



UMA AVALIAÇÃO CRÍTICA SOBRE TÉCNICAS BASEADAS EM PCA PARA DETECÇÃO DE FALHAS EM PROCESSOS DA INDÚSTRIA QUÍMICA

Fernando Elias de Freitas Fadel

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia Química, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia Química.

Orientador: José Carlos Costa da Silva Pinto

Rio de Janeiro
Junho de 2018

UMA AVALIAÇÃO CRÍTICA SOBRE TÉCNICAS BASEADAS EM PCA PARA
DETECÇÃO DE FALHAS EM PROCESSOS DA INDÚSTRIA QUÍMICA

Fernando Elias de Freitas Fadel

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO
ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE
ENGENHARIA (COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE
JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A
OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA
QUÍMICA.

Examinada por:

Prof. José Carlos Costa da Silva Pinto, D.Sc.

Prof. Maurício Bezerra de Souza Jr., D.Sc.

Dr. Cristhian Alberto Celestino Cortez, D.Sc.

RIO DE JANEIRO, RJ – BRASIL

JUNHO DE 2018

Fadel, Fernando Elias de Freitas

Uma Avaliação Crítica Sobre Técnicas Baseadas em PCA para Detecção de Falhas em Processos da Indústria Química/Fernando Elias de Freitas Fadel. – Rio de Janeiro: UFRJ/COPPE, 2018.

XI, 111 p. 29, 7cm.

Orientador: José Carlos Costa da Silva Pinto

Dissertação (mestrado) – UFRJ/COPPE/Programa de Engenharia Química, 2018.

Referências Bibliográficas: p. 94 – 108.

1. Modelagem Estatística. 2. Monitoramento Estatístico de Processos. 3. Detecção de falhas. I. Pinto, José Carlos Costa da Silva. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Química. III. Título.

*Dedico este trabalho a minha
esposa, companheira e cúmplice,
com quem aprendi a sonhar.
Também dedico a meu filho por
sua alegria contagiante que me
faz seguir em frente a cada dia.*

Agradecimentos

A Deus, pelo dom da vida e pela capacidade para concluir este trabalho.

A minha esposa, pela cumplicidade, suporte e paciência; por ter sido uma das maiores incentivadoras deste projeto.

A meu filho, por me ajudar a relaxar e a rir.

A meus pais e irmãos, porque, apesar de distantes, sempre estarão no meu coração.

A todos os colegas e professores do Programa de Engenharia Química do Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa em Engenharia, por terem me recebido e acolhido de maneira extraordinária.

De forma especial, gostaria de agradecer a meu orientador, José Carlos Costa da Silva Pinto, por me encorajar a trilhar novos caminhos, pelos bons momentos de conversa, pela paciência e pelo constante incentivo, além de me ensinar a ser um pesquisador. Obrigado!

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

UMA AVALIAÇÃO CRÍTICA SOBRE TÉCNICAS BASEADAS EM PCA PARA DETECÇÃO DE FALHAS EM PROCESSOS DA INDÚSTRIA QUÍMICA

Fernando Elias de Freitas Fadel

Junho/2018

Orientador: José Carlos Costa da Silva Pinto

Programa: Engenharia Química

O objetivo desta dissertação foi estudar, desenvolver e implementar sistemas de monitoramento de processos químicos. Para compreender as limitações das técnicas de modelagem baseadas em dados, adotaram-se como ponto de partida as técnicas clássicas, como a Análise de Componentes Principais (PCA) e técnicas correlatas. Buscou-se construir uma visão crítica dos métodos clássicos usados para monitoramento de processos. Modelos baseados em dados são bastante limitados, já que modelos previamente aprendidos não se adaptam necessariamente bem a novas regiões de operação, cujos dados não estavam disponíveis durante a fase de modelagem. Além disso, é possível concluir que as técnicas baseadas em dados disponíveis atualmente, como PCA e suas variantes, não apresentam capacidade de realmente modelar o comportamento dinâmico de um processo. Esta foi a principal motivação para o estudo dos Gráficos de Recorrência. A partir do conceito de recorrência, foi desenvolvida uma nova técnica para monitoramento de processos com múltiplos pontos operacionais. Foi proposta uma nova carta de controle para acompanhamento de processos com múltiplos pontos de operação, fundamentada na Norma de Frobenius. Muitos outros aspectos relacionados ao tema de monitoramento, tais como seleção de variáveis, pré-processamento e remoção de valores espúrios, não foram abordados no trabalho, mas são de fundamental importância para a aplicação industrial. Finalmente, é importante ressaltar que, embora a metodologia de monitoramento proposta tenha sido inspirada no conceito de recorrência, ela, de fato, não é capaz de reconstruir o espaço de fases do processo.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

A CRITICAL ASSESSMENT OF PCA BASED TECHNIQUES FOR FAULT DETECTION IN INDUSTRIAL CHEMICAL PROCESSES

Fernando Elias de Freitas Fadel

June/2018

Advisor: José Carlos Costa da Silva Pinto

Department: Chemical Engineering

The aim of this dissertation was to study, develop and implement chemical process monitoring systems. In order to understand the limitations of Data Driven modeling techniques, classical techniques based on Principal Component Analysis (PCA) and related procedures were adopted as a starting point. It was sought to build a critical view of the classical methods of process monitoring. Data Driven models are quite limited, since previously learned models do not fit necessarily well new operating regions, whose data were not available during the modeling phase. In addition, it is possible to conclude that the techniques based on data currently available, such as PCA and its variants, do not have the capacity to actually model the dynamic behavior of a process. This was the main motivation for the study of Recurrence Plots. Based on the concept of recurrence, a new technique was developed to monitor processes with multiple operational points. A new control chart was proposed to monitor processes with multiple operation points, based on the Frobenius Norm. Many other aspects related to process monitoring, such as variable selection, preprocessing and removal of spurious values, were not addressed in this work, but are of fundamental importance for industrial application. Finally, it is important to note that, although the proposed monitoring methodology has been inspired by the concept of recurrence, it, in fact, is not able to reconstruct the phase space of the process.

Sumário

Lista de Abreviaturas	x
1 Introdução	1
1.1 Objetivos	3
1.2 Organização	3
2 Revisão Bibliográfica	4
2.1 Processos Químicos Industriais	4
2.1.1 Característica dos Dados da Indústria de Processo	5
2.1.2 Causas Comuns de Falhas em Processos Químicos	7
2.2 Monitoração de Processos	9
2.2.1 Metodologia de Monitoração de Processos	9
2.2.2 Requisitos de um Sistema de Monitoramento	18
2.2.3 Panorama de Evolução da Monitoração de Processo	20
2.3 Comentários Finais	24
3 Análise de Componentes Principais	25
3.1 PCA Clássico ou Estático	25
3.1.1 Geral	25
3.1.2 Seleção do Número Componentes Principais	29
3.1.3 Índices de Detecção de Falhas \mathbf{T}^2 de Hotelling e \mathbf{Q}	33
3.2 PCA Dinâmico - DPCA	34
3.2.1 Geral	34
3.2.2 Procedimento de Seleção do Atraso	36
3.3 PCA Não-Linear - KPCA	47
3.3.1 Geral	47
3.3.2 Funções Núcleo Mais Utilizadas	50
3.4 Outras Variantes de PCA	52
3.4.1 PCA Robusto	52
3.4.2 PCA Variante no Tempo	53
3.4.3 PCA Probabilístico	53

3.5	Outras Técnicas de Análise de Subespaços	53
3.5.1	Análise de Componentes Independentes	54
3.5.2	Abordagem Linear Local	54
3.5.3	Análise de Subespaços Estacionários	54
3.5.4	Identificação de Sistemas Híbridos	55
3.6	Comentários Finais	55
4	Avaliação da Técnica DPCA Para Monitoramento de Processos	56
4.1	O CSTR com Reação de Van der Vusse	56
4.2	Identificação do Modelo Para Monitoramento	59
4.3	Monitoramento de Falhas Via DPCA	68
4.4	Comentários Finais	70
5	Detecção de Falhas Via Matriz de Recorrência	73
5.1	Introdução	73
5.2	Metodologia de Monitoramento Via Matriz de Recorrência	75
5.2.1	Comparando o Algoritmo Proposto à Técnica de PCA	79
5.2.2	Exemplo 5.1 - Prova de Conceito	81
5.3	Estendendo a Metodologia Proposta	84
5.3.1	Exemplo 5.2 - Revisitando o Reator Van der Vusse	85
5.4	Comentários Finais	89
6	Conclusões	91
6.1	Conclusões	91
6.2	Sugestões de Trabalhos Futuros	92
	Referências Bibliográficas	94
A	Terminologia Usada na Área de Monitoramento de Processos	109

Lista de Abreviaturas

ARIMA	Modelo Auto-Regressivo Integrado de Médias Móveis, p. 15
CVA	Análise de Variáveis Canônicas, p. 14
DPCA	Análise Dinâmica de Componentes Principais, p. 13
DPLS	Mínimos Quadrados Parciais Dinâmico, p. 15
FDA	Análise Discriminante de Fisher, p. 14
GMM	Modelo de Mistura de Gaussianas, p. 15
HAZOP	Estudo de Perigos e Operabilidade, p. 16
ICA	Análise de Componentes Independentes, p. 13
IID	Amostras Independentes e Identicamente Distribuídas, p. 7
KDE	Estimação de Densidade por Funções Núcleo, p. 67
KICA	Análise de Componentes Independentes por Funções Núcleo, p. 15
KPCA	Análise de Componentes Principais por Funções Núcleo, p. 13
KPLS	Mínimos Quadrados Parciais por Funções Núcleo, p. 15
LOPA	Análise de Camadas de Proteção, p. 16
MLPCA	Análise de Componentes Principais por Máxima Verossimilhança, p. 15
NNPE	Estimação de Parâmetros por Redes Neurais, p. 15
PCA	Análise de Componentes Principais, p. 13
PLS	Mínimos Quadrados Parciais, p. 14
PRR	Regressão por Redução de Posto, p. 14

SDCD	Sistema Digital de Controle Distribuído, p. 22
SIF	Função Instrumentada de Segurança, p. 16
SIS	Sistema Instrumentado de Segurança, p. 16
SVDD	Descrição de Dados por Vetor Suporte, p. 15
SVM	Máquina de Vetores Suporte, p. 15

Capítulo 1

Introdução

Durante a década de 1970, houve uma série de acidentes graves em plantas petroquímicas ao redor do mundo. "Bophal" na Índia, "Sevesco" na Itália e "Flixborough" na Inglaterra foram três dos mais graves. De acordo com investigação conduzida pela Environmental Protection Agency – EPA, pela Occupational Safety and Health Administration – OSHA e pelo U.S. Chemical Safety and Hazard Investigation Board, nos acidentes ocorridos em refinarias e petroquímicas sempre estão presentes três causas recorrentes: análise de perigos de processo inadequada, utilização de equipamentos inadequados ou mal projetados e indicação inadequada de condições de processo (BOSCO, 2009). Mesmo com os enormes recursos destinados à segurança operacional, equipamentos mais confiáveis e regulamentações cada vez mais restritivas, acidentes continuam a ocorrer na indústria petroquímica.

De acordo com FEITAL (2011), um dos maiores desafios da atualidade enfrentado pelos engenheiros de processo é o desenvolvimento, implantação e manutenção de sistemas capazes de monitorar a condição de processos químicos industriais. Para KRISHNANNAIR (2010), a detecção e diagnóstico de falhas e distúrbios que impactem negativamente as plantas de processo e a qualidade dos produtos tornaram-se críticos para obtenção de excelência operacional. Ao longo dos anos essa área recebeu diferentes denominações, como "controle estatístico de processo", "detecção e isolamento de falhas", "gestão de situações anormais" e "monitoramento da condição de processo" (FEITAL, 2011). Embora as plantas de processo modernas possuam sistemas de controle extremamente sofisticados, a monitoração da condição de processos químicos ainda não se desenvolveu plenamente (SHU *et al.*, 2016).

Além dos grandes cenários acidentais, sabe-se que a maioria dos eventos de falha são não catastróficos e um maior esclarecimento quanto às causas destes eventos é fundamental para priorizar e implementar ações corretivas de manutenção, como propõe a norma ABNT-NBR-ISO-14224 (2011). Apesar desta norma se referir a coleta e intercâmbio de dados de confiabilidade e manutenção, os benefícios por ela

descritos poderiam também se estender aos sistemas de monitoração da condição de processos químicos:

[...] melhorias sustentáveis na confiabilidade, gerando maior lucratividade, confiabilidade e segurança. Outros benefícios [...] incluem melhorias na tomada de decisões, reduções de falhas catastróficas, reduções de impactos ambientais, maior eficiência em benchmarking e análise de tendências de desempenho e aumento da disponibilidade de unidades de processo.

Como exemplo prático de aplicação de sistemas de monitoração da condição, cita-se a experiência da indústria nuclear, que tem implantado e testado sistemas para monitorar a condição de processos com o intuito de alterar a estratégia de calibração de instrumentos, como pode ser observado no trecho abaixo, extraído de HINES e SEIBERT (2006).

Nas últimas duas décadas, a Nuclear Power Utilities investigou e desenvolveu métodos visando alterar sua filosofia de manutenção, antes baseada em manutenção periódica, para uma filosofia de manutenção baseada em condição. Muitos desses métodos fazem uso de novas tecnologias desenvolvidas para verificar a condição de equipamentos durante a operação da planta. Especificamente, foram desenvolvidas técnicas para monitorar a condição dos sensores e malhas de controle. Essas técnicas são comumente denominadas Métodos de Monitoração em Tempo Real (On-Line Monitoring Methods - OLM) (p. 1, tradução nossa).

Somado a todos os fatos já relatados, hoje as organizações estão diante da próxima revolução industrial - A Revolução da Indústria 4.0. Este conceito descreve a visão de uma fábrica inteligente em que toda a cadeia produtiva e seus respectivos processos possuem capacidade de se comunicar em rede e se ajustar para atender a demanda corrente, garantindo a eficiência operacional e o uso racional dos recursos ambientais. A produção será mais flexível e mais rápida, distribuída em várias fábricas totalmente integradas com auxílio de redes de comunicação. Todos esses fatores fomentarão o diálogo franco entre produtor e consumidor, permitindo satisfazer os desejos de cada cliente individualmente (FRIEDRICH-EBERT-STIFTUNG, 2017).

Nesse cenário de integração extrema, a abundância de informação disponível para análise e monitoração será muito maior que a existente hoje. É nesse contexto de comunicação integrada, tanto entre processos produtivos no chão de fábrica quanto entre produtor e consumidor, que as técnicas de detecção de falhas deverão se inserir de agora em diante. Todos estes fatores e situações mostram que é mais do que necessário o investimento em pesquisa, desenvolvimento, implantação e suporte de sistemas de monitoração de processos químicos, justificando assim a presente dissertação.

1.1 Objetivos

Há, evidentemente, uma série de desafios práticos durante a elaboração de sistemas de monitoramento, como a dificuldade de modelagem do processo ou equipamento, o sensoriamento inadequado e a integridade dos dados disponíveis (FEITAL, 2011). O objetivo principal deste trabalho é estudar, desenvolver e implementar Sistemas de Monitoramento de Falhas Baseado em Dados (*Data Driven*). Todas as simulações foram desenvolvidas no programa científico para computação numérica Scilab (SCILAB ENTERPRISES, 2017).

Objetivos secundários desta dissertação são: i) apresentar uma introdução da área de monitoramento, com foco em camadas de proteção e segurança de processos; ii) fornecer o ferramental necessário para modelagem de processos, usando somente dados registrados e técnicas *Data Driven*; iii) oferecer idéias e referências para implementações futuras de sistemas de monitoramento; e iv) apresentar novos procedimentos baseados em dados para detectar possíveis falhas de processo.

1.2 Organização

No Capítulo 2 apresentam-se uma revisão bibliográfica da área de monitoração de processos da indústria química, a classificação básica dos processos químicos e suas características, além da metodologia e requisitos de um sistema de monitoração. No Capítulo 3 encontra-se uma revisão sobre conceitos de Análise de Componentes Principais, Análise Dinâmica de Componentes Principais, Análise de Componentes Principais por Funções Núcleo e outras técnicas correlatas. Discute-se em particular como estas técnicas se inserem no problema de monitoração de processos. No Capítulo 4 investiga-se a aplicação da Análise Dinâmica de Componentes Principais para monitoração de um reator contínuo em que ocorre reação de Van der Vusse, com o objetivo de apresentar uma visão crítica sobre a técnica. No Capítulo 5 desenvolve-se uma nova metodologia para monitoramento de processos, amparada no conceito de recorrência. Finalmente, no Capítulo 6 são apresentadas as principais conclusões do trabalho e algumas sugestões para trabalhos futuros.

Capítulo 2

Revisão Bibliográfica

Este capítulo apresenta uma revisão bibliográfica sobre a área de monitoração de processos e sobre como ela se insere na filosofia de segurança operacional de uma planta industrial. O capítulo foi dividido em duas seções. Na Seção 2.1 faz-se uma revisão dos tipos de processos químicos, com apresentação das características encontradas nos dados de processo e levantamento das causas típicas de falhas em unidades industriais. Na Seção 2.2 faz-se uma apresentação sobre Metodologias de Monitoração de Processos, levantam-se requisitos que um sistema de monitoração deve atender e, para introduzir o Capítulo 3, apresenta-se o panorama da evolução da monitoração de processos por meio da análise cronológica dos principais livros publicados sobre o tema.

2.1 Processos Químicos Industriais

SEBORG *et al.* (2010) definiram processo como:

A conversão de matéria prima em produtos por meio de operações químicas e físicas. Na prática, o termo processo é empregado para designar tanto as operações de processamento, quanto os equipamentos usados no processamento (p. 2, tradução nossa).

Espera-se que um processo produza um ou mais produtos de interesse, com vazões e características especificadas, garantindo a segurança das pessoas, dos equipamentos e do meio ambiente. Embora existam várias possíveis classificações, a literatura usualmente agrupa os processos em função do tipo de operação (contínua, batelada ou semi-batelada) e em função da relação temporal apresentada pelas variáveis de processo (permanente ou transiente). FELDER e ROUSSEAU (2005) propuseram as seguintes definições:

- i) **Processo Batelada (*batch process*)** - no início do processo, a matéria prima é carregada no sistema e, somente após o término do processo, o produto é removido. Deste modo, não existe transferência de massa através das fronteiras do sistema entre o instante em que a matéria prima é carregada e o produto é removido.
- ii) **Processo Contínuo (*continuous process*)** - durante todo o tempo de processamento existe fluxo de matéria prima e produto através das fronteiras do sistema.
- iii) **Processo Semi-Batelada (*semibatch process*)** - qualquer processo que não seja classificado como batelada ou contínuo.
- iv) **Processo em Regime Permanente (*steady state process*)** - quando as variáveis de processo, como temperatura, pressão, volume, vazão, dentre outras, não variam com o tempo, exceto devido a pequenas flutuações ao redor de valores médios.
- v) **Processo em Regime Transiente (*transient process*)** - quando qualquer uma das variáveis de processo varia com o tempo.

2.1.1 Característica dos Dados da Indústria de Processo

O objetivo desta seção é descrever os principais desafios apresentados pela indústria química para a monitoração de processo. Uma planta de processo pode apresentar uma ou mais das características relacionadas abaixo e, embora nos últimos anos tenham sido realizadas várias contribuições para lidar com cada uma das características discutidas, a experiência do presente autor mostra que nenhuma das estratégias de monitoramento encontradas na literatura é capaz de tratá-las de forma integrada.

Dimensionalidade Elevada - Com o aumento da complexidade das plantas de processo, ocorreu uma rápida explosão do número de variáveis medidas e a consequente disponibilização de um grande volume de dados para os sistemas de monitoração (WANG *et al.*, 2005). Encontrar formas adequadas para manipular tal volume de informações e métodos apropriados para realizar a seleção das variáveis relevantes constitui um dos grandes desafios enfrentados pelas técnicas de monitoração baseadas em dados (GE *et al.*, 2013). A dimensionalidade elevada é uma das características da área hoje chamada de Ciência dos Dados (Big Data ou Data Science) (BELL, 2014)

Variáveis de Processo com Distribuição Não-Gaussiana - Mudanças nas vazões de carga processada, ciclos variados de enchimento e esvaziamento, distúrbios não medidos, recirculação de material ao longo das plantas de processo, entre outros fenômenos, tendem a gerar dados com distribuição não-gaussiana (LIU *et al.*, 2008). Porém, as técnicas de controle estatístico multivariável, como por exemplo a Análise de Componentes Principais (PCA), se apoiam no fato de que as variáveis de processo apresentam distribuição gaussiana para calcular as estatísticas de monitoração, bem como seus respectivos intervalos de confiança (GE *et al.*, 2013). Esse aspecto é fundamental para a compreensão de muitas das limitações das técnicas hoje existentes, havendo forte motivação para o desenvolvimento e uso de procedimentos baseados em dados de forma não paramétrica; ou seja, que não fazem uso de uma distribuição de referência (GIBBONS e CHAKRABORTI, 2003).

Não Linearidade - Todos os processos físicos reais apresentam algum grau de não linearidade (OGUNNAIKE e RAY, 1994). A maioria das técnicas convencionais de monitoração de processo baseadas em dados foi concebida para tratar sistemas lineares, de modo que a monitoração de processos fortemente não lineares constitui uma tarefa bastante complexa. Para processos moderadamente não lineares, trabalhando em uma região estreita de operação, modelos lineares ainda podem ser satisfatórios. Grande parte dos procedimentos modernos de controle e análise de dados explora as características não lineares dos processos reais (GE *et al.*, 2013).

Variante no Tempo e Com Múltiplos Modos Operacionais (Multimodo) - Quando um sistema apresenta um comportamento dinâmico que se altera ao longo do tempo, diz-se que ele é variante no tempo. A mudança de comportamento está associada a diversos fatores, como, por exemplo, a degradação de desempenho de válvulas de controle, a incrustação em trocadores de calor, a desativação de catalisadores, ou até mesmo as variações das condições ambientais (RICE *et al.*, 2010). GE *et al.* (2013) enfatizaram que variações de especificação da carga, variações lentas nos *setpoints*, envelhecimento dos diversos sistemas que compõem um processo e efeitos de sazonalidade fazem com que o ponto operacional de um processo varie constantemente. Deste modo, a monitoração de processos que variam lentamente ou que apresentem múltiplos modos operacionais constitui uma tarefa complexa, uma vez que as técnicas de monitoração convencionais admitem usualmente que a planta possui apenas um ponto de operação estável.

Variáveis de Processo Autorrelacionadas - Do ponto de vista estatístico, uma série temporal pode ser decomposta em uma componente determinística e uma componente estocástica. Cada observação pode ser caracterizada por

uma distribuição de probabilidades e o grau de influência entre observações em diferentes instantes de tempo pode ser caracterizado pela função de autocovariância ou autocorrelação (CHOUDHURY *et al.*, 2008). Na realidade, estas funções permitem a verificação da hipótese de independência, fundamental para estudos de modelagem estocástica (LE BOUDEC, 2010). Porém, a ocorrência de observações autocorrelacionadas no tempo é uma característica intrínseca da análise de muitas séries temporais, que viola a suposição básica de que as amostras são independentes e identicamente distribuídas (IID), usualmente adotada em inferência estatística. Esse talvez seja um dos pontos menos estudados e menos compreendidos da área de monitoramento de processos, sendo que a possível existência de autorrelação é raramente avaliada e ainda mais raramente usada para interpretação dos dados disponíveis.

2.1.2 Causas Comuns de Falhas em Processos Químicos

Falhas podem ser classificadas em função do tipo, comportamento temporal e alcance (ISERMANN, 2006). Quanto ao tipo, podem ser classificadas como sistemáticas (quando ocorrem de forma persistente em uma determinada condição de operação) ou aleatórias (quando ocorrem ao acaso e não podem ser associadas a uma causa específica); em termos de comportamento temporal, como abrupta (quando ocorrem de forma inesperada e rápida), permanente (quando não podem ser corrigidas e se mantêm), transiente (quando a falha muda com o tempo e com as variações do processo), intermitente (quando a falha se repete em períodos de tempo pré-determinados, de forma periódica ou aperiódica) ou incipiente (quando a falha acontece de forma gradual ao longo do tempo, causadas por processos lentos de degradação, provocando desvio crescente na condição do processo); finalmente, com relação ao alcance, como local (quando afeta uma determinada seção do processo) ou global (quando afeta todo o processo). De acordo com JELALI (2006) as causas típicas de falhas em processos podem ser agrupadas nas categorias descritas a seguir:

Projeto Inadequado de Processo - Nesta categoria estão enquadradas as questões relacionadas à concepção de processo, seleção de materiais, configuração, dimensionamento e capacidade de equipamentos. É importante, por exemplo, entender que, durante a fase de projeto, a concepção de processo e da estrutura de controle deve ser feita de forma integrada. Segundo Ziegler e Nichols (1943), conforme citado por JELALI (2006), normalmente um controlador ruim consegue um desempenho aceitável em processos facilmente controláveis; por outro lado,

o melhor controlador, utilizado em conjunto com um processo mal projetado, dificilmente atingirá o desempenho esperado. Dessa forma, o projeto apropriado do processo é fundamental para prevenção e minimização de falhas.

Falhas de Componentes e Equipamentos - Nessa categoria se enquadram panes ou mau funcionamento de sensores, atuadores e equipamentos de processo. As causas de falha podem estar relacionadas à fabricação, instalação, manutenção, serviço fora das condições de projeto ou erro de operação. Para se ter uma idéia, as válvulas de controle, que são os atuadores mais corriqueiramente encontrados na indústria de processo, podem apresentar diversos problemas, tais como subdimensionamento, superdimensionamento, corrosão do assento, defeito do diafragma, vazamento, histerese, agarramento, banda morta e saturação (CHOUDHURY *et al.*, 2008). Com relação aos sensores, os principais tipos de falha são erro de calibração, viés (quando um sensor tende a apresentar desvios sistematicamente para uma determinada região de medida), deriva e congelamento do sinal de saída (SOBHANI-TEHRANI e KHORASANI, 2009).

Sintonia Inadequada de Controladores e/ou Falta de Manutenção - Admitindo-se que o tipo de controlador está adequado para a planta, duas situações indesejadas podem acontecer e contribuir com uma falha ou mau desempenho: o controlador nunca foi sintonizado ou foi sintonizado a partir de um modelo não representativo do processo real. Além disso, frequentemente os controladores só são sintonizados durante o comissionamento da planta, não sendo ressintonizados ao longo dos anos de operação. Todavia, ao longo da vida útil de uma unidade de processo, podem acontecer alterações nas propriedades da matéria prima processada, variações nos *setpoints* para atender diferentes requisitos operacionais e envelhecimento dos diversos componentes e equipamentos, exigindo a ressintonia dos controladores.

Projeto Inadequado de Estrutura de Controle - Pareamento inapropriado de variáveis manipuladas e controladas, possíveis iterações entre variáveis, acoplamento de malhas de controle, graus de liberdade insuficientes, presença de não linearidades e tempo morto são as causas mais frequentes de problemas de estrutura de controle (JELALI, 2006). Nesse caso, a falha pode advir de incapacidade da estrutura de controle remover perturbações de forma eficiente.

2.2 Monitoração de Processos

2.2.1 Metodologia de Monitoração de Processos

Com o aumento da complexidade das plantas de processo, a necessidade de implementar estratégias de monitoração se torna cada vez mais crítica. Na realidade, a função primária dos sistemas de monitoração é auxiliar o pessoal responsável pela operação a compreender e identificar o estado atual do processo e alertá-los em caso de padrões de operação indesejados ou não esperados (sintomas), de forma a evitar possíveis panes ou mau-funcionamentos (KÜHNERT, 2013). A Figura 2.1 ilustra com auxílio de diagramas de bloco como a função de monitoração de processo se insere em uma planta industrial. O sistema de monitoração recebe informações dos sensores da planta e dos controladores e auxilia a operação nas tomadas de decisão. Como nem sempre é possível observar uma falha diretamente, a determinação da falha muitas vezes é feita por meio da análise dos sintomas observados.

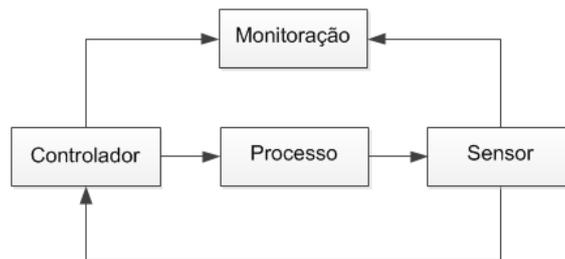


Figura 2.1: Conexão entre Processo e Monitoração

Conforme sugerido pela literatura, o termo Monitoração de Processos envolve quatro procedimentos básicos: detecção, identificação, diagnóstico e restabelecimento (RUSSELL, 2012). A Figura 2.2 ilustra a sequência de ativação de cada um destes procedimentos.

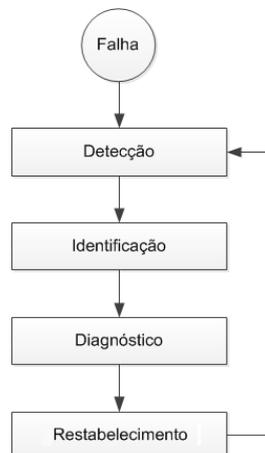


Figura 2.2: Fluxograma das Etapas de Monitoração de Processos

A detecção da falha consiste em determinar o mais rapidamente possível se existe ou não uma falha presente no processo. A detecção precoce é fundamental para garantir que o pessoal de operação e acompanhamento das plantas tenha tempo hábil para tomar as ações corretivas necessárias, evitando perdas de produção e acidentes. Tão logo uma falha tenha sido acusada, deve ser feita uma busca para identificar as variáveis de processo que estejam mais relacionadas à falha ocorrida, auxiliando a operação a focar a atenção nos subsistemas mais prováveis de serem os causadores do problema. Atinge-se assim a etapa de diagnóstico, cuja preocupação é determinar os tipos de falhas que ocorreram. É a partir desta etapa que são determinadas as ações corretivas ou mitigadoras para eliminar ou reduzir os efeitos da falha. Na etapa de restabelecimento, as ações determinadas durante o diagnóstico são efetivamente implementadas, com o objetivo de remover os efeitos causados pela falha. É importante destacar que nem sempre todas estas etapas são implementadas na prática.

Embora não exista total concordância entre os pesquisadores da área sobre como agrupar as diversas técnicas de monitoração de processo, a Figura 2.3 apresenta a classificação mais comumente encontrada na literatura (CHEN, 2017; DING, 2014; VENKATASUBRAMANIAN *et al.*, 2003a,b,c).

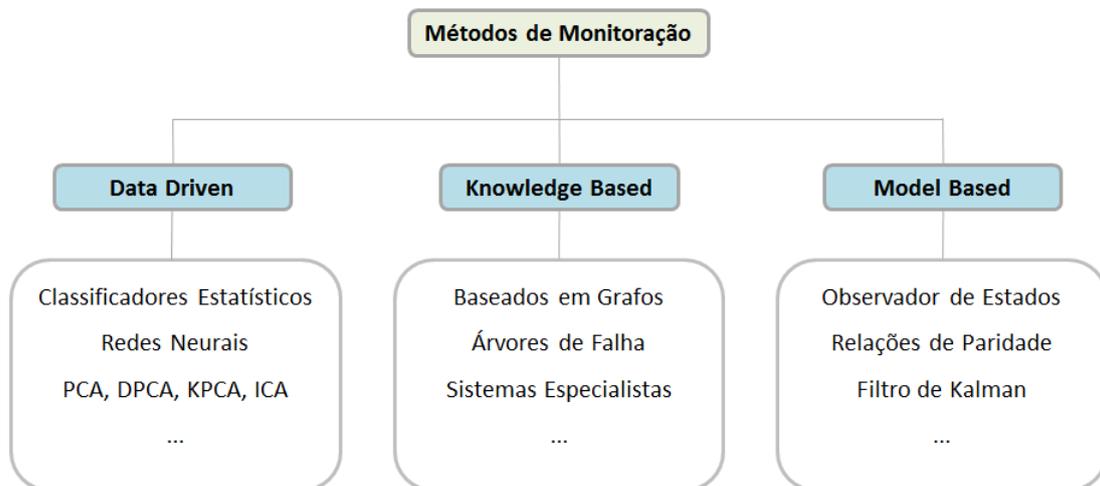


Figura 2.3: Classificação dos Métodos de Monitoração

Adicionalmente, DING (2014) e CHEN (2017) propuseram outras duas categorias para classificar as técnicas de monitoração de processo:

- Redundância Física - quando são usados um ou mais componentes de *hardware* idênticos e uma falha é detectada quando se verificam diferenças entre os sinais de saída destes componentes. Este esquema apresenta alta confiabilidade e permite fácil identificação e diagnóstico da falha; porém, apresenta maior custo;

- Técnicas de Processamento de Sinais - que admitem que os sinais de processo carregam informações sobre as falhas e que tal informação se manifesta por meio de sintomas, de maneira que o processamento adequado destes sinais pode revelar a existência de uma falha.

Modelos computacionais baseados em modelos de processo podem ser usados para projeto de equipamentos, otimização de unidades, monitoramento da condição de processo, treinamento de operadores, desenvolvimento de estratégias de controle, estudo de procedimentos de partida, parada e resposta a emergências, dentre outros (SKOGESTAD, 1991). Os métodos baseados em modelo (ou *Model Based* na Figura 2.3) se apoiam na modelagem por primeiros princípios, fundamentada em equações matemáticas de conservação que refletem o conhecimento dos fenômenos químicos, físicos e biológicos disponíveis acerca do processo de interesse. Nesta abordagem, o monitoramento é realizado por meio do acompanhamento dos parâmetros de ajuste do modelo fenomenológico por estimação de parâmetros em tempo real (CÂMARA *et al.*, 2017; COIMBRA *et al.*, 2017; PINTO *et al.*, 2017; SOUZA *et al.*, 2011). A vantagem deste tipo de modelo é que, através deles, é possível compreender o comportamento do processo, já que os modelos usualmente são aplicáveis em largas faixas de operação; porém, o desenvolvimento do modelo pode ser caro e demorado (SEBORG *et al.*, 2010). Tais desvantagens despertaram o interesse da indústria e da comunidade científica para o estudo e desenvolvimento dos métodos baseados em dados (ou *Data Driven* na Figura 2.3), que englobam métodos de processamento de sinais, modelagem estatística e de aprendizagem de máquina.

Apesar de fazerem parte da categoria *Data Driven*, os métodos estatísticos nada mais são que modelos estatísticos estudados em disciplinas de Análise de Dados e Modelagem Multivariada. Os modelos estatísticos mais conhecidos são a Análise de Componentes Principais (PCA) e Métodos Correlatos, a Análise Discriminante Linear ou Análise Discriminante de Fisher (FDA), a Análise de Componentes Independentes (ICA), a Análise de Correlação Canônica (CVA), a Regressão por Redução de Posto (PRR) e os Mínimos Quadrados Parciais (PLS) (KRÜGER e XIE, 2012). As três últimas técnicas buscam encontrar relações de causa e efeito entre as variáveis, sendo denominadas de dois blocos (*dual-block*), já que elas dividem os dados amostrais em um conjunto de preditores, ou variáveis de causa, e em um conjunto de variáveis resposta, ou variáveis de efeito (KRÜGER e XIE, 2012). Em oposição, os três primeiros métodos são ditos de bloco único (*single-block*). Dentre estas técnicas, as que mais têm recebido atenção na literatura são a Análise de Componentes Principais e Métodos Correlatos e o método dos Mínimos Quadrados Parciais.

A Análise de Componentes Principais (PCA) é provavelmente a técnica de análise multivariada mais antiga e mais conhecida (JOLLIFFE, 2002). Introduzida

por Pearson em 1901, ela só foi massivamente usada após o desenvolvimento dos computadores, assim como muitas outras técnicas multivariadas. Da mesma maneira que outras técnicas de análise multivariada, a técnica de PCA considera duas ou mais variáveis aleatórias como uma entidade única e tenta produzir um resultado global, considerando o grau de relacionamento entre as variáveis (JACKSON, 1991). Talvez as variantes mais conhecidas da Análise de Componentes Principais sejam sua versão dinâmica (KU *et al.*, 1995) e sua versão não linear (SCHÖLKOPF *et al.*, 1998); entretanto, existe uma infinidade de métodos cuja origem é a Análise de Componentes Principais, sendo que muitas publicações científicas são dedicadas exclusivamente ao tema.

A Análise Dinâmica de Componentes Principais (DPCA) surgiu para tratar a questão de variáveis correlacionadas no tempo, já que os métodos mencionados anteriormente são incapazes de lidar com esta característica, apresentando um grande número de falsos alarmes (KU *et al.*, 1995). A abordagem adotada na técnica de DPCA é similar à identificação de um modelo auto-regressivo integrado de médias móveis (ARIMA), ajustado aos dados de uma série temporal (DE KETELAERE *et al.*, 2015). A técnica PCA e outros métodos correlatos ainda recebem muita atenção de pesquisadores e acadêmicos graças à facilidade de implementação e ao baixo poder de processamento requerido para a execução dos algoritmos. Além disso, existe grande interesse no desenvolvimento de técnicas para ajuste dos parâmetros do modelo, como por exemplo a determinação do atraso das variáveis usadas na técnica de DPCA (VANHATALO *et al.*, 2017). Maiores detalhes sobre essas técnicas serão apresentados no Capítulo 3.

Mais recentemente, técnicas como a Análise de Componentes Independentes (ICA), Modelos de Mistura de Gaussiana (GMM), Máquinas de Vetor de Suporte (SVM), Descrição de Dados por Vetor Suporte (SVDD) e Métodos de Agrupamento de Dados também têm chamado a atenção da comunidade científica como alternativas para a monitoração de processos, como pode ser observado a partir dos trabalhos de FEITAL *et al.* (2013), GE *et al.* (2013) e DENG e TIAN (2014). A Tabela 2.1, adaptada a partir de GE *et al.* (2013), apresenta alguns dos métodos de monitoramento *Data-Driven* agrupados em função das características de processo, discutidas na Seção 2.1.1.

Tabela 2.1: Relação entre Método de Detecção e Características de Processo. Adaptado de GE *et al.* (2013)

Não-Gaussianidade	Não Linearidade	Multimodo	Dinâmico
ICA	KPCA	PCA Adaptativo	DPCA
GMM	KPLS	PLS Recursivo	DPLS
SVDD	KICA	PCA de Janela Móvel	Séries Temporais
Estimação de Densidade	Curvas Principais	PCA Multi-Modelo	PCA no Espaço de Estados
SVM de Uma Classe	Redes Neurais	Combinação Bayesiana	Multiway DPCA
Multiway ICA	SVDD	MLPCA	Multiway DPLS
Multiway GMM	Multiway KPCA	MLFA	ICA no Espaço de Estados
Abordagem Local	Multiway KICA	Aprendizagem Local	Multiway NNPE
	Subespaço Linear	External Analysis	CVA
	GMM	Métodos Robustos	Métodos 2D
	Métodos Bayesianos		
	Multiway GMM		
	Kernel FDA		

Na Tabela 2.1 são apresentados os principais métodos, mas certamente esta é uma área em constante evolução e novas ferramentas de monitoração têm surgido continuamente ao longo dos anos. Uma desvantagem dos modelos *Data Driven* já relatada na literatura está relacionada à baixa capacidade de extrapolação. Com o intuito de tornar os métodos adaptativos, muito esforço tem sido realizado, conforme pode ser observado nos trabalhos de WANG *et al.* (2005), JIN *et al.* (2006), HE e YANG (2008), JENG (2010), KADLEC *et al.* (2011) e CHOUAIB *et al.* (2013). O termo adaptativo é empregado aqui de forma genérica, já que, com a capacidade computacional disponível hoje, é muito provável que seja possível reconstruir o modelo em tempo real em várias aplicações, ao invés de simplesmente adaptá-lo.

Para realizar a monitoração de processo a partir dos métodos de modelagem *Data Driven*, é conveniente construir uma carta de controle. Tais cartas foram desenvolvidas por Walter A. Shewhart nos Laboratórios da Bell em 1924 e definem uma área denominada de limite de controle, delimitada por uma linha superior, o limite superior de controle (LSC), e por uma linha inferior, o limite inferior de controle (LIC), além de uma linha central (LC) (INSTITUTO ADOLFO LUTZ, 2013). Desde que os pontos amostrais estejam dispostos entre os limites de controle de forma aleatória, considera-se que o processo está sob controle. Se um ou mais pontos estiverem dispostos fora dos limites de controle, há evidência estatística de que o processo está "fora de controle" e que investigação e ações corretivas devem ser tomadas para detectar e eliminar as causas.

De uma maneira geral, nenhum dos métodos de monitoração é completamente eficaz se for implementado de forma isolada. Por isso, geralmente deve ser utilizado um sistema de monitoração híbrido, empregando várias técnicas de monitoração de forma hierárquica, para alcançar um monitoramento de processo mais efetivo. A Figura 2.4 exemplifica a concepção de um sistema de monitoração que segue esta filosofia (JOHNSON CONTROLS, 2013). Na Figura 2.4, o sentido da seta indica aumento de complexidade das estratégias.



Figura 2.4: Hierarquia de Monitoração de Processos. Adaptado de JOHNSON CONTROLS (2013)

O nível mais básico de detecção está fundamentado na monitoração com auxílio de alarmes, cujo objetivo é chamar a atenção dos operadores para faltas consideradas críticas do ponto de vista de eficiência e segurança. Usualmente, mas não necessariamente, associado ao Sistema de Alarmes existe um Sistema Instrumentado de Segurança (SIS), projetado para responder automaticamente nas condições de operação da planta que podem ser perigosas, caso nenhuma ação seja tomada. No SIS são implementadas uma ou mais Funções Instrumentadas de Segurança (SIF), com a função de garantir ou manter um estado seguro de um processo ou equipamento por meio de ações automáticas específicas frente a uma determinada falha. Não necessariamente para cada alarme existe uma SIF associada, já que tal decisão depende do Estudo de Perigos e Operabilidade (HAZOP) e da Análise de Camadas de Proteção (LOPA). Para maiores detalhes sobre as técnicas de HAZOP e LOPA sugere-se consultar os trabalhos de MARSZAL e SCHARPF (2002) e AMERICAN INSTITUTE OF CHEMICAL ENGINEERS (2001).

No caso da indústria de processos, o ciclo de vida de um SIS deve seguir o procedimento prescrito pela IEC-61511 (IEC-61511, 2017), que trata da especificação, projeto, instalação, operação e manutenção do SIS, buscando garantir que o sistema esteja operacional em caso de uma demanda de segurança. Esta norma foi desenvolvida como uma implementação específica para o setor de processos da IEC-61508 (IEC-61508, 2010). Por sua vez, a monitoração de processo por meio de alarmes está consolidada e conta com a norma ANSI/ISA-18.2, que estabelece um fluxo de trabalho para projetar, implementar, operar e manter um sistema de alarme moderno. A ANSI/ISA-18.2 define alarme como qualquer meio auditivo ou visual que indique uma condição anormal associada à operação do processo ou do equipamento e que exige uma ação em um tempo restrito. O objetivo básico da norma é aumentar a segurança operacional, uma vez que já foi observado que sistemas de alarme deficientes contribuíram para acidentes no passado. Para atender à definição, um alarme deve ser projetado considerando-se as limitações humanas e deve apresentar as seguintes características (MOTEIRO *et al.*, 2011):

- **Relevância** - deve ter importância operacional definida; se nenhuma resposta está associada ao sinal gerador do alarme, este sinal não deve ser um alarme;
- **Singularidade** - uma mesma informação não deve ser representada por dois

diferentes alarmes, evitando duplicidade de procedimentos de resposta que pode confundir e sobrecarregar o operador;

- **Tempo de Resposta Adequado** - nenhum alarme deve ser apresentado com muita antecedência em relação a sua resposta ou muito tarde para que uma medida corretiva seja executada;
- **Grau de Importância** - todo alarme deve possuir uma prioridade, facilitando assim a tomada de decisões do operador;
- **Clareza** - a mensagem do alarme deve ser de fácil compreensão e chamar atenção para as informações mais importantes que se deseja transmitir.

No próximo nível da hierarquia de monitoração estão as regras, que podem ser entendidas como alarmes avançados, e cujo objetivo é identificar deficiências não críticas de processo (JOHNSON CONTROLS, 2013). Assim, embora o processo ainda se mantenha em operação, a eficiência operacional fica comprometida. A vantagem das regras reside no fato de que elas podem representar tanto princípios fundamentais quanto lógicas e experiência dos especialistas de processo. Um sistema de regras é relativamente simples de implementar, o que favorece o custo. Suas principais limitações são: baixa capacidade de adaptação; dificuldade de gerenciamento das regras, caso sejam necessárias muitas regras para englobar todo o espectro de falhas; falhas só podem ser detectadas após sua ocorrência; grande número de falsos alarmes, em caso de regras mal concebidas (JOHNSON CONTROLS, 2013).

Ao contrário dos alarmes e das regras, que apresentam apenas dois estados, ativo ou inativo, os índices de desempenho indicam de forma contínua o quão distante um processo está do seu ponto de operação ótimo. Para facilitar a compreensão do pessoal de operação, podem ser criados intervalos indicando, por exemplo, se o sistema está em operação normal, se existe um desvio considerável do ponto ótimo de operação ou se ações de correção devem ser tomadas. Índices de desempenho de malhas de controle, equipamentos e sistemas podem ser combinados para fornecer uma visão geral do estado do processo. Os índices de desempenho são usados, por exemplo, nos sistemas de avaliação de malhas de controle, com pode ser observado nos trabalhos de KEMPF (2003) e JELALI (2006). A Figura 2.5 ilustra como os índices de desempenho variam ao longo do tempo.

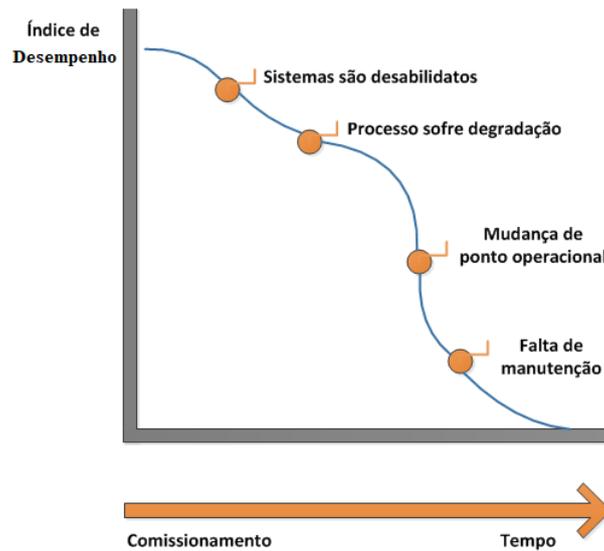


Figura 2.5: Variação dos Índices de Desempenho ao Longo do Tempo. Adaptado de JOHNSON CONTROLS (2013)

No topo da estrutura estão os sistemas de monitoração de processo que utilizam diversos tipos de análise estatística. Nesta categoria podem ser usadas três abordagens diferentes (JOHNSON CONTROLS, 2013). A primeira abordagem baseia-se na comparação estatística entre os dados atuais e dados operacionais históricos. A vantagem desta abordagem se deve ao fato de que variações que ocorrem ao longo do tempo podem ser detectadas dentro de um limite de confiança. Porém, falhas que sempre estiveram presentes no processo não serão detectadas. Deste modo, a unidade de processo deve ser comissionada antes da aquisição dos dados históricos de referência e é interessante que um especialista de processo auxilie na tarefa de seleção e validação de dados. A segunda abordagem é conhecida como avaliação entre pares e consiste em comparar sistemas ou equipamentos que desempenhem serviços semelhantes. A terceira abordagem é a mais complexa, por envolver modelagem estatística. Funciona relativamente bem para sistemas de grande escala, já que a planta de processo é tratada de forma integrada; contudo, falhas que sempre estiveram presentes no sistema não serão detectadas.

Do ponto de vista de segurança operacional, o sistema de monitoramento de processo pode ser compreendido como mais uma camada na estratégia de proteção da planta industrial (MOTEIRO *et al.*, 2011), como apresentado na Figura 2.6. O objetivo primário é proteger a planta, as pessoas e o meio ambiente da liberação de material tóxico ou inflamável, evitando que se atinja um cenário acidental, a partir do qual não há retorno à condição normal de operação (MOTEIRO *et al.*, 2011). Cada uma destas camadas se sobrepõe à camada anterior, formando um conjunto de barreiras. As camadas iniciais são preventivas, projetadas para evitar a liberação de material. Caso não seja possível evitar a falha, cabe ao sistema de mitigação e

ao plano de resposta à emergência reduzir os impactos causados pela liberação de material. Neste contexto, o sistema de monitoramento, seguido de uma ação do operador, é encarado como uma camada preventiva e deve ser o mais independente possível das demais camadas, para evitar as falhas de modo comum.

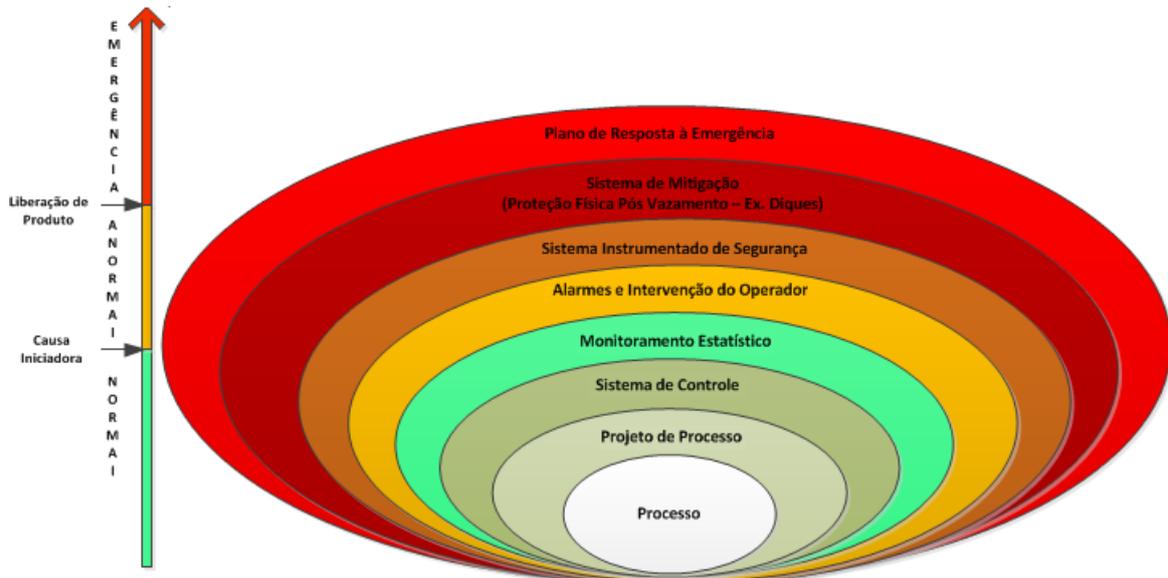


Figura 2.6: Estratégia de Proteção da Planta Industrial. Adaptado de MOTEIRO *et al.* (2011)

Como o pessoal responsável pela operação das plantas de processo desempenha um série de tarefas concomitantemente, é importante que existam meios para se priorizar o tratamento de falhas. Tal priorização poderia, por exemplo, ser feita a partir da classificação das consequências de uma falha, conforme apresentado na Tabela 2.2, extraída da ABNT-NBR-ISO-14224 (2011). É importante perceber que na determinação da prioridade são ponderados diferentes aspectos relacionados à eficiência operacional: segurança, meio ambiente, perdas de produção e danos a equipamentos e instalações.

Tabela 2.2: Classificação de Consequências de Falha. Adaptado de ABNT-NBR-ISO-14224 (2011)

	Catastrófica	Severa	Moderada	Menor
	Falha que resulta em morte ou perda no sistema	Severa lesão, doença ou dano no sistema principal	Menor lesão, doença ou dano no sistema	Inferior à menor lesão, doença ou dano no sistema
Segurança	I - Perda de vidas - Sistemas críticos à segurança vital inoperantes	V - Lesões pessoais graves - Potencial de perda de funções de segurança	X - Lesões que requerem tratamento médico - Efeito limitado nas funções de segurança	XIII - Lesões que não requerem tratamento médico - Efeito menor sobre a função de segurança
Ambiental	II - Poluição maior	VI - Poluição significativa	X - Alguma poluição	XIV - Poluição inexistente ou desprezível
Produção	II - Parada extensa na produção/operação	VII - Parada na produção acima do nível aceitável	XI - Parada na produção abaixo do nível aceitável	XV - Pequena parada na produção
Operacionais	IV - Custo de manutenção muito alto	VIII - Custo de manutenção acima do normalmente aceitável	XII - Custo de manutenção abaixo do normalmente aceitável	XVI - Custo de manutenção baixo

2.2.2 Requisitos de um Sistema de Monitoramento

Ao projetar um sistema de monitoração de processos, é importante ter em mente alguns requisitos de referência. Tais requisitos também são importantes, já que permitem comparar o desempenho de sistemas concebidos com diferentes arquiteturas. Os requisitos apresentados aqui seguem as recomendações de VENKATASUBRAMANIAN *et al.* (2003c).

Detecção e Diagnóstico Precoce - A detecção precoce é fundamental para garantir que o pessoal de operação e acompanhamento das plantas tenha tempo hábil para tomar as ações corretivas necessárias, evitando perdas de produção e acidentes. Porém, existe um compromisso entre falsos alarmes e detecção precoce.

Isolabilidade - Deseja-se de forma geral que um sistema de monitoramento tenha capacidade de distinguir entre diversos tipos de falhas. Aqui existe um compromisso entre a capacidade de distinguir entre falhas e rejeitar incertezas de modelagem, uma vez que as variáveis medidas apresentam ruído superposto ao sinal do processo e o modelo tem precisão finita, podendo induzir conclusões equivocadas sobre a origem do problema observado.

Robustez - O sistema de monitoramento deve ter capacidade de detectar falhas mesmo na presença de ruído e incertezas, inevitáveis durante a operação real.

Identificação de Novas Falhas - A função primária de um sistema de monitoração é determinar, dada as condições atuais de operação, se o processo está operando em regime normal. Em caso de um regime anormal de operação, o sistema de monitoração deve ter capacidade de distinguir entre falhas existentes e novas falhas.

Estimativa de Nível de Confiança - É importante definir uma métrica que indique o nível de confiança do diagnóstico fornecido pelo sistema de monitoramento, para que seja possível avaliar a relevância do diagnóstico proposto.

Adaptabilidade - Como já discutido na Seção 1.2.4, os processos são variantes e podem apresentar múltiplos modos operacionais. Além disso, as unidades de processo estão sujeitas a variações das condições ambientais ou mesmo atualizações de equipamentos. Deste modo, o sistema de diagnóstico deve ser capaz de se adaptar às mudanças.

Capacidade de Explicar o Diagnóstico - Deseja-se que o sistema de monitoração forneça alguma explicação de como a falha se originou e se propagou até atingir o estado atual.

Esforço de Modelagem - Para garantir que o sistema de monitoramento esteja em operação rapidamente, o esforço de modelagem deve ser o menor possível, sem, no entanto, comprometer a qualidade e robustez da análise.

Capacidade de Identificar Múltiplas Falhas - Várias falhas podem estar presentes simultaneamente em um processo, de forma que é importante ter capacidade de identificá-las; entretanto, este é um problema complexo, já que existem interações entre as diferentes falhas.

2.2.3 Panorama de Evolução da Monitoração de Processo

A Tabela 2.3 apresenta uma lista em ordem cronológica dos principais livros publicados sobre o tema de monitoração de processos. Ao que parece, os primeiros trabalhos sobre o tema eram baseados em modelagem fenomenológica; porém, hoje existe uma tendência, cada vez mais forte, de desenvolver e implementar técnicas para monitoração *Data Driven*. Na indústria em geral, e em particular na indústria de processo, o aumento do poder computacional fez com que os sistemas digitais de controle distribuído (SDCD) ganhassem rápida penetração no chão de fábrica. Associado a eles, a indústria viu surgir uma série de padrões de comunicação para redes de campo, que se polpularizaram rapidamente, fazendo o número de pontos de medição crescer enormemente, gerando enormes quantidades de dados que geralmente são armazenados nos sistemas historiadores de processo. Assim, é natural indagar se seria possível extrair conhecimento a partir de dados históricos.

O interesse pelas técnicas *Data Driven* também vem sendo impulsionado pelo movimento de Transformação Digital. Na indústria do futuro, as redes de comunicação presentes no chão de fábrica estarão totalmente integradas às redes corporativas, conectando a cadeia produtiva à estrutura gerencial das organizações. Dadas as velocidades de comunicação cada vez maiores, além da disponibilidade de grande quantidade de dados históricos, será possível acessar o chão de fábrica em tempo real. Num mundo de competição extrema, as empresas serão forçadas a fazer uso de tecnologia para melhorar seu desempenho operacional. A tecnologia terá um papel essencial nas organizações, mas isso só será possível na medida em que existam sistemas inteligentes confiáveis e capazes de extrair informação útil a partir de dados, de modo a auxiliar o pessoal responsável pela supervisão, manutenção e estratégia organizacional nas tomadas de decisão.

Observando a Tabela 2.3, é impossível não perceber o caráter multidisciplinar da área, justificando a abundância de livros e artigos que existem sobre o tema. Outra tendência claramente identificável está relacionada ao interesse crescente por estudos relacionados ao conceito de tolerância a falha. Somente entre os anos de 2013 e 2016 surgiram quatro novas publicações sobre o tema. A partir do momento em que é possível identificar prematuramente a presença de uma falha em um determinado sistema, é possível pensar e planejar ações para restabelecimento do modo normal de operação. Assim, os sistemas de controle de processo seriam flexíveis o suficiente

para permitir sua adaptação ou reconfiguração para restaurar o padrão normal de operação, evitando riscos operacionais e perdas monetárias.

Entretanto, é paradoxal e frustrante observar que, apesar da vasta literatura existente sobre monitoração de processo, do enorme esforço gasto na concepção de novos métodos de monitoração e da quantidade massiva de dados de processo disponíveis em historiadores ou em tempo real, tais técnicas ainda não atingiram seu potencial pleno. Esta constatação foi uma das principais motivações para o desenvolvimento do presente trabalho. Buscando investigar as causas que têm impedido as técnicas de monitoração *Data Driven* de atingirem uma maior aplicabilidade na indústria e levando em conta as restrições de tempo para o desenvolvimento da pesquisa, as investigações foram realizadas a partir das técnicas baseadas na Análise de Componentes Principais.

As técnicas baseadas em Análise de Componentes Principais são consideradas clássicas na área de monitoração *Data Driven*, apresentando baixo requisito computacional, sendo de implementação relativamente rápida e estando disponíveis em grande quantidade de trabalhos sobre o tema. Acredita-se que esta escolha em particular não seja um limitador para o trabalho, já que várias outras técnicas *Data Driven* são semelhantes à técnica de PCA. Por exemplo, para o aprendizado da grande maioria de técnicas *Data Driven*, é necessário que esteja disponível um conjunto de dados para treinamento que reflita as condições normais de operação da planta industrial, sendo que para realizar a detecção de falhas, geralmente é usado algum tipo de estatística. Para dar prosseguimento ao assunto, no Capítulo 3 é desenvolvida uma breve revisão sobre Análise de Componentes Principais, técnicas correlatas e sobre as principais estatísticas usadas para o estabelecimento dos limiares de detecção de falhas.

Tabela 2.3: Principais Livros Publicados Sobre Monitoração de Processos em Ordem Cronológica

Data	Referência	Título
1978	(HIMMELBLAU, 1978)	- Fault Detection and Diagnosis in Chemical and Petrochemical Processes.
1993	(BASSEVILLE e NIKIFOROV, 1993)	- Detection of Abrupt Changes: Theory and Application.
1994	(SIMPSON, 1994)	- SystemTest and Diagnosis.
1998	(GERTLER, 1998)	- Fault Detection and Diagnosis in Engineering.
1999	(CHEN e PATTON, 1999)	- Robust Model-Based Fault Diagnosis for Dynamic Systems.
2000	(GUSTAFSSON, 2000)	- Adaptative Filtering and Change Detection.
	(NARASIMHAN e JORDACHE, 2000)	- Data Reconciliation and Gross Error Detection: An Intelligent Use of Process Data.
	(PATTON <i>et al.</i> , 2000)	- Issues of Fault Diagnosis for Dynamic Systems.
2001	(CHIANG <i>et al.</i> , 2001)	- Fault Detection and Diagnosis in Industrial Systems.
2004	(CAMPBELL e NIKOUKHAH, 2004)	- Auxiliary Signal Design for Failure Detection.
2006	(BLANKE e SCHRÖDER, 2006)	- Diagnosis and Fault-Tolerant Control.
	(ISERMANN, 2006)	- Fault-Diagnosis Systems: An Introduction from Fault Detection to Fault Tolerance Diagnosis.
2008	(CHOUDHURY <i>et al.</i> , 2008)	- Diagnosis of Process Nonlinearities and Valve Stiction: Data Driven Approaches.
2009	(SOBHANI-TEHRANI e KHO-RASANI, 2009)	- Fault Diagnosis of Nonlinear Systems Using a Hybrid Approach.

Tabela 2.3 – Continuação da Página Anterior

Data	Referência	Título
	(NOURA, 2009)	- Fault-Tolerant Control Systems: design and practical application
2011	(ALWI <i>et al.</i> , 2011)	- Fault Detection and Fault-Tolerant Control Using Sliding Modes.
	(MESKIN e KHORASANI, 2011)	- Fault Detection and Isolation: Multi-Vehicle Unmanned Systems.
	(POULIEZOS e STAVRAKAKIS, 2011)	- Real Time Fault Monitoring of Industrial Processes.
2012	(KRÜGER e XIE, 2012)	- Statistical Monitoring of Complex Multivariate Processes: With Applications In Industrial Process Control.
	(RUSSELL, 2012)	- Data-Driven Methods for Fault Detection and Diagnosis In Chemical Processes.
	(SAM Y e GU, 2012)	- Fault Detection and Flight Data Measurement: Demonstrated on Unmanned Air Vehicles Using Neural Networks.
2013	(ALDRICH e AURET, 2013)	- Unsupervised Process Monitoring and Fault Diagnosis With Machine Learning Methods .
	(GE e SONG, 2013)	- Multivariate Statistical Process Control: Process Monitoring Methods and Applications.
	(KÜHNERT, 2013)	- Data-Driven Methods for Fault Localization In Process Technology.
	(MARTINEZ-GUERRA, 2013)	- Fault Detection and Diagnosis In Nonlinear Systems: A Differential and Algebraic Viewpoint.
	(ZOLGHADRI, 2013)	- Fault Diagnosis and Fault-Tolerant Control and Guidance for Aerospace Vehicles.
2014	(DING, 2014)	- Data-Driven Design of Fault Diagnosis and Fault-Tolerant Control Systems.
	(WITCZAK, 2014)	- Fault Diagnosis and Fault-Tolerant Control Strategies for Non-Linear Systems.
2016		

Tabela 2.3 – Continuação da Página Anterior

Data	Referência	Título
	(LI, 2016)	- Fault Detection and Fault-Tolerant Control for Nonlinear Systems.
2017	(CHEN, 2017)	-Data-Driven Fault Detection for Industrial Processes: Canonical Correlation Analysis and Projection Based Methods.

2.3 Comentários Finais

A área de Monitoramento de Processos constitui uma disciplina de caráter multidisciplinar, estando basicamente sob o domínio das áreas de modelagem fenomenológica, modelagem estatística e aprendizagem de máquina. O aumento da complexidade das plantas de processo e a crescente pressão que a indústria tem sofrido para melhoria da eficiência operacional têm impulsionado o desenvolvimento e a implantação de novas técnicas de monitoramento de processo. Graças à disponibilidade de dados históricos e de dados em tempo real, os trabalhos voltados para técnicas puramente *Data Driven* têm recebido crescente atenção.

Devido à complexidade das plantas de processo atuais, a experiência acumulada pelo autor indica que nenhuma das técnicas de monitoramento, implementadas de forma isolada, seja completamente eficaz. Assim, é interessante implementar sistemas híbridos, baseados em várias estratégias de monitoramento. Sendo possível detectar uma falha, torna-se também factível conceber mecanismos de tolerância, para restabelecimento do modo normal de operação da planta. Esta parece ser uma das novas tendências da área de monitoramento de processo.

Com o objetivo de compreender melhor a aplicabilidade e as limitações das técnicas clássicas de monitoramento de processos, no Capítulo 3 discute-se a técnica de Análise de Componentes Principais e algumas variantes mais conhecidas. No final do Capítulo 3 apresenta-se um panorama geral de algumas técnicas mais modernas de análise de subespaços, candidatas ao desenvolvimento de sistemas de monitoramento de processos.

Capítulo 3

Análise de Componentes Principais

Este capítulo apresenta os fundamentos da técnica de Análise de Componentes Principais (PCA) e traça um panorama de métodos correlacionados que foram desenvolvidos a partir das ideias oriundas desta técnica. O capítulo está dividido em seções que descrevem técnicas baseadas em PCA Clássico ou Estático, PCA Dinâmico, PCA Não-Linear, Outras Variantes de PCA e Outras Técnicas de Análise de Subespaços.

3.1 PCA Clássico ou Estático

3.1.1 Geral

Com o objetivo de reduzir a dimensionalidade de uma massa de dados, a técnica de PCA busca extrair um conjunto de componentes descorrelacionadas entre si, denominadas componentes principais (PCs) (VANHATALO *et al.*, 2017). Na realidade, as componentes principais são uma combinação linear das variáveis originais. A formulação matemática deste problema pode seguir uma abordagem estatística ou geométrica (VIDAL *et al.*, 2016).

Admitindo que as variáveis medidas em um processo são o resultado de uma mistura de sinais fundamentais independentes e não-observados deste mesmo processo, de suas vizinhanças e do próprio sistema de medição, o objetivo fundamental da técnica é encontrar os sinais fundamentais a partir da mistura (FEITAL, 2011). As variáveis medidas podem ser agrupadas em uma matriz \mathbf{Y} que contém N amostras de cada uma das M variáveis de processo e representada como,

$$Y = \begin{bmatrix} y_1 & \dots & y_M \end{bmatrix} \in R^{N \times M} \quad (3.1)$$

em que $\mathbf{y}_i \in \mathbf{R}^N$, $i = 1, 2, \dots, M$ são vetores coluna que contêm as N amostras de cada uma das M variáveis de processo. Definindo o vetor linha $\bar{\mathbf{Y}}$, que contém as médias \bar{y}_i das M variáveis de processo, como apresentado em na Equação (3.2), e o vetor coluna $\mathbf{1}_N \in \mathbf{R}^N$, como apresentado na Equação (3.3), é possível escrever a matriz de desvios \mathbf{X} , formada pelos vetores de desvio $\mathbf{x}_i \in \mathbf{R}^N$, $i = 1, 2, \dots, M$ associados a cada uma das M variáveis originais, como na Equação (3.4):

$$\bar{Y} = [\bar{y}_1 \quad \dots \quad \bar{y}_M] \in R^M \quad (3.2)$$

$$\mathbf{1}_N = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \in R^N \quad (3.3)$$

$$X = Y - \mathbf{1}_N \bar{Y} = [x_1 \quad \dots \quad x_M] \in R^{N \times M} \quad (3.4)$$

Informações relativas ao grau de associação que existe entre as variáveis de processo podem ser obtidas com auxílio das estatísticas de covariância ou correlação, calculadas a partir da matriz \mathbf{X} . Define-se então a matriz de covariância amostral deste conjunto de dados como,

$$S_{xx} = \frac{X^T X}{N - 1} \sim \Sigma_{xx} \in R^{M \times M} \quad (3.5)$$

sendo Σ_{xx} a matriz de covariância populacional. Seguindo uma abordagem estatística, a análise de componentes principais utiliza um critério de máxima variância para encontrar uma nova base para o espaço vetorial da matriz de observações (\mathbf{X}). Em última análise, a pergunta que se quer responder é se existe outra base, que seja uma combinação linear dos dados originais, capaz de melhor representar os dados. Trata-se, então, de uma mudança de base realizada por intermédio de uma projeção das variáveis observadas como descrito pelas seguintes equações,

$$t_j = P^T (x_j^L)^T \in R^d \quad (3.6)$$

em que $\mathbf{t}_j \in \mathbf{R}^d$, $j = 1, 2, \dots, N$ é um vetor coluna resultante da transformação do vetor coluna $(\mathbf{x}_j^L)^T \in \mathbf{R}^M$, $j = 1, 2, \dots, N$ formado pela transposta da j -ésima linha

da matriz de desvios \mathbf{X} , $\mathbf{P}^T = \begin{bmatrix} p_1^T \\ \vdots \\ p_d^T \end{bmatrix} \in \mathbf{R}^{d \times M}$ é a matriz de transformação e

$\mathbf{p}_i \in \mathbf{R}^M$, $i = 1, 2, \dots, \mathbf{d}$ são os vetores que compõem a matriz de transformação. Se todos os \mathbf{N} vetores de medição forem transformados simultaneamente, é possível reescrever a Equação 3.6 como,

$$T = XP \in R^{N \times d} \quad (3.7)$$

$$T = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix} \in R^{N \times d} \quad (3.8)$$

em que \mathbf{T} é a matriz de dados transformados. Normalmente $\mathbf{d} < \mathbf{M}$, já que é comum admitir a hipótese de que é possível representar os dados originais em um subespaço de dimensão menor que a do espaço original. Logo, a Equação 3.9 descreve a relação entre os dados originais e os dados transformados.

$$X = TP^T + E \quad (3.9)$$

Na Equação 3.9, o primeiro termo após a igualdade é denominado espaço principal e o segundo termo é denominado espaço residual. O resíduo está relacionado à variância não explicada pelo modelo. Como para cada \mathbf{p}_i é possível escrever:

$$Var(Xp_i) = E[(Xp_i)^2] = E[p_i^T X^T X p_i] = p_i^T S_{xx} p_i \quad (3.10)$$

o critério de máxima variância torna-se:

$$\max_{p_i} (\lambda_i = p_i^T S_{xx} p_i) \in R \quad (3.11)$$

em que $\lambda_i \in \mathbf{R}$ e $\mathbf{p}_i \in \mathbf{R}^M$ representam, respectivamente, o valor e o vetor característico i da matriz Σ_{xx} e \mathbf{E} representa o operador esperança matemática. Como a matriz Σ_{xx} é real e definida positiva, todos seus valores característicos são positivos e seus vetores característicos são ortogonais entre si (VIDAL *et al.*, 2016). Assim, é possível concluir que o critério de máxima variância implica em componentes principais independentes entre si e de máxima variância. Ampliando a Equação (3.10), conforme apresentado abaixo,

$$Var(XP) = E[(XP)^2] = P^T E[X^T X] P = P^T S_{xx} P \quad (3.12)$$

demonstra-se que é possível encontrar todos os componentes da matriz de transformação \mathbf{P} pela solução de um problema de valor característico da seguinte

forma (VIDAL *et al.*, 2016):

$$S_{xx} = P\Lambda P^T = p_1\lambda_1p_1^T + \dots + p_d\lambda_dp_d^T + \dots + p_M\lambda_Mp_M^T \quad (3.13)$$

em que $\mathbf{P}^T\mathbf{P} = \mathbf{I}$, $\mathbf{P}\mathbf{P}^T = \mathbf{I}$ e $\lambda_1 \geq \dots \lambda_M$. A decomposição apresentada na Equação (3.13) é conhecida como decomposição espectral (JOHNSON e WICHERN, 2007).

Como a técnica de PCA tem por objetivo explicar a variabilidade dos dados amostrais, quando as variáveis de entrada apresentam unidades de medida diferentes, os componentes principais podem ser dominados por algumas variáveis específicas. Mesmo quando os dados apresentam unidades de medida iguais, as variâncias podem variar significativamente. Para estes casos, a solução usualmente proposta consiste em normalizar os dados em relação à covariância e aplicar a técnica de PCA na matriz normalizada $\mathbf{X}_{\text{normalizado}}$ (JACKSON, 1991). A Equação (3.14) ilustra este procedimento, sendo as matrizes \mathbf{X} e \mathbf{S}_{xx} definidas como anteriormente:

$$X_{norm}^T X_{norm} = X^T S_{xx}^{-1} X \quad (3.14)$$

Aplicar a técnica de PCA em \mathbf{X} ou em $\mathbf{X}_{\text{normalizado}}$ produz resultados diferentes, já que, no primeiro caso, a resolução do problema de valor característico é feita a partir da matriz de covariâncias e, no segundo caso, a partir da matriz de correlações. A normalização pode ser entendida como uma mudança de escala das variáveis de processo que transforma a matriz de covariâncias original na matriz de correlações. A equação para normalização de cada um dos vetores de desvio pode ser escrita como:

$$x_i^{normalizado} = \frac{x_i}{\sigma_{x_i}} \quad (3.15)$$

A variância e a covariância entre os vetores de desvio podem ser escritas, respectivamente, como:

$$\sigma_{x_i} = \frac{x_i^T x_i}{N - 1} \quad (3.16)$$

$$\sigma_{x_{ij}} = \frac{x_i^T x_j}{N - 1} \quad (3.17)$$

Normalizando dois vetores de desvios quaisquer, tem-se:

$$x_i^{normalizado} = \frac{x_i}{\sigma_{x_i}} \quad (3.18)$$

$$x_j^{normalizado} = \frac{x_j}{\sigma_{x_j}} \quad (3.19)$$

$$(3.20)$$

Finalmente, ao se calcular a covariância entre os dois vetores normalizados, prova-se que matriz de covariâncias original é transformada, a partir da normalização, na matriz de correlações, como apresentado abaixo.

$$\sigma_{x_{ij}}^{\text{normalizado}} = \frac{1}{N-1} \frac{x_i^T x_j}{\sigma_{x_i} \sigma_{x_j}} = \frac{\sigma_{x_{ij}}}{\sigma_{x_i} \sigma_{x_j}} \quad (3.21)$$

3.1.2 Seleção do Número Componentes Principais

O principal propósito em muitas aplicações envolvendo PCA é reduzir a dimensionalidade do problema de forma a torná-lo mais tratável. Porém, a seleção do número de componentes principais a serem mantidos é uma das principais dificuldades práticas que envolve esta técnica. Posto de outra maneira, deseja-se substituir o conjunto original de variáveis observadas em um processo por outro conjunto de menor dimensão, sem porém acarretar perda significativa de informação que comprometa a modelagem do problema. FEITAL (2011) acrescentou o seguinte acerca do assunto:

A técnica PCA pode ser interpretada como uma expansão da matriz de covariâncias dos dados em série de Karhunen-Loève (JOLLIFFE, 2002), para representação do processo modelado; ou seja, uma técnica não-paramétrica. O truncamento desta série é análogo à escolha do número de sinais relativos ao processo ou ao número de dimensões principais. Na formulação da PCA, considera-se que os primeiros termos da série representam os sinais intrínsecos do processo, enquanto que os últimos termos podem representar os sinais das vizinhanças e do sistema de medição ou serem nulos. Portanto, de acordo com esta interpretação, a geração do modelo é definida unicamente pela determinação do número de termos da série de Karhunen-Loève.

Para abordar a questão da seleção do número de componentes principais a serem retidos, vários critérios foram propostos na literatura. As técnicas mais conhecidas são basicamente critérios empíricos, baseados na variância do conjunto original de variáveis observadas; ou seja, nos valores característicos da matriz de covariância. Enquadram-se neste conjunto de técnicas a Percentagem de Variância Acumulada, o Valor Característico Médio, a Análise Paralela e o "Teste do Joelho" (*Scree Test*). A Tabela 3.1 apresenta as equações que descrevem estes critérios matematicamente, sendo \mathbf{n}_c o número de componentes selecionado pelo critério e δ , $\bar{\lambda}$ e λ^* parâmetros definidos pelo usuário para aplicação dos critérios, representando, respectivamente, o limite da variância acumulada, a média das variâncias e as variâncias dos dados não correlacionados.

Tabela 3.1: Critérios de Seleção de \mathbf{n}_c

Denominação	Critério
Porcentagem de Variância Acumulada	$\mathbf{n}_c = \min_i [(\frac{\sum_{c=1}^i \lambda_c}{\sum \lambda}) \mathbf{100} > \delta]$
Valor Característico Médio	$\mathbf{n}_c = \min_i (\lambda_i > \bar{\lambda})$
Análise Paralela	$\mathbf{n}_c = \min_i (\lambda_i > \lambda_i^*)$
Teste do "Joelho" (<i>Scree Test</i>)	$\mathbf{n}_c = \min_i (\lambda_i \gg \lambda_{i+1})$

Para o critério de porcentagem de variância acumulada, de acordo com JOLLIFFE (2002), o valor de corte deve ser ajustado em algum ponto entre 70% e 90%, mantendo \mathbf{n}_c componentes principais, em que \mathbf{i} é o menor inteiro para o qual o critério tona-se maior que o valor de corte ajustado. Esta é uma regra prática que preserva nas \mathbf{n}_c primeiras componentes principais a maior parte da informação disponível na matriz de covariância.

Muitas vezes, o critério do "teste do joelho" é apresentado graficamente com auxílio de uma função decrescente que mostra a variância explicada por cada um dos valores característicos. Neste gráfico, denominado *SCREE PLOT*, os autovalores são apresentados no eixo das ordenadas, enquanto no eixo das abscissas é apresentada a ordem de cada autovalor de forma decrescente. Ao observar o gráfico, avalia-se o número de componentes principais a serem retidos e isso acontece em torno do "joelho ou cotovelo", o ponto em que a função muda sua inclinação e tende a um platô (LEDESMA *et al.*, 2015). A partir deste ponto, geralmente não há necessidade de manter mais nenhum componente principal.

O "teste do joelho" costuma ser usado em conjunto com a regra de Kaiser. Se todas as variáveis observadas fossem independentes, então os componentes principais seriam iguais a estas variáveis e todas teriam variância unitária, admitindo que a análise de PCA foi realizada a partir da matriz de correlação. Portanto, qualquer componente principal com variância menor que a unidade contém menos informação que qualquer variável original, justificando a eliminação. JOLLIFFE (2002) defende que seria mais indicado descartar apenas os componentes com variância menor que 0,7. A regra de Kaiser apresenta algumas variantes, inclusive para quando se trabalha com a matriz de covariâncias.

A análise paralela é uma técnica de simulação de Monte Carlo que auxilia na

determinação do número de componentes principais a serem retidos na modelagem por PCA. A técnica consiste em comparar os autovalores extraídos da matriz de correlação com os obtidos a partir de variáveis normais não correlacionadas. Do ponto de vista computacional, a análise paralela depende de uma simulação de Monte Carlo, uma vez que os autovalores "esperados" são obtidos simulando amostras aleatórias normais com o mesmo tamanho de amostra e número de variáveis que os dados originais (LEDESMA e VALERO-MORA, 2007). Um componente principal é retido se o autovalor associado for maior que a média daqueles obtidos a partir dos dados aleatórios não correlacionados.

Deve-se perceber que a etapa de redução do modelo é uma etapa comum à maior parte dos modelos empíricos baseados em dados. Nesses casos, sempre há risco de superestimar o modelo do processo, tornando o procedimento ineficaz para todos os fins práticos. Portanto, de forma similar ao procedimento de PCA, outros procedimentos fazem também uso de heurísticas que permitem definir e projetar o tamanho ideal do modelo analítico. Essa questão será abordada novamente em um exemplo do próximo capítulo.

Exemplo 3.1 - Classificação do Conjunto de Dados Íris

O objetivo do Exemplo 3.1 é aplicar a técnica de PCA para a classificação de um conjunto de dados. Do ponto de vista formal, o problema de identificação de falhas pode ser visto como um problema de classificação, em que os dados são separados em dois conjuntos distintos: dados normais e dados com falhas. O conjunto de dados íris (FISHER, 1936) é talvez o banco de dados mais conhecido que se encontra na literatura de reconhecimento de padrões. O conjunto de dados contém 3 classes, cada uma com cinquenta amostras, em que cada classe se refere a uma espécie da planta íris. Uma das classes é linearmente separável das outras duas. Porém, as duas restantes não podem ser separadas por métodos lineares. As três classes no conjunto de dados íris são: íris setosa, íris versicolor e íris virginica. Todas as variáveis do banco de dados foram medidas em centímetros e apresentam a seguinte descrição: comprimento sepal, largura sepal, comprimento da pétala e largura da pétala. Portanto, estão disponíveis $150 \times 4 = 600$ medidas no banco de dados.

A Figura 3.1 a) apresenta a versão gráfica do teste do "joelho" para o conjunto de dados íris. Nota-se que não existe um platô claramente definido nesta figura, deixando claro que o critério do "joelho" apenas não é suficiente para determinar o número de componentes a serem retidos para modelagem deste conjunto de dados. Assim, de certa forma mostra-se que o teste do "joelho" pode não ser adequado para identificação precisa dos modelos mais apropriados para uso. Este é um aspecto relevante do problema de modelagem proposto, dado que o desempenho dos procedimentos heurísticos não pode ser formalmente assegurado. Além disso,

os autovalores encontrados são 2,91, 0,92, 0,14 e 0,02. Assim, pela regra de Kaiser, o conjunto de dados avaliado pode ser representado em um espaço bidimensional, justificando o uso de dois componentes principais. A Figura 3.1 b) apresenta a projeção dos dados originais em um espaço formado por dois componentes principais, ficando evidente que apenas dois componentes principais são suficientes para separar a espécie íris setosa das demais. No entanto, as espécies íris versicolor e íris virgínica não são linearmente separáveis. Isso mostra que o método de PCA pode apresentar dificuldades para identificar a estrutura mais fina da informação disponível.

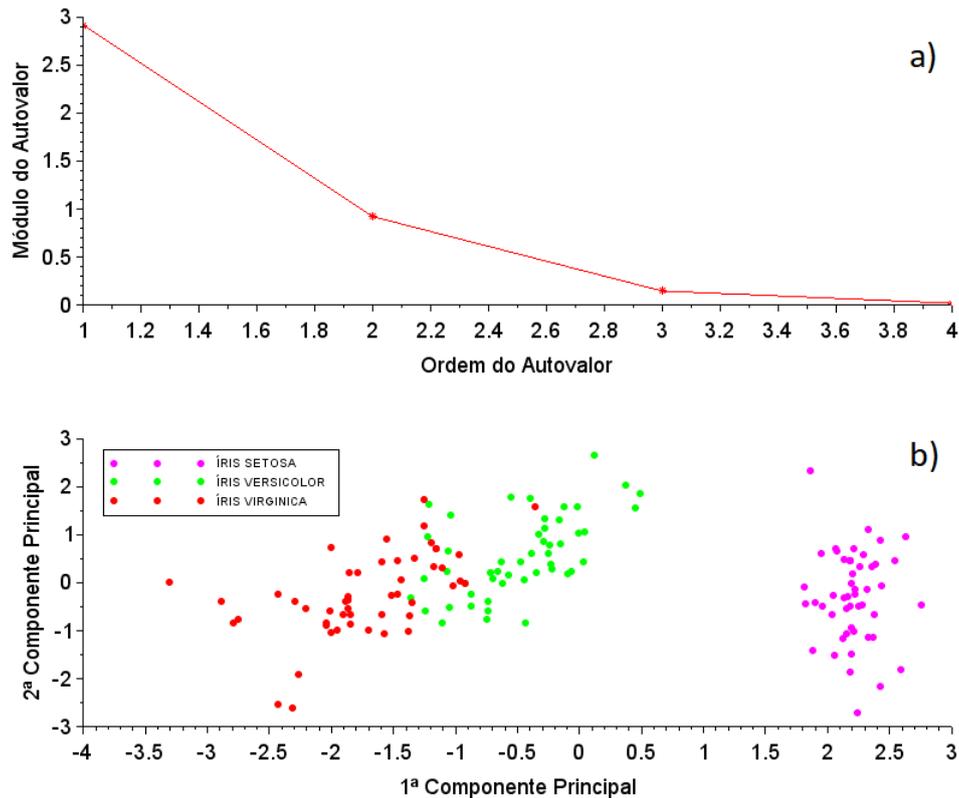


Figura 3.1: Conjunto de dados Íris: a) Gráfico SCREE e b) Gráfico de Projeções

Traçando um paralelo com a área de monitoramento de processos, se o banco de dados contivesse três tipos distintos de falhas de um equipamento qualquer, seria impossível separá-las umas das outras, já que duas destas falhas sempre seriam projetadas na mesma região do espaço bidimensional. Embora exista uma infinidade de técnicas *Data Driven*, a escolha de que técnica aplicar parece depender fortemente dos dados e, portanto, demanda experiência e até mesmo investigação prévia. Dito de outra forma, assim como na modelagem fenomenológica, para a aplicação de técnicas *Data Driven* também é necessário o auxílio de um especialista que conheça as características da planta de processo para que se obtenha sucesso durante a etapa de modelagem. De certa forma, essa é uma contradição intrínseca do processo de modelagem, dado que o uso de técnicas baseadas em dados não deveria depender de

forma alguma do conhecimento do processo (ou do domínio, como propõe o jargão da área).

3.1.3 Índices de Detecção de Falhas T^2 de Hotelling e Q

A utilidade da técnica de PCA para detecção de falhas advém do fato de que ela transforma um conjunto de variáveis correlacionadas em um novo conjunto de variáveis não correlacionadas, mais fáceis de serem monitoradas por meio de cartas de controle (JACKSON, 1991). As duas estatísticas mais utilizadas para gerar cartas de controle a partir da técnica de PCA são a estatística T^2 de Hotelling e a estatística Q . A Estatística T^2 está relacionada ao grau de afastamento de um determinado vetor de medição em relação ao vetor médio. Trata-se de uma medida de variabilidade em relação à média para o caso multivariável, considerada como uma generalização do Teste t de Student. Aplicada em conjunto com a técnica de PCA, a estatística T^2 mede a variabilidade dos componentes principais. Para um vetor qualquer $(\mathbf{x}_j^L)^T \in \mathbf{R}^M, \mathbf{j} = 1, 2, \dots, \mathbf{N}$ esta estatística é definida como (JACKSON, 1991):

$$T^2 = (x_j^L)S^{-1}(x_j^L)^T = (t_j)^T P^T S^{-1} t_j P = (t_j)^T \Lambda^{-1} t_j \quad (3.22)$$

O limiar de detecção para esta estatística pode ser calculado a partir de uma distribuição do tipo F , com auxílio da seguinte equação (JACKSON, 1991):

$$T_{max}^2 = \frac{n_c(p-1)}{n-n_c} F_\alpha(n_c, p-n_c) \quad (3.23)$$

em que n_c é o número de componentes principais, p é o número de observações usadas para gerar o modelo e F_α é o percentil da distribuição F com n_c e $p - n_c$ graus de liberdade e nível de significância α . Ao usar a Equação (3.23), admite-se que os componentes principais possuem distribuição normal e que o número de amostras seja suficientemente grande para que seja possível estimar as médias e covariâncias populacionais a partir dos dados amostrais. Essas são questões muito relevantes e raramente discutidas com profundidade, já que ambas as hipóteses podem estar erradas, levando a limiares de detecção enviesados e a geração de falhas de detecção (falsos alarmes). Esse também é um problema comum à aplicação de outros procedimentos analíticos baseados em dados, já que é usualmente necessário fazer uso de uma distribuição estatística paramétrica, quase sempre a função Gaussiana, para discriminar a falha do comportamento usual. De forma ideal, seria conveniente ser capaz de propor análises de forma independente de distribuições paramétricas e baseadas integralmente nos próprios dados disponíveis.

A estatística Q ou SPE (Erro Quadrático de Predição) está relacionada à

incapacidade do modelo para representar novas observações e permite detectar falhas no espaço residual. Ela fornece uma medida da diferença entre uma observação e sua projeção no espaço constituído pelos componentes principais que foram retidos. Esta estatística pode ser definida como (JACKSON, 1991):

$$Q = (x_j^L - \hat{x}_j^L)(x_j^L - \hat{x}_j^L)^T = x_j^L(I - PP^T)(x_j^L)^T \quad (3.24)$$

O limiar de detecção para esta estatística pode ser calculado a partir de uma distribuição normal padrão usando o seguinte conjunto de equações:

$$\theta_1 = \sum_{i=n_c+1}^p \lambda_i \quad (3.25)$$

$$\theta_2 = \sum_{i=n_c+1}^p \lambda_i^2 \quad (3.26)$$

$$\theta_3 = \sum_{i=n_c+1}^p \lambda_i^3 \quad (3.27)$$

$$h = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2} \quad (3.28)$$

$$Q_{max} = \theta_1 \left[\frac{N_\alpha \sqrt{2\theta_2 h^2}}{\theta_1} + \frac{\theta_2 h(h-1)}{\theta_1^2} + 1 \right]^{\frac{1}{h}} \quad (3.29)$$

$$(3.30)$$

em que N_α é o percentil da distribuição normal padrão e α é o nível de significância pretendido. Em muitos trabalhos, o limiar de detecção desta estatística é calculado a partir de uma distribuição do tipo χ^2 , contudo este procedimento é incorreto e deve ser evitado (JACKSON, 1991). Como no caso anterior, a hipótese de normalidade é raramente questionada, o que pode gerar problemas para caracterização apropriada dos limiares de detecção. Desta forma, as mesmas observações feitas nos parágrafos anteriores a respeito do uso de distribuições paramétricas podem ser reforçadas uma vez mais.

3.2 PCA Dinâmico - DPCA

3.2.1 Geral

Para identificar um modelo de séries temporais que se ajuste aos dados observados, a técnica de DPCA combina a capacidade da técnica de PCA de reduzir a dimensionalidade de um conjunto de variáveis e a capacidade de um modelo

auto-regressivo integrado de média móvel (ARIMA) de lidar com autocorrelação (DE KETELAERE *et al.*, 2015). Ao contrário do que é feito na identificação de modelos ARIMA, a técnica de DPCA realiza a extração do modelo e a redução da dimensionalidade dos dados de forma integrada. A diferença entre as técnicas DPCA e PCA se deve ao fato de que, antes de aplicar a decomposição em valores característicos, a matriz desvios \mathbf{X} ou a matriz de desvios normalizada $\mathbf{X}_{\text{normalizado}}$ passa por uma etapa de ampliação, também conhecida como *embedding*, que resulta numa matriz ampliada ou matriz de trajetórias \mathbf{X}_{amp} (ELSNER e TSONIS, 1996).

$$X_{\text{normalizado}} = \begin{bmatrix} x_1(k) & \dots & x_M(k) \end{bmatrix} \in R^{N \times M} \quad (3.31)$$

↓
A
M
P
L
I
A
Ç
Ã
O
↓

$$X_{\text{amp}} = \begin{bmatrix} x_1(k) & x_1(k-1) & \dots & x_1(k-L) & \dots & x_M(k) & x_M(k-1) & \dots & x_M(k-L) \end{bmatrix} \quad (3.32)$$

Nas Equações (3.31) e (3.32), os vetores coluna $\mathbf{x}_i \in \mathbf{R}^N$, $i = 1, 2, \dots, M$ formados pelas N amostras de cada uma das M variáveis de processo passaram a ser representados como $\mathbf{x}_i(\mathbf{k}) \in \mathbf{R}^N$, $i = 1, 2, \dots, M$ apenas para indicar a dependência temporal. O parâmetro L é conhecido como ordem de atraso, ou simplesmente atraso, e a operação representada por $\mathbf{x}_i(\mathbf{k} - L)$ é equivalente a atrasar o vetor original $\mathbf{x}_i(\mathbf{k})$ em L amostras. A ordem de atraso é um valor inteiro entre $1 \leq L \leq N - 2$. Para entender como a matriz ampliada é construída, considere a submatriz abaixo, escrita para apenas uma das variáveis observadas:

$$X'_{\text{amp}} = \begin{bmatrix} x_1(k) & x_1(k-1) & \dots & x_1(k-L) \end{bmatrix} \quad (3.33)$$

Se, por exemplo, o número de amostras fosse $N = 6$ e o atraso $L = 2$, a primeira variável observada $\mathbf{x}_1(\mathbf{k})$ e sua matriz ampliada \mathbf{X}'_{amp} seriam escritas, respectivamente, como:

$$x_1(k) = \begin{bmatrix} x_1^1 \\ x_1^2 \\ x_1^3 \\ x_1^4 \\ x_1^5 \\ x_1^6 \end{bmatrix} \quad (3.34)$$

$$X'_{amp} = \begin{bmatrix} x_1^1 & x_1^2 & x_1^3 \\ x_1^2 & x_1^3 & x_1^4 \\ x_1^3 & x_1^4 & x_1^5 \\ x_1^4 & x_1^5 & x_1^6 \end{bmatrix} \quad (3.35)$$

O número de colunas da matriz \mathbf{X}_{amp} será sempre igual a $\mathbf{L} + \mathbf{1}$ e o número de linhas sempre igual a $\mathbf{N} - \mathbf{L}$. Para o exemplo acima, o resultado esperado é $\mathbf{2} + \mathbf{1} = \mathbf{3}$ colunas e $\mathbf{6} - \mathbf{3} + \mathbf{1} = \mathbf{4}$ linhas. Procedimento similar é usado para construir modelos e analisar dados quando outras técnicas são adotadas. Neste sentido, a análise proposta tem grande generalidade, dado que outras técnicas estão sujeitas às mesmas avaliações descritas nas próximas seções.

3.2.2 Procedimento de Seleção do Atraso

Para extrair as relações estáticas e dinâmicas dos dados com auxílio da técnica de DPCA, além de determinar o número de componentes principais, também é preciso determinar o atraso (KU *et al.*, 1995). O procedimento apresentado na Figura 3.2 foi proposto por KU *et al.* (1995) como uma estratégia para determinar o atraso. A ideia é determinar o número de atrasos \mathbf{L} que permite obter direções principais que resultem em um modelo que satisfaça os critérios de modelagem propostos anteriormente.

Na Figura 3.2, \mathbf{M} representa o número de variáveis de processo, \mathbf{j} é um inteiro representando o \mathbf{j} -ésimo componente principal, \mathbf{L} é o número de atrasos determinado para a modelagem e \mathbf{r} é a quantidade de restrições impostas pela equações de balanço do processo. A relevância de um componente principal deve ser avaliada por meio de algum critério como, por exemplo, a Regra de Kaiser. Logo, o procedimento pressupõe que um método adequado para determinar o número de componentes principais a serem retidos, foi escolhido previamente. Os autores sugerem que seja usada a análise paralela em conjunto com os gráficos de correlação das projeções para determinar o número de componentes principais. A quantidade de restrições \mathbf{r} é calculada pela diferença entre o número de variáveis observadas e o número de componentes principais a serem retidos. Geralmente o comprimento da janela é igual a 1 ou 2 e indica a ordem do sistema dinâmico (KU *et al.*, 1995); no entanto, o comprimento de janela certamente depende da adequação do intervalo de amostragem.

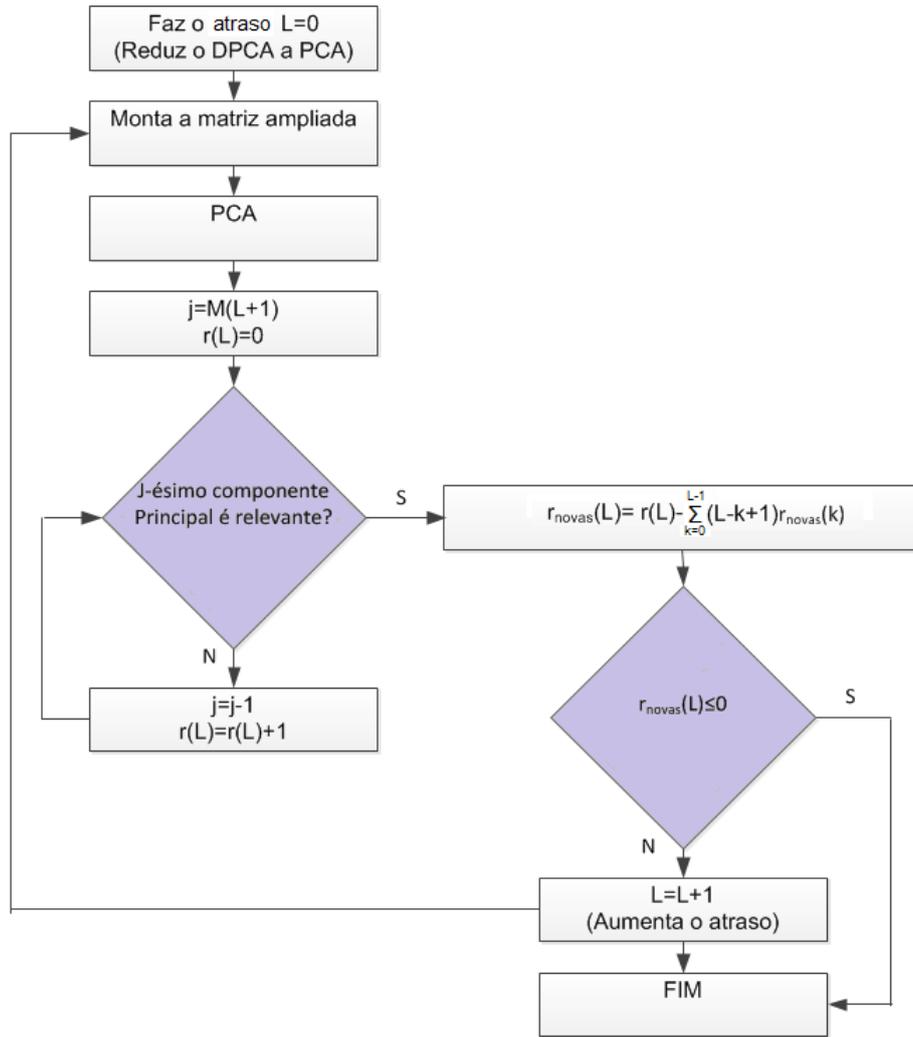


Figura 3.2: Procedimento para Seleção do Atraso para Modelagem com a Técnica DPCA

O procedimento de ampliação, ou *embedding*, é bem estabelecido no estudo de sistemas dinâmicos. Nessa disciplina, a construção da matriz ampliada costuma ser chamada de *delay reconstruction* (KANTZ e SCHREIBER, 2004). O objetivo é tentar reconstruir o espaço de fases do sistema a partir de uma única série temporal obtida por meio da medição de alguma variável de interesse. Existem inúmeros artigos que tratam de critérios relacionados à construção da matriz aumentada, como por exemplo os trabalhos de HEGGER *et al.* (1999), JUDD e MEES (1998), HIRATA *et al.* (2006) e FAES *et al.* (2012).

Nas áreas de processamento de sinais e análise de séries temporais, a técnica DPCA costuma ser conhecida como Análise Espectral Singular (SSA). Nesse contexto, o interesse é decompor uma série temporal em componentes fundamentais para detectar tendência, sazonalidade e fazer previsão. Na Análise Espectral Singular, para construir a matriz \mathbf{X}_{amp} , especifica-se o parâmetro conhecido como comprimento de janela \mathbf{W} , ao invés do atraso \mathbf{L} . O comprimento de janela é definido

simplesmente como o número de colunas desejado para a matriz \mathbf{X}_{amp} . Assim, estes dois parâmetros estão relacionados por meio da seguinte equação: $\mathbf{L} = \mathbf{W} - \mathbf{1}$. ELSNER e TSONIS (1996) e GOLYANDINA *et al.* (2001) trataram em detalhes os fundamentos matemáticos que norteiam a Análise Espectral Singular no contexto de séries temporais.

Um problema muito pouco discutido nesse campo diz respeito à frequência de amostragem de variáveis com tempos característicos de resposta distintos. A presunção do método é que o sistema apresente frequência de amostragem compatível com a do comportamento dinâmico real do processo, o que pode ser difícil de aferir "a priori" em sistemas complexos e compostos por muitas variáveis distintas. Por isso o método de DPCA, apesar da natureza empírica da formulação, requer conhecimento de domínio para que seja bem implementado. Por exemplo, a amostragem de baixa frequência pode ser incompatível com a observação do comportamento dinâmico do processo, enquanto que a amostragem de alta frequência pode induzir a escolha equivocada do atraso e tornar o modelo sensível a erros de medição, por conta da baixa variação do sinal dinâmico determinístico. Esses pontos têm sido pouco discutidos no contexto dos modelos dinâmicos empíricos.

Exemplo 3.2 - Detecção de Falhas em Sinal Senoidal

Como a técnica de DPCA também pode ser encarada como uma ferramenta de análise de sinais, foi sugerido seu uso no contexto de detecção de falhas por meio da análise de sinais de vibração, sendo possível estabelecer um valor mínimo para o comprimento de janela \mathbf{W} (KILUNDU *et al.*, 2011):

$$W_{min} = \frac{fs}{f} \quad (3.36)$$

em que fs é a frequência de amostragem e f é a frequência mínima com que se espera observar determinado fenômeno de interesse. Contudo, como ajustar o comprimento de janela num cenário mais geral de detecção de falhas, quando não se sabe a frequência de interesse? Uma vez mais pode ser observado que o conhecimento do domínio é relevante para a modelagem, mesmo quando se usam procedimentos baseados em dados, o que gera uma aparente contradição das abordagens propostas.

Com o objetivo de avaliar a capacidade de detecção de falhas da técnica DPCA em um problema simples e tentar encontrar uma resposta para a pergunta anterior, foi gerado um sinal senoidal de frequência igual a **60Hz**. Para treinamento do modelo, foram gerados **300** pontos, amostrados a uma frequência **16** vezes maior que a frequência fundamental do sinal. Para tentar entender como ajustar o comprimento da janela, foi gerado um sinal de falha que é a soma do sinal original

com uma componente de tendência, como mostrado na Figura 3.3.

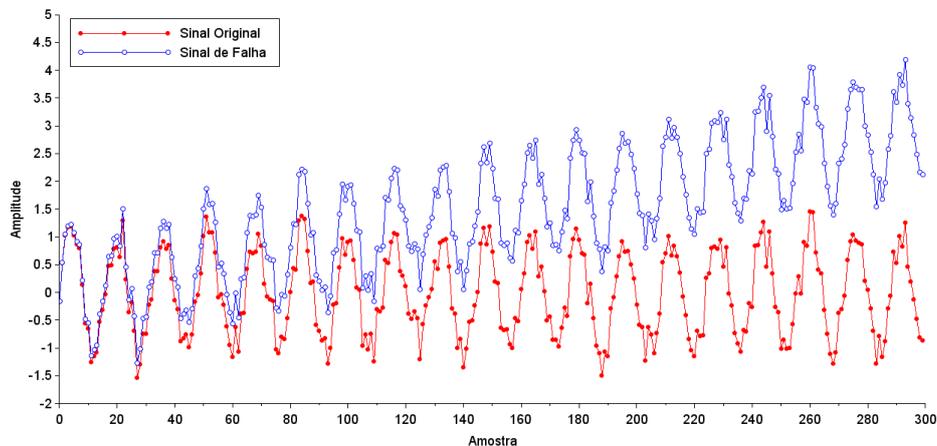


Figura 3.3: Sinal Original e Sinal de Falha no Exemplo de Aplicação da Técnica de DPCA

A técnica DPCA foi executada para diversos comprimentos de janela e os percentuais de falhas detectadas pelas estatísticas \mathbf{T}^2 e \mathbf{Q} foram calculados (Tabela 3.2). Observando a primeira e quarta colunas da Tabela 3.2, fica claro que, a partir de um certo comprimento de janela $\mathbf{W} = 8$, o número de componentes principais retido fica constante e é igual a **2**.

Tabela 3.2: Percentual de Detecção de Falhas em Função do Tamanho da Janela (\mathbf{W})

\mathbf{W}	% Detecção (\mathbf{T}^2)	% Detecção (\mathbf{Q})	n_c	Dim. Esp. Residual
2	50, 50	5, 02	1	1
4	54, 88	1, 35	1	3
8	56, 66	84, 98	2	6
14	5, 92	95, 82	2	12
22	10, 04	97, 13	2	20
32	0	100	2	30

É muito interessante notar que o percentual de detecção da estatística \mathbf{T}^2 atinge um ponto ótimo, a partir do qual ele decresce até zero (Figura 3.4). Já o percentual de detecção da estatística \mathbf{Q} começa em um valor muito baixo e cresce continuamente até atingir **100%** (Figura 3.4). Então, para pequenos comprimentos de janela, a estatística \mathbf{T}^2 parece ser mais relevante para a detecção de falhas, enquanto que, para grandes comprimentos de janela, a estatística \mathbf{Q} parece ser a mais relevante.

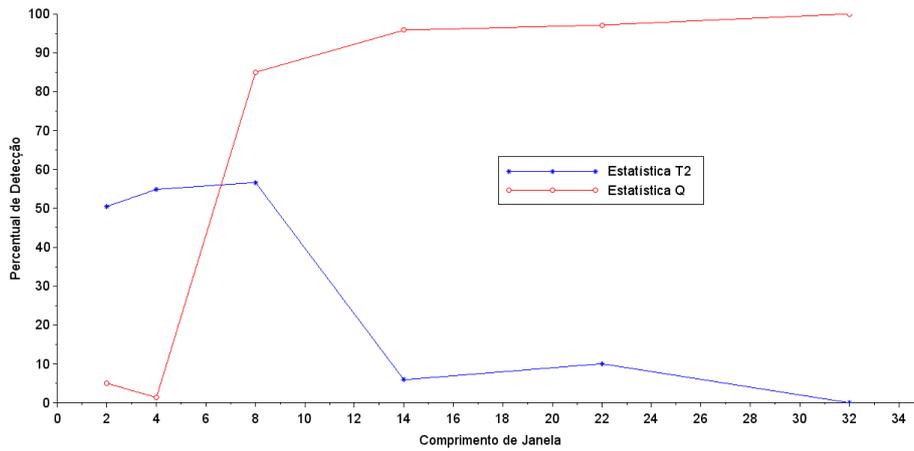


Figura 3.4: Percentual de Detecção

Talvez, a explicação óbvia para comportamento observado para a estatística **Q** esteja relacionada ao fato de que, a partir do ponto em que o número de componentes principais fica constante, o aumento do comprimento de janela também aumenta a dimensão do espaço residual, espaço em que a estatística opera. Como a informação contida no sinal original está distribuída no espaço principal formado pelos componentes principais e no espaço residual, à medida que o comprimento de janela cresce, mais informação é transferida do espaço principal, que nesse caso é sempre bidimensional, para o espaço residual, justificando o comportamento da estatística **T²**. A sequência de gráficos apresentada na Figura 3.5 parece comprovar este raciocínio. Nesta figura, a projeção do sinal original nos componentes principais está representada pelos pontos vermelhos, enquanto a projeção do sinal de falha está representada pelos pontos azuis. À medida que o comprimento de janela cresce, a projeção dos dois sinais se torna cada vez mais próxima. Isto parece comprovar o fato de que mais informação está sendo transferida do espaço principal para o espaço residual.

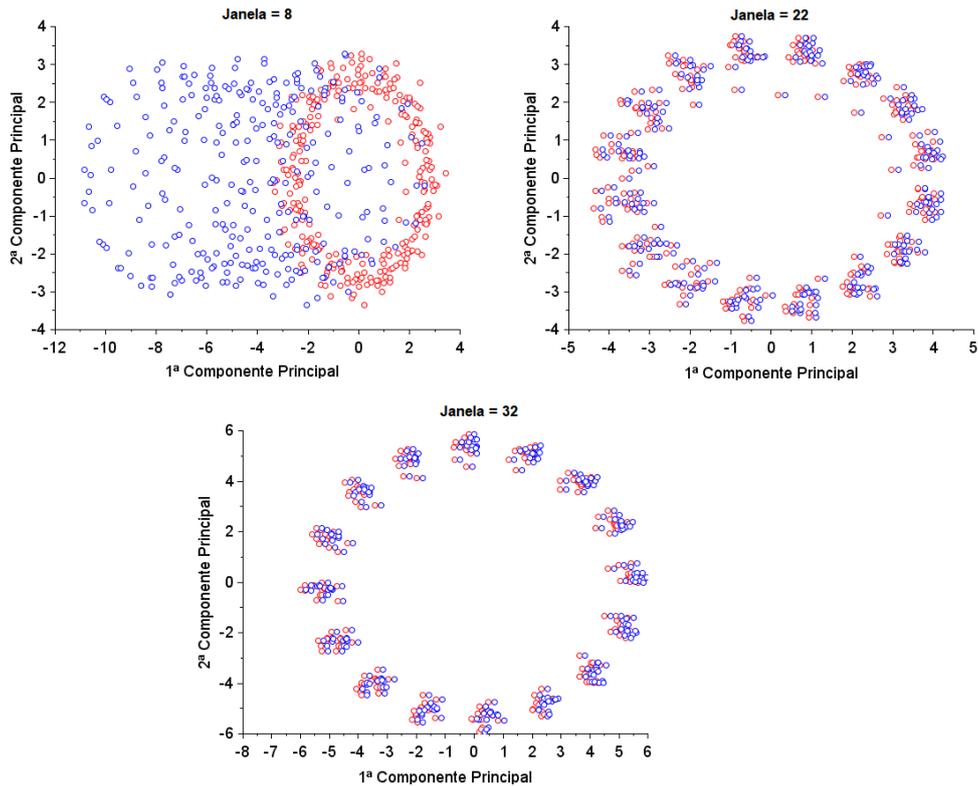


Figura 3.5: Projeções nos Componentes Principais no Exemplo de Aplicação da Técnica de DPCA

Para avaliar qual das estatísticas consegue detectar uma falha mais rapidamente no problema proposto, a Figura 3.6 mostra em um único gráfico os logaritmos dos valores padronizados das estatísticas T^2 e Q , $\log\left(\frac{T^2}{T_{\max}}\right)$ e $\log\left(\frac{Q}{Q_{\max}}\right)$, para um comprimento de janela igual a 8. Nota-se que a estatística Q , além de detectar a falha mais rapidamente, exibe um comportamento menos oscilatório e, portanto, mais fácil de ser interpretado.

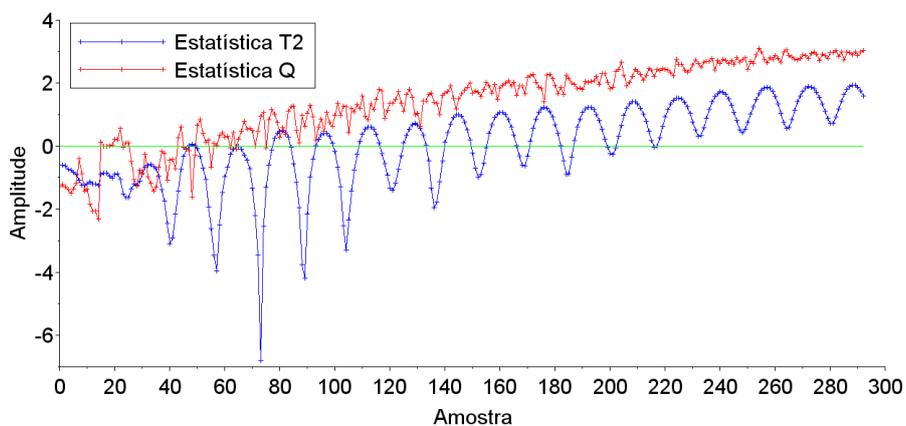


Figura 3.6: Comparação das Estatísticas de Detecção no Exemplo de Aplicação da Técnica de DPCA

O exemplo proposto mostra com clareza que o desempenho da técnica DPCA

para fins de detecção de falhas depende fortemente da correta frequência de amostragem, das variáveis de processo e da estatística particular usada para monitoramento do estado do processo, o que parece justificar o uso de múltiplas estatísticas (ou alarmes) para o monitoramento de processo, o que tem sido muito pouco investigado na literatura de referência. É curioso observar que a incorreta amostragem do processo pode resultar em dificuldades para interpretar a normalidade de operação, o que tem sido pouco discutido no âmbito dos procedimentos de identificação de falhas.

Exemplo 3.3 - Modelagem do Sistema de Três Tanques

O sistema apresentado na Figura 3.7 é composto por três tanques idênticos e conectados em série (GUERFEL *et al.*, 2009). As variáveis de entrada são determinadas pelas vazões Q_1 e Q_2 , enquanto as variáveis de saída H_1 , H_2 e H_3 representam respectivamente as alturas do nível de líquido no primeiro, no segundo e no terceiro tanque. Os tanques estão interconectados pela base por meio de tubulações, de modo que os tanques dois e três estão ligados por intermédio da válvula V_3 e o tanque dois está ligado ao meio exterior por intermédio da válvula V_2 .

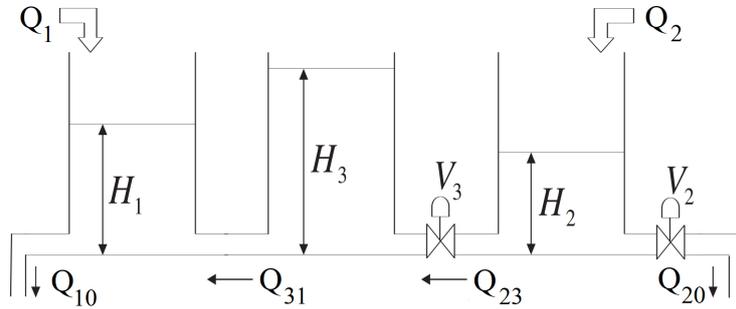


Figura 3.7: O Sistema de Três Tanques em Série. Adaptado de GUERFEL *et al.* (2009).

As equações que modelam o sistema são apresentadas abaixo para um período de amostragem de um segundo (GUERFEL *et al.*, 2009). Nestas equações, $A = 0,01539\text{m}^2$ é a área da seção reta dos tanques e os coeficientes de descarga $k_1 = 1,816 \times 10^{-4} \frac{\text{m}^{5/2}}{\text{s}}$, $k_2 = 9,804 \times 10^{-5} \frac{\text{m}^{5/2}}{\text{s}}$, $k_{23} = 1,005 \times 10^{-4} \frac{\text{m}^{5/2}}{\text{s}}$ e $k_{31} = 7,804 \times 10^{-5} \frac{\text{m}^{5/2}}{\text{s}}$ foram determinadas pelas características do processo.

$$H_1(k) = \frac{Q_1(k) + Q_{31}(k) - Q_{10}(k)}{A} - H_1(k-1) \quad (3.37)$$

$$H_2(k) = \frac{Q_2(k) - Q_{23}(k) - Q_{20}(k)}{A} - H_2(k-1) \quad (3.38)$$

$$H_3(k) = \frac{Q_{23}(k) - Q_{31}(k)}{A} - H_3(k-1) \quad (3.39)$$

$$Q_{31}(k) = k_{31}f(H_3(k) - H_1(k)) \quad (3.40)$$

$$Q_{23}(k) = k_{23}f(H_2(k) - H_3(k)) \quad (3.41)$$

$$Q_{10}(k) = k_1\sqrt{H_1(k)} \quad (3.42)$$

$$Q_{20}(k) = k_2\sqrt{H_2(k)} \quad (3.43)$$

$$f(x) = \text{sign}(x)\sqrt{|x|} \quad (3.44)$$

Uma análise dos graus de liberdade mostra que o sistema possui 5 parâmetros (\mathbf{A} , \mathbf{k}_1 , \mathbf{k}_2 , \mathbf{k}_{23} , \mathbf{k}_{31}) e 9 variáveis (\mathbf{H}_1 , \mathbf{H}_2 , \mathbf{H}_3 , \mathbf{Q}_1 , \mathbf{Q}_2 , \mathbf{Q}_{10} , \mathbf{Q}_{20} , \mathbf{Q}_{23} e \mathbf{Q}_{31}). Como existem sete equações, restam apenas duas variáveis livres ou distúrbios, definidas aqui como as vazões \mathbf{Q}_1 e \mathbf{Q}_2 por conta da manipulação real do processo. Para simular o processo, foram gerados sinais do tipo binário pseudo aleatório (PRBS) para excitar as vazões \mathbf{Q}_1 e \mathbf{Q}_2 com amplitudes variáveis, respectivamente, entre $[3,20 \times 10^{-5}, 6,71 \times 10^{-5}]m^3/s$ e $[5,73 \times 10^{-5}, 9,57 \times 10^{-5}]m^2/s$ (GUERFEL *et al.*, 2009), conforme apresentado na Figura 3.8. As condições iniciais usadas por GUERFEL *et al.* (2009) para simular o processo foram $\mathbf{H}_1(\mathbf{k}-1) = 0,147m$, $\mathbf{H}_2(\mathbf{k}-1) = 0,276m$ e $\mathbf{H}_3(\mathbf{k}-1) = 0,195m$. A Figura 3.9 mostra o comportamento das variáveis de saída \mathbf{H}_1 , \mathbf{H}_2 e \mathbf{H}_3 , quando submetidas aos sinais de excitação. Antes de aplicar o algoritmo proposto por KU *et al.* (1995), para determinar o atraso, as variáveis de entrada e saída do sistema (\mathbf{Q}_1 , \mathbf{Q}_2 , \mathbf{H}_1 , \mathbf{H}_2 e \mathbf{H}_3) foram contaminadas com ruído branco de média zero e desvio padrão igual a 5% do desvio da variável original, com o objetivo de simular o ruído de medição.

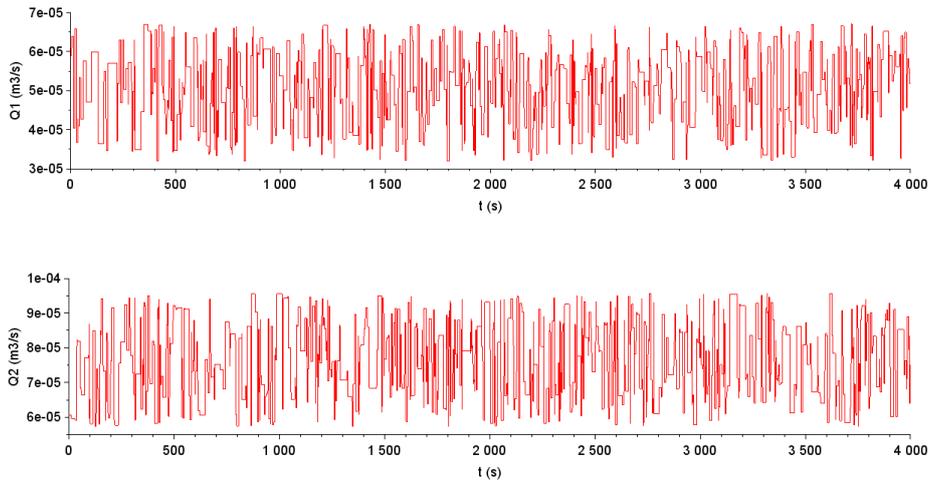


Figura 3.8: Sinais de Vazão Usados no Exemplo dos Tanques

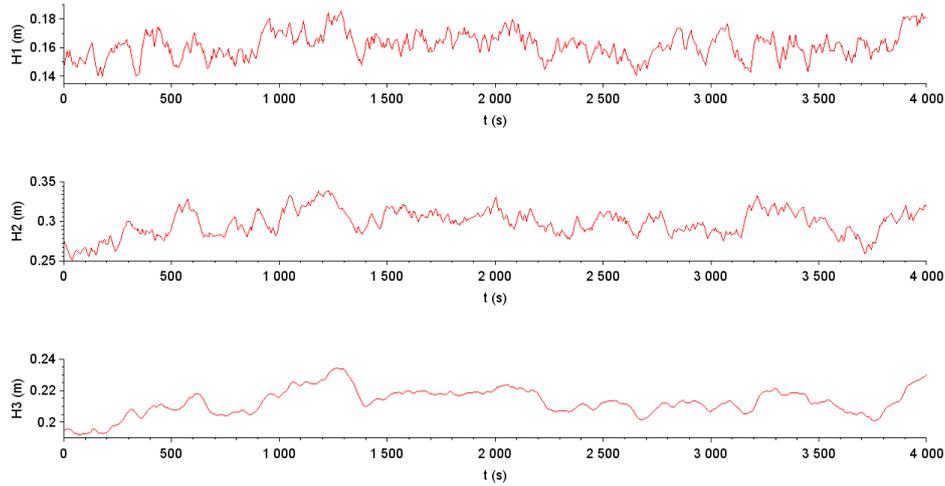


Figura 3.9: Alturas dos Níveis de Líquido nos Tanques

A Figura 3.10 apresenta os gráficos *SCREE* e os gráficos da variância explicada, quando o atraso varia entre 0 e 3. Nos gráficos *SCREE* a reta azul determina o limiar igual a 1,0 e a reta verde determina o limiar igual a 0,7, usados para avaliar se determinado valor característico está associado a uma relação linear. As Tabelas 3.3 e 3.4 apresentam os resultados obtidos a partir da aplicação do critério de seleção do atraso, conforme procedimento proposto por KU *et al.* (1995). Nos resultados apresentados da Tabela 3.3 foi considerado um limiar igual a 0,7 e na Tabela 3.4 foi considerado um limiar igual a 1,0. Para o limiar igual a 0,7, a Tabela 3.3 indica que o atraso seria igual a 2, sendo que respectivo gráfico da Figura 3.10 indica que seriam necessários 4 componentes principais, explicando aproximadamente 90% da variância.

Tabela 3.3: Seleção do atraso com limiar de 0,7 no Exemplo dos Tanques

Atraso (L)	r(L)	$\sum_{i=0}^{L-1} (L - i + 1)r_{\text{novas}}(i)$	r _{novas}
0	2	-	2
1	6	$2(1 + 1) = 4$	2
2	11	$2(2 + 1) + 2(2 - 1 + 1) = 10$	1
3	16	$2(3 + 1) + 2(3 - 1 + 1) + 1(3 - 2 + 1) = 16$	0

Para o limiar igual a 1,0 a Tabela 3.4 indica que o atraso seria igual a 1 e o respectivo gráfico da Figura 3.10 indica que seriam necessários 3 componentes principais, explicando aproximadamente 85% da variância.

Tabela 3.4: Seleção do atraso com limiar de 1,0 no Exemplo dos Tanques

Atraso (L)	$r(L)$	$\sum_{i=0}^{L-1} (L - i + 1)r_{\text{novas}}(i)$	r_{novas}
0	3	-	3
1	7	$3(1 + 1) = 6$	1
2	11	$3(2 + 1) + 1(2 - 1 + 1) = 11$	0

Portanto, conclui-se que a falta de um procedimento universal para a seleção do número de componentes principais implica em resultados diferentes para determinar o atraso, embora, para os dois limiares usados, a variância explicada pelo modelo seja aproximadamente a mesma. Esse constitui um Calcanhar de Aquiles dos métodos de DPCA, já que a definição dos atrasos depende dos procedimentos numéricos aplicados, guardando pouca conexão com o processo real, o que tem sido usualmente negligenciado na maior parte das aplicações. No fundo, este é um parâmetro de sintonia do modelo, sendo possível obter modelos distintos com desempenhos similares. Nesse caso, uma vez mais, o conhecimento do domínio pode ser muito útil e aplicado "a priori" para a definição dos atrasos, que em última instância refletem a dinâmica real do processo.

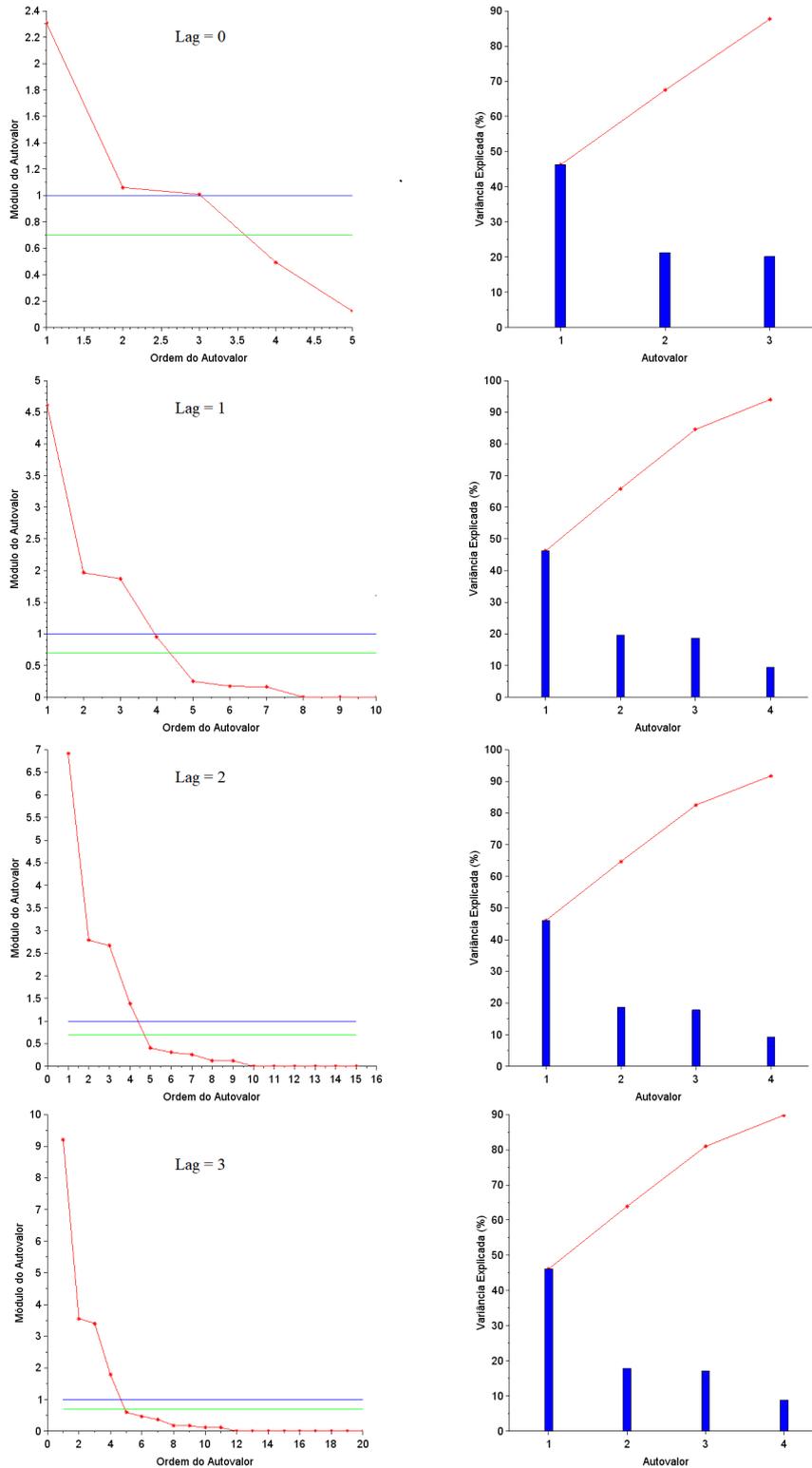


Figura 3.10: Gráfico SCREE (esquerda) e Gráfico de Variância Explicada (direita)

3.3 PCA Não-Linear - KPCA

3.3.1 Geral

De acordo com VIDAL *et al.* (2016), a técnica de PCA busca ajustar um subespaço linear de dimensão reduzida a uma coleção de pontos. O autor acrescenta que um subespaço linear pode não ser capaz de capturar não linearidades presentes nos dados em aplicações práticas. O desenvolvimento de técnicas de modelagem de processos fortemente não lineares ainda constitui um campo de pesquisa bastante desafiador e uma das fragilidades dos modelos baseados em PCA. Assim como a técnica de DPCA, a técnica de KPCA está fundamentada na ampliação da matriz de desvios \mathbf{X} ou da matriz de desvios normalizada $\mathbf{X}_{\text{normalizado}}$; entretanto, tal ampliação é realizada por meio de um mapeamento não linear das variáveis originais. O mapeamento proposto por SCHÖLKOPF *et al.* (1998) pode ser escrito como VIDAL *et al.* (2016),

$$\phi(\cdot) : R^M \rightarrow R^D \quad (3.45)$$

$$x_j^L \rightarrow \phi(x_j^L) \quad (3.46)$$

em que $\mathbf{x}_j^L \in \mathbf{R}^M$, $j = 1, 2, \dots, N$, como definido anteriormente, é a j -ésima linha da matriz de desvios \mathbf{X} . O mapeamento proposto nada mais é que uma transformação não linear do espaço original, com o objetivo de torná-lo um espaço linear para aplicação da técnica de PCA (VIDAL *et al.*, 2016). Em vários aspectos, a técnica se assemelha a uma expansão funcional de um sinal não linear, usando uma base funcional especificada como referência. Se todas as linhas da matriz de desvios \mathbf{X} forem transformadas simultaneamente, a matriz ampliada e a matriz de correlação no espaço transformado podem ser escritas, respectivamente, como,

$$\Phi = \begin{bmatrix} \phi(x_1^L) \\ \vdots \\ \phi(x_N^L) \end{bmatrix} \in R^{N \times D} \quad (3.47)$$

$$S_\Phi = \frac{\Phi^T \Phi}{N - 1} \in R^{D \times D} \quad (3.48)$$

Escrito dessa forma, a técnica de PCA ganha um formato que permite se aproximar de muitas outras técnicas não lineares de análise de dados e ganha estrutura formal que permite representar funções não lineares com a precisão desejada, bastando para isso que ϕ constitua uma base apropriada do espaço de funções. Na área de aprendizado de máquina, a etapa de ampliação é conhecida

como extração de características e o espaço resultante da transformação do espaço original é conhecido como espaço de características. Para VIDAL *et al.* (2016), uma potencial dificuldade associada à técnica de KPCA resulta do fato de que, geralmente, a matriz ampliada Φ possui um número elevado de colunas em relação ao número de linhas. Conseqüentemente, a matriz S_ϕ apresenta dimensões elevadas, dificultando sua decomposição em autovalores. Entretanto, por meio da Equação (3.49) que descreve o problema de valor característico, SCHÖLKOPF *et al.* (1998) observou que cada autovetor $\mathbf{u}_i \in \mathbf{R}^D$, $i = 1, 2, \dots, D$ pertence ao espaço linha da matriz Φ , como apresentado na Equação (3.50).

$$\lambda_i u_i = S_\Phi u_i \rightarrow \lambda_i u_i = \frac{\Phi^T \Phi}{N-1} u_i \quad (3.49)$$

$$u_i = \Phi^T \left(\frac{\lambda^{-1} \Phi u_i}{N-1} \right) \rightarrow u_i = \Phi^T w_i \in \text{Range}(\Phi) \quad (3.50)$$

$$w_i = \left(\frac{\lambda^{-1} \Phi u_i}{N-1} \right) \quad (3.51)$$

Rescrevendo a Equação (3.49) em função dos vetores \mathbf{w}_i , conclui-se que eles são os vetores característicos da matriz $\Phi\Phi^T$, com os mesmos valores característicos λ_i associados a \mathbf{u}_i .

$$\lambda_i u_i = S_\Phi u_i \rightarrow \lambda_i \Phi^T w_i = \frac{\Phi^T \Phi}{N-1} \Phi^T w_i \rightarrow \lambda_i w_i = \frac{\Phi\Phi^T}{N-1} w_i \quad (3.52)$$

Expressando $\Phi\Phi^T$ a partir do mapeamento definido pelas Equações (3.45) e (3.46), observa-se que as componentes da matriz $\Phi\Phi^T$ podem ser obtidas a partir do produto escalar entre vetores do espaço característico.

$$\Phi\Phi^T = \begin{bmatrix} \phi(x_1^L)\phi(x_1^L)^T & \dots & \phi(x_1^L)\phi(x_N^L)^T \\ \vdots & \ddots & \vdots \\ \phi(x_N^L)\phi(x_1^L)^T & \dots & \phi(x_N^L)\phi(x_N^L)^T \end{bmatrix} \in R^{N \times N} \quad (3.53)$$

Diferentemente da matriz $\Phi^T\Phi$, que possui dimensão $\mathbf{R}^{D \times D}$, a matriz $\Phi\Phi^T$ possui dimensão $\mathbf{R}^{N \times N}$. As componentes desta última matriz podem ser calculadas com auxílio das chamadas funções núcleo (*kernel trick*), cujo objetivo é encontrar o produto escalar entre dois vetores do espaço característico, sem que seja necessário computar o mapeamento individual de cada vetor. Portanto, usando a nomenclatura de funções núcleo, a Equação (3.53) pode ser escrita como:

$$K = \Phi\Phi^T = \begin{bmatrix} k_{11} & \dots & k_{1N} \\ \vdots & \ddots & \vdots \\ k_{N1} & \dots & k_{NN} \end{bmatrix} \in R^{N \times N} \quad (3.54)$$

$$k_{ij} = \phi(x_i^L)\phi(x_j^L)^T \in R \quad (3.55)$$

A matriz \mathbf{K} é simétrica e positiva semidefinida; porém, de acordo com (SCHÖLKOPF *et al.*, 1998), não há garantias que ela esteja centralizada em relação à média. O procedimento de centralização em relação à média pode ser escrito de forma matricial como (?),

$$\tilde{\Phi} = \Phi - \frac{\mathbf{1}_{NxN}}{N}\Phi = (I - \frac{\mathbf{1}_{NxN}}{N})\Phi \quad (3.56)$$

em que $\tilde{\Phi}$ é a matriz ampliada já centralizada, \mathbf{I} é a matriz identidade e $\mathbf{1}_{NxN}$ é uma matriz de dimensão $\mathbf{N} \times \mathbf{N}$, com todas as componentes iguais à unidade. A partir da Equação (3.56), a centralização da matriz \mathbf{K} pode ser obtida do seguinte modo:

$$\tilde{K} = \tilde{\Phi}\tilde{\Phi}^T = (I - \frac{\mathbf{1}_{NxN}}{N})\Phi\Phi^T(I - \frac{\mathbf{1}_{NxN}}{N})^T \quad (3.57)$$

Já que $\mathbf{K} = \Phi\Phi^T$, a matriz $\tilde{\mathbf{K}}$ pode a ser escrita como:

$$\tilde{K} = \tilde{\Phi}\tilde{\Phi}^T = (I - \frac{\mathbf{1}_{NxN}}{N})K(I - \frac{\mathbf{1}_{NxN}}{N})^T \quad (3.58)$$

A projeção de um vetor transformado e centralizado $\tilde{\phi}(\mathbf{x}_j^L)$ em um vetor característico \mathbf{u}_i pode ser calculada a partir da seguinte equação:

$$t_j = \mathbf{u}_i^T \tilde{\phi}(x_j^L)^T \quad (3.59)$$

Mas como $\mathbf{u}_i = \tilde{\Phi}^T \mathbf{w}_i$, a Equação (3.59) pode ser reescrita da seguinte forma:

$$t_j = \mathbf{w}_i^T \tilde{\Phi} \tilde{\phi}(x_j^L)^T = \mathbf{w}_i^T (I - \frac{\mathbf{1}_{NxN}}{N}) \Phi \tilde{\phi}(x_j^L)^T \quad (3.60)$$

De forma similar ao que foi feito na Equação (3.56), pode-se determinar $\tilde{\phi}(\mathbf{x}_j^L)$ da seguinte maneira:

$$\tilde{\phi}(x_j^L)^T = \phi(x_j^L)^T - \Phi^T \frac{\mathbf{1}_{Nx1}}{N} \quad (3.61)$$

Finalmente, isto permite escrever a Equação (3.60) como:

$$t_j = \mathbf{w}_i^T (I - \frac{\mathbf{1}_{NxN}}{N}) (\Phi \phi(x_j^L)^T - \Phi \Phi^T \frac{\mathbf{1}_{Nx1}}{N}) \quad (3.62)$$

$$t_j = w_i^T \left(I - \frac{1_{NxN}}{N} \right) \begin{bmatrix} k_{1j} \\ \vdots \\ k_{Nj} \end{bmatrix} - K \frac{1_{Nx1}}{N} \quad (3.63)$$

3.3.2 Funções Núcleo Mais Utilizadas

As funções núcleo apresentam propriedades matemáticas específicas, de modo que o desenvolvimento de novas funções núcleo ainda constitui uma área de interesse. De modo geral, as duas funções não lineares mais empregadas na prática são a função núcleo Gaussiana e a Polinomial, descritas, respectivamente, nas formas (RAYKAR e ANKUR, 2002):

$$k_G(x_i^L, x_j^L) = e^{-\frac{\|x_i^L - x_j^L\|^2}{\sigma^2}} \quad (3.64)$$

$$k_P(x_i^L, x_j^L) = [x_i^L(x_j^L)^T]^n \quad (3.65)$$

As Equações (3.64) e (3.65) representam o produto escalar dos dois vetores que compõem o argumento da função núcleo. Portanto, com auxílio de uma função núcleo, é possível calcular o produto escalar entre dois vetores, sem que seja necessário computar o mapeamento individual de cada vetor, o que é muito mais eficiente do ponto de vista computacional para a implementação do algoritmo de KPCA (VIDAL *et al.*, 2016).

Exemplo 3.4 - Modelagem Não Linear Via KPCA

Com o intuito de avaliar a capacidade de modelagem não linear da técnica KPCA, foram gerados dois conjuntos de pontos a partir das equações do círculo de raio unitário amostrado a cada 1 grau (Figura 3.11). O primeiro conjunto de pontos foi usado para treinar o algoritmo e o segundo conjunto de pontos foi usado para validá-lo. De modo a simular o ruído de medição, antes da aplicação do algoritmo, cuja função núcleo utilizada foi a Gaussiana, os pontos gerados foram contaminados com ruído gaussiano de média zero e variância igual a 3% do raio do círculo. Na Figura 3.11, os pontos usados no treinamento estão em vermelho e os pontos de validação estão em amarelo.

Uma análise dos graus de liberdade mostra que o exemplo proposto possui 3 variáveis (as coordenadas \mathbf{x} e \mathbf{y} e o ângulo θ) e 1 parâmetro (o raio do círculo). Como existem duas equações, resta apenas uma variável livre, definida aqui como o ângulo θ . Assim, parece razoável usar apenas um componente principal na modelagem. Para facilitar a visualização, após o treinamento os pontos apresentados na Figura 3.11 foram projetados no espaço bidimensional, formado pelos dois

primeiros componentes principais. Para tentar compreender a utilidade da técnica de KPCA para detecção de falhas, a reta azul interna ao círculo e a reta verde externa ao círculo (Figura 3.11) também foram projetadas e os resultados são apresentados na Figura 3.12.

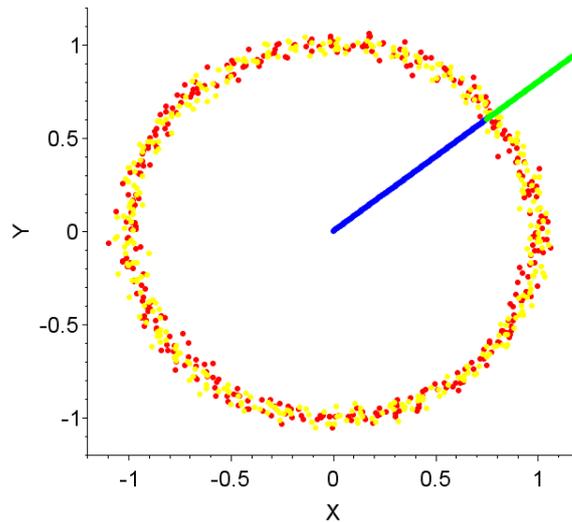


Figura 3.11: Pontos de Treinamento (Vermelho) e Pontos de Validação (Amarelo) usados no Exemplo KPCA

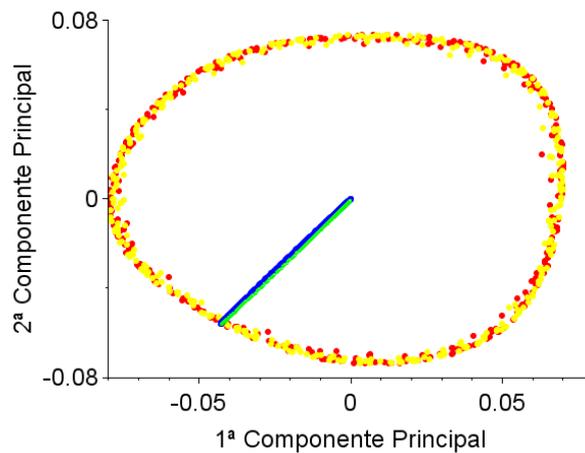


Figura 3.12: Projeções do Modelo Construído com Dois Componentes Principais no Exemplo KPCA

Tanto a reta interna quanto a reta externa foram projetadas uma sobre a outra; porém, observa-se que os pontos mais distantes do círculo foram projetados nas proximidades da origem. Este resultado é consequência da escolha da função núcleo gaussiana, já que a exponencial de vetores distantes da média é essencialmente igual a zero. Para a estatística \mathbf{T}^2 , isto implica em um limite inferior de detecção, não

em um limite superior, como normalmente acontece nas técnicas de PCA e DPCA. Este comportamento está ilustrado na Figura 3.13.

É importante destacar que, para o cálculo das projeções, é necessário que a matriz \mathbf{K} esteja disponível. À medida que mais pontos são usados no treinamento do algoritmo, esta matriz cresce, o que pode constituir um problema grave para aplicações práticas, quando o conjunto de dados de treinamento se torna relativamente grande.

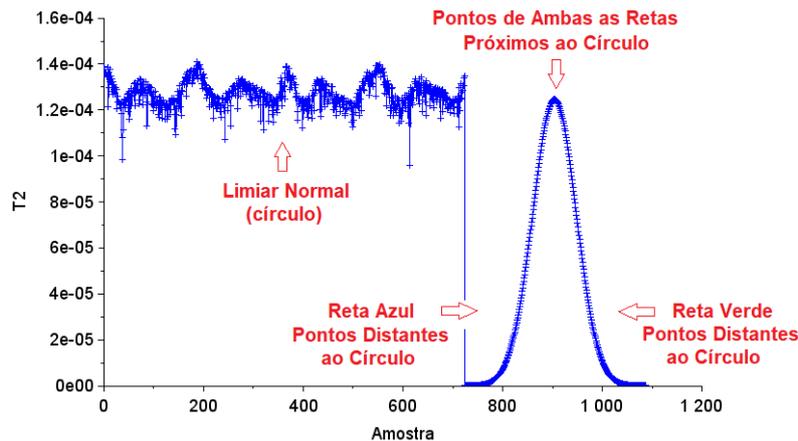


Figura 3.13: Estatística T^2 no Exemplo KPCA

O exemplo proposto mostra que a escolha da função núcleo usada na modelagem pode influenciar o cálculo do limiar de detecção da estatística de monitoramento de falhas. Também é relevante destacar que a escolha da função núcleo não é trivial, sendo este um problema pouco abordado na literatura de detecção de falhas. Além disso, no exemplo proposto, o conhecimento do domínio foi uma vez mais relevante, pois contribuiu com a definição do número de componentes a serem retidos.

3.4 Outras Variantes de PCA

O propósito desta seção é fornecer uma visão geral de outras técnicas baseadas em PCA, em razão do potencial de aplicação para o monitoramento de processos.

3.4.1 PCA Robusto

Reiteradamente em aplicações de interesse prático, os dados amostrais estão incompletos (*Missing Data*). Em muitos casos, os dados foram corrompidos por erros grosseiros ou sistemáticos e não se sabe, "a priori", exatamente quais pontos amostrais foram afetados. Para tratar estas situações, foram desenvolvidas várias extensões denominadas robustas, como por exemplo a técnica PCA Robusto

(CANDÈS *et al.*, 2011) e PCA por Covariância Ponderada (KRIEGEL *et al.*, 2008). Acredita-se que estas técnicas poderiam ser usadas para pré processamento de dados, permitindo que técnicas convencionais possam ser empregadas em cenários de dados corrompidos ou incompletos.

3.4.2 PCA Variante no Tempo

Para cenários em que os dados apresentam variações lentas ao longo do tempo, foi proposto um conjunto de técnicas com a finalidade de adaptar o modelo construído previamente para torná-lo aderente aos dados mais recentes. As técnicas mais conhecidas, segundo esta abordagem, são PCA móvel ponderado exponencialmente (Exponentially Weighted Moving PCA, EWMPCA) (WOLD, 1994), PCA Robusto Recursivo (JIN *et al.*, 2006), PCA Por Janela Móvel (Moving Window PCA, MWPCA) (WANG *et al.*, 2005), PCA Por Janela Móvel Variável (HE e YANG, 2008). KADLEC *et al.* (2011) fizeram uma revisão dos diferentes tipos de mecanismos de adaptação empregados nesta classe de algoritmos. Contudo, a aplicação prática destes métodos parece ser bastante limitada, uma vez que poucos autores trataram da questão de quando disparar o mecanismo de adaptação. A violação do limiar de detecção das estatísticas usadas na identificação de falhas não implica necessariamente na ocorrência de uma falha. Tal violação pode simplesmente indicar um valor raro ou não usual. Neste caso, o emprego de mecanismos de adaptação pode inutilizar o modelo construído previamente, comprometendo a detecção de falhas.

3.4.3 PCA Probabilístico

O PCA por Máxima Verossimilhança (Maximum Likelihood PCA) incorpora informações sobre erros de medição para desenvolver modelos PCA ótimos no sentido de máxima verossimilhança (WENTZELL *et al.*, 1997). De acordo com FEITAL (2011), foi a primeira proposta amplamente conhecida a lidar com os problemas de erros de modelagem (resíduo) dependentes, diferentemente distribuídos e heterocedásticos. Assim, esta parece ser uma técnica promissora, já que várias hipóteses usadas em modelagem estatística convencional são relaxadas e tratadas por meio do princípio de máxima verossimilhança.

3.5 Outras Técnicas de Análise de Subespaços

Apesar de não serem baseadas em PCA, esta seção apresenta um conjunto de técnicas para redução de dimensionalidade e modelagem multivariável.

3.5.1 Análise de Componentes Independentes

O propósito da Análise de Componentes Independentes (Independent Component Analysis) é encontrar uma representação linear de dados não-gaussianos, de modo a assegurar que os componentes fundamentais extraídos dos dados originais sejam estatisticamente independentes (HYVÄRINEN e OJA, 2000). Para alcançar este objetivo, a técnica utiliza no algoritmo proposto alguma métrica para refletir o grau de não gaussianidade de uma variável aleatória. Existem implementações que se baseiam em uma medida clássica do grau de achatamento de uma distribuição de probabilidades em relação à distribuição normal, conhecida como curtose (MARTINEZ e MARTINEZ, 2008). Outras abordagens operam com a minimização de uma grandeza denominada informação mútua, desenvolvida com auxílio de conceitos da Teoria da Informação (LANGLOIS *et al.*, 2010). Acredita-se que a utilidade desta técnica para o monitoramento de processo se deve ao fato de que a técnica é capaz de operar com dados provenientes de distribuições não gaussianas. Portanto, caso a análise dos dados mostre que uma ou mais variáveis apresentam distribuição não gaussiana, uma possibilidade seria aplicar a Análise de Componentes Independentes em conjunto com outras técnicas, como por exemplo a técnica de PCA.

3.5.2 Abordagem Linear Local

Assim como a técnica de PCA, a Abordagem Linear Local (Locally Linear Embedding) (ROWEIS, 2000) é uma técnica para redução de dimensionalidade de dados multivariáveis, classificada como um algoritmo de aprendizagem não supervisionada. A técnica tenta descobrir estruturas não lineares presentes nos dados, explorando simetrias locais através de reconstruções lineares. Assim como na técnica de PCA, sua utilidade para detecção de falhas advém do fato de que a técnica é capaz de reduzir a dimensionalidade dos dados originais. Todavia, diferentemente da técnica de PCA, a técnica de Abordagem Linear Local possui a vantagem de ser não linear.

3.5.3 Análise de Subespaços Estacionários

Na Análise de Subespaços Estacionários (Stationary Subspace Analysis) um conjunto de dados multivariados é decomposto em componentes estacionários e não estacionários (VON BÜNAU *et al.*, 2009). Compreender os efeitos não estacionários constitui um dos principais desafios na análise de dados, já que em muitas aplicações os dados podem ser encarados como uma mistura de sinais estacionários e não estacionários. Especialmente no caso de detecção de falhas em processo da indústria

química, esta técnica parece ser bastante promissora, já que ela permite decompor os dados originais em componentes estacionários e não estacionários. O monitoramento de processos em regime transiente ainda é um dos grandes desafios para a disciplina de detecção de falhas baseada puramente em dados.

3.5.4 Identificação de Sistemas Híbridos

O termo engloba uma classe de modelos usados para descrever processos contínuos que ocasionalmente apresentam comportamento descontínuo, causado por mudanças na dinâmica (VIDAL *et al.*, 2016). Trata-se de uma categoria de modelos que engloba vários submodelos, cujos parâmetros são identificados simultaneamente com a função de chaveamento que determina quando cada um dos submodelos está ativo. Os submodelos poderiam, por exemplo, ser modelos Autoregressivos com Entradas Exógenas (Modelos ARX) (PILLONETTO, 2016).

3.6 Comentários Finais

No presente capítulo apresentou-se uma revisão breve de métodos de monitoramento de processos químicos, tais como PCA e suas variantes. Como a área de monitoramento de processos é bastante ampla, tanto em número de publicações quanto em número de métodos de detecção, pode-se dizer que as técnicas discutidas nas Seções 3.1, 3.2 e 3.3 são apenas uma introdução às estratégias clássicas de monitoramento de processos. Entretanto, técnicas mais modernas, como as apresentadas nas Seções 3.4 e 3.5, receberam pouca atenção dos pesquisadores da área de monitoramento de processos ou mesmo da indústria.

Além de apresentar as técnicas clássicas, cada um dos exemplos abordados no Capítulo 3 serviu para construir uma visão crítica da área. Vários pontos pouco abordados na literatura foram apresentados, tais como questões relativas à frequência de amostragem, estatísticas de detecção de falhas e ajuste de parâmetro dos modelos (número de componentes retidos, atraso e função núcleo). Outro ponto bastante interessante e contraditório, intrínseco ao processo de modelagem das técnicas baseadas em dados, está relacionado à importância de conhecimento acerca do domínio para garantir que o modelo do processo seja representativo. Este ponto constitui um paradigma relevante para o desenvolvimento de sistemas baseados em dados, já que muitas vezes o usuário é induzido a pensar que o conhecimento do domínio é dispensável.

Capítulo 4

Avaliação da Técnica DPCA Para Monitoramento de Processos

Este capítulo apresenta um estudo a respeito da aplicação de um sistema de monitoramento para detecção de condições anormais em um reator contínuo, em que ocorre a bem conhecida reação de Van der Vusse (TRIERWEILER, 1997). Foram utilizados dados obtidos por meio de simulação em condições normais e anormais. A modelagem usada para o desenvolvimento do sistema de monitoramento foi baseada na técnica de DPCA. Para estabelecer o limiar de detecção de falhas das estatísticas \mathbf{T}^2 e \mathbf{Q} , foi utilizada, além das equações apresentadas no Capítulo 3, a técnica de Estimativa de Densidade Por Funções Núcleo. O objetivo principal deste Capítulo é apresentar um ensaio crítico das condições que permeiam a implementação desses sistemas de monitoramento em um problema real.

4.1 O CSTR com Reação de Van der Vusse

A garantia do desempenho de operação de um reator é de grande importância para a viabilidade econômica das plantas industriais, já que, além dos produtos de interesse, os reatores também geram subprodutos potencialmente perigosos para a saúde e para o meio ambiente (TRIERWEILER, 1997). No esquema da reação de Van de Vusse, ciclopentanol (B) é produzido a partir do ciclopentadieno (A), com produção dos subprodutos ciclopentanodiol e dicitlopentadieno (D) por meio de um mecanismo que envolve uma reação consecutiva e uma reação paralela. De acordo com TRIERWEILER (1997), o mecanismo cinético desta reação é apresentado na forma:



A Figura 4.1 apresenta um diagrama esquemático do reator. Admitindo que a densidade da mistura reacional ρ e o volume do reator V são constantes, tal sistema pode ser modelado pelo seguinte conjunto de equações não lineares (TRIERWEILER, 1997):

$$\frac{dC_a}{dt} = \frac{F}{V}(C_{a,inp} - C_a) - K_1(T)C_a - K_3(T)C_a^2 \quad (4.3)$$

$$\frac{dC_b}{dt} = -\frac{F}{V}C_b + K_1(T)C_a + K_2(T)C_b \quad (4.4)$$

$$\frac{dC_c}{dt} = -\frac{F}{V}C_c + K_2(T)C_b \quad (4.5)$$

$$\frac{dC_d}{dt} = -\frac{F}{V}C_d + \frac{K_3(T)C_a^2}{2} \quad (4.6)$$

$$\frac{dT}{dt} = \frac{F}{V}(T_{inp} - T) + \frac{k_w A_r}{V \rho C_p}(T_k - T) - \frac{K_1(T)C_a \Delta H_1}{\rho C_p} - \frac{K_2(T)C_b \Delta H_2}{\rho C_p} - \frac{K_3(T)C_a^2 \Delta H_3}{\rho C_p} \quad (4.7)$$

$$\frac{dT_k}{dt} = \frac{1}{m_k C_{pk}} [Q_k + k_w A_r (T - T_k)] \quad (4.8)$$

$$K_i(T) = K_{i0} e^{\frac{E_i}{T[^\circ C] - 273,15}} \quad (4.9)$$

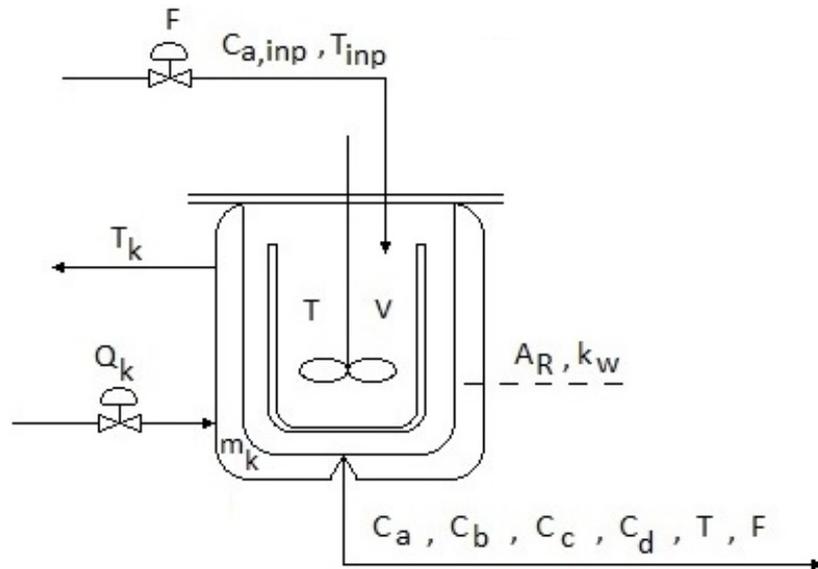


Figura 4.1: Representação Esquemática do Processo Estudado com a Reação de Van der Vusse. Adaptado de TRIERWEILER (1997)

As concentrações das substâncias A, B, C e D são respectivamente C_a , C_b , C_c

e \mathbf{C}_d . A temperatura do reator é representada por \mathbf{T} e a temperatura da jaqueta de resfriamento é representada por \mathbf{T}_k . Admitiu-se que as constantes cinéticas \mathbf{K}_i dependem da temperatura por meio da Equação de Arrhenius (Equação (4.9)). Os valores dos parâmetros físico-químicos usados para simular o modelo são apresentados nas Tabelas 4.1 e 4.2 (TRIERWEILER, 1997).

Tabela 4.1: Parâmetros Cinéticos da Reação de Van der Vusse (TRIERWEILER, 1997)

Reação	Fator de Colisão	Unidade \mathbf{K}_{i0}	\mathbf{E}_i [k]	$\Delta\mathbf{H}_i$ [kJ/mol]
$A \xrightarrow{K_1} B$	$(1,287 \pm 0,04) \times 10^{12}$	h^{-1}	-9758,3	$4,2 \pm 2,36$
$B \xrightarrow{K_2} C$	$(1,287 \pm 0,04) \times 10^{12}$	h^{-1}	-9758,3	$-(11,0 \pm 1,92)$
$2A \xrightarrow{K_3} D$	$(9,043 \pm 0,27) \times 10^9$	litros/mol h	-8560	$-(41,85 \pm 1,41)$

Tabela 4.2: Parâmetros do Reator Usado na Simulação de Reação da Reação de Van der Vusse (TRIERWEILER, 1997)

Parâmetro	Símbolo	Valor	Unidade
Densidade	ρ	$0,9342 \pm 0,0004$	[kg/l]
Capacidade Calorífica	\mathbf{C}_p	$3,01 \pm 0,04$	[kJ/kg K]
Coef. Tranf. Calor	k_w	4032 ± 120	[kJ/ m ² h K]
Área de Troca da Jaqueta	\mathbf{A}_r	0,215	[m ²]
Volume do Reator	\mathbf{V}_r	10	[l]
Massa do Refrigerante	m_k	5,0	[kg]
Coef. Tranf. Calor Refrigerante	\mathbf{C}_{pk}	$2,0 \pm 0,05$	[kJ/ kg K]

Uma análise dos graus de liberdade do reator mostra que o sistema possui 16 parâmetros (\mathbf{V} , \mathbf{K}_{10} , \mathbf{E}_1 , $\Delta\mathbf{H}_1$, \mathbf{K}_{20} , \mathbf{E}_2 , $\Delta\mathbf{H}_2$, \mathbf{K}_{30} , \mathbf{E}_3 , $\Delta\mathbf{H}_3$, k_w , \mathbf{A}_r , ρ , \mathbf{C}_p , \mathbf{C}_{pk} e m_k), 6 estados (\mathbf{C}_a , \mathbf{C}_b , \mathbf{C}_c , \mathbf{C}_d , \mathbf{T} e \mathbf{T}_k) e 4 variáveis livres ($\mathbf{C}_{a,inp}$, \mathbf{T}_{inp} , \mathbf{F} e \mathbf{Q}_k). Contudo, a vazão de entrada (\mathbf{F}) e o calor trocado com a jaqueta (\mathbf{Q}_k) são variáveis manipuladas para controlar, respectivamente, a concentração de B (\mathbf{C}_b) e a temperatura do reator (\mathbf{T}). Assim, restam apenas duas variáveis livres ou distúrbios: a temperatura de entrada (\mathbf{T}_{inp}) e a concentração do reagente A na corrente de entrada ($\mathbf{C}_{a,inp}$). Com o intuito de realizar a simulação do reator em malha fechada, para obtenção de dados e a consequente monitoração de processo, foi usado o controlador QDMC apresentado no trabalho de OLIVEIRA e LOPES (2014). A função básica do QDMC é garantir o controle da concentração de ciclopentanol (\mathbf{C}_b) e da temperatura do reator (\mathbf{T}), manipulando a taxa de diluição ($\mathbf{f} = \mathbf{F}/\mathbf{V}$) e

a taxa de remoção de calor do reator (Q_k), mesmo na presença de distúrbios (T_{inp} , $C_{a,inp}$).

O controlador QDCMC usado na simulação do reator é um tipo de Controlador Preditivo Baseado em Modelo (MPC). Segundo BEQUETTE (2003), o Controle Preditivo Baseado em Modelo (MPC) é uma das técnicas de controle avançado mais utilizadas na indústria, já que ela se adapta bem a sistemas multivariáveis com atraso e é capaz de lidar com restrições de forma bastante natural. O termo MPC é empregado em uma ampla gama de métodos de controle que utilizam um modelo matemático da planta para obter as ações de controle através da minimização de uma função objetivo (CAMACHO e BORDONS, 2007). O MPC usado neste trabalho tem sua formulação baseada no modelo de resposta ao degrau (FSR) e, ao utilizar este modelo, admite-se que um sistema real não reage instantaneamente a variações nas variáveis manipuladas e que a resposta transitória tende a zero para um tempo maior que o tempo de acomodação do processo. O horizonte de predição do MPC foi ajustado em 60, o horizonte de controle foi ajustado em 5 e foram usados 101 termos na resposta ao degrau.

4.2 Identificação do Modelo Para Monitoramento

Como discutido no Capítulo 2, a área de monitoramento de falhas é bastante ampla em termos de modelos e estratégias de monitoramento. Porém, dadas as características dos dados da indústria de processo, algumas técnicas se adaptam melhor para o desenvolvimento de modelos de monitoramento. Processos químicos são dinâmicos, apresentando correlação tanto entre variáveis diferentes quanto entre uma mesma variável e sua versão atrasada no tempo. Grande parte das vezes, alterações nas matrizes de correlações acontecem antes que um desvio operacional de alguma variável de processo possa ser efetivamente detectado (KOURTI, 2003), justificando o uso de técnicas como a Análise Dinâmica de Componentes Principais para a construção de um modelo para monitoramento. A técnica de DPCA é uma das técnicas de modelagem dinâmica multivariável mais conhecida e mais estudada na literatura; por esta razão, ela foi a técnica usada para o monitoramento do reator apresentado na seção anterior e para construção de uma análise crítica sobre as condições de implementação do sistema de detecção de falhas.

Para identificar o modelo DPCA e testá-lo para detecção de falhas, foi gerado um conjunto de dados de simulação de aproximadamente 80h de operação, registrado com um período de amostragem de 35s. Com o objetivo de excitar o sistema, os *set-points* da temperatura no reator e da concentração de ciclopentanol receberam, cada um, três perturbações do tipo degrau, com amplitudes variáveis, como apresentado na Figura 4.2.

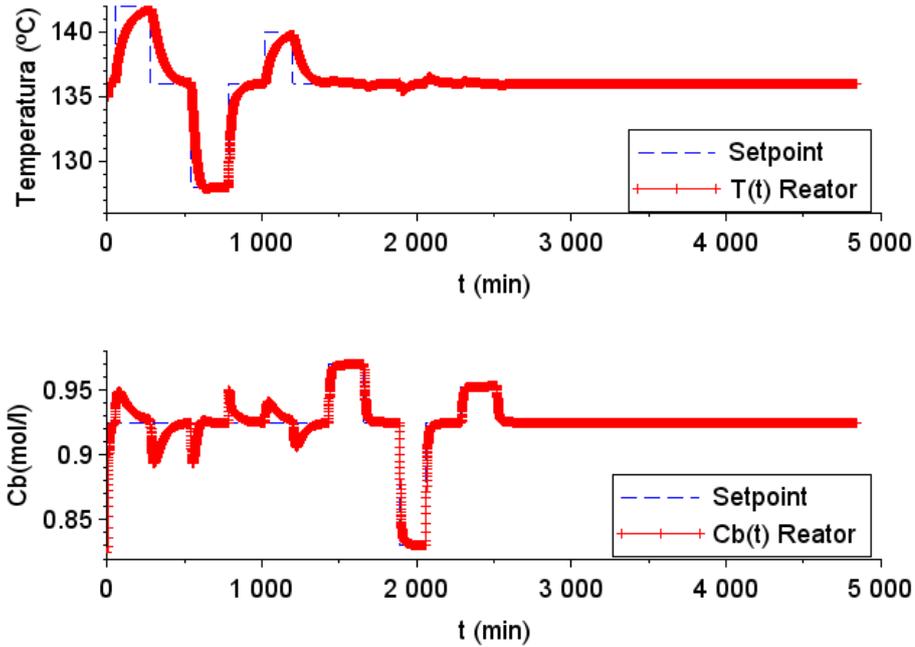


Figura 4.2: Perturbações Usadas na Simulação do Reator de Van der Vusse

Para modelagem com a técnica DPCA foram usadas as variáveis listadas na Tabela 4.3. Na análise não foram incluídas a vazão de entrada nem o calor trocado com a jaqueta, já que estas são variáveis manipuladas. Para representar o ruído dos sensores e do processo, cada uma das variáveis da Tabela 4.3 foi contaminada com ruído gaussiano de média zero com distribuições independentes e variância de **0,001** (C_a , C_b e C_d), **0,01** (C_c), **0,1** ($C_{a,inp}$) e **0,5** (T , T_k e T_{inp}). A evolução temporal das variáveis contaminadas com ruído é mostrada na Figura 4.3, os histogramas são mostrados na Figura 4.4 e a matriz de correlações é apresentada por meio de um mapa de calor em escala de cinza na Figura 4.5.

Tabela 4.3: Variáveis Usadas na Modelagem DPCA do Reator de Van der Vusse

Variável	Descrição	Média	Desv. Pad.	Unidade
C_a	Concentração de A	1,29	0,057	[mol/l]
C_b	Concentração de B	0,93	0,025	[mol/l]
C_c	Concentração de C	2,33	0,096	[mol/l]
C_d	Concentração de D	0,28	0,008	[mol/l]
$C_{a,inp}$	Concentração de A na Corrente de Entrada	5,10	0,099	[mol/l]
T	Temperatura do Reator	409,19	2,50	[K]
T_k	Temperatura da Jaqueta	404,63	3,54	[K]
T_{inp}	Temperatura da Corrente de Entrada	403,16	0,50	[K]

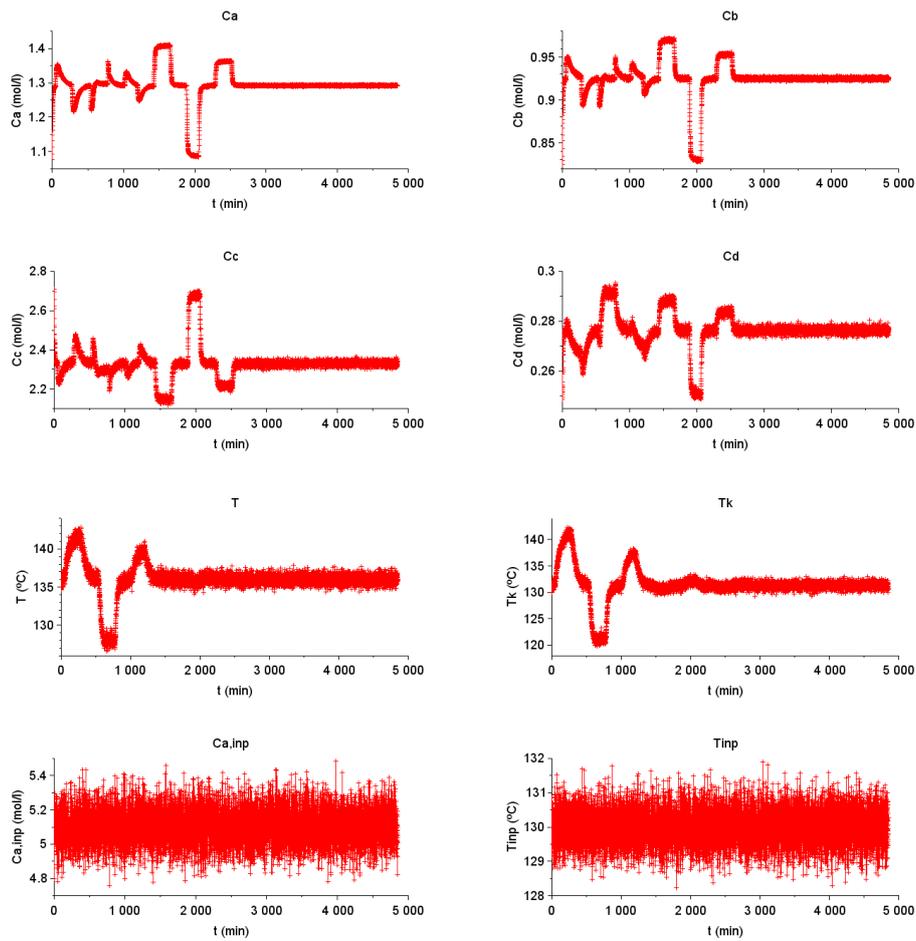


Figura 4.3: Evolução temporal das variáveis usadas na modelagem do Reator de Van der Vusse

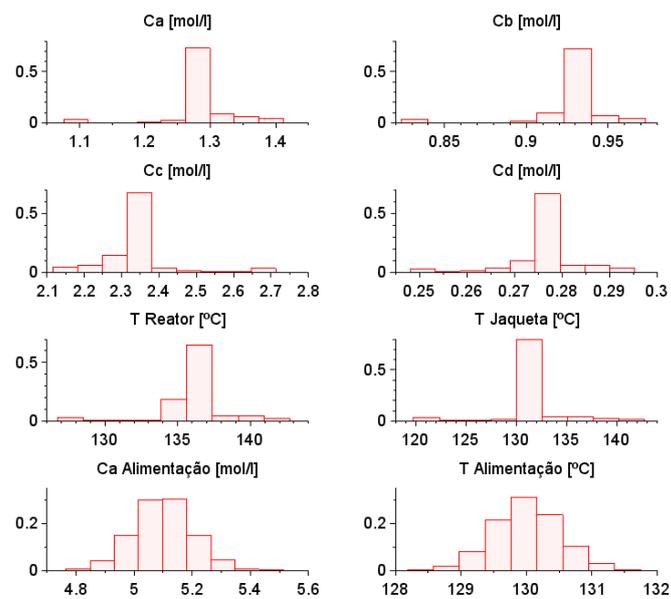


Figura 4.4: Histogramas das Variáveis Usadas no Modelo DPCA do Reator de Van der Vusse

No mapa de calor, a correlação máxima é representada pela cor branca, de modo que cores claras representam altas correlações. Com auxílio do mapa de calor, fica evidente que existem correlações significativas entre várias das variáveis monitoradas, o que corrobora a escolha de métodos de monitoramento como PCA e suas extensões (DE KETELAERE *et al.*, 2015). Nota-se que as concentrações C_a , C_b , C_c e C_d são fortemente correlacionadas, em razão do mecanismo cinético da reação de Van der Vusse. Também é possível perceber que o controlador é capaz de desacoplar relativamente bem a temperatura das concentrações. Isto era esperado já que o controle QDMC, diferentemente de uma estratégia de controle proporcional, derivativo e integral, trata a planta de forma integrada. Nota-se ainda a relação que existe entre a temperatura da jaqueta e a temperatura do reator. Finalmente, é importante comentar que, como as variáveis de distúrbio $C_{a,inp}$ e T_{inp} apresentam valores praticamente constantes em torno da média, elas não apresentam correlações significativas com as demais variáveis.

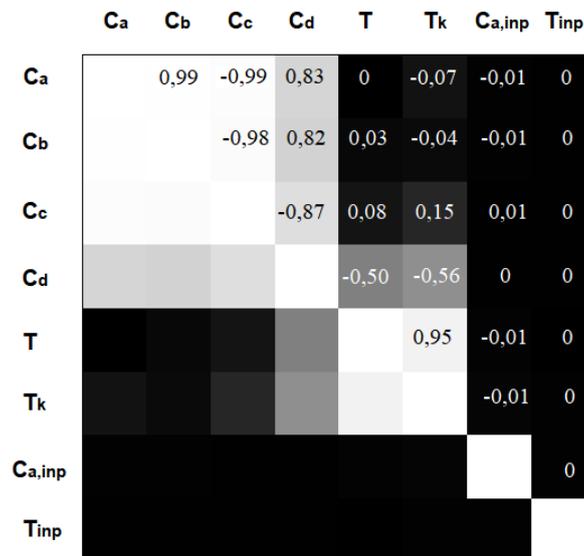


Figura 4.5: Mapa de Calor da Matriz de Correlações das Respostas das Variáveis do Reator de Van der Vusse

A Figura 4.6 apresenta os gráficos *SCREE* e os gráficos da variância explicada para o atraso variando entre 0 e 2. Nos gráficos *SCREE*, a reta azul determina o limiar igual 1,0 e a reta verde determina o limiar igual a 0,7, usados para avaliar se determinado valor característico pode ser considerado significativo.

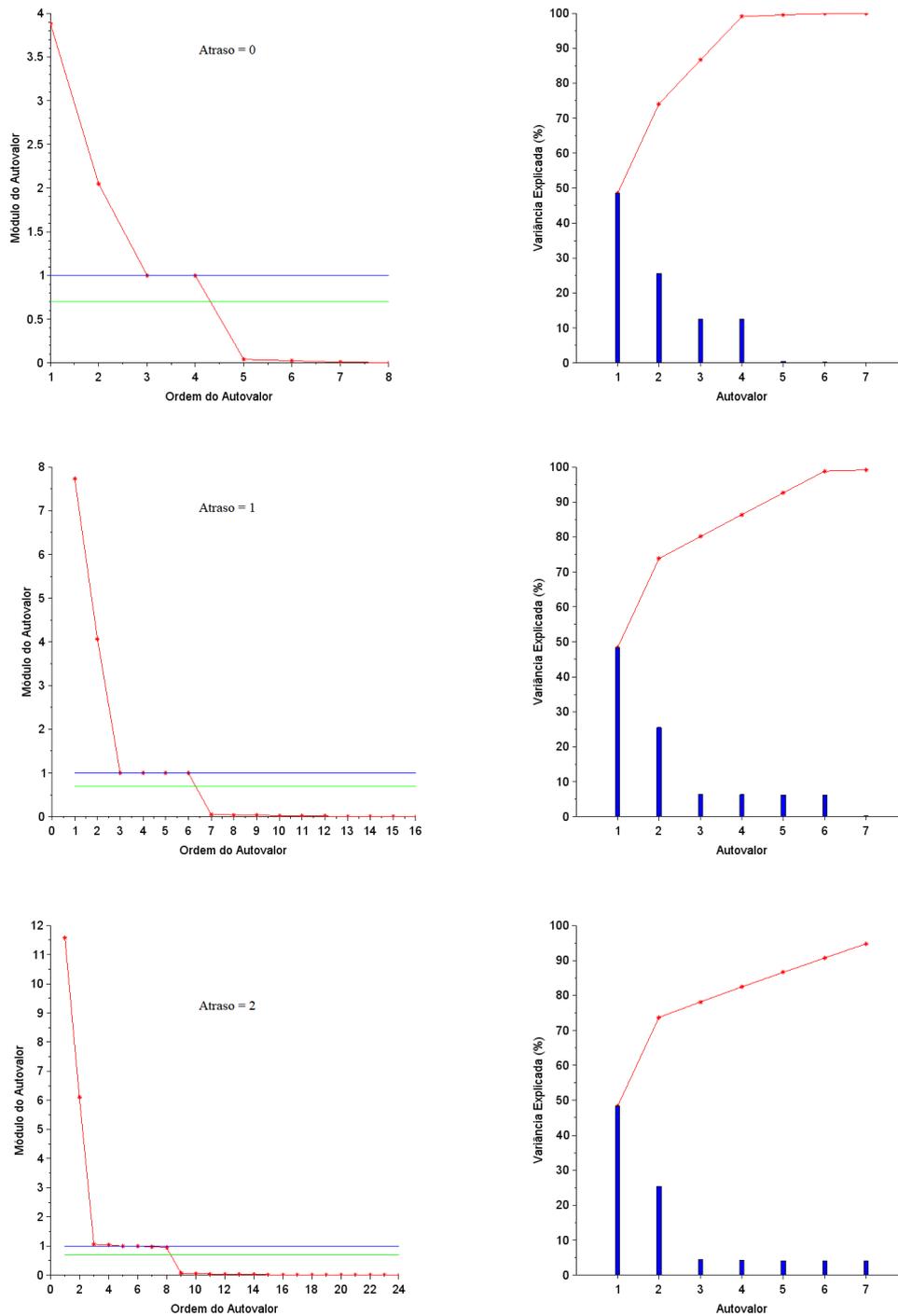


Figura 4.6: Gráfico SCREE (Esquerda) e Gráfico de Variância Explicada (Direita) para o Modelo DPCA do Reator de Van der Vusse

As Tabelas 4.4 e 4.5 apresentam os resultados obtidos a partir da aplicação do critério de seleção do atraso, conforme procedimento proposto por KU *et al.* (1995). Nos resultados apresentados da Tabela 4.4 foi considerado um limiar igual a 0,7 e na Tabela 4.5 foi considerado um limiar igual a 1,0. Para o limiar igual a 0,7, a Tabela 4.4 indica que o atraso seria igual a 1, enquanto o respectivo gráfico na Figura 3.10 indica que seriam necessários 6 componentes principais, explicando

aproximadamente 95% da variância. Para o limiar igual a 1,0, a Tabela 4.5 indica que o atraso também seria igual a 1,0, enquanto o respectivo gráfico na Figura 4.6 indica que seriam necessários 2 componentes principais, explicando aproximadamente 75% da variância.

Tabela 4.4: Seleção do atraso para o Reator de Van der Vusse com limiar de 0,7

Atraso (L)	$r(L)$	$\sum_{i=0}^{L-1} (L - i + 1)r_{\text{novas}}(i)$	r_{novas}
0	4	-	4
1	10	$4(1 + 1) = 8$	2
2	16	$4(2 + 1) + 2(2 - 1 + 1) = 16$	0

Tabela 4.5: Seleção do atraso para o Reator de Van der Vusse com limiar de 1,0

Atraso (L)	$r(L)$	$\sum_{i=0}^{L-1} (L - i + 1)r_{\text{novas}}(i)$	r_{novas}
0	6	-	6
1	14	$6(1 + 1) = 12$	2
2	22	$6(2 + 1) + 2(2 - 1 + 1) = 22$	0

Para a modelagem do reator, apesar do atraso determinado a partir dos dois limiares ser o mesmo e igual a um, chama a atenção o fato de que o número de componentes principais é diferente e, logicamente, o percentual de variância explicada. A variância explicada para o critério de seleção igual a 1,0 (75% com 2 componentes principais) é bem inferior que a variância explicada pelo critério de seleção de 0,7 (95% com 6 componentes principais). A diferença é obviamente devida à enorme diferença do número de direções principais. Assim, é natural questionar qual critério adotar na modelagem. Ou, caso seja adotado o limiar igual a 1,0, se não seria razoável empregar mais que dois componentes principais, com a finalidade de aumentar o percentual de variância explicada. Estes são dilemas típicos enfrentados na modelagem a partir de técnicas *Data Driven*, já que via de regra é necessário selecionar valores de ajuste para os parâmetros dos algoritmos com auxílio de heurísticas. Os resultados apresentados reforçam o fato de que o atraso no fundo acaba sendo tratado como um parâmetro de sintonia do modelo. No restante deste capítulo adotou-se o atraso igual a 1 e foram usados sete componentes principais, de modo a explicar 99% da variabilidade dos dados.

Para investigar as distribuições das estatísticas \mathbf{T}^2 e \mathbf{Q} , foi usada a técnica

de Estimação de Densidade Por Funções Núcleo (KDE) (SILVERMAN, 1998). As distribuições obtidas são mostradas na Figura 4.7. Os limites de detecção das duas estatísticas foram calculados a partir da respectiva função densidade de probabilidades e a partir das equações apresentadas no Capítulo 3. A Tabela 4.6 resume os resultados encontrados.

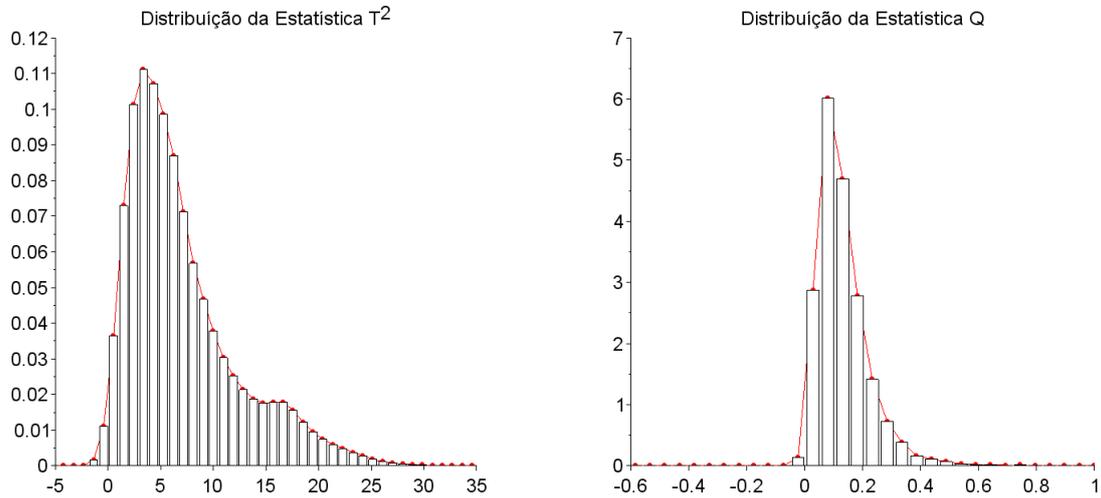


Figura 4.7: Funções Densidade de Probabilidade das Estatísticas \mathbf{T}^2 e \mathbf{Q} no Problema de Modelagem DPCA do Reator Van der Vusse

Note que a estatística \mathbf{T}^2 , diferentemente de uma distribuição \mathbf{F} , é multimodal, o que talvez explique o motivo pelo qual o limite de detecção calculado pela técnica de KDE seja 25% maior que o obtido a partir de uma distribuição \mathbf{F} . Na prática, esta diferença significa um número menor de alarmes ao monitorar o reator. Acredita-se que a multimodalidade da estatística \mathbf{T}^2 seja consequência das excitações do tipo degrau usadas para estimular o reator.

Tabela 4.6: Limites de Detecção das Estatísticas \mathbf{T}^2 e \mathbf{Q} no Problema de Modelagem DPCA do Reator Van der Vusse

\mathbf{T}^2_{\max} (Dist. \mathbf{F})	\mathbf{T}^2_{\max} (KDE)	\mathbf{Q}_{\max} (Dist. \mathbf{N})	\mathbf{Q}_{\max} (KDE)
14,09	17,60	0,30	0,28

Nas Figuras 4.8 e 4.9 apresentam-se as cartas de controle das estatísticas \mathbf{T}^2 e \mathbf{Q} . Nas figuras, a linha vermelha contínua é o limite de detecção, calculado admitindo-se independência temporal e normalidade das projeções. Deste modo, no caso da estatística \mathbf{T}^2 , o limite de detecção foi calculado a partir de uma distribuição \mathbf{F} e, no caso da estatística \mathbf{Q} , a partir de uma distribuição \mathbf{N} . A linha verde pontilhada é o limite de detecção obtido com auxílio da técnica KDE. A curva azul representa

os dados de treinamento e a curva magenta representa os dados de teste. Com o objetivo de minimizar o número de falsos alarmes, os limites de detecção foram calculados para um nível de significância de 95%.

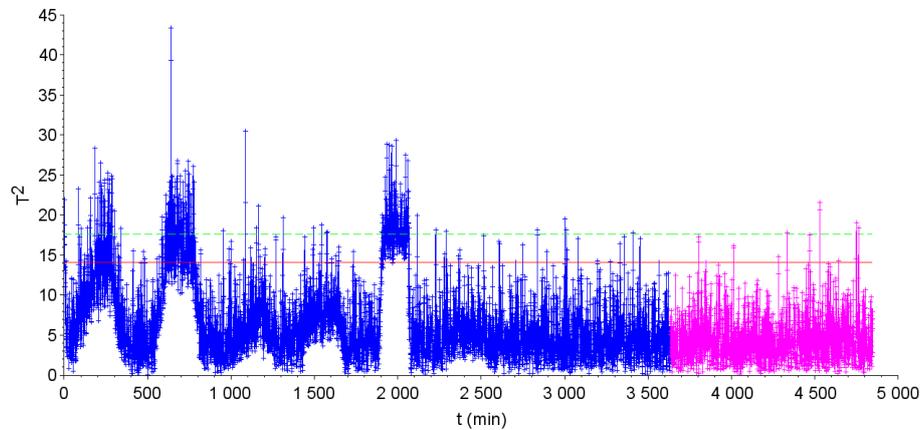


Figura 4.8: Gráfico de Controle T^2 do Problema de Modelagem DPCA do Reator Van der Vusse

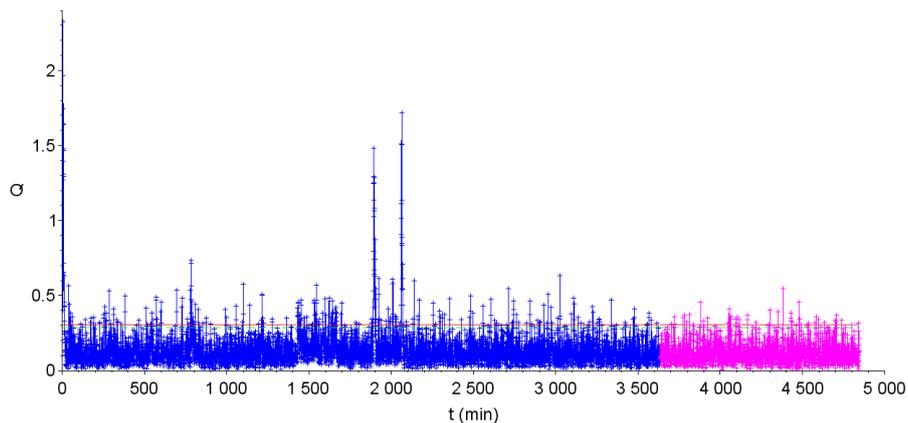


Figura 4.9: Gráfico de Controle Q do Problema de Modelagem DPCA do Reator Van der Vusse

A Tabela 4.7 apresenta uma comparação entre o número de alarmes falsos na fase de treinamento. Esperava-se que o percentual de alarmes falsos ficasse em torno de 5%, já que o nível de significância usado nos cálculos dos limites de detecção foi de 95%. Não obstante, observa-se um percentual de 10,76% de falsos alarmes para a estatística T^2 , quando o limite de detecção foi calculado admitindo-se uma distribuição F . Esse resultado pode ser explicado porque a estatística T^2 apresenta uma distribuição multimodal (Figura 4.7), fato este que não é levado em conta quando se calcula o limite de detecção por meio de uma distribuição F . Diferentemente da estatística T^2 , a estatística Q é semelhante a sua distribuição teórica, de modo que o limiar de detecção calculado pela técnica de KDE é bem próximo do limiar de detecção calculado pelas equações clássicas.

Tabela 4.7: Alarmes Falsos (Significância de 95%) no Problema de Modelagem DPCA do Reator Van der Vusse

T^2 (Dist. F)	T^2 (KDE)	Q (Dist. N)	Q (KDE)
12,13%	4,74%	5,00%	4,47%

Analisando a Figura 4.10, fica realmente evidente a natureza multimodal dos dados. Nesta figura, os pontos usados no treinamento foram projetados nos dois primeiros componentes principais, sendo possível identificar sete regiões específicas de operação. Como o modelo obtido com a técnica DPCA é construído a partir da matriz dos desvios calculados em relação à média, pode-se dizer que o modelo não é ótimo para nenhum ponto de operação específico, o que por si só constitui uma desvantagem da técnica. Muitas técnicas têm sido propostas na literatura para atacar esse problema particular, como uma soma de modelos que podem ser usados individualmente em cada *cluster* de dados ou como uma soma de modelos construídos localmente para ter um caráter global e multimodal. Deve ficar claro que essas abordagens pretendem sempre memorizar as condições eventualmente observadas e as poucas trajetórias que as conectam.

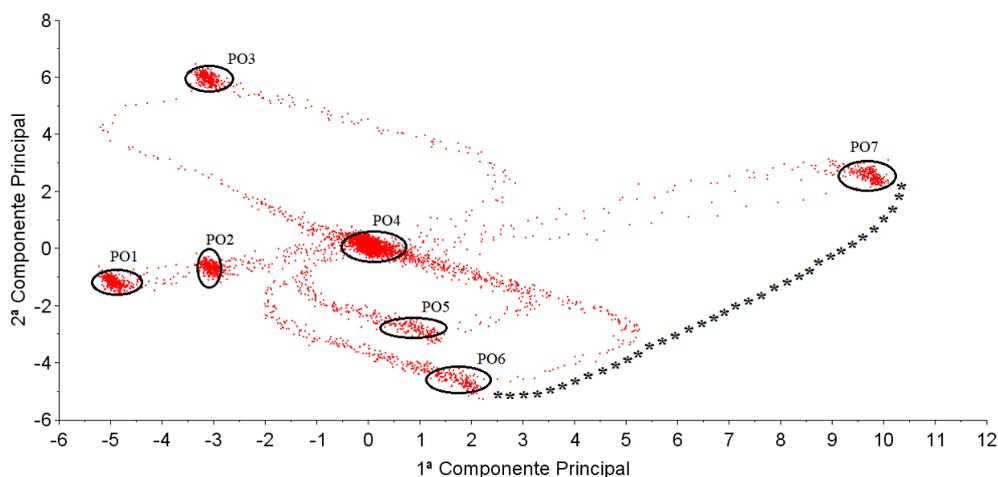


Figura 4.10: Projeções nos Dois Primeiros Componentes Principais no Problema de Modelagem do Reator Van der Vusse

Por exemplo, a trajetória marcada na Figura 4.10 com "asteriscos" constitui uma transição possível e não realizada no conjunto de dados que o modelo seria incapaz de aceitar como válida. No caso dos modelos multimodais as trajetórias dinâmicas que caracterizam as transições poderiam ser consideradas como espúrias, por serem pouco frequentes. Esses problemas tornam lícito o questionamento sobre o tipo de modelagem dinâmica que de fato é possível estabelecer com a técnica de DPCA em sistemas de grande dimensão, em que as transições podem ser muito rarefeitas no espaço de fases, tornando os modelos pouco representativos da riqueza dinâmica

que o processo pode apresentar. Nesse contexto, e uma vez mais, o conhecimento do domínio pode se revelar útil para validar trajetórias que o simples acúmulo e registro dos dados pode não permitir.

4.3 Monitoramento de Falhas Via DPCA

Com o objetivo de verificar a capacidade de detecção de falhas por intermédio da técnica de DPCA, foram simuladas duas condições de falha da seguinte maneira: a falha 1 foi gerada por mudanças progressivamente maiores na média da temperatura do reator, simulando uma falha no sensor que mede tal grandeza; enquanto a falha 2 simulou uma desativação do catalisador, obtida por meio da manipulação da constante pré-exponencial de Arrhenius da reação de formação de ciclopentanol (B) a partir do ciclopentadieno (A). Como tempo de detecção, considerou-se o momento a partir do qual todos os valores calculados para determinada estatística violaram o limite superior obtido com a técnica de KDE. A falha 1 teve início em $t = 3636\text{min}$ e foi efetivamente detectada pela estatística \mathbf{Q} em $t = 3955\text{min}$ e pela estatística \mathbf{T}^2 em $t = 4049\text{min}$ (Figura 4.11). Nota-se também que desde o início da falha ambas estatísticas passaram a apresentar valores sistematicamente crescentes.

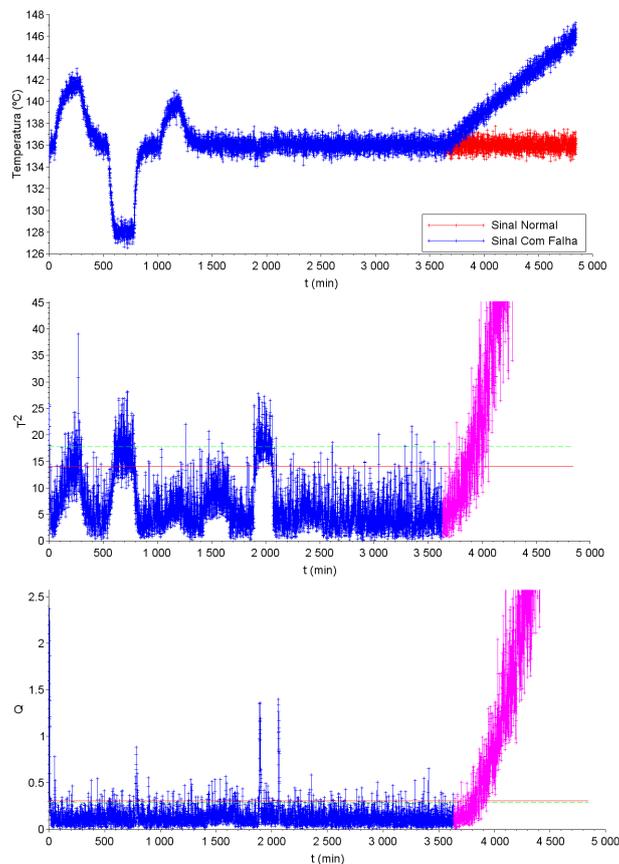


Figura 4.11: Gráficos de Controle no Cenário de Falha 1 no Problema DPCA do Reator Van de Vusse

A Figura 4.12 mostra que a falha 2, desativação do catalisador, teve início em $t = 535\text{min}$ e resultou numa redução gradual da constante K_{10} em aproximadamente 15,5% durante um período de 1860min . Na Figura 4.12, os gráficos das estatísticas foram ampliados para facilitar a visualização. A estatística Q detectou a falha em $t = 691\text{min}$ e a estatística T^2 em $t = 1368\text{min}$. Desde o início da desativação do catalisador, ambas as estatísticas passaram a apresentar valores sistematicamente maiores.

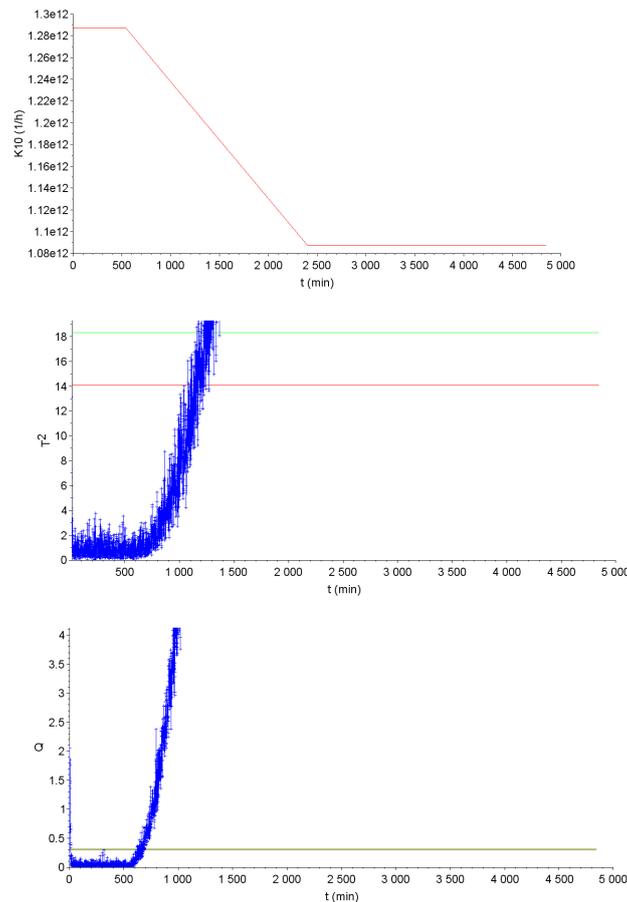


Figura 4.12: Gráficos de Controle no Cenário de Falha 2 no Problema DPCA do Reator Van der Vusse

Analisando o comportamento dos gráficos de controle obtidos para cada uma das falhas apresentadas, a estatística Q foi a que resultou no menor tempo de detecção. Isto indica que o modelo identificado pela técnica de DPCA perde a validade perante os dois cenários de falha. Tal afirmação é mais facilmente compreendida ao se considerar que a estatística Q está relacionada ao erro quadrático de predição e, portanto, indica uma falta de ajuste do modelo para os dois cenários de falha propostos.

Os processos da indústria química são frequentemente caracterizados por múltiplos modos operacionais, causados por variações nas matérias-primas, especificações do produto, *setpoints* ou estratégias de fabricação. Para avaliar a

resposta do modelo frente a uma região de operação nova, o *setpoint* da concentração de ciclopentanol foi alterado para 65% do valor de referência (Figura 4.13).

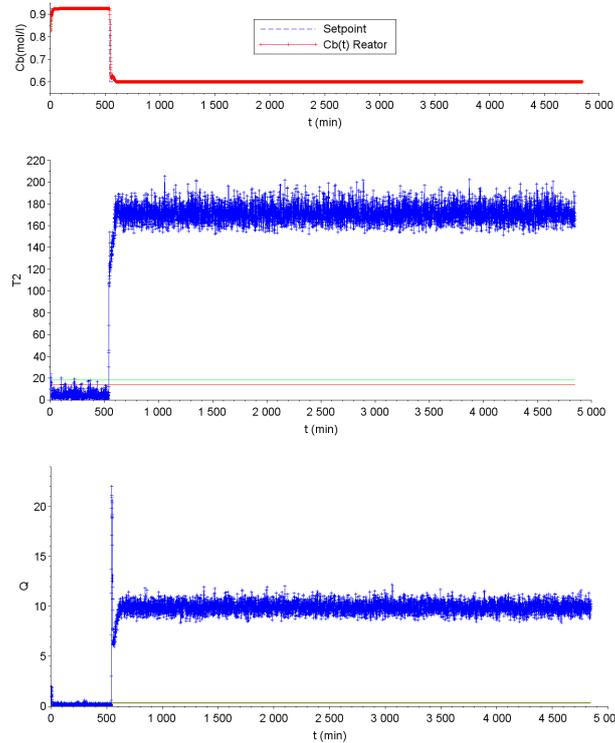


Figura 4.13: Gráficos de Controle no Cenário de Alteração do *Setpoint* de Temperatura no Problema DPCA do Reator de Van der Vusse

Com base na Figura 4.13, fica evidente que ambas as estatística T^2 e Q apresentaram uma mudança abrupta de patamar, violando os limites de controle. Este resultado mostra que a técnica de DPCA apresenta pouca capacidade de extrapolação, já que uma mudança de *setpoint* não deve ser classificada como falha. Assim, conclui-se que de fato a técnica não é capaz de modelar o espaço de fases do reator.

4.4 Comentários Finais

A implementação de procedimentos de monitoramento de processos para a melhoria da eficiência operacional ainda constitui um grande desafio nas aplicações do mundo real. Para a implementação de métodos *Data Driven* na indústria, existem dois cenários distintos possíveis. No primeiro, existe um banco de dados que contém informações históricas de operação do processo. No segundo, o banco de dados não está disponível, sendo necessário realizar experimentos para coletar dados típicos de operação. No estudo realizado neste capítulo, com base em um CSTR onde ocorre a Reação Van der Vusse, adotou-se o segundo cenário. Para unidades de processo de grande escala, que apresentam muitas malhas de controle, este segundo

cenário implica na realização de um elevado número de experimentos para a coleta de dados, o que pode demandar uma quantidade significativa de tempo e custo, além de impactar a segurança da planta. Aspectos como a política operacional da unidade também devem ser levados em conta no planejamento dos experimentos de coleta de dados para a modelagem.

Normalmente existe pouca margem de variação em torno dos *setpoints*, já que grandes perturbações podem causar prejuízos financeiros decorrentes da perda de especificação de produtos ou mesmo riscos relacionados à segurança de pessoas, equipamentos e meio ambiente. Também é importante destacar que a excitação do tipo degrau é um sinal de características pobres para a maioria dos métodos de identificação (AGUIRRE, 2015). Com a adoção crescente dos controladores preditivos baseados em modelo, as plantas de processo têm operado cada vez mais próximas das restrições operacionais, favorecendo a produção perto do ótimo operacional, mas deixando pouca margem de variação dos *setpoints* para realização de experimentos de identificação. Como um exemplo extremo, destaca-se a planta de polipropileno baseada no processo LIPP (DA SILVA ROSA *et al.*, 2012), em que as restrições de *setpoint* de algumas variáveis, como temperatura do reator, são muito rígidas, não sendo admissíveis variações superiores a 0,5 °C na condição de operação.

Uma das conclusões deste capítulo é que os métodos *Data Driven* podem ser entendidos como técnicas de identificação de sistemas e, como tal, envolvem ajustes de parâmetros do modelo e decisões relacionadas à seleção de variáveis de entrada e saída, seleção do método de modelagem que melhor se adapta aos dados disponíveis e onde e como excitar a planta, a fim de obter dados que sejam representativos da dinâmica. Conhecer as características dos dados é um pré-requisito para a seleção de técnicas *Data Driven* mais adequadas ao problema (APPEL *et al.*, 2016). Portanto, assim como na modelagem fenomenológica, na modelagem *Data Driven* é preciso conhecer o sistema de interesse, seja com a ajuda de um especialista ou com o auxílio de técnicas de análise de dados. Por conseguinte, ressalta-se a aparente contradição, dada pela necessidade de conhecimento do domínio para implementar as técnicas. A principal desvantagem dos modelos *Data Driven* advém do fato de que modelos previamente aprendidos não se adaptam bem à regiões operacionais desconhecidas, cujos dados não estavam disponíveis durante a fase de modelagem.

Mesmo quando os dados de modelagem englobam uma multiplicidade de pontos operacionais, o desenvolvimento de modelos multimodais requer algoritmos bem mais complexos que os estudados nos capítulos anteriores, como pode ser observado nos trabalhos de AFZAL *et al.* (2017) e YANG *et al.* (2017). A quantidade de parâmetros de ajuste aumenta bastante e, não raramente, a etapa de modelagem envolve alguma estratégia de agrupamento prévio dos dados (*Cluster Analysis*),

como observado nos trabalhos de FEITAL *et al.* (2013) e ZHANG e ZHAO (2017). Assim, é mais que justificável o investimento em desenvolvimento, implantação e suporte de novas tecnologias de monitoramento de procesos químicos.

Para indústria química em geral, mas especialmente na área farmacêutica e de química fina, o monitoramento de processos multimodais é fundamental, dada a crescente competição causada pelo surgimento de novas potências emergentes. Só para se ter noção da dimensão das mudanças que a indústria tem enfrentado, a maior parte do crescimento da indústria química nos últimos 25 anos foi impulsionada pela Ásia, que atualmente responde por metade das vendas de produtos químicos no mundo. Se a tendência atual continuar, já em 2030 os asiáticos serão responsáveis por dois terços do mercado (ATKEARNEY, 200-?). Neste cenário, a indústria química precisará se adaptar para operar com matérias-primas cujas propriedades apresentem um maior grau de incerteza e diversificar sua produção, buscando se tornar mais competitiva, flexível e ágil (PROCESSNET, 2006). Entretanto, para aplicação imediata na indústria, este parece ser um dos principais desafios técnicos enfrentados para o desenvolvimento de modelos *Data Driven*, já que as técnicas disponíveis hoje são de fato incapazes de distinguir uma falha de uma mudança normal no ponto de operação das plantas de processo.

As técnicas *Data-Driven* disponíveis atualmente, como PCA e suas variantes, não apresentam capacidade de realmente modelar o comportamento dinâmico de um processo, no sentido de aprender as restrições impostas pela dinâmica, permitindo ao modelo determinar quais são os estados possíveis de serem atingidos a partir do estado atual no qual a planta se encontra. Tudo indica que, para monitorar uma planta de processo de modo mais assertivo, é preciso distinguir entre mudanças normais do ponto de operação e falhas que impliquem em comportamentos dinâmicos anormais. Neste sentido, parece ser primordial desenvolver técnicas que sejam capazes de modelar o espaço de fase do processo de interesse. Esta foi a principal motivação para o estudo dos Gráficos de Recorrência apresentados no Capítulo 5.

Um aspecto particularmente interessante e observado diz respeito à sensibilidade da técnica de detecção à heurística de construção do modelo. No caso avaliado a redução do número de componentes principais, resultante da manipulação do limiar de importância dos valores característicos, tornou a técnica de DPCA menos capaz de detectar os desvios do processo, aumentando consideravelmente os tempos de detecção de falha. Esse fato parece decorrer do aumento dos valores das estatísticas \mathbf{T}^2 e \mathbf{Q} durante a fase de treinamento, resultante da modelagem menos precisa dos dados. Esse vínculo existente entre a heurística de construção do modelo e o desempenho da detecção de falhas merece maior atenção em trabalhos futuros.

Capítulo 5

Detecção de Falhas Via Matriz de Recorrência

Diante dos problemas associados às técnicas *Data-Driven* disponíveis atualmente para modelar o comportamento dinâmico de um processo, neste capítulo foi desenvolvida uma nova estratégia de monitoramento inspirada no conceito dos Gráficos de Recorrência. Ao desenvolver a metodologia proposta neste capítulo, buscou-se torná-la o mais independente possível do ponto de operação da planta de processo. O capítulo está dividido em quatro seções, que fazem uma introdução ao Método dos Gráficos de Recorrência, apresentam Metodologia de Monitoramento Via Matriz de Recorrência, discutem uma extensão da metodologia proposta e propõem alguns comentários Finais.

5.1 Introdução

O método dos Gráficos de Recorrência (Recurrence Plots - RP), proposto pela primeira vez por ECKMANN *et al.* (1987), é considerado uma ferramenta eficaz para analisar sinais não-lineares e não estacionários produzidos por sistemas dinâmicos. A materialização de um RP é uma matriz de recorrência, que representa os instantes de tempo em que os estados de um sistema se repetem no espaço de fases (WEBBER JR. e MARWAN, 2014). Nos sistemas reais, sujeitos a ruído, é praticamente improvável que o estado de um sistema se repita, sendo uma recorrência definida como um estado \mathbf{x}_j suficientemente próximo a outro estado \mathbf{x}_i . A matriz de recorrência converte as trajetórias multidimensionais do espaço de fase em uma representação bidimensional de recorrências. Matematicamente a matriz de recorrência pode ser definida como (WEBBER JR. e MARWAN, 2014):

$$R(i, j) = \Theta(\varepsilon - \|x_i - x_j\|) \quad (5.1)$$

em que $\mathbf{x}_i \in \mathbf{R}^m$, $\mathbf{i}, \mathbf{j} = 1, 2, \dots, \mathbf{N}$ são vetores que contêm as variáveis de estado do sistema dinâmico, ε é o raio de uma hipersfera que define um limiar de proximidade entre dois vetores do espaço de estado, Θ é a função de Heaviside e $\|\cdot\|$ é uma norma matemática, como por exemplo a distância Euclidiana. Quando a norma entre dois vetores do espaço de estado \mathbf{x}_i e \mathbf{x}_j for menor que ε , $\mathbf{R}(\mathbf{i}, \mathbf{j}) = \mathbf{1}$; caso contrário $\mathbf{R}(\mathbf{i}, \mathbf{j}) = \mathbf{0}$. A matriz de recorrência pode ser exibida visualmente ao se representar $\mathbf{R}(\mathbf{i}, \mathbf{j}) = \mathbf{1}$ por um ponto preto e $\mathbf{R}(\mathbf{i}, \mathbf{j}) = \mathbf{0}$ por um ponto branco. Por definição, $\mathbf{R}(\mathbf{i}, \mathbf{i}) = \mathbf{1}$, implicando que a diagonal principal da matriz de recorrência é formada por uma linha preta. A matriz de recorrência é simétrica, já que $\|\mathbf{x}_i - \mathbf{x}_j\| = \|\mathbf{x}_j - \mathbf{x}_i\|$. De uma certa forma, a matriz de recorrência é uma representação binária da matriz de correlações, indicando pontos próximos e distantes.

WEBBER JR. e MARWAN (2014) defendem que os padrões estruturais da matriz de recorrência revelam características da evolução temporal das trajetórias de um sistema dinâmico. Além de lidar bem com ruído, esta técnica pode ser aplicada em dados não estacionários e em conjuntos de dados relativamente pequenos. Na Figura 5.1, a coluna da esquerda mostra uma série de sinais típicos, enquanto a coluna da direita apresenta o gráfico de recorrências equivalente a cada um destes sinais. Todos os gráficos de recorrência são apresentados como uma imagem de 200 por 200 pixels, já que os sinais a partir do qual as imagens foram geradas possuem 200 amostras. RPs com estruturas periódicas, orientadas diagonalmente, são característicos de sistemas oscilantes, como pode ser observado com auxílio do RP do sinal senoidal da Figura 5.1a. Os RP homogêneos são característicos de sistemas estacionários e autônomos, em que os tempos de acomodação são pequenos em comparação com o tempo apresentado no RP (Figura 5.1b). Áreas brancas ao redor da diagonal principal da matriz de recorrência são causadas por comportamentos não estacionários e podem indicar, por exemplo, sistemas com parâmetros variando lentamente com o tempo (Figura 5.1d). Assim, transições lentas na dinâmica fazem com que os ângulos superior esquerdo e inferior direito dos RPs apresentem áreas brancas.

Admitindo-se que toda a discussão precedente acerca da matriz de recorrência seja válida, a técnica constitui uma ferramenta interessante para a detecção de falhas em sistemas dinâmicos, dispensando ferramentas de modelagem mais convencionais. Na próxima seção, propõe-se uma metodologia para monitoramento de processos inspirada no conceito de matriz de recorrência.

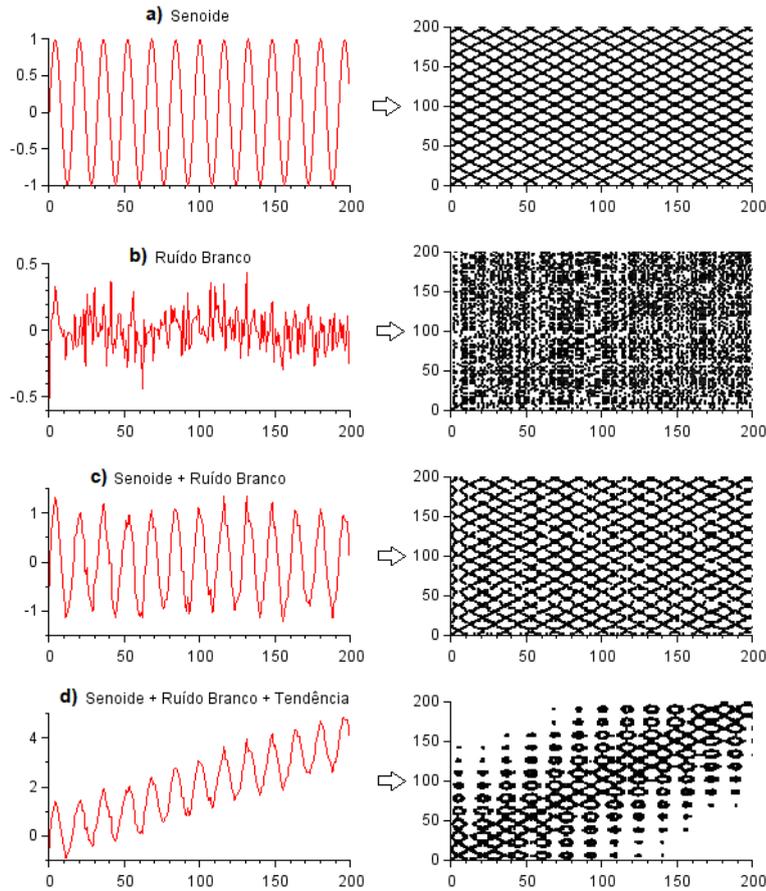


Figura 5.1: Gráficos de Recorrência de Alguns Sinais Típicos

5.2 Metodologia de Monitoramento Via Matriz de Recorrência

Trabalhos relacionado ao monitoramento de processos via matriz de recorrência são escassos. Todavia, a técnica já vem sendo usada para análise de sinais de vibração (ZHOU e ZHANG, 2015), análise de sinais do sistema cardiovascular (ZBILUT *et al.*, 2002) e análise de sinais cerebrais (WEBBER JR. e MARWAN, 2014), dentre outros. A grande maioria dos trabalhos citados se baseia numa série de grandezas calculadas a partir da matriz de recorrência, cujo objetivo é quantificar as estruturas nela existentes (WEBBER JR. e MARWAN, 2014). Tal extração de variáveis costuma ser conhecida como Quantificação de Recorrência. A Quantificação de Recorrência implica em um passo intermediário para obtenção de atributos que propiciem a posterior detecção de falhas, gerando uma sobrecarga computacional desnecessária nos algoritmos. De acordo com a Equação (5.2), a metodologia de monitoramento proposta nesta seção trabalha com matrizes formadas pelas distâncias euclidianas dos vetores \mathbf{y}_i e \mathbf{y}_j , cujas componentes são as M variáveis medidas na planta, ao invés

de operar diretamente com a matriz de recorrência. Trata-se de uma simplificação, cujo objetivo é reduzir o número de parâmetros e simplificar o algoritmo, já que não é necessário ajustar o valor de ε e nem calcular a função de Heaviside.

$$R(i, j) = \|y_i - y_j\| \quad (5.2)$$

Ao operar com matrizes baseadas na distância Euclidiana, espera-se que o algoritmo fique independente do ponto de operação, resolvendo assim uma das grandes dificuldades associadas à técnicas *Data-Driven* e aos sistemas multimodais. A Figura 5.2 fornece uma explicação de como a distância Euclidiana é útil para tornar o algoritmo livre de influências do ponto operacional. No diagrama de dispersão da Figura 5.2, os círculos vermelhos pertencem ao ponto operacional OP1 e os círculos azuis pertencem ao ponto operacional OP2. Cada círculo vermelho e azul está associado um vetor posição, cuja referência é a origem do sistema de coordenadas determinado pelos eixos **Y1** e **Y2**. Ao subtrair dois vetores com posições quaisquer, pertencentes a um determinado ponto operacional, os vetores resultantes, representados pelas setas vermelhas e azuis na figura, se tornam independentes da origem do sistema de coordenadas. Uma vez que o algoritmo proposto opera com distâncias Euclidianas, ou seja, com o módulo dos vetores resultantes da subtração de dois vetores em uma posição qualquer, a independência do ponto operacional fica garantida.

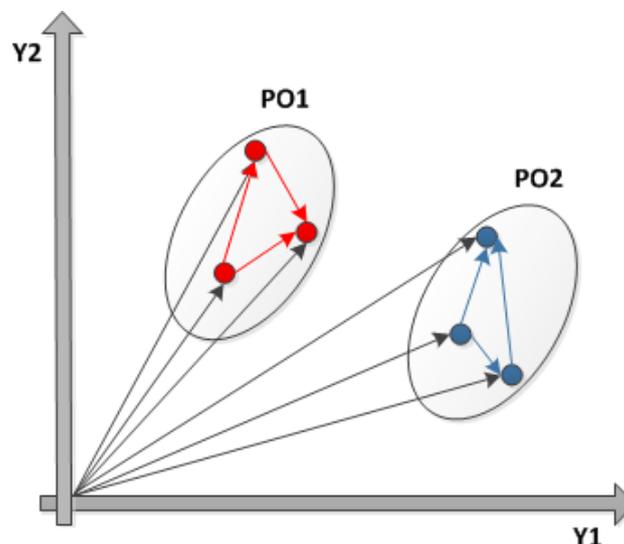


Figura 5.2: Análise da Independência do Ponto Operacional em um Modelo de Distâncias

Dito isto, o primeiro passo para implementação da metodologia de monitoramento a partir do conceito de recorrência consiste em gerar um conjunto de matrizes de distâncias, de modo a representar o padrão normal de operação de uma planta com auxílio de um modelo de referência. A partir daí, a monitoração de

processos pode ser realizada comparando-se novas matrizes de distância às matrizes de referência. Até este ponto, nada foi mencionado sobre como estabelecer um método para comparar uma matriz de distâncias, obtida a partir de dados normais de operação, a uma segunda matriz de distâncias, obtida *a posteriori*, durante a operação subsequente da planta. Buscando preencher esta lacuna, pode-se dizer, a partir de conceitos da Álgebra Linear, que as matrizes de distâncias associadas à condição normal de operação da planta encontram-se em um espaço vetorial e que cada uma destas matrizes é um vetor deste espaço. Sob quais circunstâncias seria possível admitir que dois vetores quaisquer deste espaço estão próximos ou afastados um do outro? A resposta que a Álgebra Linear fornece para tal questão advém do conceito de norma e implica em uma medida de distância que permite comparar vetores e matrizes. Na realidade, o conceito de norma é uma generalização da distância Euclidiana. No caso específico deste trabalho, para permitir comparar duas matrizes de distâncias quaisquer \mathbf{R}_1 e \mathbf{R}_2 adotou-se a Norma de Frobenius (Equação 5.3).

$$NF(R_1, R_2) = \sqrt{\text{Traço}[(R_1 - R_2)(R_1 - R_2)^T]} \quad (5.3)$$

A construção do modelo referência pode ser realizada de acordo com o esquema apresentado na Figura 5.3. Duas janelas, \mathbf{W}_1 e \mathbf{W}_2 , defasadas entre si em \mathbf{D} amostras, percorrem as séries temporais que representam o padrão normal de operação da planta. Antes de avançar sua posição, as duas janelas são tratadas como uma só, sendo convertidas em uma matriz de distâncias (Equação 5.2). No passo seguinte, as duas janelas avançam, de modo a percorrer os sinais temporais, gerando uma sequência de matrizes de distâncias. Na Figura 5.3, quando as duas janelas avançaram, não foi usada qualquer superposição entre elas, mas isto é possível e na verdade indicado, já que o objetivo é usar ao máximo a informação disponível nas séries temporais para gerar o modelo de referência.

Após a obtenção de um conjunto de matrizes de referência, todas as matrizes são comparadas entre si com base na Norma de Frobenius (Equação 5.3). Para cada matriz de distâncias comparada com todas as outras, computa-se a média das normas. De modo a calcular um limiar de detecção para o algoritmo, emprega-se a técnica de KDE para encontrar a função de distribuição de probabilidades da norma média de Frobenius. Até este ponto, todo o esforço realizado teve o objetivo de construir um modelo de referência e determinar um limiar para a métrica de detecção de falhas. A partir de agora, o algoritmo pode operar no modo *online*, permitindo que novas matrizes de distâncias sejam comparadas com o modelo de referência, como mostrado na Figura 5.4. Observando a Figura 5.4, fica claro que o

modelo de referência é uma base de dados e uma função densidade de probabilidades. Na base de dados estão armazenadas as matrizes de distâncias, extraídas de acordo com o procedimento apresentado na Figura 5.3. Portanto, é possível dizer que a técnica proposta é "puramente" *Data Driven*.

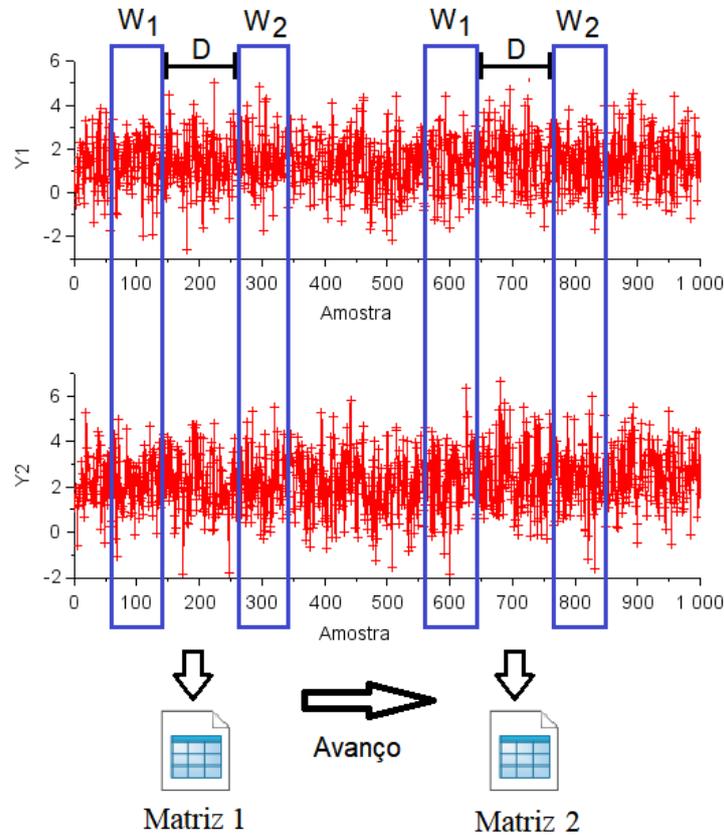


Figura 5.3: Construção do Modelo de Referência por Deslizamento de Janela

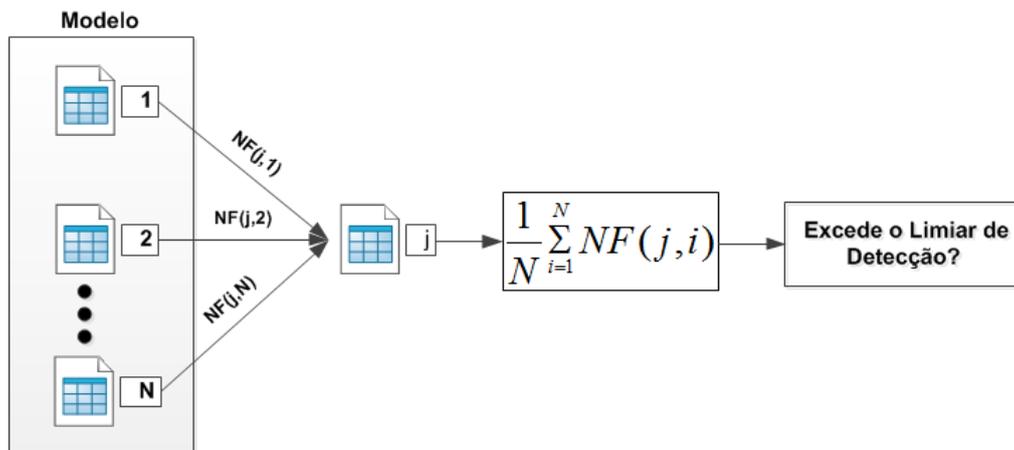


Figura 5.4: Modo *Online*

Matematicamente, a metodologia proposta pode ser generalizada com auxílio dos conceitos de espaço vetorial e transformação de espaços, como ilustrado na Figura

5.5. No espaço vetorial fonte estão contidos os vetores $\mathbf{Y}_i \in \mathbf{R}^M$ com as M variáveis medidas na planta. Por meio da transformação Γ , o espaço fonte é convertido no espaço do modelo, o espaço das matrizes de distância $\mathbf{Z}_i \in \mathbf{R}^{W \times W}$. Uma segunda transformação Λ associa vetores compreendidos no espaço do modelo com vetores do espaço das métricas de detecção de falhas $\mathbf{M}_i \in \mathbf{R}$.

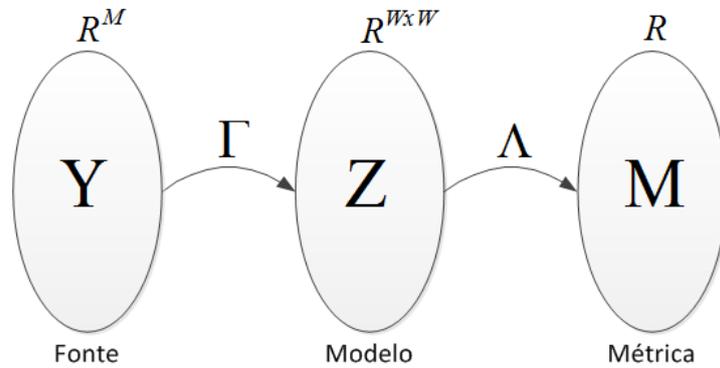


Figura 5.5: Generalização Matemática da Metodologia Proposta

5.2.1 Comparando o Algoritmo Proposto à Técnica de PCA

Além da visão tradicional de que a técnica de PCA é útil para reduzir a dimensionalidade de um conjunto de dados, NARASIMHAN e BHATT (2015) alegam que ela é uma ferramenta de identificação de modelos lineares estacionários. Os autores ainda sugerem que a técnica de PCA é capaz de determinar as relações lineares (restrições) existentes entre as variáveis do sistema a ser identificado. Para compreender o que isto significa, considere o sistema composto por um conjunto de tubulações interligadas apresentado na Figura 5.6, cujas vazões são \mathbf{F}_1 , \mathbf{F}_2 , \mathbf{F}_3 , \mathbf{F}_4 , \mathbf{F}_5 e \mathbf{F}_6 (NARASIMHAN e BHATT, 2015).

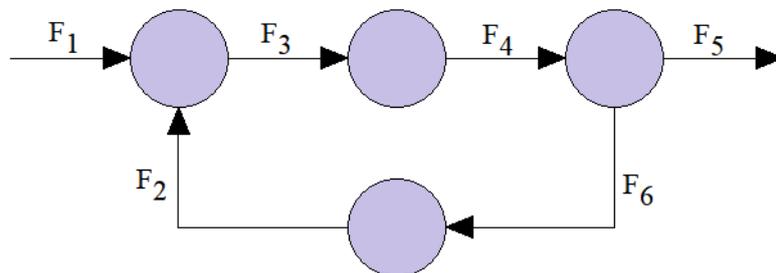


Figura 5.6: Esquema do Conjunto de Tubulações. Adaptado de NARASIMHAN e BHATT (2015)

A partir do momento em que o sistema opera em estado estacionário, não existe acúmulo de fluido em nenhuma das tubulações, sendo possível relacionar as vazões

através da seguinte equação:

$$AF = 0 \quad (5.4)$$

$$A = \begin{bmatrix} 1 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & -1 \\ 0 & -1 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (5.5)$$

$$F = [F_1 \ F_2 \ F_3 \ F_4 \ F_5 \ F_6]^T \quad (5.6)$$

Este sistema apresenta seis variáveis, relacionadas por meio de quatro equações, resultando em dois graus de liberdade. Deste modo, ao aplicar a técnica de PCA seria necessário reter apenas duas componentes principais, gerando uma representação bidimensional dos dados originais. Como os componentes principais são ortogonais entre si, os vetores característicos associados aos menores valores característicos são ortogonais ao espaço principal bidimensional identificado e guardam as relações descritas na matriz \mathbf{A} . Assim, idealmente, para realizar a identificação seria necessário que os dados fossem extraídos de, pelo menos, seis estados estacionários distintos, como apresentado na equação abaixo.

$$\text{Nº de Estados Estacionários} = \dim(\text{Esp. Principal}) + \dim(\text{Esp. Residual}) \quad (5.7)$$

Alternativamente, o algoritmo proposto é capaz de construir um modelo de referência a partir de um único estado estacionário (Figura 5.3), facilitando a aplicação em processos industriais, já que não é necessário perturbar o processo para fazer a identificação do sistema. Para fechar esta seção, na Tabela 5.1 apresenta-se uma comparação entre a metodologia proposta e a técnica de PCA.

Tabela 5.1: Comparação entre PCA e a Metodologia Proposta

PCA	Nova Metodologia
Técnica de modelagem estatística	Técnica de modelagem estatística
A modelagem depende de vários estados estacionários	A modelagem depende em tese de um único estado estacionário
Ótima quando as variáveis de entrada apresentam distribuição normal	Independente da distribuição das medidas das variáveis de entrada
Depende do ponto operacional	Independente do ponto operacional
Multivariável	Multivariável

5.2.2 Exemplo 5.1 - Prova de Conceito

Considere o seguinte sistema composto por duas entradas e duas saídas (AFZAL *et al.*, 2017):

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1,0 & 0,2 \\ 0,3 & 1,0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} \quad (5.8)$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = N\left(\begin{bmatrix} 1 & 2 \end{bmatrix}, \begin{bmatrix} 1,3 & 0,3 \\ 0,3 & 1,5 \end{bmatrix}\right) \quad (5.9)$$

em que \mathbf{e}_1 e \mathbf{e}_2 representam dois ruídos independentes, com distribuições gaussianas, média $\mu = \mathbf{0}$ e variância $\sigma = \mathbf{10}^{-3}$. Para obter dados normais de operação do sistema, foram simuladas 2000 amostras, como apresentado na Figura 5.7.

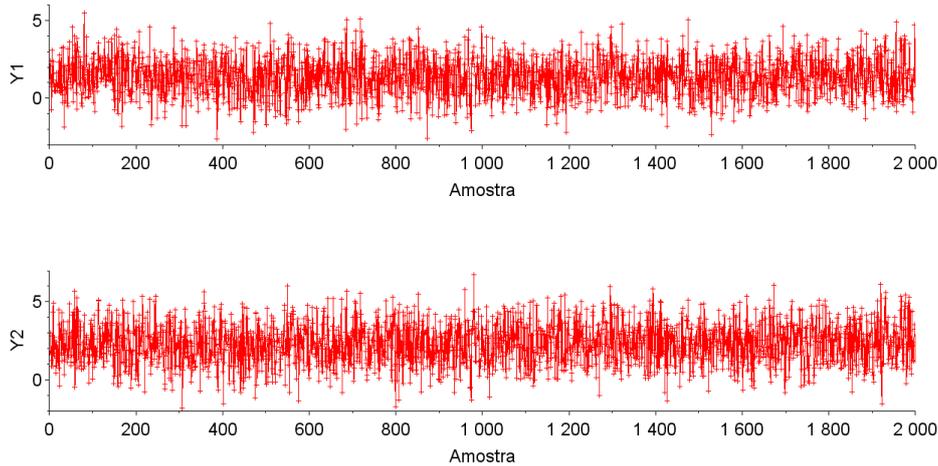


Figura 5.7: Evolução Temporal das Variáveis \mathbf{Y}_1 e \mathbf{Y}_2 do Exemplo 5.1

O banco de dados de referência é formado por 59 imagens e foi construído de acordo com o procedimento da Figura 5.3. Cada uma das janelas \mathbf{W}_1 e \mathbf{W}_2 é composta por 16 amostras, gerando uma matriz de recorrência de 32×32 pontos. As janelas estão defasadas entre si em $\mathbf{D} = \mathbf{350}$ amostras e avançam de 16 em 16 amostras até percorrer todos os 2000 pontos de simulação. Após construção do modelo de referência, cada uma das matrizes foi comparada às demais através da Norma de Frobenius, de acordo com o descrito na Seção 5.2. A aplicação da técnica de KDE produziu a distribuição de probabilidades apresentada na Figura 5.8. O limiar para detecção de falhas foi calculado em 0,96, para um nível de significância de 95%.

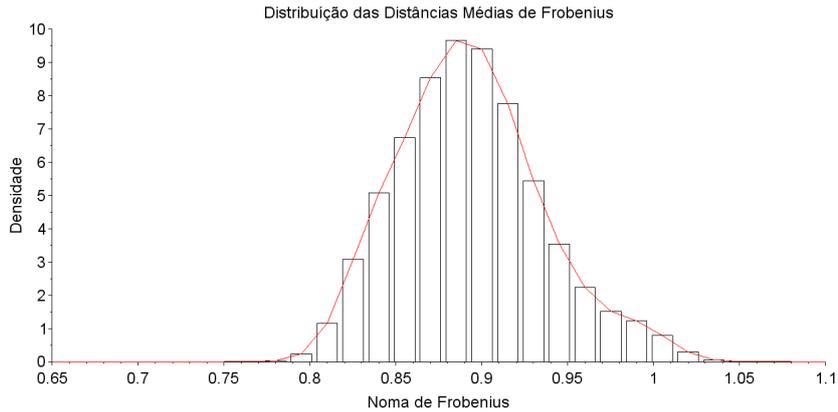


Figura 5.8: Distribuição das Distâncias Médias de Frobenius no Exemplo 5.1

Para verificação da capacidade de detecção de falhas do algoritmo proposto, foi gerada uma série de sinais anômalos (Figura 5.9). Na Figura 5.9 a primeira coluna traz o gráfico temporal dos sinais de falhas, a segunda coluna apresenta o diagrama de dispersão que relaciona as variáveis \mathbf{y}_1 e \mathbf{y}_2 e a terceira coluna apresenta a carta de controle resultante e o limiar de detecção. Nos diagramas de dispersão, os pontos vermelhos representam o padrão usado para o treinamento do algoritmo.

Os gráficos da primeira linha da Figura 5.9 foram obtidos a partir de uma alteração na variabilidade da variável \mathbf{y}_2 . Observando a respectiva carta de controle, fica claro que o algoritmo proposto é capaz de detectar alterações de variabilidade. Os gráficos da segunda linha da Figura 5.9 apresentam uma tendência lenta aplicada no sinal \mathbf{y}_1 . Na carta de controle equivalente, é possível notar uma mudança de comportamento, embora as violações do limite de controle aconteçam de maneira oscilatória. Para melhor investigar a capacidade de detecção de tendências, os gráficos da terceira linha apresentam uma tendência rápida aplicada também no sinal \mathbf{y}_1 ; porém, a partir de um determinado instante, o sinal \mathbf{y}_1 estabiliza em um novo patamar de operação. Observando a carta de controle equivalente, fica claro que uma tendência mais acentuada é facilmente detectada. Nota-se também que, a partir do momento em que o sinal \mathbf{y}_1 estabiliza, a carta de controle passa a indicar comportamento normal, o que é bastante interessante e indica que a carta de controle proposta é independente do ponto de operação, desde que no novo ponto de operação o desempenho do sistema seja compatível com o aprendido. Na última linha da Figura 5.9, após aplicar uma tendência rápida no sinal \mathbf{y}_1 , a nuvem de pontos é rotacionada até atingir uma nova posição. A partir da carta de controle, ficam evidentes as violações do limite de controle durante a rotação; entretanto, assim que a rotação cessa e o sistema passa a operar em um novo ponto de operação, a carta de controle volta a indicar operação normal.

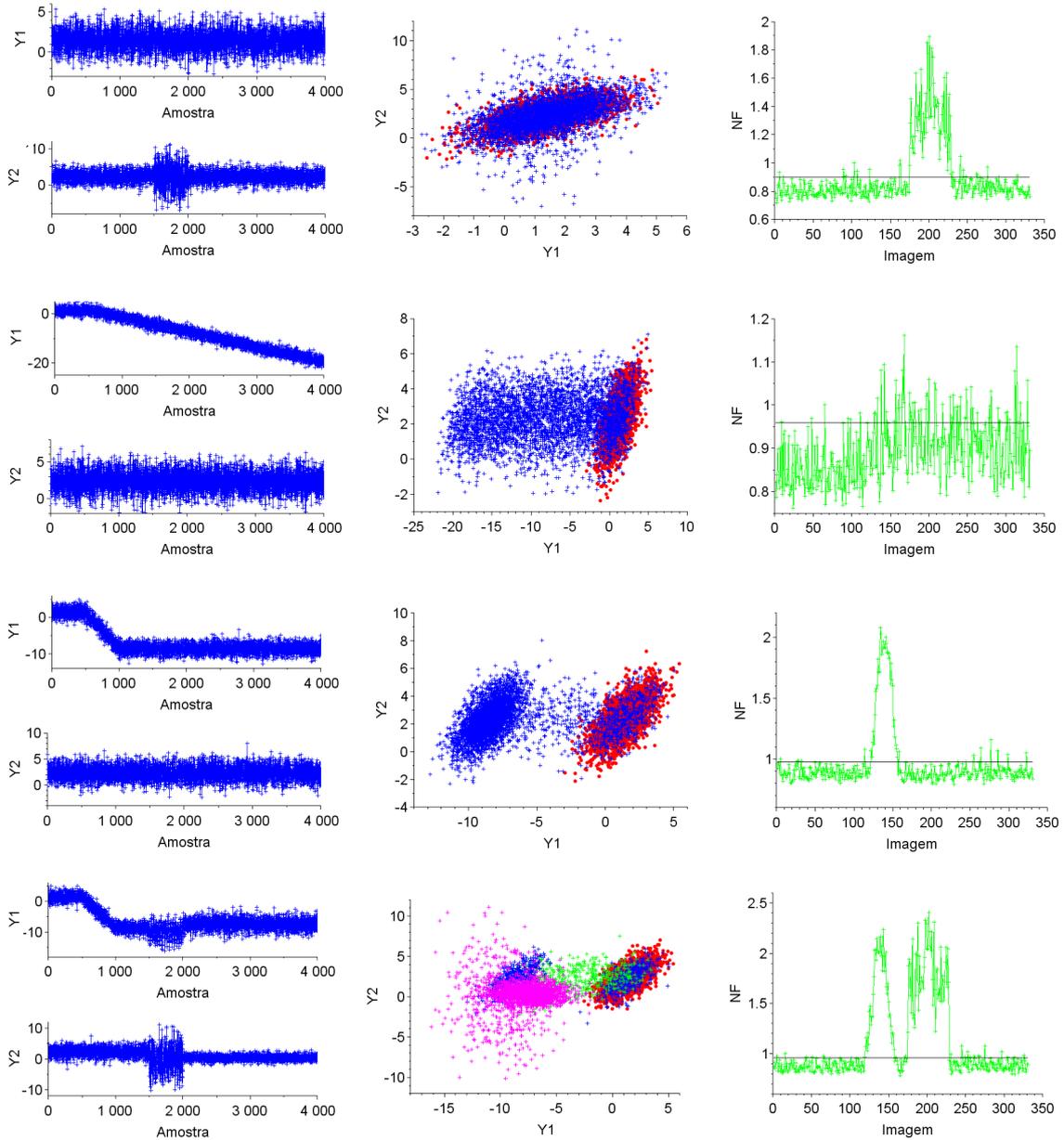


Figura 5.9: Avaliação da Capacidade de Detecção de Falhas do Algoritmo Proposto. a) alteração na variabilidade da variável y_2 , b) tendência lenta aplicada no sinal y_1 , c) tendência rápida aplicada no sinal y_1 , seguida de estabilização em novo patamar, d) tendência rápida aplicada no sinal y_1 , seguida por rotação da nuvem de pontos

Analisando a Figura 5.9, uma questão pertinente a ser levantada diz respeito a que fatores influenciam a capacidade de detecção de tendência do algoritmo proposto. A principal suspeita é que a defasagem D entre as janelas W_1 e W_2 impacte diretamente a capacidade de detecção de tendência. Na Figura 5.10 repetiu-se o teste de tendência lenta, já apresentado na segunda linha da Figura 5.9 para $D = 350$; porém, a defasagem entre as janelas foi aumentada para $D = 600$. Os resultados comprovam que a escolha adequada do parâmetro D é essencial para permitir que o algoritmo proposto seja capaz de detectar tendência. Portanto, o

parâmetro D é um parâmetro de sintonia relevante da técnica.

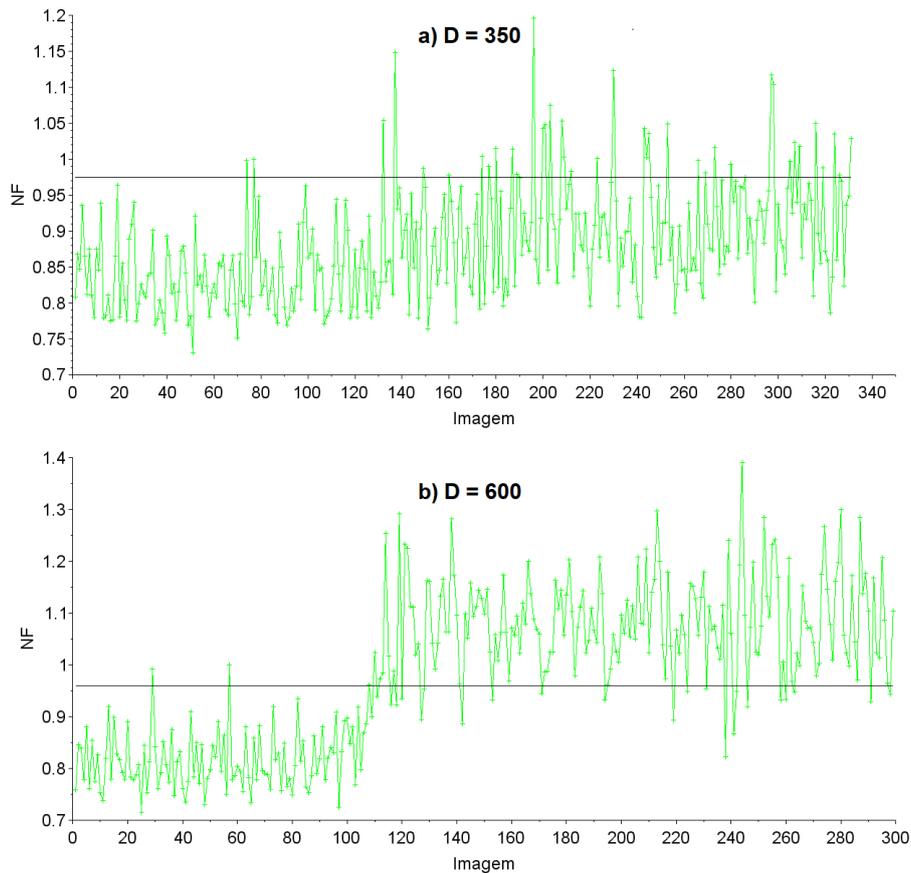


Figura 5.10: Resposta do Algoritmo Proposto Para Tendência Lenta. a) $D = 350$, b) $D = 600$

5.3 Estendendo a Metodologia Proposta

Toda a discussão da Seção 5.2 teve o objetivo de desenvolver um modelo independente do ponto de operação da planta, capaz de identificar comportamentos não estacionários. Contudo, se o sistema de controle for capaz de fazer com que a planta volte a operar com desempenho compatível com o aprendido, a carta de controle também retorna ao patamar normal. Para facilitar o diagnóstico do comportamento do processo, uma segunda carta de controle pode ser obtida ao tornar fixos os pontos que constituem a janela W_1 (Figura 5.3). A janela W_2 permanece móvel e a geração das matrizes de distâncias e o cálculo do limiar de detecção seguem o mesmo procedimento já descrito na Seção 5.2. Usando a mesma notação matemática da Seção 5.2, é possível generalizar esta segunda carta de controle como apresentado na Figura 5.11.

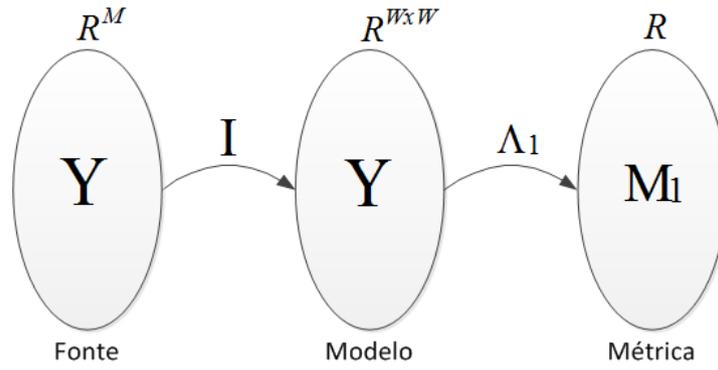


Figura 5.11: Generalização Matemática da Metodologia Proposta

A finalidade desta segunda carta de controle é verificar o afastamento do processo em relação ao ponto de operação aprendido. O afastamento do ponto operacional pode ser causado por uma variação de *setpoint* ou por distúrbios decorrentes de degradação da instrumentação da planta, alterações de processo ou mesmo por ações dos operadores (JIN, 2006). Matematicamente, o ponto de operação pode ser definido como uma função dos *setpoints* \mathbf{SP} e das variáveis de distúrbio \mathbf{d} (FEITAL *et al.*, 2013).

$$OP = f(SP, d) \quad (5.10)$$

Conseqüentemente, se o sistema de monitoramento proposto tiver acesso aos *setpoints* da planta, assim que for detectada uma mudança do ponto de operação, é possível julgar se a causa foi um distúrbio e se o sistema de controle foi capaz de reagir a ele, permitindo que a planta volte a operar de acordo com o desempenho aprendido. Para testar a abordagem proposta, na seção seguinte serão apresentadas algumas simulações envolvendo o reator Van der Vusse.

5.3.1 Exemplo 5.2 - Revisitando o Reator Van der Vusse

Com o objetivo de verificar a capacidade de monitoramento da metodologia proposta, ela foi aplicada ao CSTR com a reação de Van der Vusse. As janelas \mathbf{W}_1 e \mathbf{W}_2 foram configuradas com 16 amostras cada uma, avançaram de 16 em 16 amostras e estavam defasadas entre si em $\mathbf{D} = 350$ amostras. O modelo gerado possui 451 matrizes de referência. Os cenários testados foram os mesmos do Capítulo 4: mudança no *setpoint* da concentração de ciclopentanol no reator e desativação do catalisador. Na simulação de variação do *setpoint*, ilustrada na Figura 5.12, o *setpoint* da concentração de ciclopentanol no reator foi alterado para 65% do valor de referência. A curva verde na Figura 5.12-b é capaz de detectar a alteração do ponto operacional; porém, tão logo o controlador atinge o desempenho "normal" aprendido, a curva azul na Figura 5.12-b retorna à normalidade. Isto significa que o sistema

enfrentou um transiente e passou a trabalhar em um novo ponto operacional, com desempenho similar ao aprendido durante a fase de treinamento do algoritmo.

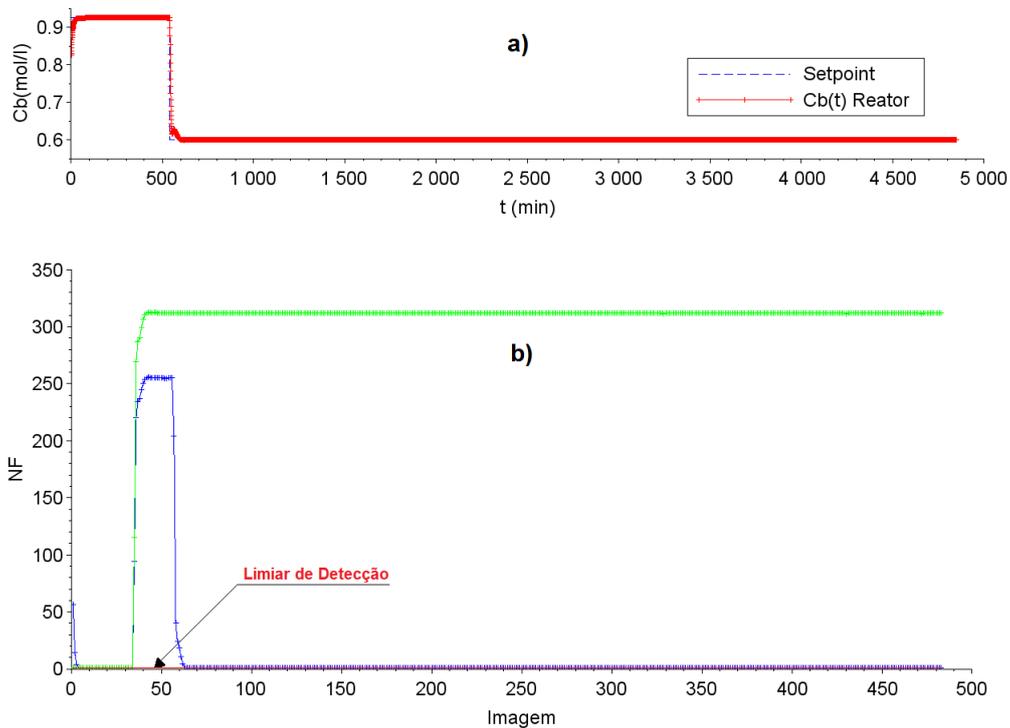


Figura 5.12: Resposta do Algoritmo Proposto Frente a Alterações de *Setpoint*. a) variação do *setpoint*, b) carta de controle da distância média de frobenius

Como a concentração de ciclopentanol no reator é controlada manipulando a taxa de diluição, o controlador é capaz de compensar a alteração de ponto operacional de modo a reequilibrar a operação do reator. Consequentemente, o padrão registrado pela curva azul tem a aparência de um pulso, cuja duração é determinada pela capacidade dos controladores de fazer com que as variáveis controladas atinjam o *setpoint*. O padrão registrado pela curva verde é similar à função degrau e indica que o sistema está operando em um ponto diferente do aprendido. Se o algoritmo de monitoramento proposto tiver acesso aos *setpoints* da planta, é possível julgar se a causa da mudança de ponto operacional foi um distúrbio, permitindo que o pessoal de operação tome ações corretivas, caso seja necessário.

No cenário de desativação do catalisador (Figura 5.13), a evolução temporal dos gráficos na carta de controle é bastante diferente em relação à simulação de variação do *setpoint*. A desativação do catalisador acontece de forma contínua (Figura 5.13-a) durante um período de **1860min**. Nessa situação, as variáveis controladas não conseguem retornar aos *setpoints* e o sistema de controle da planta tende a compensar os efeitos da desativação, sem no entanto atingir o desempenho nominal aprendido.

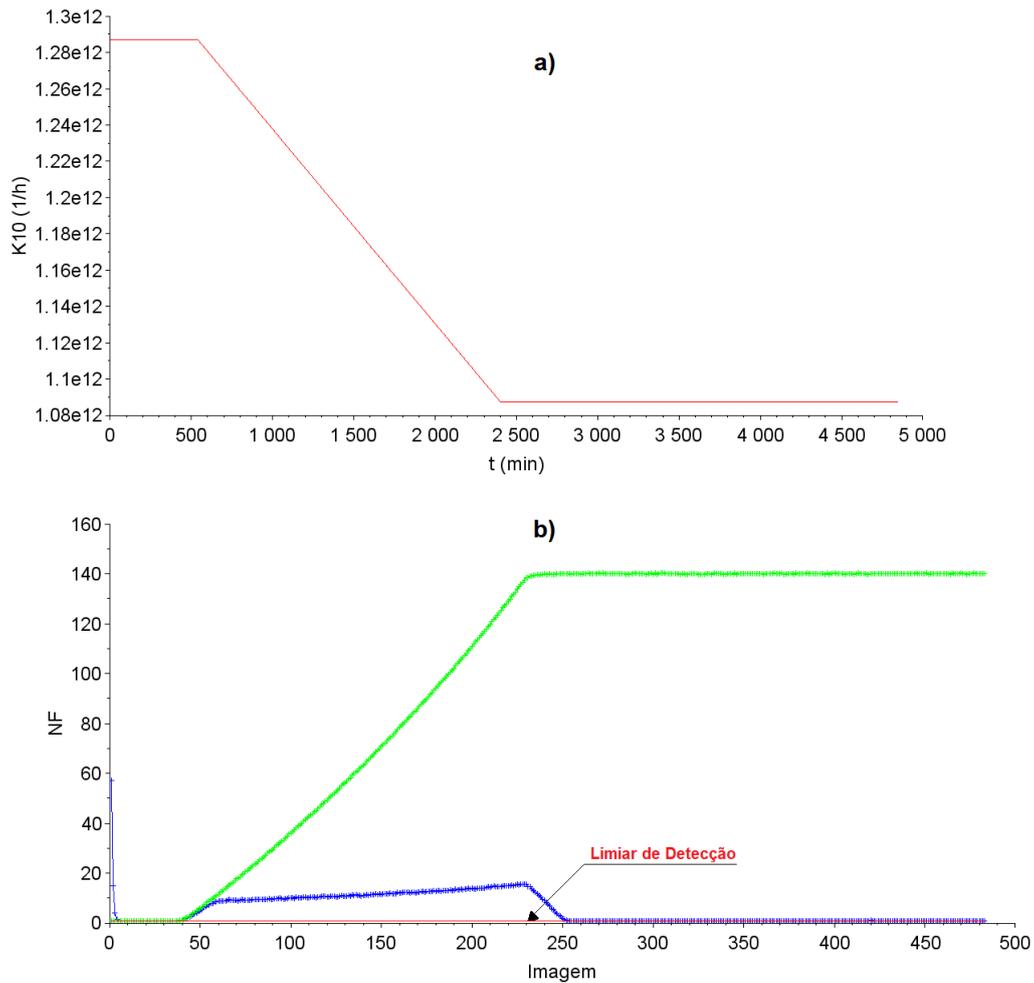


Figura 5.13: Resposta do Algoritmo Proposto Frente a Desativação do Catalisador. a) perfil de desativação, b) carta de controle da distância média de frobenius

Como resultado, os gráficos verde e azul apresentam pontos que crescem monotonicamente (Figura 5.13-b). Contudo, tão logo a desativação cessa e o controlador atinge o desempenho aprendido, a curva azul retorna à normalidade e a curva verde estabiliza em um novo patamar. Nota-se também que, apesar da curva verde indicar que um novo ponto operacional foi atingido, a curva azul só retorna ao patamar normal algum tempo depois, o que permite determinar com precisão o ponto a partir do qual seria viável implementar mecanismos adaptativos ao algoritmo de detecção de falhas, caso fosse necessário. Isto é importante, já que, quando o algoritmo de detecção de falhas se adapta a distúrbios, sua capacidade de detecção fica comprometida. JIN *et al.* (2006) é um dos poucos trabalhos que trata desta questão; entretanto, sua abordagem é bem mais complexa que a desenvolvida nesta dissertação, como pode ser observado na Figura 5.14.

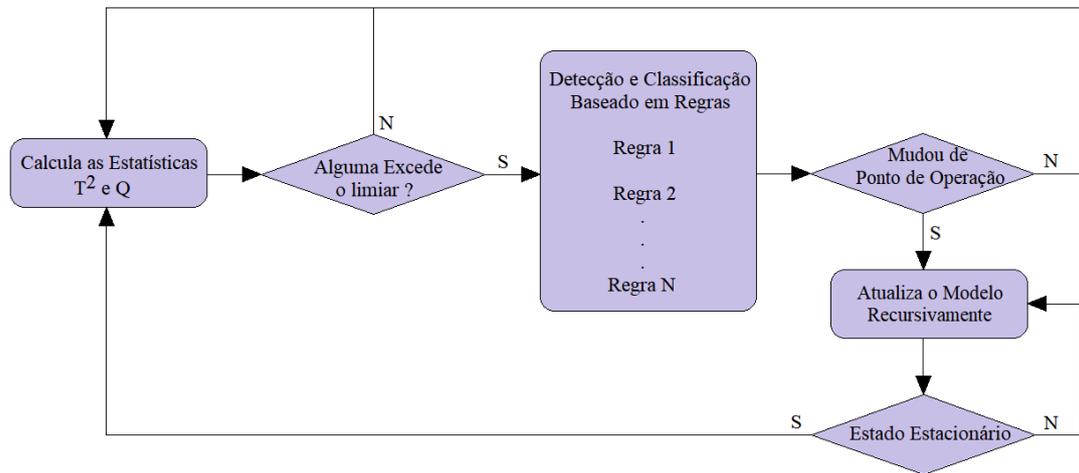


Figura 5.14: Algoritmo Proposto por JIN *et al.* (2006) para Implementação do PCA Robusto

Na abordagem proposta por JIN *et al.* (2006), a determinação do ponto a partir do qual é viável implementar mecanismos adaptativos é implementada em duas etapas. Na primeira, verifica-se se a estatística \mathbf{T}^2 ou se a estatística \mathbf{Q} excedeu o limite de detecção. Em caso positivo, o algoritmo segue para a segunda etapa, em que um conjunto de regras é avaliado para confirmar se realmente o processo mudou de ponto de operação. Caso tenha sido constatada a mudança de ponto operacional, o algoritmo entra na fase adaptativa, que termina tão logo o estado estacionário seja atingido.

Adicionalmente aos dois cenários já discutidos, apresentados nas Figuras 5.12 e 5.13, foi realizada uma terceira simulação para avaliar a resposta da metodologia proposta em situações de aumento da variabilidade da temperatura do reator. O aumento de variabilidade teve início em $t = 560\text{min}$ e terminou em $t = 2400\text{min}$ (Figura 5.15-a). Como o sistema está sob efeito de uma perturbação aleatória, tanto o gráfico azul quanto o verde violam o limiar de detecção de forma oscilatória (Figura 5.15-b). Ao contrário do que aconteceu nos outros dois cenários, o padrão de violação de ambos os gráficos é muito parecido.

Para finalizar esta seção, é importante ressaltar que a técnica proposta trabalha com blocos de dados, comparando uma nova matriz de recorrência às matrizes de referência, de modo que cada distância média desenhada na carta de controle se refere à uma nova matriz que foi comparada às matrizes de referência que constituem o modelo. Assim, na carta de controle, o eixo das abscissas contém valores inteiros que designam matrizes de recorrência obtidas durante a fase de operação *online* do algoritmo, constituindo portanto um novo paradigma de monitoração.

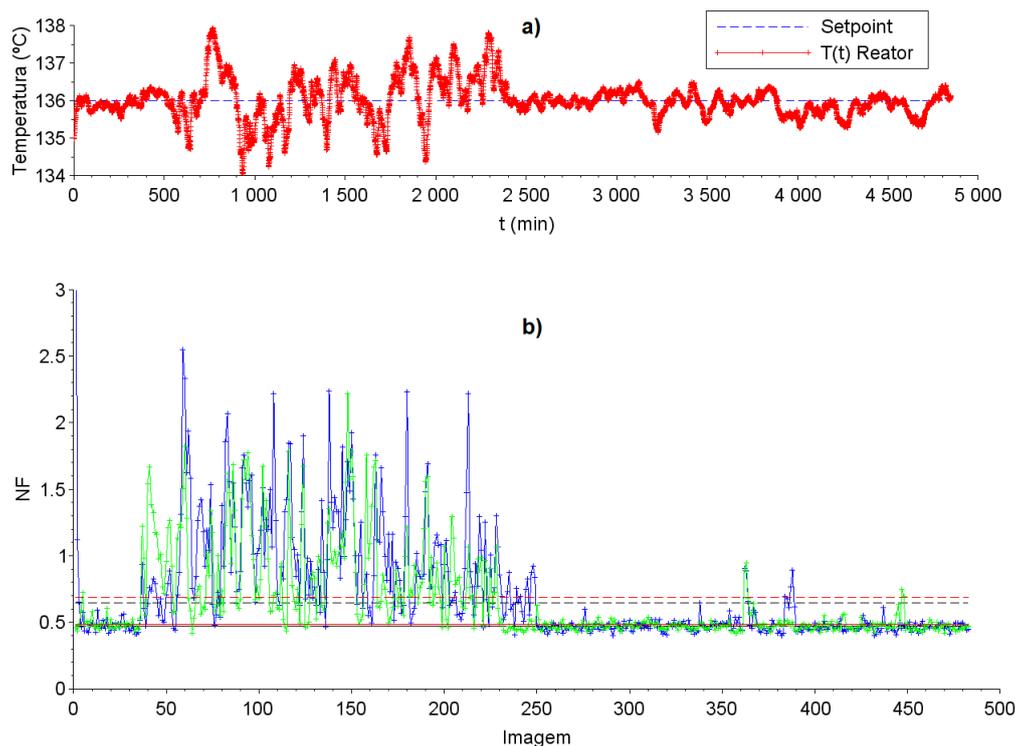


Figura 5.15: Resposta do Algoritmo Proposto Frente a Alterações de Variabilidade na Temperatura do Reator. a) temperatura do reator, b) carta de controle da distância média de frobenius

5.4 Comentários Finais

Embora vários distúrbios conduzam a desvios operacionais, sua caracterização pode ser significativamente diferente, fazendo com que a carta de controle proposta apresente diferentes padrões de resposta. Mudanças no *setpoint* do tipo degrau, tendência e variações aleatórias são distúrbios típicos de interesse para o monitoramento de processos. Portanto, para facilitar a interpretação por meio da abordagem de monitoramento proposta, a Tabela 5.2 apresenta uma síntese dos resultados obtidos em função dos diferentes tipos de distúrbios simulados na operação do reator Van der Vusse.

No caso de mudanças de *setpoint*, o controlador leva o processo para o novo ponto operacional e, se o desempenho alcançado é compatível com o padrão de treinamento, a carta de controle retorna ao patamar normal, apesar dos desvios evidentes da condição de operação. Este mecanismo apresenta o padrão apresentado na Figura 5.12 e foi discutido anteriormente quando o *setpoint* da concentração de ciclopentanol no reator foi alterado. No padrão registrado na carta de controle, a curva azul tem a aparência de um pulso e a curva verde tem a aparência de um degrau.

Admitindo que a defasagem \mathbf{D} entre as janelas foi configurada adequadamente,

distúrbios do tipo tendência também provocam alterações na carta de controle proposta. Conforme mostrado na Figura 5.13, o padrão de alarmes observado, tanto para a curva azul quanto para a curva verde, se desenvolve segundo um comportamento monotonicamente crescente, indicando que o processo está se desviando continuamente da sua condição de operação original, com o controlador trabalhando para compensar a perturbação. Já para as variações aleatórias (Figura 5.15), o controlador não consegue responder a flutuações tão abruptas e ambas as curvas apresentam um comportamento aleatório que viola o limiar de detecção, como mostrado na Figura 5.15.

Tabela 5.2: Resposta Para Distúrbios Típicos

	Curva Azul (Dinâmica)	Curva Verde (Afastamento da Média)
Degrau	Pulso	Degrau
Tendência	Cresce Monotonicamente	Cresce Monotonicamente
Variação Aleatória	Varia Aleatoriamente	Varia Aleatoriamente

Embora a metodologia de monitoramento de processo proposta tenha sido inspirada no conceito de recorrência, ela, de fato, não é capaz de reconstruir o espaço de fase do sistema de interesse. Todavia, as simplificações propostas na matriz de recorrência permitiram que o esforço de modelagem necessário para o treinamento do modelo e a construção da carta de controle seja pequeno, o que é bastante vantajoso para a indústria pois o sistema de monitoramento pode ser colocado em operação rapidamente. Outras vantagens da metodologia proposta estão relacionadas à independência do ponto operacional e à independência em relação às distribuições estatísticas das variáveis monitoradas, além de ser uma técnica multivariável.

Tudo indica que a metodologia proposta também poderia ser usada para gerar uma base de dados de modelos, para, por exemplo, permitir a identificação das variáveis de processo que estejam mais relacionadas à falhas ocorrida. Esta pode ser uma característica interessante para auxiliar o pessoal de operação a identificar os subsistemas afetados por determinada falha. Outro aspecto interessante de ser colocado diz respeito à monitoração, quando o processo está operando em regime transiente; embora o desenvolvimento apresentado neste capítulo tenha sido baseado no regime estacionário, acredita-se que a metodologia possa ser adaptada para monitoramento de processos durante transientes, permitindo que o sistema de monitoramento modele a dinâmica do processo.

Capítulo 6

Conclusões

Este capítulo apresenta algumas conclusões gerais sobre o tema de monitoramento de processos químicos e discute algumas sugestões para trabalhos futuros.

6.1 Conclusões

O monitoramento de processos químicos constitui um dos grandes desafios enfrentados pela indústria moderna. A área de monitoramento de processos é bastante ampla e conta com contribuições de diversos campos de pesquisa, tais como: modelagem fenomenológica, modelagem estatística e aprendizagem de máquina e reconhecimento de padrões. Devido ao aumento do poder computacional, da capacidade de armazenamento e da integração das redes de comunicação desde o chão de fábrica até as redes corporativas, existe um forte interesse na indústria por métodos de monitoramento *Data Driven*.

O objetivo deste trabalho foi investigar os métodos clássicos de monitoramento de processos, baseados em PCA, para desenvolver uma visão crítica da área que pudesse contribuir para o desenvolvimento de metodologias mais assertivas de monitoramento de processos. Neste sentido, no Capítulo 5 foi implementada uma nova metodologia de monitoramento inspirada na ideia dos gráficos de recorrência. Foi proposta uma carta de controle para acompanhamento de processos com múltiplos pontos de operação, fundamentada na norma de Frobenius. Muitos outros aspectos relacionados ao tema de monitoramento, tais como seleção de variáveis, preprocessamento e remoção de valores espúrios, não foram abordados neste trabalho mas são de fundamental importância para a aplicação industrial.

Embora muitos trabalhos tenham contribuído para o avanço do tema, a aplicação de métodos de monitoramento mais complexos, baseados em modelagem estatística, ainda é relativamente pequena na indústria. Em parte, este fato parece ser consequência da complexidade das plantas e da falta de adaptabilidade

dos métodos estatísticos em relação a variações do ponto operacional. Muitos algoritmos de monitoramento possuem versões adaptativas; entretanto, poucos autores discutem mecanismos para isolar distúrbios e evitar que o mecanismo de adaptação incorpore-os ao modelo, comprometendo, assim, a capacidade de detecção de falhas. Além disso, a maioria dos métodos de monitoramento foi concebida para lidar com processos lineares, enquanto que os processos reais possuem algum grau de não linearidade. Destaca-se também que o ajuste dos parâmetros dos modelos nem sempre é trivial, dependendo de conhecimento do domínio, o que de certa forma constitui uma contradição intrínseca do processo de modelagem *Data Driven*. Ademais, questões relativas à frequência de amostragem das variáveis de processo, à não normalidade dos dados e à seleção das estatísticas de detecção têm sido raramente abordadas na literatura.

6.2 Sugestões de Trabalhos Futuros

A função primária de um sistema de monitoramento é auxiliar o pessoal responsável pela operação das plantas de processo a compreender e identificar o estado atual das plantas e alertá-los em caso de padrões de operação indesejados ou não esperados, de forma a evitar possíveis panes ou maus-funcionamentos. Para cumprir este objetivo, é fundamental que existam métodos capazes de auxiliar a identificação das variáveis de processo relacionadas à falha ocorrida. O desenvolvimento de propostas de análise deste tipo de problema é fundamental para aplicação de sistema de monitoramento na indústria. Como nenhuma estratégia de monitoramento é eficaz por si só, é fundamental investir em pesquisas e no desenvolvimento de abordagens que combinem vários métodos de monitoramento.

Embora a metodologia de monitoramento de processo desenvolvida no Capítulo 5 tenha sido inspirada no conceito de recorrência, ela, de fato, não é capaz de reconstruir o espaço de fase de um sistema dinâmico. Ao que tudo indica, para monitorar uma planta de processo de modo mais assertivo é necessário distinguir entre mudanças normais do ponto de operação de falhas que impliquem em comportamentos dinâmicos anormais. Assim, parece razoável pensar em incorporar à metodologia proposta mecanismos para a modelagem dinâmica dos processos, de forma a permitir que o algoritmo seja capaz de monitorar plantas em regime transiente.

A técnica de Estimção de Densidade por Funções Núcleo (KDE) foi utilizada no Capítulo 5 para estimar a função densidade de probabilidades (PDF) da Norma Média de Frobenius. Posteriormente, a função densidade de probabilidades foi integrada, de modo a determinar o limiar de detecção de falhas do algoritmo proposto. Uma possível alternativa para este procedimento seria utilizar o teste de

Kolmogorov-Smirnov para Duas Amostras (GIBBONS e CHAKRABORTI, 2003). O teste de Kolmogorov-Smirnov consiste de um procedimento não paramétrico, usado para comparar distribuições de probabilidade. Por meio dele é possível determinar se duas amostras são provenientes de uma mesma distribuição de probabilidades ou população. A vantagem desta abordagem está no fato de que o cálculo de valores médios e do limiar de detecção podem ser evitados, já que o teste de Kolmogorov-Smirnov opera diretamente com as amostras.

Referências Bibliográficas

- ABNT-NBR-ISO-14224, 2011. “Indústrias de Petróleo e Gás Natural - Coleta e Intercâmbio de Dados de Confiabilidade e Manutenção Para Equipamentos”. .
- AFZAL, M. S., TAN, W., CHEN, T., 2017, “Process Monitoring for Multimodal Processes With Mode-Reachability Constraints”, *IEEE Transactions on Industrial Electronics*, v. 64, n. 5 (maio), pp. 4325–4335. ISSN: 0278-0046, 1557-9948. doi: 10.1109/TIE.2017.2677351. Disponível em: <<http://ieeexplore.ieee.org/document/7869281/>>.
- AGUIRRE, L. S., 2015, *Introdução à Identificação de Sistemas – Técnicas Lineares e Não-Lineares Aplicadas a Sistemas Reais*. 4ª ed. Belo Horizonte, UFMG. ISBN: 978-85-423-0079-6.
- ALDRICH, C., AURET, L., 2013, *Unsupervised process monitoring and fault diagnosis with machine learning methods*. New York, Springer. ISBN: 978-1-4471-5184-5.
- ALWI, H., EDWARDS, C., TAN, C. P., 2011, *Fault detection and fault-tolerant control using sliding modes*. Advances in industrial control. New York, Springer. ISBN: 978-0-85729-649-8.
- American Institute of Chemical Engineers (Ed.), 2001, *Layer of protection analysis: simplified process risk assessment*. A CCPS concept book. New York, Center for Chemical Process Safety of the American Institute of Chemical Engineers. ISBN: 978-0-8169-0811-0.
- APPEL, J., COLOMBO, C., DÄTWYLER, U., et al., 2016, “Comprehensive Analysis Competence and Innovative Approaches for Sustainable Chemical Production”, *CHIMIA International Journal for Chemistry*, v. 70, n. 9 (set.), pp. 621–627. ISSN: 00094293. doi: 10.2533/chimia.2016.621. Disponível em: <<http://www.ingentaconnect.com/content/10.2533/chimia.2016.621>>.

- ATKEARNEY, 200-? “Chemical Industry Vision 2040: A European Perspective”. Disponível em: <<https://www.atkearney.com/documents/20152/434123/Chemical%2BIndustry%2BVision%2B2030%2BA%2BEuropean%2BPerspective.pdf/7a5e8ff7-071b-1b9e-f890-cd9ca3562e61>>.
- BASSEVILLE, M., NIKIFOROV, I. V., 1993, *Detection of abrupt changes: theory and application*. Prentice Hall information and system sciences series. Englewood Cliffs, N.J, Prentice Hall. ISBN: 978-0-13-126780-0.
- BELL, J., 2014, *Machine learning: hands-on for developers and technical professionals*. Indianapolis, IN, John Wiley and Sons. ISBN: 978-1-118-88906-0.
- BEQUETTE, B. W., 2003, *Process control: modeling, design, and simulation*. Prentice-Hall international series in the physical and chemical engineering sciences. Upper Saddle River, N.J, Prentice Hall PTR. ISBN: 978-0-13-353640-9.
- BLANKE, M., SCHRÖDER, J., 2006, *Diagnosis and fault-tolerant control*. 2nd ed ed. Berlin ; New York, Springer. ISBN: 978-3-540-35652-3. OCLC: ocm71336524.
- BOSCO, F., 2009, “Por que acidentes (ainda) acontecem?” *Petro & Química*, v. 1, n. 317. Disponível em: <http://www.editoravalete.com.br/site_petroquimica/>.
- CAMACHO, E. F., BORDONS, C., 2007, *Model Predictive control*. London, Springer London. ISBN: 978-0-85729-398-5. Disponível em: <<http://link.springer.com/book/10.1007/978-0-85729-398-5>>.
- CAMPBELL, S. L., NIKOUKHAH, R., 2004, *Auxiliary signal design for failure detection*. Princeton series in applied mathematics. Princeton, N.J, Princeton University Press. ISBN: 978-0-691-09987-3.
- CANDÈS, E. J., LI, X., MA, Y., et al., 2011, “Robust principal component analysis?” *Journal of the ACM*, v. 58, n. 3 (maio), pp. 1–37. ISSN: 00045411. doi: 10.1145/1970392.1970395. Disponível em: <<http://portal.acm.org/citation.cfm?doid=1970392.1970395>>.
- CHEN, J., PATTON, R., 1999, *Robust model-based fault diagnosis for dynamic systems*. N. 3, Kluwer international series on Asian studies in computer and information science. Boston, Kluwer Academic Publishers. ISBN: 978-0-7923-8411-3.

- CHEN, Z., 2017, *Data-driven fault detection for industrial processes: canonical correlation analysis and projection based methods*. Research. Wiesbaden, Springer Vieweg. ISBN: 978-3-658-16756-1 978-3-658-16755-4.
- CHIANG, L. H., BRAATZ, R. D., RUSSELL, E., 2001, *Fault detection and diagnosis in industrial systems*. Advanced textbooks in control and signal processing. London ; New York, Springer. ISBN: 978-1-85233-327-0.
- CHOUAIB, C., MOHAMED-FAOUZI, H., MESSAOUD, D., 2013, “Adaptive kernel principal component analysis for nonlinear dynamic process monitoring”, *Control Conference (ASCC), 2013 9th Asian*, (jun.), pp. 1–6. doi: 10.1109/ASCC.2013.6606291. Disponível em: <<http://ieeexplore.ieee.org/document/6606291/>>.
- CHOUDHURY, M. A. A. S., SHAH, S. L., THORNHILL, N. F., 2008, *Diagnosis of process nonlinearities and valve stiction: data driven approaches*. Advances in industrial control. Berlin ; London, Springer. ISBN: 978-3-540-79223-9 978-3-540-79224-6.
- CÂMARA, M., SOARES, R., FEITAL, T., et al., 2017, “Numerical Aspects of Data Reconciliation in Industrial Applications”, *Processes*, v. 5, n. 4 (out.), pp. 56. ISSN: 2227-9717. doi: 10.3390/pr5040056. Disponível em: <<http://www.mdpi.com/2227-9717/5/4/56>>.
- COIMBRA, J., MELO, P., PRATA, D., et al., 2017, “On-Line Dynamic Data Reconciliation in Batch Suspension Polymerizations of Methyl Methacrylate”, *Processes*, v. 5, n. 3 (set.), pp. 51. ISSN: 2227-9717. doi: 10.3390/pr5030051. Disponível em: <<http://www.mdpi.com/2227-9717/5/3/51>>.
- DA SILVA ROSA, I., MELO, P. A., PINTO, J. C., 2012, “Bifurcation Analysis of the Bulk Propylene Polymerization in the LIPP Process”, *Macromolecular Symposia*, v. 319, n. 1 (set.), pp. 41–47. ISSN: 10221360. doi: 10.1002/masy.201100246. Disponível em: <<http://doi.wiley.com/10.1002/masy.201100246>>.
- DE KETELAERE, B., HUBERT, M., SCHMITT, E., 2015, “Overview of PCA-Based Statistical Process-Monitoring Methods for Time-Dependent, High-Dimensional Data”, *Journal of Quality Technology*, v. 47, n. 4 (out.), pp. 318–335. ISSN: 0022-4065, 2575-6230. doi: 10.1080/00224065.2015.11918137. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/00224065.2015.11918137>>.

- DENG, X., TIAN, X., 2014, “Multimode Process Fault Detection Using Local Neighborhood Similarity Analysis”, *Chinese Journal of Chemical Engineering*, v. 22, n. 11-12 (nov.), pp. 1260–1267. ISSN: 10049541. doi: 10.1016/j.cjche.2014.09.022. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S1004954114001293>>.
- DING, S. X., 2014, *Data-driven design of fault diagnosis and fault-tolerant control systems*. Advances in industrial control. London Heidelberg New York Dordrecht, Springer. ISBN: 978-1-4471-6410-4 978-1-4471-6409-8.
- ECKMANN, J.-P., KAMPHORST, S. O., RUELLE, D., 1987, “Recurrence Plots of Dynamical Systems”, *Europhysics Letters (EPL)*, v. 4, n. 9 (nov.), pp. 973–977. ISSN: 0295-5075, 1286-4854. doi: 10.1209/0295-5075/4/9/004. Disponível em: <<http://stacks.iop.org/0295-5075/4/i=9/a=004?key=crossref.09fbeb6883f90a0adb050fbd7323bcd5>>.
- ELSNER, J. B., TSONIS, A. A., 1996, *Singular spectrum analysis: a new tool in time series analysis*. New York, Plenum Press. ISBN: 978-0-306-45472-1. OCLC: 247608520.
- FAES, L., NOLLO, G., PORTA, A., 2012, “Non-uniform multivariate embedding to assess the information transfer in cardiovascular and cardiorespiratory variability series”, *Computers in Biology and Medicine*, v. 42, n. 3 (mar.), pp. 290–297. ISSN: 00104825. doi: 10.1016/j.combiomed.2011.02.007. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S001048251100031X>>.
- FEITAL, T., KRUGER, U., DUTRA, J., et al., 2013, “Modeling and performance monitoring of multivariate multimodal processes”, *AIChE Journal*, v. 59, n. 5 (maio), pp. 1557–1569. ISSN: 00011541. doi: 10.1002/aic.13953. Disponível em: <<http://doi.wiley.com/10.1002/aic.13953>>.
- FEITAL, T. D. S., 2011, *Monitoramento da Condição de Processos Químicos Industriais*. Tese de Doutorado, Universidade Federal do Rio de Janeiro, Rio de Janeiro, nov.
- FELDER, R. M., ROUSSEAU, R. W., 2005, *Elementary principles of chemical processes*. 3rd ed., 2005 ed. with integrated media and study tools ed. Hoboken, NJ, Wiley. ISBN: 978-0-471-68757-3 978-0-471-71877-2.
- FISHER, R. A., 1936. “Conjunto de Dados Íris”. Disponível em: <<https://archive.ics.uci.edu/ml/datasets/Iris>>. Banco de dados da UCI para reconhecimento de padrões.

- FRIEDRICH-EBERT-STIFTUNG, 2017. “Social Innovation Policy for Industry 4.0”. Disponível em: <<http://library.fes.de/pdf-files/wiso/11479.pdf>>.
- GE, Z., SONG, Z., 2013, *Multivariate statistical process control: process monitoring methods and applications*. Advances in industrial control. London ; New York, Springer. ISBN: 978-1-4471-4512-7.
- GE, Z., SONG, Z., GAO, F., 2013, “Review of Recent Research on Data-Based Process Monitoring”, *Industrial & Engineering Chemistry Research*, v. 52, n. 10 (mar.), pp. 3543–3562. ISSN: 0888-5885, 1520-5045. doi: 10.1021/ie302069q. Disponível em: <<http://pubs.acs.org/doi/abs/10.1021/ie302069q>>.
- GERTLER, J., 1998, *Fault detection and diagnosis in engineering systems*. New York, Marcel Dekker. ISBN: 978-0-8247-9427-9.
- GIBBONS, J. D., CHAKRABORTI, S., 2003, *Nonparametric statistical inference*. N. v. 168, Statistics, textbooks and monographs. 4th ed., rev. and expanded ed. New York, Marcel Dekker. ISBN: 978-0-8247-4052-8.
- GOLYANDINA, N., NEKRUTKIN, V. V., ZHIGLJAVSKY, A. A., 2001, *Analysis of time series structure: SSA and related techniques*. N. 90, Monographs on statistics and applied probability. Boca Raton, Fla., Chapman & Hall/CRC. ISBN: 978-1-58488-194-0. OCLC: 248162405.
- GUERFEL, M., OTHMAN, K. B., BENREJEB, M., 2009, “On the structure determination of a dynamic PCA model using sensitivity of fault detection”, *IFAC Proceedings Volumes*, v. 42, n. 11, pp. 958–963. ISSN: 14746670. doi: 10.3182/20090712-4-TR-2008.00157. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S1474667015304006>>.
- GUSTAFSSON, F., 2000, *Adaptive filtering and change detection*. Chichester New York, Wiley. ISBN: 978-0-471-49287-0. OCLC: ocm44101938.
- HE, X. B., YANG, Y. P., 2008, “Variable MWPCA for Adaptive Process Monitoring”, *Industrial & Engineering Chemistry Research*, v. 47, n. 2 (jan.), pp. 419–427. ISSN: 0888-5885, 1520-5045. doi: 10.1021/ie070712z. Disponível em: <<http://pubs.acs.org/doi/abs/10.1021/ie070712z>>.

- HEGGER, R., KANTZ, H., SCHREIBER, T., 1999, “Practical implementation of nonlinear time series methods: The TISEAN package”, *Chaos: An Interdisciplinary Journal of Nonlinear Science*, v. 9, n. 2 (jun.), pp. 413–435. ISSN: 1054-1500, 1089-7682. doi: 10.1063/1.166424. Disponível em: <http://aip.scitation.org/doi/10.1063/1.166424>.
- HIMMELBLAU, D. M., 1978, *Fault detection and diagnosis in chemical and petrochemical processes*. Chemical engineering monographs ; v. 8. Amsterdam ; New York : New York, Elsevier Scientific Pub. Co. ; distributors for the US and Canada, Elsevier North-Holland. ISBN: 978-0-444-41747-3.
- HINES, J. W., SEIBERT, R., 2006, *Technical Review of On-Line Monitoring Techniques for Performance Assessment: State-of-the-Art*. Relatório Técnico NUREG/CR-6895, University of Tennessee, Knoxville, jan. Disponível em: <https://www.nrc.gov/reading-rm/doc-collections/nuregs/contract/cr6895/v1/>.
- HIRATA, Y., SUZUKI, H., AIHARA, K., 2006, “Reconstructing state spaces from multivariate data using variable delays”, *Physical Review E*, v. 74, n. 2 (ago.). ISSN: 1539-3755, 1550-2376. doi: 10.1103/PhysRevE.74.026202. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevE.74.026202>.
- HYVÄRINEN, A., OJA, E., 2000, “Independent component analysis: algorithms and applications”, *Neural Networks*, v. 13, n. 4-5 (jun.), pp. 411–430. ISSN: 08936080. doi: 10.1016/S0893-6080(00)00026-5. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/S0893608000000265>.
- IEC-61508, 2010. “Functional Safety of Electrical/Electronic/Programmable Electronic Safety-Related Systems” . .
- IEC-61511, 2017. “Functional Safety - Safety Instrumented Systems for the Process Industry Sector” . .
- INSTITUTO ADOLFO LUTZ, 2013. “Manual para Elaboração de Cartas de Controle para Monitoramento de Processos de Medição Quantitativos em Laboratórios de Ensaio”. Disponível em: http://www.ial.sp.gov.br/resources/editorinplace/ial/2016_3_19/manual-carta-controle_ial_2013.pdf.

- INSTITUTO NACIONAL DE METROLOGIA, QUALIDADE E TECNOLOGIA (INMETRO), 2012. “Vocabulário Internacional de Metrologia - VIM”. Disponível em: <http://www.inmetro.gov.br/inovacao/publicacoes_avulsas.asp>.
- ISERMANN, R., BALLÉ, P., 1997, “Trends in the application of model-based fault detection and diagnosis of technical processes”, *Control Engineering Practice*, v. 5, n. 5 (maio), pp. 709–719. ISSN: 09670661. doi: 10.1016/S0967-0661(97)00053-1. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0967066197000531>>.
- ISERMANN, R., 2006, *Fault-diagnosis systems: an introduction from fault detection to fault tolerance*. Berlin ; New York, Springer. ISBN: 978-3-540-24112-6.
- JACKSON, J. E., 1991, *A user's guide to principal components*. Wiley series in probability and mathematical statistics. New York, Wiley. ISBN: 978-0-471-62267-3.
- JELALI, M., 2006, “An overview of control performance assessment technology and industrial applications”, *Control Engineering Practice*, v. 14, n. 5 (maio), pp. 441–466. ISSN: 09670661. doi: 10.1016/j.conengprac.2005.11.005. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0967066105002479>>.
- JENG, J.-C., 2010, “Adaptive process monitoring using efficient recursive PCA and moving window PCA algorithms”, *Journal of the Taiwan Institute of Chemical Engineers*, v. 41, n. 4 (jul.), pp. 475–481. ISSN: 18761070. doi: 10.1016/j.jtice.2010.03.015. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S1876107010000532>>.
- JIN, H. D., LEE, Y.-H., LEE, G., et al., 2006, “Robust Recursive Principal Component Analysis Modeling for Adaptive Monitoring”, *Industrial & Engineering Chemistry Research*, v. 45, n. 2 (jan.), pp. 696–703. ISSN: 0888-5885, 1520-5045. doi: 10.1021/ie050850t. Disponível em: <<http://pubs.acs.org/doi/abs/10.1021/ie050850t>>.
- JOHNSON, R. A., WICHERN, D. W., 2007, *Applied multivariate statistical analysis*. 6th ed ed. Upper Saddle River, N.J, Pearson Prentice Hall. ISBN: 978-0-13-187715-3.
- JOHNSON CONTROLS, 2013. “Fault Detection And Diagnostics - Enabling techno-comissioning to ease building operation

and improve performance”. Disponível em: <<http://buildingefficiencyinitiative.org>>.

JOLLIFFE, I. T., 2002, *Principal component analysis*. Springer series in statistics. 2nd ed ed. New York, Springer. ISBN: 978-0-387-95442-4.

JUDD, K., MEES, A., 1998, “Embedding as a modeling problem”, *Physica D: Nonlinear Phenomena*, v. 120, n. 3-4 (set.), pp. 273–286. ISSN: 01672789. doi: 10.1016/S0167-2789(98)00089-X. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S016727899800089X>>.

KADLEC, P., GRBIĆ, R., GABRYS, B., 2011, “Review of adaptation mechanisms for data-driven soft sensors”, *Computers & Chemical Engineering*, v. 35, n. 1 (jan.), pp. 1–24. ISSN: 00981354. doi: 10.1016/j.compchemeng.2010.07.034. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0098135410002838>>.

KANTZ, H., SCHREIBER, T., 2004, *Nonlinear Time Series Analysis*. 2 ed. Reino Unido, Cambridge University Press.

KEMPF, A. D. O., 2003, *Avaliação de Desempenho de Malhas de Controle*. Dissertação de Mestrado, Universidade Federal do Rio de Janeiro, Porto Alegre 2003.

KÜHNERT, C., 2013, *Data-driven methods for fault localization in process technology*. N. Fraunhofer-Inst. für Optronik, Systemtechnik und Bildauswertung IOSB Karlsruhe 15, Karlsruher Schriften zur Anthropomatik / Lehrstuhl für Interaktive Echtzeitsysteme, Karlsruher Institut für Technologie. Print on demand ed. Karlsruhe, KIT Scientific Publishing. ISBN: 978-3-7315-0098-8.

KILUNDU, B., DEHOMBREUX, P., CHIEMENTIN, X., 2011, “Tool wear monitoring by machine learning techniques and singular spectrum analysis”, *Mechanical Systems and Signal Processing*, v. 25, n. 1 (jan.), pp. 400–415. ISSN: 08883270. doi: 10.1016/j.ymssp.2010.07.014. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0888327010002505>>.

KOURTI, T., 2003, “Multivariate dynamic data modeling for analysis and statistical process control of batch processes, start-ups and grade transitions”, *Journal of Chemometrics*, v. 17, n. 1 (jan.), pp. 93–109. ISSN: 0886-9383, 1099-128X. doi: 10.1002/cem.778. Disponível em: <<http://doi.wiley.com/10.1002/cem.778>>.

- KRÜGER, U., XIE, L., 2012, *Statistical monitoring of complex multivariate processes: with applications in industrial process control*. Statistics in practice. Chichester, West Sussex ; Hoboken, N.J, Wiley. ISBN: 978-0-470-02819-3.
- KRIEGEL, H.-P., KRÖGER, P., SCHUBERT, E., et al., 2008, “A General Framework for Increasing the Robustness of PCA-Based Correlation Clustering Algorithms”. In: Ludäscher, B., Mamoulis, N. (Eds.), *Scientific and Statistical Database Management*, v. 5069, Springer Berlin Heidelberg, pp. 418–435, Berlin, Heidelberg. ISBN: 978-3-540-69476-2 978-3-540-69497-7. doi: 10.1007/978-3-540-69497-7_27. Disponível em: <http://link.springer.com/10.1007/978-3-540-69497-7_27>.
- KRISHNANNAIR, S., 2010, *Multiscale Process Monitoring With Singular Spectrum Analysis*. Dissertação de Mestrado, Universidade de Stellenbosch, dez.
- KU, W., STORER, R. H., GEORGAKIS, C., 1995, “Disturbance detection and isolation by dynamic principal component analysis”, *Chemometrics and Intelligent Laboratory Systems*, v. 30, n. 1 (nov.), pp. 179–196. ISSN: 01697439. doi: 10.1016/0169-7439(95)00076-3. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/0169743995000763>>.
- LANGLOIS, D., CHARTIER, S., GOSSELIN, D., 2010, “An Introduction to Independent Component Analysis: InfoMax and FastICA algorithms”, *Tutorials in Quantitative Methods for Psychology*, v. 6, n. 1 (mar.), pp. 31–38. ISSN: 1913-4126. doi: 10.20982/tqmp.06.1.p031. Disponível em: <<http://www.tqmp.org/RegularArticles/vol06-1/p031>>.
- LE BOUDEC, J.-Y., 2010, *Performance evaluation of computer and communication systems*. Computer and communication sciences. 1. ed ed. Lausanne, EPFL Press. ISBN: 978-1-4398-4992-7 978-2-940222-40-7.
- LEDESMA, R. D., VALERO-MORA, P., 2007, “Determining the Number of Factors to Retain in EFA: an easy-touse computer program for carrying out Parallel Analysis”, *Practical Assessment, Research & Evaluation*, v. 12. ISSN: 1531-7714. Disponível em: <<http://pareonline.net/pdf/v12n2.pdf>>.
- LEDESMA, R. D., VALERO-MORA, P., MACBETH, G., 2015, “The Scree Test and the Number of Factors: a Dynamic Graphics Approach”, *The Spanish Journal of Psychology*, v. 18. ISSN: 1988-2904. doi: 10.1017/sjp.2015.

13. Disponível em: <http://www.journals.cambridge.org/abstract_S113874161500013X>.
- LI, L., 2016, *Fault detection and fault-tolerant control for nonlinear systems*. Research. Wiesbaden, Springer Vieweg. ISBN: 978-3-658-13020-6 978-3-658-13019-0.
- LIU, X., XIE, L., KRUGER, U., et al., 2008, “Statistical-based monitoring of multivariate non-Gaussian systems”, *AIChE Journal*, v. 54, n. 9 (set.), pp. 2379–2391. ISSN: 00011541, 15475905. doi: 10.1002/aic.11526. Disponível em: <<http://doi.wiley.com/10.1002/aic.11526>>.
- MARSZAL, E. M., SCHARPF, E. W., 2002, *Safety integrity level selection: systematic methods including layer of protection analysis*. Research Triangle Park, NC, Instrumentation, Systems, and Automation Society. ISBN: 978-1-55617-777-4.
- MARTINEZ, W. L., MARTINEZ, A. R., 2008, *Computational statistics handbook with MATLAB*. Boca Raton, FL, Chapman & Hall/CRC. ISBN: 978-1-4200-1086-2. Disponível em: <<http://public.ebib.com/choice/publicfullrecord.aspx?p=1633195>>. OCLC: 946554757.
- MARTINEZ-GUERRA, R., 2013, *Fault detection and diagnosis in nonlinear systems: a differential and algebraic viewpoint*. New York, Springer. ISBN: 978-3-319-03046-3.
- MESKIN, N., KHORASANI, K., 2011, *Fault detection and isolation: multi-vehicle unmanned systems*. New York, Springer. ISBN: 978-1-4419-8392-3.
- MOTEIRO, G. P., NAEGELI, G., COSTA, L. A., et al., 2011. “Análise de Camadas de Proteção (LOPA) e Sistemas Instrumentados de Segurança (SIS). Curso de Segurança de Processo.” .
- NARASIMHAN, S., BHATT, N., 2015, “Deconstructing principal component analysis using a data reconciliation perspective”, *Computers & Chemical Engineering*, v. 77 (jun.), pp. 74–84. ISSN: 00981354. doi: 10.1016/j.compchemeng.2015.03.016. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0098135415000873>>.
- NARASIMHAN, S., JORDACHE, C., 2000, *Data reconciliation & gross error detection an intelligent use of process data*. Houston, Gulf Pub. Co. ISBN: 978-1-61583-657-4 978-0-88415-255-2 978-0-08-050371-4. Disponível em: <<http://www.knovel.com/knovel2/Toc.jsp?BookID=3423>>.

- Noura, H. (Ed.), 2009, *Fault-tolerant control systems: design and practical application*. Advances in industrial control. Dordrecht, Springer. ISBN: 978-1-84882-652-6.
- OGUNNAIKE, B. A., RAY, W. H., 1994, *Process dynamics, modeling, and control*. Topics in chemical engineering. New York, Oxford University Press. ISBN: 978-0-19-509119-9.
- OLIVEIRA, F. V., LOPES, G. P., 2014, *Modelagem e Simulação de Controle de um Reator Contínuo de Poliestireno*. Trabalho de Conclusão de Curso, Universidade de São Paulo, São Paulo. Disponível em: <<http://sites.poli.usp.br/p/augusto.neiva/TCC/TCCs-finais-2014/2014-09.pdf>>.
- PATTON, R., CLARK, R., FRANK, P. M., 2000, *Issues of fault diagnosis for dynamic systems*. London; New York, Springer. ISBN: 978-3-540-19968-7 978-1-84996-995-6.
- PILLONETTO, G., 2016, “A new kernel-based approach to hybrid system identification”, *Automatica*, v. 70 (ago.), pp. 21–31. ISSN: 00051098. doi: 10.1016/j.automatica.2016.03.011. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0005109816300917>>.
- PINTO, J. C., SOUZA JUNIOR, M. B., FEITAL, T., et al., 2017, *Relatório Técnico Final do Projeto de Pesquisa e Desenvolvimento: PEQ-17997 - Monitoramento da Condição de Medidores e Equipamentos de Processo Via Reconciliação de Dados e Estimação de Parâmetros*. Relatório técnico, UJRJ, Rio de Janeiro, jan.
- POULIEZOS, A. D., STAVRAKAKIS, G. S., 2011, *Real time fault monitoring of industrial processes*. Dordrecht; London, Springer. ISBN: 978-90-481-4374-0.
- PROCESSNET, 2006. “Modular Plants: Flexible Chemical Production by Modularization and Standardization - Status Quo and Future Trends”. dez. Disponível em: <https://dechema.de/dechema_media/modularplants.pdf>.
- RAYKAR, V., ANKUR, A., 2002, *CMSC878R Project Report: Fast Kernel Principal Component Analysis for the Polynomial and Gaussian Kernels*. Relatório técnico, University of Maryland.

- RICE, R. C., JYRINI, R. R., COOPER, D. J., 2010. “Performance Monitoring Fundamentals: Demystifying Performance Assessment Techniques”. Disponível em: <<https://controlstation.com/category/white-papers/>>.
- ROWEIS, S. T., 2000, “Nonlinear Dimensionality Reduction by Locally Linear Embedding”, *Science*, v. 290, n. 5500 (dez.), pp. 2323–2326. ISSN: 00368075, 10959203. doi: 10.1126/science.290.5500.2323. Disponível em: <<http://www.sciencemag.org/cgi/doi/10.1126/science.290.5500.2323>>.
- RUSSELL, E. L., 2012, *Data-driven methods for fault detection and diagnosis in chemical processes*. Place of publication not identified, Springer. ISBN: 978-1-4471-1133-7.
- SAMY, I., GU, D.-W., 2012, *Fault detection and flight data measurement: demonstrated on unmanned air vehicles using neural networks*. N. 419, Lecture notes in control and information sciences. London, Springer. ISBN: 978-3-642-24051-5 978-3-642-24052-2.
- SCHÖLKOPF, B., SMOLA, A., MÜLLER, K.-R., 1998, “Nonlinear Component Analysis as a Kernel Eigenvalue Problem”, *Neural Computation*, v. 10, n. 5 (jul.), pp. 1299–1319. ISSN: 0899-7667, 1530-888X. doi: 10.1162/089976698300017467. Disponível em: <<http://www.mitpressjournals.org/doi/10.1162/089976698300017467>>.
- SCILAB ENTERPRISES, 2017. “Scilab: Programa para Computação Numérica”. Disponível em: <<https://www.scilab.org/>>.
- SEBORG, D. E., EDGAR, T. F., MELLICHAMP, D. A., et al., 2010, *Process Dynamics and Control*. 3 ed. Estados Unidos, John Wiley & Sons.
- SHU, Y., MING, L., CHENG, F., et al., 2016, “Abnormal situation management: Challenges and opportunities in the big data era”, *Computers & Chemical Engineering*, v. 91 (ago.), pp. 104–113. ISSN: 00981354. doi: 10.1016/j.compchemeng.2016.04.011. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0098135416301077>>.
- SILVERMAN, B. W., 1998, *Density estimation for statistics and data analysis*. N. 26, Monographs on statistics and applied probability. Boca Raton, Chapman & Hall/CRC. ISBN: 978-0-412-24620-3.
- SIMPSON, W. R., 1994, *System test and diagnosis*. Place of publication not identified, Springer-Verlag New York. ISBN: 978-1-4613-6163-3.

- SKOGESTAD, S., 1991. “Modelling and Dynamic Simulation for Process Control: Issues In Dynamic Modelling Usign Distillation as An Examples”. ago.
- SOBHANI-TEHRANI, E., KHORASANI, K., 2009, *Fault diagnosis of nonlinear systems using a hybrid approach*. N. 383, Lecture notes in control and information sciences. New York, Springer. ISBN: 978-0-387-92906-4 978-0-387-92907-1.
- SOUZA, P. N., SOARES, M., AMARAL, M. M., et al., 2011, “Data Reconciliation and Control in Styrene-Butadiene Emulsion Polymerizations”, *Macromolecular Symposia*, v. 302, n. 1 (abr.), pp. 80–89. ISSN: 10221360. doi: 10.1002/masy.201000063. Disponível em: <<http://doi.wiley.com/10.1002/masy.201000063>>.
- TRIERWEILER, J. O., 1997, *A systematic approach to control structure design*. N. 97,1, Schriftenreihe des Lehrstuhls für Anlagensteuerungstechnik der Universität Dortmund. Als ms. gedr ed. Aachen, Shaker. ISBN: 978-3-8265-3031-9.
- VANHATALO, E., KULAHCI, M., BERGQUIST, B., 2017, “On the structure of dynamic principal component analysis used in statistical process monitoring”, *Chemometrics and Intelligent Laboratory Systems*, v. 167 (ago.), pp. 1–11. ISSN: 01697439. doi: 10.1016/j.chemolab.2017.05.016. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0169743917300734>>.
- VENKATASUBRAMANIAN, V., RENGASWAMY, R., KAVURI, S. N., 2003a, “A review of process fault detection and diagnosis”, *Computers & Chemical Engineering*, v. 27, n. 3 (mar.), pp. 313–326. ISSN: 00981354. doi: 10.1016/S0098-1354(02)00161-8. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0098135402001618>>.
- VENKATASUBRAMANIAN, V., RENGASWAMY, R., KAVURI, S. N., et al., 2003b, “A review of process fault detection and diagnosis”, *Computers & Chemical Engineering*, v. 27, n. 3 (mar.), pp. 327–346. ISSN: 00981354. doi: 10.1016/S0098-1354(02)00162-X. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S009813540200162X>>.
- VENKATASUBRAMANIAN, V., RENGASWAMY, R., YIN, K., et al., 2003c, “A review of process fault detection and diagnosis”, *Computers & Chemical Engineering*, v. 27, n. 3 (mar.), pp. 293–311. ISSN: 00981354. doi: 10.1016/S0098-1354(02)00160-6. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0098135402001606>>.

- VIDAL, R., MA, Y., SASTRY, S. S., 2016, *Generalized principal component analysis*. N. volume 40, Interdisciplinary applied mathematics. New York, NY, Springer. ISBN: 978-0-387-87810-2.
- VON BÜNAU, P., MEINECKE, F. C., MÜLLER, K.-R., 2009, “Stationary Subspace Analysis”. In: Adali, T., Jutten, C., Romano, J. M. T., et al. (Eds.), *Independent Component Analysis and Signal Separation*, v. 5441, Springer Berlin Heidelberg, pp. 1–8, Berlin, Heidelberg. ISBN: 978-3-642-00598-5 978-3-642-00599-2. doi: 10.1007/978-3-642-00599-2_1. Disponível em: <http://link.springer.com/10.1007/978-3-642-00599-2_1>.
- WANG, X., KRUGER, U., IRWIN, G. W., 2005, “Process Monitoring Approach Using Fast Moving Window PCA”, *Industrial & Engineering Chemistry Research*, v. 44, n. 15 (jul.), pp. 5691–5702. ISSN: 0888-5885, 1520-5045. doi: 10.1021/ie048873f. Disponível em: <<http://pubs.acs.org/doi/abs/10.1021/ie048873f>>.
- WEBBER JR., C. L., MARWAN, N., 2014, *Recurrence quantification analysis: theory and best practices*. New York, Springer. ISBN: 978-3-319-07154-1.
- WENTZELL, P. D., ANDREWS, D. T., HAMILTON, D. C., et al., 1997, “Maximum likelihood principal component analysis”, *Journal of Chemometrics*, v. 11, n. 4 (jul.), pp. 339–366. ISSN: 0886-9383, 1099-128X. doi: 10.1002/(SICI)1099-128X(199707)11:4<339::AID-CEM476>3.0.CO;2-L. Disponível em: <<http://doi.wiley.com/10.1002/%28SICI%291099-128X%28199707%2911%3A4%3C339%3A%3AAID-CEM476%3E3.0.CO%3B2-L>>.
- WITCZAK, M., 2014, *Fault Diagnosis and Fault-Tolerant Control Strategies for Non-Linear Systems*, v. 266, *Lecture Notes in Electrical Engineering*. Cham, Springer International Publishing. ISBN: 978-3-319-03013-5 978-3-319-03014-2. Disponível em: <<http://link.springer.com/10.1007/978-3-319-03014-2>>.
- WOLD, S., 1994, “Exponentially weighted moving principal components analysis and projections to latent structures”, *Chemometrics and Intelligent Laboratory Systems*, v. 23, n. 1 (abr.), pp. 149–161. ISSN: 01697439. doi: 10.1016/0169-7439(93)E0075-F. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/0169743993E0075F>>.
- YANG, J., ZHANG, M., SHI, H., et al., 2017, “Dynamic learning on the manifold with constrained time information and its application for dynamic process

monitoring”, *Chemometrics and Intelligent Laboratory Systems*, v. 167 (ago.), pp. 179–189. ISSN: 01697439. doi: 10.1016/j.chemolab.2017.04.009. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0169743916304002>>.

ZBILUT, J. P., THOMASSON, N., WEBBER, C. L., 2002, “Recurrence quantification analysis as a tool for nonlinear exploration of nonstationary cardiac signals”, *Medical Engineering & Physics*, v. 24, n. 1 (jan.), pp. 53–60. ISSN: 13504533. doi: 10.1016/S1350-4533(01)00112-6. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S1350453301001126>>.

ZHANG, S., ZHAO, C., 2017, “Stationarity test and Bayesian monitoring strategy for fault detection in nonlinear multimode processes”, *Chemometrics and Intelligent Laboratory Systems*, v. 168 (set.), pp. 45–61. ISSN: 01697439. doi: 10.1016/j.chemolab.2017.07.010. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S016974391730343X>>.

ZHOU, C., ZHANG, W., 2015, “Recurrence Plot Based Damage Detection Method by Integrating Control Chart”, *Entropy*, v. 17, n. 5 (abr.), pp. 2624–2641. ISSN: 1099-4300. doi: 10.3390/e17052624. Disponível em: <<http://www.mdpi.com/1099-4300/17/5/2624>>.

ZOLGHADRI, A., 2013, *Fault diagnosis and fault-tolerant control and guidance for aerospace vehicles*. New York, Springer. ISBN: 978-1-4471-5312-2.

Apêndice A

Terminologia Usada na Área de Monitoramento de Processos

Após exaustiva consulta a diversas fontes bibliográficas, verificou-se que a terminologia utilizada na área de monitoramento de processos varia bastante de autor para autor, de forma que o objetivo desta seção é apresentar as principais definições e termos da área. Assim como proposto por FEITAL (2011), a maioria dos termos aqui descritos foi extraída de ISERMANN e BALLÉ (1997); entretanto, alguns outros termos foram adicionados para facilitar o entendimento do texto.

Falha (*fault*) - Desvio não permitido de, pelo menos, uma propriedade característica ou parâmetro do sistema em relação a uma condição aceitável, usual ou padrão. Portanto, a falha não é um desvio qualquer do processo, mas um desvio particular, indesejado e prejudicial à operação, sendo por isso mesmo não permitido.

Pane (*failure*) - Interrupção permanente da habilidade do sistema de realizar uma função requerida sob condições de operações especificadas. Trata-se, portanto, de uma falha especialmente grave, por ser permanente.

Pane de Modo Comum (*Common Cause Failure*) - Pane em mais de um dispositivo, componente ou sistema, em decorrência de uma mesma causa direta, num período de tempo relativamente curto e não sendo tais falhas uma consequência da outra. Como exemplos de causa comum, pode ser citada a ação de atmosfera corrosiva, a interferência eletromagnética, a vibração mecânica, o entupimento das tomadas de instrumento, a perda de alimentação elétrica, a perda de pressão pneumática ou hidráulica, incêndio, explosão, descarga atmosférica, procedimento inadequado de fabricação, instalação, condicionamento, operação, manutenção, treinamento inadequado, vício ou limitação de projeto. A pane de modo comum, portanto, tem característica de falha global.

Mau Funcionamento (*malfunction*) - É a intermitência no cumprimento de uma função desejada do sistema.

Erro (*error*) - É o desvio entre o valor medido ou computado (de uma variável de saída) e o "valor verdadeiro"¹, especificado ou teoricamente correto.

Distúrbio (*Disturbance*) - É uma entrada desconhecida (e não controlada) que age sobre o sistema.

Perturbação (*Perturbation*) - É a entrada que age sobre o sistema e que resulta em um desvio temporário do estado corrente.

Resíduo (*Residual*) - É um indicador de falha, baseado no desvio entre medições e valores computados com auxílio de equações ou modelos.

Sintoma (*Symptom*) - É a mudança do comportamento normal de uma quantidade observável.

Detecção de Falhas (*Fault Detection*) - É o procedimento usado para a determinação das falhas presentes no sistema e do tempo de detecção.

Isolamento de Falhas (*Fault Isolation*) - É o procedimento usado para a determinação do tipo, localização e tempo de detecção de uma falha, sendo uma etapa posterior à etapa de detecção de falhas.

Identificação de Falhas (*Fault Identification*) - É o procedimento usado para a determinação do tamanho e do comportamento variante no tempo de uma falha, sendo uma etapa posterior à etapa isolamento de falhas.

Diagnóstico de Falhas (*Fault Diagnosis*) - É o procedimento que compreende as etapas de isolamento e identificação de falhas.

Monitoramento (*Monitoring*) - É a tarefa contínua e em tempo real de determinação das condições de um sistema físico, por meio de registro de informações, reconhecendo e indicando anomalias em seu comportamento.

¹De acordo com o Vocabulário Internacional de Metrologia (INSTITUTO NACIONAL DE METROLOGIA, QUALIDADE E TECNOLOGIA (INMETRO), 2012), o valor verdadeiro é impossível de ser conhecido.

Supervisão (*Supervision*) - É o processo de monitoramento de um sistema físico e a realização de ações apropriadas para manter a operação no caso de falhas.

Redundância Analítica (*Analytical Redundancy*) - É a aplicação simultânea de duas ou mais maneiras (não necessariamente idênticas) para determinar uma variável, sendo que uma delas pode utilizar um modelo matemático do processo de forma analítica.

Cenário (*Scenario*) - É uma sequência de eventos não propositalis que culmina em um dano.

Dano (*Harm*) - É a consequência concretizada ou resultado final de um evento perigoso sobre seres humano, meio ambiente e/ou patrimônio, expresso em termos de número de fatalidades, prejuízos ao meio ambiente, destruição de propriedades, perda de produção, dentre outros.

Risco (*Risk*) - É a combinação da frequência esperada de ocorrência de um cenário com a severidade da consequência do mesmo.

Frequência da Consequência (*Frequency of Consequence*) - É a frequência esperada de ocorrência da consequência indesejada.

Severidade da Consequência (*Severity*) - É uma medida quantitativa de danos a pessoas, meio ambiente e patrimônio.