

# **Ferramentas de Automação e *Machine Learning* para Operação Segura e Ecoeficiente de Processos de Combustão**

**Andréa Pereira Parente**



Tese de Doutorado apresentada ao Programa de Pós-Graduação em Engenharia de Processos Químicos e Bioquímicos da Escola de Química da Universidade Federal do Rio de Janeiro.

Orientadores

Maurício Bezerra de Souza Jr., D.Sc.

Andrea Valdman, D.Sc.

Rossana Odette Mattos Folly, D.Sc.

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO  
ESCOLA DE QUÍMICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PROCESSOS QUÍMICOS E BIOQUÍMICOS  
2019

**FERRAMENTAS DE AUTOMAÇÃO E MACHINE  
LEARNING PARA OPERAÇÃO SEGURA E ECOEFICIENTE DE  
PROCESSOS DE COMBUSTÃO**

*Andréa Pereira Parente*

Tese de doutorado submetida ao corpo docente do curso de Pós-Graduação em Engenharia Processos Químicos e Bioquímicos da Escola de Química da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários para a obtenção do grau de Doutora em Ciências.

Aprovado por:

---

Maurício Bezerra de Souza Jr., D.Sc.  
(orientador – presidente da banca)

---

Andrea Valdman, D.Sc.  
(orientadora)

---

Rossana Odette Mattos Folly, D.Sc.  
(orientadora)

---

Andrea Medeiros Salgado, D.Sc.

---

Flávio Vasconcelos da Silva, D.Sc.

---

Márcia Peixoto Vega Domiciano, D.Sc.

---

Marta Cristina Picardo, D.Sc.

Rio de Janeiro, RJ – Brasil

Parente, Andréa Pereira,

Ferramentas de automação e *machine learning* para operação segura e ecoeficiente de processos de combustão / Andréa Pereira Parente.

xix, 239 p.; il.

Tese (doutorado) – Universidade Federal do Rio de Janeiro, Escola de Química, Programa de Engenharia de Processos Químicos e Bioquímicos, Rio de Janeiro, 2019.

Orientadores: Maurício Bezerra de Souza Jr.; Andrea Valdman; Rossana Odette Mattos Folly.

1. Automação. 2. *Machine Learning*. 3. Combustão. 4. Tese. (Doutorado – UFRJ/EQ).

I. De Souza Jr., Maurício Bezerra (Orient). II. Valdman, Andrea (Orient) III. Folly, Rossana Odette Mattos (Orient). IV. Título.

À Eliza e a rede de balançar.

À Ivette e as mil e uma noites.

*Eu cá com meus botões de carne e osso*

*Eu falo e ouço*

*Eu penso e posso.*

*Gilberto Gil, Cérebro Eletrônico*

## **AGRADECIMENTOS**

Ficam aqui registrados meus mais sinceros agradecimentos aos meus pais, Paulo e Eliza, e também aos meus irmãos, Ana e André, e ao meu cunhado, André, por terem garantido que eu poderia ser o que quisesse na vida. Ainda que professora/cientista maluca seja a minha opção.

Aos amigos queridos da universidade e da vida. A minha escrita é com certeza atravessada pelo privilégio de ter vocês por perto. Em especial a Tiça, Massad, Isa, Cacá e Nakao, que foram amorosamente meio revisoras-meio terapeutas.

Ao Fundão, pelos 13 anos de dores e delícias.

Aos meus 3 orientadores: Andrea Valdman, por quase sempre responder as minhas perguntas com outras perguntas, dividindo um pouco do seu dom da curiosidade comigo. Rossana Folly, por me incentivar a apertar o botões e puxar os fios para ver o que acontece, dividindo um pouco do seu dom da coragem comigo. Maurício Bezerra, por mostrar tantas maneiras diferentes de enxergar um problema de modelagem e a vida, dividindo um pouco dos seus dons do otimismo e da generosidade comigo.

Ao apoio financeiro da Conselho Nacional de Pesquisa e Desenvolvimento – CNPq.

Resumo da Tese de Doutorado apresentada ao curso de Pós-Graduação em Engenharia de Processos Químicos e Bioquímicos da Escola de Química/UFRJ como parte dos requisitos necessários para obtenção do grau de Doutora em Ciências.

## **FERRAMENTAS DE AUTOMAÇÃO E *MACHINE LEARNING* PARA OPERAÇÃO SEGURA E ECOEFICIENTE DE PROCESSOS DE COMBUSTÃO**

Andréa Pereira Parente

Julho, 2019

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

Prof. Andrea Valdman, D.Sc.

Prof. Rossana Odette Mattos Folly, D.Sc.

Embora as plantas de processamento químicas sejam altamente automatizadas, a automação do gerenciamento de situações anormais ainda não ocorreu. Este trabalho integrou tecnologias digitais e analógicas para desenvolver ferramentas de automação para suporte a operação em uma estrutura integrada que mescla soluções matemáticas – implementadas por meio de algoritmos em linguagem C e *softwares* comerciais – com sistemas industriais para aquisição de dados e comunicação – SCADA e redes *Fieldbus*. Foram desenvolvidos modelos de inferência baseados em técnicas de *machine learning* para predição de variáveis de difícil medição, com média dos erros quadráticos de 0,006% v/v, e modelos de classificação para detecção e diagnóstico de falhas, com índice global de acertos superior a 70%. Além disso, uma estratégia de controle ativo tolerante a falhas foi implementada para automatizar ações remediadoras, prevenindo perdas ambientais e econômicas. Este projeto também propôs alternativas para lidar com obstáculos para a implementação em larga escala destas ferramentas de automação nos ambientes industriais: técnica para manutenção *online* dos sensores virtuais com tempo de resposta de até 1 min e 80% de acurácia para atualização dos modelos de predição em novas condições operacionais e simulações de Monte Carlo com análises de *cluster* para expansão massiva de dados de operação para superar a falta de dados representativos de situações de falha. As ferramentas de automação propostas foram validadas com dados de processo reais com potencial para minimizar o consumo de água e reagentes químicos e reduzir as emissões poluentes e as paradas operacionais não-planejadas. Embora as ferramentas desenvolvidas tenham sido implementadas em processos de combustão específicos, elas foram construídas em interfaces homem-máquina amigáveis e replicáveis, podendo ser estendidas a outros casos de aplicação.

Abstract of a Thesis presented to curso de Pós-Graduação em Engenharia de Processos Químicos e Bioquímicos - EQ/UFRJ as partial fulfillment of the requirements for the degree of Doctor of Sciences.

## **AUTOMATION AND MACHINE LEARNING TOOLS TOWARDS SAFE AND ECO-EFFICIENT COMBUSTION PROCESSES**

Andréa Pereira Parente

July, 2019

Supervisors: Prof. Maurício Bezerra de Souza Jr., D.Sc.  
Prof. Andrea Valdman, D.Sc.  
Prof. Rossana Odette Mattos Folly, D.Sc.

Although chemical processes are highly automatic nowadays, the automation of abnormal situation management is yet to be accomplished. This work integrated digital and analog technologies to develop automation tools to assist operation personnel in a framework that merges mathematical solutions – implemented by algorithms in C programming language and commercial software – with industrial data acquisition and communication systems – SCADA and Fieldbus networks. Inference models based on machine learning techniques were built to predict difficult-to-measure process variables, with mean square error of 0,006% v/v, and to detect and diagnose fault operations, with global correct rate above 70%. An active fault tolerant control strategy was also implemented to automate remedial actions to prevent environmental and economic losses. This project also proposed alternative solutions to cope with issues that might be obstacles to wide range industrial applications of such automation tools: online soft sensors maintenance technique with time response of up to 1 min and 80% of accuracy for predictions model update in new nominal conditions; and Monte Carlo simulation associated to cluster analysis for massive data expansion to surpass the lack of data in fault scenarios. The automation tools were validated with real process data and have the potential to minimize water and reagents consumption and to reduce pollutants emissions and unplanned shutdowns. Even though they were tested in specific combustion processes they were built based on friendly and replicable human-machine interfaces and may be extended to other applications.



# SUMÁRIO

<b>Parte I - Fundamentos.....</b>	<b>1</b>
<b>Capítulo I – Introdução.....</b>	<b>2</b>
I.1 Motivação .....	2
I.2 Objetivo .....	5
I.2.1 Objetivo Geral .....	5
I.2.2 Objetivos Específicos .....	5
I.3 Organização.....	5
<b>Capítulo II – Revisão Bibliográfica .....</b>	<b>9</b>
II.1 Automação de Processos Industriais .....	9
II.1.1 Visão Geral .....	9
II.1.2 Tecnologia Digital e Redes .....	12
II.1.3 Novas Tecnologias Digitais .....	14
II.2 Inteligência Artificial e Redes Neurais .....	17
II.2.1 Visão Geral .....	17
II.2.2 Sensores Virtuais.....	23
II.2.3 Problemas de Classificação.....	27
II.3 Gerenciamento de Situações Anormais .....	29
II.3.1 Visão Geral .....	29
II.3.2 FDD: Detecção e Diagnóstico de Falhas .....	30
II.3.3 CPM: Monitoramento do Desempenho de Malhas de Controle .....	36
II.3.4 FTC: Controle Tolerante a Falha .....	43
II.3.5 Gerenciamento de Alarmes .....	51

II.4 Caldeiras.....	55
II.4.1 Visão Geral .....	55
II.4.2 Emissão de Poluentes.....	56
II.4.3 Monitoramento, Diagnóstico e Alarmes em Processos de Combustão .....	59
II.4.4 Filosofia de Controle.....	62

## **Parte II - Aplicações em Processos de Combustão Semi-industriais**

.....**68**

### **Capítulo III – Caldeira Semi-industrial de Produção de Vapor.69**

III.1 Caldeira de Produção de Vapor .....	69
III.2 Instrumentação Instalada e Automação Disponível.....	71
III.2.1 Controlador Dedicado.....	71
III.2.2 Controlador Lógico Programável .....	72
III.2.3 Transmissores Fieldbus .....	72
III.2.4 Analisador de Gases Online.....	72
III.2.5 Analisador de Gases Virtual .....	75
III.2.6 Elementos Atuadores .....	75
III.2.7 Sistema Supervisório SCADA.....	76
III.2.8 Malhas de Controle Contínuas.....	79

### **Capítulo IV– Sensor Virtual para Monitoramento de Emissões de Gases: Aplicação em SCADA Industrial .....80**

IV.1 Resumo .....	80
IV.2 Metodologia Sumarizada.....	80
IV.3 Principais Resultados e Discussões .....	83
IV.4 Lista das Principais Contribuições .....	84
IV.5 Artigo Anexo .....	84

**Capítulo V– Ferramenta de Calibração Online para Sensores Virtuais:  
Validação em Planta Semi-industrial .....92**

V.1 Resumo .....	92
V.2 Metodologia Sumarizada .....	92
V.3 Principais Resultados e Discussões .....	94
V.4 Lista das Principais Contribuições.....	96
V.5 Artigo Anexo .....	96

**Capítulo VI– Estratégia de Controle Tolerante a Falha Baseada em  
Sistemas de Diagnóstico de Falhas .....122**

VI.1 Resumo .....	122
VI.2 Metodologia Sumarizada.....	122
VI.3 Principais Resultados e Discussões .....	124
.....	125
VI.4 Lista das Principais Contribuições .....	126
VI.5 Artigo Anexo .....	126

**Parte III - Aplicações em Processos de Combustão Industriais133**

**Capítulo VII– Caldeira Industrial de Recuperação Kraft .....134**

VII.1 Processo de Produção Kraft .....	134
VII.2 Caldeira de Recuperação Kraft .....	135
VII.2.1 Descrição Física .....	135
VII.2.2 Formação de Partículas .....	138

**Capítulo VIII– Monitoramento de Partículas em Caldeira de  
Recuperação Kraft por Machine Learning.....141**

VIII.1 Resumo .....	141
VIII.2 Metodologia Sumarizada .....	141

VIII.3 Principais Resultados e Discussões .....	142
VIII.4 Lista das Principais Contribuições.....	144
VIII.5 Artigo Anexo .....	144
<b>Capítulo IX– Expansão Massiva de Dados Usando Técnica de Monte Carlo Aplicada a Processo Industrial .....</b>	<b>163</b>
IX.1 Resumo .....	163
IX.2 Metodologia Sumarizada.....	164
IX.3 Principais Resultados e Discussões .....	165
IX.4 Lista das Principais Contribuições .....	166
IX.5 Artigo Anexo .....	166
<b>Parte IV - Epílogo .....</b>	<b>186</b>
<b>Capítulo X– Considerações Finais .....</b>	<b>187</b>
X.1 Conclusões .....	187
X.2 Sugestões para Trabalhos Futuros .....	190
<b>Bibliografia.....</b>	<b>192</b>
<b>Apêndice .....</b>	<b>215</b>
A. P&ID .....	215
B. Painel Elétrico CLP .....	216
B.1 Otimização da Instalação Elétrica do Painel do CLP.....	216
B.2 Diagrama Elétrico do CLP.....	217
C. Ampliação da Arquitetura de Automação .....	218
D. Outros Trabalhos Publicados .....	220
D.1 Publicados .....	220

## ÍNDICE DE FIGURAS

Figura 1 Síntese gráfica da organização da tese. ....	8
Figura 2 Número de publicações em AI aplicada a engenharia química.....	19
Figura 3 Representação esquemática do neurônio.....	20
Figura 4 Razões para uso infrequente das ferramentas de apoio à operação.....	40
Figura 5 Gráficos de tendência de variável de processo e saída de controle e prováveis causas-raízes .....	41
Figura 6 Histograma de frequência de tendências faltosas.....	42
Figura 7 Histograma de frequência de tipos de falha .....	42
Figura 8 Características desejáveis e ainda em desenvolvimento em ferramentas CPM..	43
Figura 9 Representação esquemática de sistema de controle PFTC.....	46
Figura 10 Representação esquemática de sistema de controle AFTC .....	47
Figura 11 Tempo de recuperação x Tempo disponível no sistema .....	48
Figura 12 Total de emissões de CO <sub>2</sub> antrópicas por setor em Mt no ano de 2017 .....	58
Figura 13 Consumo de energia por fonte nas indústrias.....	58
Figura 14 Curva de Perdas da Combustão.....	64
Figura 15 Rendimento Térmico Típico de Caldeiras.....	66
Figura 16 Emissão de Poluentes na Combustão .....	66
Figura 17 Produtos da Combustão.....	67
Figura 18 Central de utilidades do Laboratório de Engenharia Química da UFRJ .....	69
Figura 19 Arquitetura do sistema de automação da planta piloto .....	71
Figura 20 Analisador de gases <i>online</i> .....	73
Figura 21 Representação esquemática do sensor virtual .....	75
Figura 22 Sinóptico de operação da caldeira .....	78
Figura 23 Arquitetura de automação da planta piloto.....	81
Figura 24 Fluxograma de informações do funcionamento do sensor virtual .....	82
Figura 25 Validação <i>online</i> do sensor virtual em tempo real a) Medições e predições de CO <sub>2</sub> b) Resíduos de predição e acurácia.....	83
Figura 26 Validação <i>online</i> do sensor virtual para predições de CO <sub>2</sub> .....	96

Figura 27 Fluxograma de decisão FDD.....	123
Figura 28 Diagrama em blocos da estratégia AFTC .....	124
Figura 29 Resposta dinâmica a uma perturbação no set-point sob controle clássico.....	125
Figura 30 Resposta dinâmica a uma perturbação no set-point sob controle AFTC .....	125
Figura 31 Mapa auto-organizável 20x5 a) SOM quantitativo b) SOM qualitativo .....	143
Figura 32 Fluxo de informações da rotina de expansão massiva de dados .....	164
Figura 33 Fluxo de informações do desenvolvimento e validação da rede neuronal artificial.....	165
Figura 34 Montagem original do painel elétrico do CLP .....	216
Figura 35 Montagem do painel do elétrico do CLP após intervenção.....	216
Figura 36 Visão detalhada do sistema de automação da planta piloto .....	218
Figura 37 Fluxo de dados no sistema de gerenciamento de informações.....	219

## ÍNDICE DE TABELAS

Tabela 1 Rede Digitais Disponíveis no Mercado e Propriedades.....	14
Tabela 2 Inteligência Artificial em Aplicações Industriais .....	20
Tabela 3 Terminologia padrão usada em sistemas de diagnóstico de falhas.....	31
Tabela 4 Ferramentas e pacotes de CPM comerciais disponíveis no mercado .....	39
Tabela 5 Limites máximos para a emissão de poluentes em caldeiras à óleo .....	57
Tabela 6 Limites máximos para a emissão de poluentes em caldeiras à óleo .....	57
Tabela 7 Especificações Técnicas da Caldeira .....	70
Tabela 8 Características das Variáveis Disponibilizadas pelo Analisador <i>In situ</i> .....	73
Tabela 9 Tempos Característicos do Ciclo de Análises dos Gases Residuais .....	73
Tabela 10 Variáveis do Analisador <i>In Situ</i> Disponibilizadas no SCADA.....	74
Tabela 11 Variáveis Contínuas Disponibilizadas na Instrumentação <i>Fieldbus</i> .....	77
Tabela 12 Variáveis Discretas Disponibilizadas no Sistema SCADA da Caldeira.....	77
Tabela 13 Desempenho final da recalibração .....	95
Tabela 14 Matriz de confundimento do modelo RNA com dados expandidos .....	166
Tabela 15 Matriz de confundimento do modelo RNA com dados reais.....	166

# NOMENCLATURA

## LISTA DE ABREVIATURAS

<b>Sigla</b>	<b>Descrição</b>
AFTC	Active Fault Tolerant Control
AI	Artificial Intelligence
AM	Asset Management
ANN	Artificial Neural Network
ANEEL	Agência Nacional de Energia Elétrica
APE	Autoprodutores
ASCM	Mapa de Similaridade de Alarmes por Cores
AVC	Acidente Vascular Cerebral
CE	Cross Entropy
CEMS	Continuous Emission Monitoring Systems
CESMII	Clean Energy Smart Manufacturing Innovation Initiative
CLP	Controlador Lógico-Programável
CLM	Closed Loop Monitoring
CONAMA	Conselho Nacional do Meio Ambiente
CPM	Control Performance Monitoring
CPS	Sistemas Cyber-Físicos
CPU	Central Processing Unit
DTW	Dynamic time warping
EEMUA	Engineering Equipment and Materials Users Association
EPA	Environmental Protection Agency
EPE	Empresa de Pesquisa Energética
EQ	Escola de Química
F/I	Conversor do Protocolo Fieldbus para Corrente Elétrica
FB/P	Conversor do Protocolo Fieldbus para Sinal de Pressão
FB	Fieldbus Protocol
FDD	Fault Detection and Diagnosis
FF	Fieldbus Foundation



FPR	False Positive Ratio
FTC	Fault Tolerant Control
GE	General Electric
GHG	Greenhouse Gas
HDAP	Gráfico de Alarmes de Alta Densidade
HMI	Human-Machine Interface
HVAC	Heating, Ventilation and Air Conditioning
IFAC	International Federation of Automatic Control
IHM	Interface Homem-Máquina
INPI	Instituto Nacional da Propriedade Industrial
ISP	Tamanho Intermediário das Partículas
IoT	Internet das Coisas
ISA	International Society of Automation
LADEQ	Laboratório de Engenharia Química
LAN	Local Area Network
LS	Least Square
MDL	Mecanismo de Desenvolvimento Limpo
MLP	Multilayer perceptron
MMA	Ministério do Meio Ambiente
MME	Ministério das Minas e Energia
MPC	Model predictive controller
MSE	Median Square Error
MSPCA	Multi-Scale Principal Component Analysis
NIST	Instituto Nacional de Padrões e Tecnologia
NNS	Nearest Neighbors Search
OPC	Object Linking and Embedding for Process Control
PCA	Principal Component Analysis
P&I	Piping and Instrumentation
PF	Profibus
PFTC	Passive Fault Tolerant Control
PI	Controle Proporcional e Integral

PID	Proporcional, Integral, derivativo
PLC	Programmable Logic Controller
PLS	Partial least square
PNUD	Programa das Nações Unidas para o Desenvolvimento
PIMS	Plant Information Management System
PROMINP	Programa de Mobilização da Indústria
PRONAR	Programa de Controle da Poluição do Ar
RBF	Função de Base Radial
ROC	Receiver Operating Characteristic
RNA	Redes Neurais Artificiais
SANN	Statistica Automated Neural Networks
SCADA	Supervisory Control and Data Acquisition
SDCD	Sistemas Digitais de Controle Distribuído
SISO	Single Input, Single Output
SOM	Self-Organized Maps
SOS	Sum of Square Error
SP	Set-Point
SSE	Soma dos Erros Quadráticos
TI	Tecnologia da Informação
TPR	True Positive Ratio
UI	User Interface
UFRJ	Universidade Federal do Rio de Janeiro
USPTO	United States Patent and Trademark Office
VBA	Virtual Basic for Applications

## LETRAS LATINAS

<b>Símbolo</b>	<b>Descrição</b>
$a_i$	saída predita da $i$ ésima variável
A	evento anormal
c	tempo de amostragem propriamente dito
g	fator de correção para aumentar o número de padrões

$l$	tempo de limpeza ou purga
$mse_{ft}$	parâmetro de desempenho média do erro quadrático
$m_{reg}$	coeficiente angular da reta de regressão
$N$	evento normal
$n$	intervalo de tempo no qual um somatório é tomado
$n_{out}$	número de variáveis de saída da rede MLP
$NP$	número total de parâmetros da rede MLP
$NPa$	número mínimo necessário de padrões para treinamento
$NT$	número total de amostras
$Palt$	parâmetro de desempenho alternartivo
$r^2$	coeficiente de regressão linear
$rpp$	razão ideal entre o número de amostras e o número de parâmetros
$T$	valor alvo
$T$	intervalo de aquisição de dados da estação SCADA
$tc$	tempo de coleta
$t_f$	tempo final de coleta de dados de treino do sensor virtual
$t_i$	tempo inicial de coleta de dados de treino do sensor virtual
$w$	peso sináptico da rede neuronal
$W$	número total de pesos da rede MLP

## LETRAS GREGAS

<b>Símbolo</b>	<b>Descrição</b>
$\alpha$	fator de correção
$\sigma$	desvio padrão
$\Theta$	número total de bias da rede MLP

# **PARTE I - FUNDAMENTOS**

# Capítulo I – Introdução

---

Este capítulo introduz o assunto abordado nesta tese com foco na motivação para a pesquisa e desenvolvimento deste trabalho. Também estão incluídos os principais objetivos e contribuições deste projeto. Em seguida, é apresentada de forma sucinta e didática a estrutura organizacional deste documento.

## I.1 Motivação

O atual paradigma do modo de produção em que vivemos está se tornando insustentável (ALKAYA, BOGURCU, *et al.*, 2015). Os níveis de globalização e de consciência ambiental contemporâneos exigem que a indústria se torne ecoeficiente<sup>1</sup>, utilizando linhas de produção com alta flexibilidade de operação que permitam adaptações rápidas, racionalizando água e energia e reduzindo a emissão de poluentes. Além disso, as exigências em relação ao desempenho de plantas de processo têm se tornado cada vez mais difíceis de serem atendidas: concorrência mais forte, regulamentações ambientais e de segurança mais severas, e rápidas mudanças no cenário econômico. Assim, a missão de uma planta industrial é manter a lucratividade sem desconsiderar valores sociais e ambientais.

Um fator complicador reside no fato de que as plantas modernas se tornaram mais difíceis de operar em razão da tendência a processos complexos e altamente integrados. Para estas plantas, é difícil evitar que perturbações sejam propagadas de uma unidade a outra (SEBORG, EDGAR e MELLICHAMP, 2016). Neste contexto, os parques industriais tornam-se intrinsecamente dependentes de uma instrumentação de medição efetiva capaz de supervisionar variáveis críticas do processo por meio de diagnósticos em tempo real.

Embora estratégias de controle avançado venham sendo progressivamente aplicadas em escala industrial (CAMPOS, TEIXEIRA, *et al.*, 2009), do ponto de vista da arquitetura de automação, elas são o segundo nível de uma pirâmide suportada pelos sistemas de controle regulatório clássicos. Um sistema de controle avançado nunca alcançará os benefícios esperados

---

<sup>1</sup> – entrega de bens e serviços com preços competitivos que satisfaçam as necessidades humanas e tragam qualidade de vida, reduzindo progressivamente impactos ambientais dos bens e serviços, através de todo o ciclo de vida, em linha com a capacidade estimada da Terra em suportar (WBCSD,2000).

se for constantemente desligado pelos operadores. Assim, é fundamental que os instrumentos de medição, atuação e controladores PID sejam altamente confiáveis uma vez que seu desempenho é condição *sine qua non* para o sucesso das camadas de controle superiores. Neste sentido, há uma tendência ao uso de estratégias de controle operacional, que combinam a otimização de malhas *feedback* com previsão de índices operacionais, auto sintonia de controladores e auto recuperação em condições operacionais de falha (CHAI, QIN e WANG, 2014).

As plantas industriais modernas têm centenas de malhas de controle e cada vez menos engenheiros para manter os sistemas em operação. Bauer e colaboradores (2016) reportam que cada engenheiro de controle é responsável por 450 malhas, em média. Portanto, há uma urgência por ferramentas que realizem diagnóstico e análise de problemas associados aos controles regulatórios (CAMPOS, TEIXEIRA, *et al.*, 2009). Existe ainda uma lacuna em aplicações que utilizem as principais funcionalidades inovadoras disponíveis na atual instrumentação digital. Embora ferramentas de detecção e diagnóstico venham sendo extensivamente estudadas pela comunidade acadêmica, elas ainda não foram amplamente aplicadas na indústria (YU e JIANG, 2015). Faltam ferramentas e sistemas de gerenciamento que maximizem os ganhos das plantas digitais (BARATEIRO, JEFFERYS, *et al.*, 2011) e considerem as dificuldades práticas inerentes ao monitoramento, controle e supervisão de complexos industriais.

Em paralelo, a dificuldade em detectar e diagnosticar em tempo hábil, de forma precisa e quantificada, a causa de uma falha ocorrida no processo e tomar medidas cabíveis quase que instantaneamente é evidenciada pela ocorrência de acidentes de alto impacto social e ambiental descritos por Shu *et al.* (2016), Valdman (2013) e Pitblado (2011). O uso de inteligência artificial, em especial, das técnicas do tipo *machine learning*, aparece como uma alternativa para as tarefas de monitoramento e de diagnóstico em um contexto em que a digitalização dos dados está *in voga* com a Indústria 4.0 (GARDELS, 2018). Entretanto, a área dos processamentos químicos não é um verdadeiro reduto “*big data*” ao compararmos com os campos de finanças, jogos e linguagem, por exemplo (VENKATASUBRAMANIAN e VIELE, 2019). Nós lidamos constantemente com dados incompletos e ruidosos e dados de falhas de processos são especialmente difíceis de adquirir. É preciso automatizar as etapas de pré-processamento dos dados de processos reais e desenvolver técnicas de aumento artificial do volume (*data augmentation*) de dados representativos dos estados nominais de operação e das falhas de processo (GAO, DENG e YUE, 2019).

O uso de sistemas de detecção e diagnóstico em associação a monitoramento em malha fechada visa prover ferramentas que ajudem na identificação de malhas que estejam operando mal e permitam a implementação de alternativas de reconfiguração *online* das malhas de controle, que podem incluir re-sintonia ou ainda mudança de estratégia adotada, tornando as estratégias de controle industrial aplicadas tolerantes à falha. Muito embora as plantas de processos químicos modernas sejam altamente automatizadas, a automação do gerenciamento de situações anormais ainda não foi realizada (SHU, MING, *et al.*, 2016). Segundo Campos *et al.* (2009) o maior desafio é que estas ferramentas auxiliares incorporem uma camada de inteligência na arquitetura de automação para ajudar a operação na tomada de ações remediadoras e corretivas. Campos *et al.* (2009) ressaltam, ainda, que a integração entre diferentes ferramentas é uma preocupação relevante e um desafio.

Um dos equipamentos mais encontrados em complexos industriais é a caldeira que gera vapor, utilidade usada em diversos níveis de pressão com inúmeros objetivos em diferentes processos físico-químicos. As caldeiras envolvem um processo de combustão e utilizam, em sua maioria, combustíveis fósseis. Sobre os processos de combustão, estudos sistemáticos vêm sendo conduzidos com o intuito de aperfeiçoar a eficiência e reduzir os decorrentes níveis de poluição atmosférica (PARENTE, VALDMAN, *et al.*, 2013), que são legalmente limitados por agências de proteção ao meio-ambiente nacionais (CONAMA, 2007) e internacionais (EPA, 2013).

Da perspectiva de controle e automação, existe uma lacuna histórica entre teoria e prática que precisa ser estreitada. A experimentação é um passo importante para aprender sobre os processos reais e pode ser reveladora. Investigar aplicações proporciona o conhecimento de quais são os gargalos de desempenho, robustez e confiabilidade e de como encurtar o ciclo de projeto desde a concepção até a implantação (ÁSTRÖM e KUMAR, 2014).

Em síntese, o setor de processamentos químicos está passando por uma inédita transição - que apresenta desafios e oportunidades significativas na modelagem e na tomada de decisões automatizadas. Impulsionada pela convergência de componentes baratos, de poderosas plataformas de comunicação e também pela automação cada vez maior de operações globalmente integradas, a revolucionária indústria 4.0 converge para o uso e gerenciamento de quantidades massiva de dados e de conhecimentos diversos. É neste ponto que a inteligência artificial (AI) baseada particularmente nas técnicas de aprendizagem de máquina (no inglês, *machine learning*) desempenha um papel importante (VENKATASUBRAMANIAN, 2018).

## **I.2 Objetivo**

### ***I.2.1 Objetivo Geral***

Este trabalho tem como objetivo desenvolver ferramentas integradas em uma estrutura unificada para monitoramento, diagnóstico e controle de processos tolerante a falhas visando o gerenciamento de processos de combustão industrial sob os pontos de vista operacional e ambiental.

### ***I.2.2 Objetivos Específicos***

- (A) Revisar a literatura sobre o gerenciamento das situações anormais.
- (B) Propor e validar ferramenta de automação para inferência de variáveis de difícil medição.
- (C) Aplicar e avaliar ferramenta de automação para implementação de modelos de inferência em plataformas industriais.
- (D) Aplicar e avaliar ferramenta de automação para manutenção dos modelos de inferência em plataformas industriais.
- (E) Propor e validar ferramenta para detecção e diagnóstico de falhas de instrumentos analógicos.
- (F) Propor e implementar estratégia de controle tolerante a falhas.
- (G) Propor e validar ferramenta de automação para diagnóstico de fenômeno de dinâmica lenta.
- (H) Propor e validar técnica para aumento massivo de dados (*data augmentation*).

## **I.3 Organização**

Os capítulos desta tese foram agrupados em quatro partes da seguinte maneira:

A parte I inclui os capítulos I e II e trata das discussões iniciais.

O capítulo I, Introdução, apresentou as motivações que levaram à execução deste projeto e incluiu o principal objetivo deste trabalho bem como os seus objetivos específicos. Aqui também se incluiu uma breve explanação sobre a organização deste documento.



O capítulo II, de Revisão Bibliográfica, apresenta o estado da arte de temas relevantes para o desenvolvimento do projeto bem como fundamentos teóricos pertinentes. Trata do gerenciamento de eventos anormais e detalha tópicos que incluem sistemas de detecção e diagnóstico de falhas, monitoramento de desempenho de malhas fechadas, controle tolerante a falha e gerenciamento de alarmes. Aborda os assuntos controle automático de processos e automação de processos por meio de uma perspectiva histórica. Trata das técnicas de inteligência artificial aplicadas a processos de engenharia química com destaque para as ferramentas de *machine learning* do tipo redes neuronais e seus principais usos na atualidade. Contempla, ainda, as práticas usuais de controle industrial de caldeiras e a conjuntura atual brasileira em termos de emissões atmosféricas.

A parte II apresenta as aplicações de ferramentas de automação e de *machine learning* em processos de combustão semi-industriais.

O capítulo III, Caldeira semi-industrial de Produção de Vapor, apresenta as instalações e processos nos quais parte das aplicações experimentais deste trabalho foram realizadas. Descreve o funcionamento da planta piloto da central de utilidades do Laboratório de Engenharia Química (LADEQ) da Universidade Federal do Rio de Janeiro e detalha as características da instrumentação utilizada e da automação disponível.

Os capítulos IV a VI apresentam o resumo de artigos que fazem parte da estrutura da tese. O resumo (e não a cópia do artigo) foi preferido de modo a se evitar conflitos de direitos autorais com as editoras.

O capítulo IV trata do desenvolvimento de uma ferramenta de monitoramento das emissões atmosféricas em um processo de combustão usando um modelo de inferência baseado em técnicas de aprendizagem de máquina. Ele inclui também a proposição de uma solução de automação para viabilizar a implementação *online* em plataformas industriais.

O capítulo V descreve o desenvolvimento e validação de uma ferramenta de automação dedicada a atualização *online* de modelos de inferência baseados em redes neuronais artificiais do tipo *multilayer-perceptron*.

O capítulo VI apresenta a proposição e implementação em uma planta piloto de uma estratégia de controle tolerante a falha ativa, que inclui reconfiguração *online* entre elementos redundantes.

A parte III desta tese apresenta as aplicações de ferramentas de automação e de *machine learning* em processos de combustão industriais. De maneira análoga a parte II, os capítulos VIII a IX resumem artigos que compõe a estrutura deste trabalho.

O capítulo VII, Caldeira Industrial de Recuperação Kraft, mostra o processo de produção de papel e celulose via rota Kraft com foco no funcionamento da caldeira de recuperação e na formação de partículas que levam a falhas de operação.

O capítulo VIII propõe uma técnica para classificação do estado de operação do processo de produção de papel e celulose com base em redes neuronais do tipo mapa auto-organizável.

O capítulo IX mostra o uso de uma solução matemática para aumento massivo do volume de dados com a intenção de lidar com o problema dos dados não uniformemente distribuídos no desenvolvimento de modelos de detecção e diagnóstico de falhas.

A parte IV traz as discussões finais deste projeto.

O capítulo X, Considerações Finais, apresenta as principais conclusões deste trabalho e lista algumas sugestões para o desenvolvimento de trabalhos futuros.

A Figura 1 sintetiza a organização deste documento, em que as conexões relacionam os objetivos específicos aos capítulos e aplicações em que são desenvolvidos.

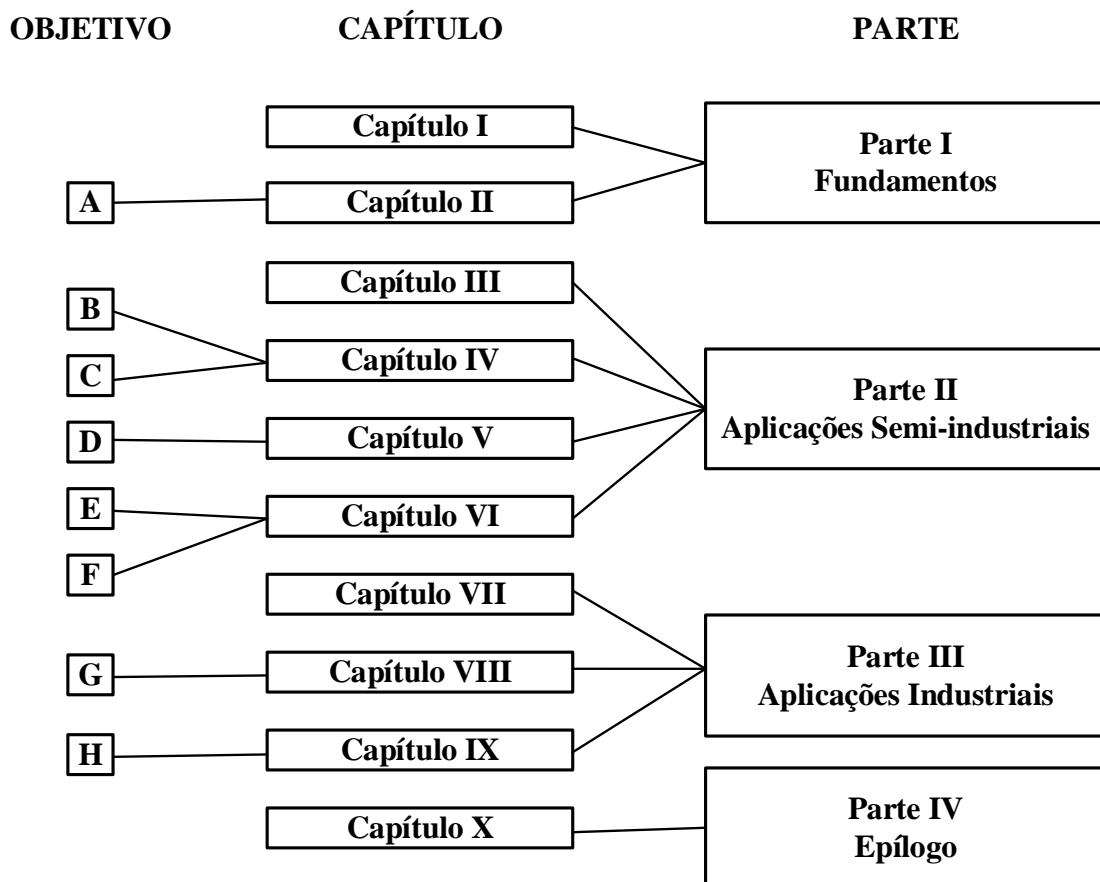


Figura 1 Síntese gráfica da organização da tese.

## Capítulo II – Revisão Bibliográfica

---

Este capítulo apresenta o estado da arte de temas significativos para o desenvolvimento do projeto bem como fundamentos teóricos relevantes, como, por exemplo, desenvolvimento e manutenção de sensores virtuais. Trata do gerenciamento de eventos anormais como uma grande área e detalha tópicos pertinentes que incluem sistemas de detecção e diagnóstico de falhas, monitoramento de desempenho de malhas fechadas, controle tolerante a falha e gerenciamento de alarmes. Aborda o assunto automação de processos por meio de uma perspectiva histórica. Contempla, ainda, as práticas usuais de controle industrial de caldeiras e a conjuntura atual brasileira em termos de emissões atmosféricas.

### II.1 Automação de Processos Industriais

#### *II.1.1 Visão Geral*

O sistema de automação é crítico para a operação das plantas de processo. Nos dias de hoje, ele executa funções de controle e outras funções avançadas que incluem, mas não se limitam, otimização, agendamento e planejamento. A automação dos processos industriais garante a medição de parâmetros apropriados, a análise das situações operacionais, a exploração de oportunidades mais lucrativas, o cálculo e a tomada de ações de controle. Em suma, o sistema de automação mantém o pessoal da planta informado permitindo que seus conhecimentos e habilidades sejam explorados, que as situações anormais sejam identificadas e abordadas e que os processos e os negócios sejam integrados (SAMAD, MCLAUGHLIN e LU, 2007).

Até a década de 40, as plantas de processo eram atuadas manualmente por muitos operadores, com poucos instrumentos mecânicos elementares que realizavam controle local. A evolução natural da indústria com o desenvolvimento de equipamentos maiores e processos mais complexos impactou diretamente nas atividades de comissionamento, operação e manutenção, demandando formas de controle mais precisas. Surgiram, então, os instrumentos pneumáticos, permitindo a transmissão de informações sobre as variáveis do processo através de tubulações específicas até certa distância. Assim, os controladores puderam ficar reunidos em uma mesma sala: a sala de controle do processo (GUTIERREZ e PAN, 2008).

A introdução da eletrônica analógica ocorreu durante as décadas de 50 e 60 e possibilitou a ampliação das distâncias entre os elementos de campo e a sala de controle, simplificando a transmissão de informações e contribuindo para a disseminação de sistemas de controle automático (GUTIERREZ e PAN, 2008).

Em 1970, surgiram os primeiros CLPs (controlador lógico programável), criados para substituir sistemas automáticos que utilizavam relés, temporizadores e sequenciadores mecânicos (GUTIERREZ e PAN, 2008). Estes equipamentos ganharam força a partir de 1980 e, ainda hoje, existem em um grande número de processos industriais, principalmente processos de fabricação de automóveis e de manufatura, baseados em informações discretas (VALDMAN, 2013).

Em 1975, os primeiros SDCDs (Sistemas Digitais de Controle Distribuído) foram lançados, um sistema integrado no qual um nível supervisor de controle monitora múltiplos subsistemas integrados. Inicialmente, esta tecnologia foi lançada para substituir a tecnologia anterior, baseada em CLPs (ABDI, 2009). No entanto, devido às limitações de cada uma das tecnologias citadas, em 1985 começaram a ser utilizados padrões de TI (tecnologia da informação), com a consequente integração de CLPs e SDCDs.

Entre os anos 70 e 80 surgiu a instrumentação eletrônica digital e com ela o grau de automação das instalações industriais foi ampliado. A partir dos anos 80, o computador digital se popularizou, agregando funções de troca de informação em rede, integrando muitas das funções que eram antes dispersas pelos vários elementos de um sistema de automação (SEBORG, EDGAR e MELLICHAMP, 2016). Os sistemas supervisórios tornaram-se uma ferramenta importante, ampliando suas funções de armazenamento de dados históricos (NIMMO, 1999). Nesta época, foram desenvolvidos os primeiros sistemas de Supervisão e Aquisição de Dados (SCADA - Supervisory Control and Data Acquisition) incorporando uma interface homem-máquina (IHM) mais amigável com o usuário (VALDMAN, 2013).

O surgimento da plataforma Windows da Microsoft no final dos anos 80 e pressões do mercado, principalmente o *boom* da utilização de equipamentos de informática em todos os setores de desenvolvimento da automação industrial, conduziram os fabricantes a também adotar plataformas operacionais como parte inerente aos seus equipamentos (VALDMAN, 2013). Os fabricantes de produtos de automação industrial rapidamente aderiram à era de comunicação em redes, propondo a padronização do barramento de campo.

No final dos anos 90, surgem os equipamentos de instrumentação baseados em redes industriais, incorporando protocolos de comunicação em tecnologia digital. Destacam-se o padrão *Fieldbus Foundation* (FF), proposto pelos fabricantes dos Estados Unidos e o padrão Profibus (PF), proposto pelos fabricantes europeus (GUTIERREZ e PAN, 2008).

O aumento da capacidade de processamento dos computadores e componentes microeletrônicos expandiu as aplicações de sistemas de controle de processos automáticos, que passaram a incorporar também o conhecimento de especialistas sobre os processos por meio de técnicas de inteligência artificial. Este salto tecnológico se reflete na miniaturização dos dispositivos e no aumento da sua sensibilidade e precisão além da diminuição do seu tempo de resposta (GUTIERREZ e PAN, 2008).

Nos dias de hoje observa-se uma tendência à padronização de elementos e de protocolos de comunicação reduzindo os custos associados à implantação dos sistemas de controle e aumentando a interoperabilidade entre diferentes fabricantes. Além disso, há uma convergência tecnológica com origens na utilização da tecnologia digital que resulta na interligação dos sistemas de monitoramento e controle de processos com os sistemas de gestão integrada das empresas e o acesso a informações e relatórios de controle pela internet e terminais portáteis (GUTIERREZ e PAN, 2008).

A instrumentação atual oferecida pelo mercado de automação para atender às demandas específicas de processos industriais contínuos utiliza, principalmente, a tecnologia de redes industriais (VALDMAN, 2013). Os novos parques industriais químicos e petroquímicos já exigem essa tecnologia como sendo padrão para seus projetos de instrumentação e, tendo em vista o salto na quantidade de informação e de inteligência em relação às tecnologias anteriores, as indústrias que foram instaladas sem esse padrão estão gradativamente migrando e adaptando a instrumentação de suas plantas, dando origem às plantas de instrumentação híbrida.

Uma tendência tecnológica sugerida pela Honeywell® (SAMAD, MCLAUGHLIN e LU, 2007) é a de que em um futuro não muito distante um novo dispositivo possa ser fisicamente ligado a uma rede e ser automaticamente descoberto pelo sistema e autoconfigurado, com mínima supervisão humana. Um estudo setorial acerca da automação de processos industriais relatado pelo BNDES (GUTIERREZ e PAN, 2008) sugere que novas tecnologias vêm sendo alvo de investimentos, como, por exemplo, a comunicação sem fio (*wireless*) e a adoção de instrumentos baseados em ultrassom, expandindo ainda mais a área de atuação do controle, indo além dos

processos industriais e das plantas de infraestrutura para abranger também sistemas prediais e de logística, procedimentos médicos e automação doméstica, para citar alguns.

A tecnologia *wireless* é agora amplamente vista como um agente de mudança em automação industrial, especialmente valioso para atualizações em plantas já existentes onde o custo para adicionar fiação é proibitivo (SAMAD, MCLAUGHLIN e LU, 2007). Computadores em plantas agora estão conectados à internet e em algumas aplicações via *wireless*. Os dispositivos conectados à rede agora também apresentam servidores web incorporados. Através de qualquer conexão com a internet, os dispositivos de campo podem servir páginas da web com valores de variáveis de processo, parâmetros e telas de configuração de maneira que tais dispositivos podem ser configurados e monitorados por um navegador em qualquer lugar através de uma conexão segura.

Outra tendência emergente se reflete em uma mudança estrutural da arquitetura de automação. Muitas funcionalidades são entregues por meio de *softwares* na forma de aplicações do tipo *stand-alone*: monitoramento, estimações, controle, otimização, visualização de tendências, pontes de integração com outras bases de dados, para citar algumas. Há um surgimento, entretanto, de arquiteturas orientadas para serviço. Nesta nova abordagem, as funcionalidades ficam localizadas em um servidor remoto e são acessadas via internet (SAMAD, MCLAUGHLIN e LU, 2007). Neste contexto segurança cibernética é um tópico *in voga*. Embora não se tenha notícia de acidentes graves causados por invasão cibernética à rede de automação, Byres (2002) reporta a ocorrência de pilhérias localizadas.

### ***II.1.2 Tecnologia Digital e Redes***

Apesar de sua ampla aceitação e aplicação nas indústrias de processos, os transmissores de sinal 4-20 mA já não satisfazem os requisitos necessários aos modernos sistemas de automação. O sinal analógico é apenas um carreador de informação em um fluxo unidirecional do sensor para o sistema de automação ou para o atuador.

A tecnologia de transmissão em redes digitais viabiliza um meio de interconexão de dispositivos de campo a custos baixos ao substituir a conexão ponto a ponto pela conexão multiponto. Assim, os instrumentos para controle se comunicam por meio de protocolos digitais puros, tornando possível a transmissão de informações processadas auxiliares. Ao contrário da

transmissão analógica, em que a informação é transmitida através da modulação de uma corrente elétrica, a informação passa a ser transmitida através de uma sequência de pulsos elétricos de mesma largura, representando dígitos binários de 0 ou 1 (MORRIS, 2001). Este formato oferece informações mais robustas e de maior qualidade de sinal uma vez que a amplitude da onda elétrica responsável pelo ruído é muito menor do que a amplitude da onda elétrica responsável pelo pulso contendo a informação de interesse (GERON, 1998).

Segundo Morris (2001), *Fieldbus* é o termo genérico utilizado para designar os protocolos de alta velocidade, baseado em redes de transmissão digital, que suportam a comunicação bidirecional de informação entre um determinado número de equipamentos inteligentes em uma rede local. As redes industriais mais amplamente disponíveis podem ser classificadas em três categorias principais: *Fieldbus* tradicionais, redes baseadas em Ethernet e redes sem fio (JÄMSÄ-JOUNELA, 2007).

Uma comparação das principais propriedades das redes atualmente mais amplamente disponíveis no mercado em cada uma das três categorias principais é dada na Tabela 1.

Um levantamento das tendências no mercado de automação industrial realizado pela Intech e divulgado pelo BNDES (GUTIERREZ e PAN, 2008) revela que desde 2008 os protocolos Profibus e *Foundation Fieldbus* são os mais abrangentes e que 80% das empresas entrevistadas já usavam instrumentos que utilizam protocolos de comunicação baseados em redes industriais afirmando uma tendência de mercado consolidada.

De um ponto de vista de arquitetura, a tecnologia *Fieldbus* tornou o sistema de controle ainda mais distribuído ao incorporar dispositivos cada vez mais sofisticados em que os transmissores podem ter algoritmos de compressão e escalonamento internos e os atuadores podem incluir processadores nos quais os cálculos de controle podem ser executados (SAMAD, MCLAUGHLIN e LU, 2007).

Além disso, os equipamentos de campo com comunicação microprocessada são capazes de indicar falha em tempo real e disponibilizar diagnóstico preventivo baseado em dados de operação do equipamento e avaliação estatística dos mesmos. As capacidades de auto validação, inerentes aos próprios equipamentos, tornam esta tecnologia de instrumentação atraente para ser usada em associação a ferramentas de detecção e diagnóstico de falhas e de controle tolerante a falha em um *framework* para gerenciamento de situações anormais.



Tabela 1 Rede Digitais Disponíveis no Mercado e Propriedades

Tipo	Rede Digital	Fabricante	Velocidade de Transmissão	Número de Equipamentos	Aplicações
Fieldbus Tradicional	CAN	Bosch	10 kb/s–1 Mb/s, 31,25 kb/s, 1Mb/s, 2,5 Mb/s (5Mb/s fibra óptica)	Máx 32	Controle Automotivo e Industrial, Sistemas Embutidos
	WorldFIP(1158-2)	Schneider		Máx 256 (64/s)	Manufatura, Distribuído em Tempo Real
	Profibus - DP	Siemens	9,6kb/s-12Mb/s	Máx 126	Automação de Manufatura, Ambiente Genérico
	Profibus - PA	Siemens	31,25kb/s	Más 32/s	Controle de Processo, Ambiente Perigoso
	Control Net Interbus	Rock well Auto	5Mb/s	Máx 99	Manufatura
	Control Net Interbus	Phoenix contact	500kb/s	Máx 512	Manufatura, IO remoto, Tecnologia de Manufatura
Fieldbus Tradicional	Asi	Siemens	167kb/s	Máx 62	Barramento de Sensor
	Foundation Fieldbus HI	Fieldbus Foundation	31,25kb/s	Máx 32	Processo
Ethernet	Ethernet/IP	Rockwell Auto	10, 100 Mb/s, 1Gb/s	-	Manufatura, Tecnologia de Sistemas de Manufatura
	Foundation Fieldbus HSE	Fieldbus Foundation	100 Mb/s	-	Manufatura
	Ether CAT	Beckhoof	100 Mb/s	Até 65.535 devices	IO Descentralizado, Controle de Movimento
	Powerlink	B&R	100 Mb/s	-	Controle de Movimento
	Profinet	Siemens	100 Mb/s	-	Manufatura, Objetos de Automação Distribuídos, Comunicação entre Sistemas, Tecnologia de Manufatura
	Sercos III	Bosch Rexroth	100 Mb/s	Máx 254	Controle de Movimento Rápido
Wireless	Modbus TCP	Schneider	10, 100 Mb/s, 1Gb/s	-	Comunicação entre Sistemas
	IEEE 802.11(a/b/g)	-	11 Mb/s, 54 Mb/s	-	SOHO
	Bluetooth, IEEE 802.15.1	-	1 Mb/s	Até 7 escravos/piconet	WPAN, Tecnologia de Substituição de Cabos
	IEEE 802.15.4/ZigBee	-	20, 40, 250 kb/s	Suporte redes grandes	WPAN, Para Sensores e Dispositivos de Controle, Rede de Sensores Sem Fio
	Ethernet (IEEE 802.3)	-	10, 100 Mb/s, 1 Gb/s	Quase ilimitado	-

Fonte: Jämsä-Jounela, 2007 - modificada

### II.1.3 Novas Tecnologias Digitais

As três primeiras revoluções industriais resultaram em mudanças radicais nos processos de produção, introduzindo mecanização, eletricidade e tecnologia de informação. Desde as máquinas a vapor até a produção automatizada elétrica e digital, estas mudanças levaram ao aumento da eficiência e da produtividade. Não obstante, o atual paradigma do modo de produção em que vivemos está se tornando insustentável (ALKAYA, BOGURCU, *et al.*, 2015).

O impacto da produção industrial sobre o meio ambiente em termos de aquecimento global e poluição ambiental é grave. O consumo de recursos não-renováveis, como petróleo e

carvão, aumenta. De um ponto de vista global, a indústria sofre de uma oferta cada vez menor de força de trabalho em razão do envelhecimento da população (WANG, WAN, *et al.*, 2016). Neste contexto, é indispensável que os processos industriais alcancem alta flexibilidade e eficiência associados ao baixo consumo de energia e redução de custos.

Conseqüentemente, uma iniciativa estratégica denominada "Industrie 4.0" (Indústria 4.0) foi proposta e adotada pelo governo alemão como parte do "*High-Tech Strategy 2020 Action Plan*" ("Recommendations for implementing the strategic initiative INDUSTRIE 4.0", 2013). Estratégias similares também foram propostas por outros países industrializados, por exemplo, "Internet Industrial" pelos Estados Unidos e "Internet +" pela China (WANG, WAN, *et al.*, 2016), estabelecendo a quarta revolução industrial.

A revolução industrial 4.0 se caracteriza por um conjunto de tecnologias que permitem a fusão do mundo físico, digital e biológico (MINISTÉRIO DA INDÚSTRIA, COMÉRCIO E SERVIÇOS, 2019). Sob a Indústria 4.0, o processo de produção consistirá em troca de informações, máquinas controladas e unidades de operação atuando de forma autônoma e inteligente (QIN, LIU e GROSVENOR, 2016), impulsionado por dados e por inteligência artificial (VENKATASUBRAMANIAN e VIELE, 2019).

Segundo o plano de ações proposto pelo governo alemão, a fábrica inteligente é a característica mais importante da Indústria 4.0. Nela, os itens relevantes para a indústria, como materiais, sensores, máquinas, produtos, cadeia de suprimentos e clientes, podem ser conectados. Isto significa que estes objetos vão trocar informações e ações de controle uns com os outros de forma independente e emancipada (QIN, LIU e GROSVENOR, 2016).

Do ponto de vista do engenheiro de controle, a fábrica inteligente pode ser vista como um sistema de malha fechada dupla. Um laço consiste em recursos físicos e nuvem, enquanto o outro laço consiste em terminais de controle de supervisão e nuvem (WANG, WAN, *et al.*, 2016).

Este sistema será possível com o auxílio de tecnologias de ponta emergentes na última década orientadas para os sistemas Cyber-Físicos (CPS) como os sensores sem fio, a computação em nuvem, a análise de Big Data (FLOURIS, GIATRAKOS, *et al.*, 2016), a internet móvel e, sobretudo, a Internet das Coisas (IoT) (WANG, WAN, *et al.*, 2016).

A Internet das Coisas (IoT) derruba as barreiras entre o mundo real e o mundo virtual ao transformar todos os objetos físicos em objetos inteligentes (OUADDAH, MOUSANNIF, *et al.*, 2017). Atzori, Iera e Morabito (2017) apresentaram a evolução da Internet das Coisas desde suas

ideias raízes até a geração tecnológica atual. Mas, em uma abordagem mais ampla que perpassa pelos objetos do dia-a-dia, nossas roupas, móveis, carros, luzes de casa ou até mesmo potes de café terão sua própria conta nas mídias sociais, interagirão com outras redes, enviarão dados para a nuvem e permitirão a agregação de dados de diferentes fontes e aspectos da nossa vida. Os medidores inteligentes deduzirão quando tomamos banho, os carros saberão quando não estamos indo para o trabalho, os dispositivos médicos saberão nosso peso, e os celulares saberão como estamos nos sentindo (OUADDAH, MOUSANNIF, *et al.*, 2017). Se esta realidade soa desconfortável do ponto de vista de privacidade e segurança para as pessoas, quando se trata da segurança dos processos industriais o controle de acesso é condição *sine qua non* e vem sendo estudado nas pesquisas recentemente divulgadas acerca de segurança cibernética. Li, Han e Jin (2016) e Ouaddah *et al.* (2017) são alguns exemplos.

A empresa especializada em segurança cibernética Symantec estima que as perdas financeiras das empresas, que incluem reparos e tempo de recuperação, causadas pelos crimes virtuais chegam a US\$385 bilhões anuais (SYMANTEC). A invasão dos sistemas com intenções nocivas de violação de dados é tipicamente carregada através de *malware*, que é, por definição, qualquer código computacional capaz de causar danos ao sistema (JANG-JACCARD e NEPAL, 2014). As vítimas de *malware* não se limitam aos usuários finais, mas incluem servidores, equipamentos em redes digitais e sistemas SCADA (JANG-JACCARD e NEPAL, 2014). As ameaças atuais incluem, além do roubo de dados econômicos e de propriedade intelectual (PARENT e CUSACK, 2016), sequestro do sistema inteiro através de uma modalidade denominada *ransomware*, em que o acesso é completamente bloqueado e devolvido mediante pagamento de resgate (MANSFIELD-DEVINE, 2016).

De acordo com a empresa CISCO, especializada em tecnologia da informação, em 2020 haverá 50 bilhões de dispositivos conectados (EVANS, 2011). Isto explica a insurgência de uma série de trabalhos dedicados a discutir as vulnerabilidades dos sistemas modernos como em Jang-Jaccard e Nepal (2014). Amin, Schwartz e Sastry (2013) publicaram um estudo em que a segurança dos sistemas de controle em redes foi abordada por meio da teoria de jogos, em que cada equipamento inteligente é considerado um jogador. Modi *et al.* (2013) publicaram uma pesquisa sobre invasores que afetam a disponibilidade, confiabilidade e integridade dos recursos e serviços em nuvem e discutiram técnicas de identificação e prevenção contra estas ameaças.

Muitos outros estudos recentes focam na resiliência industrial dos sistemas de infraestrutura crítica, como as plantas de energia e os sistemas de distribuição da rede elétrica, diante de ataques cibernéticos. Yusta, Correa e Lacal-Aránzaga (2011), Bekara (2014) e Langer *et al.* (2016) são alguns exemplos. Um obstáculo para a evolução das ferramentas de segurança cibernética para processos industriais reside no fato de que testar as soluções de segurança em fábricas reais é quase impossível; porque isto significa desacelerar ou mesmo desligar operações.

Uma série de publicações se preocupa em conceituar os requerimentos que caracterizam a indústria 4.0, como Oesterreich e Teuteberg (2016) e Stock e Seliger (2016), mas ainda não há um consenso padrão entre os autores. Outra tendência é a preocupação em estabelecer uma abordagem organizacional para a arquitetura das Fábricas Inteligentes através de metodologias (WANG, WAN, *et al.*, 2016) e *frameworks* (QIN, LIU e GROSVENOR, 2016).

## **II.2 Inteligência Artificial e Redes Neurais**

### ***II.2.1 Visão Geral***

Desde a sua concepção, o termo inteligência artificial (AI) foca em entender como os computadores podem imitar o cérebro humano em um contexto de tomar decisões (KHARGONEKAR e DAHLEH, 2018). Em uma definição não-rigorosa, proposta por Rich, em 1983, “AI é o estudo de como fazer as máquinas fazerem coisas em que, no momento, as pessoas são melhores”.

Historicamente, o termo AI engloba robótica, como os carros autônomos; processamento de linguagem e modelagem semântica, como a assistente virtual Siri, da Apple®; visão e processamento de imagens, como o reconhecimento facial e a primeira imagem de um buraco negro; jogos, como o Deep Blue, o AlphaGo, da Google® e o Watson, da IBM® – que venceram partidas de xadrez, GO e quiz; os sistemas especialistas, como, por exemplo, o MYCIN, desenvolvido pela Universidade de Stanford, para diagnosticar doenças contagiosas e as técnicas de aprendizagem de máquina, como as redes neurais e as análises de clusters (VENKATASUBRAMANIAN, 2018). Dentre os possíveis desdobramentos para aplicação da AI, os sistemas especialistas e as técnicas de aprendizagem de máquina, particularmente as redes neurais, são os mais propensos a serem diretamente e imediatamente aplicados em problemas de engenharia química.

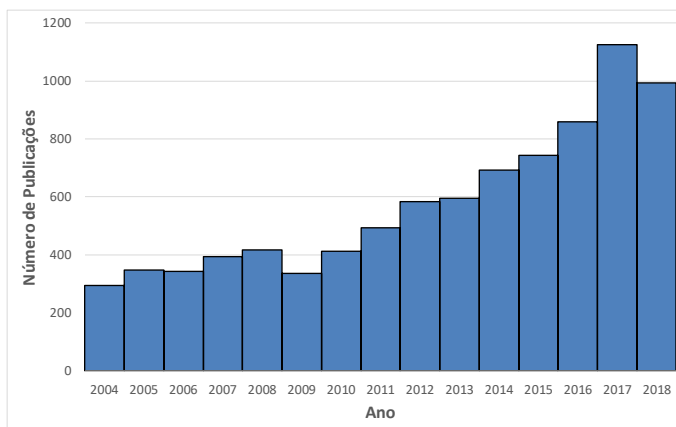
Um sistema especialista é um programa de computador que usa conhecimento profundo para resolver problemas complexos e que, normalmente, exigem especialistas humanos. Sistemas especialistas incorporam conhecimento por meio de relacionamentos e monitoram as relações e as inferências que invocam. Portanto, o conhecimento do sistema é explícito e acessível ao usuário. Um sistema especialista pode explicar por que precisa de certas informações e como chega a certas conclusões.

O aprendizado de máquina (em inglês, *machine learning*) é um método de análise de dados que automatiza a construção de modelos analíticos. É um ramo da inteligência artificial baseado na ideia de que sistemas podem aprender com dados, identificar padrões e tomar decisões com o mínimo de intervenção humana.

Uma rede neuronal é um tipo de aprendizado de máquina que representa um paradigma computacional composto de vários nós simples altamente interconectados, tal qual o sistema nervoso humano, que processam informações por meio da sua resposta dinâmica a estímulos externos. As redes neuronais aprendem por meio de exemplos, em uma espécie de lógica indutiva, sendo, portanto, capazes de aplicar o que aprenderam a um novo cenário de entrada e prever a saída apropriada.

Em analogia ao cérebro humano como um modelo para a arquitetura de processamento na inteligência artificial, um sistema especialista simula as funções de raciocínio qualitativo do hemisfério esquerdo, enquanto uma rede neural simula as capacidades de modelagem quantitativa do lado direito do cérebro (BAUGHMAN e LIU, 1995).

Nos últimos tempos, o interesse acadêmico sobre o tema AI em engenharia química tem crescido exponencialmente, como mostra a Figura 2, que reflete o número de publicações na base de dados científicos *ScienceDirect* com as palavras chave “artificial intelligence” e “chemical engineering”. Há também grande expectativa sobre o potencial comercial das aplicações de AI, atraindo somas significativas de investimentos globalmente (GARDELS, 2018).



**Figura 2** Número de publicações em AI aplicada a engenharia química

Apesar das atenções recentes, a aplicação de AI em problemas da indústria química não é recente. Venkatasubramanian (2019) reporta implementações bem-sucedidas de sistemas especialistas e de redes neurais desde os anos 80. Entretanto, havia uma série de barreiras que limitaram os impactos dos desenvolvimentos da época. Do ponto de vista tecnológico, faltava potência computacional, capacidade de armazenamento de dados, velocidade de comunicação e transmissão de informações e ambientes de programação amigáveis. Em paralelo, as atenções do mercado se voltavam para as aplicações de otimização e controle preditivo baseado em modelo, que não requeriam grandes volumes de dados nem apresentavam alta exigência computacional.

Em contrapartida, nos últimos anos, houve um salto tecnológico em que as limitações de recursos foram majoritariamente superadas. No contexto da Indústria 4.0, em ambientes do tipo *big data*, os ganhos das aplicações de técnicas como as de *machine learning* já são palpáveis. Além disso, as barreiras organizacionais e psicológicas foram sendo gradativamente diminuídas. Uma vez que as pessoas já aceitam prontamente recomendações de assistentes virtuais como Google, Alexa e Yelp para uma variedade de tarefas, as empresas iniciam uma verdadeira corrida para incorporar os recursos AI nas suas instalações, evidenciada, por exemplo, pelo número de palestras voltadas para aplicações industriais apresentadas em conferências recentes. A Tabela 2 mostra alguns exemplos.

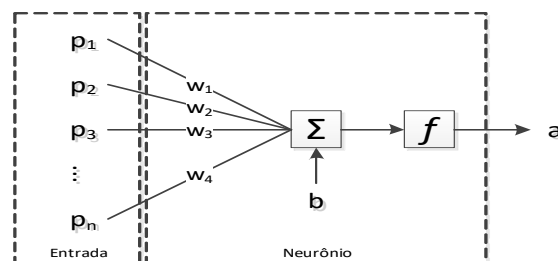
**Tabela 2 Inteligência Artificial em Aplicações Industriais**

<b>Tema</b>	<b>Autor</b>	<b>Evento</b>	<b>Empresa</b>	<b>Ano</b>
Industria 4.0 no Agronegócio	Bergamaschi, M.	COBISA	Lide Agronegócios	2019
A Transformação Digital na Braskem	Morales, D.	COBISA	Braskem	2019
The Digital Twin in the Chemical Industry	Giron, B.	PSE	Aveva	2019
Gradient Boosting Virtual Analyzer for Real-Time Process Monitoring in a Polymer Slurry Process	Neumann, G.	PSE	Braskem/Radix	2019
Digital Twin applied in the sugar industry	Soares, R.	PSE	OptimaTech	2019

Venkatasubramanian (2019), entretanto, chama a atenção para uma questão limitadora das técnicas de aprendizagem de máquina que é justamente a necessidade de um volume alto de dados. Mesmo que nos dias de hoje a quantidade de dados aquisitados seja muito maior do que no passado, a área dos processamentos químicos não é um verdadeiro reduto “*big data*” ao compararmos com os campos de finanças, jogos e linguagem, por exemplo. Nós lidamos constantemente com dados incompletos e ruidosos.

Dentre as técnicas de aprendizagem de máquina, destacam-se as redes neurais, que constituem uma metodologia com capacidade de aprender e reconhecer padrões operacionais (BAGAJEWICZ, 2001). Estes modelos são identificados por uma arquitetura, especificada pela característica funcional dos neurônios, pela topologia da rede e pelo algoritmo de treinamento utilizado (HOSKINS e HIMMELBLAU, 1988).

Os elementos básicos de um neurônio podem ser vistos na Figura 3, que apresenta a transformação dos estímulos  $p$  em uma informação, através dos sinais de entrada, da respectiva ponderação dos pesos  $w$  pelas sinapses, da adição ocorrida dentro do elemento principal e da restrição na amplitude de saída de acordo com a função de ativação do mesmo.

**Figura 3 Representação esquemática do neurônio**

Aplicada a um processo industrial, a rede neuronal deve ser capaz de fornecer respostas baseadas em um conhecimento adquirido, fornecido em uma etapa de aprendizado prévia. Em linhas gerais, ela é constituída de uma camada de entrada, recebendo as informações do processo, uma ou mais camadas escondidas, processando internamente as informações por meio de duas funções de ativação, produzindo as sinapses, e uma camada de saída, disponibilizando o resultado da análise (VALDMAN, 2010).

O objetivo de uma rede neuronal é mapear um conjunto de padrões de entrada em um conjunto de padrões de saída. A rede realiza esse mapeamento aprendendo primeiro com uma série de exemplos anteriores, definindo conjuntos de correspondências de entrada e saída para o sistema fornecido. Em seguida, a rede aplica o que aprendeu a um novo padrão de entrada para prever a saída apropriada. Os algoritmos de aprendizagem podem ser divididos em três classes: algoritmos supervisionados, não-supervisionados e reforçados (MOHAMMADI e MINAEI, 2019). Para o algoritmo supervisionado, um conjunto de pares de dados de entrada-saída (geralmente dados experimentais) são fornecidos para a rede. Na aprendizagem não-supervisionada, é apresentado um conjunto de entradas para a rede e a própria rede descobre propriedades úteis e organiza as entradas em *clusters* ou categorias. Na aprendizagem reforçada, existe um sistema de recompensas para acertos e erros e as conexões de rede são modificadas com base nas informações recebidas dos laços de realimentação.

Algumas das mais tradicionais aplicações de redes neurais incluem:

- Classificação: redes neurais podem ser usadas para identificação de imagens em fotos de satélite, para distinguir submarinos de rochas em um sonar e identificação de doenças em eletrocardiogramas;

- Redução de ruído: na presença de sinais ruidosos, é possível a substituição destes pelos valores determinados pela rede treinada;

- Predição: um uso muito comum é a predição de valores de variáveis dado um histórico de valores (MASTERS, 1993).

Dentre as arquiteturas de redes neurais artificiais com maior aplicabilidade, destacam-se as redes *perceptrons* multicamadas (MLP), as redes de função de base radial (RBF) e a rede de mapeamento auto-organizável (SOM).

A arquitetura mais utilizada para a resolução de problemas de predição é a chamada *Perceptrons* Multicamadas (MLP, em inglês *Multilayer Perceptron*), na qual as redes podem



apresentar uma ou mais camadas ocultas. Nessa arquitetura, a rede é caracterizada pela quantidade de nós em cada camada assim como pela quantidade de camadas e o tipo de conexão existente entre os neurônios. Ela apresenta um fluxo de informação do tipo *feedforward* e seu treinamento é supervisionado, sendo normalmente utilizado o algoritmo de retropropagação (*backpropagation*) (BAUGHMAN e LIU, 1995).

Outra arquitetura de rede neuronal relevante, mais utilizada para problemas de classificação de dados, é a RBF (do inglês *Radial Basis Function*) que também é fundamentada no fluxo *feedforward* e, assim como as redes MLP, apresenta um treinamento do tipo supervisionado. As redes RBF apresentam apenas três camadas, ou seja, apenas uma camada oculta.

O processo de treinamento das redes RBF é fundamentado na formação de agrupamentos (*clusters*) dos padrões de entrada, formado no espaço multidimensional das entradas. Cada agrupamento possui um centro e é possível, portanto, conhecer a distância de um determinado conjunto de dados de entrada do centro de cada “*cluster*”. Dessa forma, os neurônios calculam a distância geométrica entre o vetor de entrada (de um conjunto de dados) e um vetor referência, que representa a localização do centro da função de base radial no espaço das entradas. Essa distância é, então, usada para classificar o vetor de entrada, apontando a qual agrupamento ou “*cluster*” ele pertence. Assim, além dos pesos sinápticos, as redes RBF admitem outros dois parâmetros: o centro da função radial e sua dispersão. A função de base radial mais usada é a função gaussiana. As funções radiais apresentam uma natureza local, ou seja, elas geram saídas ativas apenas na vizinhança de seu centro de ativação (HAYKIN, 2001).

Por fim, outra arquitetura de grande importância é a de mapeamento auto-organizável ou SOM (do inglês, *Self-Organized Map*) que se destaca por ser uma rede do tipo não-supervisionada, ou seja, que não recebe os valores de saída desejada, mas apenas os dados de entradas durante o treinamento.

O princípio de funcionamento dessa rede é baseado na organização do córtex humano, em que cada região é excitada por sensações diferentes e específicas. Dessa forma, em uma rede de mapeamento auto-organizável, entradas diferentes excitam neurônios diferentes de tal modo que neurônios que se encontram próximos uns aos outros são ativados por padrões (entradas) apresentando similaridades (BAUGHMAN e LIU, 1995).

Em outras palavras, a rede de mapeamento auto-organizável realiza a modelagem da distribuição de probabilidade dos vetores de entrada e agrupa, em um mapa, vetores similares. Assim, vetores próximos geometricamente estão localizados em pontos próximos no mapa e são separados em uma mesma classe.

A rede SOM é uma rede de camada única em forma de grade uni ou bidimensional, na qual os neurônios se posicionam lado a lado, formando vizinhanças. O treinamento da rede é dividido basicamente em três etapas. Primeiramente, na etapa de competição, os neurônios são estimulados pelos dados de entrada e competem entre si de tal forma que o neurônio vencedor (seguindo uma abordagem do tipo “o vencedor leva tudo” ou do inglês “*winner-takes-all*”) seja aquele mais similar e que melhor represente a informação de entrada. Para determinação da similaridade entre neurônio e entrada, são empregadas funções discriminantes, sendo a mais usada a distância euclidiana entre os pontos que representam, respectivamente, a entrada e o neurônio. Em seguida, ocorre o chamado processo de cooperação, no qual o neurônio vencedor ativa somente os neurônios vizinhos. Na última etapa, no processo de adaptação sináptica, os pesos sinápticos desses neurônios excitados são ajustados (DE SOUZA JR, 2015).

Em resumo, as entradas, que podem apresentar  $n$  dimensões, são projetadas em um mapa de dimensão inferior (uni ou bidimensional) e a rede pode ser interpretada como uma transformação não-linear permitindo essa organização. Quando uma nova entrada é apresentada à rede após o treinamento, ela é representada no mapa auto-organizável e incorporada ao *cluster* (agrupamento) mais próximo.

### ***II.2.2 Sensores Virtuais***

Sensores virtuais ou *soft sensors*, são modelos de inferência que fazem uso das informações disponibilizadas por sensores *online*, tais como pressão, vazão, temperatura, nível e etc., para prever variáveis de processo que não podem ser medidas automaticamente ou que só podem ser medidas a custos altos, esporadicamente e com muitos atrasos. Este é o caso das variáveis associadas à qualidade dos produtos que são tipicamente analisadas em laboratório. Eles são ferramentas importantes para inúmeras indústrias de processo como Papel e Celulose, Refinarias, Polímeros, Cimento e Tratamento de Resíduos, para citar algumas (SOUZA, ARAÚJO e MENDES, 2016).

Existem três tipos de abordagem para o desenvolvimento de sensores virtuais no que diz respeito ao modelo usado: físicos, estatísticos e baseados em inteligência artificial (tais como redes neuronais e lógica *fuzzy*). Dentre eles, os modelos empíricos (também conhecidos como *data-driven models*), quando bem aplicados, são os que geram os melhores resultados uma vez que conseguem acompanhar a dinamicidade dos processos (FORTUNA, 2007).

Os dados históricos para os sensores virtuais desenvolvidos a partir de dados empíricos podem ter origem em um sistema de supervisão e aquisição de dados (SCADA) ou ainda em análises e anotações de laboratório. As principais questões concernentes ao desenvolvimento de um sensor virtual incluem o pré-processamento dos dados para lidar com dados ausentes, dados aquisitados a taxas distintas e detecção de *outliers*. Além disso, a seleção dos dados de entrada bem como a escolha da estrutura de modelo adequada e da técnica de treinamento empregada devem ser consideradas. A validação do modelo também é uma etapa crucial no desenvolvimento de um sensor preditivo para julgar se o modelo construído é capaz de informar as variáveis-alvos com desempenho satisfatório. Existe ainda uma etapa posterior, denominada manutenção do modelo, para lidar com mudanças no processo. Esta etapa representa um tópico de pesquisa ainda em aberto nos dias de hoje (SOUZA, ARAÚJO e MENDES, 2016).

Aplicações reais em sistemas complexos normalmente direcionam a escolha do modelo para a classe de modelos baseados exclusivamente em dados, uma vez que profundos conhecimentos físico-químicos acerca do processo são requeridos quando uma abordagem fenomenológica é preferida. Dentre os modelos do tipo caixa-preta, é preciso decidir entre modelos lineares e modelos não-lineares. Embora a tomada de decisão seja intrinsecamente dependente da aplicação, muitos autores sugerem que um modelo linear sempre deve ser a primeira tentativa deixando as soluções mais complexas para os casos em que esta abordagem não é satisfatória.

Qin (1997) apontou que modelos baseados em redes neuronais artificiais são particularmente úteis no desenvolvimento de sensores virtuais e González e colaboradores (2003) reportaram que modelos baseados em redes neuronais artificiais têm sido extensivamente empregados na solução de problemas da indústria petroquímica por apresentarem predições mais acuradas. Alguns exemplos são González (2003), que focou na predição do comportamento de gases de reservatório, Chai *et al.* (2014) que usaram técnicas de aprendizagem profunda para prever temperaturas de flash e de ebulição em uma refinaria, Razi e colaboradores (2013) bem

como Bispo e colaboradores (2016) que trabalharam na predição de propriedades reológicas de fluidos de perfuração.

Kalogirou (2003) publicou um extenso levantamento da aplicação de inteligência artificial em controle de processos de combustão. Fazendo referência particularmente aos sistemas de combustão que incluem caldeiras, fornos e incineradores, observou-se que em mais de 51% dos casos de aplicação utilizou-se métodos de redes neuronais artificiais.

Outras aplicações de sensores virtuais baseados em redes neuronais artificiais (ANN) relatadas na literatura incluem o trabalho de Kamat e Madhavan (2016) que desenvolveu sensores preditivos para uma ampla gama de aplicabilidade que englobam o gerenciamento das baterias de veículos exclusivamente elétricos e híbridos além do controle de combustão em veículos a combustível no setor automotivo; parâmetros de qualidade na indústria de cimento e predição de pH em processamentos químicos.

A primeira geração de sensores virtuais baseados em dados se desenvolveu por meio da modelagem *offline* usando dados históricos arquivados (KADLEC, GRBIC e GABRYS, 2011). A coleção destes dados, entretanto, contém informação limitada, que corresponde a um período limitado e que é possivelmente limitada a um conjunto de regiões operacionais do espaço de estados.

Portanto, em novos cenários que não estão descritos nos dados armazenados a tendência é que o sensor virtual apresente desempenho degradado. Estes novos cenários podem ocorrer em razão da natureza dos processos por inúmeros motivos: mudanças de matéria-prima; depósito e incrustação nos equipamentos e tubulações; abrasão de componentes mecânicos; alterações da atividade do catalisador; mudanças no ambiente externo, para citar alguns (KADLEC, GABRYS e STRANDT, 2009). Neste contexto, a área hoje denominada manutenção do modelo tem a intenção de manter uma boa resposta mesmo quando há mudanças no processo.

Equipar um sensor virtual com capacidade adaptativa requer duas tarefas a serem executadas. Primeiro, a necessidade de adaptação deve ser reconhecida através do monitoramento do desempenho do modelo. Isto pode ser realizado por meio da comparação da saída do modelo com a informação adquirida a partir de análise laboratorial ou de um analisador real. Em aplicações práticas, este passo é muitas vezes ignorado e os modelos são adaptados em intervalos periódicos ou pela implementação de fatores de esquecimento constantes (KADLEC, GRBIC e GABRYS, 2011).

Uma vez que a necessidade de adaptação foi identificada, a adaptação real deve ser realizada. De modo geral, isto é feito por meio da atualização recursiva dos parâmetros do modelo em uma etapa denominada retreinamento.

Em Souza, Araújo e Mendes (2016) este é um problema interpretado como desvio de conceito (*concept drift*), em que as propriedades estatísticas da variável alvo mudam com o tempo. Neste trabalho, os autores sugerem três abordagens clássicas para lidar com a adaptação: seleção de amostras (*sample selection*), ponderação de amostras (*sample weighting*) e aprendizagem em grupo (*ensemble learning*).

Nos casos em que o pré-processamento de dados é realizado na etapa de desenvolvimento do modelo para lidar com problemas como *outliers*, dados ausentes, etc., a adaptação do modelo também deve incluir os métodos de pré-processamento *online*.

Os primeiros sensores virtuais adaptativos foram baseados nas técnicas de janela móvel para a seleção de amostras, em que os dados dentro da janela são usados para a atualização recursiva dos parâmetros do modelo enquanto os dados fora da janela são descartados. As atualizações podem ser feitas através dos métodos de mínimos quadrados (LS), da análise de componentes principais (PCA) e dos mínimos quadrados parciais (PLS), para citar alguns (KADLEC, GRBIC e GABRYS, 2011).

Na estratégia de ponderação da amostra, atribui-se peso para as amostras de acordo com sua idade (a importância das amostras diminui ao longo do tempo). O aprendizado/adaptação dos parâmetros é geralmente feito por aprendizagem exponencial recursiva (SOUZA, ARAÚJO e MENDES, 2016).

Métodos do tipo *ensemble* encontram diferentes variáveis-alvos a serem preditas no histórico e aprendem um modelo para cada um destes objetos. A previsão final é uma combinação do conjunto dos modelos disponíveis em um arranjo que leva em conta a probabilidade de cada modelo ser responsável pelos dados a serem previstos (SOUZA, ARAÚJO e MENDES, 2016). É um método de alta demanda computacional.

Alguns exemplos de sensores virtuais adaptativos particularmente modelados por meio de redes neurais artificiais relatados na literatura podem ser encontrados em Liukkonen e colaboradores (2013), Lee e colaboradores (2005) e em Wu, He e Chen (2000), para citar alguns. Além disso, Kadlec, Grbic e Gabrys (2011) revisam inúmeros casos de estudo de sensores virtuais adaptativos baseados em diferentes estruturas de modelo.

### ***II.2.3 Problemas de Classificação***

Os problemas de classificação para a indústria de processamento químico focam em detectar e diagnosticar falhas e em categorizar características de sistemas em estado estacionário e em regime transiente.

As maiores aplicações para categorização são em bioprocessos. Por exemplo, Simon e Karim (1999), usaram redes neurais para classificar as fases de crescimento celular *lag*, exponencial e estacionária com a intenção de aumentar a produção de  $\beta$ -lactamase. Gadkar *et al.* (2005) também usaram redes neurais para classificar os estados de um processo fermentativo. Em seu trabalho, eles desenvolveram uma rede neuronal recorrente e notaram que, diante de adaptação *online* dos pesos, é possível usá-la para controle automático do processo.

Uma outra área de aplicação crescente das técnicas de aprendizagem de máquina em problemas de categorização é a de desenvolvimento de materiais (VENKATASUBRAMANIAN, 2018). Verpoort *et al.* (2018), por exemplo, usaram redes neurais para detectar e identificar medições incorretas de algumas propriedades. Seu trabalho foi validado em casos de estudo que incluíram diferentes ligas e materiais poliméricos. Outro caso é o trabalho de Wu e colaboradores (2018), que usou redes neurais para classificar as variáveis de projeto de baterias de lítio, atingindo especificações de potência e energia.

Outras inúmeras aplicações em problemas de classificação residem na área médica. O trabalho de Hosni *et al.* (2019) revisa e resume uma série de exemplos para detecção e classificação dos estágios do câncer de mama. O estudo de Garg e colaboradores (2019) propõe a automação da classificação de um subtipo isquêmico de AVC usando redes neurais para processamento de linguagem natural.

Entretanto, para a indústria de processos químicos as aplicações de detecção e diagnóstico de situações anormais de operação são as mais abrangentes. Koronakia *et al.* (2019) e Wu e Zhao (2018) são alguns exemplos. Vale ressaltar que a tarefa de classificar os estados de operação da plantas de processamento, apesar de extremamente importante do ponto de vista de segurança e eficiência operacional, não é trivial. Os processos químicos são extremamente complexos, integrados e não-lineares. Flutuações podem ser propagadas pela planta de processo, mas não necessariamente se refletem em mudanças claras de condições operacionais, dependendo do grau de interação e correlação das variáveis de processo.

Métodos de classificação estatísticos são uma alternativa para lidar com estas questões. Tong e Haizhenyuxuhua (2019) e Zhan, Li e Yang (2019) são alguns exemplos. Entretanto, de acordo com Suewatanakul (1993), as redes neuronais superam o desempenho da abordagem estatística convencional na maioria das aplicações em engenharia química possivelmente em razão das suas propriedades para organizar e perceber características em conjuntos de dados imprevisíveis e/ou imprecisos (HOSKINS e HIMMELBLAU, 1988).

As redes neuronais classificadoras de padrões mapeiam padrões de entrada distintos em suas respectivas classes de saída e podem ser divididas em classificadores paramétricos e não paramétricos (BAUGHMAN e LIU, 1995). Um classificador paramétrico tem uma distribuição de probabilidade associada às respostas de saída que estão sendo classificadas e é treinado prevendo os melhores parâmetros para a distribuição estatística. Já o classificador não paramétrico não possui uma distribuição de probabilidade para as respostas de saída, mas, em vez disso, prevê a classe de saída comparando-a com os pontos de treinamento mais próximos.

O vetor de saída de uma rede neuronal de classificação é booleano, com zero indicando que o padrão de entrada não está dentro da classe específica e um indicando que está dentro de uma classe específica. A saída da rede neuronal, na verdade, é um valor numérico entre 0 e 1, e pode ser visto como a probabilidade de que o padrão de entrada corresponda a uma classe específica (BAUGHMAN e LIU, 1995). O espaço de resposta de uma rede de classificação pode ser interpretado graficamente por meio das regiões de decisão e de transição. Nas regiões de decisão, todos os pontos estão associados a apenas uma classe de saída. Já nas regiões de transição, as faixas de classificação se misturam e podemos apenas fazer inferências sobre a classificação de saída.

A taxa de classificação incorreta e a robustez do sistema são as duas principais características do modelo que usamos para selecionar o tipo e a configuração de rede neuronal ideal. Diz-se que um sistema é robusto quando a taxa de erros de classificação tem desvios mínimos com pequenas variações nas variáveis de entrada (BAUGHMAN e LIU, 1995). As ferramentas de detecção e diagnóstico de falhas são discutidas em maiores detalhes a seção II.3.2 deste trabalho.

## II.3 Gerenciamento de Situações Anormais

### II.3.1 Visão Geral

Embora o controle altamente automatizado de processos tenha contribuído fortemente para a eficiência de produção, qualidade de produtos e segurança de processos e ambiental, a automação da área de gerenciamento de situações anormais ainda não ocorreu. De acordo com Venkatasubramanian (2005), o Reino Unido, por exemplo, perde 27 bilhões de dólares por ano em razão de situações anormais e Vásquez e colaboradores (2016) reportaram que as perdas da indústria petroquímica nos Estados Unidos chegam a 20 bilhões de dólares anuais. A tomada de decisões confiáveis e em tempo hábil feita com suporte de decisão automático para os operadores em circunstâncias de operação anormais em indústrias de processamento químico deveria ser uma função indispensável nas fábricas inteligentes. O conceito de fábricas inteligentes foi proposto por Weyer e Schmitt (2015) e abrange indústrias com ciclos de produção mais curto e altamente flexíveis e especializadas, que almejam segurança de trabalho aprimorada e competitividade em escala global.

Gerenciamento de situações anormais inclui, mas não se restringe à rápida detecção de um cenário atípico. Ferramentas de diagnóstico das causas de origem, bem como recomendações de ações remediadoras que promovam o retorno à operação normal, devem fazer parte do sistema (VENKATASUBRAMANIAN, RENGASWAMY e KAVURI, 2003b). Estudos recentes reportam que alguns dos acidentes severos na indústria química podem ser atribuídos à deficiência no diagnóstico de eventos anormais em tempo hábil (SHU, MING, *et al.*, 2016).

Embora os sistemas de controle de processos sejam modernos e altamente automatizados, acidentes ainda ocorrem em razão de erro humano e mau funcionamento de equipamentos. Ainda que existam esforços genuínos devotados à produção de componentes mais confiáveis e operação mais segura, não há menções na literatura de equipamentos que nunca falham nem de seres humanos que nunca cometem erros. Ferramentas de supervisão *online* para acompanhamento do comportamento do processo em tempo real, incluindo sistemas de detecção e diagnóstico de falhas, podem ser usadas em associação ao monitoramento em malha fechada, provendo recursos que auxiliem os operadores na identificação de malhas de controle de desempenho degradado e sugiram ações remediadoras que podem incluir re-sintonia ou reconfiguração *online*.



Estas ferramentas auxiliares deveriam ser capazes de detectar falhas de instrumentação tais como problemas de calibração, mau dimensionamento, ruídos de medição, medidas fora de escala e falta de resolução, por exemplo; não-linearidades associadas a mudanças de ponto operacional; má sintonia traduzida em oscilação e instabilidade ou ainda problemas de estratégia de controle como acoplamento e perda de graus de liberdade. Algumas destas ferramentas estão disponíveis no mercado, mas a maioria delas requer conhecimentos técnicos e especialistas sobre a natureza do processo e profissionais bem treinados capazes de interpretar as informações disponibilizadas e tomar ações corretas que variam desde substituição de uma válvula de controle até re-sintonia ou implementação de uma nova estratégia de controle (CAMPOS, TEIXEIRA, *et al.*, 2009).

A maior complexidade e o grau de integração das modernas indústrias de processos químicos significa que o potencial de perdas econômicas é maior enquanto as tarefas de diagnóstico e localizações de falhas é mais difícil. Consequentemente, o gerenciamento de situações anormais é visto como o próximo desafio no desenvolvimento e implementação de novos sistemas de controle (JÄMSÄ-JOUNELA, 2007).

### ***II.3.2 FDD: Detecção e Diagnóstico de Falhas***

O monitoramento contínuo do processo associado a ferramentas de detecção e diagnóstico de falhas pode contribuir fortemente para a excelência operacional dos processos de produção fornecendo informações vitais para atividades organizacionais como o gerenciamento de ativos, otimizando intervenções de manutenção por meio de manutenção preventiva e redução de custos de armazenamento de peças sobressalentes e alertando sobre anormalidades que poderiam resultar em custos de reparo, desligamentos não planejados e, até mesmo, acidentes.

Cerca de 40 anos atrás o professor Himmelblau (1978) publicou o primeiro livro de detecção e diagnóstico de falhas na indústria química.

A Tabela 3, extraída de Valdman (2013), resume a terminologia padrão usada para descrever as instâncias de um sistema de diagnóstico de falhas dos dias de hoje.

Tabela 3 Terminologia padrão usada em sistemas de diagnóstico de falhas

	Análise do Sinal	Descrição
Ocorrência	<b>Evento de Falha (Failure)</b>	Interrupção total ou parcial, em determinadas condições operacionais, da capacidade do sistema em executar uma função
	<b>Mau funcionamento (Malfunction)</b>	Irregularidade intermitente, ou não, em relação ao cumprimento de uma função desejada do sistema
Estado	<b>Estado de Falha (Fault)</b>	Desvio não permitido de pelo menos uma propriedade ou parâmetro característico do sistema, em relação à condição normal de operação
Sintoma	<b>Erro (Error)</b>	Desvio entre um valor medido, ou calculado, e o seu valor verdadeiro, especificado previamente ou teoricamente calculado
Tipo	<b>Distúrbio (Disturbance)</b>	Ocorrência de uma condição de entrada desconhecida e fora de controle interferindo no sistema
	<b>Perturbação (Perturbation)</b>	Ocorrência de uma condição de entrada interferindo no sistema, resultando em um desvio temporário do estado corrente
<b>Funções</b>		<b>Descrição</b>
Detecção	<b>Detecção de Falha</b>	Determinação da presença de uma falha no sistema e do tempo de detecção
Diagnóstico	<b>Isolamento da Falha</b>	Determinação do tipo, a localização e o tempo de ocorrência da falha, identificando o estado de falha existente
	<b>Identificação da Falha</b>	Detalhamento da falha (definição do tamanho, da causa, do comportamento com o tempo, dentre outros)
Correção	<b>Correção da Falha</b>	Tomada de ações apropriadas, de forma automática ou não, para restabelecer a capacidade do sistema de desempenhar a função requerida

Fontes: ISERMAN e BALLÉ, 1997; SARTORI, 2010; SARTORI *et al.*, 2012.

Nos anos 80, o comitê técnico de detecção de falhas, supervisão e segurança para progressos técnicos (SAFEPROCESS) da Federação Internacional de Controle Automático (IFAC) criou um simpósio que passou a acontecer a cada 3 anos com o objetivo de fortalecer o contato entre academia e indústria na área (SHU, MING, *et al.*, 2016). Em 1994, a empresa Honeywell organizou um consórcio de gerenciamento de eventos anormais com o objetivo de discutir incidentes e acidentes custosos à indústria química tais como paradas não planejadas, explosões e altos índices de emissões de poluentes. Nos dias de hoje, empresas como SASOL, BP, ExxonMobil, Shell, UOP e Centro de Soluções Humana além de universidades como o Centro Mary Kay O'Connor de Segurança de Processos, universidades da Califórnia, de Los Angeles, da Pensilvânia, de Alberta e a universidade Nanyang de tecnologia compõem o consórcio que é associado ao Instituto Nacional de Padrões e Tecnologia (NIST) e que publicou entre 2008 e 2010 guias para projetos de telas de operação eficazes, práticas de gestão de alarme eficientes e boas práticas processuais (BULLEMER, KIFF e THARANATHAN, 2011). Este consórcio foi precursor do Instituto de Inovação *Clean Energy Smart Manufacturing* (CESMII),

fundado em 2016, com foco no desenvolvimento de tecnologias capazes de capturar, compartilhar e processar em tempo real as quantidades crescentes de informações disponíveis nas plantas industriais (OFFICE OF ENERGY EFFICIENCY AND RENEWABLE ENERGY, 2016).

Muito embora segurança de processos e gerenciamento de situações anormais tenham sido alvo de esforço global desde os anos 90, não há redução evidente no número de acidentes relatados nos Estados Unidos ou na Europa (PITBLADO, 2011). Técnicas de detecção e diagnóstico de falhas em tempo real podem ser soluções viáveis a serem usadas em um contexto cultural de prevenção de acidentes.

Não obstante, ainda existem muitos desafios do ponto de vista das aplicações industriais em larga escala. Por exemplo, uma planta moderna pode apresentar centenas ou milhares de sensores e atuadores. Em razão do alto número de informações medidas, o sistema de gerenciamento de dados tipicamente armazena as informações após compressão que leva à perda de informações dinâmicas. Outros fatores que influenciam a qualidade dos dados industriais são justamente a falta de dados e presença de *outliers*, que deveriam ser filtrados automaticamente.

Outra questão relevante é a capacidade de classificar os dados coletados inicialmente em dados sem falhas e dados em circunstância de falha; tarefa não trivial em plantas complexas que exige intensa colaboração da operação e intensa inspeção dos dados (ZHANG, JEINSCH, *et al.*, 2011).

A análise das causas-raízes das falhas é condição *sine qua non* para que medidas corretivas sejam tomadas. Todavia, as plantas modernas são caracterizadas por alto grau de automação e alta integração energética e de matéria. Como consequência, uma única falha pode se propagar em diferentes direções, impactar em inúmeros equipamentos e causar distúrbios e oscilações ao longo da planta dificultando a identificação da origem da falha (ZHANG, JEINSCH, *et al.*, 2011).

Quando se trata de dados operacionais em malha fechada o desafio é ainda maior porque apresentam dificuldades inerentes para modelagem (CHIUSO e PICCI, 2005), análise de causas-raízes e diagnóstico (GERTLER e CAO, 2004).

Ainda sobre as dificuldades de implementação industrial das ferramentas FDD destaca-se os fenômenos de natureza lenta como depósito e corrosão de equipamentos. Estas falhas, custosas à indústria, exigem intervenções de manutenção como limpeza e regeneração (ZHANG,

JEINSCH, *et al.*, 2011). Neste sentido, maiores informações acerca do progresso do evento (depósito ou corrosão) permitiriam planejamento otimizado das atividades de manutenção e permanecem um tópico de pesquisa em aberto.

Devido à ampla extensão do problema de diagnóstico de falhas em processo e às dificuldades de solução em tempo real, várias abordagens foram desenvolvidas ao longo dos anos. Existem métodos que requerem modelos de processo precisos. Há também os métodos que dependem de modelos qualitativos tais como os modelos de causalidade. Na outra extremidade do espectro, existem os métodos que não presumem qualquer tipo de modelo e contam apenas com informações de histórico do processo, como algoritmos de redes neurais e métodos estatísticos multivariáveis. Além disso, dado o conhecimento do processo, existem diferentes técnicas de busca que podem ser aplicadas para realizar o diagnóstico. Venkatasubramanian, Rengaswamy e Kavuri (2003a, b, c) resumiram e revisaram as principais características das técnicas descritas na literatura.

Historicamente, no entanto, a maioria das aplicações envolvendo ferramentas de diagnóstico baseadas em modelos está voltada para sistemas aeroespaciais, elétricos e mecânicos (ZHANG e JIANG, 2008). Em razão das dimensões e complexidade das indústrias de processos químicos e da quantidade crescente de dados disponibilizados, as técnicas baseadas em dados passados se mostram vantajosas e vêm sendo alvo de pesquisas em FDD (SHU, MING, *et al.*, 2016). Alguns exemplos são Zhang *et al.* (2017), que usaram uma extensão da análise estatística de componentes principais (MSPCA) para detectar e diagnosticar falhas em processo químico não-linear e Santos *et al.* (2014) que usaram redes neurais para detectar e prever a magnitude do vazamento em linhas de gás.

Outra tendência atual concernente às pesquisas de sistemas FDD é a integração de diferentes metodologias. Em seu trabalho de revisão Venkatasubramanian, Rengaswamy, e Kavuri (2003a, b, c) compararam diversas técnicas: observadores, sistemas especialistas, análise de tendências qualitativas, análise de componentes principais, redes neurais, para citar algumas, e concluíram que nenhuma metodologia isolada atende a todos os requisitos desejados para uma ferramenta FDD simultaneamente; sugerindo, portanto, que algumas destas técnicas são complementares. Don e Khan (2019), por exemplo, combinaram, o conhecimento sobre o processo com a abordagem baseada em dados. Desta forma, um modelo oculto de Markov

detecta as situações anormais enquanto uma rede neuronal Bayesiana faz o diagnóstico. A metodologia proposta foi validada com o caso de estudo de referência *Tennessee Eastman*.

Para que seja possível comparar diferentes abordagens para detecção e diagnóstico é útil identificar um conjunto de características desejáveis em um sistema FDD. De maneira análoga ao *trade-off* desempenho x robustez largamente conhecido na literatura clássica de controle, existe uma compensação inerente entre velocidade de detecção e desempenho tolerável do sistema. Um sistema projetado para detectar mudanças abruptas rapidamente será sensível às influências de alta frequência, tais como ruídos, e poderá apresentar alto índice de falsos alarmes durante operação normal. É desejável também que a ferramenta apresente alta isolabilidade, que é a capacidade de distinguir entre diferentes falhas. Não obstante, um classificador com alto grau de isolabilidade tipicamente não apresenta boa rejeição às incertezas de modelo (VENKATASUBRAMANIAN, RENGASWAMY e KAVURI, 2003a).

É pertinente que a ferramenta FDD tenha a habilidade de perceber uma nova falha ao invés de diagnosticá-la erroneamente em uma classe de mau funcionamento conhecida ou em operação normal. Além disso, é importante, embora não seja uma tarefa trivial em razão da natureza altamente interativa das falhas, que o sistema consiga identificar múltiplas falhas (VENKATASUBRAMANIAN, RENGASWAMY e KAVURI, 2003a).

Outras características relevantes incluem estimativa de erro de classificação para aumentar a confiança do usuário nas recomendações do sistema automatizado, pouco esforço computacional e baixa necessidade de armazenamento, bem como baixo esforço de modelagem para que aplicações em tempo real sejam viáveis.

A diversidade operacional e técnica das tarefas envolvendo monitoramento, controle e diagnóstico de processos dificulta o desenvolvimento de um sistema único (VALDMAN, 2013) e os sistemas híbridos vêm se tornando cada vez mais populares (KADLEC, GABRYS e STRANDT, 2009).

Recentemente, muitos trabalhos de sistemas FDD híbridos como Tong *et al.* (2014), Yu e Rashid (2013), Perk *et al.* (2011), Gosh *et al.* (2011) Ng e Srinivasan (2010) foram publicados.

Nos últimos 15 anos, o número de publicações de sistemas contendo as palavras “*fault detection*” no título totaliza mais de 43000 trabalhos. Estes dados foram extraídos da base de busca *ScienceDirect*, disponível gratuitamente para os alunos da UFRJ.

Na Escola de Química/UFRJ trabalhos de Carelli e de Souza Jr. (2009a, 2009b) e de Souza Jr, Campos e Tunalá (2009) aplicando técnicas de detecção e diagnóstico de falhas a processos simulados foram desenvolvidos. Além disso, Mendonça e colaboradores (2009) implementaram ferramentas de apoio em uma unidade de processamento de gás natural e Castro (2009) utilizou dados reais de uma unidade de hidrotreamento de diesel para o desenvolvimento de uma ferramenta FDD.

Apesar das numerosas publicações dos últimos anos, as aplicações dos sistemas de detecção e diagnóstico de falhas propostos em casos reais ainda são escassas e estão refletidas no pequeno número de patentes depositadas também nos últimos 15 anos nas bases INPI, Instituto Nacional da Propriedade Industrial, e USPTO, *United States Patent and Trademark Office*, todas de livre acesso, totalizando 26 patentes brasileiras e 752 internacionais.

Embora as ferramentas FDD venham sendo estudadas há quase 40 anos, elas ainda não são consideradas tecnologias maduras. É possível que a maior barreira para implantação destes sistemas em escala industrial seja a falta de amostragem de falhas históricas. Ainda que haja armazenamento destes dados, é muito comum que não existam descrições dos eventos no livro de registros da sala de controle. E ainda que haja armazenamento de dados e que os eventos tenham sido relatados, é improvável que todas as falhas possíveis tenham ocorrido em uma única planta de processos. Em uma refinaria clássica, são coletadas dezenas de milhares de medidas a intervalos de tempo de 1s. Dessa forma, os custos de armazenamento associados ao pouco uso de dados históricos não justifica arquivamentos por mais de 2 ou 3 anos (SHU, MING, *et al.*, 2016). Este estado da arte atual das indústrias químicas em relação ao armazenamento de dados contrasta com a tendência mundial liderada pelas gigantes Google e Amazon, que consideram dados ativos de alto valor agregado (QIN, 2014). Ainda assim, as atenções dos provedores de serviços e produtos comerciais têm ajudado a propagar as abordagens baseadas em dados de histórico (ZHANG, JEINSCH, *et al.*, 2011).

O estado da arte atual das aplicações industriais de sistemas FDD em larga escala evidencia a urgência por soluções de engenharia mais amigáveis que tornem a implementação do sistema sustentável. O desenvolvimento de algoritmos FDD clássicos inclui muitas etapas: aquisição de dados, pré-processamento dos dados, modelagem, aplicação dos métodos FDD, validação do algoritmo e projeto da interface homem-máquina. As plantas vigentes estão frequentemente sujeitas a modificações em razão de mudanças de carga, inovações tecnológicas,

substituição de máquinas e equipamentos. Logo, é plausível que os algoritmos FDD precisem ser atualizados ocasionalmente para se adequarem às novas condições de trabalho (ZHANG, JEINSCH, *et al.*, 2011). Neste sentido, ferramentas unificadas que reduzam os esforços de engenharia e a necessidade de conhecimento *a priori* para projeto e reprojeção dos sistemas FDD poderiam levar a uma maior aceitação do ponto de vista industrial.

Alternativamente, a tecnologia digital permitiu que sensores e atuadores equipados com sistemas de autodiagnóstico interno integrem progressivamente as indústrias de processo. Em paralelo, instrumentos de tecnologias já consolidadas como equipamentos analógicos tendem a ser usados continuamente. Embora estes instrumentos não apresentem auto validação intrínseca, eles podem ter suas capacidades expandidas por meio de uma arquitetura de automação híbrida. Assim, o uso de características inerentes à própria instrumentação industrial pode ser uma solução viável para o problema de detecção e diagnóstico de falhas em tempo real.

### ***II.3.3 CPM: Monitoramento do Desempenho de Malhas de Controle***

Nas indústrias de processos químicos, os controladores de estrutura proporcional-integral-derivativa (PID) são os mais extensivamente usados em razão de sua simplicidade, robustez e ampla gama de aplicabilidade (JENG, TSENG e CHIU, 2014).

Miller e Desborough (2002) reportam que 97% dos controladores industriais são do tipo PID e Koivo e Tanttu (1991, apud O'Dywer, 2006, p. 1) sugerem que apenas de 5 a 10% das malhas de controle não podem ser reguladas por controladores PI ou PID de única entrada, única saída (SISO). Muito embora técnicas de controle avançado possam proporcionar melhorias significativas aos processos, um controlador PID bem projetado e sintonizado é satisfatório para a maioria avassaladora das malhas de controle industrial (CHEN e SEBORG, 2002).

Apesar da importância dos sistemas de controle, pesquisas industriais reportam que 60% dos controladores apresentam desempenho pobre, levando a produtos de baixa qualidade além de perda de produção. Dados ainda mais alarmantes identificam que pelo menos 30% dos controladores aumentam a variabilidade da variável de processo comparado ao uso da operação em manual e que 36% dos processos são operados em malha aberta (DESBOUROGH e MILLER, 2002).

Estas estatísticas desapontadoras justificam a necessidade de avaliar o desempenho das malhas de controle em concordância com uma famosa citação de Peter Drucker sobre os

estratagemas do mundo dos negócios: se você não pode medir você não pode gerenciar. Harris, em 1989, foi o primeiro pesquisador a abordar o monitoramento de malhas fechadas de maneira quantitativa, dando origem ao índice de Harris – que compara a variância da variável de processo controlada à mínima variância realizável. Desde então o campo de pesquisa que hoje é intitulado monitoramento do desempenho de malhas de controle (CPM) se expandiu em ferramentas e procedimentos que incluem isolamento e identificação de falhas e análises de distúrbios e causas-raízes na planta (BAUER, HORCH, *et al.*, 2016).

O mau funcionamento em malhas de controle se reflete no aumento da variabilidade da variável de processo e tem como consequência imediata redução na operabilidade dos equipamentos, aumento dos custos e diminuição da qualidade do produto. As técnicas de CPM têm a intenção, portanto, de responder a algumas das seguintes questões: o controlador está saudável; isto é, está realizando o seu serviço de maneira satisfatória? Em caso negativo, por quê? Qual o potencial de melhoria em relação a um desempenho de referência? Como os dados de processo podem ser úteis nesta tarefa (JELALI, 2004)?

As causas-raízes de mau funcionamento das malhas de controle são muitas. Sintonia inadequada do controlador associada à falta de manutenção é uma delas. Estudos industriais indicam que 80% dos controladores de estrutura PID são mal sintonizados e que 25% deles operam com parâmetros padrão, sugeridos pelos fabricantes (OVERSCHEE e MOOR, 2000). Em muitos outros casos é prática comum sintonizar o controlador uma vez, na fase de comissionamento da automação da planta, e deixá-lo com a sintonia original por muitos anos ou décadas ainda que tenha havido mudanças de ponto operacional ou modificações estruturais na planta (JELALI, 2004).

Além disso, do ponto de vista industrial, existem poucas pessoas responsáveis pela manutenção dos sistemas de automação. Em média, cada engenheiro de controle é responsável por 450 malhas (BAUER, HORCH, *et al.*, 2016) de maneira que tipicamente os controladores são ajustados até serem considerado “bom o suficiente”, mas não ótimos. Estima-se que o tempo de meia-vida de uma malha de controle de bom desempenho seja de seis meses (BIALKOWSKI, 1993). Frequentemente, os controladores são sintonizados de maneira conservativa para garantir estabilidade mesmo em mudanças operacionais em sistemas não-lineares, o que leva a controladores de comportamento lento. Spinner e colaboradores (2014) atribuem o alto número



de malhas de controle de desempenho degradado à falta de mão de obra qualificada e falta de conhecimento sobre o tema combinados com processos cuja natureza varia com o tempo.

Outra fonte relevante de problemas em malha fechada reside nas falhas e defeitos dos componentes do laço como sensores e atuadores, tipicamente associados a atrito. Com o advento dos sensores inteligentes, é possível que as falhas de sensores se tornem cada vez menos relevantes, já que as informações de detecção e diagnóstico podem ser enviadas diretamente às equipes de manutenção (BAUER, HORCH, *et al.*, 2016). Um problema mais sério está associado ao projeto inadequado dos próprios componentes do processo. Em concordância com as afirmações de Ziegler e Nichols (1943), processo e controladores se comportam como uma unidade de maneira que o crédito bem como o demérito do desempenho do sistema deve ser atribuído às duas instâncias.

Uma estrutura inapropriada de controle também é fonte causadora de diminuição do desempenho de malhas controladas (JELALI, 2004). Pareamento inadequado entre variáveis de entrada e saída, negligência com a natureza interativa das variáveis de processo bem como dos controladores, graus de liberdade insuficientes, presença de fortes não-linearidades e ausência de compensador de tempo de atraso são, por exemplo, algumas das questões possíveis que requerem reprojeto das estratégias de controle consideradas.

Bauer e colaboradores (2016) publicaram uma pesquisa industrial realizada por meio de questionários endereçados a engenheiros de controle líderes e experientes em diferentes indústrias. A maioria dos entrevistados trabalha nas indústrias de processos químicos e de óleo e gás, que são redutos tradicionais de CPM. A intenção era capturar a prevalência de técnicas de monitoramento de desempenho de malhas de controle identificando as que são, de fato, aplicáveis em ambiente industrial, bem como os desafios e tópicos de pesquisa em aberto. Muitos resultados interessantes foram gerados. A pesquisa conclui que, em linhas gerais, o número de instalações de pacotes e ferramentas de monitoramento de desempenho de controladores aumentou nos últimos anos.

O campo de CPM maturou ao ponto de haver inúmeros algoritmos e pacotes comerciais oferecidos na forma de produtos e de serviços disponíveis no mercado (JELALI, 2004). Muitos autores como Ogawa (1998), Horch (2000), Jämsä-Jounela e colaboradores (2003) e Paulonis e Cox (2003), para citar alguns, desenvolveram estes sistemas, ferramentas e pacotes. Jelali (2004) explicita as principais características de alguns destes produtos comerciais e cita casos relatados

na literatura de aplicação. Honeywell e DuPont são exemplos de empresas que desenvolveram suas próprias soluções para monitoramento de desempenho em malha fechada. A primeira oferece às indústrias de processo um serviço de internet chamado *LoopScout*<sup>TM</sup> (MILLER, TIMMONS e DESBOROUGH, 1998 apud JELALI, 2004, p.441). A segunda vende um pacote intitulado *Performance Surveyor*<sup>TM</sup> (HOO, PIVOSO, *et al.*, 2003). A Tabela 4 resume as soluções comerciais disponibilizadas por algumas das empresas provedoras de equipamentos para processos.

Os métodos de CPM podem ser agrupados de acordo com três critérios: métodos matemáticos e algoritmos, natureza da tendência no domínio do tempo ou tipo de falha ou mau funcionamento associado. Qin (1998) apresenta uma descrição detalhada de alguns dos métodos usuais.

**Tabela 4 Ferramentas e pacotes de CPM comerciais disponíveis no mercado**

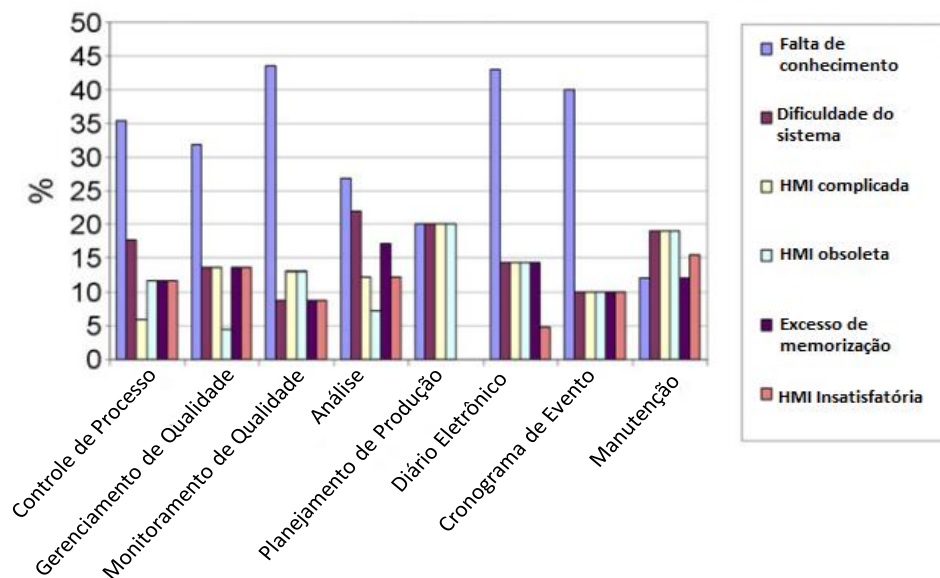
<b>Empresa</b>	<b>Acrônimo</b>
Matrikon	ProcessDoctor
ExpertTune	PlantTriage
ProControl Technology	PCT Loop Optimizer Suite (PCT LOS)
ABB	Optimize Loop Performance Manager (LPM)
Honywell	Loop Scout
Emerson Process Management	EnTech Toolkit, Delta V Inspect
ControlSoft	INTUNE
KCL	KCL-Control-Performance Analysis (KCL-CoPA)
OSIsoft	PI ControlMonitor
AspenTech	Aspen Watch
Control Arts Inc.	Control Monitor
Invensys	Loop Analyst
PAS	ControlWizard
Metso Automation	LoopBrowser
PAPRICAN	LoopMD
DuPont	Performance Surveyor
PETROBRAS	BR-Tuning

Fonte: Jelalli (2004), - modificada

Todavia, independente da natureza do método, certas características desejáveis o tornam mais ou menos atraente. Por exemplo, a habilidade de execução automática sem intervenção manual dos operadores ou engenheiros além de ausência total de interferência na operação normal da planta sugerindo que a aquisição de dados seja feita de maneira não-invasiva. Ainda sobre os dados, é altamente recomendável que a ferramenta de monitoramento faça uso de dados

brutos. Dados comprimidos, suavizados ou quantizados podem levar a resultados não confiáveis. Baixo índice de erros tanto por detecção de falsos alarmes quanto por detecções verdadeiras não realizadas são altamente desejáveis uma vez que a falta de confiança da operação nos resultados poderia fazer com que estes fossem ignorados ou ainda que a ferramenta fosse desligada. Diagnóstico de possíveis causas-raízes das malhas de desempenho degradado acoplado a sugestões de medidas corretivas também são acessórios que tornam a ferramenta mais interessante. A apresentação dos resultados ao usuário de maneira intuitiva e apropriada através da interface homem-máquina é um dos fatores chave para a aceitação do pacote. Dessa forma, uma quantidade de informação exacerbada associada ao aumento de complexidade que requeira maior conhecimento para interpretação dos resultados e tomada de decisão são fortemente desaconselhados (JELALI, 2004).

Pikkusaari-Saikkonen (2004, apud Jämsä-Jounela, 2007, p.218) reportaram uma pesquisa industrial em que cem operadores foram ouvidos. O estudo mostra que apenas 30-40% dos sistemas de suporte a operação instalados são usados de fato. As principais justificativas para o uso infrequente das ferramentas de apoio à operação são mostradas na Figura 4, corroborando o requerimento de que a interface homem-máquina seja amigável e de que a ferramenta seja de simples manuseio.



**Figura 4 Razões para uso infrequente das ferramentas de apoio à operação**

Fonte: Jämsä-Jounela (2007), - modificada

Dentre os métodos matemáticos, os que englobam estatísticas de operação foram considerados os mais úteis na pesquisa realizada por Bauer e colaboradores (2016). Estas estatísticas incluem tempo em modo manual versus operação em automático, número de intervenções do operador, quantidade de mudanças de sintonia além de laços em estado de saturação. Em linhas gerais, quanto mais simples o método maior o número de especialistas inclinados a usá-lo.

As técnicas de análise de tendência não são recentes. Nos anos 90, os engenheiros de controle procuravam pelas mesmas características nos dados que nos dias de hoje: saídas de controle nos seus limites, laços oscilatórios e malhas em operação manual. A Figura 5 mostra gráficos de tendência típicos associados à sua causa provável.

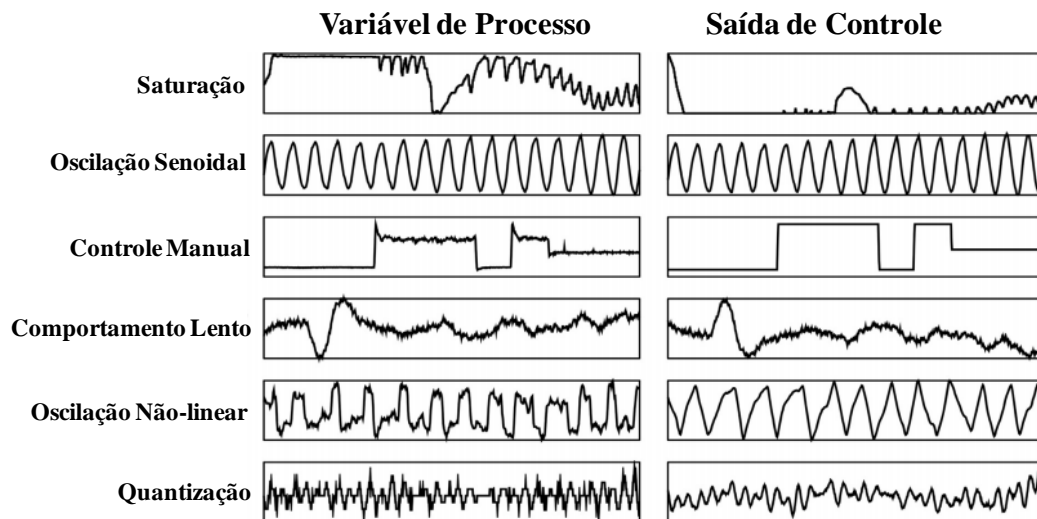


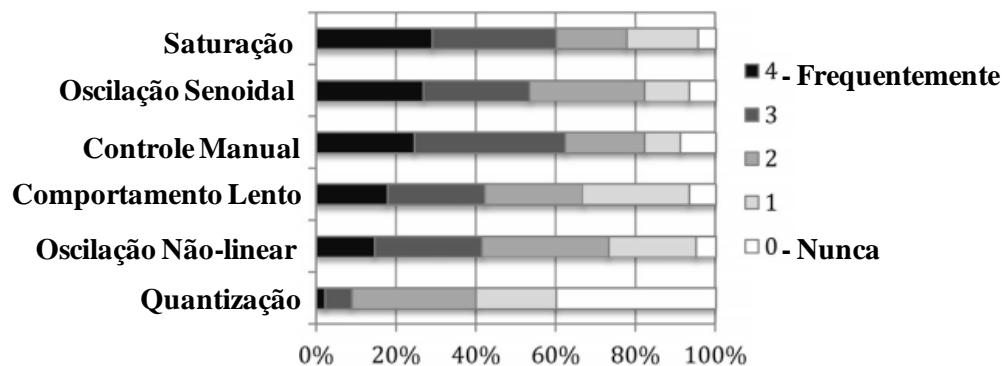
Figura 5 Gráficos de tendência de variável de processo e saída de controle e prováveis causas-raízes

Fonte: Bauer *et al.* (2016) - modificada

Saturação, por exemplo, se manifesta como resultado de um controlador mal sintonizado, atuador mal dimensionado ou pelo fenômeno *wind-up* causado pela ação integral de um controlador de estrutura PID. Já as oscilações senoidais podem ter origem em problemas de sintonia, tipicamente com controladores de ganho excessivamente altos, em falhas do atuador como atrito nas válvulas ou ainda em distúrbios externos de natureza oscilatória (HÄGGLUND, 2005). Malha de controle em operação manual é um sinal claro que o operador não confia no controlador por inúmeros motivos que devem ser investigados, desde problemas de sintonia e malhas com interação até necessidade de treinamento das equipes de operação. Comportamento *sluggish*, ou lento, é tipicamente associado a ajustes de parâmetros de controle muito

conservadores enquanto oscilações não-lineares provavelmente significam falha em um dos componentes da malha como, por exemplo, na válvula de controle. Quantização é um comportamento bastante atípico que provavelmente significa problemas de calibração no sensor tornando a leitura ruidosa.

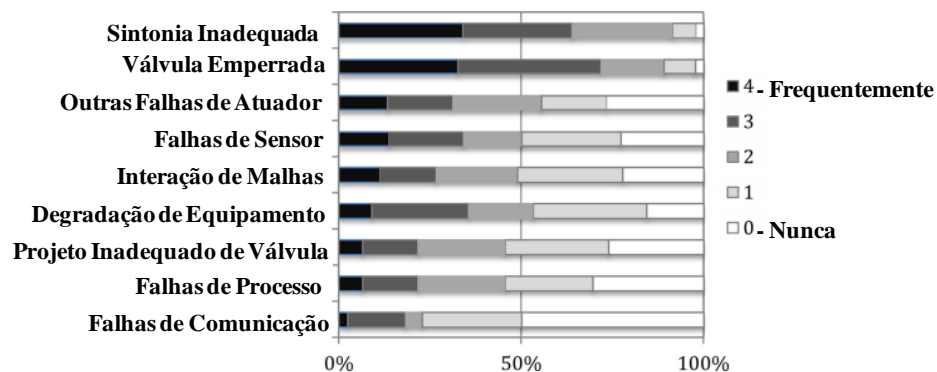
A Figura 6, também extraída de Bauer e colaboradores (2016), mostra as tendências de falha mais frequentes de acordo com os especialistas interpelados. A quantização se revela um fenômeno raro enquanto a saturação é bastante recorrente. A oscilação também é um fenômeno frequente e de grande impacto no consumo de energia e desperdício de matéria prima vem sendo estudada, por exemplo, por Dambros *et al.* (2019), que discutem os desafios da detecção das oscilações em plantas de processo.



**Figura 6** Histograma de frequência de tendências faltosas

Fonte: Bauer *et al.* (2016) - modificada

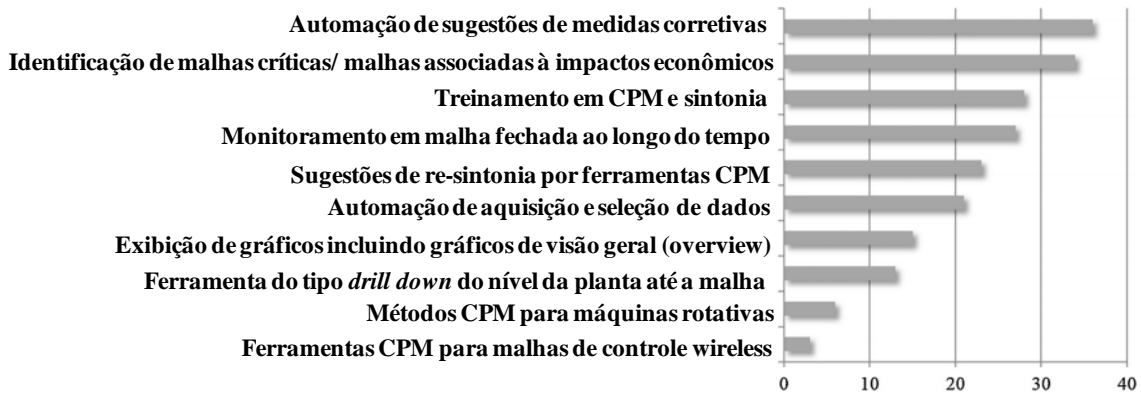
Quanto ao tipo de falha, a pesquisa reporta as mais frequentes, conforme Figura 7. É notável o fato de que embora a ciência de controle de processos tenha evoluído, as mesmas falhas clássicas continuam a ocorrer.



Fonte: Bauer *et al.* (2016) - modificada

**Figura 7** Histograma de frequência de tipos de falha

É indiscutível que ferramentas de monitoramento que detectam e informam às equipes de operação e manutenção sobre malhas de controle de desempenho deteriorado são um fator chave para tornar a operação industrial cada vez mais eficiente. Mas, existe um próximo passo que contempla a tomada de ações remediadoras de maneira automática caracterizando, por exemplo, estratégias de reconfiguração *online* ou ainda sugestões automáticas para a interferência na operação. A pesquisa industrial discutida reporta a demanda dos engenheiros especialistas indagados por ferramentas com esta capacidade, tornando evidente a motivação para este tópico de pesquisa ainda em aberto (Figura 8). Além disso, destaca-se a urgência por tratamento de dados de histórico automatizado bem como avaliação de desempenho no decorrer do tempo em detrimento da tomada de índices em intervalos de tempo discretos.



**Figura 8 Características desejáveis e ainda em desenvolvimento em ferramentas CPM**

Fonte: Bauer *et al.* (2016) - modificada

### ***II.3.4 FTC: Controle Tolerante a Falha***

Novas abordagens para o projeto de controladores incluem sistemas de controle tolerantes a falha (FTC) que possuem a habilidade de acomodar as falhas dos componentes automaticamente enquanto mantêm estabilidade e desempenho aceitável (ZHANG e JIANG, 2008).

De um ponto de vista histórico, um número significativo de pesquisas voltadas para sistemas de controle tolerantes a falhas foi motivado pela indústria de aviação que tinha por objetivo introduzir capacidades auto reparadoras que permitissem aos pilotos aterrissar mesmo na ocasião de situações anormais (CHANDLER, 1984; ETERNO, WEISS, *et al.*, 1985; MONTGOMERY e PRICE, 1976; MONTGOMERY e CAGLAYAN, 1976). Zhang e Jiang (2008) sugeriram que este esforço foi impulsionado por acidentes nos voos comerciais da Delta

(voo 1080 em abril de 1972) e da American Airlines (voo 191 em maio de 1979), ocorridos no final dos anos 70.

Nos anos 90, foram divulgados artigos científicos focados em revisão como, por exemplo, o artigo clássico de Blanke e colaboradores (1997), que apresenta uma visão holística sobre os sistemas FTC e propõe uma aplicação com arquitetura de automação em três camadas: controle, detecção e supervisão. Os autores consideram todo o projeto de controladores FTC desde a concepção de engenharia até a implementação estrutural incluindo exemplos ilustrativos.

Como um marco do reconhecimento da importância da área de FTC destaca-se o *workshop* em Controladores Reestruturáveis promovido pelo centro de pesquisas da NASA situado na Virgínia, EUA, em setembro de 1982, e a série trienal realizada pelo IFAC em Detecção de Falhas *Online* e Supervisão em Indústrias de Processos Químicos iniciada em 1992 (ZHANG e JIANG, 2008).

Recentemente, as questões concernentes aos controladores tolerantes a falha ganharam maiores atenções de comunidades acadêmicas e industriais mais diversificadas, possivelmente em razão das maiores demandas em relação à segurança e confiabilidade que um sistema controlado pode oferecer. As aplicações incluem setores de aviação (ALWI, EDWARDS e TAN, 2009; ALWI e EDWARDS, 2008; JIANG e ZHANG, 2006; KALE e CHIPPERFIELD, 2005), aeroespacial (VARMA e KUMAR, 2010; (PIRMORADI, SASSANI e SILVA, 2009), energia nuclear (MA e JIANG, 2011) e plantas de energia (YE, WANG, *et al.*, 2001; NORA, SAUTER, *et al.*, 2000; PEDERSEN, RITCHER, *et al.*, 2016), para citar algumas.

Embora os sistemas FTC venham sendo exaustivamente pesquisados ao longo dos anos, com mais de 11000 trabalhos científicos publicados nos últimos 15 anos contendo as palavras “*fault tolerant control*” no título (base de dados *ScienceDirect*), eles ainda não são uma tecnologia madura, de metodologia bem estabelecida e terminologia padronizada (ZHANG e JIANG, 2008) e ainda não foram implementados na indústria (YU e JIANG, 2015). Há registro de apenas 14 patentes internacionais depositadas desde 2002 na base USPTO e nenhuma patente brasileira.

O controle tolerante a falhas se distingue em métodos passivos e ativos. Para os primeiros, o efeito de falhas no comportamento do processo é descrito como incerteza de modelagem. A abordagem de natureza passiva independe de ferramentas FDD e não contempla controladores reconfiguráveis, apresentando limitações de habilidade de acomodar falhas (DOCHAIN,

MARQUARDT, *et al.*, 2006). Esta categoria, também denominada sistemas de controle confiáveis ou sistemas com integridade, leva a controladores robustos e, às vezes lentos, com desempenho deteriorado mesmo em circunstâncias normais de operação já que o mesmo controlador é usado em todos os cenários (ZHANG e JIANG, 2008).

Os métodos ativos consistem em realizar a reconfiguração *on-line* com base nas informações fornecidas por um sistema de diagnóstico; reagindo às falhas em uma tentativa de assegurar estabilidade e desempenho aceitáveis (DOCHAIN, MARQUARDT, *et al.*, 2006). Estes sistemas também podem ser identificados como reconfiguráveis, reestruturáveis, auto reparáveis ou ainda autoprojetáveis; e neles, o controlador compensa pelos impactos das falhas ou selecionando uma lei de controle pré-programada ou sintetizando um novo controlador *online* (ZHANG e JIANG, 2008).

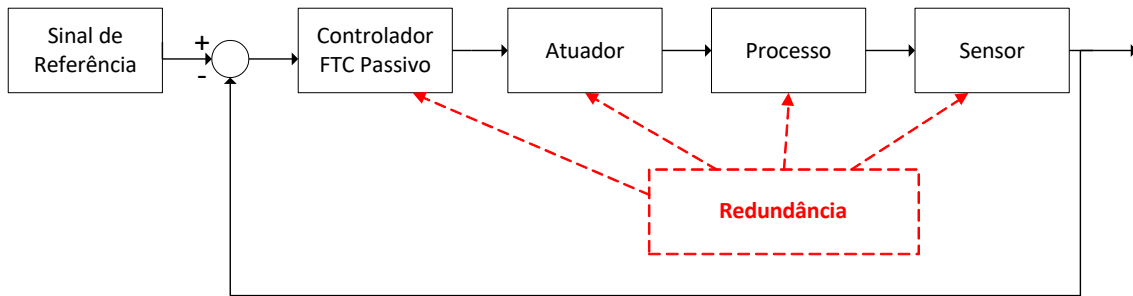
Jiang e Yu (2015) reportam que a comunidade FTC foi inspirada por um conceito proposto por Neumann (1956) que sugeria que sistemas se tornam mais confiáveis por meio da duplicação dos componentes mesmo quando os componentes são menos confiáveis individualmente. Deste modo, independente da natureza da técnica FTC, redundância de sistemas é o que é comum a ambas as abordagens que diferem na maneira como a redundância é utilizada (JIANG e YU, 2012). Aqui, redundância inclui, mas não se limita a sensores e atuadores sobressalentes. Em uma abordagem mais ampla, refere-se também à maneira como o sistema gerencia os recursos restantes para acomodar a condição de falha.

#### **II.3.4.1 PFTC: Controle Tolerante a falha Passivo**

Independentemente do método de projeto, o controlador passivo prioriza robustez em detrimento do desempenho ótimo. Seu *design* busca um controlador fixo no universo de soluções admissíveis para a região de interseção entre cada condição operacional normal; isto é, livre de falha e cada condição operacional em falha.

A Figura 9 ilustra a estrutura genérica de um sistema de controle do tipo PFTC. Porque um único controlador é sempre comissionado independente da presença de falhas, não há necessidade de um esquema FDD, de estratégias de reconfiguração nem de um sistema de gerenciamento de redundâncias.





**Figura 9** Representação esquemática de sistema de controle PFTC

Fonte: Yu e Jiang (2015) - modificada

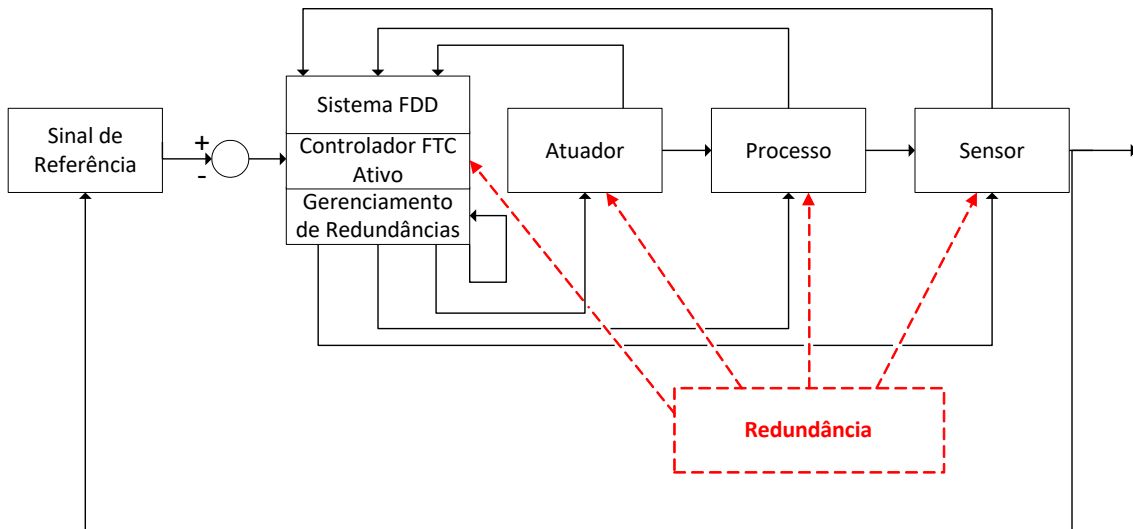
Os trabalhos pioneiros que fazem referência aos sistemas de controle passivos datam dos anos 80 em que questões como confiabilidade na presença de falhas (SILJAK, 1980), integridade de sistemas e as dificuldades inerentes aos sistemas multivariáveis (MACFARLANE, 1980) e estabilidade do ponto de vista matemático foram abordadas (FUJITA e SHIMEURA, 1988; SHIMEURA e FUJITA, 1985). Nos anos 90, os esforços se voltaram para a definição matemática de redundância (OSDER, 1999).

Nos dias de hoje, os problemas envolvendo a diminuição do conservadorismo dos controladores passivos bem como o projeto para sistemas lineares atrasados e sistemas não-lineares têm atraído as atenções da comunidade acadêmica. Alguns desenvolvimentos recentes com diferentes abordagens de projeto que incluem controle *fuzzy*, modo de controle deslizante (*sliding*), métodos baseados na teoria de Lyapunov, entre outros, estão descritos em Yu e Jiang (2015).

#### **II.3.4.2 AFTC: Controle Tolerante a falha Ativo**

Aqui, ativo representa o fato de que ações corretivas são tomadas por um mecanismo de reconfiguração para que o sistema de controle se adapte em resposta às falhas detectadas.

A Figura 10 mostra a arquitetura de um sistema AFTC genérico que consiste em uma ferramenta FDD, um controlador reconfigurável e um mecanismo de reconfiguração gerente. Todas as unidades devem operar em harmonia, sob uma estrutura ativa, balanceada e em tempo real.



**Figura 10** Representação esquemática de sistema de controle AFTC

Fonte: Yu e Jiang (2015) - modificada

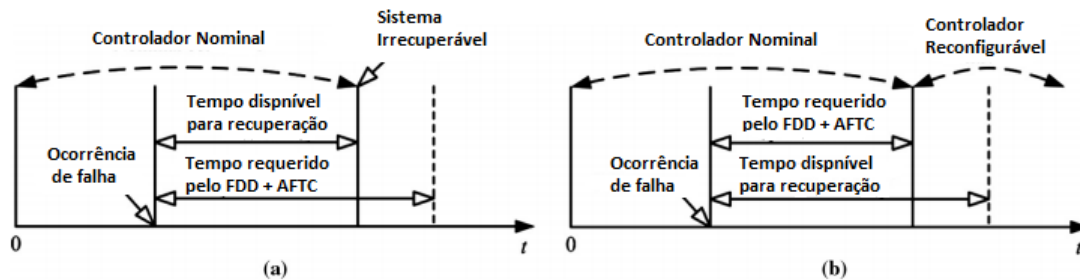
Conforme ilustrado, informações originadas no sistema de monitoramento são transmitidas ao controlador ativo ou reconfigurável por meio da ferramenta FDD. A unidade FDD deve possuir a habilidade de informar o real estado de todos os componentes do sistema da maneira mais acurada e mais rápida possível ao mesmo tempo em que falsos alarmes de sinais ruidosos e espúrios são rejeitados. Embora falsos alarmes e atrasos na detecção tenham potencial para levar todo o sistema à instabilidade, eles são inevitáveis, devendo, portanto, ser levados em consideração nos sistemas FTC.

Depois de detectada, isolada e identificada, a falha é informada ao mecanismo de reconfiguração que atua gerenciando as redundâncias que incluem sensores e atuadores sobressalentes e o controlador reconfigurável.

A transição entre os controladores deve ser suave para prevenir que transientes abruptos levem variáveis críticas para fora dos seus limites físicos (BLANKE, KINNAERT, *et al.*, 2006). Do ponto de vista de segurança, há certas fronteiras rígidas que não devem ser atravessadas em nenhum instante do processo. Por exemplo, se o avião atinge o chão o processo é irreparável (YU e JIANG, 2015). Um transiente desfavorável durante o processo de reconfiguração do controlador pode também levar a danos secundários de componentes saudáveis. Dessa forma, algumas técnicas para lidar com o gerenciamento de transientes vem sendo estudadas como relatado em Guler e colaboradores (2003), que propuseram o uso de sistemas híbridos, em Cieslak e colaboradores (2015), que trabalharam com a estabilidade exponencial global de

Lyapunov e em Bustan e colaboradores (2014), que exploraram estruturas de controle adaptativas.

Depois que uma falha ocorre existe um tempo limitado para que o sistema tolerante a falha reaja e promova ações remediadoras denominado tempo crítico de reação. No entanto, a ferramenta FDD exige um tempo de processamento para realizar o diagnóstico que é tipicamente dependente da natureza da falha (BLANKE, KINNAERT, *et al.*, 2006). Além disso, os mecanismos de reconfiguração também requerem tempo de processamento para gerenciar as redundâncias do sistema e implantar a ação remediadora e, nos casos em que algoritmos de síntese de controladores são usados, estes levam algum tempo para convergir para uma solução. Se o tempo real de reação do sistema AFTC é maior do que de o tempo de reação crítico, o processo pode entrar em um estado instável e irrecuperável conforme mostrado no caso (a) da Figura 11. O caso (b) da Figura 11 mostra a situação que o tempo disponível no sistema é suficiente para recuperação do processo (JIANG e YU, 2012).



**Figura 11 Tempo de recuperação x Tempo disponível no sistema**

Fonte: Yu e Jiang (2015) - modificada

A abordagem AFTC presume que há um espaço de solução admissível em que um controlador capaz de lidar com cada um dos diferentes cenários faltosos existe. Algoritmos de síntese que buscam a solução ótima com base em critérios de desempenho e restrições são tipicamente usados para encontrar este controlador, embora soluções sub-ótimas sejam suficientes (YU e JIANG, 2015).

Zhang e Jiang (2008) classificaram os métodos de projeto AFTC quanto às ferramentas de concepção matemáticas, às abordagens de projeto, aos mecanismos de reconfiguração ou aos tipos de sistemas a que se destinam.

Não obstante, independentemente do método de projeto, os controladores tolerantes a falha de natureza ativa são intrinsecamente ferramentas *online* e em tempo real, que devem satisfazer aos seguintes requerimentos: a reconfiguração deve ser realizada sob restrições de tempo real; o controlador reconfigurável deve ser projetado automaticamente com pouca intervenção humana; os métodos selecionados devem prover uma solução ainda que não seja a solução ótima (ZHANG e JIANG, 2008).

De maneira análoga às questões que envolvem as ferramentas FDD nenhum método sozinho pode ser considerado o melhor de todos e é mais plausível que uma combinação de diferentes métodos seja usada (ZHANG e JIANG, 2008). Yu e Jiang (2012) publicaram uma abordagem híbrida em que controladores passivos e ativos eram usados para aproveitar as vantagens inerentes aos dois métodos. A ideia era ganhar tempo para a concepção de um controlador ativo ótimo enquanto o controlador passivo garantia operação segura ainda que com desempenho degradado.

Em linhas gerais, a implementação de controle tolerante a falha é uma tarefa complexa e interdisciplinar que envolve inúmeras especialidades de engenharia tais como modelagem e identificação, matemática aplicada, estatística aplicada, análise de confiabilidade e de risco, computação, automação, controle, processamento de sinais, instrumentação além de conhecimento de *hardware* e *software*.

Em paralelo, o sucesso do sistema FTC depende de muitos fatores como, por exemplo, a velocidade e a acurácia do sistema de detecção de falhas, a disponibilidade dos atuadores “saudáveis” remanescentes, do gerenciamento de redundância analítica e de *hardware*, da estratégia de controle adotada, do projeto do controlador reconfigurável e da integração de todos os componentes do sistema. Todas estas considerações podem ajudar a justificar a pouca implementação dos sistemas FTC na indústria.

Além disso, por motivos históricos aliados à complexidade do problema, as técnicas de detecção e diagnóstico de falhas – ferramentas indispensáveis para os sistemas FTC ativos - e os sistemas FTC propriamente ditos vêm sendo estudados separadamente. Como consequência imediata, as ferramentas FDD podem não disponibilizar as informações necessárias para que um sistema reconfigurável opere em subseqüência e, por outro lado, os sistemas FTC podem presumir disponibilidade de um sistema de detecção e diagnóstico de falhas perfeito e muito rápido dificultando as aplicações em tempo real (ZHANG e JIANG, 2008). Um estudo

comparativo entre técnicas FTC passivas e ativas publicado em 2012 evidencia, por meio de uma aplicação, a influência da acurácia e do tempo de atraso das informações de falha para os controladores AFTC (JIANG e YU, 2012). Dada a importância da questão, muitas estratégias para lidar com informação imprecisa das ferramentas FDD surgiram nos últimos tempos: inversão dinâmica (JIN, KO e RYOO, 2008), técnicas de controle deslizante (*sliding*) (HAMAYUN, EDWARDS e ALWI, 2014; ALWI e EDWARDS, 2008), função de Lyapunov-Krasovskii (SHEN, JIANG e COCQUEMPOT, 2012), teoria de controle robusto (SERON e DONA, 2014), para citar algumas. Outros trabalhos relatados, como o de Richter *et al.* (2011) e Sencio (2017), fazem uso de métodos de atuador virtual para lidar com informação incompleta das ferramentas FDD e também para contornar problemas de ausência de redundância de instrumentos de atuação, que pode levar a um número de graus de liberdade insuficiente em sistemas controlados.

Zhang e Jiang (2008), com base em seu artigo de revisão, ensaiaram apontar alguns tópicos de pesquisa que ainda devem ser explorados para que os sistemas tolerantes a falha possam alcançar a indústria de processos de maneira mais abrangente. Por exemplo, é irrefutável que redundância é o ingrediente chave dos sistemas de operação segura. Porém redundância apenas de *hardware*; isto é, de equipamentos, é tipicamente custosa e pouco atraente. Uma alternativa é a consideração de redundância analítica em que sinais oriundos de modelos matemáticos são usados e cuja consequência imediata é a diminuição da dependência de equipamentos sobressalentes. Este conceito abre caminho para as pesquisas voltadas para arquitetura e gerenciamento de redundâncias.

Outro aspecto relevante é a natureza dinâmica dos sistemas de controle e do ambiente em tempo real das operações de detecção e diagnóstico de falhas e de reconfiguração de controladores de maneira que a integração entre as ferramentas FDD e FTC ganha destaque. Vale ressaltar, entretanto, que o tempo hábil para acomodação da falha e recuperação do processo é limitado. Nesta linha, é importante considerar o papel das redes de comunicação nos sistemas de controle distribuído em que diferentes *softwares* e equipamentos de diversos fabricantes, além de modos múltiplos de operação, devem coexistir em harmonia e um único sistema.

Além disso, com o rápido avanço nos campos de microeletrônica e tecnologias mecatrônicas, atuadores inteligentes e sensores com propriedades de autodiagnóstico estão cada vez mais disponíveis. Estes componentes *smart* fornecem medidas e atuações mais confiáveis,

além de prover informações muito mais detalhadas, que auxiliam no diagnóstico e até mesmo na predição de falhas nos sistemas e subsistemas (YU e JIANG, 2015). Além disso, a capacidade de processamento *on-board* e as habilidades de comunicação em rede reduzem drasticamente os custos, bem como os problemas de comunicação via cabos. Esta instrumentação inteligente vai impactar de maneira avassaladora na implementação e na estrutura propriamente dita dos sistemas AFTC e nos sistemas de segurança com habilidades inerentes que ainda estão para ser completamente exploradas (BOYES, 2012; ZHANG e JIANG, 2008). Esforços recentes visando incorporar componentes tolerantes a falha em sistemas FTC foram feitos por (PARENTE, VALDMAN, *et al.*, 2016; ISERMANN, 2011; ZUG, DIETRICH e KAISER, 2011; MUENCHHOF, BECK e ISERMANN, 2009).

O objetivo primário dos sistemas FTC reside no aumento da segurança e da confiabilidade do sistema. Por segurança do sistema, entende-se a habilidade de evitar qualquer dano à vida humana, aos equipamentos e ao meio-ambiente. Já confiabilidade está relacionada à capacidade de o sistema realizar determinadas tarefas durante um determinado período sob um determinado conjunto de condições operacionais. Estudos que consideram segurança como objetivo de projeto foram revisados em (YU e JIANG, 2015). Pesquisas que tratam confiabilidade como critério-objetivo para o projeto de sistemas FTC também foram relatados em Guenab, Webber *et al* (2011) e em Li e Zhao (2008). Contudo, medidas quantitativas dos ganhos da implementação como, por exemplo, recuperabilidade ou reconfigurabilidade, capazes de responder a perguntas como “Esta técnica aumenta a segurança e a confiabilidade do sistema?” ainda estão em discussão e podem justificar e motivar os esforços voltados para a implementação industrial de tais sistemas.

### ***II.3.5 Gerenciamento de Alarmes***

Um sistema de alarmes pode ser definido como a coleção de *hardware* e *software* que pode fornecer estados de alarme, comunicando-os aos operadores e gravando as alterações destes estados (SOARES, PINTO e DE SOUZA JR., 2016). Embora eles sejam considerados ativos críticos para as plantas de processo, em particular refinarias de petróleo, petroquímicas e plantas de energia (ROTHENBERG, 2009), seu funcionamento apresenta desempenho pobre, pouco eficiente em circunstâncias de anormalidade (HU, AFZAL, *et al.*, 2015) e muito aquém das

diretrizes industriais (WANG e CHEN, 2014), como, por exemplo, a norma ISA 18.2 (ISA, 2009).

Problemas comuns estão associados aos alarmes ruidosos, que não requerem ação corretiva específica do operador (EEMUA, 2013) e podem ser do tipo intermitente, fugaz, de longa duração e redundante (HU, AFZAL, *et al.*, 2015). De acordo com as entidades reguladoras, EEMUA (2013) e ISA (2009), a taxa de alarmes deveria ser de no máximo seis por hora para cada operador. Não obstante, pesquisas industriais reportaram que, em 1997, cada operador já recebia, em média, uma notificação a cada dois minutos (HSE, 1997). Ainda que um sistema de detecção e diagnóstico de falhas tenha sido instalado para auxiliar a operação, é comum que o operador se sinta sobrecarregado de alarmes, o que pode levar ao reconhecimento dos alarmes sem a checagem adequada da circunstância que o originou.

Outra ocorrência com raízes na crescente complexidade e alta integração dos processos industriais é a avalanche de alarmes, que já é reconhecida como um problema de segurança (WANG, LI e HUANG, 2015), e foi identificada como a causa-raiz de incidentes significativos como, por exemplo, o desastre nuclear da central *Three Mile Island* (SHU, MING, *et al.*, 2016). A avalanche de alarmes acontece quando o número de alarmes notificados surge massivamente na interface homem-máquina e excede a velocidade de capacidade de resposta dos operadores. As normas ISA-18.2 (2009) E EEMUA-191 (2013) sugerem a métrica de dez alarmes em um período de dez minutos como uma regra para identificar as avalanches, que podem ocorrer em operação normal ou em transientes como partidas e paradas; mas são particularmente críticas em circunstâncias de falha. Dados industriais em operação anormal mostram que, após um grande distúrbio na planta, ocorrem cerca de noventa alarmes no primeiro minuto e mais de setenta alarmes nos dez minutos subsequentes (HSE, 1997).

Assim, ao invés de auxiliar a operação, alarmes mal configurados têm potencial para confundir os operadores, distraindo sua atenção dos alarmes reais e aumentando o risco de condições perigosas em processos industriais.

Métodos sistemáticos para configuração e avaliação dos sistemas de alarme vêm atraindo a atenção das comunidades acadêmica e industrial, com a intenção de racionalizá-los para aumentar a sua eficiência, confiabilidade e qualidade. Alguns destes métodos foram desenvolvidos com base nos índices taxa de falsos alarmes, taxa de alarmes não detectados e atraso médio de alarme (CHENG, IZADI e CHEN, 2013a; XU, WANG e CHEN, 2012;

ADNAN, IZADI e CHEN, 2011). Outros métodos de concepção de alarmes são baseados em modelos, como, por exemplo, o trabalho de Figueredo *et. al* (2016), em que o projeto de alarmes foi abordado através de redes neuronais e aplicado a uma plataforma *offshore*.

Outras pesquisas se voltaram para investigação dos alarmes ruidosos, como em Kondaveeti *et al.* (2013), que propuseram um índice de intermitência para prover medidas quantitativas e em Naghoosi *et al.* (2011), que fizeram uso de propriedades estatísticas de variáveis de processo para o projeto de sistemas de alarme.

Os alarmes intermitentes, em particular, representam 50% das notificações de alarme (WANG e CHEN, 2014) e têm sido largamente abordados. Uma variedade de técnicas foi desenvolvida para remover alarmes intermitentes com base em dados de alarme. Wang e Chen (2014), por exemplo, propuseram uma abordagem de detecção e remoção *online*.

O gerenciamento de alarmes correlacionados é outra questão relevante. Yang e colaboradores (2013) detectaram alarmes correlacionados com base na distribuição de correlação de atrasos. Além disso, coeficientes de similaridade binária foram usados para quantificar as correlações em Kondaveeti *et al.* (2012). Soares, Pinto e de Souza Jr. (2016) aplicaram protocolos de gerenciamento de alarmes baseados nas ferramentas estatísticas de análise de correlação, análise de componentes principais e análise de *cluster* em dados reais de processo de uma planta de gás natural.

Outras aplicações em casos reais recentemente relatadas incluem o trabalho de Hu e colaboradores (2015), que fez uso de ferramentas de visualização para gerenciamento de alarmes em uma planta de extração de areia betuminosa. Já Cai e colaboradores (2015) usaram de tecnologia de processamento de informação em cenários anormais para lidar com a avalanche de alarmes tanto em um experimento simulado quanto em uma planta petroquímica real.

A detecção de incidentes industriais baseada em dados de alarme é outro tópico de grande interesse das pesquisas atuais. Com a finalidade de encontrar causas-raízes ou ainda assistir na prevenção de avalanches de alarme, algumas abordagens como o algoritmo dinâmico DTW e o algoritmo Smith-Waterman modificado foram utilizados para comparar avalanches de alarme e encontrar padrões de sequência (AHMED, IZADI, *et al.*, 2013 ;CHENG, IZADI e CHEN, 2013b). Zhu e colaboradores (2014) sugeriram uma estratégia de controle de alarmes para os transientes dos processos químicos. Em seu trabalho, para gerar alarmes úteis em situações de falha, um sistema de detecção e diagnóstico de falhas baseado em DTW é proposto.



Além disso, algumas ferramentas de visualização como, por exemplo, o gráfico de alarmes de alta densidade (HDAP) e o mapa de similaridade de alarmes por cores (ASCM) também surgiram como suporte na detecção de alarmes ruidosos (KONDAVEETI, IZADI, *et al.*, 2012).

Em 2014, Adhitya e colaboradores focaram a atenção no aspecto humano durante o desenvolvimento de vários métodos e ferramentas para uma melhor gestão de alarmes e puderam prever o tempo de ocorrência de alarmes críticos antes de serem realmente detectados.

Com o aumento da competitividade em escala global, a responsabilidade de cada operador da planta de processo tem aumentado ao longo dos anos. Assim, com a implementação dos sistemas de monitoramento, tarefas como interpretação dos resultados de diagnóstico, visualização, arquivamento de dados e geração de relatórios devem ser facilitadas para que a operação faça bom uso das ferramentas auxiliares (ZHANG, JEINSCH, *et al.*, 2011). Gerenciamento de alarmes, portanto, é parte essencial de um sistema de gerenciamento de situações anormais.

Não obstante, ainda há pouca discussão em torno de técnicas que compreendam os transientes operacionais como partidas e paradas e mudanças de pontos de operação, sendo Marcellos (2013) um dos exemplos.

Hu e colaboradores (2015) reportaram sugestões de trabalhos futuros que englobem análises de avalanche com base em conhecimento de processo associados aos dados de histórico. Desta maneira, informações de ações da operação e experiência de engenharia podem ser incorporadas aos mecanismos de interpretação de padrões com a finalidade de conseguir prever falhas em estágios iniciais da avalanche. Outro campo de pesquisa ainda não muito desenvolvido envolve a aplicação de ferramentas de gerenciamento de alarmes em tempo real, em que técnicas de detecção e remoção de alarmes ruidosos possam ser implementadas *online*. Esta automação reduziria significativamente o fardo da operação.

Zhu *et al.* (2014) sugeriram, ainda, que embora técnicas voltadas para processos multimodos tenham sido discutidas, estas ainda não são capazes de indicar as falhas fundamentais em um sistema básico de alarmes. Ademais, o estudo de sistemas de alarme dinâmicos, com a habilidade de reconhecer diferentes pontos operacionais também é uma área de interesse ainda não maturada.

## II.4 Caldeiras

### II.4.1 Visão Geral

As caldeiras são comumente utilizadas na indústria para suprir as turbinas geradoras de energia elétrica com vapor ou para fornecer calor para os processos químicos. Não é exagerado afirmar que toda indústria de processo químico tem vapor como principal fonte de aquecimento em diversas de suas operações. Em uma planta industrial utiliza-se vapor em diversos níveis de pressão com vários objetivos (CAMPOS e TEIXEIRA, 2006):

- Diminuir a pressão parcial de hidrocarbonetos e facilitar a separação residual ou *stripper*;
- Aumentar a velocidade e a turbulência em um escoamento;
- Fornecer energia aos processos unitários e reatores químicos por meio de trocadores de calor.

Quanto à participação do vapor na geração de energia elétrica vale ressaltar que embora a matriz energética brasileira seja majoritariamente hidráulica, 3004 termoeletricas estão em operação e são acionadas por picos de demanda como artifício para evitar apagões em épocas de estiagem (ANEEL, 2019). Além das centrais de serviço público existe ainda a geração dos autoprodutores (APE). A autoprodução clássica agrega as mais diversas instalações industriais que produzem energia para consumo próprio, a exemplo dos setores de Papel e Celulose, Siderurgia, Açúcar e Álcool, Química, entre outros, além do Setor Energético. Neste último, destacam-se os segmentos de exploração, refino e produção de petróleo (EPE, 2018). Em 2017, a geração de energia elétrica no Brasil em centrais de serviço público e autoprodutores atingiu 588,0 TWh. Deste montante, a autoprodução participou com 16,5% do total produzido, atingindo um total de 96,6 TWh (EPE, 2018). A geração elétrica a partir de não renováveis representou 20,8% do total nacional.

O parque gerador térmico brasileiro até o final da década de 1990, principalmente das regiões Sul, Sudeste e Centro-Oeste, era constituído por usinas térmicas a vapor, também chamadas de convencionais, movidas a óleo combustível e a carvão, enquanto o Norte do país possuía termelétricas movidas basicamente a óleo diesel. As que continuam em atividade têm em média 40 anos de operação (ANEEL, 2008).

Embora tenham surgido projetos de geração utilizando turbinas a gás movidas a gás natural, estas ofereciam desvantagens como custo da energia gerada relativamente elevado, devido ao combustível muitas vezes importado e atrelado ao dólar, e eficiência mais baixa em relação à tecnologia de ciclo combinado. Sendo assim, no ano 2000, o governo brasileiro escolheu as usinas termelétricas de ciclo combinado (geração a vapor e a gás) como principal tecnologia para geração elétrica a ser implantada no país de forma emergencial. Assim, a termelétricidade passou a ter importância estratégica, ganhando participação mais significativa na matriz energética brasileira (EPE, 2016).

Tradicionalmente, os equipamentos de grande porte utilizados na geração de energia, incluindo turbinas e caldeiras, são importados. Entretanto, um estudo divulgado pelo PROMINP (2006) que avalia a competitividade da indústria brasileira de bens e serviços para a indústria de petróleo e gás sugere que os setores de caldeiraria experimentaram uma fase de ressurgimento, ainda que existam lacunas tecnológicas particularmente associadas aos problemas de escalonamento que direcionam os investimentos para a importação.

#### ***II.4.2 Emissão de Poluentes***

O Programa de Controle da Poluição do Ar (PRONAR) existe desde 1989, mas foi em 2007, por meio da resolução número 382, que o Conselho Nacional do Meio Ambiente (CONAMA) regularizou oficialmente os limites máximos de poluentes atmosféricos para fontes fixas, incluindo os processos de geração de calor a partir da combustão externa de óleo combustível. Com base nesta resolução, que admite óleo combustível como qualquer derivado líquido obtido de material fóssil, os limites máximos estão estabelecidos de acordo com a capacidade nominal do equipamento emissor e estão descritos na Tabela 5.

Para sistemas de potência menor que 10MW, fica como encargo do órgão ambiental licenciador aceitar a avaliação periódica de monóxido de carbono, que não deve exceder o valor de 80mg/Nm<sup>3</sup> (CONAMA, 2007).

Agências de proteção ao meio-ambiente internacionais também se preocuparam em limitar as emissões atmosféricas de fontes estacionárias sendo a *Environmental Protection Agency* (EPA), dos Estados Unidos, a mais expressiva. EPA optou por estabelecer diferentes camadas de restrição de acordo com o tempo de vida dos equipamentos analisados, sendo os limites mais rigorosos para os equipamentos mais novos mostrados na Tabela 6.

**Tabela 5 Limites máximos para a emissão de poluentes em caldeiras à óleo**

Potência Térmica Nominal (MW)	Material Particulado (mg/Nm <sup>3*</sup> )	Óxidos de Nitrogênio (mg/Nm <sup>3*</sup> )	Óxidos de Enxofre (mg/Nm <sup>3*</sup> )	Monóxido de Carbono (mg/Nm <sup>3*</sup> )
Menor que 10	300	1600	2700	80
Entre 10 e 70	250	1000	2700	80
Maior que 70	100	1000	1800	80

\* Em base seca e 3% de excesso de oxigênio

Fonte: CONAMA (2007)

**Tabela 6 Limites máximos para a emissão de poluentes em caldeiras à óleo**

Potência Térmica Nominal	Ano	Monóxido de Carbono *	Hidrocarbonetos Não-Metano NMHC*	Óxidos de Nitrogênio*	Material Particulado*
kW < 8 (hp < 11)	2008	8.0 (6.0)	-	-	0.4 (0.3)
8 ≤ kW < 19 (11 ≤ hp < 25)	2008	6.6 (4.9)	-	-	0.4 (0.3)
19 ≤ kW < 37 (25 ≤ hp < 50)	2008	5.5 (4.1)	-	-	0.3 (0.22)
	2013	5.5 (4.1)	-	-	0.03 (0.022)
37 ≤ kW < 56 (50 ≤ hp < 75)	2008	5.0 (3.7)	-	-	0.3 (0.22)
	2013	5.0 (3.7)	-	-	0.03 (0.022)
56 ≤ kW < 130 (75 ≤ hp < 175)	2012-2014	5.0 (3.7)	0.19 (0.14)	0.40 (0.30)	0.02 (0.015)
130 ≤ kW ≤ 560 (175 ≤ hp ≤ 750)	2011-2014	3.5 (2.6)	0.19 (0.14)	0.40 (0.30)	0.02 (0.015)

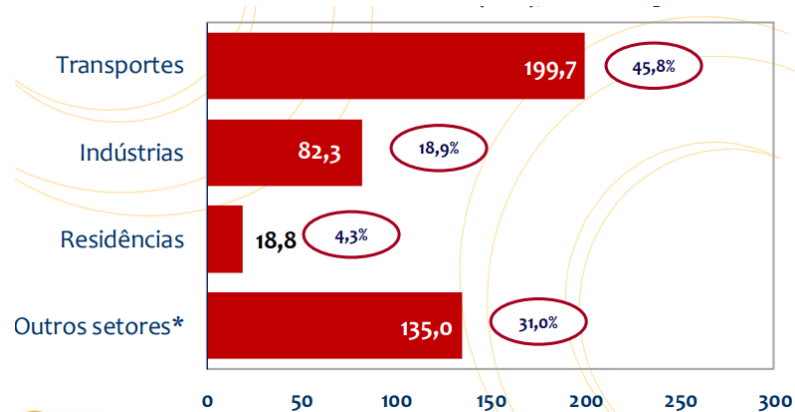
\* g/kWh (g/bhp-hr)

Fonte: EPA (2013)

Nos dias de hoje, segundo dados divulgados pelo Ministério das Minas e Energia (MME), estudados pela Empresa de Pesquisa Energética (EPE) e divulgados no relatório síntese do Balanço Energético Anual (2018) a intensidade de emissão de carbono na economia brasileira se reflete em 0,15kg de CO<sub>2</sub>/US\$ppp<sup>2</sup>. Em 2017, o total de emissões antrópicas associadas à matriz energética brasileira atingiu 435,8 milhões de toneladas de dióxido de carbono equivalente

<sup>2</sup> ppp – paridade do poder de compra

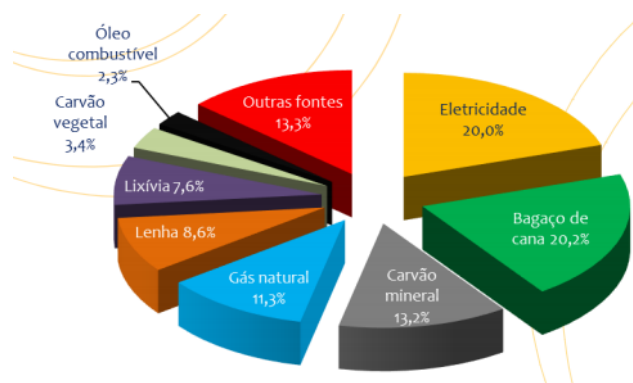
(MtCO<sub>2</sub>-eq), sendo o setor industrial o segundo maior gerador de resíduos. A Figura 12 ilustra o total de emissões de dióxido de carbono por setor em 2017.



**Figura 12** Total de emissões de CO<sub>2</sub> antropicas por setor em Mt no ano de 2017

Fonte: EPE (2018)

Embora muitos estudos tenham surgido com a intenção de mensurar a quantidade de poluentes emitidos e a eficiência térmica do uso de combustíveis alternativos frente aos combustíveis de origem fóssil como Freitas (2013), Amaral *et al.* (2016) e Jeng *et al.* (2016), para mencionar alguns, os combustíveis fósseis ainda são a grande fonte energética das indústrias brasileiras. A Figura 13 mostra que os óleos combustíveis não renováveis são a origem energética de 2,3% da energia consumida no país.



**Figura 13** Consumo de energia por fonte nas indústrias

Fonte: EPE (2018)

Um estudo sobre o potencial de geração de energia divulgado pelo Ministério do Meio Ambiente (MMA) exemplifica a preocupação do Brasil com o gerenciamento de emissões

atmosféricas e, em particular, de dióxido de carbono. O Brasil se destaca no cenário internacional como um importante ator ligado ao Mecanismo de Desenvolvimento Limpo (MDL), um dos instrumentos do Protocolo de Quioto criados para ajudar os países desenvolvidos a alcançar suas metas de redução de emissões de carbono e incentivar financeiramente os países em desenvolvimento. Em termos do potencial de reduções de emissões associado aos projetos de MDL, em 2010, o Brasil ocupava a terceira posição, sendo responsável pela redução de 375.889.172 tCO<sub>2</sub>eq, o que corresponde a 6% do total mundial para o primeiro período de obtenção de créditos (PNUD, 2010).

#### ***II.4.3 Monitoramento, Diagnóstico e Alarmes em Processos de Combustão***

O monitoramento dos gases residuais dos processos de combustão é uma demanda mundial irrefutável. Não obstante, a confiabilidade dos analisadores de gases reais é questionável em razão dos fenômenos de deposição e desgaste que levam a valores medidos imprecisos e até mesmo a ausência de dados. Estes analisadores requerem recalibração frequente sugerindo a necessidade de redundância para possibilitar o monitoramento contínuo e em tempo real destas variáveis críticas mesmo nos períodos de manutenção. Redundância de *hardware* é raramente a solução adotada devido aos altos custos de investimento e manutenção. Portanto, redundância analítica baseada em modelos surge como alternativa atraente desde que os valores estimados sejam confiáveis (KORPELA, BJÖRKQVIST, *et al.*, 2014).

Um dos primeiros usos de sucesso dos sensores virtuais foi em uma usina a gás de 100 MW nos Estados Unidos, na década de 90. Nesta ocasião, o sensor foi aplicado no monitoramento contínuo de monóxido de carbono, óxido de nitrogênio e oxigênio nos gases de combustão em detrimento do sensor real justamente visando redução dos custos de manutenção e aumento da confiabilidade (COHN, 2006).

Em 2001, Chong e colaboradores usaram uma rede neuronal para prever a emissão de gases residuais (oxigênio, monóxido de carbono e óxidos de nitrogênio) da queima de carvão em uma caldeira. Tronci *et al.* (2002) reportaram uma aplicação em escala piloto, em que um analisador virtual monitorava a emissão de gases residuais (oxigênio, dióxido de carbono e óxidos de nitrogênio) de um forno em uma planta de geração de energia.

Petersson e Holmberg (2005) propuseram o desenvolvimento de um analisador híbrido para a detecção de oxigênio, dióxido de carbono e hidrocarbonetos em caldeiras de pequeno

porte. Valdman (2011) desenvolveu um analisador de gases virtual para uma caldeira de porte semi-industrial capaz de prever a concentração de seis componentes diferentes (oxigênio, monóxido de carbono, dióxido de carbono, óxido e dióxido de nitrogênio e dióxido de enxofre).

Liukkonen *et al.* (2012) e Zhou *et al.* (2012) usaram redes neurais para prever a concentração de óxidos nitrogenados e de oxigênio em plantas industriais, sendo o primeiro para aplicação com queima de madeira e o segundo, carvão natural. Em 2013, Liukkonen *et al.* publicaram um estudo sobre o monitoramento via sensor virtual de materiais particulados em processos de combustão de biomassa.

Uma publicação recente de Wanga *et al.* (2016) enfatizou o uso dos sensores virtuais preditivos das variáveis temperatura e composição dos gases de exaustão como uma ferramenta de avaliação da eficiência de um forno de reaquecimento. Neste trabalho, a eficiência de combustão de uma planta de laminação de aço é otimizada por ações de controle que levam em conta os valores previstos de oxigênio na chaminé.

Nikula, Ruusunen e Leiviskä (2016) apresentam uma técnica de monitoramento da eficiência real de uma caldeira baseada em dados de histórico. Na abordagem apresentada, o desempenho da caldeira é monitorado usando um gráfico de controle de processo estatístico através de um *framework* para a diferença entre as eficiências esperadas e reais.

Quando se trata de detecção e diagnóstico de falhas em processos de combustão, algumas das publicações recentes se concentram em equipamentos não-estacionários, como veículos automotivos a gasolina e a diesel. Chen e Randall (2016), por exemplo, usaram redes neurais artificiais treinadas por meio de dados de falhas simuladas para detectar e identificar problemas no motor de combustão por meio de sinais de vibração.

Outros estudos focam em diagnóstico de sistemas de tecnologia de aquecimento, ventilação e condicionamento de ar residenciais (HVAC), que incluem caldeiras convencionais ou caldeiras de condensação. Yu, Woradechjumroen e Yu (2014) apresentaram revisão de diversos destes estudos motivados, principalmente, pelo desperdício de energia. Recentemente, Baldi *et al.* (2017) desenvolveram um conjunto de ferramentas para detecção e diagnóstico de caldeiras de condensação baseadas em uma abordagem dinâmica. As ferramentas incluíram um estimador de parâmetros bimodal (com condensação e sem condensação) para detectar desvios da eficiência da caldeira em relação aos valores nominais de operação; um sensor virtual para

predizer a vazão mássica de água de alimentação ao equipamento e filtros para detectar falhas do atuador e do sensor.

Um caso de aplicação de um método quantitativo FDD em uma planta de energia movida a carvão foi demonstrado em Fu *et al.* (2016). O trabalho usou análise matemática de exergia para localizar e quantificar uma falha induzida na turbina a vapor. Em outra aplicação de diagnóstico em plantas geradoras de energia, técnicas que usam redes neurais artificiais para detecção precoce de vazamento em caldeiras de leito fluidizado foram utilizadas (ROSTEK, MORYTKO e JANKOWSKA, 2015).

Em 2012, Xing-wei (2012) reportou um esforço para detectar e diagnosticar falhas em sopradores de ar, equipamentos rotativos amplamente usados em processos de combustão e, por consequência, muito comum em plantas de energia térmica. A ferramenta de diagnóstico apresentada se baseia em um problema de classificação com solução em algoritmos de técnicas de aprendizado supervisionado para reconhecimento de padrões.

Koiwanit *et al* (2016) propuseram um sistema especialista para monitoramento e diagnóstico de altos níveis de emissão de amônia em processos de pós-combustão ligados a captura de CO<sub>2</sub>. O sistema inclui recomendações de ações remediadoras cada vez que o diagnóstico do processo indica uma situação anormal. Embora o sistema seja automatizado, ele se baseia na entrada manual de dados de concentração de gases residuais.

A interface homem-máquina e, em particular, os sistemas de alarmes, podem ser identificados como o gargalo para a melhoria do desempenho da operação das plantas de processo, especialmente em situações emergenciais (LIU, NODA e NISHITANI, 2010). Não obstante, poucos estudos da literatura se dedicam ao estudo de caso de alarmes em processos de combustão.

Técnicas de modelagem e simulação do comportamento humano vêm sendo aplicadas com a intenção de auxiliar nas etapas de avaliação e projeto dos sistemas de alarmes, particularmente quando sujeitos a operação anormal. Liu, Noda e Nishitani (2010) e Jin, Yamashita e Nishitani (2004) desenvolveram modelos quantitativos para este fim e testaram seus modelos postulados em casos de estudo de plantas simuladas de caldeiras.

Em 2016, Wang e Chen (2016) publicaram um estudo sobre alarmes ruidosos de longa duração em processos industriais. Neste trabalho, os autores usaram dados históricos de operação de uma termoelétrica para identificar as principais causas de origem desta classe de alarmes e



propuseram técnicas sistemáticas para lidar com eles com base em um mecanismo dinâmico de geração de alarmes.

#### ***II.4.4 Filosofia de Controle***

##### **II.4.4.1 Práticas Usuais**

De um modo geral, as caldeiras geram vapor de pressão moderada a alta que será utilizado, por exemplo, para acionar turbinas de compressores, bombas, geradores de eletricidade entre outros equipamentos complexos que necessitam de operação estável (CAMPOS e TEIXEIRA, 2006). Portanto, a pressão de vapor é uma variável que deve ser mantida em uma faixa de variação estreita para evitar o comprometimento dos equipamentos subsequentes. Outra possibilidade é operar fornecendo vazão de vapor controlada, prática usual no caso de um coletor ou *header* comum a várias caldeiras. (CAMPOS e TEIXEIRA, 2006).

As caldeiras costumam ter os dois controles configurados no sistema digital de automação com uma chave seletora que permite ao operador escolher o modo mais adequado para diferentes condições de operação. Em ambos os casos, o controlador selecionado atua, tipicamente, manipulando a vazão de combustível ou, para estratégias mais complexas como o controle por limite cruzado, as vazões de combustível e de ar de combustão injetadas nos queimadores (CAMPOS e TEIXEIRA, 2006).

Além dos controles de produção de vapor da caldeira, outro controle fundamental para garantir estabilidade na operação é aquele que mantém o nível de água do tubulão dentro dos limites desejados variando a vazão da água de alimentação e atuando na válvula de controle de nível (BEGA, 1998).

Estas malhas procuram eliminar a interação existente entre o sistema de controle de nível e o de combustão. Pulsações na vazão da água de alimentação podem se refletir em perturbações na pressão de vapor, ocasionando variações na taxa de aquecimento, sem que tenham ocorrido, de fato, variações de demanda de vapor. Por outro lado, variações na taxa de aquecimento provocam os fenômenos de expansão (*swell*) e contração (*shrink*), dificultadores do controle de nível (BEGA, 1998).

O fenômeno de expansão volumétrica é caracterizado pela falsa elevação de nível da caldeira em que o aumento súbito da demanda de vapor reduz a pressão do tambor. Durante este transiente, a água saturada pode aumentar o seu volume em até 30% (HOUTZ, 2015). Neste

ínterim, ocorre a expansão das bolhas de vapor na fase líquida e, com isso, o nível de água no tambor aumenta embora a massa de água no tambor não tenha se alterado de fato.

No sentido oposto, isto é, quando a carga é diminuída, a concentração de bolhas tende a reduzir, neste ponto haverá um aumento da densidade do meio e conseqüentemente, redução de volume. Este é o fenômeno de contração (NIKIEL, 2009).

Tanto a expansão volumétrica quanto a contração resultam em ações equivocadas de controle ao provocar resposta do controlador contrária à que seria adequada. Estes fenômenos são tipicamente contornados com uma estratégia de controle a três elementos, em que além do próprio nível do tubulão as variáveis vazão de vapor produzido e vazão de água de alimentação são consideradas (NIKIEL, 2009).

Nas caldeiras de pequeno porte o controle de nível é tipicamente auto operado. Nas caldeiras de grande porte, o controle a três elementos permite minimizar os efeitos do *swell*. No entanto, os efeitos combinados no transiente de partida são muito drásticos, fazendo com que, de modo geral, as unidades realizem a partida em malha aberta (OLIVEIRA, 2017).

Em 2012, Patan e Korbicz (2012) usaram um controlador do tipo MPC equipado com um bloco de diagnóstico de falhas para o controle de nível em um estudo de caso de uma caldeira. Este bloco funciona por meio do cálculo de resíduos de um conjunto de modelos parciais projetados na forma de redes neuronais recursivas. Além disso, cada vez que uma falha de sensor de nível é isolada, o sistema estima o tamanho da falha e, com base nesta estimativa, o controlador é alimentado com um valor predito da variável de processo medida e a falha é acomodada.

#### **II.4.4.2 Controle de Combustão**

O controle da poluição atmosférica e o controle do rendimento térmico das caldeiras e fornalhas são normalmente realizados pelos operadores de forma independente, fazendo com que a relação entre excesso de ar, rendimento térmico e emissão de poluentes seja mal compreendida. O excesso de ar influencia tanto a eficiência térmica quanto o nível de emissões atmosféricas refletido nas concentrações dos gases de chaminé  $CO_x$ ,  $SO_x$  e  $NO_x$ . (PINHEIRO e VALLE, 1995).

Para assegurar a combustão completa, é necessário excesso de ar para garantir a existência de oxigênio em quantidade suficiente até o final da chama, que se sobreponha as deficiências do queimador.

A literatura é divergente a respeito do excesso de ar a ser utilizado (SALUM, 2011). Bega (1998) afirma que o excesso de ar adequado a cada instalação particular deve ser determinado a partir de testes experimentais já que o valor ótimo é função do tipo de combustível, tipo de queimador, características de preparação do combustível, tipo de câmara de combustão, porcentagem de carga máxima, da malha de controle utilizada, entre outros fatores. Pinheiro e Valle (1995) sugeriram que valores usuais de excesso de ar para óleo combustível variam de 5 a 15%.

O excesso de ar é fator determinante para a eficiência e a segurança da operação ao influenciar o volume, a temperatura e a pressão, além da natureza e concentração dos produtos. Insuficiência de vazão de ar poderia acarretar combustão incompleta levando à formação de monóxido de carbono, fuligem e fumaça, além de favorecer o acúmulo de combustível não queimado causando riscos de explosões. Grandes excessos de ar diminuem a temperatura da chama e aumentam as perdas de calor devido à entalpia dos gases efluentes, reduzindo a eficiência térmica da caldeira e até mesmo apagando a chama do queimador (PARENTE, VALDMAN, *et al.*, 2013). O valor ótimo do percentual de excesso de ar, conforme pode ser observado na Figura 14, é aquele onde as duas influências estão em equilíbrio.

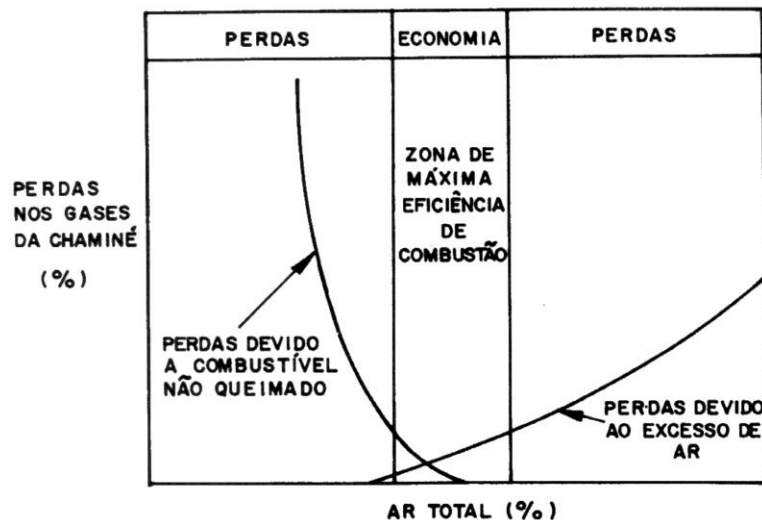


Figura 14 Curva de Perdas da Combustão

O controle da combustão por meio da relação ar/combustível é normalmente realizado por meio da manipulação da vazão de ar uma vez que a energia gerada pela caldeira refletida nos valores de vazão de vapor é função da quantidade de combustível introduzido (PINHEIRO e VALLE, 1995). Tradicionalmente, a vazão de ar é regulada em malha aberta de maneira que uma relação ar/combustível seja estabelecida e fixada. Em muitos casos, esta relação é insatisfatória (PINHEIRO e VALLE, 1995).

As maiores dificuldades encontradas para o controle da combustão real residem no fato de que o coeficiente de excesso de ar afeta a eficiência da caldeira e os níveis de emissões de poluentes de modos diferentes e antagônicos. As Figuras 15 e 16 representam valores típicos de eficiência e níveis de emissão em função do excesso de ar. As curvas dependem do equipamento considerado e as faixas de operação podem variar bastante, entretanto, seu aspecto geral é comum aos processos de combustão.

Em linhas gerais, a zona de combustão referente à faixa A é inaceitável, devido aos elevados teores de CO nos produtos da combustão. A zona de combustão correspondente à faixa B é considerada ideal devido às baixas emissões de CO e NO<sub>x</sub> e a alta eficiência da combustão. Entretanto, para operar na faixa B é comumente necessário ajustar os queimadores ou modificar radicalmente o processo de combustão. Na maioria das instalações de combustão, a faixa B só é realizável fazendo uso de artifícios como combustão em vários estágios ou recirculação de gases (PINHEIRO e VALLE, 1995). Na zona de combustão C, os níveis de emissão de NO<sub>x</sub> são elevados e a operação só se torna viável com o tratamento dos gases efluentes (PINHEIRO e VALLE, 1995). Na zona de combustão D, embora os níveis de emissão de CO e NO<sub>x</sub> encontrem-se dentro dos limites o rendimento térmico é baixo.

Para otimizar o excesso de ar é necessário um controle de combustão mais acurado, que pode ser obtido por meio da implementação de uma malha de controle *feedback*, em que a variável controlada seja um dos produtos de combustão (PINHEIRO e VALLE, 1995). Além disso, em face da importância do excesso de ar, é recomendável a utilização de analisadores em linha, uma vez que estes instrumentos possibilitam a monitoração contínua dos gases de combustão (BEGA, 1998).

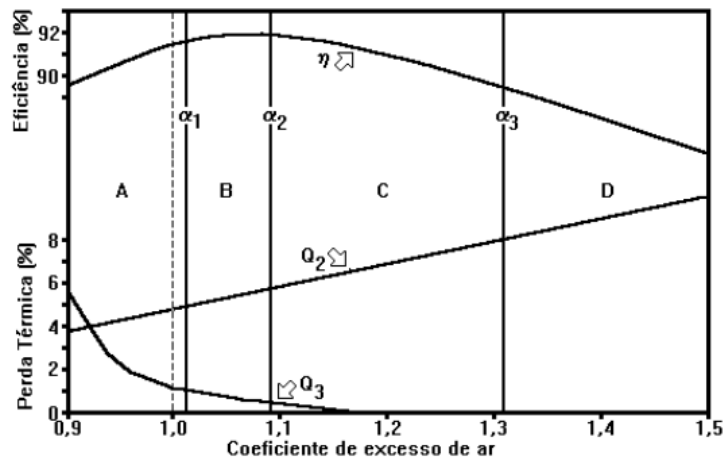


Figura 15 Rendimento Térmico Típico de Caldeiras

Fonte: Pinheiro e Valle, 1995.

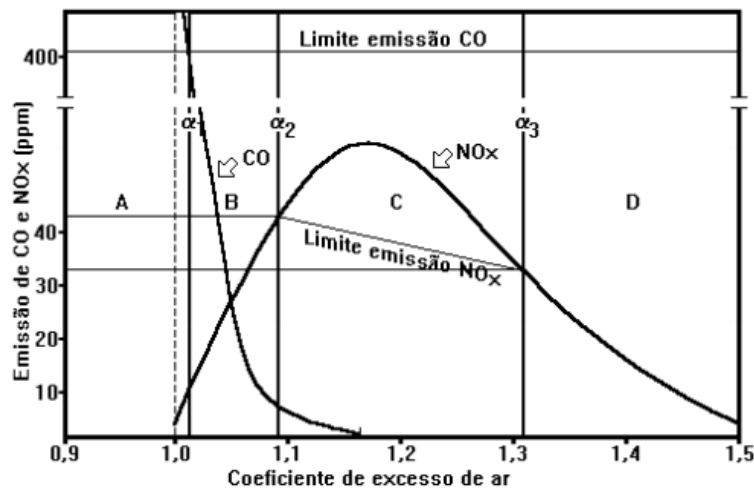


Figura 16 Emissão de Poluentes na Combustão

Fonte: Pinheiro e Valle, 1995.

O teor de oxigênio residual é a variável tradicionalmente utilizada na monitoração e controle de combustão, pois o excesso de ar é função direta do percentual de oxigênio existente. Entretanto, dióxido de carbono também pode ser utilizado com esta finalidade. Sua desvantagem é que o *set-point* deve ser ajustado para cada combustível usado (BEGA, 1998). A Figura 17 ilustra o comportamento típico dos produtos da combustão em relação à quantidade de ar fornecida.

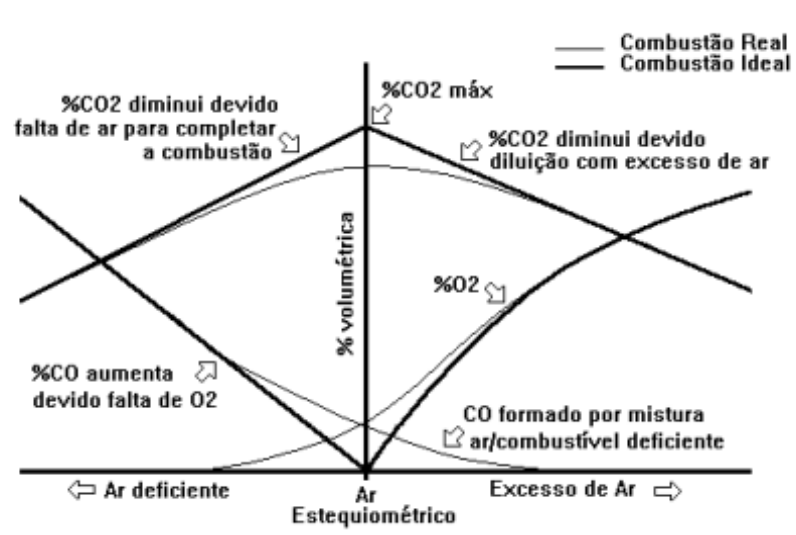


Figura 17 Produtos da Combustão

Fonte: Pinheiro e Valle, 1995.

**PARTE II - APLICAÇÕES EM PROCESSOS DE  
COMBUSTÃO SEMI-INDUSTRIAIS**

## Capítulo III – Caldeira Semi-industrial de Produção de Vapor

---

Este capítulo descreve a central de utilidades do Laboratório de Engenharia Química da Universidade Federal do Rio de Janeiro. Estão incluídas as características de instrumentação disponíveis e de automação relevantes para o trabalho, que incluem a tecnologia de instrumentação disponível, as malhas de controle implementadas e as ferramentas de detecção e diagnóstico existentes.

### III.1 Caldeira de Produção de Vapor

A caldeira flamotubular vertical de porte semi-industrial, modelo CV-VDM-500 que faz parte da planta piloto da Central de Utilidades do Laboratório de Engenharia Química da UFRJ (LADEQ/EQ/UFRJ) está apresentada na Figura 18.



Figura 18 Central de utilidades do Laboratório de Engenharia Química da UFRJ



A planta é composta por quatro tanques de armazenamento de combustível, um tanque de mistura de combustível e alimentação à caldeira, um tanque de armazenamento de água e uma linha de vapor saturado. O combustível oriundo do tanque de alimentação é bombeado até o queimador, onde é atomizado e misturado a uma corrente de ar, que é bombeada através de um soprador. A partida da queima é feita por um centelhador e a caldeira possui vazão de diesel combustível fixa. Para manipular a vazão de ar foi instalado um inversor de frequência, que permite alterações na quantidade de excesso de ar fornecido à queima (PARENTE, VALDMAN, *et al.*, 2013).

O tanque de armazenamento de água, com capacidade igual a 500 litros, fornece água de reposição para a caldeira de acordo com a demanda de vapor por meio de uma bomba. A água retida dentro da caldeira é aquecida até o estado de vapor saturado. O vapor produzido pode ser enviado a uma coluna de destilação e a um trocador de calor que compõem a Central de Operações da planta piloto ou descartado para a atmosfera através de um silenciador e um tanque de retenção de condensado. A Tabela 7 mostra as especificações técnicas da caldeira geradora de vapor.

**Tabela 7 Especificações Técnicas da Caldeira**

Modelo	Flamotubular: CV-VDM-500
Capacidade de produção de vapor (com água a 20°C)	500kg/h
Potência térmica nominal	0,3MW
Estado do vapor	Saturado
Pressão de trabalho máxima	8,00 kgf/cm <sup>2</sup>
Combustível principal	Diesel
Vazão média de combustível	~19kg/h
Tiragem	Forçada por insuflação
Superfície de aquecimento	22,80 m <sup>2</sup>
Câmara de combustão	Número de passes: 01 Número de tubos (Ø: 1 ½ "e L: 1,94m): 94
Volume médio de água em operação normal	0,6m <sup>3</sup>
Vaporização específica por metro quadrado de superfície de aquecimento	25kgv/h
Características da chaminé	Seção da passagem: circular Dimensão: 300mm de diâmetro Área média: 900cm <sup>2</sup>

O apêndice A deste documento contém o diagrama P&I da planta, que mostra o fluxograma do processo e a instrumentação instalada.

### III.2 Instrumentação Instalada e Automação Disponível

O sistema de automação é de tecnologia híbrida, baseado em equipamentos analógicos (4 a 20 mA) e redes digitais *Fieldbus*. A automação disponível tem dispositivos digitais para medição de variáveis de processo contínuas, um controlador lógico-programável (CLP) para converter variáveis discretas e sinal analógico em protocolo *Fieldbus*, um controlador dedicado que opera a lógica de intertravamento e uma estação SCADA para registrar valores de engenharia e estados de diagnóstico.

A Figura 19 apresenta a arquitetura de automação implementada, ampliada e adaptada na planta piloto em trabalhos anteriores (VALDMAN 2010, VALDMAN 2013), incluindo as principais instâncias envolvidas na instrumentação da caldeira.

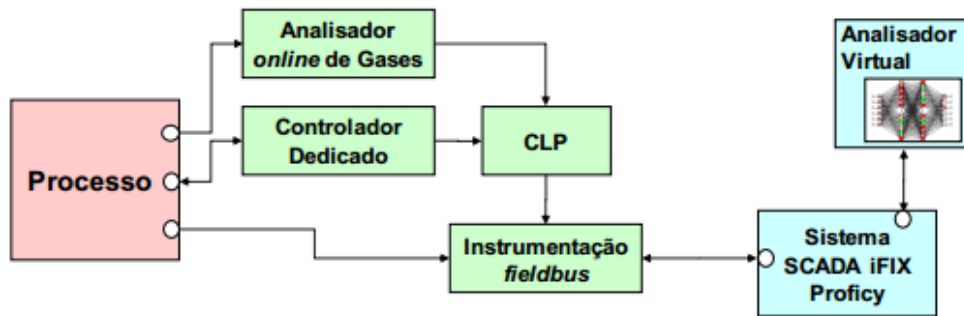


Figura 19 Arquitetura do sistema de automação da planta piloto

Fonte: Valdman., 2013

#### III.2.1 Controlador Dedicado

Acoplado à caldeira geradora de vapor existe um controlador dedicado, responsável pela lógica de intertravamento de segurança, que pode interromper a alimentação de combustível e desligar a caldeira.

Este intertravamento é disparado em condições de operação específicas: quando a pressão excede o limite de segurança, quando a chama do queimador apaga ou quando ocorre falta de água na caldeira. Estas informações são transmitidas ao controlador por meio de variáveis discretas disponíveis.

O controlador dedicado é do tipo liga-desliga e, em paralelo ao diagnóstico de falhas concernentes ao controle da queima, realiza controle do nível de água da caldeira enviando sinais para a bomba de alimentação de água em função das chaves de nível alto e baixo.

O controlador dedicado, fabricado pela Landis & Staefa, foi fornecido e instalado junto com os equipamentos físicos da caldeira e não permite interferência externa em seu funcionamento.

### ***III.2.2 Controlador Lógico Programável***

O controlador lógico-programável (CLP) é o equipamento responsável pela conversão das variáveis discretas ao protocolo de instrumentação padrão *Fieldbus*. Possui um módulo de fonte de alimentação, um módulo de CPU, um módulo com variáveis discretas de entrada e um módulo de conversão *Fieldbus*. Funciona em paralelo ao controlador dedicado e permite que as variáveis discretas sejam monitoradas *online*.

### ***III.2.3 Transmissores Fieldbus***

A instrumentação digital possibilita o acompanhamento das variáveis contínuas do processo. As informações mais detalhadas fornecidas pelos instrumentos de tecnologia digital podem ser usadas para detectar falhas no sistema.

Existem diversos transmissores *Fieldbus* de temperatura, nível, pressão e vazão instalados na planta. Além dos medidores, a planta está equipada com um conversor F/I e com um posicionador *Fieldbus*.

### ***III.2.4 Analisador de Gases Online***

A planta piloto conta com um analisador de gases comercial de tecnologia analógica da marca TESTO 350M/XL, um instrumento para medições *in situ*, que é composto por seis células de medição de concentração de gases, um sensor infravermelho, uma sonda e uma bomba interna

de sucção. A sonda deste equipamento associada à bomba de sucção interna foi instalada na linha de saída dos gases residuais. O analisador de gases é mostrado na Figura 20.



**Figura 20** Analisador de gases *online*

A Tabela 8 apresenta as características de cada uma das variáveis que podem ser disponibilizadas pelo analisador.

**Tabela 8** Características das Variáveis Disponibilizadas pelo Analisador *In situ*

Canal	Gás Analisado	Sensor	Faixas de Medida	Precisão
1	Oxigênio (O <sub>2</sub> )	Célula Eletroquímica	0,0 a 25,0%v/v	0,2%v/v
2	Monóxido de Carbono (CO)	Célula Eletroquímica	0 a 10000 ppm	até 99ppm: 10ppm; de 100 a 2000ppm: 0,5% do valor lido
3	Dióxido de Carbono (CO <sub>2</sub> )	Célula de Infravermelho	0,0 a 50,0%v/v	0,3% + 1% do valor lido
4	Monóxido de Nitrogênio (NO)	Célula Eletroquímica	0 a 3000 ppm	até 99ppm: 5%
5	Dióxido de Nitrogênio (NO <sub>2</sub> )	Célula Eletroquímica	0 a 500 ppm	até 99ppm: 5ppm
6	Dióxido de Enxofre (SO <sub>2</sub> )	Célula Eletroquímica	0 a 5000 ppm	até99ppm: 5ppm

O analisador possui uma rotina cíclica de análise caracterizada por quatro parâmetros temporais apresentados na Tabela 9.

**Tabela 9** Tempos Característicos do Ciclo de Análises dos Gases Residuais

Parâmetro	Magnitude (s)	Descrição
Tempo de Análise	300	Período ao longo do qual as análises são realizadas
Tempo de Rinsagem	300	Período ao longo do qual as análises são interrompidas para limpeza das câmaras dos sensores
Tempo de Amostragem	2	Intervalo de tempo máximo entre análises consecutivas
Intervalo para Médias	60	Intervalo de tempo considerado para o cálculo de médias aritméticas das análises instantâneas

O ciclo de funcionamento do equipamento é de 10 minutos totais. O analisador realiza automaticamente um procedimento de limpeza para manutenção das células eletroquímicas, que são elementos sensores. Este procedimento engloba etapas de rinsagem e purga e dura 5 minutos. Durante o período de limpeza, as medições são interrompidas. Portanto, são disponibilizadas medições efetivas durante os 5 minutos restantes.

A dinâmica de análise exigida pelo analisador tornava estas variáveis de medição de concentração dos gases residuais difíceis de serem utilizadas em tempo real em um sistema de controle. Entretanto, a arquitetura de automação da planta foi ampliada (VALDMAN, 2013), tornando viável a utilização do analisador integrado ao sistema de modo *online* ao disponibilizar suas análises em tempo real à rede *Fieldbus* e ao sistema SCADA. O sinal elétrico (mA) das variáveis medidas pelo analisador de gases é convertido ao protocolo *Fieldbus* via controlador lógico programável (CLP). Adotou-se uma lógica *ladder* no CLP para contornar as discontinuidades oriundas do tempo de purga. A lógica permite o armazenamento do último valor válido lido pelo sensor até que a operação do analisador volte ao normal.

Vale ressaltar, entretanto, que para as medidas concernentes à concentração de CO<sub>2</sub>, o elemento sensor é infravermelho e dispensa, portanto, a necessidade de purga para limpeza bem como as suas inconveniências. Desta forma, as medições referentes ao CO<sub>2</sub> são ininterruptas.

A Tabela 10 apresenta as variáveis configuradas no sistema para receber o sinal do analisador *in situ* a cada segundo.

**Tabela 10 Variáveis do Analisador *In Situ* Disponibilizadas no SCADA**

<b>TAG</b>	<b>Descrição</b>	<b>Unidades de Engenharia</b>	<b>Zero</b>	<b>Span</b>
XIT1041O2	Concentração de O2 nos gases residuais	%v/v O2	0	15
XIT1041CO2	Concentração de CO2 nos gases residuais	%v/v CO2	11	17

### III.2.5 Analisador de Gases Virtual

A ferramenta virtual para predição da composição dos gases residuais resultantes da combustão consiste em um modelo de inferência heurístico, baseado em redes neuronais.

A rede é do tipo *multilayer perceptron* (MLP) com 6 neurônios na camada de entrada; 10 neurônios na camada escondida e função de ativação de tangente hiperbólica; e 6 neurônios na camada de saída e função de ativação identidade.

O sensor virtual, representado na Figura 21, que está incorporada na camada de automação da planta piloto, foi desenvolvido em trabalhos anteriores (VALDMAN,2010; VALDMAN, VALDMAN, *et al.*,2011) .

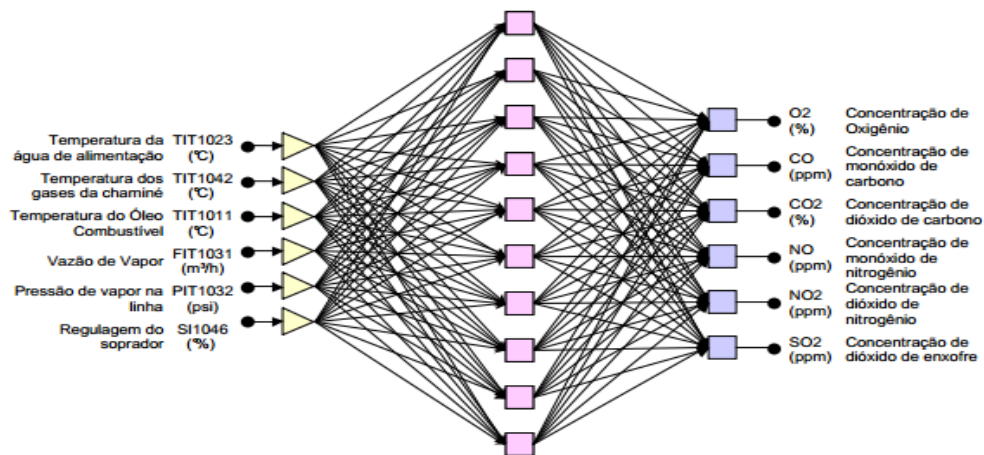


Figura 21 Representação esquemática do sensor virtual

Fonte: Valdman., 2010

### III.2.6 Elementos Atuadores

No processo estudado, um inversor de frequência da marca WEG, modelo cfw 08, modifica a velocidade de rotação do motor do soprador de ar a fim de regular a vazão de ar para combustão. Este conjunto é identificado como o elemento final de controle de uma malha de controle de gases residuais implementada previamente (PARENTE, 2014). O inversor de frequência foi integrado à rede FB utilizada em trabalhos anteriores (PARENTE, VALDMAN, *et al.*, 2013). Assim, pode ser acionado remotamente através de um sinal *Fieldbus* que é convertido ao sinal analógico padrão de corrente (4 a 20mA) através de um conversor *Fieldbus/Corrente* (F/I) instalado na caldeira. Este conversor, modelo FI-302 da Smar, possui capacidade de

processamento e pode executar algumas funções inteligentes, como, por exemplo, malhas de controle de realimentação (PARENTE, VALDMAN, *et al.*, 2013).

A planta também conta com um posicionador *Fieldbus* para válvulas de controle linear, do tipo “falha fechada” ou “ar para abrir”, como elemento final de controle de uma malha de demanda de vapor. O posicionador digital funciona com um algoritmo de controle PI interno, em que o valor percentual de abertura de válvula solicitado (pelo operador, em operação manual; ou pelo controlador, em operação automática) é o set-point da malha de controle e o valor percentual real de abertura é a variável de processo retornada.

### ***III.2.7 Sistema Supervisório SCADA***

O sistema de supervisão e aquisição de dados (SCADA) usado na planta piloto é o Ifix Proficy®, da empresa GE Intelligent Plataforms, dimensionado para monitorar simultaneamente variáveis de processo, que podem ser contínuas, discretas ou calculadas. Ele é o extremo final de uma cadeia de informações, constituindo uma ferramenta de gerenciamento na operação online via microcomputadores.

A estação SCADA foi configurada para monitorar e armazenar em histórico todas as variáveis de processo disponibilizadas pela instrumentação instalada, com tempo de amostragem de dados igual a um segundo. A interface gráfica, que permite o acompanhamento do processo em tempo real, a consulta a dados de histórico e a interferência na operação foi desenvolvida em trabalhos anteriores (VALDMAN, 2010; PARENTE, 2014) e modificada para incluir as ferramentas propostas neste trabalho.

As variáveis de processo contínuas disponíveis na instrumentação *Fieldbus* da caldeira são mostradas na Tabela 11. As variáveis discretas ou de estado são armazenadas com os valores zero ou um. O valor um indica equipamento ligado ou em alarme enquanto o valor zero indica equipamento desligado ou operação normal. A Tabela 12 apresenta as variáveis de processo discretas disponibilizadas pelo controlador lógico programável da caldeira.

A Figura 22 apresenta a tela principal de operação da caldeira, que contém o sinóptico da planta e as variáveis de processo relevantes. As variáveis contínuas são mostradas por meio de caixas de diálogo enquanto a visualização das variáveis discretas se dá através de símbolos gráficos com cores padronizadas: vermelho para o *status* um e verde para o *status* zero. Recursos

de telas para apresentação de gráficos de tendência, telas de alarmes para diagnóstico de falhas e telas específicas para emissão de relatórios através de planilhas eletrônicas também estão disponíveis no sistema supervisor.

**Tabela 11 Variáveis Contínuas Disponibilizadas na Instrumentação *Fieldbus***

<b>TAG</b>	<b>Descrição</b>	<b>Unidades de Engenharia</b>	<b>Zero</b>	<b>Span</b>
LIT1013	Nível do tanque de alimentação de combustível	L	37,50	353,00
LIT1021	Nível do tanque de alimentação de água	%	0,00	100,00
LIT1026	Nível de água da caldeira	%	0,00	100,00
FIT1024	Vazão de água de alimentação	m <sup>3</sup> /h	0,00	5,20
FIT1031	Vazão de vapor	kg/h	0,00	500,00
PIT1032	Pressão na linha de vapor	psi	0,00	100,00
TIT1011	Temperatura do combustível	°C	0,00	100,00
TIT1023	Temperatura da água de alimentação	°C	0,00	100,00
TIT1042	Temperatura dos gases residuais	°C	0,00	400,00
TMFIT1031	Totalização do vapor produzido	kg/h	0,00	999999,00
SI1046	Vazão de ar fornecida à caldeira	%	0,00	100,00

**Tabela 12 Variáveis Discretas Disponibilizadas no Sistema SCADA da Caldeira**

<b>TAG</b>	<b>Descrição</b>
BB1011	Bomba de alimentação do combustível aos tanques de estocagem
BB1012	Bomba de alimentação do combustível ao tanque de armazenamento principal
BB1013	Bomba de alimentação do combustível ao tanque de mistura
MT1014	Motor de agitação do tanque de mistura
PSH1026	Alarme – pressostato de alta
LSLL1026	Alarme – nível baixo de água
BB1015	Bomba de alimentação de combustível a caldeira
BB1021	Bomba de alimentação de água a caldeira
MT1016	Soprador
XSL1012	Alarme – falha de chama
XSH1011	Queimador ligado



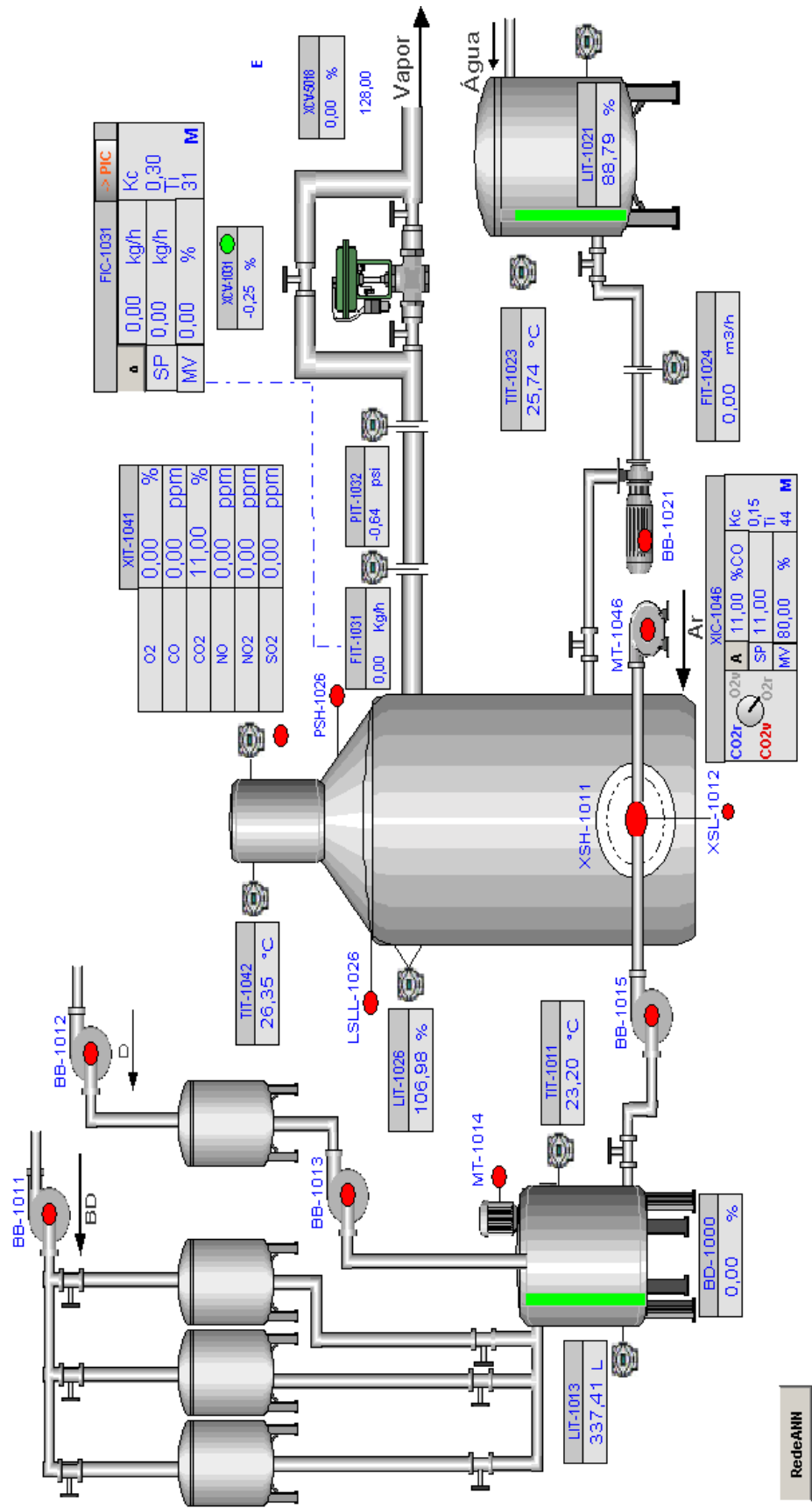


Figura 22 Sinóptico de operação da caldeira

RedeANN

### ***III.2.8 Malhas de Controle Contínuas***

A caldeira conta com uma malha de controle contínuo com modo de operação do tipo proporcional e integral sob estratégia *feedback* para regular a emissão de gases residuais implementada e sintonizada na rede digital *Fieldbus* em estudos anteriores (PARENTE, 2014). O controle *feedback* seletor permite, então, ajustar as variáveis concentração residual de oxigênio e gás carbônico na chaminé da caldeira ao manipular o percentual de vazão de ar para combustão.

De maneira análoga à malha de emissão de gases residuais, existe uma malha de controle de demanda de vapor previamente implementada e sintonizada na rede digital *Fieldbus* (PARENTE, 2014). Esta malha de controle de realimentação de estrutura PI pode controlar seletivamente pressão e vazão na linha de saída de vapor da caldeira, em que a variável manipulada é o percentual de abertura da válvula de vapor. Destaca-se a presença do conversor FB/P associado ao posicionador *Fieldbus*, elemento final de controle, que converte o sinal digital de saída de controle recebido do bloco seletor a sinal de pressão de ar.

## Capítulo IV– Sensor Virtual para Monitoramento de Emissões de Gases: Aplicação em SCADA Industrial

---

Este capítulo resume o artigo “Soft Sensor for Exhaust Gas Emissions Monitoring: Application in Industrial SCADA”, publicado nos anais de Congresso Brasileiro de Instrumentação, Sistemas e Automação.

### IV.1 Resumo

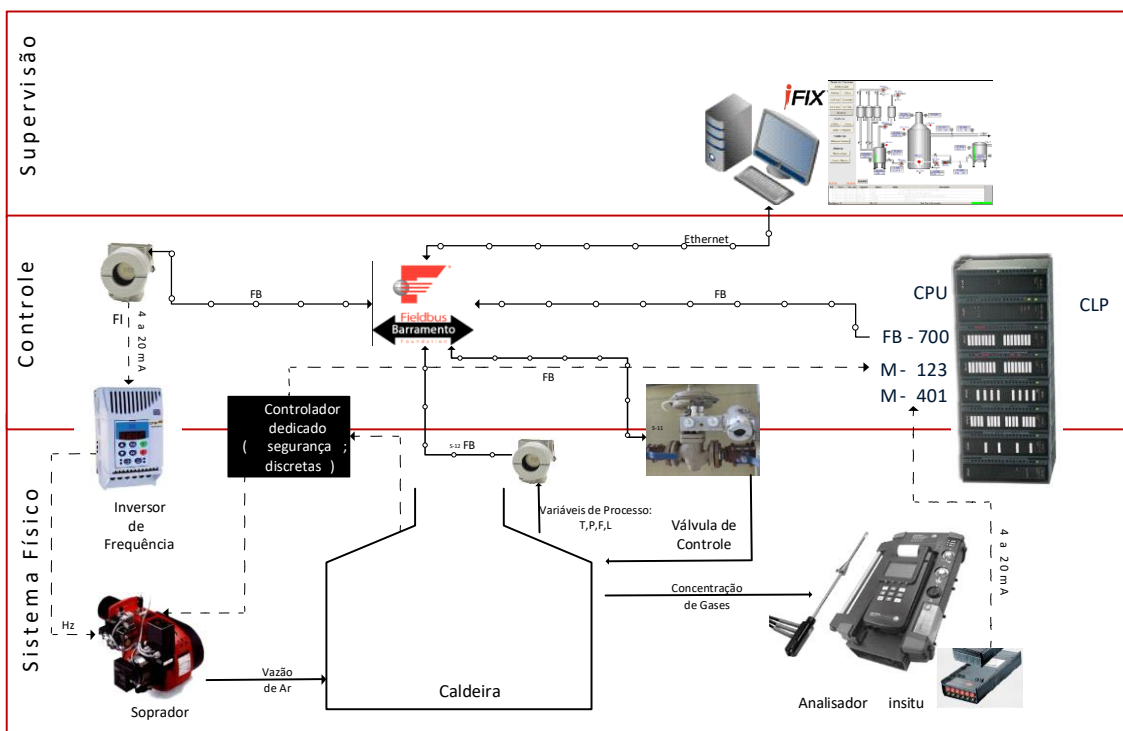
Este trabalho apresenta a implementação *online* de um sensor virtual em uma plataforma industrial SCADA. O sensor virtual é baseado em dados e em técnicas de aprendizagem de máquina. Ele foi testado e validado em uma planta piloto de tecnologia de automação híbrida e redes industriais, que incluem instrumentos de protocolo de comunicação analógico, discreto e digital *Fieldbus*. O sensor por inferência pode ser usado para monitorar os níveis de emissões de CO<sub>2</sub> para que a operação seja consistente e confiável em relação à eficiência de produção e segurança ambiental. A solução de automação foi desenvolvida em uma interface homem-máquina (IHM) amigável e genérica. Desta forma, os operadores podem criar, atualizar e usar sensores virtuais locais baseados em redes neuronais artificiais (RNA) do tipo *multilayer perceptron* (MLP). Embora a estrutura tenha sido implementada em um processo de combustão específico, ela pode ser generalizada para diferentes sensores virtuais em processos distintos.

### IV.2 Metodologia Sumarizada

Os testes experimentais foram realizados em uma caldeira de porte semi-industrial, que faz parte das instalações do Laboratório de Engenharia Química (LADEQ) da UFRJ. A caldeira produz vapor saturado para alimentar a central de operações do LADEQ e opera com vazão constante de diesel combustível e vazão regulável de ar de combustão.

O sistema de automação mescla diferentes tecnologias: tem equipamentos baseados no protocolo de comunicação digital *Fieldbus* para monitorar as variáveis de processo contínuas, um controlador lógico CLP para converter variáveis discretas e sinais analógicos em protocolo *Fieldbus*, um controlador dedicado para intertravamento de segurança e uma estação SCADA para registrar valores de engenharia e estados de diagnóstico. As concentrações dos gases

exaustos pela chaminé da caldeira são monitoradas por um analisador de gases comercial. O analisador possui uma rotina de análise caracterizada por um procedimento de limpeza cíclico, com cinco minutos de purga e rinsagem para cada cinco minutos de medições efetivas. A Figura 23 apresenta a arquitetura de automação da planta piloto. Na camada física estão os equipamentos de processo e o instrumentos de medição e atuação. Na camada de controle estão as malhas de intertravamento de segurança do CLP e as malhas de controle contínuo. A estação SCADA incorporada com os recursos de *machine learning* está na camada de supervisão. Todas as informações do processo fluem para uma rede industrial digital dedicada do tipo LAN.

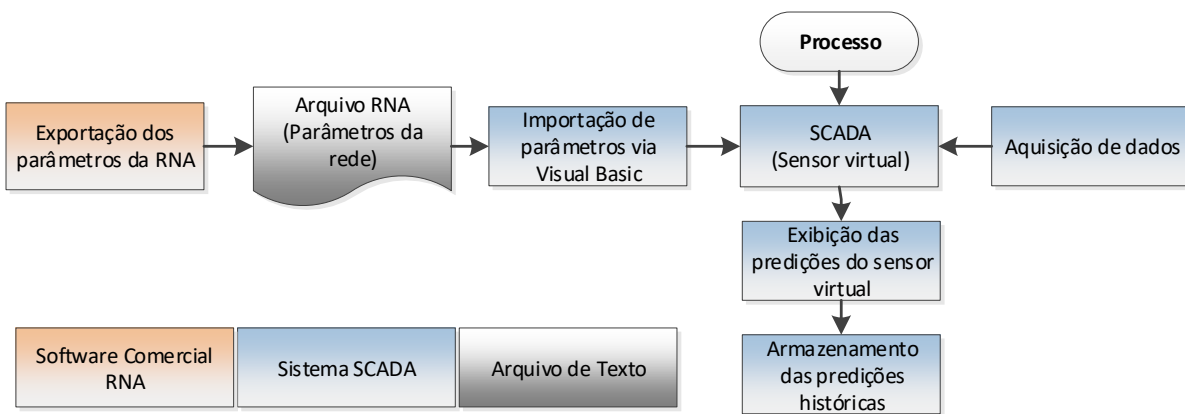


**Figura 23** Arquitetura de automação da planta piloto.

Para superar as questões inerentes ao funcionamento do sensor físico, foi desenvolvido um sensor virtual baseado em rede neuronal artificial (RNA) do tipo MLP por meio do *software* comercial Statistica®. As entradas do modelo de inferência foram definidas com base em análise estatística de sensibilidade global e incluem a temperatura do gás de chaminé, as temperaturas de alimentação de água e combustível, a pressão do vapor produzido e o *set-point* da vazão de ar de combustão.

A implementação do sensor virtual para aplicações em tempo real no sinóptico de operação da estação SCADA consistiu no desenvolvimento de uma interface de configuração, que permite que os operadores criem e atualizem sensores virtuais, e uma interface de monitoramento. A interface de configuração foi desenvolvida usando recursos VBA e deve ser preenchida com os parâmetros de inicialização do modelo. A interface de monitoramento exibe os valores preditos no sinóptico da operação em tempo real e também informações de diagnóstico, aumentando a confiabilidade do sensor virtual.

A Figura 24 resume o funcionamento do sensor virtual *online*. Primeiro, o modelo de inferência deve ser desenvolvido por meio de um pacote comercial (neste trabalho, o STATISTICA®). Em seguida ele tem suas características exportadas na forma de um arquivo txt para o sistema SCADA, que as associa aos parâmetros de inicialização configurados e ativa os cálculos da rede neuronal. Em paralelo, o sistema SCADA consulta as variáveis de processo de entrada (valores dos neurônios de entrada), processa a RNA e exibe e arquiva os valores preditos (valores do neurônio de saída).



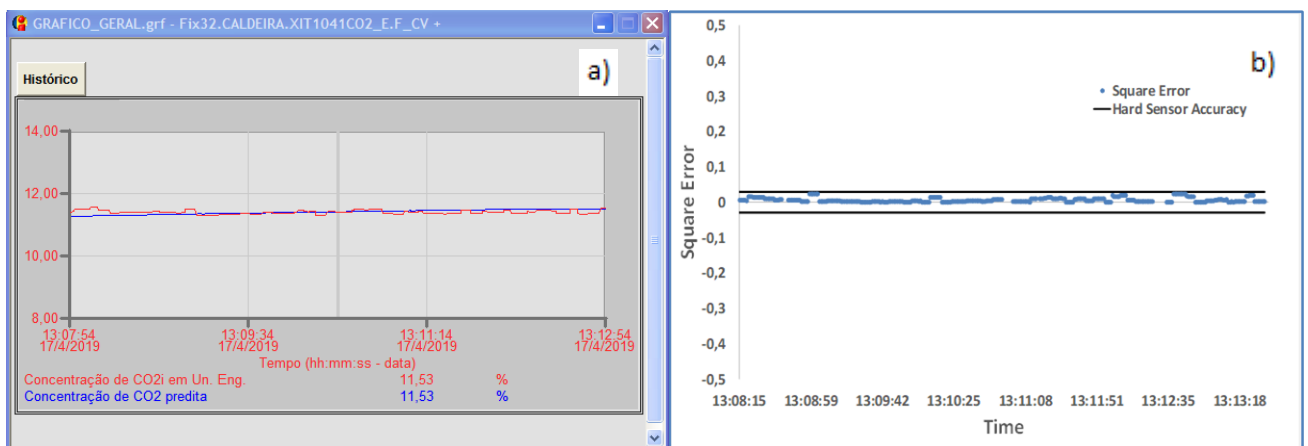
**Figura 24 Fluxograma de informações do funcionamento do sensor virtual**

Foi feito um teste experimental para validar a implementação do sensor virtual em tempo real na estação SCADA. Este teste corresponde à fase de generalização, em que dados inéditos são apresentados ao sensor virtual de CO<sub>2</sub> de uma maneira *online* por meio da ferramenta desenvolvida. As condições operacionais do teste foram operação em malha aberta com produção média de 300 kg/h de vapor saturado.

### IV.3 Principais Resultados e Discussões

A topologia da rede neuronal para inferência de CO<sub>2</sub> (%v/v), apresenta 5 neurônios na camada de entrada, 10 neurônios na camada escondida e 1 neurônio na camada de saída, com funções de ativação logística e tangente hiperbólica. Para o desenvolvimento do sensor foram usados 9333 padrões, sendo 70% deles para a fase de treinamento pelo algoritmo Bayesiano, 15% para teste e 15% para validação offline. Os padrões correspondem a diferentes regimes de operação sob diversas condições de temperatura ambiente e são constituídos por transientes de partida, perturbações de *set-point* de vazão de ar de combustão além de diferentes estados estacionários com a intenção de que a ampla varredura de condições operacionais guardasse a natureza complexa do processo de combustão. O sensor virtual apresentou coeficiente de correlação ( $r^2$ ) de 0,991 e média dos erros quadráticos (MSE) de 0,02 tanto para a fase de teste quanto para a fase de validação *offline*.

A Figura 25 apresenta os resultados da validação *online* do sensor virtual. Na Figura 25a as linhas de tendência das concentrações de CO<sub>2</sub> em tempo real são exibidas. A linha azul representa as previsões do modelo de inferência e a linha vermelha os valores alvo, medidos pelo sensor físico. O MSE foi de 0,006, cinco vezes menor do que a acurácia do sensor físico informada pelo fabricante (0,03% v/v). A Figura 25b mostra que durante toda a etapa de validação *online* do sensor virtual os resíduos de predição nunca foram superiores a acurácia do sensor real sugerindo que a ferramenta virtual pode ser usada em ambientes industriais como instrumento redundante ou até mesmo substituir o sensor físico durante os seus ciclos de manutenção.



**Figura 25** Validação *online* do sensor virtual em tempo real a) Medições e previsões de CO<sub>2</sub> b) Resíduos de predição e acurácia.

#### **IV.4 Lista das Principais Contribuições**

- Integra software comercial e SCADA industrial, facilitando implementação em plantas de processo.
- Monitoramento em tempo real de emissões atmosféricas, comprometimento com padrões ambientais.
- Sensor virtual que pode funcionar como redundância ou até mesmo substituir sensor físico.
- IHM amigável e generalizável para outros processos.

#### **IV.5 Artigo Anexo**

# Soft Sensor for Exhaust Gas Emissions Monitoring: Application in Industrial SCADA

Parente AP<sup>1\*</sup>, Valdman A<sup>2</sup>, Folly ROM<sup>2</sup> de Souza Jr. M B<sup>2</sup>.

1 Universidade Federal do Rio de Janeiro, Escola de Química, Engenharia de Processos Químicos e Bioquímicos

2 Universidade Federal do Rio de Janeiro, Escola de Química, Departamento de Engenharia Química

\* Autor correspondente: parente@eq.ufrj.br; Tel.: +55(21) 3938-7315

*RESUMO: This paper presents the online implementation of an exhaust gas concentration soft sensor in an industrial SCADA platform. The soft sensor is a data-driven model based on machine learning techniques. It was tested and validated in a pilot plant of hybrid automation technologies and networks that include analog, discrete and Fieldbus protocol-based instruments. It may be used to monitor CO<sub>2</sub> emission levels in order to achieve reliable and consistent operation regarding production efficiency and environmental safety. The automation solution is a friendly and generic framework wherein plant personnel can create, update and use local MLP ANN soft sensors. Even though the framework was implemented in a combustion process it can be generalized for different soft sensors and distinct processes.*

**Palavras-chave:** Soft Sensor, Gas Emission, Industrial SCADA

## 1. Introduction

Human society has long been relying on fossil fuels as energy sources. Despite the research endeavor to increase the use of renewable raw material, ~90% of energy is still produced from fossil fuel combustion (TAIRA, 2019). Nonetheless, in the last decades, process production has been overrun by environmental issues that are clearly stated in the Kyoto Protocol and, more recently, in the Paris Agreement. In this effort, it was agreed to unleash actions and investment that reduce greenhouse gas (GHG) emissions, mainly CO<sub>2</sub>. In addition, local environmental protection agencies like EPA (USA) and CONAMA (Brazil) legally limit air pollution arising from stationary sources. Thus, advanced intelligent systems for online and reliable measurements of such critical variables as exhaust gas concentrations are particularly useful and state a commitment with environmental standards.

Exhaust gas concentrations are difficult to measure in real time. Classic offline monitoring techniques do not meet process control demands because of the significant delay on laboratory testing and analysis. Hardware-based continuous emission monitoring systems (CEMS) such as online flue gas analyzers incur in high operational costs and need to be frequently cleaned and purged, causing data interruptions. Hence, soft sensors appear as a good alternative for the acquisition of critical process variables and process monitoring.

Soft sensor is a computer-aided estimation system that merges measurement signals and a model to predict a process variable (LIUKKONEN, HÄLIKKÄ, *et al.*, 2012). The term “soft sensor” is a combination of the words “software” and “sensor” because they are as a matter of fact computer programs that deliver equivalent information of their hardware counterparts. Fortuna (2007) listed some reasons for using soft sensors: they are a low-cost alternative to expensive hardware sensors; they can operate in parallel with hardware sensors - offering a back-up or a redundant measurement; they can be implemented on existing hardware such as microcontrollers and they allow online estimation of data. Data-driven soft sensors are



based on the data measured within the processing plants, and thus describe the real process conditions. Industrial archived process data are abundant and represent a convenient source of information that might be used for optimization and improvement of productivity. For example, in a typical refinery, thousands of measurements of process variables are collected at intervals of 1s.

The first works proposing the use of virtual sensors intended for monitoring the air appeared at the end of last century (DONG, MCACVOY e CHANG, 1995; QIN, YUE e DUNIA, 1997). Table 1 summarizes some of the late applications of soft sensor techniques in gas monitoring.

**Table 1 - Summary of soft sensor applications in gas monitoring**

Authors	Year	Subject
Ito E. H. C. et. al.	2018	Model-based soft sensor for gas composition in a distillation column
Liukkonen M. et.al.	2012	Data-driven adaptative soft sensor for NO <sub>x</sub> emission in a CFB boiler
Liukkonen M. et.al.	2013	Data-driven adaptative soft sensor for combustion of biomass
Smrekar J. et al.	2013	Linear and non-linear approaches for NO <sub>x</sub> emission predictions in a coal boiler

Despite being extensively investigated by the academic community, soft sensors have yet to be widely applied in industry due to some unaddressed development and maintenance issues (KADLEC, GABRYS e STRANDT, 2009). Nonetheless, the use of soft-sensing technology within automation facilities as a low - cost redundant measuring device may provide tools to aid plant personnel in identifying and diagnosing poorly performing sensors. They work as a back-up of the hardware measuring device and may even replace the physical sensor leading to fault tolerant processes. This paper presents a study case wherein a CO<sub>2</sub> soft sensor was implemented and validated in an industrial framework. It can perform online supervision tasks and follow the process behavior in real time. Even though the framework was implemented in a combustion process it was built as a friendly and generic Human Machine Interface (HMI). Thus, plant personnel may create, update and use local and distinct MLP soft sensors.

## 2. The Pilot Plant

The experimental tests were based on a semi-industrial pilot plant that is part of the utility facilities installed in the Chemical Engineering Laboratory (LADEQ) at School of Chemistry, Rio de Janeiro Federal University. The plant has a feed water tank, four fuel storage tanks, a fuel feed tank and a vertical fire-tube boiler that produces saturated steam to feed LADEQ's operation facilities. The boiler operates with diesel fuel, produces saturated steam and has a fixed fuel flow rate of 19kg/h and a regulated damper associated with a frequency inverter that allows air flow rate adjustments.

The automation system has Fieldbus protocol-based devices for measuring continuous process variables, a PLC controller for converting discrete-time variables and analog digital signals into Fieldbus protocol, a dedicated controller which operates interlock logic and an Ifix Proficcy® SCADA station to register engineering values and diagnosis status. It is hence a hybrid architecture automation system that merges different industrial technologies and networks. The main continuous process variables available are fuel temperature and feed tank level; feed water temperature, tank level and flow rate; boiler water level, steam production flow rate and pressure; flue gas emission temperature. In parallel, exhaust emission concentrations are continuously monitored by a Testo 350XL analog flue gas analyzer to evaluate the emission levels of the semi-industrial boiler combustion process. Among the hard sensor concentration measurements there is a periodic inherent out-of-service purge process when the flue gas analyzer signal drops. To surpass the inherent issues of the hard-measuring device there is a CO<sub>2</sub> exhaust gas concentration soft sensor, described in section 3.

Figure 1 presents the final automation architecture. It includes the physical layer – where the process equipment and the sensors and actuators such as the flue gas analyzer, the digital transmitters, the frequency inverter and the boiler itself are. The continuous and the interlock safety (PLC) control loops are part of the control level. The SCADA station embedded with computer-aided resources such as the CO<sub>2</sub> soft sensor is at the supervision layer. The automation system merges different technologies. The digital transmitters information is based on Fieldbus communication protocol and use the ethernet technology to flow through the automation system. The discrete and analog signals - as the flue gas analyzer measurements - flow through the PLC. It works as an interface so that the process status information and the CO<sub>2</sub> concentration measured values get into the digital network, which is a dedicated local area network (LAN).

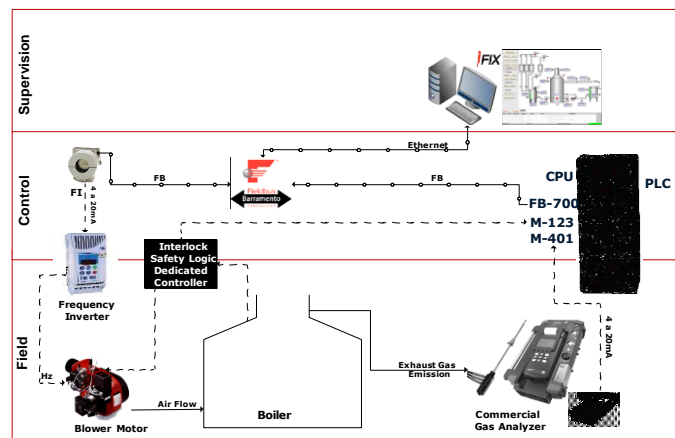


Figure 1 - Automation architecture layers.

### 3. Soft Sensor Development

The CO<sub>2</sub> exhaust gas concentrations soft sensor is based on a multilayer-perceptron (MLP) artificial neural network (ANN) model, generally described by Equation 1, wherein  $s$  is the output signal;  $f$  is the activation function;  $w$  is the synaptic weight of each neuron connection;  $\theta$  is value of the activation bias of each neuron;  $p$  is the index of the pattern vector;  $j$  is the index of each neuron;  $k$  is the index of each layer and  $n_k$  is the number of neurons in the previous layer.

$$s_{pj,k+1} = f\left[\left(\sum_{i=1}^{n_k} (w_{jik} \cdot s_{pi,k})\right) + \theta_{j,k+1}\right] \quad (1)$$

The model inputs were selected based on statistical global sensitivity analysis that shall be large ( $\gg 1$ ) for important variables. The soft sensor's input and output variables are presented in Table 2 and they are related by an MLP ANN with 5 inputs, 10 neurons on the hidden layer and 1 neuron on the output. This soft sensor was developed using the commercial software Statistica® with 9333 input-output patterns, being 70% of them used for training via the Bayesian training algorithm, 15% for model recall and 15% for model validation. The best network topology was calculated considering 1 hidden layer and the logistic and tangent-hyperbolic activation functions were used in the hidden and the output layer, respectively. This MLP presented correlation coefficient ( $r^2$ ) of 0,991 and a mean square error (MSE) of 0,02 for both recall and validation data suggesting that there was no overfitting.

**Table 2 - Input process variables of the soft sensor.**

Measured Variable	Input/Output	Unit	Instrument Range	
			Min	Max
Flue gas temperature	Input	°C	0.00	400.00
Water tank temperature	Input	°C	0.00	100.00
Fuel tank temperature	Input	°C	0.00	100.00
Vapor pressure	Input	psi	0.00	100.00
Blower flow setpoint	Input	%	0.00	100.00
CO <sub>2</sub> concentration	Output	%v/v	8.00	14.00

#### 4. Soft Sensor Online Implementation

A configuration interface was developed using embedded VBA resources (forms and scripts) within the SCADA system to allow operational personnel to create and use local MLP soft sensors. An adjustable Human Machine Interface (HMI) configuration tool (Figure 2a) was designed to activate the prediction tool. It must be filled with the following initialization parameters:

- The name of the model to be used as an online soft sensor: a reference to a text file containing the ANN model (e.g.: *3\_MLP5-10-1.ann*);

- Number of inputs, hidden and output neurons;
- Lower and upper values of each process variable in the input and output layers;
- Weight values for each neuron connection;
- Activation bias values for each neuron;
- Activation functions;
- Identification of the tags in the SCADA system that monitors the respective process data used by the input neurons;
- Identification of the tags in the SCADA system that receives the respective process data predicted by the output neurons.

Once the initialization parameters are defined, the SCADA system processes the predefined MLP algorithm to predict the output variable. On an automatic cycle basis, the soft sensor receives the 5 specified measured process variables (input variables to the soft sensor), processes the neural network model and predicts the CO<sub>2</sub> emission concentration released through the chimney (output variable of the soft sensor). The HMI was developed to allow one or more soft sensors at a time in the same SCADA system. Each of them must be parameterized and activated. Figure 2b shows the monitoring pop-up window of the pollutant soft sensor implemented in the SCADA station. In analogy to the online measured data, this architecture allows the predicted output data to be stored by the automation system with the same computational resources. Thus, the predicted values can be displayed on process synoptic, on data trends and historical data sets, including inherent alarm SCADA tools, without further special computational scripts. The SCADA interface was configured to exhibit diagnosis information in the synoptic screen to increase the reliability of the soft sensor. Whenever the prediction values equal the maximum and minimum instrument ranges, CO<sub>2</sub> concentrations are displayed in color red to identify an error. This scenario suggests that the virtual sensor is out of his trained range; in other words, the process is operating under unknown conditions and the model must be re-trained.

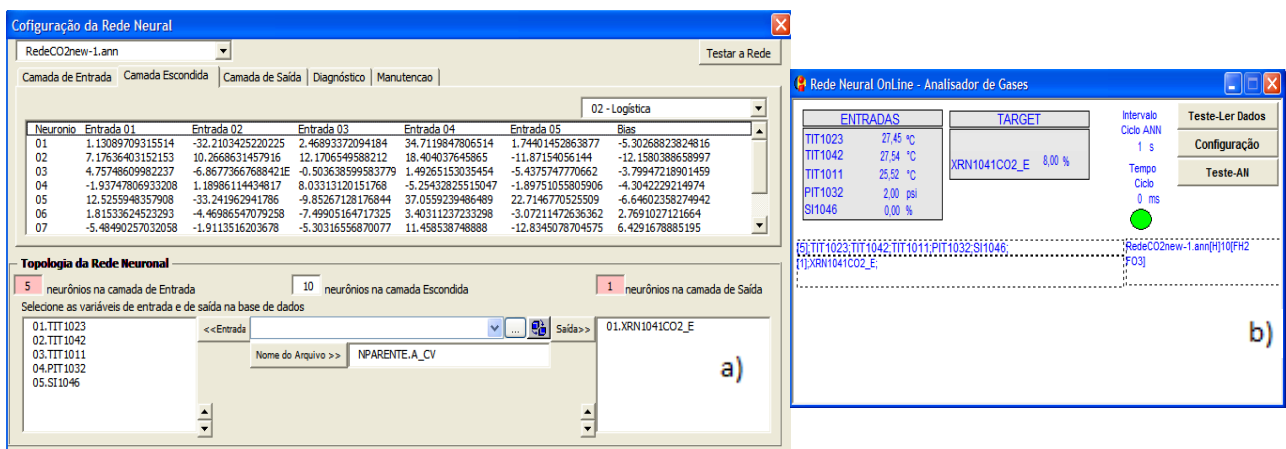


Figure 2 - HMI of the MLP ANN in SCADA a) Configuration window. b) Monitoring window.

Figure 3 illustrates the information flow diagram of the online soft sensor functioning. At first, the model must be developed using the commercial neural network package. Once the model is available the software exports its characteristics and parameters to a text file. The SCADA system uses this file along with other initialization parameters to activate the ANN calculations. In parallel, The SCADA station itself consults the input process variables (input neuron values) on time intervals of 10 seconds, performs the

model calculations and displays the predicted values (output neuron values) on the pop-up monitoring window meanwhile such predicted values are archived.

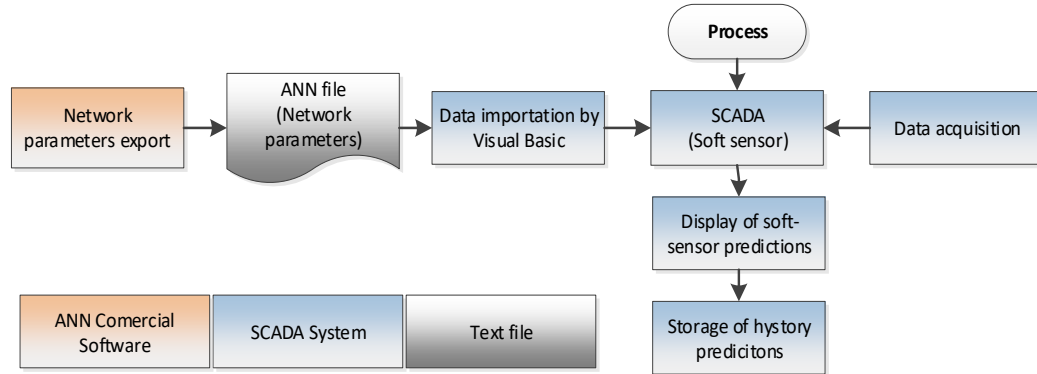


Figure 3 - Information flow diagram of the online soft sensor functioning.

## 5. Soft Sensor Online Validation

An experimental test intended to validate the implementation of the real-time soft sensor in the industrial SCADA framework was conducted. It represented the generalization phase; wherein new data was presented to the CO<sub>2</sub> soft sensor in an online manner through the developed framework. The operational condition was an open loop steady-state experiment with diesel fuel and 80% of combustion air flow rate and average steam production of 300kg/h.

Figure 4a shows the real-time CO<sub>2</sub> concentration data trend exhibited on the synoptic screen of the SCADA system during the test. To enhance performance visualization of the pollutant monitoring system the online gas analyzer target variable measurement (red line) and the soft sensor's prediction output (blue line) were compared in time trends. The final performance index (*mse*) of 0,006 was 5 times smaller than the accuracy of the hard sensor stated by the manufacturer (0,03 %v/v) proving it possible to use the model prediction framework in an industrial environment. Figure 4b shows that during the online validation test the soft sensor's prediction residuals were never higher than the hard sensor's accuracy. The online soft sensor might be used as a redundant measuring device or it may even replace the hard sensor in maintenance cycles.

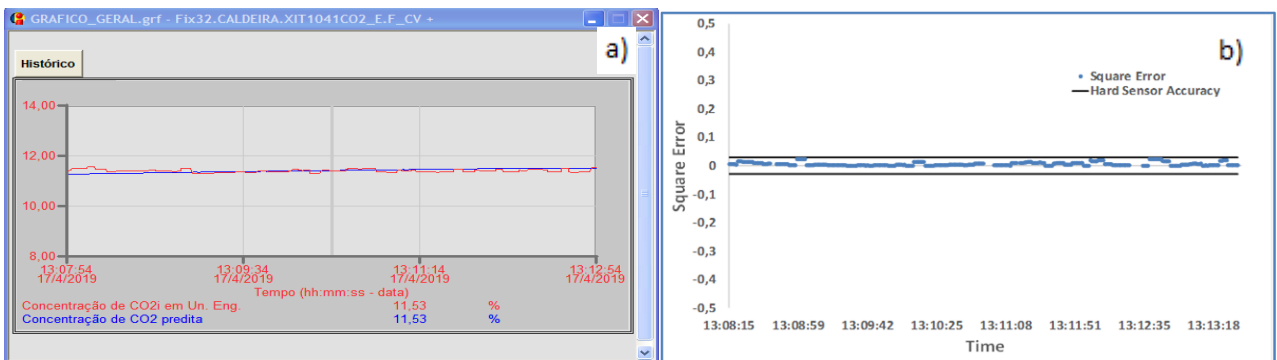


Figure 4 - Online validation of the CO<sub>2</sub> soft sensor a) CO<sub>2</sub> Measurements and prediction x time on the SCADA synoptic screen. b) Residuals and hard sensor accuracy x time.

## 6. Conclusion

This work proved it possible to use machine learning techniques in industrial SCADA platforms in order to monitor exhaust gas emission levels, stating a commitment with environmental standards. The developed soft sensor can predict CO<sub>2</sub> exhaust concentrations automatically in an online manner with *MSE* performance index of 0,006; which is 5 times smaller than the hard sensor's accuracy. Such virtual sensor may hence be used as a redundant measuring device or it may even replace the hard sensor in maintenance cycles.

A friendly Human Machine Interface (HMI) was implemented and tested so that plant personnel may create, update and use local MLP soft sensors lacking specialized developers and without great computational effort. Even though the framework was implemented in a combustion process it can be generalized for different soft sensors and distinct processes once the initialization steps can be set for other variables and independent ANN topologies.

## 7. References

DONG, D.; MCACVOY, T. J.; CHANG, L. J. Emission monitoring using multivariable soft sensors. **Process Am Control Conference**, n. 1, p. 761-765, 1995.

FORTUNA, L. **Soft sensor for monitoring and control of industrial processes**. Londres: Springer, 2007.

ITO, E. H. C. *et al.* Development of a gas composition soft sensor for distillation columns: A simplified model based and robust approach. **Computer Aided Chemical Engineering**, v. 44, p. 661-666, 2018.

KADLEC, P.; GABRYS, B.; STRANDT, S. Data-driven soft sensors in the process industry. **Computers and Chemical Engineering**, v. 33, p. 795-814, 2009.

LIUKKONEN, M. *et al.* Dynamic soft sensors for NO<sub>x</sub> emissions in a circulating fluidized bed boiler. **Applied Energy**, 2012. 483-490.

LIUKKONEN, M. *et al.* Adaptive soft sensor for fluidized bed quality: Applications to combustion of biomass. **Fuel Processing Technology**, v. 105, p. 46-51, 2013.

QIN, S. J.; YUE, H.; DUNIA, R. Self-validating inferential sensors with application to air emission monitoring. **Industrial Engineering Chemical Research**, v. 36, p. 1675-1685, 1997.

SMREKAR, J.; POTOČNIK, P.; SENEGAČNIK, A. Multi-step-ahead prediction of NO<sub>x</sub> emissions for a coal-based boiler. **Applied Energy**, v. 106, p. 89-99, 2013.

TAIRA, K. NO<sub>x</sub> emission profile determined by in-situ gas monitoring of iron ore sintering during packed-bed coke combustion. **Fuel**, 2019. 244-250.

# Capítulo V– Ferramenta de Calibração Online para Sensores Virtuais: Validação em Planta Semi-industrial

---

Este capítulo resume o artigo “An Online Calibration Tool for Soft Sensors: Development and Experimental Tests in a Semi-Industrial Boiler Plant”, submetido ao periódico *Brazilian Journal of Chemical Engineering*.

## V.1 Resumo

Sensores virtuais para predição aparecem como uma alternativa para o monitoramento de variáveis difíceis de medir. No entanto, seu uso é limitado por problemas de precisão sistemática que podem ser creditados a mudanças dos estados nominais de operação, sugerindo a necessidade de atualização frequente do modelo de inferência. Este trabalho apresenta uma ferramenta de recalibração online genérica para ser usada em sensores virtuais do tipo *multilayer perceptron* (MLP). A ferramenta de recalibração foi desenvolvida como uma estrutura única, integrando um sistema de aquisição de dados industriais e algoritmos matemáticos para a etapa de treinamento do modelo. A estrutura foi validada em um sensor virtual de uma caldeira semi-industrial para prever os níveis de emissão de poluentes em tempo real, apresentando respostas de tempo de recalibração de até 1 minuto, desempenho global de  $r^2$  acima de 80% e interface homem-máquina (IHM) amigável.

## V.2 Metodologia Sumarizada

A ferramenta de recalibração *online* para sensor virtual é baseada em técnica de janela móvel. O processo de calibração se inicia por uma decisão do operador e pode ser monitorado por meio de uma IHM amigável e a arquitetura de treinamento, permitindo dois tipos de mecanismos de recalibração. O modelo pode ser treinado do zero, usando o conjunto de dados contidos na janela, ou o procedimento de calibração pode incorporar uma etapa de “atualização” para incluir as amostras dos dados mais recentes no modelo.

Para realizar o procedimento de calibração, uma planilha eletrônica incorporada com scripts VBA (Excel®), um algoritmo em linguagem de programação C (Matlab®) e um sistema SCADA



(Ifix®) são necessários. Em suma, a estação SCADA armazena os dados históricos, a planilha eletrônica isola parte dos dados coletados para fins de calibração e o algoritmo trata os dados e executa a fase de treinamento. O resultado é um arquivo com novos valores de pesos e *biases*<sup>3</sup> que são automaticamente usados pela ferramenta de predição *online*.

O núcleo do procedimento de calibração é a fase de treinamento, que também deve incluir métodos de pré-processamento de dados. Esta fase envolve a aplicação de filtros para lidar com dados espúrios, normalização dos padrões de treinamento para o intervalo [0,1] e o próprio treinamento. O algoritmo de treinamento usado foi Regularização Bayesiana (TAMBOURGHI, FISCHER e FILETI, 2006; FORESEE e HAGAN, 1997; MACKAY, 1991), que requer uma estimativa inicial dos parâmetros a serem otimizados. A ferramenta de calibração desenvolvida possui um atributo booleano de inicialização que permite o uso (ou não) de um conjunto anterior de pesos e *biases* - um recurso que possibilita a redução do tempo de treinamento e também a preservação das informações passadas. Os critérios de parada são erro quadrático médio (MSE) mínimo de  $1 \times 10^{-5}$ , coeficiente de determinação ( $r^2$ ) mínimo de 0,90 e tempo máximo de treinamento de 30s. Após um treinamento bem sucedido, o novo conjunto de parâmetros é reunido em um arquivo em linguagem C padrão e transferido para o sistema SCADA, o que torna o novo modelo disponível e pronto para ser usado na arquitetura de automação da planta piloto.

Para aprimorar a interface homem-máquina da ferramenta de manutenção foram desenvolvidas algumas janelas do tipo *pop-up* classificadas em monitoramento, decisão e desempenho. As caixas de diálogo de monitoramento notificam o usuário sobre o fim de um evento específico. Elas trazem uma mensagem de texto, como, por exemplo “O treinamento terminou”, e um botão de clique “ok”. As janelas de decisão aparecem quando é preciso fazer escolhas. Um exemplo é uma consulta se o operador deseja ver um relatório sobre a aplicação dos filtros de pré-processamento de dados. Elas requerem a interação do operador por meio de um conjunto de botões Sim/Não/Cancela. As interfaces de desempenho permitem a avaliação da qualidade do treinamento por meio de gráficos de dispersão entre valores preditos e valores alvo. Todas as janelas *pop-up* apresentam modos de operação automático e manual. Isto significa que para modelos MLP bem conhecidos pelo operador ele pode decidir reduzir o número de interrupções, permitindo que a recalibração seja completamente automática.

---

<sup>3</sup> valor constante que corresponde ao limite de excitação de cada neurônio.



A aplicação experimental deste trabalho foi desenvolvida em uma unidade piloto que é parte da Central de Utilidades do Laboratório de Engenharia Química da UFRJ (LADEQ/EQ/UFRJ). A planta conta com uma caldeira de porte semi-industrial, que produz vapor saturado. O sistema de automação é composto por transmissores *Fieldbus Foundation* dedicados à medição das variáveis de processo contínuas e por uma estação para monitoramento e supervisão do processo (SCADA).

A planta piloto conta também com um analisador de gases comercial de tecnologia analógica da marca TESTO 350M/XL, um instrumento para medições *online* e *in situ*, que é composto por seis células de medição de concentração de gases. Este apresenta características de operação específicas, que incluem alto custo de manutenção e intervalos cíclicos fora de serviço para limpeza interna ou purga.

Em paralelo, o sistema de automação desta planta inclui um sensor virtual baseado em rede MLP para predição de concentração dos mesmos gases de combustão medidos pelo analisador de gases físico. O sensor virtual que está incorporado na camada de automação da planta piloto, foi desenvolvido em trabalhos anteriores (VALDMAN, VALDMAN, *et al.*, 2011) e apresenta 6 neurônios na camada de entrada; 10 neurônios na camada escondida e função de ativação de tangente hiperbólica; e 6 neurônios na camada de saída e função de ativação identidade. Os pesos sinápticos de cada conexão neuronal e os *biases* de ativação de cada neurônio são os parâmetros que devem ser otimizados através do procedimento de recalibração.

### V.3 Principais Resultados e Discussões

A ferramenta de recalibração foi ativada duas vezes durante os testes experimentais e realizou diversas interações de treinamento até que um dos critérios de parada fosse alcançado. Estes critérios são MSE de  $1 \times 10^{-5}$ , mínimo coeficiente de correlação de 0,90 and tempo de treinamento máximo de 30s.

Para a recalibração 1 pesos e *biases* antigos foram usados como estimativa inicial, mas apenas novas amostras de dados foram usadas para o retreinamento do modelo. Para a recalibração 2, além de usar pesos e *biases* antigos como suposição inicial, padrões da recalibração 1 foram reunidos com uma nova coleção de padrões para treinamento. Enquanto as recalibrações foram realizadas, a caldeira operou com diesel combustível na condição nominal de

operação, produzindo em média 300 kg/h de vapor saturado. Foi uma operação em malha aberta e, portanto, sujeita a perturbações de processo naturais.

Os parâmetros de desempenho final da etapa de validação *offline* após a recalibração 2 estão mostrados na Tabela 13. Resultados satisfatórios são aqueles que apresentam valores de  $r^2$  próximos a 1 e valores de MSE próximos de 0, dentro de uma tolerância. A precisão de predição para as concentrações de CO, SO<sub>x</sub> e NO<sub>x</sub> é menor devido a razões específicas: estes componentes não são rastreáveis porque aparecem em quantidades muito pequenas. Não obstante, o desempenho médio global da calibração foi aceitável com índice final de desempenho (MSE) de  $4,3414 \times 10^{-3}$  e coeficiente de determinação ( $r^2$ ) de 0,9490.

**Tabela 13 Desempenho final da recalibração**

Componente	$r^2$	<i>mse</i>
O <sub>2</sub>	0.9570	$5.4646 \times 10^{-4}$
CO	0.8897	$2.2198 \times 10^{-3}$
CO <sub>2</sub>	0.9485	$2.6103 \times 10^{-3}$
NO	0.8173	$8.7843 \times 10^{-3}$
NO <sub>2</sub>	0.8449	$1.0087 \times 10^{-2}$
SO <sub>2</sub>	0.8775	$1.8808 \times 10^{-3}$

Como o sensor virtual e a ferramenta de recalibração estão implementados em uma estrutura unificada na estação SCADA, assim que o modelo de predição era atualizado pela rotina de manutenção o sensor virtual era imediatamente usado em uma aplicação em tempo real na planta piloto. Para melhorar a visualização do desempenho do procedimento de recalibração, os valores de medição do analisador de gás *online* (linha vermelha) e os valores preditos pelo sensor virtual (linha azul) foram comparados em linhas de tendência temporais, uma para cada componente de interesse. A Figura 26 mostra estes resultados para o CO<sub>2</sub>, para fins ilustrativos. Após cada ciclo de recalibração (1 e 2), a capacidade de previsão do modelo foi melhorada. O sensor virtual estimou valores de concentração continuamente, enquanto o sensor físico não mediu as variáveis de interesse durante o procedimento de purga, o que fica evidente pela mudança abrupta na variável de processo medida, apresentando valores estáticos iguais a uma situação de *burnout*<sup>4</sup>.

<sup>4</sup> o *burnout* pode ocorrer quando a leitura do sensor está fora do range ou quando o circuito do sensor está aberto.

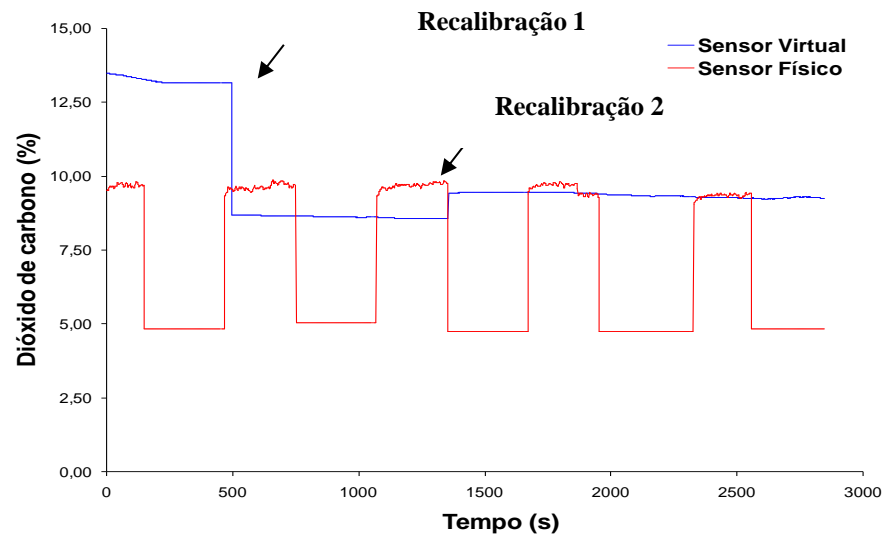


Figura 26 Validação *online* do sensor virtual para predições de CO<sub>2</sub>

#### V.4 Lista das Principais Contribuições

- Atualização de modelo automática em menos de 1 minuto, incluindo etapas de pré-processamento de dados.
- Mescla diferentes plataformas de software: soluções matemáticas e IHM amigável.
- Opções customizadas de treinamento: método matemático, metas de desempenho, parâmetros de filtragem e de coleta de dados.
- Estrutura generalizável para diferentes sensores virtuais e processos distintos.

#### V.5 Artigo Anexo

# An Online Calibration Tool for Soft Sensors: Development and Experimental Tests in a Semi-Industrial Boiler Plant

Andréa Pereira Parente, Andrea Valdman, Rossana Odette M Folly, Maurício Bezerra de Souza Jr., Ana Maria Frattini Fileti.

Corresponding author: Andréa Pereira Parente  
Address: Universidade Federal do Rio de Janeiro / Centro de Tecnologia  
Av. Athos da Silveira Ramos, 149  
Escola de Química  
Bloco E - Sala E-201 - Ilha do Fundão  
Rio de Janeiro - Brasil - CEP 21941-909  
Caixa Postal 68542  
Phone: 5521 39387315  
E-mail: [parente@eq.ufrj.br](mailto:parente@eq.ufrj.br)

## ABSTRACT

Soft sensors with real time prediction capabilities appear as a profitable solution for hard-to-measure variables whenever hard sensors are difficult to apply or subjected to high operational costs. Nonetheless, the use of soft sensors within industrial applications is still not widespread because of the systematic accuracy issues that can be introduced with process plant deviations from nominal operation states. Soft sensors models need to be constantly updated to avoid degradation of its prediction potential. This study presents an innovative view on a well-known artificial neural network (ANN) calibration method by developing a generic online calibration tool that can be used in independent data-driven soft sensors based on ANN multi-layer perceptron (MLP) models. The maintenance framework has been fully tested in a semi-industrial boiler plant to predict real time pollutants emission levels, presenting recalibration time responses up to 1 minute, overall  $r^2$  performance above 80% and an intuitive Human-Machine-Interface.

Keywords:  
model maintenance, soft sensor, gas emission, ANN, SCADA

## 1. Introduction

Certain process variables such as exhaust gas concentrations are difficult to measure in real time and, therefore, they are classically monitored by offline laboratory analysis (Andrijić *et al.*, 2018; Lin, *et al.*, 2007). Composition sensors with small real time sampling frequencies are not commonly encountered in industrial process plants, due to high operational costs of specific hard sensors. For instance, flu gas analyzers must be frequently purged, causing data interruptions. Despite the challenges to monitor such variables, they are critical assets because they reveal the levels of pollutants emissions.

In the last decades, process production has been overrun by environmental issues that are clearly stated in the Kyoto Protocol and, more recently, in the Paris Agreement. In addition, local environmental protection agencies like EPA (USA) and CONAMA (Brazil) legally limit air pollution arising from stationary sources. In this context, online and reliable measurements of such critical variables are particularly useful and state a commitment with environmental standards.

To cope with this issue Parkinson *et al.* (2019) developed a wireless hard sensor for continuous monitoring of indoor air quality. Tao *et al.* (2018) also proposed a sensor to monitor greenhouse gas emissions in components manufacturing. Nonetheless, in recent years, soft sensors - a computer-aided estimation system - have gained attention in process industry to provide online estimates of difficult-to-measure parameters in a real-time manner. Some examples are Clerqc *et al.* (2019), Santos *et al.* (2018), Kamat and Madhavan (2016) and Grbić *et al.* (2013).

Notwithstanding, soft sensors applications are limited by model performance. The built model is usually representative of the process nominal states observed during the training phase meanwhile the current process dynamics may have changed suggesting the need of maintenance. Even if a highly accurate soft sensor is developed, its predictive performance may

eventually drop in case of drifted operating conditions, set-point changes, catalyst deactivation, seasonal effects, variances of raw materials, process fouling, abrasion of mechanical components, equipment aging and process faults (Kadlec, Grbić and Gabrys, 2011; Kaneko and Funatsu, 2013; Chen, *et al.*, 2015). Kano and Ogawa (2010) reported a trial survey that addresses that “accuracy deterioration due to changes in process characteristics” accounted for 29% of the problems in soft sensing.

Besides, a prediction error greater than a threshold is typically identified as an abnormal situation. According to Kaneko and Funatsu (2013) there is no effective practice to judge if the reason of the abnormal situation is the degradation of a soft sensor model. Thus, soft sensors should be periodically updated using the newest data available to avoid performance degradation.

Many techniques developed for online adaptation to cope with the issue of degradation in soft sensor models are extensively discussed (Kadlec, Grbić and Gabrys, 2011; Chen, *et al.*, 2015; Jin, *et al.*, 2015). One of them is to utilize all the available data for model update through recursion or moving window updates. Moving window techniques are employed to update the soft sensor model sequentially using a set of selected data points in a block-wise manner. Some examples of the use of regression methods are Cao and Lou (2014), Kadlec and Gabrys (2010) and Lu *et al.* (2014).

Other possible solutions include the work of Zheng *et al.* (2018) that developed a just-in-time semi-supervised nonlinear soft sensing method to perform model update. Their methodology integrates the just-in-time learning, extreme learning machine and the graph Laplacian regularization. Lu and Chiang (2018) also proposed a maintenance strategy to deal with soft-sensor performance degradation using the robust mean and variance estimators of the variable’s inputs and outputs. The authors validated their technique in industrial case studies.

While this adaptive update process can be automated, there are specific challenges for industrial implementation mainly because of the existence of different software manufacturers in a same industrial process network.

In addition, the convergence and stability of these update algorithms depend on the data being free of faults or outliers, which is seldom the case in industrial systems. Data that are collected during process operation often contain spurious points because of different disturbances, malfunctions, degradation and errors in sensors and data acquisition system. These impurities that affect model quality negatively appear as high frequency noise, sudden abrupt changes, drifts, offsets, trends, low frequency disturbances and missing values (Souza, Araújo and Mendes, 2016).

To remove these samples data pre-processing has been incorporated in model building procedure (Warne, Prasad and Maguire, 2004; fortuna, 2007; Kadlec, 2009). In online applications data need to be pre-processed in an online manner. This consists of several operations such as outlier detection, missing value replacement and data scaling.

Another critical issue is that there are two parameters with great influence on the model performance: the size of the adaptation window or *window size*, and the size of the adaptation intervals between the updates or *step size*, that shall be set correctly in order to guarantee that the process dynamics do not change within the span of one moving window (Chen, *et al.*, 2015).

Online application of these methods also requires modern algorithm execution platforms, which has slowed their adoption industry wide according to Lu and Chiang (2017). The update scheme should be robust, simple and effective. Otherwise the model maintenance should fail gracefully when the mismatches between the trained model and process data are too significant to reconcile (Lu and Chiang, 2017).

The main scope of this study is to develop a generic online calibration tool for a soft sensor. The framework was designed and tested in a combustion semi-industrial environment with a soft sensor that predicts pollutants real-time emission levels in the pilot plant. The importance of this

soft sensor is its ability to continuously estimate concentration values; meanwhile the hard sensor did not measure the interest variables during the inherent out-of-service periodic cleaning procedure.

The maintenance procedure can be applied to any data-driven soft sensor based on an Artificial Neural Network Multilayer-Perceptron model. It integrates mathematical model training solutions with supervision and data acquisition (SCADA) software through a friendly human-machine interface (HMI). In addition, a mathematical function is proposed to calculate the size of the adaptation window on a time basis enabling the automation of data collection. An alternative performance goal is also suggested intended to enhance the model validation step.

## 2. The Online Prediction Tool

The proposed online calibration tool is intended to work in association with local MLP soft sensors, which are the online prediction tools. Hence, a configuration interface was developed using embedded VBA resources (forms and scripts) within the SCADA system to allow operational personnel to create and use MLP soft sensors, generally described by Equation 1:

$$s_{pj,k+1} = f\left[\left(\sum_{i=1}^{n_k} w_{jik} \cdot s_{pi,k}\right) + \theta_{j,k+1}\right] \text{ Equation 1}$$

Where:

- $s$  is the output signal;
- $f$  is the activation function
- $w$  is the synaptic weight of each neuron connection
- $\theta$  is value of the activation bias of each neuron

$p$  is the index of the pattern vector

- $j$  is the index of each neuron
- $k$  is the index of each layer
- $n_k$  is the number of neurons in the previous layer



The  $w$  synaptic weights and  $\theta$  activation biases are the parameters that shall be adjusted through the calibration procedure. An adjustable human-machine interface (HMI) configuration tool was designed to activate the prediction tool. It must be filled with the following initialization parameters:

- The name of the model to be used as an online soft sensor: a reference to the file containing the ANN model (e.g.: *3\_MLP6-10-6.ann*);
- Number of inputs, hidden and output neurons;
- Lower and upper values of each process variable in the input and output layers;
- Weight values for each neuron connection;
- Activation bias values for each neuron;
- Activation functions;
- Configuration of the tags in the SCADA system that monitors the respective process data used by the input neurons;
- Configuration of the tags in the SCADA system that receives the respective process data predicted by the output neurons.

Once the initialization parameters are defined, the SCADA system processes the predefined MLP algorithm to predict the output variables. The HMI was developed to allow one or more soft sensors at a time in the same SCADA system. Each of them must be parameterized and activated. In analogy to the online measured data, this architecture allows the predicted output data to be stored by the automation system with the same computational resources. Thus, the predicted values can be displayed on process synoptic, on data trends and historical data sets, including inherent alarm SCADA tools, without further special computational scripts.

### 3. The Online Calibration Tool

The soft-sensor online calibration tool is based on a block-wise moving window technique. It was designed and tested in association with a soft sensor developed on previous studies (Valdman, *et al.*, 2011) that is described in section 4. Such calibration tool is an industrial framework that must be used for periodical model maintenance.

The calibration process can be monitored through a friendly HMI and the training architecture allows two types of re-training mechanisms. The model may be trained from scratch using the set of data contained in the window or the calibration procedure may incorporate an “updating” step to include the latest data samples into the model.

To perform the calibration proceeding, an electronic sheet embedded with VBA scripts, an algorithm in programming language C and a SCADA system are required. This work used the commercial software Excel®, Matlab, with neural networks toolbox included, and Ifix Proficy®.

In short terms, the SCADA station holds the historical data, the electronic sheet isolates part of the collected data for calibration purposes and the C algorithm pre-processes the data and performs the training phase. The result is a file with new values of weights and biases that is automatically used by the online prediction tool, previously developed in the SCADA system and described in section 2. The calibration routine flow diagram is shown in Figure 1.

It is important to determine the adequate number of samples for the training phase. Inappropriate setting could lead to performance degradation instead of improvement (Kadlec, Gabrys and Strand, 2009). Based on a literature report that evaluated the ideal mathematical relationship between the number of samples and the number of weights of an ANN (Maier and Dandy, 1998) and also on a commercial software praxis that suggests that the number of samples must exceed the number of inputs from 10 to 40 times (Neuralware, 2010) the following expressions are proposed for the number of patterns.

$$NP\alpha = NP \times rpp \times \alpha \text{ Equation 2}$$

$$NP = W + \theta \text{ Equation 3}$$

Where:

- $NP\alpha$  is the minimum quantity of samples for the training step;
- $NP$  is the total number of parameters of the MLP network;
- $rpp$  is the ideal ratio between the number of samples and the number of parameters of the MLP network (used as 10 in this work based on previous experiences);
- $W$  is the total number of weights of the MLP network;
- $\theta$  is the total number of biases of the MLP network;
- The constant  $\alpha$  is a correction factor (used as 1.1 to increase the number of samples in 10% in this work).

The automated data collection is available in the SCADA system on a time basis. Thus, the minimum number of samples is converted into the window size,  $ws$ , via the following logical rule (Figure 2). The expressions consider that the equipment that provides the patterns for the supervised learning procedure has a sampling time,  $c$ , and an inherent periodic out-of-service time,  $l$ . Variable  $T$  represents the SCADA data acquisition sampling time.

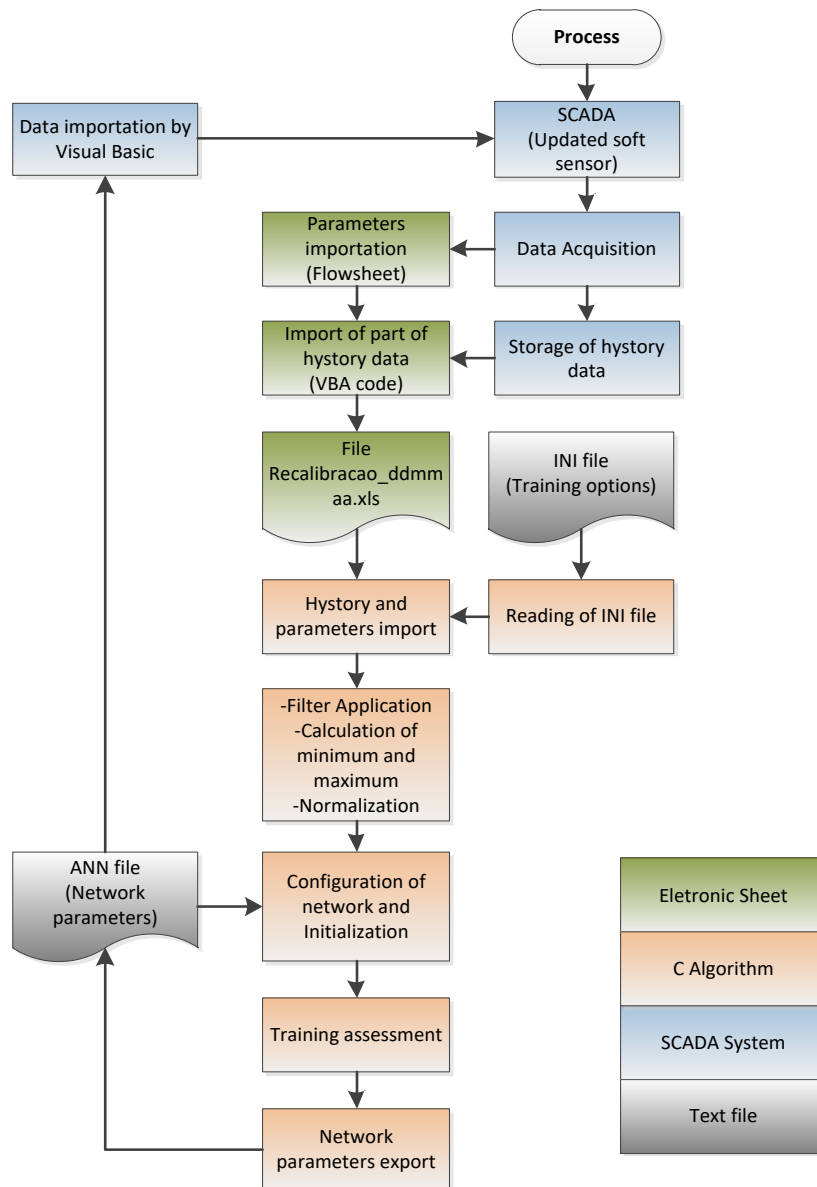


Figure 1. Calibration routine flow diagram.

$$\text{If } \frac{NP_a \cdot T}{c} = \text{Integer} \left( \frac{NP_a \cdot T}{c} \right)$$

$$ws = \left[ \text{Integer} \left( \frac{NP_a \cdot T}{c} \right) - 1 \right] \cdot (c + l) + l$$

else

$$ws = \left[ \text{Integer} \left( \frac{NP_a \cdot T}{c} \right) \right] \cdot (c + l) + l$$

$$+ \text{rest} \left( \frac{NP_a \cdot T}{c} \right)$$

Figure 2. Logical diagram for sample acquisition on a time basis.

The kernel of the calibration procedure is the training phase, which is performed by the C algorithm that is embedded with a Neural Network toolbox. The model adaptation has also to include the adaptation of the data pre-processing methods. This phase involves the application of filters and constraints to deal with spurious data, data set scaling and configuration of the network and the training itself. The removal of bad samples from the historical data set is a four-step procedure based on conditional logical rules:

- Out-of-service filter: Among hard sensor devices there might be typical periodic inherent out-of-service cleaning process. The detection and elimination of not useful data is done by means of a concomitant analysis of a discrete diagnosis variable, which returns 1 for out-of-service conditions and 0 for normal operation, and the assessment of entire rows of samples. The out-of-service device filter removes data when all the values of an entire row of samples are missing.
- Transient constraint: Removing variable values that are consistent with start-up transients improves the training speed. The transient constraint eliminates all samples when a designated process variable is below a set-point value that is representative of a nominal steady-state. In this work the transient constraint was defined by the steam flow rate of 200 kg/hr.
- Void filter: Communication errors between the field instrument and the supervisory system can happen because of random acquisition failure. The missing values displayed as an empty cell are also cut out.
- Outlier filter: In order to make the data set free of values that are not consistent with most of the data collection the  $3\sigma$  rule limit check is applied. It is based on mean and standard deviation calculations within the window size to discard such spurious samples.

After filtering, data scaling is performed so that the sets are normalized to the range [0,1]. It is a task intended to deal with variables of different magnitudes during the learning procedure. Hence, the maximum and minimum values of each process variable are consulted at each data collection.

The training algorithm is executed immediately after the data pre-processing steps and this work used the Levenberg-Marquardt with Bayesian regularization method (Tambourghi, Fischer and Fileti, 2006; Foresee and Hagan, 1997; MacKay, 1991). The numerical method of the training requires an initial estimate of the parameters to be optimized. The developed calibration tool has an initialization Boolean attribute that allows the use (or not) of a previous set of weights and biases, a feature that enables reduction of the training time and the preservation of the past information.

In order to guide the iterations, a configurable performance function is used as an assessment parameter. Standard performance goals are the coefficient of determination ( $r^2$ ) and the mean square error (MSE), described in Equation 4:

$$mse = \frac{1}{NT} \sum_{i=1}^{NT} (t_i - a_i)^2 \text{ Equation 4}$$

Where:

$NT$  is the total number of samples;

$t_i$  is each output sample;

$a_i$  is each output prediction of the  $i$ 'th sample variable.

This work suggests the use of an alternative performance index ( $P_{alt}$ ), described by Equation 5. It considers the mean square error and includes the slope of the straight line that relates predicted and measured values in dispersion plots.

$$P_{alt} = \frac{1}{m_{reg}} \sqrt{\frac{MSE}{n_{out}}} \text{ Equation 5}$$

Where:

- $MSE$  is the mean square error of each predicted variable;
- $m_{reg}$  is the slope of the straight line fitted in dispersion plots for each output;
- $n_{out}$  is number of outputs of the prediction model.

When the training phase is accomplished, the final mean square error, the coefficient of determination and the alternative index are calculated to evaluate the result of the training. The stopping criteria are normalized minimum  $MSE$  performance goal of  $1 \times 10^{-5}$ , minimum correlation coefficient of 0.90 and maximum training time of 30s.

In case the training result is not satisfactory the calibration tool transmits a training error message and the operator can delete the calculated parameters and perform a new model adaptation increasing the number of samples. After a successful training, the new set of parameters is gathered in a standard C file and transferred to the SCADA system, which makes the new model available and ready to be used at the automation architecture of the pilot plant.

The calibration framework has a few preset Boolean parameters that allow or not pop-up windows meanwhile the procedure is performed. This debug option enables a fully automated or an interactive system, where the operation staff can make decisions and monitor the recalibration process step-by-step through an HMI.

#### 4. The Pilot Plant

The experimental applications of this work were based on a pilot plant of semi-industrial size that is part of the utility facilities installed in the Chemical Engineering Laboratory (LADEQ) at School of Chemistry, Rio de Janeiro Federal University. Concisely, the plant has a feed water tank, four fuel storage tanks, a fuel feed tank and a vertical fire-tube boiler that produces

saturated steam to feed LADEQ's operation facilities. The boiler has a fixed fuel flow rate of 19kg/hr, and a regulated damper associated with a frequency inverter that allows air flow rate adjustments. Table 1 shows the technical specifications of the boiler.

**Table 1. Technical specifications of the boiler.**

Type	Fire-tube, vertical
Fuel	Diesel
Nominal boiler load (MW)	0.30
Maximum work pressure (kgf/cm <sup>2</sup> )	8.00
Steam produced	Saturated

The automation system has Fieldbus protocol-based devices for measuring continuous process variables, a PLC controller for converting discrete-time variables and analog digital signals into Fieldbus protocol, a dedicated controller which operates interlock logic and an Ilix Proficy® SCADA station to register engineering values and diagnosis status.

The main continuous process variables available are fuel temperature and feed tank level; feed water temperature, tank level and flow rate; boiler water level, steam production flow rate and pressure; flue gas emission temperatures.

In addition, exhaust emission concentrations are continuously analyzed by a Testo 350XL analog flue gas analyzer to evaluate the emission levels of a semi-industrial boiler combustion process. Among the hard sensor concentration measurements there is a periodic inherent out-of-service cleaning process when the flue gas analyzer signal differs from the usual data. The purge procedure is cyclical and lasts 5 minutes for every 5 minutes of useful data. Because the gas analyzer is not a widespread industrial equipment and also because the measurements come with a high operational cost, the flue gas concentrations of O<sub>2</sub>, CO<sub>2</sub>, CO, NO, NO<sub>2</sub> and SO<sub>2</sub> constitute the variables estimated by the soft sensor, in other words, the output or target variables. Its features are listed in Table 2.



**Table 2. Process variables predicted as soft sensor outputs.**

Measured Variable	Unit	Instrument Range		Accuracy
		Min	Max	
Oxygen concentration	%	0.0	25.00	0,2% up to 99 ppm: 10ppm
Carbon monoxide concentration	ppm	0	10000	from 100 to 2000ppm: 5% of read value
Carbon dioxide concentration	%	0.0	50.00	0,3% + 1% of read value
Nitrogen monoxide concentration	ppm	0	3000	up to 99 ppm: 5ppm
Nitrogen dioxide concentration	ppm	0	500	up to 99 ppm: 5%
Sulfur dioxide concentration	ppm	0	5000	up to 99 ppm: 5%

This soft sensor is based on an ANN model. The model inputs of the soft sensor are presented in Table 3 and they are monitored by pressure, temperature and flow rate digital transmitters. The set-point of the air blower flow rate is also an input for the built model.

**Table 3. Input process variables of the soft sensor.**

Measured Variable	Unit	Instrument Range	
		Min	Max
Flue gas temperature	°C	0.00	400.00
Water tank temperature	°C	0.00	100.00
Fuel tank temperature	°C	0.00	100.00
Vapor pressure	psi	0.00	100.00
Vapor flow rate	kg/hr	0.00	500.00
Blower flow setpoint	%	0.00	100.00

The inputs and outputs of the soft sensor are related by a multilayer-perceptron (MLP) neural network with 6 inputs, 10 neurons on the hidden layer and 6 neurons on the output. To develop this soft sensor Valdman *et al.* Valdman and co-workers (2011) used 502 input-output patterns, being 80% of them used for training via the Bayesian training algorithm and 20% for validation of the ANN model. The best network topology was calculated considering 1 hidden layer and the tangent-sigmoid and linear activation functions were used in the hidden and the output layer, respectively. The topology of the neural network is presented in Figure 3.

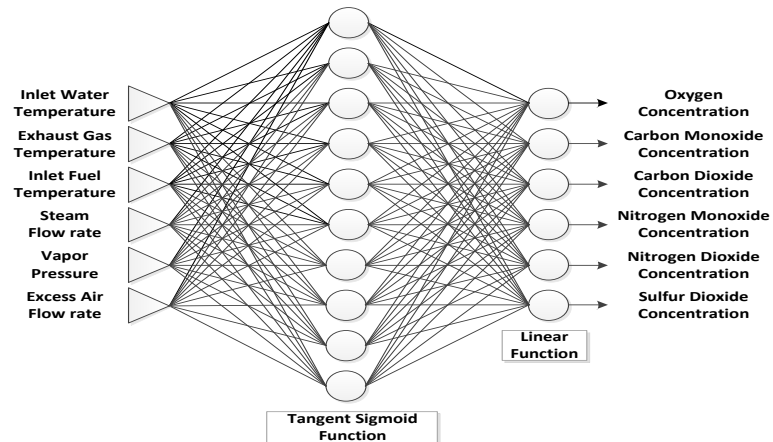


Figure 3. Topology of the soft sensor.

## 5. Results and Discussion

This section presents the main results of the experimental tests intended to validate the developed calibration tool. In Section 5.1 the general features of the tool are discussed. Section 5.2 discusses the initialization features that make the calibration tool a generic framework. Sections 5.3 and 5.4 show the results of the experimental validation tests.

The tests were performed in an online manner and included a start-from-scratch calibration, where no previous knowledge of the process was used, and two sequential recalibration procedures.

For the start-from-scratch calibration the operational condition was an open loop experiment with diesel fuel. Aiming at generating samples for the first calibration of the soft sensor the combustion air flow rate was subjected to step changes leading to 30, 20 and 10% of excess air for combustion. This test lasted 100 minutes and was used to assess and validate the automated data pre-processing steps.

For recalibration 1 old weights and biases were used as initial guess but only new data samples were used for model training. For recalibration 2 not only old weights and biases were used as initial guess but also data samples from recalibration 1 were gathered with a new collection of samples for training. In order to enhance performance visualization of the *online* validation test hard sensor measurements and soft sensor predictions were acquired at intervals

of 1s and compared. While the recalibrations were performed, the boiler plant was operating with diesel fuel at nominal steady-state condition, producing in average 300kg/hr of saturated steam. It was an open loop operation, hence subjected to natural process disturbances.

Table 4 summarizes the tests parameters.

**Table 4. Experimental Tests Parameters**

<b>Test</b>	<b>Duration</b>	<b>Number of Samples</b>	<b>Previous parameters?</b>	<b>Sample Accumulation?</b>
Calibration	100 min	5953	No	No
Recalibration 1	500s	2200	Yes	No
Recalibration 2	1357s	2300	Yes	Yes

### *5.1 Overview of the calibration tool*

The calibration routine was successfully developed: the data transfer worked properly, and the training process lasted up to 1 minute. Aiming at the enhancement of the human-machine interface (HMI) some auxiliary pop-up windows were developed. They are:

- Monitoring windows;
- Decision windows;
- Performance windows.

The monitoring dialog boxes are intended to notify the user of the end of a specific event. They consist of text information and an “ok” button. When important choices are necessary, the decision dialog boxes appear. They include the set of buttons Yes/No/Cancel. The performance windows are useful to evaluate the training quality comparing the predicted output and the target output values. The list of the pop-up windows for the user interface is shown on Table 5.

All the designed dialog boxes allow automatic and manual operation mode. Thus, for well-known MLP models the operator can suitably reduce the number of interruptions by setting automatic calibration.

**Table 5. List of windows developed.**

<b>Window</b>	<b>Class</b>
XLS file reading end	Monitoring
Filter application report	Decision
Configured Network notification end	Monitoring
Training end	Monitoring
General regression chart	Performance
Regression chart for each variable	Performance
Training results report and assessment	Decision
ANN file writing end	Monitoring

### 5.2 Initialization of the calibration tool

The calibration tool initialization procedure fulfills the framework with flexible features, allowing customized options through configurable Boolean flags in an .ini-type file. The initialization parameters shown in Table 6 enable users to decide whether to start training from scratch or use old weights and biases as an initial guess. In addition, the training can incorporate or not previous training samples and thus increase the training data window size, raising the operational range of the soft sensor and improving the predictions quality. It is also possible to choose if pop-up dialog boxes are necessary during the calibration process or if an automated – and faster – strategy is preferred.

**Table 6. List of initialization parameters**

<b>Initialization Parameter</b>	<b>Description</b>
PreviousTData	Add previous training data?
Infbox	Use informative boxes?
Qstbox	Use questions boxes?
UI	Show user interface
PreviousValues	Use previous values as initial guess?
PlotRegression	Show dispersion plots?

The commercial neural network toolbox allows the user to modify the default training method available in the commercial software database (BEALE, HAGAN e DEMUTH, 2012). Hence, the maintenance procedure is inherently generic and may be applied to different neural networks architectures and various training algorithms. The performance indexes and goals (Table 7) that

are used to evaluate the calibration process and define the stopping training criteria are also configurable.

**Table 7. List of performance parameters**

<b>Performance Parameter</b>	<b>Description</b>
Goal	Minimum MSE Performance
minR2	Minimum R2
Alt	Alternative performance index
Time	Maximum time of training

### *5.3 Automated data pre-processing*

The automated data pre-processing methods were individually tested. They were able to normalize the samples and dealt with spurious and out of range data successfully.

In this experiment just two of the four filters reduced the number of samples: the out-of-service filter and the outliers filter. The out-of-service filter reduced the number of samples in more than 50% because of the flue gas analyzer inherent cleaning process. It takes 5 minutes to purge impurities for each cycle of measurement that lasts 10 minutes. The transient filter did not remove any sample, indicating that within the historical data window there was not a shutdown – which would be evidenced by low steam production. The void filter did not remove any sample either, indicating the absence of communication errors. The sample reduction of the outlier's filter is considerably small which demonstrates a historical data of high quality. The filter application results can be seen in Table 8 and the percentage reduction is expressed in relation to the initial number of samples.

**Table 8. Number of samples before and after application of filters.**

	<b>Initial number of samples</b>	<b>Final number of samples</b>	<b>Percentage reduction</b>
Purge filter	5953	2811	52.70 %
Flow filter	2811	2811	0.00 %
Void filter	2811	2811	0.00 %
Outliers filter	2811	2560	4.20 %
All filters	5953	2560	57.00 %

### 5.4 Model Prediction

The calibration tool was activated three times during the experimental tests and performed several training interactions until a prediction of the target variables within previously stated tolerance criteria was achieved, allowing an updated visualization of the predicted values at the HMI of the SCADA system. The calibration was a start-from-scratch procedure wherein no previous knowledge about the process for initial guesses was used. The recalibrations 1 and 2 were maintenance procedures intended for model update.

The calibration technique uses 80% of the data in the moving window for training and 20% for offline validation. In the soft sensor maintenance routine, the performance indexes  $r^2$  and MSE for each output variable in the offline validation phase are calculated and presented to the user. The user interface for monitoring the results of the calibration are exhibited in a pop-up window (Figure 4).

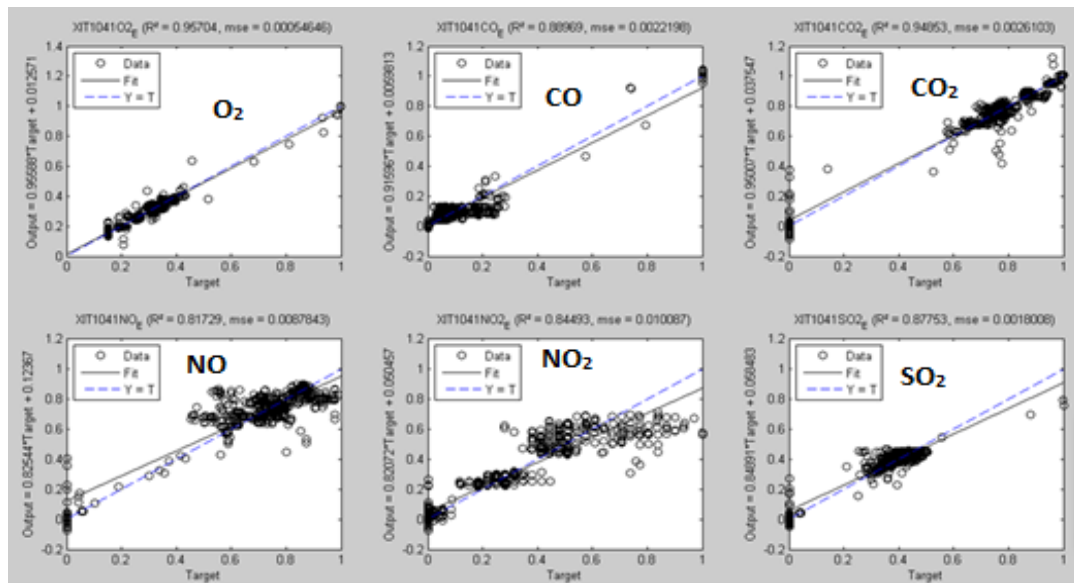


Figure 4. User Interface (UI) – pop-up window to monitor the results of the offline validation phase.

Satisfactory results are those that present values of  $r^2$  close to 1 and MSE values close to 0 within a tolerance. The alternative performance parameter ( $P_{alt}$ ) shall tend to 0. The final performance parameters of the calibration test are listed on Table 9.

**Table 9. Performance parameters for each output variable.**

<b>Component</b>	<b><math>r^2</math></b>	<b><i>mse</i></b>	<b><math>P_{alt}</math></b>
O <sub>2</sub>	0.9570	5.4646x10 <sup>-4</sup>	0.0100
CO	0.8897	2.2198x10 <sup>-3</sup>	0.0210
CO <sub>2</sub>	0.9485	2.6103x10 <sup>-3</sup>	0.0204
NO	0.8173	8.7843x10 <sup>-3</sup>	0.0464
NO <sub>2</sub>	0.8449	1.0087x10 <sup>-2</sup>	0.0500
SO <sub>2</sub>	0.8775	1.8808x10 <sup>-3</sup>	0.0204

The prediction accuracy for CO, SO<sub>x</sub> and NO<sub>x</sub> concentrations are the lowest due to specific reasons. The fuel characterization showed that there are only traces of nitrogen (0,08%*m/m*) and sulfur (0,0529%*m/m*) compounds, with concentrations that are at least one thousand times smaller than the carbon concentration (86,22%*m/m*). Consequently, the patterns of SO<sub>x</sub> and NO<sub>x</sub> compounds for the soft sensor learning phase are limited to values that are close to the hard sensor's minimum range. In addition, combustion NO<sub>x</sub> is generated as a function of air to fuel ratio and is more pronounced when the mixture is on the fuel-lean side of the stoichiometric ratio (EPA, 1999). Thermal NO<sub>x</sub> formation is only significant for combustion temperature above 1,300°C (2,370°F) (EPA, 1999). The boiler nominal operating conditions consist of great excess air that produces an air-fuel mixture rich in oxygen and reduces the combustion chamber temperature.

The CO accuracy prediction of 0,8897 is also explained by the boiler process operation with large amounts of excess air that range from 10 to 30% and results on fewer CO samples for the

learning process due to complete combustion of the fuel. Thus, emission levels are concentrated on a minimum span region (up to 20%).

Nonetheless, the overall calibration performance was acceptable as individual components MSE indices such as O<sub>2</sub> e CO<sub>2</sub> were 10 times smaller than the accuracy of the hard sensor stated by the manufacturer. The final global performance index (MSE) was  $4.3414 \times 10^{-3}$  and the coefficient of determination ( $r^2$ ) was 0.9490.

Since the calibration tool and the prediction tool were implemented in a unique framework, as the model was updated by the calibration routine it was immediately used in a real-time application at the pilot plant. To enhance performance visualization of the online validation of the recalibration procedure the online gas analyzer target variables measurements and the soft sensor's prediction output were collected and compared. These outputs (actual measurement values and predicted values) were compared in time trends, one for each component of interest.

After each recalibration cycle (recalibration 1 and recalibration 2) the model's prediction ability improved. Prediction for O<sub>2</sub> (Figure 5) and CO<sub>2</sub> (Figure 6) concentrations for the application phase are exhibited. They are the controlled variables in typical combustion control loops. The soft sensor was able to estimate concentration values continuously meanwhile the hard sensor did not measure the interest variables during the purge procedure. Purge is evidenced by the abrupt change in the process variable measured, presenting static values equal to a burnout situation for the specific analyzer channel.



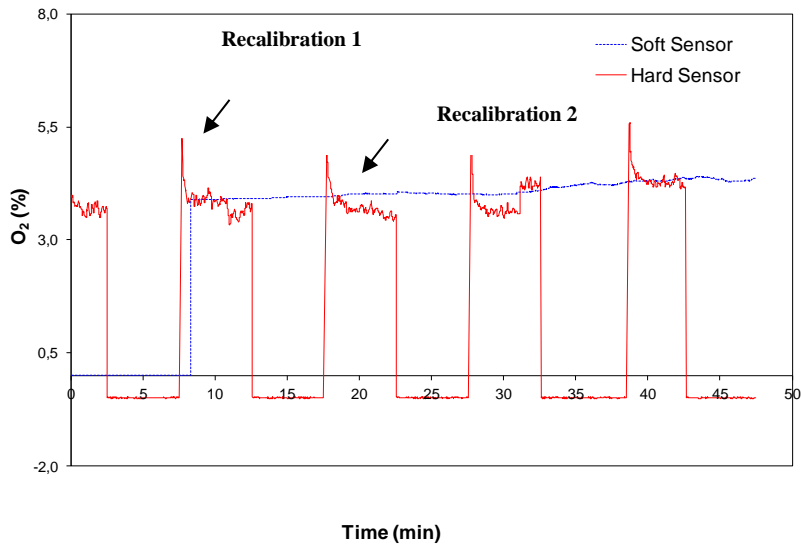


Figure 5. Online validation of the soft sensor for  $O_2$  prediction.

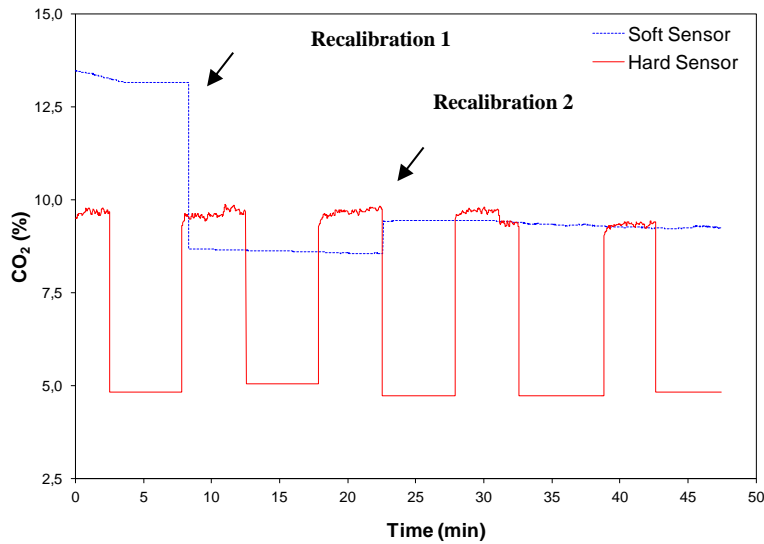


Figure 6. Online validation of the soft sensor for  $CO_2$  prediction.

## 6. Conclusion

The soft sensor calibration tool was successfully developed. The algorithm merges different software platforms in a framework that includes consolidated mathematical solutions and industrial data acquisition systems through a friendly human-machine interface. Electronic

sheets and SCADA systems are most commonly part of the automation systems installed in chemical and biochemical industrial process plants.

The developed calibration tool can perform a model maintenance procedure automatically in less than 1 minute, including the data pre-processing steps - a typically time-consuming manual procedure. Also, the mathematical function of the window size on a time basis enabled the collection of enough patterns straight from the raw process data in an online manner.

The calibration routine is initialized onsite by operators: periodically, as a preventive maintenance procedure or as a result of a systematic noticed prediction failure.

The main points of the development were tested and evaluated: initialization custom options, dialog boxes, filters and prediction capability. The calibration tool was hence validated through experimental tests in a semi-industrial pilot plant. The overall  $r^2$  performance was above 80%. An alternative expression for performance assessment that includes not only the standard mean square error but also the slope of the straight line that relates predicted and measured values is suggested.

Even though the framework was implemented in a combustion process it can be generalized for different soft sensors and distinct processes once the initialization steps can be set for other variables and independent ANN topologies. The customized options include different training methods, performance goals, adaptable filtering parameters and adjustable data collection that may or may not incorporate previous training samples and old weights and biases values as initial guess. Such features contribute for a flexible calibration tool.

## **7. Acknowledgments**

The authors gratefully acknowledge support from the Federal University of Rio de Janeiro for providing software licenses of Matlab® (C Algorithm), Excel® (Electronic Sheet embedded with VBA) and Ifix Proficy® (SCADA).

## 8. Reference

- Andrijić, Ž. U.; Cvetnić, M.; Bolf, N. Soft Sensor Models for a Fractionation Reformate Plant Using Small and Bootstrapped Data Sets, v. 35, p. 745-756, 2018.
- Beale, M. H.; Hagan, M. T.; Demuth, H. B. Neural Network Toolbox – User's Guide. The Mathworks. [S.l.]. 2012.
- Cao, P.; Luo, X. Modeling for soft sensor systems and parameters updating online. Journal of Process Control, v. 24, p. 975-990, 2014.
- Chen, K., Castillo, I., Chiang, L. H. & Yu, J., Soft Sensor Model Maintenance: A Case Study in Industrial Processes, IFAC-PapersOnLine, 48(8), 427–432 (2015).
- Clercq, D. D. *et al.* Machine learning powered software for accurate prediction of biogas production: A case study on industrial-scale Chinese production data. Journal of Cleaner Production, v. 218, p. 390-399, 2019.
- EPA. Nitrogen Oxides (NO<sub>x</sub>), Why and How They Are Controlled. Clean Air Technology Center. [S.l.], p. 57. 1999.
- Fortuna, L., Soft sensor for monitoring and control of industrial processes. Springer, Londres (2007).
- Grbić, R., Slisković, D. & Kadlec, P., Adaptive soft sensor for online prediction and process monitoring based on a mixture of Gaussian process models, Computers and Chemical Engineering, 58, 84-97 (2013).
- Kamat, S.; Madhavan, K. Developing ANN based Virtual / Soft Sensors for Industrial Problems. IFAC- PapersOnLine, v. 49(1), p. 100-105, 2016.
- Jin, H., Chen, X., Yang, J., Wang, L., Wu, L., Online local learning based adaptive soft sensor and its application to an industrial fed-batch chlortetracycline fermentation process, Chemometrics and Intelligent Laboratory Systems, 143, 58-78 (2015).
- Kadlec, P., On robust and adaptive soft sensors, AIChE Journal (2009).
- Kadlec, P., Gabrys, B., Adaptive on-line prediction soft sensing without historical data, Neural Networks (IJCNN), (2010).
- Kadlec, P., Gabrys, B., Strandt, S., Data-driven soft sensors in the process industry, Computers and Chemical Engineering, 33, 795-814, (2009).
- Kadlec, P., Grbić, R., Gabrys, B., Review of adaptation mechanisms for data-driven soft sensors, Computers and Chemical Engineering, 35, 1-245, (2011).
- Kaneko, H., Funatsu, K., Classification of the degradation of soft sensor models and discussion on adaptive models. AIChE J, 59, 2339-2347, (2013).

Kano, M., Ogawa, M., The state of the art in chemical process control in Japan: Good practice and questionnaire survey, *Journal of Process Control*, 20 , 969–982, (2010).

Lin, B., Recke, B., Knudsen, J. K. H. & Jørgensen, S. B., 2007. A systematic approach for soft sensor development. *Computers and Chemical Engineering*, Volume 31, p. 419–425.

Lu, B.; Chiang, L. Semi-supervised online soft sensor maintenance experiences in the chemical industry. *Journal of Process Control*, v. 67, p. 23-34, 2018.

Lu, B., Stuber, J., Edgar, T. F., Integrated online virtual metrology and fault detection in plasma etch tools, *Industrial Engineering Chemical Research*, 53, 5172–5181, (2014).

Maier, H. R., Dandy, G. C., The effect of internal parameters and geometry on the performance of back-propagation neural networks: an empirical study, *Environmental Modelling and Software*, 13, 193-209, (1998).

Neuralware, Support: Frequently Asked Questions, [Online], (2010). Available at: [http://www.neuralware.com/support\\_faqs.jsp](http://www.neuralware.com/support_faqs.jsp) [Accessed 30 05 2010].

Parkinson, T.; Parkinson, A.; Dear, R. Continuous IEQ monitoring system: Context and development. *Building and Environment*, v. 149, p. 15-25, 2019.

Santos, B. F. D. *et al.* Monitoring of Biosurfactant Production by *Bacillus subtilis* Using Beet Peel as Culture Medium via the Development of a Neural Soft-Sensor in an electronic spreadsheet. *Brazilian Journal of Chemical Engineering*, v. 35, p. 1678-4383, 2018.

Souza, F. A. A., Araújo, R., Mendes, J., Review of soft sensor methods for regression applications, *Chemometrics and Intelligent Laboratory Systems*, 152, 69-79, (2016).

Tambourghi, E. B.; Fischer, G. A.; Fileti, A. M. F. Neural modeling for cytochrome b5 extraction. *Process Biochemistry*, v. 41, p. 1272–1275, 2006.

Tao, X. *et al.* Greenhouse gas emission monitoring system for manufacturing prefabricated components. *Automation in Construction*, v. 93, p. 361-374, 2018.

Valdman, A., Folly, R., de Souza Jr, M. B., Valdman, B., A systematic Methodology on Developing an Online Soft Sensor Based on Neural Networks for Monitoring Boiler Gas Emissions, *AIDIC Conference series*, 10, 353-361, (2011).

Warne, K., Prasad, G., Rezvani, S., Maguire, L., Statistical and computational intelligence techniques for inferential model development: a comparative evaluation and a novel proposition for fusion, *Engineering Application of Artificial Intelligence*, 17(8), 871-885, (2004).

Zheng, W. *et al.* Just-in-time semi-supervised soft sensor for quality prediction in industrial rubber mixers. *Chemometrics and Intelligent Laboratory Systems*, v. 180, p. 36-41, 2018.

## Capítulo VI- Estratégia de Controle Tolerante a Falha Baseada em Sistemas de Diagnóstico de Falhas

---

Este capítulo resume o artigo “Enhanced Control Strategy of a Semi-industrial Boiler Based on Failure Diagnosis Systems: A Study Case on Hybrid Automation Systems”, publicado nos anais de International Conference on Modelling, Simulation and Identification.

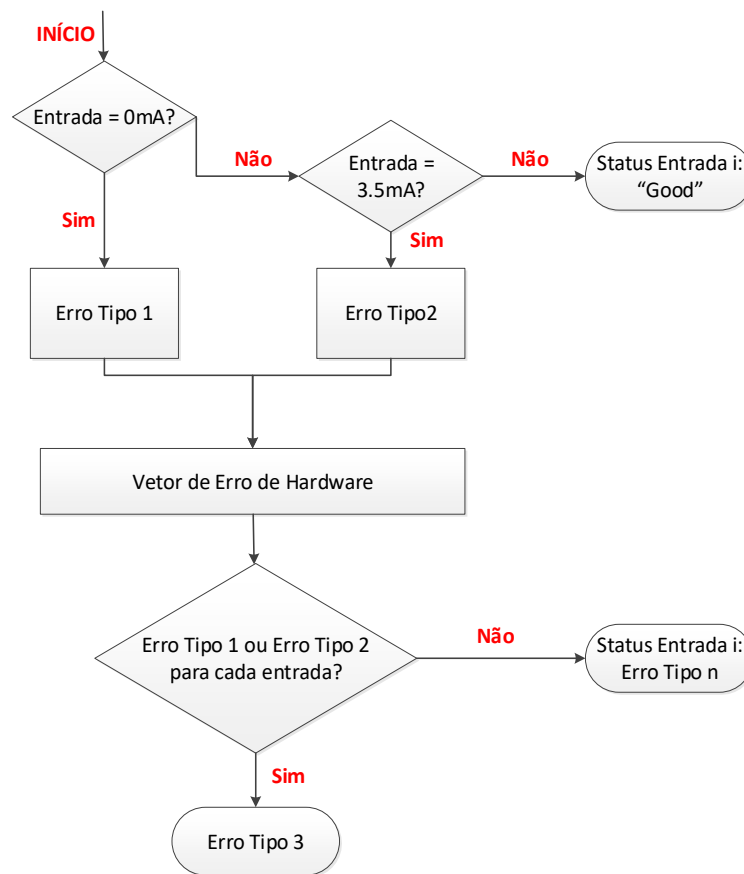
### VI.1 Resumo

Apesar da importância do controle automático de processos, muitas malhas de controle industrial operam aquém do desempenho esperado. Este estudo teve como objetivo implementar uma estratégia de controle aprimorado em uma caldeira semi-industrial a fim de minimizar os riscos operacionais devido a falhas de sensores e gerenciar os níveis de emissões de poluentes atmosféricos. O trabalho foi desenvolvido em uma planta piloto que faz parte das instalações da central de utilidades do Laboratório de Engenharia Química (LADEQ) da Universidade Federal do Rio de Janeiro. As aplicações experimentais foram realizadas em uma planta composta por uma caldeira flamotubular vertical que gera até 500 kg/h de vapor de água e um sistema de automação híbrido, baseado em dispositivos analógicos, redes industriais digitais *Fieldbus* e sistemas de supervisão e aquisição de dados (SCADA). A malha de gases de combustão controla os níveis de emissão de CO<sub>2</sub>, medidos por um analisador de gás *in situ*, e manipula o fluxo de ar de combustão, atuado por um soprador de ar associado a um inversor de frequência. Uma estrutura de controle seletiva auxiliar foi desenvolvida para ampliar as capacidades do laço de controle contínuo padrão. Assim, uma estratégia de controle de realimentação PI tolerante a falhas (FTC) foi implementada e uma estrutura independente de diagnóstico de detecção de falhas (FDD) foi desenvolvida para identificar e isolar as falhas que fazem parte da estratégia da FTC.

### VI.2 Metodologia Sumarizada

A ferramenta FDD é uma estrutura auxiliar para ajudar a operação a identificar e isolar falhas antes que elas evoluam para defeitos. Os tipos de falha são apresentados na interface homem-máquina da estação SCADA de maneira que ações corretivas possam ser tomadas enquanto a estratégia de controle tolerante a falha (FTC) mantém operação segura e confiável em malha fechada.

A ferramenta FDD (Figura 27) é processada no CLP, que funciona como interface entre o analisador de gases (4-20 mA) e a rede industrial digital (*Fieldbus*). O primeiro passo para a detecção e o diagnóstico da falha consiste em checar se existem entradas elétricas desconectadas identificadas por leituras de 0 mA. Este é o erro tipo 1, rotulado como “erro de comunicação”, que é provavelmente causado por fio solto ou mau contato. Então, cada variável de entrada do analisador de gases é escaneada para verificar se o seu valor está abaixo do valor de fundo de escala (3,5 mA), caracterizando um erro tipo 2 ou de leituras “fora da faixa”. Quando todas as entradas estão fora de escala o erro tipo 3 é diagnosticado e identificado como “erro de detecção de *hardware*”.



**Figura 27 Fluxograma de decisão FDD**

A estratégia de controle FTC (Figura 28) é do tipo ativa, baseada em reconfiguração *online*. Os canais 1 e 2 do analisador de gases transmitem a mesma variável medida,

concentração %v/v de CO<sub>2</sub>, para emular instrumentos de medição redundantes. A estrutura seletiva auxiliar decide pelas medições válidas, estendendo as capacidades da malha de controle de realimentação PI padrão. Assim, se a variável de processo medida e transmitida pelo canal 1 não for confiável, a ferramenta FDD detecta a situação de falha e a variável controlada pela realimentação muda automaticamente para a variável medida pelo próximo canal.

Para validar a metodologia foram realizados testes em que uma falha de sensor do tipo 1 foi provocada soltando o fio do canal 1. Estes experimentos foram feitos enquanto o processo era submetido a uma perturbação do tipo degrau no *set-point* de 1%v/v.

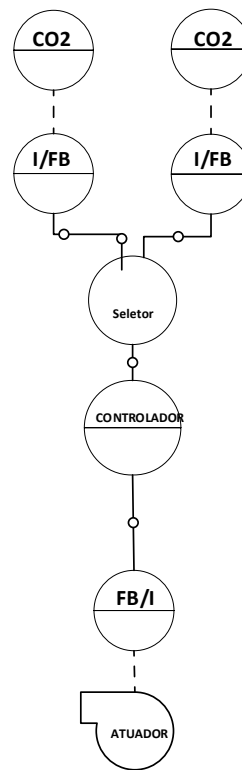


Figura 28 Diagrama em blocos da estratégia AFTC

### VI.3 Principais Resultados e Discussões

A ferramenta FDD foi capaz de perceber, isolar e indicar a falha criada, exibindo esta informação no sinóptico de operação do sistema SCADA. As Figuras 29 e 30 comparam a operação sob uma estratégia de controle clássica e não-tolerante a falha com a estratégia tolerante a falha proposta. Na Figura 29, sob controle clássico, a falha de sinal (em  $t = 600s$ ) interrompeu a operação automática e a malha de controle passou a operar no modo manual com vazão de ar de combustão fixa. Isto poderia aumentar a variabilidade da variável controlada e levar o processo

para uma região operacional de baixa eficiência térmica. Já na Figura 30, sob controle FTC, a estratégia muda automaticamente e instantaneamente para um sinal de medição confiável, garantindo operação segura e eficiente que, portanto, é capaz de gerenciar a quantidade de gases residuais da combustão.

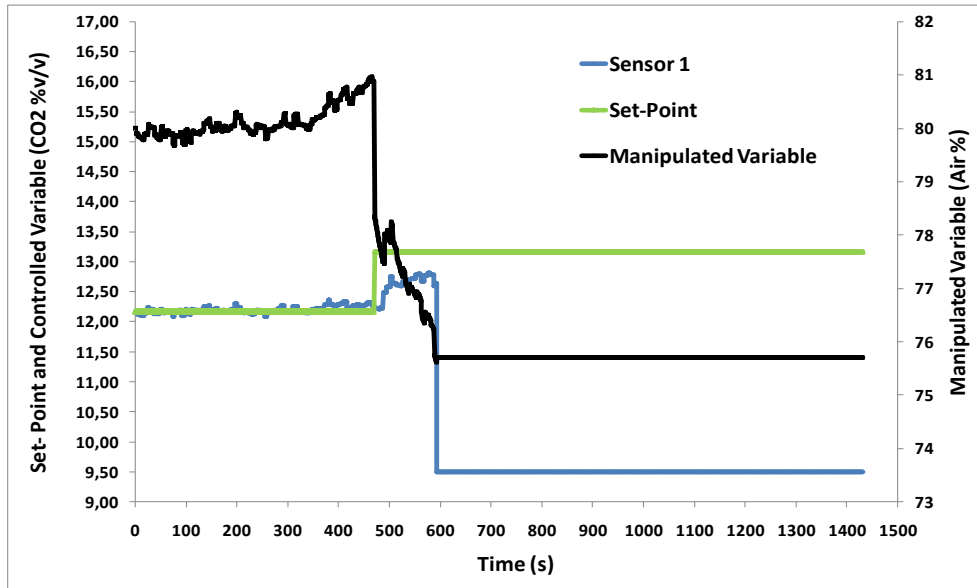


Figura 29 Resposta dinâmica a uma perturbação no set-point sob controle clássico

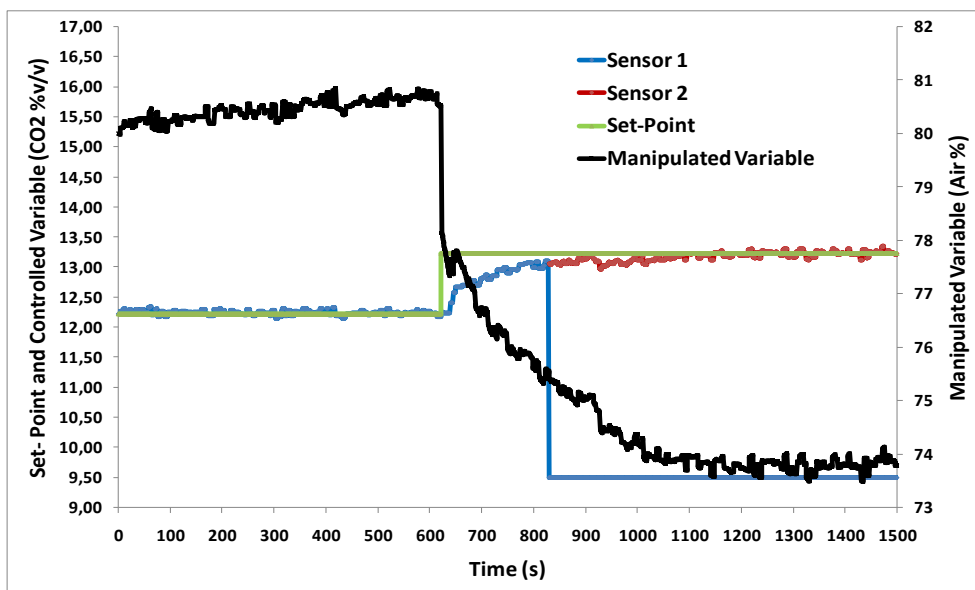


Figura 30 Resposta dinâmica a uma perturbação no set-point sob controle AFTC



#### **VI.4 Lista das Principais Contribuições**

- Integração de tecnologias analógicas e digitais em um sistema de automação híbrida
- Ferramenta FDD
- Estratégia de controle FTC ativa com reconfiguração *online*
- Operação em malha fechada confiável, flexível e ecoeficiente
- Gerenciamento das emissões atmosféricas

#### **VI.5 Artigo Anexo**

# ENHANCED CONTROL STRATEGY OF A SEMI-INDUSTRIAL BOILER BASED ON FAILURE DIAGNOSIS SYSTEMS: A STUDY CASE ON HYBRID AUTOMATION SYSTEMS

Andréa Pereira Parente, Andrea Valdman, Mauricio Bezerra de Souza Jr., Rossana Odette Mattos Folly  
Federal University of Rio de Janeiro (UFRJ)  
Av. Horacio Macedo, 149, Centro de Tecnologia, Bl E/sl. E-209, Rio de Janeiro, 21949-900/Brazil  
parente@eq.ufrj.br, avaldman@eq.ufrj.br, mbsj@eq.ufrj.br, rossana@eq.ufrj.br

## ABSTRACT

This study aimed at implementing an enhanced control strategy in a semi industrial boiler concerning the pollutants emission management in order to minimize operational risks due to sensor failures. The proposed control strategy is fault tolerant and it uses failure diagnosis systems in association with closed-loop online monitoring. The current work was developed based on a pilot plant that is part of the utility facilities installed in the Chemical Engineering Laboratory (LADEQ) at School of Chemistry, Rio de Janeiro Federal University. The experimental applications were carried out in a plant consisting of a vertical fire-tube boiler which generates 500 kg/h of water steam and a hybrid automation system based on analog devices and industrial Fieldbus networks. The flue gas loop controls the exhaust gas emission levels of CO<sub>2</sub> measured by an *in situ* gas analyzer manipulating the flow of combustion air. An auxiliary selective control structure was developed to extend the capabilities of continuous standard control loop. Thus, a switch selector automatically decides which of the redundant values measured by a commercial gas analyzer is reliable.

## KEY WORDS

FTC FDD automation boiler control semi-industrial instrumentation experimental application study case combustion process pilot plant CO<sub>2</sub>

## 1. Introduction

Despite the importance of process control, many industrial control loops operate poorly. Industrial surveys report that more than 60% percent of controllers provide faulty performance, leading to bad product quality and loss of production. Even more alarming data state that at least 30% of control loops increase the variability of the process variable compared to the use of manual control, and another 36% of processes operate in open loop [1]. In order to achieve production and performance goals, process industries have been increasing the use of

online and real time instrumentation technology within their automation facilities. Fault diagnosis analysis and intelligent framework approaches are some of the modern issues arisen. Online supervision tasks capable of following the process behavior in real time, including fault detection and diagnosis (FDD) systems may be used in association with closed-loop monitoring, providing tools to aid plant personnel in identifying poorly performing loops and suggesting remedial actions which may include controller retuning or online reconfiguration.

Combustion based equipment are widely integrated in industrial complexes, where fossil fuels are still dominant [2]. Boiler units are critical in chemical processes because the steam flow rate, temperature and pressure, affect the performance of the whole process plant. Furthermore, systematic studies have been conducted with the aim of reducing levels of air pollution that arise out of the combustion processes [3] focusing on environmental concerns and efficiency improvements [4] [5].

According to Dochain, Marquardt *et al.* (2006) stricter environmental laws and workers protection as well as energy saving measures call for a more widespread use of on-line monitoring and reconfiguration systems.

This paper presents a study case wherein an active fault tolerant control (FTC) feedback PI strategy is implemented in a semi-industrial boiler plant. The proposed strategy considers fault detection analysis and integrates industrial instruments based on Fieldbus foundation communication equipment and SCADA software. The loop controls the exhaust gas emission levels of carbon dioxide and manipulates the flow of combustion air. This work also presents an independent fault detection diagnosis (FDD) framework developed to identify and isolate the faults that are part of the FTC strategy.

## 2. The Pilot Plant

The current work was developed based on a pilot plant that is part of the utility facilities installed in the

Chemical Engineering Laboratory (LADEQ) at School of Chemistry, Rio de Janeiro Federal University. Concisely, the plant has a feed water tank, four fuel storage tanks, a fuel mixture and feed tank, a semi-industrial vertical fire-tube boiler that produces saturated vapor. The boiler has a fixed fuel flow rate of 19kg/h and a regulated damper associated with a frequency inverter that allows air flow rate adjustments. Table 1 shows the boiler's technical specifications.

Table 1 - Boiler Technical Specifications

Boiler model	CV-
Steam production	500
Steam	satur
Nominal heating	0.3
Maximum working	8.00k
Main fuel	diesel
Fuel flow rate	~19k

The automation system has Fieldbus devices for measuring continuous process variables, a PLC controller for converting discrete-time variables and analog signal into Fieldbus protocol, a dedicated controller which operates interlock logic and a SCADA station to register engineering values and diagnosis status.

The main continuous process variables available are fuel temperature and feed tank level; feed water temperature, tank level and feed water flow rate; boiler water level, steam production flow rate and pressure; flue gas emission temperatures. Exhaust emission concentrations which include Carbon Dioxide are continuously analyzed by a Testo 350XL flue gas analyzer.

### 3. FDD Framework

The developed fault detection and diagnosis framework is an auxiliary structure intended for aiding plant personnel in identifying and isolating faults before they develop into operational failures. The available fault types are presented in the SCADA interface so that operational remedial actions can be taken meanwhile the FTC strategy (section 4) ensures safety and reliable closed loop operation.

The FDD's hybrid architecture automation includes a PLC that is the electronic interface between the commercial gas analyzer (4-20mA measuring data) and the digital network. In order to minimize data traffic and maximize speed in the framework, the FDD system is processed in the PLC. The logic operations are performed by Ladder Logic programming language using contact and conductor

arrangements, reproducing electrical connections, along with numerical operations performed using function blocks.

The first step of fault detection and diagnosis consists in checking whether there are disconnected inputs (0 mA readings). This verification is done by means of a comparison block that indicates to the SCADA system that there is a type 1 error, identified by "communication error" probably caused by wire loose or intermittent contact.

Then each input is checked to determine if its value is below full scale (3.5 mA), characterizing a type 2 error that is related to "out of range" readings. When all the inputs are out of range a type 3 error is identified as "hardware error detection". Error type 1 shall not interfere in hardware error diagnosis.

The fault detection and diagnosis logic can be seen in the flowchart shown in Figure 1.

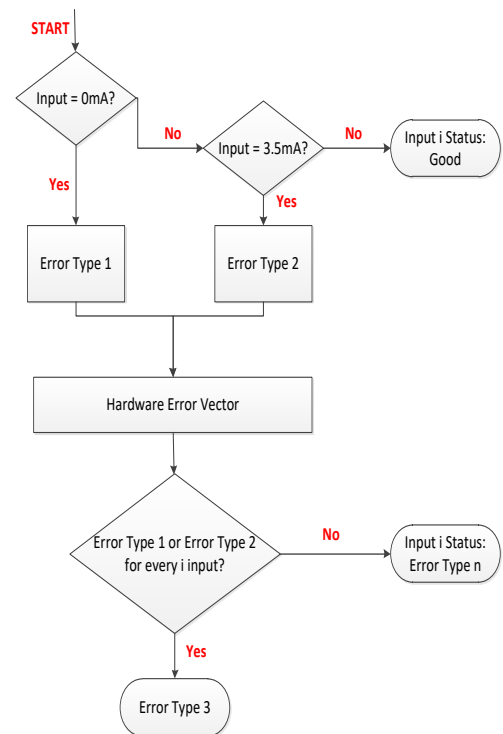


Figure 1 FDD logical decision flowchart

Table 2 summarizes the faults that can be recognized.

Table 2 - Error types in the FDD system

Error Type	Diagnosis Status
1	communication error
2	value out of range
3	hardware error

Relevant changes were made in the SCADA station so that the human-machine interface exhibits the FDD system. The process variable values read from the analyzer are displayed, using different colors to indicate each error type (blue symbolize normal operation; red symbolizes specific input or channel error).

#### 4. FTC Strategy

Fault tolerant control may be distinguished in passive and active methods. For the first ones, the effect of faults on the process behavior is described as modeling uncertainties which leads to robust and sometimes sluggish controllers. Active methods consist in performing online reconfiguration based on the information provided by a diagnostic system [6]. An active FTC strategy was proposed for the waste gas emission control loop wherein a PI feedback controller was implemented. The strategy includes fault diagnosis analysis to promote online reconfiguration and presents a hybrid architecture automation that integrates industrial instruments based on Fieldbus foundation communication equipment, analogic devices and SCADA software. Channels 1 and 2 of the commercial gas analyzer transmit the same process variable, Carbon Dioxide concentration in %v/v, in order to simulate redundant measuring devices. Thus, the auxiliary selective control structure can decide to use valid measured values extending the capabilities of continuous standard control loop. Figure 2 presents the instrumentation diagram, including main FB function blocks.

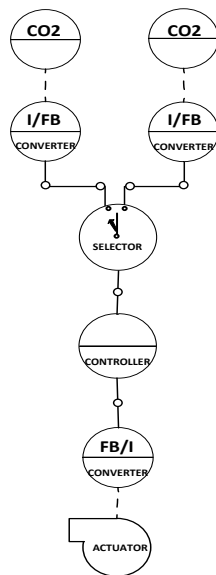


Figure 2 Instrumentation Diagram

The choice of the control loop in operation can be done manually by the operator or automatically by the selection criteria configured in Fieldbus instrumentation itself. The selection criterion adopted and implemented by means of the configuration software was “first input signal classified as good”; in accordance with the fault diagnosis detection framework described in section 3.

Thus, if the process variable measured in the first input channel is not reliable, the FDD Framework detects a faulty situation and the feedback loop controlled variable will automatically change for the process variable measured in the next channel. If both measuring elements are functioning properly the prioritization configuration suggests that the first channel is preferred.

The developed FTC strategy was hence configured in an industrial Fieldbus device (FI301, SMAR) which converts the digital signal from the controller to analogic signal to the frequency inverter, the loop actuator. The controlled variable is the residual concentration of carbon dioxide in the exhaust station of a steam-raising boiler and the manipulated variable is the damper engine speed.

The loop configuration in the digital Fieldbus network is shown in Figure 3. The variables measured by the commercial gas analyzer are transmitted to a selector block. The selected signal is sent to the PID block, which compares the measured value with the desired value and transmits the calculated control signal to the actuator element. The parameters "BKCAL\_OUT" and "BKCAL\_IN" connected to prevent the phenomenon called reset windup. Thus, the integral action does not increase unlimitedly in circumstances in which the actuator element reaches saturation. The connection "TRK\_VAL" is also noteworthy. It is associated with the output tracking algorithm also called bumpless transfer which avoids abrupt changes in error when the control action is switched from manual to automatic.

#### 5. Experimental Validation

Experimental tests were conducted to validate the FTC strategy proposed and the FDD framework developed.

The methodology consisted of inducing a type 1 sensor failure to the first input channel. To enhance performance visualization, the fault was created while the process was submitted to a step change in set-point of 1%v/v. Controller tuning was made by means of on-site adjustments that are object of study of previous work [6].

The first test performed was intended to validate the FDD man-machine interface. Figure 4 shows the

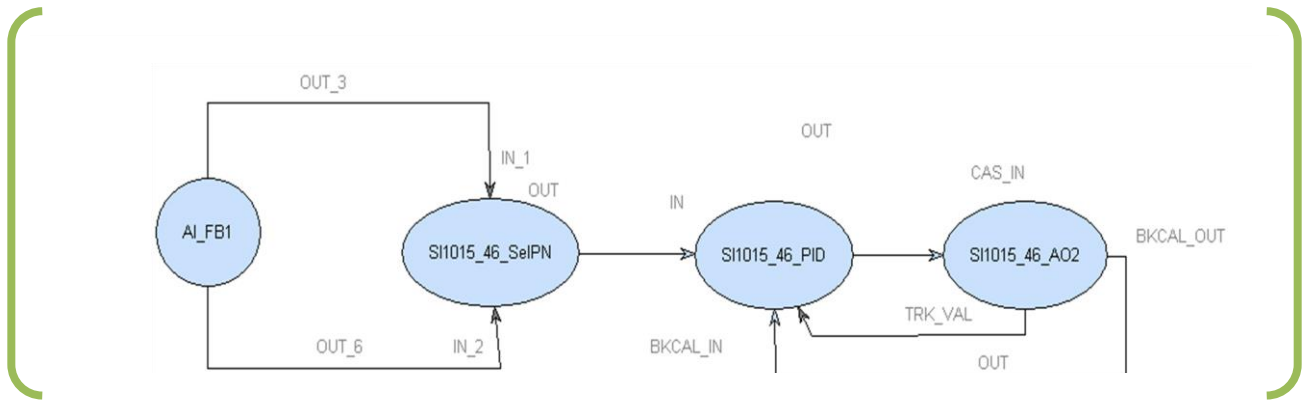


Figure 3 Proposed FTC strategy configuration in the digital Fieldbus network

abrupt change in the process variable measured by channel 1, indicating a malfunction such as a sensor drop or poor contact. The FDD framework was able to notice, isolate and indicate the created fault and Figure 5 shows how the SCADA exhibits this diagnosis information in the synoptic screen. Channel 1 CO<sub>2</sub> concentration value indicates the minimum range configured value and color red identifies an error. The second indication presents the correct CO<sub>2</sub> concentration value, measured by channel 2. Two other experiments were performed aiming to validate the FTC strategy proposed. The online reconfiguration performance result shown in Figure 7 is compared to non-fault tolerant control strategy subjected to the same induced error presented in Figure 6.

In Figure 6 faulty operation under non-tolerant automatic control is shown. Bad signal quality identified on time equal to 600s interrupts automatic mode changing the control loop to manual. Open loop operation with a fixed flow of combustion air could increase the controlled variable variability. This situation could lead the process to low combustion and thermal efficiency operational locus. Figure 7 reveals how the fault tolerant control strategy implemented automatically and instantaneously changed the measured variable that is part of the closed loop to a reliable signal ensuring a safe and efficient operation hence managing the amount of residual combustion gases.

### 6. Conclusion

Digital technology allowed self-validating sensors and actuators equipped with internal FDD systems to progressively integrate process industries. As information will be traveling over data networks, tolerance with respect to dropped or lost sensors shall be considered. Since analogic technologies are consolidated and tend to be continuously used, hybrid architecture automation systems must be considered. This work integrates digital and analogic technologies leading to a hybrid automation system with active fault tolerant control loop. A fault detection and diagnosis framework able to detect faults before they develop into failures was developed providing the information needed to include online reconfiguration mechanisms able to take appropriate measures to reach safe operating mode, hence decreasing the risk of pollution or casualties. The proposed control strategy includes prioritization mechanisms, fault detection and diagnosis and online reconfiguration, creating reliable plant operation. The high degree of flexibility induced by the developed system allows quick adjustments, ensuring an energy-efficient operation, minimizing water consumption and reducing pollutant emissions.

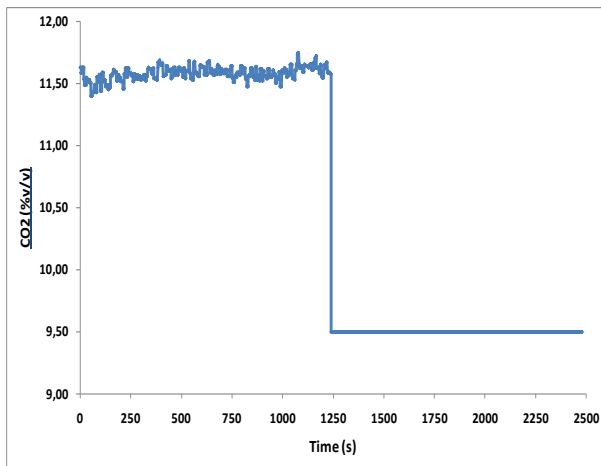


Figure 4 Abrupt status change in measured/controlled variable CO<sub>2</sub> (%v/v)

CO <sub>2</sub>	9,50	%
CO <sub>2</sub>	12,52	%

Figure 5 FDD in the SCADA system: man-machine interface

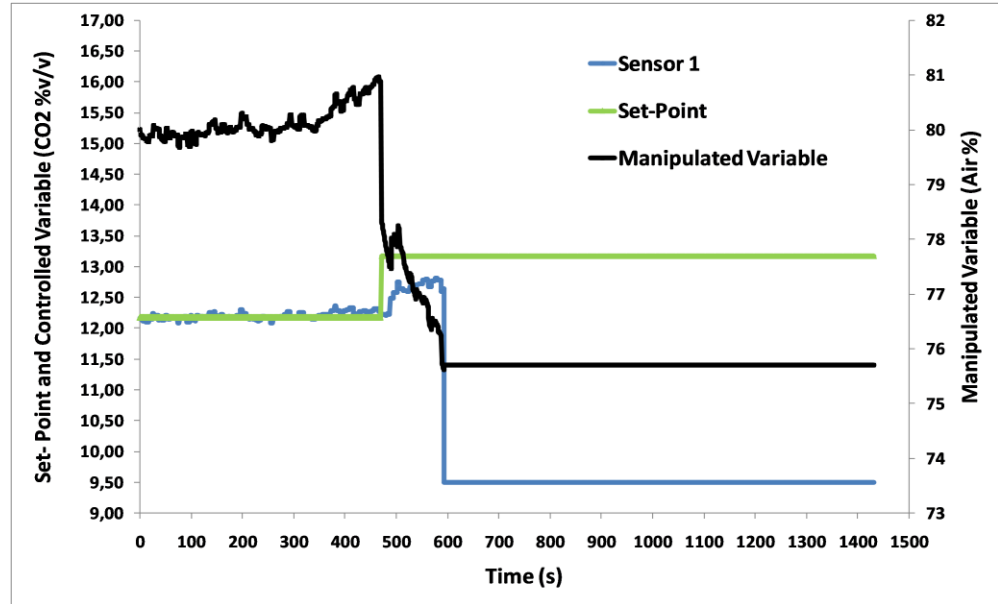


Figure 6 Induced faulty operation under non-tolerant automatic control

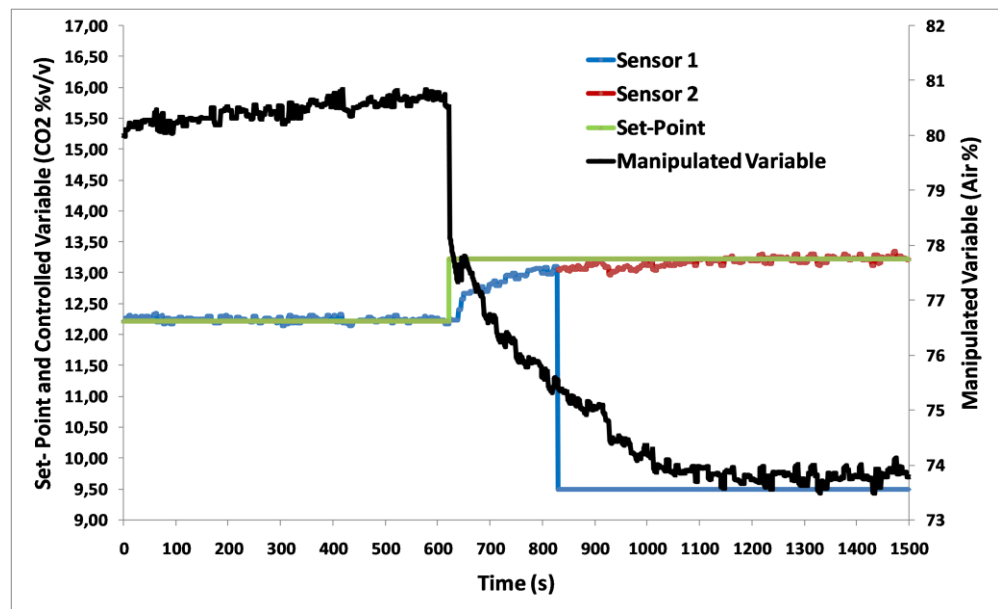


Figure 7 Induced faulty operation under active fault tolerant automatic control

## References

- [1] L. Desborough, R. Miller, Increasing customer value of industrial control performance monitoring-Honeywell's experience, Sixth International Conference on Chemical Process Control AICHE, 2002.
- [2] T. Drapela, M. Pavlas, P. Popela, J. Boran, P. Stehlik, Energy conception of an integrated system - Analysis of available data and its processing, Chemical Engineering Transactions, 18, 2009, 635-640.
- [3] A.D. Sanches, M.C.M Perez, G. Alfonso, E.R. Dominguez, I. Herrera, V. Nunez, Diagnosis of the air quality in a zone affected by combustion gases sources, Chemical Engineering Transactions, 21, 2010, 199-204.
- [4] I. Hippinen, P. Houhonen, L. Silvill, J. Federley, J. Hakala, J. Manninen, P. Ahtila, Methods for industry to measure and improve the energy efficiency of utility systems, Chemical Engineering Transactions, 21, 2010, 349-354.
- [5] A. Valdman, B. Valdman, R.O.M. Folly, M.B. Souza Jr, Online Monitoring of Exhaust Gas Emissions of a Boiler with Diesel/Biodiesel Fuel Blends, Chemical Engineering Transactions, 24, 2011, 277-282.
- [6] A. P. Parente, Estratégia de controle de emissão de gases residuais aplicados a caldeiras de pequeno porte, Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Tecnologia de Processos Químicos e Bioquímicos, EQ, Universidade Federal do Rio de Janeiro, 2014.
- [7] D. Dochain, W. Marquart, S.C. Won, O. Malik, M. Kinnaert, Monitoring and control of processes and power systems: Towards new paradigms Status report prepared by IFAC Coordinating committee on Process and Power Systems, Annual Reviews in Control, 30, 2006, 69-79.

**PARTE III - APLICAÇÕES EM PROCESSOS DE  
COMBUSTÃO INDUSTRIAIS**



## Capítulo VII– Caldeira Industrial de Recuperação Kraft

---

Este capítulo apresenta o processo de produção de uma importante fábrica de Papel e Celulose brasileira, cujo nome foi omitido por questões de confidencialidade.

### VII.1 Processo de Produção Kraft

O processo de obtenção de polpa de celulose pelo processo Kraft pode ser dividido em dois ciclos principais: a deslignificação da madeira e a recuperação de reagentes.

O princípio básico da deslignificação da madeira pelo processo de produção Kraft é a digestão, ou seja, o cozimento da matéria-prima celulósica sob pressão com auxílio de sulfeto de sódio, hidróxido de sódio, além de outros sais de sódio. A mistura de reagentes alimentada no digestor recebe o nome de licor branco.

Ao final do cozimento, tem-se uma polpa marrom que passará por diversas etapas de processamento. As principais correntes obtidas serão a polpa celulósica lavada que seguirá para a etapa de branqueamento, e o licor negro, que será concentrado por evaporação e alimentado na caldeira de regeneração Kraft. O licor negro é constituído por reagentes não consumidos, lignina retirada da madeira, matéria orgânica derivada da degradação da lignina e outros elementos residuais do cozimento (SILVA, 2019).

O licor negro concentrado, chamado de licor forte, é pré-aquecido e alimentado em uma caldeira de recuperação. Na caldeira, a matéria orgânica, como, por exemplo, a lignina, é queimada para geração de vapor para uso na planta enquanto os sais inorgânicos são recuperados para reciclo (SILVA, 2019).

Os reagentes químicos são fundidos na caldeira e recuperados, sendo conduzidos a um tanque de dissolução para obtenção do licor verde. O licor verde, rico em carbonato de sódio e sulfeto de sódio (CASTRO, 2009), é levado para caustificadores. O líquido resultante da etapa de caustificação é o próprio licor branco, composto de hidróxido de sódio e sulfeto de sódio, que é bombeado para a etapa de digestão.

A polpa celulósica lavada, que ainda contém até 5% de lignina (CASTRO, 2009), segue para uma etapa de branqueamento, ou seja, um tratamento físico-químico para a continuação da retirada de lignina usando agentes alvejantes, para obtenção de uma polpa estável.

Após o branqueamento, pode ser necessária a remoção de impurezas em uma etapa suplementar de depuração em centrífugas. Caso a pasta seja o produto final da unidade industrial em questão, é realizada uma etapa de secagem compreendendo uma drenagem ou uma prensagem, seguida da passagem por secadores a vapor ou a ar quente. Caso o papel seja o produto final da unidade, incluem-se etapas de refino e desfibrção seguidas do preparo da receita final em uma mesa plana com operações de formação, prensagem, secagem, calandra e enrolamento.

## **VII.2 Caldeira de Recuperação Kraft**

O licor negro, subproduto do digestor de deslignificação, é uma solução altamente alcalina. Ele é obtido, como visto anteriormente, após o processamento da polpa marrom que separa o licor negro da polpa de celulose. O licor negro tem composição complexa: além de conter a lignina e os açúcares extraídos da madeira, material orgânico combustível, apresenta os sais inorgânicos, que resultam da reação de deslignificação. As propriedades e a composição do licor negro variam em cada planta de acordo com as condições operacionais estratégicas de cada empresa. Essas diferenças são atribuídas, principalmente, aos seguintes fatores: condições operacionais do digestor, espécie da madeira utilizada, razão entre as quantidades de reagentes químicos alimentados ao digestor e a quantidade de madeira, propriedades do licor branco, rendimento da polpa e eficiência da lavagem do material descarregado do digestor.

A recuperação dos sais inorgânicos contidos no licor negro é crítica para reduzir os custos e o impacto ambiental do processo (GALADIANO, 2006). Para esta finalidade, foram desenvolvidas caldeiras de recuperação que permitem recuperar sulfeto de sódio e o carbonato de sódio. Com isso, o descarte dos subprodutos é reduzido e há um ganho importante em termos ambientais. Além da recuperação dos sais, a caldeira permite aproveitar o poder calorífico da lignina para produzir vapor, que pode ser utilizado para geração de energia elétrica e para suprir o consumo de vapor da própria planta.

### ***VII.2.1 Descrição Física***

As caldeiras de recuperação são essencialmente caldeiras aquatubulares de grande porte, em que a seção de combustão é bem diferente das caldeiras convencionais usadas para queima de combustíveis líquidos ou gasosos.

Uma caldeira de recuperação Kraft típica se divide basicamente em duas seções. Na primeira seção, denominada fornalha, ocorre a queima do licor negro. A segunda seção é uma região de transferência de calor por convecção, onde fica localizado um sistema de troca térmica para aproveitamento de calor, que compreende os seguintes componentes: superaquecedores, feixe convectivo, tubulões e economizadores. Entre as duas seções, é formado um estreitamento chamado “nariz”.

### **A Fornalha**

A fornalha tem diversas funções: é nela que se remove a umidade remanescente do licor negro e que se faz a combustão dos compostos orgânicos e a redução dos sais de sódio de forma a recuperar o sulfeto de sódio em forma fundida.

A queima do licor negro, mais precisamente a etapa de pirólise, se dá na fornalha da caldeira e acarreta a formação do chamado “*char*”, que consiste nos produtos sólidos residuais da pirólise do licor negro. O “*char*” é formado por uma mistura de material carbonáceo e de sais inorgânicos, que compreende, predominantemente,  $\text{Na}_2\text{CO}_3$ ,  $\text{Na}_2\text{S}$  e  $\text{Na}_2\text{SO}_4$ , e que se depositam no fundo da fornalha (GRACE, CAMERON e CLAY, 1985). Uma mistura de sais inorgânicos fundidos, chamada de “*smelt*”, também se encontra no fundo da fornalha e, em conjunto com o “*char*”, forma um leito ou “*char bed*”.

A fornalha é a região da caldeira com seção transversal retangular. A fornalha de uma caldeira de recuperação Kraft possui entradas para alimentação de licor negro concentrado, que se distribuem em suas quatro paredes, e injetores de ar de combustão localizados em diferentes alturas. A maioria das caldeiras de recuperação apresenta entradas de ar de combustão em três níveis diferentes, denominados alimentadores de ar primário, ar secundário e ar terciário.

No nível primário, o ar é alimentado na parte inferior da fornalha, pouco acima da superfície do leito (ou “*char bed*”). As entradas de ar primário são localizadas nas quatro paredes e fornecem oxigênio para a queima de material proveniente das partes superiores da fornalha. Nessa região, tem-se predominantemente um ambiente redutor.

O ar primário tem por função queimar o “*char*”, da superfície do leito formado no fundo da fornalha. O “*char*” é formado em etapas, conforme as partículas aspergidas do licor negro vão sofrendo queima. O fluxo de ar primário é responsável, também, pela topologia do leito de “*char*”, que favorece a recuperação dos sais fundidos. O ar injetado nessa região não penetra no

leito, e, por isso, ocorrem reações de redução dos sais inorgânicos, propiciando a formação de sulfeto de sódio. O piso da fornalha em algumas configurações é inclinado para facilitar a remoção dos sais fundidos do “*smelt*” (FORNETTI e LEGAULT, 1998). A massa de sais inorgânicos fundida, formada, predominantemente, de sulfeto de sódio ( $\text{Na}_2\text{S}$ ) e carbonato de sódio ( $\text{Na}_2\text{CO}_3$ ), é vazada pelo fundo da caldeira em um tanque de dissolução contendo água, obtendo-se assim o chamado licor verde, que compreende basicamente  $\text{Na}_2\text{CO}_3$ ,  $\text{Na}_2\text{S}$  e que, após a etapa de caustificação, é transformado em licor branco retornando ao processo de deslignificação.

No nível secundário, acima do nível primário, o ar é, em geral, alimentado através das paredes frontal e traseira da fornalha, para definir adequadamente os perfis de fluxos dos gases que circulam e a altura do leito de “*char*”. Nesta região, tem-se também, predominantemente, um ambiente redutor. A principal função do ar secundário é a queima dos gases que são formados no leito de “*char*”. Ele influencia o padrão de escoamento dos gases na fornalha e, conseqüentemente, o desempenho da combustão.

No nível terciário, acima do nível secundário, o ar é alimentado de forma semelhante ao ar secundário, em um ponto acima da entrada do licor negro. Nessa região, tem-se predominantemente um ambiente oxidante. A função do ar terciário é fornecer oxigênio suplementar para queima de gases e materiais arrastados pela corrente gasosa.

É inevitável que exista o arraste de pequenas partículas sólidas pelos gases de exaustão, além da formação de sais na parte superior da fornalha, que seguem o fluxo de gases de exaustão. Evidentemente, este arraste é indesejável, pois pode causar diversos problemas, tais como deposição na superfície externa dos tubos de água, chegando a causar bloqueio da passagem dos gases e redução da eficiência da troca térmica, além do custo de remoção destas partículas no separador eletrostático na saída da caldeira (CARMO, 2017).

### **Fluxo de água na caldeira**

O sistema de troca térmica consiste em diversos elementos tubulares com a função de gerar vapor a partir da água de alimentação de caldeira. A água é proveniente de uma unidade de desmineralização para que o teor de sais, que podem causar incrustações no equipamento, seja minimizado. No sistema de troca térmica da caldeira, a água segue um fluxo contrário em relação aos gases de exaustão. A água de alimentação é pré-aquecida em trocadores de calor e

introduzida nos economizadores, onde ocorre o aproveitamento final da energia contida nos gases de exaustão. A água segue, então, para o tubulão de vapor, ou tubulão superior.

O tubulão de vapor é um vaso horizontal que se localiza em um ponto alto da caldeira entre o superaquecedor e o economizador. Ele é o principal componente do sistema de água, por funcionar simultaneamente como reservatório da água e de vapor saturado. O seu nível, crítico para a segurança operacional, é mantido por meio de estratégias de controle e condições de intertravamento rigorosas. No Brasil, o controle automático de nível é exigido pela NR-13.

Abaixo do tubulão de vapor e paralelo a ele, encontra-se o tubulão de água, ou tubulão inferior. Os tubulões se comunicam por um feixe de tubos chamado de feixe convectivo.

Do tubulão de água saem diversos tubos que vão formar os diversos componentes do sistema na área da fornalha: as paredes de água, que envolvem a fornalha, o nariz, que separa a fornalha – em que predomina a transferência de calor por radiação – da seção dos tubos de água – em que predomina a transferência de calor por convecção - e as cortinas, que são tubos localizados na região dos superaquecedores. Todos os tubos de água (feixe convectivo, paredes de água, nariz e cortinas) recebem calor e geram um fluxo de água parcialmente vaporizada que retorna para o tubulão de vapor, onde a água líquida e o vapor se separam.

No topo do tubulão de vapor existem saídas de vapor que alimentam os superaquecedores, nos quais o vapor é superaquecido e enviado a seu destino na planta. No caso de caldeiras modernas, que operam a pressões elevadas, em geral, o vapor é enviado a turbinas para geração de energia elétrica. Já o vapor de média ou baixa pressão pode ser utilizado nos demais processos da planta.

### ***VII.2.2 Formação de Partículas***

O acúmulo de partículas pode ser causado tanto pela vaporização e posterior condensação de compostos alcalinos quanto por fragmentos de partículas de “*char*” que são arrastados pelo gás de exaustão. O progresso das incrustações é acompanhado, principalmente, pelo monitoramento das seguintes variáveis operacionais: queda de pressão, temperatura do gás de exaustão ao longo da caldeira e temperatura do vapor produzido (VAKKILAINEN, 2005). O material particulado depositado nas superfícies de troca térmica é retirado por ramonadores, que injetam vapor d’água. Porém, algumas incrustações podem se tornar resistentes a esse mecanismo

de remoção. Quando há parada de operação da caldeira, os ramonadores podem ejetar água líquida ao invés de vapor para lavagem.

Para que o acúmulo de material particulado ocorra sobre as superfícies de troca térmica, as partículas devem ser capazes de aderir às superfícies. Para isso, é necessário que haja uma certa quantidade de fase líquida. Assim, destacam-se quatro temperaturas importantes na deposição de materiais particulados: a dita “primeira” temperatura de fusão, abaixo da qual todo o material está na fase sólida; a “segunda” temperatura de fusão, acima da qual todo o material inorgânico se encontra na fase líquida e duas temperaturas intermediárias, a saber: a chamada “*sticky temperature*”, que é a temperatura na qual o material apresenta características de aderência aos tubos e paredes e contém entre 15% e 20% de líquido; e a temperatura de deformação radical, na qual 70% do material inorgânico está líquido e apresenta uma grande fluidez e pouca capacidade de aderência (CARMO, 2017).

Além do processo de deposição de partículas sobre as superfícies do sistema de troca térmica, há ainda outro fenômeno importante chamado de sinterização. A sinterização se caracteriza pela aglomeração e pela união de partículas vizinhas com densificação e aumento da rigidez. O processo de sinterização dificulta a remoção das partículas pelos ramonadores e ocorre de forma mais rápida para partículas menores. Os depósitos de material particulado consistem majoritariamente em compostos alcalinos solúveis em água (99%), principalmente  $\text{Na}_2\text{SO}_4$  e  $\text{Na}_2\text{CO}_3$ , de uma pequena quantidade de  $\text{NaCl}$  e de compostos de enxofre reduzidos, tal como  $\text{Na}_2\text{S}$ . Há ainda a presença de compostos de potássio (CARMO, 2017).

Os depósitos de partículas nas regiões superiores da caldeira são compostos basicamente de três tipos de material: partículas de arraste (em inglês, *carryover*), fumo (em inglês, *fume*) e partículas de tamanho intermediário ou ISP do inglês “*intermediate sized particle*”, que diferem entre si na composição e no tamanho.

As partículas de arraste são formadas por impactação inercial de partículas maiores do “*smelt*” ou partículas do licor negro parcialmente queimadas, que foram arrastadas pelo gás de exaustão. O tamanho das partículas de arraste varia de 0,01 a 3 mm e, quando depositadas, elas apresentam uma coloração rosada (TRAN, 2007). Por causa de sua origem, as partículas de arraste possuem uma composição próxima à do “*smelt*” oxidado.

Os depósitos de fumos são originários de compostos alcalinos de sódio e potássio do licor negro que vaporizam e seguem o fluxo de gás de exaustão. As partículas de fumos, com

aparência branca, costumam apresentar uma distribuição uniforme de tamanho, que varia de 0,1 a 1  $\mu\text{m}$  e a sua composição varia com a temperatura do leito de “*char*”, a sulfidez, a concentração de voláteis no licor e com a posição na caldeira (TRAN, 2007).

As partículas ISP apresentam coloração branca ou acinzentada e tamanhos variando de 1 a 100  $\mu\text{m}$  e são provenientes do material fundido arrastado pelo gás. Sua composição é de difícil determinação. Elas podem ser formadas pela fragmentação de gotículas de licor negro durante a combustão, pelo escape de partículas do material fundido durante a queima do “*char*”, ou pela aglomeração de fumos que foram depositados, sofreram sinterização e retornaram ao fluxo de gás (TRAN, 2007). As partículas ISP podem ainda ser resultantes da formação de partículas irregulares nos evaporadores devido à presença de impurezas minerais, compreendendo compostos de silício, magnésio ou cálcio (VAKKILAINEN, 2005).

## Capítulo VIII- Monitoramento de Partículas em Caldeira de Recuperação Kraft por Machine Learning

---

Este capítulo resume o artigo “Monitoramento de Partículas em Caldeira de Recuperação Kraft por Machine Learning”, publicado no periódico O Papel.

### VIII.1 Resumo

A caldeira de recuperação é um equipamento-chave na produção de papel e celulose pelo processo Kraft. Ela permite minimizar a quantidade de rejeitos e maximizar o aproveitamento energético nas plantas de processamento. A deposição de material particulado formado na caldeira representa um importante problema operacional, que deve ser monitorado e controlado para evitar paradas operacionais não planejadas. O processo de formação de partículas nesse equipamento é um fenômeno de dinâmica lenta e complexa e modelos fenomenológicos são de difícil desenvolvimento. Uma alternativa promissora, que se insere no contexto da Indústria 4.0, é a aplicação de técnicas de *Machine Learning* para modelagem de tal fenômeno. Este trabalho propõe uma metodologia para o emprego de redes neurais artificiais no monitoramento da formação de material particulado em uma caldeira de recuperação. Foram utilizadas redes do tipo mapeamento auto-organizável (SOM) para o desenvolvimento de modelos classificatórios a partir de dados históricos de um ano de operação em uma planta industrial. Foi obtido um mapa auto-organizável de topologia 20x5, que permitiu a identificação de regiões de operação normal, regiões de transição e regiões de maior probabilidade de operação anormal. Este método pode ser usado na geração de recomendações referentes à segurança operacional, sinalizando condições operacionais com maior potencial de risco.

### VIII.2 Metodologia Sumarizada

As redes de mapas auto organizáveis (SOM) têm aprendizagem não supervisionada, ou seja, não precisam de exemplos de saída do processo para realizar sua tarefa. Elas são capazes de reconhecer características nos dados de entrada, permitindo o seu agrupamento e posterior classificação. O treinamento das redes SOM se dá em três etapas: competição, cooperação e



adaptação sináptica e gera uma rede de camada única em forma de grade uni ou bidimensional, na qual os neurônios se posicionam lado a lado, formando vizinhanças. A camada única é, portanto, uma representação da saída de rede de mapeamento auto-organizável.

A abordagem proposta foi o uso de redes SOM para a classificação dos dados em função do número de partículas em um estudo de caso com dados reais de uma importante fábrica de papel e celulose brasileira. Os dados representam 12 meses de operação de uma caldeira de recuperação Kraft, que passou por condições de operação anormal caracterizada por elevada formação de partículas em 5,8% dos casos. A variável de processo *Epart* representa o número de partículas formadas por minuto e os seus valores são inferidos pela média aritmética do número de partículas formadas nos lados direito e esquerdo da entrada do superaquecedor da caldeira, monitorados por câmeras. As variáveis de processo que são as entradas do modelo são parâmetros de monitoramento do ar de combustão bem como características do licor negro alimentado à caldeira de recuperação: vazões, temperaturas e pressões de ar primário e secundário e também a vazão de ar terciário da fornalha; vazão volumétrica, concentração de sólidos e temperatura e pressão na alimentação do licor negro à caldeira.

O estudo do uso de redes para classificação se baseou na separação dos dados em 7, 5, 4, 3 e 2 classes. Utilizou-se a ferramenta *Automated Neural Networks* (SANN) do pacote *Data Mining* do *software* STATISTICA® para avaliar redes bidimensionais de mapeamento auto-organizável com diferentes topologias. Variou-se o número de neurônios de 4 a 180 e 80% dos dados foram usados para treinamento e 20% deles para validação. A rede gera um mapa topológico, escolhido retangular, em que cada neurônio corresponde a um elemento, e cada conjunto de dados é associado a um neurônio.

### **VIII.3 Principais Resultados e Discussões**

A principal análise consistiu em identificar regiões em que há uma predominância de dados de operação anormal (Classe 2), em que o número de partículas formadas é superior a 200 partículas/min. Os dados de operação normal (Classe 1) são caracterizados por valores de formação de partícula por minuto inferior a 200.

Foram identificados os neurônios vencedores para cada conjunto de dados. Para cada neurônio, foi calculado o percentual de casos de cada classe que nele foram agrupados. Assim, foi possível buscar regiões do mapa em que houve grande concentração de determinada classe.

A partir das porcentagens calculadas para os dados de classe 2 em cada neurônio, foram construídos mapas evidenciando esses valores de modo a permitir a identificação visual de regiões com maior incidência de casos de classe 2. Uma escala de cor foi utilizada para facilitar a visualização, cores mais escuras correspondem a zonas com mais casos de classe 2, conforme mostrado na Figura 31a para a rede de topologia 20x5.

Os neurônios com porcentagens nulas foram ativados somente por dados de classe 1. A Figura 31b fornece uma visualização qualitativa do mapa gerado pela rede. As regiões em verde indicam elevada probabilidade de operação sem problemas de formação de partículas. As regiões em vermelho, pelo contrário, indicam possibilidade de emissão acima de 200 partículas/min. Os quatro neurônios majoritários estão destacados com vermelho mais escuro. As regiões amarelas podem ser entendidas como faixas de transição.

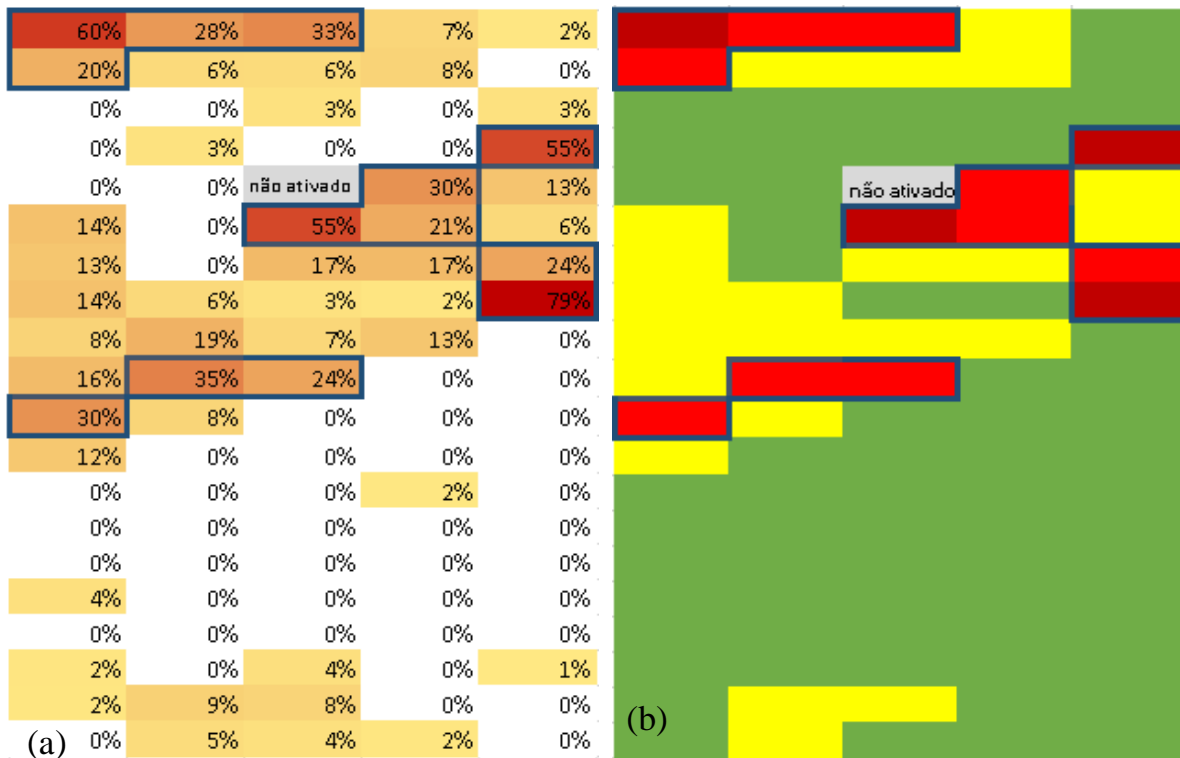


Figura 31 Mapa auto-organizável 20x5 a) SOM quantitativo b) SOM qualitativo

O modelo gerado a partir de uma rede de mapeamento auto-organizável permitiu a geração de uma ferramenta visual para auxiliar o operador na identificação de regiões de maior probabilidade de operação anormal.

#### **VIII.4 Lista das Principais Contribuições**

- Técnicas de aprendizagem de máquina para lidar com fenômenos de difícil modelagem.
- Ferramenta para visualização dos estados de operação da caldeira.
- Pode ser usado na geração de recomendações referentes à segurança operacional.
- Metodologia de classificação generalizável para outros processos.

#### **VIII.5 Artigo Anexo**

## Monitoramento de Partículas em Caldeira de Recuperação Kraft por *Machine Learning*

Elisa C. Werneck do Carmo<sup>1</sup>, Andréa P. Parente<sup>1</sup>, Maurício B. de Souza Jr.<sup>1\*</sup>, Flavio Waltz<sup>2</sup>

<sup>1</sup>EQ, UFRJ; <sup>2</sup>Radix Engenharia & Software S.A.

\*Autor correspondente, UFRJ, Escola de Química, CT, Av. Athos da Silveira Ramos, 149, Bloco E - Sala E-201 - Ilha do Fundão, Rio de Janeiro - Brasil - CEP 21941-909, Caixa Postal 68542, e-mail: mbsj@eq.ufrj.br

**RESUMO:** A caldeira de recuperação é um equipamento chave nas plantas de produção de papel e celulose pelo processo Kraft. A deposição de material particulado formado na caldeira representa um importante problema operacional, devendo ser monitorada e controlada para prevenção de paradas operacionais não programadas. O processo de formação de partículas neste equipamento é complexo e modelos fenomenológicos são de difícil desenvolvimento. Uma alternativa promissora, que se insere no contexto da Indústria 4.0, é a aplicação de técnicas de aprendizado de máquinas, ou "*Machine Learning*", para modelagem de tal fenômeno. O objetivo do presente trabalho foi propor uma metodologia para o emprego de redes neurais artificiais no monitoramento da formação de material particulado em uma caldeira de recuperação. Foram utilizadas redes do tipo mapeamento auto-organizável para o desenvolvimento de modelos classificatórios, a partir de dados históricos de um ano de operação em uma planta industrial. Foi obtido um mapa auto-organizável que permitiu a identificação de regiões de maior probabilidade de operação anormal. Este método pode ser usado para identificar condições com maior potencial de risco operacional.

**PALAVRAS-CHAVES:** monitoramento; redes neurais; caldeira de recuperação; processo Kraft; *machine learning*.

### 1. INTRODUÇÃO

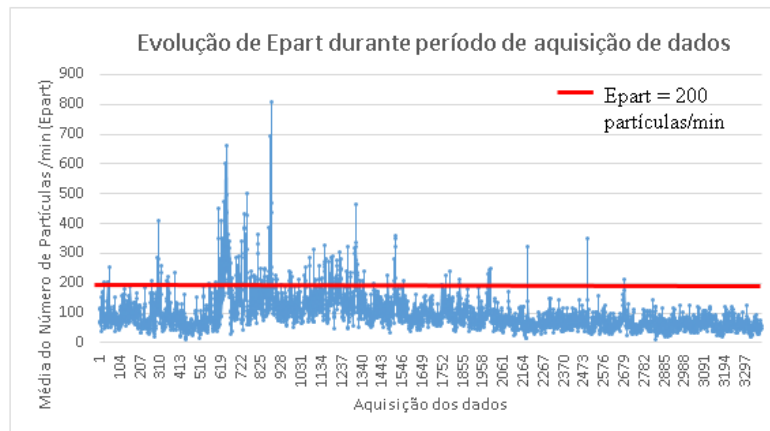
A necessidade das indústrias de aumentar a sua produtividade, dinamizar a sua produção, adaptando-as às exigências dos clientes e ao aumento de sua competitividade está levando à quarta revolução, chamada de Indústria 4.0. Esta revolução industrial se baseia no conceito de *smart factories*, com elevada automação, eficiência e integração.

A Indústria 4.0 engloba tecnologias dotadas da capacidade de interagir com o ambiente em seu entorno e de resolver problemas, buscando soluções e tomando decisões de forma autônoma. Diversos conceitos são essenciais para compreender a complexidade e os potenciais desta nova revolução industrial, destacando-se o de "*Machine Learning*" ou aprendizagem de máquinas.

As técnicas de “*Machine Learning*” são técnicas de inteligência artificial, baseadas no aprendizado a partir de exemplos, que são usadas no reconhecimento de padrões. Desta forma, máquinas são capazes de aprender, de se adaptar e de agir de forma autônoma a partir dos dados de processo aos quais elas têm acesso (NILSSON *et al.*, 2005). O presente estudo propõe o uso de “*Machine Learning*” na busca de soluções inovadoras para um dos desafios enfrentados pela indústria de papel e celulose: o monitoramento da formação de partículas em uma caldeira de recuperação.

Durante a queima de licor negro nas caldeiras de recuperação ocorre a formação de material particulado, que pode ser arrastado pelos gases de exaustão e depositado nas superfícies de troca térmica. Esta deposição provoca perda de eficiência de troca térmica e até mesmo bloqueio do escoamento dos gases, tornando necessária a realização de paradas não programadas. Assim, a quantidade de partículas formadas é um parâmetro que auxilia na avaliação da operação da caldeira e o seu monitoramento tem o potencial de reduzir a frequência de paradas para manutenção.

A Figura 1 ilustra, com os dados empregados no presente estudo, a formação de material particulado, variável Epart, ao longo de um ano, em uma caldeira industrial. Detalhes sobre a obtenção destes dados serão apresentados adiante. Considerando que valores de Epart acima de 200 partículas/min correspondem a condições operacionais anormais. (COSTA *et al.*, 2005), percebe-se que este limite, representado pela linha vermelha horizontal, é ultrapassado em vários momentos da operação da caldeira.



**Figura 1: Evolução de Epart durante o período de aquisição de dados.**

O objetivo geral do presente estudo é, portanto, propor uma metodologia com base em técnicas de “*Machine Learning*”, mais especificamente redes neurais, para prever a formação de material particulado em uma caldeira industrial de recuperação Kraft a partir de dados operacionais. Foram adotados aqui mapas auto organizáveis – redes SOM ou “*Self Organizing Maps*” – a fim de fazer a análise de *cluster* ou agrupamentos dos dados.

Assim, o mapa deverá inferir, tendo por base dados operacionais da caldeira, se a formação de partículas está elevada (operação anormal) ou não (operação normal). Dessa forma, mesmo quando os sensores empregados na contagem de partículas estiverem em falha ou em manutenção, o operador terá um apoio operacional para avaliar a condição do processo.

Este artigo está organizado como segue. A motivação e os objetivos são apresentados nesta seção 1. Na seção 2 são descritos brevemente os fundamentos da caldeira de recuperação Kraft, com ênfase na formação de material particulado, e os dados empregados no estudo. Na seção 3 é descrita a metodologia de diagnóstico da condição operacional baseada em redes SOM. A seção 4 discute os resultados e a seção 5 traz as conclusões do trabalho. Por fim, o Apêndice 1 expõe as equações empregadas no treinamento das redes.

## **2. A CALDEIRA DE RECUPERAÇÃO KRAFT**

### **2.1 O PROCESSO**

O processo Kraft é o mais empregado na produção de polpa celulósica (COSTA, 2004). Nele, a caldeira de recuperação é responsável pela geração de vapor e energia elétrica para a planta a partir da queima do licor negro, assim como pela recuperação de reagentes não consumidos na digestão, o que permite a redução de impactos ambientais.

O licor negro, subproduto do digestor de deslignificação, é uma solução altamente alcalina. Ele é obtido após a o processamento da polpa marrom que separa o licor negro da polpa de celulose. Sua composição é complexa: além de conter a lignina

e os açúcares extraídos da madeira, material orgânico combustível, apresenta os sais inorgânicos, que resultam da reação de deslignificação (BOUCARD, 2014).

As propriedades e a composição do licor negro variam em cada planta de acordo com as condições operacionais estratégicas de cada empresa. Estas diferenças são atribuídas, principalmente, aos seguintes fatores: condições operacionais do digestor, espécie da madeira utilizada, razão entre as quantidades de reagentes químicos alimentados ao digestor e a quantidade de madeira, propriedades do licor branco, rendimento da polpa e eficiência da lavagem do material descarregado do digestor.

A recuperação dos sais inorgânicos contidos no licor negro é crítica para reduzir os custos e o impacto ambiental do processo (GALDIANO, 2006). Para esta finalidade, foram desenvolvidas caldeiras de recuperação. Com isso, o descarte dos subprodutos é reduzido e há um ganho importante em termos ambientais. Uma caldeira de recuperação Kraft típica se divide basicamente em duas seções. Na primeira seção, denominada fornalha, ocorre a queima do licor negro. A segunda seção é uma região de transferência de calor por convecção, onde fica localizado um sistema de troca térmica para aproveitamento de calor, e compreende os seguintes componentes: superaquecedores, feixe convectivo, tubulões e economizadores. O circuito de água de uma caldeira de recuperação é muito semelhante ao das caldeiras aquatubulares utilizadas em geração de vapor em termoelétricas e na indústria de petróleo e gás. Assim, o detalhamento da segunda seção não será feito aqui, podendo ser encontrado em COSTA (2004).

A fornalha é uma seção muito específica da indústria de papel e celulose. É nela que se remove a umidade remanescente do licor negro e que se faz a combustão dos compostos orgânicos e a redução dos sais de sódio de forma a recuperar o sulfeto de sódio em forma fundida.

Pode-se dividir a fornalha em 3 zonas para melhor compreensão dos processos físico-químicos que nela ocorrem. O licor negro, contendo sólidos em suspensão, é alimentado na caldeira, na 1ª zona, através de uma série de bocais de aspersão na forma de gotículas com tamanhos controlados. A 1ª zona da fornalha é uma região de redução, onde ocorrem, predominantemente, as etapas de secagem, pirólise e queima do carbono fixo do licor negro. A zona 2, localizada abaixo da primeira, é uma região de

redução, compreendendo entradas de ar primário e ar secundário e saídas dos sais inorgânicos recuperados na forma de uma massa fundida. A zona 3, localizada acima da 1ª zona, é uma região de oxidação, com injeção de ar terciário.

É inevitável que exista o arraste de pequenas partículas sólidas pelos gases de exaustão, além da formação de sais na parte superior da fornalha, que seguem o fluxo de gases de exaustão. Este arraste pode causar diversos problemas, tais como deposição na superfície externa dos tubos de água, chegando a causar bloqueio da passagem dos gases e redução da eficiência da troca térmica, além do custo de remoção destas partículas no separador eletrostático na saída da caldeira. O progresso das incrustações é acompanhado, principalmente, pelo monitoramento das seguintes variáveis operacionais: queda de pressão, temperatura do gás de exaustão ao longo da caldeira e temperatura do vapor produzido (COSTA, 2004).

A quantidade de partículas que sai da fornalha pode ser medida por instrumentos localizados na região de entrada dos superaquecedores. Em algumas instalações são utilizados câmeras e processamento de imagens para contar o número de partículas que deixam a fornalha por unidade de tempo. Os dados usados no presente estudo são referentes a uma caldeira dotada de tal sistema de medição, em que existem dois sensores localizados no lado direito e no lado esquerdo da fornalha pouco abaixo dos superaquecedores.

Estes sensores apresentam baixa confiabilidade, dadas as condições rigorosas na caldeira. Assim, é necessária uma abordagem de inferência para o caso de falhas ou manutenção nos sensores. Sensores virtuais podem ser usados para inferir o número de partículas a partir das condições de operação da caldeira. As RNA se constituem em uma excelente alternativa para o desenvolvimento desses sensores (WALTZ *et al.*, 2017).

## **2.2 OS DADOS OPERACIONAIS**

Foram usados os dados operacionais tratados por Costa (2004) de uma caldeira de recuperação Kraft de uma fábrica de papel e celulose brasileira. Os dados correspondem a 12 meses de operação, incluem 12 variáveis de processos e 3.381



pontos operacionais (ou conjuntos de dados). O nome da empresa será omitido por razões de confidencialidade.

A obtenção dos dados foi realizada através do histórico do sistema de supervisão e aquisição de dados da indústria. Todas as variáveis empregadas na análise estão apresentadas na Tabela 1.

O número de partículas, formadas fornecido pela variável Epart, foi exibido na Figura 1. A medição é feita por câmeras e posterior processamento de imagem. A quantidade de partículas, de acordo com os dados disponíveis, variou entre 9,62 e 806,12 (valores mínimo e máximo, respectivamente). O valor médio de Epart é igual a 96,99 e o desvio padrão a 61,01.

A análise dos dados mostrou que a operação apresenta grandes variações do valor de Epart, assim como das demais variáveis, e que os valores de Epart não são distribuídos uniformemente. A distribuição não uniforme é esperada, já que os dados foram obtidos durante operação rotineira da caldeira com objetivo de acompanhar o processo e não de realizar uma varredura experimental de todos os valores possíveis de Epart. A menor quantidade de pontos com Epart elevado (acima de 200) também é esperada e evidencia que a caldeira funciona, na maior parte do tempo, dentro das condições desejadas em relação à formação de material particulado.

A Tabela 2 fornece os valores máximos e mínimos, a média e o desvio padrão de cada variável. As variáveis que apresentaram maiores valores de desvio padrão são a pressão de alimentação do ar secundário, Par2, e a quantidade de partículas, Epart. Ademais, foi possível observar que mudanças significativas ocorreram em relação às variáveis Var1 e Var2 a partir, aproximadamente, do conjunto de dados de número 884 e em relação à variável Par2, que sofreu um aumento a partir dos dados de número 1380, caracterizando uma mudança na forma de operação da caldeira.

**Tabela 1: Variáveis de processo. Adaptado de COSTA, 2004.**

<b>TAG</b>	<b>Descrição</b>	<b>Local de medição</b>	<b>Unidades</b>
Vvlu	Vazão volumétrica de licor negro alimentado à caldeira.	Alimentação de licor negro.	m <sup>3</sup> /h
x	Concentração de sólidos na alimentação do licor negro.	Alimentação de licor negro.	%
Tlu	Temperatura de alimentação do licor negro.	Alimentação de licor negro.	°C
Var1	Vazão de ar primário.	Alimentação de ar primário.	t/h
Var2	Vazão de ar secundário.	Alimentação de ar secundário.	t/h
Var3	Vazão de ar terciário.	Alimentação de ar terciário.	t/h
Par1	Pressão de alimentação do ar primário.	Alimentação de ar primário.	mmca
Par2	Pressão de alimentação do ar secundário.	Alimentação de ar secundário.	mmca
Tar1	Temperatura de alimentação do ar primário.	Alimentação de ar primário.	°C
Tar2	Temperatura de alimentação do ar secundário.	Sistema de alimentação de ar secundária.	°C
PI	Pressão de alimentação do licor negro.	Sistema de alimentação de licor negro.	Kgf/cm <sup>2</sup>
Epart	Número médio de partículas por minuto.	Média entre as quantidades medidas na entrada da região dos superaquecedores no sentido do fluxo dos gases do lado direito e esquerdo.	partículas/ min

**Tabela 2: Estatísticas das variáveis de entrada.**

	Vvlu	x	Tlu	Var1	Var2	Var3	Par1	Par2	Tar1	TAr2	PI	Epart
Valor Mínimo	25,09	70,71	127,15	82,35	99,64	18,02	12,72	123,31	120,37	109,34	1,76	9,62
Valor Máximo	74,73	87,39	143,46	156,1	143,61	93,07	247,78	510,71	165,23	205,94	4,31	806,12
Média	63,81	81,46	140,23	122,09	129,14	65,82	112,91	324,84	153,36	141,24	3,91	96,99
Desvio padrão	3,96	2,24	1,07	14,49	8,13	7,37	20,53	103,08	4,61	13,73	0,23	61,01

### 3. METODOLOGIA

#### 3.1 CONCEITOS PRELIMINARES: AS REDES SOM

A natureza complexa da formação de material particulado na caldeira dificulta o desenvolvimento de modelos matemáticos convencionais a partir de métodos analíticos e fenomenológicos, tornando interessante, como alternativa, o uso de redes neurais artificiais (RNA), ou em inglês *Artificial Neuronal Networks* (ANN).

Costa *et al.* (2005) modelaram a formação de material particulado de uma caldeira industrial a partir de redes neurais multicamadas. Foram testadas redes preditivas e classificadoras. O presente trabalho propõe uma abordagem alternativa, pelo uso das redes SOM. Essas redes diferem das redes multicamadas, por terem uma aprendizagem não supervisionada, ou seja, não precisam de exemplos de saída do processo para realizar sua tarefa, mas são capazes de reconhecer características nos dados de entrada que permitem seu agrupamento e posterior classificação. Isso confere às redes SOM um enorme potencial na identificação de condições novas, podendo assistir o operador na automação do monitoramento.

Como as redes multicamadas são bastante conhecidas na literatura (DE SOUZA JR., 1993) e não serão empregadas neste trabalho, elas não serão revisadas aqui. Os fundamentos da rede adotada neste estudo são apresentados na sequência.

O princípio de funcionamento do mapeamento auto-organizável ou SOM (do inglês, *Self-Organized Map*) é baseado na organização do córtex humano, em que cada região é excitada por sensações diferentes e específicas (ROSA e VALLE, 2012). Dessa forma, em uma rede de mapeamento auto-organizável, entradas diferentes

excitam neurônios diferentes de tal modo que neurônios que se encontram próximos uns aos outros são ativados por padrões (entradas) apresentando similaridades.

Ao contrário das redes multicamadas, a SOM é uma rede de camada única em forma de grade uni ou bidimensional, na qual os neurônios se posicionam lado a lado, formando vizinhanças. A camada única é, portanto, uma representação da saída de rede de mapeamento auto-organizável. O treinamento das redes SOM é apresentado no Apêndice 1.

### 3.2 CLASSIFICAÇÃO DO NÚMERO DE PARTÍCULAS

A abordagem proposta foi o uso de redes SOM para a classificação dos dados em função do número de partículas. O estudo do uso de redes para classificação se baseou na separação dos dados em 7, 5, 4, 3 e 2 classes. As classes foram estabelecidas conforme proposto por Costa (2004) e suas definições são apresentadas na Figura 2.

Faixa Epar	Quantidade de dados	Classe e (conjuntos de dados na classe)				
		A	B	C	D	E
0 - 50	481	1 (481)	1 (481)	1 (481)	1 (2236)	
50 - 100	1755	2 (1755)	2 (1755)	2 (1755)		1 (3184)
100 - 150	710	3 (710)	3 (710)	3 (710)	2 (948)	
150 - 200	238	4 (238)	4 (238)			
200 - 250	106	5 (147)				
250 - 300	41			4 (435)		
300 - 350	25	6 (34)	5 (197)		3 (197)	2 (197)
350 - 400	9					
>400	16	7 (16)				

**Figura 2: Definição das classes estabelecidas e estudadas na classificação do número de partículas.**

### 3.3 ANÁLISE DE CLUSTER

Neste trabalho, utilizou-se a ferramenta *Automated Neural Networks* (SANN) do pacote *Data Mining* do software STATISTICA® versão 8.0. Foram avaliadas redes bidimensionais de mapeamento auto-organizável com diferentes topologias. Variou-se o número de neurônios de 4 a 180, assim como a sua disposição na rede e foram testadas redes com as seguintes topologias: 2x2, 3x3, 1x100, 3x40, 3x60, 4x10, 5x7, 5x8, 5x20, 6x6, 10x10 e 20x5 (em que o primeiro número indica o número de linhas e o segundo o número de colunas da matriz formada pelos neurônios, e o produto destes valores fornece o total de neurônios da rede). Foram adotados os valores recomendados pelos desenvolvedores do software para os demais parâmetros da rede, que se encontram resumidos na Tabela 3.

**Tabela 3: Parâmetros utilizados para a geração de redes de mapeamento auto-organizável.**

Parâmetros	Valores definidos
Número de redes geradas a cada análise	1
Amostra de treinamento	80% do conjunto de dados
Amostra de teste	20% do conjunto de dados
Amostra de validação	0% do conjunto de dados
Número de neurônios	De 4 a 180
Entradas	Vvlu, x, Tlu, Var1, Var2, Var3, Par1, Par2, Tar1, Tar2, PI
Ciclos (épocas) de treinamento	1.000
Taxas de aprendizado	De 0,1 a 0,02
Condições de parada de treinamento	A partir de 10 ciclos, quando a melhora na performance é de $10^{-7}$
Vizinhança	Inicia-se com 3 neurônios e finaliza com 0
Inicialização dos pesos	Distribuição normal, valores aleatórios

A rede gera um mapa topológico, escolhido retangular, em que cada neurônio corresponde a um elemento, e cada conjunto de dados é associado a um neurônio.

#### 4. RESULTADOS

Estudou-se o problema utilizando redes neuronais de mapas auto organizáveis em duas etapas. A análise das redes geradas se baseou na capacidade de agrupar dados com características similares em regiões visualmente identificáveis.

Em uma primeira etapa, para a avaliação de cada rede, foi realizada uma análise baseada na divisão dos dados em duas classes: a classe 1 inclui os conjuntos de dados em que  $E_{part}$  é menor ou igual a 200 e a classe 2 reúne os conjuntos de dados em que  $E_{part}$  é maior que 200 (caso E na Figura 2). A análise consistiu em identificar regiões em que há uma predominância de dados de classe 2, ou seja, de operação anormal.

Para tal, foram identificados os neurônios vencedores para cada conjunto de dados. Para cada neurônio, foi calculado o percentual de casos de cada classe que nele foram agrupados. Assim, foi possível buscar regiões do mapa em que houve grande concentração de determinada classe. Por exemplo, na rede SOM 20x5, o neurônio 1 foi o vencedor 42 vezes, das quais 17 correspondem a dados de classe 1 e 25 da classe 2. A partir destes resultados, calculou-se a porcentagem de dados de classe 2 que foram agrupados no neurônio 1. Neste caso, 60% dos dados agrupados no neurônio 1 são de classe 2 e, portanto, este neurônio foi considerado como sendo majoritariamente pertencente a uma região da “classe 2”.

A partir das porcentagens calculadas para os dados de classe 2 em cada neurônio, foram construídos mapas evidenciando estes valores de modo a permitir a identificação visual de regiões com maior incidência de casos de classe 2. Uma escala de cor foi utilizada para facilitar a visualização, cores mais escuras correspondem a zonas com mais casos de classe 2.

A Figura 3 apresenta o mapa criado para a rede 20x5. Nela, cada valor percentual corresponde a um neurônio e a contagem do número do neurônio é realizada a partir da primeira linha (na primeira linha estão localizados os neurônios de 1 a 5, na segunda, de 6 a 10, e assim por diante). Análises análogas foram feitas para as demais configurações da rede., mas não serão exibidas aqui por questão de espaço e porque os resultados para o mapeamento 20x5 foram considerados satisfatórios e ilustrativos.

Os neurônios com porcentagens nulas foram ativados somente por dados de classe 1. Existem 4 neurônios que agrupam majoritariamente dados de classe 2 (mais de 50 %) e há pelo menos 3 regiões que agrupam porcentagens superiores a 20%. Os quatro neurônios majoritários da rede 20x5 capturam 27% dos dados da classe 2.

A Figura 4 fornece uma visualização qualitativa do mapa gerado pela rede, em que as diferentes regiões são indicadas com cores diferentes. As regiões em verde indicam elevada probabilidade de operação sem problemas de emissão de partículas. As regiões em vermelho, pelo contrário, indicam possibilidade de emissão acima de 200 partículas/min. Os quatro neurônios majoritários estão destacados com vermelho mais escuro. As regiões amarelas podem ser entendidas como faixas de transição.

60%	28%	33%	7%	2%
20%	6%	6%	8%	0%
0%	0%	3%	0%	3%
0%	3%	0%	0%	55%
0%	0%	não ativado	30%	13%
14%	0%	55%	21%	6%
13%	0%	17%	17%	24%
14%	6%	3%	2%	79%
8%	19%	7%	13%	0%
16%	35%	24%	0%	0%
30%	8%	0%	0%	0%
12%	0%	0%	0%	0%
0%	0%	0%	2%	0%
0%	0%	0%	0%	0%
0%	0%	0%	0%	0%
4%	0%	0%	0%	0%
0%	0%	0%	0%	0%
2%	0%	4%	0%	1%
2%	9%	8%	0%	0%
0%	5%	4%	2%	0%

**Figura 3: Mapa auto-organizável com 100 neurônios dispostos em uma matriz 20x5, indicando regiões de grande incidência de casos de classe 2.**

A região em verde da Figura 4 concentra 72% de todos os conjuntos de dados, dos quais apenas 0,6% pertencem à classe 2. A região em vermelho agrupa somente 9% dos dados de operação, 61% dos quais pertencem à classe 2. Isto mostra que se um dado inédito for agrupado pela rede nesta região, haverá grande probabilidade de

ele pertencer à classe 2. A região em amarelo agrupa 19% dos dados operacionais, dos quais 10% pertencem à classe 2. Percebe-se, então, que o mapa permite classificar os dados, porém existe uma faixa de indeterminação. Por um lado, uma parte dos dados de classe 2 não são indicados na região em vermelho (falsos negativos) e, por outro, alguns dados de classe 1 são classificados como operação anormal (falsos positivos).

Desta forma, o modelo gerado a partir de uma rede de mapeamento auto-organizável permitiu a geração de uma ferramenta visual para auxiliar o operador na identificação de regiões de maior probabilidade de operação anormal.



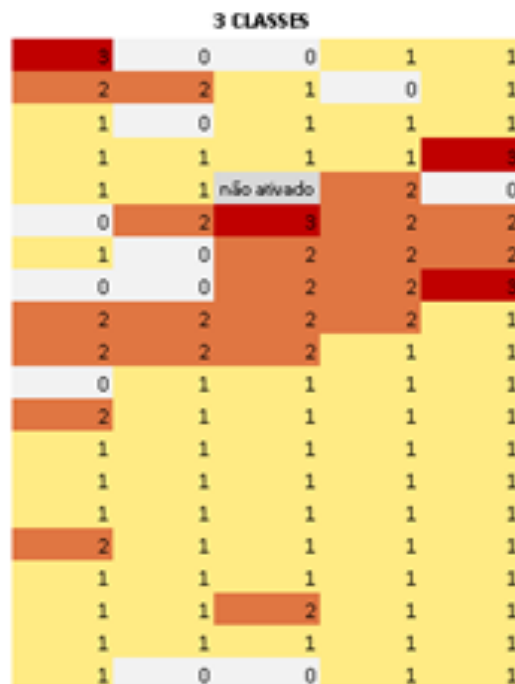
**Figura 4: Ferramenta para detecção de condições anormais baseada na rede 20x5.**

Em uma segunda etapa da análise, a rede teve seu comportamento avaliado quando os dados foram separados em 3, 4, 5 e 7 classes. O procedimento utilizado para este estudo foi similar ao reportado acima, entretanto, como a complexidade da análise aumenta com o número de classes, ao invés de serem criados mapas contendo os valores percentuais de uma determinada classe na posição de cada neurônio, optou-se por indicar qual a classe que ativou majoritariamente aquele neurônio. Foi



estabelecido um valor percentual mínimo de 50% para que uma classe fosse considerada majoritária em um dado neurônio. Caso nenhuma classe fosse preponderante, o neurônio recebia o valor “0”. Essa metodologia é chamada de “proporção mínima com vizinhança por Voronoi” (DE SOUZA Jr., 2003). Foram obtidos resultados como os ilustrados na Figura 5, para classificação em três classes (caso D na Figura 2).

Comparando o mapa da Figura 5 com o da Figura 4, percebe-se que a região de transição (em amarelo) é bastante semelhante à região do mapa de três classes, onde predomina a classe 2. Portanto, o resultado para duas e três classes pode ser entendido como o mesmo resultado expresso de forma diferente. Entretanto, a análise com base em três classes apresenta uma desvantagem pela presença de zonas indefinidas, indicadas com o número zero na Figura 5. A presença de zonas indefinidas ocorre porque é mais difícil um neurônio concentrar mais de 50 % de acertos numa dada classe para um número de classes maior.



**Figura 5: Comparação da rede de topologia 20x5 para análise em 3 classes.**

Para a análise em mais classes, há um aumento da complexidade do mapa com poucos ganhos em relação a melhoria dos resultados; por esta razão esses mapas não serão exibidos aqui.

## 5. CONCLUSÕES

Foi obtida uma rede neuronal artificial de topologia 20x5 em um mapa auto-organizável (SOM) que permite a identificação de regiões de operação normal, regiões de transição e regiões de maior probabilidade de operação anormal, ou seja, com emissão de material particulado acima de 200 partículas/min em uma caldeira de recuperação Kraft. A região de operação normal concentra 72% de todos os conjuntos de dados, dos quais apenas 0,6% apresentam valores acima de 200 partículas/min. A região de maior probabilidade de operação anormal agrupa somente 9% dos dados de operação, 61% dos quais apresentam emissão de partículas superior ao limite operacional. A existência de dados de operação anormal localizados em regiões de operação normal gera falsos negativos, enquanto a situação inversa gera falsos positivos.

Embora a modelagem com base em redes SOM não tenha sido capaz de segregar completamente todos os pontos anormais de operação, o que limita seu uso para o controle do processo, este método pode ser usado para gerar recomendações referentes à segurança operacional, indicando que a operação pode estar comprometida.

Os resultados poderiam ser melhorados caso se dispusesse de uma distribuição uniforme dos dados nas classes, pois o número de exemplos de operação normal é muito maior do que os de falhas, o que prejudica o treinamento da rede neuronal. No entanto, a metodologia proposta pode ser empregada para bases de dados que podem ser aumentadas à medida que novos padrões de dados forem sendo armazenados.

## 6. REFERÊNCIAS BIBLIOGRÁFICAS

BOUCARD, H. **Contributions to the understanding of the hydrothermal processes: application to black liquor**. 2014. 309 f. Tese (Doutorado em Engenharia Química e de Processo) - École des Mines d'Albi-Carmaux, Universidade de Toulouse, Toulouse, França, 2014.

CASTRO F. C. C.; CASTRO M. C. F. Mapas Auto-organizados de Kohonen - SOM. In: **Redes Neurais Artificiais**: Curso de Engenharia Elétrica. PUCRS, Faculdade de Engenharia, Departamento de Engenharia Elétrica, Rio Grande do Sul, 2014. cap. 7, p. 1-19.

COSTA, A. O. S. **Descrição matemática da caldeira de recuperação Kraft**. 2004, 150 f. Tese (Doutorado em Engenharia Química) - COPPE/UFRJ, 2004.

COSTA, A. O. S.; SOUZA Jr., M. B.; BISCAIA Jr., E. C; LIMA, E. L. Monitoring particulate material formation in a kraft furnace recovery boiler. **Tappi Journal**, Atlanta, Estados Unidos da América, v. 4, n. 4, p. E33-47, 2005.

DE SOUZA Jr., M. B. **Redes neuronais multicamadas aplicadas à modelagem e controle de processos químicos**. 1993, 194 f. Tese (Doutorado em Engenharia Química) - COPPE/UFRJ, Rio de Janeiro, 1993.

DE SOUZA Jr., M. B. **Redes neurais com Statistica®**. In: Curso de treinamento para a Versão 6. Escola de Química da UFRJ. UFRJ, Rio de Janeiro, 2003, apostila, 39 p.

GALDIANO, G. P. **Inventário do ciclo de vida do papel offset produzido no Brasil**. 2006. 280 f. Dissertação (Mestrado em Engenharia) - Escola Politécnica da USP, Departamento de Engenharia Química, São Paulo, 2006.

NILSSON, N. J. **Introduction to machine learning**: An early draft of a proposed textbook. Robotics Laboratory, Department of Computer Service. Stanford University, Stanford, Califórnia, Estados Unidos da América, 2005, p. 188. Disponível em: <<http://robotics.stanford.edu/people/nilsson/mlbook.html>>. Acesso em: 10 jul. 2017.

ROSA, R. A. F.; VALLE, M. E. Um estudo da aplicação de redes neurais auto-organizáveis para a identificação autônoma de fonemas portugueses. **Revista Ciências Exatas e Naturais**, v. 17, n. 2, p. 199-218, jul./dez. 2012.

WALTZ, F.; DE SOUZA Jr. M. B.; MORADA, T.; FILHO L. Indústria 4.0 no aumento dos resultados da indústria de papel e celulose. **Webinar ABTCP: Indústria 4.0**. 2017. Disponível em: <<http://www.radixeng.com.br/docs/news/ApresentacaoABTCP.pdf>>. Acesso em: 20 jun. 2017.

WASSERMAN, P. D., **Neural Computing: Theory and Practice**, Van Nostrand, Reinhold, 1989.

## APÊNDICE 1 – TREINAMENTO DA REDE SOM

O treinamento das redes SOM segue o treinamento de Kohonen (WASSERMAN, 1989) e compreende três etapas: competição, cooperação e adaptação sináptica (ROSA e VALLE, 2012). Uma etapa preliminar é necessária para a inicialização dos pesos e do raio da gaussiana. Wasserman (1989) sugere uma inicialização com valores aleatórios.

Após inicialização dos pesos, é começado o processo competitivo que tem por objetivo a definição do neurônio vencedor. Para tal, um padrão de entrada é alimentado à rede e é realizada a busca pelo neurônio que minimiza a distância em relação ao vetor de entrada. Assim, as distâncias são calculadas entre o vetor de entrada ( $\underline{x}$ ) e o vetor de pesos ( $\underline{w}_j$ ) correspondente a cada neurônio. Geralmente, utiliza-se a distância euclidiana. A distância ( $D_j$ ) do neurônio  $j$  em relação à entrada  $\underline{x}$  de dimensão ( $N \times 1$ ), sendo  $w_{ij}$  o peso entre a entrada  $i$  e o neurônio  $j$ , e  $n$  o número da iteração, é dada por:

$$D_j(n) = \|\underline{x} - \underline{w}_j\| = \sqrt{\sum_{i=1}^N (x_i(n) - w_{ij}(n))^2} \quad 1)$$

Para encontrar o neurônio vencedor, deseja-se encontrar a distância mínima, ou seja, o neurônio vencedor  $v(x)$  é dado pela Equação (2):

$$v(x) = \operatorname{argmin}_j \|\underline{x} - \underline{w}_j\| \quad 2)$$

Determinado o neurônio vencedor, dá-se o processo cooperativo. Uma função  $h_{j,v}$  define a vizinhança do neurônio vencedor, centrada nele próprio. Ela deve ser simétrica e atingir seu valor máximo no neurônio vencedor. Sua amplitude deve diminuir monotonicamente até ser nula conforme a distância lateral entre o vencedor e demais neurônios aumenta. Um exemplo de função que atende a essas exigências é a gaussiana. O centro da gaussiana é nulo ( $\mu=0$ ) para o neurônio vencedor. Para que o raio da vizinhança diminua com o tempo de treinamento, pode-se utilizar um desvio-

padrão  $\sigma$  dinâmico, que diminui com o tempo. O decaimento pode ser linear ou exponencial.

Segue-se o processo adaptativo para o ajuste dos pesos dos neurônios na vizinhança topológica do neurônio vencedor (CASTRO e CASTRO, 2014):

$$\underline{w}_j(n+1) = \underline{w}_j(n) + \eta(n) \times h_{j,v(\underline{x})}(n) \times (\underline{x}(n) - \underline{w}_j(n)) \quad 3)$$

Uma vez que os pesos tenham sido ajustados, retorna-se ao início do processo iterativo com a apresentação de um novo padrão e as etapas de competição, cooperação e adaptação são repetidas até que não haja mais modificações significativas no mapa.

**AGRADECIMENTO:** Professor Maurício B. de Souza Jr. agradece ao CNPq (Processo No. 3106682016-6).

# Capítulo IX- Expansão Massiva de Dados Usando Técnica de Monte Carlo Aplicada a Processo Industrial

---

Este capítulo resume o artigo “Massive Expansion of Data Using Monte Carlo Technique Applied to a Paper and Pulp Industry Case”, submetido ao periódico *Process*.

## IX.1 Resumo

Modelos baseados em dados, como redes neurais artificiais (RNA), entre outras técnicas de aprendizado de máquina, são baseados nos dados que são medidos nas plantas de processamento e que, portanto, descrevem as condições reais de processo. Os dados de processos industriais armazenados em histórico representam uma fonte conveniente de informações e podem ser usados para melhorar a segurança e a eficiência das plantas por meio de detecção e diagnóstico de falhas (FDD). No entanto, a maioria dos dados utilizados para geração de modelos é representativa dos estados nominais do processo e, portanto, não é suficiente para resolver problemas de classificação destinados a determinar condições anormais de processo. Este trabalho propõe o uso de técnicas para expandir massivamente a quantidade dos padrões de dados reais originais, evitando experimentos que possam comprometer a segurança do processo. A metodologia proposta usa a técnica de Monte Carlo para aumentar artificialmente o número de entradas do modelo. Em seguida, os dados simulados gerados são comparados aos dados reais disponíveis por meio da busca por vizinhos mais próximos (NNS) por distâncias geométricas para classificar consistentemente os padrões gerados em status normal ou de falha. Finalmente, uma rede neuronal de múltiplas camadas é treinada com os dados simulados. A metodologia sugerida foi validada com um estudo de caso em que 3381 pontos de dados reais de uma indústria de papel e celulose foram expandidos para monitorar a formação de partículas em uma caldeira de recuperação - um equipamento crucial nas fábricas de produção Kraft de papel e celulose. Apenas 5,8% dos dados do processo originais eram exemplos de condições de falha, em que o número de partículas formadas é superior a 200 partículas por minuto. A deposição do material particulado formado é crítica e pode levar a perdas de eficiência e paradas operacionais não planejadas devendo, portanto, ser monitorada e controlada. A nova coleta de dados ampliada e balanceada aumentou o desempenho de classificação da rede neuronal, permitindo seu uso futuro em uma ferramenta FDD.

## IX.2 Metodologia Sumarizada

A técnica de Monte Carlo para expansão massiva de dados pode ser descrita em 3 etapas (Figura 32). No passo 1  $n$  variáveis de entrada aleatórias são geradas usando um algoritmo em linguagem C. Então os valores da variável de saída são calculados através da função de performance nos  $n$  casos. Este trabalho utilizou uma função das distâncias Euclidianas entre os dados da planta de processo e os dados simulados e os vizinhos mais próximos buscam definir se a operação estava em condições normais ou em falha.

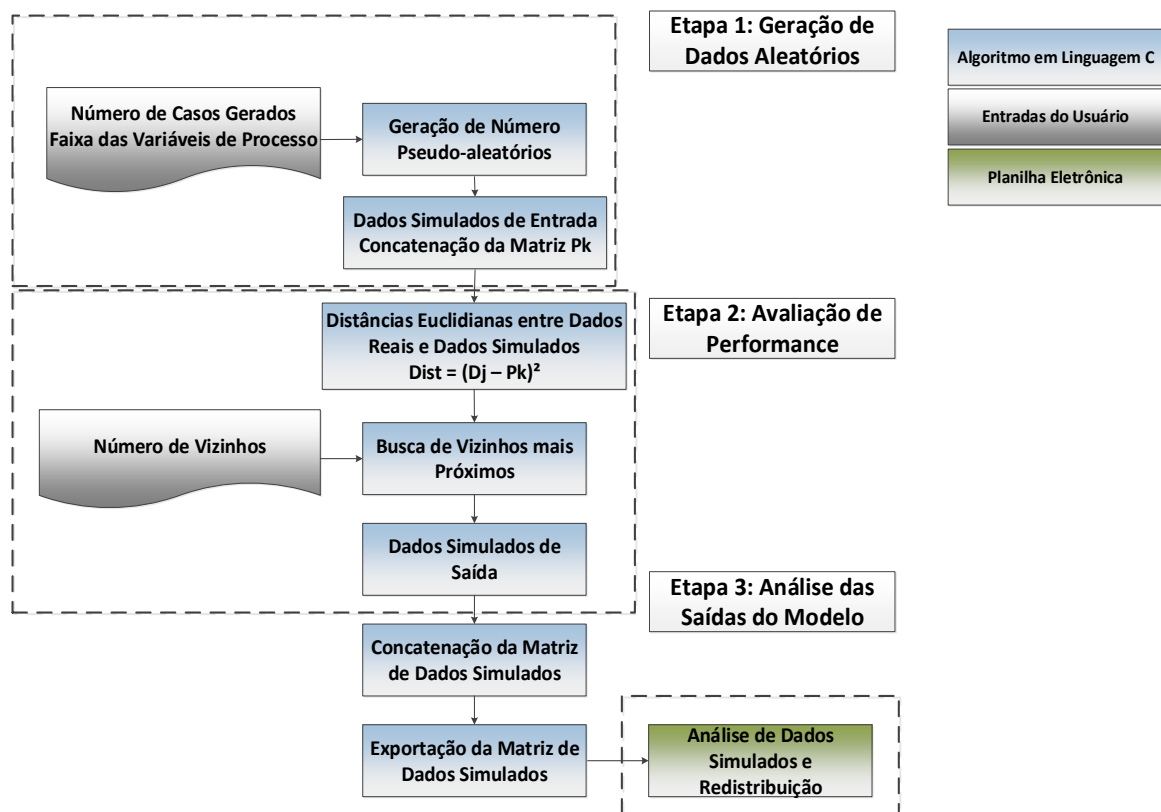


Figura 32 Fluxo de informações da rotina de expansão massiva de dados

As características estatísticas dos dados simulados (saídas do modelo de Monte Carlo) são observadas e se tornam as entradas de um modelo de classificação destinados à detecção e diagnóstico de falhas do processo (Figura 33). Se o conjunto de dados de saída mantiver a distribuição original desequilibrada, ele poderá ser amostrado artificialmente para que a nova coleta de dados exiba uma distribuição uniforme.

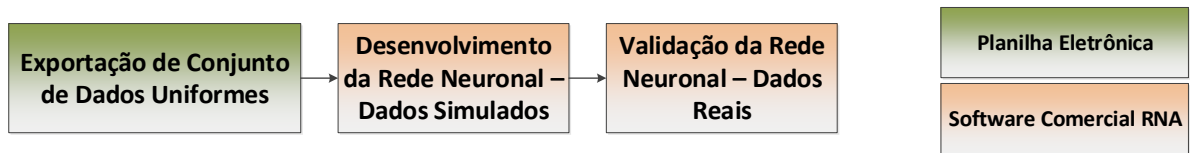


Figura 33 Fluxo de informações do desenvolvimento e validação da rede neuronal artificial

### IX.3 Principais Resultados e Discussões

A simulação de Monte Carlo gerou 160.000 dados não correlacionados e não distribuídos uniformemente. Deste montante, 5502 dados foram re-amostrados para garantir que os status de operação normal e em falha tenham a mesma probabilidade de ocorrência. Esta nova matriz de dados balanceados com onze variáveis de processo associadas a características e 1 variável de processo associada a classe ou status da operação se tornou a matriz de padrões de treinamento para o modelo de detecção e diagnóstico de falhas (FDD) baseado em rede neuronal (RNA) de função de base radial (RBF).

O melhor modelo de RNA - construído com dados simulados – tem topologia de 11 neurônios na camada de entrada, 27 neurônios na camada escondida e 2 neurônios na camada de saída, com funções de ativação gaussiana e softmax. Ele apresenta um índice de desempenho global de 85,6% e taxa de erro de 19,6% e 16,7% para as classes normal e de falha durante as fases de treinamento e de teste.

A fase de validação do modelo utilizou casos reais de dados do processo. A Tabela 14 apresenta a matriz de confundimento do modelo FDD. Os resultados mostraram que à medida que a ferramenta de classificação se torna mais precisa em estados de falha, o número de falsos alarmes aumenta. A ferramenta FDD foi capaz de classificar corretamente os status normal e de falhas em quase 70% dos casos. A análise de características de operação ROC, com área de aproximadamente 0,7, sugere que o classificador é um modelo robusto para uma ampla faixa de prevalência de classes.

A Tabela 15 foi incluída para fins de comparação entre o modelo classificador desenvolvido com dados expandidos e o modelo construído com dados reais de processo não balanceados, ambos seguindo a mesma metodologia de treinamento e refinamento. A melhoria do desempenho do modelo desenvolvido a partir da metodologia proposta neste trabalho é evidente. Embora a rede treinada exclusivamente com dados reais tem alto desempenho global, uma



inspeção das taxas de acerto para as classes individuais mostra que esta ferramenta classificadora não é capaz de diagnosticar operação em falha e, portanto, não seria útil em uma aplicação real.

**Tabela 14 Matriz de confundimento do modelo RNA com dados expandidos**

<b>Topologia RNA: RBF 11-27-2</b>	Classe 1	Classe 2	Global
<b>Corretos (%)</b>	69,88	62,94	69,48
<b>Incorretos (%)</b>	30,12	37,06	30,52

**Tabela 15 Matriz de confundimento do modelo RNA com dados reais**

<b>Topologia RNA: RBF 11-21-2</b>	Classe 1	Classe 2	Global
<b>Corretos (%)</b>	99,238	11,1111	94,212
<b>Incorretos (%)</b>	0,762	88,8889	5,788

#### **IX.4 Lista das Principais Contribuições**

- Solução robusta matemática para conjunto de dados de distribuição não-uniforme.
- Mescla diferentes plataformas de software: Algoritmo C, planilha eletrônica e software comercial para desenvolvimento de RNA.
- Construído para estudo de caso, mas facilmente generalizado para diferentes processos.

#### **IX.5 Artigo Anexo**

## **Massive Expansion of Data Using Monte Carlo Technique Applied to a Paper and Pulp Industry Case**

### **Abstract**

Data-driven models such as artificial neural network (ANN), among other machine learning techniques, are based on the data measured within the processing plants, and thus describe the real process conditions. Industrial archived process data represent a convenient source of information that can be used for safety and efficiency improvement such as early or even predictive fault detection and diagnosis (FDD). Nonetheless, most of the data used for model generation are representative of the process nominal states and therefore are not enough for classification problems intended to determine abnormal process conditions. This work proposes the use of techniques to massively expand the amount of the original real data standards, dismissing the need of experiments that could jeopardize process safety. The proposed methodology uses the Monte Carlo technique to artificially increase the number of model inputs. Then, the nearest neighbor search (NNS) by geometric distances is utilized to consistently classify the generated patterns in normal or failure statuses comparing to the real data available. Finally, a multilayered neural network is trained with the simulated data. The suggested methodology was validated with a study case wherein 3381 pulp and paper industrial data points were expanded to monitor the formation of particles in a recovery boiler - a key equipment in Kraft's pulp and paper production plants. Only 5.8% of the original process data were examples of faulty conditions, wherein the number of particles is greater than 200 particles per minute. The deposition of the formed particulate material is critical and might lead to efficiency losses and unscheduled operational shutdowns. Therefore, it shall be monitored and controlled. The methodology applied to the original data set generated uniformly distributed samples. The new expanded and balanced data collection leveraged the classification performance of the neural network, allowing its future use in a FDD application.

### **Keywords**

Monte Carlo   Data-driven   FDD   Study Case   Pulp and Paper Industry

## Introduction

Even though process control systems have made modern chemical processes highly automatic, the automation of abnormal situation management is yet to be accomplished. According to Venkatasubramanian (2005), UK loses 27 billion dollars per year due to abnormal situations and Vásquez and co-workers (2016) report that the economic losses of petrochemical industry in USA are up to 20 billion dollars per year.

Continuous process monitoring in association with fault detection and diagnosis (FDD) tools might contribute greatly to operational excellence by optimizing maintenance interventions and avoiding unplanned shutdowns and even preventing accidents.

Several approaches have been developed over the years to cope with the FDD issues in industrial processes. Venkatasubramanian, Rengaswamy, and Kavuri (2003a, b, c) summarize and review the main characteristics of the techniques described in the literature. Nonetheless, the size and complexity of the chemical process industries and the increasing amount of data available in the digital architecture era endorse the use of the data-based FDD techniques such as neural network algorithms and multivariate statistical methods (SHU, MING, *et al.*, 2016). Ragabab and co-workers (2018), Casanova-Peláez and co-workers (2012) and Leiviskä (2009) are some examples of data-driven FDD models in chemical processes applications.

There are still many challenges until FDD become widespread in industrial applications. One issue is that most of the historical data used for model generation are representative of the process nominal states. Therefore, the data collection contains limited information, which corresponds to a limited period and is possibly limited to a set of operational states space regions. In industrial processes, low-data of faulty conditions or non-uniform data distributions are common because, of course, processes are not allowed to operate in faulty scenarios in order to decrease the maintenance costs and avoid eventual accidents. Under unbalanced data distributions it would be difficult for the classification algorithm to identify the operational status. The categories with low-data such as unusual faults would be easily be neglected and submerged by the statuses with large amounts of samples (GAO, DENG e YUE, 2019).

Venkatasubramanian (2019) highlights that one drawback of machine learning data centered techniques is exactly the need of tremendous amount of data. Even though we collect much more data now than we did in the past, our domain is not a true “big data” domain like finance, vision, game playing and speech. We also usually deal with incomplete and noisy data sets. Clinical

decision making (AKBILGIC e DAVIS, 2019) and civil and structural engineering (AMICO, MYERS, *et al.*, 2019) are also subjected to such data issues.

Data augmentation seems to be an alternative for chemical engineering applications where computer simulations can generate such voluminous and reliable data. Huang and co-workers (2019), for example, used data augmentation techniques for marine organisms' detection and recognition, Zaifeng *et al.* (2019) focused on image analysis with augmented data and Gao *et al.* (2019) used the data augmentation approach to deal with unbalanced data sets for FDD in industrial benchmarks.

This work proposes the use of the Monte Carlo simulation to artificially increase the amount of the original real data collection, leading to an expanded and widespread data set that is representative of nominal steady states and faulty conditions. The methodology used geometric distances and the nearest neighbors search to preserve the phenomenological characteristics of the original data set in the augmented categorical data collection. The suggested technique was validated in an industrial study case wherein 3381 pulp and paper process data standards were expanded to monitor the formation of particles in a recovery boiler - a key equipment in Kraft's pulp and paper production plants. The new expanded and balanced data collection was used to develop an artificial neural network model to classify operational status in normal or failure, allowing its future use in a FDD application.

### **Pulp and Paper Industry**

The pulp and paper industry is one of the largest industries in the world probably because paper has many powerful benefits to human society through education, communication, security and hygiene. World's paper production was around 406 million tons in 2015 and it is expected to reach 482 million tons by 2030. Brazil is the fourth largest producer of pulp in the world and the ninth largest producer of paper (IBÁ, 2016). Nevertheless, in recent years, pulp and paper mills are facing challenges concerning energy efficiency mechanisms and management of the resulting pollutants, considering the environmental feedbacks and the competitive markets (ENVIRONMENTAL PAPER NETWORK, 2018).

More than half of the paper used globally is for packaging (cartonboard and containerboard). There have been substantial reductions in consumption of printing and writing paper since 2010, which globally represents about a quarter of paper use by volume. The most rapid growth is in

sanitary paper (tissue), although it accounts for less than 10% of global volume at the present (ENVIRONMENTAL PAPER NETWORK, 2018).

The Kraft process, which uses sodium hydroxide (NaOH) and sodium sulphide (Na<sub>2</sub>S) to pulp wood, is the dominant production process in the paper industry (SINGH and CHANDRA, 2019). In this process, the wood is dissolved, and together with the spent pulping chemicals, forms a liquid stream called weak black liquor (BL). The BL is washed out from the pulp and it is sent to the Kraft recovery boiler, where the inorganic pulping chemicals are recovered for reuse meanwhile the dissolved organics are used as a fuel to make steam and power. For every ton of pulp produced, the Kraft pulping process produces about 10 tons of weak black liquor or about 1.5 tons of black liquor dry solids that need to be processed through the chemical recovery process (TRAN and POYRY, 2016).

Every year 1.3 billion tons of BL are processed in recovery boilers worldwide making it possible to recover 15 million tons of cooking chemicals, reducing the amount of waste, and producing 700 million tons of steam at elevated pressure, which makes BL the fifth most important fuel in the world. (Reeve,2002; Lofstedtet al.,2016; SINGH and CHANDRA, 2019).

Despite its importance for increasing process efficiency and reducing environmental impacts, the Kraft recovery process is not easy to operate. Recovery boilers typical problems include fouling of heat transfer tubes and plugging of flue gas passages by fireside deposits, which cause low steam production, blackouts, air emissions and lead to unscheduled operational shutdowns (TRAN and POYRY, 2016). Many of the recovery boiler's issues might be caused by particles formation and deposition, a slow dynamic phenomenon that is difficult to monitor and predict.

Slow dynamics phenomena are difficult to track, and they usually demand maintenance interventions like cleaning and regeneration procedures (ZHANG, JEINSCH, *et al.*, 2011). In this sense, the number of particles formed inside the boiler is a parameter that assists the evaluation of the operation and its control has the potential of reducing the frequency of maintenance interventions.

The complex nature of the formation of particulate material inside the boiler makes it difficult to develop conventional mathematical models from analytical and phenomenological methods, making the use of Artificial Intelligence and Machine Learning techniques an alternative to address this problem.

Artificial neural networks (ANNs) are processing techniques that use empirical information to generate complex system models through the identification and generalization of patterns found in a given set of data. They are endowed with the capacity of learning from examples and are part of the Artificial Intelligence and Machine Learning methods.

### **Case Study**

This study presents real operational data of an important Brazilian pulp and paper mill, whose name was omitted for confidentiality. The data correspond to twelve months of operation of the recovery boiler of a Kraft production process and include 12 process variables and 3,381 cases, acquired through the industrial SCADA system (COSTA, DE SOUZA JR., *et al.*, 2005).

Variable Epart is the number of particles formed inside the boiler. Its value is the average of the number of particles that are dragged by the gas flow in the furnace, at the entrance of the superheater region, on the right and left sides. Measurements are made by means of image processing captured by two cameras.

The operational data collected was preprocessed to deal with spurious data. Whenever the difference between the two camera sensors was greater than a 100 particles per minute the operational case was withdrawn. The Epart values ranged from 9.62 up to 806.12 particles per minute. Epart values of up to 200 particles per minute correspond to normal operating conditions. This threshold is exceeded sometimes, characterizing fault operating condition. There are 197 operational cases wherein the Epart values were greater than 200 particles per minute (5.8% of the total).

Figure 1 shows real process operating conditions wherein failure status is characterized by average number of particles greater than 200/min. The Epart values are not evenly distributed. Non-uniform distribution is expected because the data points were gathered during routine operation in order to monitor the process and not to perform an experimental scan of all possible values of Epart. The low number of points with high Epart values (above 200) shows that the boiler works, most of the time, within the desired operating conditions.

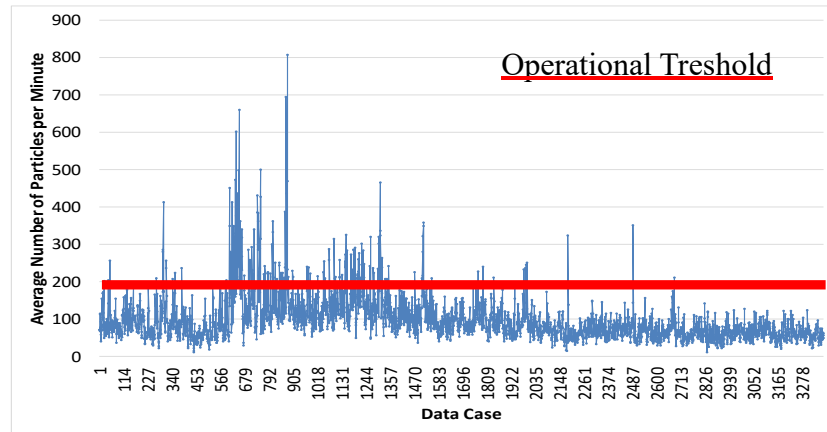


Figure 1 - Average number of particles/min in regular operation.

However, when the data is poorly distributed, the network training process can be impaired, a relevant issue in process control and FDD applications. For instance, if a certain class presents enough training standards, the network will classify with high accuracy rate the data belonging to this specific class. However, if some class is not well trained, the network will be unable to perform generalization and classify future entries belonging to it. Thus, it is possible that the network presents a high overall performance, once it can classify most of the operational cases correctly. Yet, it performs poorly for one of the classes, which is obviously not desired. In process control a network that represents well the normal operating situations, but which cannot model an abnormal situation due to an operational disturbance might lead to unsafe situations.

## Methodology

Monte Carlo simulation is a powerful statistical analysis tool that avoids physical experiments and rather conducts a large number of computerized experiments. It is named after the city of Monte Carlo in Monaco, which is famous for gambling, because it involves generating chance variables and exhibits random behaviors. It is particularly suitable for solving complex engineering problems because it can deal with many random variables, various distribution types, and highly nonlinear engineering models.

Monte Carlo stochastic technique might be depicted in 3 steps as shown in Figure 2. In step 1  $n$  pseudo-random input variables with normal distribution are generated using an algorithm in C programming language (see Equation 1). Rand is a pseudo-random number ranged between 0 and

1. The minimum and maximum values are limited to the real process data range and the number of cases generated are defined by the user. This step led to a large matrix with 11 columns – one for each process variable – and  $n$  lines.

$$P = D_{Min} + (D_{Max} - D_{Min}) * rand \quad \text{Equation 1}$$

Then the values of the output variable are calculated through the performance function at the  $n$  cases: a combination of geometric distances measurements with clustering analysis that establish a decision rule to classify the  $n$  cases depending on the process features. This work used square Euclidian distances between the process plant data and the simulated data, and the nearest neighbors search to define whether the operation was under normal or faulty conditions. The geometric distances were weighted by the inverse of the maximum values of the process variables to avoid influence of their orders of magnitude, according to Equation 2.

$$Dist_{kj} = \sum \frac{(D_{j,i} - P_k)^2}{D_{jMax}} ; k = 1:n \quad i = 1:l \quad \text{Equation 2}$$

Wherein,

$n$  – number of cases generated

$l$  – length of real process data matrix

$D_j$  – real process data matrix

$P_k$  – simulated process data matrix

$D_{jmax}$  – real process data maximum values

The nearest neighbors search clustered five operational conditions alike to the simulated case being evaluated. The output variable (class) was hence determined by the arithmetic mean of the nearest neighbors' classes, keeping the original real data representativity in the augmented categorical data collection.

If the output data set retains the original unbalanced distribution it is possible to be resample it so that the new data collection exhibits uniform distribution.



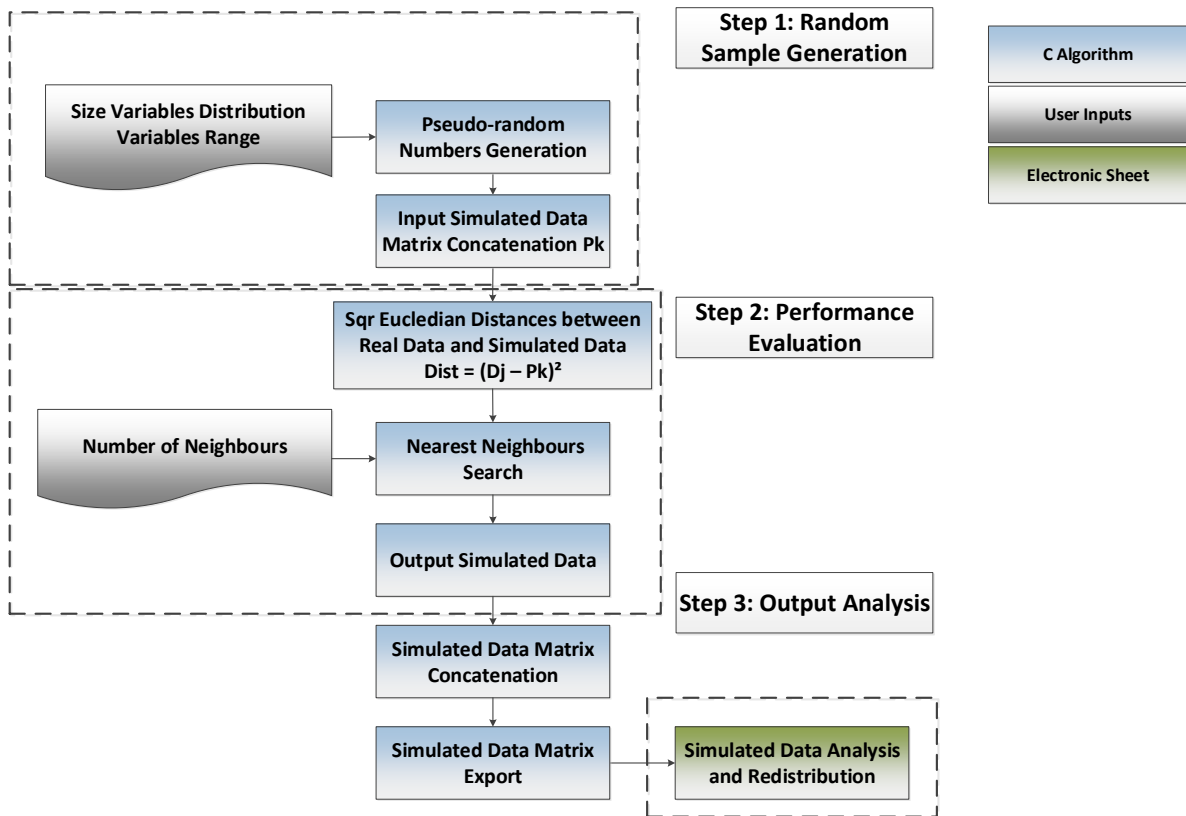


Figure 2 - Massive data expansion routine flow diagram

The statistical characteristics of the experiments (model outputs) such as data correlation are observed and used to develop classification models intended for process fault detection and diagnosis (FDD) (Figure 3).

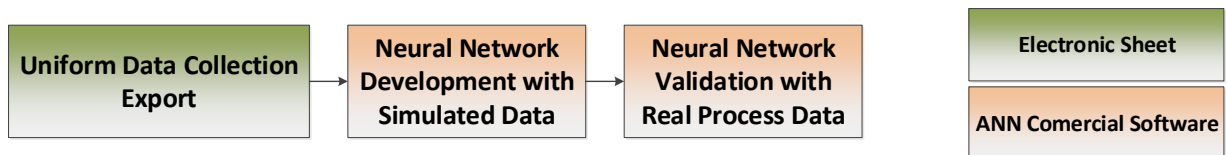


Figure 3 - Artificial neural network development and validation flow diagram

Uniformly distributed simulated data sets that differed in size were used to develop Radial Basis Function (RBF) neural networks using the commercial software STATISTICA®.

The simulated data matrix is the input-output pattern collection that will be used for the fault detection and diagnosis model development and its size will affect the quality of the FDD predictions. A large number of features (or input variables) might lead to numeric difficulties in

obtaining a good class estimate. Given the finiteness of the training set the classifier could be biased, fitted to the training set. It is only when the number of cases,  $n$ , is sufficiently larger than the number of features,  $d$ , that the classifier is able to generalize.

The choice of an adequate dimensionality ratio ( $n/d$ ) was based on power analysis and sample size estimation. If the pattern matrix size is too small, the empirical model will lack the precision to provide reliable answers to the questions it is investigating. If it is too large, time and resources will be wasted, often for minimal gain (SÁ, 2007).

Briefly, power analysis is the hypothesis test that computes the probability of finding an effect that is there (FINKELSTEIN, 2009). In order to increase the power of a test for a fixed level of significance, one is compelled to increase the sample size. This work used a minimum acceptable power of 90% and a confidence level of 5%. The mean values of the population for the significance test were the integer values 1 and 2 - markers attributed to each process class or status. The sample size estimations results are displayed graphically and allowed the determination of the pattern matrix size used for the RBF FDD model development.

The RBF classification tool is a non-parametric method, and therefore it makes no assumptions about the underlying data distribution. Table 1 summarizes the ANN model development parameters.

Table 1 - Neural network development parameters

<b>Parameter</b>	<b>Values</b>
Number of Neural Networks developed	1,000
Training Sample	80% of the data set
Test Sample	20% of the data set
Number of hidden neurons	10 to 100
Input Variables	Vvlu, x, Tlu, Var1, Var2, Var3, Par1, Par2, Tar1, Tar2, PI
Number of neural networks retained	10

At each test, corresponding to the generation of 1,000 new neural networks, the architectures presenting the best performances were selected and used for refining the search. This process was repeated until there was no significant performance improvement. Overall

performance evaluation for both training and test phases were measured by the functions sum of squared error (SOS) and cross entropy (CE), described in equations 3 and 4. The percentage of correct classifications for each class is presented as a confusion matrix.

$$E_{SOS} = \sum_i^N (s_i - t_i)^2 \quad \text{Equation 3}$$

$$E_{CE} = -\sum_i^N t_i \times \ln\left(\frac{s_i}{t_i}\right) \quad \text{Equation 4}$$

wherein

$s_i$  is the output value predicted by the ANN,

$t_i$  is the target output value,

$N$  is the training data set size

Global statistical sensitivity analysis might be used to apportion the importance of each input variable in the ANN model performance by the ratio of a new error (maintaining the evaluated variable constant at its average value) and the original error. If a variable has a great influence on the neural network result, the ratio will be a large positive number.

The best ANN model concerning training and recall performance was validated with real process data. The results were displayed by the confusion matrix and the ROC (receiver operating characteristic) curve.

The ROC curve is an analysis tool for two-class problems that are intended to detect rarely occurring events such as process faults (SÁ, 2007). Figure 4 shows the canonical classification matrix for this situation, with N being the normal event and A being the abnormal process condition. Based on a given decision rule, the true classes are exhibited along the rows and the predicted or decided classifications are displayed along the columns.

		Decision	
		A	N
Reality	A	<i>a</i>	<i>b</i>
	N	<i>c</i>	<i>d</i>

Figure 4 - The canonical classification matrix for two-class discrimination of an abnormal event (A) from the normal event (N).

The ROC analysis method depicts the true positive ratio (TPR, in equation 5) - or sensitivity - versus the false positive ratio (FPR, in equation 6) - or the complement of the specificity - for every possible decision threshold (SÁ, 2007).

$$TPR = \frac{a}{a+b} \quad \text{Equation 5}$$

$$FPR = 1 - \frac{d}{c+d} \quad \text{Equation 6}$$

A classification method with high sensitivity will rarely miss the abnormal event when it occurs. A classifier with a high specificity will have a very low rate of false alarms, which means classifying a normal event as abnormal. A decision method is considered highly accurate if it simultaneously has a high sensitivity (rarely misses the abnormal event when it occurs) and a high specificity (has a low false alarm rate).

Nonetheless, there is a compromise to be made between sensitivity and specificity, graphically presented in the ROC curve. For the classification method to perform optimally for a large range of prevalence situations, the ROC curve shall be very near the perfect curve, i.e., with an underlying area of 1.

## Results and Discussion

Monte Carlo simulation generated 160,000 new patterns for the FDD model development. It created specific faulty situations that are either impossible or too expensive to be forced to happen in the real world and too complex to be theoretically modeled by first principles equations. The data simulation process lasted about 8 hours. Computational time was measured in

an Intel® Core™ i7-8550U processor running at 2GHz in the Windows 10 operating system. The algorithm was embedded with a monitoring pop-up window (Figure 5).

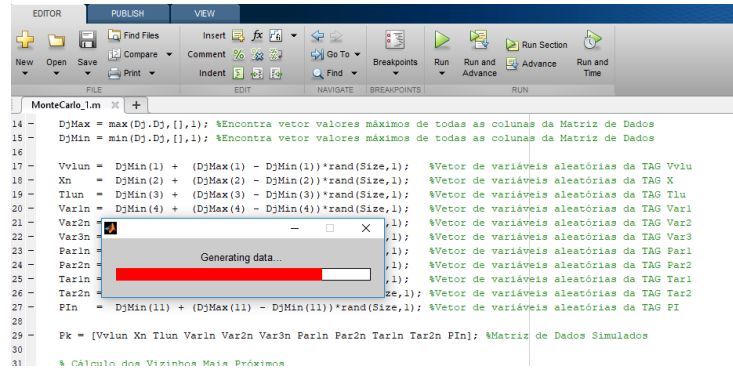


Figure 5 - Monte Carlo simulation monitoring window

Despite the large number of simulated cases the data distribution retained its original unbalanced characteristics of unequal prevalence of the process statuses. The pattern collection was, hence, re-sampled in order to present uniform distribution, which means that any class status has the same probability to happen.

The simulated data matrix size was defined based on power analysis and sample size estimation exhibited in Figure 6. Results show that for a minimum power goal of 90% there must be about 2500 samples for each output class group. It is also inferred that classification results are particularly sensitive to the sample size.

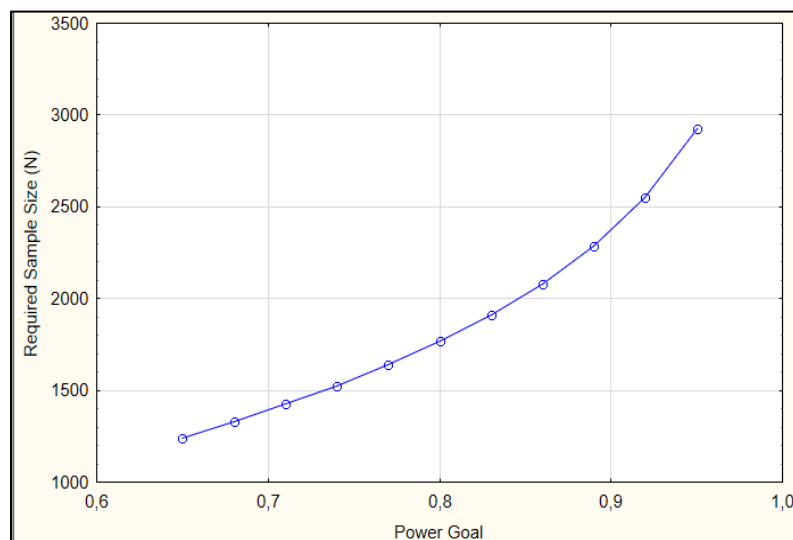


Figure 6 - Sample size versus power for confidence level of 0.05.

The input-output pattern matrix that resulted from the Monte Carlo simulation with 5502 cases was exported to the commercial software STATISTICA® to develop the ANN FDD classification model. Table 2 shows the input variables that are significant for classification purposes according to global sensitivity statistical criteria. The sensitivity test found that all features contribute to class discrimination and that none of them has remarkably greater impact than the others.

Table 2 - ANN model input variables

<b>TAG</b>	<b>Description</b>	<b>Units</b>
Vvlu	Black licor feed flow rate	m <sup>3</sup> /h
x	Solids concentration in black licor feed	%
Tlu	Black licor feed temperature	°C
Var1	Primary air flow rate	t/h
Var2	Secondary air flow rate	t/h
Var3	Tertiary air flow rate	t/h
Par1	Primary air feed pressure	mmca
Par2	Secondary air feed pressure	mmca
Tar1	Primary air feed temperature	°C
Tar2	Secondary air feed temperature	°C
Pl	Black licor feed pressure	Kgf/cm <sup>2</sup>

However, discarding features with no aptitude for class discrimination might not guarantee that there is no redundancy. Figure 7 shows the scatterplots matrix of the variable's correlations. None of the correlations coefficients was greater than 0.9 and so all the 11 process variables were used in the input layer of the FDD ANN model.

Model development had a total running time of 60 hours of computational effort. The best ANN model – built with simulated data – has 11 neurons on the input layer, 27 neurons on the hidden layer and 2 neurons on the output layer. The activation functions for the hidden and output layers are, respectively, gaussian and softmax, described in equations 7 and 8 for the jth

neuron of layer  $k+1$ . The ANN model exhibits overall performance index of 85.6% and error rate of 19.6% and 16.7% for normal and failure classes during training and recall phases.

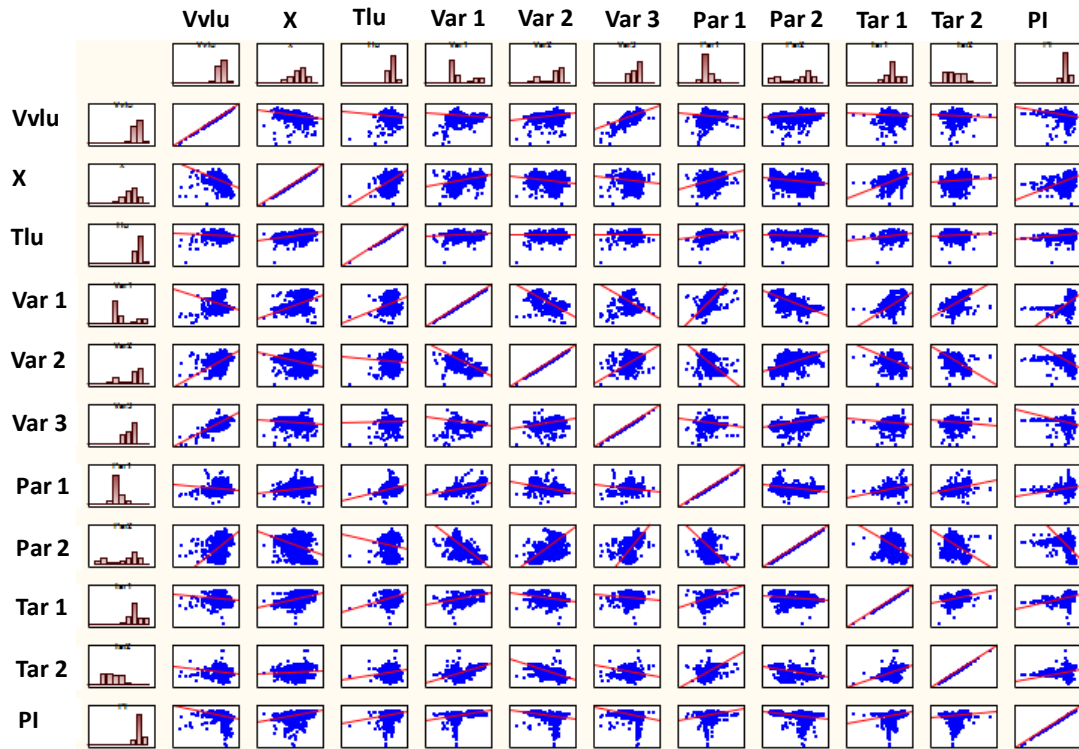


Figure 7 - Real Process data matrix of scatterplot correlation

$$\prod_{i=1}^{n_k} e^{-\left(\frac{(s_{pi,k} - \mu_{ji})^2}{2\sigma_{j1}^2}\right)} \quad \text{Equation 7}$$

$$f(\lambda_{pj,k+1}) = \frac{e^{\lambda_{pj,k+1}}}{\sum_{m=1}^{N_k} e^{\lambda_{pm,k+1}}} \quad \text{Equation 8}$$

In the model validation phase the classifier was tested using a set of independent real process data cases. Its confusion matrix is shown in Table 3 and the sensitivity and specificity values are, respectively, 0.63 and 0.70. Results showed that as the classification tool becomes more accurate in failure states – or more sensitive - the number of false alarms increases - or less specific -, increasing the overall classification error. One may wonder if this methodology that compels equal prevalence of the input-output pattern matrix was beneficial after all. However, risk assessment must not be exclusively influenced by wrong decisions. The classification risks

associated to the losses of a false positive alarm - a normal operational condition diagnosed as failure - are smaller than the losses regarding false negative detections - missing an abnormal operation. The latter might lead to huge economical losses due to unplanned shutdowns.

Table 3 - ANN model confusion matrix (expanded data)

<b>ANN Topology: RBF 11-27-2</b>	Class 1	Class 2	Overall
<b>Total</b>	3184.00	197.00	3381.00
<b>Correct</b>	2225.00	124.00	2349.00
<b>Incorrect</b>	959.00	73.00	1032.00
<b>Correct (%)</b>	69.88	62.94	69.48
<b>Incorrect (%)</b>	30.12	37.06	30.52

It is noteworthy that the variance of the error estimate is predominantly influenced by the finiteness of the validation data set. A larger and proper selection of the raw process data could lead to a classification method with higher accuracy since the real process data could be used in association with the simulated data in the training and validation phases by means of sample partition methodologies.

Figure 8 displays the ROC curve of the RBF FDD model developed. It presents underlying area of 0.69 and feature threshold of 0.52. The shape of the curve is off diagonal, making of it an informative classifier. Considering that the classifier is applied to a scenario where the prevalence of the abnormal situation is low the decision maker would be expected to operate in the lower left part of the ROC curve (small threshold) in order to keep the false positive ratio (FPR) as small as possible. Given the high prevalence of the normal situation, the high rate of false alarms obtained might be explained by the median threshold of the classifier. Nonetheless the high underlying curve area suggests that the classification method is robust because it performs well for a large range of prevalence situations.



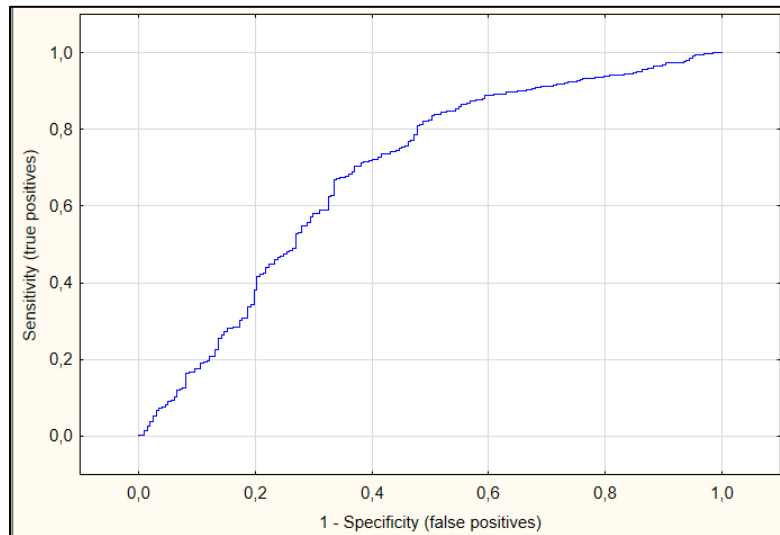


Figure 8 - Receiver operating characteristic (ROC) curve for validation samples

Table 4 was included for purposes of comparison between the classifier model developed with augmented data and the model constructed with real unbalanced process data, both following the same methodology for training and refinement. The improvement of the performance of the model developed from the methodology proposed in this work is evident. Although the network exclusively trained with real data has high overall performance, an inspection of the error rates for individual classes shows that this classifier tool is not able to diagnose fault operation and therefore would be of no use in a real application. The sensitivity value of 0.11 shows that this model will rarely detect an abnormal situation when it occurs. But the specificity value of 0.99 shows that if the classification tool indeed indicates an abnormal situation it will have a high probability of being a true positive scenario.

Table 4 - ANN model confusion matrix (real process data)

<b>ANN Topology: RBF 11-21-2</b>	Class 1	Class 2	Overall
<b>Total</b>	2232.000	135.0000	2367.000
<b>Correct</b>	2215.000	15.0000	2230.000
<b>Incorrect</b>	17.000	120.0000	137.000
<b>Correct (%)</b>	99.238	11.1111	94.212
<b>Incorrect (%)</b>	0.762	88.8889	5.788

## Conclusion

Early or even predictive process fault detection and diagnosis (FDD) models – especially those based on machine learning techniques - seem to be a pillar for the automation of the abnormal situation management to be accomplished. Nonetheless, real process data are typically unbalanced and only representative of normal process operation.

This work proved it possible to use the Monte Carlo technique to massively increase the amount of the original real data standards – a robust mathematical solution to deal with unbalanced data sets. The methodology used geometric distances and the nearest neighbors search to keep the original phenomenological characteristics of the original data set in the categorical augmented data collection. The simulation dismissed the need of experiments that could jeopardize process safety. 160,000 artificial process cases were generated and a new uniformly distributed data collection of 5502 patterns was used to train a multilayered radial basis neural network in order to classify the process operation is normal and failure statuses.

The RBF FDD model developed was validated with real process plant data of a Brazilian pulp and paper mill in order to monitor the formation of particles in a recovery boiler - a key equipment in Kraft's pulp and paper production plants. The FDD tool was able to correctly classify the normal and failures status in almost 70% of the times, with sensitivity of 0.63 and specificity of 0.70. The ROC curve underlying area of about 0.7 suggests that it is a robust model for a wide range of class prevalence.

The proposed solution merges different software platforms: Algorithm in programming language C, Electronic Sheet and Commercial ANN developer. And even though it was built for a study case it may be easily generalized for different processes.

## References

- Akbilgic, O., & Davis, R. L. (2019). The Promise of Machine Learning: When Will it be Delivered? *Journal of Cardiac Failure*, 25, 484-485.
- Amico, B. D., Myers, R. J., Sykese, J., Voss, E., Cousins-Jenvey, B., Fawcett, W., . . . Pomponi, F. (2019). Machine Learning for Sustainable Structures: A Call for Data. *Structures*, 19, 1-4.

Casanova-Peláez, P. J., Cruz-Peragón, F., Palomar-Carnicero, J. M., Dorado, R., & López-García, R. (2012). RBF-ARX model of an industrial furnace for drying olive pomace. *Energy Conversion and Management*, 64, pp. 106 -112.

Costa, A. O., de Souza Jr., M. B., Biscaia Jr., E. C., & Lima, E. L. (2005). Monitoring particulate material formation in a kraft furnace recovery boiler. *Tappi Journal*. *Tappi Journal*, 4, 33-47.

Environmental Paper Network. (2018). The State of the Global Paper Industry. Retrieved from [https://environmentalpaper.org/wp-content/uploads/2018/04/StateOfTheGlobalPaperIndustry2018\\_FullReport-Final-1.pdf](https://environmentalpaper.org/wp-content/uploads/2018/04/StateOfTheGlobalPaperIndustry2018_FullReport-Final-1.pdf)

Finkelstein, M. O. (2009). *Basic Concepts of Probability and Statistics in the Law*. New York: Springer.

Gao, X., Deng, F., & Yue, X. (2019). Data augmentation in fault diagnosis based on the Wasserstein generative adversarial network with gradient penalty. *Neurocomputing*.

Huang, H., Zhou, H., Yang, X., Zhang, L., Qi, L., & Zang, A.-Y. (2019). Faster R-CNN for marine organisms detection and recognition using data augmentation. *Neurocomputing*, 337, 372-384.

Indústria Brasileira de Árvores (IBÁ). (2016). *Relatório Anual*. Brasília. Retrieved from [http://iba.org/images/shared/Biblioteca/IBA\\_RelatorioAnual2016\\_.pdf](http://iba.org/images/shared/Biblioteca/IBA_RelatorioAnual2016_.pdf)

Leiviskä, K. (2009). Elman Network in Kappa Number Prediction. *IFAC Proceedings Volumes*, 42, pp. 477- 482.

Lofstedt, J., Christian Dahlstrand, A. O., Meuzelaar, G., Sawadjoon, S., Galkin, M. V., Agback, P., . . . Samec, J. S. (2016). *Green Diesel from Kraft Lignin in Three Steps*. Chemsuschem Communications.

Ragabab, A., El-Koujok, M., Poulin, B., Amazouz, M., & Yacout, S. (2018, April). Fault diagnosis in industrial chemical processes using interpretable patterns based on Logical Analysis of Data. *Expert Systems with Applications*, 95, pp. 368-383.

Reeve, D. W. (2002). *The Kraft Recovery Cycle*. Tappi Kraft Recovery Operations Short Course.

Sá, J. P. (2007). *Applied Statistics Using SPSS, STATISTICA, MATLAB and R*. New York: Springer.

Shu, Y., Ming, L., Cheng, F., Zhang, Z., & Zhao, J. (2016). Abnormal situation management: Challenges and opportunities in the big data era. *Computers and Chemical Engineering*, 91, 104-113.

Singh, A. K., & Chandra, R. (2019, June). Pollutants released from the pulp paper industry: Aquatic toxicity and their health hazards. *Aquatic Toxicology*, 211, pp. 202 - 216.

Tran, H., & Pöyry, E. K. (2016, February). The Kraft Chemical Recovery Process. 1-8.

Vásquez, J. W., Travé-Massuyès, L., Subias, A., Jimenez, F., & Agudelo, C. (2016). Alarm management based on diagnosis. *IFAC-Papers online*, 126-131.

Venkatasubramanian, V. (2005). Prognostic and diagnostic monitoring of complex systems for product lifecycle management: challenges and opportunities. *Computers and Chemical Engineering*, 29, 1253-1263.

Venkatasubramanian, V., & Viele, S. R.-P. (2019). Artificial Intelligence in Process Systems Engineering: Quo Vadis? I Congresso Brasileiro em Engenharia de Sistemas em Processos. Rio de Janeiro.

Venkatasubramanian, V., Rengaswamy, R., & Kavuri, S. N. (2003a). A review of process fault detection and diagnosis: part I: quantitative model-based methods. *Computers and Chemical Engineering*, 27, 293-311.

Venkatasubramanian, V., Rengaswamy, R., & Kavuri, S. N. (2003b). A review of process fault detection and diagnosis: part II: qualitative models and search strategies. *Computers and Chemical Engineering*, 27, 313-326.

Venkatasubramanian, V., Rengaswamy, R., & Kavuri, S. N. (2003c). A review of process fault detection and diagnosis: part III: process history based methods. *Computers and Chemical Engineering*, 27, 327-346.

Zaifeng, S., Minghe, L., Qingjie, C., Huizhenga, R., & Tao, L. (2019). A data augmentation method based on cycle-consistent adversarial networks for fluorescence encoded microsphere image analysis. *Signal Processing*, 161, 195-202.

Zhang, P., Jeansch, T., Ding, S. X., & Liu, P. (2011). Process Monitoring and Fault Diagnosis - Status and Applications. *Proceedings of the 18th World Congress*. Milano, Italy: The International Federation of Automatic Control.

## **PARTE IV - EPÍLOGO**

## Capítulo X- Considerações Finais

---

Este capítulo apresenta as conclusões finais do trabalho e sugere futuras potenciais áreas de pesquisa no campo das ferramentas para automação do gerenciamento das situações anormais em processos de combustão.

### X.1 Conclusões

Este trabalho integrou tecnologias digitais e analógicas que conduzem a um sistema de automação híbrido, que, em uma estrutura integrada, mesclam soluções matemáticas robustas - por meio de algoritmos em linguagem C e *softwares* comerciais - com sistemas industriais para aquisição de dados e comunicação - como o SCADA e as redes industriais *Fieldbus*.

No Capítulo IV, técnicas de inteligência artificial foram implementadas em uma plataforma SCADA industrial para monitorar os níveis de emissão atmosférica de uma fonte estacionária, afirmando um compromisso com os padrões ambientais. O sensor virtual desenvolvido é capaz de prever as concentrações de CO<sub>2</sub> em tempo real com índice de desempenho MSE igual a 0,006%v/v, que é 5 vezes menor do que a acurácia do sensor físico. Isto sugere que o sensor virtual pode ser usado como instrumento redundante ou até mesmo substituir o sensor físico durante os seus ciclos de manutenção. Desenvolveu-se, também, uma interface homem-máquina na qual os operadores podem criar, atualizar e usar sensores virtuais locais do tipo MLP sem precisar de desenvolvedores especialistas ou de um alto esforço computacional.

O Capítulo V apresentou uma ferramenta de manutenção *online* como solução para o problema da atualização dos modelos de inferência, que podem se tornar obsoletos diante de mudanças das condições operacionais dos processos. O algoritmo desenvolvido para a ferramenta de calibração integra diferentes plataformas em uma estrutura unificada que inclui soluções matemáticas consolidadas, planilhas eletrônicas e um sistema industrial SCADA por meio de uma interface homem-máquina amigável. A ferramenta de calibração coletou dados brutos de processo, realizou as etapas de pré-processamento de dados e atualizou o modelo de inferência automaticamente em menos de 1 minuto, com índice de desempenho global  $r^2$  maior do que 80%. A ferramenta desenvolvida é flexível, dotada de opções de customização que incluem diferentes métodos de treinamento, índices de desempenho alternativos e filtros de pré-processamento

adaptáveis. Além disso, informações históricas podem ser preservadas por meio do acúmulo de padrões de treinamento e do uso de pesos e *biases* antigos como chute inicial para novas calibrações.

Já no Capítulo VI foi desenvolvida uma estrutura de detecção e diagnóstico de falhas capaz de detectar falhas antes de se desenvolverem em defeitos, fornecendo as informações necessárias para incluir mecanismos de reconfiguração *online* capazes de tomar medidas apropriadas para alcançar um modo operacional seguro, diminuindo o risco de poluição ou acidentes. Em associação a esta ferramenta FDD foi proposta uma estratégia de controle que incluiu mecanismos de priorização, detecção e diagnóstico de falhas e reconfiguração *online*, levando a uma operação confiável da planta. Quando submetida a uma circunstância experimental de falha forçada, respondeu imediatamente realizando a reconfiguração do sinal de medição da variável controlada para um sinal de boa qualidade e levou o processo ao *set-point* de maneira rápida e estável. O alto grau de flexibilidade induzido pelo sistema desenvolvido permite ajustes rápidos, garantindo uma operação eficiente em termos energéticos, minimizando o consumo de água e reduzindo as emissões poluentes.

O Capítulo VII usou redes neuronais para fins de detecção de uma falha de processo caracterizada pela elevada formação de partículas em uma caldeira de recuperação Kraft. Foi obtida uma RNA de topologia 20x5 em um mapa auto-organizável (SOM), que permite a identificação de regiões de operação normal, regiões de transição e regiões de maior probabilidade de operação anormal. A região de operação normal concentra 72% de todos os conjuntos de dados, dos quais apenas 0,6% são valores classificados de maneira errada. A região de maior probabilidade de operação anormal agrupa somente 9% dos dados de operação, 61% dos quais apresentam emissão de partículas superior ao limite operacional. Embora a modelagem com base em redes SOM não tenha sido capaz de segregar completamente todos os pontos anormais de operação, o que limita seu uso para o controle do processo, este método pode ser usado para gerar recomendações referentes à segurança operacional, indicando que a operação pode estar comprometida. Os resultados sugeriram que o índice de acerto da ferramenta FDD poderia ser melhorado caso se dispusesse de uma distribuição uniforme dos dados nas classes, pois o número de exemplos de operação normal é muito maior do que os de falhas, o que prejudica o treinamento da rede neuronal.

O Capítulo VIII, então, usou a simulação de Monte Carlo para aumentar a quantidade de dados de processo massivamente. A técnica se mostrou uma solução matemática robusta para lidar com dados não-uniformemente distribuídos. 160.000 casos foram simulados estocasticamente e distâncias Euclidianas quadradas associadas à busca pelos vizinhos mais próximos foram usadas para classificar os dados simulados em estado normal ou falha. A nova coleção de dados balanceada, contendo 5502 padrões foi usada para treinar redes neurais artificiais de base radial para desenvolver um modelo de classificação das condições de processo. O modelo desenvolvido foi validado com dados de processo reais de uma planta de papel e celulose para monitorar a formação de partículas em uma caldeira de recuperação Kraft, acertando em 70% dos casos.

Em suma, foram desenvolvidos modelos de inferência para predição de variáveis de difícil medição e modelos de classificação para detecção e diagnóstico de falhas operacionais com base em técnicas de *machine learning*. Além disso, uma estratégia de controle ativo tolerante a falhas foi implementada com a intenção de automatizar ações corretivas e remediadoras e evitar que falhas se transformem, em última instância, em acidentes com perdas ambientais, econômicas e humanas.

Soluções matemáticas robustas foram propostas para lidar com algumas questões que permanecem em discussão na literatura atual e que, tipicamente, impedem a implementação em larga escala destas ferramentas de automação nos ambientes industriais. Uma delas é o problema da atualização dos modelos de predição para novas condições operacionais – cuja proposta de solução incorpora uma técnica para manutenção *online* dos sensores virtuais. Outra questão é a falta de dados de circunstâncias de falha nos processos industriais – cuja solução inclui simulações estocásticas para expansão massiva de dados de operação.

As ferramentas de automação propostas foram validadas com dados de processo reais. Todas as técnicas aqui desenvolvidas levaram ao aumento do grau de flexibilidade dos sistemas, permitindo monitoramento em tempo real que pode resultar em ajustes rápidos e tomada de ações corretivas e remediadoras, garantindo uma operação segura e eficiente, minimizando consumo de água e reagentes e reduzindo as emissões poluentes e as paradas operacionais. Embora as ferramentas desenvolvidas tenham sido implementadas em processos de combustão específicos, elas foram construídas em interfaces homem-máquina amigáveis e replicáveis, podendo ser estendidas a outros casos de aplicação.



## X.2 Sugestões para Trabalhos Futuros

- *Data augmentation* para modelos baseados em dados

Neste trabalho, redes neurais de base radial foram desenvolvidas com base em técnicas de *data augmentation*; isto é, a combinação de dados industriais com dados simulados para geração de modelos. Sugere-se a exploração desta técnica para geração de modelos de inferência de outras topologias, como por exemplo, as redes SOM e MLP.

- Redes neurais recursivas no tempo para predição

Redes neurais baseadas em janelas temporais móveis podem ser capazes de lidar com os transientes operacionais e as instabilidades de um processo. Sugere-se o desenvolvimento de redes neurais de arquitetura recursiva no tempo para aplicações em otimização e controle adaptativo de processos. Estas redes são voltadas para modelagem, predição e controle de sistemas variantes no tempo.

- Ferramentas FDD para malha fechada FTC

Desenvolvimento de um sistema de detecção e diagnóstico remoto de falhas não só para o processo ou para os sensores, mas também para os atuadores e controladores. Se as não conformidades serão isoladas e diagnosticadas em associação a monitoramento em malha fechada em estação SCADA, as alternativas de reconfiguração *online* das malhas de controle por meio do gerenciamento de redundâncias e re-sintonia dos controladores podem tornar as estratégias de controle industrial cada vez mais tolerantes a falha.

- Gerenciamento de alarmes

Neste contexto de gerenciamento das situações anormais, propõe-se também o desenvolvimento de uma metodologia para gerenciamento de alarmes baseados em técnicas de racionalização e priorização. A hierarquização dos alarmes tem com fim último tornar a operação mais segura ao criar sistemas de alarme menos sobrecarregados e aumentar o desempenho da interface homem-máquina.

- Ampliação das arquiteturas de automação: gerenciamento de ativos

A abordagem deste trabalho se volta para a integração das diferentes tecnologias disponíveis nos sistemas de automação industrial. Complementando este sistema integrado, sugere-se um sistema de gerenciamento dos ativos em uma camada de automação de hierarquia superior do tipo PIMS para gerência das informações fornecidas pelas diversas instâncias de instrumentação utilizadas no parque industrial. Por exemplo, nestes sistemas cartas de controle estatístico poderiam ser usadas para acompanhar a performance global das plantas de processo. A alternativa proposta aumentaria a transparência operacional da unidade, permitindo o acesso dos órgãos reguladores e da sociedade civil a estes dados em tempo real, evidenciando o seu comprometimento ambiental e social.

- Combinação de redes neuronais artificiais e sistemas especialistas

Assim como as redes neuronais artificiais, os sistemas especialistas também encontram inspiração biológica. Enquanto as primeiras aprendem socraticamente, com exemplos, os sistemas especialistas aprendem didaticamente, com regras lógicas. A combinação das duas técnicas poderia não só prever e classificar as condições operacionais como também gerar recomendações de ações remediadoras e interferência de manutenção como suporte automatizado a operação.

## Bibliografia

---

“RECOMMENDATIONS for implementing the strategic initiative INDUSTRIE 4.0”, 2013. Disponivel em: <[http://www.acatech.de/fileadmin/user\\_upload/Baumstruktur\\_nach\\_Website/Acatech/root/de/Material\\_fuer\\_Sonderseiten/Industrie\\_4.0/Final\\_report\\_\\_Industrie\\_4.0\\_accessible.pdf](http://www.acatech.de/fileadmin/user_upload/Baumstruktur_nach_Website/Acatech/root/de/Material_fuer_Sonderseiten/Industrie_4.0/Final_report__Industrie_4.0_accessible.pdf)>. Acesso em: 28 Janeiro 2017.

### Unsupported source type (Misc) for source Tra16.

ABDI. **Panorama setorial de eletrônica para automação**. Agência Brasileira de Desenvolvimento Industrial - Centro de Gestão de Estudos Estratégicos. Brasília. 2009.

ADHITYA, A.; CHENG, S.; LEE, Z. Quantifying the effectiveness of an alarm management system through human factors studies. **Computers and Chemical Engineering**, n. 67(4), p. 1-12, 2014.

ADNAN, N. A.; IZADI, I.; CHEN, T. On expected detection delays for alarm systems with deadbands and delay-timers. **Journal of Process Control**, n. 21, p. 1318-1331, 2011.

AHMED, K. et al. Similarity analysis of industrial alarm flood data. **IEEE Transactions on Automation Science and Engineering**, n. 10, p. 452-457, 2013.

AKBILGIC, O.; DAVIS, R. L. The Promise of Machine Learning: When Will it be Delivered? **Journal of Cardiac Failure**, v. 25, p. 484-485, 2019.

ALKAYA, E. et al. Adaptation to climate change in industry: improving resource efficiency through sustainable production applications. **Water Environment Research**, v. 87, p. 14-25, 2015.

ALWI, H.; EDWARDS, C. Fault tolerant control using sliding modes with online control allocation. **Automatica**, v. 44, p. 1859-1866, 2008.

ALWI, H.; EDWARDS, C.; TAN, C. P. Sliding mode estimation schemes for incipient sensor faults. **Automatica**, v. 45, p. 1679-1685, 2009.

AMARAL, B. S. et al. Comparative profile of pollutants generated by a stationary engine fueled with diesel, biodiesel, and ethanol. **Journal of Aerosol**, v. 100, p. 155-163, 2016.

AMICO, B. D. et al. Machine Learning for Sustainable Structures: A Call for Data. **Structures**, v. 19, p. 1-4, 2019.

AMIN, S.; SCHWARTZ, G. A.; SASTRY, S. S. Security of interdependent and identical networked control systems. **Automatica**, v. 49, p. 186-192, 2013.

ANDRIJIĆ, Ž. U.; CVETNIĆ, M.; BOLF, N. Soft Sensor Models for a Fractionation Reforamte Pnat Using Small and Bootstrapped Data Sets, v. 35, p. 745-756, 2018.

ANEEL. **Atlas da Energia Elétrica do Brasil**. 2<sup>a</sup>. ed. [S.l.]: [s.n.], 2008.

ANEEL. Capacidade de Geração do Brasil. **BIG - Banco de Informações de Geração**, 2019. Disponível em: <<http://www2.aneel.gov.br/aplicacoes/capacidadebrasil/capacidadebrasil.cfm>>. Acesso em: 22 Junho 2019.

ÁSTRÖM, K. J.; KUMAR, P. R. Control: A perspective. **Automatica**, v. 50, p. 3-43, 2014.

ATZORI, L.; IERA, A.; MORABITO, G. Understanding the Internet of Things: definition, potentials, and societal role of a fast evolving paradigm. **Ad Hoc Networks**, v. 56, 2017.

BAGAJEWICZ, M. J. Process Plant Instrumentation: design and upgrade. **Technomic Publishing**, Pensylvania, 2001.

BALDI, S. et al. Real-time monitoring energy efficiency and performance degradation of condensing boilers. **Energy Conversion and Management**, v. 136, p. 329-339, 2017.

BARATEIRO, C. E. R. B. et al. **The Automation of Industrial Plants and the EPC Contracts (Engineering, Procurement and Construction)**. ISA Automation Weekv2011: Installation, Operations, & Maintenance. [S.l.]: [s.n.]. 2011.

BAUER, M. et al. The current state of control loop performance monitoring - A survey of application in industry. **Journal of Process Control**, v. 38, p. 1-10, 2016.

BAUGHMAN, D. R.; LIU, Y. A. **Neural Networks in Bioprocessing and Chemical Engineering**. [S.l.]: Academic Press, Inc, 1995.

BEALE, M. H.; HAGAN, M. T.; DEMUTH, H. B. **Neural Network Toolbox – User’s Guide**. The Mathworks. [S.l.]. 2012.

BEGA, E. A. **Intrumentação Aplicada ao Controle de Caldeiras**. 2<sup>a</sup>. ed. Rio de Janeiro: JR Editora Técnica Ltda., 1998.

BEKARA, C. **Security Issues and Challenges for the IoT-based Smart Grid**. *Procedia Computer Science: International Workshop on Communicating Objects and Machine to Machine for Mission Critical Applications (COMMCA-2104)*. [S.l.]: [s.n.], 2014. p. 532-537.

BIALKOWSKI, W. L. Dreams vs. reality: A view from both sides of the gap. **Pulp and Paper Canada**, v. 94, p. 19-27, 1993.

BISPO, V. D. D. S. et al. Development of an ANN-based soft-sensor to estimate the apparent viscosity of water-based drilling fluids. **Journal of Petroleum Science and Engineering**, p. 1-5, 2016.

BLANKE, M. et al. Fault-tolerant control systems - A holistic view. **Control Engineering Practice**, v. 5, p. 693-702, 1997.

BLANKE, M. et al. **Diagnosis and fault-tolerant control**. [S.l.]: Springer, 2006.

BOYES, W. Control special report: Ten steps to avoid unnecessary plant shutdowns. **Schaumburg, II: ControlGlobal**, 2012.

BULLEMER, P. T.; KIFF, L.; THARANATHAN, A. Common procedural execution failure modes during abnormal situations. **?**, v. 24, p. 814-818, 2011.

BUSTAN, D.; SANI, S.; PARIZ, N. Adaptive fault-tolerant spacecraft attitude control design with transient response control. **IEEE/ASME Transactions on Mechatronics**, v. 19(4), p. 1404-1411, 2014.

BYRES, E. Can't happen at your site? **In Tech Magazine**, Fevereiro 2002.

CAI, Z. et al. Comprehensive alarm information processing technology with application in petrochemical plant. **Journal of Loss Prevention in the Process Industries**, v. 38, p. 101-113, 2015.

CAMPOS, M. C. M. D.; TEIXEIRA, H. C. G. **Controles Típicos de Equipamentos e Processos Industriais**. 1. ed. [S.l.]: Edgard Blucher, 2006.

CAMPOS, M. et al. Challenges and problems with advanced control and optimization technologies. **IFAC Proceedings Volumes**, v. 42, p. 1-8, 2009.

CAO, P.; LUO, X. Modeling for soft sensor systems and parameters updating online. **Journal of Process Control**, v. 24, p. 975-990, 2014.

CARELLI, A. C.; JR., M. B. D. S. Controller Performance Monitoring and Diagnosis Applied to a Diesel Hydrotreating Reactor. **International Symposium on Advanced Control of Chemical Processes**, Istanbul, v. 1, n. Adchem Preprints, IFAC, p. 101-106, 2009a.

CARELLI, A. C.; JR., M. B. D. S. Stochastic and Deterministic Performance Assesment of PID and MPC Conbtrollers: Application to a Hydrotreater Reactor. **Computer-Aided Chemical Engineering**, v. 27, p. 1635-1640, 2009b.

CARMO, E. C. W. D. **Técnicas de Aprendizado de Máquinas aplicadas ao monitoramento de partículas em caldeira de recuperação Kraft**. UFRJ. Rio de Janeiro, p. 122. 2017.

CASANOVA-PELÁEZ, P. J. et al. RBF-ARX model of an industrial furnace for drying olive pomace. **Energy Conversion and Management**, 64, 2012. 106 -112.

CASTRO, H. F. **Processos Químicos Industriais II**. Universidade de São Paulo. Lorena, p. 1-30. 2009.

CASTRO, R. M. **Desenvolvimento de Ferramentas de Apoio Operacional para uma Unidade de Hidrotratamento**. Escola de Química, UFRJ. [S.l.]. 2009.

CHAI, T.; QIN, S. J.; WANG, H. Optimal operational control for complex industrial processes. **Annual Reviews in Control**, v. 38, p. 81-92, 2014.

CHANDLER, P. R. **Self-repairing flight control system reliability and maintainability program - Executive overview**. Proceedings of the IEE national aerospace and eletronics conference. [S.l.]: [s.n.]. 1984. p. 586-590.

CHEN, D.; SEBORG, D. E. PI/PID Controller Design Based on Direct Synthesis and Disturbance Rejection. **Industrial & Engineering Chemistry Research**, v. 41, p. 4807-4822, 2002.

CHEN, J.; RANDALL, R. B. Simulation, intelligent diagnosis of bearing knock faults in internal combustion engines using vibration. **Mechanism and Machine Theory**, v. 104, p. 161-176, 2016.

CHEN, K. et al. Soft Sensor Model Maintenance: A Case Study in Industrial Processes. **IFAC-PapersOnLine**, v. 48, n. 8, p. 427-432, 2015.

CHENG, Y.; IZADI, I.; CHEN, T. Optimal alarm signal processing: filter design and performance analysis. **Transactions on Automation Science and Engineering**, n. 10, p. 446-451, 2013a.

CHENG, Y.; IZADI, I.; CHEN, T. Pattern matching of alarm flood sequences by a modified Smith-Waterman algorithm. **Chemical Engineering Research and Design**, n. 91, p. 1085-1094, 2013b.

CHIUSO, A.; PICCI, G. Consistency analysis of some closed-loop subspace identification methods. **Automatica**, v. 41, p. 377-391, 2005.

CHONG, A.; WILCOX, S.; WARD, J. Prediction of gaseous emissions from a chain grate stoker boiler using neural networks of ARX structure. **IEE Proceeding of Science Measurement Technology**, v. 148(3), p. 95-102, 2001.

CIESLAK, J.; EFIMOV, D.; HENRY, D. Transient management of a supervisory fault-tolerant control scheme based on dwell-time conditions. **Journal of Adaptive Control and Signal Processing**, v. 29(1), p. 123-142, 2015.

CLERCQ, D. D. et al. Machine learning powered software for accurate prediction of biogas production: A case study on industrial-scale Chinese production data. **Journal of Cleaner Production**, v. 218, p. 390-399, 2019.

COHN, P. E. **Analisadores industriais: no processo, na área, na supervisão de poluentes e na segurança**. Rio de Janeiro: Interciência, 2006.

CONAMA, C. N. D. M. A. **Resolução CONAMA no382**. [S.l.]: [s.n.], 2007.

COSTA, A. O. S. et al. Monitoring particulate material formation in a kraft furnace recovery boiler. *Tappi Journal*. **Tappi Journal**, Atlanta, v. 4, p. 33-47, 2005.

DAMBROS, J. W. V.; TRIERWEILER, J. O.; FARENZENA, M. **A Short Review and Challenges on Oscillation Detection in Process Industries**. I Congresso Brasileiro em Engenharia de Sistemas em Processos. Rio de Janeiro: [s.n.]. 2019.

DE SOUZA JR, M. B. **Redes Neurais Artificiais**. UFRJ. [S.l.]. 2015.

DE SOUZA JR, M. B.; CAMPOS, M. C. M. M.; TUNALA, L. F. Dynamic Principal Component Analysis Applied to the Monitoring of a Diesel Hydrotreating Unit. In: FERRARINI, L.; VEBER, C. **Modeling, Control, Simulation and Diagnosis of Complex Industrial and Energy Systems**. [S.l.]: The International Society of Automation (ISA), 2009. Cap. Chapter 4, p. 75-96.

DESBOUROGH, L.; MILLER, R. **Increasing customer value of industrial control performance monitoring-Honeywell's experience**. Sixth International Conference on Chemical Process Control, AIChE Symposium Series Number 326. [S.l.]: [s.n.]. 2002.

DOCHAIN, D. et al. Monitoring and control of process and power systems: Towards new paradigms. Status report prepared by IFAC Coordinating committee on Process and Power Systems. **Annual Reviews in Controle**, v. 30, p. 69-79, 2006.

DONG, D.; MCACVOY, T. J.; CHANG, L. J. Emission monitoring using multivariable soft sensors. **Process Am Control Conference**, n. 1, p. 761-765, 1995.

DOWNS, J. J.; SKOGESTAD, S. An industrial and academic perspective on plantwide control. **Annual Reviews in Control**, v. 35, p. 99-110, 2011.

EEMUA. **Alarm systems - A guide to design, management and procurement**. EEMUA-191. London. 2013.

ENVIRONMENTAL PAPER NETWORK. **The State of the Global Paper Industry**. [S.l.]. 2018.

EPA. **Nitrogen Oxides (NO<sub>x</sub>), Why and How They Are Controlled**. Clean Air Technology Center. [S.l.], p. 57. 1999.

EPA. Environmental Protection Agency National Emission Standards for Hazardous Air Pollutants for Major Sources: Industrial Commercial and Institutional Boilers and Process Heaters Final Rule, Federal Register Rules and Regulation, v. 78, n. 21, 2013.

EPE. **Energia Temelétrica: Gás Natural, Biomassa, Carvão, Nuclear**. [S.l.]: [s.n.], 2016.

EPE. **Balanco Energético Anual: Relatório Síntese**. MME. [S.l.]. 2018.

EPE. **Balanco Energético Nacional: Relatório Final**. MME. [S.l.]. 2018.

ETERNO, J. S. et al. **Design issues for fault tolerant-restructurable aircraft control**. Proceedings of the IEE national aerospace and electronics conference. [S.l.]: [s.n.]. 1985. p. 900-905.

EVANS, D. How the Next Evolution of the Internet Is Changing Everything. **CISCO - White Paper**, 2011. Disponível em: <[http://www.cisco.com/c/dam/en\\_us/about/ac79/docs/innov/IoT\\_IBSG\\_0411FINAL.pdf](http://www.cisco.com/c/dam/en_us/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf)>. Acesso em: 02 março 2017.

FIGUEREDO, A. K. M. et al. **SISTEMA INTELIGENTE DE ANÁLISE DE DADOS DISCRETOS DE ALARMES PARA DIAGNÓSTICO DE FALHAS NO TRATAMENTO DA ÁGUA OLEOSA EM PLATAFORMA OFFSHORE**. XXI CONGRESSO BRASILEIRO DE AUTOMÁTICA. Vitória: [s.n.]. 2016.

FINKELSTEIN, M. O. **Basic Concepts of Probability and Statistics in the Law**. New York: Springer, 2009. 174 p.

FLOURIS, I. et al. Issues in complex event processing: Status and prospects in the Big Data era. **The Journal of Systems and Software**, p. 1-20, 2016.



- FORESEE, F. D.; HAGAN, M. T. Gauss-Newton approximation to bayesian learning. **Proceedings of International Conference on Neural Networks**, p. 1-6, 1997.
- FORNETTI, M. A.; LEGAULT, T. A. **Combustion system for a black liquor recovery boiler**. The Mead Corporation. Hayton. 1998.
- FORTUNA, L. **Soft sensor for monitoring and control of industrial processes**. Londres: Springer, 2007.
- FREITAS, E. D. S. **Estudo da Eficiência e da Emissão de Caldeira Operada a Diesel e Biodiesel**. UFRJ. [S.l.], p. 92. 2013.
- FU, P. et al. Performance degradation diagnosis of thermal power plants: A method based on advanced exergy analysis. **Energy Conversion and Management**, v. 130, p. 219-229, 2016.
- FUJITA, M. A.; SHIMEURA, E. Integrity against arbitrary feedback-loop failure in linear multivariable control systems. **Automatica**, v. 24(6), p. 765-772, 1988.
- GADKAR, K. G.; MEHRA, S.; GOMES, J. On-line adaptation of neural networks for bioprocess control. **Computers and Chemical Engineering**, p. 1047-1057, 2005.
- GALADIANO, G. P. **Inventário do ciclo de vida do papel offset produzido no Brasil**. Escola Politécnica da Universidade de São Paulo. São Paulo, p. 280. 2006.
- GALLAGHER, N. et al. Development and benchmarking of multivariate statistical process control tools for a semiconductor etch process: Improving robustness through model updating. **IFAC ADCHEM'97Citeseer**, 1997. 78–83.
- GAO, X.; DENG, F.; YUE, X. Data augmentation in fault diagnosis based on the Wasserstein generative adversarial network with gradient penalty. **Neurocomputing**, 2019.
- GARDELS, N. The great artificial intelligence duopoly. **The Washington Post**, 2018.
- GARG, R. et al. Automating Ischemic Stroke Subtype Classification Using Machine Learning and Natural Language Processing. **Journal of Stroke and Cerebrovascular Diseases**, v. 28, p. 2045-2051, 2019.
- GERON, L. C. A Utilização de Fieldbus para Controle de Caldeiras. **IBP**, Rio de Janeiro, Outubro 1998.
- GERTLER, J.; CAO, J. PCA-based fault diagnosis in the presence of control and dynamics. **AIChE Journal**, v. 50(2), p. 388-402, 2004.

GONZÁLEZ, A.; STARTZMAN, M. A. B. R. Improved neural-network model predicts dewpoint pressure of retrograde gases. **Journal of Petroleum Science and Engineering**, v. 37, p. 183-194, 2003.

GOSH, K.; NATARAJAN, S.; SRINIVASAN, R. Hierarchically distributed fault detection and identification through Dempster-Shafer evidence fusion. **Ind. Wng. Chem. Res.**, v. 50, p. 9249-9269, 2011.

GRACE, T. M.; CAMERON, J. H.; CLAY, D. T. **Char Burning, Project 3473-6**. The Institute of Paper Chemistry. Appletown. 1985.

GRBIC, R.; SLISKOVIC, D.; KADLEC, P. Adaptive soft sensor for online prediction and process monitoring based on a mixture of Gaussian process models. **Computers and Chemical Engineering**, v. 58, p. 84-97, 2013.

GUENAB, F. et al. Design of a fault tolerant control system incorporating reliability analysis and dynamic behaviour constraints. **International journal of Systems Science**, v. 42(1), p. 219-233, 2011.

GULER, M. et al. Transition management for reconfigurable hybrid control systems. **IEEE Control Systems Magazine**, v. 23(1), p. 36-49, 2003.

GUTIERREZ, R. M. V.; PAN, S. S. K. Complexo Eletrônico: Automação do Controle. **BNDES Setorial**, Rio de Janeiro, v. 28, p. 189-232, Setembro 2008.

HÄGGLUND, T. Industrial implementation of on-line performance monitoring tools. **Control Engineering Practice**, v. 13, p. 1383-1390, 2005.

HAMAYUN, M. T.; EDWARDS, C.; ALWI, H. Augmentation scheme for fault-tolerant control using integralsliding modes. **IEEE Transactions on system Control Technology**, v. 22(1), p. 307-313, 2014.

HARRIS, T. J. Assessment of control loop performance. **Canadian Journal of Chemical Engineering**, v. 67, p. 856-861, 1989.

HAYKIN, S. S. **Redes Neurais -Princípios e Práticas**. 2ª. ed. [S.l.]: Bookman Companhia, 2001. 900 p.

HIMMELBLAU, D. M. **Fault Detection and Diagnosis in Chemical and Petrochemical Processes**. Nova Iorque: Elsevier Scientific Publishing Company, 1978.

HOO, K. A. et al. Process and controller performance monitoring: Overview with industrial applications. **International Journal of Adaptive Control and Signal Processing**, v. 17, p. 635-662, 2003.

HORCH, A. **Condition monitoring of control loops, PhD thesis**. Royal Institute of Technology. Stockholm, Sweden. 2000.

HOSKINS, J. C.; HIMMELBLAU, D. Artificial Neural Network Models of knowledge representation in chemical engineering. **Computer Chemical Engineering**, v. 12, p. 881-890, 1988.

HOSNI, M. et al. Reviewing ensemble classification methods in breast cancer. **Computer Methods and Programs in Biomedicine**, v. 177, p. 89-112, 2019.

HOUTZ, A. D. **DynamicShrink**. [S.l.]: [s.n.], 2015.

HSE. **The explosions and fires at Texaco Refinery, Milford Haven, 24 July 1994**. HSE Books. [S.l.]. 1997.

HU, W. et al. An application of advanced alarm management tools to an oil sand extraction plant. **IFAC-Papers online**, p. 641-646, 2015.

HUANG, H. et al. Faster R-CNN for marine organisms detection and recognition using data augmentation. **Neurocomputing**, v. 337, p. 372-384, 2019.

INDÚSTRIA BRASILEIRA DE ÁRVORES (IBÁ). **Relatório Anual**. Brasília, p. 100. 2016.

ISA. **Management of alarm systems for the process industries**. ANSI/ISA 18-2. [S.l.]. 2009.

ISERMANN, R. Fault diagnosis applications: Model-based condition monitoring: Actuators, drivers, machinery, plants, sensors, and fault-tolerant systems. **Springer-Verlag**, 2011.

ITO, E. H. C. et al. Development of a gas composition soft sensor for distillation columns: A simplified model based and robust approach. **Computer Aided Chemical Engineering**, v. 44, p. 661-666, 2018.

JÄMSÄ-JOUNELA, S. L. et al. Evaluation of control performance: Methods, monitoring tool and applications in a flotation plant. **Minerals Engineering**, v. 16, p. 1069-1074, 2003.

JÄMSÄ-JOUNELA, S.-L. Future trends in process automation. **Annual Reviews in Control**, v. 31, p. 211-220, 2007.

JANG-JACCARD, J.; NEPAL, S. A survey of emerging threats in cybersecurity. **Journal of Computer and System Sciences**, v. 80, p. 973 – 993, 2014.

JELALI, M. An overview of control performance assessment technology and industrial applications. **Control Engineering Practice**, v. 14, p. 441-466, 2004.

JENG, J.-C.; TSENG, W.-L.; CHIU, M.-S. A one-step tuning method for PID controllers with robustness specification using plant-step-response. **Chemical Engineering Research and Design**, v. 92, p. 545-558, 2014.

JENG, P. et al. Effects of alternative fuels on the combustion characteristics and emission products from diesel engines: A review. **Renewable and Sustainable Energy Reviews**, p. 1-12, 2016.

JIANG, J.; YU, X. Fault-tolerant control systems: A comparative study between active and passive approaches. **Annual Reviews in Control**, v. 36, p. 60-72, 2012.

JIANG, J.; ZHANG, Y. M. Accepting performance degradation in fault tolerant control system design. **IEEE Transactions on Control Systems Technology**, v. 14, p. 284-192, 2006.

JIN, H. et al. Online local learning based adaptive soft sensor and its application to an industrial fed-batch chlortetracycline fermentation process. **Chemometrics and Intelligent Laboratory Systems**, v. 143, p. 58-78, 2015.

JIN, J.; KO, S.; RYOO, C. K. Fault tolerant control for satellites with four reaction wheels. **Control Engineering Practice**, v. 16(10), p. 1250-1258, 2008.

JIN, Y.; YAMASHITA, Y.; NISHITANI, H. Human modeling and simulation for plant operations. **Computers and Chemical Engineering**, v. 28, p. 1967-1980, 2004.

KADLEC, P. On robust and adaptive soft sensors. **AIChE Journal**, 2009.

KADLEC, P.; GABRYS, B. Adaptive on-line prediction soft sensing without historical data. **Neural Networks (IJCNN)**, 2010.

KADLEC, P.; GABRYS, B.; STRANDT, S. Data-driven soft sensors in the process industry. **Computers and Chemical Engineering**, v. 33, p. 795-814, 2009.

KADLEC, P.; GRBIC, R.; GABRYS, B. Review of adaptation mechanisms for data-driven soft sensors. **Computers and Chemical Engineering**, v. 35, p. 1-24, 2011.

KADLEC, P.; GRBIĆ, R.; GABRYS, B. Review of adaptation mechanisms for data-driven soft sensors. **Comput. Chem. Eng.**, 35, 2011. 1-245.

KALE, M. M.; CHIPPERFIELD, A. J. Stabilized MPC formulations for robust reconfigurable flight control. **Control Engineering Practice**, v. 13, p. 771-788, 2005.

KALOGIROU, S. A. Artificial intelligence for the modelling and control of combustion processes: a review. **Progress in Energy and Combustion Science**, v. 29, p. 515-566, 2003.

KAMAT, S.; MADHAVAN, K. Developing ANN based Virtual / Soft Sensors for Industrial Problems. **IFAC- PapersOnLine**, v. 49(1), p. 100-105, 2016.

KANEKO, H.; FUNATSU, K. Classification of the degradation of soft sensor models and discussion on adaptive models. **AIChE J**, v. 59, p. 2339-2347, 2013.

KANO, M.; OGAWA, M. The state of the art in chemical process control in Japan: Good practice and. **Journal of Process Control**, 20 , 2010. 969–982.

KHARGONEKAR, P. P.; DAHLEH, M. A. Advancing systems and control research in the era of ML and AI. **Annual Reviews in Control**, v. 45, p. 1-4, 2018.

KOIVO, H. N.; TANTTU, J. T. **Tuning of PID controllers**: survey on SISO and MIMO techniques. Proceeding of the IFAC Intelligent Tuning and Adaptive Control Symposium. Singapura: [s.n.]. 1991. p. 75-80.

KOIWANIT, J. et al. An expert system for monitoring and diagnosis of ammonia emissions from the post-combustion carbon dioxide capture process system. **International Journal of Greenhouse Gas Control**, v. 26, p. 158-168, 2016.

KONDAVEETI, S. R. et al. Graphical tools for routine assessment. **Computers and Chemical Engineering**, n. 46, p. 39-47, 2012.

KONDAVEETI, S. R. et al. Quantification of alarm chatter based on run length distributions. **Chemical Engineering Research and Design**, n. 91, p. 2550-2558, 2013.

KORONAKIA, E. D. et al. Classification of states and model order reduction of large scale Chemical Vapor Deposition processes with solution multiplicity. **Computers and Chemical Engineering**, v. 121, p. 148-157, 2019.

KORPELA, T. et al. **Online monitoring of flue gas emissions in power plants having multiple fuels**. Proceedings of the 19th World Congress. Cape Town: [s.n.]. 2014. p. 1355-1360.

LANGER, L. et al. From old to new: Assessing cybersecurity risks for an evolving smart grid. **Computers and Security**, v. 62, p. 165-176, 2016.

LEE, M. et al. Application of a moving-window adaptive neural network to the modeling of a full-scale anaerobic filter process. **Industrial and Engineering Chemistry Research**, v. 44(11), p. 3973–3982, 2005.

LEIVISKÄ, K. Elman Network in Kappa Number Prediction. **IFAC Proceedings Volumes**, 42, 2009. 477- 482.

LI, F.; HAN, Y.; JIN, C. Practical access control for sensor networks in the context of the Internet of Things. **Computer Communications**, v. 89-90, p. 154-164, 2016.

LI, H.; ZHAO, Q. Design of fault tolerant control for MTTF. **International Journal of robust and Nonlinear Control**, v. 18(16), p. 1551-1574, 2008.

LIN, B. et al. A systematic approach for soft sensor development. **Computers and Chemical Engineering**, 31, 2007. 419–425.

LIPTÁK, B. G. **Instrument engineers' handbook. Process measurement and analysis**. [S.l.]: CRC Press, 2003. 1031 p.

LIU, X.; NODA, M.; NISHITANI, H. Evaluation of plant alarm systems by behavior simulation using a virtual subject. **Computers and Chemical Engineering**, v. 34, p. 374-386, 2010.

LIUKKONEN, M. et al. Dynamic soft sensors for NO<sub>x</sub> emission in a circulating fluidized bed boiler. **Applied Energy**, v. 97, p. 483-490, 2012.

LIUKKONEN, M. et al. Dynamic soft sensors for NO<sub>x</sub> emissions in a circulating fluidized bed boiler. **Applied Energy**, 2012. 483-490.

LIUKKONEN, M. et al. Adaptive soft sensor for fluidized bed quality: Applications to combustion of biomass. **Fuel Processing Technology**, v. 105, p. 46-51, 2013.

LU, B.; CHIANG, L. Control Semi-supervised online soft sensor maintenance experiences in the chemical industry. **Journal of Process Control**, 2017.

LU, B.; CHIANG, L. Semi-supervised online soft sensor maintenance experiences in the chemical industry. **Journal of Process Control**, v. 67, p. 23-34, 2018.

LU, B.; STUBER, J.; EDGAR, T. F. Integrated online virtual metrology and fault detection in plasma etch tools. **Ind. Eng. Chem. Res.**, 53, 2014. 5172–5181.

MA, J. P.; JIANG, J. Applications of fault detection and diagnosis methods in nuclear power plants: A review. **Progress in Nuclear Energy**, v. 53, p. 255-266, 2011.

MACFARLANE, A. G. J. **Complex variable methods for linear multivariable feedback systems**. Londres: Taylor and Francis Ltd., 1980.

MACKAY, D. J. C. A practical bayesian framework for backprop networks. **Neural Computation**, p. 1-11, 1991.

MAIER, H. R.; DANDY, G. C. The effect of internal parameters and gemoetry on the performance of back-propagation neural networks: an empirical study. **Environmental Modelling and Software**, 13, 1998. 193-209.

MANSFIELD-DEVINE, S. Ransomware: taking businesses hostage. **Network Security**, v. 10, p. 8-17, 2016.

MARCELLOS, C. F. C. **Controle Estatístico Aplicado a Processos Contínuos**. UFRJ. [S.l.], p. 110. 2013.

MASTERS, T. Practical Neural Network Recipes in C++. **Academic Press**, San Diego, 1993.

MENDONÇA, E. S.; CAMPOS, M. C. M. M.; JR., M. B. D. S. Desenvolvimento de Ferramentas de Apoio a Tomada de Decisões em Unidades de Recuperação de Líquido de Gás Natural. **Petro e Química**, v. 313, p. 68-72, 2009.

MIHIRAN GALAGEDARAGE DON, F. K. Dynamic process fault detection and diagnosis based on a combined approach of hidden Markov and Bayesian network model. **Chemical Engineering Science**, v. 201, p. 82-96, 2019.

MILLER, R. M.; TIMMONS, C. F.; DESBOROUGH, L. D. **CITGO's experience with controller performance monitoring**. Proceedings of the NPRA computer conferenve. San Antonio,USA: [s.n.]. 1998.

MINISTÉRIO DA INDÚSTRIA, COMÉRCIO E SERVIÇOS. Agenda Brasileira para a Indústria 4.0: O Brasil preparado para os desafios do futuro. **ABDi**, 2019. Disponível em: <<http://industria40.gov.br/>>. Acesso em: 21 junho 2019.

MODI, C. et al. A survey of intrusion detection techniques in Cloud. **Journal of Network and Computer Applications**, v. 36, p. 42–57, 2013.

MOHAMMADI, V.; MINAEI, S. Artificial Intelligence in the Production Process. In: \_\_\_\_\_ **Engineering Tools in the Beverage Industry**. [S.l.]: [s.n.], 2019. p. 27-63.

MONTGOMERY, R. C.; CAGLAYAN, A. K. Failure accommodation in digital flight control systems by Bayasian decision theory. **Journal of Aircraft**, v. 13, p. 69-75, 1976.

MONTGOMERY, R. C.; PRICE, D. B. Failure accommodation in digital flight control systems for nonlinear aircraft dynamics. **Journal of Aircraft**, p. 76-82, 1976.

MORRIS, A. S. **Measurement & Instrumentation Principles**. Oxford: Reed Educational and Professional Publishing, 2001. 491 p.

MOTA, A. M. **Calibração Online de Sensor Virtual Baseado em Redes Neurais**. UFRJ. [S.l.]. 2013.

MUENCHHOF, M.; BECK, M.; ISERMANN, R. Fault-tolerant actuators and drivers - Structures, fault detection principles and applications. **Annual Reviews in Control**, v. 33(2), p. 136-148, 2009.

NAGHOOSI, E.; IZADI, I.; CHEN, T. Estimation of alarm chattering. **Journal of Process Control**, n. 21, p. 1243-1249, 2011.

NEUMANN, J. V. Probabilistic logic and synthesis of reliable organisms. **Annals of Mathematics Studies**, v. 34, p. 43-98, 1956.

NEURALWARE. Support: Frequently Asked Questions. **Neuralware**, 2010. Disponível em: <[http://www.neuralware.com/support\\_faq.jsp](http://www.neuralware.com/support_faq.jsp)>. Acesso em: 30 maio 2010.

NG, Y. S.; SRINIVASAN, R. Multi-agent based collaborative fault detection and identification in chemical processes. **Eng. Appl. Artif. Intell.**, v. 23, p. 934-949, 2010.

NIKIEL, J. M. **Sistema de instrumentação e automação de caldeira com combustíveis alternativos (Estudo da influência do uso da mistura diesel-biodiesel no comportamento de uma caldeira)**. EQ/UFRJ. Rio de Janeiro, p. 115. 2009.

NIKULA, R.-P.; RUUSUNEN, M.; LEIVISKÄ, K. Data-driven framework for boiler performance monitoring. **Applied Energy**, v. 183, p. 1374-1388, 2016.

NIMMO, I. Future of Supervisory Systems in Process Industries: Lessons for discrete manufacturing. **Annual reviews in Control**, v. 23, p. 45-52, 1999.

NORA, H. et al. Fault-tolerant control in dynamic systems: Application to a winding machine. **IEEE Control Magazine**, v. 20, p. 117-124, 2000.

O'DWYER, A. **Handbook of PI and PID tuning rules**. 2<sup>a</sup>. ed. [S.l.]: Imperial College Press, 2006.

OESTERREICH, T. D.; TEUTEBERG, F. Understanding the implications of digitisation and automation in the context of Industry 4.0: A triangulation approach and elements of a research agenda for the construction industry. **Computers in Industry**, v. 83, p. 121-139, 2016.



OFFICE OF ENERGY EFFICIENCY AND RENEWABLE ENERGY. **Clean Energy Smart Manufacturing Innovation Institute (CESMII)**, 2016. Disponível em: <<https://www.energy.gov/eere/amo/clean-energy-smart-manufacturing-innovation-institute-cesmii>>. Acesso em: 22 Junho 2019.

OGAWA, S. **A data analysis and graphical representation system for control loop performance assessment**. Proceedings of the TAPPI process control conference. Vancouver, Canada: [s.n.]. 1998.

OLIVEIRA, D. E. D. S. **Controle de Realimentação a 3 elementos em uma Caldeira Industrial**. EQ/UFRJ. [S.l.]. 2017.

OSDER, S. Practical review of redundancy management application and theory. **Journal of guidance, Control and Dynamics**, v. 22(1), p. 12-21, 1999.

OUADDAH, A. et al. Access control in the Internet of Things: Big challenges and new opportunities. **Computer Networks**, v. 112, p. 237-262, 2017.

OVERSCHEE, P. V.; MOOR, B. D. **RaPID: the end of heuristic PID tuning**. Preprints of Proceeding of PID'00: IFAC Workshop on digital control. Terrassa: [s.n.]. 2000. p. 687-692.

PARENT, M.; CUSACK, B. Cybersecurity in 2016: People, technology, and processes. **Business Horizon**, v. 59, p. 567 - 569, 2016.

PARENTE, A. et al. Estratégias de Controle de Emissão de Gases Residuais Aplicadas a Caldeiras de Pequeno Porte. **7º Congresso Brasileiro de Pesquisa e Desenvolvimento em Petróleo e Gás**, Rio de Janeiro, 2013.

PARENTE, A. et al. **Enhanced control strategy of a semi-industrial boiler based on failure diagnosis systems: a study case on hybrid automation systems**. Proceedings of the 15th IATED International Conference. [S.l.]: [s.n.]. 2016. p. 367-372.

PARENTE, A. P. **Estratégias de Controle de Emissão de Gases Residuais Aplicadas a Caldeiras de Pequeno Porte**. EQ/UFRJ. Rio de Janeiro. 2014.

PARENTE, A. P. et al. **Controle da emissão de gases residuais: um estudo de caso em uma caldeira semi-industrial de automação híbrida**. XXI Congresso Brasileiro de Engenharia Química. Fortaleza: [s.n.]. 2016.

PARKINSON, T.; PARKINSON, A.; DEAR, R. Continuous IEQ monitoring system: Context and development. **Building and Environment**, v. 149, p. 15-25, 2019.

PATAN, K.; KORBICZ, J. Sensor fault estimation in the framework of model predictive control. Boiler case study. **8th IFAC Symposium on Fault Detection**, Mexico City, Agosto 2012.

PAULONIS, M. A.; COX, J. W. A practical approach for large-scale controller performance assessment, diagnosis and improvement. **Journal of Process Control**, v. 13, p. 155-168, 2003.

PEDERSEN, A. S. et al. Fault tolerant emergency control to preserve power system stability. **Control Engineering Practice**, v. 53, p. 151-159, 2016.

PERK, S.; TEYMOUR, F.; CINAR, A. Adaptive agent-based system for process fault diagnosis. **Ind. Chem. Eng. Res.**, v. 50, p. 9138-9155, 2011.

PETERSSON, H.; HOLMBERG, M. Initial studies on the possibility to use chemical sensors to monitor and control boilers. **Sensors and Actuators**, v. B 111-112, p. 487-493, 2005.

PIKKUSAARI-SAIKKONEN, J. **Operation's support at the paper mill**. MSc. Espoo: Helsinki University of Technology, Department of Chemical Technology. [S.l.], p. 103. 2004.

PINHEIRO, P. C. C.; VALLE, R. M. Controle de Combustão: Otimização do Excesso de Ar. **II Congresso de Equipamentos e Automação da Indústria Química e Petroquímica**, Rio de Janeiro, 1995. 157-162.

PIRMORADI, F. N.; SASSANI, F.; SILVA, C. W. Fault detection and diagnosis in a spacecraft attitude determination system. **ACTA Astronautica**, v. 65, p. 710-729, 2009.

PITBLADO, R. Global process industry initiative to reduce major accident hazards. **?**, v. 24, p. 57-62, 2011.

PITBLADO, R. Global process industry initiatives to reduce major accident hazards. **Journal of Loss Prevention in the Process Industries**, v. 24, p. 57-62, 2011.

PNUD. **Estudo sobre o Potencial de Geração de Energia a partir de Resíduos de Saneamento (lixo, esgoto), visando incrementar o uso de biogás como fonte alternativa de energia renovável**. MMA. [S.l.]. 2010.

PROMINP. **Competitividade da indústria brasileira de bens e serviços para a indústria de Petróleo e Gás**. [S.l.]. 2006.

QIN, J.; LIU, Y.; GROSVENOR, R. A Categorical Framework of Manufacturing for Industry 4.0 and Beyond. **Changeable, Agile, Reconfigurable & Virtual Production**, v. 52, p. 173-178, 2016.

QIN, S. Control performance monitoring - a review and assessment. **Computer and Chemical Engineering**, v. 23, p. 173-186, 1998.

QIN, S. J. Neural networks for intelligent sensors and control – practical issues and some solutions. **Neural Systems for Control**, p. 213, 1997.

QIN, S. J. Process data analytics in the era of big data. **AIChE**, v. 60, p. 3092-3100, 2014.

QIN, S. J.; YUE, H.; DUNIA, R. Self-validating inferential sensors with application to air emission monitoring. **Industrial Engineering Chemical Research**, v. 36, p. 1675-1685, 1997.

RAGABAB, A. et al. Fault diagnosis in industrial chemical processes using interpretable patterns based on Logical Analysis of Data. **Expert Systems with Applications**, 95, April 2018. 368-383.

RAZI, M. M. et al. Artificial neural network modeling of plastic viscosity, yield point, and apparent viscosity for waterbased drillingfluids. **J. Dispers. Sci. Technol.**, v. 34, p. 822-827, 2013.

REEVE, D. W. **The Kraft Recovery Cycle**. Tappi Kraft Recovery Operations Short Course. [S.l.]. 2002.

RICH, E. **Artificial Intelligence**. [S.l.]: McGraw Hill, 1983. 436 p.

RICHTER, J. H. et al. Reconfigurable control of piecewise affine systems with actuator and sensor faults: Stability and tracking. **Automatica**, v. 47(4), p. 678-691, 2011.

ROSTEK, K.; MORYTKO, Ł.; JANKOWSKA, A. Early detection and prediction of leaks in fluidized-bed boilers using artificial neural networks. **Energy**, v. 89, p. 914-923, 2015.

ROTHENBERG, D. **Alarm management for process control**. New York: Momentum Press, 2009.

SÁ, J. P. M. D. **Applied Statistics Using SPSS, STATISTICA, MATLAB and R**. New York: Springer, 2007. 519 p.

SALUM, A. D. **Eficiência Energética em Sistema de Combustão de Caldeira**. Universidade Federal da Bahia. Salvador, p. 54. 2011.

SAMAD, T.; MCLAUGHLIN, P.; LU, J. System architecture for process automation: Review and trends. **Journal of Process Control**, v. 17, p. 191-201, 2007.

SANTOS, B. F. D. et al. MONITORING OF BIOSURFACTANT PRODUCTION BY *Bacillus subtilis* USING BEET PEEL AS CULTURE MEDIUM VIA THE DEVELOPMENT

OF A NEURAL SOFT-SENSOR IN AN ELECTRONIC SPREADSHEET. **Brazilian Journal of Chemical Engineering**, v. 35, p. 1678-4383, 2018.

SANTOS, R. B. et al. Detection and on-line prediction of leak magnitude in a gas pipeline using an acoustic method and neural network data processing. **Brazilian Journal of Chemical Engineering**, v. 31, p. 145-153, 2014.

SEBORG, D. E.; EDGAR, T. F.; MELLICHAMP, D. A. **Process Dynamics and Control**. 4ª Edição. ed. [S.l.]: John Wiley & Sons, Inc, 2016.

SENCIO, R. R. **Reconfigurable Control of Processes Subjected to Actuator Faults: a Two-layer MPC-based Approach**. UNICAMP. [S.l.], p. 79. 2017.

SERON, M. M.; DONA, J. A. D. Robust actuator fault compensation accounting for uncertainty in the fault estimation. **International journal of adaptive Control and Signal Processing**, v. 28(12), p. 1440-1453, 2014.

SHEN, Q. K.; JIANG, B.; COCQUEMPOT, V. Adaptive fault tolerant synchronization with unknown propagation delays and actuator faults. **International Journal of Control, Automation and Systems**, v. 10(5), p. 883-889, 2012.

SHIMEURA, E.; FUJITA, M. A. A design method for linear state feedback systems possessing integrity based on a solution of a Riccati-type equation. **International Journal of Control**, v. 42(4), p. 887-899, 1985.

SHU, Y. et al. Abnormal situation management: Challenges and opportunities in the big data era. **Computers and Chemical Engineering**, v. 91, p. 104-113, 2016.

SILJAK, D. D. Reliable control using multiple control systems. **International Journal of Control**, v. 31(2), p. 303-329, 1980.

SILVA, F. S. D. **Estudo de Aplicação da Lignina com Enfoque em Fibra de Carbono**. UFRJ. Rio de Janeiro, p. 245. 2019.

SIMON, L.; KARIM, M. N. Growth-Phase Classification Using Backpropagation and Probabilistic Neural Networks. **IFAC Proceedings Volumes**, v. 32, p. 7568-7572, 1999.

SINGH, A. K.; CHANDRA, R. Pollutants released from the pulp paper industry: Aquatic toxicity and their health hazards. **Aquatic Toxicology**, 211, June 2019. 202 - 216.

SMREKAR, J.; POTOČNIK, P.; SENEGAČNIK, A. Multi-step-ahead prediction of NO<sub>x</sub> emissions for a coal-based boiler. **Applied Energy**, v. 106, p. 89-99, 2013.

SOARES, V. B.; PINTO, J. C.; DE SOUZA JR., M. B. Alarm management practices in natural gas processing plants. **Control Engineering Practice**, p. 185-196, 2016.

SOUZA, F. A. A.; ARAÚJO, R.; MENDES, J. Review of soft sensor methods for regression applications. **Chemometrics and Intelligent Laboratory Systems**, v. 152, p. 69-79, 2016.

SOUZA, F. A. A.; ARAÚJO, R.; MENDES, J. Review of soft sensor methods for regression applications. **Chemometrics and Intelligent Laboratory Systems**, 152, 2016. 69-79.

SPINNER, T.; SRINIYASAN, B.; RENGASWAMY, R. Data-based automated diagnosis and iterative retuning of proportional-integral (PI) controllers. **Control Engineering Practice**, v. 29, p. 23-41, 2014.

STOCK, T.; SELIGER, G. **Opportunities of Sustainable Manufacturing in Industry 4.0**. 13th Global Conference on Sustainable Manufacturing - Decoupling Growth from Resource Use. [S.l.]: [s.n.]. 20016. p. 536 – 541.

SUENATANAKULL, W. **A Comparison of Fault Detection and Classification Using Artificial Neural Networks**. University of Texas. Austin. 1993.

SYMANTEC. Internet Security Threats Report. **Symantec**. Disponível em: <<http://www.symantec.com/threatreport/>>. Acesso em: 02 Março 2017.

TAIRA, K. NO<sub>x</sub> emission profile determined by in-situ gas monitoring of iron ore sintering during packed-bed coke combustion. **Fuel**, 2019. 244-250.

TAMBOURGHI, E. B.; FISCHER, G. A.; FILETI, A. M. F. Neural modeling for cytochrome b5 extraction. **Process Biochemistry**, v. 41, p. 1272–1275, 2006.

TAO, X. et al. Greenhouse gas emission monitoring system for manufacturing prefabricated components. **Automation in Construction**, v. 93, p. 361-374, 2018.

TONG, C. et al. Fault detection and isolation in hybrid process systems using a combined data-driven and observer-design methodology. **AIChE**, v. 60, p. 2805-2814, 2014.

TONG, T. L. C.; HAIZHENYUXUHUASHI. Statistical monitoring for non-Gaussian processes based on MICA-KDR method. **ISA Transactions**, p. 1-33, 2019.

TRAN, H. N. **Recovery Boiler Fireside Deposits and plugging Prevention**. St. Petersburg, p. 537-572. 2007.

TRONCI, S.; BARATTI, R.; SERVIDA, A. Monitoring pollutant emissions in a 4.8MW power plant through neural network. **Neurocomputing**, v. 43, p. 3-15, 2002.

VAKKILAINEN, E. K. **Kraft recovery boilers: Principles and Practice**. Helsinki: [s.n.], 2005. 246 p.

VALDMAN, A. **Sistema de Automação para Monitoramento Online de Gases Residuais e Diagnóstico de uma Caldeira Operada com Misturas Diesel/Biodiesel**. UFRJ. Rio de Janeiro, p. 244. 2010.

VALDMAN, A. **Estrutura Unificada Baseada em Sensor Virtual e Tecnologia Fieldbus para Monitoramento, Diagnóstico e Controle de Caldeiras**. Universidade Federal do Rio de Janeiro. Rio de Janeiro, p. 224. 2013.

VALDMAN, A. et al. A systematic Methodology on Developing an Online Soft Sensor Based on Neural Networks for Monitoring Boiler Gas Emissions. **AIDIC Conference series**, 10, 2011. 353-361.

VALDMAN, A. et al. Online monitoring of exhaust gas emissions of a boiler with diesel/biodiesel fuel blends. **Chemical Engineering Transactions**, v. 24, p. 277-282, 2011.

VARMA, S.; KUMAR, K. D. Fault tolerant satellite attitude control using solar radiation pressure based on nonlinear adaptative sliding mode. **ACTA Astronautica**, v. 66, p. 710-729, 2010.

VÁSQUEZ, J. W. et al. Alarm management based on diagnosis. **IFAC-Papers online**, p. 126-131, 2016.

VENKATASUBRAMANIAN, V. Prognostic and diagnostic monitoring of complex systems for product lifecycle management: challenges and opportunities. **Computers and Chemical Engineering**, v. 29, p. 1253-1263, 2005.

VENKATASUBRAMANIAN, V. The Promise of Artificial Intelligence in Chemical Engineering: Is It Here, Finally? **AIChE Journal**, p. 1-13, 2018.

VENKATASUBRAMANIAN, V.; RENGASWAMY, R.; KAVURI, S. N. A review of process fault detection and diagnosis: part I: quantitative model-based methods. **Computers and Chemical Engineering**, v. 27, p. 293-311, 2003a.

VENKATASUBRAMANIAN, V.; RENGASWAMY, R.; KAVURI, S. N. A review of process fault detection and diagnosis: part II: qualitative models and search strategies. **Computers and Chemical Engineering**, v. 27, p. 313-326, 2003b.

VENKATASUBRAMANIAN, V.; RENGASWAMY, R.; KAVURI, S. N. A review of process fault detection and diagnosis: part III: process history based methods. **Computers and Chemical Engineering**, v. 27, p. 327-346, 2003c.

VENKATASUBRAMANIAN, V.; VIELE, S. R.-P. G. **Artificial Intelligence in Process Systems Engineering: Quo Vadis?** I Congresso Brasileiro em Engenharia de Sistemas em Processos. Rio de Janeiro: [s.n.]. 2019.

VERPOORT, P. C.; DONALD, P. M.; CONDUIT, G. J. Materials data validation and imputation with an artificial neural network. **Computational Materials Science**, v. 147, p. 176-185, 2018.

WANG, J.; CHEN, T. Online reduction of chattering alarms due to random noise and oscillation. **Proceedings of the 19th World Congress**, p. 10604-10609, 2014.

WANG, J.; CHEN, T. Main causes of long-standing alarms and their removal by dynamic state-based alarm systems. **Journal of Loss Prevention in the Process Industries**, p. 106-119, 43 2016.

WANG, J.; LI, H.; HUANG, J. A data similarity based analysis to consequential alarms of industrial processes. **Journal of Loss Prevention in the Process Industries**, v. 35, p. 29-34, 2015.

WANG, S. et al. Towards smart factory for industry 4.0: a self-organized multi-agent system with big data based feedback and coordination. **Computer Networks**, v. 10, p. 158-168, 2016.

WANG, J.-G. et al. Soft-sensing method for optimizing combustion efficiency of reheating furnaces. **Journal of the Taiwan Institute of Chemical Engineers**, p. 1-11, 2016.

WARNE, K. et al. Statistical and computational intelligence techniques for inferential model development: a comparative evaluation and a novel proposition for fusion. **Engineering application of Artificial Intelligence**, 17, n. 8, 2004. 871-885.

WASSERMAN, P. D. **Neural Computing: Theory and Practice**. [S.l.]: Van Nostrand Reinhold, 1989.

WBCSD. **Measuring Eco-Efficiency. A Guide to Reporting Company Performance**. World Business Council for Sustainable Development. [S.l.]. 200.

WEYER, S.; SCHIMITT, M. Towards Industry 4.0-standardization as the crucial challenge for highly modular, multi-vendor production systems. **IFAC-PapersOnline**, v. 48-3, p. 579-584, 2015.

WU, B. et al. Application of artificial neural networks in design of lithium-ion batteries. **Journal of Power Sources**, v. 395, p. 128-136, 2018.

WU, H.; ZHAO, J. Deep convolutional neural network model based chemical process. **Computers and Chemical Engineering**, v. 115, p. 185-197, 2018.

WU, K.; HE, X.; CHEN, B. On-line training method of ANN in DMS for side draw quality of refinery fractionator. **Computers and Chemical Engineering**, v. 24, p. 1585-1589, 2000.

XING-WEI, W. The Fault Diagnosis of Blower Ventilator Based-on Multi-class Support Vector Machines. **2012 International Conference on Future Electrical Power and Energy Systems**, Shenyang, 2012. 1193-1200.

XU, J.; WANG, J.; CHEN, T. Performance assessment and design for univariate alarm systems based on FAR, MAR, and AADD. **Transactions on Automation Science and Engineering**, n. 9, p. 296-307, 2012.

YANG, F. et al. Improved correlation analysis and visualization of industrial alarm data. **ISA Transactions**, v. 51, p. 499-506, 2012.

YANG, Z.; WANG, J.; CHEN, T. Detection of correlated alarms based on similarity coefficients of binary data. **IEEE Transactions on Automation Science and Engineering**, n. 10, p. 1014-1025, 2013.

YE, L. et al. Control/maintenance strategy fault tolerant mode and reliability analysis for hydro power plant stations. **IEEE Transactions on Power Systems**, v. 16, p. 340-345, 2001.

YU, J.; RASHID, M. M. A novel dynamic Bayesian network-based networked process monitoring approach for fault detection, propagation identification, and root cause diagnosis. **AIChE**, v. 59, p. 2348-2365, 2013.

YU, X.; JIANG, J. Hybrid fault-tolerant flight control system design against partial actuator failures. **IEEE Transactions on Control System Technology**, v. 20(4), p. 871-886, 2012.

YU, X.; JIANG, J. A survey of fault-tolerant controllers based on safety-related issues. **Annual Reviews in Control**, v. 39, p. 46-57, 2015.



YU, Y.; WORADECHJUMROEN, D.; YU, D. A review of fault detection and diagnosis methodologies on air-handling units. **Energy and Buildings**, v. 82, p. 550-562, 2014.

YUSTA, J. M.; CORREA, G. J.; LACAL-ARÁNTGUI, O. Methodologies and applications for critical infrastructure protection: State-of-the-art. **Energy Policy**, v. 39, p. 6100-6119, 2011.

ZHANG, S. et al. A data augmentation method based on cycle-consistent adversarial networks for fluorescence encoded microsphere image analysis. **Signal Processing**, v. 161, p. 195-202, 2019.

ZHAN, C.; LI, S.; YANG, Y. Improved process monitoring based on global–local manifold analysis and statistical local approach for industrial process. **Journal of Process Control**, v. 75, p. 107-119, 2019.

ZHANG, H. et al. Fault detection and diagnosis of chemical process using enhanced KECA. **Chemometrics and Intelligent Laboratory Systems**, v. 161, p. 61-69, 2017.

ZHANG, P. et al. **Process Monitoring and Fault Diagnosis - Status and Applications**. Proceedings of the 18th World Congress. Milano, Italy: The International Federation of Automatic Control. 2011.

ZHANG, Y.; JIANG, J. Bibliographical review on reconfigurable fault-tolerant control systems. **Annual Reviews in Control**, v. 32, p. 229-252, 2008.

ZHENG, W. et al. Just-in-time semi-supervised soft sensor for quality prediction in industrial rubber mixers. **Chemometrics and Intelligent Laboratory Systems**, v. 180, p. 36-41, 2018.

ZHOU, H. et al. Modeling NO<sub>x</sub> emissions from coal-fired utility boilers using vector regression with ant colony optimization. **Engineering applications of artificial intelligence**, v. 25, p. 147-158, 2012.

ZHU, J.; ZHAO, Y. S. J.; YANG, F. A dynamic alarm management strategy for chemical process transitions. **Journal of Loss Prevention in Process Industries**, n. 30, p. 207-218, 2014.

ZIEGLER, J. G.; NICHOLS, N. B. Process lags in automatic control circuits. **Transactions of the ASME**, v. 65, p. 433-444, 1943.

ZUG, S.; DIETRICH, A.; KAISER, J. An architecture for a dependable distributed sensor system. **IEEE Transactions on Instrumentation and Measurement**, v. 60(2), p. 408-419, 2011.



## B. Painel Elétrico CLP

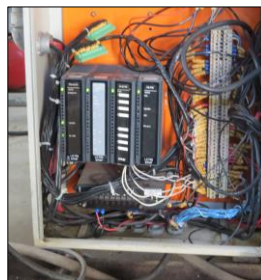
### *B.1 Otimização da Instalação Elétrica do Painel do CLP*

O controlador lógico-programável (CLP) é um equipamento modular que consiste em cartões de processamento agrupados sequencialmente em uma estrutura denominada *rack*. O CLP da planta piloto contém, além dos cartões fonte e CPU, essenciais ao seu funcionamento, cartões de entrada analógica, entrada discreta e um cartão de conversão para o protocolo *Fieldbus*, que disponibiliza as variáveis discretas da planta (Tabela 12) e as variáveis medidas pelo analisador de gases real (Tabela 8) na rede digital *Fieldbus*.

Na montagem original, mostrada na Figura 34, o rack do controlador lógico se encontrava na posição horizontal, que é desfavorável à circulação natural de ar e prejudica fortemente a refrigeração dos componentes eletrônicos. A otimização de toda a parte elétrica foi, então, realizada com o objetivo de evitar o superaquecimento do painel que poderia levar ao mau funcionamento do controlador e em última instância, a curtos elétricos e é mostrada na Figura 35.



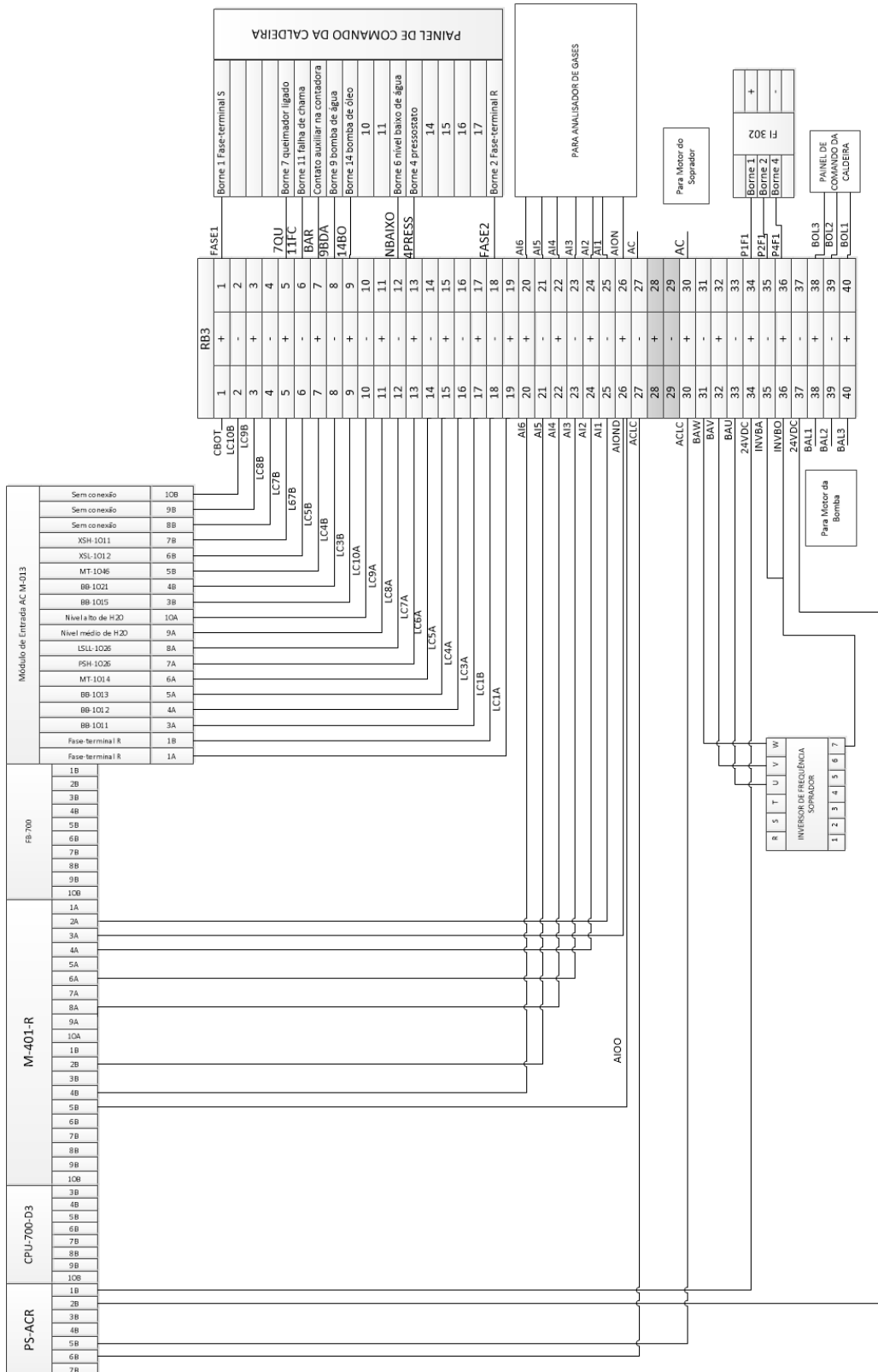
**Figura 34 Montagem original do painel elétrico do CLP**



**Figura 35 Montagem do painel do elétrico do CLP após intervenção**

Esta etapa incluiu a remontagem do barramento elétrico, que compreende instalações mecânicas e elétricas do painel, além da criação da documentação pertinente e padronizada para facilitar atualizações e modificações futuras.

## B.2 Diagrama Elétrico do CLP



### C. Ampliação da Arquitetura de Automação

O sistema de automação da planta foi desenvolvido através da sobreposição de três camadas: física, controle e supervisão. Este projeto envolve cada uma destas camadas e se propõe a incluir a quarta camada de gerenciamento, conforme mostrado na Figura 36. Desta maneira, as informações fluem em um ciclo fechado que se inicia na aquisição de dados pelos instrumentos de campo, atravessa as malhas de controle lógicas e contínuas, chega até o monitoramento das variáveis de processo e integra uma estrutura unificada de gerenciamento das informações.

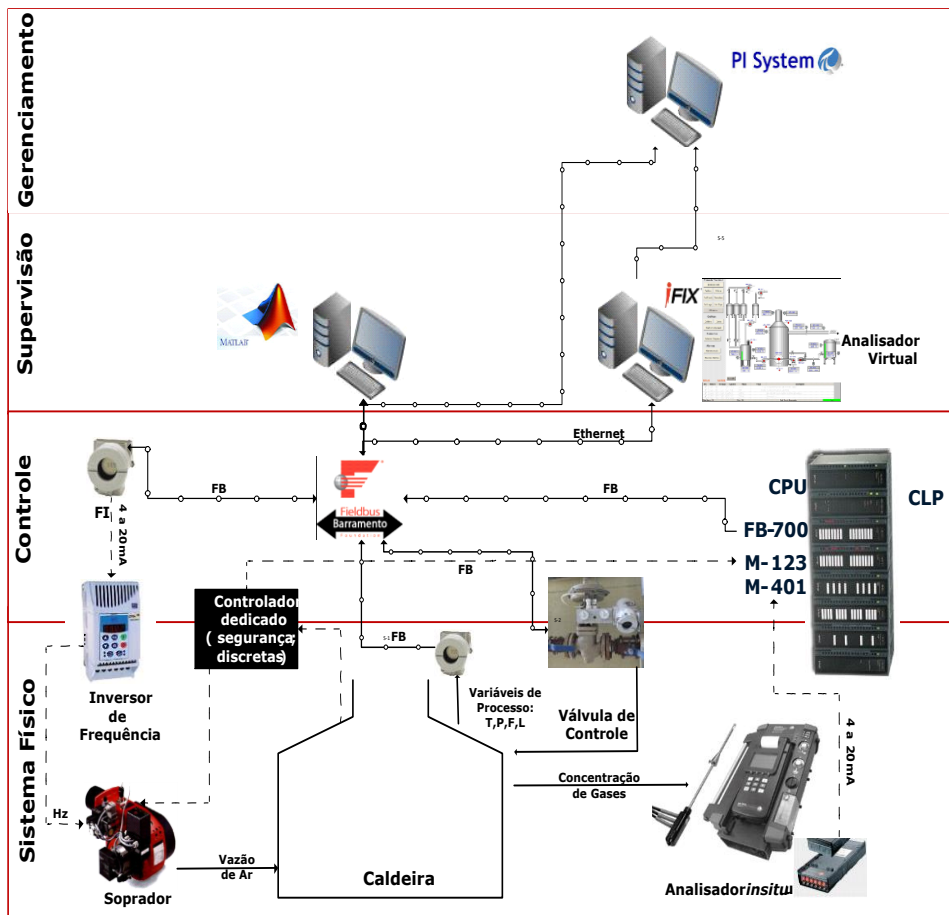


Figura 36 Visão detalhada do sistema de automação da planta piloto

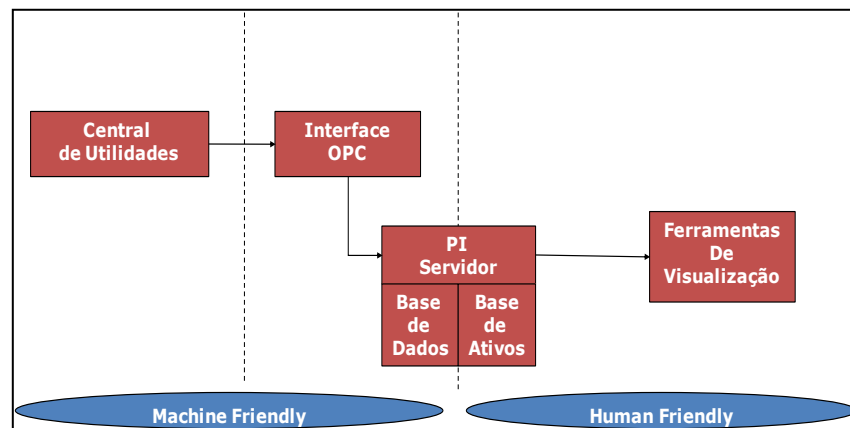
Fonte: Valdman, 2013 – modificada

A ampliação da arquitetura do sistema de automação existente contempla a inclusão de uma camada hierarquicamente superior para fins de gerenciamento da informação em uma

estação PIMS (*Plant Information Management System*). A esta camada compete realizar análises e gerar relatórios sem consumir recursos computacionais dedicados à estação de aquisição e supervisão de dados. Esta nova camada de infraestrutura de dados pode, ainda, integrar dados de diferentes origens para fins de análise, monitoramento e diagnóstico transformando-os em informação útil para decisões de otimização em negócios e em engenharia, em consonância com uma tendência industrial de gerenciamento de ativos.

Neste trabalho, o acompanhamento do processo piloto do ponto de vista de desempenho e de segurança operacional e ambiental será implementado nesta camada; de maneira que possa ser generalizado para outros processos de combustão em um modelo replicável.

Este novo nível será implementado utilizando o software industrial *PI System*® e a aquisição de dados do processo será efetuada em paralelo à aquisição de dados da estação SCADA. Deste modo, o servidor dedicado ao sistema de gerenciamento receberá e arquivará os dados de processo através de uma interface OPC. A partir daí, os dados fluem para a base de ativos, que transforma as *tags* do processo em informações mais amigáveis e contextualizadas para não-especialistas em automação e disponibiliza estes ativos para ferramentas de visualização. O fluxo de informação descrito está representado na Figura 37.



**Figura 37 Fluxo de dados no sistema de gerenciamento de informações**

## D. Outros Trabalhos Publicados

### *D.1 Publicados*

**PARENTE, A. P.**; VALDMAN, A. ; FOLLY, R. O. M. ; SOUZA JR., M. B. . Controle da emissão de gases residuais: um estudo de caso em uma caldeira semi-industrial de automação híbrida. In: XXI Congresso Brasileiro de Engenharia Química, 2016, Fortaleza. Anais do XXI Congresso Brasileiro de Engenharia Química, 2016. v. 1.

**PARENTE, A. P.**; VALDMAN, A. ; FOLLY, R. O. M. ; NASCIMENTO, I. C. S. Power Consumption and Pollutants Emissions in Stationary Sources: A Study Case on a Semi-Industrial Boiler. In: 10th World Congress of Chemical Engineering, 2017, Barcelona. Proceedings of the 10th WCEE, 2017.

**PARENTE, A. P.**; VALDMAN, A.; FOLLY, R. O. M.; SOUZA JR., M. B. Desenvolvimento e Aplicação de Ferramenta de Recalibração Online para Sensores Virtuais In: XXII Congresso Brasileiro de Engenharia Química, 2018, São Paulo. XXII Congresso Brasileiro de Engenharia Química., 2018.

**PARENTE, A. P.**; VALDMAN, A.; FOLLY, R. O. M. ; DE SOUZA JR., M. B.; NASCIMENTO, I. C. S. Automação de Processos Industriais: do Pneumático à Indústria 4.0. Revista Processos Químicos, n. 24, p. 101-108, 2018.