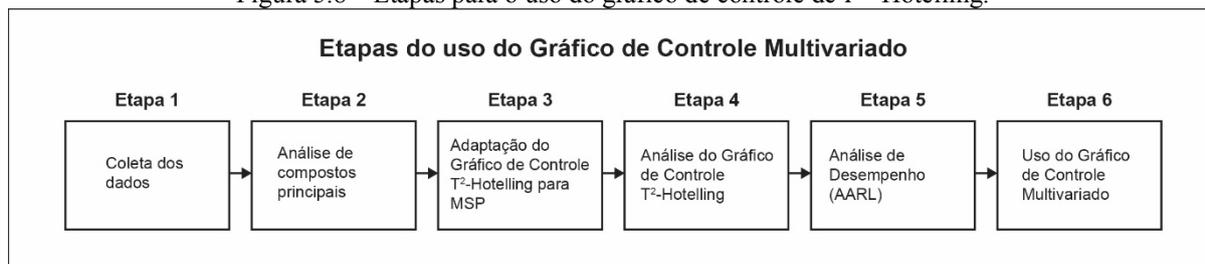
Na Tabela 5.6, pode-se observar que diferentes AARLs são encontradas se o desvio na média ocorrer para t positivo ou negativo. AARLs menores ocorrem para pequenos desvios positivos na média (t positivo). Do outro lado, para t negativo, encontramos valores maiores de AARL. Este é um resultado especial para a função de densidade de probabilidade de mistura finita. Ainda, a Tabela 5.6 mostra desvio positivo na média quando o processo está fora de controle. A linha em destaque apresenta o comportamento do valor mais próximo ao teórico — de 370,4 —, que indica a ocorrência de um alarme falso, indicando que o processo estaria fora de controle, a cada 370,4 amostras. Nessa proposta de gráfico, o LC — três-sigma — seria 377,2. Ou seja, haverá um alarme falso entendendo que o processo estaria fora de controle a cada 377,2 amostras. Assim, podemos notar, por essa linha em destaque, que os desvios positivos têm maior desempenho — ou maior poder de teste — que o gráfico de controle tradicional de Shewhart.

No próximo item será apresentada a aplicação do gráfico multivariado, no caso da APC e do T^2 de Hotelling, como segunda proposta para o MSP.

5.2 PROCEDIMENTO PARA O CASO DA ACP E DO T^2 DE HOTELLING

Como já mencionado, foram coletados dados reais com 40 itens referentes às 5 posições do cilindro e, posteriormente, os resultados foram expandidos utilizando a técnica *Bootstrap* para 5000 elementos (1000 de cada posição). Por isso, foi possível observar o Processo de Múltiplos Fluxos (MSP) por meio de suas posições. Contudo, um estudo foi realizado utilizando técnicas de estatística multivariada, como a ACP e o T^2 de Hotelling, no intuito de concentrar as informações em apenas um gráfico de controle e visando o monitoramento otimizado do processo. Para tal, o processo foi dividido em duas fases de monitoramento (Fase I e Fase II), sendo a primeira feita com os 40 dados originais e a segunda com os 5000 dados originais expandidos.

De modo a desenvolver o método T^2 de Hotelling em gráficos de controle, foram estipuladas seis etapas, conforme consta na Figura 5.8. Assim, primeiramente, foi feita a coleta dos dados seguindo a Figura 5.2. Em seguida, realizou-se uma Análise de Componentes Principais (ACP) e foi feita a adaptação do gráfico de controle T^2 de Hotelling para Processos de Múltiplos Fluxos (MSPs). Posteriormente, analisou-se o desempenho do modelo de gráfico de controle multivariado com a AARL e, por fim, utilizou-se o modelo proposto.

Figura 5.8 – Etapas para o uso do gráfico de controle de T^2 -Hotelling.

Fonte: figura elaborada pela autora.

5.2.1 Fase I: Amostra Mestre

Na Fase I, primeiramente, a ACP foi realizada com amostra mestre de 40 elementos, considerando 8 itens dos 5 fluxos ou 5 posições de fabricação no cilindro, portanto identificou-se que cada fluxo foi composto de 8 unidades, sendo eles subdivididos. Foi analisada a característica de qualidade peso em quilogramas. Nota-se na Tabela 5.7 que as médias das posições nas extremidades — posição 1 (P1) e posição 5 (P5) — são menores em relação às posições centrais — posição 2 (P2), posição 3 (P3) e posição 4 (P4). Observa-se, também, que o desvio padrão segue o mesmo padrão das médias.

Tabela 5.7 – Estatística descritiva e matriz de correlação da amostra mestre.

	Média	Desv.Padrão	P1	P2	P3	P4	P5
P1	542,62	35,84864	1,000000	0,635709	-0,625309	0,887908	0,882711
P2	556,12	24,37468	0,635709	1,000000	-0,648012	0,859625	0,659989
P3	554,75	15,21043	-0,625309	-0,648012	1,000000	-0,677248	-0,420118
P4	557,00	17,11307	0,887908	0,859625	-0,677248	1,000000	0,801669
P5	519,25	21,76334	0,882711	0,659989	-0,420118	0,801669	1,000000

Fonte: tabela elaborada pela autora.

Então, as equações a seguir foram estabelecidas para demonstrar a estrutura de correlação linear entre as posições no cilindro. Iniciou-se ajustando da média menor para a maior, assumindo que $X_i \sim N(\mu_i, \sigma_i)$, onde X_i é uma v.a. oriunda da posição i com média μ_i e desvio-padrão σ_i , como segue:

$$\begin{aligned}
\text{Posição 5: } X_{5i} &= \mu_5 + \sigma_5 \\
\text{Posição 1: } X_{1i} &= \mu_5 + (\mu_1 - \mu_5) + \sigma_1 \\
\text{Posição 3: } X_{3i} &= \mu_5 + (\mu_3 - \mu_1) + \sigma_3 \\
\text{Posição 2: } X_{2i} &= \mu_5 + (\mu_2 - \mu_3) + \sigma_2 \\
\text{Posição 4: } X_{4i} &= \mu_5 + (\mu_4 - \mu_2) + \sigma_4
\end{aligned} \tag{5.5}$$

As equações acima mostram a correlação empírica entre as medidas de cada posição. Por exemplo, as medidas da posição 1 é média da posição 5, mais a diferença entre as médias das posições 1 e 5, com desvio-padrão σ_1 . Se não há diferença entre as médias da posição 1 e 5, e supondo que $\sigma_5 = \sigma_1$, então $X_{1i} = \mu_5 + \sigma_5$, ou seja, não há diferença entre os fluxos.

Verificou-se que 89,97% da variância total poderia ser explicada por duas Componentes Principais (CPs), soma formada por 77,26% da primeira componente — com autovalor maior que 1 — e por 12,71% da segunda componente — com autovalor menor que 1. Isso pode ser observado na Tabela 5.8. Sendo assim, essas 2 componentes principais resumem a variância amostral total, ou seja, pode trabalhar com um único que gráfico de controle para todas as posições no cilindro ou variáveis do ACP, já que a primeira componente (CP1) representa todas as posições do cilindro.

Tabela 5.8 – Variância explicada da amostra mestre.

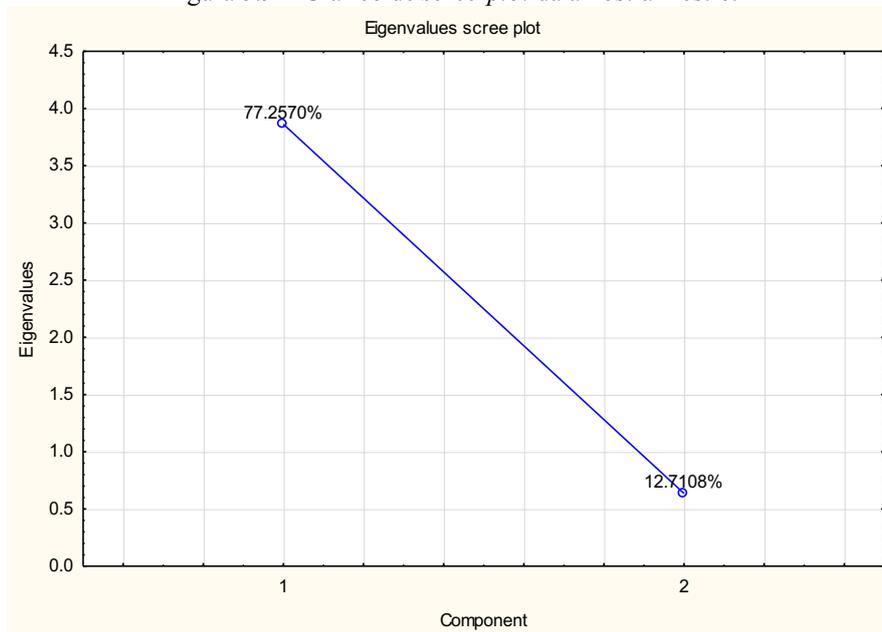
	Autovalores	Variância Total %	Autovalores Cumulativos	Cumulativo %
1	3,86	77,26	3,86	77,26
2	0,64	12,71	4,50	89,97

*Assumiu-se duas CPs para obter uma explicação da variância superior a 80%
 Fonte: tabela elaborada pela autora.

Por meio das Tabelas 5.8 e 5.9, junto à Figura 5.9, foi possível analisar as componentes principais. Conclui-se que a primeira componente (CP1) está explicada por todas variáveis, ou seja, pelas cinco posições dos produtos na esteira, com destaque para primeira e

para a quarta. Já a segunda componente (CP2) está explicada pelas variáveis das posições 3 e 5 dos produtos na esteira.

Figura 5.9 – Gráfico de *scree-plot* da amostra mestre.



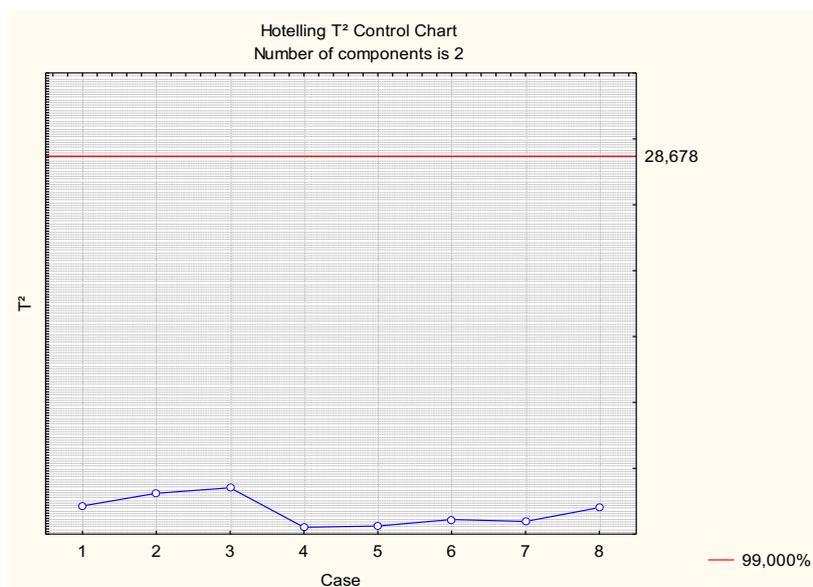
Fonte: figura elaborada pela autora.

Tabela 5.9 – Componentes Principais (CPs) e suas cargas fatoriais em relação às posições no cilindro da amostra mestre.

Variáveis	CP1	CP2
Posição 1	0,924511	0,223179
Posição 2	0,866784	-0,188643
Posição 3	-0,753779	0,591674
Posição 4	0,968185	0,016259
Posição 5	0,866747	0,446993

Fonte: figura elaborada pela autora.

Em relação ao gráfico T^2 de Hotelling representado na Figura 5.10, observa-se que o processo se encontra sob controle estatístico para o Processo de Múltiplos Fluxos (MSP), analisados sob a ótica de suas posições no cilindro e com Limite Superior de Controle (LSC) de 28,68.

Figura 5.10 – Gráfico de Controle de T^2 - Hotelling da amostra mestre.

Fonte: figura elaborada pela autora.

5.2.2 Fase II: Dados Expandidos

Na segunda fase, foi realizada, primeiramente, a ACP com a amostra mestre expandida utilizando a técnica *Bootstrap*. Ao todo foram 5000 elementos, sendo 1000 de cada uma das cinco posições, e estas são referentes ao cilindro de corte, que possui cinco facas em toda sua extensão e superfície. No presente estudo as cinco posições no cilindro ou fluxos representam as variáveis. Com as posições representando as variáveis, foi possível utilizar a técnica estatística T^2 de Hotelling nos gráficos de controle multivariado e, assim, construir um gráfico de controle único para analisar esse tipo de processo, o Processo de Múltiplos Fluxos (MSP).

Nota-se, por meio da Tabela 5.10, que o padrão se mantém como na Fase 1, com a amostra mestre. Observa-se, também, que as médias das posições das extremidades — posição 1 (P1) e posição 5 (P5) — são menores em relação às médias das posições centrais —

posição 2 (P2), posição 3 (P3) e posição 4 (P4) — e que o desvio padrão segue a mesma regra das médias.

Tabela 5.10 – Estatística Descritiva da Fase II

	N	Média	Mínimo	Máximo	Desv. Padrão
P1	1000	543,27	423,08	668,87	42,00
P2	1000	556,84	450,87	655,51	31,74
P3	1000	554,41	495,41	611,21	19,19
P4	1000	556,04	482,37	621,55	20,46
P5	1000	518,65	461,33	586,38	21,41

Fonte: tabela elaborada pela autora.

Já na Tabela 5.11, verificou-se que 80,62% da variância total poderia ser explicada por duas Componentes Principais (CPs), soma formada por 66,70% da primeira componente — com autovalor maior que 1 — e por 13,92% da segunda componente — com autovalor menor que 1.

Tabela 5.11 – Variância explicada da amostra expandida

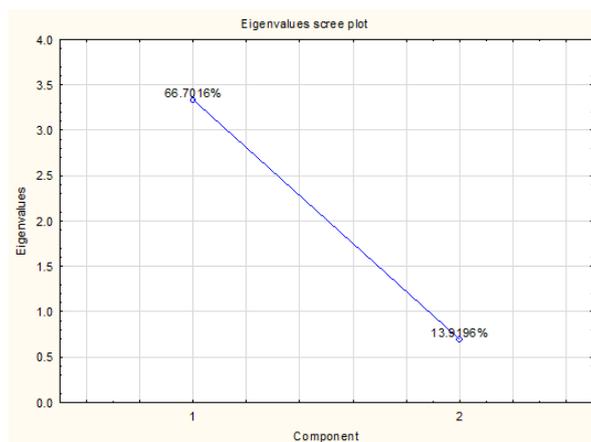
	Autovalores	Variância Total %	Autovalores Cumulativos	Cumulativo %
1	3,34	66,70	3,34	66,70
2	0,70	13,92	4,03	80,62*

*Assumiu-se duas CPs para obter uma explicação da variância superior a 80%

Fonte: tabela elaborada pela autora.

Analisando as Componentes Principais (CPs) na Figura 5.11 e na Tabela 5.12, mostra-se que a primeira componente (CP1) é explicada pelas 5 variáveis que representam as 5 posições dos produtos no cilindro, principalmente pela posição 5, 3 e 4, mas somente está correlacionada negativamente com a posição 5. Além disso, é possível observar que a segunda componente (CP2) é explicada pela variável referente à posição 1 dos produtos no cilindro. Em comparação com a amostra original, em termos de ACP, a análise da amostra expandida foi próxima a ela, considerando que as possíveis diferenças que foram causadas pelo erro aleatório de gerá-la.

Figura 5.11 – Gráfico de Scree-plot da amostra expandida.



Fonte: figura elaborada pela autora.

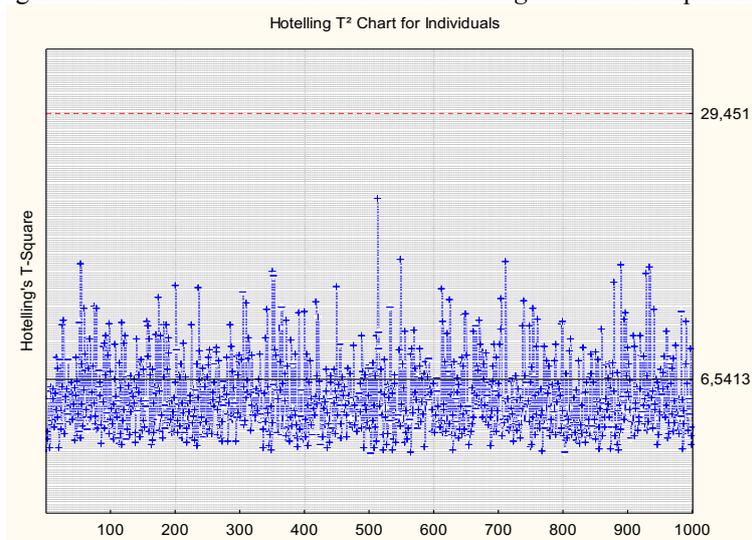
Tabela 5.12 – Componentes Principais e suas Cargas Fatoriais em relação as posições no cilindro.

Variáveis	CP1	CP2
Posição 1	0,633313	0,765644
Posição 2	0,769542	-0,276771
Posição 3	0,867717	-0,121151
Posição 4	0,832615	-0,110900
Posição 5	-0,946372	0,078662

Fonte: tabela elaborada pela autora.

5.2.3 Desempenho do Gráfico T^2 -Hotelling Proposto

Primeiramente, foi plotado o gráfico de controle multivariado de T^2 de Hotelling, utilizando os dados da amostra expandida, por meio do *software Statistica 13.0*, para monitorar o MSP, observando que o processo está sob controle estatístico. Isso é apresentado na Figura 5.12, bem como os valores dos Limites: LSC = 29,45 e LIC = 6,54.

Figura 5.12 – Gráfico de controle T^2 de Hotelling da amostra expandida.

Fonte: figura elaborada pela autora.

Posteriormente, foi realizado um estudo em que se monitorou se o processo estava sob controle ou não por meio da alteração dos números de desvios padrão (t), variando de 1 a 3. A Tabela 5.13 mostra cada posição no cilindro e a variação do t . Sendo assim, realizou-se um cálculo da média final, utilizado posteriormente para estimar a AARL, e nele utilizou-se o $\mu = \mu_0 + t\sigma$. Novamente, optou-se pelo gráfico de controle T^2 de Hotelling com os dados da amostra expandida para monitorar o processo com múltiplos fluxos. É possível observar, na Tabela 5.13, que as médias finais das posições do cilindro aumentam conforme aumenta o t .

Tabela 5.13 – Médias finais das posições do cilindro em relação aos t da amostra expandida

t	P1	P2	P3	P4	P5
0,00	542,62	556,12	554,75	557,00	519,25
0,15	548,00	559,78	557,03	559,57	522,51
0,25	551,58	562,21	558,55	561,28	524,69
0,50	560,54	568,31	562,36	565,56	530,13
0,75	569,51	574,40	566,16	569,83	535,57
1,00	578,47	580,49	569,96	574,11	541,01
1,25	587,43	586,59	573,76	578,39	546,45
1,50	596,39	592,68	577,57	582,67	551,90
1,75	605,36	598,78	581,37	586,95	557,34
2,00	614,32	604,87	585,17	591,23	562,78
2,25	623,28	610,96	588,97	595,50	568,22
2,50	632,24	617,06	592,78	599,78	573,66
2,75	641,20	623,15	596,58	604,06	579,10

3,00 650,17 629,24 600,38 608,34 584,54

Fonte: tabela elaborada pela autora.

Então, outra observação, referente à variação do t que representa o número de desvio padrão e a sua interferência no comportamento do processo, foi feita por meio da construção de outras tabelas (Tabelas 5.14 a 5.20). Elas foram formuladas utilizando o gráfico de controle T^2 de Hotelling para valores individuais, com Limites de Controle alterados pela variação do $\alpha_{Nominal}$. Também observou-se a variação do t a partir das médias finais encontradas na Tabela 5.13. Os Limites Superiores de Controle (LSCs) dos gráficos de controle T^2 de Hotelling analisados nas Tabelas 5.14 a 5.20 estão apresentados no Apêndice 2 em forma de tabela para consultas.

Na Tabela 5.14, pode-se notar que o gráfico de controle indica erroneamente que o processo está sob controle se $0 < t \leq 0,25$. Porém, quando $t \geq 1$, o gráfico indica que o processo não se encontra mais sob controle. Além disso, com a AARL=2,994, tem-se que o gráfico indica acertadamente que o processo está fora de controle em média a cada três amostras. Sendo assim, os outros t estudados nos gráficos de controle T^2 mostram que, quanto maior o t , maior é o poder de detecção do gráfico em indicar que o processo está fora de controle.

Tabela 5.14 - Comparação dos números de desvios-padrão, t , para análise do desempenho dos gráficos de controle T^2 em termos do AARL, com amostra expandida para $\alpha_{Nominal} = 0,135\%$.

t	Pontos fora	$1 - P(L\hat{C}L_x \leq \bar{x} \leq U\hat{C}L)$	AARL
0	0	0	∞
0,15	0	0	∞
0,25	1	0,001	1000
0,5	6	0,006	166,667
0,75	77	0,077	12,987
1	334	0,334	2,994
1,25	704	0,704	1,420
1,5	945	0,945	1,058
1,75	994	0,994	1,006
2	1000	1	1

2,25	1000	1	1
2,5	1000	1	1
2,75	1000	1	1
3	1000	1	1

Fonte: tabela elaborada pela autora.

Já no gráfico de controle T^2 com $\alpha_{Nominal} = 0,27\%$, ilustrado na Tabela 5.15, os valores se alteram. Entretanto, praticamente o mesmo padrão de comportamento do gráfico de controle anterior é mantido (Tabela 5.14). Ainda assim, é possível observar que os valores da AARL possuem maior capacidade de detecção de alarmes falsos e o gráfico de controle T^2 com $\alpha_{Nominal} = 0,27\%$ é mais apto que o anterior, de $\alpha_{Nominal} = 0,135\%$, a indicar que o processo está fora de controle.

Tabela 5.15 - Comparação dos números de desvios-padrão, t , para análise do desempenho dos gráficos de controle T^2 em termos do AARL, com amostra expandida para $\alpha_{Nominal} = 0,27\%$.

t	Pontos fora	$1 - P(L\bar{C}L_{\bar{X}} \leq \bar{X} \leq U\bar{C}L)$	AARL
0	0	0	∞
0,15	0	0	∞
0,25	1	0,001	1000
0,5	18	0,018	55,556
0,75	125	0,125	8,000
1	416	0,416	2,404
1,25	774	0,774	1,292
1,5	968	0,968	1,033
1,75	999	0,999	1,001
2	1000	1	1
2,25	1000	1	1
2,5	1000	1	1
2,75	1000	1	1
3	1000	1	1

Fonte: tabela elaborada pela autora.

Para o gráfico de controle T^2 com $\alpha_{Nominal} = 2,00\%$, como mostra a Tabela 5.16, os valores se alteram, indicando realmente um processo fora de controle. Além disso, os

valores da AARL apresentam maior poder de detecção de alarmes falsos e o gráfico é mais apto que os anteriores a indicar que o processo está fora de controle.

Tabela 5.16 - Comparação dos números de desvios-padrão, t , para análise do desempenho dos gráficos de controle T^2 em termos do AARL, com amostra expandida para $\alpha_{Nominal} = 2,00\%$.

t	Pontos fora	$1 - P(L\hat{C}L_{\bar{x}} \leq \bar{x} \leq U\hat{C}L)$	AARL
0	2	0,002	500
0,15	6	0,006	166,667
0,25	16	0,016	62,500
0,5	88	0,088	11,364
0,75	332	0,332	3,012
1	698	0,698	1,433
1,25	939	0,939	1,065
1,5	994	0,994	1,006
1,75	1000	1	1
2	1000	1	1
2,25	1000	1	1
2,5	1000	1	1
2,75	1000	1	1
3	1000	1	1

Fonte: tabela elaborada pela autora.

A Tabela 5.17 mostra o gráfico de controle T^2 com $\alpha_{Nominal} = 2,25\%$. É possível notar que os valores se alteram, mas o padrão de comportamento do gráfico de controle anterior (Tabela 5.16) é praticamente mantido. No entanto, é possível observar que, para o gráfico de controle T^2 com $\alpha_{Nominal} = 2,25\%$, há maior capacidade de detecção de alarmes falsos. Ademais, vê-se que o gráfico indica que o processo está fora de controle de forma mais adequada, em comparação com o anterior (Tabela 5.16, $\alpha_{Nominal} = 2,00\%$) e, portanto, apresenta um melhor desempenho.

Tabela 5.17 - Comparação dos números de desvios-padrão, t , para análise do desempenho dos gráficos de controle T^2 em termos do AARL, com amostra expandida para $\alpha_{Nominal} = 2,25\%$.

t	Pontos fora	$1 - P(L\hat{C}L_{\bar{x}} \leq \bar{x} \leq U\hat{C}L)$	AARL
0	3	0,003	333,333
0,15	7	0,007	142,857
0,25	18	0,018	55,556
0,5	104	0,104	9,615

0,75	357	0,357	2,801
1	717	0,717	1,395
1,25	946	0,946	1,057
1,5	994	0,994	1,006
1,75	1000	1	1
2	1000	1	1
2,25	1000	1	1
2,5	1000	1	1
2,75	1000	1	1
3	1000	1	1

Fonte: tabela elaborada pela autora.

Para o gráfico de controle T^2 com $\alpha_{Nominal} = 2,5\%$, como é ilustrado na Tabela 5.18, os valores se alteram. No entanto, vemos novamente que ele apresenta praticamente o mesmo padrão de comportamento do gráfico de controle anterior. No entanto, no gráfico de controle T^2 com $\alpha_{Nominal} = 2,5\%$, há maior poder de detecção de alarmes falsos. Além disso, ele se mostra mais eficiente em indicar que o processo está fora de controle do que o gráfico de controle T^2 anterior.

Tabela 5.18 - Comparação dos números de desvios-padrão, t , para análise do desempenho dos gráficos de controle T^2 em termos do AARL, com amostra expandida para $\alpha_{Nominal} = 2,50\%$.

t	Pontos fora	$1 - P(L\hat{C}L_x \leq \bar{X} \leq U\hat{C}L)$	AARL
0	3	0,003	333,333
0,15	7	0,007	142,857
0,25	22	0,022	45,455
0,5	110	0,110	9,091
0,75	377	0,377	2,653
1	734	0,734	1,362
1,25	948	0,948	1,055
1,5	995	0,995	1,005
1,75	1000	1	1
2	1000	1	1
2,25	1000	1	1
2,5	1000	1	1
2,75	1000	1	1
3	1000	1	1

Fonte: tabela elaborada pela autora.

De modo a encontrar um gráfico de controle T^2 com melhor desempenho, observou-se, ainda, mais duas variações de $\alpha_{Nominal}$, sendo elas: (i) $\alpha_{Nominal} = 8,00\%$, na Tabela 5.19; e (ii) $\alpha_{Nominal} = 9,00\%$, na Tabela 5.20. A Tabela 5.19 demonstra que o gráfico de controle T^2 com $\alpha_{Nominal} = 8,00\%$ tem maior poder de detecção de alarmes falsos e indica com mais qualidade quando o processo está fora de controle, isso em comparação com os gráficos de controle T^2 anteriores.

Tabela 5.19 - Comparação dos números de desvios-padrão, t , para análise do desempenho dos gráficos de controle T^2 em termos do AARL, com amostra expandida para $\alpha_{Nominal} = 8,00\%$.

t	Pontos fora	$1 - P(\widehat{LCL}_{\bar{x}} \leq \bar{x} \leq \widehat{UCL})$	AARL
0	26	0,026	38,462
0,15	44	0,044	22,727
0,25	76	0,076	13,158
0,5	254	0,254	3,937
0,75	595	0,595	1,681
1	871	0,871	1,148
1,25	986	0,986	1,014
1,5	1000	1	1
1,75	1000	1	1
2	1000	1	1
2,25	1000	1	1
2,5	1000	1	1
2,75	1000	1	1
3	1000	1	1

Fonte: tabela elaborada pela autora.

No gráfico de controle T^2 com $\alpha_{Nominal} = 9,00\%$, como exige a Tabela 5.20, os valores também se alteram, mas, novamente, é possível notar que o padrão de comportamento do gráfico de controle anterior praticamente se mantém. Contudo, o gráfico de controle T^2 com $\alpha_{Nominal} = 9,00$ tem maior poder de detecção de alarmes falsos e é mais habilitado para indicar que o processo está fora de controle do que os gráficos de controle T^2 analisados anteriormente.

Tabela 5.20 - Comparação dos números de desvios-padrão, t , para análise do desempenho dos gráficos de controle T^2 em termos do AARL, com amostra expandida para $\alpha_{Nominal} = 9,00\%$.

t	Pontos fora	$1 - P(L\bar{C}L_{\bar{x}} \leq \bar{X} \leq U\bar{C}L)$	AARL
0	33	0,033	30,303
0,15	50	0,050	20,000
0,25	82	0,082	12,195
0,5	275	0,275	3,636
0,75	620	0,620	1,613
1	885	0,885	1,130
1,25	988	0,988	1,012
1,5	1000	1	1
1,75	1000	1	1
2	1000	1	1
2,25	1000	1	1
2,5	1000	1	1
2,75	1000	1	1
3	1000	1	1

Fonte: tabela elaborada pela autora.

Para a análise do desempenho do gráfico usando a AARL, $t = [0; 0,15; 0,25; 0,5; 0,75; 1,0; 1,25; 1,5; 1,75; 2,0; 2,25; 2,5; 2,75; 3,0]$ com a variação do $\alpha_{Nominal} = p$, foi alterado para $p = [0,00135; 0,0027; 0,02; 0,0225; 0,025; 0,08; 0,09]$, como mostrado nas tabelas anteriores. No sentido de visualizar com mais qualidade, comparativamente, o resultado é mostrado na Tabela 5.21. Vê-se que o melhor desempenho da AARL está relacionado ao $\alpha_{Nominal} = 9\%$ ou $p = 0,0900$.

Tabela 5.21 - Valores de AARL para diferentes valores de t e p ($\alpha_{Nominal}$).

p	t													
	0	0,15	0,25	0,5	0,75	1	1,25	1,5	1,75	2	2,25	2,5	2,75	3
0,00135	∞	∞	1000	166,667	12,987	2,994	1,420	1,058	1,006	1	1	1	1	1
0,00270	∞	∞	1000	55,556	8,000	2,404	1,292	1,033	1,001	1	1	1	1	1
0,02000	500,000	166,667	62,500	11,364	3,012	1,433	1,065	1,006	1	1	1	1	1	1
0,02250	333,333	142,857	55,556	9,615	2,801	1,395	1,057	1,006	1	1	1	1	1	1
0,02500	333,333	142,857	45,455	9,091	2,653	1,362	1,055	1,005	1	1	1	1	1	1
0,08000	38,462	22,727	13,158	3,937	1,681	1,148	1,014	1	1	1	1	1	1	1
0,09000	30,303	20,000	12,195	3,636	1,613	1,130	1,012	1	1	1	1	1	1	1

Fonte: tabela elaborada pela autora.

6 ANÁLISE COMPARATIVA DAS PROPOSTAS DE GRÁFICOS PARA UM PROCESSO DE MÚLTIPLOS FLUXOS (MSP)

Primeiramente, o gráfico de distribuições de probabilidade de mistura finita foi comparado com o parâmetro K do gráfico de controle tradicional de Shewhart adicionalmente, visando demonstrar suas diferenças em termos de desempenho. Em seguida, o gráfico de distribuições de probabilidade de mistura finita foi comparado com o gráfico de controle T^2 de Hotelling, com intuito de demonstrar os desempenhos, as vantagens e as desvantagens operacionais dos métodos propostos. Essa sessão está dividida em dois subitens e também tem o objetivo de responder a questão de pesquisa levanta no presente estudo.

6.1 ANÁLISE COMPARATIVA DO GRÁFICO DE CONTROLE DE MISTURAS COM O PARÂMETRO K DO GRÁFICO DE CONTROLE TRADICIONAL

Um Processo de Múltiplos Fluxos (MSP) é tradicionalmente ilustrado por um gráfico de controle de grupo (MONTGOMERY, 2013). A abordagem mostrada em Montgomery (2013) supõe que cada fluxo tem o mesmo σ com distribuição de probabilidade normal. Essa perspectiva traz como proposição que cada fluxo tenha o mesmo tamanho, os mesmos gráficos de média máxima e mínima após o cálculo de \bar{X} e o mesmo desvio padrão ou amplitude. No entanto, esse procedimento parte do pressuposto que a distribuição de probabilidade é normal e que cada fluxo tem mesma média e mesmo desvio padrão — o que, no nosso caso, não é verdade. A análise de desempenho foi realizada usando um gráfico para cada fluxo — assim, usamos quatro gráficos, o que é muito caro — e usando a abordagem com mistura. Portanto, nesta seção, fez-se o desempenho da análise comparativa do gráfico de controle de mistura com a abordagem tradicional de múltiplos fluxos para situações genéricas.

A distribuição de mistura proposta neste trabalho converge para o gráfico de controle tradicional de Shewhart para \bar{X} , para o caso do μ desconhecido e para quando o parâmetro $\delta_1 = \delta_2 = \dots = \delta_l = 0$. Sendo que δ representa diferença das médias entre os fluxos, proveniente da ANOVA. Desta forma, a probabilidade de uma estadia de amostragem entre os limites de controle é:

$$P(\widehat{LCL}_{\bar{X}} \leq \bar{X} \leq \widehat{UCL}) = \pi_1 \left[\Phi\left(\frac{w}{\sqrt{m}} + K - t\sqrt{n}\right) - \Phi\left(\frac{w}{\sqrt{m}} - K - t\sqrt{n}\right) \right] + \pi_2 \left[\Phi\left(\frac{w}{\sqrt{m}} + K - t\sqrt{n}\right) - \Phi\left(\frac{w}{\sqrt{m}} - K - t\sqrt{n}\right) \right] + \dots + \pi_l \left[\Phi\left(\frac{w}{\sqrt{m}} + K - t\sqrt{n}\right) - \Phi\left(\frac{w}{\sqrt{m}} - K - t\sqrt{n}\right) \right] \quad (6.1)$$

Na equação acima, l representa o número de fluxos. Em sequência, se os fluxos também tiverem o mesmo μ e tiverem seus σ com distribuição de probabilidade normal, $P(\widehat{LCL}_{\bar{X}} \leq \bar{X} \leq \widehat{UCL})$. Assim, foi proposto um modelo que pode ser usado em situações genéricas tanto para se obter o número da Função de Distribuição de Probabilidade (FDP) de diferentes fluxos, quanto no caso dos fluxos se reduzirem em um, após a ação de melhoria.

Observe que o termo usado no modelo $K\sqrt{1 + \frac{1}{l}\sum_l \delta_l^2}$, expanda K na magnitude de $\sqrt{\frac{1}{l}\sum_l \delta_l^2}$; outro lado, K é expandido por δ .

A mesma análise foi feita com o objetivo de entender os efeitos do t que representa o número de desvio padrão na média, no desempenho do gráfico de controle de mistura. Também fez-se o uso do modelo de ANOVA (Análise de Variância) para construir o modelo de controle para a mistura de distribuições. Na Tabela 6.1, é mostrada a AARL para conjuntos de t considerando que $X_i \sim N(\mu_i, \sigma_i)$. O caso especial é para $\delta_i = 0$, que resulta no tradicional gráfico de controle \bar{X} . É simples entender a AARL, mostrado na Tabela 6.1, para a seguinte restrição:

$$\sum_i^l \delta_i = 0. \quad (6.2)$$

Observe na Tabela 6.1, onde $K = 3,0$ usado tradicionalmente no gráfico de controle de Shewhart e $K = 3,585$, que o fluxo único tem, respectivamente, a AARL = 319,7 e 2496,8. Quando há mais de um fluxo, valores menores de AARL são encontrados na situação de controle — na qual $t = 0$. Para isso, no gráfico de controle de mistura, foi proposto um $K > 3,00$ para uma AARL maior que 370,4. A Figura 6.1 mostra a curva das AARLs em gráficos de controle de mistura para diferentes valores de t , isto é, quando o procedimento se encontra fora do estado de controle. Pode-se observar que a curva do gráfico de controle de mistura não representa um gráfico de controle similar tradicional. A vantagem é que o gráfico de mistura apresenta melhor desempenho no estado sob controle do que com o parâmetro K do gráfico de Shewhart \bar{X} tradicional que trabalha apenas com fluxo único. Nota-se que, para $t > 1,0$, o uso do gráfico de controle de mistura possui maior capacidade para detectar uma causa especial em relação ao parâmetro K do gráfico de Shewhart \bar{X} tradicional.

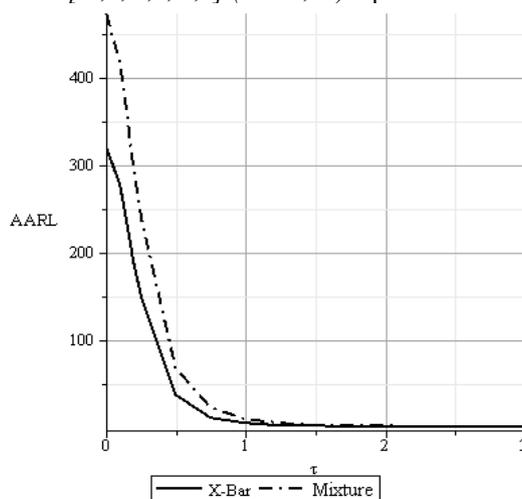
Tabela 6.1 - Desempenho do AARL para diferentes conjuntos de t para $K = 3,00$ e para $K = 3,59$.

t				AARL	
1	2	3	4	K=3,0	K=3,59
0	0	0	0	319,7	2496,8
-0,5	0,5	0	0	87,1	473,2
0,5	-0,5	0	0	87,1	473,2
0	0	-0,5	0,5	87,1	473,2
0	0	0,5	-0,5	87,1	473,2
-1	0,5	0,5	0	33,7	155,6
1	-0,5	-0,5	0	33,8	155,7
0	-1	0,5	0,5	33,7	155,6
0	1	-0,5	-0,5	33,8	155,7
0,5	0	-1	0,5	33,7	155,6
-0,5	0	1	-0,5	33,8	155,7
-0,25	0,25	0	0	198,5	1314
1	-0,5	-0,25	-0,25	33,3	146,2
1	-0,5	-0,5	0	33,7	155,6
0	-1,25	1,25	0	16,9	80,4

Fonte: tabela elaborada pela autora.

Logo, pode-se concluir que é possível que o gráfico de controle de mistura proposto tenha o mesmo poder de detecção de uma causa especial do que o gráfico de \bar{X} para cada fluxo (desvio na média geral). O método proposto, usando a distribuição de probabilidade de mistura finita, propôs o uso de um gráfico único para monitorar múltiplos fluxos. O modelo é simplesmente a extensão do gráfico de controle tradicional de Shewhart para fluxos exclusivos.

Figura 6.1 - Análise comparativa entre gráfico de controle de mistura para ajuste específico $[-0,5, 0,5, 0,0]$ ($K = 3,59$) e parâmetro K tradicional de Shewhart \bar{X} .



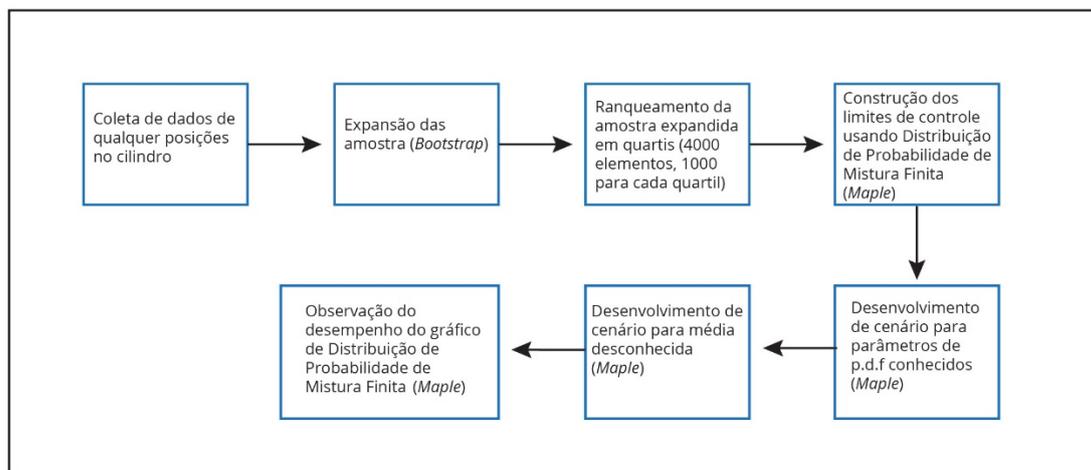
Fonte: figura elaborada pela autora.

6.2 ANÁLISE COMPARATIVA DO GRÁFICO DE CONTROLE DE MISTURAS COM O GRÁFICO DE CONTROLE T^2 DE HOTELLING

Em termos operacionais, os gráficos de controle propostos — o gráfico de controle de distribuições de probabilidade de mistura finita e o gráfico de controle T^2 de Hotelling — foram ilustrados nos fluxogramas representados respectivamente pela Figura 6.2 e pela Figura 6.3, conforme seus processos de construção. Pode-se observar maior facilidade operacional no gráfico de controle T^2 de Hotelling, construído mais facilmente pelo *software Statistica 13.0*, em termos de simplicidade matemática e automação, dado que não há necessidade de programar ou utilizar linguagens de programação. Também é possível notar, na construção do

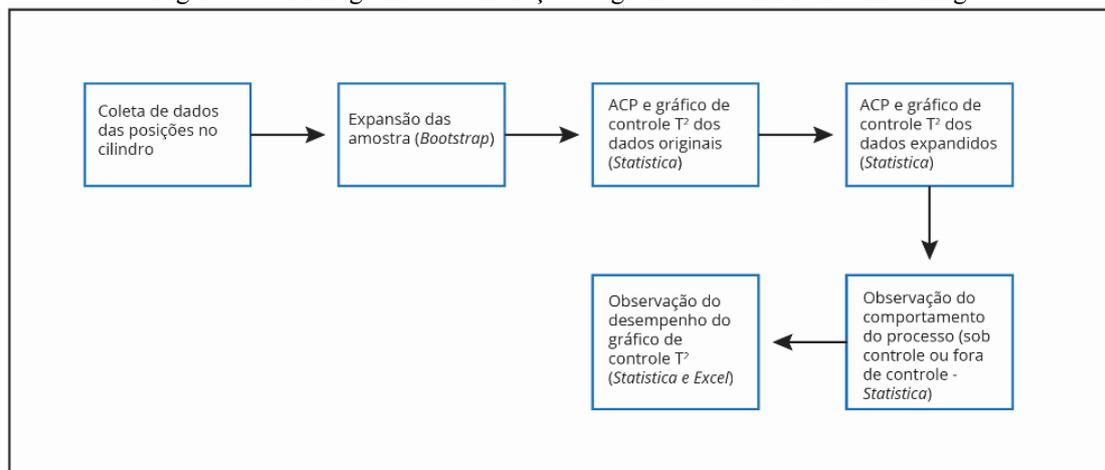
gráfico de controle T^2 de Hotelling, que há maior foco e atenção nas posições do cilindro, no âmbito do Processo de Múltiplos Fluxos (MSP). Este gráfico considera as interdependências entre os fluxos ou posições no cilindro, diferentemente do que é feito com as posições dos quartis apresentadas no gráfico de controle de distribuições de probabilidade de mistura finita. A vantagem operacional apresentada nos dois modelos consiste na facilidade de monitoramento do MSP utilizando somente um gráfico de controle.

Figura 6.2 - Fluxograma da construção do gráfico de controle de distribuições de probabilidade de mistura finita.



Fonte: figura elaborada pela autora.

Figura 6.3 - Fluxograma da construção do gráfico de controle T^2 - Hotelling.



Fonte: figura elaborada pela autora.

Já em termos de desempenho dos gráficos de controle propostos, analisados por meio da AARL, o gráfico de controle de distribuições de probabilidade de mistura finita e o gráfico de controle T^2 de Hotelling foram comparados nas Tabela 6.2 e na Figura 6.4, utilizando os resultados apresentados no capítulo 4. O α real foi calculado e, assim, representado por α_R para $t = 0$. Além disso, o $\alpha_{Nominal}$ foi representado por p .

Tabela 6.2 - Desempenho da AARL para diferentes conjuntos de t e p , sendo $K = 3,59$ para gráfico de controle de mistura finita, e $p=[0,00135$ a $0,09]$ para gráfico de controle T^2 de Hotelling.

p/t	0	0,25	0,5	1,0	2,0
Mistura Finita					
	377,2 / $\alpha_R = 0,00265$	97,15	33,07	8,271	4,059
T^2-Hotelling					
0,00135	∞	1000	166,667	2,994	1
0,00270	∞	1000	55,556	2,404	1
0,02000	500,000 / $\alpha_R = 0,00200$	62,500	11,364	1,433	1
0,02250	333,333 / $\alpha_R = 0,00300$	55,556	9,615	1,395	1
0,02500	333,333 / $\alpha_R = 0,00300$	45,455	9,091	1,362	1
0,08000	38,462 / $\alpha_R = 0,02600$	13,158	3,937	1,148	1
0,09000	30,303 / $\alpha_R = 0,03300$	12,195	3,636	1,130	1

* p = valor nominal do erro tipo I dos gráficos T^2 -Hotelling

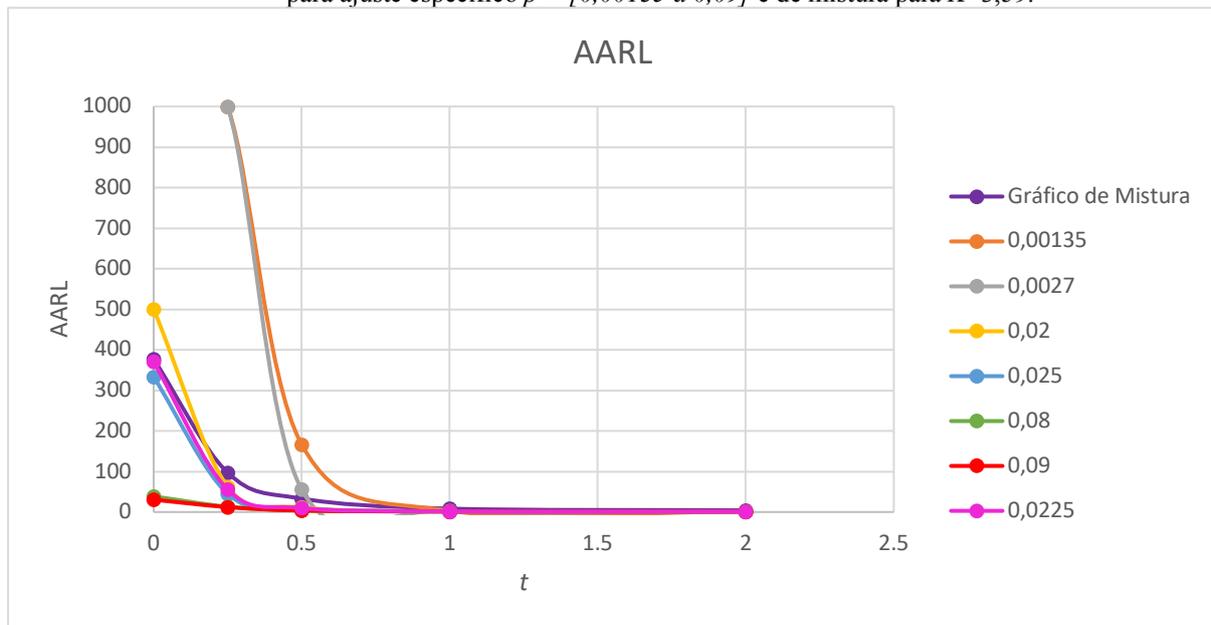
* t = número de desvios padrões de deslocamento da média nominal.

Fonte: tabela elaborada pela autora.

Na Tabela 6.2, os valores em negrito representam o valor teórico (0,00027) e usual, contendo três desvios padrão para, primeiramente, o gráfico de controle de distribuições de probabilidade de mistura finita e, posteriormente, para o gráfico de controle T^2 de Hotelling. Os outros valores dizem respeito aos gráficos de controle T^2 de Hotelling, com ajustes de $p = [0,00135$ a $0,09]$, sendo p o valor de α nominal. Observa-se que, a partir do $t = 0,25$, os valores da AARL são menores no gráfico de controle de T^2 -Hotelling destacado do que no gráfico de controle de mistura finita. Isto observando os valores de AARL, respectivamente, nos gráficos de controle de T^2 -Hotelling e de mistura finita: para $t = 0,25$, tem-se as AARLs de 55,556 e 97; para $t = 0,5$, as AARLs são de 9,615 e 33,07; para $t = 1,00$, as AARLs têm valores de 1,395 e 8,271; para $t = 2,00$, as AARLs são de 1,000 e 4,059. Sendo assim, é possível notar que o

gráfico de controle T^2 de Hotelling tem melhor poder de teste, ou seja, $(1-\beta)$, $AARL=1/(1-\beta)$ que o gráfico de distribuições de probabilidade de mistura finita, pois os valores das AARLs são menores para o gráfico de controle T^2 de Hotelling.

Figura 6.4 - Análise comparativa entre gráfico de controle T^2 de Hotelling para ajuste específico $p = [0,00135 a 0,09]$ e de mistura para $K=3,59$.



Fonte: tabela elaborada pela autora.

A Figura 6.4 indica, também, que o gráfico de controle T^2 de Hotelling tem um melhor poder de teste que o gráfico de distribuições de probabilidade de mistura finita, principalmente a partir do valor de $p = 0,02$. Em relação ao valor teórico que se destaca na Tabela 6.2, pode-se comprovar o melhor poder de teste do gráfico de controle T^2 de Hotelling quando comparado ao gráfico de mistura. Sendo assim, observa-se que o gráfico de controle T^2 de Hotelling proposto pode ter melhor poder para detectar uma causa especial do que o gráfico de mistura para cada fluxo (desvio na média geral). O método proposto usando o gráfico de controle T^2 de Hotelling propõe o uso de um gráfico único para monitorar múltiplos fluxos. O modelo é mais simples que o do gráfico de mistura finita e do que do gráfico de controle

tradicional de Shewhart para fluxos exclusivos, pois, com apenas um único gráfico de controle, já é possível monitorar o processo com um bom desempenho de detecção de causa especial.

6.3 ANÁLISE GERAL

O caso analisado mostrou duas aplicações: um gráfico de controle com distribuições de probabilidade de mistura finita e um gráfico de controle T^2 -Hotelling em um processo de uma empresa do setor de alimentos, cujo produto consiste em massa de pastel. Observou-se que o gráfico de mistura proposto é mais complexo, pois envolve mais parâmetros a serem considerados. No entanto, pode ajudar os profissionais a monitorar em um único gráfico o processo de múltiplos fluxos. Já o gráfico de controle T^2 -Hotelling é mais simples e também pode ser usado para monitorar o processo com um único gráfico quando comparado ao gráfico de controle de distribuições de probabilidade de mistura finita.

Um modelo matemático foi desenvolvido para o processo de múltiplos fluxos, abordando a distribuições de probabilidade de mistura finita, para média conhecida e desconhecida. Foram propostos também novos limites para esse gráfico de controle. Ele fornece uma nova abordagem para o monitoramento de múltiplos fluxos quando a produção possui vários fluxos e os dados pertencem às distribuições de probabilidade de mistura finita, com várias subpopulações referentes aos quartis que foram encontrados a partir do ranquiamento dos dados expandidos. Para análise do desempenho gráfico, utilizou-se o AARL.

Conforme o discutido por Vicentin *et al.* (2018), o uso especial para o proposto gráfico de controle de mistura ocorre quando os clientes recebem lotes de produtos continuamente de um fornecedor e precisam monitorar estatisticamente a qualidade do produto com múltiplos fluxos. Nessa situação, observamos que diferentes AARL são encontradas e ocorrem desvios na média, tanto positivos quanto negativos. A AARL menor ocorre para um

pequeno desvio negativo na média. Por outro lado, para valores maiores de t , ocorrem menores AARL para desvio positivo na média geral do processo. Este é um resultado especial para a função de densidade de mistura de probabilidade. Este resultado mostra que, na mesma situação, o gráfico de controle é mais sensível para desvios negativos ou positivos na média comparada ao parâmetro K do gráfico tradicional de Shewhart \bar{X} . A limitação da pesquisa é aplicável em um único caso e requer verificação em outros casos com problemas semelhantes.

Já o modelo matemático que foi desenvolvido para o processo de múltiplos fluxos, abordando ACP e T^2 -Hotelling é mais simples que o de mistura finita. Foram ainda analisadas as correlações das 5 posições dos produtos no cilindro (no final da linha de produção), dando destaque para as posições no cilindro como múltiplos fluxos para construir esse gráfico de controle. Ele fornece uma nova abordagem para o monitoramento de múltiplos fluxos quando a produção possui vários fluxos e os dados pertencem às diferentes posições de fabricação, representando várias subpopulações. Para análise do desempenho gráfico, utilizou-se o AARL para comparar com o gráfico de distribuições de probabilidade de mistura finita.

O uso especial para o proposto gráfico de controle de T^2 -Hotelling ocorre quando os clientes recebem lotes de produtos continuamente de um fornecedor e precisam monitorar estatisticamente a qualidade do produto com múltiplos fluxos. Nessa situação, observamos que diferentes AARL são encontradas. Por outro lado, para valores maiores de δ , ocorrem menores AARL para desvio positivo na média geral do processo. Este resultado mostra que, na mesma situação, o gráfico de controle T^2 -Hotelling é mais sensível comparado ao gráfico de mistura finita. Já a vantagem e semelhança entre os gráficos de controle de modelos de mistura finita e T^2 -Hotelling está na facilidade de utilizar somente um gráfico para o monitoramento de MSP.

O modelo proposto de controle de T^2 -Hotelling é teoricamente mais simples do que o gráfico de distribuições de probabilidade de mistura finita e possui melhor desempenho, sendo, então, mais fácil sua implantação na prática, mas pode encontrar erros similares do tipo

I e tipo II quando em relação aos outros designs de gráficos de controle. Sendo assim, mais estudos são necessários para comparar o desempenho desse método proposto com outros projetos de gráfico de controle para múltiplos fluxos.

7. CONCLUSÕES

Este trabalho contribui com o conhecimento científico da área, bem como beneficia os profissionais que trabalham com processos produtivos de múltiplos fluxos, uma vez que propõem uma solução alternativa aos gráficos de Shewhart, incluindo os gráficos de T^2 -Hotelling, entendendo esse último como uma extensão do gráfico de Shewhart. Considerando que o objetivo deste trabalho foi construir um único gráfico de controle utilizando a abordagem de distribuições de probabilidade de mistura finita para o monitoramento de MSP e compará-lo com o gráfico de controle de T^2 -Hotelling em termos de AARL, entende-se que esse objetivo foi alcançado, pois após a sua construção ainda pode se realizar uma comparação empírica do desempenho, em termos de AARL, do gráfico de controle de distribuições de probabilidade de mistura finita em relação ao gráfico de controle multivariados de T^2 -Hotelling para o monitoramento de processo de múltiplos fluxos, em que foi possível observar sua eficácia e bom desempenho em relação a sua variabilidade entre os fluxos. Visto que para Zwetsloot e Woodall (2019) o desempenho em relação a sua variabilidade (variâncias) entre os fluxos seria muito importante para o monitoramento adequado do processo, sem mascarar os efeitos de causas especiais e sem perder informações sobre o processo, porém ainda é pouco estudada. Já Epprech (2015), Ahangar e Chimka (2015) evidenciam que nos processos de múltiplos fluxos não foi encontrada nenhuma solução plenamente satisfatória. Portanto é necessário desenvolver métodos para tais aplicações específicas e análises comparativas mais detalhadas do desempenho de métodos existentes, como foi realizado nesse estudo.

Como o descrito na introdução, para nortear o objetivo proposto, foram formuladas as seguintes questões de pesquisa: Qual dessas duas abordagens, mistura finita e multivariada, têm melhor desempenho em termos do erro tipo I (gerar menor falso alarme)? Qual delas é

mais fácil de ser implementada? Que foram respondidas ao longo da tese e retomada nessa conclusão.

Como previsto, realizou-se uma análise comparativa entre os gráficos de mistura finita e multivariada, pois foi possível mensurar os erros tipo I e II e seus respectivos AARL, para diferentes combinações de parâmetros desses gráficos, tais como número de amostra e tamanho de amostra extraídas na execução da fase I. Observou-se que o gráfico de controle de T^2 -Hotelling obteve melhor desempenho, em termo dos erros tipo I e II, gerando menor falso alarme. Para o caso do erro tipo I gerou menor falso alarme e para o caso erro tipo II obteve melhor detecção do processo fora de controle. Como foi possível observar nas seções precedentes, a partir de $t \geq 0,25$ os valores de AARL foram menores do gráfico de controle de T^2 -Hotelling com relação ao gráfico de controle de mistura finita, ou seja, ele detecta mais rapidamente um processo fora de controle. Em outras palavras melhora-se o poder de teste (menos erro tipo II). Por exemplo, para $t=0,25$ a AARL de 55,556 para o de T^2 -Hotelling e AARL de 97,15 para mistura finita; $t=0,5$ a AARL de 9,615 para o de T^2 -Hotelling e AARL de 33,07; $t=1,00$ a AARL de 1,395 para o de T^2 -Hotelling e AARL de 8,271; $t=2,00$ a AARL de 1,000 para o de T^2 -Hotelling e AARL de 4,059.

Nota-se também que os dois gráficos são aplicáveis, sendo o gráfico de controle de T^2 -Hotelling aparentemente mais fácil de ser implantado por sua menor complexidade matemática. No entanto, o gráfico de controle de mistura finita é mais inovador e seu o procedimento de amostragem ser mais fácil, pois é realizada uma amostragem aleatória, não importando a posição dos dados e a determinação das subpopulações é feito depois que coletou os dados, na fase I. Pode ser implementado com apenas um gráfico de controle e ainda é possível observar as variâncias entre os fluxos para o monitoramento de MSP. Entretanto, essa pesquisa tem como limitação a aplicação em um único caso ilustrativo com dados reais. Assim,

para pesquisas futuras, pode-se verificar em outros casos com problemas semelhantes às propostas desses gráficos de controle que foram apresentadas e comparações com o modelo ANOVA de um fator.

REFERÊNCIAS

ABBAS, N.; RIAZ, M.; AHMAD, S.; ABID, M.; ZAMAN, B. On the Efficient Monitoring of Multivariate Processes with Unknown Parameters. *Mathematics*, v. 8, n. 5, p. 823-834, 2020.

ABBASI, S. A.; ADEGOKE, N. A. Auxiliary-information-based efficient variability control charts for Phase I of SPC. *Quality and Reliability Engineering International*, p.1-16, 2020.

ABDULAZIZ, A. G.; RIBAS, C. S.; WEHEBA, G. S. Application of Group Control Charts for Multiple Parts Manufacturing, *Journal of Management & Engineering Integration*, v. 12, n. 2, p. 41-48, 2019.

ABU-SHAWIESH, M. O. A. A control chart based on robust estimators for monitoring the process mean of a quality characteristic, *International Journal of Quality & Reliability Management*, v. 26, n. 5, p. 480-496, 2009.

ADAMS, C. P. Finite mixture models with one exclusion restriction, *Econometrics Journal*, v. 00, p. 1-16, 2016.

AHANGAR, N. E.; CHIMKA, J. R. Attribute Control Charts with Optimal Limits. *Quality and Reliability Engineering International*, v. 32, p. 1381–1391, 2016.

AHSAN, M.; MASHURI, M.; KUSWANTO, H.; PRASTYO, D. D. Intrusion Detection System using Multivariate Control Chart Hotelling's T² based on PCA. *International Journal on Advanced Science, Engineering and Information Technology*, v. 8, n. 5, p. 1905–1911, 2018.

ALBERS, W.; KALLENBERG, W. C. M. Empirical non-parametric control charts: estimation effects and corrections. *Journal of Applied Statistics*. v. 31, n. 3, p. 345-360, 2004.

ALEVIZAKOS, V.; KOUKOUVINOS, C. A progressive mean control chart for monitoring time between events. *Quality and Reliability Engineering International*, v. 36, n. 1, 2019.

ALMEIDA, F. A.; GOMES, G. F.; GAUDÊNCIO, J. H. D.; GOMES, J. H. F.; PAIVA, A. P. A new multivariate approach based on weighted factor scores and confidence ellipses to precision evaluation of textured fiber bobbins measurement system. *Precision Engineering*, v. 60, p. 520–534, 2019.

ALT, F. B. *Multivariate quality control*. In: JOHNSON, N. L.; KOTZ, S. *Encyclopedia of Statistical Science*. Nova York: Wiley, 1985.

AMIN, R. W.; WOLFF, H., BESENFELDER, W.; BAXLEY, R. EWMA Control Charts for the Smallest and Largest Observations. *Journal of Quality Technology*, v. 31, n. 2, p. 189–206, 1999-2018.

ANTONY, J.; PREECE, D. *Understanding, Managing and Implementing Quality*. Londres: Routledge, 2002.

APSEMIDIS, A.; PSARAKIS, S.; MOGUERZA, J. M. A Review of Machine Learning Kernel Methods in Statistical Process Monitoring. *Computers & Industrial Engineering*, v.142, n.106376, 2020.

ASLAM, M.; WU, C. W.; AZAM, M.; JUN, C. H. Mixed acceptance sampling for product inspection using process capability index, *Quality Engineering*, v. 26, n. 4, p. 450-459, 2014.

AUTOMOTIVE INDUSTRY ACTION GROUP (AIAG). *Fundamental Statistical Process Control: Reference Manual*, Automotive Industry Action Group. Southfield, Michigan, 1991.

BERK, J.; BERK, S. *Administração da qualidade total: o aperfeiçoamento contínuo: teoria e prática*. São Paulo: IBRASA, 1997.

BAKER, R. C.; BROBST, R. W. Conditional double sampling, *Journal of Quality Technology*, v. 10, n. 4, p. 150-154, 1996.

BENJAMINI, Y.; YOSEF H. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, v. 57, n. 1, p. 289–300, 1995.

BERSIMIS, S.; PSARAKIS, S.; PANARETOS, J. Control Charts : An Overview. *Quality and Reliability Engineering International*, v. 23, p. 517–543, 2007.

BIANCHINI, D. F.; BISOGNIN, C.; SOARES, D. S. Uma proposta didática para o ensino de estatística: o uso do excel para representação gráfica. *RENOTE: revista novas tecnologias na educação, Porto Alegre*, v. 13, n. 2, 2015. Disponível em: <https://seer.ufrgs.br/renote/article/view/614310>. Acesso em: 07 de janeiro de 2019.

BOONE, J. M.; CHAKRABORTI, S. Two simple shewhart-type multivariate nonparametric control charts. *Applied Stochastic Models in Business and Industry*, v. 28, n. 2, p.130–40, 2012.

BOTHE, D. R. *Process Capability Indices for Multiple Stream Processes*. In *Encyclopedia of statistics in quality and reliability*. New York: Wiley, 2008.

BOYD, D. F. Applying the group chart for XN and R. *Industrial Quality Control*, v. 7, p. 22–25, 1950.

BROWN, A. R.; SCHAFFER, J. R. A nonparametric CUSUM control chart for multiple stream

processes based on a modified extended median test. *Communications in Statistics - Theory and Methods*, p. 1–14, 2020.

BRYMAN, A. *Research Methods and Organization Studies*. London: Routledge, 1989.

CASTAGLIOLA, P.; CELANO, G.; FICHERA, S. Comparison of the \bar{X} -Chart and the t Chart when the Parameters are Estimated. *Quality Technology & Quantitative Management*, v. 10, p.1-16, 2013.

CASTAGLIOLA, P.; MARAVELAKIS, P.; PSARAKIS, S.; VÄNNMAN, K. Monitoring Capability Indices using Run Rules. *Journal of Quality in Maintenance Engineering*, v. 15, n. 4, p. 358-370, 2009.

CASTILLO, E.; MONTGOMERY, D. C. Short-Run Statistical Process Control: Q-Chart Enhancements and Alternatives, *Quality and Reliability Engineering International*, v. 10, p. 87-97, 1994.

CHAKRABORTI, S.; GRAHAM, M. A. Nonparametric statistical process control. Hoboken, NJ: Wiley, 2019.

CHAKRABORTI, S.; GRAHAM, M. A. Nonparametric (distribution-free) control charts: An updated overview and some results. *Quality Engineering*, p. 1–22, 2019.

CHAKRABORTI, S.; QIU, P.; MUKHERJEE, A. Introduction to the special issue: Nonparametric statistical process control charts. *Quality and Reliability Engineering International*, v. 3, n. 1, p.1–2, 2015.

CHAKRABORTI S.; SPARKS R. S. *Statistical Process Monitoring and the Issue of Assumptions in Practice: Normality and Independence*. In: Koutras M., Triantafyllou I. *Distribution-Free Methods for Statistical Process Monitoring and Control*. Springer, Cham., 2020, p. 137-155.

CHAUVEAU, D.; HOANG V. T. Nonparametric mixture models with conditionally independent multivariate component densities, *Computational Statistics*, v. 103, p. 1-6, 2016.

CHEN, N.; ZI, X.; ZOU, C. A distribution-free multivariate control chart. *Technometrics*, n. 58, n. 4, p. 448–59, 2016.

CHEN, Y. K.; HSIEH, K. L. Hotelling's T^2 charts with variable sample size and control limit. *European Journal of Operation Research*. v. 182, p. 1251-1262, 2007.

COELHO, M. L. I.; GRAHAM, M. A.; CHAKRABORTI, S. Nonparametric signed-rank control charts with variable sampling intervals. *Quality and Reliability Engineering International*, v. 33, n. 8, p. 2181–92, 2017.

COSTA, A. F. B. Joint Xbar and R Charts with Variable Parameters. *IIE Transactions*, v. 30, p. 505-514, 1998.

COSTA, A. F. B. \bar{X} Chart with Variable Sample Size and Sampling Intervals, *Journal of Quality Technology*, v. 29, n.2, p. 197-204, 1997.

CRESWELL, J. W.; CRESWELL, D. J. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. 5 ed. London: Sage Publications, 2017.

CRESWELL, J. W.; CLARK, V. L. P. *Designing and Conducting Mixed Methods Research*. 2 ed. Londres: Sage, 2006.

DASTOORIAN, R.; WELLS, L. Gauge capability studies for high-density data: SPC Phase 0. *Procedia Manufacturing*, 48, 105–113, 2020.

de Almeida, F. A.; Santos, A. C. O.; de Paiva, A. P.; Gomes, G. F.; Gomes J. H. F. Multivariate Taguchi loss function optimization based on principal components analysis and normal boundary intersection. *Engineering with Computers*, (2020).

DIXON, P. Assessing the evidence for response time mixture distributions. *Behavior Research Methods (Online)*, v. 44, n. 3, p. 706-724, 2012.

DOĞU, E. Identifying the time of a step change with multivariate single control charts. *Journal of Statistical Computation and Simulation*, v. 85, n. 8, p. 1529-1543, 2015.

DOĞU, E.; KIM, M. J. Self-starting single control charts for multivariate processes: a comparison of methods. *Production*, v. 30, 2020.

DUNN, O. J. Estimation of the means of dependent variables. *The Annals of Mathematical Statistics*, v. 29, p. 1095–1111, 1958.

ELG, M.; OLSSON, J.; DAHLGAARD, J. J. Implementing statistical process control: an organizational perspective, *International Journal of Quality & Reliability Management*, v. 25, n. 6, p. 545-560, 2008.

EPPRECHT, E. K. Statistical Control of Multiple-Stream Processes: A Literature Review, *Frontiers in Quality Statistical Control*, v. 11, p. 49-64, 2015.

EPPRECHT, E. K.; BARBOSA, L. F. M.; SIMÕES, B. F. T. SPC of multiple stream processes—a chart for enhanced detection of shifts in one stream. *Produção*, v. 21, n. 2, p. 242–253, 2011.

EPPRECHT, E. K.; LOUREIRO, L. D.; CHAKRABORTI, S. Effect of the amount of Phase I data on the Phase II performance of S^2 and S Control Charts. *Journal of Quality Technology*, v. 47, n. 2, p. 139-155, 2015.

EPPRECHT, E. K.; SIMÕES, B. F. T. Statistical Control of Multiple-Stream Processes-A Literature Review. Paper presented at the 11th International Workshop on Intelligent Statistical Quality Control, 20-23 Aug, Sydney, Australia, 2013. Disponível em: <http://wp.csiro.au/isqc2013/abstracts-bios/#abstracts> . Acesso em: 05 January 2016.

EVERITT, B. S.; HAND, D. J. *Finite Mixture Distributions*. Dordrecht: Springer Netherlands, 1981.

EXLER, R. B.; LIMA, C. J. B. Controle Estatístico de Processos (CEP): uma ferramenta para melhoria da qualidade. *ReAC*. v. 4, n. 3, p. 78-92, 2012.

FIRKA, D. Statistical, technical and sociological dimensions of design of experiments. *The TQM Journal*, v. 23, n. 4, p. 435-445, 2011.

FLOTT, L. W. Introduction to Control Charts. *Metal Finishing*, v. 110, n. 6, p. 36-38, 2012.

FONSECA, J. J. S. *Metodologia da pesquisa científica*. Fortaleza: UEC, 2002.

FRANCO, B. C.; CELANO, G.; CASTAGLIOLA, P.; COSTA, A. F. B. Economic design of Shewhart control charts for monitoring autocorrelated data with skip sampling strategies. *International Journal of Production Economics*, v. 151, p. 121-130, 2014.

GAY, L. R.; DIEHL, P. L. *Research Methods for Business and Management*. New York: Macmillan, 1992.

GOEDHART, R.; SILVA, M. M.; SCHOONHOVEN, M.; EPPRECHT, E. K.; CHAKRABORTI, S.; DOES, R. J. M. M.; VEIGA FILHO, A. L. Shewhart Control Charts for Dispersion Adjusted for Parameter Estimation. *IISE Transactions*, v. 49, n. 8, p. 838-848, 2017.

GOEDHART, R.; SCHOONHOVEN, M.; DOES, R. J. M. M. Nonparametric control of the conditional performance in statistical process monitoring. *Journal of Quality Technology*, p. 1-15, 2019.

GRASSO, M.; COLOSIMO, B. M.; SEMERARO, Q.; PACELLA, M. A Comparison Study of Distribution-Free Multivariate SPC Methods for Multimode Data. *Quality and Reliability Engineering International*, v. 31, n. 1, p. 75-96, 2014.

GRAVES, S. B.; MURPHY, D. C.; RINGUEST, J. L. Acceptance sampling versus redundancy as alternative means to achieving goals for system reliability, *International Journal of Quality & Reliability Management*, v. 16, n. 4, p. 362-370, 1999.

GRIMSHAW, S. D.; BRYCE, R. G.; MEADE, D. J. Control limits for group charts, *Quality Engineering*, v. 12, n. 2, p. 177-184, 1999.

GU, F.; HALL, P.; MILES, N. J. Performance evaluation for composites based on recycled polypropylene using principal component analysis and cluster analysis. *Journal of Cleaner Production*, v. 115, p. 343–353, 2016.

HADIAN H.; RAHIMIFARD, A. Multivariate statistical control chart and process capability indices for simultaneous monitoring of project duration and cost. *Computers & Industrial Engineering*, v. 130, p. 788-797, 2019.

HAIR JR, J. F.; BLACK, W. C.; BABIN, B. J.; ANDERSON, R. E. *Multivariate Data Analysis*. 8 ed. Boston: Cengage, 2019.

HALL, P.; ZHOU, X.-H. Nonparametric estimation of component distributions in a multivariate mixture, *The Annals of Statistics*, v. 31, n. 1, p. 201-224, 2003.

HOU, H.-M.; KUNG, J.-F.; HSU, Y.-B.; YAMAZAKI, Y.; MARUYAMA, K.; TOYOSHIMA, Y.; CHEN, C. Prediction of ppm level electrical failure by using physical variation analysis, *Metrology, Inspection, and Process Control for Microlithography XXX*, v. 9778, 2016.

HUANG, J.; NUALART, D.; VIITASAARI, L. A central limit theorem for the stochastic heat equation. *Stochastic Processes and Their Applications*, 2020.

HUANG, M.; WANG, Y.; SHIRINKAM, S.; ALAEDDINI, A.; YANG, K. Application of Multivariate Control Chart Techniques to Identifying Nonconforming Pallets in Automotive Assembly Plants, *SAE Technical Paper*, v. 2020-01-0477, 2020.

ISO QS-9000. Quality System Requirements QS-9000. International Organization for Standardization, 1998. Disponível em: <<http://ciiias.files.wordpress.com/2007/11/qs9000.pdf>>. Acesso em: 12 de novembro de 2019.

JARDIM, F. S.; CHAKRABORTI, S.; EPPRECHT, E. K. Effect of the Amount of Phase I Data on the Conditional Performance of the Phase II X Chart. *Technical Memorandum*, v. 01, 2017.

JARDIM, F. S.; CHAKRABORTI, S.; EPPRECHT, E. K. Chart with estimated parameters: the conditional ARL distribution and new insights. *Production and Operations Management*, v. 28, n. 6, p. 1545-1557, 2019.

JARDIM, F. S.; CHAKRABORTI, S.; EPPRECHT, E. K. Two perspectives for designing a phase II control chart with estimated parameters: The case of the Shewhart Chart. *Journal of Quality Technology*, p. 1–20, 2019.

JENSEN, W. A.; JONES-FARMER, L. A.; CHAM, P. C.W.; WOODALL, W. H. Effects of parameter estimation on control chart properties: a literature review, *Journal of Quality Technology*, v. 38, p. 349-364, 2006.

JIRASETTAPONG, P.; ROJANAROWAN, N. A guideline to select control charts for multiple stream processes control, *Engineering Journal*, v. 15, n. 3, p. 1-14, 2011.

JOLLIFFE, I. T.; CADIMA, J. Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, v. 374, n. 2065, 2016.

JOLLIFFE I. T. *Principal Component Analysis*, Series: Springer Series in Statistics, 2. ed., Springer, NY, 2002.

JOHNSON, R. A.; WICHERN, D. W. *Applied Multivariate Statistical Analysis*. 6 ed. Prentice Hall, 2007.

JURAN, J. M. *A qualidade desde o projeto: novos passos para o planejamento da qualidade em produtos e serviços*. Tradução: Nivaldo Montigelli Jr. São Paulo: Pioneira Thimson Learning, 1997.

KAFFEL, R. H.; PRIGENT, J. R. Optimal positioning in financial derivatives under mixture distributions, *Economic Modelling*, v. 52, n.a, p. 115-124, 2016.

KANG, J. H.; YU, J.; KIM, S. B. Adaptive nonparametric control chart for time-varying and multimodal processes. *Journal of Process Control*, v. 37, p. 34–45, 2016.

KAYID, M.; IZADKHAH, S. A new extended mixture model fresidual life time distributions, *Operations Research Letters*, v. 43, n. 2, p. 183-1, 2015.

KEEFE, M. J.; WOODALL, W. H.; JONES-FARMER, L. A. The Conditional In-Control Performance of Self-Starting Control Charts. *Quality Engineering*, v. 27, n. 4, p. 488–499, 2015.

KIM, S. M.; DO, T. T.; OECHTERING, T.; PETERS, G. On the entropy computation of large complex Gaussian mixture distributions, *IEEE Transactions on Signal Processing*, v. 63, n. 17, p. 4710-4723, 2015.

KITCHENHAM, B.; BRERETON, O. P.; BUDGEN, D.; TURNER, M.; BAILEY, J.; LINKMAN, S. Systematic literature reviews in software engineering—a systematic literature review. *Information and software technology*, v. 51, n. 1, p. 7–15, 2009.

KOH, K. L.; AHAD, N. A. Normality for Non-normal Distributions. *Journal of Science and Mathematics Letters*, v. 8, n. 2, p. 51-60, 2020.

LAKATOS, E. M.; MARCONI, M. A. *Fundamentos de Metodologia Científica*, 8 ed. São Paulo: Atlas, 2017.

LANNING, J. W.; MONTGOMERY, D. C.; RUNGER, G. C. Monitoring a Multiple Stream Filling Operation Using Fractional Samples, *Quality Engineering*, v. 15, n. 2, p. 183-195, 2002.

LEE, D.; CHA, S. K.; LEE, A. H. A Performance anomaly detection and analysis framework for DBMS development. *IEEE Transactions on Knowledge and Data Engineering*, v. 24, n. 8, p. 1345-1360, 2012.

LEE, M.; WANG, L.; SHEN, H.; HALL, P.; GUO, G.; MARRON, J. S. Least squares sieve estimation of mixture distributions with boundary effects. *Journal of the Korean Statistical Society*, v. 44, n. 2, p. 187-201, 2014.

LEI, X.; MACKENZIE, C. Distinguishing between common cause variation and special cause variation in a manufacturing system: A simulation of decision making for different types of variation. *International Journal of Production Economics*, v. 220, 2019.

LEVINSON, W. A. *Statistical process control for real-world applications*. Florida: CRC Press, 2011.

LI, S.; TANG, L. S.; NG, S. H. Nonparametric CUSUM and EWMA control charts for detecting mean shifts. *Journal of Quality Technology*, v. 42, n. 2, p. 209–26. 2010.

LIU, X.; MACKAY, R. J.; STEINER, S. H. Monitoring Multiple Stream Processes. *Quality Engineering*, v. 20, n. 3, p. 296–308, 2008.

LIU, J.; LIU, S. J.; WONG, D. S. H. Process Fault Diagnosis Based on Bayesian Inference. *Computer Aided Chemical Engineering*, v. 32, p. 751-756, 2013.

LOWRY, C. A.; MONTGOMERY, D. C. A review of multivariate control charts. *IIE Transactions*, v. 27, p. 800-810, 1995.

MACGREGOR, J. F.; JAECKLE, C.; KIPARISSIDES, C.; KOUTOUDI, M. Process monitoring and diagnosis by multi-block PLS methods. *AIChE Journal*, v. 40, n. 5, p. 826–838, 1994.

MAGALHÃES, M. S.; EPPRECHT, E. K.; COSTA, A. F. B. Economic Design of a Vp Xbar Chart. *International Journal of Production Economics*, v. 74, p. 191-200, 2001.

MANLY, B. J. F. ALBERTO, J. A. N. *Multivariate Statistical Methods: A Primer*. 4 ed. Boca Raton: CRC Press, 2017.

MAPLESOFT. How Maple™ Compares to Mathematica. Disponível em: <https://www.maplesoft.com/products/maple/compare/HowMapleComparestoMathematica.pdf>. Acesso em 08 de abril de 2020.

MARQUEZ, S. R.; VIVAS, J. J. Multivariate SPC methods for controlling manufacturing processes using predictive models – A case study in the automotive sector. *Computers in Industry*, v. 123, 2020.

MARTINS, R. A. *Abordagens quantitativas e qualitativas*. In: MIGUEL, Paulo A. M (org). Metodologia de pesquisas em engenharia de produção e gestão de operações. Rio de Janeiro: Elsevier, p. 45-61, 2010.

MARSHALL, C.; BEST, N.; BOTTLE, A.; AYLIN, P. Statistical issues in the prospective monitoring of health outcomes across multiple units. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, v. 167, n. 3, p. 541–559, 2004.

MASMOUDI, A.; MASMOUDI, A.; PUECH, W. A New Semiparametric Finite Mixture Model-Based Adaptive Arithmetic Coding for Lossless Image Compression, *Circuits Syst Signal Process*, v. 35, n. 4, p. 1163-1186, 2016.

MASON, R. L.; YOUNG, J. C. *Multivariate Statistical Process Control with Industrial Application*. Philadelphia: Society For Industrial And Applied Mathematics, 2001.

MASON, R. L.; TRACY, N. D.; YOUNG, J. C. A practical approach for interpreting multivariate T^2 control chart signals. *Journal of Quality Technology*, v. 29, n.4, 1997.

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

MASTRANGELO, C. M.; RUNGER, G. C.; MONTGOMERY, D. C. Statistical process monitoring with principal components. *Quality and Reliability Engineering International*, v. 12, p. 203-210, 1996.

MCLACHLAN, G. J.; LEE, S. X.; RATHNAYAKE, S. I. Finite Mixture Models. *Annual Review of Statistics and Its Application*, v. 6, n. 1, 2019.

MEI, Y. Efficient scalable schemes for monitoring a large number of data streams. *Biometrika*, v. 9, n. 2, p. 419–433, 2010.

MEI, Y. Is Average Run Length to False Alarm Always an Information Criterion?, *Sequential Analysis*, Vol. 27, pp. 351-376, 2008.

MENECES, N. S.; OLIVERA, S. A.; SACCONI, C. D.; TESSORE, J. Statistical Control of Multiple-Stream Processes: A Shewhart Control Chart for Each Stream. *Quality Engineering*, v. 20, n. 2, p. 185–194, 2008.

MEREDITH, J. Alternative research paradigms in operations. *Journal of Operations Management*, v. 8, n. 4, p. 297–326, 1989.

MIGUEL, P. A. C.; FLEURY, A.; MELLO, C. H. P.; NAKANO, D. N.; TURRIONI, J. B.; HO, L. L.; *et al.* *Metodologia de pesquisa em engenharia de produção e gestão de operações*. Rio de Janeiro: Elsevier, 2010.

MINNOTTE, M. C. Nonparametric testing of the existence of modes. *Ann. Statist.*, v. 25, p. 1646–1660, 1997.

MONTGOMERY, D. C. *Introduction to Statistical Quality Control*, 7 ed. Hoboken, NJ: John Wiley & Sons, Inc, 2013.

MORABITO, R.; PUREZA, V. Modelagem e simulação. In: CAUCHICK MIGUEL, P.A.C. *et al.* *Metodologia de pesquisa em engenharia de produção e gestão de operações*. Rio de Janeiro: Elsevier, 2010. p. 165-192.

MORTELL, R. R.; RUNGER, G. C. Statistical process control for multiple stream processes. *Journal of Quality Technology*, v. 27, n. 1, p. 1-12, 1995.

NELSON, L. S. Control Chart for Multiple Stream Processes. *Journal of Quality Technology*, v. 18, n. 4, p. 255–256, 1986.

NENES, G.; TASIAS, K. A.; CELANO, G. A general model for the economic-statistical design of adaptive control charts for processes subject to multiple assignable causes. *International Journal of Production Research*, v. 53, n. 7, p. 2146–2164, 2014.

NEUMAN, W. L. *Social research methods: qualitative and quantitative approaches*. 6 ed. Boston: Allyn and Bacon, 2006.

NEZHAD, M. S. F.; NIAKI, S. T. A. A Max-EWMA approach to monitor and diagnose faults of multivariate quality control processes. *Int J Adv Manuf Technol*, v. 68, p. 2283–2294, 2013.

NOOR, S.; NOOR-UL-AMIN, M.; ABBASI, S. A. Bayesian EWMA control charts based on Exponential and transformed Exponential distributions, *Quality Reliability Engineering International*, v. 37, n. 4, p. 1-21, 2020.

OPRIME, P. C.; LIZARELLI, F. L.; PIMENTA, M. L.; ACHCAR, J. A. Acceptance X-bar chart considering the sample distribution of capability indices, C_p and C_{pk} . *International Journal of Quality & Reliability Management*, v. 36, n. 6, p. 875-894, 2019.

OPRIME, P. C.; TOLEDO, J. C. GONZÁLEZ, M. O. A.; CHAKRABORTI, S. Método de estimativa dos limites da carta de controle não paramétrica que monitora simultaneamente a média e variância. *Gest. Prod., São Carlos*, v. 23, n. 1, p. 146-164, 2016.

OTT, E. R.; SNEE, R. D. Identifying Useful Differences in a Multiple-Head Machine. *Journal of Quality Technology*, v. 5, n. 2, p. 47–57, 1973.

PENG, X.; LI, Z.; ZHONG, W.; QIAN, F.; TIAN, Y. Concurrent Quality-Relevant Canonical Correlation Analysis for Nonlinear Continuous Process Decomposition and Monitoring. *Industrial & Engineering Chemistry Research*, v. 59, p. 8757-8768, 2020.

PETER HE, Q.; WANG, J. Statistics Pattern Analysis: A Statistical Process Monitoring Tool for Smart Manufacturing. *Computer Aided Chemical Engineering*, v. 44, p. 2071–2076, 2018.

PETERSEN, K.; FELDT, R.; MUJTABA, S.; MATTSSON, M. Systematic mapping studies in software engineering. In *Proceedings of the 12th International Conference on Evaluation and Assessment in Software Engineering*, Swindon, UK. BCS Learning & Development Ltd, p. 68–77, 2008.

PIOVESANA, A.; SENIOR, G. How Small Is Big: Sample Size and Skewness. *Assessment*, v. 25, n. 6, 793–800, 2016.

PHALADIGANON, P.; KIM, S.B.; CHEN, V.C.; JIANG, W. Principal component analysis-based control charts for multivariate nonnormal distributions, *Expert Systems Applications*, v. 40, n. 8, p. 3044-3054, 2013.

POLANSKY, A. M. A general framework for constructing control charts. *Quality and Reliability Engineering International*. v. 21, n. 6, p. 633-653, 2005.

PRABHU, S. S.; MONTGOMERY, D. C.; RUNGER, G. C. A Combined Adaptive Sample Size and Sampling Interval \bar{X} Control Scheme. *Journal of Quality Technology*, v. 26, n. 3, p. 164–176, 1994.

RALAIVAOSAONA, D.; WAGNER, S. A central limit theorem for additive functionals of increasing trees, *Combinatorics, Probability and Computing*, v. 28, n. 4, p. 618-637, 2019.

RAMAKER, H.-J.; VAN SPRANG, E. N. M.; WESTERHUIS, J. A.; SMILDE, A. K. The effect of the size of the training set and number of principal components on the false alarm rate in statistical process monitoring. *Chemometrics and Intelligent Laboratory Systems*, v. 73, n. 2, p. 181–187, 2004.

RATO, T. J.; REIS, M. S. Improved Fault Diagnosis in Online Process Monitoring of Complex Networked Processes: a Data-Driven Approach. *Computer Aided Chemical Engineering*, v. 40, Pages 1681-1686, 2017.

RUNGER, G.; ALT, F. B.; MONTGOMERY, D. Controlling multiple stream processes with principal components. *International Journal of Production Research*, v. 34, n. 11, p. 2991-2999, 1996.

- SÁ, J. *Applied Statistics Using Spss, STATISTICA, Matlab and R*. Berlin: Springer, 2007.
- SALAH, B.; ZOHEIR, M.; SLIMANE, Z.; JURGEN, B. Inferential sensor-based adaptive principal components analysis of mould bath level for breakout defect detection and evaluation in continuous casting. *Applied Soft Computing*, v. 34, p. 120–128, 2015.
- SALEH, N. A.; MAHMOUD, M. A.; KEEFE, M. J.; WOODALL, W. H. The difficult in designing Shewhart \bar{x} and \bar{X} Control Charts with estimated parameters. *Journal of Quality Technology*, v. 47, n. 2, p. 127-138, 2015.
- SALO, F.; NASSIF, A. B.; ESSEX, A. Dimensionality Reduction with IG-PCA and Ensemble Classifier for Network Intrusion Detection. *Computer Networks*, v. 148, p. 164-175, 2019.
- SCHOONHOVEN, M.; DOES, R. J. M. M. A Robust Standard Deviation Control Chart. *Technometrics*, v. 54, n. 1, p. 73–82, 2012.
- SHAMS, M. B.; ELKAMEL, A.; MOORTHY, K.; RAFINEJAD, G.; SAXENA, A. Optimal design of T2 monitoring chart for chemical processes. *International Journal of Process Systems Engineering*, v. 3, p. 232–247, 2015.
- SHANG, Y; TSUNG, F; ZOU, C. Statistical process control for multistage processes with binary outputs. *IIE Transactions*, v. 45, n. 9, p. 1008-1023, 2013.
- SHAO, Y. E.; CHANG, P.-Y.; LU, C.-J. Applying Two-Stage Neural Network Based Classifiers to the Identification of Mixture Control Chart Patterns for an SPC-EPC Process. *Complexity*, v. 2017, p. 1–10, 2017.
- SHORE, H. Non-normal populations in quality applications: a revisited perspective. *Quality and Reliability Engineering International*. v. 20, n. 4, p. 375–382, 2004.
- SIDAK, Z. Rectangular Confidence Regions for the Means of Multivariate Normal Distributions. *Journal of the American Statistical Association*, v. 62, n. 318, p. 626-633, 1967.
- SIVASAMY, A.; SUNDAN, B. A Dynamic Intrusion Detection System Based on Multivariate Hotelling's T² Statistics Approach for Network Environments. *The Scientific World Journal*, v. 2015, p. 1–9, 2015.
- SOBUE, C. E. F.; JARDIM, F. S.; CAMARGO, V. C. B.; LIZARELLI, F. L.; OPRIME, P. C. Unconditional performance of the \bar{X} chart: Comparison among five standard deviation estimators. *Quality and Reliability Engineering International*, v. 1, p. 12, 2020.
- SPIEGEL, M. R.; SCHILLER, J. J.; SRINIVASAN, A. R. *Probabilidade e Estatística*. Coleção Schaum. 3 ed. Porto Alegre: Bookman Companhia, 2013.

STATISTICA STAT SOFT, Inc. 2015. STATISTICA (data analysis software system) Version 13.0.

STOUMBOS, Z. G.; REYNOLDS, M. R. J. Robustness to non-normality and autocorrelation of individuals control charts. *Journal of Statistical Computation and Simulation*, v. 66, p. 145-187, 2000.

STUBBS, S.; ZHANG, J.; MORRIS, J. BioProcess performance monitoring using multiway interval partial least squares. *Process Systems Engineering for Pharmaceutical Manufacturing*, v. 41, p. 243-259, 2018.

STRICKER, C.; REDMAN, S. Quantal analysis based on density estimation, *Journal of Neuroscience Methods*, v. 130, p. 159-171, 2003.

Suárez, M. F. B.; Rodríguez, F. G. G. Cornerstone root causes through the analysis of the Ishikawa diagram, is it possible to find them? A first research approach, *International Journal of Quality and Service Sciences*, v. 11, n. 2, p. 302-316, 2019.

TASIAS, K. A.; NENES, G. Economic-statistical design of Vp control schemes for joint monitoring of mean and variance in the presence of multiple assignable causes. *Quality Technology & Quantitative Management*, v. 15, n. 4, p. 484-506, 2016.

TEEL, C.; PARK, T.; SAMPSON, A. R. Communications in Statistics, *Simulation and Computation*, v. 44, n. 6, p. 1545-1556, 2015.

THISSEN, U.; SWIERENGA, H.; DE WEIJER, A. P.; WEHRENS, R.; MELSSEN, W. J.; BUYDENS, L. M. C. Multivariate statistical process control using mixture modelling. *Journal of Chemometrics*, v. 19, n. 1, p. 23-31, 2005.

TIAN, W.; JIN, R.; HUANG, T.; CAMELIO, J. A. Statistical process control for multistage processes with non-repeating cyclic profiles. *IIE Transactions*, v. 49, n. 3, p. 320-331, 2016.

TITTERINGTON, D. M.; SMITH, A. F. M.; MAKOV, U. E. *Statistical analysis of finite mixture distributions*. Wiley, Universidade da Califórnia, 1985.

VAL, A. G. D.; SAWIK, B.; AGUSTÍN, A.; FAULÍN, J.; DIÉGUEZ, P. M. A SPC Strategy for Decision Making in Manufacturing Processes. *Decision Making in Manufacturing and Services*, v. 13, n. 1-2, p. 5-15, 2019.

VANHATALO, E.; KULAHCI, M.; BERGQUIST, B. On the structure of dynamic principal component analysis used in statistical process monitoring. *Chemometrics and Intelligent Laboratory Systems*, v. 167, p. 1-11, 2017.

VARIYATH, A. M.; VATTATHOO, J. Robust Control Charts for Monitoring Process Variability in Phase I Multivariate Individual Observations, *Quality and Reliability*

Engineering International, v. 30, n. 6, p. 795-812, 2014.

VICENTIN, D. S.; SILVA, B. B.; PICCIRILLO, I.; BUENO, F. B.; OPRIME, P. C. Monitoring process control chart with finite mixture probability distribution: An application in manufacture industry, *International Journal of Quality & Reliability Management.*, v. 35, n. 2, p. 335–353, 2018.

VON ROSING, M.; SCHEER, A.-W.; ZACHMAN, J. A.; JONES, D. T.; WOMACK, J. P.; VON SCHEEL, H. Phase 3: Process Concept Evolution. *The Complete Business Process Handbook*, p. 37–77, 2015.

XIANG, L.; TSUNG, F. Statistical monitoring of multi-stage processes based on engineering models. *IIE Transactions*, v. 40, n. 10, p. 957–970, 2008.

WANG, D.-S.; CHOU, C.-Y.; LIN, Y.-C. On the Asymptotic Confidence Intervals of Multiple-Stream Yield Index S_{pk}^M , *Communications in Statistics-Theory and Methods*, v. 40, n. 22, p. 3948-3958, 2011.

WANG, D.; ZHANG, L.; XIONG, Q. A non parametric CUSUM control chart based on the Mann–Whitney statistic. *Communications in Statistics - Theory and Methods*, v, 46, n. 10, p. 4713–25, 2017.

WANG, F. K. Process Yield for Multiple Stream Processes with Individual Observations and Subsamples, *Quality and Reliability Engineering International*, v. 32, n. 2, p. 335–344, 2015.

WISE, B. M.; GALLAGHER, N. B. The process chemometrics approach to process monitoring and fault detection. *J Proc Contr*, v. 6, n. 6, p. 329–348, 1996.

WOODALL, W. H. Bridging the Gap between Theory and Practice in Basic Statistical Process Monitoring. *Quality Engineering*, v.29, n. 1, p. 2–15, 2016.

WOODALL, W. H. Controversies and contradictions in Statistical Process Control. *Journal of Quality Technology*, v. 32, n. 4, p. 12-13, 2000.

WOODALL, W. H. The statistical design of quality control charts. *Journal of the Royal Statistical Society*, v. 34, p. 155-160, 1985.

WOODALL, W. H.; FALTIN, F. W. Rethinking control chart design and evaluation. *Quality Engineering*, v. 31, n. 4, p. 596–605, 2019.

WOODALL, W. H.; MONTGOMERY, D. C. Some Current Directions in the Theory and Application of Statistical Process Monitoring. *Journal of Quality Technology*, v. 46, n. 1, p. 78-95, 2014.

WOODALL, W. H.; NCUBE, M. M. Multivariate CUSUM Quality-Control Procedures.

Technometrics, v. 27, n. 3, p. 285-292, 1985.

YANG, M.; WU, Z.; LEE, K. L.; KHOO, M. B. C. The X control chart for monitoring process shifts in mean and variance, *International Journal of Production Research*, v. 50, n. 3, p. 893-907, 2012.

YU, J.; LIU, J. LRProb control chart based on logistic regression for monitoring mean shifts of auto-correlated manufacturing processes, *International Journal of Production Research*, v. 49, n. 8, p. 2301-2326, 2012.

ZHANG, H.; ALBIN, S. Determining the number of operational modes in baseline multivariate SPC data, *IIE Transactions*, v. 39, n. 12, p. 1103-1110, 2007.

ZHANG, F.; WENG, Z. Mixture Principal Component Analysis Model for Multivariate Processes Monitoring. *Journal of Advanced Manufacturing Systems*, v. 4, n. 2, p. 151-166, 2005.

ZHOU, B.; YE, H.; ZHANG, H.; LI, M. Process monitoring of iron-making process in a blast furnace with PCA-based methods. *Control Engineering Practice*, v. 47, p. 1-14, 2016.

ZHOU, W.; CHENG, C.; ZHENG, Z. Optimal design of an attribute control chart for monitoring the mean of autocorrelated processes. *Computers & Industrial Engineering*, v. 137, p. 106081, 2019.

ZWETSLOOT, I. M.; WOODALL, W. H. A review of some sampling and aggregation strategies for basic statistical process monitoring. *Journal of Quality Technology*, p. 1-16, 2019.

APÊNDICES

APÊNDICE 1. REVISÃO SISTEMÁTICA DE LITERATURA PARA MÚLTIPLOS FLUXOS

1. REVISÃO DA LITERATURA

Com o objetivo de revisar na literatura as publicações que relatam *Multiple-Stream Process* (MSP) na área de Controle Estatístico de Processo, foi realizado o processo de revisão sistemática. Para um maior rigor na execução do método, foram consideradas as diretrizes formalizadas por Petersen *et al.* (2008) e Kitchenham *et al.* (2009). A revisão sistemática, segue uma estratégia definida e maleável, visando responder a uma ou mais questões de pesquisa de forma imparcial e replicável (PETERSEN *et al.*, 2008; KITCHENHAM *et al.*, 2009).

A metodologia da revisão sistemática segundo Kitchenham (2007) descreve um método de pesquisa que segue as seguintes fases:

- definição de um protocolo, que especifica as questões de pesquisa e os métodos que serão usados para sua realização;
- estratégia de busca definida pelo pesquisador (fontes de busca, *string* de busca, período), objetivando identificar o máximo possível de estudos relevantes à questão de pesquisa;
- critérios explícitos de inclusão e exclusão de estudos primários, de forma a avaliar a necessidade destes estudos para a pesquisa;
- critérios de qualidade para a avaliação de cada estudo, pré-requisito para uma meta-análise quantitativa de dados.

As etapas da revisão sistemática podem ser resumidas segundo Kitchenham (2007) em três fases: (i) planejamento, (ii) condução e (iii) apresentação.

Na etapa de planejamento é onde caracteriza o andamento da revisão sistemática, definindo as questões de pesquisa a serem respondidas, sendo capaz de responder os objetivos e quais devem ser listados, *strings* de buscas, as fontes de pesquisas, critérios de inclusão e exclusão, como e quais critérios foram utilizados para avaliar a qualidade dos estudos, os critérios de qualidade, extração e sintetização dos dados, suas diferenças e análises documentadas (KITCHENHAM, 2007; TRAVASSOS, 2007).

A etapa de Condução da Revisão para Travassos e Biolchini (2007) divide-se em fases como a: busca primária onde utiliza as *strings* nas fontes definidas e os resultados coletados formam os estudos primários, que são estabelecidos de acordo com os protocolo da revisão e com os critérios de exclusão e inclusão obtendo os resultados que serão utilizados. Na Avaliação de qualidade dos estudos primários e a fase onde após a seleção são avaliados em critérios de qualidade estabelecidos evitando resultados tendenciosos e maximizando a validade. E por último a extração dos dados de estudos primários que são analisadas de acordo com as questões de pesquisa.

Assim na etapa final a de Apresentação da revisão são extraídos os dados e a sumarização dos resultados em cima dos dados coletados.

1.2 PLANEJAMENTO DE REVISÃO

A revisão sistemática inicia-se com o protocolo de revisão, onde descreve todos o processo e os métodos que serão aplicados. Serão definidos os objetivos, as questões de pesquisas, escopos, as fontes, busca primária e *strings*.

Esta revisão sistemática teve como objetivo de identificar abordagens de *Multiple Stream Process* na área de Controle Estatístico de Processo, onde visa analisar as publicações de artigos e teses relevantes por meio da revisão sistemática, identificando o que vem sendo estudado e desenvolvido na área.

Assim partindo do objetivo de investigação definimos as questões de pesquisa que auxiliaram na revisão sistemática.

- (QP1): – Quantas publicações relatam "*multiple stream process*" na área de controle estatístico de processo?
- (QP2): – Quais abordagens são utilizadas nessas publicações?

Com objetivo de assegurar a viabilidade da pesquisa foi definido o escopo, onde foi realizada a escolha das bases de pesquisa, fontes e descrições seguindo os seguintes critérios:

- Disponibilidade para consultas web;
- Disponibilidade de artigos em inglês ou português;
- Presença de mecanismo de busca que faça uso de palavras-chave;
- Relevância da fonte;
- Artigos duplicados (em mais de uma fonte de busca) devem ser considerados apenas a sua versão mais recente ou a mais completa;
- Artigos e teses devem mencionar as palavras-chave da pesquisa no título, resumo ou nas palavras-chave do artigo;

- fontes que possibilitam consultas de forma gratuita (também serão consideradas fontes que oferecem consultas de forma gratuita a partir do acesso pelo domínio da UFSCar)
- A pesquisa esteve restrita aos resultados publicados entre 01 de janeiro de 2010 até 2020.

Assim a seleção da fonte de pesquisa foi escolhida por meio exploração do tema de forma abrangente para verificar e definir as bases científicas a serem investigadas. Como resultado dessa etapa, definiu-se como relevante as seguintes bases: SciELO, SCOPUS e Google Acadêmico.

Na busca primária encontramos os estudos nas bases de pesquisa, onde foram feitas buscas manuais e automáticas para formulação da *string* de busca, e posteriormente aplicar nas bases. A partir disso foi elaborada a seguinte string geral de busca para selecionar publicações que relatam o uso *Multiple Stream Process* na área de controle estatístico de processo.

Inglês: (“*Multiple-Stream Process*” OR *MSP*) AND (“*Monitoring Processes Control*” “*Statistical Process Control*” OR *MPC* OR *SPC* OR *Statistical* OR *Process* OR *Control*)

Português: (“Processos de Múltiplos Fluxos” OR *PMF*) AND (“Monitoramento Estatístico de Processos” “Controle Estatístico de Processos” OR *MEP* OR *CEP* OR *Estatística* OR *Processo* OR *Controle*)

1.3 EXTRAÇÃO DE DADOS

Para realizar a extração de dados dos artigos selecionados na revisão sistemática de literatura, definiu-se os seguintes critérios:

- i. Atividades realizadas e instrumentos utilizados nos estudos analisados;
- ii. Tipo de abordagem, método e metodologia, que os autores adotaram nas pesquisas;
- iii. Resultados encontrados, analisados e discutidos;
- iv. Lições aprendidas nas pesquisas realizadas pelos autores e apontamentos futuros.

Esses critérios foram analisados no item 1.6, onde contém os títulos dos artigos selecionados para a revisão com as análises e discussões dos critérios aqui apresentados.

1.4 RESULTADOS DAS BUSCAS

Na busca primária, com refinamento das *strings*, foram encontrados 672 arquivos. Cada estudo foi avaliado pela pesquisadora em questão, a qual decidiu selecionar 6 estudos, pois no Scopus dos 570 inicialmente encontrados apenas 4 se enquadravam nos critérios de inclusão. No Google Académico, de 51 artigos encontrados apenas 3 se enquadraram, porém 2 eram repetidos e o outro uma tese, a qual também não se insere no critério de inclusão. Por fim, na base da Scielo foram encontrados inicialmente 51 estudos, os quais após o refinamento nenhum foi selecionado, devido o único que se enquadrava no critério já estar contabilizado.

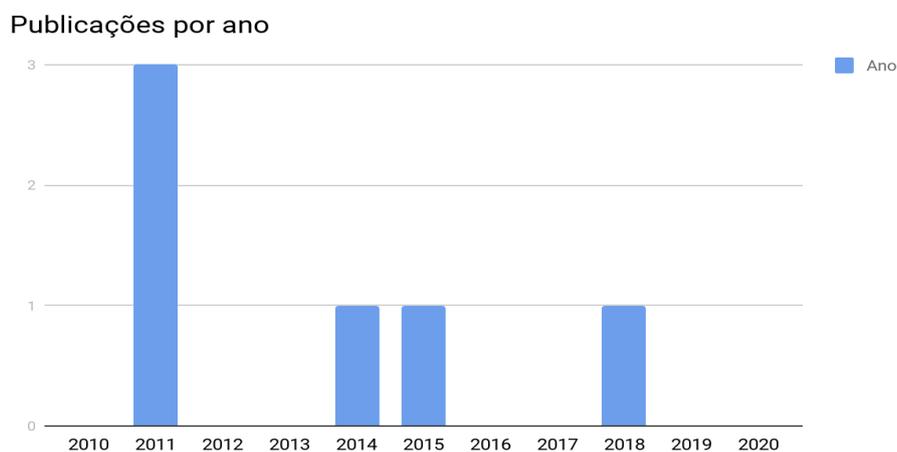
Os artigos selecionados nas diferentes etapas seguem representados no Quadro 1.

Quadro 1- Publicações incluídas e excluídas por base

	Scielo	Scopus	Google Acadêmico	Total
Artigos incluídos	0	5	1	6
Artigos excluídos	51	565	50	666
Total artigos	51	570	51	672

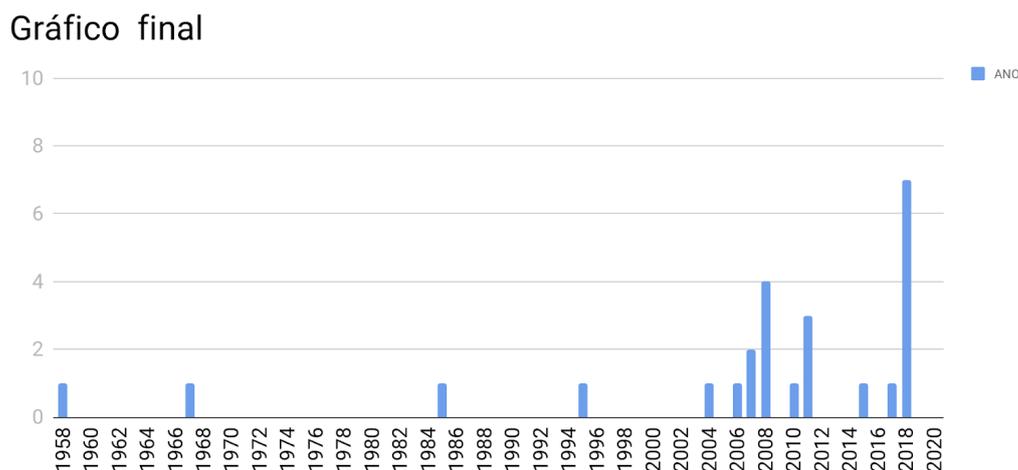
Com isso, o refinamento dos seis artigos, cujos de publicação estão apresentados a seguir (Figura 4), considerou que todos estavam de acordo com os critérios. Os quais foram: introdução, metodologia e conclusão dos mesmos.

Figura 1 - Publicações do tema por ano



Aplicando-se a técnica de *snowballing* nesses seis artigos, um deles era uma revisão sistemática. Conseguimos incluir mais 19 artigos, que se enquadraram em nossos critérios de inclusão, dos 35 que foram mencionados no artigo revisado. Totalizando 25 artigos, que se encontram na revisão sistemática. Na Figura 5, há a distribuição pelo ano de publicação desses.

Figura 2 - Gráfico final de publicações do tema por ano



1.5 DISCUSSÕES

A partir da leitura completa dos artigos, foram extraídos dados e verificadas as tendências para encontrar evidências que possam sugerir futuros estudos. Neste item, apresentam-se resumidamente os dados, a análise e discussão dos mesmos.

O objetivo foi compreender quais abordagens os autores estão utilizando e, assim, pode-se compreender que estudos foram apontados como relevantes para o tema de processos de múltiplos fluxos, a fim de modelar diferentes situações do que a literatura já expôs e compreender o monitoramento das variáveis consideradas. Ao mesmo tempo mostraram que não há como eleger o melhor modelo, já que cada situação tem suas peculiaridades, o que demanda estudos em diferentes áreas, mas ainda podem contribuir em outras áreas para as quais não foi considerado, servindo de referencial.

Após essa busca dos últimos 10 anos encontrou-se um artigo relevante¹ em que o autor realizou um mapeamento desde quando a temática que foi proposta a estudar começou a

¹ Statistical Control of Multiple-Stream Processes: A Literature Review, escrito por Eugenio K. Epprecht.

ser campo de interesse de estudos. Analisou-se outros 19 artigos que não estavam analisados nesse trabalho do autor, os quais serão discutidos a seguir.

As publicações selecionadas para análise respondem a questão QP1. Em relação a QP2, as abordagens mais adotadas são o uso de modelos que priorizam o controle por meio de gráficos de controle estatístico (análise dos parâmetros ou não) e modelagem dos sistemas por integrações e derivações variadas, distribuição normal; adequadas a cada modelo que se pretende estudar. Que discutiram após a modelagem o que é eficaz ou não, no que supera ou não o rendimento e a correlação entre os fluxos.

Alguns artigos trabalham com testes, dentre esses destaco o teste F e o de razão de verossimilhança. Outras publicações trouxeram ilustrações das simulações numéricas e discussões com o auxílio da heurística, e retângulo de confiança.

As sugestões que os trabalhos apontam é que haja mais estudos sobre o tema, buscando a partir dos artigos já realizados compor pesquisas a serem testadas no modelo pouco explorado pela literatura. Também pode-se notar que na área biológica os métodos estatísticos são usados em maior quantidade do que na área industrial.

1.6 DISCUSSÕES DOS ARTIGOS PESQUISADOS

A seguir segue as análises e discussões dos artigos que foram utilizados na pesquisa de Revisão Sistemática de Literatura para o tema da tese em questão.

1.6.1 SPC of multiple stream processes - a chart for enhanced detection of shifts in one stream (2011) - Epprecht, E. K., Barbosa, L. F. M., Simões, B. F. T.

O artigo em questão traz uma discussão acerca da eficiência dos gráficos de controle de grupos propostos por Boyd, para o processo de múltiplos fluxos. Ao trazer ao modelo a questão de a variabilidade ser comum em todos os canais comprometendo a eficácia dessa técnica. Assim, apresenta modelos alternativos que levam em conta essa questão e elucidam um modelo próprio, mostrando que há poucos estudos na área, já que há menos de duas décadas não se pesquisava a fundo sobre as poderia haver duas características de qualidade dos múltiplos fluxos: uma comum a todos os fluxos e outra individual a cada fluxo. No novo modelo apresentam um gráfico de controle de grupos modificado, o qual toma as diferenças entre valores da qualidade em cada canal, além da média dos seus valores em todos canais Os dados foram obtidos analiticamente ou por simulação, comparando os dados dos canais discrepantes. Superou-se os outros modelos no ganho de desempenho, exceto em um caso: alterações da média de magnitude inferior a um desvio-padrão, mas que nos outros modelos já não era contemplado.

1.6.2 Statistical Control of Multiple-Stream Processes: A Literature Review (2015) - Epprecht, E. K.

O artigo tem como objetivo fazer um levantamento bibliográfico sobre o que se tem pesquisado sobre o controle estatístico de processos industriais de múltiplos fluxos. Apontou que há poucas publicações na área, mesmo o tema sendo relevante, trazendo apontamentos para futuras pesquisas no campo industrial, já que em outras áreas o avanço nas pesquisas considera-se qualitativos e quantitativamente presente. O que pode vir a contribuir a área de interesse do autor. O que reforça a relevância da pesquisa bibliográfica, pois traz o que

está sendo pesquisado na área afim e como a partir dela ampliar pesquisas para áreas particulares, ou seja, aplicados a outros problemas/processos.

1.6.3 On the Asymptotic Confidence Intervals of Multiple-Stream Yield Index S_{pk}^M (2011) - Dja-Shin Wang, Chao-Yu Chou e Yu-Chang Lin

Ao considerarem a medida adimensional, índice de capacidade de um processo (CPK), uma forma simples de avaliar qualitativamente a capacidade e normalidade de um processo, sob um processo estável, trouxeram como proposta medir múltiplos fluxos, a partir do índice de rendimento (S_{pk}^M), em aplicações práticas industriais. A problemática se faz em torno da construção da confiança de intervalo para o índice de rendimento. Utilizou-se métodos de derivação e de integração por meio de software matemático, tabelas foram obtidas oferecendo dados para a construção do intervalo de confiança do S_{pk}^M e por fim aplicadas a problemas reais de um processo de produção para ilustrá-las.

1.6.4 Monitoring process control chart with finite mixture probability distribution: An application in manufacture industry (2018) – Vicentin, D. S., Silva, B.B., Piccirillo, I., Bueno, F.B., Oprime, P.C.

O estudo tem como objetivo o desenvolvimento do Monitoramento de Processos de Múltiplos Fluxos, a partir de distribuições de probabilidade de mistura finita, na indústria. O estudo foi realizado com dados extraídos de uma indústria alimentícia que produz massas à base de trigo que foram organizados em gráficos de controle. A organização se deu primeiramente com a coleta da amostra mestre, a partir de diversos lotes da produção em questão; depois por métodos gráficos verificou-se a distribuição de probabilidade da mistura finita. Após ajustar um modelo teórico a essa distribuição de cada subpopulação do lote em

questão elaborou-se um modelo estatísticos e assim delimitou-se os limites de controle para comparar os gráficos (mistura e controle tradicional de Shewhart). O monitoramento do fluxo foi feito por meio de análise gráfica. A relevância do trabalho está na tabela proposta que poder ser utilizada como ferramenta de controle de qualidade dos produtos recebidos de diferentes lotes.

1.6.5 A Guideline to Select Control Charts for Multiple Stream Processes Control (2011) - Jirasettapong, P., Rojanarowan, N.

O artigo traz que o gráfico de controle é uma das ferramentas para controlar os processos de múltiplos fluxos, de maneira a serem selecionados adequadamente a fim de minimizar a possibilidade de uma falsa decisão e ou irrelevantes sobre o processo. O estudo focou no controle considerando o número de fluxos e a correlação entre esses, a limitação do uso de gráficos e a diferença entre a média dos fluxos e o tamanho médio dos fluxos detectados. Portanto dentre os diferentes gráficos não pode-se eleger o melhor, pois depende do caso a ser analisado.

1.6.6 Controlling Multiple Stream Processes with Principal Components (2007) - Bersimis, S., Psarakis, S. and Panaretos, J.

A análise dos processos de múltiplos fluxos é feita por um gráfico de controle multivariado. Os resultados foram apresentados a partir da análise de uma distribuição qui-quadrada não central cujos parâmetros estão relacionados ao número e proporção de fluxos. Para aprimorar o desempenho do gráfico utilizou-se EWMA e PCA para projeção do esquema

de controle. Sendo que o gráfico nesse modelo foi o mais apropriado por usar a média de todos os fluxos.

1.6.7 Estimation of the Means of Dependent Variables (1958) - Dunn, O. J.

O estudo aborda a utilização de métodos para a construção de conjuntos de intervalos de confiança de distribuição normal multivariada. Surge então, a probabilidade de um alarme falso de um grupo gráfico, ocasionado pelo uso incorreto de limites no GCC.

1.6.8 Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing (1995) - Benjamini, Yoav., Yosef Hochberg.

O estudo traz uma inovação ao propor uma abordagem diferente das publicações anteriores sobre o controle do FDR, reduzindo o custo do controle da multiplicidade não precisa ser grande. Mas não apresentam como empregá-los no projeto de monitoramento para processo.

1.6.9 EWMA Control Charts for the Smallest and Largest Observations (1990-2018) - Amin, R. W., Wolff, H., Besenfelder, W., & Baxley, R.

Com base nas menores e maiores observações de cada amostra foi construído o gráfico EWMA. Esse gráfico mostra os limites de tolerância, ou seja, mostra quais parâmetros aumentam e ou diminuem, recebendo o nome de MaxMin EWMA. Para aproximar o tempo médio de execução propuseram pela cadeia Markov, o que traz eficácia ao método em mais de 100 casos testados, oferecendo informações sobre o tipo de mudança.

1.6.10 Process Capability Indices for Multiple Stream Processes (2008) - Bothe, D. R

O índice Cpk médio é uma medida de capacidade da saída de todos os fluxos. É um método para avaliar a qualidade de saída das misturas de processos múltiplos de fluxo, podendo ser utilizado em controle de qualidade de peças e tributos.

1.6.11 \bar{X} Chart with Variable Sample Size and Sampling Intervals (1997) - Costa, A. F. B.

Os gráficos \bar{X} com intervalos de amostragem variáveis e os com tamanho de amostra variável são mais rápidos que o gráfico \bar{X} tradicional na detecção de mudanças em processo. As análises foram realizadas usando o método da cadeia de Markov, o que resulta em menor esforço de amostragem, aumentando a eficiência do monitoramento.

1.6.12 Control Limits for Group Charts (1999-2007) - Grimshaw, S. D., Bryce, R. G., Meade, D. J.

O estudo traz a discussão do uso inadequado de limites de 3-sigma com GCC. A redução no número de gráficos para controle do processo veio de um ajuste os limites com a Correção de Dunn-Sidak o que ofereceu tabelas de constantes para calcular limites.

1.6.13 Effects of Parameter Estimation on Control Chart Properties: A Literature Reviewn (2006) - Jensen, W.A., Jones-Farmer, L.A., Cham, P.C.W. and Woodall, W.H.

No artigo é realizada uma revisão da literatura sobre o efeito de estimar parâmetros das propriedades do gráfico de controle. Chegando à conclusão de que muitos estudos estão voltados a pesquisas sobre determinação das propriedades de gráficos de controle quando os parâmetros são estimados e aponta para que novos estudos sejam ampliados na área.

1.6.14 Monitoring a Multiple Stream Filling Operation Using Fractional Samples (2002) - Lanning, J. W., Montgomery, D. C., Runger, G. C

Nesse estudo é explorado que o processo múltiplo de fluxo seja monitorado por parte do total de fluxos, ótimo método para indústrias como grande fluxo e velocidade de produção em que o processo não é totalmente automatizado.

1.6.15 Monitoring Multiple Stream Processes (2008) - Liu, X., MacKay, R. J., & Steiner, S. H.

Discute-se a importância de os processos de múltiplos fluxos detectarem alterações na média geral do processo e na corrente individual. O controle é proposto com teste F e no de razão de verossimilhança. Para assim detectar vieses em alguns medidores, considerando dados retrospectivos para estimar parâmetros desconhecidos, superando testes anteriores na área.

1.6.16 Statistical issues in the prospective monitoring of health outcomes across multiple units (2004) - Marshall, C., Best, N., Bottle, A., & Aylin, P.

Gráficos de controle estatístico de processo são usados para monitorar resultados médicos. Há de se ter cuidado em usar métodos industriais da mesma forma em casos

relacionados a saúde. Trazendo os testes e apresentando aplicação a abordagem nos dados anuais de mortalidade para as causas monitoradas nos hospitais da Inglaterra e País de Gales entre 1995 e 2000.

1.6.17 Efficient scalable schemes for monitoring a large number of data streams (2010)- Mei, Y.

O estudo de fluxos no contexto de monitoramento on-line teve como objetivo detectar a ocorrência de um evento no tempo real. Usando soma acumulada local e as estatísticas de cada fluxo individual para mostrar que minimizam os atrasos de detecção para cada combinação possível de fluxos. A análise é feita com as ilustrações das simulações numéricas e com auxílio da heurística.

1.6.18 Statistical Control of Multiple-Stream Processes: A Shewhart Control Chart for Each Stream (2008) - Meneces, N. S., Olivera, S. A., Saccone, C. D., & Tessore, J.

Os autores propõem o uso de gráficos separados para cada fluxo, devido facilidade por ter recursos informáticos. Compararam o desempenho com o desempenho de outros métodos. Trouxeram a discussão sobre diferentes números de fluxos e diferentes graus de correlação entre eles.

1.6.19 Statistical Process Control of Multiple Stream Processes (1995) - Mortell, R.R., Runger, G.C.

A pesquisa traz a discussão de processos com múltiplos fluxos de produtos idênticos, como uma máquina de enchimento de múltiplas cabeças, podem ser monitorados usando muitos métodos diferentes de gráficos de controle. O método mais adequado depende do número de fluxos de produtos e do tamanho esperado das mudanças nos fluxos.

1.6.20 Control Chart for Multiple Stream Processes (1986) - Nelson, L. S.

O estudo propõe um teste de execução além do usado 3-sigma, afetando apenas os fluxos, a ser usado com gráfico de grupo. Assim verificaram se as saídas dos eixos têm a mesma distribuição normal, usando ARLs no controle.

1.6.21 Identifying Useful Differences in a Multiple-Head Machine (1973-2018) - Ott, E. R., & Snee, R. D.

O artigo traz uma análise sobre melhoria do processo off-line, sobre como uma máquina com várias cabeças, com o objetivo de comparar a variabilidade e suas mudanças individuais. Esse trabalho é ímpar quanto método, pois utilizou-se semi-resíduos, resíduos completos, gráficos de dados brutos organizados de maneiras diferentes e ANOVA para separar os efeitos de tempo e de cabeça.

1.6.22 A Combined Adaptive Sample Size and Sampling Interval \bar{X} Control Scheme (1994-2018) - Prabhu, S. S., Montgomery, D. C., & Runger, G. C.

Foram estudadas as ligações de um gráfico \bar{X} com o tamanho da amostra fixo e o intervalo de amostragem variando. As amostras foram coletadas frequentemente quando a

observação atual caia nos limites de controle, sendo que, se a observação atual se aproximar da linha central, a subseção da amostra restante está atrasada como mostra o gráfico ATS6. O ATS6 é menor para o gráfico de intervalo adaptável do que nos regimes combinados considerados.

1.6.23 Rectangular confidence regions for the means of multivariate normal distributions* (1967) - Sidak, Z.

A região de confiança retangular construída é conservadora para qualquer caso de coordenadas dependentes, o que permite calcular a probabilidade de uma única coordenada. O que foi comprovado pelo fato de o retângulo de confiança ser escolhido para variáveis independentes, sendo válido para quaisquer variáveis dependentes.

1.6.24 Multivariate CUSUM Quality-Control Procedures (1985) - Woodall, W. H., & Ncube, M. M.

O estudo discute a prática de usar simultaneamente CUSUM unilaterais e bilaterais, consideraram na análise o procedimento único multivariado. Usaram um método na época novo, o de comparação de gráficos de controle multivariados, mostrando que para casos em que as características da qualidade são variáveis aleatórias normais bivariadas o MCUSUM é melhor que o T^2 .

1.6.25 Statistical monitoring of multi-stage processes based on engineering models (2008) - Xiang, L., & Tsung, F.

Muitos processos de fabricação na verdade consistem em um grande número de estágios. A modelagem de processos de vários estágios, que descreve a qualidade e a ligação entre as etapas. Para trazer novos estudos, relata o Controle Estatístico de Processo (CEP). O desenvolvimento do método na pesquisa, se deu na estimativa de probabilidade máxima. Os resultados mostram a eficiência do método nos gráficos propostos, em sua aplicação de fabricação de capô de automóveis e montagem de peças.

APÊNDICE 2. Tabela de Limite de Superior Controle para os Gráficos de Controle T^2 -Hotelling

<i>p</i>	LSC
0,00135	29,451
0,00270	27,083
0,02000	19,989
0,02250	19,556
0,02500	19,167
0,08000	14,726
0,09000	14,257

APÊNDICE 3 – Artigo com Parte dos Resultados da Tese Publicado

VICENTIN, D. S.; SILVA, B. B.; PICCIRILLO, I.; BUENO, F. B.; OPRIME, P. C. Monitoring process control chart with finite mixture probability distribution: An application in manufacture industry, *International Journal of Quality & Reliability Management.*, v. 35, n. 2, p. 335–353, 2018. Acesso em: 05 de maio de 2018. Disponível em: <https://www.emerald.com/insight/content/doi/10.1108/IJQRM-11-2016-0196/full/html>.