

UNIVERSIDADE ESTADUAL PAULISTA “JULIO DE MESQUITA FILHO”  
FACULDADE DE ENGENHARIA DE BAURU – FEB  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

MARCO AKIO IKESHOJI

DESENVOLVIMENTO DE METODOLOGIA PARA DIAGNÓSTICO DE  
FALHAS INCIPIENTES EM TRANSFORMADORES DE POTÊNCIA  
UTILIZANDO ANÁLISE DE GASES E APRENDIZADO DE MÁQUINA

BAURU/SP  
Agosto/2020

UNIVERSIDADE ESTADUAL PAULISTA “JULIO DE MESQUITA FILHO”  
FACULDADE DE ENGENHARIA DE BAURU – FEB  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

MARCO AKIO IKESHOJI

DESENVOLVIMENTO DE METODOLOGIA PARA DIAGNÓSTICO DE  
FALHAS INCIPIENTES EM TRANSFORMADORES DE POTÊNCIA  
UTILIZANDO ANÁLISE DE GASES E APRENDIZADO DE MÁQUINA

Tese apresentada ao programa de Pós-Graduação em Engenharia Elétrica da Faculdade de Engenharia de Bauru (FEB) – UNESP, para obtenção do título de Doutor em Engenharia Elétrica.

Linha de Pesquisa: Sistemas de Energia

Orientador: Prof. Dr. André Nunes de Souza  
Coorientador: Prof. Dr. Danilo Sinkiti Gastaldello

BAURU/SP  
Agosto/2020

I26d

Ikeshoji, Marco Akio

Desenvolvimento de metodologia para diagnóstico de falhas incipientes em transformadores de potência utilizando análise de gases e aprendizado de máquina / Marco Akio Ikeshoji. -- Bauru, 2020

159 p. : il., tabs.

Tese (doutorado) - Universidade Estadual Paulista (Unesp), Faculdade de Engenharia, Bauru

Orientador: André Nunes de Souza

Coorientador: Danilo Sinkiti Gastaldello

1. Transformadores de potência. 2. Análise de gases dissolvidos. 3. Aprendizado de máquina. 4. Algoritmos inteligentes. 5. Diagnóstico de falhas. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Engenharia, Bauru. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

**ATA DA DEFESA PÚBLICA DA TESE DE DOUTORADO DE MARCO AKIO IKESHOJI, DISCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA, DA FACULDADE DE ENGENHARIA - CÂMPUS DE BAURU.**

Aos 07 dias do mês de agosto do ano de 2020, às 14:00 horas, no(a) Via sistemas de videoconferência e outras ferramentas para comunicação a distância, reuniu-se a Comissão Examinadora da Defesa Pública, composta pelos seguintes membros: Prof. Dr. ANDRE NUNES DE SOUZA - Orientador(a) do(a) Departamento de Engenharia Elétrica / Faculdade de Engenharia de Bauru - UNESP, Prof. Dr. PEDRO DA COSTA JUNIOR do(a) Departamento de Engenharia Elétrica / Faculdade de Engenharia de Bauru - UNESP, Professor Associado JOAO PAULO PAPA do(a) Departamento de Computação / Faculdade de Ciências de Bauru - UNESP, Prof<sup>a</sup>. Dr<sup>a</sup>. TALITA MAZON do(a) Ministério da Ciência e Tecnologia / Centro de Tecnologia da Informação Renato Archer - CTI, Professor Titular OSVALDO RONALD SAAVEDRA MENDEZ do(a) Departamento de Engenharia de Elétrica / Universidade Federal do Maranhão - UFMA, sob a presidência do primeiro, a fim de proceder a arguição pública da TESE DE DOUTORADO de MARCO AKIO IKESHOJI, intitulada **DESENVOLVIMENTO DE METODOLOGIA PARA DIAGNÓSTICO DE FALHAS INCIPIENTES EM TRANSFORMADORES DE POTÊNCIA UTILIZANDO ANÁLISE DE GASES E APRENDIZADO DE MÁQUINA**. Após a exposição, o discente foi arguido oralmente pelos membros da Comissão Examinadora, tendo recebido o conceito final: **APROVADO** . Nada mais havendo, foi lavrada a presente ata, que após lida e aprovada, foi assinada pelos membros da Comissão Examinadora.

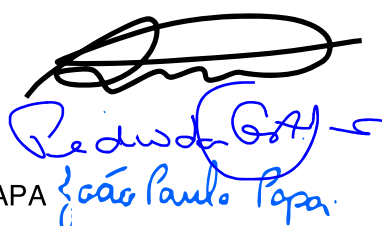
Prof. Dr. ANDRE NUNES DE SOUZA

Prof. Dr. PEDRO DA COSTA JUNIOR

Professor Associado JOAO PAULO PAPA

Prof<sup>a</sup>. Dr<sup>a</sup>. TALITA MAZON

Professor Titular OSVALDO RONALD SAAVEDRA MENDEZ



## AGRADECIMENTOS

A Deus, que sempre guiou meus passos ...

Aos meus pais, Akito (*in memorian*) e Sumiko, e irmãs Célia e Rosemeire, pelo amor e apoio em todas as fases da minha vida.

A minha esposa Elisangela e filhas Helen e Júlia, meus eternos amores, por estarem ao meu lado, alegrando-me e incentivando-me a sempre superar os desafios que surgem.

Ao meu orientador Prof. Dr. André Nunes de Souza pela confiança, paciência e profissionalismo durante a orientação deste trabalho.

Ao meu coorientador Prof. Dr. Danilo Sinkiti Gastaldello, pelo acompanhamento e efetiva participação no desenvolvimento deste trabalho.

Ao Carlos Guilherme Gonzales, pela amizade e contribuições técnicas neste trabalho.

Aos Professores Dr. João Paulo Papa e Dr. Pedro da Costa Junior pela amizade e ajuda sempre que necessária.

Ao Luís, Marcos e pesquisadores do laboratório RECOGNA, pela amizade, simpatia e as valiosas trocas de informações.

Ao Gustavo, André, Haroldo, e demais colegas do laboratório LSISPOTI, pela amizade e companheirismo.

Ao Programa de Pós-Graduação em Engenharia Elétrica, que direta ou indiretamente contribuiu com a realização deste trabalho.

IKESHOJI, Marco Akio. DESENVOLVIMENTO DE METODOLOGIA PARA DIAGNÓSTICO DE FALHAS INCIPIENTES EM TRANSFORMADORES DE POTÊNCIA UTILIZANDO ANÁLISE DE GASES E APRENDIZADO DE MÁQUINA. 2020. 159 p. Tese (Doutorado em Engenharia Elétrica) – Faculdade de Engenharia de Bauru, Unesp, Bauru, 2020.

## RESUMO

O transformador de potência é um elemento essencial para funcionamento e estabilidade do sistema elétrico. Sua saída repentina de operação por falha, ocasiona desde transtornos sociais, a prejuízos financeiros proporcionais ao tempo de atendimento e correção da interrupção no fornecimento de energia. Descobrir de modo incipiente falhas em transformadores, além de evitar esses problemas, contribui para o aumento da segurança e confiabilidade do sistema elétrico. Nesse contexto, esta tese apresenta uma metodologia que emprega os algoritmos Floresta de Caminhos Ótimos e K Vizinhos mais Próximos para auxílio na gestão da manutenção desses equipamentos que, diferentemente dos métodos convencionais e inteligentes de análise de gases que apenas classificam padrões, utilizam uma nova abordagem no diagnóstico de falhas com uso dos registros (recentes e históricos) de análises dos gases dissolvidos no óleo, tornando o método mais informativo, uma vez que permite observar a evolução dos eventos e distinguir tendências de falhas. Assim, diferentes estratégias e ações programadas de manutenção podem ser elaboradas, reduzindo-se o risco de interrupções de energia repentinas e suas consequências. Portanto, este trabalho inicialmente revê os custos estimados nos casos de falhas, apresenta as estatísticas dos tipos e localizações delas, as características básicas do sistema isolante e como os gases são gerados devido aos estresses térmicos e elétricos. Em seguida são apresentadas as técnicas normatizadas (convencionais) e inteligentes de classificações empregadas na análise de falhas com gases dissolvidos, algumas técnicas de otimização e de seleção de atributos, e uma proposta alternativa de diagnóstico de falhas. Posteriormente, é mostrado um método de geração de dados sintéticos baseado em incertezas de medições, segundo a IEC, para o caso de dados escassos, bem como são feitas simulações e análises com dados reais da concessionária de energia para validar a metodologia proposta, por meio da qual foi possível obter resultados promissores na identificação de falhas.

**Palavras-chave:** Transformadores de Potência. Análise de Gases Dissolvidos. Aprendizado de Máquina. Algoritmos Inteligentes. Diagnóstico de Falhas.

IKESHOJI, Marco Akio. DEVELOPMENT OF METHODOLOGY FOR INCIPIENT FAULT DIAGNOSIS IN POWER TRANSFORMERS USING GAS ANALYSIS AND MACHINE LEARNING. 2020. 159 p. Thesis (Doctorate) – Bauru Electrical Engineering College, Unesp, Bauru, 2020.

## ABSTRACT

The power transformer is an essential element in the operation and stability of the electric power system. Its unexpected exit from the power system causes severe social and economic problems for all involved (energy companies and users). Thus, discovering incipient transformer failures avoids the problems mentioned earlier and increases the electric power system's safety and reliability. Within this context, this thesis proposes the application of the Optimal-Path Forest and K-Nearest Neighborhood algorithms to support transformer's maintenance management. Unlike conventional and intelligent gas analysis techniques that only classify patterns, this method uses a new approach of fault diagnosis in the oil-filled transformers that work with continuous records (recent and historical data) of dissolved gas analysis, making the method more informative as it allows observing trends or the evolution of initials faults. With this information, it is possible to determine strategies and programmed actions for transformer maintenance, thus reducing the sudden outages and consequences. Therefore, this work initially reviews the estimated costs involved in the case of transformers faults, presents statistics of the types and fault locations, essentials characteristics of the insulation system and how thermal and electrical stresses generate the gases. Subsequently, the conventional and intelligent classifications techniques for analysis of incipient faults using dissolved gases are presented, as well as the new proposal for fault diagnosis. Finally, a method of generating synthetic data based on IEC measurement uncertainties in case of scarce data is presented, as well as simulations and analyzes are performed with real data from energy utility to validate the proposed approach, achieving excellent results in the fault trend analysis.

**Keywords:** Power Transformers. Dissolved Gas Analysis. Machine Learning. Intelligent Algorithms. Faults Diagnosis.

## LISTA DE FIGURAS

Figura 1 – Estrutura básica do transformador de potência a óleo .....	26
Figura 2 – Concentrações relativas de gases vs. temperatura e tipos de falhas.....	36
Figura 3 – Estrutura de tomada de decisão para transformador em operação.....	43
Figura 4 – Estrutura de decisão de contingenciamento para transformador em operação .....	44
Figura 5 – Conceito do Modelo IFTM .....	45
Figura 6 – Fluxograma de ações originadas do HI.....	46
Figura 7 – Diagrama simplificado de análise de criticidade .....	75
Figura 8 – Diagrama em blocos do método de análise e diagnóstico de falhas proposto .....	77
Figura 9 – Variações dos gráficos de falhas .....	79
Figura 10 – Espaços de características constituídos por hiperplanos e grafos de classificadores .....	86
Figura 11 – Gráfico de evolução e tendência das classes de falhas em relação ao tempo .....	87
Figura 12 – Representação gráfica ampliada da distribuição das amostras DGA via PCA .....	92
Figura 13 – Taxa de acerto média para dados de DGA parciais .....	93
Figura 14 – Componentes principais PC1 e PC2 do conjunto de amostras sintéticas de DGA .....	99
Figura 15 – Curvas de concentrações de gases e tendências de falhas no transformador TR1 .....	105
Figura 16 – TR1: (a) Vestígios carbonização óleo cilindro; (b) Aquecimento cordoalha comutador .....	106
Figura 17 – Curvas de concentrações de gases e tendências de falhas no transformador TR2 .....	107
Figura 18 – TR2: (a) Vestígios carbonização óleo cabeçote; (b) Acionamento chave comutadora .....	109
Figura 19 – Curvas de concentrações de gases e tendências de falhas no transformador TR3 .....	110
Figura 20 – TR3: (a) Vestígios carbonização óleo cabeçote; (b) Acionamento chave comutadora .....	112



Figura 21 – Curvas de concentrações de gases e tendências de falhas no transformador TR4 .....	114
Figura 22 – TR4: (a) Isolação degradada; (b) Sedimentos; (c) Porca solta; (d) Fusão metal .....	115
Figura 23 – Curvas de concentrações de gases e tendências de falhas no transformador TR5 .....	117
Figura 24 – TR5: (a) Aquecimento contato comutador; (b) e (c) Aquecimento cabo aterramento .....	119
Figura 25 – Curvas de concentrações de gases e tendências de falhas no transformador TR6 .....	120
Figura 26 – TR6: (a) Resíduo de óleo carbonizado no interior núcleo; (b) Sedimentos sobre núcleo .....	122
Figura 27 – Curvas de tendências de falhas com dados artificiais transformador TR2 .....	125
Figura 28 – Curvas de tendências de falhas com dados artificiais transformador TR3 .....	125
Figura 29 – Chave comutadora: (a) Localização; (b) Chave comutadora sem carga.....	146
Figura 30 – Buchas e derivações no transformador .....	147
Figura 31 – Triângulo de Duval: (a) N°1 falhas elétricas e térmicas; (b) N°4 falhas de baixa temperatura; e (c) N°5 falhas de alta temperatura .....	149
Figura 32 – Curvas de concentrações de gases e tendências de falhas no transformador TR7 .....	154
Figura 33 – Curvas de concentrações de gases e tendências de falhas no transformador TR8 .....	155
Figura 34 – Curvas de concentrações de gases e tendências de falhas no transformador TR9 .....	157
Figura 35 – Curvas de concentrações de gases e tendências de falhas no transformador TR10 .....	158

## LISTA DE TABELAS

Tabela 1 – Taxas de falhas em população de transformadores de 21 países.....	30
Tabela 2 – Tipos de falhas vs. relevância dos gases produzidos.....	38
Tabela 3 – Características dos algoritmos evolutivos e de inteligência coletiva.....	66
Tabela 4 – Valores limites de detecção de $S_{LIM}$ em DGA.....	72
Tabela 5 – Valores limites de concentrações de gases para amostras normais ( $\mu\text{l/l}$ ou PPM) .....	94
Tabela 6 – Taxa de acertos KNN, OPF, kOPF, NB, SVM, ANN e ELM, para dados brutos e pré-processados .....	95
Tabela 7 – Seleção de atributos para dados pré-processados para 5 tipos de falhas .....	96
Tabela 8 – Resultados obtidos para simulações após seleção dos melhores 8 atributos .....	97
Tabela 9 – Matrizes de confusões/erros para dados DNP e brutos para 2, 3 e 5 classes.....	98
Tabela 10 – Matrizes de confusões médias para dados sintéticos e diferentes classificadores .....	100
Tabela 11 – Identificação dos transformadores de potência avaliados .....	103
Tabela 12 – Matriz de confusão ou erro .....	148
Tabela 13 – Relação dos gases dissolvidos no óleo para o método de Rogers .....	151
Tabela 14 – Relação dos gases dissolvidos no óleo IEC 60599 – Transformadores.....	151
Tabela 15 – Limite L1 de concentrações de gases da IEEE C57.104 para o método Doernenburg .....	152
Tabela 16 – Relação dos gases dissolvidos no óleo para o método Doernenburg .....	152

## LISTA DE ABREVIATURAS E SIGLAS

ABNT	Associação Brasileira de Normas Técnicas
ACO	Ant Colony Optimization
ACS	Ant Colony System
AIS	Artificial Immunity System
ANFIS	Adaptive Neuro-Fuzzy Inference System
ANN	Artificial Neural Network
AS	Ant System
ASTM	American Society for Testing and Materials
BAT	Bat Algorithm
BPNN	Backpropagation Neural Network
CPANN	Counter Propagation Artificial Neural Network
CS	Cuckoo Search
DAI	Descarga de Alta Intensidade
DGA	Dissolved Gas Analysis
DP	Descarga Parcial
DPBI	Descarga Parcial de Baixa Intensidade
DPy	Degree of Polymerization
EAM	Enterprise Asset Management
ELM	Extreme Learning Machine
ERP	Enterprise Resource Planning
FCM	Fuzzy C-means
FRA	Frequency Response Analysis
KFCM	Kernel Fuzzy C-Means
kOPF	Kernel Optimum-Path Forest
K-OPF	K Optimum-Path Forest
FSVM	Fuzzy Support Vector Machine
GA	Genetic Algorithm
GWO	Gray Wolf Optimization
GRNN	Generalized Regression Neural Network
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronic Engineers
ITMS	Information Theoretic Mean Shift

KNN	K-Nearest Neighborhood
MLP	Multilayer Perceptron
MMAS	Min-Max Ant System
MST	Minimum Spanning Tree
NB	Näive Bayes
NBR	Norma Brasileira
OPF	Optimum-Path Forest
OPFL	Optimal Power Flow
OVO	One Versus One
PCA	Principal Component Analysis
PPM	Part Per Million
PSO	Particle Swarm Optimization
PV	Percentage Loss of Life
RAP	Regulatory Account Principles
RBF	Radial Basis Function
RBFN	Radial Basis Function Network
RST	Rough Set Theory
SaE	Self-adaptative Evolutionary
SCADA	Supervisory Control and Data Acquisition
SVM	Support Vector Machine
TDCG	Total Dissolved Combustible Gas
WNN	Wavelet Neural Network

## SIMBOLOGIA

$\alpha$	Parâmetro de controle do feromônio algoritmo ACO
$\alpha_p$	Passo do algoritmo CS
°C	Graus centígrados
$C_{class}$	Número total de classes ou padrões
$C_p$	Possível carbonização do papel
$C_{reg}$	Parâmetro de regularização <i>kernel</i>
$C_{st}$	Custo
$C_{ya}$	Custo ya entre nova amostra e a amostra rotulada do espaço de atributos
$C_{yb}$	Custo yb entre nova amostra e a amostra rotulada do espaço de atributos
$C_{yc}$	Custo yc entre nova amostra e a amostra rotulada do espaço de atributos
$CH_4$	Metano
$C_2H_2$	Acetileno
$C_2H_4$	Etileno
$C_2H_6$	Etano
CO	Monóxido de carbono
CO <sub>2</sub>	Dióxido de carbono
$c_1$	Constante de aceleração cognitiva PSO
$c_2$	Constante de aceleração social PSO
d	Distância euclidiana
D1	Descarga de baixa energia
D2	Descarga de alta energia
$\delta$	Função delta Dirac
$e$	Função exponencial
FN	Número de amostras falso negativo
FP	Número de amostras falso positivo
$\phi$	Função <i>kernel</i>
$\gamma$	Coefficiente <i>kernel</i>
H <sub>2</sub>	Hidrogênio
i	Variável inteira
j	Variável inteira
K	Número de vizinhos mais próximos

$K_C$	Número de grupos ( <i>clusters</i> ) de diferentes classes
kV	Quilovolt ou 1000 Volts
$L_{sup}$	Limite superior binário (1) para seleção de atributos
$L_{inf}$	Limite inferior binário (0) para seleção de atributos
$L_1$	Limite de concentração dos gases resultante da degradação do óleo
$\lambda$	Coefficiente Levy para caminhada aleatória
$m$	Números de atributos
$\mu\text{l/l}$	Microlitro por litro
$\mu\text{s}$	Microsegundos
$n$	Número de amostras conjunto treinamento
$\mathcal{N}$	Número máximo de atualizações
$nn$	Vizinho mais próximo
NT	Número total de amostras por classe
$N_2$	Nitrogênio
$O_2$	Oxigênio
$O_{PO}$	Superaquecimento do papel ou óleo
$p_a$	Probabilidade de ocorrência entre [0,1]
$P_1$	Posição 1 da amostra no espaço de atributos
$P'_1$	Posição intermediária 1' da amostra no espaço de atributos
$P_2$	Posição 2 da amostra no espaço de atributos
$P'_2$	Posição intermediária 2' da amostra no espaço de atributos
$P_3$	Posição 3 da amostra no espaço de atributos
$P'_3$	Posição intermediária 3' da amostra no espaço de atributos
$P_4$	Posição 4 da amostra no espaço de atributos
$P'_4$	Posição intermediária 4' da amostra no espaço de atributos
PC1	Componente principal 1
PC2	Componente principal 2
PC3	Componente principal 3
$R_C$	Reações catalíticas
$R_1$	Razão entre metano e hidrogênio
$R_2$	Razão entre acetileno e etileno
$R_3$	Razão entre acetileno e metano
$R_4$	Razão entre etano e acetileno

$R_5$	Razão entre etileno e etano
$R_6$	Razão entre etano e metano
$\Re$	Números reais
$\rho$	Taxa de evaporação do feromônio algoritmo ACO
$s$	Qualquer amostra do conjunto de teste
$s'$	Qualquer nova amostra do conjunto de teste
$S$	Conjunto de amostras protótipos
$S_G$	Gases parasitas ( <i>Stray gassing</i> )
$S_{LIM}$	Limite de sensibilidade de detecção analítico
$\sigma_{boolean}$	Limite booleano (0,1)
$t$	Qualquer amostra do conjunto de treino
$T$	Temperatura
$t_p$	Intervalo de tempo
$t'$	Qualquer nova amostra do conjunto de treino
$T1$	Falha térmica em temperatura $T < 300^\circ\text{C}$
$T2$	Falha térmica em temperatura entre $300^\circ\text{C} < T < 700^\circ\text{C}$
$T3$	Falha térmica em temperatura $T > 700^\circ\text{C}$
$\theta$	Conjunto total de classes (ou rótulos) de amostras
$\tau_{inicial}$	Feromônio inicial algoritmo ACO
$U$	Classe de tensão
$VP$	Número de amostras verdadeiro positivo
$VN$	Número de amostras verdadeiro negativo
$\vec{v}$	Vetor de atributos de amostra individual
$\vec{V}$	Conjunto de vetores de atributos
$x$	Amostra do conjunto de treino
$\mathfrak{X}$	Conjunto de soluções para população de partículas
$x_{des}$	Amostra desconhecida
$x_p$	Eixo das abcissas da representação do espaço de atributos
$y$	Rótulo da amostra do conjunto de treinamento
$\mathcal{Y}$	Vetor binário das soluções para população de partículas
$\mathcal{Y}$	Vetor binário equivalente para população de partículas
$y_{desc}$	Classe (ou rótulo) da amostra desconhecida
$y_p$	Eixo das ordenadas da representação do espaço de atributos

$Z$	Conjunto total de amostras
$Z_{desc}$	Conjunto de amostras desconhecidas
$Z_1$	Conjunto de amostras de treino do algoritmo de aprendizado de máquina
$Z_2$	Conjunto de amostras de teste ou validação do algoritmo de aprendizado de máquina
$\omega_{inicial}$	Coefficiente de inércia inicial algoritmo PSO
$\omega_{correção}$	Coefficiente de correção algoritmo PSO



## SUMÁRIO

<b>CAPÍTULO 1 – INTRODUÇÃO .....</b>	<b>20</b>
1.1. Contexto .....	20
1.2. Relevância do tema.....	21
1.3. Objetivos da tese.....	22
1.4. Metodologia da pesquisa.....	23
1.5. Originalidade .....	23
1.6. Publicações relacionadas à tese .....	24
1.7. Organização do documento .....	25
<b>CAPÍTULO 2 – ESTUDO SOBRE FALHAS EM TRANSFORMADORES.....</b>	<b>26</b>
2.1. Caracterização dos transformadores de potência .....	26
2.2. Estimativa de custos que envolvem a falha do transformador .....	27
2.3. Tipo, localização e estatística de falha nos transformadores.....	28
2.4. Caracterização do sistema isolante.....	31
2.4.1. O papel isolante.....	32
2.4.2. O óleo mineral isolante.....	33
2.5. Degradação do sistema isolante e formação dos gases.....	34
2.6. Análises e aquisições de dados <i>online</i> e <i>offline</i> .....	39
2.7. Análise de criticidade envolvendo falhas em transformadores.....	41
2.8. Considerações gerais do capítulo .....	47
<b>CAPÍTULO 3 – ESTUDOS SOBRE TÉCNICAS DE DIAGNÓSTICO DE FALHAS EM TRANSFORMADORES E APRENDIZADO DE MÁQUINA .....</b>	<b>49</b>
3.1. Métodos e técnicas normatizadas e convencionais.....	49
3.1.1. Gás chave .....	49
3.1.2. Triângulo de Duval.....	50
3.1.3. Razões dos gases combustíveis .....	50

3.1.3.1. Rogers.....	51
3.1.3.2. Doernenburg.....	51
3.1.3.3. IEC 60599 e NBR 7274.....	52
3.1.4. Outras técnicas convencionais de DGA.....	52
<b>3.2. Métodos e técnicas de inteligência computacional.....</b>	<b>53</b>
3.2.1. Lógica <i>Fuzzy</i> .....	54
3.2.2. Redes Neurais Artificiais (ANN).....	56
3.2.3. Máquina de Aprendizado Extremo (ELM).....	58
3.2.4. Máquina de Vetor Suporte (SVM).....	59
3.2.5. Nãive Bayes (NB).....	61
3.2.6. K Vizinhos mais Próximos (KNN).....	61
3.2.7. Floresta de Caminhos Ótimos (OPF).....	63
<b>3.3. Métodos de otimização.....</b>	<b>64</b>
<b>3.4. Pré-processamento de dados.....</b>	<b>66</b>
3.4.1. Limpeza de dados ( <i>outliers</i> ).....	66
3.4.2. Construção de atributos.....	68
3.4.3. Seleção de atributos.....	69
3.4.4. Dados sintéticos de DGA.....	71
<b>3.5. Considerações gerais do capítulo.....</b>	<b>72</b>
<b>CAPÍTULO 4 – PROPOSTA DE ABORDAGEM NO DIAGNÓSTICO DE FALHAS EM TRANSFORMADORES.....</b>	<b>75</b>
<b>4.1. Limpeza do conjunto de dados de gases.....</b>	<b>79</b>
<b>4.2. Construção e seleção de atributos.....</b>	<b>80</b>
<b>4.3. Reconhecimento de padrões com KNN.....</b>	<b>81</b>
<b>4.4. Reconhecimento de padrões com OPF.....</b>	<b>82</b>
<b>4.5. Avaliação de conjunto com dados desbalanceados.....</b>	<b>83</b>
<b>4.6. Geração de dados sintéticos de DGA.....</b>	<b>84</b>

<b>4.7.</b>	<b>Proposta de metodologia de análise de tendência de falhas em transformadores</b>	<b>85</b>
<b>4.8.</b>	<b>Considerações gerais do capítulo</b>	<b>90</b>
<b>CAPÍTULO 5 – RESULTADOS EXPERIMENTAIS E DISCUSSÕES</b>		<b>92</b>
<b>5.1.</b>	<b>Base de dados de DGA</b>	<b>92</b>
<b>5.2.</b>	<b>Pré-processamento do conjunto de dados de treinamento do DGA</b>	<b>94</b>
<b>5.3.</b>	<b>Análise do histórico de DGA usando a abordagem desenvolvida e dados reais</b>	<b>102</b>
5.3.1.	Transformador TR1	104
5.3.2.	Transformador TR2	107
5.3.3.	Transformador TR3	110
5.3.4.	Transformador TR4	113
5.3.5.	Transformador TR5	116
5.3.6.	Transformador TR6	120
<b>5.4.</b>	<b>Sugestão de abordagem no tratamento da interpretação de diagnóstico proposto</b>	<b>122</b>
<b>5.5.</b>	<b>Análise do histórico de DGA usando a abordagem desenvolvida e dados sintéticos</b>	<b>124</b>
<b>5.6.</b>	<b>Considerações gerais do capítulo</b>	<b>126</b>
<b>CAPÍTULO 6 – CONCLUSÃO E TRABALHOS FUTUROS</b>		<b>128</b>
<b>6.1.</b>	<b>Aspectos gerais</b>	<b>128</b>
<b>6.2.</b>	<b>Futuros desdobramentos da pesquisa</b>	<b>130</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS</b>		<b>131</b>
<b>APÊNDICE A - DEFINIÇÕES</b>		<b>146</b>
<b>APÊNDICE B – TABELAS E FIGURAS DOS MÉTODOS NORMATIZADOS</b>		<b>149</b>
<b>APÊNDICE C – ANÁLISE DE OUTROS TRANSFORMADORES</b>		<b>153</b>

## CAPÍTULO 1 – INTRODUÇÃO

### 1.1. Contexto

O crescimento socioeconômico de um país determina o aumento na demanda de energia e conseqüentemente, da infraestrutura que a transporta, seja na forma de implantação de novas subestações e de linhas de transmissões e distribuições, e nas melhorias das redes elétricas existentes das quais os transformadores fazem parte.

Os transformadores de potência a óleo são ativos de altos valores agregados, fundamentais para viabilização e funcionamento da infraestrutura de transmissão e distribuição de energia elétrica (BUSTAMANTE et al., 2019; USDE, 2012). Na ocorrência de falha com eles, impactos socioeconômicos negativos nas formas de multas, processos, parada de produção, insegurança, contaminação ambiental, entre outros, são gerados para fornecedores e consumidores de energia (ANEEL, 2018ab; ALMEIDA; DELGADO; MOREIRA, 2003, FILHO, 2012; LIMA; SAAVEDRA; MIRANDA, 2015; SUN; HUANG; HUANG, 2012;). Esses aspectos têm influenciado nos estudos e determinações de metodologias, com abordagens preditivas e preventivas de manutenção, que visem reduzir as causas e efeitos de uma saída de serviço não programada. Assim, conhecer a saúde do transformador e suas possíveis falhas iniciais, tem sido objeto de grande interesse entre pesquisadores e empresas de energia elétrica, como pode ser visto em Cheng e Yu (2018), sobretudo do ponto de vista do aumento de segurança e continuidade do fornecimento da energia.

As falhas nos transformadores podem ter diversas origens, logo, vários também são os ensaios e testes disponíveis para detectá-los. Dentre essas origens, os estresses térmicos e elétricos, provocados por determinados tipos de eventos, que resultam na degradação do sistema isolante óleo e papel, permitem serem diagnosticados em seu estágio inicial (IEEE, 2019; IEC, 2015). Tais ensaios ou testes podem ser de natureza elétrica (relação de transformação, isolamento, entre outros), físico-químico (rigidez dielétrica, impurezas, entre outros) e dos gases dissolvidos no óleo (BECHARA, 2010; SOKOLOV, 2005), sendo este último uma excelente fonte de informação para avaliar a saúde operacional do transformador e o seu sistema isolante (KAUR; BRAR; LEENA, 2019; LIEPNIECE; VITOLINA; MARKS, 2017; DUVAL, 2012), com a vantagem adicional de ser possível realizar a análise com o equipamento em serviço no sistema elétrico.

Nesse contexto, diversas técnicas e algoritmos têm sido desenvolvidos com o auxílio de inteligência computacional, para obter um diagnóstico seguro e eficaz sobre a saúde do transformador a óleo (CHENG; YU, 2018), entretanto, sua aplicação imediata nas concessionárias de energia ainda não é comum, em virtude do tradicional uso dos métodos analíticos normatizados, como por exemplo, razões dos gases, gás chave, percentual relativo gráfico (triângulo, pentágono), entre outros, na avaliação de transformadores (NETO; ASSUNÇÃO; ASSUNÇÃO, 2009).

## **1.2. Relevância do tema**

Mesmo que os transformadores apresentem elevado grau de confiabilidade, a vida útil desses equipamentos é naturalmente limitada (há degradação natural do sistema isolante) e sua falha desencadeia uma série de consequências que nem sempre, as concessionárias de energia estão preparadas em recursos materiais e tempo de reparo para uma solução imediata. Assim, diante de um elevado número de transformadores existentes nas instalações das concessionárias de energia, reduzir a possibilidade de ocorrências de eventos de falhas nesse tipo de equipamento contribui diretamente para o aumento da segurança e confiabilidade do sistema elétrico, além de evitar custos adicionais indesejados aos consumidores e fornecedores de energia.

As causas das falhas em transformadores podem ter origem na ação de agentes externos (curto-circuito externo, sobretensões, descargas elétricas), internos (corrosão, falhas em elemento móveis, conexões), defeito de fabricação, envelhecimento, falta de manutenção e muitos outros. Descartando-se a possibilidade de uma falha repentina devido a uma ação externa ou até mesmo interna, alguns desses fatores podem ser monitorados e seus comportamentos analisados e diagnosticados quanto ao aparecimento de falha incipiente, permitindo assim avaliar o estado do equipamento e minimizar essas ocorrências, com a correção em tempo hábil (BECHARA, 2010).

Sob essa perspectiva, desenvolveu-se a motivação para a realização de uma investigação de técnicas atuais relacionadas ao aprendizado de máquina associados ao diagnóstico de falhas em transformadores, buscando articulá-las e assim propor um método de análise que auxilie num diagnóstico mais seguro de modo a:

- Contribuir para o aumento da confiabilidade do sistema de transmissão e distribuição de energia;

- Reduzir custos não previstos, ocasionados por uma parada não planejada do transformador;
- Proporcionar um aumento de informações que contribuam para o planejamento antecipado de novas ações de manutenção; e
- Melhorar a eficácia na classificação de falhas, empregando os algoritmos de aprendizado de máquina.

### 1.3. Objetivos da tese

Desenvolver uma metodologia que auxilie no diagnóstico de falhas incipientes dos transformadores de potência a óleo, empregando algoritmos de aprendizado de máquina e análise dos gases dissolvidos. Logo, para o desenvolvimento da tese foram definidos os seguintes objetivos específicos:

- Estudar os tipos de falhas, ou seja, como surgem, seus indicadores e influências na saúde do transformador, sob o ponto de vista de análise dos gases dissolvidos no óleo isolante;
- Estudar as técnicas e metodologias mais empregadas no diagnóstico de falhas que envolvem os transformadores; e
- Propor uma metodologia que envolva esses algoritmos e as análises dos gases que permitam observar e avaliar as tendências das falhas incipientes.

Para atingir os objetivos específicos, as atividades de estudos e pesquisas foram organizadas em:

- Revisão bibliográfica relacionada ao tema da tese;
- Levantamento e obtenção de dados;
- Análise de algoritmos de aprendizado de máquina: Redes Neurais Artificiais (ANN – *Artificial Neural Network*), Máquina de Aprendizado Extremo (ELM – *Extreme Learning Machine*), Máquinas de Vetor Suporte (SVM – *Support Vector Machine*), Floresta de Caminhos Ótimos (OPF – *Optimum-Path Forest*), Bayes (NB – *Näive Bayes*) e K vizinhos mais próximos (KNN – *K Nearest Neighborhood*);
- Estudo de métodos de otimizações associados aos algoritmos de aprendizado de máquinas; e
- Simulações computacionais com dados reais e análise de resultados.

#### 1.4. Metodologia da pesquisa

Para o desenvolvimento desta tese foram pesquisadas as seguintes bases de dados:

- Scopus Elsevier;
- *Institute of Electrical and Electronic Engineers – IEEE*;
- Catálogo de teses e dissertações da Capes;
- Portal Periódico Capes; e
- Instituto Brasileiro de Informação em Ciência e Tecnologia – IBICT.

Os parâmetros empregados foram descritores e análises de resumos e títulos. Dentre os descritores empregados na pesquisa, incluiu-se também a tradução para o português de: *Power Transformer, Dissolved Gas Analysis, Machine Learning, Artificial Intelligence, Pattern Recognition, Intelligent algorithm, Fault Diagnosis, Optimum-Path Forest, Forecasting*. Assim, a partir dos descritores individuais, pesquisou-se as seguintes combinações:

- *Power Transformer and Fault Diagnosis*;
- *Power Transformer and Dissolved Gas Analysis*;
- *Power Transformer and Machine Learning*;
- *Power Transformer and Artificial Intelligence*;
- *Power Transformer and Intelligence Algorithm*;
- *Power Transformer and Optimum-Path Forest*;
- *Power Transformer and Forecasting*; e
- *Dissolved Gas Analysis and Forecasting*.

#### 1.5. Originalidade

Os estudos realizados sobre diagnósticos de falhas incipientes em transformadores que empregam análises dos gases dissolvidos no óleo têm estado em constante pesquisa, por meio da busca de:

- Métodos científicos que melhorem o desempenho na classificação de falhas;
- Conhecimento do grau de confiança na classificação com uso da estatística;
- Previsão da evolução dos gases e falhas;
- Na estimativa da vida útil remanescente; e
- Análise de criticidade e risco.

Nesta tese, uma forma de análise empregando algoritmos classificadores baseados na teoria dos grafos é apresentada. Diferentemente dos métodos convencionais de interpretação de DGA (*Dissolved Gas Analysis*), ou de algoritmos de aprendizado de máquina que atribuem somente os rótulos das classes de falhas (única classe para cada amostra) que ocorrem com os transformadores, a metodologia aqui proposta, tem como originalidade o uso das informações contidas no espaço de atributos, estruturado na etapa de treinamento dos algoritmos baseados em grafos, para construir o resultado do diagnóstico de falha com base na avaliação do histórico ou de amostras sucessivas de gases produzidos.

No caso, o método proposto explora o espaço de atributos de falhas e a heurística de classificação dos algoritmos OPF e KNN para conversão das amostras dos gases em gráficos de falhas e assim, permite fazer a análise comportamental do transformador durante sua vida operacional. Com esse método torna-se possível conhecer, além do estado atual do transformador, as tendências do seu comportamento com as evoluções dos gases e os possíveis cenários de falhas, não restrito somente a um tipo de falha específico, mas de todas as classes de falhas e com o sequenciamento das chances de ocorrências determinadas pelos seus custos (relação de conectividade entre as amostras avaliadas e as falhas reais) apresentados.

Tal abordagem permite auxiliar o agente especialista em manutenção de transformadores, fornecendo-lhe uma resposta complementar de falhas mais informativa para o planejamento de novas ações, como por exemplo, realizar outros procedimentos de teste específicos complementares, programar a realização de inspeção e manutenção, a fim de reduzir as chances de saída inesperada do transformador do sistema elétrico por falha.

## 1.6. Publicações relacionadas à tese

Segue as produções científicas de resultados parciais relacionadas ao tema e/ou ferramentas de inteligência computacional.

- Marco Akio Ikeshoji, André Nunes de Souza, Pedro da Costa Junior, Danilo Sinkiti Gastaldello, João Paulo Papa, Carlos Guilherme Gonzales. Análise de Incertezas em Amostras de Gases Usadas no Diagnóstico de Falhas em Transformadores de Potência. VIII Simpósio Brasileiro de Sistemas Elétricos, 2020, Santo André-SP.
- Marco Akio Ikeshoji, Luis Claudio Sugi Afonso, Danilo Sinkiti Gastaldello, André Nunes de Souza, Pedro da Costa Junior, Carlos Guilherme Gonzales, João Paulo Papa. Qualidade no Fornecimento de Energia: Contribuições de uma Ferramenta Inteligente



para Gestão de Falhas Incipientes em Transformadores. XIII Conferência Brasileira sobre Qualidade de Energia Elétrica, 2019, São Paulo-SP.

- Marco Akio Ikeshoji, André Nunes de Souza, Pedro da Costa Junior, Danilo Sinkiti Gastaldello, Murilo Vargas da Silva, Juliana A. G. Maginador, Desenvolvimento de uma Ferramenta de Diagnóstico de Falhas em Transformadores Utilizando a Abordagem de Aprendizagem de Máquina Extrema. Induscon, 2018, São Paulo-SP.
- Marco Akio Ikeshoji, André Nunes de Souza e Danilo Sinkiti Gastaldello. Análise de Padrões dos Gases Dissolvidos no Óleo e sua Influência na Classificação de Falhas em Transformadores. IX Seminário da Pós-Graduação em Engenharia Elétrica, 2019, Bauru-SP (em correção).

### **1.7. Organização do documento**

No Capítulo 2 são apresentadas as revisões bibliográficas sobre a importância do transformador, os custos relacionados a saída de operação do sistema, e os tipos e estatísticas de falhas nos elementos que compõem o transformador, caracterizando também o sistema isolante (óleo e papel), os mecanismos de degradação e formação dos gases internos, tipos de ensaios e meios de aquisições de dados, bem como as motivações do uso das análises dos gases dissolvidos no óleo, para o diagnóstico de falhas incipientes em transformadores.

No Capítulo 3 são apresentadas as revisões bibliográficas dos principais métodos e técnicas de análise e interpretação de gases dissolvidos, envolvendo os conhecidos métodos analíticos normatizados e de aprendizado de máquina; alguns modelos de análises crítica e de risco envolvendo transformadores; formas de pré-processamento de dados; de otimização, construção e seleção de atributos; e geração de dados sintéticos, com base nas incertezas de medições da IEC.

No Capítulo 4 são apresentados a metodologia de tratamento e seleção dos melhores atributos dos dados, bem como a geração de dados sintéticos, todos empregados para melhorar a classificação nos algoritmos OPF e KNN, e então é proposta a metodologia de classificação e análise gráfica da evolução de falha incipiente em transformadores.

No Capítulo 5 são preparados os dados de treino dos algoritmos OPF e KNN e realizados seus treinamentos, e então simuladas as amostras do histórico de gases de alguns dos transformadores de uma concessionária de energia, cuja identidade será preservada, para apresentação, análise e validação do método.

No Capítulo 6 é feita a conclusão e uma proposta para trabalhos futuros.

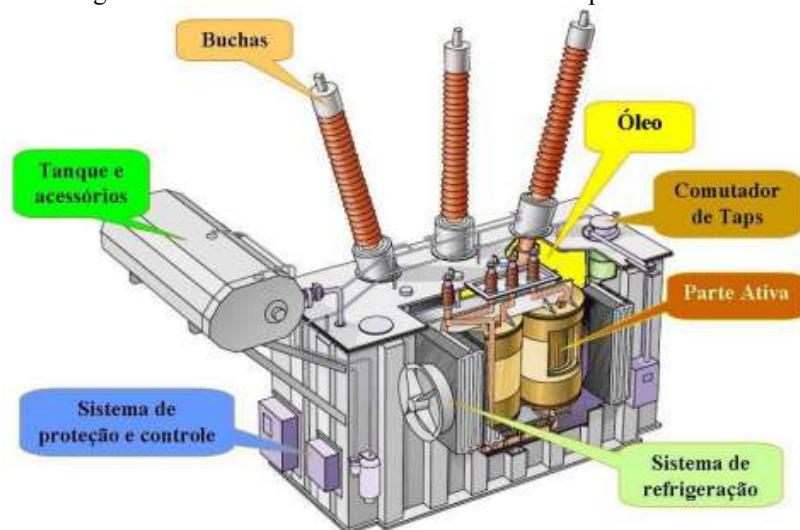
## CAPÍTULO 2 – ESTUDO SOBRE FALHAS EM TRANSFORMADORES

### 2.1. Caracterização dos transformadores de potência

O transformador de potência é uma máquina elétrica que converte diferentes níveis de tensões. Pode ser do tipo monofásico ou trifásico, elevador ou abaixador de tensão, isolado ou não, a seco ou a óleo, e de nível de potência conforme os requisitos do projeto elétrico no qual se insere. Em se tratando de transformadores a óleo, historicamente, iniciou-se o emprego do óleo mineral em conjunto com o papel na manufatura dos transformadores em 1890, abrindo assim o caminho para elevar seus níveis de tensões e potências (VISHAL et al., 2011). A combinação desses materiais resultou na melhora das características de isolamento elétrica e dissipação térmica dos transformadores (IPTC, 1994), determinando, até nos dias atuais, uma ampla utilização (WANG et al., 2018). Nesse contexto, a expectativa da vida útil dos transformadores tornou-se também dependente do estado da isolamento de ambos materiais e, portanto, mantê-los em boas condições significa retardar o envelhecimento precoce do equipamento e o surgimento de falhas internas.

Basicamente, sua estrutura construtiva principal (Figura 1) pode ser caracterizada em duas partes: uma ativa constituída do núcleo de material ferromagnético e enrolamentos de cobre, e a outra composta pelos elementos: tanque principal que acomoda o conjunto ativo, trocadores de calor ou radiadores, buchas isoladoras, tanque de expansão do óleo e comutador de tensão associado (ambos dependentes do modelo do equipamento).

Figura 1 – Estrutura básica do transformador de potência a óleo



Fonte: Azevedo et al. (2019)

Estimar a vida útil e determinar a ocorrência de falhas incipientes nos transformadores exigem análises de diversos fatores envolvidos que compreendem: o projeto (características construtivas) e materiais empregados, presença de impurezas e umidade nos materiais isolantes, o tempo de operação, os esforços mecânicos (curto-circuito) e térmicos (sobrecarga, harmônicos) em que são submetidos, entre outros (BECHARA, 2010; SOKOLOV, 2005). Porém, como nem sempre todos esses dados encontram-se disponíveis para análise, isso contribui para que diferentes estratégias sejam desenvolvidas, a fim de auxiliarem nos diagnósticos e tomadas de decisões de manutenção, como é o caso do emprego da inteligência computacional.

## **2.2. Estimativa de custos que envolvem a falha do transformador**

A saída repentina de operação do transformador ocasiona além de transtornos de ordem social, elevados prejuízos financeiros às concessionárias e aos setores produtivos locais. Tais eventos estão relacionados aos aspectos (ANEEL, 2018ab; SUN; HUANG; HUANG, 2012; ALMEIDA; DELGADO; MOREIRA, 2003; FILHO, 2012): danos ou perda material, energia não fornecida, imagem da empresa, perda de vidas, danos ambientais, litígio jurídico, entre outros. Em um estudo de caso realizado entre o período de 1997 a 2001, sobre prejuízos ocasionados por falhas envolvendo 94 transformadores de 25 MVA, Bartley (2013) relatou prejuízos da ordem de US\$ 123 milhões com a interrupção dos negócios e de US\$ 163 milhões com perda de material das concessionárias (75%) e particulares (25%).

Quanto ao custo somente de transformadores de potência, ele é determinado por uma série de fatores que dependem (USDE, 2012): das características técnicas, do fabricante ou marca, das condições do mercado e do local de instalação da fábrica ao ponto de entrega. Estima-se que o custo por MVA situe-se entre US\$11 e US\$39 mil para transformadores elevadores da geração, e entre US\$8,5 e US\$35 mil para transmissão, ambos com o prazo de fabricação e entrega variando entre 10 e 20 meses. Segundo Scattigio, Pompili e Calacara (2018), grandes transformadores podem representar cerca de 60% do investimento total de uma subestação. Outras estimativas de custo são apresentadas em Filho (2012) e Bonfin (2015), entretanto, o que se pode concluir é que o tempo e custo de produção são altos.

Em uma enorme população de transformadores, manter a alta confiabilidade do sistema elétrico envolve um custo substancial com a manutenção desses equipamentos, e em face à pressão de manter baixo o preço da energia, tem se tornado um grande desafio para as concessionárias aumentarem (ou até preservarem) seus lucros. Desse modo, a abordagem

tradicional de manutenção baseada no tempo de serviço tem sido substituída por abordagens de manutenção preditiva e preventiva (BOSS et al., 2004), com uso de métodos e técnicas do tipo *online* ou *offline* de diagnósticos e monitoramento (ASSUNÇÃO, 2007; LUOPAJÄRVI, 2010; TENBOHLEN et al., 2016), ou composta por ambos (ILGEVICIUS; BABIZKI, 2015).

### **2.3. Tipo, localização e estatística de falha nos transformadores**

As falhas que ocorrem com os transformadores de potência podem ter diversas origens, isto é, desde o término da sua vida útil devido às próprias ações naturais decorrentes dos fenômenos térmicos, hidrolíticos e oxidativos inerentes ao seu funcionamento normal, até àquelas resultantes de interferências externas operacionais e ambientais conforme segue (BECHARA, 2010; SOKOLOV, 2005):

- Falha do dielétrico que ocorre quando a resistência dielétrica do meio isolante é rompida por tensões transitórias que excedem os limites especificados em projeto e produz sobreaquecimento, ruptura do dielétrico, geração de arco elétrico, curto-circuito com deformação mecânica do enrolamento, entre outros;
- Degradação ou envelhecimento do material isolante sólido e líquido que ocorre de modo natural sob a influência da temperatura, oxigênio, umidade e acidez, porém, quando excedido certos limites, atuam como catalizadores acelerando o processo de envelhecimento levando à falha precoce do equipamento. A presença do enxofre corrosivo no óleo, leva à falência da celulose e grandes danos materiais e financeiros;
- Deformação mecânica dos enrolamentos (radial ou axial ao raio da espira) por esforços mecânicos provocados por curto-circuito externos, que podem ocasionar o rompimento dos materiais isolantes sólidos internos;
- Falhas associadas ao circuito magnético (indução magnética) que geram aquecimento em determinados pontos internos (por circulação de corrente), com ou sem geração de faísca, e ocorre devido à ausência ou posicionamento inadequado da blindagem magnética, parafusos de prensagem do núcleo frouxos, deslocamento da estrutura interna e perda do ponto de aterramento;
- Falhas nas buchas que ocorrem devido à degradação ou contaminação do sistema isolante interno, ao aumento da temperatura interna por sobrecarga, à falha no seu processo de fabricação e no uso de materiais defeituosos. Geralmente, resultam em graves danos ao transformador, quando ocorre a explosão do corpo isolante;

- Falha em cabos e conexões que são associadas ao superaquecimento localizado. Em conexões, condiciona-se tal efeito às deficiências de contato entre os terminais de ligação entre elementos internos e externos;
- Falhas em comutadores geralmente são decorrentes de problemas mecânicos, manutenção inadequada, degradação do óleo isolante e limites de operação com tensão de impulso de transferência acima do especificado (em comutadores sob carga);
- Erros de projeto, falha no processo de produção e fabricação, uso de materiais defeituosos e até instalação incorreta do equipamento. Não é sempre que as falhas geradas por esses fatores são detectadas nos ensaios de fábrica ou no campo durante o processo de comissionamento, e sim somente após a entrada em operação no sistema elétrico;
- Falhas em acessórios e componentes que podem ser caracterizadas como simples, envolve elementos com funções restritas ao funcionamento do transformador, ou seja, requer pequenas e simples intervenções, tais como, substituição de relés, sensores e reparos de vazamentos, ou complexas, a partir do momento que envolve a parada do equipamento e maior tempo para restabelecimento da operação como falhas das buchas ou comutador de tensão; e
- Falta ou manutenção inadequada do transformador, que acelera seu envelhecimento e reduz sua vida útil.

O levantamento estatístico dos índices ou taxas de falhas que ocorrem com os transformadores do sistema elétrico contribui com informações que permitem aos fabricantes melhorarem seus equipamentos, processos e especificações, e as concessionárias melhorarem seus procedimentos de manutenção, determinar políticas de reservas de equipamentos ou acessórios, realizarem planejamentos, entre outros (BASTOS et al., 2013). Muitos especialistas relacionam a “curva da banheira” (*bathhtub curve*) como fator de justificativa da ocorrência de falha, ou seja, há um risco maior da falha no transformador ocorrer nos anos iniciais, após entrada em funcionamento e próximo do fim da sua vida útil, contudo, estatisticamente na prática, isso nem sempre tem se demonstrado verdadeiro (SOKOLOV, 2005), portanto, para Jarman et al. (2009), é inseguro basear a previsão da falha apenas em dados estatísticos de históricos, uma vez que se trata de um fenômeno probabilístico que depende de três variáveis de controle principais: projeto, envelhecimento e fatores ou eventos externos.

Fabricantes estimam um tempo de vida útil dos transformadores entre 25 e 40 anos, podendo atingir mais de 50 anos, desde que realizados acompanhamento e manutenção adequados (CARDOSO, 2005; GORGAN et al., 2012). Com avanço dos anos de operação e proximidade do fim do tempo de vida útil do transformador, diversos métodos (GORGAN et al., 2012; HILLARY et al., 2017; HONG; MEEKER; MCCALLEY, 2009) que visam estimar seu tempo restante de vida, têm sido pesquisados e desenvolvidos, de maneira que evitem a saída repentina ou prematura do sistema.

Uma investigação global envolvendo 21 países, realizado entre 1996 e 2010 pelo CIGRÉ sobre estatística de falhas em transformadores a óleo, resultou em uma taxa de falha geral dentro de 1% (TENBOHLEN et al., 2016), conforme mostra a Tabela 1. Tal investigação contou com a colaboração de 56 empresas de energia de 21 países, dos quais foram observados somente as falhas maiores, caracterizadas pela retirada do serviço por mais de sete dias para efeito de investigação, trabalhos de reparação ou substituição. Considerando-se essas falhas maiores em transformadores de subestações e de geração acima de 100 kV (536 e 119 unidades respectivamente), os maiores índices percentuais de falhas se localizaram nos enrolamentos com 39,8%, seguidos de falhas nos comutadores com 27,5%, buchas 17,3%, e cabos ou conexões (internas), na saída dos enrolamentos principais com 8,2%.

Tabela 1 – Taxas de falhas em população de transformadores de 21 países

Classe de tensão	Equipamentos (unidade)	Maiores falhas (unidade)	Taxa geral de falhas (%)
$69\text{kV} \leq U < 100\text{kV}$	2962	144	0,94
$100\text{kV} \leq U < 200\text{kV}$	10932	280	0,43
$200\text{kV} \leq U < 300\text{kV}$	4272	186	0,50
$300\text{kV} \leq U < 500\text{kV}$	3233	142	0,60
$500\text{kV} \leq U < 700\text{kV}$	434	27	0,57
$U \geq 700\text{ kV}$	348	10	0,33
Total	22181	799	0,53

Fonte: Adaptado pelo autor de Tenbohlen et al. (2016)

Para Metwally (2011), as taxas de falhas encontradas para os transformadores elevadores da geração e subestação, para 100% da concentração do mercado global em três faixas de tensões, foram de 1,35% e 0,05% na classe de 60-100kV, de 1,55% e 0,80% na classe de 100-300kV e de 1,80% e 4,15% na classe de 300-700kV, determinando uma média geral de falhas de 1,6%. Nas localizações das principais falhas, cerca de 30% eram nos enrolamentos, 40% no comutador e 14% nas buchas. Já em Murugan e Ramasamy (2015) analisou-se as taxas de falhas nos transformadores em aproximadamente 842 subestações da

Índia, com classes de tensões compreendidas entre 33-400 kV, por aproximadamente quatro anos, sendo observadas cerca de 196 falhas distribuídas entre 13% nas buchas, 10% no comutador, 14% no enrolamento, 41% na isolação (sólida e líquida) e o restante dividido entre núcleo, tanque, radiador, entre outros.

No Brasil, o grupo de trabalho GTA2-02 do CIGRÉ-Brasil (BASTOS et al., 2013) realizou o levantamento das taxas de falhas e defeitos dos transformadores (elevadores da geração e de subestações) e reatores em 12 grandes concessionárias de energia. Foram estabelecidas premissas para o fornecimento de dados tais como: equipamentos com classe de tensão acima de 138kV, todos a óleo, sem restrição de potência, com e sem comutadores, fabricados após 1980 e falhas apresentadas após 1997. Entretanto, devido às dificuldades com a falta de informações nos dados registrados entre 1997 e 2003, foram estimadas apenas as taxas de falhas e defeitos entre 2004 e 2009, resultando em uma taxa média de 1,25%, considerando todos os equipamentos. Para o caso de apenas eventos de falhas em transformadores, a taxa média foi de 0,8%. Quanto à localização das principais falhas, aproximadamente 19,5% foram relacionadas a ocorrências nos enrolamentos, 14,3% no comutador e 14% nas buchas. A idade dos transformadores no período da análise era de 75% acima de 10 anos, 14% entre 5 e 10 anos e o restante, com menos de 5 anos. Em 2012 foi feito um levantamento complementar de taxas de falhas e defeitos envolvendo os anos de 2010 e 2011 e a ocorrência mostrou-se maior em equipamentos com mais de 10 anos.

Analisando-se todos esses casos, pode-se concluir que as localizações da grande maioria das falhas concentram-se nos enrolamentos, comutadores, cabos e conexões internas, permitindo assim direcionar o foco das pesquisas para reduzir os índices de falhas sobre esses elementos que têm em comum o contato com o óleo isolante. As análises e manutenção das buchas recebem tratamento separado dos transformadores e, portanto, não são abordadas nessa tese.

#### **2.4. Caracterização do sistema isolante**

A vida útil do transformador está diretamente ligada ao estado do sistema de isolação formado pelo óleo e isolante sólido (celulose). Enquanto o primeiro permite ser substituído ou reprocessado, o segundo, em caso de degradação extrema, decreta o fim do uso do equipamento.

De acordo com Neto, Assunção e Assunção (2009), o material isolante sólido mais comum empregado na fabricação do transformador é de base celulósica (papel *kraft* e

*pressboard*) termicamente estabilizada, visto que este tipo de material apresenta além de boas propriedades dielétricas e mecânicas, disponibilidade comercial e baixo custo. Há também outros tipos de materiais como a fibra de aramida (NOMEX<sup>®</sup>) que possui excelentes características de estabilidade térmica, contudo, devido ao seu alto custo, é destinado apenas a aplicações especiais, e combinações híbridas formadas por isolantes sólidos de alta temperatura e celulose, que são distribuídos estrategicamente entre as partes internas do transformador, atendendo aos diferentes gradientes de temperaturas de trabalho. Quanto ao óleo isolante, esse pode ser de origem mineral (ROUSE, 1998), natural (RYCROFT, 2014), sintético (BORSI; GOCKENBACH, 2005; GOCKENBACH; BORSI, 2008) ou uma mistura desses óleos (PERRIER; BEROUAL; BESSEDE, 2006), cada qual obtido por meio de um processo específico, cujas propriedades se diferenciam quanto a sua natureza física, química e dielétrica.

Nesta pesquisa, o tipo de óleo analisado é mineral, por ser o isolante líquido mais empregado na fabricação de transformadores e outros tipos de equipamentos elétricos (LIAO et al., 2011; ROUSE, 1998) e por ter o menor custo (COOPER POWER SYSTEMS, 2000). Complementarmente, novas tecnologias envolvendo o óleo mineral e sua associação com a nanotecnologia têm sido desenvolvidas, de modo a conferir melhora nas suas propriedades física, química e dielétrica, conforme pode ser visto no primeiro transformador de 1100kV produzido pela Siemens e colocado em operação em Tianjin, na China (WANG et al., 2018).

#### 2.4.1. O papel isolante

A degradação do papel isolante ocorre com a elevação térmica da sua temperatura de trabalho. Quanto maior essa temperatura, mais acelerada é a sua decomposição. Essa degradação também é significativamente influenciada pelo tipo e composição do material isolante, presença do oxigênio, umidade (BANDARA; EKANAYAKE; SAHA, 2015; LUNDGAARD et al., 2004; MTETWA, 2008) e nível de acidez do óleo (ASSUNÇÃO, 2007).

Lelekakis, Martins e Wijaya (2012) analisaram os efeitos de diferentes composições de umidade, oxigênio e temperatura na degradação do papel *Kraft*, com e sem tratamento térmico, e concluíram que em concentrações de oxigênio e umidade no óleo abaixo de 4%, a degradação do *Kraft* tratado era menor do que o não tratado, e a partir de 4%, praticamente não havia diferença.



O nível de degradação do papel isolante pode ser determinado medindo-se o grau de polimerização (DPy - *Degree of Polymerization*) com o método viscosimétrico (ABNT, 2009; ASTM, 2016). Embora seja o método mais empregado pelas indústrias de transformadores (KRAUSE et al., 2014), há necessidade da retirada do equipamento de serviço e seu respectivo desmonte, para obter as amostras (papel) das principais partes (ativa e não ativa) para avaliação. Contudo, visando contornar esse inconveniente, métodos indiretos como a medida de umidade presente no papel e óleo utilizando-se curvas de equilíbrio (ASSUNÇÃO, 2007; DU et al., 1999) e análise do gás 2-furfuraldeído dissolvido no óleo (BABURAO et al., 2007; MARTINS, 2007) foram desenvolvidas com objetivo de estimar o DPy do papel isolante e, conseqüentemente, a vida útil remanescente do transformador. Valores de DPy para papel novo após o processo de secagem variam entre 1000 e 1300, e abaixo de 200, são considerados como fim da vida útil do papel (PRZYBYLEK, 2018; CHEIM et al., 2012).

#### 2.4.2. O óleo mineral isolante

O óleo mineral isolante usado em diversos tipos de equipamentos de alta tensão tem como função refrigerar, melhorar as características de isolação elétrica e extinguir arco elétrico. De acordo com Santos (2013), o óleo pode ser de: base parafínica cujas propriedades químicas são estáveis e resistentes as alterações, e que se tornam menos susceptíveis ao envelhecimento (oxidação), porém possui uso restrito em baixas temperaturas, visto que tende a se sedimentar; ou de base naftênica cujas propriedades químicas lhe conferem excelentes características dielétricas e de fluidez (refrigeração) mesmo em baixas temperaturas, no entanto podem acelerar a degradação de materiais sintéticos e elastômetros. O uso do óleo mineral parafínico tem crescido motivado pela baixa reserva mundial do petróleo do tipo naftênico e o desenvolvimento de tecnologias que permitem contornar seu problema de trabalho em baixas temperaturas (BONFIN, 2015).

A Resolução da Agência Nacional do Petróleo ANP N° 25, de 9 de setembro de 2005, classificava o óleo isolante em dois tipos: de base naftênica sem restrição de classe de tensão no emprego em transformadores; e de base parafínica aplicáveis até à classe de tensão de 145 kV. Posteriormente, em sua Resolução N°36, de 5 de dezembro de 2008, retirou-se o limite de tensão do óleo de base parafínica.

Para que os óleos isolantes desempenhem sua função de modo confiável e seguro, é necessário que atendam determinados requisitos físicos, químicos e elétricos recomendado em normas relacionadas. De acordo com a norma brasileira NBR 10576 (ABNT, 2017), se um

determinado grau de degradação for excedido, de algum modo haverá comprometimento da margem de segurança operacional do equipamento e assim, deve-se considerar o risco de ocorrer uma falha. Dentre os principais ensaios recomendados para análise do óleo mineral isolante, podem ser destacados: aspecto visual, cor, ponto de fulgor, ponto de fluidez, densidade, viscosidade, ponto de anilina, tensão interfacial, índice de neutralização ou acidez, fator de potência ou fator de perdas dielétricas, rigidez dielétrica e teor de água no óleo. Os valores de referência para início do controle do óleo isolante, bem como do óleo em uso, podem ser encontrados nessa norma e na Resolução ANP N°36, de 5/12/2008.

## **2.5. Degradação do sistema isolante e formação dos gases**

A formação de determinados tipos de gases dissolvidos no óleo isolante do transformador deve-se à decomposição da celulose e do próprio óleo em decorrência do envelhecimento natural e presença de distúrbios de natureza elétrica e térmica que ocorrem no equipamento durante toda a sua vida operacional.

De acordo com a norma IEEE C57.104, ao efeito térmico, associam-se as perdas geradas nos condutores e elementos internos (conexões, mau contato entre elementos móveis, entre outros), ao elétrico, às descargas parciais, descargas de baixa intensidade e arco elétrico. Nesse contexto, as vantagens oferecidas pela análise de gases dissolvidos no óleo isolante contribuem para (IEEE, 2019):

- Avaliar e acompanhar o estado da saúde do transformador novo ou reparado (gerenciamento de risco);
- Prever antecipadamente o desenvolvimento de falhas;
- Identificar o tipo de falha;
- Programar o período de inspeção e manutenção dos equipamentos; e
- Melhorar a segurança e confiabilidade do sistema elétrico.

A distinção no diagnóstico de falha entre a natureza térmica e elétrica nos transformadores tem relação com o grau de urgência no atendimento e correção da falha. Para Duval (2012), os diagnósticos quanto aos níveis de criticidade devem ser caracterizados pelo tipo de falha envolvida, localização da falha (celulose ou no óleo) e quantidade de gases produzidos, conforme segue:

- Falhas mais perigosas: arco de alta energia no óleo e celulose, arco de baixa energia no papel (ou celulose), pontos quentes no papel para temperatura  $T > 300^{\circ}\text{C}$ ;

- Falhas menos perigosas: arco de baixa energia no óleo, pontos quentes no óleo para temperatura  $T > 300^{\circ}\text{C}$  e pontos quentes no papel em baixa temperatura; e
- Falhas não perigosas: pontos quentes no óleo para  $T < 300^{\circ}\text{C}$ , formação de gases (*stray gassing*) no óleo em baixa temperatura, descarga parcial corona (a menos que forme nível elevado de hidrogênio), reações catalíticas da celulose com água e envelhecimento do papel.

A quantificação das concentrações e tipificação dos gases dissolvidos no óleo são obtidas por meio do método de análise cromatográfica gasosa (ASTM, 2017; IEC, 2005). Sua aceitabilidade está embasada na alta precisão, repetibilidade e possibilidade de leitura de todos os gases (BAKAR; ABU-SIADA; ISLAM, 2014) tornando-se assim, o método mais informativo (LONDO et al., 2015) sobre os gases presentes. Entretanto, em virtude do seu custo elevado, outros métodos de medições de gases experimentais têm sido propostos como alternativas para contornar essa questão, como por exemplo, a espectroscopia fotoacústica (SKELLY, 2012), espectroscopia infravermelha (CHEN et al., 2008) e Raman (WANG et al., 2016), no entanto, algumas restrições quanto aos limites mínimos mensuráveis dos gases, sensibilidade às vibrações e variações de temperatura e pressão estão presentes nesses testes e merecem atenção na aplicação.

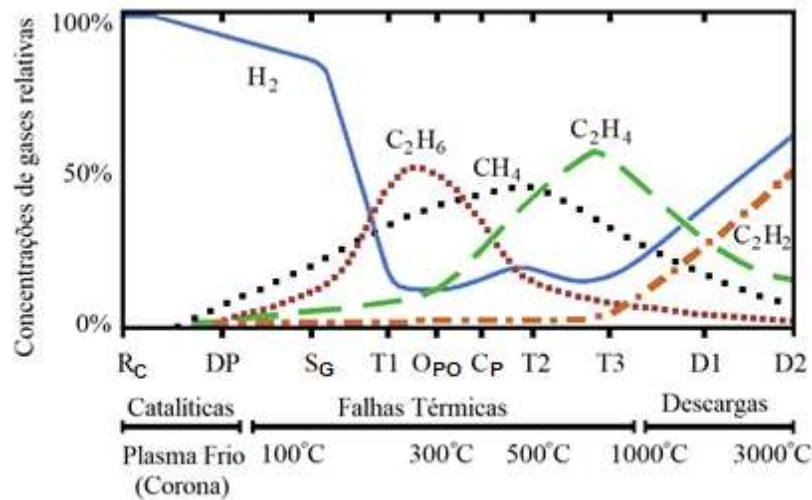
Quanto as relações dos eventos de falhas associados à degradação da celulose e óleo, podem ser assim resumidas (IEEE, 2019; IEC, 2015):

- A falha térmica no óleo ocorre quando a cadeia de hidrocarbonetos é quebrada devido ao estresse térmico e elétrico, tendo como consequência a produção dos gases combustíveis  $\text{H}_2$  (hidrogênio),  $\text{CH}_4$  (metano),  $\text{C}_2\text{H}_2$  (acetileno),  $\text{C}_2\text{H}_4$  (etileno) e  $\text{C}_2\text{H}_6$  (etano). A quantidade e composição dos tipos de gases dissolvidos no óleo dependem das possíveis combinações entre hidrocarbonetos, distribuição de energia e temperatura nas proximidades da falha, e do tempo de duração do evento. Acima de  $500^{\circ}\text{C}$  a concentração de  $\text{C}_2\text{H}_4$  se sobrepõe ao  $\text{CH}_4$  e  $\text{C}_2\text{H}_6$ . Já o gás  $\text{C}_2\text{H}_2$  requer temperatura  $> 800^{\circ}\text{C}$  para formação em grande quantidade, contudo traços podem aparecer em temperatura  $< 800^{\circ}\text{C}$ . A formação do *stray gassing* ocorre em temperaturas  $< 200^{\circ}\text{C}$  com a formação de  $\text{H}_2$ ,  $\text{CH}_4$  e  $\text{C}_2\text{H}_6$ . Já a carbonização do óleo ocorre em temperaturas entre  $500$  e  $800^{\circ}\text{C}$ .
- Na falha térmica na celulose, os principais gases produzidos a partir da decomposição térmica da celulose são o  $\text{CO}$  (monóxido de carbono) e  $\text{CO}_2$  (dióxido de carbono). A quebra da cadeia polimérica inicia-se em temperatura mais baixa que a do óleo e passa

a ser significativa a partir de temperatura  $>105^{\circ}\text{C}$  e com carbonização total acima de  $300^{\circ}\text{C}$ . O uso de ambos os gases como indicadores de falha no papel deve ser acompanhado da interpretação de outros fatores, como por exemplo, valores de  $\text{CO}>1000$  PPM (Parte Por Milhão) e baixa razão  $\text{CO}_2/\text{CO}<3$  acompanhada de formação significativa dos hidrocarbonetos pode indicar presença de falha.

A Figura 2 mostra as concentrações relativas de gases aproximadas em relação à temperatura e tipos de falhas relacionados.

Figura 2 – Concentrações relativas de gases vs. temperatura e tipos de falhas



Fonte: Adaptado pelo autor de IEEE (2019)

Onde:

- R<sub>C</sub> – Reações catalíticas entre aço galvanizado e água;
- DP – Descargas parciais;
- S<sub>G</sub> – *Stray gassing* (gases parasitas) para temperatura  $T < 200^{\circ}\text{C}$ ;
- O<sub>PO</sub> – Superaquecimento do papel ou óleo;
- C<sub>P</sub> – Possível carbonização do papel;
- T1 – Falha térmica em temperatura  $T < 300^{\circ}\text{C}$ ;
- T2 – Falha térmica em temperatura entre  $300^{\circ}\text{C} < T < 700^{\circ}\text{C}$ ;
- T3 – Falha térmica em temperatura  $T > 700^{\circ}\text{C}$ ;
- D1 – Descarga de baixa energia; e
- D2 – Descarga de alta energia.

De acordo as normas 60599 (IEC, 2015) e C57.104 (IEEE, 2019), as falhas podem ser assim caracterizadas:

- Descarga Parcial – descarga elétrica localizada que ocorre parcialmente em isolantes (sólido ou líquido) entre condutores, e pode ou não ocorrer nas adjacências do mesmo, com duração inferior a 1  $\mu$ s. Descarga corona é um tipo de descarga parcial (ionização) que ocorre em meio gasoso (interno, as bolhas que podem surgir no óleo) entre condutores separados por meio de isolamento líquida ou sólida. A descarga parcial pode formar partículas sólidas de cera (fragmentos polimerizados) no papel dificultando a dissipação do calor;
- Descarga elétrica “disruptiva” – pode ser caracterizada como a passagem de um arco elétrico que pode provocar o colapso do isolamento. Dependendo da intensidade de energia envolvida pode ser classificado como descarga elétrica de baixa intensidade ou descarga elétrica de alta intensidade;
- Descarga elétrica de baixa energia – descarga elétrica que envolve baixa intensidade de energia e pode ocorrer no óleo e celulose decorrente de: arco ou centelhamento em conexões ruins entre diferentes potenciais ou flutuantes, entre condutores dos enrolamentos e anéis de proteção; descargas entre elementos de fixação (frouxos), entre bucha e tanque; caminhos/trilhas em superfície de blocos isolantes de madeira, através do óleo, entre outros. Como consequências podem ser encontradas perfurações ou carbonização da superfície do papel, e partículas de carbono no óleo;
- Descarga elétrica de alta energia – descarga elétrica que envolve alta intensidade de energia e pode ocorrer no óleo e celulose provocados por arco elétrico proveniente de curto-circuito entre condutores dos enrolamentos, terminais, enrolamento e núcleo, bucha e tanque, entre outros. Como resultados desse evento, pode haver grande destruição e carbonização do papel, intensa carbonização do óleo, fusão de metais, e atuação da proteção;
- Falha térmica  $T < 300^{\circ}\text{C}$  – aquecimento do óleo e/ou papel abaixo de  $300^{\circ}\text{C}$  e ocorre mudança da coloração do papel para marrom. Normalmente, são provocados por sobrecargas ocasionadas por aumento do consumo em momentos de emergências (por exemplo, saída de operação de fontes geradoras interligadas), por bloqueio no fluxo do óleo entre os enrolamentos, e fluxo de dispersão sobre o suporte de prensagem do núcleo. Caracterizado como subtipo dessa falha, há o superaquecimento do papel ou óleo para  $T < 250^{\circ}\text{C}$  sem que haja perda das propriedades elétricas isolantes do papel, e  $T < 200^{\circ}\text{C}$  (*stray gassing*) que são gases que podem ser produzidos em razão da

instabilidade química dos óleos minerais e incompatibilidade de materiais (por exemplo, uso de passivadores no óleo);

- Falha térmica  $300^{\circ}\text{C} < T < 700^{\circ}\text{C}$  – falhas térmicas no óleo e/ou papel acima de  $300^{\circ}\text{C}$  que provoca a carbonização do papel. Podem originar-se em conexões defeituosas, conexões dos cabos, circulação de corrente em fiações de aterramento, parafusos da estrutura de prensagem do núcleo, grampos de proteções magnéticas, e outros elementos internos ao transformador; e
- Falha térmica  $T > 700^{\circ}\text{C}$  – que pode ser evidenciada quando há grande quantidade de carbonização do óleo presente, mudança na cor do metal ( $> 800^{\circ}\text{C}$ ) envolvido ou fusão do metal ( $> 1000^{\circ}\text{C}$ ). Podem ser ocasionadas por circulação de elevados níveis de corrente no tanque e no núcleo, ou correntes de menor intensidade nas paredes do tanque gerado por fluxo magnético desbalanceado, e curto-circuito entre lâminas do núcleo.

A Tabela 2 associa os eventos de falhas à produção dos gases mais relevantes encontrados.

Tabela 2 – Tipos de falhas vs. relevância dos gases produzidos

Causas da geração do gás		H <sub>2</sub>	CH <sub>4</sub>	C <sub>2</sub> H <sub>6</sub>	C <sub>2</sub> H <sub>4</sub>	C <sub>2</sub> H <sub>2</sub>	CO	CO <sub>2</sub>
Falha Elétrica	Óleo DP corona	X	O					
	Óleo Arco + Faísca (DP)	X			O	X		
	Celulose							O
Falha Térmica	Óleo Baixa Temperatura	O	X	X				
	Óleo Média/Alta Temperatura	O	O		X			
	Óleo 700°C ou mais	O			X	O		
	Celulose						X	O
Envelhecimento	Óleo	O		X				
	Celulose						O	X

DP - Descarga Parcial, X – Muito Relevante, O – Média Relevância

Fonte: Adaptado pelo autor de LEE et al. (2013)

A presença de O<sub>2</sub> (oxigênio) e N<sub>2</sub> (nitrogênio) dissolvido no óleo, embora possa ser utilizado como relação entre gases não combustíveis para avaliar se o sistema de preservação do óleo está em contato com o ar atmosférico (FILHO, 2012; MILASCH, 1984), ambos não são considerados subprodutos de degradação da isolamento (JAKOB; DUKARM, 2015). Em um estudo realizado pelo grupo de trabalho da IEEE (JAKOB; DUKARM, 2015) envolvendo a razão inversa entre os gases N<sub>2</sub>/O<sub>2</sub>, concluiu que quanto maior era a quantidade de oxigênio

dissolvido no óleo, menor era a concentração dos gases combustíveis presentes e, recentemente, em 2019 esta consideração passou a ser empregada na norma IEEE C57.104.

Outros tipos de gases como os compostos furânicos também podem ser obtidos com a análise do óleo, contudo sua existência é associada à degradação da celulose (BANDARA; EKANAYAKE; SAHA, 2015; MTETWA, 2008; WATTAKAPAIBOON; PATTANADECH, 2016) e estimativa da vida útil remanescente do transformador (BABURAO et al., 2007; GORGAN et al., 2012).

## 2.6. Análises e aquisições de dados *online* e *offline*

O aumento na quantidade de transformadores determina a necessidade do aumento da preocupação com a confiabilidade operacional do sistema elétrico, logo, diferentes recursos de caráter *online* ou *offline* podem ser empregados para extrair as informações necessárias para acompanhar a saúde desses equipamentos.

A abordagem *online* é caracterizada pela não retirada do transformador de operação, ou seja, os parâmetros de interesse para avaliação do estado do equipamento são obtidos em serviço. A seguir são relacionados alguns exemplos de análises *online* empregados na avaliação e diagnósticos dos transformadores (BECHARA, 2010).

- Análise físico-química do óleo isolante que visa qualificar o seu estado (adequado ou não) e a presença de impurezas (enxofre corrosivo e sedimentos);
- Análise de gases dissolvidos no óleo isolante para determinação de possíveis falhas incipientes e do teor de furfuraldeído para avaliar o estado da isolação sólida e estimar a vida útil remanescente do transformador;
- Análise termográfica da superfície externa do transformador com foco no sistema de refrigeração e conexões de terminais externos; e
- Emprego de monitoramento com coleta de dados via sensores instalados nas principais partes e elementos auxiliares (como por exemplo, bobina, comutadores sob carga, motoventiladores, entre outros) do transformador para análise e diagnóstico em tempo real.

Em relação à forma de análise de gás dissolvido no óleo isolante dos transformadores de potência, há dois tipos a considerar:

- Monitoramento contínuo do transformador quando há sensores e sistema supervisorio instalados permitindo assim, realizar análise e diagnósticos de falhas em tempo real

(BUSTAMANTE et al., 2019). Entretanto, apesar das tecnologias de supervisão em tempo real já estarem disponíveis, devido ao seu alto custo de implantação, torna-se impeditivo seu emprego generalizado em transformadores de menor potência sendo, portanto, destinados a equipamentos de altas potências ou considerados críticos (BENEDET, 2008). Atualmente, novas tecnologias estão em desenvolvimento, como por exemplo, nano sensores que detectam a presença dos diferentes tipos e composições de gases formados no interior dos transformadores, para que alcancem um custo suficientemente baixo e tornem viável a aplicação em larga escala da supervisão *online* contínua. Em Sun, Ohodnicki e Stewart (2017), são analisados as tecnologias atuais e tendências futuras no desenvolvimento dos sensores e nano sensores para leitura de diferentes tipos de gases; e

- Monitoramento intermitente por meio de medições e avaliações periódicas de amostras *in loco* (com uso de equipamentos portáteis) ou em laboratórios para diagnóstico incipiente de falhas. Este tipo de acompanhamento reserva-se aos diagnósticos derivados de informações em sua grande maioria das análises físico-químicas e de gases. Além disso, trata-se do meio mais utilizado na análise do óleo isolante, tendo em vista o custo elevado do monitoramento contínuo.

Na abordagem *offline* os ensaios são realizados com os transformadores desconectados do sistema elétrico, sendo essa condição normalmente empregada em: inspeções programadas; caso de suspeita de falha existente; equipamento retirado da instalação e sob manutenção; ou em ensaios de tipo e entrada em operação. Isso se deve às características dos próprios métodos de análises, por questões de segurança e de redução de ruídos, que podem interferir no resultado do diagnóstico. Dentre as técnicas *offline* emergentes para avaliação do sistema de isolamento, podem ser citadas (PAULINO, 2014abc; NETO; ASSUNÇÃO; ASSUNÇÃO, 2009; AZEVEDO, 2009):

- Medida da capacitância e fator de dissipação (ou potência) para avaliar a degradação do sistema isolante do equipamento;
- Medida da tensão de retorno e método de polarização (carga) e despolarização (descarga) da corrente são usados para estimar o teor de umidade na isolamento sólida;
- Medidas da resistência de isolamento e do índice de polarização são usados como indicadores de contaminação e umidade presente no papel isolante;



- Método de descargas parciais (DP), utilizado para avaliar a existência de falhas ou defeitos internos ao transformador, produzindo descargas elétricas, que podem romper o dielétrico entre dois condutores adjacentes ou não. Há equipamentos que medem a DP por meio do método acústico e com o transformador *online*, entretanto, em virtude da presença de ruídos de diversas origens que podem interferir na localização correta da falha, é recomendado seu uso em laboratório, muito embora existam estudos para filtragem; e
- Medição da resposta em frequência (FRA - *Frequency Response Analysis*) e impedância terminal são usadas para identificar possíveis alterações geométricas das bobinas causadas por movimentações ocorridas na fase de transporte do equipamento, ou por esforços mecânicos, devido a curtos-circuitos ou descargas atmosféricas.

## 2.7. Análise de criticidade envolvendo falhas em transformadores

Como visto, diversas técnicas para avaliar e diagnosticar falhas em transformadores estão disponíveis para uso, que associadas ao conhecimento de outras informações (como por exemplo, grau da falha diagnosticada, ocorrência de eventos recentes na subestação, histórico de manutenção, histórico de operação, problemas crônicos em modelo, idade, entre outros), podem constituir ferramentas de análises de falhas, de manutenção preditiva, ou de gestão de ativos que apoiam a tomada de decisão de retirada do equipamento de operação, uma vez que há impactos financeiros envolvidos.

Tais ferramentas permitem avaliar as condições críticas operacionais dos equipamentos e, assim, propor soluções que evitem ou minimizem os prejuízos financeiros decorrentes de falhas e redução da vida útil do equipamento, seja por meio da decisão de atrasar ou retirar imediatamente o transformador do sistema para manutenção ou preservação do ativo, ou admitir a sobrecarga em situações de emergências conhecendo-se os riscos envolvidos, ou até promover sua substituição.

Como objetivos da análise de criticidade dos transformadores, podem ser encontrados modelos que consideram as análises de variáveis envolvendo somente as análises de gases dissolvidos e físico-química, a modelos que usam informações do banco de dados dos ativos da manutenção e operação (carregamento, temperatura, entre outros). Já a nível de gestão de risco, além de considerar a criticidade do equipamento, observa-se a influência dos fatores econômicos operacionais, de modo a buscar preservar a continuidade no fornecimento do serviço. Segundo Sousa, Saavedra e Lima (2018), com a análise de risco é possível determinar

as características críticas operacionais e fornecer soluções que possam prevenir ou minimizar os prejuízos financeiros e, assim, permitir a adoção de diferentes abordagens para melhorar a confiabilidade do transformador.

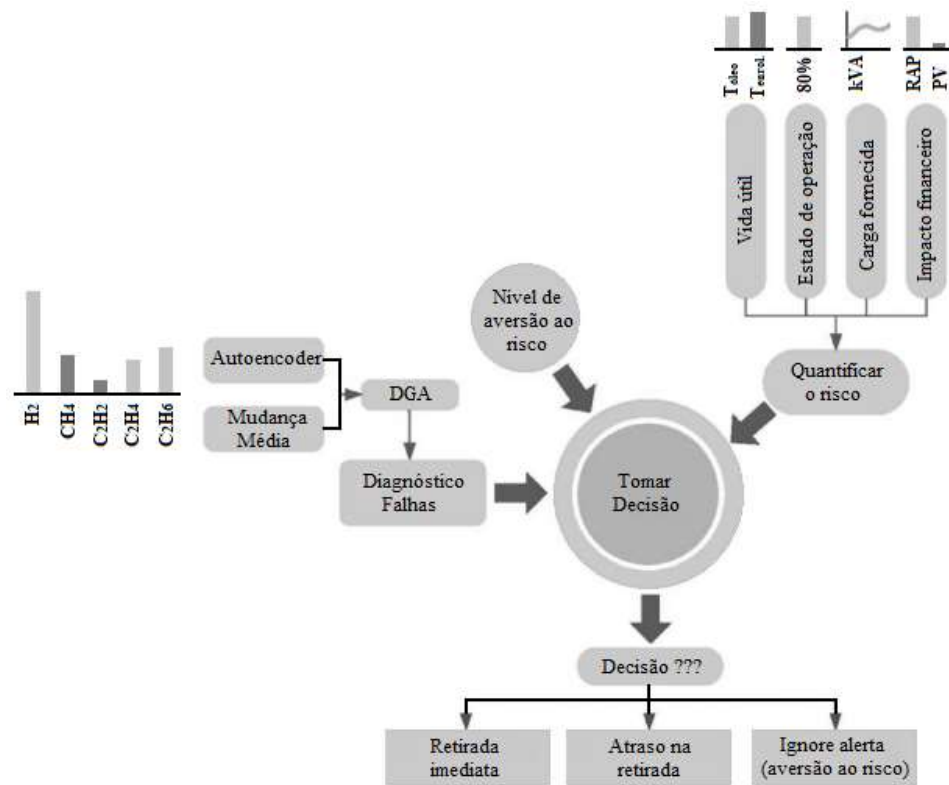
Para Lima, Saavedra e Miranda (2015), há necessidade de avaliar os riscos relacionados a um atraso no atendimento de curto prazo para melhor planejamento, ou de longo prazo, ou até a retirada imediata (mais custosa financeiramente) e, dessa forma, propuseram uma estrutura em dois níveis para diagnóstico de falhas e tomada de decisão para transformadores de potência. No primeiro nível, monitora-se a existência de falhas incipientes nos transformadores (via redes neurais associativas) e no segundo, são avaliados e estimados os riscos envolvidos observando-se a vida útil, estado operacional em serviço, carregamento (carga fornecida) e o impacto financeiro, sempre procurando preservar as demandas e continuidades do serviço.

Logo, com essa metodologia torna-se possível analisar diferentes cenários de riscos como por exemplo, somente com sobrecarga, somente com indicativo de falhas, e com ambas sobrecarga e indicativo de falhas, permitindo avaliar os impactos resultantes sobre os transformadores e, assim, auxiliar o operador para a tomada de decisão quanto à seleção de quais equipamentos podem ser desligados, ou mantidos em serviço por um maior tempo no atendimento à emergência, a fim de proporcionar uma menor perda financeira para a concessionária.

A Figura 3 mostra as principais etapas do processo de decisão, em que é possível observar:

- O diagnóstico de falhas com base na análise de gás dissolvido (DGA), uso de redes neurais associativas e mudança média;
- O estimador de risco que avalia a vida útil, a condição operacional, o carregamento e impacto financeiro;
- O grau de aversão ao risco atribuído pelo operador; e
- A decisão envolvendo a retirada ou adiamento do equipamento de serviço.

Figura 3 – Estrutura de tomada de decisão para transformador em operação

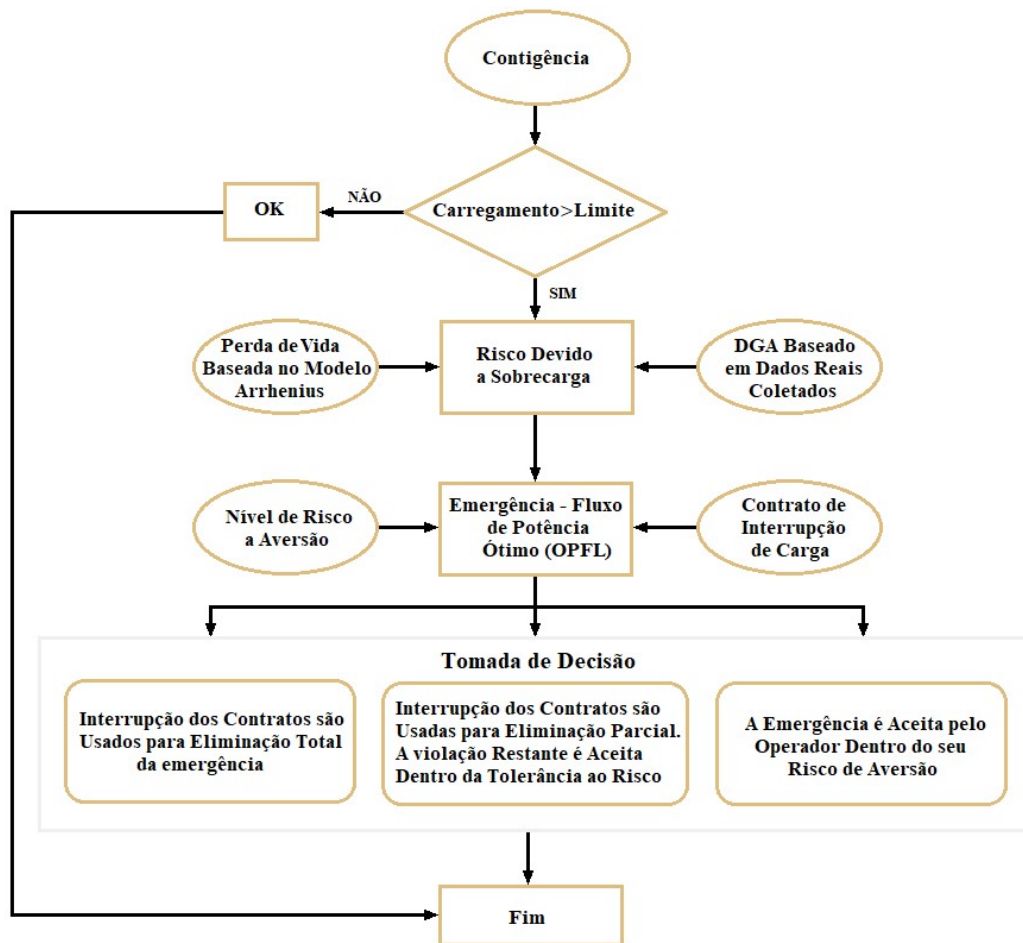


Fonte: Adaptado pelo autor de Lima, Saavedra e Miranda (2015)

Já Sousa, Saavedra e Lima (2018) propuseram uma metodologia de gestão de risco para a tomada de decisão de contingenciamento que visa ao atendimento em operações emergenciais do transformador. O modelo considera como fatores de análise: os custos do contrato de interrupção que permite aceitar ou não a sobrecarga (neste último caso, em troca de compensação financeira), o risco relacionado à perda da vida útil, a existência de falhas incipientes (via Doernenburg) e o grau de aversão do operador ao risco.

A Figura 4 mostra as etapas desse processo, em que se analisa inicialmente, a necessidade de contingenciamento dos recursos, observando-se de início, o limite admissível do carregamento. Caso ele seja maior do que o valor admitido pelo transformador, analisa-se o risco de redução da vida útil, decorrente da sobrecarga e a probabilidade de existência de falha durante o período do contingenciamento, que junto com o risco de aversão definido pelo operador e o custo envolvido pelo contrato de interrupção de carga, define-se o fluxo de potência ótimo (OPFL - *Optimal Power Flow*) e então, se toma a decisão quanto ao atendimento à emergência. Por meio dessa metodologia, torna-se possível encontrar um ponto de equilíbrio entre o risco assumido e o custo envolvido orientando o operador pela melhor decisão de contingenciamento.

Figura 4 – Estrutura de decisão de contingenciamento para transformador em operação



Fonte: Adaptado pelo autor de Sousa, Saavedra e Lima (2018)

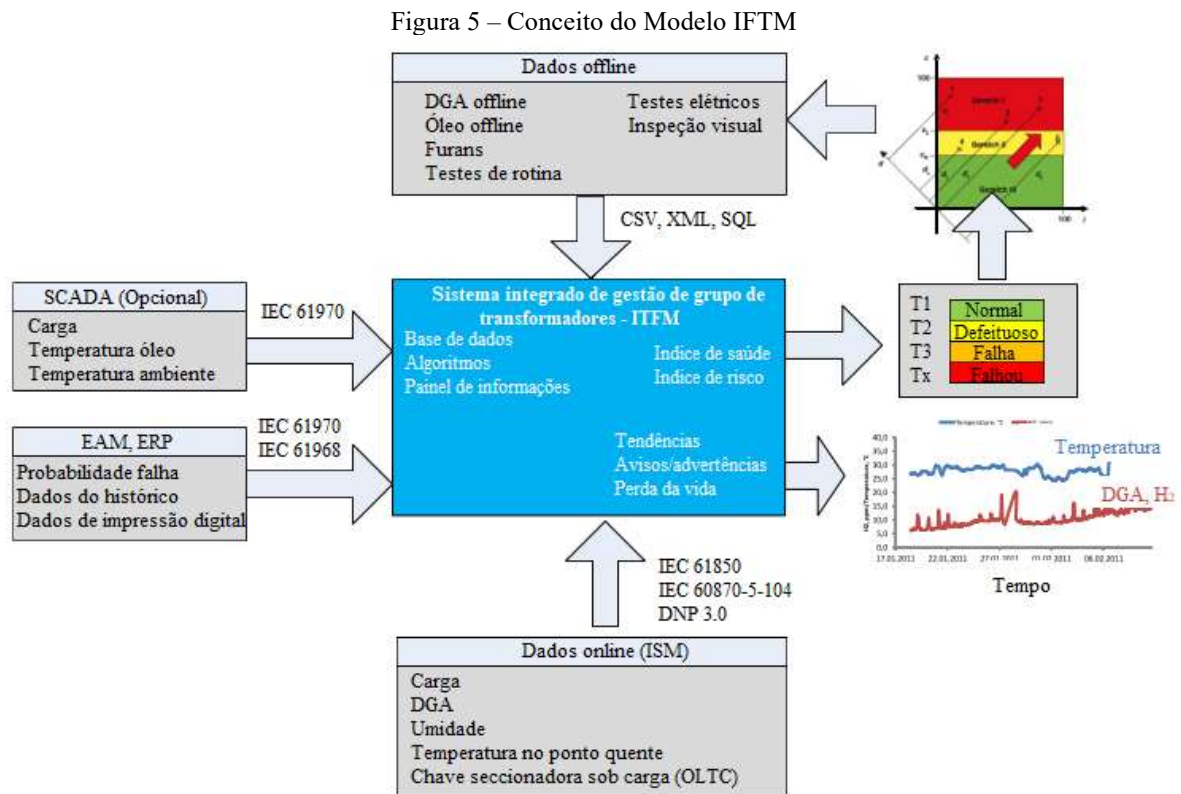
Ilgevicus e Babizki (2015) propuseram um modelo de gestão de falhas em transformadores denominado de ITFM (*Integrated Transformer Fleet Management*), que consideram o índice de saúde (HI – *Healthy Index*) e criticidade relacionados a eles, atendendo as diretrizes das normas: inglesa PAS 55, alemã DIN VDE 0109, e IEC 61970. A estrutura ITFM proposta para condição de avaliação do transformador (Figura 5) possui três níveis:

- No primeiro, o monitoramento e diagnóstico são sem interrupção (*online*), neste nível são observados dados baseados no tempo, como por exemplo, temperatura do óleo, carregamento, DGA e compostos furânicos, análise físico-química do óleo, e dados do histórico em serviço;
- No segundo, em que há necessidade de uma breve interrupção do equipamento para medições *offline* de rotina para diagnóstico, como por exemplo, medidas elétricas; e

- No terceiro, quando a interrupção é maior para medições *offline* avançadas, como por exemplo, análise da resposta em frequência, corrente de polarização e despolarização, espectroscopia no domínio da frequência, entre outros.

Em ambos os níveis, 1 e 2, todas as análises são pontuadas para a composição de 8 grupos de risco (mecânico, térmica, dielétrica, seccionadoras, buchas, sistema de resfriamento, tanque/reservatório e acessórios, e histórico de confiabilidade), que definirá o fator do HI (cores e números) para priorização das ações de manutenções nos transformadores.

Ainda de acordo com Ilgevicus e Babizki (2015), para a determinação da priorização, o HI também considera sobre o sistema de decisão os fatores econômicos envolvidos com a interrupção no fornecimento de energia, a estabilidade do sistema, e as condições de reparos ou penalidades existentes, que são convertidos em informações significativas para o gestor responsável pelos ativos e/ou manutenção. Essa metodologia fornece diferentes índices de HI que permitem distinguir e definir as prioridades de atendimento no grupo de transformadores, segundo um valor de HI limite definido pelo operador (por exemplo,  $HI > 30\%$ ), como risco de falha antes que esse evento ocorra, de fato.



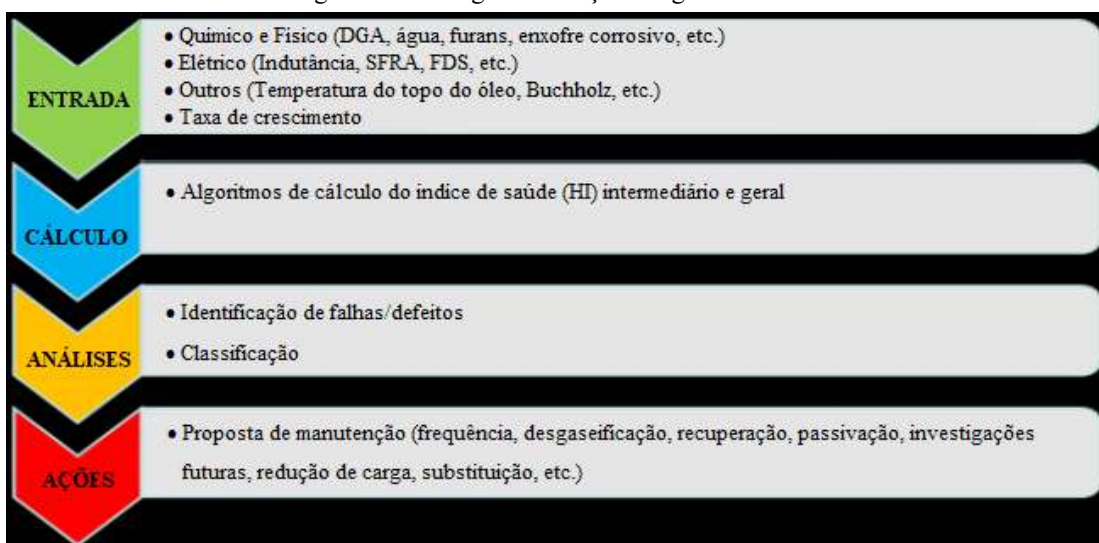
Fonte: Adaptado pelo autor de Ilgevicus e Babizki (2015)

Em Scatiggio, Pompili e Calacara (2018) é apresentada a abordagem do índice de saúde (HI) empregada pela empresa Terna, a principal operadora de sistemas de transmissão de eletricidade na Europa. Na metodologia proposta, a avaliação do HI do transformador (Figura 6) ocorre em 4 etapas:

- Na primeira são coletados (modo *online* ou *offline*) diversos parâmetros elétricos, físicos, químicos, entre outros, e então reunidos em 4 grupos distintos de risco (dielétricos, térmicos, mecânicos e condições do óleo);
- Na segunda, são realizados os cálculos individuais e geral do HI avaliando o resultado com base em diversas normas (da empresa, IEC, IEEE);
- Na terceira é constituída uma classificação do grupo de transformadores para definir prioridades de manutenção e substituição; e
- Na última etapa examina-se individualmente as condições do equipamento, com sugestões de correção ou atenuantes, para aumentar sua confiabilidade e vida útil.

Cinco classificações foram definidas para o HI do equipamento, sendo elas: bom, razoável, ruim, muito ruim e crítica, cada qual com limites de valores definidos para o estado em questão. O uso do HI no sistema da Terna a partir de 2010, contribuiu para a redução drástica do número de falhas menores, não sendo contabilizadas as falhas mais graves, uma vez que sempre estavam associadas a eventos de características imprevisíveis (descargas atmosféricas e falhas abruptas).

Figura 6 – Fluxograma de ações originadas do HI



Fonte: Adaptado pelo autor de Scatiggio, Pompili e Calacara (2018)

Outras aplicações de avaliação de risco em transformadores de potência, como proposto por Aizpurua et al. (2018), apresenta uma formulação do HI baseada em probabilidades envolvendo o aprendizado de máquina na análise de DGA, valores absolutos dos gases-chave e a taxa de evolução. Arshad, Islam e Khaliq (2014) propõem o uso de lógica *fuzzy* na gestão e tomada de decisões envolvendo transformadores de potência, em que são observados seis grupos de criticidade (térmica, elétrica, estado operacional, integridade do enrolamento, qualidade do óleo e isolamento) para identificar e integrar os pontos críticos e fornecer a informação de classificação entre excelente até muito pobre (valor entre 0 e 1) para o gerenciamento do ativo.

Assim, no contexto de análise de criticidade dos trabalhos aqui referenciados, como parte integrante das metodologias, há necessidade de conhecer a existência de falhas incipientes dentro do processo de análise, uma vez que impactam diretamente no índice de classificação ou tomada de decisão, e o conhecimento dessas falhas podem ser obtidos com as informações provenientes da análise e interpretação dos gases dissolvidos no óleo isolante.

## **2.8. Considerações gerais do capítulo**

Compreende-se que os transformadores de potência são elementos considerados essenciais ao funcionamento do sistema elétrico e sua saída repentina de operação ocasiona problemas socioeconômicos aos consumidores e concessionárias, porém cabe ao último, fazer a gestão do fornecimento de energia de modo seguro e com qualidade a partir de um melhor controle de falhas sobre esses equipamentos.

As origens das falhas são inúmeras e ocorrem em diferentes pontos e elementos que compõem a estrutura do transformador. No uso da análise do óleo isolante na detecção de falhas, o foco é direcionado aos elementos internos do equipamento, por estarem em contato com o óleo, em que são observados os dielétricos (do isolante líquido e sólido), cabos e conexões, comutadores, erros de projeto e montagem inadequada.

No contexto das análises do óleo isolante e sua relação com o diagnóstico de falhas, pode-se verificar diversas técnicas e ensaios para avaliar a saúde do transformador, cada qual atuando de um modo específico para identificar e diagnosticar determinados tipos de falhas.

A análise físico-química e de impurezas tem forte relação com a preservação da vida útil do transformador. Manter suas características dentro das especificações técnicas é requisito mínimo para que a condição de preservação prevaleça. De outro modo, o não

tratamento do óleo pode, a curto ou longo prazo, levar o equipamento à condição de falha, cujo momento exato de sua ocorrência é de difícil previsão.

A análise dos gases dissolvidos (DGA) é uma técnica estabelecida (DUVAL, 2012; TENBOHLEN et al., 2016), normatizada (IEC, 2015; IEEE, 2019), e é um dos tipos de análises que permite diagnosticar a evolução de uma falha interna incipiente, e assim, antecipar e corrigir eventuais problemas com ações de manutenção preditiva e preventiva em seu estágio inicial de formação. Para esse tipo de análise, não há necessidade de retirar o transformador de serviço, além de possibilitar medições de gases e verificar o estado dele em tempo real com uso de equipamento de monitoramento contínuo acoplado. Portanto é sobre essa perspectiva que o presente trabalho aborda as análises dos gases combustíveis e seus benefícios no diagnóstico de falhas incipientes em transformadores de potência.

Em complemento, tem-se observado um crescente desenvolvimento de ferramentas de análise de criticidade que contribuem para o gerenciamento dos ativos quanto à tomada de decisão e criação de estratégias de operacionais, de análise de falhas e manutenção preditiva dos transformadores, dentre os quais a DGA é parte integrante desse processo. Quanto aos classificadores de falhas, alguns dos trabalhos mostrados não definem o uso de algoritmos ou técnicas específicas, logo, podem ser associados às metodologias de análise de criticidade, os algoritmos de aprendizado de máquinas ou os normatizados que apresentem melhor eficácia na identificação de falhas.



## CAPÍTULO 3 – ESTUDOS SOBRE TÉCNICAS DE DIAGNÓSTICO DE FALHAS EM TRANSFORMADORES E APRENDIZADO DE MÁQUINA

### 3.1. Métodos e técnicas normatizadas e convencionais

A análise de gás dissolvido (DGA) no óleo isolante é uma ferramenta amplamente empregada no diagnóstico do estado operacional dos transformadores (EECKHOUDT et al., 2017) como forma de identificar a presença de possíveis falhas incipientes e assim nortear as ações necessárias, sejam para realizar ensaios auxiliares de verificação, ou para programar inspeção e manutenção do equipamento.

Entretanto, devido ao comportamento não linear na produção dos gases dissolvidos no óleo, relacionados aos estresses térmicos e elétricos, reações químicas internas entre materiais e outros, não é uma tarefa simples diagnosticar de modo preciso o tipo de falha e, portanto, diversas técnicas de interpretações e diagnósticos convencionais encontram-se disponíveis para uso, conforme mostrado em Cheng e Yu (2018), e Faiz e Solemani (2017). É importante observar que a DGA não é uma ferramenta adequada para uso e prevenção de falhas instantâneas (SUN; HUANG; HUANG, 2012).

As normas IEC 60599 (IEC, 2015) e IEEE C57.104 (IEEE, 2019) relacionam diversos métodos convencionais de diagnóstico de DGA que podem envolver a análise de valores individuais (percentuais relativos) dos gases combustíveis, suas razões e evoluções das concentrações no tempo, conforme apresentados brevemente neste capítulo.

#### 3.1.1. Gás chave

A análise do gás chave é considerada a técnica mais simples de diagnóstico de falhas uma vez que se baseia no maior valor percentual relativo do gás produzido (denominado gás chave) em decorrência da falha relacionada. Na decomposição térmica do óleo e celulose tem-se a predominância dos gases  $C_2H_4$  e  $CO$  respectivamente, para descarga elétrica de baixa energia predomina-se o  $H_2$  e em descargas de alta energia (arco) o gás  $C_2H_2$ . Como esse método não usa correlação numérica entre o gás chave e o tipo de gás envolvido, para um diagnóstico mais preciso, é necessário ser experiente nessa área (GRAY, 2010; SUN; HUANG; HUANG, 2012).

### 3.1.2. Triângulo de Duval

Triângulo de Duval (DUVAL, 2008a) trata-se de uma técnica gráfica de diagnóstico de falha que emprega os percentuais relativos de três tipos de gases em suas arestas formando um triângulo com regiões mapeadas por falhas identificadas (descarga parcial, falha elétrica de alta e baixa intensidade de energia, três tipos de falhas térmicas e um misto de falha térmica com elétrica). A partir dos valores percentuais relativos dos gases calculados, sobre a representação gráfica do triângulo traçam-se retas concorrentes, que definem um ponto comum sobre a região de falha conhecida (mais informações, consulte o Apêndice B1).

Essa mesma técnica possui cinco variações para dois tipos de óleos. Para óleo mineral: Triângulo 1 – trata-se da variação mais empregada, o Triângulo 2 – para uso em equipamentos com comutador sob carga, Triângulos 4 e 5 – são complementares ao 1 na análise de falhas térmicas de baixa e alta temperatura, respectivamente. Para óleo não mineral – Triângulo 3.

Quanto aos tipos de gases envolvidos, os Triângulos 1, 2 e 3 usam os gases  $\text{CH}_4$ ,  $\text{C}_2\text{H}_2$  e  $\text{C}_2\text{H}_4$ , o Triângulo 4 os gases  $\text{H}_2$ ,  $\text{CH}_4$  e  $\text{C}_2\text{H}_6$ , e o Triângulo 5, os gases  $\text{CH}_4$ ,  $\text{C}_2\text{H}_4$  e  $\text{C}_2\text{H}_6$ . Como esta técnica não possui diagnóstico para condição de operação normal do equipamento, é importante conhecer os limites dos valores típicos dos gases identificados como normais, principalmente em transformadores que operam há muito tempo no sistema elétrico com concentrações de gases elevadas, pois sem esse conhecimento prévio, falsos diagnósticos serão obtidos quando analisadas as amostras. Em Duval (2008b) é demonstrado como determinar os valores típicos para os gases avaliados.

### 3.1.3. Razões dos gases combustíveis

Os métodos que envolvem as razões dos gases podem relacionar até seis tipos de gases combustíveis ( $\text{H}_2$ ,  $\text{CH}_4$ ,  $\text{C}_2\text{H}_2$ ,  $\text{C}_2\text{H}_4$ ,  $\text{C}_2\text{H}_6$  e  $\text{CO}$ ) para identificar os tipos de falhas nos transformadores (IEEE, 2019; IEC, 2015). As combinações adotadas para os métodos descritos a seguir, são:  $R_1 = \text{CH}_4/\text{H}_2$ ,  $R_2 = \text{C}_2\text{H}_2/\text{C}_2\text{H}_4$ ,  $R_3 = \text{C}_2\text{H}_2/\text{CH}_4$ ,  $R_4 = \text{C}_2\text{H}_6/\text{C}_2\text{H}_2$ ,  $R_5 = \text{C}_2\text{H}_4/\text{C}_2\text{H}_6$  e  $R_6 = \text{C}_2\text{H}_6/\text{CH}_4$ .

Para valores de gases nulo ou considerados não detectáveis no denominador da razão, recomenda-se o uso do limite inferior de detecção do gás dos equipamentos de medições ou valores especificados conforme normas aplicáveis (ABNT, 1981; IEC, 2005; USDI, 2000).

### 3.1.3.1. Rogers

No método original de Rogers são empregadas quatro relações de gases ( $R_1$ ,  $R_2$ ,  $R_5$  e  $R_6$ ) para diagnosticar até doze possíveis tipos de falhas nos transformadores (GEDAM; SWAMI; THOSAR, 2015; SARMA; KALYANI, 2004), contudo a norma IEEE C57.104 (IEEE, 2019) apresentou algumas readequações em relação ao método original, com objetivo de simplificar seu uso. No caso, passou-se a considerar apenas três relações de gases (excluiu-se  $R_6$ , visto que se referia apenas ao limite de decomposição por temperatura), eliminou-se o uso da codificação para classificação dos tipos de falhas e reduziu o número de classes de falhas, juntando rótulos em grupos específicos de falhas relacionadas, ampliando assim as possíveis combinações dessa variáveis (mais informações, consulte o Apêndice B2).

Os tipos de falhas diagnosticáveis passaram a ser: normal, descarga de baixa energia, descarga de alta energia, e três tipos de falhas térmicas (para diferentes faixas de temperaturas), e podem ser obtidas após calcular as razões dos gases e aplicar as restrições das faixas limites dos gases que levam à identificação. Uma nova análise de gases e investigação deve ser realizada, caso exista combinações de razões que resultem em falhas não reconhecidas (sem diagnóstico).

### 3.1.3.2. Doernenburg

Neste método são usados os valores obtidos de quatro relações de gases ( $R_1$ ,  $R_2$ ,  $R_3$  e  $R_4$ ) para diagnosticar as possíveis falhas nos transformadores, entretanto há necessidade de que sejam validadas por meio do atendimento obrigatório a três condições de concentrações mínimas de gases envolvidos (IEEE, 2019): a primeira condição é que no mínimo um entre os gases  $H_2$ ,  $CH_4$ ,  $C_2H_2$  e  $C_2H_4$  tenha o dobro do valor da concentração limite dos gases  $L_1$  especificados pela IEEE C57.104 (maiores detalhes - Apêndice B); a segunda é que o gás  $C_2H_6$  ou  $CO$  apenas exceda o seu valor de concentração limite  $L_1$ ; e se apenas uma (ou nenhuma) das condições anteriores for atendida, o transformador é classificado como normal, caso contrário, passa-se para a terceira e última condição de validação, em que ao menos um dos gases contidos em cada uma das razões exceda a concentração limite  $L_1$  e, não sendo essa última condição atendida, recomenda-se a análise de uma nova amostra ou usar outros meios de investigação, como por exemplo, testes complementares e levantamento de informações de ocorrências.

Somente após as amostras serem validadas pelas três restrições de uso, calcula-se as razões e aplica-se às restrições das faixas limites dos gases para obter as classificações das falhas dentre três possíveis: descarga parcial, arco e falhas térmicas (mais informações, consulte o Apêndice B4). Nesse método, devido às restrições de faixas classificatórias serem incompletas, ele pode apresentar muitos resultados de falhas não identificadas (SUN; HUANG; HUANG, 2012).

#### 3.1.3.3. IEC 60599 e NBR 7274

As normas NBR 7274 (ABNT, 2012) e IEC 60599 (IEC, 2015) abordam métodos similares nas caracterizações dos diagnósticos de falhas de transformadores e compartilham das mesmas três razões ( $R_1$ ,  $R_2$  e  $R_5$ ) de gases do método Rogers da IEEE, porém diferenciam-se entre si em alguns valores limites das faixas de análise e diagnóstico.

Nesses métodos, para os cálculos das razões dos gases, destacam-se as recomendações em que gases com valores de concentrações  $0 \mu\text{l/l}$  (ZERO microlitro/litro) ou abaixo do valor do limite de sensibilidade de detecção analítico dos equipamentos de medições (ABNT, 1981; IEC, 2005) devem ser substituídos pelo valor  $S_{LIM}$  e ao menos uma das concentrações dos gases envolvidos deve estar acima do valor típico e da taxa evolução dos gases (valor de referência da norma).

O valor típico refere-se, por exemplo, ao valor em que 90% das amostras (equipamentos de um mesmo tipo) estão abaixo das concentrações dele, entretanto tal valor pode ser definido pelo especialista. Após readequar os dados conforme recomendações, calculam-se as razões e aplicam-se as restrições das faixas limites dos gases para classificar as falhas (mais informações, consulte o Apêndice B2).

#### 3.1.4. Outras técnicas convencionais de DGA

Com base nos dados de falhas de DGA da IEC TC10, cinco tipos de gases combustíveis e a análise de 15 combinações entre seis razões de gases (como por exemplo,  $R_1$ - $R_2$ ,  $R_1$ - $R_3$ ,  $R_1$ - $R_4$ , ...,  $R_5$ - $R_6$ ), Kim et al. (2013) desenvolveram um método de diagnóstico de falhas em transformadores que selecionou três razões que melhor representavam as falhas e também restrições das faixas limites dos gases para classificá-las.

Lee et al. (2013) propuseram dois métodos gráficos de identificação de falhas baseados no uso dos percentuais relativos de quatro tipos de gases combustíveis. O primeiro

método deu origem a um quadrilátero, com regiões de falhas elétricas e térmicas mapeadas, que permitiu graficamente classificar novas amostras. No segundo método, empregou-se quatro combinações com o percentual relativo de dois tipos de gases cada, formando assim quatro triângulos mapeados com as regiões das falhas elétricas e térmicas.

Duval e Lamarre (2014) desenvolveram um método gráfico denominado “Pentágono” com uso do percentual relativo de cinco tipos de gases. Foram elaborados dois pentágonos mapeados com regiões de falhas elétricas e térmicas, esse último incrementado com novas classes de falhas térmicas de baixa temperatura. Eles podem ser usados de modo individual nos processos de identificações de falhas, ou combinados com as variações do triângulo de Duval, para permitir distinguir falhas térmicas de menor intensidade ( $T < 250^{\circ}\text{C}$ ).

Baseado no método do pentágono, Gouda, El-hoshy e El-tamaly (2017) desenvolveram o método gráfico do heptágono. Esse modelo considerou sete tipos de gases (seis gases combustíveis mais o  $\text{CO}_2$ ) para a realização do diagnóstico.

Zhao et al. (2014) propuseram um método de diagnóstico que utiliza o conteúdo relativo de quatro tipos de gases combustíveis dissolvidos no óleo combinados. Com base nesses gases, definiram-se inicialmente dez razões entre eles e então avaliaram suas relações com os tipos de falhas térmicas e elétricas encontradas. Após a análise, foram escolhidas seis razões com fortes conexões com os tipos de falhas e determinado um conjunto de regras restritivas e classificatórias para tais razões.

Todos os métodos citados anteriormente foram comparados com os métodos normatizados (triângulo de Duval, Rogers, Doernenburg, IEC, entre outros) e apresentaram uma melhor eficácia na classificação de falhas. Ressalta-se que, como todos esses métodos são desenvolvidos com base em amostras de gases que representam apenas as falhas, eles não preveem a condição de normalidade. Logo, um conhecimento prévio sobre o estado das amostras como normal é pré-requisito para o uso da maioria desses métodos.

### **3.2. Métodos e técnicas de inteligência computacional**

Embora algumas das técnicas convencionais estejam presentes em normas (TENBOHLEN et al., 2016), frequentemente elas são associadas às lacunas que existem nas identificações dos tipos de falhas ou combinações de gases, levando a resultados inconsistentes ou ambíguos e, portanto, ainda não há técnica única universalmente aceita (VENKATASAMI; LATHA, 2016). Essa é uma das razões que tem impulsionado o desenvolvimento de diversas pesquisas com diferentes tipos de algoritmos de aprendizado de

máquina, não somente para melhorar a assertividade dessas técnicas analíticas convencionais compensando suas incertezas, como por exemplo, causa e efeito das falhas, julgamento dos dados dos gases, medições inadequadas, entre outros (CHENG; YU, 2018), mas também para melhorar a eficiência entre os próprios algoritmos de aprendizado de máquina, seja por meio de associações entre diferentes classificadores, associações com algoritmos de otimização, emprego de métodos estatísticos, uso diferentes heurísticas e outros.

Nesse contexto, diversas técnicas ou algoritmos classificadores, como por exemplo, Floresta de caminhos ótimos (OPF), K vizinhos mais próximos (KNN), Rede Neural Artificial (ANN), Lógica *Fuzzy*, Nãive Bayes (NB), Máquina de Vetor Suporte (SVM), Máquina de Aprendizado Extremo (ELM), e muitas outras têm sido apresentadas.

Adicionalmente, no âmbito de desenvolvimento dessas técnicas inteligentes, existem variações que buscam ainda melhorar os resultados das classificações de padrões, por meio: do emprego de recursos de outras técnicas associadas (SONG; XIU-YING; WEN-XU, 2011; AMORA et al., 2012); métodos de otimização (SHAWE-TAYLOR; SUN, 2011; NUR; HAIZAN; OSMAN, 2014; SUGUNA; THANUSHKODI, 2010) e associações híbridas (LI et al., 2017; SINGH; UPADHYAY, 2017).

Neste tópico de métodos e técnicas de inteligência computacional são abordadas resumidamente algumas técnicas de inteligência computacional e, entre elas, algumas das mais usadas nos estudos que envolvem a classificação de falhas em transformadores, contudo maiores detalhes podem ser obtidos, consultando-se as bibliografias relacionadas.

### 3.2.1. Lógica *Fuzzy*

Trata-se de uma técnica já consolidada e empregada comercialmente em diversos tipos de aplicações (GUPTA, 2017). A teoria *fuzzy* busca lidar com as incertezas subjetivas (ambiguidade, imprecisão), por meio de representações matemáticas, empregando a teoria de conjuntos e funções de associações (ou pertinências), e assim tornar a decisão lógica de uma máquina mais próxima da humana, ou seja, uma decisão não baseada apenas em decisões binárias (certo ou errado), mas com a existência de valores intermediários (por exemplo, talvez) permitindo assim representar melhor a realidade.

Basicamente, a relação entre a lógica *fuzzy* e o DGA pode ser estabelecida a partir de um conjunto de dados de DGA dos transformadores, em que se define as variáveis de trabalho e regras *fuzzy*, e então aplica-se o processo de “fuzzificação” e “defuzzificação” para obter a classificação dos padrões (YUAN, 1997). Ibrahim, Sayed e El-zahab (2014) propuseram uma

lógica *fuzzy*, empregando como variáveis principais de trabalho três razões de gases e classes de falhas da IEC 60599, uso da função de pertinência Gaussiana com método de composição *Max* e *Min*, e sistema de inferência Mamdani. Em Su (2016), o método *fuzzy* proposto utiliza quatro razões de gases como variáveis principais do problema e também as classes de falhas IEC 60599, função de pertinência Demi-Caught e vetor de diagnóstico, determinado também pela regra de composição *Max* e *Min*.

Em Soni e Chaudhari (2016) são avaliados diversos métodos de falhas (IEC 60599, Doernenburg, Rogers, Cigré, razão CO/CO<sub>2</sub>) combinados com lógica *fuzzy* para contornar as limitações (diferentes classificações, dependência do especialista, falta de diagnóstico) que eles apresentam. As variáveis principais avaliadas trata-se das razões dos gases presentes nos métodos citados e que combinam até sete tipos de gases associados em conjunto com a função de pertinência triangular. Cada método normatizado possui sua regra *fuzzy* e saída (“defuzzificação”) própria para a classificação das falhas.

Khan, Equbal e Islam (2015) propuseram três métodos (IEC 60599, Doernenburg e Rogers) combinados com a lógica *fuzzy* tradicional e a adaptativa ANFIS (*Adaptive Neuro-Fuzzy Inference System*) que utiliza regra de aprendizagem híbrida composta de ANN e *fuzzy* para a análise lógica e computacional. Na comparação entre a lógica tradicional *fuzzy* e ANFIS, o último apresentou resultado superior nos diagnósticos das falhas.

Ma, Ekanayake e Saha (2012) propuseram o uso de técnicas híbridas, com o objetivo de minimizar a influência dos *outliers* e ruídos (presentes no conjunto de dados dos gases usados no treinamento da rede) conjugando *fuzzy* SVM (FSVM - *Fuzzy Support Vector Machine*) com a técnica *fuzzy* c-médias (FCM – *Fuzzy C-means*), devido à sua eficácia em lidar com esses tipos de incertezas, e com *fuzzy* C-Médias *kernel* (KFCM – *Kernel Fuzzy C-means*). Em ambas as técnicas, as amostras de DGAs identificadas como *outliers* ou ruídos recebem um peso menor e passaram a ser tratadas como menor importância no processo de treinamento da rede.

Para Cheng e Yu (2018), as seguintes vantagens e desvantagens podem ser associadas à teoria *fuzzy*: bom para resolver problemas de incertezas e imprecisão; como as relações entre fenômenos, causas, tipos e mecanismo de falha são incertas, é difícil ajustar a função de pertinência; as informações das amostras precisam ser completadas na tabela de regras *fuzzy*; e podem ser aplicadas em conjunto com diversas outras técnicas, por exemplo, teoria dos conjuntos aproximados (RST - *Rough Set Theory*), probabilísticas, ANN, entre outras.

### 3.2.2. Redes Neurais Artificiais (ANN)

A ANN trata-se de uma representação matemática do funcionamento de uma rede neural biológica humana (NETO; ASSUNÇÃO; ASSUNÇÃO, 2009). Como classificador de padrões, é possível fazer o mapeamento de um conjunto de dados de interesse e assim extrair informações sobre características e propriedade dos fenômenos que os envolvem (IKESHOJI et al., 2018), tendo como resultado a geração de um espaço de características ou atributos de  $m$ -dimensão, constituído de regiões definidas por classes distintas que são separadas por hiperplanos. As classificações de novas amostras são feitas com base nas diferentes regiões do espaço de atributos que elas ocupam. Basicamente, a estrutura de uma ANN é constituída de camadas de neurônios, pesos, bias e funções de ativações, conectadas entre si e necessita de etapa de treinamento e teste para a apresentação dos resultados de classificações.

A escolha da configuração da rede não é uma tarefa simples, pois existem diversas variáveis de projeto (topologia, estratégia de aprendizado, dados de interesse e funções de ativações) que devem ser alinhadas com a aplicação requerida. Na estratégia de aprendizado supervisionado, a topologia de uso mais comum é *FeedForward Backpropagation*, completamente conectados e com três camadas (entrada, intermediária e saída) (MIIKKULAINEN, 2011).

Em trabalhos que envolvem classificação de falhas em transformadores, a maioria deles empregam a estratégia supervisionada cujos rótulos são os padrões de falhas encontrados nas inspeções dos equipamentos. Na etapa de treinamento, o conjunto de dados de treino é apresentado à rede, e por meio de processos iterativos são ajustados os pesos para que a maioria das amostras sejam corretamente identificadas. Na etapa de teste, apresenta-se o conjunto de dados de teste para a rede treinada e os resultados obtidos são comparados com os rótulos verdadeiros das falhas para verificar a capacidade de generalização.

No contexto de redes neurais empregadas na melhoria do diagnóstico de falhas envolvendo DGA, Farooque, Wani e Khan (2016) propuseram uma ANN MLP (*Artificial Neural Network Multilayer Perceptron*) com única camada escondida para resolver o método gráfico do pentágono de Duval. Já Zakaria, Johari e Musirin (2013), usando essa mesma topologia de rede neural, porém com duas camadas escondidas e três relações de gases da IEC 60599 como dados de entrada, relataram boa generalização da rede no diagnóstico de falhas dos transformadores avaliados.



Malabanan e Nerves (2015) associaram o algoritmo sistema de imunidade artificial (AIS - *Artificial Immunity System*) e a rede neural de base radial (RBFN - *Radial Basis Function Network*) para contornar as limitações de diagnóstico dos métodos convencionais.

Illias et al. (2015) propuseram combinações híbridas entre ANN e enxame de partículas (PSO - *Particle Swarm Optimization*). O resultado foi comparado com várias outras técnicas de classificação e demonstraram taxas elevadas de assertividades na predição de falhas.

Li et al. (2019) desenvolveram uma metodologia híbrida para o diagnóstico de falhas que reúne rede neural de regressão generalizada (GRNN - *Generalized Regression Neural Network*), algoritmo de busca Cuco (CS - *Cuckoo Search*) e teoria dos conjuntos aproximados (RST), em que o RST reduz a alta dimensão dos dados de entrada, preservando os melhores atributos para submissão à GRNN, e o CS otimiza o fator de suavização da GRNN para melhorar o desempenho da rede.

Miranda, Castro e Lima (2012) propuseram um conjunto de redes neurais associativas, cada qual sintonizada para classificar um tipo de falha específica, por meio das amostras de DGA, e para as poucas amostras reais dos dados disponíveis que formavam grupos (*clusters*) de falhas, empregou-se o algoritmo deslocamento médio teórico da informação (ITMS - *Information Theoretic Mean Shift*) para gerar os dados sintéticos que foram usados no treinamento da rede, e para validação do método proposto, empregou-se os próprios dados reais com cerca de 20% de ruído.

Zhao e Zhang (2013) apresentaram uma variação de rede neural que emprega a contra propagação (CPANN - *Counter Propagation Artificial Neural Network*) otimizada por algoritmo genético (GA - *Genetic Algorithm*) que utiliza cinco gases combustíveis e identificação das classes de falhas, similar à IEC 60599.

No caso de redes probabilísticas envolvendo DGA, Niu, Li e Shi (2010) propuseram o emprego de RBFN probabilístico que utiliza monitoramento *online* contínuo dos gases do transformador para diagnóstico de cinco tipos de falhas, como resultado são demonstradas as vantagens desse tipo de monitoramento na identificação de mudanças nas concentrações dos gases e reconhecimento de falhas. Song, Xiu-Ying e Wen-Xu (2011) desenvolveram um trabalho similar com a mesma rede e com três razões de gases, porém aplicados à análise de DGA com uso de banco de amostras existentes. Já Wang, Wang e Wang (2008) propuseram um sistema híbrido probabilístico envolvendo PSO, BPNN e RBFN. O algoritmo PSO foi empregado para determinar a direção ótima global (população de soluções candidatas) e a

BPNN para encontrar a melhor solução, e então aplicou-se o RBFN probabilístico para melhorar a classificação de falhas.

Para Cheng e Yu (2018), a ANN apresenta as seguintes vantagens e desvantagens no diagnóstico de falhas em transformadores: alta capacidade de processamento de informações; robustez e tolerância a falhas; boa generalização; pode mapear sistemas altamente não lineares e incertos; desempenho limitado pelas amostras de treinamento; oscilação no resultado pode ocorrer; e difícil de entender seu processo de análise intermediário e dedução da rede.

### 3.2.3. Máquina de Aprendizado Extremo (ELM)

Com base na ANN *feedforward* clássica de uma camada oculta, e visando acelerar o processo de treinamento supervisionado desse tipo de rede, com manutenção da capacidade de generalização e desempenho, Huang, Zhu e Siew (2006) propuseram um algoritmo denominado de Máquina de Aprendizado Extremo (ELM). A rapidez no treinamento da ELM deve-se ao uso de um processo analítico na determinação dos pesos entre a camada oculta e a sua saída, diferentemente da ANN, que utiliza um processo iterativo. A sistemática de classificação de padrões é similar à ANN, em que as classes são separadas por um hiperplano.

Buscando melhorar a capacidade de generalização e a eficácia na classificação de padrões da ELM convencional, Huang et al. (2012) propuseram mudança nesse algoritmo, introduzindo a regularização e o uso *kernel* (alta dimensionalidade). Quanto à parametrização, não há necessidade de configurar o número de neurônios da camada escondida devido à generalização (dados de entradas levado a uma maior dimensão) da ANN *feedforward*, entretanto, há necessidade de escolher uma função de ativação e ajustar o coeficiente de regularização e o parâmetro *kernel* associado à função de ativação empregada.

No quesito diagnóstico de transformadores com análise de DGA, Ikeshoji et al. (2018) compararam o desempenho e eficácia de alguns dos métodos normatizados, ANN, ELM, e ELM *kernel* na classificação de falhas em transformadores.

Para melhorar a eficácia na classificação das falhas com uso de métodos híbridos, Zhang e Yuan (2015) aplicaram o método PSO para determinar os melhores valores para os coeficientes de regularização e *kernel* gaussiano. Já Huang, Wang e Tian (2018) propuseram o uso do algoritmo otimização do lobo cinzento (GWO – *Gray Wolf Optimization*) para ajuste dos parâmetros *kernel* do ELM híbrido, resultando em melhor taxa de acertos, quando comparado aos algoritmos BPNN (*Backpropagation Neural Network*) e ELM convencional.

Venkatasami e Latha (2016), objetivando contornar as limitações dos métodos normatizados e prover melhor eficácia da classificação de falhas, propuseram a associação entre ELM, utilizando a entalpia (energia da reação) dos gases dissolvidos no óleo do transformador como parâmetros de entrada da rede e comparou com o caso de entrada de dados somente com gases, tendo como resultado uma maior taxa de acertos do método envolvendo entalpia.

Li et al. (2016) apresentaram uma abordagem na classificação de falhas usando a teoria evolucionária diferencial auto adaptável (SaE - *Self-adaptive Evolutionary*) associada ao ELM (denominada SaE-ELM), e o reescalonamento dos valores dos gases usando a função arco tangente, o SaE otimiza os pesos de entrada da rede bem como modifica os parâmetros da camada oculta da ELM e a função arco tangente assume papel similar à normalização.

Como vantagens e desvantagens da ELM, podem ser destacadas (LIN et al., 2015): Rapidez no treinamento, boa capacidade de generalização, entretanto limitado à função Gaussiana.

#### 3.2.4. Máquina de Vetor Suporte (SVM)

Proposto por BOSER, GUYON e VAPNIK (1992), o algoritmo classificador de padrões máquina de vetor suporte, também se utiliza de uma fronteira de decisão para realizar a classificação dos padrões e requer etapas de treinamento e teste. O SVM busca maximizar, com base em determinados vetores (denominados de suporte) localizados no espaço de atributos, a margem da fronteira de decisão entre diferentes classes, para que seja a maior possível (PARIKH; SHAH, 2016) e assim, melhorar o processo de classificação.

De acordo com Lorena e Carvalho (2003), o aprendizado do SVM é fundamentado em princípios da Teoria do Aprendizado Estatístico, por meio do qual se estabelecem condições matemáticas que contribuem para determinar um classificador adequado a partir dos dados de treinamento. Para o caso de classificação linear, o SVM dispõe de dois tipos de margens que definem a fronteira de separação, a rígida que não admite dados de treinamento entre as margens, e a suave que permite que alguns dados permaneçam entre as margens bem como que ocorram alguns erros de classificações. A diferença entre ambos, é que a margem suave lida melhor com ruídos, *outliers* e aplicações com dados não linearmente separáveis (emprego limitado).

Na classificação não linear, o teorema de Cover (HAYKIN, 2009) permitiu ao SVM adotar a estratégia de mapear os dados de entrada para um espaço de alta dimensionalidade e

assim aumentar a probabilidade de separar linearmente as diferentes classes existentes.

Em aplicações envolvendo DGA, Neto, Assunção e Assunção (2009) analisaram e compararam a eficácia entre alguns dos métodos normalizados e os classificadores SVM e ANN, sendo apresentados os melhores resultados para ambos algoritmos de aprendizado de máquina. Dhini et al. (2019) empregaram uma configuração SVM binária multiestágio para classificar as falhas do transformador e concluíram que o uso de algoritmos de aprendizado de máquina é eficaz na classificação de falhas.

Singh e Upadhyay (2017) propuseram a associação do algoritmo *K-means* para reduzir e extrair os dados mais informativos e o SVM para classificar, sendo reportado menor tempo de treinamento e boa generalização. Zhu, Zheng e Wang (2008) apresentaram um método baseado em agrupamento (*clusters*), que usa a distância mínima euclidiana para definir grupos binários de falhas em transformadores (por exemplo, normal e falha, falha térmica e elétrica) e uma árvore de SVM binário para classificar.

Sahri e Yusof (2015) propuseram um modelo híbrido com algoritmo genético (GA) para melhorar o conjunto de dados de treinamento com a eliminação das características redundância e irrelevância para classificação do SVM, obtendo uma melhora nas classificações de falhas de dois conjuntos de diferentes bancos de dados. Kari et al. (2018) usaram a combinação de algoritmos GA com SVM em que, recursos de filtragem são inicialmente, aplicados para classificar 48 tipos de dados de entrada da rede extraídos de sete tipos de gases, posteriormente diferentes métodos de combinações de atributos são usados para avaliar e constituir diversos subconjuntos de dados de dimensões reduzidas, e então o GA é usado para otimizar os parâmetros SVM e escolher o subconjunto que apresenta a melhor classificação.

Shiyong et al. (2019) propuseram um modelo híbrido aplicando o algoritmo *Apriori* que encontra uma correlação entre diferentes atributos num mesmo evento (por exemplo, correlação dos tipos de gases em relação as falhas), e assim estabelecer pesos para uso na função do SVM e consequente classificação.

Wu, Tang e Wei (2014) apresentaram um método que usa o teste Kolmogorov-Smirnov (função de distribuição de probabilidade) para ranquear os três melhores atributos (analisados sobre os métodos de Rogers, Duval, total de gases combustíveis e cinco razões de gases) que representam as falhas normais, térmicas, elétricas de alta e baixa intensidade, e então foram estruturados conjuntos de dados para treinamento e teste do SVM binário multiestágio com seus parâmetros ajustados via PSO. Benmahamed, Teguar e Boubakeur (2017) propuseram o uso do PSO para otimizar os parâmetros de ajuste do SVM e o vetor dos

gases determinados pelo Pentágono de Duval.

Como vantagens e desvantagens do SVM, Bhavsar e Ganatra (2012) destacam: robustez e precisão, insensível ao número de dimensões, menos propenso a *overfitting* que outras técnicas, requer algumas dezenas de amostras para treinamento, encontra a melhor fronteira de separação entre duas classes, possui alto custo computacional no treinamento, e o requisito de memória cresce com o quadrado do número de dados de treinamento.

### 3.2.5. Nãive Bayes (NB)

Trata-se de um classificador probabilístico baseado na hipótese que os atributos envolvidos são independentes entre si, ou seja, o Nãive (ingênuo) Bayes ignora as possíveis dependências entre as correlações dos atributos de entrada, e converte um problema multivariado em um grupo de problemas univariados, ou seja, basicamente o classificador assume que dado um valor alvo de uma determinada instância, a probabilidade de um evento ocorrer é resultante dos produtos das probabilidades para os atributos individualizados (ISLAM et al., 2007).

Benmahamed et al. (2018) avaliaram a eficácia do algoritmo Nãive Bayes na classificação de falhas para diversas composições de gases: valores absolutos, relativos e razões, obtendo o melhor resultado com a composição de gases relativos. Yongqiang et al. (2005) propuseram um método empregando redes bayesianas e três razões de gases envolvendo CO, CO<sub>2</sub>, O<sub>2</sub>, N<sub>2</sub>, C<sub>2</sub>H<sub>2</sub> e H<sub>2</sub>, e comparou sua eficácia com o método convencional de análise de três razões de gases combustíveis.

De acordo Jadhav e Channe (2016), as vantagens de Nãive Bayes (NB) podem ser relacionadas ao bom desempenho, baixo custo computacional e por descartar as características irrelevantes, melhora o desempenho de classificação. Como desvantagens, o algoritmo assume que as amostras são mutuamente independentes e por essa razão alguns conjuntos de dados podem apresentar baixa assertividade na classificação de padrões quando comparado com outros classificadores, e para o algoritmo obter bons resultados de classificação, é necessário o uso de muitas amostras.

### 3.2.6. K Vizinhos mais Próximos (KNN)

Esse algoritmo foi proposto inicialmente por Fukunaga e Narendra (1975). O algoritmo supervisionado KNN tradicional também é considerado um classificador não

paramétrico, simples de usar e implementar, e possui desempenho competitivo em muitos domínios (GOU et al., 2012; HU et al., 2016).

O algoritmo KNN para classificar o padrão de uma nova amostra, considera K elementos de uma determinada classe mais próxima para atribuir o rótulo, por exemplo, seja o caso mais simples de  $K=1$  vizinho, o algoritmo 1NN atribui o rótulo do único elemento mais próximo para amostra, isto é, aquele com a menor distância. Para o número de vizinhos  $K>1$ , o classificador considera o número de vizinhos (sistema de voto) escolhidos para rotular a amostra, ou seja, a classe com maior número de vizinhos mais próximos determina o rótulo, nesse caso, torna-se mais robusto e menos sensível a ruídos e *outliers* (SOUZA; LOTUFO; RITTNER, 2012).

Em uma comparação entre os classificadores KNN e OPF, Souza, Lotufo e Rittner (2012) concluíram que para condição de  $K=1$ , ambos os algoritmos têm comportamento equivalente quando todas as amostras de treinamento são usadas como protótipos, suas fronteiras de decisões para diferentes classes também são similares, porém, com OPF apresentando-se mais suave. E para  $K>1$ , o KNN é menos sensível a ruído presente nos dados por avaliar outras amostras próximas e, assim, teoricamente, deveria apresentar melhor desempenho que o 1NN, entretanto não é o que tem se visto na prática, em alguns casos. Em outro trabalho, Souza, Rittner e Lotufo (2014) comparam os algoritmos KNN com o K-OPF (*K Optimum-Path Forest*) e também confirmam o comportamento semelhante entre os dois.

Para melhorar o desempenho do KNN existem diversas propostas de técnicas, como por exemplo, ponderar a distância (GOU et al., 2012), ponderar os K vizinhos próximos, usar conjunto de dados sintéticos, e usar o classificador *Means* (SARMA et al., 2011). A métrica de distância mais empregada é a euclidiana, entretanto outras métricas como a Minkowsky, Manhattan, Coseno, Chi-square, entre outras, também podem ser usadas (HU et al., 2016).

Em aplicações que envolvem DGA, Benmahamed, Tegar e Boubakeur (2017) apresentaram uma combinação entre KNN e o pentágono de Duval (usa as coordenadas de determinação do centróide que define a falha), obtendo uma melhora na classificação de falhas com relação à aplicação somente do método do pentágono. Em outro trabalho, Benmahamed et al. (2018) avaliaram o desempenho do algoritmo KNN para duas diferentes unidades (microlitro/litro e % relativo) de vetores de gases e razões envolvendo três métodos normalizados, demonstrando uma melhor relação entre KNN e triângulo de Duval.

Shintemirov, Tang e Wu (2009) propuseram o uso de técnica de amostragem aleatória (*bootstrapping*) para aumento de amostras de DGA para quatro classes de rótulos (normal, superaquecimento, descarga de alta e de baixa energia) e programação genética para extração

das melhores características para uso em classificadores de falhas.

Quanto às vantagens e desvantagens do algoritmo KNN, para Jadhav e Channe (2016) podem ser relacionadas: fácil de entendimento e implementação, maior robustez a ruído para  $K > 1$ , funciona bem com muitos rótulos (multiclasse), possui restrições de memória, sensível estrutura local de dados (considera recursos irrelevantes, impactando no aumento do tempo computacional), é considerado algoritmo de aprendizado lento, uma vez que memoriza os dados do conjunto e avalia todas as amostras do espaço de atributos.

### 3.2.7. Floresta de Caminhos Ótimos (OPF)

O OPF (*Optimum-Path Forest*) é um classificador de padrões baseado em grafos e simples de usar, pois não há necessidade de parametrização, pode ser aplicado em problemas de multiclasse, não assume forma ou “separabilidade” das classes e fornece bons resultados na classificação dos conjuntos de amostras avaliados (PAPA, 2008). Exemplos de aplicação podem ser encontrados em diversas áreas, como na saúde (PAPA et al., 2008), reconhecimento de imagens (SILVA, 2016) e de padrão (SOUZA et al., 2012).

No contexto de avaliação e melhoria de desempenho do OPF, Barbosa, Costa e Neto (2018) apresentaram uma técnica que utiliza GA para descartar as amostras irrelevantes de grandes conjuntos de dados e evitar o crescimento do custo computacional, preservando ou até melhorando os resultados. Papa et al. (2012) desenvolveram outra técnica que melhora o desempenho do classificador OPF, por meio da não necessidade de acessar todos os nós na fase de classificação, bem como por descartar dados irrelevantes do conjunto de treinamento com pequena perda na eficácia. Fernandes e Papa (2017) apresentaram uma técnica para melhorar o aprendizado do OPF, usando uma “bolsa” de classificadores e medidas de confiança das amostras. Basicamente, cada amostra usada na fase de treinamento recebe um grau (peso) de confiança, calculado por seu desempenho individual (taxa de reconhecimento) sobre um conjunto de validação denominado de “bolsa” classificadora, posteriormente aplica-se a nova função custo do OPF, que considera o grau de confiança para treinamento da rede.

Visando abordar questões de redução de dimensionalidade dos atributos tornando-o mais simples para análise e melhorar o desempenho do OPF, Rodrigues et al. (2014) propuseram o uso do OPF em conjunto com o algoritmo de otimização do Morcego (BAT - *Bat Algorithm*). No caso, o problema da seleção dos atributos é modelado pela técnica de otimização BAT e empregado a taxa de acertos do OPF sobre o conjunto de validação como função de adequação a ser maximizada. Resultados mostram que são gerados conjuntos de

atributos mais compactos e significativos, e que em alguns dos casos (entre os conjuntos de dados avaliados), até melhoraram a classificação quando comparado com os conjuntos de dados originais contendo todos atributos.

Em direção oposta à redução de dimensionalidade, Afonso, Pereira e Papa (2017) propuseram o uso do *kernel* com o OPF, para encontrar amostras adequadas em espaços dimensionais superiores, em que a hipótese de amostras linearmente separáveis é maior, no caso o *kernel* é incorporado na fase de treinamento e classificação do algoritmo, e seu resultado comparado com o SVM e OPF padrão.

Com objetivo de monitorar o comportamento dos consumidores de energia (residenciais, comerciais e industriais) por meio do consumo caracterizado por perdas não técnicas nas redes de distribuição elétrica (por exemplo, desvio ou ligações clandestinas), Fernandes et al. (2019) propuseram um classificador probabilístico baseado em OPF que determina a probabilidade de ocorrência desse evento, logo, tornando o resultado mais informativo do que apenas classificá-lo com um rótulo.

No contexto de diagnóstico de falhas em transformador via análise de DGA, SOUZA et al. (2012) empregaram o OPF como classificador de falhas, comparando seu desempenho com os classificadores SVM, Bayes e ANN. Ikeshoji et al. (2019) propuseram uma abordagem na análise de falhas em transformadores empregando gráficos de tendência de três tipos de falhas (normal, térmica e elétrica) constituídos a partir da classificação das amostras sucessivas de um conjunto de dados contendo o histórico de registro das análises de DGA.

### **3.3. Métodos de otimização**

Em busca de melhorar a eficácia dos algoritmos inteligentes no processo de classificações de falhas, ou a estimativa de previsão da vida útil restante dos transformadores, tem-se observado uma forte tendência na sua associação com técnicas estatísticas e de otimizações, como pode ser visto no estudo bibliográfico relacionado aos métodos classificadores, envolvendo DGA abordados nesta tese. Os problemas de otimização são caracterizados por contemplarem inúmeras variáveis, nas quais cada uma ou a combinação delas apresenta uma solução ótima.

Nesse contexto, as ferramentas de computação evolucionária e inteligência coletiva têm sido amplamente aplicadas na resolução de problemas nas áreas financeiras, médica, engenharia, entre outras (PARSOPOULOS; VRAHATIS, 2010). Dentre essas técnicas, podem ser relacionadas algumas amplamente empregadas ou em crescimento:



- Proposto por Dorigo, Maniezzo e Coloni em 1996 (DORIGO; BIRATTARI; STUTZLE, 2006), o algoritmo de otimização por colônia de formigas (ACO) artificiais trata-se de uma técnica do ramo da inteligência computacional associada à inteligência coletiva que imita o comportamento social das formigas de uma colônia e a colaboração entre elas, para determinar o melhor caminho na busca pelo alimento, ou seja, a solução ótima para o problema da área ao qual o algoritmo é relacionado;
- Nesta mesma linha de inteligência coletiva, o algoritmo meta-heurístico busca Cuco, desenvolvido por Yang e Deb (2009), imita o comportamento parasita e agressivo adotado por algumas espécies de pássaros Cucos na fase de reprodução como forma de busca da solução ótima para o problema. No caso, três regras foram criadas para o desenvolvimento do algoritmo (SONEJI; SANGHVI, 2012; LI et al., 2018): um ovo por cuco e depositado em ninho aleatório, os ninhos com os melhores ovos passam para próxima geração, e o número de ninhos é fixo com o ovo sujeito à descoberta pelos pássaros hospedeiros com uma probabilidade  $p_a \in [0,1]$ ;
- Já o algoritmo de otimização por enxame de partículas PSO (KENNEDY; EBERHART, 1995) é um método de otimização estocástico associado às técnicas evolucionárias e simula o comportamento social de um bando de pássaros em revoada, com movimento local aleatório, porém globalmente determinado na solução do problema. Diferentemente dos algoritmos que usam operadores genéticos para evoluir a base de dados, o PSO realiza a evolução a partir de cooperação e competição, entre si, por diversas gerações como pode ser visto em Xue, Zhang e Browne (2013); e
- Proposto por Holland em 1970, o algoritmo genético (GA) é um algoritmo evolutivo de busca e otimização de soluções baseado nos princípios evolutivos da seleção natural e genética, isto é, a partir de uma população aleatória inicial de indivíduos, somente aqueles mais aptos, que representam as possíveis soluções do problema dão origem às novas gerações de indivíduos (HALDURAI et al., 2016). Tais princípios são (YANG, 2014): seleção do mais apto, cruzamento com outros indivíduos, e mutação.

Cheng e Yu (2018) realizaram uma investigação sobre diversas técnicas envolvendo abordagens inteligentes e DGA, sendo verificado que a combinação de técnicas compensa algumas das limitações do uso individual dos algoritmos classificadores. A Tabela 3 mostra algumas delas, indicando suas virtudes e uso combinados.

Tabela 3 – Características dos algoritmos evolutivos e de inteligência coletiva

Algoritmos evolutivos e inteligência coletiva	Características em análise de DGA / Uso combinado
Algoritmos Genéticos (GA)	<ul style="list-style-type: none"> <li>– Melhor capacidade de otimização global</li> <li>– SVM, WNN (<i>Wavelet Neural Network</i>) e BPNN</li> </ul>
Otimização por colônia de formigas (ACO)	<ul style="list-style-type: none"> <li>– Maior taxa de exatidão de diagnósticos</li> <li>– SVM, RBFN</li> </ul>
Otimização por exame de partículas (PSO)	<ul style="list-style-type: none"> <li>– Elevada precisão na otimização</li> <li>– SVM, WNN, ANN</li> </ul>
Algoritmo de Busca Cuco (CS) (YANG; DEB, 2014)	<ul style="list-style-type: none"> <li>– Rápida convergência e eficiência, capacidade de busca local e global.</li> <li>– SVM, BPNN</li> </ul>
Algoritmos híbridos (com inteligência de coletiva)	<ul style="list-style-type: none"> <li>– Rápida busca, forte otimização global e evita cair em ótimos locais</li> <li>– SVM, BPNN, ANN, RBFN</li> </ul>

Fonte: Adaptado pelo autor de Cheng e Yu (2018)

### 3.4. Pré-processamento de dados

O emprego de dados para classificação de classes ou padrões determina a necessidade de análise prévia deles, uma vez que pode ter impacto negativo no resultado da avaliação final. Nesse caso, o pré-processamento é uma importante etapa anterior ao processo da classificação de padrões e pode compreender a limpeza, transformação e seleção dos atributos, de modo que resulte na melhora da classificação preditiva dos dados avaliados. De acordo com Van Den Broeck et al. (2005), o tratamento de dados é visto como uma atividade suspeita que está à margem da manipulação de dados, entretanto, é necessário saber lidar com os erros contidos nas fontes e seus efeitos nos resultados dos estudos.

Nesse tópico são abordados resumidamente, estudos que envolvem o pré-processamento de dados, contudo maiores detalhes podem ser obtidos consultando-se as bibliografias relacionadas.

#### 3.4.1. Limpeza de dados (*outliers*)

Em mineração de dados, o estudo sobre detecção de *outlier* é aplicado em diversas áreas de atuação, incluindo (GUPTA et al., 2014): análise de dados em alta dimensão, incertezas nos dados, rede de dados, dados de séries temporais, entre outros. Tallón-Ballesteros e Riquelme (2014) destacam dois aspectos relevantes na importância da preparação (limpeza) dos dados. O primeiro, é que os dados reais são incompletos,

inconsistentes e contêm erros ou *outliers*, e o segundo é que dados de qualidade produzem consequentemente padrões de referência de alta qualidade. Aguinis, Gottfredson e Joo (2013) realizaram uma profunda revisão bibliográfica sobre técnicas de detecção de *outliers* e como resultado, encontraram 14 definições para o termo, 39 técnicas para identificação e 20 maneiras de manipulação. Com base nesse estudo, empregou-se a definição de amostras de dados que estão distantes de outras, por motivo de incertezas, ou seja, valores destoantes que não fazem parte da população de interesse analisada, caracterizada por erro de manipulação, procedimento de amostragem inadequado, erro de preparação dos dados e no registro de informações sobre ocorrências.

Na detecção de *outliers* em DGA, para Khalyasmaa et al. (2016) os valores máximo e mínimo que destoam dos demais dados são considerados, como quantidade de concentrações de gases inadequadas para análise de falhas em transformadores, pois além de diferir acentuadamente dos demais valores de análises obtidos antes ou depois deles, seu uso pode levar a uma decisão incorreta no reconhecimento das falhas, e portanto devem ser excluídos do conjunto de dados. No entanto, não é uma tarefa simples limpar os dados dos gases, visto que há muitas incertezas inseridas e que podem envolver desde o processo de formação desses gases no óleo, até a sua análise final em laboratório (IEC, 2015; SUN; HUANG; HUANG, 2012), bem como também por procedimentos operacionais e anormalidades externas que influenciam a razão gás-óleo, como por exemplo, a regeneração, concentrações de gases residuais após correção de falhas, mistura com o óleo restante do vaso de expansão do comutador de tensão sob carga, entre outros (KHALYASMAA et al., 2016).

Em um estudo realizado pelo *Institute for Interlaboratories Study* para avaliação da competência de 24 laboratórios de 14 países da Europa que realizam a análise de gás dissolvido no óleo de transformadores, Jafar, Visser e Starink (2011) relataram que o valor percentual médio de *outliers* encontrados entre as medições realizadas por esses diferentes laboratórios no período de 2007 e 2010 foi de 5%, e concluíram que uma variação entre 3,5 a 7% eram normais, mesmo com todo o cuidado que envolve o processo, desde a extração até medição. A IEC 60599 (IEC, 2015) relaciona incerteza nas medições dos gases tipicamente entre  $\pm 15\%$ , podendo chegar a  $\pm 30\%$  para baixos valores de concentrações (<5 vezes a medição analítica). Nesse contexto de tratamento de dados de DGA, diversos trabalhos envolvendo o uso de algoritmo de máquinas e classificação de falhas em transformadores foram desenvolvidos.

Qin et al. (2019) propuseram um método de detecção de *outlier* na análise de gases em monitoramento contínuo (o que reduz a possibilidade de maior introdução de incertezas)

empregando a transformação *Wavelet* dos sinais para identificação de ocorrências de anomalias (com uso de uma janela deslizante para leitura de intervalos) operando em conjunto com o algoritmo fator ponderado de *outlier* locais. O sinal avaliado, resultante da soma dos gases CH<sub>4</sub>, C<sub>2</sub>H<sub>2</sub>, C<sub>2</sub>H<sub>4</sub> e C<sub>2</sub>H<sub>6</sub>, permitiu identificar as ocorrências de anomalias durante o período de aquisição dos dados, sendo possível observar principalmente em pontos extremos dos sinais que destoam-se dos demais, por serem esporádicos e em seus valores absolutos, características comuns aos *outliers*.

Ma, Ekanayake e Saha (2012) propuseram a eliminação de *outliers* nas amostras de gases dos transformadores, empregando os algoritmos de agrupamento FCM e KFCM junto ao classificador máquina de vetor suporte FSVM. No caso, o FCM e KFCM atribuem diferentes pesos para as amostras, em função das suas distâncias com relação aos centros dos grupos a que pertencem, e posteriormente, submete-se ao FSVM para realizar a classificação. Já Li et al. (2017) empregaram o algoritmo de agrupamento *K-medoids* para eliminar as amostras do conjunto de dados que não contribuíam para classificação das falhas e assim melhoraram a eficiência e assertividade do algoritmo SVM.

Já Subroto et al. (2017) empregaram o método (estatístico) Thompson Tau para filtrar *outliers* existentes nos gases dissolvidos e posteriormente combinaram a técnica FCM com os métodos normalizados, para criar um sistema de pesos e assim melhorar a assertividade da classificação de falhas, entretanto no caso de conjuntos de dados com grande dispersão, esse método pode promover elevada redução das amostras contidas nele e assim reduzir a confiança no resultado.

A norma IEC 60599 (2015) emprega o percentual de 90% em seus valores de referências típicos para condição de normal e de crescimento de gases, e a IEEE C57.104 (IEEE, 2019), a distribuição normal entre 90 e 95% na determinação dos valores limites dos gases.

#### 3.4.2. Construção de atributos

O espaço de atributos é definido e constituído segundo as características contidas no conjunto de dados, e sua representação é um dos fatores de elevado impacto no resultado e desempenho final do algoritmo classificador, principalmente naqueles que possuem poucas parametrizações a serem feitas. Na mineração de dados, a transformação dos dados pode compreender atividades de agregação, discretização de variáveis contínuas, construção de dados, entre outros, cujo objetivo final é descobrir informações de melhor qualidade para

representação dos espaços de atributos dos classificadores. De acordo com Otero et al. (2003), construir atributos é produzir novos dados a partir dos originais, para que as regularidades nos dados sejam mais facilmente detectadas e melhorem a precisão preditiva do classificador. Quanto aos tipos de métodos de construções, eles podem ser orientados a hipóteses, orientados a dados, orientados ao conhecimento, e orientados à multiestratégia (ARCISZEWSKI; MICHALSKI; WNEK, 1995). O primeiro método constrói novos atributos a partir de hipóteses criadas anteriormente, o segundo, a partir da relação entre os atributos, o terceiro tem seu domínio construído a partir do conhecimento do especialista, e no último, duas ou mais estratégias são usadas para construir os atributos.

Em se tratando de gases dissolvidos no óleo, as falhas de origem térmica e elétrica podem produzir até sete tipos de gases como visto no Capítulo 2. A partir deles, é possível observar a construção dos dados que estão presentes nos métodos convencionais por meio do uso do percentual relativo e razões (como por exemplo, %H<sub>2</sub>, %CH<sub>4</sub>, C<sub>2</sub>H<sub>2</sub>/C<sub>2</sub>H<sub>4</sub> e CH<sub>4</sub>/H<sub>2</sub>, respectivamente) entre as concentrações dos gases, com o objetivo de fornecer melhores resultados nos diagnósticos de falhas.

Em trabalhos envolvendo aprendizado de máquinas, Samirmi, Tang, e Wu (2013) avaliaram a influência de diferentes combinações dos gases (construídos 21 atributos a partir de seis gases combustíveis) nos diagnósticos de falhas, empregando métodos estatísticos (teste t, Kolmogorov-Smirnov e teste divergência Kullback Leibler), para a escolha dos vetores de gases associado ao algoritmo KNN usado como ferramenta de validação na melhora do desempenho.

Já Amora et al. (2012) propuseram a expansão dos atributos de entradas também em 21 atributos a partir dos sete gases e o uso do algoritmo da árvore de decisão, a fim de constituir regras de classificação para os limites dos atributos selecionados.

Malik e Mishra (2016) utilizaram o algoritmo componentes principais (PCA – *Principal Component Analysis*) para redução e seleção dos melhores atributos entre 21 variações, construídos a partir de sete combinações entre gases, como vetor de entrada do algoritmo ELM, e o resultado obtido foi comparado aos vetores de atributos definidos pelos métodos convencionais.

### 3.4.3. Seleção de atributos

De acordo com Raudys e Jain (1991), a quantidade finita de amostras requer cuidado e atenção na seleção dos vários parâmetros de avaliação, pois podem afetar o desempenho dos

algoritmos classificadores de padrões. A seleção de atributos ou de recursos, trata-se da escolha de um subconjunto de atributos relevantes para uso em algoritmos de aprendizado de máquina (YU; LIU, 2003).

De acordo com (LADHA; DEEPA, 2011), a seleção de atributos é um tópico importante na mineração de dados, especialmente na análise de conjuntos de atributos de alta dimensão, pois reúne as melhores características que representam o problema em um subconjunto de dados e descarta aqueles menos importantes, reduzindo então a dimensão do conjunto original. Tais atributos de natureza nominal, contínua ou discreta, podem ser caracterizados como relevantes, irrelevantes e redundantes.

A seleção de atributos difere da redução de dimensão, pois a primeira apenas inclui ou exclui os atributos sem modificá-los, já a segunda usa a combinação dos atributos para gerar novos atributos de análise (como por exemplo, a transformação ortogonal da análise de componentes principais). Nesse contexto, a seleção de atributos pode contribuir para melhorar o desempenho do algoritmo de aprendizado de máquina, facilitar a compreensão e interpretação dos dados, reduzir o espaço de armazenamento de dados e o custo computacional (SÁNCHEZ-MAROÑO; ALONSO-BETANZOS; CASTILLO, 2005).

Existem três categorias de algoritmos de seleção de atributos, o modelo *wrapper*, *filter* e *embedded* (LIU; ZHOU; LIU, 2019; LADHA; DEEPA, 2011) a saber:

- A seleção de atributos *wrapper* emprega um algoritmo de aprendizado de máquina pré-determinado e usa seu desempenho para definir os atributos escolhidos. A cada novo subconjunto de atributos avaliado, o modelo *wrapper* precisa aprender uma hipótese ou realizar uma classificação e, nesse caso, tende a encontrar os atributos mais adequados ao algoritmo de aprendizado de máquina usado, conduzindo a um desempenho superior no aprendizado, entretanto com elevado custo computacional se comparado ao modelo *filter* (YU; LIU, 2003);
- A seleção de atributos *filter* baseia-se na análise das características gerais dos dados de treinamento sem envolvimento de nenhum algoritmo de aprendizado (YU; LIU, 2003). Nesse método, geralmente empregam-se técnicas estatísticas para avaliar e pontuar o relacionamento entre cada variável de entrada e a de saída, tais pontuações (maiores valores) são usadas para determinar os atributos de entrada para serem usadas no modelo, como por exemplo, em um classificador (KUHN; JOHNSON, 2013); e
- A seleção de atributos *embedded*, é considerada como uma combinação entre dois métodos anteriores e, também, é associada ao uso de algoritmos classificadores. Contudo essa ligação é maior no *embedded* do que no *wrapper*. Este método usa

algoritmos que contêm intrinsecamente sua própria capacidade interna de selecionar recursos (LIU; ZHOU; LIU, 2019), ou seja, aprendem quais recursos contribuem para melhorar o desempenho do modelo, sendo bastante comum o uso da regularização (ou penalização).

Na análise de DGA, diversos trabalhos (KARI et al., 2018; LI et al., 2016; SAHRI; YUSOF, 2015; SAMIRMI; TANG; WU, 2013) foram desenvolvidos com o objetivo de determinar o melhor subconjunto de atributos, para que resultassem em melhor desempenho dos algoritmos, seja por meio da redução do tempo computacional, da melhora na taxa de acerto da classificação dos padrões, e outros.

#### 3.4.4. Dados sintéticos de DGA

A questão associada à quantidade de dados empregadas em algoritmos de reconhecimento de padrões relaciona-se com o grau de confiança dos resultados, ou seja, uma maior quantidade proporciona uma maior confiança sobre os resultados de classificação (JAIN; CHANDRASEKARAN, 1982). Entretanto, como nem sempre esses dados podem estarem disponíveis, ou serem de fácil obtenção, gerar dados artificiais pode ser a solução para contornar esse problema de poucas amostras. Isso tem sido bastante empregado no reconhecimento de imagens e com resultados promissores (LI et al., 2018; EKBATANI; PUJOL; SEGUI, 2017; MAYER et al., 2018).

Em diagnóstico de transformadores, empregando a análise de DGA, Miranda, Castro e Lima (2012) usaram o algoritmo de ITMS para aumentar a quantidade de dados para treinamento da rede, e com esse mesmo fim de geração de dados. Shintemirov, Tang e Wu (2009) empregaram a técnica da amostragem aleatória (*bootstrapping*) para aumentar o número de dados.

De acordo com Viana (2014), a geração de dados sintéticos é uma técnica que permite criar dados a partir dos dados originais, por intermédio do uso de modelos que reproduzam as características mais próximas deles. Logo, explorar as informações de determinadas características que influenciam os dados reais disponíveis, pode ser uma estratégia para geração dados sintéticos. Nesse contexto, nas aquisições das medidas de DGA dos transformadores, existem incertezas que são introduzidas por equipamentos (exatidão e precisão nas medições, controle de calibração) e procedimentos envolvidos na fase de extração dessa informação (cuidados no manuseio, transporte e preservação das amostras), conforme pode ser visto em Duval e Dukarm (2005), Lelekakis et al. (2011) e NBR7070

(ABNT, 1981), e que podem ser exploradas no caso de necessidade de aumento de dados. Para a IEC 60599 (IEC, 2015), é possível recalcular os valores dos gases, empregando os valores percentuais de incertezas considerados na norma e os limites de detecções analíticos contidos na IEC 60567 (IEC, 2005) conforme Tabela 4.

Tabela 4 – Valores limites de detecção de  $S_{LIM}$  em DGA

Gases (Concentrações a 20°C)	Teste de aceitação ( $\mu\text{l/l}$ )	Teste em serviço ( $\mu\text{l/l}$ )
Hidrogênio	2	5
Hidrocarbonetos	0,1	1
Monóxido de Carbono	5	25
Dióxido de Carbono	10	25
Oxigênio	500	500
Hidrogênio	2000	2000

Fonte: Adaptado pelo autor de IEC (2005)

Onde  $S_{LIM}$  é a variável do limite de detecção, para valores de medidas dos gases acima  $10 \times S_{LIM}$  a precisão típica é de 5% sobre os valores absolutos e de 10% na razão entre esses. Para valores abaixo de  $10 \times S_{LIM}$ , a precisão típica decresce rapidamente até 20% em  $5 \times S_{LIM}$  e até 40% na razão dos gases IEC 60599 (IEC, 2007), sendo esse percentual aumentado na versão mais recente de 2015 ( $\pm 15\%$  em  $10 \times S_{LIM}$  e  $\pm 30\%$  em  $5 \times S_{LIM}$ ) o que torna esses valores mais conservadores.

### 3.5. Considerações gerais do capítulo

Embora sejam empregados por empresas especializadas em manutenção de transformadores e em concessionárias de energia para diagnósticos de falhas, alguns dos métodos normatizados (Rogers, Duval, Gás Chave, Doernenburg, entre outros) frequentemente são associados às questões de que nem todos podem prever todas as situações de falhas ou combinações de gases e, portanto, levam a resultados inconsistentes ou ambíguos. Nesse contexto, para reduzir as lacunas das falhas e relações de gases não previstas, várias metodologias e técnicas com bases nos métodos convencionais de análise (como por exemplo, pentágono e heptágono), método hierárquico (*fuzzy*) e de aprendizado de máquina (SVM, ANN, OPF e outros) têm sido propostas.

Em se tratando dos métodos de aprendizado de máquina, eles estão sendo associados a recursos: de otimização para escolha das características mais significativas das amostras ou para ajuste de parâmetros de regularização (ou penalização) e função da ferramenta de



classificação; uso de algoritmos de filtragem de ruídos e *outliers* contidos nos dados; de geração de dados sintéticos em caso de dados escassos; de aumento da dimensão do espaço de atributos e seleção de atributos, e muitos outros, com o objetivo de melhorar a taxa de acertos na classificação de falhas. Entretanto, vale lembrar que os métodos normatizados continuam sendo empregados como referências nas análises de falhas para a emissão de relatórios técnicos sobre a saúde do equipamento, relegando aos outros métodos analíticos ou de aprendizado de máquinas, um papel auxiliar nesse cenário de aplicação real.

Dentre os métodos ou técnicas de classificações de falhas envolvendo DGA nos transformadores, as abordagens que envolvem apenas as rotulações ou classificações desses eventos podem determinar diferentes procedimentos nas fases de diagnóstico e decisão de inspeção e manutenção, em detrimento daqueles aplicáveis aos métodos probabilísticos ou que determinam as tendências de falhas. As abordagens que envolvem técnicas probabilísticas e de análise de tendências de falhas podem ser consideradas mais informativas, uma vez que a informação disponibilizada pode quantificar a possibilidade de ocorrência do evento ou sinalizar uma tendência de desempenho, com base nas informações antecedentes e atuais dos gases. Essas abordagens permitem adotar estratégias mais apropriadas para acompanhar e tomar medidas eficazes para a manutenção do equipamento, e conseqüentemente, aumentar a segurança do sistema elétrico.

A aplicação dos métodos em DGA apenas como ferramentas classificadoras de falhas pode ser feita sobre qualquer amostra, cuja composição dos tipos dos gases seja compatível com o método em questão, porém devido aos erros de classificações e resultados ambíguos que apresentam, a confiança com base nesse tipo de abordagem de apenas classificar as amostras sem conhecimento do contexto geral que envolve o equipamento, pode tornar-se frágil. Logo, como meio de melhorar essa confiança no diagnóstico de falhas, as concessionárias normalmente fazem o acompanhamento da evolução dos gases dissolvido no óleo, isto é, em caso de detecção do crescimento anormal da quantidade dos gases, o especialista da empresa aplica diversas ferramentas classificadoras normatizadas e avalia todo o contexto (histórico de manutenção e operação, ocorrências de eventos externos, problemas crônicos do modelo, e outros) para decidir a realização e o momento de inspeção do equipamento.

A questão acerca dessa abordagem de evolução dos gases, no caso de constatação de falha após abertura do equipamento, normalmente associa-se à falha existente aos valores de concentrações dos gases encontrados na última medida antes da abertura, entretanto seu início pode anteceder a esse momento, uma vez que depende do tempo em que a ação de inspeção

foi realizada. A esse contexto podem ser relacionados as seguintes situações: primeiramente como os classificadores não observam os dados dos gases e falhas antecedentes, isso limita o conhecimento sobre o comportamento e saúde do equipamento ao longo do tempo de serviço, impedindo que uma ação de manutenção fosse tomada antecipadamente, para evitar possíveis consequências mais graves; a outra refere-se aos dados que constituem os conjuntos de amostras de gases e sua relação com as falhas serem caracterizados apenas por valores que confirmam o evento, o que reduz (ou até omitem) as informações antecedentes que poderiam ser usadas para melhorar os métodos ou técnicas de diagnóstico de falhas.

Tendo em vista o objetivo de oferecer um método mais informativo sobre o diagnóstico de falhas em transformadores com base no uso de classificadores de padrões, propôs-se então o desenvolvimento de uma abordagem metodológica que reúne as funções de um classificador que permite acompanhar a evolução e tendência de falhas.

## CAPÍTULO 4 – PROPOSTA DE ABORDAGEM NO DIAGNÓSTICO DE FALHAS EM TRANSFORMADORES

A Figura 7 mostra um diagrama em blocos simplificado de um exemplo de análise de criticidade, e a partir do cruzamento e análise dessas informações, toma-se a decisão sobre o que fazer com o transformador. Ao histórico de operação, são associadas informações do nível de carregamento, temperatura, tensão, entre outros. À manutenção, relacionam-se registros anteriores de análise físico-química e DGA, manutenções realizadas, presença de ruídos fora do normal, registro de ocorrência com grupos de transformadores de mesmo modelo, entre outros. Os eventos referem-se a ocorrências recentes na subestação (por exemplo, descargas atmosféricas, atuação do relé de Buchholz). Aos problemas crônicos são associados problemas de projeto do equipamento e que podem ser associados ao grupo de mesmo modelo.

Figura 7 – Diagrama simplificado de análise de criticidade



Fonte: Elaborado pelo autor

Nesta tese, a abordagem proposta de diagnóstico de falhas é usada para identificar a tendência, bem como classificar as falhas incipientes. O uso dessa informação e o conhecimento sobre o histórico de ocorrências, operação e serviços por parte do gestor, permite a ele identificar as criticidades envolvidas com o equipamento e então tomar decisões oportunas de manter o equipamento em serviço, ou programar a retirada para manutenção.

Diferentemente de métodos probabilísticos, que podem quantificar a ocorrência dos tipos de falhas para cada amostra tornando a análise mais informativa, ou método de análise de regressão para estimativa de previsão de falha, a maioria dos métodos de classificações determina apenas uma classe de falha para cada amostra analisada. Por esse motivo, há

necessidade de desenvolver novos recursos, que busquem melhorar cada vez mais as taxas de acertos, e assim contornar as características que apenas rotulam as classes individualmente. Neste tópico propõe-se uma metodologia de análise de DGA mais informativa, que usa o espaço de características dos algoritmos classificadores KNN e OPF na construção de gráficos que refletem o histórico do comportamento do estado do transformador, e a cada amostra de DGA avaliada, em relação às diferentes classes de falhas existentes.

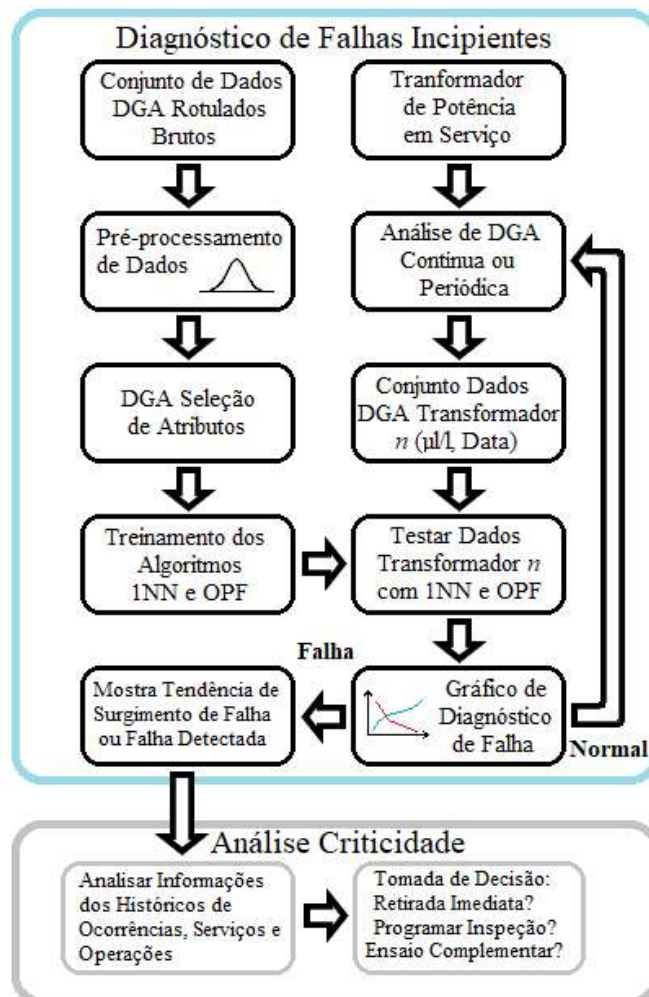
A escolha dos algoritmos KNN e OPF, em relação aos outros algoritmos abordados nesta tese, deve-se às características de simplicidade no uso, adequam-se à solução de problemas multiclasse, não necessitam de parametrizações complexas, e principalmente porque ambos utilizam o parâmetro distância entre as amostras presentes no espaço de atributos como um dos seus recursos no processo de classificação e, portanto estabelece uma relação de proximidade física com a falha. Contudo, uma vez que esses algoritmos buscam extrair informações dos fenômenos térmicos e elétricos que envolvem o conjunto de dados dos gases, menos parametrizações (relacionadas às variáveis de entrada do algoritmo classificador), em contrapartida, significa a necessidade de que o conjunto dos dados dos gases de entrada seja preparado para que alcance um melhor resultado classificatório, e isso pode ser feito com o emprego de técnicas de pré-processamento de dados (por exemplo, eliminação de *outliers*, seleção de atributos, e outros) que melhoram a qualidade informativa dos dados analisados.

No processo de classificação do KNN e OPF, a relação de distância entre diferentes classes avaliadas, associado à heurística do processo de classificação de novas amostras, recebe a denominação “custo” e se refere à intensidade da conectividade entre as classes. Um menor valor de custo significa que cada nova amostra está mais fortemente conexa a uma determinada amostra (que compõe o espaço de atributos) com sua respectiva classe. Essa característica de determinação do custo permite, então, relacionar um conjunto de amostras extraídas periodicamente, como por exemplo, as análises de DGA obtidas por meio da rotina de serviços e manutenção (preditiva e preventiva) da concessionária de energia, e assim explorar e conhecer o comportamento das amostras no espaço de atributos dos algoritmos e no tempo. É sobre essa perspectiva que se propõe o desenvolvimento dessa abordagem.

O processo de análise de falhas nesta tese compreende inicialmente aplicar o pré-processamento de dados do conjunto de treinamento de DGA rotulado para reduzir a presença de possíveis *outliers*, posteriormente faz-se a escolha dos melhores atributos por intermédio do aumento da dimensão de cada amostra dos gases e então, a seleção com uso de algoritmos de otimização. Para isso, foram avaliadas as técnicas: distribuição normal de probabilidade,

algoritmos clusterizadores (*K-means* e *K-medoids*) e critério de seleção com valores típicos da IEC 60599 (2015) na limpeza dos dados; e para a escolha dos melhores atributos, empregou-se os algoritmos CS, PSO, GA e ACO associado à técnica de seleção de atributos binária, proposta por Salesi e Cosma (2017). Sobre o conjunto de dados resultante, realizou-se o treinamento dos algoritmos OPF e KNN, e então a classificação das novas amostras de gases (*online* ou periódica) para cada transformador avaliado. Também foi analisada a geração de dados artificiais com base nas incertezas das medições de DGA, para o caso da necessidade de aumento do número de amostras em conjuntos com baixa quantidade de dados. A Figura 8 mostra o diagrama em blocos do processo de análise e diagnóstico de falhas da metodologia proposta.

Figura 8 – Diagrama em blocos do método de análise e diagnóstico de falhas proposto



Fonte: Elaborado pelo autor

Após classificar e determinar os custos das novas amostras submetidas à avaliação são gerados os gráficos de tendências de falhas para os diagnósticos, e em caso de evolução (convergência) em direção à falha ou falha diagnosticada, é então realizada a análise de

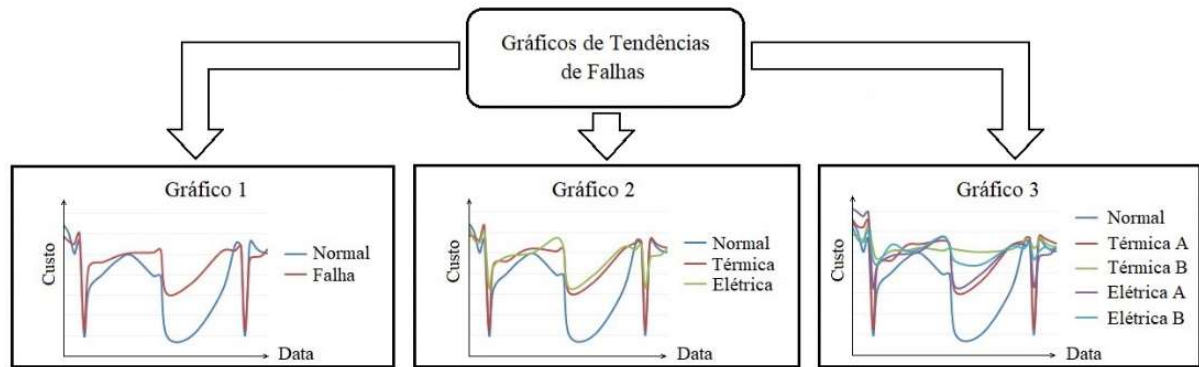
criticidade com base em outros parâmetros técnicos para tomar a decisão quanto a realizar outros testes específicos complementares, avaliar o adiamento ou a retirada imediata do transformador do sistema elétrico, para inspeção e manutenção. À medida que as concentrações de gases se alteram, naturalmente as curvas de caracterização das falhas também se modificam, uma vez que as amostras analisadas movimentam-se no espaço de atributos, buscando o seu melhor posicionamento. Se o conjunto de amostras contém informações de todo o período de serviço do transformador, os gráficos irão representar essa situação em seu eixo gráfico temporal.

Os gráficos gerados para análise são disponibilizados considerando três modos de visualização de falhas (Figura 9), ou seja, envolvendo os estados: normal e falha; normal, falha térmica e falha elétrica; normal e falhas (térmica e elétrica) específicas. Além de facilitar a visualização da evolução das falhas, com essa abordagem, as taxas de acertos tendem a ser diferentes e assim contribuem com o diagnóstico e análise, ou seja, um menor número de classes de falhas tende a apresentar uma melhor classificação, se comparado a um número maior de classes.

Além disso, esse recurso permite estimar os possíveis tipos de falhas, não restrito a uma classe, mas observando uma sequência de falhas determinadas pelos custos apresentados do menor (falha principal diagnosticada) para o maior (falhas secundária, terciária, e outras), o que significa avaliar as proximidades dos diferentes tipos de falhas que podem estar envolvidos. Essa situação suaviza os possíveis erros de classificações a que os algoritmos estão sujeitos, quando são feitas análises que identificam apenas falha única (por exemplo, falha térmica A, ou falha elétrica B) o que leva também à interpretação única, logo, conhecer as proximidades com outras falhas, permite avaliar a possibilidade de ocorrência de outras, que podem estar envolvidas.

A validação da metodologia foi realizada analisando os resultados de relatórios de inspeções de falhas encontradas em grandes transformadores da transmissão de energia elétrica e avaliação do especialista da concessionária. O gráfico de classes de falhas térmicas e elétricas específicas pode acomodar um número maior de classes, uma vez que dependerá somente da quantidade de rótulos falhas disponíveis.

Figura 9 – Variações dos gráficos de falhas



Fonte: Elaborado pelo autor

#### 4.1. Limpeza do conjunto de dados de gases

A retirada de *outliers* do conjunto de dados é uma etapa do pré-processamento empregada para reduzir erros de classificações induzidos por essas amostras, principalmente quando os algoritmos classificadores são mais sensíveis à presença deles. Visto que os dados de DGA obtidos aleatoriamente nas bases e fontes de dados consultadas não acompanham registros e informações sobre as particularidades das falhas associadas, principalmente em ocorrências que envolvam concentrações atípicas de gases para permitir serem excluídos do conjunto de dados, como por exemplo, valores de concentrações muito baixas que podem ser interpretadas como condição normal, torna-se necessário realizar a limpeza do conjunto.

As técnicas empregadas na detecção de *outliers* foram a condução de análise com e sem *outlier*, em que se compara a taxa de acertos na classificação das amostras para ambas as situações, o agrupamento  $K_C$  (*K-means* e *K-medoid*), considerando os maiores grupos de modo a totalizar 95% dos dados, a análise de distribuição normal (percentil 95) e o percentual limite de 95% (valor típico) adotado sobre a quantidade de dados normais, de acordo com a IEC 60599 (2015). Todos os dados foram manipulados com base na análise de fenômeno único de interesse (taxa de acertos empregando algoritmos classificadores), e ajustes de valores adequados dentro de um intervalo confiável de valores, com a eliminação do resto das amostras. Os valores limites dos gases para corte de possíveis *outliers* existentes no conjunto de falhas, foram determinados sobre os dados de amostras de transformadores considerados operando normalmente no sistema elétrico.

Os algoritmos não supervisionados *K-means* e *K-medoids* são clusterizadores que buscam encontrar semelhanças em conjuntos de dados numéricos e separá-los em  $K_C$  grupos (*clusters*) distintos. Diferenciam-se na heurística de classificação, em que o primeiro faz o agrupamento das amostras com base na minimização do custo, por meio da mudança média

dos centroides dos grupos (HUANG, 1998), e o segundo usa o objeto (*medoid*) contido na base de dados de cada grupo, como sendo o ponto central do agrupamento (BUDIJI; LEISCH, 2019). Para a definição do valor adequado de grupos  $K_C$  nesses algoritmos, foi utilizado o método *Elbow* (SYAKUR et al., 2018) que usa o percentual da variância dos dados em função do número de grupos  $K_C$ , ou seja, o algoritmo testa diferentes valores de  $K_C$  (a partir de  $K_C=2$  e com passo unitário) e a partir do momento em que passa a não haver um aumento expressivo da variância, diz-se ter encontrado o valor do  $K_C$  ideal.

#### 4.2. Construção e seleção de atributos

Na etapa de aumento ou construção de atributos dos dados, a partir das concentrações dos cinco gases de interesse, são gerados novos atributos com base no valor relativo do total de gases e nas razões dos gases utilizados nos métodos normalizados de diagnóstico de falhas da IEEE C57.104 (2019) e IEC 60599 (2015), resultando no aumento do número de atributos para cada vetor amostra do conjunto de dados que foram normalizados pela média e desvio padrão.

A seleção de atributos empregada (*wrapper*) para a escolha do melhor conjunto de dados, foi tratada como um problema binário discreto e modelada como um problema de rede booleana  $m$ -dimensional, conforme proposto por Salesi e Cosma (2017), ou seja, diferentemente de alguns algoritmos de otimização que geram soluções em um domínio contínuo, as soluções da seleção de atributos são apresentadas e atualizadas nos vértices de um hipercubo. Em outras palavras, os vetores binários que representam essa solução serão limitados no intervalo  $[0,1]$ , isto é, as novas soluções serão geradas e atualizadas, somente se estiverem dentro dos limites superior ( $L_{sup}=1$ ) e inferior ( $L_{inf}=0$ ). Assim, para constituir o valor de adequação de uma solução numérica gerada entre  $(0,1)$ , ela é convertida em um vetor de binários equivalente, usando as equações (1) e (2) e, então, o vetor binário gerado é submetido à função de adequação responsável pela seleção dos atributos no espaço contínuo a serem avaliados. O valor igual a 1 determina o gás que fará parte do subconjunto a ser avaliado, de outro modo para valor 0, o gás não será avaliado.

$$y_i^j(t) = \begin{cases} 1, & \text{se } x_i^j(t) \geq \sigma_{boolean} \\ 0, & \text{se } x_i^j(t) < \sigma_{boolean} \end{cases} \quad (1)$$



$$y_i(t) = \bigcup_{i=1}^{\mathcal{N}} y_i^j(t) \quad (2)$$

Onde  $\sigma_{boolean} \in (0,1)$  é o limite booleano, nos algoritmos de inteligência coletiva  $x_i^j(t)$  representa as soluções para  $j$ -ésima população (por exemplo, de partículas no PSO e ovos no CS) e  $i$ -ésima variável de atualização (por exemplo, posição, ninhos, feromônio, respectivamente) no instante  $t_p$  (sendo  $t_p=0$  a solução inicial),  $y_i^j(t)$  corresponde ao vetor binário das soluções de  $x_i^j(t)$ ,  $Y_i(t)$  é o vetor binário equivalente no mesmo instante e  $\mathcal{N}$ , o número de máximo de atualizações. A cada iteração, novas soluções no espaço contínuo são geradas e verificadas, se houver violações dos limites  $L_{sup}$  e  $L_{inf}$ , essas são corrigidas para dentro dos limites, desse modo, as soluções originais permanecem no espaço contínuo desses algoritmos e são convertidas em binários sempre que necessário.

Os algoritmos de otimização associados à técnica de seleção de atributos empregada foram enxame de partículas (PSO), algoritmo genético (GA), busca cuco (CS) e colônia de formigas (ACO) e a função *fitness* adotada foi a maior taxa de acertos determinada pelo algoritmo INN, tendo em vista o melhor resultado apresentado.

### 4.3. Reconhecimento de padrões com KNN

A variação do classificador KNN de uso mais comum é por voto majoritário (WLODARCZAK, 2019), ou seja, o rótulo de toda nova amostra submetida à classificação será determinado pelo número de  $K$  amostras de um mesmo padrão mais próximo. Um dos parâmetros usados para classificação é a distância entre as amostras (arco entre os nós) no espaço de atributos, e nesse caso, a métrica euclidiana é de uso mais comum. Para a condição  $K=1$ , a decisão de classificar determinada amostra é simplificada para o único padrão mais próximo, entretanto, para  $K>1$ , o algoritmo é menos sensível a ruídos e *outliers* nos dados, conforme visto no Capítulo 3.

Funcionalmente, considere  $Z_1 = \{(x_i, y_i)\}_{i=1}^n$  o conjunto de treinamento do algoritmo, onde  $n$  é o número de amostras,  $x_i \in \mathfrak{R}^m$  é o vetor de treinamento no espaço  $m$ -dimensional e  $y_i$  os rótulos de suas classes correspondentes. Seja uma amostra  $x_{desc}$  no espaço de atributos para determinação de sua classe (ou rótulo)  $y_{desc}$  desconhecida, inicialmente um conjunto alvo de  $K$  vizinhos com rótulos semelhantes, durante a análise de  $x_{desc}$  é identificado. Representado por  $Z_{desc} = \{(x_i^{nn}, y_i^{nn})\}_{i=1}^K$ , o conjunto é organizado na ordem crescente em

termos de valores de distância euclidiana (equação 3), ou seja, distância entre as amostras  $x_{desc}$  e  $x_i^{nn}$ . Finalmente, a previsão do rótulo atribuído para a amostra  $x_{desc}$  é determinado pela maioria dos K vizinhos mais próximos (equação 4) (GOU et al., 2012).

$$d(x_{desc}, x_i^{nn}) = \sqrt{(x_{desc} - x_i^{nn})^T (x_{desc} - x_i^{nn})} \quad (3)$$

$$y_{desc} = \arg \max_y \sum_{(x_i^{nn}, y_i^{nn}) \in Z_{desc}} \delta(y = y_i^{nn}), \quad i = 1, \dots, K \quad (4)$$

Onde  $y$  é o rótulo da classe,  $y_i^{nn}$  é o rótulo da classe para  $i$ -ésimo vizinho mais próximo entre seus K vizinhos, e a função delta Dirac  $\delta(y = y_i^{nn})$  assume o valor 1 se  $y = y_i^{nn}$ , ou 0, caso contrário. Nessa tese, a distância entre a única amostra mais próxima de outra amostra no espaço de atributos do 1NN ( $K=1$ ), será denominada de “custo” no processo de reconhecimento e classificação de padrões. Essa denominação comum ao algoritmo OPF foi empregada para simplificar a explicação e entendimento dos gráficos de tendência de diagnósticos do método proposto.

#### 4.4. Reconhecimento de padrões com OPF

No algoritmo OPF proposto por Papa (2008), o problema do reconhecimento de padrão é tratado como particionamento de um grafo em um dado espaço de atributos, em vários subgrafos (árvores de caminhos ótimos) que constituem diferentes regiões de influências com base nos padrões existentes. Esse particionamento ocorre por meio de um processo de competição entre amostras chaves (nós protótipos) que são extraídas de um conjunto de dados rotulados de interesse. Os nós protótipos são considerados os melhores representantes dos padrões existentes e formam as raízes das árvores de caminhos ótimos e, ao conjunto dessas árvores, dá-se o nome de floresta de caminhos ótimos.

Funcionalmente, seja  $Z = Z_1 \cup Z_2$  um conjunto de dados  $\theta$  rotulados formados por um conjunto de treino  $Z_1$  e teste  $Z_2$  (onde  $Z_1 \cap Z_2 = \emptyset$ ). Cada amostra do conjunto  $Z$  refere-se a um vetor de atributos  $\vec{v}(s)$  do conjunto  $\vec{V} \in \mathfrak{R}^m$  (onde  $m$  é o número de atributos) e representa um nó do grafo completo desse conjunto. Os arcos que ligam todos os nós do grafo são ponderados por uma dada função distância  $d(s, t)$ , como por exemplo a euclidiana, *chi-square*, entre outras. Considere  $S$  o conjunto de amostras protótipos formado pelo algoritmo

da árvore de espalhamento mínimo (MST – *Minimum Spanning Tree*) na fase de treino com  $Z_1$  e que contenha todos os rótulos dos padrões existentes, o OPF cria um classificador com particionamento discreto ótimo no espaço de atributos, capaz de atribuir o rótulo  $\theta(s)$  correto para qualquer amostra  $s \in Z_2$  usando os parâmetros de  $S$ ,  $\vec{v}$  e distância euclidiana  $d(s, t)$  entre as amostras  $s$  e  $t$  (sendo  $t \in Z_1$ ). Na fase de teste, quando uma nova amostra  $s'$  é inserida no espaço de atributos, a sua classificação é determinada pela função custo ótimo  $Cst(s')$  da equação (5). Neste caso, a amostra  $s'$  receberá o rótulo da amostra qualquer  $t' \in Z_1$  que lhe oferecer o menor custo.

$$Cst(s') = \min\{Cst(s'), \{\max\{Cst(t'), d(s', t')\}\}, \forall t' \in Z_1 \quad (5)$$

Em sua variação *kernel* (AFONSO; PEREIRA; PAPA, 2017), denominada kOPF (*Kernel Optimum-Path Forest*), o mapeamento do espaço de alta dimensão é feito calculando-se um vetor de características, em que cada componente ou atributo tem o valor da distância da amostra a ser mapeada com uma diferente amostra de treino, usando uma certa função *kernel*  $\phi$  determinadas pelas equações (6), (7) e (8). Nesse caso, há de se destacar o aumento do tempo computacional na fase de treinamento para conjunto com grande quantidade de dados.

$$\text{Identidade: } \phi(s, t) = ||s, t|| \quad (6)$$

$$\text{RBF: } \phi(s, t) = e^{-(\gamma ||s, t||^2)} \quad (7)$$

$$\text{Sigmoide: } \phi(s, t) = \tanh(\gamma ||s, t|| + C_{reg}) \quad (8)$$

Onde  $||s, t||$  se refere à distância euclidiana entre  $s$  e  $t$ ,  $C_{reg}$  o parâmetro de regularização no intervalo entre  $[-32, 32]$  e  $\gamma$  o coeficiente *kernel* entre  $[0, 32]$ , com passo 2.

#### 4.5. Avaliação de conjunto com dados desbalanceados

Um bom resultado nos classificadores de padrões, quando submetidos ao treinamento e teste com conjuntos de dados de diferentes quantidades de amostras por classe, principalmente com a presença de uma classe predominantemente maior sobre as demais, nem sempre significa boa capacidade de generalização do algoritmo classificador, ou seja, o classificador pode acertar toda a classe dominante e errar todas as outras classes menores.

Conforme proposto por Papa et al. (2009), o cálculo das taxas de acertos nos algoritmos utilizados nesta tese considera os diferentes tamanhos das classes que podem conter os bancos de dados, ou seja, considera os impactos dos erros de classificações que ocorrem nas menores classes no cálculo da taxa final. Assim, seja  $Z_2$  um conjunto de dados em avaliação formado por uma quantidade  $NT_i$  (onde,  $i = 1, 2, 3, \dots, C_{class}$ ) de amostras de cada classe  $i$ , o percentual de acerto das classes desbalanceadas é dada pela equação (9):

$$Taxa\ de\ acertos = \left( 1 - \frac{\sum_{i=1}^{C_{class}} \frac{FP_i}{|Z_2| - |NT_i|} + \frac{FN_i}{|NT_i|}}{2 C_{class}} \right) \times 100\% \quad (9)$$

Onde  $FN$  e  $FP$  referem-se às classificações falso negativo e falso positivo de cada classe respectivamente, e  $|Z_2|$  é a quantidade de amostras no conjunto  $Z_2$  e  $C_{class}$  o total do número de classes.

#### 4.6. Geração de dados sintéticos de DGA

Para o caso de haver poucos dados para classificações de padrões, neste tópico se propõe um modo de gerar dados artificiais explorando as incertezas que são inseridas no processo da análise de DGA, assim baseado nas informações de incertezas da IEC 60599 (IEC, 2005), considerou-se uma aproximação linear dos valores percentuais atribuídos a elas, resultando na Tabela 5 para determinação dos valores sintéticos.

Tabela 5 – Relação entre  $S_{LIM}$  e precisão típica

Valores de intervalo de $S_{LIM}$ ( $\mu\text{l/l}$ )	Precisão aplicada (%)
Gases $\geq 10 \times S_{LIM}$	5
$9 \times S_{LIM} \leq \text{Gases} < 10 \times S_{LIM}$	8
$8 \times S_{LIM} \leq \text{Gases} < 9 \times S_{LIM}$	11
$7 \times S_{LIM} \leq \text{Gases} < 8 \times S_{LIM}$	14
$6 \times S_{LIM} \leq \text{Gases} < 7 \times S_{LIM}$	17
$5 \times S_{LIM} < \text{Gases} < 6 \times S_{LIM}$	19
Gases $\leq 5 \times S_{LIM}$	20

Fonte: Elaborado pelo autor

Assim, a partir do conjunto de dados reais, gera-se um conjunto de dados sintéticos aleatório de falhas com os percentuais relacionados na Tabela 5, e nesse caso, a amostra real é assumida como valor de referência central para a geração de novas amostras. Os valores nulos

dos dados das amostras de gases reais foram substituídos pelos valores de  $S_{LIM}$  (Tabela 4). Para avaliar o impacto e qualidade das amostras produzidas, foram analisadas as matrizes de confusões resultantes de diversas simulações com alguns dos algoritmos inteligentes abordados.

#### 4.7. Proposta de metodologia de análise de tendência de falhas em transformadores

Entre os algoritmos ANN, SVM, ELM, KNN e OPF, a classificação dos padrões de falhas supervisionadas ocorre em duas etapas distintas: treinamento e teste. Na fase de treinamento, o algoritmo extrai as informações dos fenômenos e propriedades contidas nos dados das amostras dos gases e constitui um espaço de atributos  $m$ -dimensional. Esse espaço é então dividido em várias regiões, onde cada região é formada por uma classe ou padrão de falha existente. A separação dessas classes ocorre por meio de hiperplanos definidos por uma função específica, como por exemplo, presentes nos algoritmos ANN, SVM e ELM, ou por intermédio de uma estrutura formada por grafos construídos a partir da heurística e relação de adjacências entre as amostras dos gases, como as utilizadas nos algoritmos KNN e OPF. Na fase de teste, a classificação dos padrões de falhas das novas amostras ocorre com base na região do espaço em que são alocadas, ou por meio da relação de conectividade com as amostras das diferentes classes existentes na estrutura dos grafos.

Uma vez que o espaço de atributos do classificador, constituído sobre dados rotulados, é usado para reconhecer os padrões das novas amostras distribuídas sobre essa estrutura espacial, então é possível pressupor que uma nova amostra  $A$  qualquer, dotada dos mesmos atributos de outra amostra  $B$ , utilizada na constituição desse espaço, irá dividir a mesma posição e receberá o mesmo rótulo da região ocupada, ou conforme afinidade com a amostra rotulada mais próxima. Se houver mudanças nos valores das concentrações dos gases, essa mesma amostra  $A$  irá se posicionar em outro local da região espacial e, portanto, receberá o rótulo de acordo com a heurística do algoritmo.

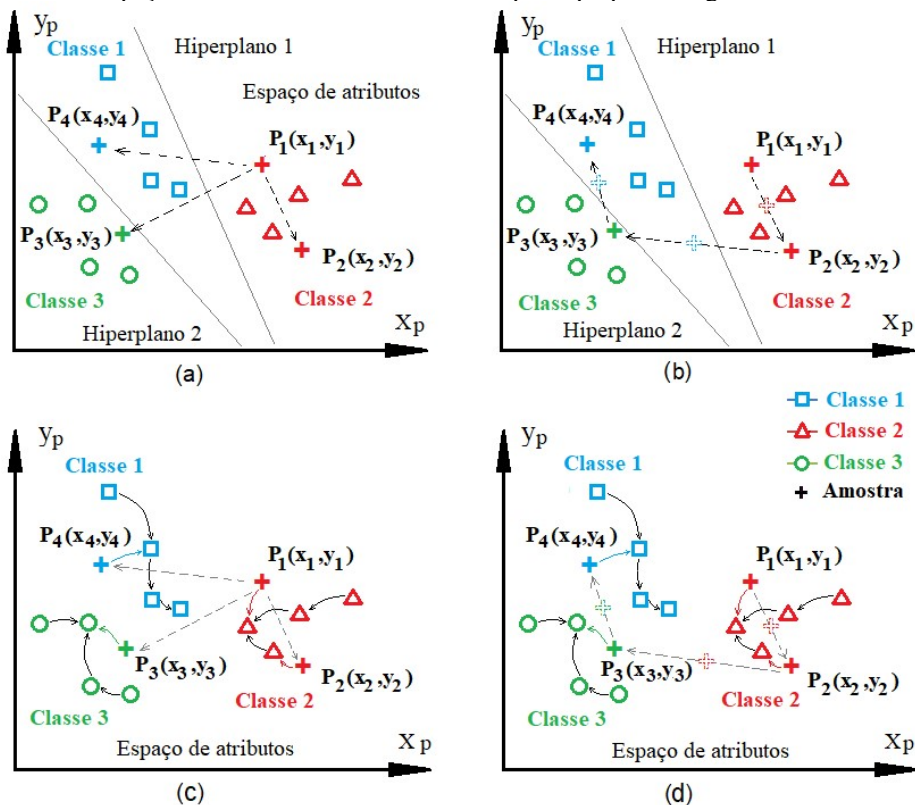
Portanto, seja uma representação simplificada de um espaço bidimensional de atributos para apenas três rótulos, conforme mostra a Figura 10, em que  $P_1(x_1, y_1)$  é a posição inicial de uma nova amostra inserida nesse espaço, e  $P_2(x_2, y_2)$ ,  $P_3(x_3, y_3)$  e  $P_4(x_4, y_4)$  são as posições hipotéticas que essa amostra pode assumir quando se varia os atributos, ou no caso as coordenadas  $x$  e  $y$ . Se os valores assumidos por  $x$  e  $y$  são discretos, as mudanças entre as posições  $P_1$  a  $P_4$  ocorrem de forma discreta (Figuras 10(a) e (c)) no classificador. Se os valores forem contínuos, a mudança de posição será suave, determinando um caminho (por

exemplo, a linha tracejada) que passaria por todas as posições entre  $P_1$  a  $P_4$  (Figuras 10(b) e (d)).

Para o caso de classificação de falhas utilizando DGA, significa que a mudança no valor das concentrações dos gases de uma amostra determina o seu reposicionamento dentro desse espaço de características entre  $P_1$  e  $P_4$ . Tais variações nas concentrações dos gases são resultantes, por exemplo, de falhas que podem evoluir de menor para maior gravidade, ou dependendo do tipo da falha até cessar.

Considerando que o crescimento dos gases é um importante indicador de diagnóstico de falhas incipientes (IEC, 2015; IEEE, 2019), o monitoramento contínuo com intervalo de tempo regular (por exemplo, semanal, mensal, trimestral, outros) entre as análises de DGA em um transformador, também permite conhecer o caminho percorrido por essas leituras sucessivas ao longo do tempo sobre o espaço de atributos dos algoritmos classificadores. Desde que seja possível conhecer a distância e a ordem cronológica dessas amostras em relação ao hiperplano que separa as diferentes classes ou rótulos (Figura 10(b)), ou a intensidade da relação de conectividade entre as amostras próximas (Figura 10(d)), torna-se possível, além de determinar o rótulo das amostras, conhecer a tendência ou comportamento de variação dos gases em relação às falhas.

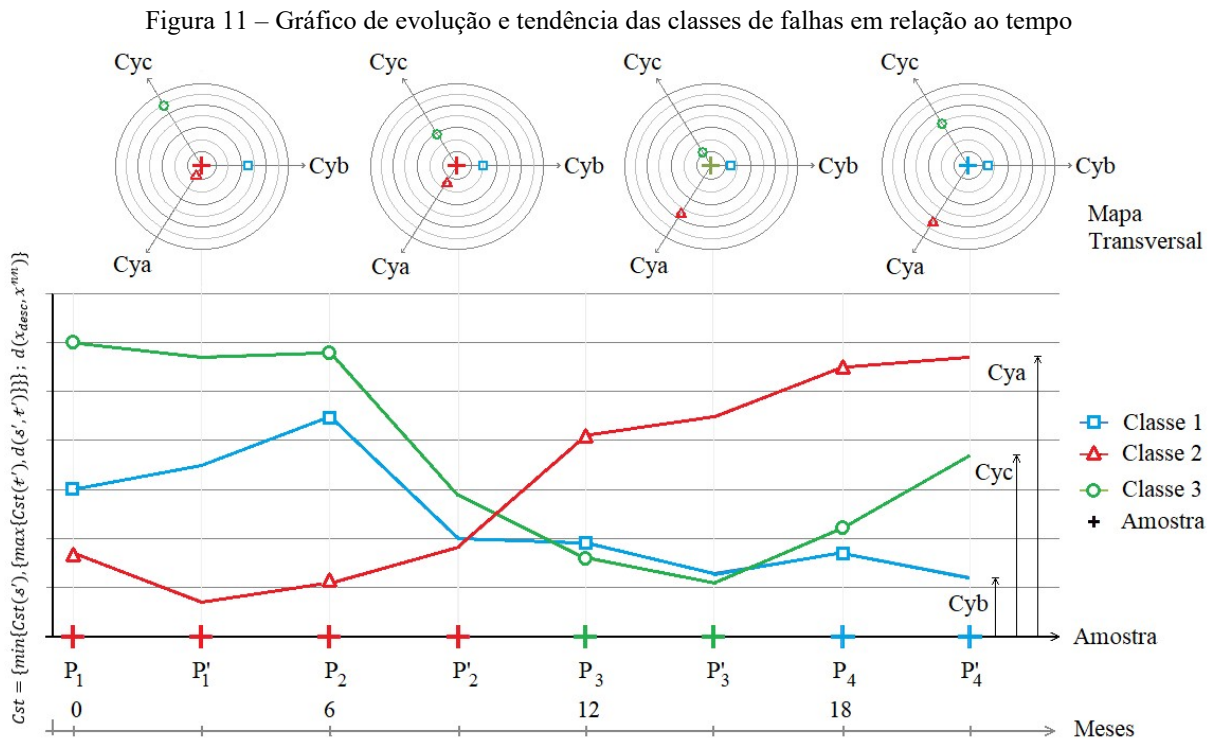
Figura 10 – Espaços de características constituídos por hiperplanos e grafos de classificadores



Fonte: Elaborado pelo autor

Onde: (a) trata-se do espaço de atributos com classes separadas por hiperplanos e valores  $P_1$  a  $P_4$  discretos; (b) espaço de atributos com classes separadas por hiperplanos e valores  $P_1$  a  $P_4$  contínuos; (c) espaço de atributos com classes separadas em grafos e valores  $P_1$  a  $P_4$  discretos; e (d) espaço de atributos com classes separadas em grafos e valores  $P_1$  a  $P_4$  contínuos.

Assim, desde que o conjunto de teste seja formado por amostras de análises sucessivas de gases, determinadas pelos diferentes períodos ou datas de medições em rotinas de manutenção preditiva, a apresentação gráfica do comportamento do estado de um dado transformador com a evolução dos gases pode ser hipoteticamente representado pela curva de tendência de falhas (Figura 11) nas datas de análise de DGA. Observa-se que o período regular e menor entre leituras, melhorará a percepção e o acompanhamento das variações.



Fonte: Elaborado pelo autor

Onde: o eixo das abcissas “Amostra” representa as sucessivas medições ou amostras de DGA do transformador que foram submetidas à análise dos algoritmos 1NN e OPF nas datas exemplos para  $P_1, P_1', P_2, \dots, P_4'$  (no exemplo, mensais); o eixo das ordenadas representa o  $Cst = \{\min\{Cst(s'), \{\max\{Cst(t'), d(s', t')\}\}\}; d(x_{desc}, x^{nn})\}$  que utiliza como parte de sua heurística de classificação a métrica distância euclidiana entre as amostras do eixo “Amostra” com as diferentes classes existentes no espaço de atributos dos algoritmos, ou seja,

$C_{ya}$ ,  $C_{yb}$  e  $C_{yc}$  referem-se à conectividade (custo) entre a nova amostra classificada e as amostras das diferentes classes. Quanto menor o valor do custo, maior será a proximidade e conectividade com determinada classe (como pode ser visto no mapa de representação transversal), ou seja, cada nova amostra do transformador submetida à classificação, recebe o rótulo da amostra existente no espaço de atributos que lhe oferecer o menor custo, como por exemplo,  $P_1$  é rotulado como Classe 2, pois,  $C_{ya} < C_{yb} < C_{yc}$ .

Analisando-se o comportamento das falhas, entre o intervalo  $P_1$  e  $P_2'$ , observa-se que as variações das concentrações de gases das amostras do transformador movimentaram-se no espaço de atributos, no entanto, mantiveram o rótulo da Classe 2. Note que, entre as amostras nas datas  $P_2$  e  $P_2'$  há uma sinalização de aproximação (convergência) entre os três rótulos e isso significa que a amostra medida em  $P_2'$  está caminhando em direção a uma região de transição (fronteira de decisão com classes diferentes muito próximas) em que a qualquer mudança nas concentrações de gases pode levar a uma mudança de rótulo e conseqüentemente, da classificação da falha, como pode ser visto logo a seguir em  $P_3$ , e assim por diante.

Outra informação disponibilizada pelo gráfico (Figura 11) é a visualização da proximidade de diferentes tipos de falhas incipientes durante a vida operacional do transformador, ou seja, torna-se possível estimar a tendência de situações de falhas não restritas somente a uma classe (classe 1), mas observando a sequência de falhas determinadas pelos custos apresentados do menor (falha principal – classe 1) para o maior (falhas secundária – classe 3 e terciária – classe 2). Isso permite conhecer as proximidades de diferentes tipos de falhas, que podem estar relacionadas ao evento em questão, principalmente nos casos de erros na classificação da falha principal que os algoritmos podem cometer, logo, tal informação de proximidade entre falhas amplia o horizonte de informação envolvendo-as e as concentrações de gases relacionadas, tornando-a assim, a análise mais informativa.

A importância dessas curvas de tendências de falhas está associada à possibilidade de avaliar as ocorrências de falhas a curto e médio prazo, e suas contribuições relacionam-se com os seguintes fatores:

- Permite entender o comportamento operacional do transformador com relação às falhas incipientes no período (tempo) observado, bem como identificar mudanças que ocorrem (previsão). Nesse caso, com essa informação torna-se possível antecipar ações que evitem a parada do equipamento, ou seja, indica o risco variável no tempo e assim possibilita também tomar a decisão com base no tempo;



- Permite avaliar e comparar o comportamento operacional de grupos de transformadores de mesmo modelo em uma subestação, podendo distinguir os estados operacionais diferentes;
- Possibilita relacionar questões térmicas e possíveis impactos na vida útil; e
- Auxilia nas tomadas de decisões antecipadas, para evitar ou reduzir custos indesejados.

De acordo com a IEEE C57.104, a escolha da frequência de análise DGA pode ser determinada com base em custos financeiros decorrentes de uma parada forçada, custo de substituição do ativo, da criticidade do fornecimento do serviço, requisitos contratuais (compensação financeira), questões de segurança, entre outros, sendo de uso comum frequência anual, entretanto, se os custos decorrentes da parada forçada for muito alto, pode ser justificável o uso de monitoramento contínuo ou com frequência reduzida para mensal ou trimestral.

No Brasil, as concessionárias de transmissão de energia tem adotado uma frequência semestral de análise (aprovado em audiência pública ANEEL nº22/2014), sendo adotado também pela distribuição (exceto redes secundárias de baixa tensão) em suas subestações de média e alta tensão, pois uma janela de período muito grande em caso de presença de falha incipiente, poderia resultar em maiores danos econômicos. Uma frequência entre análises de DGA maior que duas vezes ao ano pelas concessionárias (por exemplo, quatro vezes ou trimestral), permitiria um melhor acompanhamento do estado do transformador indiferentemente do método de diagnóstico de falha adotado, e para a metodologia aqui abordada, uma melhor definição do gráfico de tendência falhas, contudo deve ser viavelmente econômica e possível de executar operacionalmente e, nesse caso, uma saída seria identificar as subestações críticas e determinar rotinas diferenciadas com período menor de análise.

Essa metodologia de diagnóstico de falha aplica-se a transformadores de médias e altas tensões (exceto de redes secundárias de baixa tensão), que compõem as subestações de transmissão, subtransmissão e distribuição, a partir de 69kV, ou que estejam sujeitas a determinação da análise de DGA, conforme Resolução Normativa nº669/2015 por parte da ANEEL. No caso da transmissão, as preocupações voltam-se sobre manter a continuidade do serviço e, portanto, desde que o transformador seja importante para manter o sistema operacional, está sujeito à manutenção preditiva e análise de DGA. No caso da distribuição, as preocupações são as mesmas quanto à manutenção do serviço, entretanto, a regulamentação que trata do assunto é menos agressiva economicamente, bem como os recursos das redes

elétricas e políticas de reserva de equipamentos disponíveis, possibilitam uma maior flexibilidade de operacional para contornar a situação.

#### **4.8. Considerações gerais do capítulo**

Com base nos algoritmos OPF e INN, é proposta então uma abordagem em análise de falhas incipientes em transformadores. Por intermédio da análise do histórico de gases dissolvidos no óleo, obtidos através de medidas laboratoriais, torna-se possível avaliar o estado operacional do transformador, observando o comportamento e a tendência das curvas de evolução das falhas ao longo do tempo de sua vida em serviço. O condicionamento e uso de dados atuais e a manutenção do histórico de DGA, naturalmente agrega ao OPF e INN características sobre uma importante informação empregada no diagnóstico de falhas que é a taxa de evolução e crescimento das concentrações dos gases, ocasionadas pelos estresses térmicos e elétricos a que o transformador está sujeito.

Tal informação tem relação próxima com o grau de confiança sobre a análise do diagnóstico, ou seja, uma taxa de crescimento dos gases elevada (por exemplo, acima de 10% ao mês (IEC, 2007) sinaliza um aumento na chance na presença de falha incipiente. Contudo, para que seja possível obter uma melhor definição dos gráficos de evoluções dos gases e tendências, devem ser previstos uma frequência de análise menor possível e ajustado nos critérios de criticidades envolvidos.

A proposta do método de diagnóstico prevê o comportamento para dois, três e cinco tipos de estados de falhas, e assim será possível além de acompanhar o comportamento da tendência de falhas específicas no transformador, conhecer os níveis de criticidade das falhas envolvidas, para uma melhor gestão da manutenção. Nesse contexto, a análise com base nos gráficos de tendências de falhas proposta nesta tese, deve seguir a rotina: o primeiro gráfico com duas classes, é usado para acompanhar a evolução e identificar a presença ou não de falhas, de modo rápido e mais preciso que os demais; o segundo gráfico permite observar se a falha possui origem térmica ou elétrica; e o terceiro gráfico distingue e classifica a falha principal e as proximidades com outras possíveis. Essa abordagem de análise e diagnóstico trata-se de um instrumento de apoio à manutenção e no auxílio na tomada de decisão de inspeção do transformador, bem como pode fazer também parte do conjunto de análise de risco que envolve sua eventual parada.

Outras possibilidades de uso que o método permite são: o acompanhamento contínuo do comportamento operacional do transformador, com uso de monitoramento remoto e,

assim, obter informações sempre atualizada do equipamento em uma janela de tempo ajustável pelo operador; pode ser utilizado como recurso para avaliar o período em que o equipamento esteve sujeito ao sobreaquecimento e assim, contribuir para avaliar a vida útil remanescente do equipamento; e permite ser integrado aos métodos de análise de criticidade e gestão de risco do transformador.

O uso desse método não descarta o conhecimento e uso informativo das concentrações e tipos dos gases presentes, uma vez que esses são indicativos que caracterizam o tipo de falha envolvida e devem continuar sendo usadas para apoiar essa análise. Não se pode também negligenciar informações acerca de ocorrências ou eventos particulares de equipamentos similares (do mesmo modelo), acionamento de dispositivos de proteção, carregamento e outros, que podem ter impactos na produção dos gases.

Quanto as bases de construções do conjunto de dados de treinamento dos algoritmos, uma vez que as informações das classificações fornecidas para as novas amostras são representantes deles, os transformadores operando normalmente no sistema elétrico com níveis elevados de gases e acima dos limites de concentrações definidos como normais, possuem grande chance de serem classificados em outra classe, visto que estão fora da faixa normal. Logo, procedimentos de desgaseificação do óleo devem ser avaliados como parte do processo do uso desta metodologia.

Equipamentos especiais que trabalhem com níveis elevados de gases, como por exemplo transformadores, que possuem chave seccionadora sob carga (taps sob carga) e compartilham do tanque de expansão de óleo, não podem fazer uso dos algoritmos treinados com conjuntos de dados que não possuem tal dispositivo, ou seja, os algoritmos devem ser treinados com conjunto de dados que considerem essas particularidades.

Em complemento, neste capítulo, apresentou-se o processo adotado para tratamentos dos dados, seleção dos melhores atributos e geração de dados sintéticos empregados na análise dos resultados experimentais.

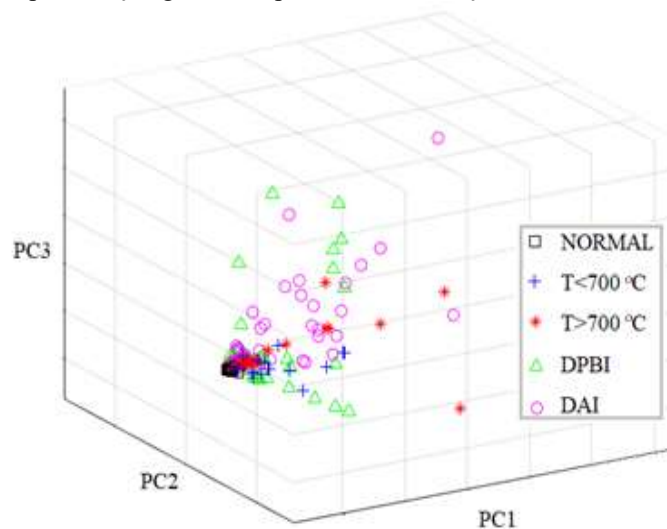
## CAPÍTULO 5 – RESULTADOS EXPERIMENTAIS E DISCUSSÕES

### 5.1. Base de dados de DGA

O conjunto de dados usados para treinamento dos algoritmos foi extraído das bases IEC TC10 (DUVAL; DEPABLO, 2001), IEEE *Dataport* (ENWEN et al., 2019), CPFL – Companhia Paulista de Força e Luz (FILHO, 2012) e artigos (EQUBAL et al., 2018; GHONEIM; WARD, 2012; SONI; CHAUDHARI, 2016), reunindo um total de 2298 amostras distribuídas em 2049 amostras normais, 63 falhas térmicas  $T < 700^{\circ}\text{C}$ , 47 falhas térmicas  $T > 700^{\circ}\text{C}$ , 70 descargas elétricas parciais e de baixa intensidade associadas (DPBI), e 69 descargas de alta intensidade (DAI), para cinco tipos de gases combustíveis ( $\text{H}_2$ ,  $\text{CH}_4$ ,  $\text{C}_2\text{H}_2$ ,  $\text{C}_2\text{H}_4$  e  $\text{C}_2\text{H}_6$ ).

A Figura 12 mostra uma projeção espacial das amostras de gases do conjunto de dados obtidos, representadas pelos seus componentes principais PC1, PC2 e PC3. O algoritmo de análise dos componentes principais (PCA) extrai as características essenciais dos dados, por meio da transformação do espaço vetorial, para reduzir a dimensão do conjunto de dados e assim permitir, como por exemplo, evidenciar as características mais importantes no reconhecimento dos padrões de classes com menos variáveis (SILVA, 2011). Os dados dos transformadores analisados e relatórios técnicos usados na validação, foram obtidos de uma concessionária de transmissão de energia elétrica.

Figura 12 – Representação gráfica ampliada da distribuição das amostras DGA via PCA

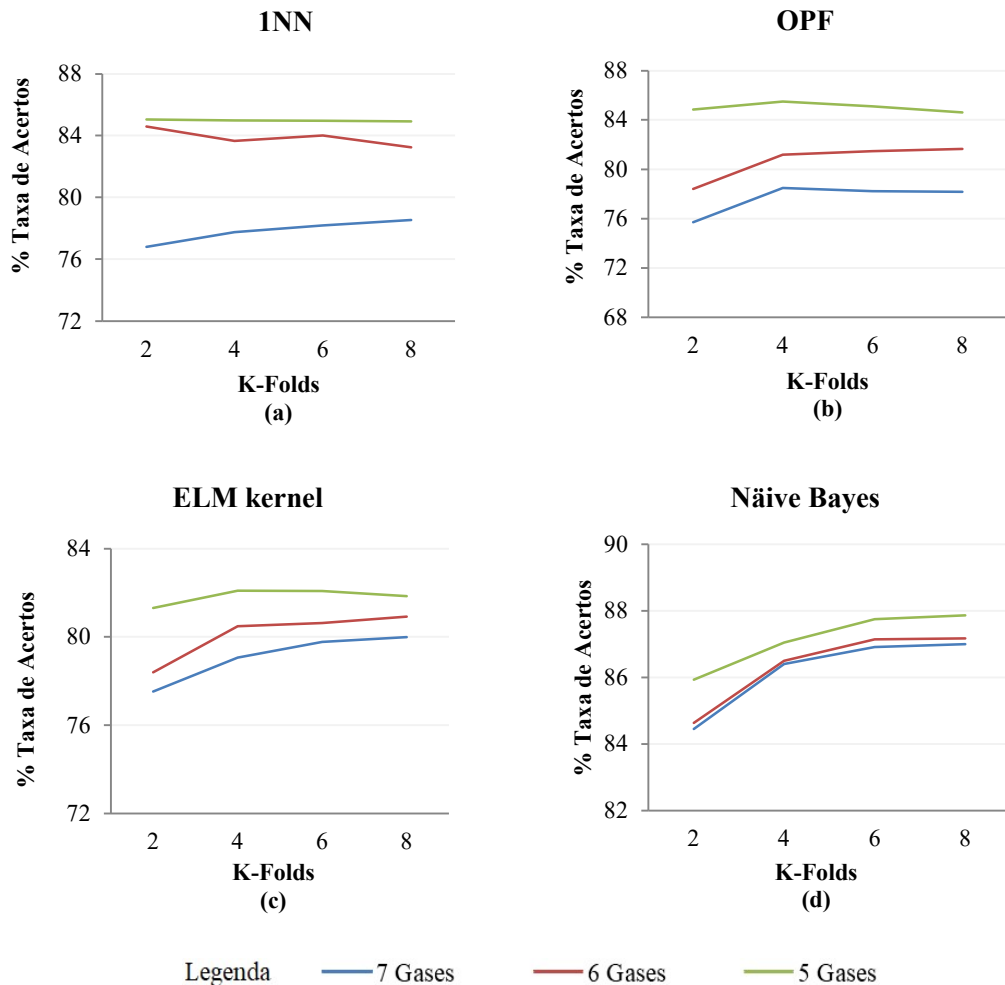


Fonte: Elaborado pelo autor

Os gases CO e CO<sub>2</sub> não foram usados na análise, em razão de serem relacionados à degradação do papel isolante (DUVAL, 2012; IEEE C57.104, 2019). Sendo tratada à parte nos métodos normatizados, bem como por algumas das fontes de dados não disponibilizarem as suas concentrações, ou por não relacionarem a um maior número de tipos de falhas, o que reduziria o número de amostras considerado.

Contudo, com objetivo de investigação, foi possível formar um conjunto com um menor número de amostras contendo 2049 normais, 57 térmicas e 74 elétricas com os 7 gases e 3 tipos de falhas (normal, elétrica e térmica) para avaliar a influência do CO e CO<sub>2</sub> sobre a taxa de acertos nos algoritmos classificadores. Como resultado das simulações para o conjunto de dados selecionado composto por H<sub>2</sub>, CH<sub>4</sub>, C<sub>2</sub>H<sub>2</sub>, C<sub>2</sub>H<sub>4</sub>, C<sub>2</sub>H<sub>6</sub>, CO e CO<sub>2</sub>, e diferentes validações cruzadas, nas Figuras 13(a) a (d) é possível observar que ambos os gases CO e CO<sub>2</sub> não contribuíram para melhora da taxa no conjunto de dados avaliado.

Figura 13 – Taxa de acerto média para dados de DGA parciais



Fonte: Elaborado pelo autor

## 5.2. Pré-processamento do conjunto de dados de treinamento do DGA

A partir da análise do conjunto com 2049 amostras normais, a Tabela 5 apresenta os valores das concentrações de gases determinados com base nas técnicas aplicadas para limpeza de *outliers*, sendo FT (Falha Térmica) e FE (Falha Elétrica) o número de amostras encontradas nos conjuntos de dados de falhas cujos gases individuais estão todos abaixo dos limites individuais em questão. As amostras de falhas identificadas podem levar ao aumento de erros de classificação, uma vez que fazem parte das amostras rotuladas que constituiriam o espaço de atributos na fase de treinamento dos algoritmos. Também é possível observar as relações entre tipos de falhas identificadas e os limites dos gases encontrados, ou seja, falhas térmicas associadas aos gases CH<sub>4</sub>, C<sub>2</sub>H<sub>4</sub> e C<sub>2</sub>H<sub>6</sub>, e as falhas elétricas aos gases H<sub>2</sub> e C<sub>2</sub>H<sub>2</sub>. O número de grupos K<sub>C</sub> determinado pelo método *Elbow* usados nos clusterizadores *K-means* e *K-medoids* foi de 8 grupos para uma variância máxima admitida de 15%.

Tabela 5 – Valores limites de concentrações de gases para amostras normais (µl/l ou PPM)

Técnica	H <sub>2</sub>	CH <sub>4</sub>	C <sub>2</sub> H <sub>2</sub>	C <sub>2</sub> H <sub>4</sub>	C <sub>2</sub> H <sub>6</sub>	FT	FE
DNP	66,8	79,4	1,2	67,1	134,6	19	0
IEC 60599	42,0	80,0	0	50,0	72,0	7	0
<i>K-means</i>	59,9	67,5	1,8	67,3	37,8	14	2
<i>K-medoids</i>	88,6	69,77	1,8	67,3	38,8	18	3

DNP – Distribuição Normal de Probabilidade, FT – Falha Térmica, FE – Falha Elétrica

Fonte: Elaborado pelo autor

A Tabela 6 apresenta os resultados de simulações após a limpeza dos *outliers*. Os algoritmos usados nessa avaliação, foram: KNN, OPF, kOPF (*kernel* OPF), NB, SVM, ANN MLP e ELM. As taxas de acertos dos algoritmos foram obtidas para as condições antes (com dados brutos) e após a limpeza no conjunto de dados, considerando cinco tipos de gases, cinco classes de falhas, 50 simulações cada, dados normalizados e validação cruzada *5-folds*.

Analisando-se a taxa de acertos dos valores obtidos para as diferentes técnicas de filtragem e classificadores, observa-se que há uma margem pequena entre eles, contudo o melhor resultado geral foi obtido para condição de limpeza via DNP (Distribuição Normal de Probabilidade) sendo, portanto adotado para continuidade do pré-processamento de dados. O novo conjunto de dados ficou então composto por 1818 amostras de condições normais, 44 amostras de falhas térmicas T<700°C, 47 amostras de falhas térmicas T>700°C, 70 amostras

de descargas elétricas parciais e/ou de baixa intensidade, e 69 amostras de descargas de alta intensidade, totalizando 2048 amostras.

Tabela 6 – Taxa de acertos KNN, OPF, kOPF, NB, SVM, ANN e ELM, para dados brutos e pré-processados

Algoritmos	Dados	Taxa de acertos (%) / Parametrização				
		Número K	1	3	5	7
KNN	Dados Brutos	80,9±2,8	77,7±2,7	77,8±2,6	77,3±2,6	76,7±2,7
	DNP <sup>1</sup>	<b>84,9±2,3</b>	<b>82,1±2,7</b>	<b>81,6±2,7</b>	<b>80,5±2,8</b>	<b>80,0±2,9</b>
	IEC 60599	81,3±2,8	78,3±3,0	77,9±3,0	77,9±2,9	77,7±2,9
	<i>K-means</i>	83,3±2,8	80,6±2,8	80,8±2,9	79,4±3,0	78,3±3,2
	<i>K-medoids</i>	83,8±2,8	80,6±2,8	80,1±2,9	79,5±3,1	78,5±3,0
OPF	Dados Brutos			79,9±2,4		
	DNP			<b>84,1±2,3</b>		
	IEC 60599			80,9±2,8		
	<i>K-means</i>			83,1±2,5		
	<i>K-medoids</i>			83,5±2,4		
kOPF <sup>2</sup>	Dados Brutos			77,7±3,1		
	DNP			<b>81,4±2,4</b>		
	IEC 60599			78,0±2,5		
	<i>K-means</i>			81,3±2,3		
	<i>K-medoids</i>			81,0±2,0		
Náive Bayes <sup>3</sup>	Dados Brutos			73,8±0,7		
	DNP			<b>79,6±1,0</b>		
	IEC 60599			75,3±0,9		
	<i>K-means</i>			77,4±0,8		
	<i>K-medoids</i>			78,1±0,7		
SVM <sup>4</sup>	Dados Brutos		81,8±0,8 / C <sub>reg</sub> =928,53 e $\gamma=0,079512$			
	DNP		<b>87,0±0,8</b> / C <sub>reg</sub> =131072 e $\gamma=0,015625$			
	IEC 60599		82,5±0,8 / C <sub>reg</sub> =22,27 e $\gamma=0,008631$			
	<i>K-means</i>		85,0±0,7 / C <sub>reg</sub> =6,69 e $\gamma=0,021173$			
	<i>K-medoids</i>		84,5±0,9 / C <sub>reg</sub> =43,74 e $\gamma=0,10576$			
ANN MLP <sup>5</sup>	Dados Brutos			81,5±1,5		
	DNP			<b>85,7±1,4</b>		
	IEC 60599			81,9±1,5		
	<i>K-means</i>			84,2±1,3		
	<i>K-medoids</i>			83,6±1,4		
ELM <sup>6</sup>	Dados Brutos		75,9±0,8 / C <sub>reg</sub> =131072 e $\gamma=0,015625$			
	DNP		<b>81,4±0,8</b> / C <sub>reg</sub> =32 e $\gamma=0,0078125$			
	IEC 60599		78,0±0,7 / C <sub>reg</sub> =1024 e $\gamma=0,0015625$			
	<i>K-means</i>		79,9±0,9 / C <sub>reg</sub> = 256 e $\gamma=0,015625$			
	<i>K-medoids</i>		80,0±0,8 / C <sub>reg</sub> = 4096 e $\gamma=0,0078125$			

NOTAS:

<sup>1</sup> Distribuição Normal de Probabilidade

<sup>2</sup> Melhor taxa de acerto: *kernel* RBF função de base radial, para C<sub>reg</sub>=1,1 e  $\gamma=1,1$

<sup>3</sup> Estimador de densidade *kernel* gaussiana

<sup>4</sup> *kernel* RBF e estratégia um contra um (OVO – *One Versus One*)

<sup>5</sup> Parametrização: 1 camada, 9 neurônios, Levenberg Marquardt, tangente hiperbólica

<sup>6</sup> *kernel* RBF

C<sub>reg</sub> Coeficiente de regularização

$\gamma$  Coeficiente *kernel*

Fonte: Elaborado pelo autor

Na etapa de construção dos dados, a partir do conjunto DNP com 5 gases combustíveis ( $H_2$ ,  $CH_4$ ,  $C_2H_2$ ,  $C_2H_4$ ,  $C_2H_6$ ) foram gerados 12 outros atributos ( $CH_4/H_2$ ,  $C_2H_2/CH_4$ ,  $C_2H_2/C_2H_4$ ,  $CH_4/H_2$ ,  $C_2H_4/C_2H_6$ ,  $C_2H_6/CH_4$ ,  $H_2/S_m$ ,  $CH_4/S_m$ ,  $C_2H_2/S_m$ ,  $C_2H_4/S_m$ ,  $C_2H_6/S_m$  e  $S_m = H_2 + CH_4 + C_2H_2 + C_2H_4 + C_2H_6$ ), sendo posteriormente normalizados pela média e desvio padrão. Os algoritmos de otimizações usados para seleção dos melhores atributos foram enxame de partículas (PSO), algoritmo genético (GA), busca cuco (CS) e colônia de formigas (ACO), tendo como função *fitness* maior taxa de acertos determinada pelo algoritmo 1NN. A Tabela 7 apresenta os resultados obtidos para simulação do conjunto DNP, expandido para 17 atributos de gases e os parâmetros usados nos ajustes dos algoritmos de otimização.

Tabela 7 – Seleção de atributos para dados pré-processados para 5 tipos de falhas

Algoritmo	Seleção de atributos				Tipos de gases/combinção
	Mín.	Máx.	Melhor	Taxa acerto médio (%)	
ACO+1NN	5	11	8	86,2±3,0	$H_2$ , $CH_4$ , $C_2H_2$ , $C_2H_4$ , $C_2H_6$ , $S_m$ , $C_2H_2/CH_4$ , $C_2H_6/C_2H_2$
PSO+1NN	5	11	7	87,0±2,7	$H_2$ , $CH_4$ , $C_2H_2$ , $C_2H_4$ , $C_2H_6$ , $S_m$ , $C_2H_2/C_2H_4$
CS+1NN	7	9	<b>8</b>	<b>87,3±2,7</b>	$H_2$ , $CH_4$ , $C_2H_2$ , $C_2H_4$ , $C_2H_6$ , $S_m$ , $C_2H_2/CH_4$ , $C_2H_2/C_2H_4$
GA+1NN	6	10	<b>8</b>		

Parâmetros de simulação: 1NN@50 simulação, CS/GA/PSO/ACO@100 simulações e 20 iterações.

ACO	Formigas=30, feromônio $\tau_{inicial}=0,1$ , evaporação $\rho=0,1$ , peso feromônio $\alpha=1$
PSO	Partículas=30, inércia $\omega_{inicial}=1$ , $\omega_{correção}=0,9$ , constantes cognitivas e sociais $C_1=1$ e $C_2=2$
CS	Ninhos=30, número de mutações=10, Coeficiente Levy $\lambda = 1$ , passo $\alpha_p=1$
GA	Genes=30, taxa mutação=0,1, taxa cruzamento=0.9

Fonte: Elaborado pelo autor

Os limites mínimo e máximo de atributos encontrados nas simulações variaram entre os intervalos de 5 a 11 tipos ou combinações de gases, sendo apresentados os melhores resultados entre 7 e 8 atributos, conforme mostra a Tabela 7. Comparando-se o desempenho entre os algoritmos na busca das soluções, observou-se uma maior consistência nos resultados obtidos para os algoritmos CS e GA, ou seja, respectivamente 55% e 33% dos resultados foram para o mesmo tipo de combinação de gases encontrada, levando ao uso do conjunto de dados composto pelos gases  $H_2$ ,  $CH_4$ ,  $C_2H_2$ ,  $C_2H_4$ ,  $C_2H_6$ ,  $S_m$ ,  $C_2H_2/CH_4$ ,  $C_2H_2/C_2H_4$ .

A Tabela 8 apresenta a simulação para outros algoritmos de aprendizado de máquina que usam diferentes processos de classificações de padrões para avaliar o risco de *overfitting* sobre os dados obtidos após a seleção dos atributos. Como pode ser observado, que a maioria dos algoritmos utilizados apresentaram melhora na taxa de acertos com os 8 atributos



afastando a possibilidade de *overfitting* sobre a composição determinada pelo uso algoritmo 1NN como fator *fitness*.

Tabela 8 – Resultados obtidos para simulações após seleção dos melhores 8 atributos

Algoritmo	Atributos	Taxa acerto média (%)	Configuração para 50 simulações, 5 classes de falhas e validação cruzada 5-fold
OPF	5	80,4±2,9	-
	8	<b>86,9±2,6</b>	
ANN MLP	5	86,2±1,7	1 Camada, 9 neurônios, Levenberg Marquardt, tangente hiperbólica
	8	<b>86,1±1,2</b>	
ELM	5	81,6±0,9	$C_{reg}=256$ (regularização), RBF <i>kernel</i> , $\gamma=0,015625$ (coeficiente <i>kernel</i> )
	8	<b>83,0±1,0</b>	
SVM	5	87,8±0,6	$C_{reg}=428,06$ , RBF <i>kernel</i> , $\gamma= 0,12522$ , estratégia um contra um (OVO)
	8	<b>88,5±0,9</b>	
Näive Bayes	5	79,6±0,8	Estimador de densidade <i>kernel</i>
	8	<b>82,0±0,6</b>	

Fonte: Elaborado pelo autor

Portanto, a partir do conjunto de dados de treinamento validado com 8 atributos de DGA, foram feitas simulações para diferentes composições de classes de falhas, conforme proposto para a abordagem no diagnóstico de falha. A Tabela 9 mostra as matrizes de confusões ou erros (mais informações, consulte o Apêndice A.3) obtidas com a média dos resultados de classificação para os conjuntos de dados brutos e seleção de atributos empregando o algoritmo 1NN e OPF. No caso, foram analisadas as taxas de acertos para os conjuntos de dados assumindo diferentes composições de rotulações, ou seja, com dois rótulos (estado normal e falha), três rótulos (estado normal, falha térmica e falha elétrica) e cinco rótulos identificado anteriormente.

Observa-se que nos casos de dados pré-processados DNP, houve melhora na taxa de acertos, se comparado aos dados brutos. A maior taxa de acerto ocorreu para duas classes demonstrando uma boa “separabilidade” entre as duas classes, normal e falha, sendo reduzido às classificações erradas. No caso com três classes de falhas, já foi possível observar o esperado, ou seja, redução nos erros de classificação térmica, visto que a limpeza no conjunto de dados DNP ocorreu exclusivamente sobre a classe de falhas térmicas, isso tem relação com os limites dos gases  $CH_4$ ,  $C_2H_4$  e  $C_2H_6$  (Tabela 5). Para cinco classes de falhas, o maior impacto ocorreu sobre os dados de falhas térmicas  $T < 700^\circ C$ , sendo em sua maioria para falhas com temperatura  $T < 300^\circ C$ , que pode envolver produção de gases de baixa temperatura (*stray gassing*), já entre as falhas elétricas os erros permaneceram próximos, e isto se deve à grande similaridade encontrada nos dados de gases das falhas de descarga de baixa

intensidade (sem contar descarga parcial incluída) e de alta intensidade. Também ocorreu o esperado com a queda da taxa de acerto média para a distinção de um número maior de classe, contudo, isso pode ser usado na estratégia de identificações de falhas, sua origem e classe específica, tornando a análise mais informativa.

Tabela 9 – Matrizes de confusões/erros para dados DNP e brutos para 2, 3 e 5 classes

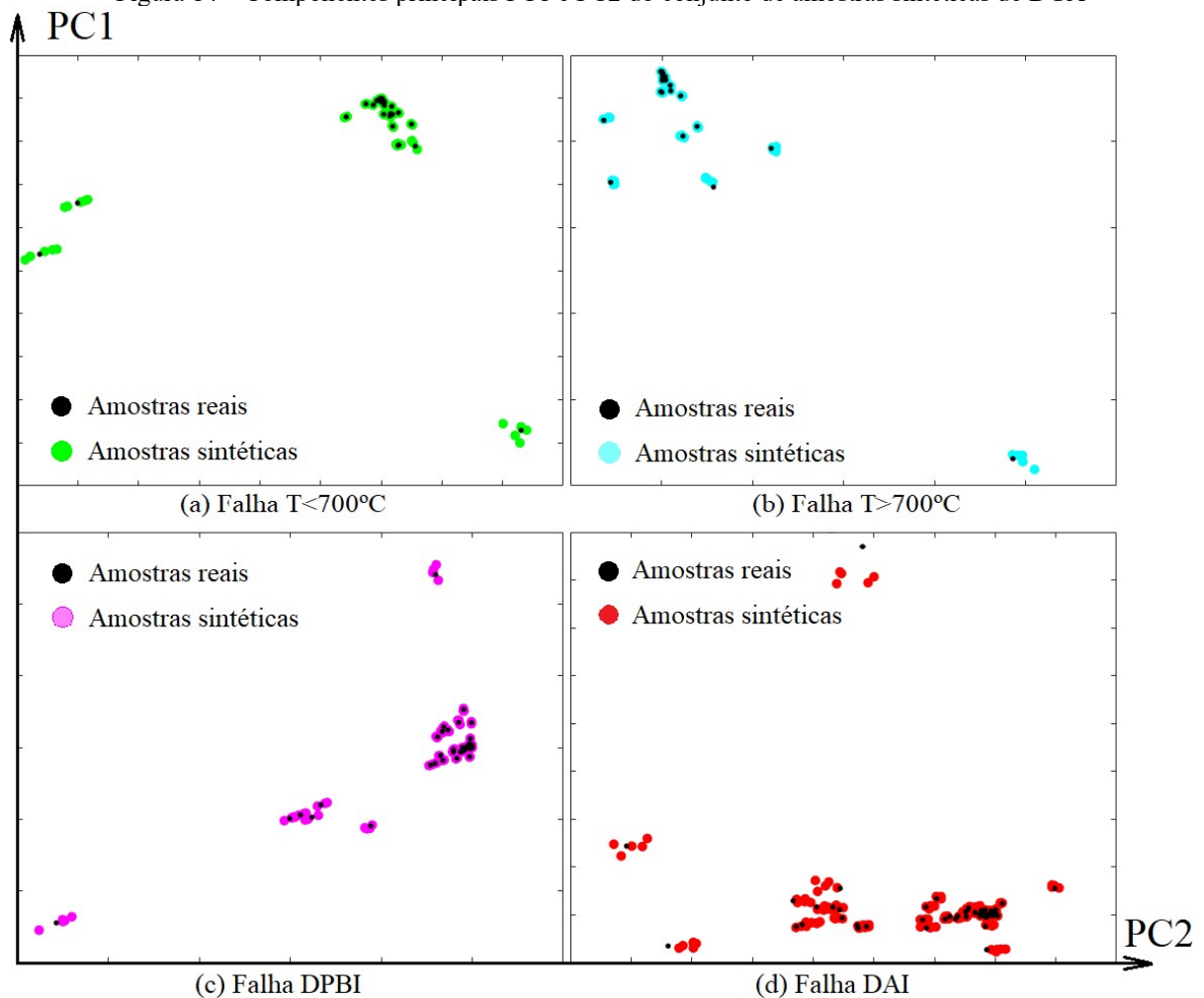
Nº Classes	Algoritmo	Dados	Qtd. Gases	Taxa Acertos (%)	Matriz de confusão								
					Falha	N	TE	-	-	-			
2	1NN	Bruto	5	91,4±2,6	N	403,6	6,2	-	-	-			
					TE	7,8	42,0	-	-	-			
		DNP	8	<b>97,8±1,5</b>	N	363,2	0,4	-	-	-			
					TE	1,9	44,1	-	-	-			
	OPF	Bruto	5	91,4±2,5	N	403,4	6,6	-	-	-			
					TE	7,9	42,1	-	-	-			
		DNP	8	<b>97,7±1,5</b>	N	363,1	0,5	-	-	-			
					TE	1,9	44,1	-	-	-			
-	-	-	-	-	Falha	N	T	E	-	-			
3	1NN	Bruto	5	88,2±2,4	N	403,6	5,5	0,7	-	-			
					T	6,3	13,7	2,0	-	-			
					E	1,4	2,3	24,1	-	-			
		DNP	8	<b>94,1±2,0</b>	N	363,1	0,3	0,2	-	-			
					T	1,6	14,5	2,1	-	-			
					E	0,3	2,5	25,0	-	-			
	OPF	Bruto	5	88,2±2,6	N	403,7	6,2	0,5	-	-			
					T	5,9	14,1	1,9	-	-			
					E	1,8	2,4	23,8	-	-			
		DNP	8	<b>94,1±1,8</b>	N	363,4	0,4	0,4	-	-			
					T	1,4	15,0	3,2	-	-			
					E	0,5	1,9	25,6	-	-			
-	-	-	-	-	Falha	N	T<700°C	T>700°C	DPBI	DAI			
5	1NN	Bruto	5	81,1±2,8	N	403,8	4,4	1,0	0,6	0,1			
					T<700°C	5,3	4,5	1,3	1,3	0,3			
					T>700°C	1,1	1,4	6,5	0,2	0,2			
					DPBI	1,2	1,6	0,2	7,4	3,6			
					DAI	0,3	0,2	0,3	2,8	10,3			
					N	363,2	0,0	0,3	0,1	0,1			
		DNP	8	<b>87,3±2,7</b>	T<700°C	0,7	5,2	1,4	1,2	0,3			
					T>700°C	0,8	0,9	7,0	0,4	0,2			
					DPBI	0,0	1,4	0,3	9,6	2,7			
					DAI	0,2	0,3	0,5	2,3	10,5			
					OPF	5	80,4±2,9	N	403,1	5,2	1,1	0,6	0,1
								T<700°C	5,3	4,7	1,2	1,4	0,3
	T>700°C	1,1	1,6	6,9				0,3	0,3				
	DPBI	1,6	1,7	0,1				7,2	3,2				
	DAI	0,1	0,3	0,2				3,4	9,9				
	N	363,6	0,1	0,2				0,1	0,0				
	DNP	8	<b>86,9±2,6</b>	T<700°C	0,5	4,5	2,1	1,4	0,4				
				T>700°C	1,0	1,0	7,7	0,2	0,1				
DPBI				0,0	1,3	0,1	10,0	2,6					
DAI				0,1	0,3	0,5	2,5	10,6					

N – Normal, T – Térmico, E – Elétrico, TE – Térmico + Elétrico, DPBI – Descarga Parcial e de Baixa Intensidade, DAI – Descarga de Alta Intensidade, Térmico = T<700°C+T>700°C, Elétrico= DPBI+DAI

Fonte: Elaborado pelo autor

No contexto de poucos dados disponíveis, a partir do conjunto de dados reais de falhas composto por 230 amostras com cinco gases, foram gerados novas amostras (sintéticas) numa proporção de cinco vezes esse valor, formando um conjunto de dados adicionais composto por 220 amostras  $T < 700^{\circ}\text{C}$ , 235 amostras  $T > 700^{\circ}\text{C}$ , 350 amostras de descargas parciais e de baixa intensidade (DPBI), e 345 amostras de descargas de alta intensidade (DAI), sendo posteriormente agregado ao conjunto real de 2048 amostras, totalizando 3198. Os 1818 dados normais não foram replicados em razão do número elevado de amostras. As Figuras 14(a) a (d) mostram a disposição das amostras de falhas artificiais geradas com base nas incertezas, representadas no espaço bidimensional, por meio de suas componentes principais PC1 e PC2, sendo possível verificar a ocorrência de um aumento de densidade próxima às amostras reais, conforme esperado. As amostras reais são representadas por círculos na cor preta e as amostras sintéticas geradas em círculo colorido conforme legenda.

Figura 14 – Componentes principais PC1 e PC2 do conjunto de amostras sintéticas de DGA



Fonte: Elaborado pelo autor

Os resultados das matrizes de confusões para a simulação dos dados sintéticos com diversos algoritmos são apresentados na Tabela 10 e se referem a 50 simulações com dados normalizados e validação cruzada 5-fold.

Tabela 10 – Matrizes de confusões médias para dados sintéticos e diferentes classificadores

1-NN						
2048 amostras reais, taxa média de acerto <b>84,8±2,6%</b> para 5 gases						
Falhas	Normal	T<700°C	T>700°C	DPBI	DAI	
Normal	<b>363,32</b>	0,01	0,11	1,16	0	
T<700°C	0,80	<b>4,93</b>	1,34	1,49	0,24	
T>700°C	0,62	1,344	<b>6,82</b>	0,27	0,35	
DPBI	0,68	1,84	0,32	<b>7,53</b>	<b>3,63</b>	
DAI	0,04	0,24	0,48	2,78	10,26	
3198 amostras, taxa média de acerto <b>99,9±0,1%</b> para 5 gases						
	Normal	T<700°C	T>700°C	DPBI	DAI	
Normal	<b>362,90</b>	0	0,20	0,50	0	
T<700°C	0	<b>52,76</b>	0	0,04	0	
T>700°C	0,02	0	<b>56,38</b>	0	0	
DPBI	0	0,28	0	<b>83,70</b>	0	
DAI	0	0	0	0,20	<b>82,80</b>	
3198 amostras, taxa média de acerto <b>99,4±0,3%</b> para 8 gases						
	Normal	T<700°C	T>700°C	DPBI	DAI	
Normal	<b>363,13</b>	0,02	0,22	0,2	0,04	
T<700°C	0,33	<b>52,06</b>	0	0,41	0	
T>700°C	0,17	0	<b>56,23</b>	0	0	
DPBI	0,04	0,80	0	<b>82,28</b>	0,88	
DAI	0	0	0	0,79	<b>82,01</b>	
OPF						
3198 amostras, taxa média de acerto <b>99,8±0,2%</b> para 5 gases						
	Normal	T<700°C	T>700°C	DPBI	DAI	
Normal	<b>363,10</b>	0	0,10	0,40	0	
T<700°C	0	<b>53</b>	0	0	0	
T>700°C	0,20	0	<b>56,20</b>	0	0	
DPBI	0	0,40	0	<b>83,60</b>	0	
DAI	0	0	0,30	0	<b>82,70</b>	
Taxa média de acerto <b>99,2±0,4%</b> para 8 gases						
	Normal	T<700°C	T>700°C	DPBI	DAI	
Normal	<b>363,6</b>	0	0,27	0,14	0	
T<700°C	0,71	<b>51,57</b>	0	0,71	0	
T>700°C	0,29	0,43	<b>56,19</b>	0	0	
DPBI	0,01	1,14	0	<b>82,14</b>	0,57	
DAI	0	0	0	0,57	<b>82,42</b>	
Náive Bayes						
3198 amostras, taxa média certo <b>87,3±0,2%</b> para 5 gases, estimador de densidade <i>kernel</i> gaussiana						
	Normal	T<700°C	T>700°C	DPBI	DAI	
Normal	<b>362,18</b>	1,4	0	0,02	0	
T<700°C	1,22	<b>41,77</b>	2,58	6,98	0,25	
T>700°C	0,48	16,48	<b>30,91</b>	7,42	1,11	
DPBI	2,72	3,14	0,24	<b>63,80</b>	14,10	
DAI	0	1,21	0,15	16,89	<b>64,55</b>	

Conclusão Tabela 10 – Matrizes de confusões médias para dados sintéticos e reais e diferentes classificadores

Taxa média certo <b>89,6±0,19%</b> para 8 gases					
Normal	<b>359,66</b>	3,92	0	0,02	0
T<700°C	1,04	<b>43,62</b>	2,34	5,80	0,06
T>700°C	0,53	16,76	<b>33,54</b>	5,80	0,06
DPBI	0,40	4,86	1,31	<b>69,60</b>	7,82
DAI	0	1,2	1,2	16,69	<b>63,71</b>

SVM					
3198 amostras, taxa média acerto <b>90,3±0,5%</b> para 5 gases, RBF kernel, $C_{reg}=0,32978$ , $\gamma = 0,0029782$ , OVO					
	Normal	T<700°C	T>700°C	DPBI	DAI
Normal	<b>362,43</b>	1,14	0,03	0	0
T<700°C	1,71	<b>34,69</b>	8,32	8,08	0
T>700°C	1,84	9,24	<b>43,41</b>	0,59	1,31
DPBI	2,60	1,20	0	<b>78,56</b>	1,64
DAI	0,01	0,62	2,08	15,08	<b>65,01</b>

Taxa média acerto <b>92,5±0,5%</b> para 8 gases					
Normal	<b>362,52</b>	0,68	0,20	0	0,20
T<700°C	0,80	<b>39,40</b>	9,4	2,92	0,28
T>700°C	1	7	<b>46,84</b>	0,56	1
DPBI	0,2	2,48	0,04	<b>72,64</b>	8,64
DAI	0	1,44	2	6,68	<b>72,68</b>

ANN MLP					
3198 amostras, taxa média acerto <b>94,6±1,1%</b> para 5 gases, 1 camada escondida, 11 neurônios, Levenberg Marquardt, tangente hiperbólica					
	Normal	T<700°C	T>700°C	DPBI	DAI
Normal	<b>363,23</b>	0,31	0,05	0,01	0
T<700°C	2,39	<b>44,46</b>	2,57	3,15	0,23
T>700°C	1,85	1,67	<b>51,34</b>	0,66	0,88
DPBI	2,95	2,29	0,05	<b>76,2</b>	2,51
DAI	1,12	0,33	2,07	7,06	<b>72,22</b>

Taxa média acerto <b>95,3±1,5%</b> para 8 gases					
Normal	<b>362,77</b>	0,33	0,41	0,27	0,07
T<700°C	2,15	<b>44,81</b>	2,81	1,97	1,06
T>700°C	0,81	1,09	<b>53,35</b>	0,71	0,43
DPBI	0,76	1,99	0,23	<b>77,50</b>	3,50
DAI	4,19	0,63	3,01	2,67	<b>72,29</b>

ELM					
3198 amostras, taxa média acerto <b>97,7±0,1%</b> para 5 gases, RBF kernel, $C_{reg}=4$ e $\gamma=0,00012207$					
	Normal	T<700°C	T>700°C	DPBI	DAI
Normal	<b>364,4</b>	0,01	0	0	0
T<700°C	1,36	<b>50,38</b>	0	1,17	0
T>700°C	3,58	0	<b>52,62</b>	0,28	0
DPBI	2,99	0,35	0	<b>80,78</b>	0,04
DAI	0,44	0	0	2,44	<b>80,17</b>

Taxa média acerto <b>97,9±0,1%</b> para 8 gases					
Normal	<b>363,93</b>	0,00	0,2	0,2	0,04
T<700°C	1,86	<b>50,88</b>	0	0,23	0
T>700°C	2,00	0	<b>54,52</b>	0	0
DPBI	5,07	0,71	0	<b>77,49</b>	0,9
DAI	0,45	0	0	0,90	<b>81,60</b>

Fonte: Elaborado pelo autor

Analisando a Tabela 10, verifica-se que o percentual de assertividade das falhas melhorou em relação aos conjuntos que possuem apenas dados reais (Tabelas 6 e 9). Isso denota que os dados gerados artificialmente, dentro da faixa de incertezas de medição da IEC 60599 (IEC, 2007), têm forte relação com os dados reais, ou seja, explica-se essa melhora devido à maior chance de sempre existir dados próximos (com maior similaridade) pertencentes à mesma classe no espaço de características em razão da maior quantidade de amostras geradas com base nos dados reais.

Essa forma de geração de dados funcionou bem nos algoritmos 1NN e OPF em razão da sua heurística de classificação, pois o percentual de acertos apresentados demonstra grande conectividade entre os dados reais e sintéticos de falhas, bem como as fronteiras desses algoritmos não faz menção à forma ou “separabilidade” das classes, como os algoritmos que utilizam hiperplanos de separação.

Quanto aos demais algoritmos, mesmo apresentando melhoria, deve-se observar que a geração dos dados artificiais pode levar a um possível aumento de amostras na fronteira de decisão entre diferentes classes, dificultando a separação entre elas como no caso nos algoritmos de ANN e SVM que usam hiperplanos como fronteiras de classificação. O aumento de dados artificiais de falhas, junto com a manutenção da quantidade de amostras normais, ocasionou mudanças nos percentuais de composição das diferentes classes e consequentemente na probabilidade (original e posterior) de ocorrência usada no algoritmo Nãive Bayes, principalmente individualizadas entre as próprias classes térmicas e as próprias classes elétricas devido às similaridades nas concentrações dos gases existentes (Tabela 10), logo, resultando em menor taxa de acerto, se comparado a outros algoritmos. Um meio de reduzir os erros de classificações decorrente do aumento de dados nas fronteiras do espaço de atributos, entre diferentes classes de padrões, é aplicar uma filtragem através do descarte das amostras artificiais classificadas erradamente.

### **5.3. Análise do histórico de DGA usando a abordagem desenvolvida e dados reais**

O Anexo B da norma C57.104 (IEEE, 2019) traz considerações gerais sobre o uso de DGA na avaliação de taxas de concentrações de gases. Os dados de DGA podem fornecer informações de tendências de mudanças nas condições de operacionais do equipamento, bem como pode ser útil para determinar os efeitos do carregamento e possíveis falhas em evolução neles. Entretanto, nesse método, é necessário estabelecer uma data base associada a uma frequência de análise (ou medição) adequada, para que seja possível observar as taxas de

variações que ocorrem nas concentrações de gases. Diagnósticos que observam a taxa de evolução dos gases necessitam de cuidados acerca do uso dos valores de concentrações entre amostragens, uma vez que tem impacto no resultado da análise, como por exemplo, influências de diferentes métodos de DGA entre laboratórios, qualidade da amostra, temperatura no momento da amostragem dos gases, entre outros. Em complemento, o crescimento elevado nas concentrações de gases observadas após medidas de confirmação de valores, determina um acompanhamento próximo do equipamento.

Nesse contexto, na metodologia de análise proposta, é recomendado definir frequências de análises adequadas, manter as qualidades das medidas e evitar a variação entre laboratórios, para que não sejam inseridas mais incertezas no processo de classificação. Nos transformadores analisados, a frequência de realização de análises para a condição normal de produção dos gases é de duas vezes ao ano, entretanto, no caso de elevação dos gases, ela é reduzida para medições mensais, semanais ou até diárias, dependendo do valor do crescimento. As variações bruscas que ocorrem nos gráficos das curvas de falhas apresentadas devem-se principalmente, ao grande intervalo de tempo existente entre as medições dos gases dissolvidos, e isso significa que para uma medição com intervalos menores e constantes, tendem a amenizar as variações bruscas, representando melhor a presença das falhas dos equipamentos.

Assim, a partir do conjunto das amostras de falhas reais, dos dados históricos de DGA dos transformadores de potência (da concessionária de energia) e dos seus respectivos relatórios de inspeções, foram realizadas simulações com os algoritmos OPF e 1NN, a fim de validar a metodologia proposta no diagnóstico desses equipamentos. A Tabela 11 relaciona os transformadores analisados, sendo observado que seus dados não fizeram parte do conjunto de treinamento (mais informações sobre outros transformadores analisados, consulte o Apêndice C).

Tabela 11 – Identificação dos transformadores de potência avaliados

Transformador Nº	Tensão (kV)	Potência (MVA)	Idade (Anos)	Quantidade DGA avaliadas (Amostras)
TR1	345/88	133,33	40	53
TR2	345/230	100	6 (reformado)	18
TR3	345/230	100	6 (reformado)	21
TR4	230/88	50	16	43
TR5	245/88	133,33	44	19
TR6	345/88	133,33	36	25

Fonte: Elaborado pelo autor

### 5.3.1. Transformador TR1

A Figura 15(a) mostra o registro de concentrações dos gases durante o serviço do transformador, compreendido entre 04/08/2000 a 30/09/2019. Na data de 01/04/2012 foi realizado o tratamento do óleo, para melhorar suas características físico-químicas e extrair os gases dissolvidos. Após retorno à operação, com a evolução dos gases a partir da análise de 10/11/2014, foi reduzido o intervalo entre as suas análises de semestral para trimestral, e após 08/12/2016 para mensal, em razão da elevada concentração dos gases encontrada. Somente após 09/04/2017 foi feita a inspeção interna do equipamento pela concessionária de energia.

Analisando-se os gráficos de evolução de falhas (Figuras 15(b) e (c)) nas datas de 04/08/2000 e 01/04/2012, a estabilidade e baixa concentração dos gases para as amostras submetidas nos algoritmos INN e OPF foram classificadas como normais (menor custo) e distante da classe falha em vista do intervalo existente entre ambas as curvas. Após o tratamento do óleo, com o crescimento das concentrações dos gases ( $\text{CH}_4$  e  $\text{C}_2\text{H}_4$ , típicos de falha térmica), a partir de 06/01/2014 é possível observar a aproximação entre as curvas normal e falha, demonstrando uma tendência na mudança de estado com a evolução dos gases, o que acabou ocorrendo entre as análises de 06/01/2014 e 10/11/2014 até a inspeção interna do transformador.

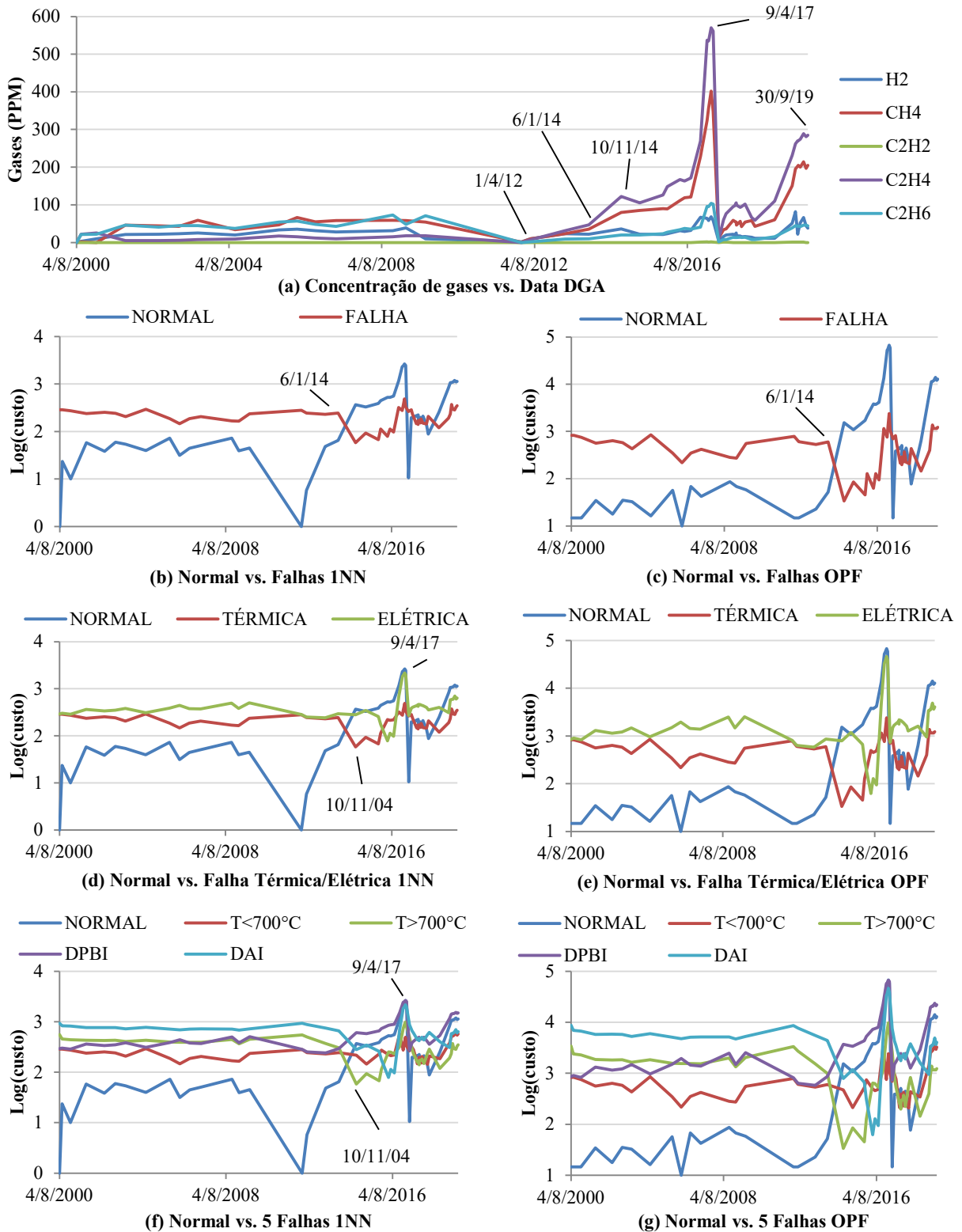
Como complemento das Figuras 15(b) e (c), os gráficos das Figuras 15(d) e (e) permitem reconhecer as características do grupo de falha envolvido, isto é, se a origem é térmica ou elétrica. Observa-se que houve a presença dos dois grupos de falhas (térmica e elétrica) entre o período de 10/11/2014 a 09/04/2017. Nesse período de sinalização de falha, os principais gases produzidos foram  $\text{CH}_4$  e  $\text{C}_2\text{H}_4$ , pequena quantidade de  $\text{H}_2$  e  $\text{C}_2\text{H}_6$ , e traços de  $\text{C}_2\text{H}_2$ , característicos de falha térmica ou elétrica, dependendo das combinações das concentrações envolvidas.

Os gráficos das Figuras 15(f) e (g) complementam todas as anteriores atribuindo um rótulo específico para cada tipo de falha, ou seja, é possível observar características das curvas de normalidade, falhas térmicas ( $T < 700^\circ\text{C}$  e  $> 700^\circ\text{C}$ ), descargas elétricas de baixa (DPBI) e descargas de alta intensidade (DAI), para o período de operação do transformador. Assim, no período de 10/11/2014 a 09/04/2017, são indicados respectivamente  $T > 700^\circ\text{C}$ , DAI e  $T < 700^\circ\text{C}$ , conforme evolução das concentrações dos gases. Em continuidade à análise gráfica, observe que mesmo após a manutenção da falha e retorno do equipamento à operação em 31/05/2017, há de se considerar um período de homogeneização dos gases residuais impregnados no papel e posterior normalização, quando então 17/11/2018 novamente, há um



crescimento dos gases e nova sinalização de falha ( $T > 700^{\circ}\text{C}$ ), nesse caso, como procedimento o transformador é mantido ainda em estado de observação, com intervalos reduzidos de DGA para avaliar o momento de inspeção, caso não haja estabilização dos gases.

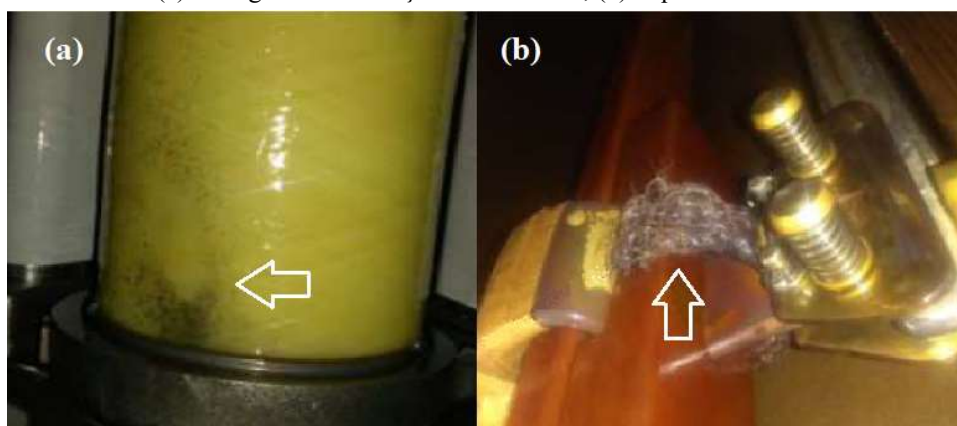
Figura 15 – Curvas de concentrações de gases e tendências de falhas no transformador TR1



Fonte: Elaborado pelo autor

Na inspeção interna (abertura) do equipamento, após 09/04/2017, foi encontrado na chave comutadora sem carga, a presença de partículas carbonizadas de óleo aderidas à superfície do cilindro de acionamento (Figura 16(a)), entretanto não se notou um ponto de ocorrência específico em seu entorno. Também encontrou-se sinal de aquecimento da cordoalha de ligação de interconexão entre as bobinas com alteração na cor do metal (Figura 16(b)), ou seja, falhas características de descargas elétricas de baixa intensidade e de sobreaquecimento  $T > 700^{\circ}\text{C}$ , respectivamente. Associando as evidências das falhas encontradas no transformador com o método proposto nesta tese, o problema de aquecimento encontrado provavelmente iniciou-se com a evolução dos gases característicos desse tipo de falha, a partir de 10/11/2014, conforme mostram as Figuras 15(b) a (g). Apesar de também haver indícios de ocorrências de descargas elétricas de baixa intensidade, como apresentados nos gráficos de falhas, a classificação obtida de alta intensidade (DAI), foi tratada como erro de classificação, ao qual os algoritmos estão sujeitos (Tabela 10). Para a falha térmica encontrada em 08/12/2016 e que antecedeu a inspeção foi  $T > 700^{\circ}\text{C}$ , migrou para  $T < 700^{\circ}\text{C}$  na data de inspeção por mudança nas concentrações dos gases  $\text{CH}_4$  e  $\text{C}_2\text{H}_4$ , levando novamente ao erro de classificação.

Figura 16 – TR1: (a) Vestígios carbonização óleo cilindro; (b) Aquecimento cordoalha comutador



Fonte: Adaptado pelo autor de concessionária de energia

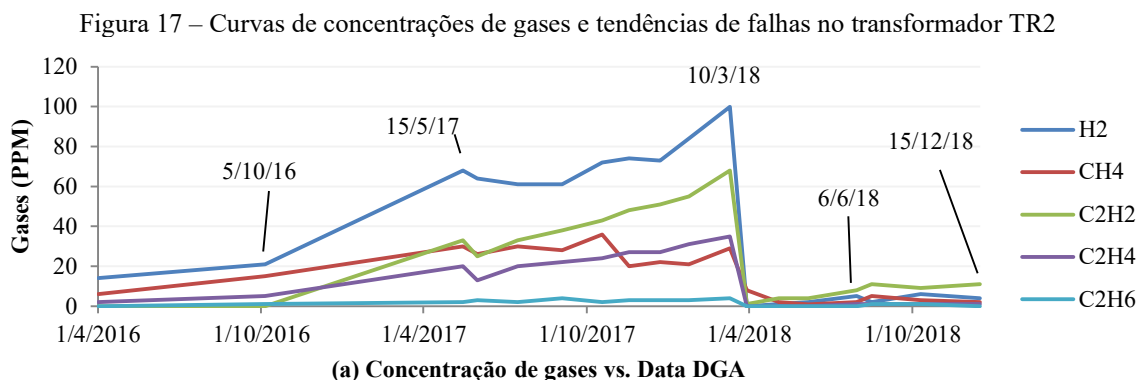
Todavia, esse método permite observar as proximidades entre os diferentes tipos de falhas e assim, extrair informações de outras possibilidades de ocorrência com base na relação com as concentrações de gases analisadas, ou seja, no intervalo entre 10/11/2014 a 09/04/2017, a falha principal (primária, pois, possui menor custo)  $T > 700^{\circ}\text{C}$  tem proximidade com a falha secundária (possui o segundo melhor custo)  $T < 700^{\circ}\text{C}$ , no trecho com falha elétrica principal DAI há proximidade com as falhas térmicas  $T < 700^{\circ}\text{C}$  e  $T > 700^{\circ}\text{C}$  (alternam-se as posições), e próximo da abertura, a falha primária é  $T < 700^{\circ}\text{C}$  e secundária  $T > 700^{\circ}\text{C}$ ,

portanto tornando assim mais informativo a análise gráfica com o evento real encontrado, ou seja, mostrando outras possibilidades de ocorrência.

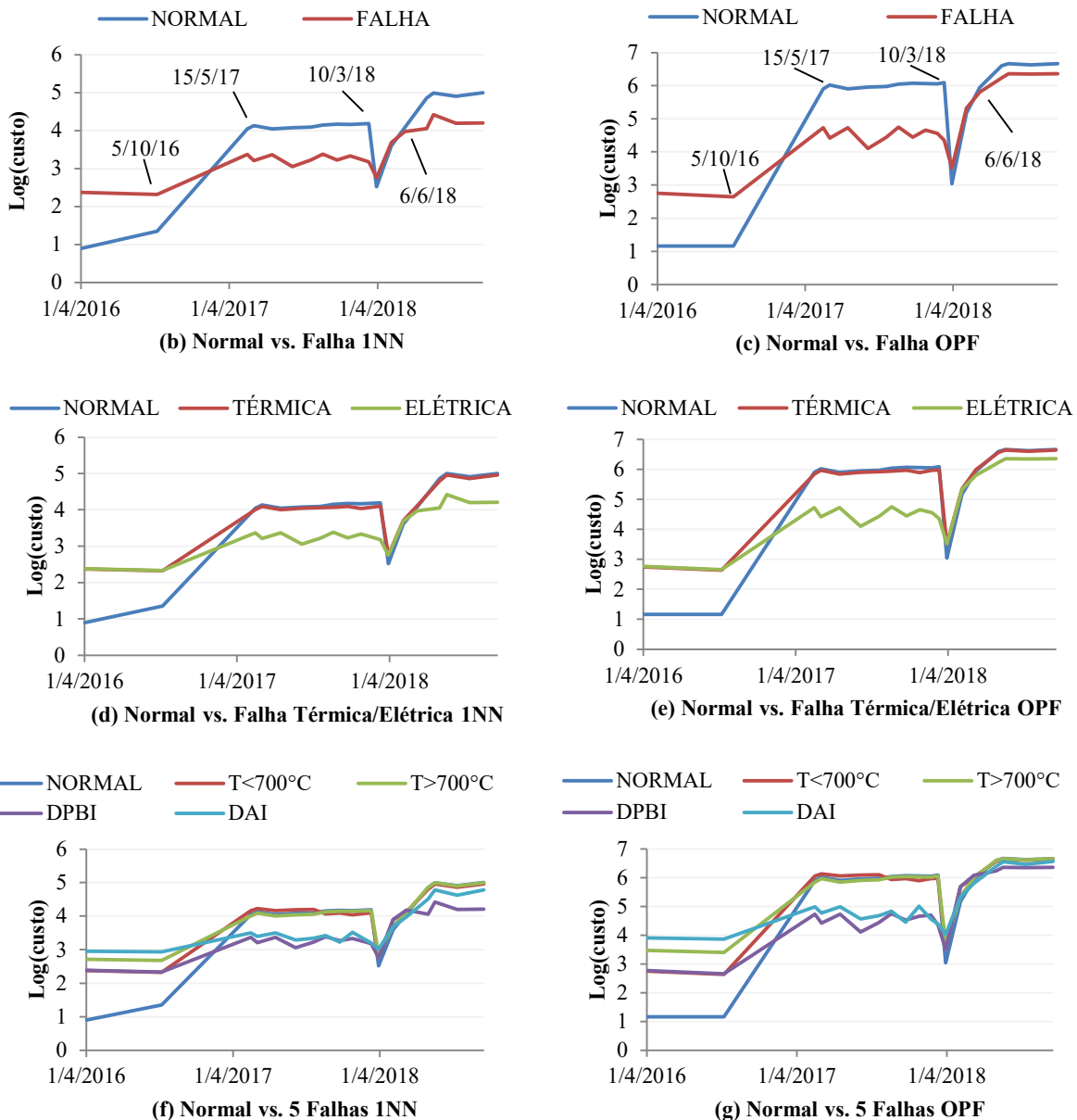
Em resumo, com o gráfico de dois estados (Figuras 15(b) e (c)) foi possível diagnosticar corretamente a tendência e sinalização de presença de falha com a evolução dos gases (Figura 15(a)), no gráfico de três estados, identificou-se que as falhas possuíam natureza térmica em sua maior ocorrência e também elétrica (Figuras 15(d) e (e)), e no gráfico com cinco estados as falhas primárias específicas em alguns momentos divergiram da real encontrada no período analisado, porém seguidas de informações que permitem conhecer a proximidade com outros tipos de falhas (por exemplo, falha secundária). Comparando-se às curvas de evoluções de falhas em ambos algoritmos 1NN e OPF, observa-se grande semelhança no comportamento, sendo encontrado pequenas diferenças em razão das diferentes heurísticas de classificação dos algoritmos.

### 5.3.2. Transformador TR2

A Figura 17(a) mostra o registro de concentrações de gases durante o serviço do transformador compreendido entre 01/04/2016 a 15/12/2018. O histórico deste transformador relata uma reforma geral com troca dos enrolamentos e melhoria do projeto em 2012, e desde então, operava normalmente no sistema elétrico. A partir das amostras de 15/05/2017, foi detectada a elevação dos gases, determinando a redução do intervalo de acompanhamento semestral para mensal, e somente após 10/03/2018, foi feita a inspeção interna do equipamento pela concessionária de energia.



Conclusão Figura 17 – Curvas de concentrações de gases e tendências de falhas no transformador TR2



Fonte: Elaborado pelo autor

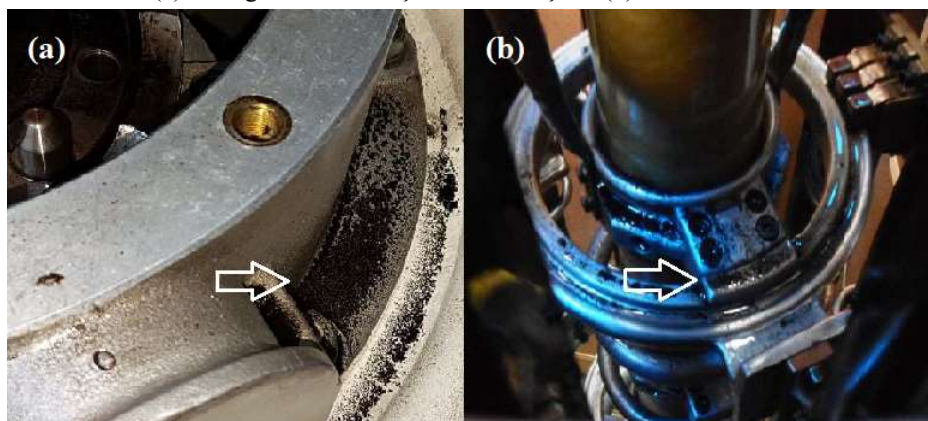
Analisando-se os gráficos de evoluções de falhas (Figuras 17(b) e (c)) para as amostras submetidas à classificação dos algoritmos 1NN e OPF, observa-se a condição de operação normal do transformador entre 01/04/2016 e 05/10/16 determinado pela baixa concentração dos gases  $H_2$ ,  $CH_4$  e  $C_2H_2$ , no entanto, nesse período há um leve aumento dos gases, levando a uma tendência de aproximação entre as curvas normal e falha.

Na análise seguinte de DGA em 15/05/2017, observa-se um aumento expressivo dos gases citados acrescido do gás  $C_2H_2$ , típicos de descargas elétricas de baixa ou alta intensidade, dependendo das combinações e concentrações encontradas com outros gases. Nessa mesma data, passa a haver indicação de falha, como mostram as Figuras 17(b) e (c),

que se mantém até 10/3/2018, antes da inspeção interna. Nas Figuras 17(d) e (e), observa-se que em todo o intervalo da presença de falha está caracterizada como elétrica, assumindo os rótulos de descargas parciais de baixa ou de alta intensidade (Figuras 17(f) e (g)), conforme os gases principais  $H_2$  e  $C_2H_2$  crescem.

Na inspeção (abertura) do equipamento após 10/3/2018, sob a suspeita de falha de descarga de baixa/alta intensidade, foram encontrados no comutador sem carga (mais informações consulte o Apêndice A.1) vestígios de carbonização do óleo, característicos dessas falhas, depositados sobre o mecanismo de manobra e nos anéis de contato (Figuras 18(a) e (b)).

Figura 18 – TR2: (a) Vestígios carbonização óleo cabeçote; (b) Acionamento chave comutadora



Fonte: Adaptado pelo autor de concessionária de energia

Tal falha foi relacionada à equalização deficiente do mecanismo de acionamento do comutador sem carga, porém não houve danos ou maiores consequências com o equipamento, sendo providenciada apenas a equalização adequada do dispositivo. Associando-se às evidências das falhas encontradas no transformador com o diagnosticado pelo método proposto nesta tese, o problema de descargas elétricas mostrou-se compatível com o encontrado e se iniciou antes de 15/05/2017, com a evolução dos gases característicos desse tipo de falha.

Nas Figuras 17(f) e (g), observa-se uma alternância entre a falha primária e secundária entre as descargas elétricas, prevalecendo uma maior ocorrência de baixa intensidade. Após correção da falha e retorno à operação (em 29/03/2018), as Figuras 17(b) a (g) passam a indicar novamente falha elétrica (DPBI), a partir de 06/06/2018 (dois meses após a manutenção) em razão do crescimento do gás  $C_2H_2$ . A expectativa é que junto com a baixa concentração de gases e com a estabilização na evolução de  $C_2H_2$ , passe a haver uma redução gradativa dos gases. Nesse período, o transformador deve ser mantido em estado de

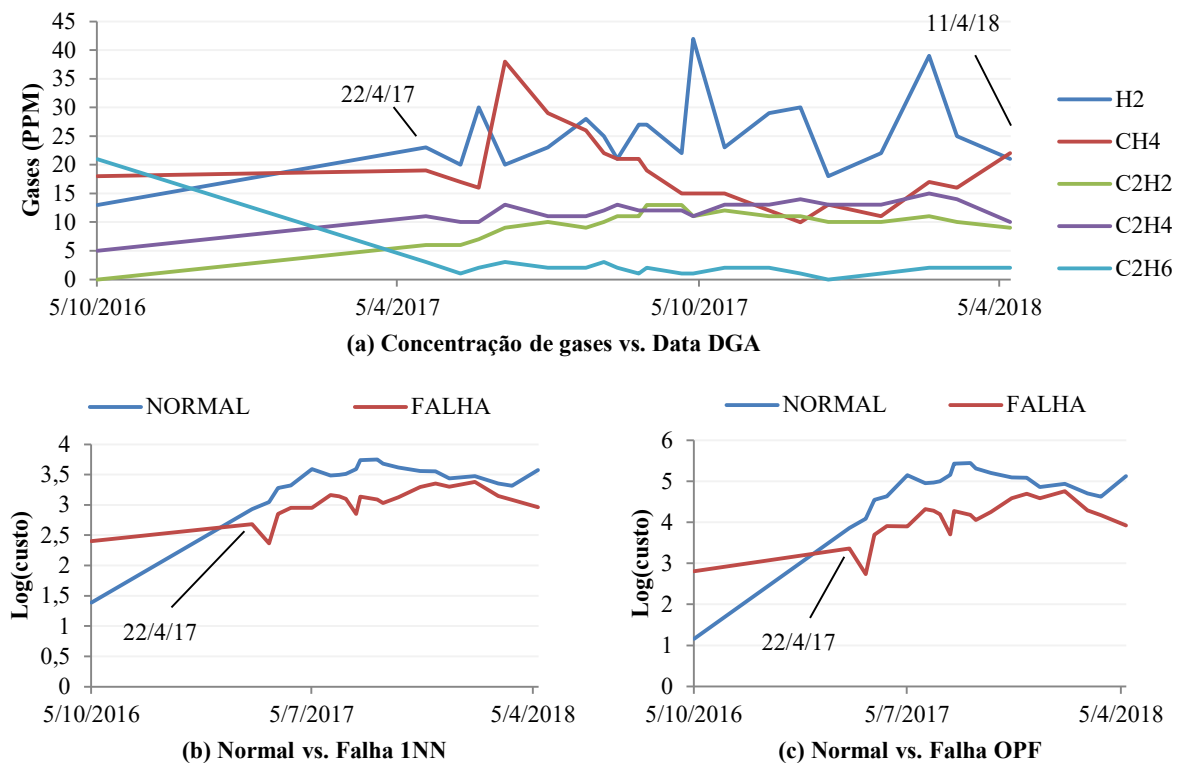
observação, porém, persistindo a sinalização em conjunto com o crescimento dos gases, uma nova inspeção deverá ser providenciada.

Comparando-se as curvas de evoluções de falhas nos algoritmos INN e OPF, observa-se que há pequenas diferenças no comportamento, sem afetar o resultado geral, isto se deve às diferentes heurísticas de classificação dos algoritmos, sendo mais suave o OPF nas fronteiras de diferentes classes.

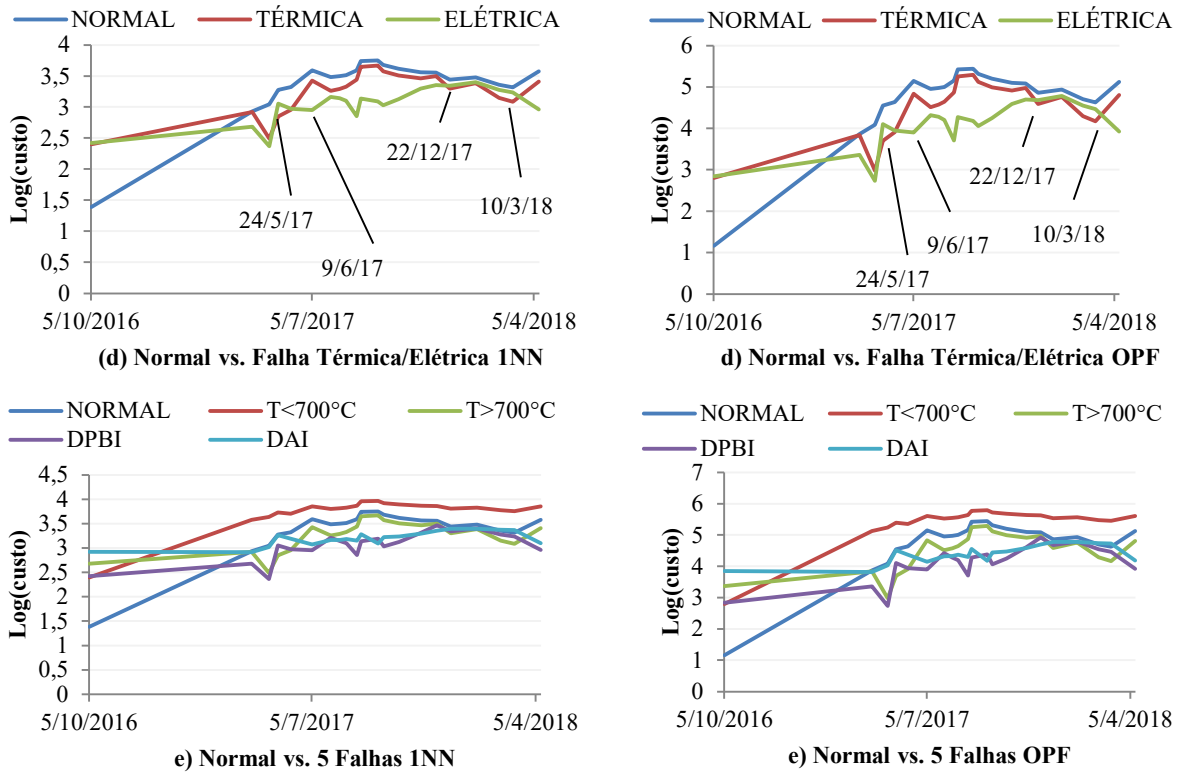
### 5.3.3. Transformador TR3

A Figura 19(a) mostra o registro de concentrações de gases durante a operação do transformador compreendida entre 05/10/2016 a 11/04/2018. Modelo similar ao TR2, esse transformador revela em seu histórico que foi reformado em 2003, para corrigir um problema com a blindagem do fluxo magnético de dispersão e desde então, operava normalmente no sistema elétrico. A partir das amostras de 22/04/2017, foi detectada a elevação de alguns gases críticos, determinando a redução do intervalo entre análises de semestral para mensal, e somente após 11/04/2018, foi realizada a inspeção interna do equipamento pela concessionária de energia.

Figura 19 – Curvas de concentrações de gases e tendências de falhas no transformador TR3



Conclusão Figura 19 – Curvas de concentrações de gases e tendências de falhas no transformador TR3



Fonte: Elaborado pelo autor

Analisando-se os gráficos de evoluções de falhas (Figuras 19(b) e (c)) para as amostras submetidas à classificação dos algoritmos 1NN e OPF, em 22/04/2017, observa-se um leve crescimento dos gases  $H_2$ ,  $C_2H_4$  e  $C_2H_2$ , típicos de descargas elétricas de baixa ou alta intensidade, dependendo das combinações e concentrações encontradas. Nessa mesma data, surge também a indicação de falha como mostra as Figuras 19(b) e (c) e que se manteve até 11/04/2018, antes da inspeção interna.

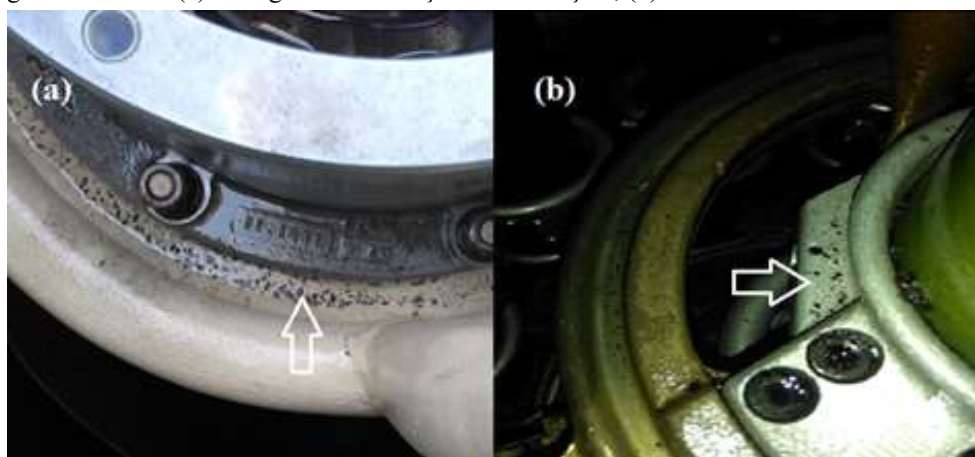
Nas Figuras 19(d) e (e), observa-se que no intervalo entre 22/04/2017 e 11/04/2018, há sinalizações de falha elétrica no início e fim desse período, e falha térmica em dois intervalos no entremeio da elétrica, ou seja, entre 24/05/2017 a 09/06/2017, e de 22/12/2017 a 10/03/2018, o que significa que as amostras de DGA com baixa concentração dos gases encontram-se numa região com amostras de ambas as classes próximas.

No reconhecimento dos tipos específicos de falhas das curvas das Figuras 19(f) e (g), as amostras rotuladas no período foram identificadas como descargas parciais de baixa e alta intensidade, e térmica  $T > 700^\circ C$ , variando as posições (primária, secundária, outras) entre si, conforme variam as diferentes concentrações de gases, envolvendo  $H_2$ ,  $CH_4$ ,  $C_2H_4$  e  $C_2H_2$ .

Na inspeção do equipamento após 11/4/2018, sob a suspeita de falha de descarga de baixa/alta intensidade, foram encontrados no comutador (sem carga), vestígios de

carbonização do óleo característicos desses tipos de falhas depositados sobre o mecanismo de manobra e nos anéis de contato (Figuras 20(a) e (b)) similar ao TR2 e, portanto, a solução adotada para correção foi a equalização do mecanismo de acionamento da chave comutadora similar ao transformador TR2.

Figura 20 – TR3: (a) Vestígios carbonização óleo cabeçote; (b) Acionamento chave comutadora



Fonte: Adaptado pelo autor de concessionária de energia

Associando-se às evidências das falhas reais encontradas com o diagnosticado pelo método proposto nesta tese, observa-se a indicação de falha térmica e elétrica alternando-se em alguns momentos do período de acompanhamento do estado do transformador, contudo, verifica-se um período de incidência elétrica maior (entre DPBI como primária e DAI como secundária ou vice-versa) no início da detecção da falha e próximo da inspeção, indicando maior possibilidade de ocorrência.

Considerando o momento da abertura após 11/04/2018, as evidências das falhas encontradas no transformador mostraram-se compatíveis com indicação da falha primária DPBI e secundária DAI. Observa-se que nos momentos em que houve as indicações de falha térmica  $T > 700^{\circ}\text{C}$  como primária, ela era acompanhada da falha secundária DPBI, e essa combinação tem relação com as variações (elevação e queda), principalmente dos gases  $\text{CH}_4$  e  $\text{H}_2$ . No período de sinalização de falha térmica entre 22/12/2017 e 21/03/2018, existem quatro classes com custos próximos com concentrações dos gases praticamente estáveis, e qualquer pequena mudança no valor de um dos gases pode modificar a condição da sinalização, visto que a posição da amostra também se altera no espaço de atributos.

Comparando-se as curvas de evoluções de falhas dos algoritmos INN e OPF, observa-se uma grande semelhança no comportamento dos gráficos.



#### 5.3.4. Transformador TR4

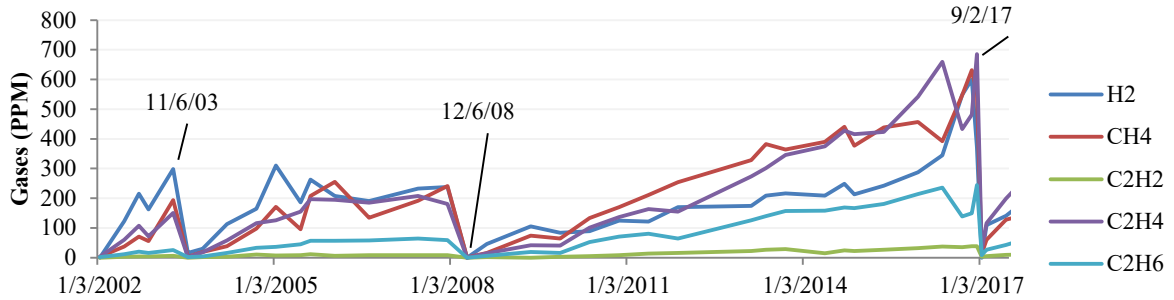
A Figura 21(a) mostra o registro de concentrações dos gases durante o período de serviço do transformador compreendido entre 25/03/2002 a 06/09/2018, em que é possível identificar três momentos de intervenções caracterizados por quedas bruscas nas concentrações dos gases: após 11/06/2003, quando houve a substituição das bobinas (queima) e troca do óleo por novo; em 12/06/2008, no tratamento termo vácuo e filtragem do óleo, após análise físico-química determina-la como inadequada; e após 09/02/2017, quando houve o crescimento contínuo da concentração dos gases e consequente suspeita de falha que levou à inspeção do equipamento por parte da concessionária de energia.

Analisando-se os gráficos de evoluções de falhas (Figuras 21(b) a (g)) para as amostras submetidas à classificação dos algoritmos 1NN e OPF, observa-se que logo após quatro meses do início de operação do transformador, a partir de 25/03/2002, houve crescimento dos gases  $H_2$ ,  $CH_4$ ,  $C_2H_4$  e aparecimento de traços de  $C_2H_2$  característicos de falha térmica ( $T > 700^\circ C$ ) ou elétrica (DPBI), dependendo das concentrações encontradas, tendo como consequência a indicação de falha (Figuras 21(b) e (c)), até a ocorrência da inspeção e manutenção após 11/06/2003.

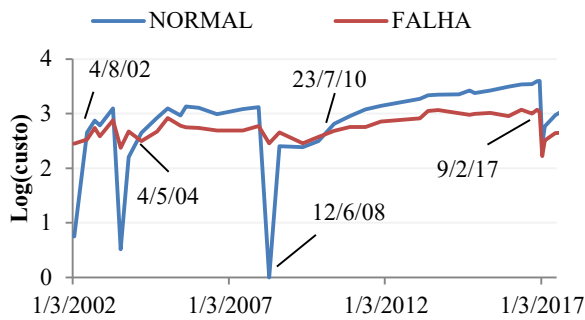
Nas Figuras 21(d) a (g), é possível identificar que a sinalização trata-se de falha térmica, especificamente primária  $T > 700^\circ C$  e secundária  $T < 700^\circ C$ , muito embora a concentração principal de  $H_2$  sobre as demais seja característica de DPBI. Em 28/09/2003, após sete meses de retorno ao serviço, houve o crescimento dos mesmos gases presentes na falha anterior e com perfil inicialmente similar, porém estabilizando-se a partir de 04/08/2005, até o tratamento do óleo. Nesse período, a partir de 04/05/2004, ocorreu também a indicação de falha térmica primária  $T > 700^\circ C$  e secundária  $T < 700^\circ C$  (Figuras 21(f) e (g)).

Posteriormente ao tratamento do óleo, o transformador trabalhou normalmente até antes de 23/07/2010, quando se iniciou o crescimento contínuo novamente de gases dos mesmos tipos das falhas anteriores, porém acrescidos de  $C_2H_6$ , até a inspeção do equipamento após 09/02/2017. Em consequência desse aumento na concentração dos gases, a partir de 23/07/2010, os gráficos de falhas (Figuras 21(b) a (g)) também passaram a sinalizar falha térmica primária  $T > 700^\circ C$  e secundária  $T < 700^\circ C$  até 14/07/2014, quando então houve uma inversão das posições entre elas. Posterior à manutenção do equipamento e volta à operação, a sinalização de falha térmica permaneceu presente e, portanto, levando o transformador ao estado de observação, para acompanhar a homogeneização dos gases e a saída da sinalização de falha.

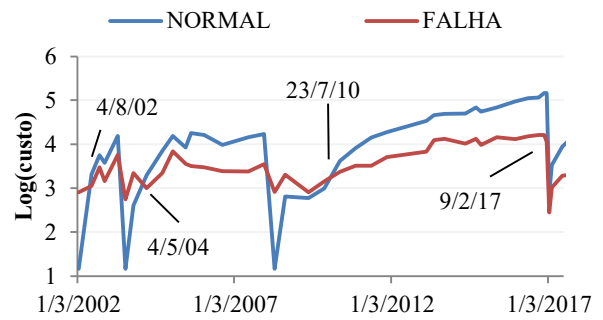
Figura 21 – Curvas de concentrações de gases e tendências de falhas no transformador TR4



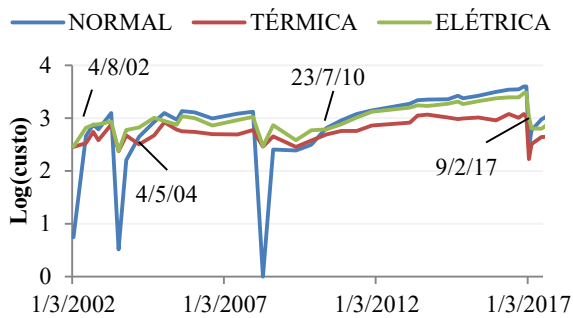
(a) Concentração de gases vs. Data DGA



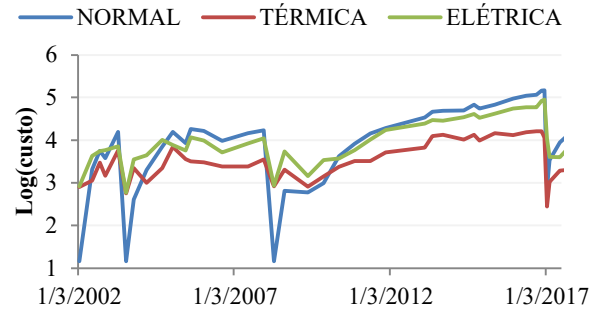
b) Normal vs. Falhas INN



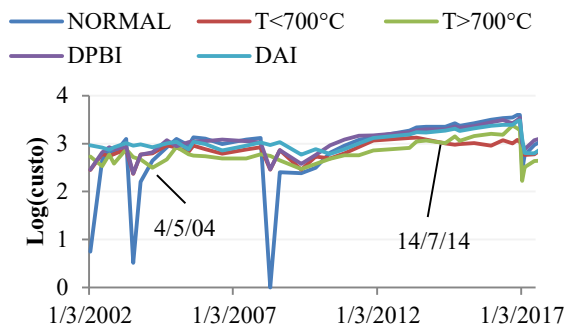
c) Normal vs. Falhas OPF



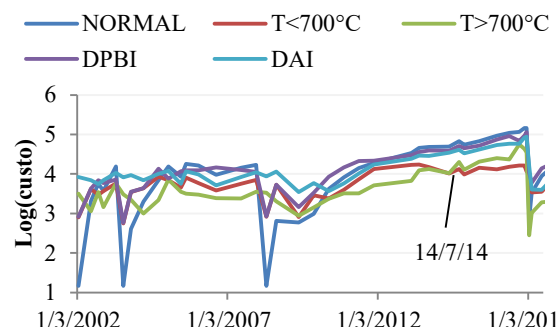
(d) Normal vs. Falha Térmica/Elétrica INN



(e) Normal vs. Falha Térmica/Elétrica OPF



f) Normal vs. 5 Falhas INN



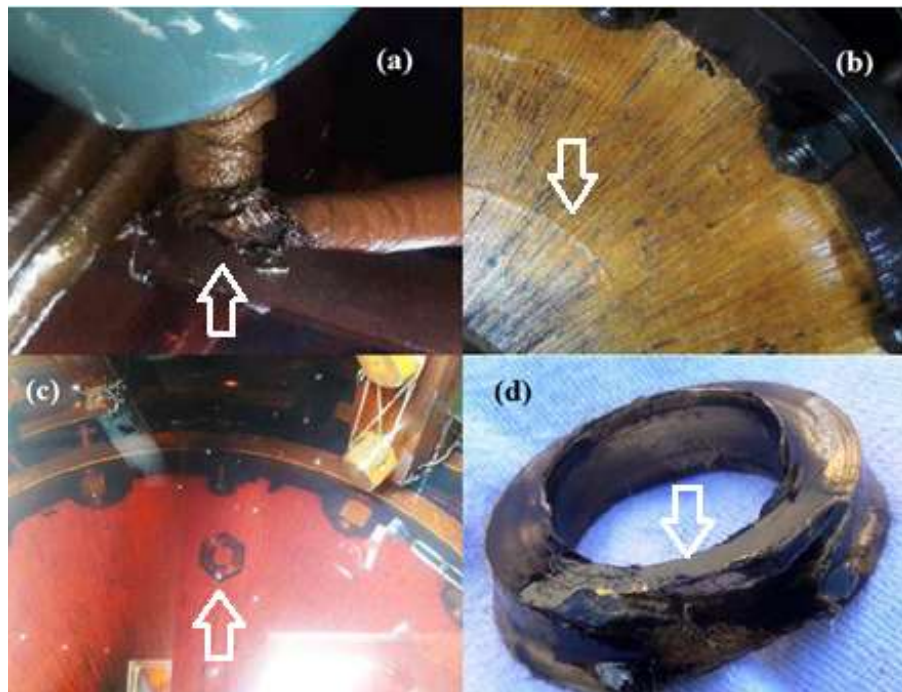
g) Normal vs. 5 Falhas OPF

Fonte: Elaborado pelo autor

Desde a entrada em operação do transformador, houve registros do equipamento ser ruidoso e que vibrava muito. Na inspeção interna (abertura) do equipamento, após 09/02/2017, sob a suspeita de falha térmica, foram encontrados diversos problemas associados

ao crescimento das concentrações de gases e que devem ter relação com o excesso de vibração do equipamento. Dentre os principais problemas encontrados, foram: a Figura 22(a) mostra a degradação da isolação (papel) em uma (de duas) das cordoalhas que apresentaram falha (buchas X1 e X2); a Figura 22(b) mostra a presença de sedimentos (metálico e papel carbonizado) depositados sobre o núcleo, que também foram encontrados sobre outros elementos internos; e as Figuras 22(c) e (d), que mostram respectivamente uma porca caída sobre o núcleo e sinal de fusão metálica da arruela do parafuso de prensagem do núcleo. Todos esses problemas têm relação com a falha térmica de média e alta temperatura, o que explicaria a presença de  $C_2H_2$ .

Figura 22 – TR4: (a) Isolação degradada; (b) Sedimentos; (c) Porca solta; (d) Fusão metal



Fonte: Adaptado pelo autor de concessionária de energia

Associando as evidências das falhas encontradas no transformador com o método de diagnóstico proposto nesta tese, na primeira ocorrência, após a manutenção de 11/06/2003, consta apenas a informação de que as bobinas foram substituídas por motivo de queima, ou seja, sem registro do tipo de falha encontrado (se a origem era elétrica ou térmica). Nesse período, os gráficos (Figuras 21(b) a (g)) sinalizaram a presença de falha térmica primária  $T > 700^\circ\text{C}$  que, mesmo em caso de erro de classificação em detrimento à possível falha elétrica, é considerada falha grave e de resposta urgente no atendimento.

Na segunda ocorrência antes do tratamento do óleo, entre 04/05/2004 e 12/06/2008 houve sinalização de falha térmica primária  $T > 700^\circ\text{C}$  (Figuras 21(d) e (e)) na presença dos

gases  $H_2$ ,  $CH_4$ ,  $C_2H_4$  e traços  $C_2H_2$ . Na terceira ocorrência de manutenção, após 12/06/2008, houve aumento expressivo de  $CH_4$ ,  $C_2H_4$ ,  $C_2H_2$  e  $C_2H_6$  (que, além de  $CO$  e  $CO_2$ , possui relação com aquecimento e carbonização do papel) acompanhado das sinalizações alternadas entre falhas térmicas  $T > 700^\circ C$  e  $T < 700^\circ C$ , entre 23/07/2010 até 09/02/2017. Nesse contexto de geração de gases, pressupõe-se que a falha térmica nas cordoalhas das buchas seja menos recente, uma vez que envolve a carbonização do papel e tenha ocorrido no segundo período antecedente à última manutenção, pois observou-se a estabilização das concentrações dos gases  $H_2$ ,  $CH_4$ ,  $C_2H_4$ ,  $C_2H_2$  (traço) com maior valor de  $CO_2$  no período, e a falha que envolve aquecimento e fusão (ponto quente) tenha se originado após do tratamento do óleo em função das características encontradas nos elementos e que contribuíram também para o crescimento dos mesmos gases no terceiro período, porém acompanhados de presença evidente de  $C_2H_2$ .

Relata-se que em novembro de 2015 foi realizado ensaio de emissão acústica para verificar a presença de descargas elétricas no interior do transformador, mediante as evoluções dos gases, porém, nada de anormal foi encontrado e, portanto indicando que as falhas anteriores a essa data provavelmente, não tinham origem elétrica.

Note que, em todos os casos, é possível associar o crescimento dos gases com a evolução e existência de falhas, sendo que nos dois últimos períodos de sinalização de falha, há certa compatibilidade entre aquelas sinalizadas com o encontrado na inspeção. Após a manutenção do equipamento e retorno ao serviço, observa-se ainda a sinalização presente de falha térmica  $T > 700^\circ C$ , necessitando de um acompanhamento próximo, até certificar-se da homogeneização dos gases no óleo e fim da sinalização, caso contrário, uma nova inspeção deve ser realizada. Para o transformador analisado neste caso, é possível verificar também o comportamento de classificação entre INN e OPF bastante similar.

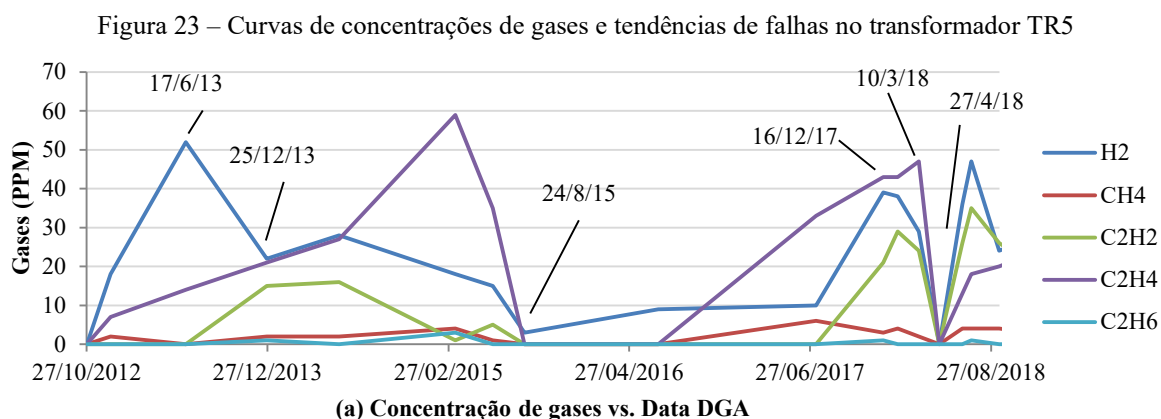
#### 5.3.5. Transformador TR5

A Figura 23(a) mostra o registro de concentrações dos gases durante a operação do equipamento compreendida entre 27/10/2012 a 15/09/2018. Na data de 24/08/2015 foi realizado um procedimento de substituição das buchas de alta e baixa tensão, e aproveitou-se para fazer a inspeção interna do equipamento, em virtude da evolução dos gases apresentada após 27/10/2012. Posteriormente, com retorno ao serviço, a partir da análise de 16/12/2017 iniciou-se o crescimento dos gases e então foi reduzido o intervalo semestral de análises de DGA para mensal, e somente após a data de 10/03/2018 foi feita a inspeção interna do equipamento pela concessionária de energia.

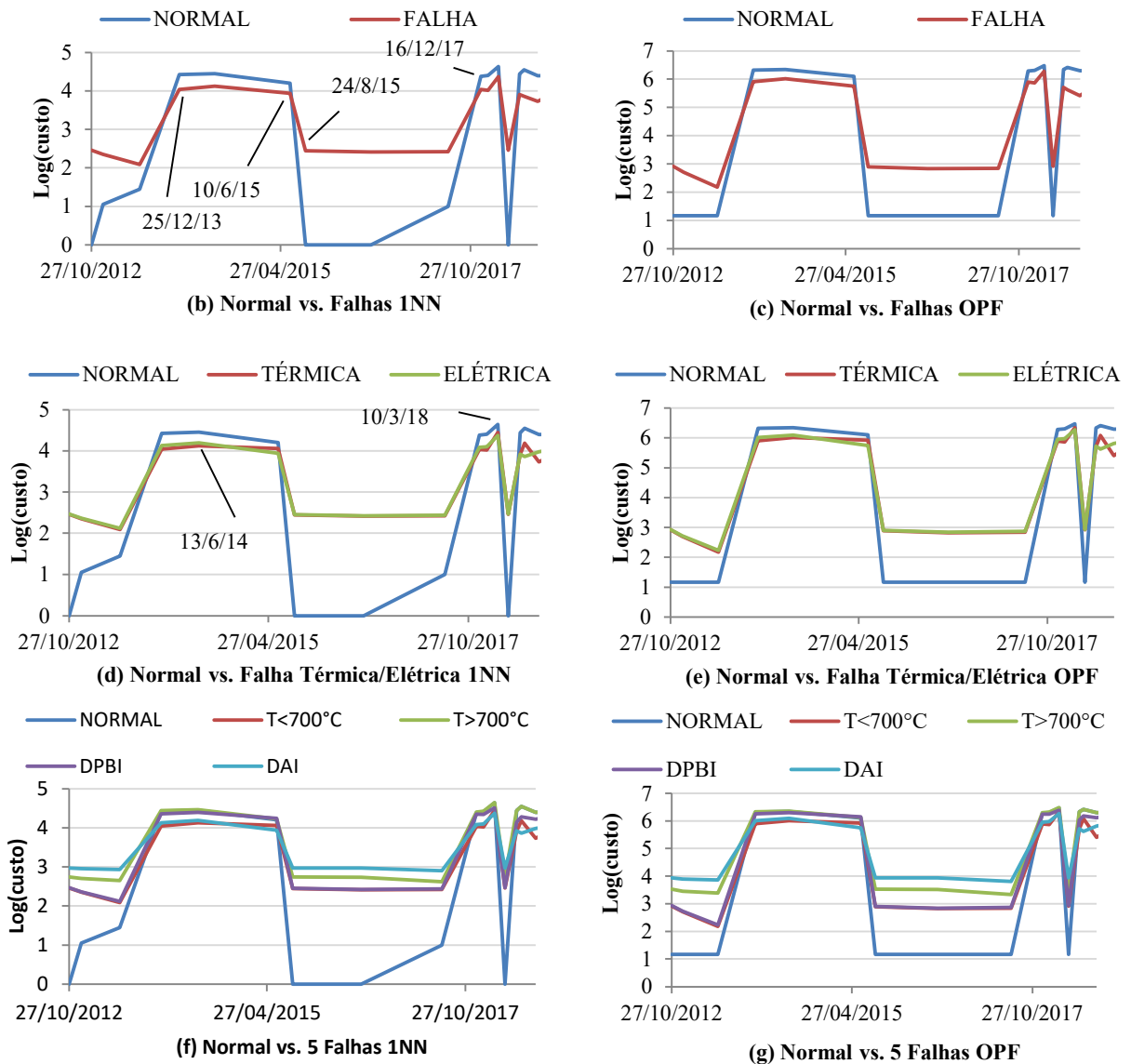
Analisando-se o gráfico de evoluções de gases (Figura 23(a)) para as amostras submetidas à classificação dos algoritmos 1NN e OPF, em 17/06/2013, verifica-se um aumento na concentração dos gases  $H_2$ ,  $C_2H_4$  e traços de  $CH_4$ , que são típicos de falha elétrica ou térmica, dependendo da combinação e concentração encontrada, porém a partir de 25/12/2013, é detectado a presença de  $C_2H_2$  (produzido a partir de  $T > 500^\circ C$  e associado as falhas mais severas), acompanhado da indicação de falha presente (Figuras 23(b) e (c)), que permaneceram até a data de 24/08/2015, quando ocorreu a manutenção das buchas de alta e baixa tensão.

Nesse período de sinalização, são indicadas falhas térmica (Figuras 23(d) e (e)) do tipo  $T < 700^\circ C$  e elétrica do tipo DAI (Figuras 23(f) e (g)), sendo a última, sinalizada em um momento pontual de composição das quantidades de concentrações de gases encontrada, ou seja, embora haja sinalização de falha elétrica DAI, as concentrações de todos esses gases (em 10/06/2015 com  $C_2H_4$  relevante,  $H_2$  média relevância e traços de  $C_2H_2$ ) encontradas, aparentavam ser características de falha térmica. Após manutenção e retorno ao serviço normal, em 16/12/2017 é detectado nova elevação nas concentrações dos gases junto com a indicação de falha térmica primária  $T < 700^\circ C$  e falha elétrica DAI secundária, que novamente alternam suas posições antes da abertura em 10/03/2018. Observa-se que por razão similar, as concentrações dos gases serem próximos ao caso anterior, seja a causa novamente de indicação de falha elétrica.

Após correção da falha em 27/04/18, verifica-se variação dos gases e sobreposições de classes falhas, entretanto, isso remete à questão do acompanhamento dos gases impregnados no papel e sua homogeneização com óleo, bem como na manutenção realizada. Nesse caso, o acompanhamento com um período de amostragem menor, deve ser feito até constatar sua normalização, caso contrário, nova inspeção deve ser programada.



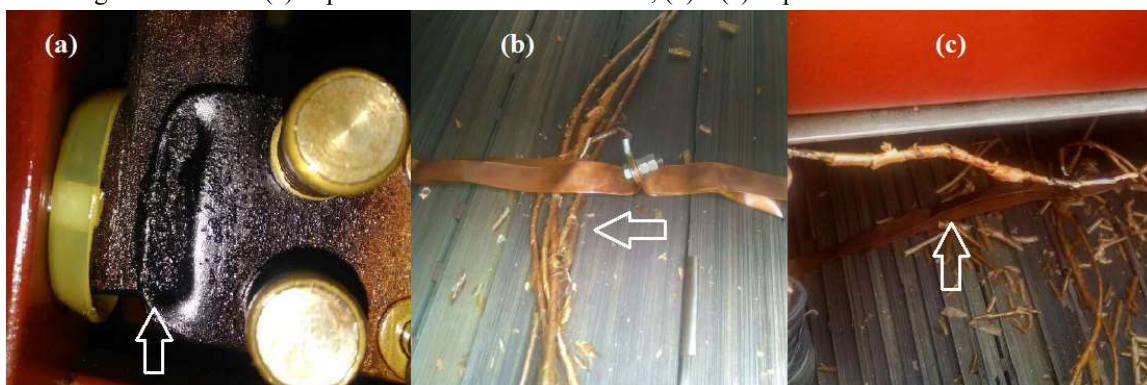
Conclusão Figura 23 – Curvas de concentrações de gases e tendências de falhas no transformador TR5



Fonte: Elaborado pelo autor

Aproveitando-se de uma parada programada para manutenção das buchas do transformador em 24/08/2015, investigou-se a elevação dos gases  $H_2$ ,  $C_2H_4$  e  $C_2H_2$  que havia ocorrido em 27/10/2012 e foi encontrado o aquecimento dos contatos da chave comutadora (típico de temperatura  $300^\circ C < T < 700^\circ C$ ) sem carga (Figura 24(a)), como responsável pela elevação dos gases, e a solução adotada para correção da falha foi o *by-pass* da chave comutadora.

Figura 24 – TR5: (a) Aquecimento contato comutador; (b) e (c) Aquecimento cabo aterramento



Fonte: Adaptado pelo autor de concessionária de energia

Na segunda inspeção, após elevação dos gases em 16/12/2017, foram verificados os elementos internos (bobinas, macaco de prensagem, entre outros) e encontrou-se o papel isolante dos cabos de aterramento do núcleo “quebradiço” e com partes carbonizadas (Figuras 24(b) e (c)), indicando aquecimento ( $300^{\circ}\text{C} < T < 700^{\circ}\text{C}$ ) por circulação de corrente, em que foi providenciada a substituição dos condutores e da isolação.

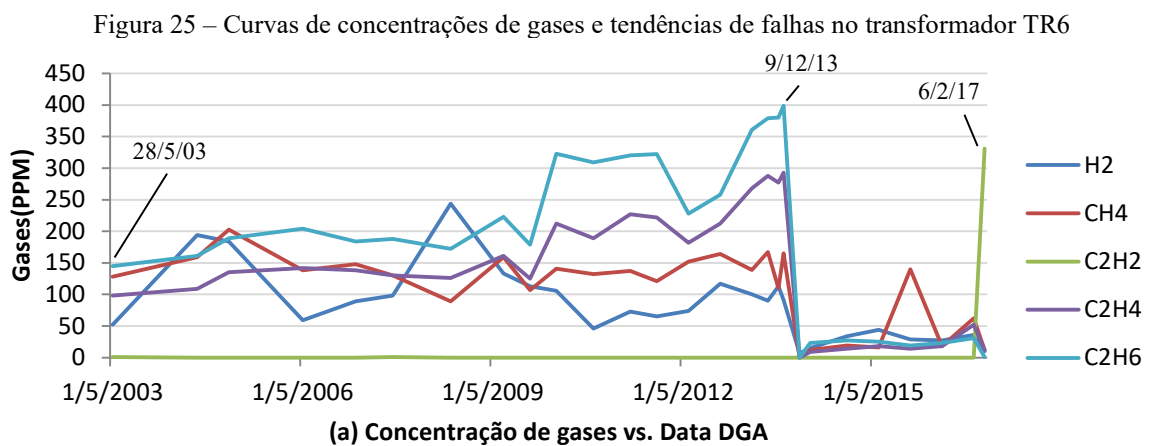
Associando as evidências das falhas encontradas no transformador com o método de diagnóstico proposto nesta tese, nos gráficos de evoluções de falhas das Figuras 23(b) a (g) observa-se que antecedente à primeira inspeção interna, mostraram a presença de falhas do tipo  $T < 700^{\circ}\text{C}$  e DAI, alternando-se como falhas principal e secundária, e após procedimento de inspeção interna durante a troca das buchas, foi encontrada falha térmica no contato da chave comutadora. Aqui novamente, os custos de ambas as falhas  $T < 700^{\circ}\text{C}$  e DAI estão próximos e pequenas mudanças nas concentrações de gases impactam na mudança de estado rapidamente, caracterizando erro de classificação na medição antes da abertura, porém acompanhado por falha secundária  $T < 700^{\circ}\text{C}$ . Na segunda inspeção em 27/04/18, a elevação da concentração de gases (após 16/12/2017) tem o comportamento semelhante ao encontrado em falha anterior, tendo como consequência também a sinalização de falha similar e suas razões. Após correção da falha, registra-se novo crescimento dos gases e sinalizações de falhas presentes, contudo na última medição de 15/09/2018, observa-se o decrescimento dos gases  $\text{H}_2$  e  $\text{C}_2\text{H}_2$ . Nesse caso, o transformador estava sendo acompanhado com um período de amostragem de dois meses, na expectativa de normalização.

Para o transformador analisado nesse caso, é possível verificar pequenas diferenças entre o comportamento de classificação do INN e OPF, entretanto isso não alterou o resultado.

### 5.3.6. Transformador TR6

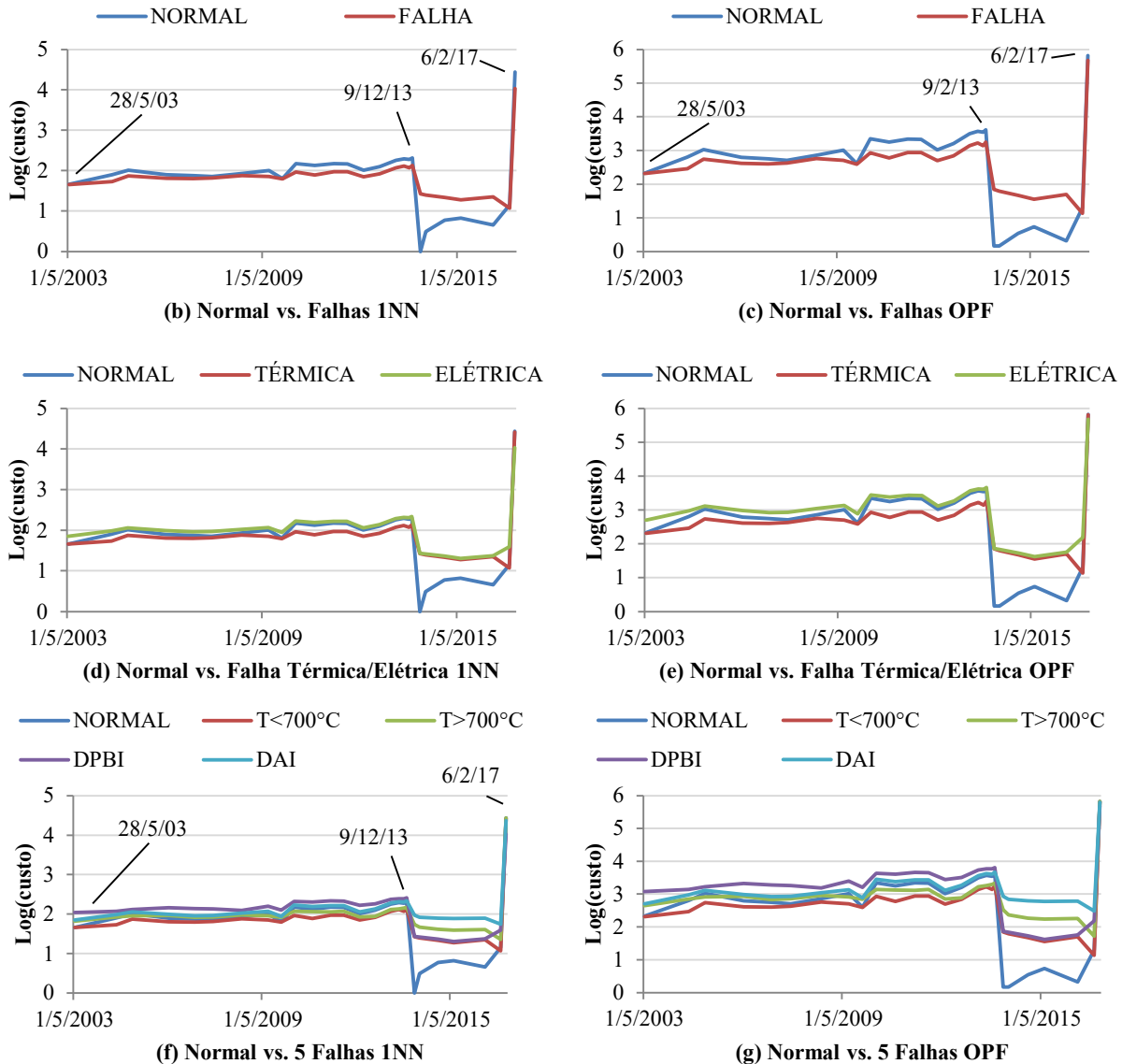
A Figura 25(a) mostra o registro de concentrações dos gases durante a operação do transformador, compreendida entre 28/05/2003 a 09/02/2017. Trata-se de um transformador que tem operado com níveis elevados de gases, contudo, mesmo com o crescimento dos gases iniciado em 28/05/2010 e depois estabilizado, somente após 05/06/2013 foi reduzido o intervalo entre as suas análises de semestral para trimestral, devido ao aumento contínuo, e então, feita a inspeção após 09/12/2013. Retornado ao serviço, o transformador trabalhou normalmente, com baixos níveis de gases até antes de 06/02/2017, quando então atuou a proteção (relé de Buchholz), desligando o equipamento e em 09/02/2017, ocorreu a inspeção do equipamento pela concessionária de energia.

Analisando-se os gráficos de evolução de falhas (Figuras 25(b) e (c)) nas datas de 28/05/2003 a 09/12/2013, a concentração de gases para as amostras submetidas nos algoritmos 1NN e OPF foram todas classificadas como falha, do tipo térmica (Figuras 25(d) e (e)), com características primária  $T < 700^{\circ}\text{C}$  e secundária  $T > 700^{\circ}\text{C}$ . Na Figura 25(a) observa-se os níveis elevados dos gases  $\text{C}_2\text{H}_6$ ,  $\text{C}_2\text{H}_4$ ,  $\text{CH}_4$  e  $\text{H}_2$ , que são típicos de falha térmica  $T < 700^{\circ}\text{C}$ . Após inspeção interna e correção da falha, o equipamento volta a operar, porém, há uma elevação abrupta do gás  $\text{C}_2\text{H}_2$ , típico de descarga elétrica de alta intensidade, conforme sinalizado pelas Figuras 25(d) e (e).





Conclusão Figura 25 – Curvas de concentrações de gases e tendências de falhas no transformador TR6

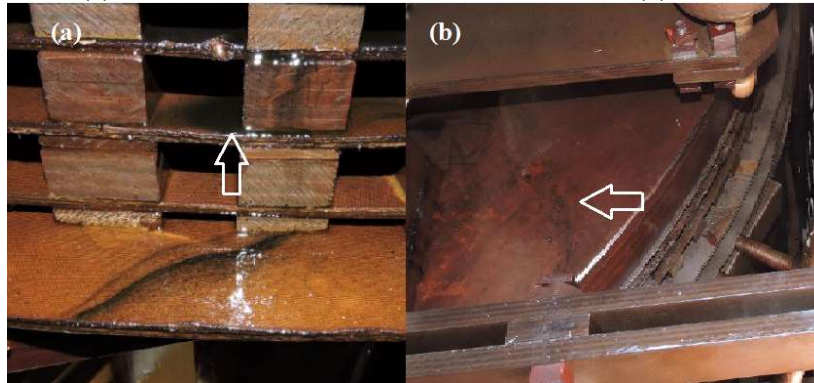


Fonte: Elaborado pelo autor

Na inspeção realizada após 09/12/2013, foram encontrados pontos quentes característicos de temperatura  $T < 700^{\circ}\text{C}$ , coerente com a produção dos gases encontrados, sendo corrigidos tais pontos e retornado à operação. Após quase três anos de retorno ao serviço, em 06/02/2017, houve a atuação do primeiro e segundo estágio do relé de Buchholz, acompanhado de confirmação de geração de gás combustível e deslocamento do óleo para o conservador, características essas de falha grave. Na análise de DGA, foram encontrados altíssimos níveis do gás acetileno por provável curto circuito elétrico (DAI), resultando em danos internos à bobina e liberação de resíduos (Figuras 26(a) e (b)), condenando a reforma e uso do equipamento após essa análise. Associando às evidências das falhas encontradas, a primeira encontrada estava condizente com o método proposto, já para a segunda falha o

método classificou somente após o desligamento do transformador como falha elétrica, uma vez que se tratou de uma falha instantânea, ou seja, os gases estavam em níveis normais e estáveis, sem crescimento, impossibilitando de prever ou apontar a tendência da falha, conforme relatado pela própria concessionária. Isso confirma que o método DGA não é adequado para identificar essas falhas abruptas.

Figura 26 – TR6: (a) Resíduo de óleo carbonizado no interior núcleo; (b) Sedimentos sobre núcleo



Fonte: Adaptado pelo autor de concessionária de energia

#### 5.4. Sugestão de abordagem no tratamento da interpretação de diagnóstico proposto

Um recurso de grande importância relacionado ao diagnóstico de tendência da falha é a possibilidade de antecipar algumas ações suplementares relacionadas à análise de criticidade e risco e, portanto, projetar diferentes cenários de ações, que permitem uma melhor avaliação dos custos e das consequências da retirada do transformador em operação.

Como visto, o mecanismo de disparo para uso dos classificadores no diagnóstico de falhas incipiente, com base na DGA, é determinado pela presença de crescimento dos gases. A norma IEC 60599 (IEC, 2007) sugere que um crescimento de 10% pode ser indicativo de falha em evolução, já a IEEE C57.104 (2008), em seu método total de gases combustíveis, (TDCG – *Total Dissolved Combustible Gas*) determina faixa de valores limites para classificar os riscos, quando não há histórico de gases.

No método proposto, a taxa de crescimento e o histórico (tempo) são naturalmente integrada ao algoritmo de aprendizado de máquina, em que o primeiro estabelece a medida de similaridade com o padrão de falha, e o segundo, o comportamento ou tendência em direção aos tipos de falhas e, portanto, do ponto de vista operacional do método é necessário estabelecer três níveis básicos de análise conforme segue:

- Condição 1: Equipamento operando em normalmente ou em condição satisfatória. As concentrações dos gases permanecem praticamente estáveis em relação ao seu histórico, e as curvas de tendências de falha acompanham esta condição, podendo ser visualizada e percebida rapidamente na curva normal *versus* falha. Nesse caso, nenhuma investigação é necessária e a frequência de análise de DGA estabelecida como adequada para o equipamento em questão pode ser mantida.
- Condição 2: Equipamento operando em normalmente ou em condição satisfatória, entretanto, as curvas de tendências de condições normal e falha apresentam uma aproximação, motivado pelo crescimento das concentrações dos gases. Nesse caso, identificar a aproximação dos tipos das possíveis falhas, se térmicas ou elétricas e suas especificidades, bem como os tipos dos gases em evolução para verificar a relação com os tipos de falhas.

Desde que o crescimento dos gases seja baixo e o tipo de gás envolvido não seja gerado por falha grave, reduzir a frequência de análise de DGA (por exemplo, se semestral para trimestral, ou menor se necessário) para acompanhar a evolução da sinalização. Caso o crescimento seja representativamente elevado, investigações adicionais devem ser promovidas, como por exemplo, imediata análise de DGA, para acompanhar a evolução e se necessário, ajustar a frequência para o caso após medição. Quanto mais rápido for identificada uma anormalidade, mais eficaz pode ser a ação de manutenção e menor será o custo envolvido.

- Condição 3: Equipamento passa a sinalizar falha presente em comparação ao histórico de medidas de DGA anterior. Iniciar investigação imediatamente, ou seja, identificar inicialmente, os tipos de falhas envolvidas e a gravidade que a acompanha, bem como os gases envolvidos e sua relação com o tipo de falha. Em caso de compatibilidade entre os gases e o tipo de falha, refazer imediatamente a DGA e considerar a retirada do transformador de operação, mediante a gravidades das falhas envolvidas. Entretanto, desde que a premissa seja a continuidade do serviço, a retirada do equipamento somente será feita com base na análise de criticidade que envolve a análise suplementar de outras fontes de informações, conforme mostrado na Figura 7. Como por exemplo, a avaliação do caso de sinalização de falha térmica de baixa temperatura envolve checar informações quanto ao carregamento do transformador ou realizar operação de redução do carregamento para avaliar a persistência da falha. Em caso de falha elétrica de alta energia, checar se não houve eventos de descarga atmosférica na subestação, atuação de proteções, entre outros. Em outras palavras, a

análise deve procurar estabelecer uma normativa de informações que permitam confirmar a falha e conseqüentemente, apoiar a tomada de decisão.

Uma questão de grande importância sobre a entrada dos dados de DGA para análise, é que esses devem ser validados antes de serem submetidos à análise, ou seja, condições especiais de análise de DGA como por exemplo, após ocorrência de substituição do óleo, tratamento termo-vácuo ou de regeneração, em que os gases contidos no papel ao se homogeneizarem com óleo, podem provocar o crescimento dos gases devido à migração, causando falsa sensação de crescimento dos gases, neste caso, devendo ser registrado o evento devidamente para fim de consulta em caso de análise de falha.

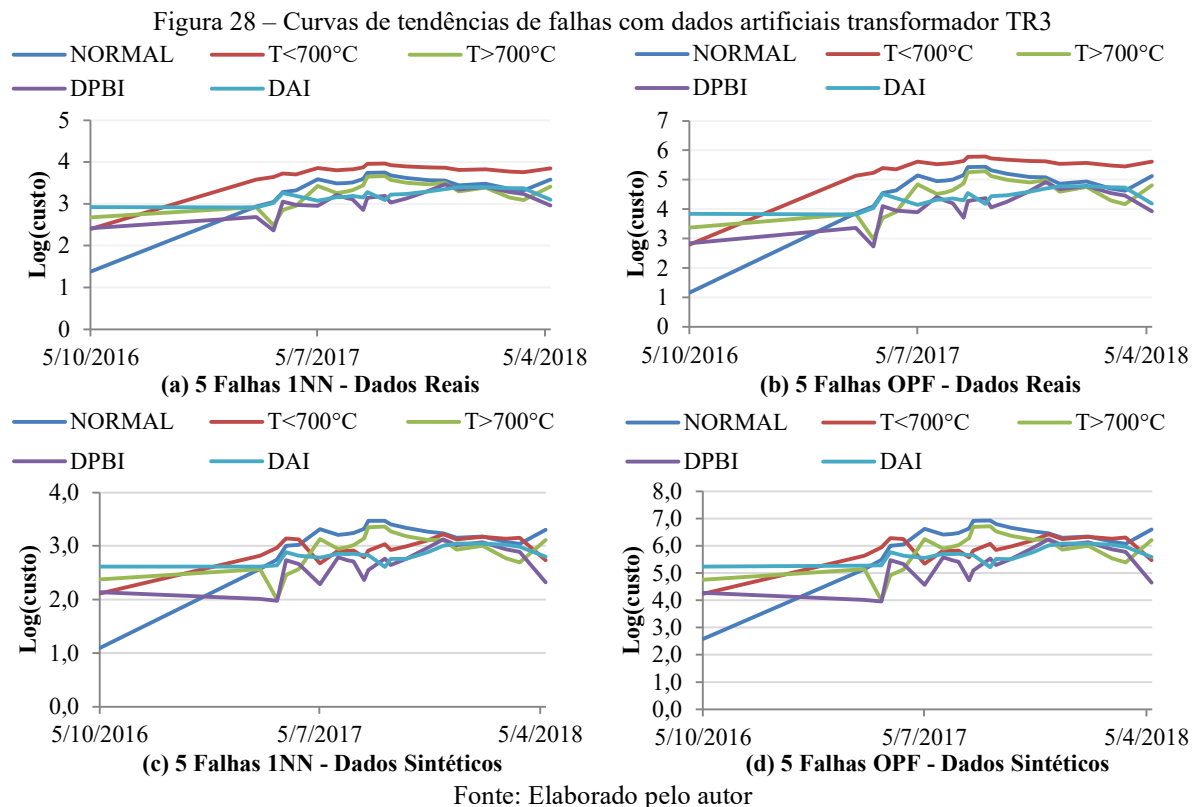
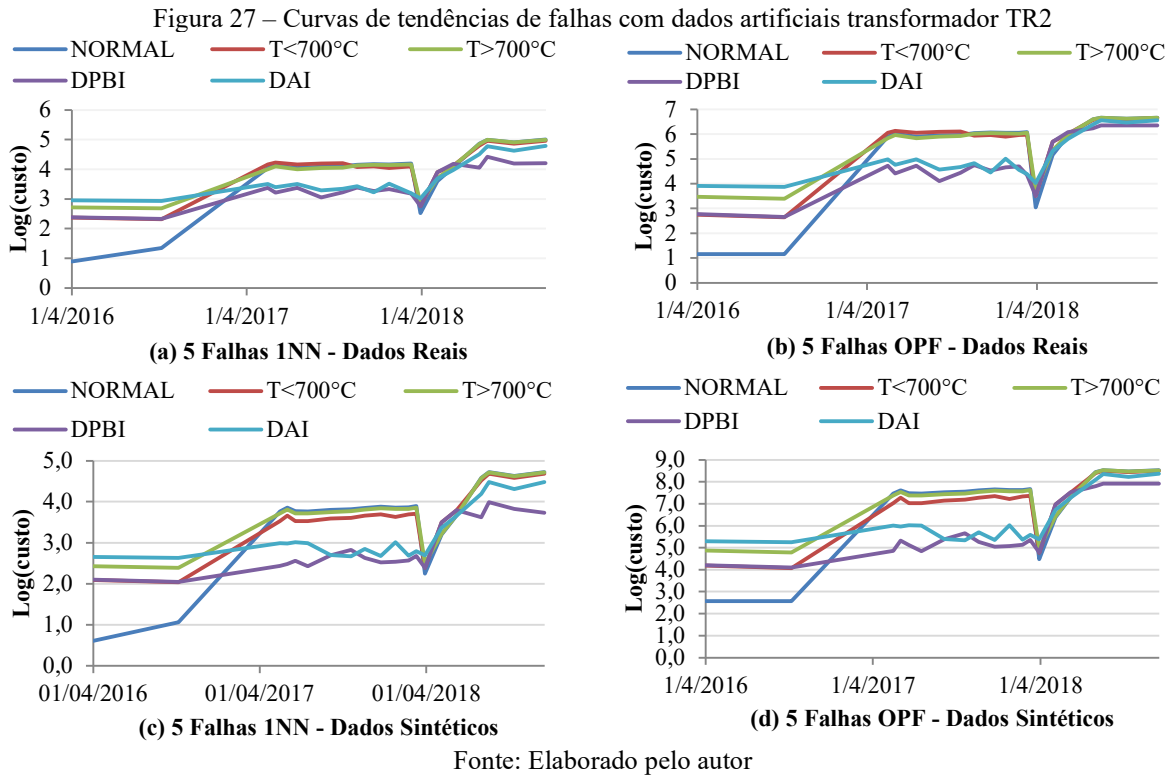
### **5.5. Análise do histórico de DGA usando a abordagem desenvolvida e dados sintéticos**

A criação de dados artificiais tem impacto na formação do espaço de atributos modificando sua estrutura usada na classificação de novas amostras quando elas são submetidas à avaliação. No entanto, isso não significa que no caso proposto há um aumento na capacidade de generalização do algoritmo no contexto do reconhecimento correto de amostras nunca apresentadas, principalmente em algoritmos que usam poucas parametrizações.

A geração de dados com base na incerteza de medição, visa criar no entorno da amostra de DGA, uma região com possíveis pontos de ocupação, aumentando assim a possibilidade de que amostras com características semelhantes sejam reconhecidas e então, classificadas. Assim, para efeito de análise do conjunto de dados de falhas sintéticos gerados, os mesmos transformadores de TR1 a TR6 foram submetidos ao diagnóstico e os resultados somente daqueles que apresentaram mudanças no comportamento das curvas de falhas são mostrados.

Nas Figuras 27(a) a (d), e 28(a) a (d), verifica-se respectivamente os comportamentos das curvas de diagnósticos de tendências de falhas (para cinco estados) no transformador TR2 e TR3 simulados nos algoritmos 1NN e OPF com dados reais e sintéticos. Comparando-se os resultados, é possível notar nas curvas de ambas as figuras com dados sintéticos, mudanças no comportamento das curvas  $T < 700^{\circ}\text{C}$  e trecho central da curva DPBI e DAI. Observa-se que essas mudanças ocorrem em regiões onde os custos das amostras reais do espaço de atributos estão mais próximos (menor deslocamento) ou menos densos (maior deslocamento), contudo no resultado das classificações com dados sintéticos das Figuras 27(c) e (d), e 28(c) e (d), praticamente mantiveram com relação aos dados reais. Como resultado de análise desses

casos, mesmo com a mudança de classificação do resultado sendo positiva na análise de um dos transformadores, os procedimentos de acompanhamento continuariam sendo os mesmos e tal melhora não significaria que isso sempre ocorreria.



## 5.6. Considerações gerais do capítulo

Baseado na relação de custo ótimo entre o vetor de características das amostras de DGA e o espaço de atributos dos algoritmos classificadores 1NN e OPF, é proposto o método de diagnóstico de falhas em transformadores abordado nesta tese. Destinado ao uso na gestão de falhas incipientes desses equipamentos, essa ferramenta permite, além de classificar padrões, relacionar a taxa de crescimento dos gases dissolvidos no óleo, apontando tendências de possíveis falhas em níveis distintos (falha principal, secundária, entre outras), com base no histórico e medidas sucessivas, tornando-a mais informativa para que seja possível adotar estratégias e ações adequadas na manutenção do transformador e, conseqüentemente, minimizar os efeitos negativos de possíveis interrupções no fornecimento de energia por paradas inesperadas.

Como resultados obtidos de aplicação e diagnóstico, pode ser visto que na maioria dos casos dos transformadores analisados, as classificações e análises de tendências de falhas apresentadas são promissoras, demonstrando boa coerência com os diagnósticos de inspeções realizadas e análise do especialista em manutenção de transformadores. Também foi possível verificar que os gráficos mostraram a tendência de ocorrência de falhas, antes mesmo de serem confirmadas, tornando possível avaliar a criticidade antecipadamente e auxiliar na tomada de decisão de manter o transformador operando, mesmo que parcialmente, até encontrar solução adequada que minimizem os inconvenientes de uma parada.

Recursos para melhorar a eficácia dos classificadores por meio do uso de técnicas híbridas foram demonstrados, como por exemplo, o uso de algoritmo de otimização na seleção de atributos, pode contribuir diretamente para melhorar o desempenho da metodologia da análise proposta, pois como parte do processo de classificação, análise do histórico e das novas amostras de gases dissolvidos no óleo, os algoritmos supervisionados se utilizam de conjuntos de dados conhecidos que contêm informações dos gases e seus respectivos rótulos (classes de falhas) para constituírem seus espaços de atributos multidimensionais, e que normalmente são mal condicionados.

Em complemento, a partir de várias bases e artigos contendo dados de falhas em transformadores envolvendo DGA, constituiu-se o conjunto de dados para treinamento dos algoritmos 1NN e OPF usados na metodologia proposta, em que inicialmente foi realizado o seu pré-processamento, por meio da eliminação dos *outliers* e escolha dos melhores atributos com o uso de algoritmos de otimização para melhorar a qualidade dos dados, e então aplicado o método. Os resultados obtidos no tratamento mostraram que esse procedimento, mesmo que

se faça uso de outras técnicas de classificação de falhas (por exemplo, ANN, SVM, e outros), é adequado para reduzir os erros de classificações que serão gerados com base no espaço de atributos construído com o conjunto de dados de treinamento, principalmente em algoritmos com maior sensibilidade a ruídos. Ressalta-se que um conjunto de dados de DGA para treinamento dos algoritmos formado apenas por amostras do grupo de transformadores, tende a agregar as características e comportamento deles, e que podem contribuir com a melhora na identificação de falhas e condições normais de operação.

Outra forma de melhorar a assertividade no treinamento do algoritmo pode ser por intermédio do uso de dados sintéticos, principalmente no caso de poucos dados disponíveis. Produzir dados sintéticos de DGA com base nas incertezas de medições dos equipamentos pode ser um recurso, contudo os valores limites utilizados devem preferencialmente, ser escolhidos com base naqueles definidos pelos laboratórios de medições, ou conforme equipamentos particulares usados nas medições dos gases, entre outros. Cuidados devem ser tomados na escolha e adoção de valores de incertezas, quando existem diferentes laboratórios realizando as análises.

Na avaliação dos dados sintéticos gerados nesta tese, é possível verificar a melhora produzida na taxa de acerto, contudo também se observa que há pequenas mudanças introduzidas nos gráficos de falhas que resultam em trocas de posições e de classes de falhas, mas que nos casos avaliados, não modificaram as análises e procedimentos de acompanhamento.

Quanto ao comportamento dos algoritmos INN e OPF, observou-se grande semelhança nos resultados apresentados, permitindo o emprego de qualquer um dos algoritmos com a metodologia proposta.

No desenvolvimento desta tese, foi possível comprovar através dos gráficos de tendências, que as informações sobre as concentrações de gases relacionadas às falhas nem sempre encontram-se alinhadas com o momento exato de ocorrência, ou seja, os registros podem conter um atraso entre o início real (concentração inicial), o momento da inspeção e medida relacionada (concentração final).

Nesta tese não se avaliou o custo computacional do treinamento dos algoritmos INN e OPF, uma vez que se trata do diagnóstico e análise de tendência de falhas incipientes em transformadores, e nesse caso, a rapidez no treinamento e classificação não é fator preponderante em relação à assertividade ou indicação correta da tendência de falha.

## CAPÍTULO 6 – CONCLUSÃO E TRABALHOS FUTUROS

### 6.1. Aspectos gerais

O problema da saída repentina dos transformadores de operação do sistema elétrico por falha pode ser reduzido, por meio da adoção de procedimentos de acompanhamento e manutenção adequada. Para isso, é necessário definir indicadores que fomentem as tomadas de decisões corretas, para que esse objetivo seja alcançado. A estatística de falhas em transformadores a óleo é um dos indicadores que permite a concessionária de energia realizar análises e planejamentos, para melhorar seus procedimentos de manutenção e determinar políticas de reservas de equipamentos e acessórios.

Os índices de falhas têm demonstrado uma grande ocorrência nos elementos internos que possuem contato com o líquido isolante e, portanto, não é por acaso que a interpretação da DGA tem sido o método mais empregado na análise de falhas internas incipientes, uma vez que já está estabelecida uma relação entre falhas e gases gerados com degradação do óleo e papel, e também porque permite ser feita com o equipamento em serviço. Presentes nas normas da IEC e IEEE, os conhecidos métodos normatizados de análise de DGA continuam sendo amplamente utilizados por empresas especializadas em manutenção de transformadores e em concessionárias de energia, mesmo que ainda sejam associados às questões que nem todos esses métodos preveem as possíveis situações de falhas e levam a resultados inconsistentes ou ambíguos, o que obriga ao usuário a ter experiência na área para análise.

Nesse contexto, o avanço da tecnologia tem contribuído com o desenvolvimento de novos recursos para aquisição e análise de DGA, bem como o uso de inteligência artificial (aprendizado de máquinas) para a criação de novas metodologias, que busquem facilitar sua interpretação, proporcionar maior precisão na classificação de falhas, ou que forneçam um diagnóstico mais informativo, ponderando informações de falhas com diferentes perspectivas de ocorrências. No estudo dos algoritmos classificadores de aprendizado de máquina, foi possível identificar recursos, como por exemplo, de tratamento de dados, uso de algoritmos de otimização, geração de dados artificiais, entre outros, que associados a esses algoritmos, contribuem com a melhora na taxa de acerto da metodologia proposta.

Sobre a perspectiva de resultados mais informativos, nesta tese, desenvolveu-se uma nova forma de análise, que permitiu extrair mais informações sobre os resultados, após classificação, que evidenciassem a proximidade entre outros tipos de falhas. Com o método de diagnóstico de falhas em transformadores proposto, foi possível classificar e relacionar a taxa



de crescimento dos gases dissolvidos no óleo, indicando tendências de falhas em evolução e em níveis de importância distintos (falha principal, secundária, outras), tornando a resposta mais informativa. O ineditismo está em capturar o dinamismo das evoluções dos gases e sua relação com as classes de falhas, permitindo conhecer as tendências de falhas incipientes, antes que ocorram, assim criando indicadores diferenciados de falhas, que podem auxiliar/contribuir com um maior suporte à tomada de decisão, no âmbito de fornecer soluções para prevenir ou reduzir os prejuízos financeiros.

Nesse contexto, essas curvas permitem conhecer e entender o comportamento operacional do transformador e indicar tendências de mudanças no seu estado, e assim ser utilizada como indicador de análise do comportamento térmico para avaliação da vida útil do transformador, bem como servir de ferramenta de comparação comportamental de família/grupos de transformadores de uma mesma subestação. Com essas informações agregadas à análise de criticidade ou gerenciamento de risco dos ativos, tornam-se uma ferramenta poderosa para a tomada de decisão.

Nesta tese foram apresentados os resultados das análises de alguns transformadores de potência com o método proposto, sendo possível constatar o comportamento da condição de operação normal dos equipamentos ao surgimento de falhas, compatível com a variação da concentração dos gases, além disso, destaca-se o fato de que o método indicou as tendências de ocorrência falhas, muito antes de serem comprovadas. Quanto ao uso, é apresentada uma sugestão de análise e interpretação das curvas de tendências de falhas em três níveis de ações de modo que seja possível avaliar as indicações de falhas existentes, e determinar ou não a retirada do equipamento de operação, contudo, à medida que for sendo utilizado ou integrado novos recursos, devem ser complementados e melhorados, com objetivo de tomar decisões mais precisas. Também são sugeridos o ajuste de alguns fatores como por exemplo, redução na frequência entre análise de DGA em subestações críticas e necessidade de validar as medidas de DGA, realizadas antes de submeter à classificação pelo algoritmo.

A metodologia aplica-se à análise de falhas em grandes transformadores de potência de média e alta tensão, a partir de 69kV, ou onde sua saída do sistema de potência seja fator de representativo custo financeiro à concessionária.

Diante do exposto, pode-se concluir que o desenvolvimento de novas metodologias para o diagnóstico de falhas incipientes em transformadores de potência, utilizando a análise DGA e aprendizado de máquina, ganha um destaque especial, levando-se em conta o grande desenvolvimento na área inteligência artificial e a possibilidade de acesso automático de dados, referentes ao funcionamento do transformador.

## 6.2. Futuros desdobramentos da pesquisa

Como continuidade de atividades futuras, vislumbra-se a possibilidade de melhorar a avaliação dos transformadores e os processos de manutenções, da seguinte forma:

- Analisar o comportamento e utilizar o emprego de novas variáveis (por exemplo, temperatura do óleo, volume, carregamento do transformador, gradiente de crescimento dos gases, umidade, análise físico-química, entre outros) em conjunto com os gases combustíveis, para melhorar o desempenho dos algoritmos inteligentes e consequentemente, o método aqui proposto;
- Analisar e desenvolver novos recursos nos algoritmos OPF e KNN para que tornem mais robustos a presença de *outliers* no conjunto de dados;
- Integrar essa metodologia à análise de criticidade de transformadores, de modo que seja possível realizar a tomada de decisão, quanto a antecipar a manutenção do equipamento, ou adia-la (desde que possível), mantendo-o mesmo em operação por um determinado período, até planejar uma solução adequada; e
- Implementar a gestão de risco envolvendo a análise de criticidade em transformadores, as necessidades operacionais e de continuidade do fornecimento do serviço.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ABB. **DPC Type De-energized tap-changer**. Disponível em: <<https://search.abb.com/library/Download.aspx?DocumentID=1ZKM013119&LanguageCode=en&DocumentPartId=&Action=Launch>>. Acesso em: 6 jul. 2020.
- ABNT. **ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS - ABNT NBR 7070: Guia para amostragem de gases e óleo em transformadores e análise dos gases livres e dissolvidos**, Rio de Janeiro: ABNT, 1981.
- ABNT. **ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS - NBR IEC 60450: Medição do grau de polimerização viscosimétrico médio de materiais celulósicos novos e envelhecidos para isolamento elétrica**, Rio de Janeiro: ABNT, 2009.
- ABNT. **ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS - NBR 7274: Interpretação da análise dos gases em transformadores em serviço**, Rio de Janeiro: ABNT, 2012.
- ABNT. **ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS - NBR 10576: Óleo mineral isolante de equipamentos elétricos - Diretrizes para supervisão e manutenção**, Rio de Janeiro: ABNT, 2017.
- AFONSO, L. C. S.; PEREIRA, D. R.; PAPA, J. P. **A Kernel-based Optimum-Path Forest Classifier**. Iberoamerican Congress on Pattern Recognition - CIARP 2017. **Anais...2017**
- AGUINIS, H.; GOTTFREDSON, R. K.; JOO, H. Best-Practice Recommendations for Defining, Identifying, and Handling Outliers. **Organizational Research Methods**, v. 16, n. 2, p. 270–301, 2013.
- AIZPURUA, J. I. et al. Towards a Comprehensive DGA Health Index. **2018 IEEE 2nd International Conference on Dielectrics, ICD 2018**, v. 1, p. 55–58, 2018.
- ALMEIDA, A.; DELGADO, J.; MOREIRA, L. **Power Quality Problems and New Solutions**. Proceedings of the International Conference on Renewable Energy and Power Quality. **Anais...2003** Disponível em: <<http://www.icrepq.com/pdfs/PL4.ALMEIDA.pdf>>
- AMORA, M. A. B. et al. Improved DGA method based on rules extracted from high-dimension input space. **Electronics Letters**, v. 48, n. 17, p. 1048–1049, 2012.
- ANEEL. Resolução Normativa nº 794. In: **Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional**. [s.l: s.n.]. v. 8p. 88.
- ANEEL. Resolução Normativa nº 669/2015. **Regulamenta os Requisitos Mínimos de Manutenção e o monitoramento da manutenção de instalações de transmissão de Rede Básica**. p. 11. Disponível em: <http://www2.aneel.gov.br/aplicacoes/audiencia/arquivo/2014/022/resultado/ren2015669.pdf>
- ANEEL. **Avaliar Aprimoramentos na Regulação de Continuidade do Fornecimento de Energia Elétrica Relatório de Análise de Impacto Regulatório**, 2018b. Disponível em: <<http://www.aneel.gov.br/documents/656877/18485189/7+Modelo+de+AIR+-+SRD+-+Continuidade+do+Fornecimento.pdf/8a3f3cc3-f5f0-6e17-5dc7-5ef70d6a6eb9>>. Acesso em:

1 jan. 2019.

ARSHAD, M.; ISLAM, S. M.; KHALIQ, A. Fuzzy logic approach in power transformers management and decision making. **IEEE Transactions on Dielectrics and Electrical Insulation**, v. 21, n. 5, p. 2343–2354, 2014.

ARCISZEWSKI, T.; MICHALSKI, R.; WNEK, J. **Constructive Induction: The Key To Design Creativity**. Proceedings of the Third International Round-Table Conference on Computational Models of Creative Design. **Anais...Queensland**: 1995

ASSUNÇÃO, T. C. B. N. **Contribuição à modelagem e análise do envelhecimento de transformadores de potência**. [s.l.] Tese (Doutorado em Engenharia Elétrica), Universidade Federal de Minas Gerais, 2007.

ASTM. **AMERICAN SOCIETY FOR TESTING AND MATERIALS - ASTM D4243-16: Standard Test Method for Measurement of Average Viscometric Degree of Polymerization of New and Aged Electrical Papers and Boards**. [s.l.] ASTM International, West Conshohocken, PA, 2016.

ASTM. **AMERICAN SOCIETY FOR TESTING AND MATERIALS - ASTM D3612-02: Standard Test Method for Analysis of Gases Dissolved in Electrical Insulating Oil by Gas Chromatography**. [s.l.] ASTM International, West Conshohocken, PA, 2017.

AZEVEDO, C. H. B. **Metodologia para a eficácia da detecção de descargas parciais por emissão acústica como técnica preditiva de manutenção em transformadores de potência imersos em óleo isolante**. [s.l.] Dissertação (Mestrado em Engenharia Elétrica), Universidade Federal de Goiás, 2009.

BABURAO, K. et al. **The experience of DP and furan in remnant life assessment of power transformer**. Proceedings of 2008 International Conference on Condition Monitoring and Diagnosis. **Anais...CMD**, 2008.

BAKAR, N.; ABU-SIADA, A.; ISLAM, S. A review of dissolved gas analysis measurement and interpretation techniques. **IEEE Electrical Insulation Magazine**, v. 30, n. 3, p. 39–49, 2014.

BANDARA, K.; EKANAYAKE, C.; SAHA, T. K. Understanding the Ageing Aspects of Natural Ester Based Insulation Liquid in Power Transformer. **IEEE Transactions on Dielectrics and Electrical Insulation**, v. 23, p. 246–257, 2015.

BARBOSA, G. S.; COSTA, L. S.; NETO, A. R. R. **A New Genetic Algorithm-based Pruning Approach for Optimum-Path Forest**. 2018 7th Brazilian Conference on Intelligent Systems. **Anais...BRACIS**, 2018.

BARTLEY, W. H. **An Analysis of International Transformer Failures, Part 1**, 2013. Disponível em: <[https://www.hsb.com/TheLocomotive/uploadedFiles/ArticleLibrary/An Analysis of International Transformer Failure, Part 1.pdf](https://www.hsb.com/TheLocomotive/uploadedFiles/ArticleLibrary/An%20Analysis%20of%20International%20Transformer%20Failure,%20Part%201.pdf)>

BASTOS, G. et al. **Avaliação de Desempenho de Transformadores de Potência e Reatores no Sistema Elétrico Brasileiro**. [s.l.] CIGRÉ-Brasil, 2013. Disponível em: <[http://www.cigre.org.br/archives/BT\\_16.pdf](http://www.cigre.org.br/archives/BT_16.pdf)>.

- BECHARA, R. **Análise de falhas de transformadores de potência.** [s.l.] Dissertação (Mestrado em Engenharia Elétrica), Escola Politécnica da Universidade de São Paulo, 2010.
- BENEDET, M. E. **Otimização de um analisador de gás dissolvido em óleo de múltiplos transformadores de potência.** [s.l.] Dissertação (Mestrado em Metrologia Científica e Industrial), Universidade Federal de Santa Catarina, 2008.
- BENMAHAMED, Y. et al. **Diagnosis of Power Transformer Oil Using KNN and Naïve Bayes Classifiers.** 2018 IEEE 2nd International Conference on Dielectrics. **Anais...IEEE**, 2018.
- BENMAHAMED, Y.; TEGUAR, M.; BOUBAKEUR, A. Application of SVM and KNN to Duval Pentagon 1 for transformer oil diagnosis. **IEEE Transactions on Dielectrics and Electrical Insulation**, v. 24, n. 6, p. 3443–3451, 2017.
- BHAVSAR, H.; GANATRA, A. A Comparative Study of Training Algorithms for Supervised Machine Learning. **International Journal of Soft Computing and Engineering**, v. 2, n. 4, p. 8, 2012.
- BONFIN, THIAGO SAMUEL. **Sistema Especialista Para Análise Integrada de Respostas Advindas de Ensaio Experimentais Executados em Transformadores de Potência.** [s.l.] Dissertação (Mestrado em Ciências), Escola de Engenharia de São Carlos da Universidade de São Paulo, 2015.
- BORSI, H.; GOCKENBACH, E. **Properties of ester liquid midel 7131 as an alternative liquid to mineral oil for transformers.** IEEE International Conference on Dielectric Liquids. **Anais...IEEE**, 2005.
- BOSER, B. E.; GUYON, I. M.; VAPNIK, V. N. **A training algorithm for optimal margin classifiers.** PROCEEDINGS OF THE 5TH ANNUAL ACM WORKSHOP ON COMPUTATIONAL LEARNING THEORY. **Anais...1992**
- BOSS, P. et al. **Guide on Economics of Transformer Management.** [s.l.] CIGRÉ, 2004.
- BRICEÑO, A. **Grupo de Flujo de carga en los sistemas eléctricos de potencia SEP.** Disponível em: <[https://grupos.emagister.com/debate/sistema\\_de\\_enfriamiento\\_de\\_los\\_transformadores\\_de\\_potencia/30499-852455/p2](https://grupos.emagister.com/debate/sistema_de_enfriamiento_de_los_transformadores_de_potencia/30499-852455/p2)>. Acesso em: 6 jul. 2020.
- BUDIAJI, W.; LEISCH, F. Simple K-Medoids Partitioning Algorithm for Mixed Variable Data. **Algorithms**, v. 12, n. 9, p. 177, 2019.
- BUSTAMANTE, S. et al. Dissolved gas analysis equipment for online monitoring of transformer oil: A review. **Sensors (Switzerland)**, v. 19, n. 19, p. 4–12, 2019.
- CARDOSO, P. M. **Adaptação de um sistema de medição de gases dissolvidos em óleo mineral isolante para monitoração de múltiplos transformadores de potência.** [s.l.] Dissertação (Mestrado em Metrologia) - Universidade Federal de Santa Catarina, p. 123, 2005.
- CHEIM, L. et al. Furan analysis for liquid power transformers. **IEEE Electrical Insulation Magazine**, v. 28, n. 2, p. 8–21, 2012.

CHEN, W. G. et al. Analysis of infrared absorption properties of dissolved gases in transformer oil. **CSEE Journal of Power and Energy Systems**, v. 16, p. 148–153, 2008.

CHENG, L.; YU, T. Dissolved gas analysis principle-based intelligent approaches to fault diagnosis and decision making for large oil-immersed power transformers: A survey. **Energies**, v. 11, n. 4, p. 1–69, 2018.

COOPER POWER SYSTEMS. **Rethinking Transformers for Safety, Performance, and Value**, 2000. Disponível em: <[http://www.electricalpartmanuals.com/part\\_manuals/pdf/transformer/Cooper\\_Power\\_Systems/Other/Bulletin-00006.pdf](http://www.electricalpartmanuals.com/part_manuals/pdf/transformer/Cooper_Power_Systems/Other/Bulletin-00006.pdf)>

DHINI, A. et al. **Intelligent Fault Diagnosis for Power Transformer Based on DGA Data Using Support Vector Machine (SVM)**. Proceedings - 2018 3rd International Conference on System Reliability and Safety, ICSRS 2018. **Anais...IEEE**, 2019.

DORIGO, M.; BIRATTARI, M.; STUTZLE, T. Ant colony optimization. **IEEE Computational Intelligence Magazine**, v. 1, n. 4, p. 28–39, nov. 2006.

DU, Y. et al. Moisture Equilibrium in Transformer Paper-oil Systems. **IEEE Electrical Insulation Magazine**, v. 15, n. 1, p. 11–20, 1999.

DUVAL, M. The Duval Triangle for Load Tap Changers, Non-Mineral Oils and Low Temperature Faults in Transformers. **IEEE Electrical Insulation Magazine**, v. 24, n. 6, p. 22–29, 2008a.

DUVAL, M. Calculation of DGA limit values and sampling intervals in transformers in service. **IEEE Electrical Insulation Magazine**, v. 24, n. 5, p. 7–13, 2008b.

DUVAL, M. **New frontiers of DGA interpretation for power transformers and their accessories**, 2012. Disponível em: <[http://www.seeei.org.il/prdFiles/2922\\_desc2.pdf](http://www.seeei.org.il/prdFiles/2922_desc2.pdf)>

DUVAL, M.; DEPABLO, A. Interpretation of gas-in-oil analysis using new IEC publication 60599 and IEC TC 10 databases. **IEEE Electrical Insulation Magazine**, v. 17, n. 2, p. 31–41, 2001.

DUVAL, M.; DUKARM, J. J. Improving the reliability of transformer gas-in-oil diagnosis. **IEEE Electrical Insulation Magazine**, v. 21, n. 4, p. 21–27, 2005.

DUVAL, M.; LAMARRE, L. The Duval Pentagon — A New Complementary Tool for the Interpretation of Dissolved Gas Analysis in Transformers. **IEEE Electrical Insulation Magazine**, v. 30, n. 6, p. 1–4, 2014.

EECKHOUDT, S.; AUTRU, S.; LERAT, L. Stray gassing of transformer insulating oils: Impact of materials, oxygen content, additives, incubation time and temperature, and its relationship to oxidation stability. **IEEE Electrical Insulation Magazine**, v. 33, n. 6, p. 27–32, 2017.

ENGINEERING WORD. **What is inside a transformer ?** Disponível em: <<https://www.engineeringworldchannel.com/transformer-definition/>>. Acesso em: 6 jul. 2020.

ENWEN, L.; WANG, L.; SONG, B. **Dissolved gas data in transformer oil---Fault Diagnosis of Power Transformers with Membership Degree**. Disponível em: <<https://www.ieee-dataport.org/documents/dissolved-gas-data-transformer-oil-fault-diagnosis-power-transformers-membership-degree>>. Acesso em: 1 mar. 2020.

EQUBAL, M. D.; KHAN, S. A.; ISLAM, T. Transformer incipient fault diagnosis on the basis of energy-weighted DGA using an artificial neural network. **Turkish Journal of Electrical Engineering and Computer Sciences**, v. 26, n. 1, p. 77–88, 2018.

FAIZ, J.; SOLEIMANI, M. Dissolved gas analysis evaluation in electric power transformers using conventional methods a review. **IEEE Transactions on Dielectrics and Electrical Insulation**, v. 24, n. 2, p. 1239–1248, 2017.

FAROOQUE, M. U.; WANI, S. A.; KHAN, S. A. **Artificial neural network (ANN) based implementation of Duval pentagon**. 2015 International Conference on Condition Assessment Techniques in Electrical Systems, **Anais...CATCON**, 2016.

FAWCETT, T. An introduction to ROC analysis. **Pattern Recognition Letters**, v. 27, n. 8, p. 861–874, 2006.

FERNANDES, S. E. N. et al. A Probabilistic Optimum-Path Forest Classifier for Non-Technical Losses Detection. **IEEE Transactions on Smart Grid**, v. 10, n. 3, p. 3226–3235, 2019.

FERNANDES, S. E. N.; PAPA, J. P. Improving optimum-path forest learning using bag-of-classifiers and confidence measures. **Pattern Analysis and Applications**, v. 22, n. 2, p. 703–716, 2017.

FILHO, G. L. **Comparação entre os critérios de diagnósticos por análise cromatográfica de gases dissolvidos em óleo isolante de transformador de potência**. [s.l.] Universidade de São Paulo, 2012.

FUKUNAGA, K.; NARENDRA, P. M. A branch and bound algorithm for computing k-nearest neighbors. **IEEE Transactions on Computers**, v. 100, n. 7, p. 750–753, 1975.

GEDAM, A. V; SWAMI, P. P. S.; THOSAR, A. A Comparative Analysis of DGA Method for the Incipient Fault Diagnosis in Power Transformer Using ANN Approach. **International Journal of Scientific & Engineering Research**, v. 6, n. 5, p. 387–391, [s.d.].

GHONEIM, S. M.; WARD, S. A. Dissolved Gas Analysis as a Diagnostic Tools for Early Detection of Transformer Faults. **Advances in Electrical Engineering Systems**, v. 1, n. 3, p. 152–156, 2012.

GOCKENBACH, E.; BORSI, H. **Natural and synthetic ester liquids as alternative to mineral oil for power transformers**. 2008 Annual Report Conference on Electrical Insulation and Dielectric Phenomena. **Anais...2008**.

GORGAN, B. et al. **Calculation of the remaining lifetime of power transformers paper insulation**. Proceedings of the International Conference on Optimisation of Electrical and Electronic Equipment. **Anais...2012**.

GOU, J. et al. A new distance-weighted k-nearest neighbor classifier. **Journal of**

**Information & Computational Science**, v. 9, n. February, p. 1429–1436, 2012.

GOUDA, O. E.; EL-HOSHY, S. H.; EL-TAMALY, H. H. Proposed heptagon graph for DGA interpretation of oil transformers. **IET Generation, Transmission & Distribution**, v. 12, p. 490–498, 2017.

GUPTA, M. et al. Outlier Detection for Temporal Data: A Survey. **IEEE Transactions on Knowledge and Data Engineering**, v. 26, n. 9, p. 2250–2267, 2014.

GUPTA, P. Applications of Fuzzy Logic in Geophysics. **International Journal of Advanced Research in Computer Science**, v. 8, n. 5, p. 1795–1800, 2017.

HALDURAI, L.; MADHUBALA, T.; RAJALAKSHMI, R. A Study on Genetic Algorithm and its Applications. **International Journal of Computer Sciences and Engineering**, v. 4, n. 10, p. 139–143, 2016.

HAYKIN, S. **Neural Networks and Learning Machines**. 3. ed. New Jersey: Pearson Education Inc., 2009.

HILLARY, W. D. A. G. et al. **A Tool For Estimating Remaining Life Time of a Power Transformer**. Moratuwa Engineering Research Conference. **Anais...IEEE**, 2017.

HONG, Y.; MEEKER, W. Q.; MCCALLEY, J. D. **Prediction of remaining life of power transformers based on left truncated and right censored lifetime data**. *Annals of Applied Statistics*. **Anais...2009**.

HU, L.-Y. et al. The distance function effect on k-nearest neighbor classification for medical datasets. **SpringerPlus**, v. 5, n. 1, 2016.

HUANG, G.-B. et al. Extreme learning machine for regression and multiclass classification. **IEEE transactions on systems, man, and cybernetics. Part B, Cybernetics**, v. 42, n. 2, p. 513–29, 2012.

HUANG, G. B.; ZHU, Q. Y.; SIEW, C. K. Extreme learning machine: Theory and applications. **Neurocomputing**, v. 70, n. 1–3, p. 489–501, 2006.

HUANG, X.; WANG, X.; TIAN, Y. **Research on Transformer Fault Diagnosis Method based on GWO Optimized Hybrid Kernel Extreme Learning Machine**. *Condition Monitoring and Diagnosis (CMD)*. **Anais...2018**.

HUANG, Z. Extensions to the k-Means Algorithm for Clustering Large Data Sets with Categorical Values. *Data Mining and Knowledge Discovery* 2, 283-304. **Data Mining and Knowledge Discovery**, v. 2, n. 3, p. 283–304, 1998.

IBRAHIM, M. M.; SAYED, M. M.; EL-ZAHAB, E. E. A. **Diagnosis of Power Transformer Incipient Faults Using Fuzzy Logic-IEC Based Approach**. 2014 IEEE International Energy Conference. **Anais...IEEE**, 2014.

IEC 60567. **International Electrotechnical Commission - IEC 60567: Oil-filled electrical equipment - Sampling of gases and of oil for analysis of free and dissolved gases - Guidance**. [s.l.] IEC, Geneva, 2005.



IEC 60599. **Mineral oil-impregnated electrical equipment in service – Guide to the interpretation of dissolved and free gases analysis.** [s.l.] INTERNATIONAL ELECTROTECHNICAL COMMISSION, Geneva, 2007.

IEC 60599. **Mineral oil-filled electrical equipment in service – Guidance on the interpretation of dissolved and free gases analysis.** 3. ed. [s.l.] INTERNATIONAL ELECTROTECHNICAL COMMISSION, 2015.

IEEE C57.104. **IEEE Guide for the Interpretation of Gases Generated in Oil-Immersed Transformers.** [s.l.] IEEE, New York, NY, 2008.

IEEE C57.104. **IEEE Guide for the Interpretation of Gases Generated in Mineral Oil-Immersed Transformers.** June 13, 2 ed. [s.l.] IEEE, 2019. v. 2019.

IKESHOJI, M. A. et al. **Desenvolvimento de uma Ferramenta de Diagnóstico de Falhas em Transformadores Utilizando à Abordagem de Aprendizagem de Máquina Extrema.** 2018 INDUSCON - Conferencia de Aplicação Industrial de Eletricidade, Conservação, Confiabilidade e Controle/Automação. **Anais...Induscon**, 2018.

IKESHOJI, M. A. et al. **Qualidade no Fornecimento de Energia: Contribuições de uma Ferramenta Inteligente para Gestão de Falhas Incipientes em Transformadores.** XIII Conferência Brasileira sobre Qualidade de Energia Elétrica. **Anais...2019.**

ILGEVICIUS, A.; BABIZKI, A. **Integrated Transformer Fleet Management (ITFM) System.** 23rd International Conference on Electricity Distribution. **Anais...2015.**

ILLIAS, H. A. et al. Transformer incipient fault prediction using combined artificial neural network and various particle swarm optimisation techniques. **PLoS ONE**, v. 10, n. 6, p. 1–16, 2015.

IPTC, W. G. R. Information on High Temperature Insulation for Liquid-Immersed Power Transformer. **IEEE Transactions on Power Delivery**, v. 9, n. 4, p. 1892–1906, 1994.

ISLAM, M. J. et al. **Investigating the performance of Naive- Bayes classifiers and K-nearest neighbor classifiers.** 2007 International Conference on Convergence Information Technology. **Anais...2007.**

JADHAV, S. D.; CHANNE, H. P. Comparative Study of K-NN, Naive Bayes and Decision Tree Classification Techniques. **International Journal of Science and Research (IJSR)**, v. 5, n. 1, p. 1842–1845, 2016.

JAFAR, M.; VISSER, R. G.; STARINK, R. **Results of Proficiency Test Dissolved Gas Analysis.** Disponível em: <<http://www.iisnl.com/pdf/iis10L03DGA.pdf>>. Acesso em: 3 jul. 2020.

JAIN, A. K.; CHANDRASEKARAN, B. Dimensionality and Sample Size Considerations in Pattern Recognition Practice. In: **Handbook of Statistics.** v. 2 ed. [s.l.] Elsevier B.V., 1982. p. 835–855.

JAKOB, F.; DUKARM, J. J. Transformer Fault Severity. **IEEE Transactions on Power Delivery**, v. 30, n. 4, p. 1941–1948, 2015.

- JARMAN, P. et al. **End-of-Life Modelling for Power Transformers in Aged Power System Networks**. 6th Southern Africa Regional Conference. *Anais...Southern Africa: Cigré*, 2009.
- KARI, T. et al. Hybrid feature selection approach for power transformer fault diagnosis based on support vector machine and genetic algorithm. *IET Generation, Transmission & Distribution*, v. 12, n. 21, p. 5672–5680, 2018.
- KAUR, A.; BRAR, Y. S.; LEENA, G. Fault detection in power transformers using random neural networks. *International Journal of Electrical and Computer Engineering (IJECE)*, v. 9, n. 1, p. 78, 2019.
- KENNEDY, J.; EBERHART, R. **Particle swarm optimization**. IEEE International Conference on Neural Networks - Conference Proceedings. *Anais...IEEE*, 1995.
- KHALYASMAA, A. I. et al. **The problems of dissolved in oil gases analysis results' interpretation in information analytical systems**. Proceedings of the 2016 International Conference and Exposition on Electrical and Power Engineering. *Anais...2016*.
- KHAN, S. A.; EQUBAL, M. D.; ISLAM, T. A comprehensive comparative study of DGA based transformer fault diagnosis using fuzzy logic and ANFIS models. *IEEE Transactions on Dielectrics and Electrical Insulation*, v. 22, n. 1, p. 590–596, 2015.
- KIM, S.-W. et al. New methods of DGA diagnosis using IEC TC 10 and related databases Part 1: Application of gas-ratio combinations. *IEEE Transactions on Dielectrics and Electrical Insulation*, v. 20, n. 2, p. 685–690, 2013.
- KRAUSE, C. et al. **The Degree of Polymerization of Cellulosic Insulation: Review of Measuring Technologies and its Significance on Equipment**. 2014 Electrical Insulation Conference. *Anais...Philadelphia, PA: 2014*.
- KUHN, M.; JOHNSON, K. **Applied Predictive Modeling**. 1. ed. [s.l.] Springer, 2013.
- LADHA, L.; DEEPA, T. Feature Selection Methods And Algorithms. *International Journal on Computer Science and Engineering*, v. 3, n. 5, p. 1787–1797, 2011.
- LEE, S.-J. et al. New methods of DGA diagnosis using IEC TC 10 and related databases Part 2: Application of gas-ratio combinations. *IEEE Transactions on Dielectrics and Electrical Insulation*, v. 20, n. 2, p. 685–690, 2013.
- LELEKAKIS, N. et al. Comparison of dissolved gas-in-oil analysis methods using a dissolved gas-in-oil standard. *IEEE Electrical Insulation Magazine*, v. 27, n. 5, p. 29–35, 2011.
- LELEKAKIS, N.; MARTIN, D.; WIJAYA, J. Ageing Rate of Paper Insulation used in Power Transformers Part 2 : Oil / paper System with Medium and High Oxygen Concentration. *IEEE Transactions on Dielectrics and Electrical Insulation*, v. 19, n. 6, 2012.
- LI, A. et al. **Cuckoo Search Optimized NN-Based Fault Diagnosis Approach for Power Transformer PHM**. International Conference on Sensing, Diagnostics, Prognostics, and Control, SDPC 2018. *Anais...IEEE*, 2018a
- LI, A. et al. An optimized GRNN-enabled approach for power transformer fault diagnosis.

**IEEE Transactions on Electrical and Electronic Engineering**, n. 2, 2019.

LI, J. et al. Optimal dissolved gas ratios selected by genetic algorithm for power transformer fault diagnosis based on support vector machine. **IEEE Transactions on Dielectrics and Electrical Insulation**, v. 23, n. 2, p. 1198–1206, 2016a.

LI, S. et al. Interpretation of DGA for transformer fault diagnosis with complementary SaE-ELM and arctangent transform. **IEEE Transactions on Dielectrics and Electrical Insulation**, v. 23, n. 1, p. 586–595, 2016b.

LI, J. et al. **A transformer fault diagnosis method based on sub-clustering reduction and multiclass multi-kernel support vector machine**. IEEE Conference on Energy Internet and Energy System Integration. **Anais...**Beijing: 2017.

LI, J. et al. A new iterative synthetic data generation method for CNN based stroke gesture recognition. **Multimedia Tools and Applications**, v. 77, n. 13, p. 17181–17205, 2018b.

LIAO, R. et al. **Study on aging characteristics of mineral oil/natural ester mixtures-paper insulation**. Proceedings - IEEE International Conference on Dielectric Liquids. **Anais...**2011.

LIEPNIECE, R.; VITOLINA, S.; MARKS, J. Study of approaches to incipient fault detection in power transformer by using dissolved gas analysis. **Energetika**, v. 63, n. 2, p. 66–74, 2017.

LIMA, S. L.; SAAVEDRA, O. R.; MIRANDA, V. A two-level framework to fault diagnosis and decision making for power transformers. **IEEE Transactions on Power Delivery**, v. 30, n. 1, p. 497–504, 2015.

LIN, S. et al. Is extreme learning machine feasible? A theoretical assessment (Part II). **IEEE Transactions on Neural Networks and Learning Systems**, v. 26, n. 1, p. 21–34, 2015.

LIU, H.; ZHOU, M.; LIU, Q. An embedded feature selection method for imbalanced data classification. **IEEE/CAA Journal of Automatica Sinica**, v. 6, n. 3, p. 703–715, 2019.

LONDO, L. et al. Hybrid Dissolved Gas-in-Oil Analysis Methods. **Journal of Power and Energy Engineering**, n. 2, p. 10–19, 2015.

LORENA, A. C.; CARVALHO, A. C. P. L. F. Uma Introdução às Support Vector Machines. **Revista de Informática Teórica Aplicada**, v. 14, n. 2, p. 25, 2007.

LUNDGAARD, L. E. et al. Aging of Oil-Impregnated Paper in Power Transformers. **IEEE Transactions on Power Delivery**, v. 19, n. 1, p. 230–239, 2004.

LUOPAJÄRVI, J. **Condition Monitoring of Power Transformer as Part of Power Plant Maintenance Process**. [s.l.] Thesis (Master Degree of Science), University of VAASA, 2010.

MA, H.; EKANAYAKE, C.; SAHA, T. Power transformer fault diagnosis under measurement originated uncertainties. **IEEE Transactions on Dielectrics and Electrical Insulation**, v. 19, n. 6, p. 1982–1990, 2012.

MALABANAN, D. R.; NERVES, A. C. **Power transformer condition assessment using an immune neural network approach to Dissolved Gas Analysis**. IEEE Region 10 Annual International Conference, Proceedings/TENCON. **Anais...IEEE**, 2015.

- MALIK, H.; MISHRA, S. **Extreme learning machine based fault diagnosis of power transformer using IEC TC10 and its related data**. 12th IEEE International Conference Electronics, Energy, Environment, Communication, Computer, Control: (E3-C3), INDICON 2015. **Anais...IEEE**, 2016.
- MARTINS, M. A. G. Furfuraldeído - Um Indicador Prático da Degradação Térmica do Papel Kraft de Transformadores. **Ciência e Tecnologia dos Materiais**, v. 19, n. 1/2, 2007.
- MAYER, N. et al. What Makes Good Synthetic Training Data for Learning Disparity and Optical Flow Estimation ? **International Journal of Computer Vision**, v. 126, n. 9, p. 942–960, 2018.
- METWALLY, I. A. Failures, monitoring and new trends of power transformers. **IEEE Potentials**, v. 30, n. 3, p. 36–43, 2011.
- MIKKULAINEN, R. Topology of a Neural Network. In: **Encyclopedia of Machine Learning**. Boston, MA: Springer US, 2011. p. 988–989.
- MILASCH, M. **Manutenção de Transformadores em Líquido Isolante**. 1. ed. [s.l.] Blucher, 1984.
- MIRANDA, V.; CASTRO, A. R. G.; LIMA, S. Diagnosing faults in power transformers with autoassociative neural networks and mean shift. **IEEE Transactions on Power Delivery**, v. 27, n. 3, p. 1350–1357, 2012.
- MTETWA, N. S. **Accuracy of Furan Analysis in Estimating the Degree of Polymerization in Power Transformers**. [s.l.] Thesis (Master Degree of Science), University of the Witwatersrand, 2008.
- MURUGAN, R.; RAMASAMY, R. Failure analysis of power transformer for effective maintenance planning in electric utilities. **Engineering Failure Analysis**, v. 55, p. 182–192, 2015.
- NETO, A. D.; ASSUNÇÃO, T. C. B. N.; ASSUNÇÃO, J. T. **Classificação dos Transformadores de Potência Empregando a Análise dos Gases Dissolvidos no Óleo Isolante**. IX Congresso Brasileiro de Redes Neurais. **Anais...Ouro Preto, MG: CBRN**, 2009.
- NIU, Q. F.; LI, W.; SHI, Y. P. **Realization of a power transformer on-line monitoring and diagnosis system based on DGA and PNN**. 2010 International Conference on Computer, Mechatronics, Control and Electronic Engineering. **Anais...IEEE**, 2010.
- NUR, A. S.; HAIZAN, N. H. M.; OSMAN, A. Artificial Neural Network Weight Optimization : A Review. **TELKOMNIKA Indonesian Journal of Electrical Engineering**, v. 12, n. March 2017, p. 6897 – 6902, 2014.
- OTERO, F. E. B. et al. **Genetic Programming for Attribute Construction in Data Mining**. 6th European Conference. **Anais...Essex, UK: 2003**Disponível em: <[http://link.springer.com/10.1007/3-540-36599-0\\_36](http://link.springer.com/10.1007/3-540-36599-0_36)>
- PAPA, J. P. **Classificação supervisionada de padrões utilizando floresta de caminhos ótimos**. [s.l.] Tese (Doutorado em Ciência da Computação), Universidade Estadual de Campinas, UNICAMP., 2008.

- PAPA, J. P. et al. **Optimum path forest classifier applied to laryngeal pathology detection**. Proceedings of IWSSIP 2008 - 15th International Conference on Systems, Signals and Image Processing. **Anais...**2008.
- PAPA, J. P. et al. Efficient supervised optimum-path forest classification for large datasets. **Pattern Recognition**, v. 45, n. 1, p. 512–520, 2012.
- PAPA, J. P.; FALCÃO, A. X.; SUZUKI, C. T. N. Supervised pattern classification based on optimum-path forest. **International Journal of Imaging Systems and Technology**, v. 19, n. 2, p. 120–131, 2009.
- PARIKH, K. S.; SHAH, T. P. Support Vector Machine – A Large Margin Classifier to Diagnose Skin Illnesses. **Procedia Technology**, v. 23, p. 369–375, 2016.
- PARSOPOULOS, K. E.; VRAHATIS, M. N. **Particle Swarm Optimization and Intelligence : Advances and Applications**. [s.l: s.n.].
- PAULINO, M. Capítulo IX Análise de resposta em frequência Diagnóstico de transformadores de potência. **O Setor Elétrico**, p. 48–55, 2014a.
- PAULINO, M. Capítulo VII Ensaio de resistência de isolamento e de rigidez dielétrica. **O Setor Elétrico**, p. 56–61, 2014b.
- PAULINO, M. Capítulo VIII Avaliação do isolamento em transformadores de potência Ensaio de perdas dielétricas e capacitância. **O Setor Elétrico**, p. 52–59, 2014c.
- PERRIER, C.; BEROUAL, A.; BESSEDE, J.-L. Improvement of power transformers by using mixtures of mineral oil with synthetic esters. **IEEE Transactions on Dielectrics and Electrical Insulation**, v. 13, n. 3, p. 556–564, 2006.
- PRZYBYLEK, P. A New Concept of Applying Methanol to Dry Cellulose Insulation at the Stage of Manufacturing a Transformer. **Energies**, v. 11, n. 7, p. 1658, 2018.
- QIN, J. F. et al. **Outlier detection for on-line monitoring data of transformer based on wavelet transform and weighted LOF** **Outlier detection for on-line monitoring data of transformer based on wavelet transform and weighted LOF**. 2019 International Conference on New Energy and Future Energy System. **Anais...**2019.
- RAUDYS, S. J.; JAIN, A. K. Small Sample Size Effects in Statistical Pattern Recognition: Recommendations for Practitioners. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 13, n. 1, p. 252–264, 1991.
- RODRIGUES, D. et al. A wrapper approach for feature selection based on Bat Algorithm and Optimum-Path Forest. **Expert Systems with Applications**, v. 41, n. 5, p. 2250–2258, 2014.
- ROUSE, T. O. Mineral Insulating Oil in Transformers. **IEEE Electrical Insulation Magazine**, v. 14, n. 3, p. 6–16, 1998.
- RYCROFT, M. Vegetable oil as insulating fluid for transformers. **Energize**, n. April, p. 37–40, 2014.
- SAHRI, Z.; YUSOF, R. **Fault diagnosis of power transformer using optimally selected**

- DGA features and SVM.** 2015 10th Asian Control Conference (ASCC). *Anais...IEEE*, 2015.
- SALESI, S.; COSMA, G. **A novel extended binary cuckoo search algorithm for feature selection.** 2017 2nd International Conference on Knowledge Engineering and Applications, ICKEA 2017. *Anais...2017*.
- SAMIRMI, F. D.; TANG, W.; WU, H. **Feature selection in power transformer fault diagnosis based on dissolved gas analysis.** 2013 4th IEEE/PES Innovative Smart Grid Technologies Europe, ISGT Europe 2013. *Anais...IEEE*, 2013.
- SÁNCHEZ-MAROÑO, N.; ALONSO-BETANZOS, A.; CASTILLO, E. **A new wrapper method for feature subset selection.** ESANN 2005 Proceedings - 13th European Symposium on Artificial Neural Networks. *Anais...2005*.
- SANTOS, F. DA C. **Resíduos Impregnados com Óleo Mineral isolante na CEMIG : avaliação das técnicas de compostagem , fitorremediação e processo oxidativo para seu tratamento.** [s.l.] Tese (Doutorado em Geografia), Universidade Federal de Uberlândia, 2013.
- SARMA, D. V. S. S. S.; KALYANI, G. N. S. **ANN Approach for Condition Monitoring of Power Transformers Using DGA.** 2004 IEEE Region 10 Conference. *Anais...IEEE*, 2004.
- SARMA, T. H. et al. An improvement to k-nearest neighbor classifier. **2011 IEEE Recent Advances in Intelligent Computational Systems**, p. 227–231, 2011.
- SCATIGGIO, F.; POMPILI, M.; CALACARA, L. Transformers Fleet Management Through the use of an Advanced Health Index. **2018 IEEE Electrical Insulation Conference, EIC 2018**, n. June, p. 395–397, 2018.
- SHAWE-TAYLOR, J.; SUN, S. A review of optimization methodologies in support vector machines. **Neurocomputing**, v. 74, n. 17, p. 3609–3618, 2011.
- SHINTEMIROV, A.; TANG, W.; WU, Q. H. Power transformer fault classification based on dissolved gas analysis by implementing bootstrap and genetic programming. **IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews**, v. 39, n. 1, p. 69–79, 2009.
- SHIYONG, L. et al. **Transformer Familial Defect Detection Method Based on SVM Improved by Apriori Algorithm - Based on Analysis of Dissolved Gases in Oil.** 2019 2nd International Conference on Electrical Materials and Power Equipment. *Anais...IEEE*, 2019
- SILVA, L. E. B. DA. **Identificação de Padrões de Contaminação em Óleo de Transformador Através da Monitoração Acústica das Descargas Parciais.** [s.l.] Tese (Doutorado em Ciências), Universidade Federal de Itajubá, 2011.
- SILVA, J. J. **Combinação de Classificadores Floresta de Caminhos Ótimos aplicados no Reconhecimento Facial.** [s.l.] Dissertação (Mestrado em Ciência da Computação), Faculdade Campo Limpo Paulista, 2016.
- SINGH, A.; UPADHYAY, G. **Dissolved Gas Analysis of power transformer using K-means and Support Vector Machine.** 1st IEEE International Conference on Power Electronics, Intelligent Control and Energy Systems. *Anais...IEEE*, 2017.

- SKELLY, D. **Photo-acoustic spectroscopy for dissolved gas analysis: Benefits and experience.** International Conference on Condition Monitoring and Diagnosis. **Anais...**2012.
- SOKOLOV, V. **Understanding failure modes of transformers.** Euro TechCon 2005. **Anais...**2005.
- SONEJI, H.; SANGHVI, R. C. **Towards the improvement of Cuckoo search algorithm.** Proceedings of the 2012 World Congress on Information and Communication Technologies, WICT 2012. **Anais...**2012.
- SONG, L.; XIU-YING, L.; WEN-XU, W. **Fault diagnosis of transformer based on probabilistic neural network.** Proceedings - 4th International Conference on Intelligent Computation Technology and Automation. **Anais...IEEE**, 2011.
- SONI, R.; CHAUDHARI, H. **An Approach to Diagnose Incipient Faults of Power Transformer Using Dissolved Gas Analysis of Mineral Oil by Ratio Methods Using Fuzzy Logic.** International conference on Signal Processing, Communication, Power and Embedded System. **Anais...SCOPES**, 2016.
- SOUZA, A. N. et al. **Fast fault diagnosis in power transformers using optimum-path forest.** INES 2012 - IEEE 16th International Conference on Intelligent Engineering Systems, Proceedings. **Anais...IEEE**, 2012.
- SOUZA, R.; LOTUFO, R.; RITTNER, L. **A comparison between optimum-path forest and  $\kappa$ -nearest neighbors classifiers.** Brazilian Symposium of Computer Graphic and Image Processing. **Anais...**2012.
- SOUZA, R.; RITTNER, L.; LOTUFO, R. A comparison between k-Optimum Path Forest and k-Nearest Neighbors supervised classifiers. **Pattern Recognition Letters**, v. 39, n. 1, p. 2–10, 2014.
- SOUSA, J. C. S.; SAAVEDRA, O. R.; LIMA, S. L. Decision Making in Emergency Operation for Power Transformers with Regard to Risks and Interruptible Load Contracts. **IEEE Transactions on Power Delivery**, v. 33, n. 4, p. 1556–1564, 2018.
- STEHMAN, S. V. Selecting and Interpreting Measures of Thematic Classification Accuracy. **Remote Sensing of Environment**, v. 62, n. 1, p. 77–89, 1997.
- SU, C. Q. **A New Fuzzy Logic Method for Transformer Incipient Fault Diagnosis.** 2016 IEEE International Conference on Fuzzy Systems. **Anais...IEEE**, 2016.
- SUBROTO, C. et al. **Artificial intelligence for DGA interpretation methods using weighting factor.** ICEMPE 2017 - 1st International Conference on Electrical Materials and Power Equipment. **Anais...IEEE**, 2017.
- SUGUNA, N.; THANUSHKODI, K. An Improved k-Nearest Neighbor Classification Using Genetic Algorithm. **International Journal of Computer Science**, v. 7, n. 4, p. 7–10, 2010.
- SUN, C.; OHODNICKI, P. R.; STEWART, E. M. Chemical Sensing Strategies for Real-Time Monitoring of Transformer Oil : A Review. **IEEE Sensors Journal**, v. 17, n. 18, p. 5786–5806, 2017.

SUN, H.-C.; HUANG, Y.-C.; HUANG, C.-M. **A Review of Dissolved Gas Analysis in Power Transformers**. 2nd International Conference on Advances in Energy Engineering. **Anais...**Elsevier, 2012.

SYAKUR, M. A. et al. **Integration K-Means Clustering Method and Elbow Method for Identification of the Best Customer Profile Cluster**. IOP Conference Series: Materials Science and Engineering. **Anais...**2018.

TALLÓN-BALLESTEROS, A. J.; RIQUELME, J. C. **Deleting or keeping outliers for classifier training?** 2014 6th World Congress on Nature and Biologically Inspired Computing, NaBIC 2014. **Anais...**2014.

TENBOHLEN, S. et al. Diagnostic measurements for power transformers. **Energies**, v. 9, n. 5, p. 1–25, 2016.

USDE. **United States Department of Energy. Large Power Transformers and the U.S. Electric Grid**, 2012. Disponível em: <[https://www.energy.gov/sites/prod/files/Large Power Transformer Study - June 2012\\_0.pdf](https://www.energy.gov/sites/prod/files/Large%20Power%20Transformer%20Study%20-%20June%202012_0.pdf)>

USDI. **United States Department of the Interior. Transformer maintenance (FIST 3-30), U.S. Bureau of Reclamation**, 2000. Disponível em: <[https://www.usbr.gov/power/data/fist/fist3\\_30/fist3\\_30.pdf](https://www.usbr.gov/power/data/fist/fist3_30/fist3_30.pdf)>

VAN DEN BROECK, J. et al. Data cleaning: Detecting, diagnosing, and editing data abnormalities. **PLoS Medicine**, v. 2, n. 10, p. 966–970, 2005.

VENKATASAMI, A.; LATHA, P. Application of Extreme Learning Machine in Fault Classification of Power Transformer. **Circuits and Systems**, v. 07, n. 10, p. 2837–2845, 2016.

VIANA, S. R. **Métodos de Geração de dados sintéticos para a criação de uso publico**. [s.l.] Dissertação (Mestrado em Modelação, Análise de Dados e Sistemas de Apoio à Decisão Orientada), Faculdade de Economia da Universidade do Porto, 2014.

VISHAL et al. **Transformer's History and its Insulating Oil**. Proceedings of the 5th National Conference. **Anais...**New Delhi: 2011.

WANG, P. et al. **Simultaneously analyze fault characteristic gases extracted from transformer oil by Raman spectroscopy**. 2016 IEEE International Conference on High Voltage Engineering and Application. **Anais...**ICHVE, 2016.

WANG, X. et al. Review of Research Progress on the Electrical Properties and Modification of Mineral Insulating Oils Used in Power Transformers. **Energies**, v. 11, p. 487–518, 2018.

WANG, X.; WANG, T.; WANG, B. **Hybrid PSO-BP based probabilistic neural network for power transformer fault diagnosis**. Proceedings - 2008 2nd International Symposium on Intelligent Information Technology Application. **Anais...**IEEE, 2008.

WATTAKAPAIBOON, W.; PATTANADECH, N. **The state of the art for dissolved gas analysis based on interpretation techniques**. International Conference on Condition Monitoring and Diagnosis. **Anais...**IEEE, 2016.



WLODARCZAK, P. **Machine Learning and its Applications**. 1. ed. [s.l.] CRC Press, 2019.

WU, Q.; TANG, W.; WEI, C. Dissolved gas analysis method based on novel feature prioritisation and support vector machine. **IET Electric Power Applications**, v. 8, n. 8, p. 320–328, 2014.

XUE, B.; ZHANG, M.; BROWNE, W. N. Particle swarm optimization for feature selection in classification: A multi-objective approach. **IEEE Transactions on Cybernetics**, v. 43, n. 6, p. 1656–1671, 2013.

YANG, X.-S. Genetic Algorithms. In: **Nature-Inspired Optimization Algorithms**. 1. ed. [s.l.] Elsevier, 2014. p. 77–87.

YANG, X.-S.; DEB, S. **Cuckoo Search via Lévy Flights**. 2009 World Congress on Nature & Biologically Inspired Computing (NaBIC). **Anais...IEEE**, 2009. Disponível em: <http://ieeexplore.ieee.org/document/5393690/>

YANG, X. S.; DEB, S. Cuckoo search: Recent advances and applications. **Neural Computing and Applications**, v. 24, n. 1, p. 169–174, 2014.

YONGQIANG, W. et al. **The fault diagnosis method for power transformer based on BN And DGA**. Proceedings of 2005 International Symposium on Electrical Insulating Materials. **Anais...IEEE**, 2005.

YU, L.; LIU, H. Feature Selection for High-Dimensional Data: A Fast Correlation-Based Filter Solution. **Proceedings, Twentieth International Conference on Machine Learning**, v. 2, p. 856–863, 2003.

YUAN, C. Developing a New Transformer Fault Diagnosis System through Evolutionary Fuzzy Logic. **IEEE Transactions on Power Delivery**, v. 12, n. 2, p. 761–767, 1997.

ZAKARIA, F.; JOHARI, D.; MUSIRIN, I. **Artificial neural network (ANN) application in dissolved gas analysis (DGA) methods for the detection of incipient faults in oil-filled power transformer**. Proceedings - 2012 IEEE International Conference on Control System, Computing and Engineering. **Anais...IEEE**, 2013.

ZHANG, L.; YUAN, J. Fault diagnosis of power transformers using kernel based extreme learning machine with particle swarm optimization. **Applied Mathematics and Information Sciences**, v. 9, n. 2, p. 1003–1010, 2015.

ZHAO, A.-X.; ZHANG, C.-T. **Propagation Neural Network Optimized by Parallel Genetic Algorithm**. IEEE International Conference of IEEE Region 10 (TENCON 2013). **Anais...IEEE**, 2013.

ZHAO, A. X. et al. **The DGA interpretation method using relative content of characteristic gases and gas-ratio combinations for fault diagnosis of oil-immersed power transformers**. Proceedings of the International Symposium on Electrical Insulating Materials. **Anais...ISEIM**, 2014.

ZHU, Y. L.; ZHENG, J. B.; WANG, F. **Transformer fault diagnosis based on Euclidean clustering and support vector machines**. Proceedings of the 2007 International Conference on Wavelet Analysis and Pattern Recognition. **Anais...IEEE**, 2008.

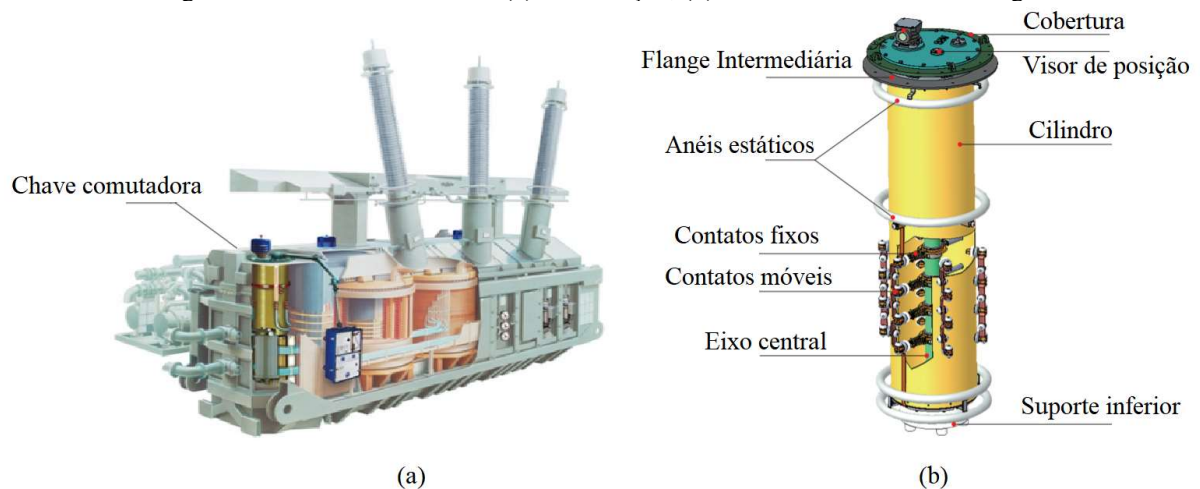
## APÊNDICE A – DEFINIÇÕES

São apresentados alguns esclarecimentos e definições de termos usados nesta tese com o objetivo de facilitar a compreensão do contexto no qual está inserido.

### A.1. Chave comutadora

A chave comutadora de derivação, é um dispositivo eletromecânico que tem a função de elevar ou abaixar as tensões de saída dos transformadores. Isso é feito trocando-se as posições das derivações (taps) existentes nos enrolamentos que alteram a relação de tensão. A chave comutadora pode ser especificada para operar sob carga quando se deseja que a comutação ocorra com o transformador em serviço, ou sem carga, necessariamente quando o transformador está a vazio. As Figuras 29(a) e (b) mostram a localização da chave no transformador e o comutador sem carga respectivamente.

Figura 29 – Chave comutadora: (a) Localização; (b) Chave comutadora sem carga

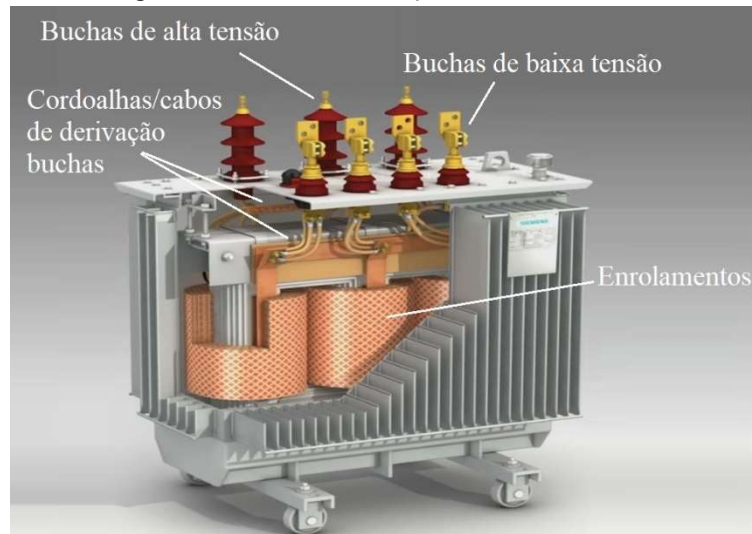


Fonte: Adaptado pelo autor de (a) Briceño (2013) e (b) ABB (2018)

### A.2. Buchas de alta e baixa tensão

As buchas (Figura 30) são elementos que possuem corpo externo em material isolante, e elemento interno condutor que faz à ligação dos terminais de entrada e saída dos enrolamentos dos transformadores, de modo seguro com o meio externo (rede elétrica). Elas podem diferir quanto ao material empregado (por exemplo, resina, porcelana, silicone), conter ou não óleo isolante, diferenciar na forma de construção, entre outros.

Figura 30 – Buchas e derivações no transformador



Fonte: Adaptado pelo autor de Engineering Word (2016)

### A.3. Matriz de confusão ou erros

A matriz de confusão ou erros, é uma medida quantitativa empregada para avaliar o comportamento e desempenho (taxa de acertos) de um algoritmo de aprendizado de máquina na classificação de padrões (ou classes), principalmente quando o conjunto de dados possuem diferentes tipos e quantidades de padrões, ou seja, são desbalanceados (por exemplo, conjunto de dados com 80 amostras do padrão *A* e 20 do padrão *B*). Este recurso é bastante empregado para avaliar a precisão no reconhecimento de imagem, como por exemplo, mapa de cobertura do solo (STEHMAN, 1997), na área da saúde (FAWCETT, 2006), entre outros.

O *layout* usado na apresentação do resultado após classificação dos padrões é na forma de uma matriz quadrada, sendo que cada linha representa as instâncias de um padrão real e cada coluna as instâncias previstas desse padrão, ou ao contrário. Na diagonal principal são colocadas as amostras das instâncias classificadas corretamente e fora dela os erros cometidos pelo classificador. Este recurso dá uma visão geral de como todas essas instâncias foram detalhadamente classificadas, e assim reconhecer facilmente os acertos e erros de todas as instâncias envolvidas e os pontos de maior ocorrência. Além disso, também permite avaliar se as estratégias usadas para melhorar o desempenho do algoritmo estão surtindo efeito. A Tabela 12 mostra um exemplo de matriz de confusão para duas instâncias em sua formação e denominações (FAWCETT, 2006).

Tabela 12 – Matriz de confusão ou erro

		Predição	
		Classe 1	Classe 2
Real	Classe 1	<b>Verdadeiro Positivo (VP)</b>	<b>Falso Negativo (FN)</b>
	Classe 2	<b>Falso Positivo (FP)</b>	<b>Verdadeiro Negativo (VN)</b>

Fonte: Adaptado pelo autor de Fawcett (2006)

Onde: VP trata-se das amostras da instância Classe 1 prevista e observada classificada corretamente; FP trata-se das amostras da Classe 1 prevista e observada classificada incorretamente como Classe 2; VN trata-se das amostras da instância Classe 2 prevista e observada classificada corretamente; e FN trata-se das amostras da Classe 2 prevista e observada classificada incorretamente como Classe 1.

Uma vez constituída a matriz de erros, torna-se possível avaliar os dados segundo algumas métricas conhecidas como:

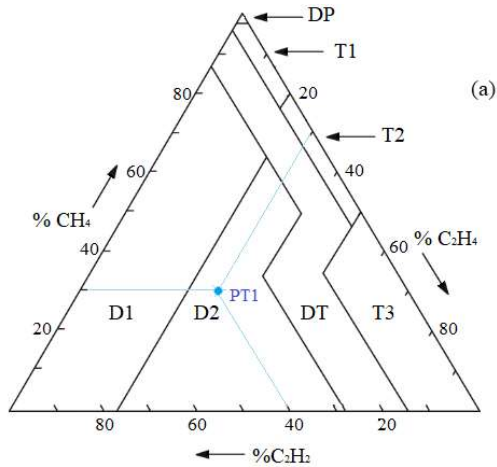
- Exatidão (*accuracy*) retorna a taxa de acertos global de classificação dada por  $(VP+VN)/(VP+VN+FP+FN)$ ;
- Precisão (*precision*) retorna os resultados mais relevantes das classes individuais dado por: Precisão classe 1 =  $VP/(VP+FN)$  e Precisão classe 2 =  $VN/(VN+FP)$ ;
- Sensibilidade (*recall*) retorna parcialmente os resultados de todas as classes relevantes dada por: Sensibilidade classe 1 =  $VP/(VP+FP)$  e Sensibilidade classe 2 =  $VN/(VN+FN)$ ;
- Pontuação F1 (*F1 score*) onde a sensibilidade e a precisão são ponderadas uniformemente pela média da harmônica das mesmas, sendo comum seu emprego para avaliar a qualidade do classificador, dado por:  $F1\ score = (2VP)/(2VP+FP+FN)$ .

Caso exista um número maior de classes, essas passam a ser inseridas como novas linhas e colunas na matriz, e os resultados de FP e FN são obtidos com as somatórias dos termos relacionados.

**APÊNDICE B – TABELAS E FIGURAS DOS MÉTODOS NORMALIZADOS**

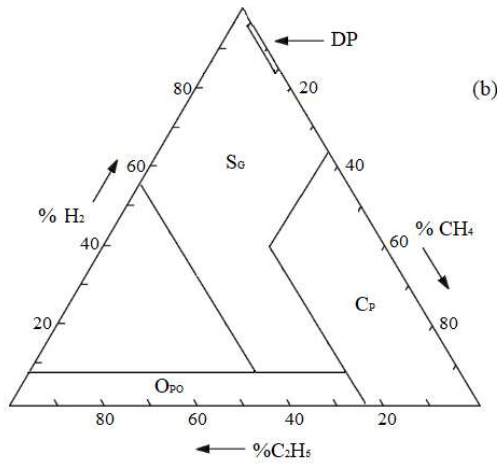
**B.1. Método dos triângulos de falhas de Duval para óleo mineral sem chave comutadora sob carga**

Figura 31 – Triângulo de Duval: (a) N°1 falhas elétricas e térmicas; (b) N°4 falhas de baixa temperatura; e (c) N°5 falhas de alta temperatura



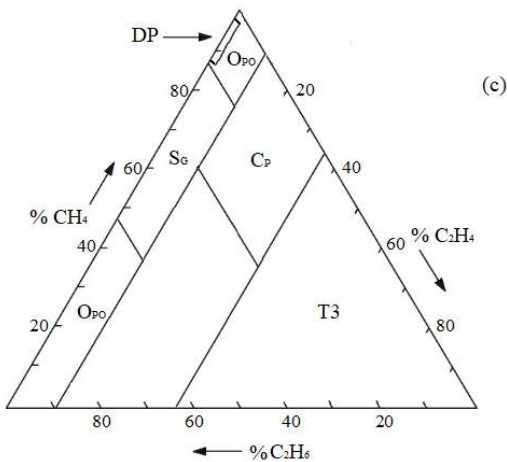
Legenda

- T1 Falha térmica  $T < 300^{\circ}\text{C}$
- T2 Falha térmica  $300 < T < 700^{\circ}\text{C}$
- T3 Falha térmica  $T > 700^{\circ}\text{C}$
- DP Descargas parciais
- DT Misto de falha elétrica e térmica
- D1 Descarga de baixa energia
- D2 Descargas de alta energia



Legenda

- DP Descargas parciais
- SG Stray gassing (gás parasita)
- CP Possível carbonização papel (pontos quentes  $T > 300^{\circ}\text{C}$ )
- OPO Superaquecimento do papel ou óleo



Legenda

- DP Descargas parciais
- SG Stray gassing (gás parasita)
- CP Possível carbonização papel (pontos quentes  $T > 300^{\circ}\text{C}$ )
- OPO Superaquecimento do papel ou óleo
- T3 Falha térmica  $T > 700^{\circ}\text{C}$

Fonte: Adaptado pelo autor de Duval (2008a)

Equações para cálculo dos valores relativos dos gases para o triângulo nº 1. Na Figura 31(a) é mostrado a localização do ponto PT1, identificado como falha D2, para os percentuais de gases %CH<sub>4</sub>=30, %C<sub>2</sub>H<sub>2</sub>=40, %C<sub>2</sub>H<sub>4</sub>=30.

$$\Delta_1 = \text{CH}_4 + \text{C}_2\text{H}_2 + \text{C}_2\text{H}_4 \text{ (valores em } \mu\text{l/l ou PPM)} \quad (10)$$

$$\% \text{CH}_4 = \frac{100 * \text{CH}_4}{\Delta_1} \quad (11)$$

$$\% \text{C}_2\text{H}_2 = \frac{100 * \text{C}_2\text{H}_2}{\Delta_1} \quad (12)$$

$$\% \text{C}_2\text{H}_4 = \frac{100 * \text{C}_2\text{H}_4}{\Delta_1} \quad (13)$$

Equações para cálculo dos valores relativos dos gases para o triângulo nº 4.

$$\Delta_4 = \text{H}_2 + \text{CH}_4 + \text{C}_2\text{H}_6 \text{ (valores em } \mu\text{l/l ou PPM)} \quad (14)$$

$$\% \text{H}_2 = \frac{100 * \text{H}_2}{\Delta_4} \quad (15)$$

$$\% \text{CH}_4 = \frac{100 * \text{CH}_4}{\Delta_4} \quad (16)$$

$$\% \text{C}_2\text{H}_6 = \frac{100 * \text{C}_2\text{H}_6}{\Delta_4} \quad (17)$$

Equações para cálculo dos valores relativos dos gases para o triângulo nº 5.

$$\Delta_5 = \text{CH}_4 + \text{C}_2\text{H}_4 + \text{C}_2\text{H}_6 \text{ (valores em } \mu\text{l/l ou PPM)} \quad (18)$$

$$\% \text{CH}_4 = \frac{100 * \text{CH}_4}{\Delta_5} \quad (19)$$

$$\% \text{C}_2\text{H}_4 = \frac{100 * \text{C}_2\text{H}_4}{\Delta_5} \quad (20)$$

$$\% \text{C}_2\text{H}_6 = \frac{100 * \text{C}_2\text{H}_6}{\Delta_5} \quad (21)$$

## B.2. Valores limites das razões dos gases do método de Rogers

Tabela 13 – Relação dos gases dissolvidos no óleo para o método de Rogers

Caso	$R_2=C_2H_2/C_2H_4$	$R_1=CH_4/H_2$	$R_5=C_2H_4/C_2H_6$	Diagnóstico de falha sugerida
0	< 0,1	Entre 0,1 e 1,0	< 0,1	Normal
1	< 0,1	< 0,1	< 0,1	Descarga de baixa energia – Descarga Parcial <sup>1</sup>
2	Entre 0,1 e 3,0	Entre 0,1 e 1,0	> 3,0	Descarga de alta energia – Arco
3	< 0,1	Entre 0,1 e 1,0	Entre 0,1 e 3,0	Falha térmica de baixa temperatura
4	< 0,1	> 1,0	Entre 0,1 e 3,0	Falha térmica < 700°C
5	< 0,1	> 1,0	> 3,0	Falha térmica > 700°C

<sup>1</sup> Existe a tendência de  $R_2$  e  $R_5$  aumentarem acima de 3 com o desenvolvimento da intensidade da descarga

Fonte: Adaptado pelo autor de IEEE C57.104 (2019)

## B.3. Valores limites das concentrações de gases normais e das razões dos gases do método de IEC 60599

Tabela 14 – Relação dos gases dissolvidos no óleo IEC 60599 – Transformadores

Caso	Diagnóstico de falha sugerida	$R_2=C_2H_2/C_2H_4$	$R_1=CH_4/H_2$	$R_5=C_2H_2/C_2H_6$
DP	Descarga parcial	NS	< 0,1	< 0,2
D1	Descarga de baixa energia	> 1,0	Entre 0,1 e 0,5	> 1,0
D2	Descarga de alta energia	Entre 0,6 e 2,5	Entre 0,1 e 1,0	> 2,0
T1	Falha térmica $T < 300^\circ\text{C}$	Não significativa	> 1,0 (porém, NS)	< 1,0
T2	Falha térmica $300 < T < 700^\circ\text{C}$	< 0,1	> 1,0	Entre 1 e 4
T3	Falha térmica $T > 700^\circ\text{C}$	< 0,2 <sup>1</sup>	> 1,0	> 4,0

Nota 1 Condição para cálculo das razões:

- Calcular quando pelo menos um dos gases está com concentração e a taxa de crescimento acima dos valores típicos definidos pelo especialista (como por exemplo, valor abaixo de 90% da concentração normal e taxa de crescimento de 10% ao mês);
- Mesmo que os valores analisados estejam abaixo dos valores típicos normais, porém, com concentrações crescentes e anormais de gases, recomenda-se aplicar o cálculo;
- Evite calcular as razões quando os valores não forem suficientemente representativos conforme IEC 60567 (IEC, 2005).

Nota 2 Padrão de degradação dos gases similar ao encontrado para descarga parcial tem sido relatadas como resultado de *stray gassing* (gás parasita) dissolvido no óleo.

Nota 3 <sup>1</sup> Um aumento da concentração de  $C_2H_2$  pode indicar pontos quentes com  $T > 1000^\circ\text{C}$ .

Nota 4 NS - Não significativa ou pode ser qualquer valor.

Fonte: Adaptado pelo autor de IEC (2015)

#### B.4. Valores limites das concentrações de gases normais e das razões dos gases do método de Doernenburg

Tabela 15 – Limite L1 de concentrações de gases da IEEE C57.104 para o método Doernenburg

<b>Tipo de gás</b>	<b>H<sub>2</sub></b>	<b>CH<sub>4</sub></b>	<b>C<sub>2</sub>H<sub>2</sub></b>	<b>C<sub>2</sub>H<sub>4</sub></b>	<b>C<sub>2</sub>H<sub>6</sub></b>	<b>CO</b>
Concentrações (µl/l ou PPM)	100	120	1	50	65	350

Fonte: Adaptado pelo autor de IEEE C57.104 (2019)

Tabela 16 – Relação dos gases dissolvidos no óleo para o método Doernenburg

<b>Diagnóstico de falha sugerida</b>	<b>R<sub>1</sub>=CH<sub>4</sub>/H<sub>2</sub></b>	<b>R<sub>2</sub>=C<sub>2</sub>H<sub>2</sub>/C<sub>2</sub>H<sub>4</sub></b>	<b>R<sub>3</sub>=C<sub>2</sub>H<sub>2</sub>/CH<sub>4</sub></b>	<b>R<sub>4</sub>=C<sub>2</sub>H<sub>6</sub>/C<sub>2</sub>H<sub>2</sub></b>
Decomposição térmica	>1,0	<0,75	<0,3	>0,4
Descarga parcial baixa energia	<0,1	Não significativa	<0,3	>0,4
Descarga de alta energia (arco)	>0,1 - <1,0	>0,75	>0,3	<0,4

Fonte: Adaptado pelo autor de IEEE C57.104 (2019)



## APÊNDICE C – ANÁLISE DE OUTROS TRANSFORMADORES

Apresenta-se a seguir as análises dos históricos de gases de outros transformadores com o método proposto.

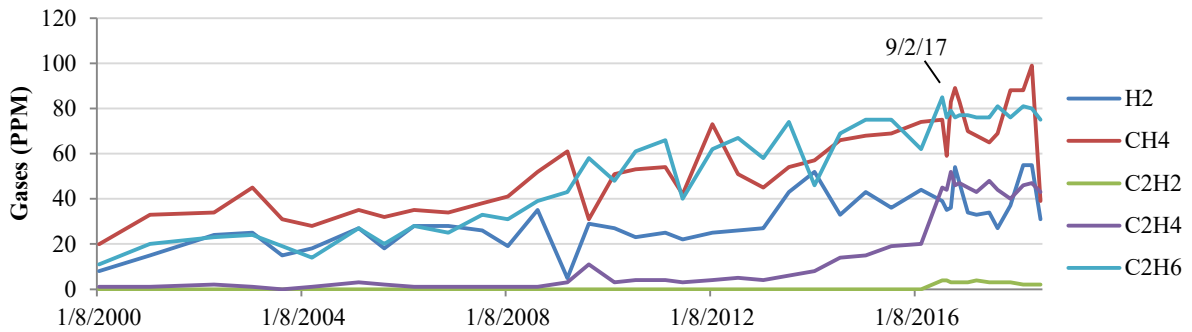
### C.1. Transformador TR7

A Figura 32(a) mostra o registro de concentrações dos gases (41 amostras) durante o período de serviço do transformador de 138/13,8kV/33,33MVA, compreendido entre 07/08/2000 e 02/01/2019. Este é um caso de equipamento operando normalmente no sistema elétrico no período citado.

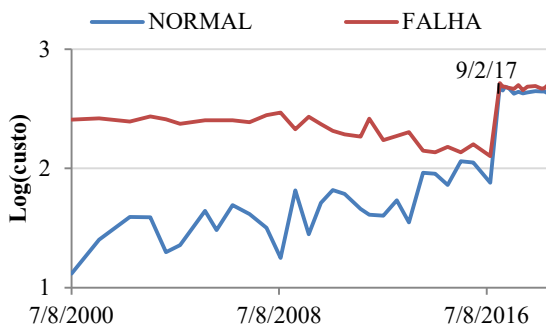
Analisando-se a Figura 32(a), é possível observar o crescimento lento e gradativo dos gases  $H_2$ ,  $CH_4$  e  $C_2H_6$ , acompanhado pela tendência de aproximação entre as curvas normal e falha (Figuras 32(b) e (c)). Com o surgimento de traços de  $C_2H_2$  em 09/02/2017, e presença principal de  $CH_4$  e  $C_2H_6$  (combinação característica de falha por aquecimento), há uma mudança no comportamento das curvas das Figuras 32(d) e (e) fazendo com que a curva normal se aproxime da falha térmica bruscamente no intervalo de 4 meses, ou especificamente a curva primária normal se aproxime da secundária falha térmica  $T > 700^\circ C$  (Figuras 32(f) e (g)), e assim qualquer mudança nas concentrações dos gases pode levar para condição de falha presente. Note que antes da presença do gás  $C_2H_2$  em 09/02/2017, a tendência indicava aproximação da curva normal com falha térmica  $T < 700^\circ C$  (Figuras 32(f) e (g)) compatível com as concentrações de gases encontradas.

Nesse período, o procedimento adotado pela concessionária foi de colocar este transformador em estado de observação, reduzindo seus intervalos de análise DGA, e observando ao mesmo tempo a evolução dos gases, para efeito da tomada de decisão de realizar ou não de inspeção. Na última medição de DGA em 02/01/2019, ocorreu o decréscimo de  $H_2$ ,  $CH_4$  e  $C_2H_2$  (Figura 32(a)) e o início de distanciamento entre a curva normal e falha, indicando tendência de normalização operacional. Como procedimento, outra medida deve ser realizada após, por exemplo, 1 mês para avaliar se o efeito da redução nas concentrações dos gases permaneceu ou se tratou de um possível ruído. As reduções nas concentrações de gases podem ser decorrentes de uma redução no carregamento do transformador, quando operando normalmente no sistema, ou pela interrupção momentânea (ou até definitiva) da falha interna.

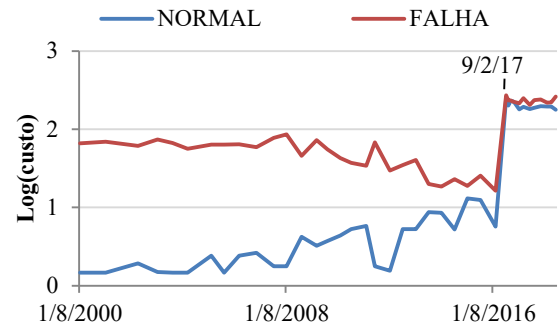
Figura 32 – Curvas de concentrações de gases e tendências de falhas no transformador TR7



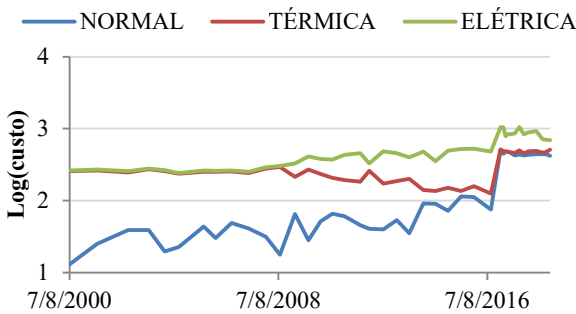
(a) Concentração de gases vs. Data DGA



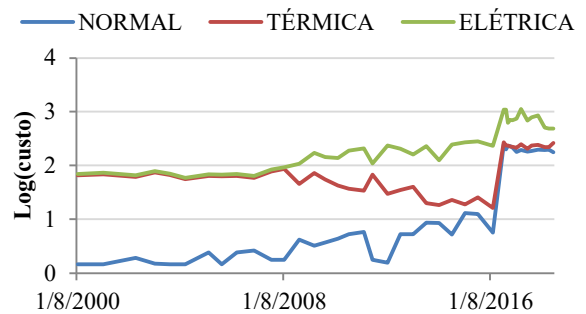
(b) Normal vs. Falhas 1NN



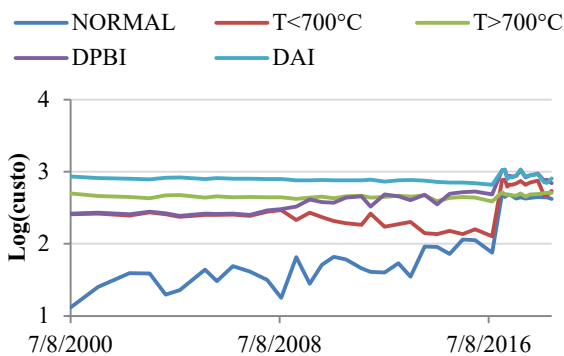
(c) Normal vs. Falhas OPF



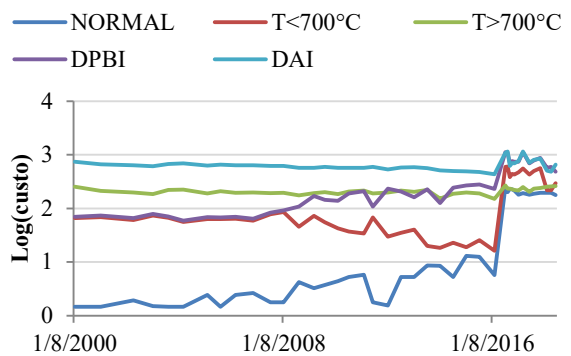
(d) Normal vs. Falha Térmica/Elétrica 1NN



(e) Normal vs. Falha Térmica/Elétrica OPF



(f) Normal vs. 5 Falhas 1NN



(g) Normal vs. 5 Falhas OPF

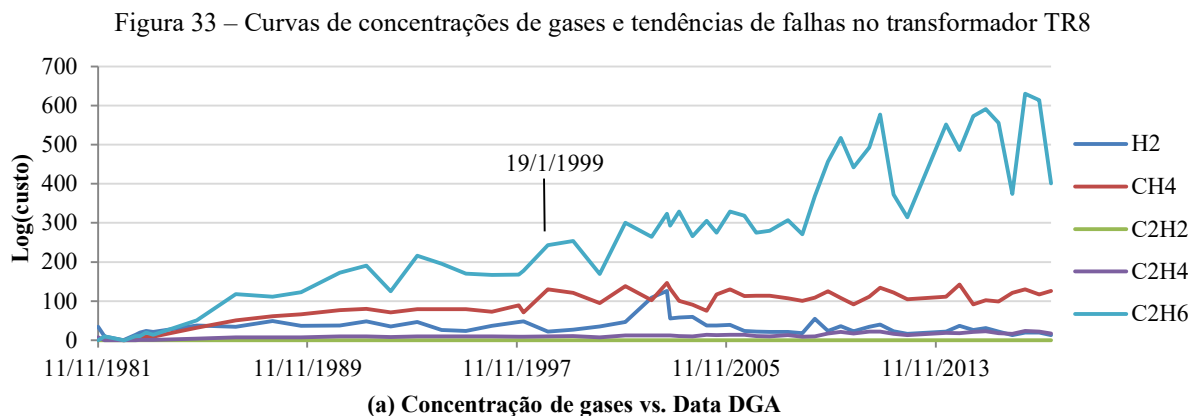
Fonte: Elaborado pelo autor

## C.2. Transformador TR8

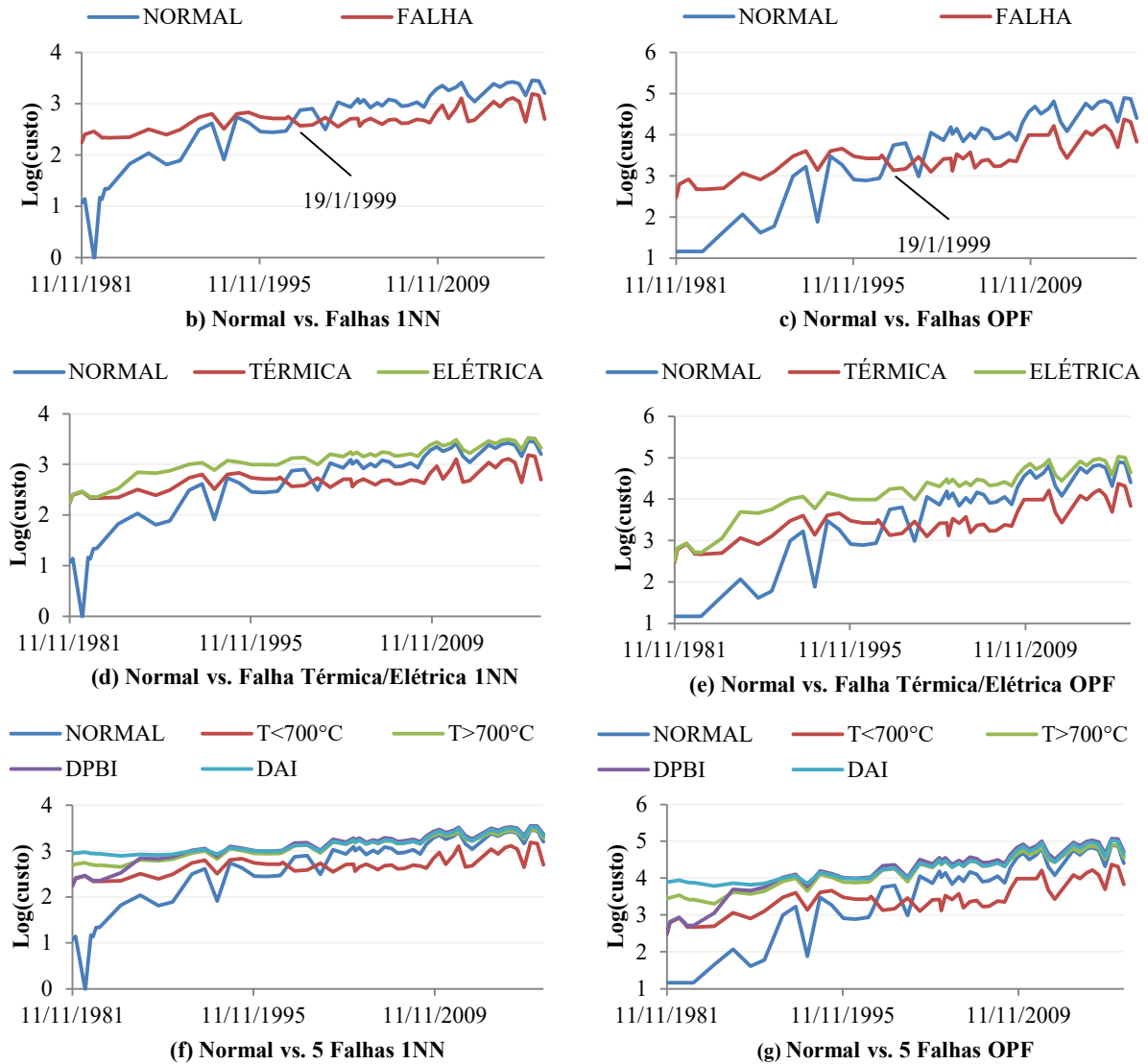
A Figura 33(a) mostra o registro de concentrações dos gases (53 amostras) durante o período operacional do transformador de 138/13,8kV/18,75MVA, compreendido entre 11/11/1981 e 08/04/2018. Este é um caso de equipamento operando normalmente no sistema elétrico no período em questão. Analisando-se a Figura 33(a), é possível observar que desde 1981 se destaca apenas o crescimento do gás  $C_2H_6$ , acompanhado de certa quantidade de  $CH_4$  e pequena de  $H_2$ , ambos estáveis. A presença desses gases normalmente é associada a degradação do próprio óleo decorrente de possíveis carregamentos que elevam a temperatura interna, porém, inferior a  $300^\circ C$ .

Na análise do gráfico de evolução das falhas, nota-se que a partir de 19/01/1999 passou a indicar presença de falha (Figuras 33(b) e (c)), do tipo térmica (Figuras 33(d) e (e)) e para temperatura como condição primária  $T < 700^\circ C$  e secundária normal (Figuras 33(f) e (g)). Note que, no conjunto de dados usados no treinamento do OPF e INN, a faixa da classe  $T < 700^\circ C$  inclui gases produzidos para  $T < 300^\circ C$  e *stray gassing* que surgem entre  $120^\circ C$  e  $200^\circ C$ , além disso, falhas térmicas para  $T < 300^\circ C$  não são caracterizadas como de alto risco.

Quanto ao diagnóstico encontrado pelo método proposto, pode ser considerado coerente com a evolução (tendência) dos gases produzidos. Partindo-se do pressuposto de que houve falha térmica, para efeito de análise das amostras após 19/01/1999, aplicou-se os métodos do Pentágono 1 e 2 de Duval (IEC, 2015) que prevê falhas térmicas de baixa temperatura (obtidos através ensaios de aquecimento realizados em laboratório) e obteve-se classificações de todas para temperatura  $T < 250^\circ C$ . Ambos OPF e INN apresentaram resultados similares, com pequenas diferenças, mas que não impactam na mudança do resultado.



Conclusão Figura 33 – Curvas de concentrações de gases e tendências de falhas no transformador TR8

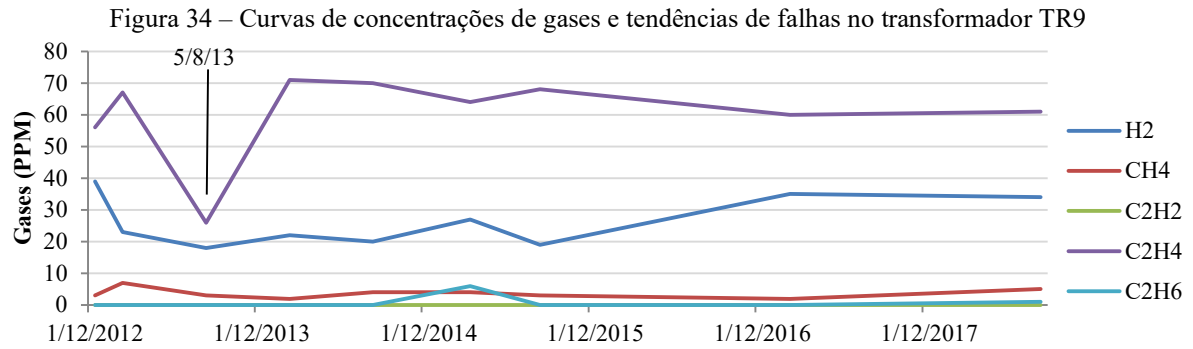


Fonte: Elaborado pelo autor

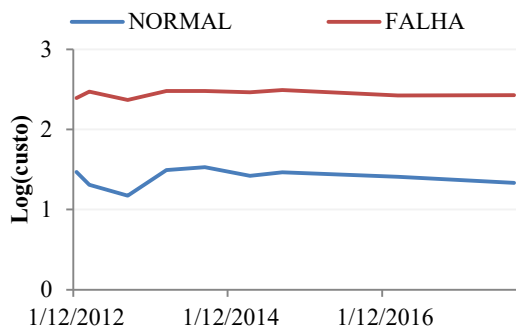
### C.3. Transformador TR9

Este transformador de 138/69kV/33MVA, encontra-se operando normalmente no sistema elétrico e sem relatos de anormalidades ou ocorrências entre o período de 26/12/2012 e 16/08/2018. Os valores dos gases (9 amostras) se encontram baixos e estáveis (Figura 34(a)), refletindo assim na condição de estado normal e distante das classes de falhas como mostram os gráficos de evoluções de falhas (Figuras 34(b) a (g)) para o período analisado. A presença do gás  $C_2H_4$  tem forte relação com o aquecimento normal do óleo. Em 05/08/2013 houve um decréscimo acentuado do valor de  $C_2H_4$ , porém, como o sentido foi de queda e os demais gases se mantiveram estáveis e baixos considerou-se como normal, contudo, uma nova medição poderia ser feita a critério da concessionária para avaliar se a medição estava

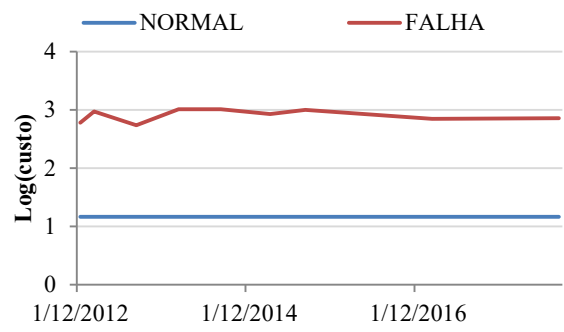
correta. As pequenas diferenças entre os gráficos de falhas entre os algoritmos OPF e 1NN, se devem a heurística de classificação empregada, contudo não impactando no resultado da análise final.



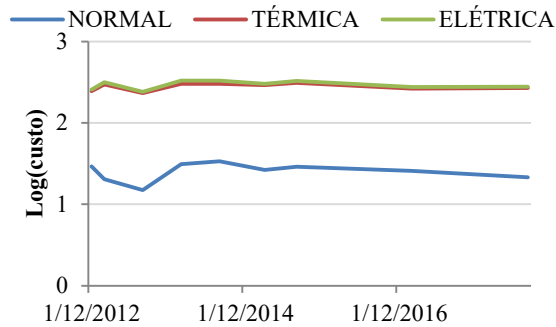
(a) Concentração de gases vs. Data DGA



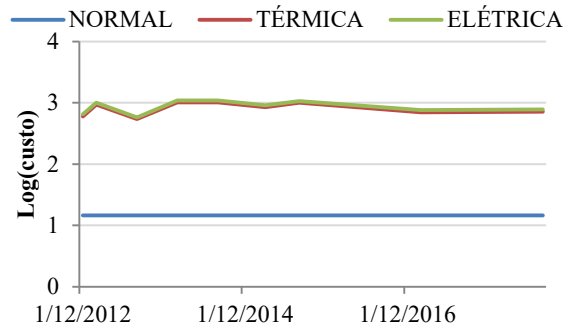
(b) Normal vs. Falhas 1NN



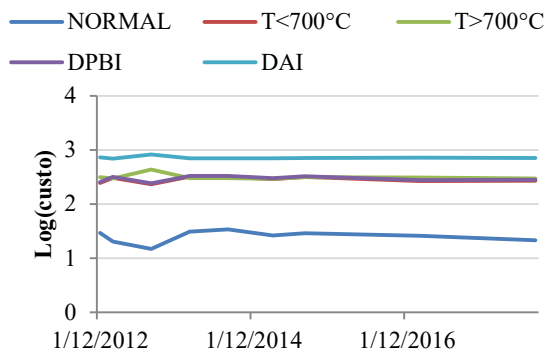
(c) Normal vs. Falhas OPF



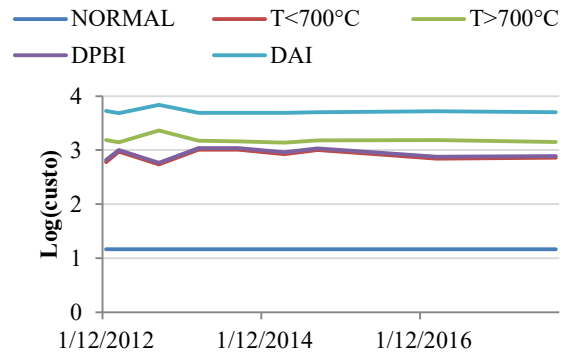
(d) Normal vs. Falha Térmica/Elétrica 1NN



(e) Normal vs. Falha Térmica/Elétrica OPF



(f) Normal vs. 5 Falhas 1NN



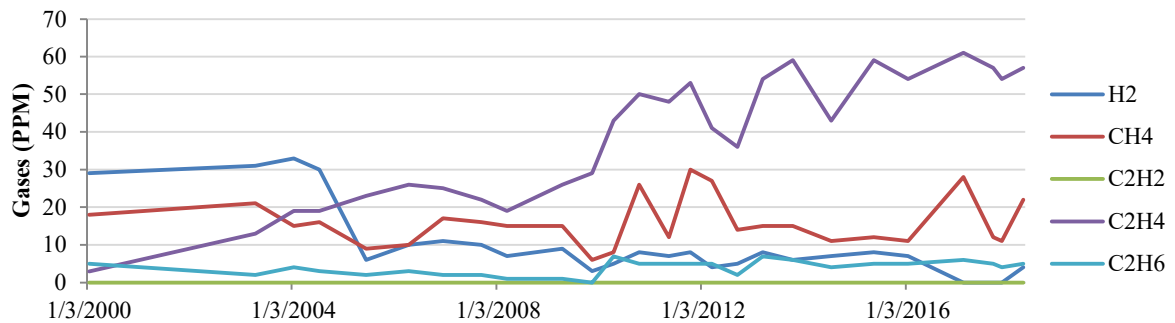
(g) Normal vs. 5 Falhas OPF

Fonte: Elaborado pelo autor

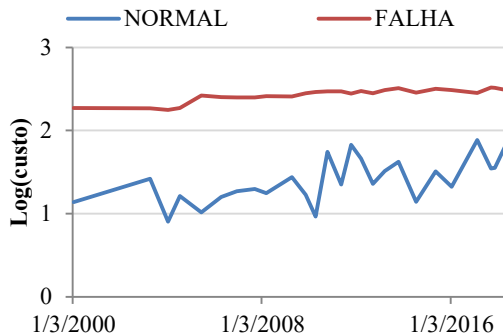
#### C.4. Transformador TR10

Este transformador de 34/13,8kV, encontra-se operando normalmente no sistema elétrico e sem relatos de anormalidades ou ocorrências entre o período de 13/03/2000 e 25/06/2018. Analisando o comportamento da evolução dos gases (26 amostras), os valores se encontram baixos e praticamente estáveis nas últimas medições (Figura 35(a)), refletindo assim na condição de estado normal nos gráficos de evoluções de falhas (Figuras 35(b) a (g)). Comparando-se os resultados entre os algoritmos 1NN e OPF, novamente as pequenas diferenças se devem a heurística de classificação, sendo possível notar uma maior suavidade na curva normal do OPF.

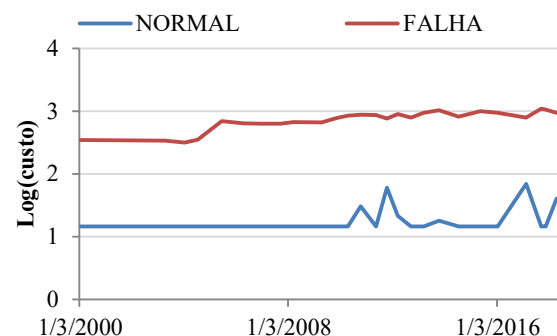
Figura 35 – Curvas de concentrações de gases e tendências de falhas no transformador TR10



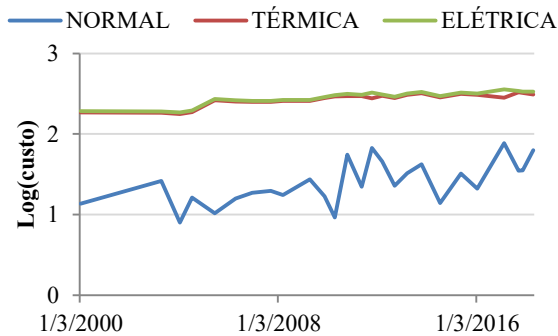
(a) Concentração de gases vs. Data DGA



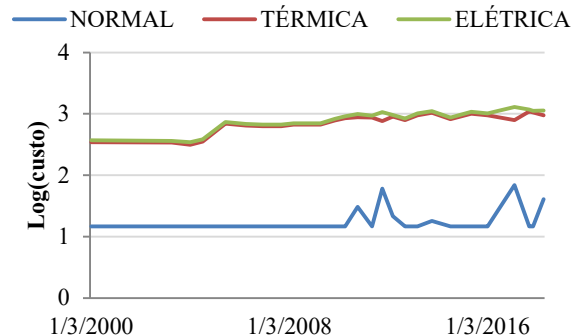
(b) Normal vs. Falhas 1NN



(c) Normal vs. Falhas OPF



(d) Normal vs. Falha Térmica/Elétrica 1NN



(e) Normal vs. Falha Térmica/Elétrica OPF

Conclusão Figura 35 – Curvas de concentrações de gases e tendências de falhas no transformador TR10

