

Victor Grobberio Trancoso

**Detecção e Diagnóstico de Falhas em  
Processos com Múltiplos Modos de Operação**

Vitória

2016

Victor Grobberio Trancoso

# **Deteccção e Diagnóstico de Falhas em Processos com Múltiplos Modos de Operação**

Projeto de Graduação apresentado ao Departamento de Engenharia Elétrica do Centro Tecnológico da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do grau de Bacharel em Engenharia Elétrica.

Universidade Federal do Espírito Santo – UFES

Centro Tecnológico

Departamento de Engenharia Elétrica

Orientador: Prof. Dr. Celso José Munaro

Vitória

2016

Victor Grobberio Trancoso

## **Detecção e Diagnóstico de Falhas em Processos com Múltiplos Modos de Operação**

Projeto de Graduação apresentado ao Departamento de Engenharia Elétrica do Centro Tecnológico da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do grau de Bacharel em Engenharia Elétrica.

Trabalho aprovado. Vitória, 05 de julho de 2016:

---

**Prof. Dr. Celso José Munaro**  
Orientador

---

**Prof. Dr. Thomas Walter Rauber**  
Examinador

---

**Prof. Dr. Patrick Marques Ciarelli**  
Examinador

Vitória  
2016

*Dedico este trabalho a todos bons exemplos que tive.*

# Agradecimentos

A Deus.

O mais sincero agradecimento a minha mãe, que é tão responsável por este trabalho quanto eu e soube ser o melhor exemplo de garra que conheço; aos meus irmãos que sempre souberam ser pacientes e colaborativos em todas as situações, inclusive naquelas mais árduas; ao meu pai pelos puxões de orelha. À Luiza que soube dar apoio moral e suporte mesmo quando o trabalho parecia ser interminável.

Tive a grande oportunidade de trabalhar com pessoas que se comprometeram com a discussão científica séria e imparcial, e que não pouparam esforços para entregar as melhores soluções. Sou muito agradecido ao Professor Celso J. Munaro pela excelente orientação que tive durante este percurso. Espero ter entregado um trabalho a altura do que foi a mim ensinado.

A Esteban Arroyo que além de amigo soube ser um exímio orientador durante o período de intercâmbio. Sou grato por toda a disposição e respeito com que me apresentou esta temática. Tenho uma dívida que espero algum dia quitar.

Ao professor Thomas W. Rauber, gostaria de agradecer por todas as conversas enriquecedoras que, sem dúvida, agregaram muito valor a este trabalho.

Aos colegas do Laboratório de Controle e Instrumentação, em especial Vinícius Belmuds, Vitor Roriz, Oscar Becerra, Alexandre Paes, agradeço as discussões ricas e esclarecedoras que tivemos, mas também todas aquelas sem conteúdo científico algum.

Existem muitas outras pessoas que, de alguma forma, contribuíram para a conclusão deste trabalho. Faltariam linhas para tantos nomes, apesar disto minha gratidão não é menor em nenhum momento.

*"Pedi, e vos será concedido;  
buscai, e encontrareis;  
batei, e a porta será aberta para vós.  
Pois todo o que pede recebe;  
o que busca encontra;  
e a quem bate, se lhe abrirá  
Mateus 7:7-8"*

# Resumo

Muitos trabalhos vêm sendo publicados na área de detecção e diagnóstico de falhas em processos industriais. Isto é catalisado por grandes interesses econômicos, que se refletem através da melhoria de qualidade e diminuição dos prejuízos provocados por estes problemas.

Entretanto, muitas destas abordagens ainda consideram que o sistema é regido por apenas um modo de operação normal, quando, na verdade, processos reais podem operar sob diversos modos a fim de atender requisitos diferentes de produção. Além disso, ao contrário dos algoritmos de detecção de falhas, pouca atenção vem sendo dada às metodologias de diagnóstico, que são fundamentais para a rapidez na tomada de ações corretivas e combate ao problema.

Este trabalho explora ferramentas da estatística multivariada, apresentando uma metodologia para detecção de falhas em processos industriais que assumem diferentes modos de operação. Para identificar a presença de anomalias serão utilizadas a estatística  $T^2$  de Hotelling e Análise de Componentes Principais (PCA). Os modos de operação de novas amostras serão identificados através da distância de Mahalanobis. O diagnóstico das falhas levará em conta as contribuições individuais das variáveis de processo para o valor da estatística. É proposta uma abordagem de treinamento dinâmica e contínua, que permite que novos padrões de operação normal e de falha sejam sempre acrescentados ao modelo de conhecimento. Também será proposta uma abordagem para a caracterização das assinaturas de falha, que buscará ampliar a capacidade de generalização dos padrões para os diversos modos de operação e diferentes intensidades assumidas por estas anomalias.

Durante a fase experimental, os algoritmos de aprendizado e de supervisão serão validados com dados de simulação de um processo industrial complexo (CSTR). Diferentes formas de execução do treinamento e variações na intensidade das falhas serão produzidas, visando verificar a robustez das assinaturas e a capacidade de aprendizado dinâmico da abordagem.

**Palavras-chave:** Detecção e diagnóstico de falhas. Estatística Multivariada. CSTR.

# Abstract

Many works have been published in the area of fault detection and diagnosis in industrial processes. This is catalyzed by large economic interests that can be reached by the improvement of quality and reduction of losses caused by these problems.

However, many of these approaches still consider that the system is regulated by only one normal operation mode, when, in fact, real processes can operate in several modes to meet different production requirements. Moreover, unlike the fault detection algorithms, little attention has been given to diagnostic approaches, which are the key to accelerate the corrective actions and solve the problem.

This work explores tools of multivariate statistics, presenting a methodology for fault detection in industrial processes that assume different operating modes. To identify the presence of anomalies are used Hotelling's  $T^2$  statistic and Principal Component Analysis (PCA). The operation modes will be identified through the Mahalanobis distance. The fault diagnostic takes into account the individual contributions of process variables for the Hotelling  $T^2$  statistic. It is proposed a dynamic and continuous training approach that allows new normal operation modes and fault patterns to be learned and added to the knowledge model. It will also be proposed an approach to characterize the failure signatures, which seek to expand the generalization capacity of the descriptor for several operating modes and different intensities assumed by the anomalies.

During the experimental phase, the learning algorithms and supervision will be validated with data simulation of a complex industrial process (CSTR). Different ways of implementing the training and variations in the intensity of failures will be produced in order to verify the robustness of signatures and dynamic learning capability approach.

**Keywords:** Fault Detection and Diagnosis. Multivariate statistics. CSTR.



# Lista de ilustrações

Figura 1 – Custo anual de manutenção com base no PIB (MILHÕES DE REAIS)	16
Figura 2 – Aplicações de um classificador e sua limitação	21
Figura 3 – As diferentes naturezas assumidas por uma falha.	24
Figura 4 – Metodologia para supervisão e recuperação do processo.	24
Figura 5 – Variações de causa especial e de causa comum	31
Figura 6 – Função de distribuição de probabilidade gaussiana univariada	33
Figura 7 – Representação gráfica dos erros do tipo 1 e do tipo 2	34
Figura 8 – Carta de Shewhart	37
Figura 9 – Elipse de controle da estatística multivariada	38
Figura 10 – Função de densidade de probabilidade gaussiana multivariada	39
Figura 11 – Elipses de controle para significância de 68% e 95%	40
Figura 12 – Torre de destilação de petróleo	49
Figura 13 – Sinais do monitoramento do reator-regenerador da torre de destilação	50
Figura 14 – Regiões de operação normal produzidas por diferentes modos	50
Figura 15 – Variáveis de um processo industrial com ocorrência de falha na 100 <sup>o</sup> amostra	52
Figura 16 – Contribuições originais das variáveis de processo e contribuição binárias	53
Figura 17 – Vetor de assinatura $f$ da falha do exemplo	53
Figura 18 – Fluxograma de treinamento	55
Figura 19 – Fluxograma de supervisão	57
Figura 20 – Modelo do CSTR definido em (Finch (1989)) e (Oyeleye (1989)).	58
Figura 21 – Processo operando normalmente em quatro modos de operação distintos.	61
Figura 22 – <i>Clusters</i> formados pelos quatro modos de operação.	62
Figura 23 – Conjuntos resultantes do treinamento das regiões de operação normal	63
Figura 24 – Comparação entre anomalia observável e não-observável no plano definido pelas primeiras duas componentes principais	64
Figura 25 – Comparação entre as assinaturas das falhas observáveis calculadas para $OP_1$ , $OP_2$ , $OP_3$ e $OP_4$	64
Figura 26 – Comparação entre as assinaturas das falhas encontradas para cada um dos modos de operação.	65
Figura 27 – Detecção e diagnóstico da Falha 6 no modo de operação $OP_4$ simulada para intensidade 50.	67
Figura 28 – Detecção e diagnóstico de Falha 7 no modo de operação $OP_1$ simulada para intensidade 300. Esta anomalia compartilha a mesma assinatura com a Falha 2 e 21.	68

Figura 29 – Ampliação do resultado do diagnóstico da Falha 7 no modo de operação $OP_1$ . É possível observar que o diagnóstico é confundido entre as Falhas 2, 7 e 21. . . . .	68
Figura 30 – Supervisão para a situação onde apenas mudanças nos modos de operação ocorrem. A planta opera em $OP_1$ , $OP_3$ , $OP_4$ e $OP_2$ respectivamente. . . . .	69
Figura 31 – Estrutura de treinamento modificada. Neste caso, as assinaturas são conhecidas somente em $OP_1$ , conforme indicado pelo campo $F_k$ . . . . .	69
Figura 32 – Resultado da supervisão para falhas não treinadas em $OP_2, OP_3$ e $OP_4$ . . . . .	70
Figura 33 – Estrutura de treinamento modificada. Neste caso, as assinaturas são conhecidas somente em $OP_1$ , conforme indicado pelo campo $F_k$ . . . . .	70

# Lista de tabelas

Tabela 1 – Indicadores de disponibilidade dos processos entre 1997 e 2011 . . . . .	16
Tabela 2 – Variáveis observadas de constantes da planta CSTR . . . . .	59
Tabela 3 – Possíveis falhas no CSTR . . . . .	60
Tabela 4 – Atributos necessários para a simulação . . . . .	60
Tabela 5 – Modos de operação projetados . . . . .	60
Tabela 6 – Teste de diagnóstico correto para falhas em diferentes intensidades . .	66

# Lista de abreviaturas e siglas

PCA	Principal Component Analysis
SPC	Statistical Process Control
UCL	Upper Control Limit
LCL	Lower Control Limit
CSTR	Continuous Stirred-Tank Reactor
ABRAMAN	Associação das empresas brasileiras de manutenção
CON	Condição de Operação Normal
ANI	Anomalia Não Identificada
TMPR	Tempo Médio Para Reparo

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>14</b>
1.1	Motivação	14
1.2	Estrutura do Trabalho	17
1.3	Estado da Arte	18
1.4	Objetivos	18
<b>2</b>	<b>DETECÇÃO DE ANOMALIAS</b>	<b>20</b>
2.1	Falhas em Processos Industriais	23
2.2	Detecção de Falhas	25
2.3	Diagnóstico de Falhas	26
2.4	Controle Estatístico de Processos - SPC	27
2.4.1	História	27
2.4.2	Melhoria de Qualidade	28
2.4.3	Ferramentas de SPC	29
<b>3</b>	<b>REVISÃO TEÓRICA</b>	<b>31</b>
3.1	Variações de Causa Comum e de Causa Especial	31
3.2	Análise Estatística Univariada	32
3.2.1	Teste-Z	34
3.2.2	Teste-t	35
3.2.3	Carta de Controle Univariado	36
3.3	Análise Estatística Multivariada	37
3.4	Estatística $T^2$ de Hotelling	41
3.5	Análise de Componentes Principais	42
3.5.1	Algoritmo	42
3.5.2	Aplicação da Análise de Componentes Principais	44
3.6	Agrupamento	45
3.7	Gráficos de contribuição	46
<b>4</b>	<b>O MÉTODO</b>	<b>48</b>
4.1	Descritores das Múltiplas regiões de operação normal	48
4.1.1	A presença de múltiplas regiões	48
4.1.2	Abordagem proposta	49
4.2	Descritores das falhas	51
4.3	Treinamento	53
4.4	Supervisão	56

<b>5</b>	<b>ESTUDO DE CASO</b> . . . . .	<b>58</b>
<b>5.1</b>	<b>Reator Tanque Continuamente Agitado - CSTR</b> . . . . .	<b>58</b>
<b>5.2</b>	<b>Experimento</b> . . . . .	<b>59</b>
5.2.1	Etapa de treinamento . . . . .	61
5.2.2	Etapa de supervisão . . . . .	66
<b>5.3</b>	<b>Conclusões</b> . . . . .	<b>71</b>
<b>5.4</b>	<b>Trabalhos futuros</b> . . . . .	<b>72</b>
	 <b>REFERÊNCIAS</b> . . . . .	 <b>73</b>

# 1 Introdução

## 1.1 Motivação

Incrementar a produção, diminuindo o tempo de parada de planta para reparos, e corrigir falhas no processo antes que a qualidade do produto final seja comprometida são questões de fundamental importância e compõe uma nova fronteira para a automação. Nesse cenário, a segurança e a confiabilidade dos processos industriais são a chave para a competitividade das companhias [Murray-Smith e Johansen \(1997\)](#).

Nas últimas décadas, foi possível observar a rápida escalada das tecnologias de instrumentação e a explosão da capacidade de armazenamento dos dispositivos de memória. Isto permitiu que muitas plantas contem com centenas ou milhares de sensores, que, além de regularem o funcionamento dos processos, também permitem descrever o funcionamento da planta, o que pode ser armazenado nos bancos de dados disponíveis. Isto abre portas para diversas estratégias de supervisão avançadas, que podem ir do monitoramento do processo à predição de falhas. Neste contexto, surgem as perguntas:

Como utilizar dados do histórico de funcionamento para detectar anomalias?  
Como a detecção e o diagnóstico de falhas poderiam impactar na produção?

Falhas possuem um alto custo e grande impacto sobre as atividades, a sua presença na cadeia produtiva diminui tempo útil de produção; aumenta custos de reparos, já que elementos não relacionados diretamente com o problema também podem ser afetados; e representam riscos à saúde dos colaboradores e ao meio ambiente. Em muitos casos, a tomada de uma política de manutenção preventiva é preferível, já que podem ser programadas e diminuem substancialmente a possibilidade de ocorrência de problemas maiores. Porém, com o aumento da produção e a demanda por produtos de qualidade sempre superior, prevenir falhas, tomando como base apenas inferências estatísticas sobre o tempo de vida útil dos elementos, tende a ser dispendioso, dado a crescente complexidade dos equipamentos e, conseqüentemente, dos seus custos. Outra forma de ação contra a falha é a abordagem corretiva. Apesar de mais barata à primeira vista, os prejuízos causados pela manutenção a posteriori podem não se justificar, além disso, no atual sistema de produção, não é mais aceitável que os equipamentos ou sistemas parem de modo não planejado. Isto impõe mudanças profundas nos paradigmas de manutenção, já que as maneiras de gestão anteriormente propostas, diga-se preventiva e corretiva, não colaboram para nenhuma vantagem competitiva das indústrias modernas.

Para alcançar bons níveis de produtividade, é necessário que o processo desenvolva boas taxas de continuidade, requerendo poucas paradas de planta por manutenção ou, o que é o pior dos casos, por ocorrência de falha. Em todo caso, o incremento da margem de lucro deve enfrentar duas pressões antagônicas. De um lado, existe a necessidade de manutenção, a fim de manter a segurança da operação e a operação em si, e do outro, existe a necessidade de máxima produção e diminuição dos custos e frequência dos reparos. Isto fica claro em [Wireman \(2004\)](#), quando descreve os custos de manutenção nos Estados Unidos já em 1979, quando alcançaram cerca de 200 bilhões de dólares, com o crescimento esperado de 10% a 15% ao ano. Neste mesmo estudo, é relatado que cerca de um terço do custo total de manutenção foi gasto desnecessariamente por razões como mau planejamento, custos de horas extra, falta de uso dos sistemas de ordem de serviço, etc. Conforme definido por [Cooke e Paulsen \(1997\)](#), uma boa estratégia de manutenção acarreta em pouquíssimas ações corretivas e na menor quantidade possível de intervenções preventivas. A produtividade de trabalho, que é um importante indicador da eficiência da indústria e se reflete em lucros, pode ser sintetizada como a razão entre o total produzido e horas trabalhadas num período, definido na equação 1.1:

$$\text{Produtividade} = \frac{\text{total produzido}}{\text{total de horas trabalhadas}} \quad (1.1)$$

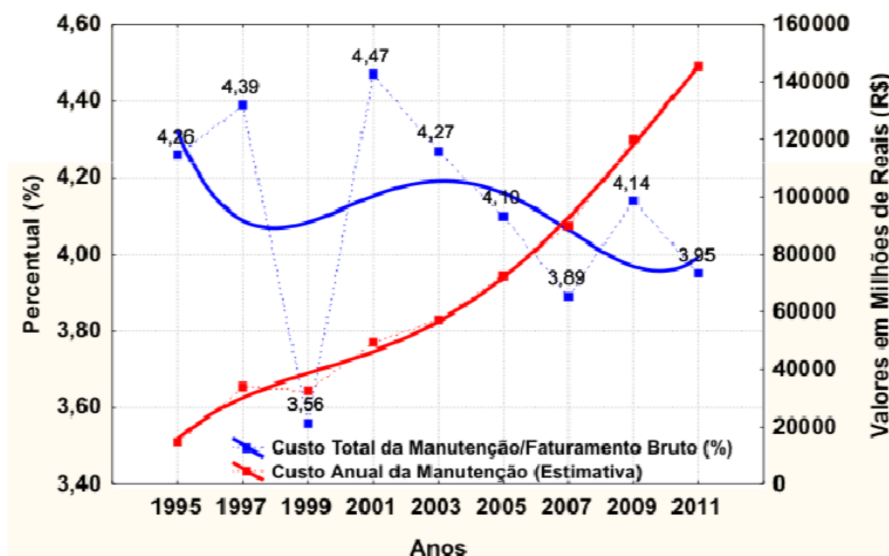
Muitas variáveis podem influenciá-la, porém, em termos gerais, destacam-se os indicadores de disponibilidade e a confiabilidade. Para tornar o sistema mais disponível é preciso que os prazos de parada e o tempo médio para reparo (TMPR) sejam minimizados. A partir de estudos de campo realizados pela Associação das Empresas Brasileiras de Manutenção (ABRAMAN) em 2011 e apresentados no 26º Congresso Brasileiro de Manutenção, Figura 1, observa-se que os custos percentuais de manutenção na indústria seguem uma tendência de queda ano após ano graças à inserção de novas tecnologias e treinamento de pessoal, que auxiliam na gestão de ativos.

Em contraste com esta queda, observa-se na Tabela 1 que a disponibilidade operacional apresenta um padrão de crescimento no últimos anos, embora não tenha sido provocada pela queda de indisponibilidade devido à manutenção, que se manteve aproximadamente constante. Isto pode indicar que os sistemas, apesar de se tornarem mais robustos contra falhas, ainda carecem de formas de identificação e diagnóstico mais precisos sobre a causa raiz; e os equipamentos, por serem mais complexos, demandam maior tempo para o reparo.

Tradicionalmente, melhorias nas operações foram abordadas através da construção de grandes modelos, que permitiam a otimização e o monitoramento dos processos. No entanto, a síntese destes modelos personalizados para cada planta específica exige muito tempo e esforço, que, muitas vezes, são difíceis de se justificar na indústria [Kourti \(2002\)](#). Durante as últimas décadas, métodos alternativos foram propostos para alcançar taxas menores de interferência. Entre aqueles propostos destaca-se a análise dos dados de processo



Figura 1 – Custo anual de manutenção com base no PIB (MILHÕES DE REAIS)



Fonte: ABRAMAN - 26º Congresso brasileiro de manutenção

Tabela 1 – Indicadores de disponibilidade dos processos entre 1997 e 2011

Ano	Indicadores de Disponibilidade (%)							
	1997	1999	2001	2003	2005	2007	2009	2011
Disponibilidade Operacional	85,82	89,3	91,36	89,48	88,2	90,82	90,27	91,3
Indisponibilidade devido à manutenção	4,74	5,63	5,15	5,82	5,8	5,3	5,43	5,44

Fonte: ABRAMAN - 26º Congresso brasileiro de manutenção

por métodos estatísticos multivariados que, em conjunto com o conhecimento sobre a operação dos operadores e engenheiros, pode avaliar a saúde da planta e a qualidade do produto final. Essa ferramenta de supervisão não toma o espaço dos profissionais da área, mas, ao contrário, servem de ferramentas para que a o monitoramento não dependa somente da sua experiência, do seu treinamento ou da sua atenção. Algumas aplicações bem sucedidas na indústria siderúrgica e química podem ser vistas em [Dudzic, Vaculik e Miletic \(2000\)](#) e [Piovoso, Kosanovich e Yuk \(1992\)](#).

Por serem destinadas à indústria, de uma forma geral, estas aplicações devem enfrentar algumas barreiras antes de se tornarem ferramentas indispensáveis para o sistema de produção. Além das dificuldades de natureza tecnológica, o emprego de muitas destas aplicações depende do entusiasmo com que são recebidas na indústria, em especial por aqueles que as usarão. O estudo de [Fuentes et al. \(2006\)](#) aponta que, apesar dos constantes investimentos na qualificação dos profissionais, níveis maiores de disponibilidade do processo

só não são alcançados, porque softwares de processamento de dados tendem a ser muito gerais e levam consigo uma carga muito grande de funções para operações econômicas. Além disto, abordagens muito complexas no campo da matemática exigem uma compreensão mais refinada e esforço para que sejam compreendidas. Por essas razões, muitos supervisórios acabam se tornando menos atraentes e não são contemplados integralmente.

Neste sentido, o desenvolvimento de supervisórios orientados ao usuário final é uma forma interessante de aumentar o poder de penetração destes produtos no mercado. Uma boa aposta para contornar as questões de aceitação do público alvo poderia ser, por exemplo, a melhoria e expansão de ferramentas as quais os operadores já estejam habituados. Uma análise rápida de aplicações pode revelar que abordagens estatísticas são bastante difundidas para o controle de qualidade e/ou de processo. Dentro deste contexto, se valer dos fundamentos estatísticos propiciaria vantagens tanto para resolução dos problemas propostos, quanto para a aceitação dos métodos pela operação.

## 1.2 Estrutura do Trabalho

Este trabalho apresentará uma abordagem para o tratamento de falhas em processos que descrevam múltiplos modos de operação. O seu desenvolvimento resultou em um artigo submetido ao congresso internacional INDUSCON - 2016 sob o título "Robust Fault Detection and Diagnosis for Multimode Processes". Em linhas gerais, este trabalho pode ser dividido em 5 capítulos.

As principais motivações são explicadas no Capítulo 1, dando ênfase ao interesse econômico dos métodos de detecção e diagnóstico de problemas, assim como a sua importância para as indústrias modernas.

O Capítulo 2 definirá as questões ligadas às anomalias, caracterizando-as e explicando importantes conceitos que serão explorados nos Capítulos posteriores. Será introduzido o conceito de falhas em processos industriais e apresentará os principais conceitos sobre métodos de controle estatístico de processo.

O Capítulo 3 fará uma revisão bibliográfica e teórica sobre o principal ferramental utilizado neste trabalho. A estatística  $T^2$  de Hotelling e a Análise de Componentes Principais, principais ferramentas deste trabalho serão explicadas matematicamente observando sua relação com outras técnicas estatísticas. Também será apresentada a ideia de contribuição das variáveis de processo para o rompimento do limiar estatístico.

O Capítulo 4 explica a formulação do *framework* de treinamento e do supervisor, elaborados para a detecção e diagnóstico de falhas. Estas duas metodologias serão detalhadas, explicando a maneira como as regiões de normalidade serão identificadas e descritas assim como as assinaturas de falha serão produzidas.

O Capítulo 5 demonstra a aplicação da abordagem para a supervisão de um processo industrial simulado. Para tanto, o processo complexo de um Tanque Reator Continuamente Agitado (CSTR) será utilizado, e operará sob diferentes modos e sob diferentes ocorrência de anomalias. A capacidade de detecção e diagnóstico será avaliada para falhas em diferentes modos de operação e diferentes intensidades. Além disso, será observado a capacidade da metodologia em identificar o modo de operação sob o qual o processo opera em cada instante.

### 1.3 Estado da Arte

Em [Maestri et al. \(2010\)](#) é apresentada uma metodologia para clusterização não-supervisionada das diferentes regiões de operação produzidas pelos diferentes modos. A clusterização se baseia na distância de Mahalanobis, sendo robusta à *outliers*. Esta abordagem é aplicada somente à detecção de falhas, que é realizada pela estatística  $T^2$  e  $Q$ . A consideração primordial é a constância da estrutura de correlação das variáveis de processo independente do modo de operação.

Em [Monroy et al. \(2012\)](#), uma abordagem de diagnóstico de falhas a partir do transitório é proposto. A detecção de não conformidade do processo é indicada pelas estatísticas  $T^2$  e  $Q$  em conjunto com a análise de componentes principais. As assinaturas de falha são descritas por redes neurais e *Support Vector Machine*. O trabalho é aplicado aos dados de simulação do processo de Tennessee Eastman e produz resultados consistentes, diminuindo as taxas de falsos diagnósticos para o processo. Entretanto, apenas um modo de operação é considerado.

No trabalho [Zheng, Qin e Chai \(2016\)](#), os autores propõem uma metodologia para a detecção de falhas em processos com múltiplos modos de operação através de estruturas latentes. Novamente, apenas a ocorrência das falhas é analisada.

Em [Portnoy et al. \(2016\)](#), uma abordagem recursiva da Análise de Componentes Principais é proposta para resolver o problema de incremento de falsos alarmes devido às mudanças naturais e lentas causadas pelo envelhecimento da planta. Apesar da aparente robustez a estas alterações, mudanças relativamente rápidas provocadas pela alteração do modo de operação representam um problema sério para esta abordagem.

Além destes trabalhos, muitos outros seguem o mesmo padrão, ou atendem à condição de múltiplos pontos de operação ou atendem à problemática do diagnóstico

### 1.4 Objetivos

Este trabalho discutirá a aplicação de métodos de estatística multivariada baseados em variáveis latentes para a detecção de anomalias, os cuidados necessários nos procedi-

mentos e as limitações desta abordagem. Esquemas para o diagnóstico serão apresentados na sequência, de forma que, ao final, será apresentada uma abordagem que contém a integração entre ambas partes na detecção e no diagnóstico.

Um estudo de caso para o modelo simulado de um Reator Tanque Continuamente Agitado (CSTR) levará em conta a capacidade da abordagem de isolamento da falha e diagnóstico. Ao final, deixa-se como legado um algoritmo estruturado e passível de expansão que permitirá a continuidade das pesquisas relacionadas a esta área no Laboratório de Controle e Instrumentação (LCI).

## 2 Detecção de Anomalias

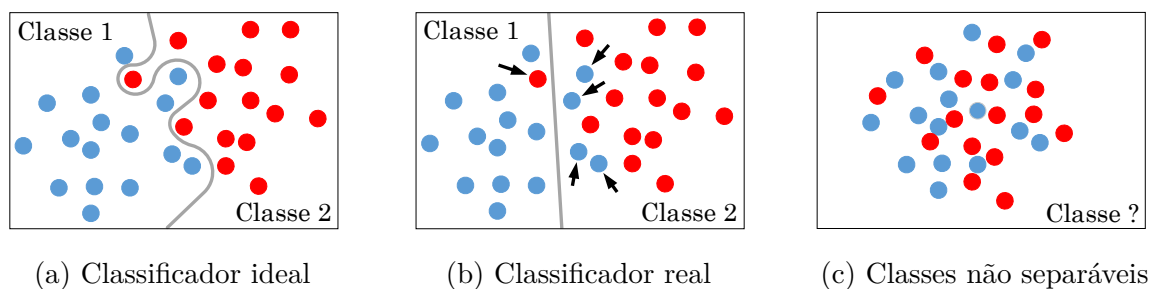
Genericamente, anomalias podem ser entendidas como padrões destoantes de um comportamento considerado normal. A detecção das anormalidades encontra utilização nos mais diversos campos de estudo, tendo várias aplicações no campo da medicina, da economia e militar [Chandola, Banerjee e Kumar \(2009\)](#). Em todo caso, as abordagens de identificação possuem grande importância, pois permitem reduzir os impactos causados pelas anormalidades através de ações corretivas no curto prazo. Exemplos disso podem ser a identificação de tumores por imagens [Spence, Parra e Sajda \(2001\)](#) e o reconhecimento de problemas em rolamentos de motores elétricos [Zarei \(2012\)](#).

Sob uma perspectiva geométrica, a condição de normalidade deve resultar em uma região bem definida no espaço das  $m$  variáveis de processo. Com a ocorrência de evento anormal, as novas amostras tenderiam a se afastar desta região inicial, produzindo outras. Isto implica que anomalias poderiam ser identificadas através da pertinência dos dados a regiões conhecidas no espaço  $m$ -dimensional. Assim, a questão de segmentar este espaço em diversos subespaços relativos a diferentes classes (ou categorias) é destinada à teoria de classificação de padrões. Neste caso, os algoritmos de classificação seriam os responsáveis por definir as fronteiras entre a região de normalidade e os desvios provocados pelas anomalias. Apesar de relativamente simples, alguns fatores tornam esta tarefa em uma atividade desafiadora. Segundo [Chandola, Banerjee e Kumar \(2009\)](#), alguns fatores podem ser entendidos como

- Definir uma região normal que englobe todo comportamento normal possível é complicado. Além disso, definir uma fronteira entre os dados de comportamento normal e anômalo é, geralmente, impreciso. Assim, as observações anômalas perto da fronteira podem ser, na verdade, dados normais e vice-versa;
- Em muitas aplicações, o comportamento normal evolui no tempo e a ideia de normalidade atual pode não ser suficiente para representar os dados no futuro;
- A disponibilidade de dados rotulados para treinamento e validação dos modelos projetados para a detecção de anomalias é, usualmente, o maior problema.

Tendo em vista o papel dos classificadores, um importante parâmetro que define a eficiência da abordagem utilizada é a sua capacidade de separação das diferentes regiões (ou classificação), isto é, quão exato as fronteiras que delimitam as diferentes classes podem ser definidas. Idealmente, os elementos de classes diferentes não deveriam ser confundidos entre si. Assumindo que as amostras na [Figura 2](#) representem toda uma população e desconsiderando questões ligadas ao *overfitting*, a [Figura 2a](#) ilustra o caso ideal. Entretanto,

Figura 2 – Aplicações de um classificador e sua limitação



Fonte: Próprio Autor

no mundo real, fronteiras podem não ser bem definidas e assim o modelo dos classificadores pode não contemplar todas as possíveis realizações. A Figura 2b ilustra este caso, indicando as falsas classificações pelas setas pretas. Note que estas classificações incorretas seriam equivalentes a dizer que uma anomalia seria normal e vice-versa.

Além das dificuldades listadas anteriormente, que impedem a elaboração de um classificador ideal, também devem ser levadas em consideração as questões ligadas à observabilidade da anomalia, como ilustrado pela Figura 2c. A observabilidade é alcançada através do monitoramento das variáveis adequadas, que geralmente estão associadas de alguma forma com o problema. Tome como um exemplo absurdo a averiguação do estado febril de um paciente através de sua estatura física. Naturalmente, por maior que fosse a sua temperatura corporal, nenhuma alteração seria vista na sua altura. Para as abordagens baseadas em dados históricos, duas fases devem preceder o uso dos classificadores: treinamento e teste respectivamente. De uma forma sintética, a etapa de treinamento é aquela responsável por estimar os parâmetros que definirão as fronteiras das regiões de normalidade. Conforme visto anteriormente, é a etapa essencial, requerendo um conjunto de treinamento rico em informação e com variáveis representativas. A eficiência do método é posta em cheque na etapa de teste, quando dados em diferentes situações são testados. É boa prática e necessário que os conjuntos de dados utilizados em ambas etapas sejam diferentes, ao risco que, se realizados com o mesmo conjunto de dados, os resultados produzidos podem não ser representativos para novas amostras.

Apesar da disponibilidade de dados, muitas vezes é difícil encontrar rótulos associados aos conjuntos, que indiquem a classe a qual pertencem. A rotulação, por ser realizada por operadores humanos, acaba sendo muito custosa e não justifica o tempo empreendido. Além disso, catalogar todas as classes possíveis para as condições de normalidade e anômalas é difícil. Ainda que os operadores possuam essa cultura, a rotulação dos dados dependeria da ocorrência dos eventos.

Segundo, [Chandola, Banerjee e Kumar \(2009\)](#), existem diferentes técnicas para o treinamento, que variam de acordo com os diferentes níveis de conhecimento transmitido pelo operador

**Treinamento Supervisionado:** considera a existência de dados rotulados tanto para a normalidade quanto para as anomalias. As regiões representativas para estas condições são definidas e, havendo rótulos também para os dados de teste, a eficiência do classificador pode ser medida. Este tipo de treinamento permite que as anomalias sejam classificadas, o que é desejável em muitas aplicações;

**Treinamento semi-supervisionado:** nesta técnica, leva em consideração os dados rotulados apenas para os casos relativos à normalidade. Assim, este treinamento requer que o operador defina o conceito de operação normal para o sistema, mas não os padrões relativos às anomalias. Em muitos casos, a identificação da anomalia é suficiente para a aplicação e, aliado à relativa facilidade de obtenção de dados de normalidade, fazem deste tipo de treinamento mais frequentes que a técnica supervisionada;

**Treinamento não-supervisionado:** seguindo a lógica, esta técnica de treinamento não leva em conta nenhum rótulo para os dados. Duas hipóteses devem ser consideradas para tanto “dados em situação de normalidade são muito mais frequentes que anômalas” e “dados em normalidade tendem a se agrupar em regiões bem definidas”. Em muitos casos, algoritmos de clusterização são usados para definir os grupos de dados que definem cada região de normalidade presente no conjunto de treinamento. Esta técnica, entretanto, leva a uma taxa maior de erros, já que a maior frequência de dados em normalidade nos conjuntos utilizados para o treinamento não exclui a possibilidade da presença de anomalias. Assim as situações de comportamento indesejado poderiam compor o modelo de normalidade, tornando-o não representativo. Além disso, muitos algoritmos de clusterização demandam que a quantidade de clusters em um conjunto qualquer seja definido a priori, requerendo alguma intervenção do operador.

Em muitos casos, é necessário que uma etapa de extração de características seja aplicada às variáveis, o que reduz muitas vezes a dimensão dos dados originais e, por consequência, torna os classificadores mais robustos. Além disto, existe a possibilidade de que a informação observada "*in-natura*" seja pobre em informação e não representativa para a descrição dos classificadores, requerendo que informação útil seja extraída para o treinamento. A ideia que norteia esta etapa é a construção de um outro espaço descrito por características extraídas dos dados, onde as anomalias sejam mais separáveis da

normalidade, incrementando, conseqüentemente, a taxa de acerto dos classificadores. A principal hipótese neste caso é que existe um sub-espaço qualquer em que as diferentes classes de dados apareçam em regiões razoavelmente distantes. Exemplos de abordagens que utilizam estes extratores para a detecção de anomalias podem ser vistos em [Bin et al. \(2012\)](#).

Existem outras técnicas de redução dimensional que não levam em conta a extração de características, mas exploram as próprias estruturas de correlação dos dados, retirando informação não evidente. As variáveis produzidas são conhecidas como variáveis latentes, onde o termo “latente” faz referência a esta não-evidência da informação coletada. Como efeito, as variáveis latentes possuem maior densidade de informação e podem descrever efeitos não observáveis diretamente nas variáveis observadas. Abordagens como a Análise de Componentes Principais (PCA) e a Redução por Mínimos Quadrados Parciais (PLS) se baseiam nestas propriedades, sendo bastante difundidas na análise multivariada de dados.

## 2.1 Falhas em Processos Industriais

O esforço empregado pela indústria durante as últimas décadas, buscando alcançar níveis de qualidade sempre superior, levou a um investimento cada vez maior em sistemas de controle em malha fechada e instrumentação. A evolução da automação industrial, neste período, deixou uma marca interessante nos dados industriais. Neste novo cenário, as malhas de controle atuam mantendo as variáveis do sistema em suas respectivas faixas de operação apesar de oscilações naturais causadas pela mudança de ponto de operação (corrigida por controle servo) ou por distúrbios no sistema (corrigida por controle regulatório). Sob regime estacionário, as variáveis observadas tenderiam a se manter constantes em torno do seu *set-point* o que, conseqüentemente, descreveria uma região densa no espaço das variáveis. Qualquer desvio desta região poderia ser visto como algo incomum e passível à investigação. Note que o conceito de anomalia é facilmente estendido à falhas em processos industriais.

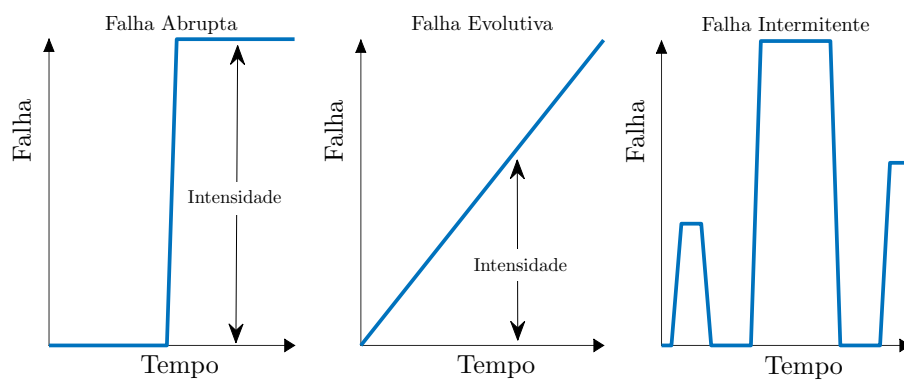
Conforme definido por [Isermann e Ballé \(1997\)](#), falhas podem ser entendidas como desvios não permitidos de pelo menos uma propriedade característica ou parâmetro do sistema em relação a sua condição aceitável, usual ou normal. Estas variações não esperadas podem provocar mudanças significativas na maneira como o sistema se comporta e o controlador, cujo projeto foi otimizado para o processo sem falhas, poderia se tornar ineficaz para a manutenção das variáveis em suas faixas de operação. Como consequência, haveria uma fuga dos dados da região de normalidade, sendo este evento passível de identificação.

De uma forma geral, falhas podem ser classificadas em três categorias: abruptas, incipientes (ou evolutivas) e intermitentes [Isermann \(1997\)](#). As graduações definidas para



a natureza das falhas faz referência à sua intensidade no tempo, como ilustra a Figura 3. Em processos industriais dominados por elementos mecânicos, o tipo de falha mais comum é a evolutiva, tendo em vista o desgaste natural sofrido pelos materiais com o tempo de uso. Devido à dependência inerente entre os subsistemas fabris criada pelo arranjo da cadeia de produção, os desgastes em uma região propagam seus efeitos, podendo ganhar proporção de um subsistema para outro à medida que o equipamento é mais exigido e mais utilizado. É importante pontuar que, no cenário industrial, a presença das derivações é levada em conta e admitida até um nível que não impeça a continuidade da operação ou que não comprometa a qualidade do produto final.

Figura 3 – As diferentes naturezas assumidas por uma falha.

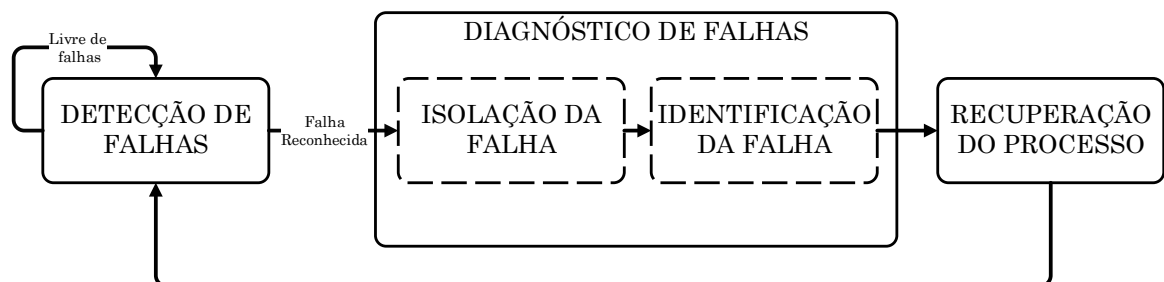


Fonte: Próprio Autor

Falhas abruptas podem ser interpretadas como uma falha evolutiva de período de evolução extremamente curto. Como o nome sugere, inserem uma falha no sistema de um instante para o outro, podem ser entendidas como a quebra ou mau-funcionamento de um sensor. De forma parecida, falhas intermitentes podem ser vistas como uma falha de natureza abrupta e pulsante no sistema. A sua presença não é constante no tempo, mas intercala períodos de funcionamento normal com períodos de mau funcionamento.

Para qualquer natureza de falha, a metodologia de recuperação do processo e

Figura 4 – Metodologia para supervisão e recuperação do processo.



Fonte: Próprio Autor

tratamento da anomalia deve seguir etapas bem definidas como ilustrado na Figura 4. O papel de cada uma das etapas e suas respectivas nomenclaturas seguirá as definições propostas em [Isermann e Ballé \(1997\)](#).

**Detecção da Falha:** identificação da ocorrência ou não de alguma falha. Por permitir que uma ação rápida seja tomada assim que uma condição anômala é reconhecida, é a mais elementar das etapas da recuperação do processo;

**Isolação da Falha:** determinação do tipo, local e instante da detecção da falha. Procede a detecção;

**Identificação da Falha:** determinação do tamanho e do comportamento temporal da falha. Procede a isolação;

**Diagnóstico da Falha:** determinação do tipo, tamanho, local e instante de ocorrência da falha, o que engloba a Isolação e a Identificação. Esta etapa é responsável por determinar a causa raiz da falha. Se corretamente identificada, ações podem ser tomadas para corrigir de uma só vez o problema.

## 2.2 Detecção de Falhas

A detecção de falhas por si só é uma ferramenta de enorme interesse para muitas indústrias. Ainda que nenhum diagnóstico seja realizado a priori, a rápida identificação da ocorrência de falhas ajuda na tomada de atitudes rápidas, abrindo as portas para planos de contingência. Dessa forma, custos de produção poderiam ser preservados, o sistema se tornaria mais seguro e haveria ganhos importantes de qualidade e eficiência do processo.

Em um contexto onde a amostragem e o armazenamento de dados dos processos já não representam mais uma grande barreira tecnológica, a chave para um esquema adequado de monitoramento torna-se a extração de informação útil para modelagem da região de normalidade do processo. Segundo [Chiang, Braatz e Russell \(2001\)](#) o monitoramento de processos pode ser dividido em três vertentes principais: métodos analíticos, métodos baseados no histórico e métodos baseados em conhecimento. Alguns fatores no contexto de aplicação podem direcionar a escolha da vertente mais adequada para o monitoramento. Exemplo disto é a complexidade dos processos industriais modernos, que torna o desenvolvimento de modelos analíticos da planta muito custosos e restritos, já que a modelagem matemática detalhada de sistemas complexos com grandes não-linearidades é um desafio considerável e muitas vezes não justificável. Enquanto isso, abordagens baseadas

em histórico de dados não produzem modelos descritivos dos processos, mas se valem dos dados coletados para criar modelos sintéticos que refletem o seu funcionamento. Entretanto, para estas abordagens, existe a necessidade de um estágio de treinamento com amostras reais. A limitação neste caso é relativa à natureza dos dados armazenados, que pode impor algumas restrições.

Grandes em tamanho, compostas de variáveis altamente correlacionadas, com muitas medições em falta e baixo teor de informação devido aos baixos índices de sinal-ruído, são algumas das características que a abordagem baseada em histórico deverá levar em conta para cumprir com o seu papel [Kourti \(2002\)](#). Neste contexto, as metodologias de estatística multivariada se destacam, já que, segundo [Dudzic, Vaculik e Miletic \(2000\)](#), esta metodologia fornece um compromisso equilibrado entre a fidelidade adequada, visão de processo, facilidade de desenvolvimento, e excelente robustez na presença de dados ruidosos ou em falta.

## 2.3 Diagnóstico de Falhas

Após identificar a presença de um comportamento anormal no processo, o passo seguinte é a identificação da falha em questão. Entretanto, o diagnóstico é uma tarefa complicada e geralmente requer conhecimentos substanciais sobre o processo, devendo ser tratada por experientes operadores de processo [Kruger e Xie \(2012\)](#). Esta necessidade de conhecimento externo torna, em muitas situações, obrigatório que o aprendizado dos algoritmos de diagnóstico seja supervisionado.

Diferentemente das metodologias de detecção, o aprendizado dos algoritmos de diagnóstico baseados em dados podem modelar a falha através de diferentes trechos do conjunto de dados. Em [He et al. \(2009\)](#), [Kulkarni, Jayaraman e Kulkarni \(2005\)](#) e [Ruiz, Nougues e Puigjaner \(2001\)](#) as falhas são diagnosticadas a partir do novo estado estacionário alcançado pelo sistema após a ocorrência da falha. Os trabalhos apresentados em [Rokach \(2008\)](#), [Su e Yang \(2008\)](#) e [Yélamos et al. \(2009\)](#) utilizam dados aleatórios após a detecção da anomalia. Em [Monroy et al. \(2012\)](#), [Brydon, Cilliers e Willis \(1997\)](#) e [Chiang, Russell e Braatz \(2000\)](#) o classificador utiliza os elementos do transitório provocado pela falha, isto é, os dados imediatamente posteriores ao surgimento da anomalia.

A presença de múltiplos modos de operação torna a análise dos padrões de falha um pouco mais complexa. A interação da anomalia com as variáveis de processo pode assumir diferentes formas e intensidades de acordo com o modo de operação em questão, isto, por sua vez, poderia produzir diferentes assinaturas de uma mesma anomalia variando somente pelo modo de operação de onde se originou. Neste contexto, o desenvolvimento de uma assinatura que aumentasse a capacidade de generalização do diagnóstico propiciaria enormes benefícios, tendo em vista que falhas poderiam ser classificadas sem que fossem

treinadas para todos os modos de operação.

## 2.4 Controle Estatístico de Processos - SPC

### 2.4.1 História

A história dos métodos de controle estatístico de processos não é nova, alguns dos seus fundamentos foram desenvolvidos no início do século XX. Segundo [Kruger e Xie \(2012\)](#), os princípios de SPC foram introduzidos por Walter A. Shewhart no final de seu emprego no Departamento de Engenharia de Inspensão da Western Electric Company entre 1918 e 1924 e de 1925 até a sua aposentadoria em 1956 nos Laboratórios Bell. Shewhart estendeu seu trabalho, levando a aplicação de controle estatístico para processos de metrologia e reafirmou a importância de métricas básicas estipuladas para ciência, indústria e comércio. Outro pioneiro neste assunto é Vilfredo Pareto. Pareto postulou que muitas falhas em sistemas são resultados de poucas causas. Essas contribuições resultaram em duas vertentes diferentes de análise, enquanto o trabalho de Shewhart pode ser visto como a “observação de um sistema”, o trabalho de Pareto serve como uma “análise de causa raiz”, se o sistema opera de forma anormal. Assim, o aspecto de controle desta metodologia atribuído à lei de Pareto indica que a melhoria do processo requer profissionais qualificados que encontrem e consertem os principais problemas, cuja presença é acusada através da observação do sistema, tal como abordado por Shewhart. A partir de 1934 as estratégias de controle apresentadas pelo trabalho de Shewhart encontraram grande sucesso na indústria através de W. Edwards Deming e Raymong T. Birge, que as aplicaram na indústria bélica americana durante a segunda guerra mundial, resultando no aumento de produtividade das fábricas de munições. No pós-guerra, o estudo dos princípios de controle estatístico de processos minguou nos Estados Unidos, sendo introduzido no Japão por Deming em meados de 1950. Deming ensinou estes conceitos a engenheiros, gerentes e estudantes. A qualidade e a confiabilidade dos produtos japoneses são atribuídos em grande parte graças à transferência criteriosa dos princípios de controle estatístico e a introdução dos métodos de Taguchi, apresentadas por Genichi Taguchi, em todos os níveis de produção, incluindo na gestão. Mais tarde, o SPC foi incorporado como a pedra angular em um contexto mais amplo originado na década de 80 e conhecido como Gestão da Qualidade Total.

Em 1901, Karl Pearson descreveu a análise de componentes principais em [Pearson \(1901\)](#). Entretanto, o procedimento geral conhecido e aplicado atualmente tem suas raízes no trabalho pioneiro de Harold Hotelling [Hotelling \(1933\)](#). A área foi bastante estudada durante as décadas de 30 e 40, enquanto a aplicação das ferramentas ficou limitada ao desenvolvimento dos computadores modernos [Jackson \(2005\)](#).

Hoje, graças aos dispositivos de processamento e memória cada vez mais velozes e poderosos, as aplicações destas técnicas tornam-se novamente promissoras e factíveis.

Segundo Vicente Falconi, um dos pioneiros no emprego da Gestão da Qualidade Total no Brasil, em Campos (2009), o futuro da análise de dados de hoje é a linha adotada pela Toyota Hino (2005) adoção de banco de dados estruturados que possam ser usados para a coleta de dados históricos que permitam a solução rápida da maioria dos problemas da empresa pela utilização da Análise Multivariada. Dessa forma, quando é possível utilizar os bancos de dados com recursos de Análise Multivariada da estatística, a solução dos problemas fica quase automatizada.

## 2.4.2 Melhoria de Qualidade

A qualidade e o preço dos produtos têm se tornado importantes fatores de escolha de produtos refletindo em questões importantes para as empresas. É notório que os custos envolvidos para a produção de um bem afeta diretamente seu preço de venda final, estreitando a margem de lucro possível. De forma equivalente, existem requerimentos que são esperados pelo comprador e que devem ser atendidos a fim de que os produtos ganhem mercado. Isto é, a qualidade pode ser vista como o cumprimento dos requerimentos dos bens produzidos pelos compradores.

Segundo Montgomery e Runger (2010), qualidade é determinada pela interação de qualidade de design e qualidade de conformidade. Por Qualidade de Design é entendido diferentes níveis de performance, confiabilidade, operacionalidade, e funções que são resultado deliberado de decisões de engenharia ou gestão. Da mesma forma, por Qualidade de Conformidade é entendido redução sistemática de variabilidade e eliminação de defeitos, até que cada unidade produzida seja idêntica e sem defeitos. Muitas vezes, há um equívoco de interpretação sobre a melhoria de qualidade, que é vista somente como o investimento em processos mais modernos ou produtos mais elaborados. Como consequência encarece os custos de produção e estreita a margem de lucro.

Melhoria de qualidade, entretanto, significa eliminação sistemática de desperdício, que implica tanto em eliminação do retrabalho quanto do mau aproveitamento das matérias primas. É natural supor que, quando combatido o desperdício, surgem processos mais eficientes e baratos. Como consequência, para a mesma capacidade produtiva instalada, há o aumento de produtividade, aumento das margens de lucros do negócio devido à redução dos custos variáveis, e há o atendimento da demanda de mercado dentro dos prazos estipulados. A melhoria de qualidade passa obrigatoriamente pelo controle de processo, que é onde as devidas correções podem ser feitas e garantem vantagens competitivas importantes.

### 2.4.3 Ferramentas de SPC

Da perspectiva de eficiência de produção, não é prático inspecionar a qualidade através do produto finalizado, tendo em vista que os produtos devem ser produzidos corretamente de uma só vez. A vantagem de utilizar técnicas de controle de processo é justamente esta, ao invés de observar a consequência sobre os produtos gerados, as fontes de variabilidade podem ser identificadas e corrigidas. Algumas ferramentas de controle estatístico de processos podem ser listadas abaixo. Muitas delas encontram grande aplicabilidade na indústria, não sendo escopo deste trabalho a descrição detalhada de cada vertente, com exceção das cartas de controle.

- Histograma;
- Diagrama de Pareto;
- Diagrama Causa-efeito;
- Diagrama de concentração de defeito;
- Cartas de controle;
- Diagrama de pontos;
- Planilha de verificação.

É importante ter em mente que, apesar de objetivar a redução total da variabilidade do sistema, as técnicas de controle estatístico não são capazes de reduzi-la a zero. A variabilidade é inerente aos sistemas, podendo somente ser reduzida, o que é alcançado pela aplicação sistemática das ferramentas de controle. Segundo [Montgomery e Runger \(2010\)](#), a carta de controle é a ferramenta mais poderosa entre aquelas que compõem as técnicas de SPC e a sua grande popularidade no meio industrial é justificada por cinco motivos:

- É comprovado que cartas de controle aumentam produtividade: Perdas de tempo e matéria prima com retrabalho são causas primárias de redução de produtividade. Uma vez reduzidas, a produtividade aumenta, os custos caem e a capacidade produtiva é ampliada;
- Cartas de controle são eficientes para prevenir defeitos: as cartas de controle são otimizadas para manter o processo em controle, o que é consistente com a filosofia de “fazer certo de uma só vez”. Se o processo produz mercadoria de qualidade desejada agora, é determinístico dizer que, caso se mantenha estável, continuará produzindo produtos idênticos;

- Cartas de controle previnem ajustes desnecessários no processo: tendo em vista que as cartas conseguem distinguir entre a variabilidade inerente do sistema e aquela decorrente de um evento especial, a ação dos operadores pode ser direcionada para a recalibragem do sistema sempre que haja necessidade. Da mesma forma, pode evitar que ajustes desnecessários degradem a performance;
- Cartas de controle proveem informação de diagnóstico: geralmente, padrões das variáveis na carta de controle contêm informação preciosa para operadores experientes, o que o permite implementar a mudança no processo e melhorar a performance;
- Cartas de controle informam sobre a capacidade do processo: proveem informações sobre o valor de importantes valores do processo e sua estabilidade ao longo do tempo. Assim, permite predizer a capacidade do sistema com base empírica, o que auxilia nos projetos de produtos e processos.

## 3 Revisão teórica

### 3.1 Variações de Causa Comum e de Causa Especial

Segundo [Ogunnaike \(1994\)](#), existem duas naturezas de variações: as de causa comum e as de causa especial. Qualquer variável observada  $x$  é afetada exclusivamente por variações de causa comum se, dado o seu valor desejado  $x_d$ , o resíduo gerado é um ruído branco.

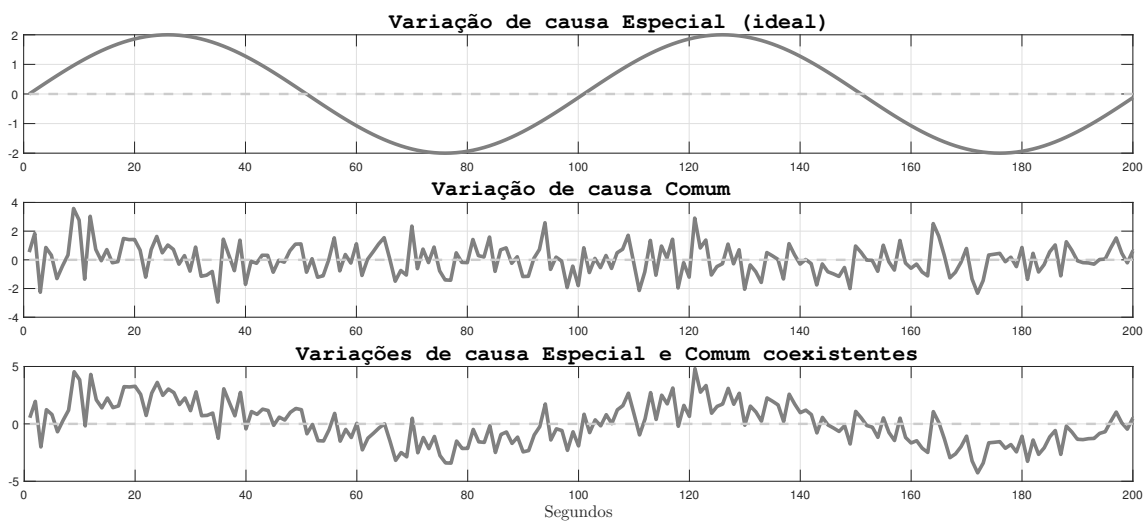
$$x - x_d = e$$

O resíduo  $e$  é, usualmente, de baixa magnitude, afeta todas as variáveis observadas ao mesmo tempo e não pode ser completamente retirada do processo. Se as condições do processo se mantêm inalteradas, as sucessivas amostras  $x$  tendem a se distribuir em torno do valor desejado sendo muitas vezes possível modelar a distribuição como uma curva gaussiana.

Por outro lado, a variabilidade de causa especial é atribuível a qualquer efeito diferente da causa comum, representando muitas vezes algum evento notável como: falhas no processo, mudanças da região de operação, diferença na qualidade das matérias primas, oscilações no sistema, etc.

Para efeito ilustrativo, a Figura 5 retrata as variações das duas diferentes causas, sendo a de causa especial retratada como uma oscilação.

Figura 5 – Variações de causa especial e de causa comum



Fonte: Próprio Autor



## 3.2 Análise Estatística Univariada

Enquanto produtos são continuamente produzidos em um processo e atividades são monitoradas, dados devem ser coletados para avaliar a estabilidade e a capacidade do processo em questão. Assim, se o processo é estacionário, um desvio deve ser provocado somente por variações de causa comum, conforme discutido anteriormente. Apesar de originalmente pensada para o controle de qualidade, a análise estatística pode ser estendida ao monitoramento do processo através das técnicas de controle estatístico de processo (SPC). Isso é válido se as questões ligadas à natureza estocástica das variáveis observadas são respeitadas. Em todo caso, se a variação dos dados se resume à de causa comum, a distribuição das amostras deve seguir um modelo gaussiano com média  $\mu$  e desvio padrão  $\sigma$ , isto é  $x \sim N(\mu, \sigma)$  descrito pela equação 3.1, que retrata a função de densidade de probabilidade gaussiana. De uma forma geral, mesmo que a distribuição das amostras não sigam uma curva gaussiana, o teorema do limite central permite dizer que, para um conjunto de amostras suficientemente grande, a média e o desvio padrão das médias amostrais convergem para a média e o desvio padrão populacionais Hogg e Craig (1970), o que acentua a importância e o poder de generalização deste modelo.

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\left(\frac{x - \mu}{\sigma}\right)^2\right) \quad (3.1)$$

Para o caso em que as propriedades da distribuição populacional fossem conhecidas, os intervalos de confiança poderiam ser calculados diretamente em termos do desvio padrão, segundo uma distância normalizada em função do desvio padrão. Assim, pode-se definir a distância estatística univariada  $z$  como

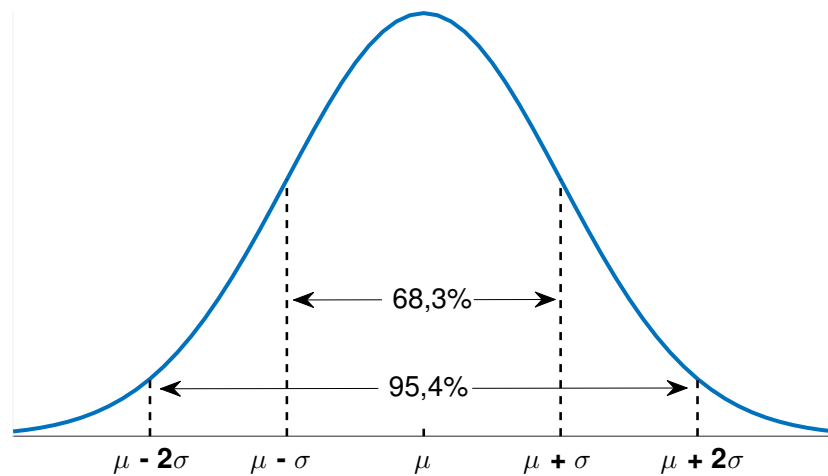
$$z = \frac{x - \mu}{\sigma} \quad (3.2)$$

A integral da função de densidade de probabilidade na Equação 3.1 retorna a probabilidade de ocorrência de observações em diferentes faixas de possíveis valores assumidos. Intuitivamente, é possível notar que, quanto mais distantes da média  $\mu$ , mais raros se tornam as realizações e, de modo análogo, quanto mais perto, mais comuns se tornam as realizações. É possível definir limiares em torno da média para identificar a presença de ocorrências não comuns a partir de um nível de significância  $\alpha$ . Isto quer dizer que a região compreendida pelos limiares devem englobar  $100(1 - \alpha)\%$  das realizações, permitindo que amostras que se projetassem fora da região fossem estudadas. Muitas vezes, as bordas são definidas em termos do desvio padrão e definem alguns valores notáveis para  $\alpha$ .

$$\text{Para } \alpha = 31,7\% : \int_{\mu - \sigma}^{\mu + \sigma} f(x) dx = P\{\mu - \sigma < x < \mu + \sigma\} = 100(1 - 0,317)\% = 68,3\%$$

$$\text{Para } \alpha = 4,6\% : \int_{\mu - 2\sigma}^{\mu + 2\sigma} f(x) dx = P\{\mu - 2\sigma < x < \mu + 2\sigma\} = 100(1 - 0,046)\% = 95,4\%$$

Figura 6 – Função de distribuição de probabilidade gaussiana univariada



Fonte: Próprio Autor

Assim, testes de hipótese estatísticos podem ser utilizados para inferir sobre a normalidade da variável analisada. De uma forma genérica, os testes assumem duas hipóteses que comparam as amostras coletadas com a distribuição conhecida descrita pelo comportamento normal. Qualquer desvio em relação aos parâmetros normais podem, então, ser identificados. A hipótese nula  $H_0$  verifica se a média amostral corresponde à média populacional para o nível de significância  $\alpha$ . Assim:

$$H_0 : \mu = x$$

A hipótese alternativa  $H_1$ , por sua vez, é assumida caso a hipótese nula não seja validada, o que significa que:

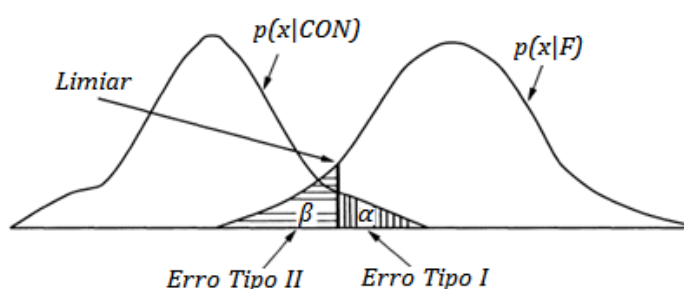
$$H_1 : \mu \neq x$$

Tanto para o controle de qualidade quanto para o controle de processos, a hipótese alternativa  $H_1$  traduz a ocorrência de eventos anormais. É importante observar que, estatisticamente, valores superiores aos limiares são previstos para a variável monitorada mesmo em casos normais. Entretanto, estes acontecimentos são geralmente muito raros, ocorrendo à taxa  $\alpha$ . De toda forma, estas amostras podem conter informações importantes e merecem ser investigada.

Existe uma importante relação de compromisso que envolve a escolha dos níveis de significância. Assuma que  $CON$  represente a situação de normalidade e  $F$  um evento anormal para uma variável  $x$  qualquer. A hipótese nula  $H_0$  traduz  $x \in CON$  enquanto a hipótese alternativa  $H_1$  indica que  $x \in F$ . A Figura 7 ilustra o caso em que alguma ocorrência anormal provoca uma alteração no valor da média populacional, transformando a função de densidade de probabilidade  $p(x|CON)$  em  $p(x|F)$ . É possível ver que a intersecção das curvas é delimitada por um limiar, resultando em duas diferentes áreas

representadas por  $\alpha$  e outra por  $\beta$ . Neste caso, amostras menores (à esquerda) que o limiar serão consideradas como evento normal  $CON$ , enquanto aquelas superiores (à direita) como evento anormal  $F$ . Note que para intervalos de confiança muito largos há a diminuição da taxa de falsos alarmes definida por  $\alpha$ , ao passo que torna o sistema menos sensível às falhas com o crescimento de  $\beta$ . O oposto é da mesma forma válido, diminuindo a região de confiabilidade ( $\alpha$  maior) o sistema identifica qualquer variação decorrente da falha ( $\beta$  menor), com o custo de aumentar significativamente a taxa de falsos alarmes. Uma regra geral é que  $\alpha$  muito pequeno implica em insensibilidade à variações, enquanto  $\alpha$  muito grande implica em altas taxas de falsos alarmes. Na prática, é necessário que o limiar escolhido seja capaz de identificar a ocorrência de eventos anormais, mas que, sobretudo, não ocasione altas taxas de falsos alarmes, isto poderia levar o supervisor ao descrédito.

Figura 7 – Representação gráfica dos erros do tipo 1 e do tipo 2



Fonte: Chiang, Braatz e Russell (2001). Figura adaptada.

onde o Erro do Tipo I significa o falso negativo, isto é, a acusação de falha em uma situação de normalidade e o Erro do Tipo II significa o falso positivo, isto é, a acusação de normalidade em uma situação de anômala.

Conforme visto, a verificação das hipóteses requer o conhecimento das propriedades populacionais, o que é uma limitação em aplicações reais, já que estes valores não são conhecidos na maioria das vezes. A inferência, nestes casos, representa uma importante parte de qualquer análise estatística, pois a obtenção de informações sobre uma população com base em amostras é fundamental para as conclusões assumidas. Técnicas bem difundidas na literatura como o teste-Z e o teste-t serão explicados a seguir. Estas metodologias são capazes de inferir sobre a normalidade dos dados tendo em vista o grau de incerteza inserido pelo não conhecimento dos parâmetros.

### 3.2.1 Teste-Z

Sejam  $x_1, x_2, \dots, x_n$  amostras aleatórias de uma variável cuja distribuição segue uma gaussiana de média  $\mu$  desconhecida e desvio padrão  $\sigma$  conhecido, por Montgomery e Runger (2010) as médias amostrais  $\bar{x}$  devem estar normalmente distribuídas com uma

média  $\mu$  e desvio padrão  $\sigma/\sqrt{n}$ , onde  $n$  é o número de amostras. É possível uniformizar a distribuição pela Equação 3.3 resultando na variável  $Z$ , que segue uma gaussiana definida por  $N(0, 1)$ .

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \quad (3.3)$$

O intervalo de confiança estimado para  $\mu$  é calculado através dos dados amostrados para um nível de significância  $\alpha$ . Assim, se  $Z \sim N(0, 1)$ :

$$P\{-z_{\alpha/2} < \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} < z_{\alpha/2}\} = 1 - \alpha$$

onde  $z_{\alpha/2}$  é o limiar superior que exclui  $100\alpha/2\%$  dos dados da distribuição  $N(0, 1)$ . Por desenvolvimento matemático, os limiares de confiabilidade superior  $UCL$  e inferior  $LCL$  podem ser definidos como:

$$LCL = \bar{X} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

$$UCL = \bar{X} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

Resultando em

$$P\{LCL < \frac{x - \mu}{\sigma/\sqrt{n}} < UCL\} = P\{\bar{X} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} < \frac{x - \mu}{\sigma/\sqrt{n}} < \bar{X} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}\} = 1 - \alpha$$

### 3.2.2 Teste-t

Assumir variância populacional conhecida é, muitas vezes, tão improvável quanto a assumir a média populacional conhecida. Se grandes conjuntos de amostras são realizáveis, isto é ( $n > 40$ ), o teste anterior se mantém válido independente das propriedades populacionais. Isto é justificado pelo teorema do limite central, em que a variância amostral converge para a variância populacional para grandes números de amostras. O teorema permite ainda dizer que a aproximação é válida para qualquer que seja a distribuição das amostras.

Porém, para o caso de poucas amostras e variância populacional desconhecida, é necessário fazer suposições sobre a distribuição dos dados originais de forma que um intervalo de confiança válido seja calculado. Uma consideração razoável em muitos casos é que os dados se distribuem normalmente. Muitas variáveis encontradas na prática se distribuem desta forma permitindo que este modelo encontre grande aplicabilidade mesmo em situações que existam desvios moderados desta distribuição.

Assuma  $x_1, x_2, \dots, x_n$  amostras aleatórias de uma variável cuja distribuição segue uma gaussiana de média  $\mu$  e desvio padrão  $\sigma$  desconhecidos. A média e o desvio padrão amostral são definidos por  $\bar{x}$  e  $S$ , respectivamente. Para definir o intervalo de confiança de  $\mu$ , pode-se substituir  $\sigma$  na equação 3.3 pelo valor conhecido  $S$ . Esta operação transforma

a variável  $Z$  de distribuição  $N(0, 1)$  na variável  $T$ . Os efeitos causados por esta troca, que pode ser interpretado como uma forma de aproximação, foram primeiramente estudados por William Sealy Gosset, que descreveu a nova distribuição de  $T$  como a distribuição t [Gosset \(1904\)](#). De fato, a distribuição t representa uma família de curvas que variam em função da quantidade de amostras  $n$  coletadas, convergindo para uma distribuição normal quando  $n \rightarrow \infty$ . Para um nível de significância  $\alpha$ , os limiares de confiabilidade superior  $UCL$  e inferior  $LCL$  podem ser definidos de forma análoga aos limiares anteriores. Seja  $t_{\frac{\alpha}{2}, n-1}$  o limiar superior de que exclui  $100\alpha/2\%$  das amostras de uma distribuição t com  $n - 1$  graus de liberdade, os limites podem ser vistos como:

$$P\{LCL < \frac{x - \mu}{S/\sqrt{n}} < UCL\}$$

Que pode ser descrita por

$$P\{X - t_{\frac{\alpha}{2}, n-1} \cdot \frac{S}{\sqrt{n}} < \mu < X + t_{\frac{\alpha}{2}, n-1} \cdot \frac{S}{\sqrt{n}}\} = 1 - \alpha$$

### 3.2.3 Carta de Controle Univariado

Uma ferramenta muito utilizada ainda hoje é a carta de controle (ou carta de Shewhart em homenagem ao seu criador) ilustrada na Figura 8. De uma forma geral, a proposta de qualquer carta de controle é a identificação da ocorrência de variações de causa especial.

Esta carta simplifica o teste de hipóteses apresentado anteriormente para uma análise gráfica. Nesta análise, os dados são plotados em função do tempo e são limitados por duas linhas horizontais, chamadas limites de controle, ou de confiabilidade, que representam o total de variabilidade admitido para a causa comum. Em situações normais, os dados devem permanecer dentro da região delimitada pelas retas. Eventualmente, algumas amostras podem cair fora desta faixa, o que é estatisticamente permitido e modelado pela significância  $\alpha$ . Para os casos em que os parâmetros populacionais não sejam conhecidos, a média e o desvio padrão amostrais de dados em controle podem ser usados para modelar os limites de confiabilidade.

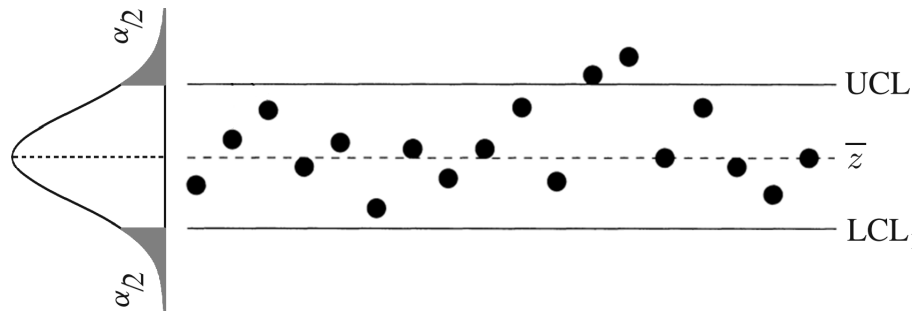
É importante mencionar que a identificação de anormalidades nos dados através da carta de controle não se limita à observação da fuga das amostras da região de confiabilidade. O conjunto de regras definido em [Electric \(1956\)](#) define uma série de verificações para diagnosticar padrões não randômicos nos dados analisados.

- Um ponto excede o limite de controle definido por  $3\sigma$ ;
- Dois ou três pontos consecutivos acima do limite definido por  $2\sigma$ ;
- Quatro ou cinco pontos consecutivos acima do limite definido por  $\sigma$ ;

- Oito pontos consecutivos de um lado da linha central.

De uma forma grosseira, os produtos responsáveis pelo alarme poderiam ser recolhidos, pela consideração de não cumprimento dos requisitos mínimos de qualidade.

Figura 8 – Carta de Shewhart



Fonte: [Chiang, Braatz e Russell \(2001\)](#). Figura adaptada.

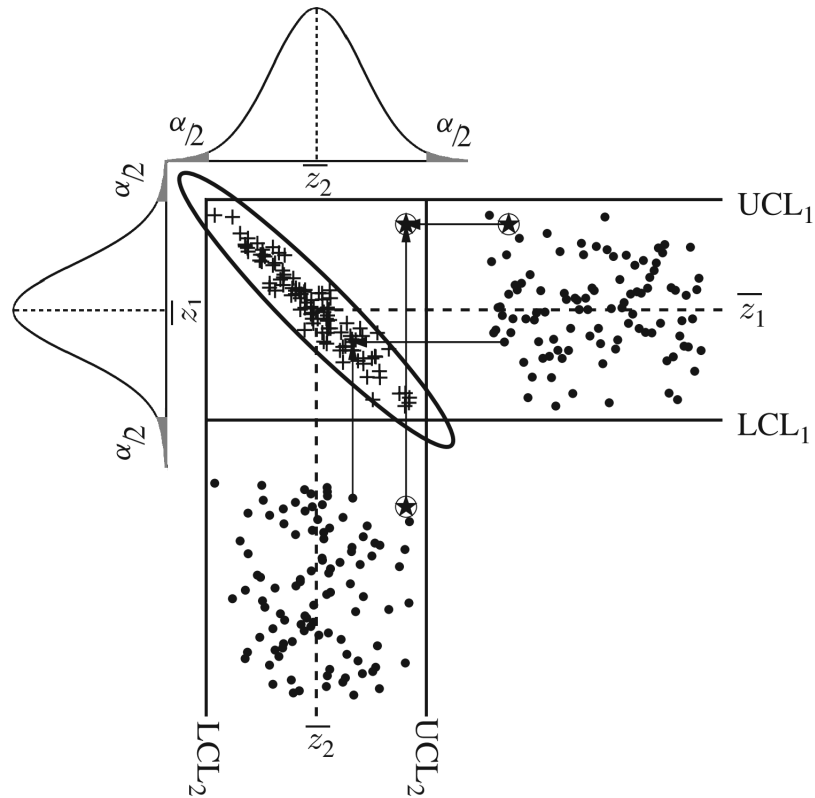
### 3.3 Análise Estatística Multivariada

Apesar da aparente robustez da análise univariada, existe uma questão que merece atenção durante a utilização desta metodologia. Variáveis poucas vezes são descorrelacionadas uma das outras, o que torna um pouco arriscado acreditar que uma propriedade por si só poderá descrever completamente a normalidade. Apesar disso, muitas indústrias ainda utilizam as Cartas de Shewhart para monitorar os produtos e as variáveis-chave do processo [Kourti \(2002\)](#).

Tendo em vista o estado de instrumentação das plantas atuais, observar somente as propriedades do produto final pode ser algo inadequado. Análises de qualidade como massa molecular de polímeros, características de fusão, etc. produzem menos informações em intervalos de tempo muito mais espaçados, sendo coletadas, em grande parte, já no final do processo através de testes laboratoriais. Esta estratégia pode não assegurar vantagens competitivas às indústrias, já que ações corretivas não podem ser tomadas a tempo de solucionar as causas das variabilidades indesejadas, aumentando o desperdício. Por outro lado, o sensoriamento das plantas gera enormes blocos de informação a cada instante de amostragem em diversos estágios da produção, que seriam perdidos caso não fossem levados em consideração por algum sistema de supervisão. Nestas situações, a correlação entre os dados é agravada devido ao encadeamento dos processos, aos sistemas realimentados e à interdependência das malhas de controle. Os efeitos da variação conjunta dos dados sobre as técnicas de controle estatístico univariado devem ser levados em conta para que as regiões de confiança sejam corretamente descritas. Supondo duas variáveis  $z_1$  e  $z_2$  que seguem uma distribuição normal e estão correlacionadas, a Figura 9 apresenta a região

no plano  $z_1$ - $z_2$  descrita pelas variáveis durante a operação em normalidade. As cartas de controle univariada são apresentadas na horizontal para  $z_1$  e na vertical para  $z_2$ , fazendo com que a projeção dos intervalos de confiança univariados  $[UCL_1, LCL_1]$  e  $[UCL_2, LCL_2]$  produzam um retângulo no plano.

Figura 9 – Elipse de controle da estatística multivariada



Fonte: Kruger e Xie (2012). Figura adaptada.

Na situação apresentada, as análises univariadas de ambas variáveis indicam que  $z_1$  e  $z_2$  estão sob normalidade, já que para todas as amostras os limites definidos foram respeitados. Entretanto, é possível observar que os dados não se organizam preenchendo uniformemente a região retangular. Devido à correlação entre as variáveis, a região definida se assemelha a uma elipse. Para as situações limite, onde existe total correlação entre os dados e onde há independência entre as variáveis, os dados formariam uma reta e um círculo, respectivamente.

Para a verificação da normalidade, o monitoramento multivariado acusaria corretamente que a amostra identificada por  $\star$  representa uma possível divergência do padrão de normalidade, que é ilustrado pelo afastamento da amostra da nuvem de pontos. Assim, tendo em vista que o contorno definido pela elipse produz um classificador mais refinado, a análise multivariada seria mais sensível à presença de anomalias para o mesmo nível de significância considerada nas análises univariadas.

A função gaussiana definida anteriormente para o caso univariado pode ser estendida

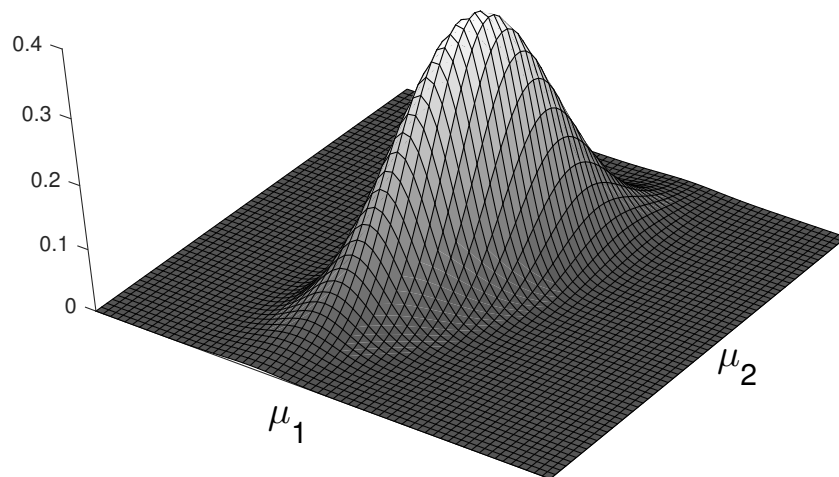
na forma da equação 3.4 para descrever a distribuição conjunta de múltiplas variáveis, que se distribuam normalmente. Neste caso, qualquer que seja a dimensão do espaço analisado, isto é, a quantidade  $m$  de variáveis observadas, a função gaussiana multivariada descreverá um elipsoide no hiperplano.

$$f(x) = \frac{1}{\sqrt{(2\pi)^m |\Sigma|}} \exp\left(-\frac{1}{2}(x - \mu)\Sigma^{-1}(x - \mu)\right) \quad (3.4)$$

onde  $\mu$  é a média populacional de dimensão  $1 \times m$  e  $\Sigma$  é a matriz de covariância das variáveis de dimensão  $m \times m$ . Para o caso bidimensional, a forma assumida pela distribuição é ilustrada pela Figura 10, usando:

$$\mu = (2, 2) \qquad \Sigma = \begin{pmatrix} 0,9 & 0,4 \\ 0,4 & 0,3 \end{pmatrix}$$

Figura 10 – Função de densidade de probabilidade gaussiana multivariada



Fonte: Próprio autor.

A notação da distância estatística univariada definida anteriormente na equação 3.2 pode ser estendida para o caso  $m$ -dimensional. Assumindo que a matriz simétrica  $\Sigma$  é positiva definida, a generalização da distância estatística pode ser sintetizada por

$$d = (x - \mu) \cdot \Sigma^{-1} \cdot (x - \mu)^T \quad (3.5)$$



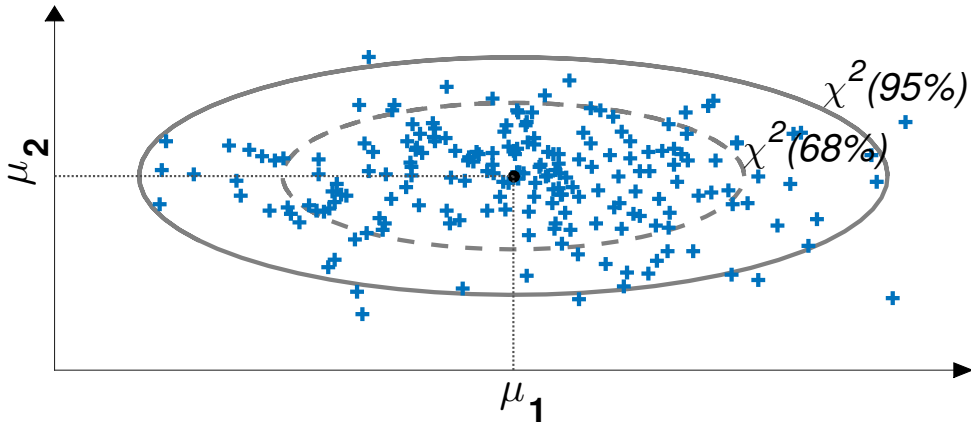
Pela análise da equação 3.5 é possível ver que curvas de contorno dos elipsóides definem contornos equiprováveis, isto é, a densidade de probabilidade é igual para toda amostra  $x_{1 \times m}$  no contorno  $C$ . Isto fica claro quando o caso bidimensional para duas variáveis descorrelacionadas  $\sigma_{12} = \sigma_{21} = 0$  é considerado. Para um contorno  $C$  qualquer, é trivial verificar que o seu formato acompanha uma elipse, conforme pode ser verificado a seguir.

$$(x - \mu) \cdot \Sigma^{-1} \cdot (x - \mu)^T = c^2$$

$$([x_1 - \mu_1, x_2 - \mu_2]) \begin{pmatrix} 1/\sigma_1^2 & 0 \\ 0 & 1/\sigma_2^2 \end{pmatrix} ([x_1 - \mu_1, x_2 - \mu_2])^T = c^2$$

$$\frac{(x_1 - \mu_1)^2}{c^2 \sigma_1^2} + \frac{(x_2 - \mu_2)^2}{c^2 \sigma_2^2} = 1$$

Figura 11 – Elipses de controle para significância de 68% e 95%



Fonte: Próprio autor.

Seja  $\phi$  uma variável aleatória definida por

$$\phi = \sum_{i=1}^m z_i^2 = \sum_{i=1}^m \left( \frac{x_i - \mu_i}{\sigma_i} \right)^2 \tag{3.6}$$

Para  $z_i$  independentes,  $\phi$  deverá seguir uma distribuição Qui-quadrada, isto é,  $\phi \sim \chi_m^2$  para  $m$  graus de liberdade. Segundo Johnson, Wichern et al. (2002), para qualquer que seja o nível de correlação entre as variáveis, o contorno  $C$  pode ser escolhido de forma que traduza a significância estatística representada pela distribuição Qui-quadrada  $\chi_m^2(\alpha)$ , onde  $\chi_m^2(\alpha)$  é o limite superior para exclusão de  $100\alpha\%$  das amostras para uma distribuição com  $m$  graus de liberdade. Dessa forma, é possível definir um contorno que contenha  $100(1 - \alpha)\%$  das amostras, que pode ser usado de forma similar a análise univariada para a definição de um limiar de confiança.

$$(x - \mu) \cdot \Sigma^{-1} \cdot (x - \mu)^T \leq c^2 = \chi_m^2(\alpha)$$

No contexto da análise multivariada, surgem demandas de diferentes campos. De um lado, existe a demanda de refinamento dos modelos e da exploração dos potenciais ganhos oferecidos pela análise das variáveis de processo, onde a baixa relação sinal-ruído dos sinais, a alta correlação e a natureza não-causal devem ser levadas em consideração. Do outro lado, existe a demanda da utilização, que objetiva um sistema de monitoramento simples de usar, de apresentar e de interpretar, tais como as cartas univariadas. A dificuldade com a exposição dos resultados é potencializada quando as variáveis descrevem espaços com centenas de dimensões, tornando impossível o monitoramento visual dos dados conforme apresentado no caso bivariável. Entretanto, visto que um limiar de confiança estatístico pode ser definido e a distância retorna um valor real para qualquer que seja a dimensão  $m$ , cartas de controle multivariada podem ser analisadas da mesma forma que as univariadas, permitindo que as demandas de utilização sejam atendidas sem grandes problemas.

Diversas extensões para as análises multivariadas são encontradas na literatura e apresentam soluções para questões como a redução espacial do espaço original e a inferência estatística quando as propriedades populacionais sejam definidas por informação amostral. A Análise de Componentes Principais (PCA) combinada a abordagens de estatística multivariada como a estatística  $T^2$  de Hotelling e residual ( $Q$ ) são frequentemente encontradas em aplicações reais, pois auxiliam no incremento da robustez dos descritores estatísticos conforme será comentado nas próximas seções.

### 3.4 Estatística $T^2$ de Hotelling

A Estatística  $T^2$  de Hotelling é uma generalização do teste-t que se estende para o caso multidimensional. De forma equivalente ao caso univariado, a consideração prévia sobre a distribuição gaussiana das amostras é necessária.

Em termos de engenharia, o objetivo desta estatística é identificar a presença de anormalidades em amostras futuras a partir de dados coletados do processo em operação estável. Se estas variáveis de processo são estacionárias e independentemente distribuídas, isto é, não há autocorrelação, a distribuição conjunta pode ser vista como  $N_m(\mu, \Sigma)$ . Entretanto, muitas vezes  $\Sigma$  não é conhecido, forçando que a inferência da covariância  $S$  e a média amostral  $\bar{X}$  sejam utilizadas. Segundo [Johnson, Wichern et al. \(2002\)](#), se  $X_1, X_2, \dots, X_n$  compõem uma região de operação normal, a análise estatística para uma observação futura  $X$  é:

$$T^2 = (X - \bar{X})S^{-1}(X - \bar{X})^T \quad (3.7)$$

Assim como no caso da distância estatística, para a análise da estatística  $T^2$  um limiar de confiabilidade pode ser definido. Porém, é importante notar que o uso da distribuição

Qui-quadrada para a definição do limiar requer que a matriz de covariância real das variáveis seja conhecida, mas que, muitas vezes, é estimada pelos dados amostrais. Segundo apontado em [Kruger e Xie \(2012\)](#) e discutido em detalhes em [Anderson \(2003\)](#), a equação 3.6 descreveria uma distribuição  $F$  caso a variância-covariância populacional não fosse conhecida. Assim, o valor do limiar de confiança para a estatística  $T^2$ , conforme definido em [MacGregor e Kourti \(1995\)](#), fica

$$T_\alpha^2 = \frac{(n^2 - 1)m}{n(n - m)} F_{m, n-m} \quad (3.8)$$

onde  $F_{n, n-m}$  é uma distribuição  $F$  com  $m$  e  $n - m$  graus de liberdade. É natural supor que o aumento do número de amostras utilizadas para a estimação dos parâmetros reais torna a covariância estimada mais próxima da populacional. Assim, para  $n \rightarrow \infty$ , 3.8 converge para uma distribuição Qui-quadrada.

$$(\mathbf{X} - \bar{\mathbf{X}}) \cdot S^{-1} \cdot (\mathbf{X} - \bar{\mathbf{X}})^T \leq T_\alpha^2 = \frac{(n^2 - 1)m}{n(n - m)} F_{m, n-m}$$

Para a hipótese nula  $H_0: T^2 \leq T_\alpha^2$ , o que indica que os dados coletados estão dentro do elipsoide definido como modelo de normalidade. Por sua vez, a hipótese alternativa  $H_1$  será aceita sempre que esse  $T^2$  exceder o valor limite definido pela estatística, indicando que as amostras fugiram do padrão de normalidade. A violação do limiar estatístico, portanto, é equivalente a detectar uma falha [Fuente et al. \(2009\)](#).

## 3.5 Análise de Componentes Principais

A análise de componentes principais (PCA) é uma técnica de decorrelação linear e redução dimensional, que preserva a maior parte da informação contida nos dados através de uma projeção linear ortogonal. O coração desta técnica está em definir um conjunto ortogonal de combinações lineares das variáveis originais que reproduzam a estrutura de correlação. Segundo [Chiang, Braatz e Russell \(2001\)](#), as aplicações de PCA são motivadas por um fator ou mais dentre três. A primeira razão é que a representação da informação em dimensões mais compactas generalizam melhor os dados independentes do conjunto de treinamento, havendo um incremento na capacidade de detecção e diagnóstico de falhas. A segunda razão é que a estrutura abstraída pela análise pode ser útil para identificação das variáveis responsáveis pela falha e/ou daquelas mais afetadas por ela. A terceira e última razão é que a PCA pode separar o espaço original dos dados em um subespaço que descreve as tendências sistemáticas do processo, e outro que contém basicamente ruído aleatório.

### 3.5.1 Algoritmo

Supondo uma matriz  $X_{n \times m}$ , cujas linhas representam as amostras e as colunas representam as leituras de cada variável observada, o problema matemático se concentra

em encontrar as direções no espaço  $m$ -dimensional sobre as quais a variância é maximizada. Assim, dado um vetor qualquer  $w_{m \times 1}$  e o  $i$ -ésimo vetor de amostras  $x_i$  da matriz  $X$ . A projeção de  $x_i$  em  $\omega$  é:

$$t_i = x_i \cdot w$$

Deseja-se um vetor  $\omega$  que satisfaça

$$w = \arg \max_{\|w\|=1} \left\{ \sum_{i=1}^n t_i^2 \right\} = \arg \max_{\|w\|=1} \left\{ \sum_{i=1}^n (x_i \cdot w)^2 \right\}$$

A consideração de todas as amostras pode ser estendida pela forma matricial

$$w = \arg \max_{\|w\|=1} \{ \|X \cdot w\|^2 \} = \arg \max_{\|w\|=1} \{ \|w^T X^T X w\| \}$$

Se  $\omega$  é um vetor não-nulo, a equação resultante é representada por

$$w = \arg \max_{\|w\|=1} \left\{ \frac{w^T X^T X w}{w^T w} \right\} \quad (3.9)$$

Conforme pode ser visto em [Chiang, Braatz e Russell \(2001\)](#), uma maneira de resolver a equação 3.9 é através da decomposição em valores singulares (SVD) da matriz  $X$

$$\frac{1}{\sqrt{n-1}} \cdot X = U \Sigma V^T \quad (3.10)$$

onde  $U$  e  $V$  são matrizes unitárias e a matriz diagonal  $\Sigma$  contém os valores singulares reais e não-negativos em ordem decrescente  $[\psi_1 > \psi_2 > \dots > \psi_m]$ . Os vetores de carregamento são as colunas da matriz  $V$  e a variância do conjunto de treinamento projetada na direção da  $k$ -ésima coluna de  $V$  é igual a  $\psi_k^2$ . Outra maneira de resolução de 3.9 é através da decomposição de autovalores da matriz de covariância  $S$ :

$$S = \frac{1}{n-1} \cdot X^T X = V \Lambda V^T \quad (3.11)$$

onde a matriz diagonal  $\Lambda = \Sigma^T \Sigma$  contém os autovalores reais não-negativos em ordem decrescente de magnitude  $[\lambda_1, \lambda_2, \dots, \lambda_m]$ , a relação entre o  $k$ -ésimo autovalor e o  $k$ -ésimo valor singular é dada por  $\lambda_k = \psi_k^2$ . As colunas da matriz  $V$ , neste caso, representam os mesmos vetores descritos anteriormente. Um desenvolvimento matemático mais profundo sobre esta análise pode ser visto em [Jackson \(2005\)](#) e [Kruger e Xie \(2012\)](#).

Note que as colunas da matriz  $V$  representam as bases ortogonais do hiperplano  $m$ -dimensional. Para que a redução espacial proposta seja possível, um subconjunto de colunas da matriz  $V$  deve ser escolhido, descrevendo a base do subespaço  $a$ -dimensional  $a < m$  do espaço original, as  $a$  colunas serão tratadas como componentes principais para futuras discussões. É importante destacar que o processo de escolha do subconjunto

privilegia os autovetores cujos autovalores associados (ou valores singulares) são os maiores. Apesar de haver um critério claro para a escolha das componentes, a decisão sobre a quantidade  $a$  não é trivial. Existem várias abordagens para a escolha da quantidade ótima de componentes principais, validação cruzada [Wold \(1978\)](#), *scree test* [Cattell \(1966\)](#), critério da informação de Akaike [Akaike \(1974\)](#), etc. Este trabalho, entretanto, usará o critério de variância acumulada percentual para a definição da quantidade de componentes mantidas. Assim, seja  $\eta$  o percentual desejado e sabendo que os autovalores traduzem a informação de variância do conjunto de dados,  $a$  pode ser definido na forma

$$\frac{\sum_{i=1}^a \lambda_i}{\sum_{i=1}^m \lambda_i} \leq \eta$$

É importante destacar que este método é sensível à ordem de grandeza das variáveis utilizadas como referência. Se mal dimensionada, algumas variáveis sem importância poderiam dominar a variância na matriz  $X$  por causa das suas unidades de medida, mascarando a variância de variáveis de menor grandeza. No geral, é comum que os dados sejam estandarizados, a fim de assumirem média zero e variância unitária [Zhang \(2000\)](#), conforme indica a equação 3.12. Cuidado deve ser tomado no escalonamento de variáveis quase constantes, já que pequenas variações seriam amplificadas.

$$x' = \frac{x - \mu_{\bar{x}}}{\sigma_{\bar{x}}} \quad (3.12)$$

### 3.5.2 Aplicação da Análise de Componentes Principais

Tendo em vista que a operação normal de um processo pode resultar em um modelo PCA próprio, esta análise pode ser utilizada em conjunto com outras abordagens para inferir sobre a normalidade de um processo. Seja  $P \in R^{m \times a}$  a matriz de componentes principais e  $X \in R^{n \times m}$  a matriz de amostras escalonadas

$$T = XP$$

onde  $T \in R^{n \times a}$  é a matriz de scores, isto é, a projeção das amostras de  $X$  no espaço projetado de menor dimensão. Como PCA é em essência uma metodologia projetiva, os scores  $T$  podem ser reprojetaados para o espaço original  $m$ -dimensional, resultando no espaço de scores definido por

$$\hat{X} = TP^T$$

Note que existe uma perda de informação inerente nestes processos de transformação, que pode ser traduzida no espaço residual  $E$

$$E = X - \hat{X}$$

A análise destes dois espaços é igualmente importante, já que o espaço de scores reflete as tendências sistemáticas do processo e o residual reflete basicamente a estrutura dos ruídos aleatórios. A estatística  $T^2$  pode ser estendida para inferir sobre a normalidade de um processo a partir dos *scores* obtidos pela projeção nas componentes principais. Neste caso, dados de normalidade do processo poderiam ser utilizados para construção de um modelo PCA da região de operação correspondente. Como consequência desta análise, a robustez do supervisor é incrementada, já que taxas menores de falsos alarmes são alcançadas pela desconsideração das componentes de baixos valores singulares [Chiang, Braatz e Russell \(2001\)](#).

Seja  $S$  a matriz de covariância definida pelas variáveis em operação normal, a estatística  $T^2$  é

$$T^2 = \frac{n}{n+1}(X - \bar{X})S^{-1}(X - \bar{X})$$

Por 3.11

$$T^2 = \frac{n}{n+1}(X - \bar{X})V(\Sigma^T\Sigma)^{-1}V^T(X - \bar{X})^T$$

Aplicando as componentes principais e os autovalores associados

$$T^2 = \frac{n}{n+1}(X - \bar{X})P\Sigma_a^{-2}P^T(X - \bar{X})^T$$

onde o limiar estatístico é definido da mesma forma que na equação 3.8 para  $a$  e  $n - a$  graus de liberdade.

## 3.6 Agrupamento

A consideração de múltiplos modos de operação em um processo industrial requer que novas abordagens sejam aplicadas para a verificação do modo atual. Isto permite que as análises estatísticas levem em consideração os parâmetros de uma determinada região descrita pelas variáveis ao invés de uma região generalista.

Em muitos processos, a identificação da região de operação pode ser realizada através da distância entre as novas amostras coletadas e as regiões já conhecidas. Diferentes métricas como a distância Euclidiana e a distância de Mahalanobis podem ser aplicadas para identificar o conjunto mais próximo, podendo ser calculadas tanto no espaço das variáveis originais quanto no espaço das componentes principais. Em contraste com a distância euclidiana, a métrica de Mahalanobis, apresentada em [Mahalanobis \(1936\)](#), leva em consideração a correlação entre os dados, uma vez que utiliza a inversa da matriz de covariância do conjunto de variáveis analisadas. Porém, segundo [Maesschalck, Jouan-Rimbaud e Massart \(2000\)](#), o cálculo desta inversa pode causar alguns problemas. Se as variáveis contêm informação redundante ou correlacionada, a colinearidade desta informação induz à singularidade ou à quase-singularidade da matriz de covariância,

impedindo sua inversão. Outra limitação é que, para o cálculo da matriz de covariância, é necessário um número de amostras superior ao número de variáveis analisadas. Note que, para contornar estas limitações, uma análise de componentes principais pode ser aplicada, de forma que a informação redundante seja excluída e que o número de variáveis seja reduzido. A formulação da distância de Mahalanobis pode ser vista em 3.13. Neste caso  $S$  e  $\bar{x}$  são a matriz de covariância e a média de um conjunto de dados respectivamente, e  $x_i$  é uma amostra qualquer.

$$dm = \sqrt{(x_i - \bar{x}).S^{-1}.(x_i - \bar{x})^T} \quad (3.13)$$

Tendo em vista a ortogonalidade dos *scores* no espaço PCA, a correlação entre as variáveis é nula. Nesta perspectiva, é fácil ver a distância de Mahalanobis como uma distância Euclidiana ponderada com base na variância dos *scores*, ou seja, nos valores singulares. Além disto, apesar da semelhança com a distância estatística definida anteriormente, esta métrica não faz qualquer suposição sobre a distribuição dos dados.

### 3.7 Gráficos de contribuição

O principal objetivo dos gráficos de contribuição é a identificação das variáveis de processo mais relevantes para a ocorrência da falha. Usualmente, em análises univariadas este mapeamento destas variáveis é realizado através da observação daquelas que excedem seus limiares de confiabilidade. Entretanto, para o caso multivariado outra abordagem é requerida. A metodologia apresentada em Chiang, Braatz e Russell (2001) leva em consideração a transformação do espaço original para o espaço PCA. Neste caso, a abordagem baseia-se na quantificação da contribuição de cada variável de processo para os *scores* individuais do espaço PCA. Assim, a contribuição total de cada variável de processo é a soma das parcelas de contribuição para todos aqueles *scores* responsáveis pelo estado de anormalidade, ou seja, aqueles que, de alguma forma, romperam um limiar de confiabilidade. O procedimento proposto a seguir é realizado sempre que houver alguma violação da estatística  $T^2$ :

1. Checar os *scores* normalizados  $(t_i/\psi_i)^2$  para a observação  $x$  e determinar os  $r \leq a$  *scores* responsáveis pelo estado de anormalidade. Para tanto, estes *scores* serão definidos como aqueles que  $(t_i/\psi_i)^2 > \frac{1}{a}(T_\alpha^2)$ . Assuma que  $t_i$  é a projeção da observação na  $i$ -ésima componente principal e  $\psi_i$  é o valor singular correspondente;
2. Calcular a contribuição de cada variável de processo  $v_j$ , que compõe o vetor de amostras  $x = [v_1, v_2, \dots, v_m]$ , para os *scores*  $t_i$  que indicam anormalidade:  $cont_{i,j} = \frac{t_i}{\psi_i} p_{i,j}(v_j - \mu_j)$ . Neste caso,  $p_{i,j}$  é o elemento  $(i, j)$  da matriz de componentes principais  $P$ ;

3. Para qualquer valor em  $cont_{i,j}$  seja negativo, fazê-lo zero;
4. Calcular o total de contribuição da  $j$ -ésima variável de processo  $v_j$ :  $CONT_j = \sum_{i=1}^r(cont_{i,j})$ ;

Tendo em vista que o vetor  $CONT$  contém a contribuição individual das  $m$  variáveis de processo, os dados podem ser normalizados, na forma da equação 3.14. A normalização traduz as informação transmitida pelas contribuições, tornando-as mais intuitivas. Esta simplificação também torna mais fácil a formulação de limiares de significância, que definem a participação das variáveis de processo para o estado de funcionamento anormal. Nestas situações, por exemplo, apenas as variáveis cujas contribuições superassem  $\theta\%$  poderiam ser analisadas pelas equipes técnicas.

$$CONT_{\%} = \frac{CONT}{\sum_{i=1}^m(CONT_i)} \quad (3.14)$$



## 4 O método

A abordagem apresentada se baseia na Análise de Componentes Principais (PCA) e na estatística multivariada  $T^2$  de Hotelling para a detecção de falhas, enquanto o diagnóstico das anomalias é realizado através do estudo das contribuições das variáveis de processo para a estatística. A consideração de múltiplas regiões de operação implica que cada uma será descrita por um modelo de normalidade próprio, estudo semelhante é realizado por [Zhao, Zhang e Xu \(2004\)](#) e também por [Maestri et al. \(2010\)](#), que modela as diferentes regiões de normalidade assumindo a mesma estrutura de correlação entre as variáveis.

A metodologia apresentada pode ser sintetizada por duas etapas fundamentais: Treinamento e Supervisão. A estrutura desenvolvida permite flexibilizar o aprendizado e a dinâmica de supervisão, de forma que um novo modo de operação ou um padrão de falha sejam aprendidos sempre que possível e o monitoramento do processo seja executado online.

### 4.1 Descritores das Múltiplas regiões de operação normal

#### 4.1.1 A presença de múltiplas regiões

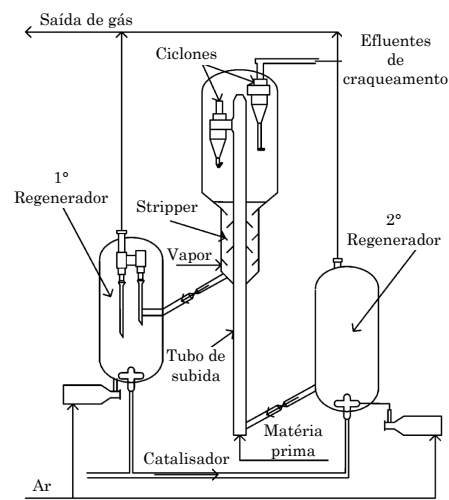
Para atender diferentes requisitos de qualidade, produzir diferentes produtos, etc. os processos industriais operam em diferentes modos de produção, que, conseqüentemente, produzem diferentes regiões no espaço das variáveis.

Apesar do sucesso alcançado pela aplicação da PCA e ferramentas estatísticas multivariadas para a detecção de situações anormais nos processos, quando estas ferramentas são aplicadas a processos com múltiplos modos de operação, muitos alarmes falsos podem surgir ainda que o processo esteja operando em condições estacionárias e normais [Zhao, Zhang e Xu \(2004\)](#). Este não é um efeito aleatório, mas causado pelo fato de que muitas das abordagens atuais são baseadas na hipótese de que o processo tenha apenas uma região de operação nominal, enquanto, na verdade, processos industriais reais possuem vários [Maestri et al. \(2010\)](#). Além disso, ainda que todos os dados fossem considerados como pertencendo a uma única região de operação normal, o volume desta região torna-se incorretamente grande. A ferramenta de monitoramento considerando tal região levaria a um aumento do número de alarmes falsos negativos e falsos positivos [Zhao, Zhang e Xu \(2004\)](#) e [Zhao, Zhang e Xu \(2006\)](#).

Como efeito ilustrativo, a questão dos múltiplas regiões pode ser observada pelo trabalho desenvolvido em [Zhao, Zhang e Xu \(2004\)](#). Dois processos industriais são apre-

sentados no artigo, sendo ambos relativos a processos químicos. O segundo estudo de caso é relativo a uma torre de destilação, de onde 74 variáveis de processo são observadas sendo categorizadas em quatro grupos: reator e regenerador; deposição de carbono; fracionador; e processo de estabilização. Uma discussão mais detalhada sobre o processo é deixado a cargo do artigo original. Apesar disso, vale ressaltar que estes processos desempenham um papel extremamente importante nas refinarias de petróleo, sendo desejado que o estado de produção nominal se preserve. O esquemático da planta é ilustrado pela Figura 12.

Figura 12 – Torre de destilação de petróleo



Fonte: Zhao, Zhang e Xu (2004)

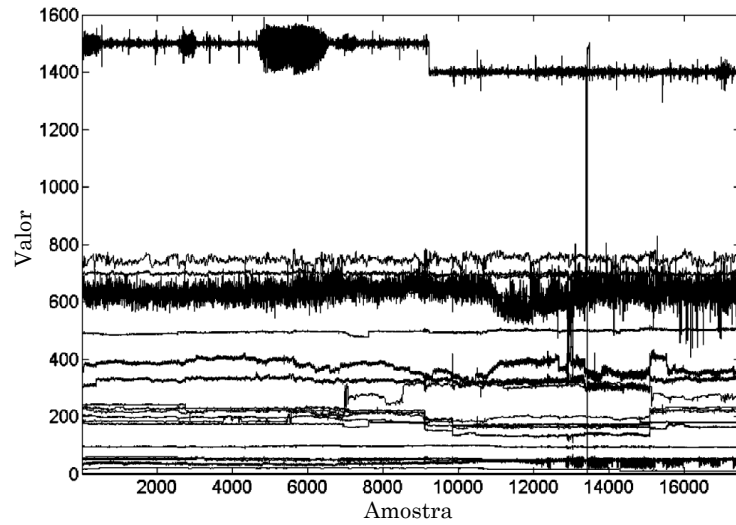
Para simplificação, apenas os sinais referentes ao monitoramento do reator-regenerador são apresentados. Os dados históricos compostos por 17500 amostras de 18 variáveis de processo sob diferentes modos de operação é ilustrado pela Figura 13.

Para tornar o espaço de 18 dimensões visualmente possível, devemos projetá-lo em um plano ou espaço tridimensional. Assim, um modelo PCA pode ser calculado, de forma que a maior parte da informação contida nos dados seja preservada. No caso apresentado, o modelo é definido por todas as 17500 amostras do conjunto de treinamento e serão mantidos apenas os dois primeiros scores  $t_1$  e  $t_2$  decorrentes da projeção nas componentes principais mais significantes. O resultado da análise apresentado pela Figura 14 ilustra a organização dos dados. É possível notar a presença de três regiões de operação bem definidas, que seriam consideradas como uma única grande região caso as ferramentas de monitoramento não fossem capazes de segmentá-la.

#### 4.1.2 Abordagem proposta

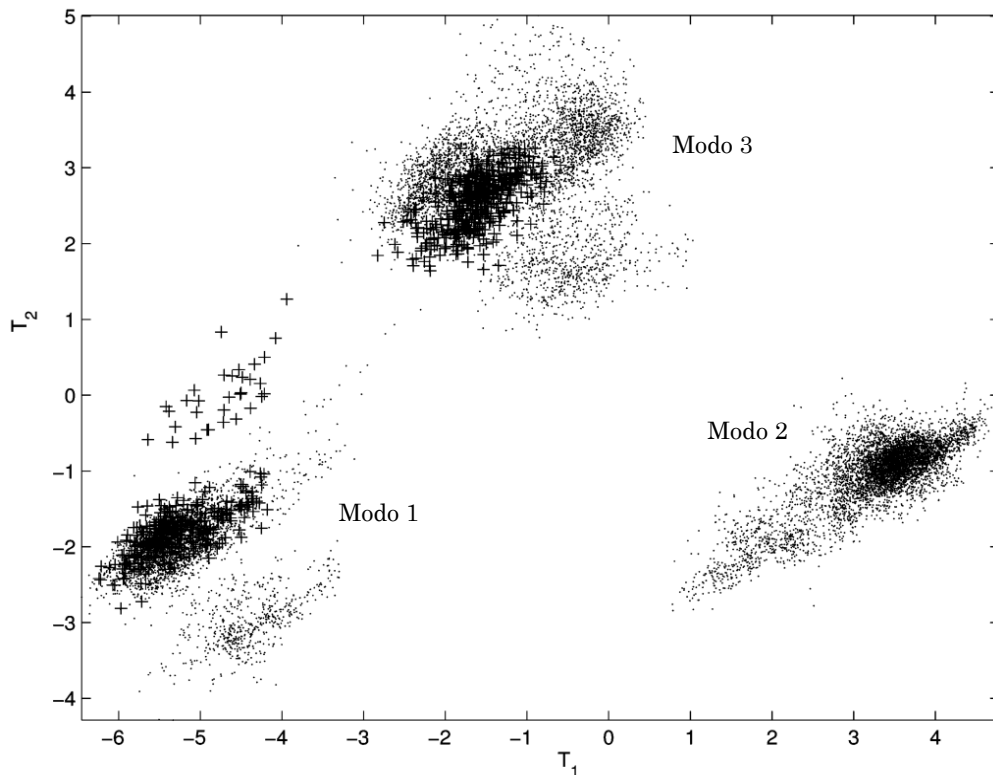
Para caracterizar as múltiplas regiões de operação, é necessário definir quais propriedades são suficientes e necessárias para descrevê-las. Assim, se as variáveis do

Figura 13 – Sinais do monitoramento do reator-regenerador da torre de destilação



Fonte: Zhao, Zhang e Xu (2004)

Figura 14 – Regiões de operação normal produzidas por diferentes modos



Fonte: Zhao, Zhang e Xu (2004)

sistema são consideradas normalmente distribuídas, o conjunto  $C_k$  que define a  $k$ -ésima região de operação normal é descrito pelos elementos

$$C_k = \{\mu_k, \sigma_k, M_k, T_{\alpha,k}^2, d_k, \Sigma_k\}$$

onde  $\mu_k$  é o centróide da região,  $\sigma_k$  é o vetor de desvio padrão das variáveis,  $\Sigma_k$  é a matriz de covariância das variáveis e  $T_{\alpha,k}^2$  é o limiar estatístico calculado através dos dados de treinamento desta região. A matriz de mapeamento  $M_k$  é definida por:

$$M_k = P.\Lambda_a^{-1}.P^T$$

onde  $P$  e  $\Lambda_a$  são as matrizes contendo as  $a$  componentes principais e dos  $a$  autovalores associados respectivamente. A variável real positiva  $d_k^2$  representa a distância quadrática de Mahalanobis máxima que uma amostra pertencente à  $k$ -ésima região de operação normal deve estar do centróide de  $C_k$ . Seja a distância  $d_x$  de uma amostra  $x$  qualquer a  $\mu_k$  é representada por

$$d_x^2 = (x - \mu_k)\Sigma_k^{-1}(x - \mu_k)^T \quad (4.1)$$

para um conjunto de treinamento  $X_{n \times m}$ , onde  $n$  é o número de amostras totais e  $m$  é o número de variáveis observadas, o operador deve definir uma quantidade  $r$  de amostras do conjunto que serão consideradas *outliers*. Para as  $n - r$  amostras restantes, encontra-se uma distância  $d_k$  que seja suficiente para englobar todos os *inliers*. Em outras palavras,  $d_k$  é definida pelo *inlier* mais afastado de  $\mu_k$ .

## 4.2 Descritores das falhas

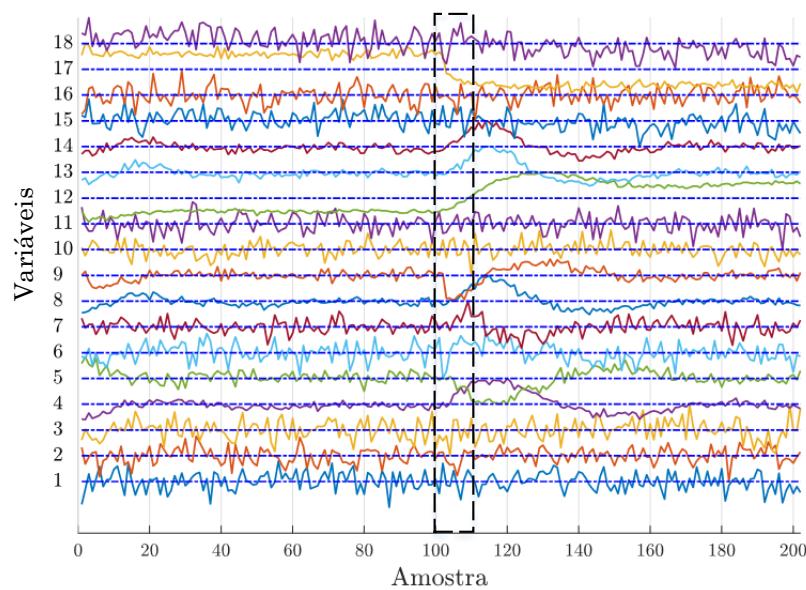
Para descrição das falhas, a contribuição individual das variáveis de processo para a estatística  $T^2$  será utilizada. Estes atributos representam uma importante característica, já que podem ser vistos como a impressão digital da falha, identificando as variáveis do processo que foram mais afetadas e/ou contribuíram para o evento [Kourti \(2002\)](#). Isto permite que a investigação se concentre em subsistemas mais específicos de uma planta grande e complexa. É importante destacar que nenhuma informação sobre a causa raiz da falha é definida pelas contribuições por si só. Entretanto, em abordagens de aprendizado supervisionado, os padrões de falhas projetados nas contribuições poderiam ser armazenados e catalogados pelo operador, permitindo que as ocorrências futuras das mesmas falhas fossem diagnosticadas através destes atributos.

Neste trabalho, o modelo proposto para a assinatura da falha leva em conta a evolução da falha na planta e as diferentes dinâmicas que cada variável do processo assume. A ideia é que os efeitos da propagação da falha pelo sistema ao longo do tempo sejam registrados, evidenciando variáveis que possam se tornar grandes contribuidoras/afetadas por conta da evolução. Conforme discutido anteriormente na [Seção 3.7](#), as contribuições são calculadas a cada instante de amostragem e, por isso, é necessário que este procedimento seja repetido por uma parcela de vezes a fim de caracterizar o comportamento temporal.

A influência de cada variável será calculada a partir do instante em que a falha é detectada, estendendo-se por um período que contenha as  $\Delta$  primeiras amostras posteriores

à falha. Isto indica que as assinaturas serão definidas com base no regime de transitório da falha. Para fins elucidativos, a Figura 15 é apresentada. Nela são ilustrados os sinais de 18 variáveis de um processo industrial, onde há a ocorrência de uma falha na 100ª amostra, que é quando o comportamento anômalo é identificado. O retângulo tracejado indica a janela de amostras retiradas do conjunto a partir das quais as contribuições serão calculadas.

Figura 15 – Variáveis de um processo industrial com ocorrência de falha na 100ª amostra

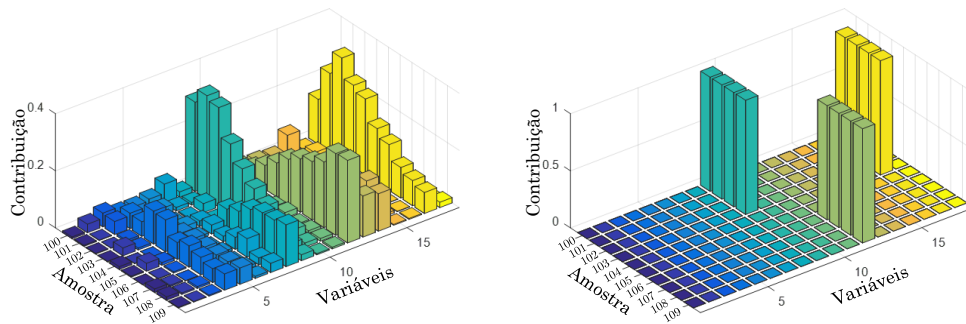


Fonte: Próprio autor.

Transcorrido o período de interesse, o valor da contribuição para cada variável em cada instante calculado, ilustrado na Figura 16a, é comparado a um limiar  $\theta$  a fim de definir a significância das contribuições no tempo. Assim, todos os valores que ultrapassarem  $\theta$  serão assinaladas por "1" e as que se mantiverem inferiores serão assinaladas por "0", como indicado na Figura 16b

Para aumentar o poder de generalização da assinatura, o conjunto de contribuições binárias dos  $\Delta$  instantes analisados pode ser sintetizado para um único vetor  $f_{1 \times m}$ , onde  $m$  é o número de variáveis de processo. Isto será realizado através da soma das  $\Delta$  contribuições de cada variável, que em binário equivale à operação "ou". A assinatura  $f$  pode ser interpretada como um indicador de quais variáveis contribuíram significativamente para a estatística durante o período analisado. A Figura 17 ilustra o resultado de  $f$  para o caso da Figura 16b

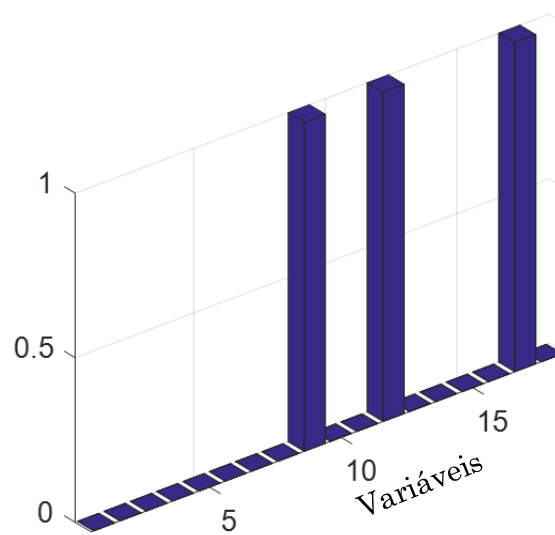
Figura 16 – Contribuições originais das variáveis de processo e contribuição binárias



(a) Contribuição das variáveis

(b) Contribuição binárias ( $\theta = 0, 2$ )

Fonte: Próprio autor.

Figura 17 – Vetor de assinatura  $f$  da falha do exemplo

Fonte: Próprio autor.

### 4.3 Treinamento

O treinamento é um *framework* de aprendizado supervisionado e contínuo que foca contemplar todas as regiões de operação descritas pela planta, associando a cada uma seu respectivo conjunto  $C$ . Além disso, esta estrutura é aplicada para o aprendizado das assinaturas das falhas no sistema. Em todo caso, a matriz  $X_{n,m}$  de dados coletados do processo deve ser acompanhada do rótulo  $\omega$ , que indicará se o conjunto se refere à operação normal (neste caso,  $\omega = 0$ ) ou se representa alguma operação em falha ( $\omega > 0$ ). O rótulo terá papel fundamental para a catalogação da falha, pois representará um índice para cada diferente falha conhecida.

A importância da propriedade de continuidade do aprendizado se deve ao fato de

que muitas das regiões de operação não são conhecidas a priori pelo operador, mas devem entrar para o modelo assim que os dados estiverem disponíveis. Isso garante flexibilidade e com uma gestão apropriada de treinamento, taxas inferiores de falsos alarmes são alcançadas.

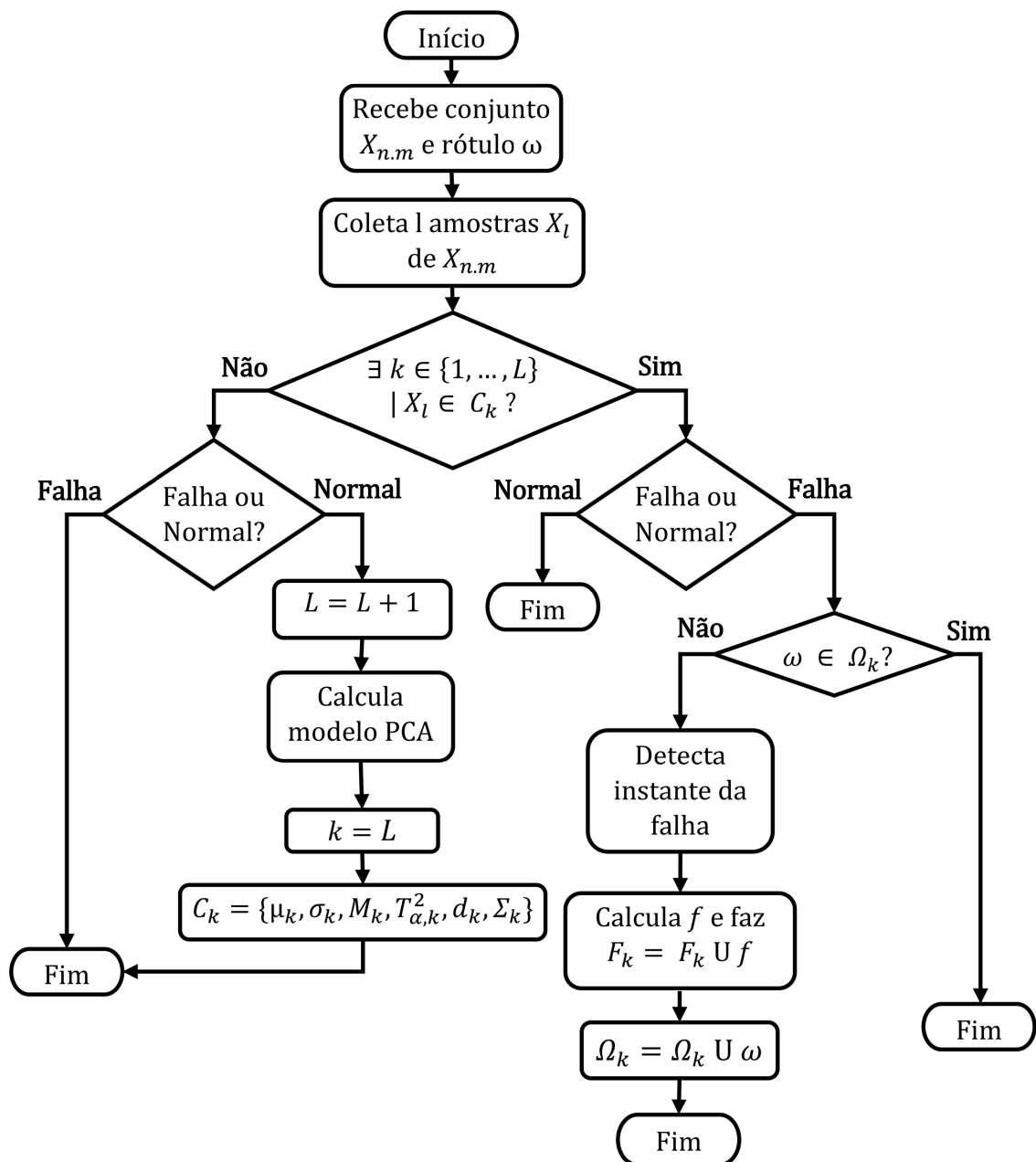
Este trabalho assumirá a hipótese de que uma mesma falha pode se comportar de maneiras diferentes de acordo com o modo de operação de onde se origina. Isso torna necessário que, para cada região de operação normal  $k$ , também seja definido um conjunto  $F_k$ , que contém os vetores de assinatura  $f$  para cada falha conhecida nesta região. Um vetor complementar  $\Omega_k$  informa os índices de falha relativo a cada elemento do conjunto  $F_k$ . Note que, para tornar esta abordagem possível, é fundamental que os conjuntos de treinamento de operações anômalas sejam precedidos de um trecho de operação normal.

A Figura 18 ilustra o fluxograma simplificado da etapa de treinamento. Suponha um conjunto de dados utilizados para o treinamento  $X_{n,m}$  composto de  $n$  amostras. Assumindo que o trecho inicial dos dados sempre representa uma operação normal, inicialmente é verificado se o conjunto sob análise corresponde a alguma região conhecida. Para isso, a pertinência das primeiras  $l$  amostras do conjunto são testadas para as  $L$  regiões de operação normal conhecidas, calculando a distância de Mahalanobis (Equação 3.5) das amostras a cada centróide. Se um determinado percentual destas amostras respeitar a desigualdade  $d_x < d_k$ , a conclusão é que estes dados devem pertencer a  $k$ -ésima região de operação normal. Caso contrário, o conjunto será julgado como relativo a uma nova região.

Em todo caso, o rótulo indicado pelo operador deverá ser analisado. Se  $X$  é dito pertencer a alguma região conhecida e o rótulo indicar uma operação normal ( $\omega = 0$ ), nada será feito, tendo em vista que essa região já foi modelada. Se o rótulo indicar operação anômala para uma falha conhecida em  $k$  ( $\omega \in \Omega_k$ ), nada será feito novamente, pois a assinatura de falha também já é conhecida. Porém, se o rótulo indicar uma nova falha em  $k$  ( $\omega \notin \Omega_k$ ), a assinatura da anomalia deverá ser capturada e armazenada. O processo é sintetizado por três etapas: identificação do início da falha através da estatística  $T^2$ ; cálculo do vetor de assinaturas  $f$  a partir do primeiro instante que  $T^2 > T_{\alpha,k}^2$  conforme a descrição da Seção 4.2; concatenação de  $f$  em  $F_k$  e  $\omega$  em  $\Omega_k$  para registrar o conhecimento.

Por outro lado, se o conjunto  $X$  é dito pertencer a uma região desconhecida, mas o rótulo indicar operação com falha, nada será feito. Isto decorre do fato de que seria impossível associar o padrão da falha a algum modo de operação. Porém, se o rótulo indicar uma operação de normalidade, os dados representarão uma nova região de operação, que será descrita por um modelo PCA próprio e os atributos comentados na Seção 4.1.2.

Figura 18 – Fluxograma de treinamento



Fonte: Próprio autor.



## 4.4 Supervisão

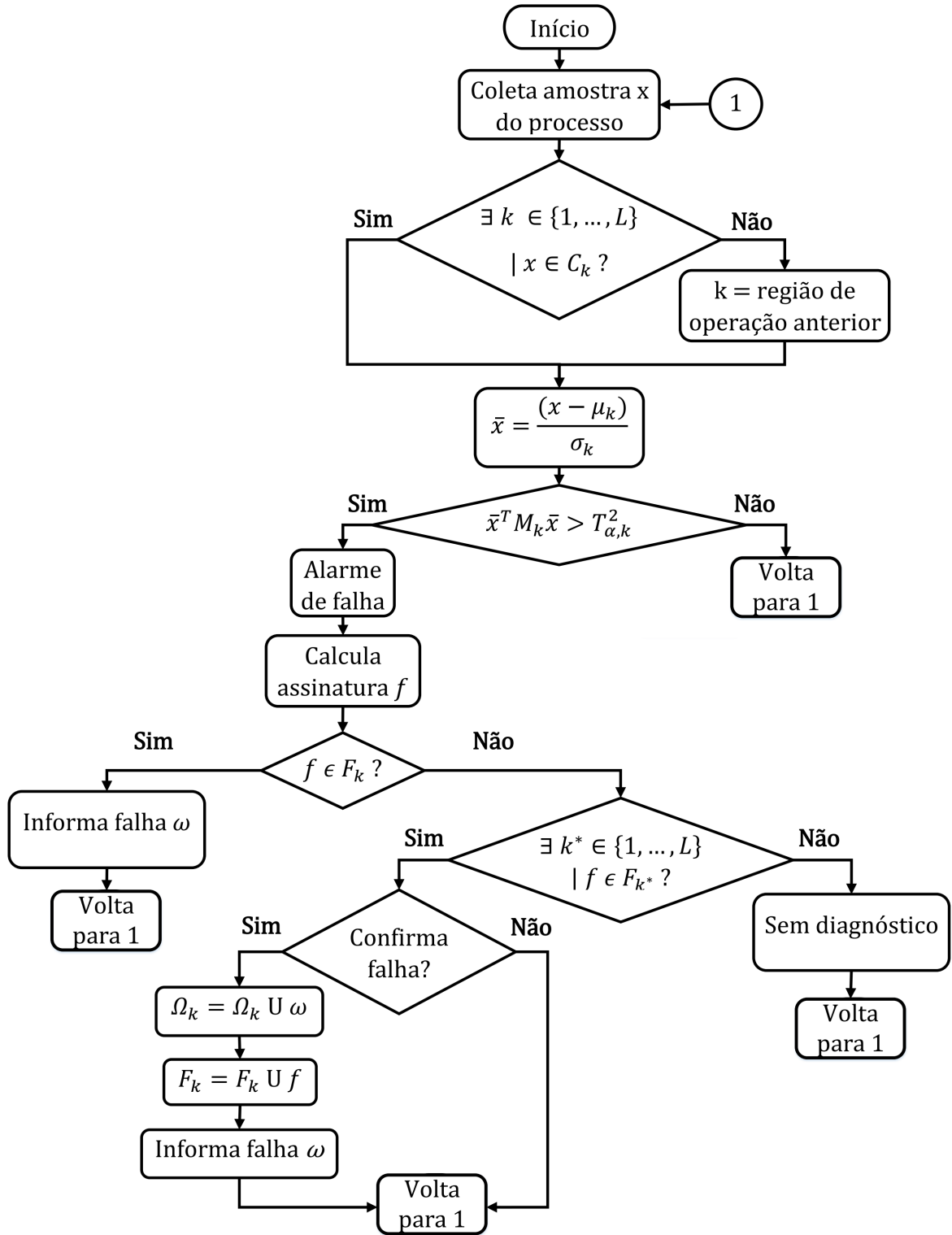
O papel da supervisão é monitorar o funcionamento do processo a partir das amostras coletadas do sistema, inferindo sobre a presença de falhas e diagnosticando-as de acordo com o conhecimento adquirido nas sucessivas etapas de treinamento anteriores.

Sempre que uma nova amostra  $x$  for coletada, a sua pertinência deverá ser verificada para as  $L$  regiões de operação normal conhecidas. Novamente, esta inferência será realizada pela distância de Mahalanobis, onde  $x$  é associado à região  $k$  que satisfaz  $d_x < d_k$ . Note que, caso exista algum desvio do comportamento normal em  $k$ , não seria possível identificar a região de operação correspondente. Assumindo que este desvio tenha sido provocado por uma falha, é necessário que o modelo de normalidade anterior seja mantido, o que permite a continuidade do monitoramento.

A partir dos parâmetros da região de operação normal  $C_k$ , a amostra  $x$  pode ser normalizada pela média e desvio padrão das variáveis de processo conforme apresentado na Equação 3.12. Tendo as variáveis escaladas, a estatística  $T^2$  do modelo PCA da região  $k$  é calculada e comparada ao limiar  $T_{\alpha,k}^2$ . Se o limiar de controle é respeitado, o processo é dito estar em operação normal e retorna para o início, onde espera a coleta da nova amostra do sistema. Entretanto, se o limiar é ultrapassado, o alarme de falha é soado e o vetor  $f$  de características começa a ser calculado. É importante pontuar que, durante os  $\Delta$  instantes de amostragem necessários para sua confecção, nenhum diagnóstico é possível, uma vez que o comportamento da falha ainda não tenha sido completamente descrito. Durante este período, o diagnóstico deverá acusar *ANI* (Anomalia não identificada). Transcorrido o tempo necessário, a assinatura  $f$  é comparada com todas aquelas no conjunto  $F_k$ . Caso, exista algum padrão em  $F_k$  idêntico ao calculado, ou seja  $f \in F_k$ , é diagnosticada a falha  $\omega$  do conjunto  $\Omega_k$  associada a este padrão e a supervisão retorna para o início, esperando uma nova amostra do sistema.

Caso não exista nenhuma coincidência para a região  $k$ , a assinatura  $f$  é comparada às assinaturas de falhas de todas as regiões de operação normal conhecidas. Neste caso existe a expectativa de que a falha descreva o mesmo comportamento mesmo originando em diferentes regiões. Caso  $f \notin F_j$  para qualquer região de operação normal  $j$ , o diagnóstico é *ANI*. Porém, se para uma região qualquer  $k^* \neq k$  a assinatura  $f \in F_{k^*}$ , uma confirmação sobre a falha identificada  $\omega$  é requerida ao operador. Se o operador responsável confirmar que o padrão  $f$  corresponde à falha  $\omega$ ,  $F_k$  é atualizado para conter a nova assinatura e o conjunto  $\Omega_k$  de falhas conhecidas é atualizado para conter  $\omega$ . Feito isto, a supervisão volta para o início e aguarda uma nova amostra.

Figura 19 – Fluxograma de supervisão



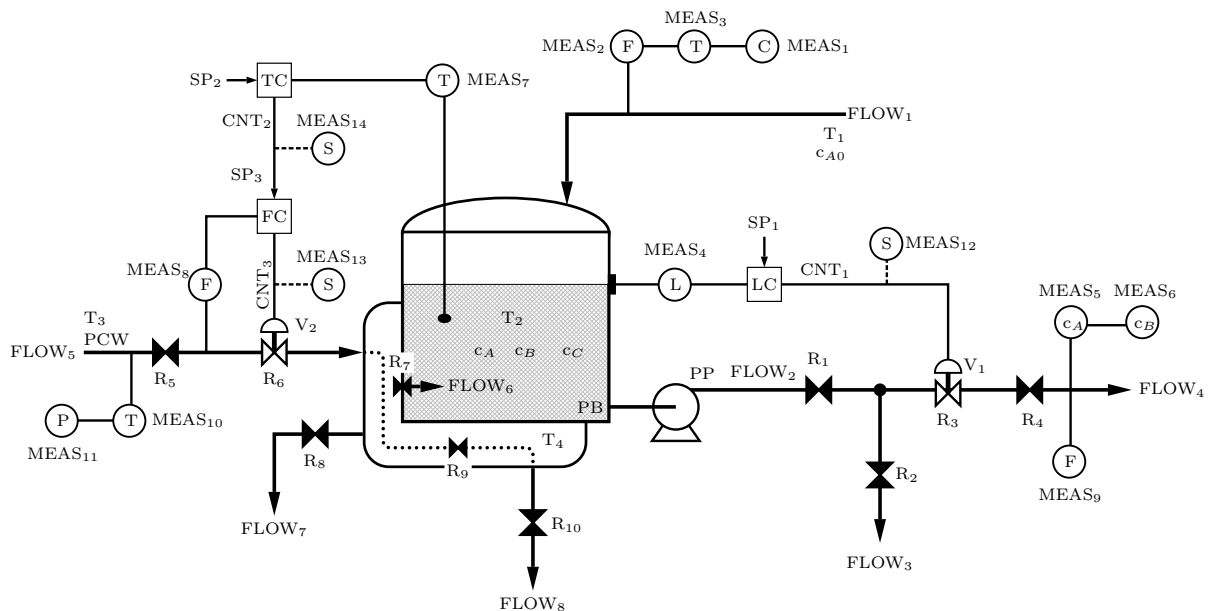
Fonte: Próprio autor.

## 5 Estudo de Caso

### 5.1 Reator Tanque Continuamente Agitado - CSTR

O objeto de estudo utilizado neste trabalho é o simulador de um sofisticado reator químico enjaquetado, onde uma reação exotérmica ocorre, ilustrado na Figura 20. A motivação para utilizar este processo se deve a sua alta não linearidade, o processo possui laços de controle interagindo entre si e existem múltiplos caminhos causais com tendências opostas entre as variáveis. Estas propriedades ficam claras durante a análise de trabalhos anteriores Finch (1989) e Oyeleye (1989). Outro aspecto que torna este processo relevante para a análise é a disponibilidade completa da sua descrição analítica, o que dá ao pesquisador uma compreensão mais profunda sobre dinâmica do processo e as consequências causadas por uma anomalia. A exposição da descrição analítica completa do CSTR e os códigos para o simulador em linguagem C e Fortran podem ser acessados em [sites.google.com/site/cstrsimulator](https://sites.google.com/site/cstrsimulator).

Figura 20 – Modelo do CSTR definido em (Finch (1989)) e (Oyeleye (1989)).



A Tabela 2 informa os nomes reais das variáveis e suas unidades, referenciando os acrônimos apresentados pela Figura 20.

A operação da planta é regida por duas variáveis:  $SP_L$  e  $SP_T$ , relativas ao controle de nível e de temperatura respectivamente. Além de permitir que diferentes modos de operação sejam definidos para o CSTR, o simulador deste processo também possibilita que

35 diferentes falhas sejam inseridas à planta, onde 14 destas falhas são falhas em diferentes sensores. Para cada uma das falhas é possível definir o seu instante de ocorrência e sua intensidade, esta última propriedade será responsável por descrever a rapidez com que a falha evolui no processo.

Todas as possíveis falhas deste processo são apresentadas com seus relativos índices e parâmetros afetados na Tabela 3. Resumidamente, os atributos necessários para inicialização das simulações são descritos pela Tabela 4.

Tabela 2 – Variáveis observadas de constantes da planta CSTR

#	Nome da Variável/Constante	Acrônimo	Unidade
1	Concentração da Alimentação	$c_{A0}$	mol/m <sup>3</sup>
2	Vazão da Alimentação	$F_1$	m <sup>3</sup> /s
3	Temperatura da Alimentação	$T_1$	K
4	Nível do Reator	$L$	m
5	Concentração do Produto $A$	$c_A$	mol/m <sup>3</sup>
6	Concentração do Produto $B$	$c_B$	mol/m <sup>3</sup>
7	Temperatura do Reator	$T_2$	K
8	Vazão do Refrigerante	$F_5$	m <sup>3</sup> /s
9	Vazão do Produto	$F_4$	m <sup>3</sup> /s
10	Temperatura de Entrada do Refr.	$T_3$	K
11	Pressão de Entrada do Refr.	PCW	Pa
12	Saída do Controlador de nível	CNT <sub>1</sub>	-
13	Saída do Controlador de Resfriamento	CNT <sub>2</sub>	-
14	Setpoint do Refrigerante	CNT <sub>3</sub>	-
15	Estoque	$r_1$	-
16	Balanco de Mol	$r_2$	-
17	Perda de Pressão da Água de Refrigeração	$r_3$	-
18	Perda de Pressão do Efluente	$r_4$	-

## 5.2 Experimento

Para a verificação da capacidade de detecção e diagnóstico de falhas da metodologia proposta, o CSTR será simulado para modos de operação distintos, definidos por diferentes valores desejados para o nível e temperatura do reator. Para este estudo, quatro modos de operação normal serão definidos, seguindo os valores adotados para as duas variáveis de processo  $SP_L$  e  $SP_T$  conforme indicado pela Tabela 5.

A inclusão das falhas para o processo de treinamento do algoritmo também seguirá diferentes estratégias. Neste caso, algumas falhas serão treinadas somente para alguns dos modos de operação. Pretende-se com isto, que seja avaliada a capacidade de generalização da assinatura das anomalias e, conseqüentemente, do diagnóstico das falhas sob circunstâncias em que ela é ou não conhecida para o modo em questão. Além disso, é de interesse que

Tabela 3 – Possíveis falhas no CSTR

Índice	Nome da Falha	Parâmetro Afetado
1	Entupimento na saída do tanque	$R_1$
2	Entupimento na jaqueta	$R_9$
3	Vazamento no ambiente da jaqueta	$R_8$
4	Vazamento da jaqueta para o tanque	$R_7$
5	Vazamento na bomba	$R_2$
6	Perda de pressão na bomba	PP
7	Incrustação da área de resfriamento da jaqueta	UA
8	Fonte externa de Calor	$Q_{ext}$
9	Reação primária da energia de ativação	$\beta_1$
10	Reação secundária da energia de ativação	$\beta_2$
11	Fluxo de alimentação anormal	$F_1$
12	Temperatura de alimentação anormal	$T_1$
13	Concentração da alimentação anormal	$c_{A0}$
14	Temperatura anormal da água de resfriamento	$T_3$
15	Pressão anormal da água de resfriamento	PCW
16	Pressão anormal dos efluentes da jaqueta	JEP
17	Pressão anormal dos efluentes do reator	REP
18	Setpoint anormal do controlador de nível	SP <sub>1</sub>
19	Setpoint anormal do controlador de temperatura	SP <sub>2</sub>
20	Agarramento da válvula de controle 1	$V_1$
21	Agarramento da válvula de controle 2	$V_2$
22	Falha(s) nos sensores (variáveis observadas)	'MEAS'

Tabela 4 – Atributos necessários para a simulação

Propriedade do processo	Variáveis
Modo de operação	$SP_L$ , $SP_T$
Falha	<i>início, intensidade, tipo</i>

seja verificada a robustez do diagnóstico em relação às diferentes intensidades que a falha pode assumir.

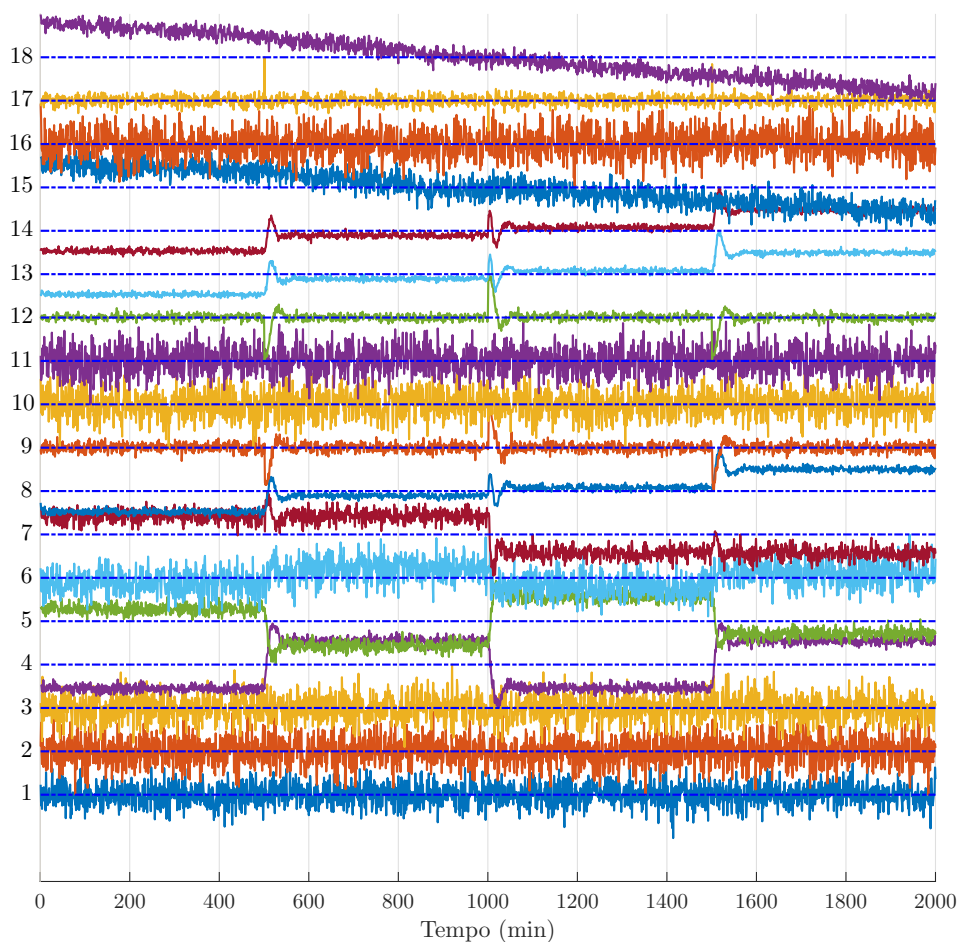
Tabela 5 – Modos de operação projetados

Modo de operação	SP <sub>L</sub>	SP <sub>T</sub>
$OP_1$	2.00	80
$OP_2$	2.15	79
$OP_3$	2.15	80
$OP_4$	2.00	79

### 5.2.1 Etapa de treinamento

A primeira etapa de aprendizado é destinada ao treinamento das regiões de operação normal. Isto permite que a detecção das falhas já seja realizável pelo supervisor. O processo é, então, simulado para quatro modos de operação sem a presença de falhas, de forma que os sinais possam ser então utilizados para o reconhecimento das regiões de normalidade. As variáveis que definem a dinâmica da planta  $SP_L$  e  $SP_T$  são alteradas a cada 500 minutos e seguem os valores listados na Tabela 5. O período de amostragem programado para a planta é de 1 minuto, resultando em 500 amostras para cada um dos modos. As variáveis de processo produzidas pela simulação são apresentadas pela Figura 21.

Figura 21 – Processo operando normalmente em quatro modos de operação distintos.



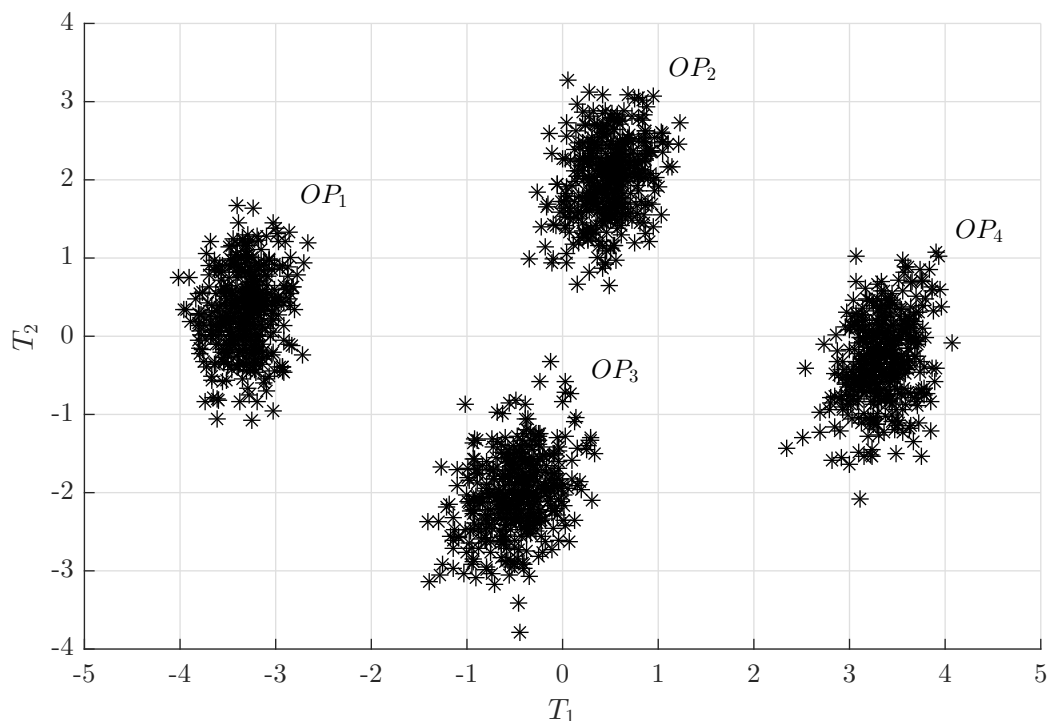
Fonte: Próprio autor.

Para o treinamento das regiões de normalidade ilustradas pela Figura 21, é importante que os trechos relativos aos transitórios sejam descartados, de forma que apenas o regime estacionário seja mantido. Para verificar a separabilidade das amostras, uma estratégia semelhante àquela empregada para produção da Figura 14 pode ser tomada. Neste caso, um modelo PCA para os dados em regime estacionário pode ser criado, de forma que os *clusters* produzidos no espaço 18-dimensional possam ser projetados no

plano definido pelos dois *scores* mais importantes, ou seja, aqueles que acumulam maior variância. É possível observar que as quatro regiões apresentam boa separabilidade e, conforme discutido anteriormente, poderiam aumentar os erros de classificação caso fossem consideradas como apenas uma grande região.

O algoritmo de treinamento é executado para cada um dos trechos relativos a  $OP_1$ ,  $OP_2$ ,  $OP_3$  e  $OP_4$ . Além destes dados, para o cálculo dos parâmetros de cada conjunto, é também necessário definirmos o nível de significância  $\alpha$  e o número de componentes principais que deverão ser mantidas. Para satisfazer a relação de compromisso entre sensibilidade e taxa de falsos alarmes,  $\alpha$  será como 1%. Além disto, o número de componentes mantidas será regulada pelo percentual de variância acumulada pelas componentes principais que, neste caso, será de 90%. Em muitos trabalhos que utilizam esta abordagem estatística para a detecção de anomalias, é comum que um filtro exponencial de média móvel seja aplicado aos dados. Isto permite que a taxa de falsos alarmes seja reduzida, tornando o supervisor mais confiável. A constante deste filtro deve variar entre 0, quando há total filtragem do sinal, e 1, quando o sinal não é filtrado. Para este trabalho, a constante de amortização será definida em 0.7, de forma que ruídos sejam filtrados e pouco atraso seja inserido aos sinais.

Figura 22 – *Clusters* formados pelos quatro modos de operação.



Fonte: Próprio autor.

Para os conjuntos de dados analisados, a análise de componentes principais resulta na manutenção de 13 componentes para  $OP_1, OP_2, OP_3$  e  $OP_4$ . As estruturas produzidas

pelo algoritmo em Matlab podem ser visualizadas na Figura 23.

Figura 23 – Conjuntos resultantes do treinamento das regiões de operação normal

Fields	Centro	t2lim	M	Std	ValSing	PC
1	1x18 do...	28.0199	18x18 ...	1x18 d...	13x13 do...	18x13 ...
2	1x18 do...	28.0199	18x18 ...	1x18 d...	13x13 do...	18x13 ...
3	1x18 do...	28.0199	18x18 ...	1x18 d...	13x13 do...	18x13 ...
4	1x18 do...	28.0199	18x18 ...	1x18 d...	13x13 do...	18x13 ...

Uma vez que a normalidade tenha sido ensinada ao supervisor, as falhas podem ser então catalogadas e referenciadas a cada uma das regiões de operação. Para a simulação, o instante de ocorrência de todas as 22 falhas foi definido para o minuto 250 e sua intensidade regulada para 200. O horizonte de simulação é fixado em 500 minutos e o período de amostragem, novamente, como 1 minuto.

Inicialmente, os quatro modos de operação serão treinados para todas as anomalias simuladas. O treinamento desta maneira possui dois objetivos

- Verificar o poder de generalização das assinaturas, isto é, se os padrões das contribuições se mantêm os mesmos para qualquer que seja a região de operação;
- Verificar se as assinaturas são robustas à variação de intensidade das falhas.

Durante a etapa de treinamento das falhas observou-se que as falhas 9, 16, 17 e 18 apresentam problemas de observabilidade nas regiões de operação projetadas e não podem ser detectadas pela análise estatística. A Figura 24a ilustra a ocorrência da falha 9 em  $OP_3$ . Note que, mesmo com a ocorrência da falha, as amostras se mantêm dentro da região de controle, indicada pelo elipsóide verde. Para efeito de comparação, a Figura 24b ilustra a falha 3 em  $OP_3$ , quando é possível observar o deslocamento das amostras da região de normalidade. Em ambos casos analisados, o plano é definido pelos dois *scores* mais importantes.

Por não ser possível observar as falhas 9, 16, 17 e 18 neste trabalho, essas anomalias não serão consideradas nas próximas análises. Dessa forma, o treinamento levará em conta somente as falhas cuja observação é possível, definindo a assinatura de cada uma delas para as regiões de operação estudadas. A significância da contribuição das variáveis de processo para a estatística  $T^2$  será determinada pelo limiar  $\theta = 0,3$  em um período de análise de  $\Delta = 5$  instantes, estes valores foram escolhidos de forma que a semelhança entre os padrões fosse maximizada. As assinaturas identificadas para cada um dos modos de normalidade são ilustradas pelas Figuras 25a, 25b, 25c e 25d.

Em muitos casos, é possível observar que a assinatura é a mesma para vários modos de operação, o que é uma característica fundamental para o aspecto de generalização dos



Figura 24 – Comparação entre anomalia observável e não-observável no plano definido pelas primeiras duas componentes principais

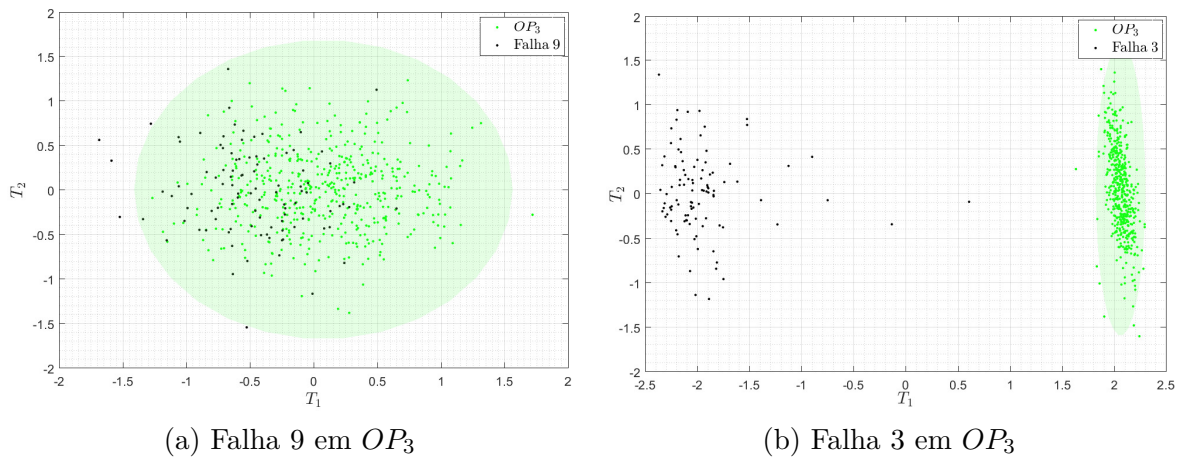
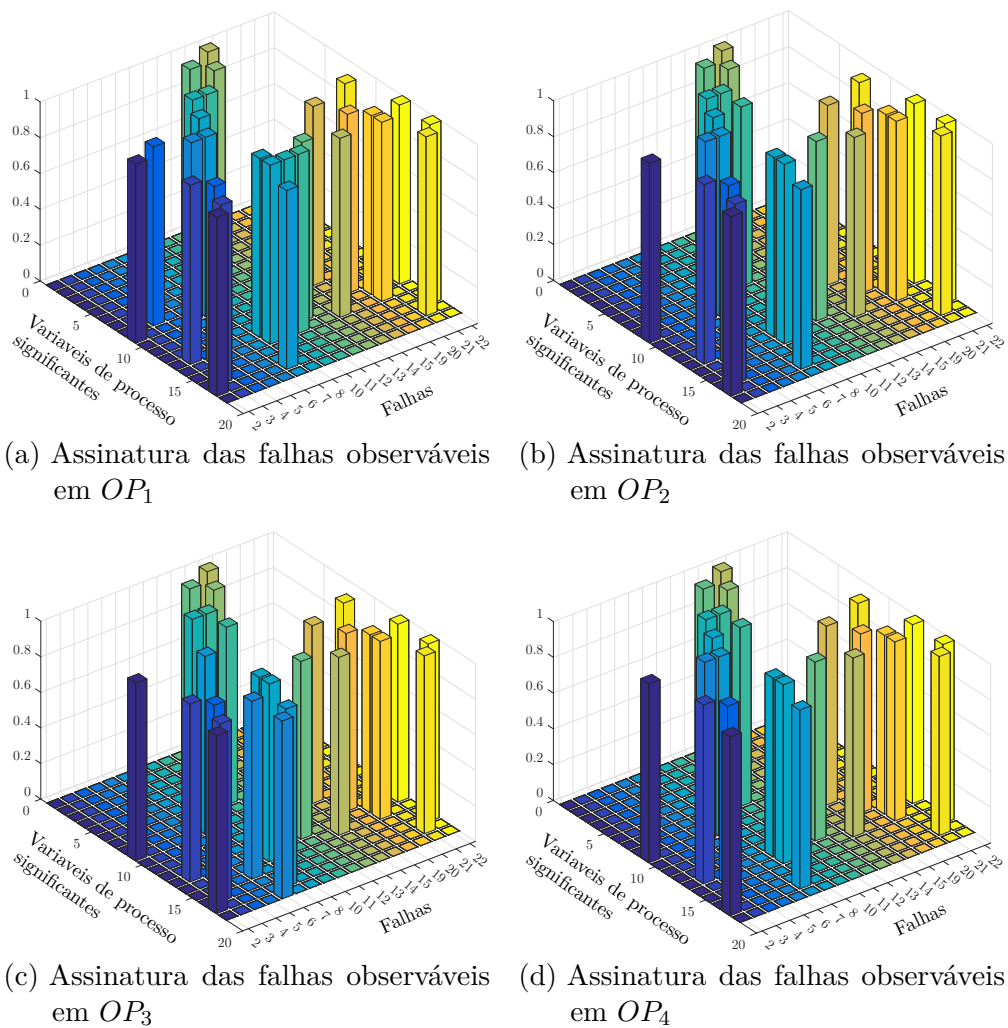


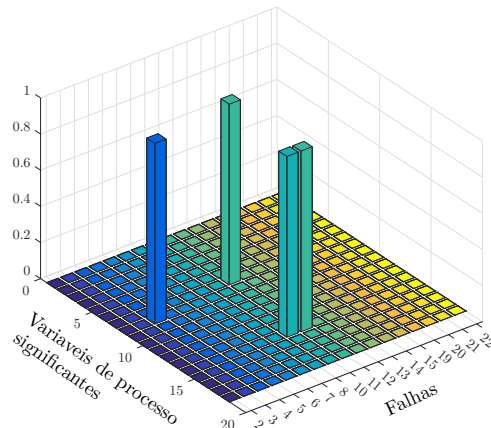
Figura 25 – Comparação entre as assinaturas das falhas observáveis calculadas para  $OP_1$ ,  $OP_2$ ,  $OP_3$  e  $OP_4$



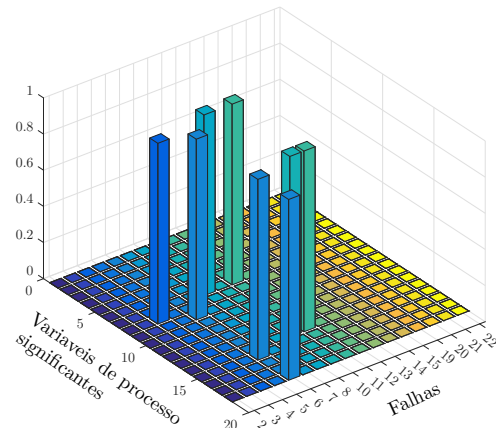
Fonte: Próprio autor.

padrões de falha. Para efeito de simplificação da análise, a Figura 26 ilustra as diferenças entre as assinaturas encontradas nos modos de operação estudados.

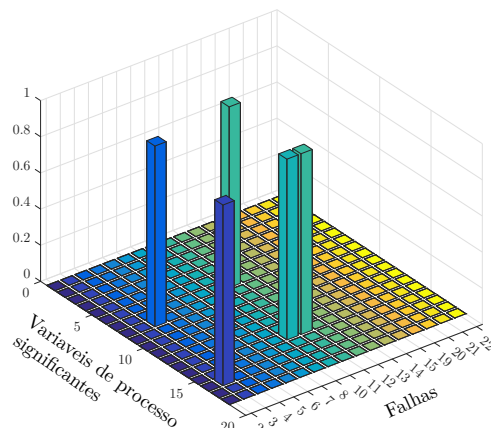
Figura 26 – Comparação entre as assinaturas das falhas encontradas para cada um dos modos de operação.



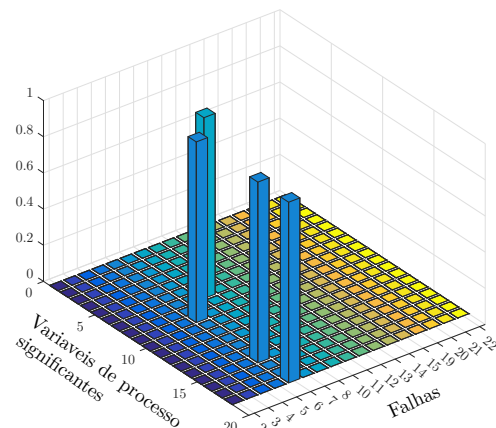
(a) Diferença entre as assinaturas das falhas em  $OP_1$  e  $OP_2$



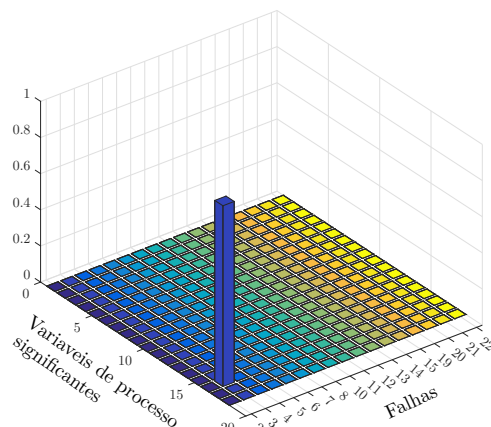
(b) Diferença entre as assinaturas das falhas em  $OP_1$  e  $OP_3$



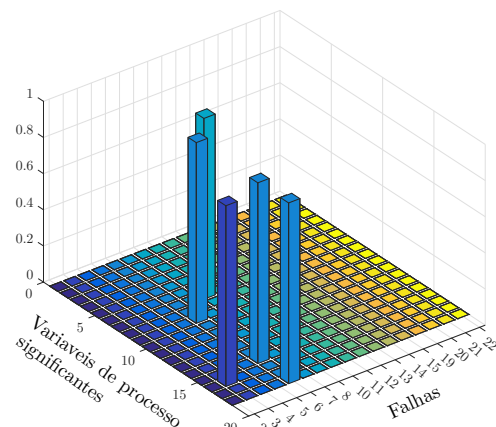
(c) Diferença entre as assinaturas das falhas em  $OP_1$  e  $OP_4$



(d) Diferença entre as assinaturas das falhas em  $OP_2$  e  $OP_3$



(e) Diferença entre as assinaturas das falhas em  $OP_2$  e  $OP_4$



(f) Diferença entre as assinaturas das falhas em  $OP_3$  e  $OP_4$

Fonte: Próprio autor.

## 5.2.2 Etapa de supervisão

A partir do treinamento realizado na etapa anterior, é possível assumir que tanto a operação normal quanto os padrões de falha já são conhecidos. Isto permite que o algoritmo de supervisão monitore o processo para outras diferentes situações que uma falha pode ocorrer. Nestes novos casos, as simulações levarão em intensidades de falha distintas daquela utilizada para o treinamento. Assim, o objetivo desta etapa é a verificação da robustez das assinaturas em relação a esta propriedade da anomalia.

As simulações da operação com falha serão realizadas para os valores de intensidade: 50 e 300. A Tabela 6 indica o resultado do diagnóstico produzido pelo algoritmo de supervisão para os novos conjuntos considerando os mesmos parâmetros utilizados para o treinamento. Neste caso, as anomalias cuja detecção é corretamente realizada são indicadas por *X*.

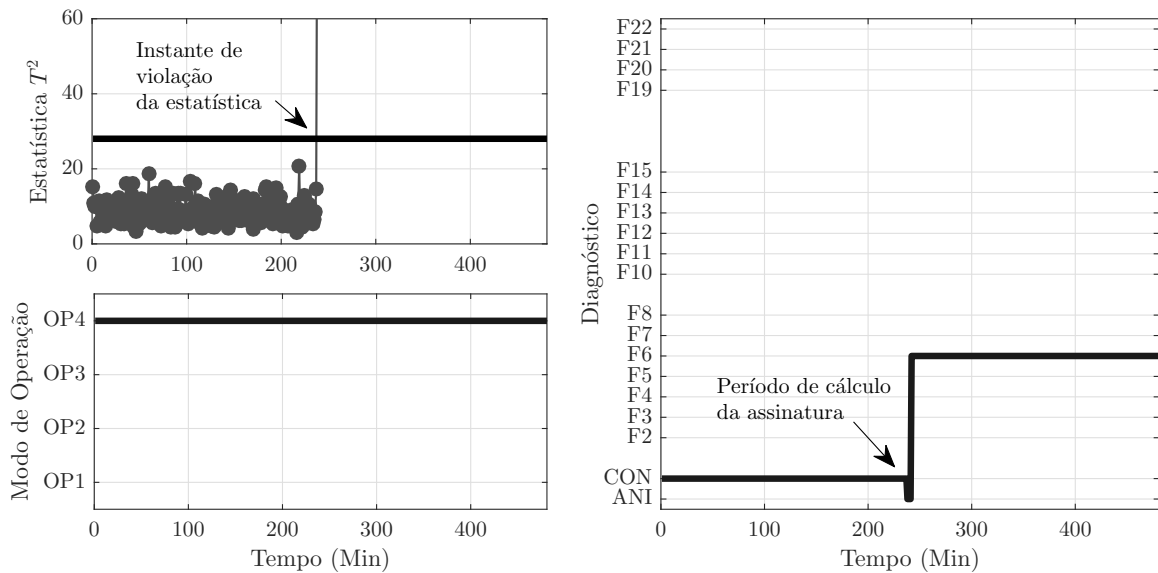
Tabela 6 – Teste de diagnóstico correto para falhas em diferentes intensidades

Falha	$OP_1$		$OP_2$		$OP_3$		$OP_4$		Acertos (%)
	intensidade		intensidade		intensidade		intensidade		
	50	300	50	300	50	300	50	300	
2	X	X		X		X		X	62,5
3		X	X		X	X	X	X	75
4			X	X	X				37,5
5	X		X	X	X	X	X	X	87,5
6	X		X	X			X	X	62,5
7	X	X	X	X	X	X	X	X	100
8		X				X	X		37,5
10	X	X		X	X	X	X	X	87,5
11						X	X		25
12		X	X	X	X	X	X	X	87,5
13	X	X	X	X	X	X		X	87,5
14	X	X	X	X	X		X	X	87,5
15	X	X	X	X	X	X	X	X	100
19	X	X	X	X	X	X	X	X	100
20	X	X	X	X	X	X	X	X	100
21	X	X	X	X	X	X	X	X	100
22						X			12,5
Acertos (%)	68,75	70,58	70,58	76,47	70,58	82,35	76,47	76,47	

O resultado da detecção e do diagnóstico para algumas falhas são ilustradas pelas Figuras 27 e 28, que são geradas de acordo com a interface elaborada para o supervisor. Nestes casos, é possível verificar a atuação do algoritmo para o diagnóstico de uma falha (Falha 6) em  $OP_4$  e que possui uma assinatura única, e para o diagnóstico de outra falha (Falha 7) em  $OP_1$  cuja assinatura seja similar a de outras falhas. As duas anomalias ocorreram no minuto 250 e se perpetuaram até o horizonte de simulação, definido como 500 minutos.

Em ambos casos, é possível observar que a detecção da região de operação ocorreu corretamente, assim como a detecção da anomalia. Entretanto, para o primeiro caso, a Falha 6 é corretamente identificada, enquanto a Falha 7 é diagnosticada juntamente com as Falhas 2 e 21, que possuem a mesma assinatura da real anomalia. Isto pode ser melhor visualizado através da Figura 29. Note que, inicialmente, o diagnóstico indica *CON* (Condição de Operação Normal), que se mantém até o primeiro instante de detecção da falha. Durante as próximas  $\Delta = 5$  amostras, o diagnóstico aponta *ANI* (Anomalia Não Identificada). Isto acontece já que o algoritmo ainda está identificado a assinatura da falha e, por isso, não poderia diagnosticar a ocorrência. Uma vez transcorrido o período necessário para a captura do padrão, a assinatura produzida é comparada com cada uma das assinaturas da região conhecida.

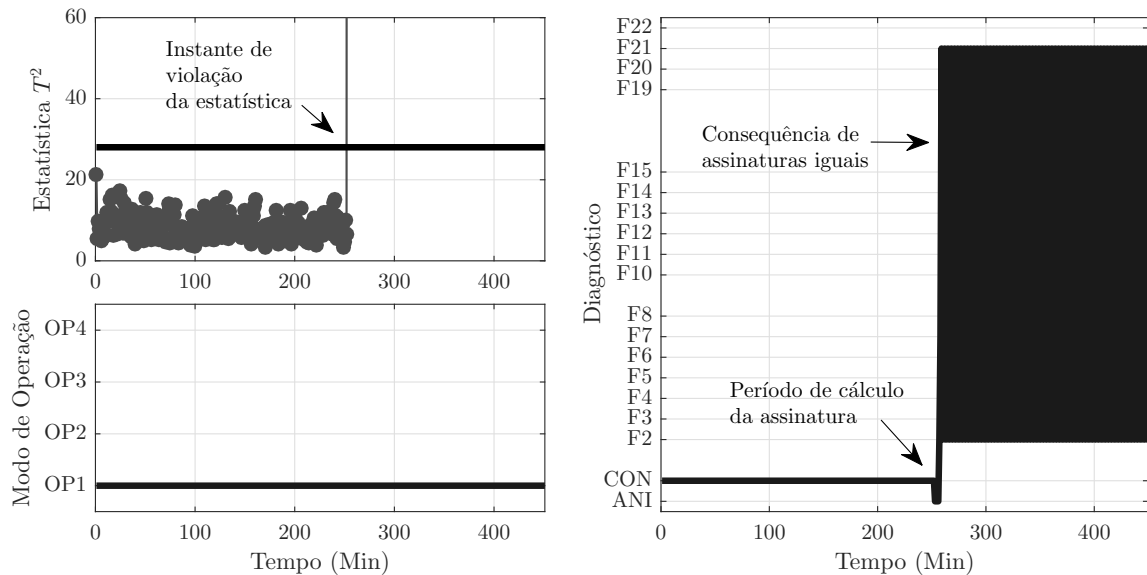
Figura 27 – Detecção e diagnóstico da Falha 6 no modo de operação  $OP_4$  simulada para intensidade 50.



Fonte: Próprio autor.

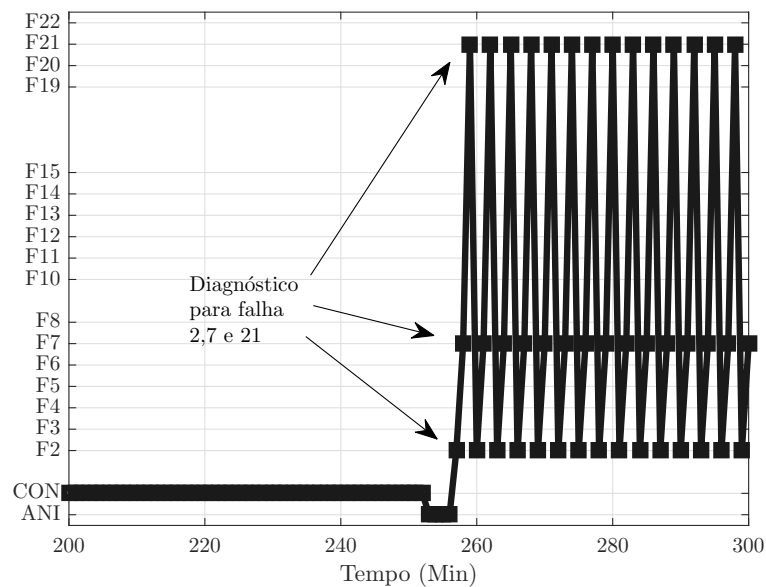
A fim de verificar a atuação do supervisor para uma situação sem falhas, em que apenas mudanças nos modos de operação ocorrem, é simulada uma operação da planta variando entre os modos conhecidos  $OP_1$ ,  $OP_3$ ,  $OP_4$  e  $OP_2$ , respectivamente. O resultado da supervisão é ilustrado pela Figura 30. Novamente, os modos de operação são corretamente identificados. O rompimento do limiar estatístico é um reflexo esperado do transitório (variação de causa especial), sendo corretamente caracterizado e perdurando até o instante em que o sistema volte ao um novo estado estacionário. Entretanto, o diagnóstico produz resultados equivocados durante o transitório. É possível observar que a resposta encontrada oscila entre o diagnóstico correto, que neste caso deveria ser *ANI*, e o diagnóstico da Falha 19.

Figura 28 – Detecção e diagnóstico de Falha 7 no modo de operação  $OP_1$  simulada para intensidade 300. Esta anomalia compartilha a mesma assinatura com a Falha 2 e 21.



Fonte: Próprio autor.

Figura 29 – Ampliação do resultado do diagnóstico da Falha 7 no modo de operação  $OP_1$ . É possível observar que o diagnóstico é confundido entre as Falhas 2, 7 e 21.



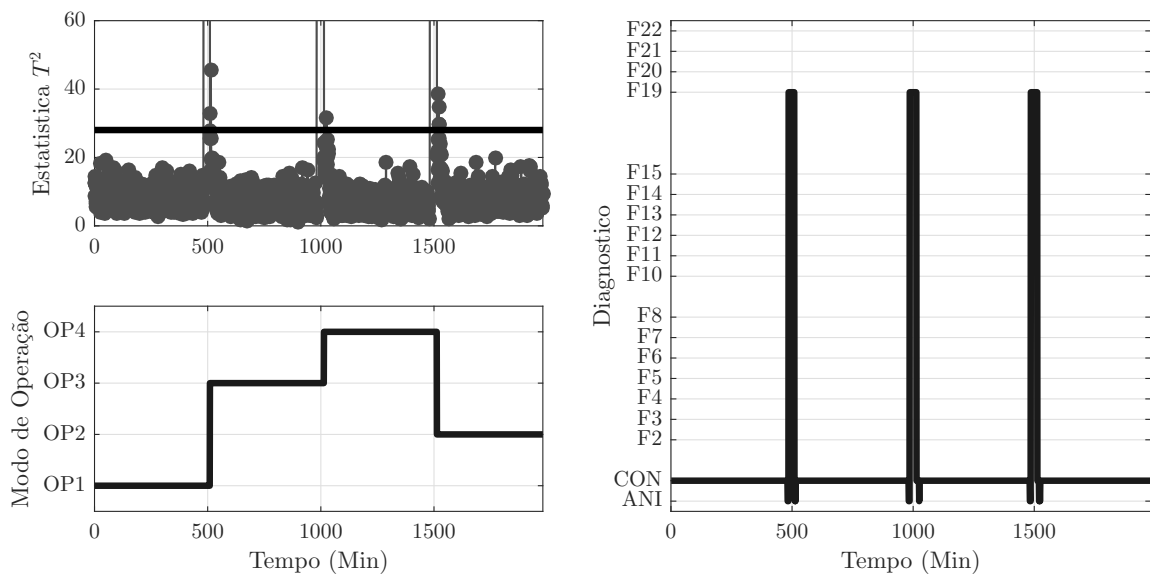
Fonte: Próprio autor.

Outra situação que ainda merece análise é a capacidade de generalização das assinaturas entre os diferentes modos de operação. Para que isto seja verificado, uma alteração nos resultados de treinamento anterior é necessária. Assim, para os modos de operação  $OP_2$ ,  $OP_3$  e  $OP_4$  serão apagadas todas as assinaturas identificadas anteriormente,

de forma que os padrões de falha sejam conhecidos apenas em  $OP_1$ . A modificação nos conjuntos é ilustrada pela Figura 31. O processo é simulado começando em  $OP_1$  onde a Falha 6 ocorre no instante 250. A operação volta à normalidade em  $OP_2$  no minuto 300, no minuto 550 a Falha 8 ataca o sistema. A planta volta a operar normalmente somente no instante 600 assumindo  $OP_3$  e assim continua por 250 minutos, quanto a Falha 5 ocorre. Por final, a operação retorna à normalidade no minuto 900, ocorrendo a Falha 12 no instante 1150. Em todos os casos, o operador confirma o aprendizado das assinaturas conforme a requisição apresentada pelo fluxograma da Figura 19. O resultado da supervisão é apresentado pela Figura 32.

É possível observar que os modos de operação foram, novamente, identificados de forma correta e a estatística indica, precisamente, o surgimento das anomalias nos instantes simulados. Muito embora o algoritmo não tenha sido treinado para as falhas

Figura 30 – Supervisão para a situação onde apenas mudanças nos modos de operação ocorrem. A planta opera em  $OP_1$ ,  $OP_3$ ,  $OP_4$  e  $OP_2$  respectivamente.



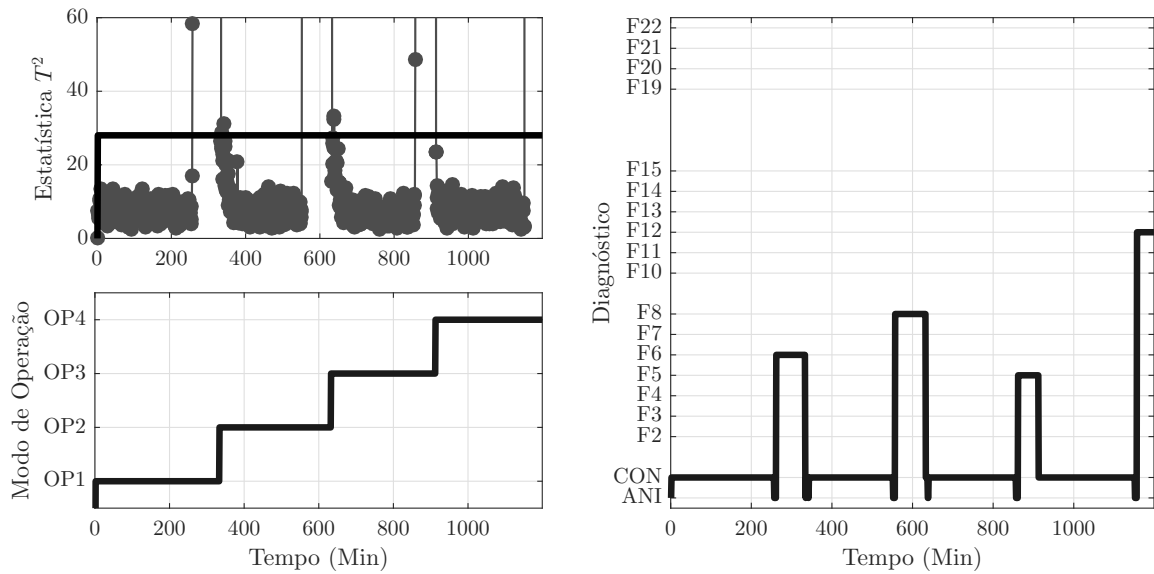
Fonte: Próprio autor.

Figura 31 – Estrutura de treinamento modificada. Neste caso, as assinaturas são conhecidas somente em  $OP_1$ , conforme indicado pelo campo  $F_k$ .

Fields	Centro	t2lim	M	Std	ValSing	PC	F_k
1	1x18 do...	28.0199	18x18 ...	1x18 d...	13x13 do...	18x13 ...	17x1 struct
2	1x18 do...	28.0199	18x18 ...	1x18 d...	13x13 do...	18x13 ...	□
3	1x18 do...	28.0199	18x18 ...	1x18 d...	13x13 do...	18x13 ...	□
4	1x18 do...	28.0199	18x18 ...	1x18 d...	13x13 do...	18x13 ...	□

Fonte: Próprio autor.

Figura 32 – Resultado da supervisão para falhas não treinadas em  $OP_2, OP_3$  e  $OP_4$ .



Fonte: Próprio autor.

em  $OP_2, OP_3$  e  $OP_4$ , o diagnóstico acusa corretamente todas as falhas simuladas e, com auxílio do operador, é possível agregar estes padrões aos modos em questão. O resultado deste aprendizado dinâmico pode ser visualizado através da Figura 33, isto permite que novos padrões possam ser acrescentados aos modos de operação sem que, necessariamente, uma anomalia ocorra nestas direções.

Figura 33 – Estrutura de treinamento modificada. Neste caso, as assinaturas são conhecidas somente em  $OP_1$ , conforme indicado pelo campo  $F_k$ .

Fields	Centro	t2lim	M	Std	ValSing	PC	F_k
1	1x18 do...	28.0199	18x18 ...	1x18 d...	13x13 do...	18x13 ...	17x1 struct
2	1x18 do...	28.0199	18x18 ...	1x18 d...	13x13 do...	18x13 ...	1x1 struct
3	1x18 do...	28.0199	18x18 ...	1x18 d...	13x13 do...	18x13 ...	1x1 struct
4	1x18 do...	28.0199	18x18 ...	1x18 d...	13x13 do...	18x13 ...	1x1 struct

### 5.3 Conclusões

Desde o seu uso aplicado na indústria, metodologias estatísticas vêm transformando possibilidades de ganho em produção e qualidade numa realidade. Diversas abordagens foram criadas no decorrer das últimas décadas e ainda continuam sendo aplicadas em muitas indústrias, tendo um papel extremamente importante para os processos.

A abordagem apresentada por este trabalho contribui para o problema de detecção de falhas em múltiplos pontos de operação e também uma solução para a questão do diagnóstico. As considerações necessárias, para tanto dizem respeito à distribuição das variáveis, que devem ser gaussianas, e à observabilidade das falhas.

Para a maioria das anomalias analisadas, foi possível ao menos detectar sua ocorrência através da estatística  $T^2$ . Além disso, a metodologia proposta para a identificação das regiões de operação através da distância de Mahalanobis produziu resultados confiáveis em todos os casos. Neste contexto, a natureza bem definida e espaçada dos *clusters* foi um importante contribuidor para a acurácia do método. Em todo caso, a inferência sobre o modo de operação e detecção das anomalias apresentaram resultados consistentes.

Entretanto, muito trabalho foi empregado para o desenvolvimento da assinatura das falhas, de forma que a sua capacidade de generalização fossem maximizadas. A generalização, neste caso, visou tanto a equivalência das assinaturas nos diferentes modos de operação, quanto a equivalência para falhas com diferentes intensidades. É importante ressaltar que os atributos do limiar de significância  $\theta$  e o período de coleta da assinatura  $\Delta$  foram parâmetros cruciais para tanto. De toda forma, o nível de significância  $\alpha$  definido para a estatística  $T^2$ , o percentual de confiabilidade adotado para a análise de componentes principais e a constante de amortização do filtro também tiveram significativa influência sobre os padrões. Isto se deve ao fato que todos estes atributos alteram, de alguma forma, a região de confiabilidade ou inserem algum atraso às variáveis de processo. Provocando, naturalmente, certas diferenças nas contribuições das variáveis para o rompimento do limiar estatístico.

As assinaturas produzidas foram bastante similares entre os modos de operação conforme ilustrado pelas Figuras 25 e 26. Além disto, as diferentes intensidades de falha também puderam ser corretamente identificadas para a maioria dos casos, conforme indicado pela Tabela 6. Em todo caso, a similaridade entre os padrões permitiu que a metodologia de aprendizado dinâmico e contínuo fosse validada. Isto possibilitou que novas falhas pudessem ser acrescentadas aos conjuntos dos modos de operação sem que, necessariamente, ocorressem antes para os modos em questão.



## 5.4 Trabalhos futuros

Tendo em vista que os atributos  $\theta$  e  $\Delta$  são fundamentais para o diagnóstico e generalização dos padrões, propõem-se como trabalho futuro a elaboração de alguma metodologia automatizada para a definição dos valores ótimos, ou seja, aqueles que apliem a capacidade de generalização dos padrões.

Como complemento, deixa-se uma proposta para a utilização da estatística  $Q$  e utilização da contribuição das variáveis para a confecção de novas assinaturas de falha, que visem diminuir a taxa de falsos diagnósticos, como visto na Figura 30.

Deverá ser realizada uma análise mais aprofundada sobre a não conformidade dos padrões das assinaturas das falhas 3, 4, 6, 8, 10 e 11 em relação aos diferentes modos de operação. Além disso, é necessário que os resultados descritos na Tabela 6 sejam melhor analisados. Os problemas ligados à observabilidade das falhas 9, 16, 17 e 18 deverão ser avaliados futuramente.

Outro aspecto que merece atenção e é deixado como trabalho futuro é a inclusão de mecanismos para o aperfeiçoamento do modelo das regiões de operação conhecidas através das técnicas de Análise de Componentes Principais Recursiva propostas em [Portnoy et al. \(2016\)](#).

Por final, é fundamental que a abordagem proposta por este trabalho também seja aplicada à supervisão de dados de operações reais.

# Referências

- AKAIKE, H. A new look at statistical model identification *IEEE Transact Automatic Control* 19: 716–722. *Find this article online*, 1974. Citado na página 44.
- ANDERSON, T. *An Introduction into Multivariate Statistical Analysis*. [S.l.]: John Wiley & Sons, 2003. Citado na página 42.
- BIN, G. et al. Early fault diagnosis of rotating machinery based on wavelet packets—empirical mode decomposition feature extraction and neural network. *Mechanical Systems and Signal Processing*, Elsevier, v. 27, p. 696–711, 2012. Citado na página 23.
- BRYDON, D.; CILLIERS, J.; WILLIS, M. Classifying pilot-plant distillation column faults using neural networks. *Control Engineering Practice*, Elsevier, v. 5, n. 10, p. 1373–1384, 1997. Citado na página 26.
- CAMPOS, V. F. *O verdadeiro poder*. [S.l.]: INDG-Instituto de Desenvolvimento Gerencia, 2009. Citado na página 28.
- CATTELL, R. B. The scree test for the number of factors. *Multivariate behavioral research*, Taylor & Francis, v. 1, n. 2, p. 245–276, 1966. Citado na página 44.
- CHANDOLA, V.; BANERJEE, A.; KUMAR, V. Anomaly detection: A survey. *ACM computing surveys (CSUR)*, ACM, v. 41, n. 3, p. 15, 2009. Citado 2 vezes nas páginas 20 e 22.
- CHIANG, L. H.; BRAATZ, R. D.; RUSSELL, E. L. *Fault detection and diagnosis in industrial systems*. [S.l.]: Springer Science & Business Media, 2001. Citado 7 vezes nas páginas 25, 34, 37, 42, 43, 45 e 46.
- CHIANG, L. H.; RUSSELL, E. L.; BRAATZ, R. D. Fault diagnosis in chemical processes using fisher discriminant analysis, discriminant partial least squares, and principal component analysis. *Chemometrics and intelligent laboratory systems*, Elsevier, v. 50, n. 2, p. 243–252, 2000. Citado na página 26.
- COOKE, R.; PAULSEN, J. Concepts for measuring maintenance performance and methods for analysing competing failure modes. *Reliability Engineering & System Safety*, Elsevier, v. 55, n. 2, p. 135–141, 1997. Citado na página 15.
- DUDZIC, M.; VACULIK, V.; MILETIC, I. On-line applications of multivariate statistics at dofasco. *Proceedings Future Trends in Automation in Mineral and Metal Processing*, p. 438–443, 2000. Citado 2 vezes nas páginas 16 e 26.
- ELECTRIC, W. *Statistical quality control handbook*. Easton: The Mack Printing Company, 1956. Citado na página 36.
- FINCH, F. E. *Automated fault diagnosis of chemical process plants using model-based reasoning*. Tese (Doutorado) — Massachusetts Institute of Technology, 1989. Citado 2 vezes nas páginas 8 e 58.

- FUENTE, M. J. et al. Fault detection and identification method based on multivariate statistical techniques. In: IEEE. *2009 IEEE Conference on Emerging Technologies & Factory Automation*. [S.l.], 2009. p. 1–6. Citado na página 42.
- FUENTES, F. F. E. et al. Metodologia para inovação da gestão de manutenção industrial. Florianópolis, SC, 2006. Citado na página 16.
- GOSSET, W. S. The application of the law of error to the work of the brewery. *Guinness Internal Note*, p. 33, 1904. Citado na página 36.
- HE, X. B. et al. Variable-weighted fisher discriminant analysis for process fault diagnosis. *Journal of Process Control*, Elsevier, v. 19, n. 6, p. 923–931, 2009. Citado na página 26.
- HINO, S. *Inside the mind of Toyota: Management principles for enduring growth*. [S.l.]: CRC Press, 2005. Citado na página 28.
- HOGG, R. V.; CRAIG, A. T. *Introduction to Mathematical Statistics: 3d Ed.* [S.l.]: Macmillan, 1970. Citado na página 32.
- HOTELLING, H. Analysis of a complex of statistical variables into principal components. *Journal of educational psychology*, Warwick & York, v. 24, n. 6, p. 417, 1933. Citado na página 27.
- ISERMANN, R. Supervision, fault-detection and fault-diagnosis methods—an introduction. *Control engineering practice*, Elsevier, v. 5, n. 5, p. 639–652, 1997. Citado na página 23.
- ISERMANN, R.; BALLÉ, P. Trends in the application of model-based fault detection and diagnosis of technical processes. *Control engineering practice*, Elsevier, v. 5, n. 5, p. 709–719, 1997. Citado 2 vezes nas páginas 23 e 25.
- JACKSON, J. E. *A user's guide to principal components*. [S.l.]: John Wiley & Sons, 2005. v. 587. Citado 2 vezes nas páginas 27 e 43.
- JOHNSON, R. A.; WICHERN, D. W. et al. *Applied multivariate statistical analysis*. [S.l.]: Prentice hall Upper Saddle River, NJ, 2002. v. 5. Citado 2 vezes nas páginas 40 e 41.
- KOURTI, T. Process analysis and abnormal situation detection: from theory to practice. *IEEE control systems*, IEEE, v. 22, n. 5, p. 10–25, 2002. Citado 4 vezes nas páginas 15, 26, 37 e 51.
- KRUGER, U.; XIE, L. *Advances in statistical monitoring of complex multivariate processes: with applications in industrial process control*. [S.l.]: John Wiley & Sons, 2012. Citado 5 vezes nas páginas 26, 27, 38, 42 e 43.
- KULKARNI, A.; JAYARAMAN, V. K.; KULKARNI, B. D. Knowledge incorporated support vector machines to detect faults in tennessee eastman process. *Computers & chemical engineering*, Elsevier, v. 29, n. 10, p. 2128–2133, 2005. Citado na página 26.
- MACGREGOR, J. F.; KOURTI, T. Statistical process control of multivariate processes. *Control Engineering Practice*, Elsevier, v. 3, n. 3, p. 403–414, 1995. Citado na página 42.
- MAESSCHALCK, R. D.; JOUAN-RIMBAUD, D.; MASSART, D. L. The mahalanobis distance. *Chemometrics and intelligent laboratory systems*, Elsevier, v. 50, n. 1, p. 1–18, 2000. Citado na página 45.

- MAESTRI, M. et al. A robust clustering method for detection of abnormal situations in a process with multiple steady-state operation modes. *Computers & chemical engineering*, Elsevier, v. 34, n. 2, p. 223–231, 2010. Citado 2 vezes nas páginas 18 e 48.
- MAHALANOBIS, P. C. On the generalized distance in statistics. *Proceedings of the National Institute of Sciences (Calcutta)*, v. 2, p. 49–55, 1936. Citado na página 45.
- MONROY, I. et al. Enhanced plant fault diagnosis based on the characterization of transient stages. *Computers & Chemical Engineering*, v. 37, p. 200 – 213, 2012. ISSN 0098-1354. Citado 2 vezes nas páginas 18 e 26.
- MONTGOMERY, D. C.; RUNGER, G. C. *Applied statistics and probability for engineers*. [S.l.]: John Wiley & Sons, 2010. Citado 3 vezes nas páginas 28, 29 e 34.
- MURRAY-SMITH, R.; JOHANSEN, T. *Multiple model approaches to nonlinear modelling and control*. [S.l.]: CRC press, 1997. Citado na página 14.
- OGUNNAIKE, B. A. *Process dynamics, modeling, and control*. [S.l.: s.n.], 1994. Citado na página 31.
- OYELEYE, O. O. *Qualitative modeling of continuous chemical processes and applications to fault diagnosis*. Tese (Doutorado) — Massachusetts Institute of Technology, 1989. Citado 2 vezes nas páginas 8 e 58.
- PEASON, K. On lines and planes of closest fit to systems of point in space. *Philosophical Magazine*, v. 2, p. 559–572, 1901. Citado na página 27.
- PIOVOSO, M.; KOSANOVICH, K.; YUK, J. Process data chemometrics. *IEEE transactions on instrumentation and measurement*, IEEE, v. 41, n. 2, p. 262–268, 1992. Citado na página 16.
- PORTNOY, I. et al. An improved weighted recursive pca algorithm for adaptive fault detection. *Control Engineering Practice*, Elsevier, v. 50, p. 69–83, 2016. Citado 2 vezes nas páginas 18 e 72.
- ROKACH, L. Genetic algorithm-based feature set partitioning for classification problems. *Pattern Recognition*, Elsevier, v. 41, n. 5, p. 1676–1700, 2008. Citado na página 26.
- RUIZ, D.; NOUGUES, J. M.; PUIGJANER, L. Fault diagnosis support system for complex chemical plants. *Computers & Chemical Engineering*, Elsevier, v. 25, n. 1, p. 151–160, 2001. Citado na página 26.
- SPENCE, C.; PARRA, L.; SAJDA, P. Detection, synthesis and compression in mammographic image analysis with a hierarchical image probability model. In: *IEEE. Mathematical Methods in Biomedical Image Analysis, 2001. MMBIA 2001. IEEE Workshop on*. [S.l.], 2001. p. 3–10. Citado na página 20.
- SU, C.-T.; YANG, C.-H. Feature selection for the svm: An application to hypertension diagnosis. *Expert Systems with Applications*, Elsevier, v. 34, n. 1, p. 754–763, 2008. Citado na página 26.
- WIREMAN, T. *Benchmarking best practices in maintenance management*. [S.l.]: Industrial Press Inc., 2004. Citado na página 15.

- WOLD, S. Cross-validatory estimation of the number of components in factor and principal components models. *Technometrics*, Taylor & Francis, v. 20, n. 4, p. 397–405, 1978. Citado na página 44.
- YÉLAMOS, I. et al. Performance assessment of a novel fault diagnosis system based on support vector machines. *Computers & chemical engineering*, Elsevier, v. 33, n. 1, p. 244–255, 2009. Citado na página 26.
- ZAREI, J. Induction motors bearing fault detection using pattern recognition techniques. *Expert systems with Applications*, Elsevier, v. 39, n. 1, p. 68–73, 2012. Citado na página 20.
- ZHANG, H. *Statistical Process Monitoring and Modeling Using PCA and PLS*. Dissertação (Mestrado) — University of Alberta, 2000. Citado na página 44.
- ZHAO, S. J.; ZHANG, J.; XU, Y. M. Monitoring of processes with multiple operating modes through multiple principle component analysis models. *Industrial & engineering chemistry research*, ACS Publications, v. 43, n. 22, p. 7025–7035, 2004. Citado 3 vezes nas páginas 48, 49 e 50.
- ZHAO, S. J.; ZHANG, J.; XU, Y. M. Performance monitoring of processes with multiple operating modes through multiple pls models. *Journal of process Control*, Elsevier, v. 16, n. 7, p. 763–772, 2006. Citado na página 48.
- ZHENG, Y.; QIN, S. J.; CHAI, T. Fault detection of multimode processes using concurrent projection to latent structures. 2016. Citado na página 18.