



Thiago Baptista Rodrigues

**Detecção, separação e classificação
de sinais de descargas parciais
em isolamentos de alta tensão**

Tese de Doutorado

Tese apresentada como requisito parcial para obtenção
do grau de Doutor pelo Programa de Pós-Graduação
em Engenharia Elétrica da PUC-Rio.

Orientador: Prof. Ricardo Tanscheit

Co-orientadora: Prof^ª. Marley Maria Bernardes Rebuszi Vellasco

Rio de Janeiro

Setembro de 2019



Thiago Baptista Rodrigues

Detecção, separação e classificação de sinais de descargas parciais em isolamentos de alta tensão

Tese apresentada como requisito parcial para obtenção do grau de Doutor pelo Programa de Pós-Graduação em Engenharia Elétrica da PUC-Rio. Aprovada pela Comissão Examinadora abaixo.

Prof. Ricardo Tanscheit

Orientador

Departamento de Engenharia Elétrica – PUC-Rio

Prof.^a Marley Maria Bernardes Rebuszi Vellasco

Co-orientadora

Departamento de Engenharia Elétrica – PUC-Rio

Prof.^a Karla Tereza Figueiredo Leite

UERJ

Prof. Carlos Roberto Hall Barbosa

Programa de Pós-Graduação em Metrologia – PUC-Rio

Prof. Hélio de Paiva Amorim Junior

CEPEL

Prof. Ronaldo Bosignoli

UERJ

Prof. José Franco Machado do Amaral

UERJ

Rio de Janeiro, 13 de setembro de 2019

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Thiago Baptista Rodrigues

Possui graduação em Engenharia Elétrica pela Universidade Federal de Juiz de Fora (2004) e mestrado em Engenharia Elétrica pela Pontifícia Universidade Católica do Rio de Janeiro (2007). Atualmente é pesquisador do Centro de Pesquisas de Energia Elétrica da Eletrobras (Eletrobras Cepel), atuando nas seguintes áreas: desenvolvimento e análise de novas tecnologias em sistemas elétricos de potência, medição e análise de descargas parciais em equipamentos elétricos, ensaios elétricos em campo ou laboratório, diagnóstico de equipamentos elétricos, e simulações matemáticas e computacionais de fenômenos elétricos.

Ficha Catalográfica

Rodrigues, Thiago Baptista

Detecção, separação e classificação de sinais de descargas parciais em isolamentos de alta tensão / Thiago Baptista Rodrigues; orientador: Ricardo Tanscheit; co-orientadora: Marley Maria Bernardes Rebutzi Vellasco. – 2019.

163 f. : il. color. ; 30 cm

Tese (doutorado) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Elétrica, 2019.

Inclui bibliografia

1. Engenharia Elétrica – Teses. 2. Descargas parciais. 3. Para-raios. 4. Transformador de corrente. 5. Gerador. 6. Classificação de padrões. I. Tanscheit, Ricardo. II. Vellasco, Marley Maria Bernardes Rebutzi. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica. IV. Título.

CDD: 621.3

Para meus pais, Francisco (*in memoriam*)
e Rosaly, meu irmão, Raphael,
e minha esposa, Rafaela

Agradecimentos

Ao programa de Pós-Graduação em Engenharia Elétrica da PUC-Rio, pela confiança em mim depositada, e à CAPES pelo auxílio financeiro concedido através da Bolsa-taxas. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Aos meus orientadores, Professor Ricardo Tanscheit e Professora Marley Vellasco, pela paciência, disponibilidade e conhecimento compartilhado, e aos Professores integrantes da Banca examinadora, pela contribuição crítica fundamental ao enriquecimento do trabalho.

Ao Centro de Pesquisas de Energia Elétrica (Cepel), pelo apoio concedido ao longo do curso, e à equipe do Projeto IMA-DP do Cepel, minha segunda família, pelo apoio e companheirismo de sempre.

A Deus, pela força concedida sempre e, principalmente, nos momentos mais difíceis, nos quais nunca me faltou.

Aos meus pais, Francisco Rodrigues (*in memoriam*) e Rosaly Coelho Baptista, pela educação que me proporcionaram e pelo apoio psicológico nas horas em que mais precisei, principalmente no início da jornada acadêmica.

À minha saudosa avó, Arinda Coelho Baptista, por tudo que me ensinou, pelo exemplo de ser humano que foi, e pelo auxílio espiritual que me concede sempre.

Por fim, agradeço à minha amada esposa, Rafaela Cristina Landeiro da Silva Rodrigues, pelo incentivo e ajuda concedidos, imprescindíveis ao longo desse processo, e também por se fazer presente sempre, o que certamente me possibilitou vencer obstáculos que surgiram pelo caminho com maior facilidade.

Resumo

Rodrigues, Thiago Baptista; Tanscheit, Ricardo (orientador); Vellasco, Marley Maria Bernardes Rebuszi (co-orientadora). **Detecção, separação e classificação de sinais de descargas parciais em isolamentos de alta tensão**. Rio de Janeiro, 2019, 163p. Tese de Doutorado - Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

A medição e classificação de descargas parciais constituem uma importante ferramenta de avaliação dos sistemas de isolamento utilizados em equipamentos de alta tensão. Após o pré-processamento dos dados, que captura, digitaliza e filtra o sinal de descargas parciais, geralmente eliminando os ruídos, existem basicamente duas etapas principais, que são a extração de características e a classificação de padrões. As descargas parciais contêm um conjunto de características discriminatórias únicas que lhes permitem ser reconhecidas. Assim, o primeiro procedimento no processo de classificação é definir quais delas podem ser utilizadas e qual o método de extração destas características. O fenômeno de descargas parciais tem uma natureza transitória e é caracterizado por correntes pulsantes com uma duração de vários nanossegundos até poucos microssegundos. Sua magnitude não é sempre proporcional ao dano causado, sendo que descargas de pequena magnitude podem levar rapidamente à evolução de um defeito. Por isso a necessidade de se entender bem este fenômeno e saber interpretar os dados. Além disso, equipamentos de alta tensão de grande porte, como motores e geradores, podem apresentar mais de uma fonte interna de descargas parciais, sendo importante separar os sinais dessas diferentes fontes antes de realizar a classificação. No caso de outros equipamentos de alta tensão de menor porte, como para-raios e transformadores de corrente de subestação, a simples detecção da presença de descargas parciais interna ao equipamento, independente do número de fontes, já é suficiente para indicar a retirada de operação destes equipamentos, dado seu baixo custo relativo e o elevado grau de importância destes para a confiabilidade do sistema onde estão inseridos. Para um diagnóstico completo e confiável de isolamentos de alta tensão, há a demanda por um sistema de análise capaz de promover com eficácia a detecção de descargas parciais internas aos equipamentos, a separação das diversas fontes de descargas parciais, no caso dos equipamentos de grande porte, bem como realizar a correta

classificação do tipo de defeito, com base principalmente na análise das características discriminantes das diferentes fontes e na assinatura dos sinais para os diferentes defeitos. Este estudo contribui para o preenchimento desta lacuna, apresentando metodologias que se mostram robustas e precisas nos testes realizados, de modo que possam efetivamente orientar os especialistas em manutenção na tomada de decisões. Para fazer isso, são propostas novas variáveis capazes de extrair informações relevantes de sinais no tempo medidos em diversos tipos de isolamentos, sendo aplicadas aqui em dados obtidos em campo e em laboratório para avaliar sua eficácia na tarefa. Essas informações são tratadas utilizando técnicas de classificação de padrões e inteligência artificial para determinar de forma automática a presença de descargas parciais, o número de fontes diferentes e o tipo de defeito nos isolamentos de alta tensão utilizados no estudo. Outra contribuição do estudo é a criação de um banco de dados histórico, baseada em processamento de imagem, com padrões de mapas de descargas parciais conhecidos na literatura em máquinas rotativas, para serem utilizados na classificação de novos mapas medidos neste tipo de equipamento.

Palavras-chave

Descargas parciais; para-raios; transformador de corrente; gerador; separação de fontes; classificação de padrões.

Abstract

Rodrigues, Thiago Baptista; Tanscheit, Ricardo (advisor); Vellasco, Marley Maria Bernardes Rebuizi (co-advisor). **Detection, separation e classification of partial discharge signals in high voltage insulations**. Rio de Janeiro, 2019, 163p. Tese de Doutorado - Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

Measurement and classification of partial discharges are an important tool for the evaluation of insulation systems used in high voltage equipments. After pre-processing of data, which captures, scans and filters the signal of partial discharges, generally eliminating noises, there are basically two main steps, which are the extraction of characteristics and the pattern classification. Partial discharges contain a set of unique discriminatory characteristics that allow them to be recognized. Thus, the first procedure in the classification process is to define which of them can be used and which is the method for extraction of those characteristics. The phenomenon of partial discharges has a transient nature and is characterized by pulsating currents with a duration of several nanoseconds up to a few microseconds. Its magnitude is not always proportional to the damage caused, and discharges of small magnitude can quickly lead to the evolution of a failure. Therefore the need to understand this phenomenon well and to know how to interpret the data. In addition, large high voltage equipments such as motors and generators may have more than one internal source of partial discharges, and it is important to separate the signals from those different sources prior to classification. In the case of smaller high voltage equipments, as surge arrester and substation current transformers, the simple detection of the presence of partial discharges inside the equipment, regardless of the number of sources, is sufficient to indicate the withdrawal of operation of the equipment, given their low relative cost and the high degree of importance of these to the reliability of the system where they are part of. For a complete and reliable diagnosis of high voltage insulations, there is a demand for an analysis system capable of effectively promoting the detection of the partial discharges internal to the equipments, the separation of the various sources of partial discharges in the case of large equipments, as well as to carry out the correct classification of the type of failure. The system should be based mainly on the analysis of the discriminating

characteristics of the different sources and the signature of the signals for the different failure. This study contributes to fill this gap by presenting methodologies that are robust and accurate in the tests performed, so that they can effectively guide maintenance specialists in decision making. To do this, new variables are proposed to extract relevant information from time signals measured in various types of insulations, being applied here in field and laboratory data to evaluate their effectiveness in the task. This information is treated using standard classification techniques and artificial intelligence to automatically determine the presence of partial discharges, the number of different sources and the type of defect in the high voltage insulations used in the study. Another contribution of the study is the creation of a historical database, based on image processing, with partial discharge map patterns known in the literature on rotating machines, to be used in the classification of new maps measured in this type of equipment.

Keywords

Partial discharges; surge arrester; current transformer; generator; separation of sources; signal classification.

Sumário

1. Introdução	19
1.1. Motivação	19
1.2. Objetivos	24
1.3. Organização do Texto	25
2. Fundamentos	26
2.1. Equipamentos de Alta Tensão	26
2.2. Descargas Parciais	30
2.2.1. Definições	30
2.2.2. Grandezas Associadas às Descargas Parciais	32
2.2.3. Representação Elétrica das Descargas Parciais	35
2.2.4. Medição de Descargas Parciais	36
2.2.5. Sinais de Descargas Parciais	38
2.2.5.1. Características Associadas ao Tempo	38
2.2.5.2. Características Associadas à Fase	40
2.2.5.3. Características Desassociadas do Tempo e da Fase	42
2.3. Reconhecimento de Padrões em Descargas Parciais	42
2.3.1. Extração de Características dos Sinais	44
2.3.1.1. Métodos Estatísticos	44
2.3.1.2. Caracterização do Sinal no Tempo	46
2.3.1.3. Processamento de Sinais	47
2.3.1.4. Processamento de Imagens	49
2.3.2. Separação das Diferentes Fontes de Descargas Parciais	51
2.3.2.1. Propriedades para uma Separação Efetiva	51
2.3.2.2. Espaços para Separação de Fontes Múltiplas	52
2.3.2.3. Seleção dos <i>Clusters</i>	53
2.3.2.4. Avaliação dos Espaços de Separação	53
2.3.3. Classificação de Padrões em Descargas Parciais	54
2.3.3.1. Classificadores Baseados em Distância	55
2.3.3.2. Classificadores Estatísticos	56
2.3.3.3. Classificadores Baseados em Redes Neurais Artificiais	57

2.3.3.4. Classificadores Baseados em Conceitos de Sistemas Fuzzy	59
2.3.3.5. Classificadores Baseados em Máquina de Vetor Suporte	59
2.4. Panoramas Atuais Relacionados ao Estudo	60
3. Dados e Técnicas Utilizados no Estudo	66
3.1. Isolamentos de Alta Tensão	66
3.2. Sistemas de Medição	71
3.3. Sinais Utilizados	73
3.4. Extração de Características dos Sinais	76
3.5. Agrupamento dos Sinais	84
3.6. Classificação dos Sinais	87
3.6.1. Isolamentos com Classificação Baseada na Assinatura do Sinal no Tempo	87
3.6.2. Isolamentos com Classificação Baseada na Assinatura do Mapa de Descargas Parciais	88
3.6.2.1. Metodologia de Classificação	89
3.6.2.2. Banco de Dados Histórico	90
4. Resultados e Discussão	96
4.1. Utilizando os Sinais dos Para-raios e Transformadores de Corrente	96
4.2. Utilização dos Sinais das Células e da Barra de Gerador	106
4.2.1. Seleção de Características e Agrupamento	106
4.2.2. Classificação dos Mapas de Descargas Parciais	113
4.3. Aplicação das Técnicas Propostas em um Sinal Real com Ruído	119
4.4. Testando a Eficácia da Classificação do Banco de Dados Histórico em Mapas de Descargas Parciais com Classe Conhecida	122
4.5. Adaptação do Banco de Dados Histórico para Mapas de Descargas Parciais Unipolares com Classe Conhecida	126
5. Conclusões e Trabalhos Futuros	129
Referências Bibliográficas	132

Anexo A - Relief	145
Anexo B - Dendrogramas	147
Anexo C - Árvores de Decisão	149
Anexo D - Máquinas de Vetor Suporte	153
D.1 - A Teoria de Aprendizado Estatístico	153
D.2 - SVMs Lineares	155
D.2.1 - SVMs com Margens Rígidas	155
D.2.2 - SVMs com Margens Suaves	159
D.3 - SVMs Não Lineares	161

Lista de Figuras

Figura 1.1 – TC de 345 kV: (a) antes e (b) após o sinistro	21
Figura 1.2 – TC de 500 kV: (a) antes e (b) após o sinistro	22
Figura 1.3 – PR de 500 kV retirado de operação: (a) arranjo para ensaio em laboratório; (b) duas colunas do PR desmontadas, sendo uma com umidade interna (direita) e outra sem (esquerda)	22
Figura 2.1 – Representação elétrica de um material isolante	28
Figura 2.2 – Diagrama fasorial representativo	28
Figura 2.3 – Distribuição de potencial no entorno de uma bucha de alta tensão com material semicondutor próximo à região de potencial nulo	29
Figura 2.4 – Processo de formação das DP e seus efeitos	31
Figura 2.5 – Circuito equivalente simplificado representando um dielétrico com cavidades	35
Figura 2.6 – Circuito de medição de DP pelo método elétrico direto	37
Figura 2.7 – Circuito de medição de DP utilizando TC de alta frequência	37
Figura 2.8 – TC de alta frequência conectado ao cabo de aterramento do equipamento	37
Figura 2.9 – Sinal no tempo de DP em uma célula de teste	38
Figura 2.10 – Sinal no tempo de DP em um gerador	38
Figura 2.11 – Sinal no tempo de DP no ar 01	39
Figura 2.12 – Sinal no tempo de DP no ar 02	39
Figura 2.13 – Sinal no tempo de DP no óleo 01	39
Figura 2.14 – Sinal no tempo de DP no óleo 02	39
Figura 2.15 – Sinais de DP sobre a frequência industrial 01	39
Figura 2.16 – Sinais de DP sobre a frequência industrial 02	40
Figura 2.17 – Sinais de DP sobre a frequência industrial 03	40
Figura 2.18 – Mapa estatístico de DP 01	41
Figura 2.19 – Mapa estatístico de DP 02	41
Figura 2.20 – Mapa estatístico de DP 03	41
Figura 2.21 – Mapa estatístico de DP 04	41

Figura 2.22 – Mapa estatístico de DP 05	41
Figura 2.23 – Modelo do pulso de DP	46
Figura 2.24 – Rede Neural <i>Backpropagation</i>	58
Figura 2.25 – Classificador baseado em lógica fuzzy	59
Figura 3.1 – Células de teste de DP e barra de gerador: (a) célula ponta-plano no ar (PPA); (b) célula com eletrodos flutuantes no ar (EFA); e (c) barra de gerador	67
Figura 3.2 – Foto ilustrativa de um gerador	68
Figura 3.3 – Foto de um dos PRs de 345 kV medidos em campo	68
Figura 3.4 – Foto de um dos TCs de 345 kV medidos em campo	68
Figura 3.5 – Tratamento necessário em cada tipo de isolamento	70
Figura 3.6 – Circuito de medição de DP utilizado nas células de teste em laboratório e no gerador em campo	71
Figura 3.7 – Circuito de medição de DP utilizando TCAF	73
Figura 3.8 – Exemplo de um TCAF conectado ao cabo de aterramento do equipamento	73
Figura 3.9 – Estrutura de uma árvore de decisão	87
Figura 3.10 – Padrão DP1 (Gap) (a) Imagem retirada da norma e (b) Mapa de DP após processamento	92
Figura 3.11 – Padrão DP2 (Ranhura) (a) Imagem retirada da norma e (b) Mapa de DP após processamento	92
Figura 3.12 – Padrão DP3 (Cavidades internas) (a) Imagem retirada da norma e (b) Mapa de DP após processamento	92
Figura 3.13 – Padrão DP4 (Delaminação interna) (a) Imagem retirada da norma e (b) Mapa de DP após processamento	93
Figura 3.14 – Padrão DP5 (Tracking de superfície) (a) Imagem retirada da norma e (b) Mapa de DP após processamento	93
Figura 3.15 – Mapa padrão célula PPA	94
Figura 3.16 – Mapa padrão célula EFA	94
Figura 3.17 – Mapa padrão barra de gerador	94
Figura 3.18 – Mapa padrão ruído	94
Figura 3.19 – DP de Gap entre 2 barras	94
Figura 3.20 – DP de Gap entre a barra e a prensa do núcleo do	94

estator	
Figura 3.21 – DP de ranhura entre a barra e o núcleo do estator	95
Figura 3.22 – DP por tracking de superfície no braço de projeção do enrolamento	95
Figura 3.23 – Corte transversal em uma barra com delaminação	95
Figura 4.1 – Sinais medidos em PRs com e sem DP, sendo:(a) com DP; e (b) sem DP	103
Figura 4.2 – Sinais medidos em TCs com e sem DP, sendo:(a) com DP; e (b) sem DP	103
Figura 4.3 – Interface do classificador de sinais de PRs e TCs desenvolvido	104
Figura 4.4 – Sinais medidos no PR identificado com DP, sendo: (a) Sinal medido no campo; e (b) sinal medido em laboratório	106
Figura 4.5 – Sinais medidos no TC identificado com DP, sendo: (a) Sinal medido no campo; e (b) sinal medido em laboratório	106
Figura 4.6 – Sinais das células plotados em relação às suas classes originais utilizando as variáveis 2, 4 e 5	109
Figura 4.7 – Dendrograma PPA+EFA com 23 variáveis	111
Figura 4.8 – Dendrograma PPA+barra com 23 variáveis	111
Figura 4.9 – Dendrograma EFA+barra com 23 variáveis	111
Figura 4.10 – Dendrograma PPA+EFA+barra com 23 variáveis	111
Figura 4.11 – Arranjo de ensaio montado em laboratório para ensaiar os 3 objetos em paralelo	114
Figura 4.12 – Taxa de acerto na classificação dos mapas de DP separados em função do número de variáveis de entrada e analisando o grau de pertinência mínimo	116
Figura 4.13 – Mapa de DP medido nos ensaio com as duas células e a barra de gerador em paralelo	117
Figura 4.14 – Mapas de DP obtidos após agrupamento do sinal original, sendo a classe de saída: (a) célula PPA; (b) célula EFA; e (c) barra de gerador	118
Figura 4.15 – Histogramas positivo e negativo do número de pontos em relação à fase – $H_n(\text{fase})$	118

Figura 4.16 – Histogramas positivo e negativo da carga máxima em relação à fase – $H_{qmax}(fase)$	119
Figura 4.17 – Histogramas positivo e negativo da carga média em relação à fase – $H_{qmed}(fase)$	119
Figura 4.18 – Mapa de DP original do gerador 1 com ruído	120
Figura 4.19 – Mapas derivados do mapa original medido em campo no gerador 1, sendo: (a) fonte 1 (DP3 – cavidades internas); (b) fonte 2 (ruído); e (c) fonte 3 (ruído)	122
Figura 4.20 – Mapa medido no acoplador 1	124
Figura 4.21 – Mapa medido no acoplador 2	124
Figura 4.22 – Mapa medido no acoplador 3	124
Figura 4.23 – Mapa medido no acoplador 4	124
Figura 4.24 – Mapa medido no acoplador 5	124
Figura 4.25 – Mapa medido no acoplador 6	124
Figura 4.26 – Mapa medido no acoplador 7	125
Figura 4.27 – Mapa medido no acoplador 8	125
Figura 4.28 – Mapa unipolar medido no gerador 3	127
Figura 4.29 – Mapa unipolar medido no gerador 4	127
Figura 4.30 – Mapa unipolar medido no gerador 5	128
Figura 4.31 – Mapa unipolar medido no gerador 6	128
Figura 4.32 – Mapa unipolar medido no motor 1	128
Figura A.1 – Algoritmo do Relief	146
Figura B.1 – Dendrograma formado após processamento dos dados	148
Figura C.1 – Exemplos de testes executados por uma árvore de decisão sobre um atributo nominal (a) e dois numéricos (b e c), sendo citado em cada folha: número de exemplos da classe indicada / número de exemplos da outra classe	151
Figura D.1 – Conjunto de treinamento binário e três diferentes hipóteses	154
Figura D.2 – Cálculo da distância d entre os hiperplanos H_1 e H_2	157
Figura D.3 – (a) Conjunto de dados não linear; (b) Fronteira não linear no espaço de entradas; (c) Fronteira linear no espaço de características	161

Lista de Tabelas

Tabela 2.1 – Resistividade de materiais isolantes e condutores	27
Tabela 2.2 – Rigidez dielétrica de alguns materiais	28
Tabela 2.3 – Resistividade de materiais semicondutores	29
Tabela 2.4 – Níveis de DP aceitáveis	33
Tabela 3.1 – Sinais medidos em laboratório nas células de teste e na barra de gerador	74
Tabela 3.2 – Sinais medidos em campo nos PRs e TCs	75
Tabela 3.3 – Distribuição dos PRs e TCs pela tensão de operação	75
Tabela 3.4 – Sinais medidos em laboratório no PR e no TC retirados de operação	75
Tabela 3.5 – Sinal no tempo medido em campo no gerador 1	75
Tabela 3.6 – Mapas de DP medidos em campo em monitoramento <i>on-line</i>	76
Tabela 3.7 – Mapas de DP medidos em campo em medições <i>off-line</i>	76
Tabela 3.8 – Resumo das variáveis utilizadas no estudo	83
Tabela 3.9 – Resumo dos defeitos encontrados em máquinas rotativas	91
Tabela 4.1 – Resultado do Relief nos PRs em relação às 23 variáveis	97
Tabela 4.2 – Resultado do Relief nos TCs em relação às 23 variáveis	97
Tabela 4.3 – Taxa de acerto na classificação dos PRs utilizando Árvore de Decisão	98
Tabela 4.4 – Taxa de acerto na classificação dos PRs utilizando SVM	99
Tabela 4.5 – Taxa de acerto na classificação dos TCs utilizando Árvore de Decisão	100
Tabela 4.6 – Taxa de acerto na classificação dos TCs utilizando SVM	101
Tabela 4.7 – Melhores percentuais de acerto atingidos nos PRs e	102

TCs	
Tabela 4.8 – Matriz de confusão da melhor classificação nos PRs (SVM/13 variáveis)	102
Tabela 4.9 – Matriz de confusão da melhor classificação nos TCs (árvore de decisão/8 variáveis)	102
Tabela 4.10 – Percentual de acerto do classificador automático nos equipamentos retirados de campo para ensaio no laboratório de AT	105
Tabela 4.11 – Resultado do Relief para as 23 variáveis nas células e na barra de gerador	108
Tabela 4.12 – Influência do limiar de junção na determinação do número de <i>clusters</i> ideal utilizando as 23 variáveis	110
Tabela 4.13 – Taxa de acerto do FCM para 3 <i>clusters</i> (classes conhecidas - objetos ensaiados separadamente e unidos depois)	112
Tabela 4.14 – Taxa de acerto na classificação dos mapas de DP separados dos sinais medidos em paralelo nos objetos em laboratório	115
Tabela 4.15 – Tipo de defeito detectado em torno de cada acoplador e grau de risco associado	123
Tabela 4.16 – Taxa de Acerto em cada acoplador	125
Tabela 4.17 – Classes reais dos mapas avaliados	127
Tabela B.1 – Dados dos indivíduos	147
Tabela B.2 – Níveis de junção	148
Tabela C.1 – Pequeno conjunto de exemplos de treinamento	150
Tabela D.1 – Funções Kernel mais comuns	163

1

Introdução

1.1

Motivação

O isolamento elétrico é uma parte significativa em todos os equipamentos de alta tensão. A análise de falhas revela que os problemas no isolamento são a causa raiz de danos a tais equipamentos em mais de 60 % dos casos [1]. Por isso, é crucial garantir que o isolamento esteja em boas condições. A medição de descargas parciais (DP) tem recebido aceitação global como sendo uma forma eficaz de diagnóstico, permitindo avaliar e monitorar sistemas de isolamento tanto na fabricação quanto na operação [2].

As DP ocorrem se o campo elétrico local for superior ao valor limite suportado pelo dielétrico naquele ponto, causando uma disrupção parcial naquela região do isolamento. Elas podem ocasionar sérios danos no isolamento e reduzir consideravelmente o tempo de vida dos equipamentos de alta tensão [3][4].

Quando a falha de isolamento ocorre, o equipamento é substituído sem nenhuma informação do tipo de DP ocorrido, visto que equipamentos comerciais de detecção de DP não fornecem qualquer informação sobre a fonte causadora [5]. Todo defeito tem a sua própria característica única de degradação e é possível utilizar este atributo especial para correlacionar os padrões de DP com o tipo de defeito para determinar a qualidade do isolamento [6].

A determinação do tipo de DP é crucial para avaliar a condição do sistema de isolamento pela gravidade que o defeito encontrado implica [7][8]. Além disso, a classificação de DP pode ser utilizada para prever qualquer falha de isolamento iminente e determinar se há necessidade de substituir o sistema de isolamento atual [9].

Os equipamentos instalados em uma subestação estão sujeitos a severas condições de solicitações dielétricas causadas por sobretensões, provenientes de manobras no Sistema Elétrico de Potência (SEP) e descargas atmosféricas (surtos). Para realizar a coordenação de isolamento, a fim de impedir que tais transientes de tensão danifiquem os equipamentos, torna-se necessária a instalação

de dispositivos adequados para limitar os níveis máximos de sobretensão, sendo o para-raio (PR) o equipamento mais adequado para executar esta função. Ele atua como limitador do nível de tensão nos terminais dos equipamentos de uma subestação, evitando desta forma que sobretensões possam solicitar os equipamentos para os quais fornece proteção [10].

Os PRs exercem papel de suma importância no SEP, pois contribuem de forma decisiva para sua operacionalidade, segurança, economia e confiabilidade. Estes equipamentos não apresentam, normalmente, dispositivos indicativos de sua condição operativa, pouco se podendo concluir quanto ao estado operacional após sua energização. A avaliação dos PRs em serviço é dificultada também pelo grande número de unidades instaladas nas subestações do sistema. A falha em um PR, além de sua perda, pode causar oscilações no sistema, danos aos equipamentos periféricos e também colocar em risco a integridade física das pessoas que, porventura, estejam trabalhando nas proximidades do PR no instante da falha.

Os transformadores de potencial e de corrente (TPs e TCs) são os responsáveis por alimentar com tensões e correntes, respectivamente, os medidores, relés de proteção e outros dispositivos de controle do tipo de corrente alternada instalados nas subestações. Eles proporcionam isolamento contra a alta-tensão do circuito primário, sendo chamados de transformadores de instrumentos (TIs), e suprem os dispositivos de controle com tensões e correntes proporcionais às que circulam pelos circuitos de potência, reduzindo estas a níveis suportáveis pelo isolamento dos instrumentos de medição. Os TCs têm o seu enrolamento primário ligado em série com o circuito de alta tensão (circuito primário) e são de extrema importância para o bom funcionamento do sistema [11].

No cenário atual, com o aumento do consumo de energia elétrica, qualquer corte na geração ou transmissão de energia elétrica, provocada por uma falha de algum equipamento de uma usina ou subestação, proporcionará um desgaste enorme na imagem da concessionária de energia, além do pagamento de multas altíssimas impostas pelos órgãos fiscalizadores. Em termos de custo¹, o valor aproximado de um PR de 500 kV gira em torno de R\$ 50mil e o de um TC de 500 kV em torno de R\$ 100mil, sendo que o valor de um transformador de potência,

¹ Informações de custo obtidas junto às empresas do Grupo Eletrobras em junho de 2019.

por exemplo, que é protegido pelos PRs contra surtos de tensão, pode ultrapassar a casa dos R\$ 5 milhões. Desta forma, uma atenção maior deve ser dada ao estado operativo dos PRs e TCs.

O Cepel (Centro de Pesquisas de Energia Elétrica) propôs e desenvolveu [10] uma nova técnica de detecção de DP em PRs e TCs utilizando sensores de alta frequência sem desligamento dos equipamentos. Em conjunto com os ensaios já consolidados de termovisão e medição da corrente de fuga, no caso dos PRs, e com a análise de gases dissolvidos no óleo, no caso dos TCs, essa técnica vem fornecendo mais subsídios para a engenharia de manutenção das concessionárias de energia. Com isto, é possível gerenciar melhor a substituição desses equipamentos em caso de indicação de falha operacional, aumentando assim a confiabilidade do SEP nacional e ajudando a evitar sinistros como os mostrados nas Figuras 1.1 e 1.2 ocorridos em TCs de 345 e 500 kV, respectivamente, retirando de operação os equipamentos com suspeita de DP antes que venham a falhar, como foi feito com o PR de 500 kV retirado de operação com a ajuda do Cepel mostrado na Figura 1.3 [10]. Neste caso, o PR formado por 4 colunas foi diagnosticado com DP no campo e levado para ensaio em laboratório, onde a suspeita de DP se confirmou e o equipamento foi aberto para verificação (autópsia), mostrando corrosão severa por entrada de umidade em uma das colunas [10].

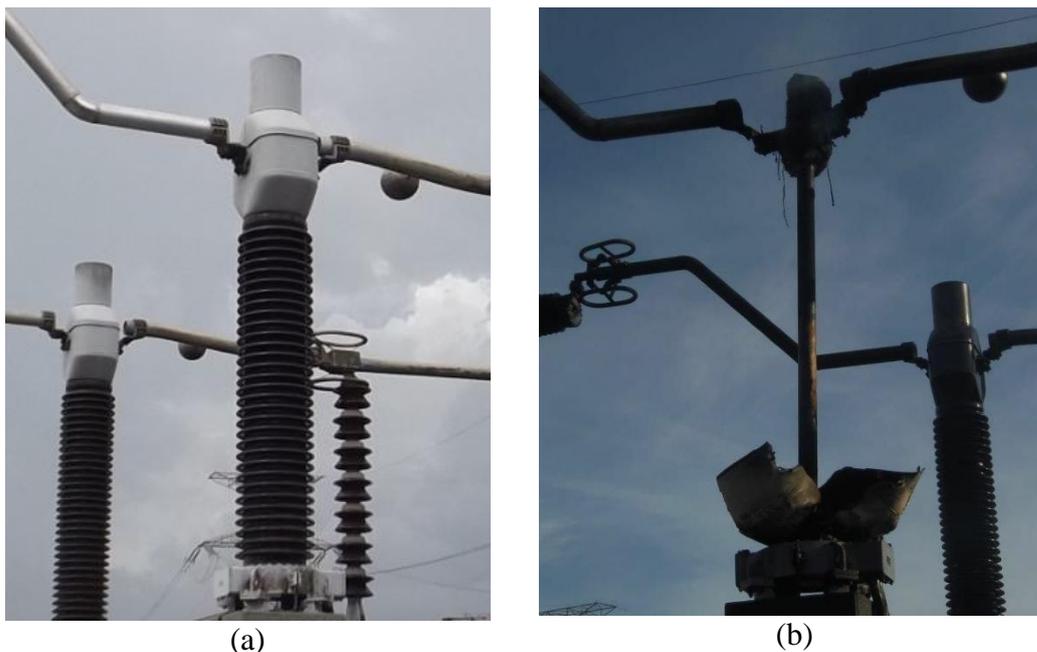


Figura 1.1 – TC de 345 kV: (a) antes e (b) após o sinistro

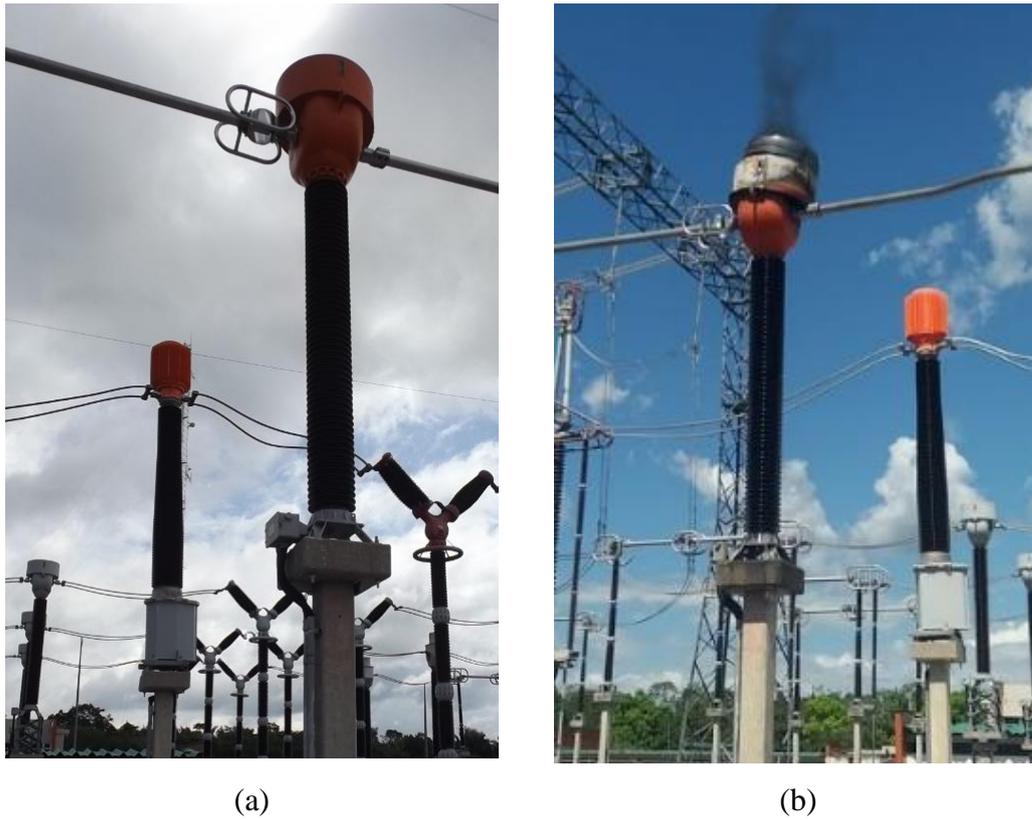


Figura 1.2 – TC de 500 kV: (a) antes e (b) após o sinistro

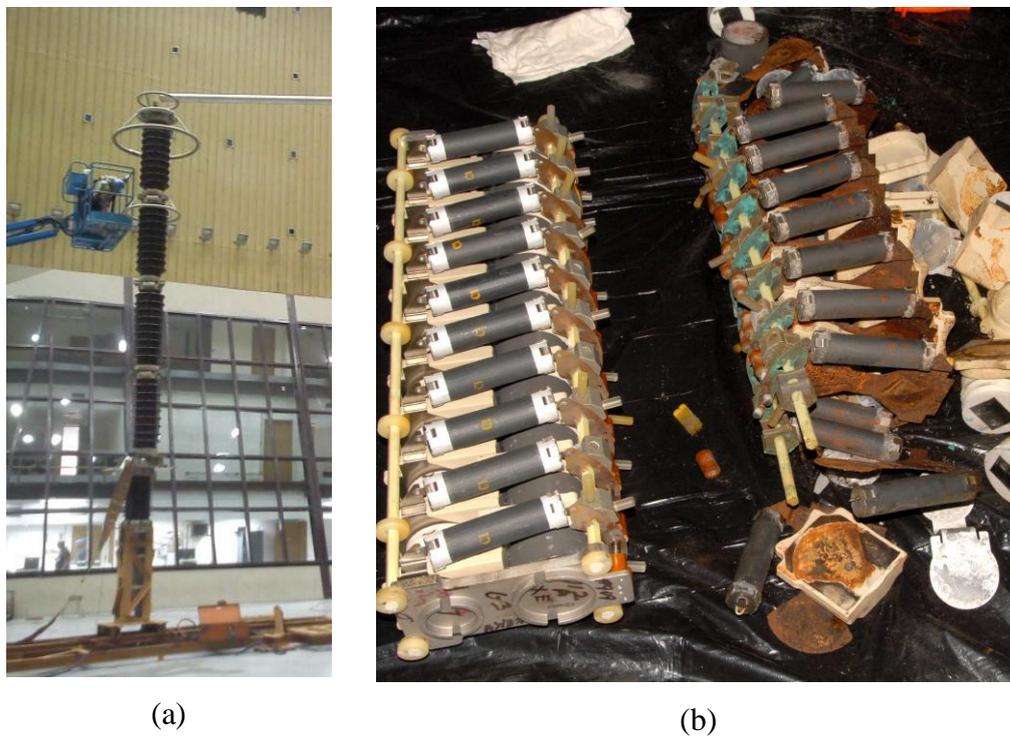


Figura 1.3 – PR de 500 kV retirado de operação: (a) arranjo para ensaio em laboratório; (b) duas colunas do PR desmontadas, sendo uma com umidade interna (direita) e outra sem (esquerda)

A indicação dos equipamentos com DP é realizada ainda pelo especialista, que analisa os sinais no tempo e os compara aos observados em laboratório com assinatura característica de DP. Ou seja, há a carência de um sistema capaz de classificar de forma automática sinais no tempo de equipamentos com e sem DP, de forma a considerar todas as nuances dos sinais por meio da extração de suas características relevantes.

Este estudo apresenta uma metodologia eficaz para a classificação automática dos sinais medidos nos PRs e TCs, avaliando as características discriminantes dos sinais que proporcionam uma melhor classificação em relação à presença de DP nos equipamentos avaliados. O classificador, baseado em árvore de decisão e SVM (“máquina de vetor suporte” ou “*support vector machine*”), pode ser utilizado nos PRs e nos TCs.

A confiabilidade do equipamento de alta tensão é afetada significativamente pela presença de defeitos em sua isolamento. Assim, a identificação e caracterização de defeitos de isolamento é um requisito fundamental para a programação da manutenção. Como defeitos de isolamento geralmente dão origem a DP, as medições de DP podem constituir uma ferramenta poderosa para a identificação e a avaliação de tais defeitos [12].

Uma investigação de DP visa a principalmente correlacionar a atividade de DP ao dano à fonte de DP relevante, a fim de permitir que os procedimentos de diagnóstico e manutenção sejam ativados. Isso requer um conhecimento adequado das características do equipamento sob teste. Na verdade, diferentes tipos de equipamentos e sistemas de isolamento podem diferir muito na capacidade de suportar atividades de DP. Mesmo assim, as informações derivadas das medições de DP podem ser inconsistentes ou enganosas se uma metodologia apropriada não for seguida.

As medições de DP são afetadas por dois problemas básicos. O primeiro é a presença de ruído e distúrbios, que podem alterar ou ocultar atividades de DP ou, em alguns casos, ser erroneamente interpretados como DP. O segundo problema é que a natureza do defeito deve ser adequadamente caracterizada, juntamente com sua tendência (com o valor da tensão de alimentação ou com o tempo de aplicação da tensão), a fim de avaliar seu dano e realizar o diagnóstico.

Os sistemas comerciais de medição de DP disponíveis hoje são carentes de um procedimento robusto que seja capaz de separar e classificar os sinais de DP.

Algumas técnicas são propostas na literatura, mas ainda há necessidade de aprimoramento e espaço para novas variáveis que auxiliem na discriminação das diferentes fontes de DP pela caracterização do sinal no tempo, e também na classificação do defeito.

Este estudo propõe uma nova metodologia para a avaliação de DP em diferentes tipos de isolamentos de alta tensão, com soluções que auxiliam na caracterização do sinal no tempo, na separação das fontes de DP e na classificação desta quanto ao tipo de defeito. Estas soluções podem ser utilizadas isoladamente ou em conjunto com técnicas já existentes.

1.2

Objetivos

O objetivo principal deste estudo é desenvolver uma sistemática precisa e eficaz para detectar e analisar sinais de DP medidos em isolamentos de alta tensão, separando as diferentes fontes de DP internas ao isolamento e classificando cada uma delas, com base em um banco de dados histórico, em relação ao tipo de defeito que representam.

Dentre os objetivos específicos deste estudo, destacam-se os seguintes:

- Analisar e determinar características dos sinais no tempo que possam ser utilizadas para discriminar sinais de PRs e TCs com e sem DP.
- Analisar e determinar características discriminantes dos sinais de DP no tempo que possam ser utilizadas para separar as diferentes fontes de DP internas ao isolamento.
- Propor uma técnica de agrupamento que permita uma boa separação das diferentes fontes de DP internas ao isolamento.
- Testar características discriminantes dos mapas de DP que possam ser utilizadas para classificar os diferentes tipos de defeitos, após a separação das fontes.
- Criar um banco de dados histórico com padrões de mapas de DP conhecidos na literatura em máquinas rotativas.

1.3

Organização do Texto

O texto apresentado aqui está dividido em 5 capítulos, conforme segue:

Capítulo 1: discorre sobre a motivação e os objetivos a serem alcançados no estudo;

Capítulo 2: apresenta um resumo dos fundamentos necessários para uma melhor compreensão do trabalho, abordando conceitos relacionados a descargas parciais;

Capítulo 3: descreve detalhes dos dados e técnicas utilizados para o desenvolvimento deste estudo;

Capítulo 4: apresenta os resultados obtidos e algumas considerações;

Capítulo 5: apresenta as conclusões obtidas no estudo e indicações para trabalhos futuros.

Anexos: apresenta detalhes de algumas técnicas utilizadas no estudo.

2

Fundamentos

2.1

Equipamentos de Alta Tensão

Os equipamentos de alta tensão são ítems indispensáveis na cadeia de funcionamento do sistema elétrico, composta pela geração, transmissão e distribuição de energia elétrica. Têm em comum a classe de tensão de operação superior a 1000 V em corrente alternada ou 1500 V em corrente contínua. Como exemplo desses equipamentos, têm-se os geradores, motores, para-raios, disjuntores, chaves seccionadoras, transformadores de potência, transformadores de potencial, transformadores de corrente, entre outros.

Os principais materiais que constituem os equipamentos de alta tensão podem ser divididos basicamente em condutores e isolantes elétricos. Os materiais condutores são responsáveis pelo transporte de energia entre os vários pontos de um sistema e são caracterizados por impedâncias série pequenas, o que permite o fluxo de altas correntes elétricas com pequenas quedas de tensão. Entre os materiais condutores, tem-se prata, cobre, ouro, alumínio, ferro e aço. Já os materiais isolantes têm a função de fornecer proteção e isolamento entre as partes energizadas e os pontos de aterramento locais, sem que haja perda nas funções operacionais dos equipamentos. São caracterizados por impedâncias muito elevadas, o que permite elevadas diferenças de potencial com fluxo reduzido de corrente elétrica. Exemplos de materiais isolantes são: papel, mica, vidro, teflon, ar, porcelana, resinas, epóxi, borrachas, óleos e SF₆ (gás Hexafluoreto de Enxofre) [13].

A resistência elétrica de um material é a capacidade de um corpo qualquer se opor à passagem de corrente elétrica mesmo quando existe uma diferença de potencial aplicada. Ela pode ser determinada pela seguinte expressão:

$$R = \rho \times L / A \qquad \text{Eq.2.1}$$

Onde: $\rho \rightarrow$ resistividade elétrica ($\Omega \cdot m$);

$L \rightarrow$ comprimento do material (m);

$A \rightarrow$ área da seção transversal do material (m^2).

Materiais isolantes e condutores apresentam diferentes ordens de grandeza para valores de resistividade elétrica característicos ($\rho_{\text{isolante}} \approx 10^{20} \times \rho_{\text{condutor}}$), conforme pode ser observado na tabela 2.1.

Tabela 2.1 – Resistividade de materiais isolantes e condutores

Tipo	Material	Resistividade elétrica ($\Omega \cdot m$)
Condutor	Prata	$16,4 \times 10^{-9}$
	Cobre	$17,2 \times 10^{-9}$
	Ouro	$24,5 \times 10^{-9}$
	Alumínio	$28,0 \times 10^{-9}$
Isolante	Papel	$0,1 \times 10^{11}$
	Mica	$5,0 \times 10^{11}$
	Vidro	$10,0 \times 10^{11}$
	Teflon	$30,0 \times 10^{11}$

Uma característica básica dos materiais isolantes é a rigidez dielétrica, definida pela suportabilidade máxima em kV/mm até a qual não há corrente circulante no material, ou seja, é o maior valor de campo elétrico aplicado a um isolante sem que ele se torne condutor. A rigidez dielétrica de um material isolante pode ser alterada com a temperatura, umidade do ar, espessura do material, forma da tensão aplicada, taxa de subida da tensão, forma dos eletrodos, entre outras. Na tabela 2.2 são apresentados alguns valores de rigidez dielétrica de determinados materiais, sob pressão de 1 atm, temperatura de 25°C, e entre eletrodos esféricos.

Um material isolante pode ser representado eletricamente por uma resistência e uma capacitância em paralelo, conforme ilustrado na Figura 2.1. Uma análise do diagrama fasorial representativo da tensão e das correntes em um material isolante, apresentado na Figura 2.2, permite compreender melhor o

comportamento destes materiais quando submetidos a uma diferença de potencial entre os seus terminais.

Tabela 2.2 – Rigidez dielétrica de alguns materiais

Material	Rigidez dielétrica (kV/mm)
Teflon	60
Baquelite	24
Papel	16
Óleo de silicone	15
Nylon	14
Quartzo	8
Ar	3

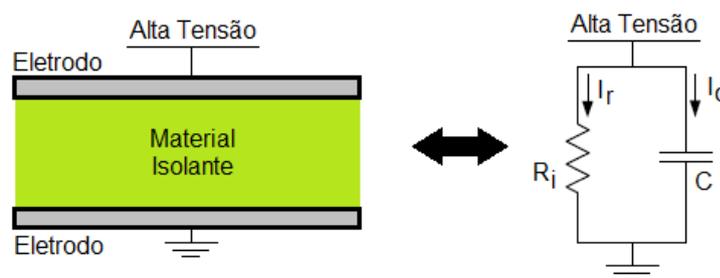


Figura 2.1 – Representação elétrica de um material isolante

Onde: $C \rightarrow$ capacitância global entre eletrodos;

$R_i \rightarrow$ resistência de isolamento;

$I_c \rightarrow$ corrente capacitiva de carga;

$I_r \rightarrow$ corrente resistiva de perdas.

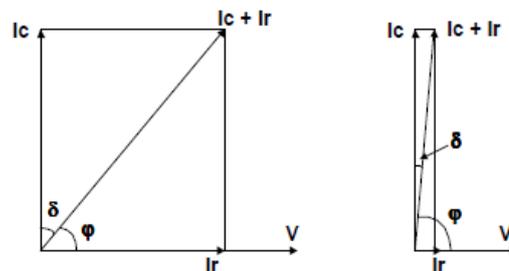


Figura 2.2 – Diagrama fasorial representativo

O “fator de perdas” (ou “tangente de delta”), dado pela razão I_r/I_c , constitui uma característica de um material dielétrico ou de um sistema de isolamento elétrico, sendo uma indicação do grau de pureza do meio dielétrico. Representa a qualidade, ou a dissipação relativa de calor no meio dielétrico quando submetido a um campo elétrico alternado. Pode ser determinado também pela razão entre a potência útil e a potência reativa do sistema capacitivo. Quanto menor for a amplitude de I_r com relação a I_c , menor será o ângulo delta (δ) entre I_c e a corrente total I_r+I_c , menor o valor de $\tan \delta$, e melhor a condição do dielétrico.

Entre os materiais condutores e isolantes, existe uma categoria intermediária com materiais denominados semicondutores. Estes materiais podem ser obtidos industrialmente via confecção de fitas, resinas, tintas, papéis, etc. São responsáveis pelas funções de isolamento progressivo entre eletrodos separados por campos elétricos intensos, permitindo sua distribuição de forma progressiva. São utilizados, por exemplo, em terminais de barras de geradores, buchas de alta tensão (Figura 2.3) e cabos de alta tensão [13]. A resistividade elétrica de alguns desses materiais é mostrada na tabela 2.3.

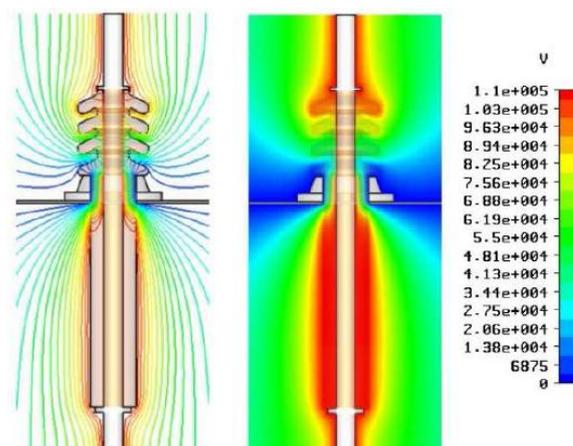


Figura 2.3 – Distribuição de potencial no entorno de uma bucha de alta tensão com material semicondutor próximo à região de potencial nulo

Tabela 2.3 – Resistividade de materiais semicondutores

Tipo	Material	Resistividade elétrica ($\Omega.m$)
Semicondutor	Carbono	40×10^{-6}
	Germânio	0,47
	Silício	640

2.2

Descargas Parciais

2.2.1

Definições

Inicialmente, para um melhor entendimento, apresentam-se alguns termos relacionados ao fenômeno das DP [13]:

Ionização – qualquer processo pelo qual um átomo torna-se eletricamente carregado.

Corona – micro descargas elétricas que ocorrem normalmente no ar em regiões de campo elétrico intenso. O termo “efeito corona” é mais utilizado no caso de descargas ocorrendo no ar em torno de um elemento condutor, assim como “streamers”, “centelhamentos”, “micro descargas” ou “descarga autossustentada”.

Ionização interna – termo utilizado pelas normas no caso de corona interno aos PRs, sendo quantificado pela medição de RIV (tensão de rádio interferência), normalmente em microvolts (μV).

Elétrons livres – elementos existentes nas áreas externas dos materiais condutores, com liberdade e movimentação contínua. Os primeiros elétrons do processo são criados por fotoionização. Quando acelerados por um campo elétrico, na fronteira com um ambiente com gás, colidem com átomos neutros produzindo a movimentação de outros elétrons e íons positivos. As colisões, sendo inelásticas, conservam a maior parte da energia dos elétrons e íons, o que facilita a manutenção do processo. Do ponto de vista elétrico, este processo corresponde à geração de correntes impulsivas de polaridades positiva e negativa, com uma grande variedade de amplitudes, formas temporais e taxas de repetição.

As DP são descargas elétricas localizadas, que não chegam a percorrer todo o caminho dentro de um material isolante colocado entre dois condutores submetidos a uma diferença de potencial. Elas são pulsos de corrente de alta frequência que ocorrem, de maneira repetitiva, no interior dos sistemas isolantes dos equipamentos de alta tensão. Pode-se dizer também que são descargas disruptivas ocorrendo em pequena parte de um material isolante. Em muitos

casos, o corona na superfície de materiais condutores e dielétricos também é medido como sendo DP [13].

Segundo a norma internacional IEC60270 [13], as DP são descargas elétricas localizadas que curto-circuitam parcialmente um material isolante que separa dois condutores e que pode estar ou não adjacente a um dos condutores. A Figura 2.4 ilustra o processo de formação das DP e seus efeitos.

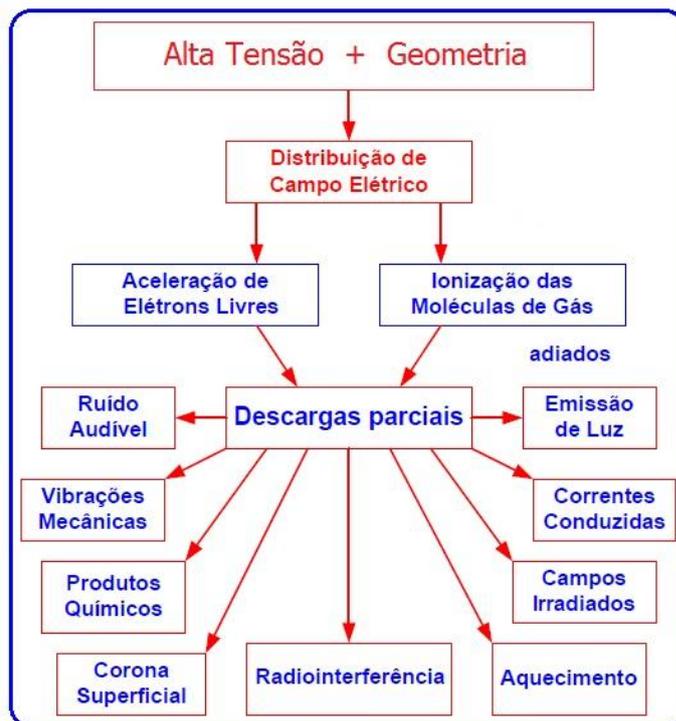


Figura 2.4 – Processo de formação das DP e seus efeitos

Com relação aos locais de ocorrência, as DP podem se dar em qualquer ponto do dielétrico, seja na junção de dois dielétricos diferentes ou adjacentes ao condutor, do lado do eletrodo aterrado, ou até mesmo ocorrer seguidamente em múltiplos pontos do sistema isolante. Em dielétricos sólidos, as DP são produzidas pela ionização de pequenas cavidades de ar no interior do dielétrico. Em dielétricos líquidos, ocorrem pela ionização de bolhas de gás no seu interior. No ar, se dão pela ionização das moléculas de ar nos pontos de maior gradiente de potencial elétrico [13].

A necessidade dos ensaios de DP vem do fato de que elas são uma fonte contínua de deterioração do material isolante, ou seja, modificam suas propriedades dielétricas, reduzindo a sua suportabilidade. Dependendo da

intensidade das DP, a vida útil do material será reduzida. A medição de DP pode indicar uma certa expectativa de vida útil dos dielétricos. É uma forma de controle da qualidade do sistema isolante, dos processos construtivos e materiais utilizados na fabricação dos equipamentos, e também uma forma de verificar a condição dielétrica após a manutenção de equipamentos.

As medições de DP servem também para identificar possíveis montagens irregulares, como complementação de outras avaliações, bem como para identificar o início de descargas internas. Elas possibilitam a verificação do estado dielétrico antes e após outros ensaios de solicitação e servem como previsão para a manutenção preventiva dos equipamentos [13].

As disrupções das DP ocorrem sempre em meio gasoso, sendo caracterizadas por pulsos discretos de curta duração e taxa de ocorrência elevada. Elas promovem a elevação de temperatura no dielétrico e não apresentam características visíveis a olho nu, ou seja, quando um equipamento é aberto, pode não ser possível confirmar a presença de DP, a menos que o defeito tenha proporções muito críticas.

2.2.2

Grandezas Associadas às Descargas Parciais

Algumas grandezas relacionadas às DP, segundo as normas ABNT NBR 6940 [15] e IEC 60270 [13], têm suas definições apresentadas a seguir:

Carga aparente (q) – medida normalmente em picocoulombs (pC), é a carga que, se injetada em um tempo muito curto nos terminais externos de um equipamento, causa a mesma queda de tensão sobre uma impedância de medição de um circuito de ensaio do que a própria carga real gerada no interior desse equipamento.

Taxa de repetição das descargas (n) – número médio de pulsos de descargas ocorrendo por uma unidade de tempo dentro de uma determinada faixa de amplitudes das descargas.

Corrente média de descarga (I) – é a soma dos valores absolutos das cargas aparentes (q_i , $i = 1, \dots, m$) durante um curto intervalo de tempo dividido por este intervalo (T_{ref}).

$$I = (|q_1| + |q_2| + |q_3| + \dots + |q_m|) / T_{ref} \quad \text{Eq.2.2}$$

Taxa quadrática (D) – é a soma dos quadrados das cargas aparentes, durante um certo intervalo de tempo, dividida por este intervalo.

$$D = (q_1^2 + q_2^2 + q_3^2 + \dots + q_m^2) / T_{ref} \quad \text{Eq.2.3}$$

Potência de descarga (P) – potência média percebida nos terminais do objeto ensaiado devido exclusivamente às DP. Equivale à soma do produto $q_i \times u_i$ (sendo u_i os valores instantâneos de tensão nos terminais do objeto sob ensaio nos instantes das descargas q_i , onde i varia de 1 a m) dividido pelo período de tempo T_{ref} em que estas descargas ocorrem.

$$P = (q_1 \times u_1 + q_2 \times u_2 + q_3 \times u_3 + \dots + q_m \times u_m) / T_{ref} \quad \text{Eq.2.4}$$

Intensidade de descargas parciais – valor medido das DP, em um objeto sob ensaio, em uma ou mais das seguintes grandezas: carga aparente, taxa de repetição das descargas, corrente média de descarga, taxa quadrática, potência de descarga. Para cada tipo de equipamento existe um nível aceitável de intensidade de DP para o qual se considera que não haverá dano, conforme apresentado na tabela 2.4 em valores de carga aparente.

Tabela 2.4 – Níveis de DP aceitáveis

Equipamento	Carga aparente (q)
De alta tensão em geral	< 10 pC
Cabos blindados de alta tensão	< ruído de fundo (1 a 2 pC)
Transformadores de potência de grande porte	< 300 pC
Motores e geradores	2000 a 5000 pC

É interessante atentar para as seguintes relações:

$$\text{Carga de 1 elétron} = 1,6 \times 10^{-19} \text{ coulombs (C)}$$

$$1 \text{ C} = 6,2 \times 10^{18} \text{ elétrons}$$

$$1 \text{ pC} = 6,2 \times 10^6 \text{ elétrons}$$

$$1 \text{ A} = 1 \text{ C/s} = 6,2 \times 10^{18} \text{ elétrons / s}$$

Tensão de início de descargas parciais (U_I) – menor tensão na qual são observadas DP quando a tensão aplicada no objeto sob ensaio é gradualmente aumentada a partir de um valor baixo no qual não são observadas DP.

Tensão de extinção de descargas parciais (U_e) – maior tensão na qual cessam as DP quando a tensão é gradualmente reduzida a partir de um valor superior à tensão de início de descarga.

Tensão de ensaio de descargas parciais – tensão aplicada em um procedimento de ensaio especificado, durante o qual o objeto sob ensaio não pode apresentar DP superiores a uma intensidade especificada.

Fator de escala do circuito de medição (ke) – fator pelo qual a indicação de um instrumento é multiplicada para se obter o valor da grandeza medida. Este fator de escala não é necessariamente igual ao do instrumento.

Frequência de corte inferior e superior (f_1 e f_2) – frequências em que a resposta de um circuito decai 3 dB em relação a um valor constante, que limita a faixa de frequência dentro da qual a resposta do circuito de medição é aproximadamente constante.

Largura de banda (Δf) – diferença entre as frequências de corte superior e inferior.

$$\Delta f = f_2 - f_1 \quad \text{Eq.2.5}$$

Tempo de resolução de pulso (T_r) – menor intervalo entre dois pulsos consecutivos que resulta em um erro de amplitude não superior a 10 % à sobreposição dos pulsos. Normalmente o tempo de resolução de pulso é inversamente proporcional à largura da banda do instrumento.

Ângulo de fase das descargas (Φ_i) e tempo de ocorrência (t_i) – posição do pulso de descarga (em “graus” e em “s”, respectivamente) em relação ao período T (em “s”) da onda de tensão aplicada à frequência industrial (60 Hz, no caso do Brasil).

$$\Phi_i = 360 \times (t_i / T) \quad \text{Eq.2.6}$$

2.2.3

Representação Elétrica das Descargas Parciais

Conforme as características do material dielétrico contido no equipamento, existirá em seu interior uma certa quantidade de cavidades de várias formas e dimensões preenchidas com ar e gases, que são os pontos fragilizados do dielétrico onde podem ocorrer as DP. Costuma-se, para efeito de simplificação e de análise, considerar uma única cavidade de contorno plano, pois as várias partes do dielétrico podem ser simuladas idealmente por capacitores de placas paralelas [13]. Assim, pode-se formar um circuito equivalente simplificado, conforme mostrado na Figura 2.5.

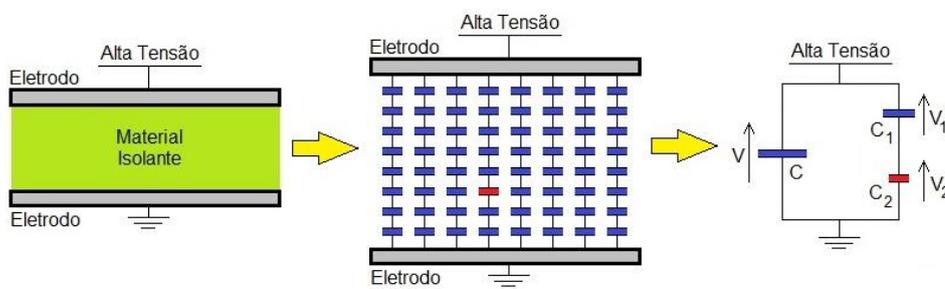


Figura 2.5 – Circuito equivalente simplificado representando um dielétrico com cavidades

Onde: $C_2 \rightarrow$ Representa a capacitância da cavidade;

$C_1 \rightarrow$ Representa a capacitância total em série com C_2 ;

$C \rightarrow$ Representa o restante da capacitância em paralelo com C_1 e C_2 .

Para cada capacitância, existe também uma resistência correspondente (conforme mostrado nas Figuras 2.1 e 2.2), que é aqui desprezada. A tensão nos terminais da cavidade (V_2), em função dos parâmetros do circuito e da tensão aplicada externamente, é dada por:

$$V_2 = V \times C_1 / (C_1 + C_2) \quad \text{Eq.2.7}$$

Uma descarga parcial significa um curto-circuito através da capacitância C_2 , o que acarretará uma diminuição instantânea da tensão nos terminais do dielétrico

do equipamento, através do surgimento de um sinal transitório de alta frequência (que caracteriza as DP) com amplitude equivalente a V_2 .

2.2.4

Medição de Descargas Parciais

As DP apresentam diversas características que permitem sua detecção e medição, podendo produzir ondas eletromagnéticas, ondas acústicas, pulsos de corrente, consumo de energia, luminescências, variações térmicas, variações químicas e vibrações mecânicas. Essas técnicas de detecção e medição podem ser divididas em dois grupos: elétricos e não-elétricos (conforme ilustrado na Figura 2.4). O método elétrico insere o circuito de medição e detecção no circuito onde as descargas estão ocorrendo, conforme mostrado nas Figuras 2.6, 2.7 e 2.8. Já os métodos não-elétricos, geralmente servem apenas de apoio aos métodos elétricos. Existem diversos métodos de detecção e medição não-elétricos, dos quais pode-se destacar:

Método óptico – detecta a ocorrência de DP pelas emissões de luz provocadas no momento de ocorrência destas.

Método acústico – capta as ondas ultrassônicas, não audíveis, através de sensores apropriados, auxiliando também na localização das DP em equipamentos de grande porte.

Método de rádio frequência – um receptor capta interferências provocadas por ondas eletromagnéticas, permitindo a identificação do local de ocorrência das DP.

Método químico – analisa a presença de produtos dissolvidos no dielétrico provenientes da ocorrência de DP.

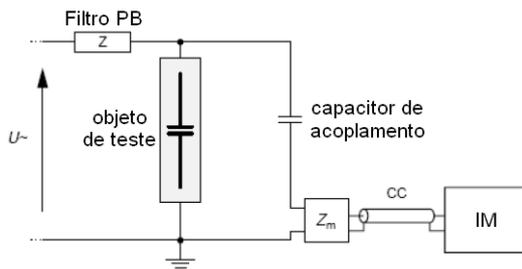


Figura 2.6 – Circuito de medição de DP pelo método elétrico direto

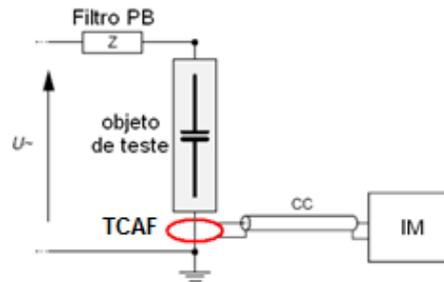


Figura 2.7 – Circuito de medição de DP utilizando TC de alta frequência



Figura 2.8 – TC de alta frequência conectado ao cabo de aterramento do equipamento

onde: Filtro PB → Filtro passa-baixa;

U → Tensão de teste;

CC → Cabo coaxial;

Z_m → Impedância de medição;

IM → Instrumento de medição;

TCAF → Transformador de corrente de alta frequência.

Nas figuras 2.6 e 2.7, o "filtro PB" é responsável por não deixar passar ao circuito de medição sinais de alta frequência oriundos da fonte de alta tensão que fornece a tensão de teste. O "capacitor de acoplamento" funciona como sensor de DP e também, juntamente com Z_m , como divisor de tensão para reduzir a tensão que chega ao IM. Z_m também funciona como filtro passa-alta, deixando passar sinais de alta frequência ao IM (que representam as DP oriundas do objeto de teste) e reduzindo a próximo de zero a amplitude dos sinais de baixa frequência (como o 60 Hz, por exemplo). Através do cabo coaxial, o IM capta os sinais de

DP para registro. Na Figura 2.7, o TCAF funciona como sensor de DP e também como filtro passa-alta.

2.2.5

Sinais de Descargas Parciais

Os dados de DP são classificados basicamente em três categorias distintas, de acordo com os parâmetros medidos: DP associadas ao tempo de duração da descarga, DP associadas à fase da tensão aplicada, e DP não associadas nem à fase e nem ao tempo [16].

2.2.5.1

Características Associadas ao Tempo

Os padrões de dados associados ao tempo (q, t) mostram o formato de um pulso individual de DP [17]. Neste caso, a tensão de teste é tratada como se fosse constante. Estes padrões de dados podem apresentar uma relação entre a física do defeito e a forma do sinal [16]. Nas figuras 2.9 a 2.14, são mostrados alguns pulsos de DP medidos em função do tempo.

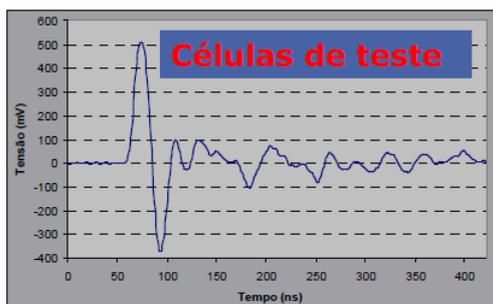


Figura 2.9 – Sinal no tempo de DP em uma célula de teste

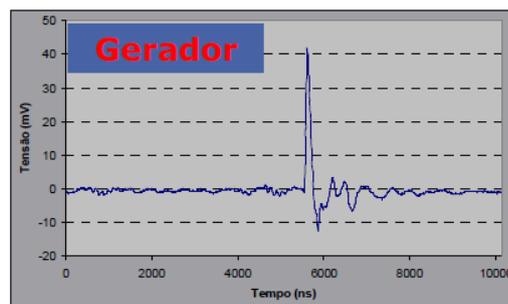


Figura 2.10 – Sinal no tempo de DP em um gerador

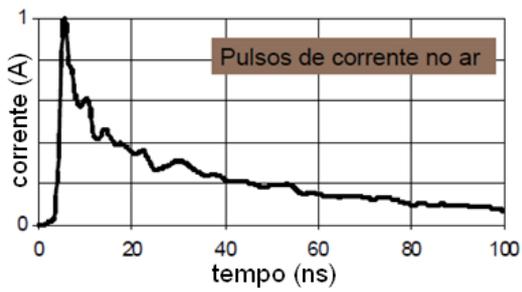


Figura 2.11 – Sinal no tempo de DP no ar 01

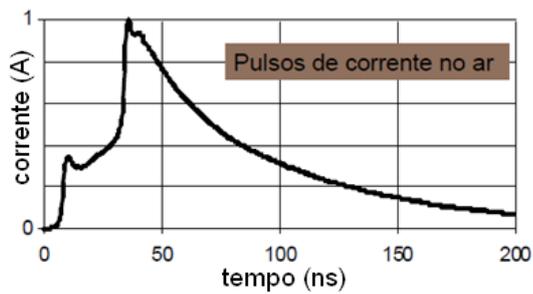


Figura 2.12 – Sinal no tempo de DP no ar 02

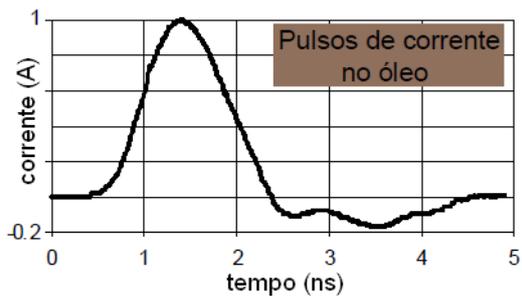


Figura 2.13 – Sinal no tempo de DP no óleo 01

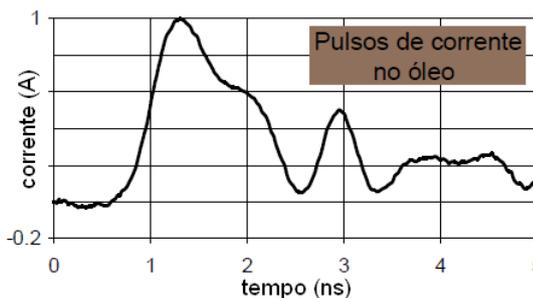


Figura 2.14 – Sinal no tempo de DP no óleo 02

Pode-se também visualizar os sinais de DP no tempo juntamente à onda de tensão de frequência industrial à qual eles estão associados, conforme mostrado nas figuras 2.15 a 2.17.

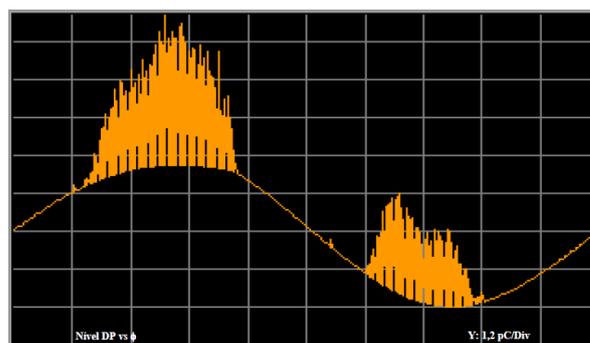


Figura 2.15 – Sinais de DP sobre a frequência industrial 01

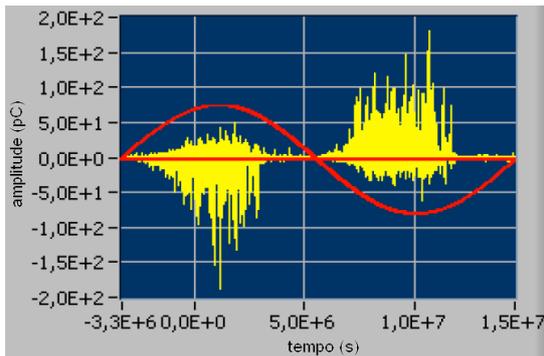


Figura 2.16 – Sinais de DP sobre a frequência industrial 02

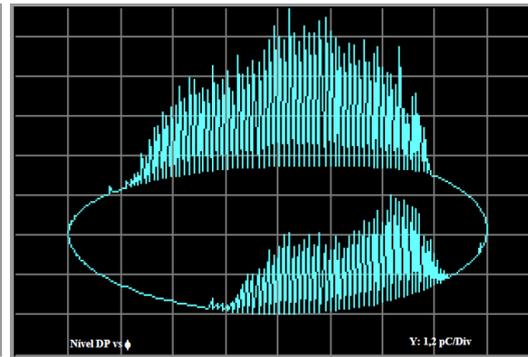


Figura 2.17 – Sinais de DP sobre a frequência industrial 03

2.2.5.2

Características Associados à Fase

Neste caso, considera-se que a tensão de teste é constante e os ângulos de fase da onda de tensão são discretizados em faixas. O detector mede a amplitude das descargas (q), relacionadas às fases da tensão de teste (φ), durante um intervalo de tempo, sendo que neste intervalo os valores das amplitudes são acumulados e podem ser expressos pela taxa de ocorrência (n). Estes dados de DP são classificados como associados à fase. Os softwares de análise exibem tais dados como distribuições univariadas em função da fase φ ou distribuições bivariadas do tipo (φ, q, n) , também denominadas “mapas estatísticos de DP” (ou simplesmente “mapas de DP”) [16]. As distribuições univariadas mais comuns são [18]: (q_a, φ) - carga média em cada janela de fase vs φ ; (q_p, φ) - pico de carga em cada janela de fase vs φ ; (n, φ) - quantidade de pulsos de DP detectadas em cada janela de fase vs φ ; e (I, φ) - corrente média da descarga em cada janela de fase vs φ .

O espaço tridimensional, formado pelos parâmetros (φ, q, n) , é a forma mais tradicional de exibição dos dados de DP com distribuição bivariada, sendo também a forma mais completa de representação gráfica do pulso de DP (figuras 2.18 a 2.22, onde a terceira dimensão, “ n ”, está associada à cor). Ou seja, nesses “mapas de DP”, cada ponto cinza equivale a um pulso de DP (de alta frequência) associado à fase da senóide onde ele ocorre (eixo horizontal) e à sua carga equivalente (eixo vertical), sendo que pulsos com mesmas fase e carga equivalente passam a ser representados no mapa em cores diferentes, respeitando

um mapa de cores indexado. Os histogramas das amplitudes dos pulsos (n vs. q) podem ser obtidos diretamente ou indiretamente através de (φ, q, n) [19].

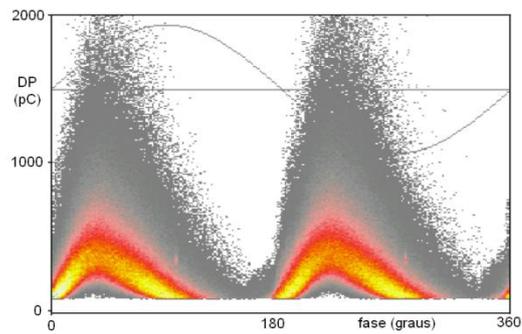


Figura 2.18 – Mapa estatístico de DP 01

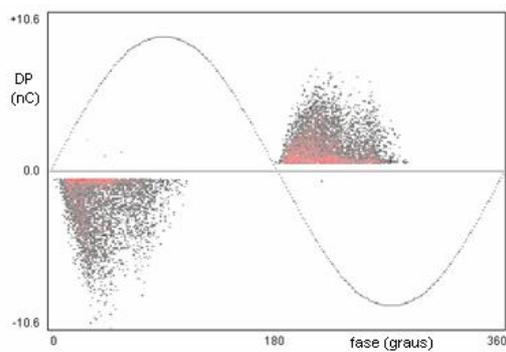


Figura 2.19 – Mapa estatístico de DP 02

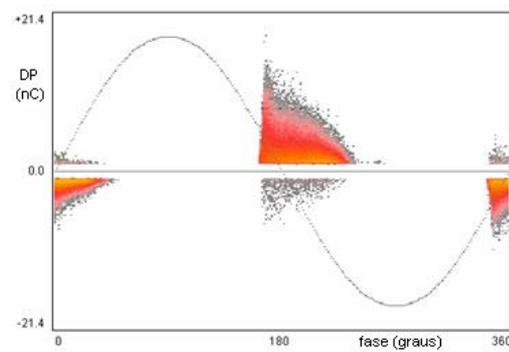


Figura 2.20 – Mapa estatístico de DP 03

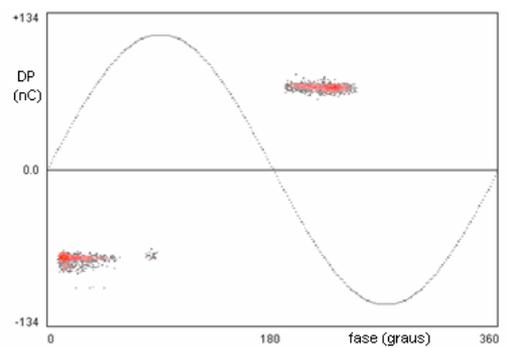


Figura 2.21 – Mapa estatístico de DP 04

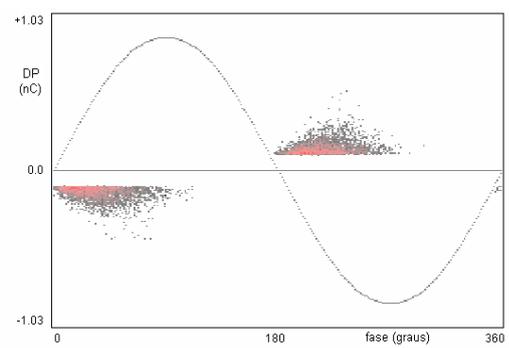


Figura 2.22 – Mapa estatístico de DP 05

2.2.5.3

Características Desassociadas do Tempo e da Fase

São parâmetros atemporais, associados a um ou mais parâmetros do sinal de DP que não seja a fase da tensão. Um exemplo de sinal de DP sem informação de tempo nem fase pode ser obtido pela medição da amplitude da DP quando se varia a tensão de teste, de maneira a se obter o efeito de histerese da DP no ciclo de teste [20]. O efeito observado de histerese é diretamente relacionado ao fato de que, uma vez que o processo de descarga parcial tenha se iniciado, a aplicação de pouca tensão é suficiente para sustentá-lo [16].

2.3

Reconhecimento de Padrões em Descargas Parciais

Diversos estudos vêm sendo desenvolvidos ultimamente com o objetivo de identificar padrões em dados de DP, seja analisando dados com informação temporal, dados com informação de fase, ou até mesmo dados sem informação de tempo ou de fase. Além desta divisão em três categorias, o problema de reconhecimento de padrões em DP pode ser dividido em duas etapas: extração de características e classificação dos padrões.

A etapa de extração de características é aquela na qual os parâmetros de um conjunto padrão são coletados, formando-se então um modelo de DP que irá compor uma base de dados. Em seguida, são coletados os mesmos parâmetros de um vetor de dados qualquer de teste e comparados com a base padrão, o que compõe a etapa de classificação [16].

No passado, grande parte das pesquisas em DP supunha a existência de apenas uma fonte de DP durante a medição. Na realidade, existem em alguns casos mais de uma fonte de DP ativa causada pela presença de vários tipos de defeitos ou ruído elétrico [21]. Tem-se observado que o objetivo de algumas pesquisas em classificação de DP está mudando seu foco de análise de fonte única para fontes múltiplas de DP [6][22]. As técnicas atuais de classificação de fonte única de DP ainda têm espaço para aprimoramentos. Isto contribuiria também para a busca de uma melhor precisão na classificação em pesquisas sobre fontes múltiplas de DP.

É importante salientar que algumas características extraídas dos sinais de DP podem ser eficazes na separação de fontes múltiplas e não o ser na classificação dos sinais com relação ao tipo de defeito que eles representam. O inverso também pode ocorrer, ou seja, características representativas na classificação do tipo de defeito e não-representativas na separação de fontes múltiplas.

A identificação de um defeito (reconhecimento de padrão) é um processo de inferência que permite endereçar corretamente um dado conjunto de sinais de DP para uma tipologia de defeito específica. Desta forma, considera-se a informação sobre o tipo de defeito como o alvo do diagnóstico, enquanto a análise dos sinais de DP é a maneira de obtê-lo. Assim, para que se tenha uma identificação robusta, seu resultado deve permanecer inalterado mesmo que os sinais de DP mudem devido a [23]:

- a) localização diferente da mesma tipologia de defeito;
- b) uso de diferentes instrumentos (ganho, largura de banda, etc);
- c) uso de diferentes tipologias de acopladores;
- d) presença de múltiplas fontes de sinal;
- e) mudança da tensão aplicada;
- f) envelhecimento do isolamento.

Uma vez que milhares de pulsos são gravados durante uma sessão de medição de DP, a informação de diagnóstico é tipicamente sintetizada em padrões de DP, histogramas, parâmetros médios ou de valor extremo obtidos considerando diferentes características dos sinais gravados. Entre eles, apenas aqueles que permitem uma correspondência confiável entre os sinais de DP e a tipologia dos defeitos podem ser utilizados no processo de inferência. A seleção de padrões de DP e parâmetros adequados para uma identificação robusta pode ser realizada com base nas seis propriedades de invariância descritas anteriormente. Desta forma se estabelece uma conexão direta entre um conjunto de sinais de DP e uma tipologia de defeito específica. Assim, a eficiência da fase de identificação deve ser avaliada em termos de [23]:

I) Ambiguidade: um único padrão pode abordar múltiplas tipologias de defeitos.

II) Variabilidade: um padrão relevante para uma tipologia de defeito específico pode alterar suas características mediante variação das condições de teste, da tensão aplicada ou durante o envelhecimento.

III) Separabilidade: os padrões de DP mantêm suas características específicas após a aplicação de um procedimento de separação.

2.3.1

Extração de Características dos Sinais

O propósito principal da extração de características é a obtenção de aspectos de entrada relevantes dos dados de DP para representar características associadas a um defeito em particular. Quaisquer padrões que possam ser reconhecidos e classificados têm um número de atributos ou características discriminatórias. Desta forma, o passo inicial no processo de reconhecimento é buscar a identificação de quais características discriminatórias devem ser selecionadas e como extrair características relevantes. Além disso, a extração de características também tenta reduzir a dimensão dos dados originais para facilitar o processamento, visto que a quantidade de dados brutos contidos em cada um dos padrões armazenados pode ser bastante grande [24][25].

Diferentes técnicas vêm sendo desenvolvidas para a escolha das características que fornecerão a informação relevante para o problema de reconhecimento de padrões em DP. Essas técnicas podem ser agrupadas, basicamente, nas seguintes categorias: métodos estatísticos, caracterização do sinal no tempo, processamento de sinais e processamento de imagens [16].

2.3.1.1

Métodos Estatísticos

Os parâmetros estatísticos foram utilizados pela primeira vez [26] em 1993 e consistem principalmente em média, variância, assimetria, curtose e fator de correlação cruzada. Estes momentos estatísticos são aplicados em distribuições univariadas de DP, da forma $(q_{\text{média}}, \varphi)$, $(q_{\text{max}}, \varphi)$ e (n, φ) , e constituem um vetor

característico [18]. As distribuições podem ser obtidas segundo uma função geral $y_i = f(x_i)$. Assim, os momentos podem ser calculados da seguinte maneira:

$$\text{Média:} \quad \mu = \frac{\sum_{i=1}^N x_i f(x_i)}{\sum_{i=1}^N f(x_i)} \quad \text{Eq.2.8}$$

$$\text{Variância:} \quad \sigma^2 = \frac{\sum_{i=1}^N (x_i - \mu)^2 f(x_i)}{\sum_{i=1}^N f(x_i)} \quad \text{Eq.2.9}$$

$$\text{Assimetria:} \quad S_k = \frac{\sum_{i=1}^N (x_i - \mu)^3 f(x_i)}{\sigma^3 \sum_{i=1}^N f(x_i)} \quad \text{Eq.2.10}$$

$$\text{Curtose:} \quad K_u = \frac{\sum_{i=1}^N (x_i - \mu)^4 f(x_i)}{\sigma^4 \sum_{i=1}^N f(x_i)} - 3 \quad \text{Eq.2.11}$$

$$\text{Fator de correlação cruzada:} \quad cc = \frac{\sum_{i=1}^N x_i^+ x_i^- - \sum_{i=1}^N x_i^+ \sum_{i=1}^N x_i^- / N}{\sqrt{\left[\sum_{i=1}^N (x_i^+)^2 - \frac{(\sum_{i=1}^N x_i^+)^2}{N} \right] \left[\sum_{i=1}^N (x_i^-)^2 - \frac{(\sum_{i=1}^N x_i^-)^2}{N} \right]}} \quad \text{Eq.2.12}$$

sendo N o número de pontos na janela de fase em cada semiciclo da tensão aplicada (positivo e negativo). Caso a atividade dos pulsos de DP ocorra nas duas metades da tensão senoidal de teste aplicada, os momentos estatísticos devem ser computados separadamente para cada uma das metades.

Apenas com a média e o desvio padrão já é possível determinar um padrão. Porém, dada a complexidade dos dados de DP, a classificação pode ficar prejudicada. Estatísticas de ordem superior, como curtose, assimetria e correlação cruzada, podem se fazer necessárias. Denomina-se curtose o “grau de achatamento” de uma distribuição de frequências, geralmente unimodal, medido em relação ao de uma distribuição normal (de Gauss) tomada como padrão. Se a distribuição tem a mesma distribuição da normal, $K_u=0$; caso contrário, $K_u \neq 0$. Já as medidas de assimetria indicam o grau de assimetria de uma distribuição de frequências unimodal em relação a uma linha vertical que passa por seu ponto mais elevado. Se a distribuição é simétrica, $S_k=0$; caso contrário, $S_k \neq 0$. Por fim, o fator de correlação cruzada indica a dissimilaridade entre os formatos das

distribuições de descarga no semiciclo positivo (q_{\max} , φ_+) e no semiciclo negativo (q_{\max} , φ_-) da onda de tensão de referência. Se ocorrer assimetria total, $cc=0$; no caso de simetria total, $cc=1$. Uma descrição detalhada de curtose, assimetria e correlação cruzada pode ser encontrada em [27][16].

2.3.1.2

Caracterização do Sinal no Tempo

Neste caso, os parâmetros que compõem o vetor de características de DP são definidos por vários atributos significativos que descrevem a forma de onda do pulso no domínio do tempo, dentre os quais se destacam os seguintes [17]:

- a) Tempo de subida (t_r): período correspondente entre 10 % e 90 % do valor de pico do pulso antes do ponto de máximo;
- b) Tempo de decaimento (t_d): período correspondente entre 90 % e 10 % do valor de pico após o ponto de máximo;
- c) Largura do pulso (t_w): intervalo de tempo entre os níveis de 50 % do valor de pico antes e após o ponto de máximo;
- d) Área sob o pulso: área correspondente do pulso entre os níveis de 10 % do valor de pico antes e após o ponto de máximo.

A Figura 2.23 mostra os pontos de cálculo dos parâmetros no pulso de DP.

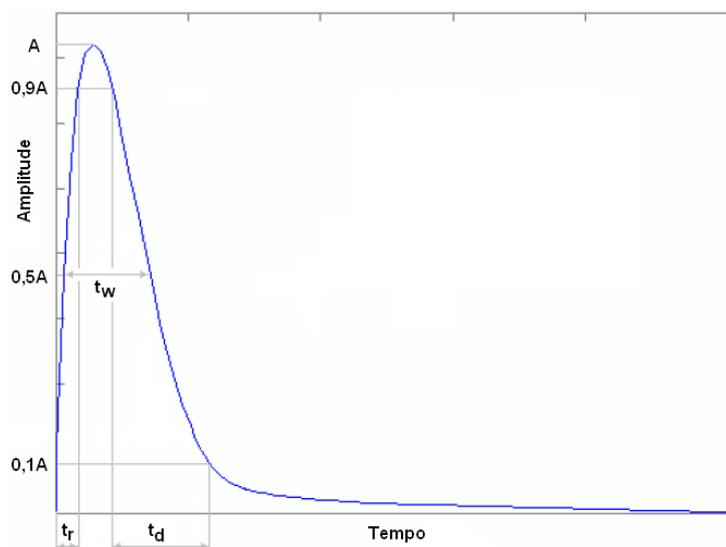


Figura 2.23 – Modelo do pulso de DP

Outros parâmetros dos pulsos de DP, que levam em consideração a carga aparente (q_i) e a tensão de teste (V_i), também podem ser avaliados [16]:

$$\begin{array}{l} \text{Corrente de} \\ \text{descarga média:} \end{array} \quad I = \frac{1}{\Delta t} \sum_{i=1}^N |q_i| \quad \text{Eq.2.13}$$

$$\begin{array}{l} \text{Razão quadrática:} \end{array} \quad D = \frac{1}{\Delta t} \sum_{i=1}^N q_i^2 \quad \text{Eq.2.14}$$

$$\begin{array}{l} \text{Potência de} \\ \text{descarga:} \end{array} \quad P = \frac{1}{\Delta t} \sum_{i=1}^N q_i V_i \quad \text{Eq.2.15}$$

$$\begin{array}{l} \text{Taxa de repetição:} \end{array} \quad n = N/\Delta t \quad \text{Eq.2.16}$$

$$\begin{array}{l} \text{Magnitude de pico} \\ \text{de descarga:} \end{array} \quad q_{\text{máx}} = \max(q_1, \dots, q_n) \quad \text{Eq.2.17}$$

$$\begin{array}{l} \text{Magnitude de} \\ \text{descarga média:} \end{array} \quad q_{\text{médio}} = \sum_{i=1}^N |q_i|/N \quad \text{Eq.2.18}$$

onde N é o número total de descargas dentro de um intervalo de tempo Δt .

2.3.1.3

Processamento de Sinais

Diversas técnicas de processamento de sinais vêm sendo aplicadas no reconhecimento de padrões em DP. Dentre as mais conhecidas, tem-se as transformadas de Fourier, *Wavelet*, Haar e Walsh.

A utilização da transformada de Fourier em sinais de DP baseia-se principalmente no fato de haver uma forte correlação entre a tensão de teste CA e os sinais de DP. Visto que a tensão de teste CA é periódica, o perfil de distribuição de carga $q(\varphi)$ tende a exibir um comportamento aproximadamente periódico. O vetor característico consiste nas componentes espectrais da distribuição de carga normalizada no domínio da frequência [28]:

$$\bar{a}_k = \frac{1}{P_0} \sum_{P=1}^{P_0} \frac{\int_{\varphi=0}^{2\pi} q_P(\varphi) \cos(k\varphi) d\varphi}{\sqrt{\int_{\varphi=0}^{2\pi} q_P^2(\varphi) d\varphi}} \quad \text{Eq.2.19}$$

$$\bar{b}_k = \frac{1}{P_0} \sum_{P=1}^{P_0} \frac{\int_{\varphi=0}^{2\pi} q_P(\varphi) \sin(k\varphi) d\varphi}{\sqrt{\int_{\varphi=0}^{2\pi} q_P^2(\varphi) d\varphi}} \quad \text{Eq.2.20}$$

sendo $q_p(\varphi)$ a carga aparente na posição de fase φ no p -ésimo ciclo ($q_p(3\pi) = q_{p+1}(\pi)$), e P_0 o número de ciclos considerados. A carga média normalizada é dada por [28]:

$$\bar{Q}(\varphi) = \frac{\bar{a}_0}{2} + \sum_{k=1}^{\infty} (\bar{a}_k \cos(k\varphi) + \bar{b}_k \sin(k\varphi)) \quad \text{Eq.2.21}$$

sendo a_0 o valor médio do sinal correspondente.

Na transformada de Fourier, as características extraídas dos sinais de DP são expressas apenas no domínio da frequência. Já as transformadas *wavelets* (TW) podem ser aplicadas para realizar análises no domínio do tempo [29]. Uma transformada *wavelet* contínua (TWC) é geralmente utilizada para obter a informação variante no tempo em detalhes, sendo a medida de similaridade entre um determinado sinal e um conjunto de funções oscilantes de decaimento rápido, conhecidas como *wavelets*. Cada membro (*wavelet*) deste conjunto é construído pela compressão ou dilatação de uma função original chamada *wavelet* mãe. A TWC de um sinal $x(t)$ com relação à *wavelet* mãe $\psi(t)$ é definida da seguinte forma:

$$TWC(a, b) = \frac{1}{\sqrt{a}} \int_{\varphi=0}^{2\pi} x(t) \psi\left(\frac{t-b}{a}\right) dt, \quad a > 0 \quad \text{Eq.2.22}$$

onde a é a dilatação ou fator de escala e b é o fator de translação.

A equação anterior mostra que o sinal unidimensional original $x(t)$ é mapeado em uma nova função em um espaço bidimensional em escala a e

translação b pela transformada *wavelet*. O coeficiente $TWC(a,b)$ representa o quão bem corresponde o sinal original $x(t)$ com a *wavelet* mãe $\psi(t)$ [30].

Em [31], as *wavelets* Malvar foram aplicadas na classificação do fenômeno de envelhecimento através da análise da degradação dos pulsos de descarga parcial. Em [32], o vetor característico é construído utilizando-se o coeficiente de aproximação e o coeficiente de detalhes em uma TWD de seis níveis, baseado na *wavelet* Daubechies 12 (db12).

A forma de cálculo dos coeficientes a_k e b_k nas aproximações por série de Fourier é uma estimativa da correlação dos coeficientes do sinal de DP $q(\varphi)$ com $\sin(\varphi)$ e $\cos(\varphi)$, bem como seus harmônicos de ordem mais elevada. Uma análise mais detalhada dos pulsos irá revelar que um pulso típico de DP não se encontra significativamente correlacionado com funções senoidais. Ao invés disso, as bordas rápidas e a duração pontual dos pulsos de DP relacionam-se muito melhor com funções como Haar e Walsh [28]. Assim, o vetor característico, utilizando-se essas funções, pode ser construído a partir dos coeficientes correspondentes e suas diferenças absolutas de uma maneira análoga àquela realizada para a aproximação da transformada de Fourier.

2.3.1.4

Processamento de Imagens

Os métodos de processamento de imagens consideram as informações fornecidas pelos parâmetros (φ, q, n) (ou seja, os mapas estatísticos de DP) como uma imagem. Desta forma, algoritmos de processamento de imagem são utilizados para avaliar as características das imagens. Dentre os métodos utilizados, destacam-se os baseados em análise de textura, em teoria de fractais e na decomposição de imagem baseada em *wavelets*.

Os métodos de análise de textura [33] avaliam a variação das imagens (mapas estatísticos de DP, no caso) em escala de cinza por meio de seus pixels. Desta forma, o mapa de DP é decomposto em pixels nas direções horizontal e vertical, referenciando a posição das DP em relação à fase da tensão de teste em que ocorrem e à sua própria amplitude, respectivamente. O número de DP em cada ponto do mapa é dado como referência aos valores da escala de cinza. Há

relatos de diferentes métodos de análise de textura [34][35][36][37], porém todos avaliam a amplitude relativa e a distribuição espacial dos pulsos de DP.

Os fractais são adequados para modelar formas complexas e fenômenos naturais em que as técnicas matemáticas existentes são consideradas insuficientes. São gerados por um processo iterativo iniciado através de uma construção por blocos, onde assumem um formato em que suas partes menores revelam grande semelhança com o inteiro quando ampliadas. Esta propriedade é chamada de auto-similaridade [38] e constitui-se no aspecto mais elementar da geometria fractal, quando o acontecimento macroscópico é constituído, inteiramente, por outros microscópicos similares. Considerando que as DP são um fenômeno natural com superfícies e formas complexas, os fractais podem ser utilizados para sua modelagem [39]. Um mapa de DP pode ser representado por meio de duas características fractais, denominadas lacunaridade e dimensão fractal, que são extraídas pelo método da contagem de caixas. Uma vez que a dimensão fractal é invariante à escala, ela pode ser utilizada para medir a rugosidade da superfície. A lacunaridade é a densidade da superfície fractal. Esses atributos podem ser utilizados para compor o vetor de características do mapa de DP [38]. Os cálculos matemáticos de lacunaridade e dimensão fractal podem ser encontrados em [40]. As características fractais foram utilizadas em [41], [42], [43], [44], [45] e [46] como atributos de entrada para classificar padrões de DP. Em 2006, por aproximação de dimensão fractal de imagens de DP, uma técnica melhorada de contagem de caixas diferencial foi criada [47]. Em [41], a técnica da compressão de imagem fractal é utilizada combinando tarefas de compressão de padrões DP com extração de características.

As transformadas *wavelets* multi-resolução têm apresentado bons resultados na decomposição de imagens e podem ser aplicadas na avaliação de mapas de DP onde exista mais de uma origem de DP envolvida, mesmo quando estas estiverem parcialmente sobrepostas no mapa de DP avaliado [48][49]. Em [48], quatro sub-imagens são criadas a partir do mapa de DP, sendo uma delas um coeficiente de aproximação (A) e as demais formam três conjuntos de coeficientes detalhados (vertical (V), horizontal (H) e diagonal (D)). Desta forma, cada sub-imagem é um componente do vetor característico, onde os coeficientes *wavelets* mostram as variações vertical, horizontal e diagonal da imagem de DP original.

2.3.2

Separação das Diferentes Fontes de Descargas Parciais

A identificação de defeitos associados às DP pode ser eficaz quando aplicada a padrões de DP que se referem a uma única fonte de DP. A presença simultânea de diferentes fontes de defeito ou ruídos em forma de pulso num isolamento de alta tensão pode resultar na gravação de uma sequência de pulsos mistos durante uma sessão de medição de DP [50]. Um procedimento de separação das diferentes fontes é então obrigatório para se evitar uma interpretação e avaliação de risco incorretas. As propriedades para a avaliação de uma separação efetiva podem ser muito úteis para selecionar o espaço mais eficiente e os algoritmos de agrupamento a serem adotados na separação dos sinais de DP [23].

2.3.2.1

Propriedades para uma Separação Efetiva

A separação pode ser definida como a classificação da sequência do sinal de pulso misto em subsequências, cada uma pertinente a uma tipologia de fonte específica (defeito), com base na suposição de que a mesma fonte gera sinais com forma semelhante, enquanto fontes diferentes são caracterizadas por diferentes formas de onda. Esta suposição é apoiada por investigações passadas que descobriram diferentes mecanismos de DP em defeitos artificiais recorrendo à análise da forma do pulso de DP [51][52]. A semelhança entre os sinais pode ser avaliada em domínios de tempo-frequência ou Função de Autocorrelação (ACF), considerando todo o conjunto das amostras digitais de cada sinal gravado ou espaços de formação, selecionando recursos que preservem as informações fornecidas pelo formato do sinal [23]. Os sinais com diferentes formas são agrupados de forma diferente nos espaços de classificação e, em geral, os diferentes grupos podem ser:

- a) Bem separados;
- b) Separados, mas cercados por pontos isolados;
- c) Sobrepostos parcialmente;

- d) Parcialmente sobrepostos com diferentes densidades;
- e) Totalmente sobrepostos.

A combinação entre o espaço de separação e o algoritmo de agrupamento constitui o estágio de separação do método de diagnóstico. A eficiência de separação pode ser avaliada considerando-se [23]:

- Classificação errada: as subclasses ainda contêm sinais mistos que geram padrões de DP mistos;
- Classificação fraca: menos classes do que fontes;
- Classificação excessiva: mais classes do que fontes;
- Classificação precisa: número de classes igual ao número de fontes.

Pode-se dizer que o espaço de separação mais eficaz é aquele capaz de fornecer uma série de grupos iguais às fontes de sinal em número, com tais grupos apresentando as maiores densidades e distâncias mais longas entre si.

2.3.2.2

Espaços para Separação de Fontes Múltiplas

Em geral, a dimensão do sinal deve ser convenientemente reduzida recorrendo-se a recursos capazes de preservar informações relacionadas à forma do sinal. Desta forma, diferentes mapas de densidade (espaços de separação) podem ser obtidos considerando-se diferentes características extraídas dos sinais resultantes do processo de redução de dimensionalidade. Para facilitar a análise, os mapas propostos na literatura ou implementados em instrumentos comerciais podem ser classificados em três grupos diferentes. Podem ser obtidos considerando-se valores globais ou médios (G1), valores médios de sub-intervalos (G2) ou componentes simples de vetores representativos dos sinais registrados avaliados nos três domínios diferentes (tempo/frequência/ACF) (G3) [23].

G1) Espaços de separação que utilizam parâmetros globais ou médios, como o tempo de subida, as quantidades de descarga, a energia do sinal, o tempo equivalente e a largura de banda equivalente, derivados dos vetores de processamento relacionados aos sinais gravados, são usados para formar espaços

de separação [21][53][54]. Um dos primeiros mapas é o chamado mapa TxF que representa o número de sinais em função de seu tempo equivalente e largura de banda equivalente [21].

G2) Nestes espaços, a amplitude média [55] ou a energia média do sinal [56], avaliadas em sub-intervalos do espectro de frequência selecionado por meio de filtros analógicos de banda estreita, são utilizadas para formar gráficos 2D/3D. A energia associada a cada nível de decomposição das ondas é avaliada por meio da Análise de Componentes Principais para formar espaços de separação em 3D [57].

G3) Neste caso, componentes simples do espectro de frequência [58] ou Função de Autocorrelação (ACF) [59] são selecionados para formar as coordenadas de um espaço de separação 3D.

2.3.2.3

Seleção dos *Clusters*

A seleção dos *clusters* representativos das diferentes fontes de defeito em um espaço de separação pode ser realizada por algoritmos de agrupamento, como o K-means ou o Fuzzy C-Means, por exemplo. Esses algoritmos baseiam-se normalmente em métricas de distância, número de vizinhos mais próximos e densidade de pontos para realizar a seleção dos *clusters*.

A extração do *cluster* também pode ser feita selecionando-se manualmente um intervalo nos eixos do espaço de separação após a inspeção visual, ou ainda via algoritmos adequados que tentam detectar os vales por declividade, detecção de joelho ou máximos locais.

2.3.2.4

Avaliação dos Espaços de Separação

A eficiência de separação de um espaço proposto pode ser avaliada de acordo com as propriedades para uma separação efetiva definidas no item 2.3.2.1. Deve se basear em resultados de separação de sinais oriundos de objetos de teste onde o número de fontes de DP é conhecido a priori [23]. Desta forma, a qualidade da separação é avaliada considerando-se:

a) A correspondência entre grupos e fontes de sinal. A classificação errada, fraca ou excessiva determina uma falha no teste.

b) Os sinais devidos a uma única fonte devem ser agrupados em um sub-espaço denso. A densidade do *cluster* é definida como o número médio de pontos por unidade de volume. Para evitar a influência dos pontos esparsos, o volume pode ser considerado como uma esfera centrada no centróide do *cluster* e seu raio é avaliado para conter um certo percentual do número total de pontos (68,3 %, por exemplo [23]). Se todos os eixos forem normalizados, as densidades avaliadas nos diferentes espaços são comparáveis.

c) Grupos de sinais devidos a diferentes fontes devem estar bem separados. A distância entre os diferentes *clusters* pode ser avaliada considerando-se a distância entre os centróides relevantes [23].

d) O número de pontos não classificados é outro índice de densidade. Quanto mais concentrado é o *cluster*, menos são os pontos esparsos não incluídos em *cluster* algum.

2.3.3

Classificação de Padrões em Descargas Parciais

Os padrões de DP podem ser classificados em várias categorias diferentes. Alguns trabalhos de investigação tentam classificar o tamanho das cavidades (ou “vazios”), em dielétricos sólidos, com base nos padrões de DP [60], enquanto outros estão interessados na determinação da fase onde ocorre a DP em linhas de transmissão trifásicas [61]. No entanto, a maioria das investigações é destinada a identificar o tipo de defeito no isolamento que ocasiona a DP. De acordo com [62], existem quatro principais tipos de defeitos que podem ser classificados com base em padrões de DP: cavidades, descargas de superfície, corona e arborescência elétrica.

Existe uma grande quantidade de classificadores para o reconhecimento de padrões [63]. São geralmente baseados em funções de decisão, distância ou verossimilhança. Os classificadores propostos para DP podem ser agrupados, basicamente, em classificadores de distância, classificadores estatísticos, classificadores baseados em redes neurais artificiais, classificadores baseados em

conceitos de sistemas fuzzy e classificadores baseados em máquina de vetor suporte.

2.3.3.1

Classificadores Baseados em Distância

Algoritmos de classificação baseados na função de distância supõem que existe uma biblioteca de referência para cada classe padrão ω_i (onde $i = 1, 2, \dots, M$ e M é o número de classes). Essa biblioteca de referência pode ser constituída a partir de um conjunto inicial de medidas de vetores característicos pertencentes a diferentes classes. Os principais algoritmos utilizados para a classificação baseada em distância são os classificadores de distância mínima e o classificador de vizinho mais próximo.

O conceito de classificação de padrões utilizando funções de distância baseia-se no fato de que o meio mais simples de avaliar a semelhança entre as características de medição, representadas por pontos no espaço Euclidiano, é a determinação de sua proximidade com os *clusters* padrões. O classificador determina a classe que está mais próxima ao vetor desconhecido (classificador de distância mínima). Existem várias maneiras de calcular a distância d_i entre o vetor característico desconhecido x e as classes de referência. Estas classes são representadas por um vetor característico médio m_i ($i=1, 2, \dots, M$), onde M denota o número de possíveis padrões de DP. As várias aproximações para computação de distância podem ser derivadas do classificador de distância generalizado:

$$d_i^z(x) = \sqrt[z]{\frac{1}{N} \sum_{j=1}^N \omega_{ij} |x_j - m_i|^z}, \quad i = 1, \dots, M \quad \text{Eq.2.23}$$

onde ω_{ij} é um fator de peso.

Fazendo-se $z=1$ e $\omega_{ij} = 1$, obtém-se o classificador de Manhattan [60]. A distância Euclidiana normalizada é obtida fazendo-se $z=2$ e $\omega_{ij}=1$ (o fator $1/N$ pode ser visto como um fator de escala) [16]. O fator ω_{ij} pode ainda ser substituído por $1/\sigma_{ij}^2$, onde σ_{ij} representa o desvio padrão do j -ésimo componente do vetor característico da i -ésima classe.

Já o classificador de vizinho mais próximo atribui ao padrão x de classe desconhecida a classe de seu vizinho mais próximo. O vizinho mais próximo é determinado por meio da distância Euclidiana. De forma mais específica, considera-se que existe um conjunto de amostras com classes conhecidas $\{s_1, s_2, \dots, s_L\}$, pertencentes às classes $\omega_1, \omega_2, \dots, \omega_M$. O padrão desconhecido x é atribuído à mesma classe ω_i de seu vizinho mais próximo pertencente ao conjunto $\{s_1, s_2, \dots, s_L\}$.

2.3.3.2

Classificadores Estatísticos

Os classificadores estatísticos são baseados em alguma consideração de origem estatística para a classificação de padrões. Em DP, existem dois algoritmos que são bastante utilizados: o classificador de Bayes e o classificador baseado na taxa de reconhecimento.

O classificador de Bayes [63] baseia-se na função de verossimilhança (ou função de densidade de probabilidade) para minimizar o risco por meio da menor probabilidade de erro na classificação. Ele utiliza propriedades estatísticas de dados previamente observados, cuja densidade de probabilidade é resumida pelo valor médio e a pela matriz de covariância. A função normal multivariada pode ser utilizada como a função densidade de probabilidade [16].

Para M classes de padrões, a função densidade normal multivariada $p(x/\omega_i)$ do vetor medido x quando o vetor pertence à classe ω_i , é dada por:

$$p(x/\omega_i) = \frac{\exp[-0,5(x - m_i)'C_i^{-1}(x - m_i)]}{(2\pi)^{N/2}|C_i|^{0,5}} \quad i = 1, 2, \dots, M \quad \text{Eq.2.24}$$

onde a densidade de cada classe é descrita pelo seu vetor médio m_i e pela matriz de covariância C_i . A função decisão para a classe ω_i é dada por:

$$d_i(x) = p(x/\omega_i)p(\omega_i) \quad \text{Eq.2.25}$$

onde $p(\omega_i)$ é a probabilidade a priori da classe ω_i . Assim, um padrão x é atribuído à classe ω_i se $d_i(x) > d_j(x)$, para todo $j \neq i$, sendo $i = 1, 2, \dots, M$.

O classificador baseado na taxa de reconhecimento tem sido utilizado com vetores característicos constituídos de operadores estatísticos (OE) no diagnóstico de fontes de DP. Esse classificador funciona da seguinte forma: para cada classe conhecida de DP (tipo de defeito), várias amostras são extraídas. A média aritmética e o desvio padrão de seus parâmetros característicos são calculados. Em seguida, obtém-se o intervalo de confiança (IC) – geralmente 95 % da distribuição normal – sobre esta média, da seguinte forma:

$$IC_1 = M_{oe} - \frac{\lambda\sigma}{\sqrt{N}} \quad IC_2 = M_{oe} + \frac{\lambda\sigma}{\sqrt{N}} \quad \text{Eq.2.26}$$

onde M_{oe} é a média aritmética do operador estatístico obtida da série de N observações de um mesmo tipo de defeito, σ é o desvio padrão da série e λ é um parâmetro de teste estatístico dependente de N .

Para um padrão desconhecido, todos os parâmetros do vetor característico são examinados segundo um intervalo de confiança (por exemplo, 95 %). Se um parâmetro encontra-se dentro do intervalo de confiança, este é computado para aquela classe de referência testada. Este procedimento é repetido para todos os parâmetros que constituem o vetor característico. A taxa de reconhecimento pode variar entre zero e o número máximo de parâmetros. Esse procedimento de pontuação é realizado para todas as classes de referência. A classe padrão que obtiver a maior quantidade de parâmetros é a vencedora.

2.3.3.3

Classificadores Baseados em Redes Neurais Artificiais

As redes neurais artificiais (RNAs) têm sido bem sucedidas em problemas de classificação de padrões de DP. Pode-se dizer que uma vantagem das RNAs sobre outros classificadores é a sua habilidade de aprender por exemplos. O conhecimento adquirido do conjunto de treinamento é extraído e armazenado nos conectores dos neurônios (“sinapses” e “bias”) durante a fase de aprendizado. Diversos tipos de estruturas de RNAs têm sido utilizadas no reconhecimento de padrões de DP, tais como rede neural *backpropagation*, mapas auto-organizáveis

de Kohonen, rede de quantização do vetor aprendizagem, rede neural *counterpropagation*, redes neurais modulares e cascadeadas.

A rede neural *backpropagation* tem sido a mais utilizada em termos de aplicações em geral nas mais diversas áreas. Ela recebe este nome pelo modo como trata a diferença entre o valor desejado na saída e o valor de saída da rede. Esta diferença (erro) é propagada na rede em sentido inverso e, nesse percurso de volta, os coeficientes da rede são atualizados [64]. A rede é composta de uma camada de entrada, pelo menos uma camada escondida e uma camada de saída. Cada camada é conectada à camada seguinte. O número de nós (neurônios) na camada de entrada é igual ao número de parâmetros selecionados para a classificação. O número de neurônios na camada de saída é igual ao número de fontes de DP a serem classificadas (classes). A determinação do número de neurônios da camada escondida depende de diversos fatores, incluindo a estatística disponível para o treinamento. Detalhes podem ser encontrados em [65][66][67].

No trabalho de reconhecimento de DP, o vetor característico, pertencente a um conjunto de sinais de DP utilizado para o treinamento da rede, é apresentado na entrada da rede de uma maneira aleatória. Dessa forma, a rede vai armazenando essas informações aprendidas em coeficientes w_{ij} (onde i é o neurônio anterior e j o neurônio posterior). O treinamento é realizado até se alcançar um índice de desempenho satisfatório. Desta forma, a rede está apta a ser utilizada para classificar vetores de classes desconhecidas apresentados em sua entrada. A Figura 2.24 mostra uma topologia típica de uma rede neural *backpropagation*.

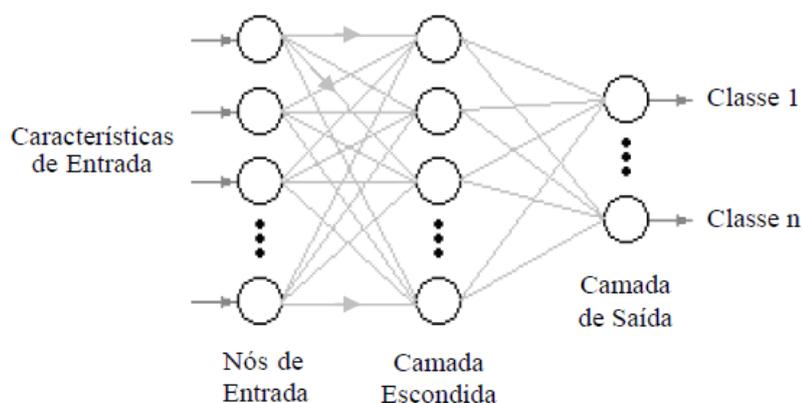


Figura 2.24 – Rede Neural *Backpropagation*

2.3.3.4

Classificadores Baseados em Conceitos de Sistemas Fuzzy

Esta classificação baseia-se principalmente no conceito de que alguns parâmetros não são facilmente quantificáveis [68]. Eles são geralmente representados como valores fuzzy, como, por exemplo: pequeno, médio e grande. Portanto, as características são mapeadas para a pertinência de um outro conjunto fuzzy baseado em seus próprios domínios. As principais etapas em lógica fuzzy são a fuzzificação das características de entrada, passando por uma máquina de inferência, e por último a defuzzificação da saída. O reconhecimento da fonte de DP é determinado com base na saída que é defuzzificada. Um diagrama esquemático do classificador baseado em lógica fuzzy é mostrado na Figura 2.25.

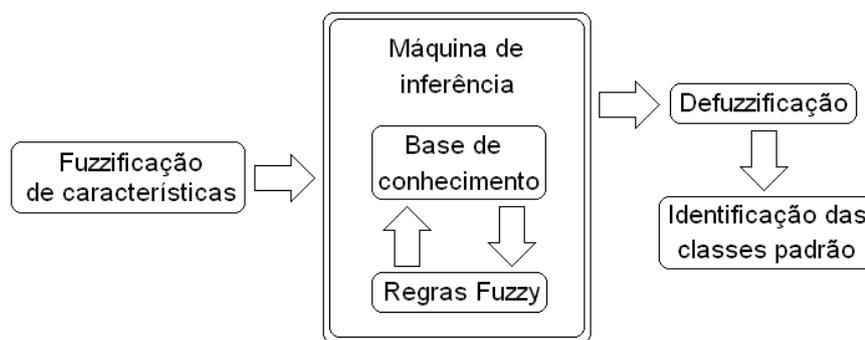


Figura 2.25– Classificador baseado em lógica fuzzy

A lógica fuzzy permite uma forma não exclusiva da classificação, fazendo com que cada padrão possa ser atribuído a mais de uma classe com um grau correspondente de pertinência [69][70]. Isto é extremamente benéfico para o monitoramento de falhas em equipamentos, onde decisões flexíveis podem ser tomadas de acordo com o arranjo particular do resultado da classificação fuzzy. Outros trabalhos são encontrados em [71][72][73].

2.3.3.5

Classificadores Baseados em Máquina de Vetor Suporte

Máquina de vetor suporte (SVM na sigla em inglês) é um grupo de máquinas de aprendizagem supervisionada que lida com as dificuldades de

classificação de padrões complexos. Utilizando uma classificação linear, a SVM mapeia os dados para um espaço dimensional maior [74]. Ele pode ser modificado para vários domínios e tarefas utilizando métodos de *kernel* com base na seleção do algoritmo de base e função de *kernel*. A SVM é adequada para reconhecimento de padrões quando se tem uma pequena quantidade de amostras, alta dimensionalidade e problemas não-lineares [75].

A SVM pode ser utilizada para encontrar as funções de um conjunto de dados de treino rotulados. Cada conjunto de dados pode ser representado por um vetor cuja dimensão depende da quantidade de padrões utilizados. A função é uma regressão ou uma função de classificação [75]. As SVMs têm muito menos parâmetros a serem modificados, fazendo com que sejam menos dependentes de procedimentos empíricos [74]. São baseadas em vários conceitos distintos, tais como máquinas lineares de aprendizagem, funções de *kernel*, espaços de recursos, otimização e aprendizagem estatística. Essas teorias são fundidas para formar o sistema de aprendizagem SVM. Foi comprovado que a SVM supera redes neurais em várias áreas [75]. Pesquisas com SVMs foram implementadas para determinação de partículas em transformadores [76][77] e também para classificar sinais de DP com diferentes fontes em cabos PILC [78].

2.4

Panoramas Atuais Relacionados ao Estudo

No estudo [79], publicado em 2016 pelo Cepel, verificou-se a possibilidade de detecção de DP em PRs instalados no campo via utilização de uma técnica não convencional: um sensor denominado "transformador de corrente de alta frequência (TCAF)" acoplado ao cabo de aterramento do equipamento para medição dos sinais drenados por este para o terra. No estudo em questão, foram realizados ensaios em laboratório, onde o ambiente é totalmente controlado, e em campo, sem o desligamento dos equipamentos, para avaliar a capacidade deste sensor na identificação de DP. Os resultados obtidos foram promissores, indicando que este tipo de procedimento poderia ser realizado no campo.

Segundo [79], após uma série de medições de PRs no campo com o TCAF, os sinais colhidos foram comparados aos medidos em laboratório em um PR com DP. Comparando a forma de onda entre os sinais registrados em um PR medido

no campo e o PR com DP avaliado em laboratório (utilizado como modelo), foi possível verificar semelhanças significativas entre os sinais. Conforme o estudo em questão, sinais que apresentam uma frente de onda lenta e um decaimento bastante amortecido, com várias oscilações, podem ser interpretados como sinais externos ao equipamento, o que não representa risco para ele. Já os sinais que apresentam frente de onda rápida e decaimento quase instantâneo são decorrentes de sinais de DP internas ao equipamento, conforme observado nas medições e avaliações no campo e também no laboratório.

O estudo [79] sugere então a implantação deste tipo de ensaio em campo como parte da manutenção preventiva, permitindo que inspeções regulares possam ser feitas em períodos definidos de tempo, e coloca como principais vantagens desta técnica as seguintes:

1) Rapidez na execução dos ensaios e na avaliação dos PRs em operação, permitindo assim a sua retirada de serviço antes de uma eventual falha.

2) Realização do ensaio com o equipamento energizado e em funcionamento normal, sem a necessidade de intervenções nos equipamentos e nem no sistema elétrico de potência.

3) Capacidade de prevenir falhas dos equipamentos, evitando assim gastos desnecessários e avarias em equipamentos vizinhos ou protegidos por estes.

4) Maior segurança às equipes de manutenção.

O estudo para PRs foi ampliado e estendido a TCs pelo Cepel, onde também se obteve sucesso na retirada de operação de TCs com suspeita de DP internas e comprovação disso em laboratório, em condições controladas de ensaio. Porém, tanto no caso dos PRs quanto no caso dos TCs, a classificação dos sinais medidos em campo é realizada apenas pelo olhar clínico do especialista, visto que não existe um sistema automático de classificação dos sinais. O trabalho aqui desenvolvido tem o objetivo de complementar a técnica, suprimindo esta carência. O banco de dados com sinais medidos ao longo de oito anos em PRs e TCs instalados em diversas subestações do Brasil é utilizado como base para investigar as características de sinais de equipamentos com DP e sem DP, como também para avaliar características discriminantes nestes sinais.

O objetivo último é a construção de um sistema automático de classificação dos sinais. Desta forma, são utilizadas técnicas baseadas em árvores de decisão e SVM (máquina de vetor suporte ou *Support Vector Machine*) para discriminar

sinais de equipamentos com DP e sem DP, com base nas melhores variáveis selecionadas e que apresentem elevado percentual de acerto no conjunto de teste, tanto para PRs quanto para TCs.

Esse sistema de classificação automática deverá então ser incorporado pelo Cepel na classificação de sinais dos equipamentos medidos em campo, o que será também muito útil para as equipes de manutenção de todo o sistema Eletrobrás, que poderão dele usufruir e acelerar a retirada de equipamentos com DP.

Em [80], é proposto que a eficiência dos espaços de separação seja avaliada considerando a densidade e a distância entre os diferentes grupos de sinais (*clusters*), supondo que um espaço de separação eficiente mostra grupos densos e bem separados de sinais. Estes dois índices são definidos da seguinte forma:

i) Os sinais devidos a uma única fonte de DP devem ser agrupados em um sub-espaço denso. A densidade de cada *cluster*, δ_σ , é definida como o número médio de pontos por unidade de volume.

$$\delta_\sigma = \frac{N_P}{V_\sigma} \quad \text{Eq.2.27}$$

Devido à natureza aleatória dos sinais de DP e para evitar a influência dos pontos esparsos, [80] considera que o volume V_σ é uma esfera centrada no centróide do *cluster* e seu raio é avaliado para conter 68,3 % do número total de pontos (N_P).

ii) Grupos de sinais devidos a fontes diferentes devem estar bem separados. A distância entre os diferentes *clusters* é avaliada considerando-se tanto a distância entre os centróides quanto a distância mínima entre sinais pertencentes a diferentes grupos.

A eficiência de um espaço de separação, segundo [80], pode então ser avaliada recorrendo-se à distância média entre os centróides do *cluster* (ii) e à densidade média (i) dos *clusters* encontrados no mesmo espaço. Ou seja, um bom espaço de separação deve mostrar classes bastante concentradas e bem distanciadas entre si, pois assim tenderá a apresentar o número correto de *clusters* que estão representados no mapa de DP.

No método de seleção automática do número ideal de *clusters* proposto aqui neste estudo, a análise baseia-se principalmente na distância relativa entre os

clusters, por meio da análise da “matriz de dissimilaridade relativa”. Seu cálculo leva em consideração, inicialmente, um número de *clusters* bem maior do que o número ideal representado na realidade. Toma-se a distância entre os dois *clusters* mais afastados entre si como referência e realiza-se a redução gradual deste número de *clusters* até se atingir uma determinada quantidade na qual a distância relativa entre eles não seja menor que um determinado percentual dos dois *clusters* mais afastados no início do processo (o “limiar de junção”).

Esse método é proposto aqui com base na observação de diferentes espaços de separação (montados com variáveis relevantes e não redundantes), em que alguns deles apresentavam *clusters* até densos, mas bem próximos entre si, e que, mesmo desta forma, podem ser separados com eficácia. Ou seja, acredita-se que a distância relativa entre os *clusters* possa representar melhor o espaço de separação quando comparada à distância absoluta.

Em [82], são apresentados resultados de investigações com o objetivo de correlacionar padrões de DP a tipologias específicas de defeitos que podem afetar o isolamento. Uma ampla coleção de mapas de DP representativos de defeitos típicos que podem afetar o isolamento de máquinas rotativas foram apresentados neste documento. Verificou-se que, em alguns casos, não existe uma correlação única entre a forma de um determinado mapa de DP e uma tipologia de defeito específica, ou seja, alguns defeitos podem gerar padrões com diferentes formas.

Os resultados apresentados em [82] mostram que cada tecnologia de isolamento exibe um conjunto específico de padrões de DP (mapas), sendo necessárias regras adicionais para identificar padrões cuja causa principal possa ser duvidosa. Assim, qualquer procedimento de identificação baseado na análise de forma do mapa de DP deve ser adaptado para cada tecnologia de isolamento específica para evitar identificações errôneas de defeitos.

No estudo desenvolvido aqui, é sugerida a análise dos histogramas de fase dos mapas de DP para caracterização destes, por meio da correlação de Pearson de mapas a classificar com mapas que componham um banco de dados histórico com padrões já classificados. Esse banco deve conter sinais representativos de cada tipo de defeito, podendo apresentar até mesmo mais de um padrão diferente para o mesmo tipo de defeito, como reforçado no estudo [82]. Quanto mais padrões, maior a chance do sistema de classificação proposto aqui acertar a classe do defeito.

É importante salientar que o sistema de classificação proposto aqui deve ser utilizado como um auxílio à classificação feita por um especialista, dada a possibilidade de ocorrência de padrões ambíguos, ou seja, padrões semelhantes para defeitos diferentes, citada também por [82]. Desta forma, pode-se avaliar no sistema proposto não só o padrão com maior correlação de Pearson entre os histogramas, mas também o segundo e os próximos, e então ponderar com outras informações disponíveis (inspeção visual, por exemplo). A simples redução do número de possibilidades de defeito já seria de grande valia para auxiliar na análise do especialista.

É importante frisar também que o banco de dados histórico deve estar voltado para o tipo específico de equipamento a que se destina classificar. Ou seja, deve-se criar um banco de dados exclusivo para ser utilizado em geradores e motores e outro exclusivo para transformadores de potência, por exemplo, dada a especificidade de cada isolamento e suas diferentes assinaturas de defeito.

Em [81], a análise de forma dos mapas de DP é proposta como uma ferramenta para teste de aceitação e controle da qualidade de barras novas e bobinas individuais para máquinas rotativas. Um novo sistema de isolamento pode ser considerado em boas condições quando é afetado apenas pela DP que ocorre em micro-cavidades distribuídas, localizadas dentro do isolamento da parede principal, e nenhum outro padrão DP for detectado até o valor máximo da tensão de teste.

Por meio do método sugerido, [81] obtém resultados comparáveis, em termos de identificação do processo de descarga, testando diferentes tipos de barras e bobinas, em diferentes níveis de tensão e utilizando diferentes instrumentos de medição de DP e acopladores capacitivos. Um procedimento de separação baseado na análise do mapa TxF (espaço de separação tempo-frequência, formado pelas variáveis 17 e 18 mencionadas no estudo desenvolvido aqui e detalhadas no capítulo 3) foi utilizado para separar as múltiplas fontes de DP e o ruído de fundo.

Dois tipos de padrões de DP foram encontrados em [81] para a mesma atividade de DP (mesmo tipo de defeito), um detectado no início e outro no final do ensaio elétrico. Estes padrões podem ser detectados e claramente identificados mesmo quando são utilizados diferentes sistemas de medição de DP com uma

sensibilidade adequada e fornecidos por um sistema de separação baseado no mapa TxF.

A eficácia do espaço de separação TxF também pode ser comprovada em [82], onde os autores criam defeitos artificiais em barras de máquinas rotativas e as submetem a ensaios em laboratório para captura dos sinais de DP. O espaço TxF é utilizado para separar as diferentes fontes dos sinais medidos nas barras.

Percebe-se que as variáveis 17 e 18, utilizadas também no estudo desenvolvido aqui, vêm se mostrando como um dos melhores espaços de separação de fontes de DP [82]. Porém, haverá casos (como os apresentados neste estudo) em que a utilização de outras variáveis pode representar uma melhor qualidade do processo de agrupamento. Vale ressaltar que optou-se aqui por avaliar o desempenho das técnicas propostas neste estudo através da comparação dos percentuais de acerto obtidos com as diferentes variáveis quando aplicadas aos bancos de dados aqui utilizados, não apresentando portanto percentuais de acerto obtidos em estudos anteriores com as variáveis 17 e 18 mencionadas, dada a especificidade de cada banco de dados utilizado pelos diferentes autores.

3

Dados e Técnicas Utilizados no Estudo

3.1

Isolamentos de Alta Tensão

São utilizados aqui dados experimentais colhidos em laboratório, especialmente para este estudo, e dados reais, colhidos em campo durante as campanhas de medição realizadas pelo Cepel para avaliar o isolamento de alguns equipamentos de alta tensão com relação à presença de DP internas. Também foram utilizados sinais de laboratório medidos em equipamentos retirados de operação após medições do Cepel constatarem possibilidade de DP internas nestes.

Os isolamentos ensaiados em laboratório (exclusivamente para este estudo) são denominados "células de teste de DP", e se propõem a representar os diferentes tipos de defeitos relacionados às DP que podem ser encontrados em equipamentos de grande porte de isolamento sólido. Existem também células para representar isolamentos líquidos, imersos em óleo mineral, principalmente, mas que não serão utilizadas aqui por retratarem natureza diferente (visto que se pretende comparar sinais no tempo das células). Ou seja, cada célula está associada fisicamente a uma ou mais partes internas do equipamento. Para este estudo foram ensaiadas duas células diferentes em laboratório, que são descritas a seguir:

a) célula ponta-plano no ar (PPA) – alta tensão aplicada na agulha (parte de cima da célula), com placa circular metálica aterrada (parte de baixo da célula), e o ar separando estas duas partes. Representam descargas que ocorrem diretamente entre a alta tensão e as partes aterradas do equipamento, como em grampos de fixação, conectores, tirantes e outras partes do equipamento que apresentem pontas salientes.

b) célula com eletrodos flutuantes no ar (EFA) – alta tensão aplicada no cilindro metálico (parte de cima da célula), com placa circular metálica aterrada (parte de baixo da célula), dois pequenos eletrodos flutuantes metálicos no meio (separados entre si e isolados da alta tensão e do terra), e o ar separando a alta

tensão do terra. Representam descargas envolvendo condutores de alta tensão e partes não aterradas do equipamento, tais como descargas entre saídas de alta tensão do equipamento, descargas em objetos condutores flutuantes e descargas provocadas por mau contato nos conectores.

Outro isolamento ensaiado em laboratório foi uma “barra de gerador” com DP, cuja foto é apresentada na Figura 3.1 juntamente com a das células. Essas barras são ligadas em série no gerador para formar a bobina do estator, como mostrado na figura 3.2.

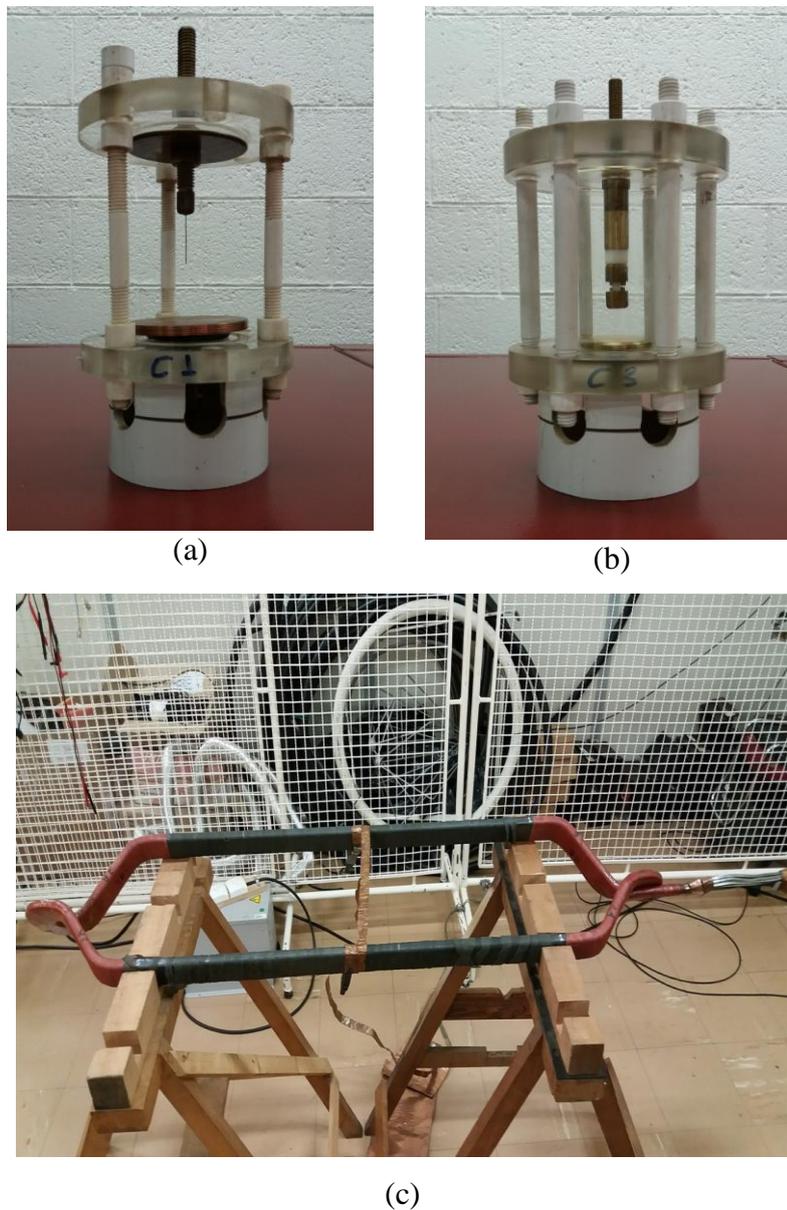


Figura 3.1 - Células de teste de DP e barra de gerador: (a) célula ponta-plano no ar (PPA);(b) célula com eletrodos flutuantes no ar (EFA); e (c) barra de gerador

Os dados reais utilizados neste estudo foram obtidos em ensaios de campo de para-raios (PRs) e transformadores de corrente (TCs) medidos em diversas subestações no Brasil, ao longo de oito anos, e também sinais de DP medidos pelo Cepel em geradores elétricos e em um motor, todos de grande porte. Foram utilizados também sinais de laboratório medidos em um PR e um TC retirados de operação após medições do Cepel constatarem possibilidade de DP internas nestes. Nas figuras 3.2, 3.3 e 3.4, são mostradas fotos ilustrativas de alguns destes equipamentos.

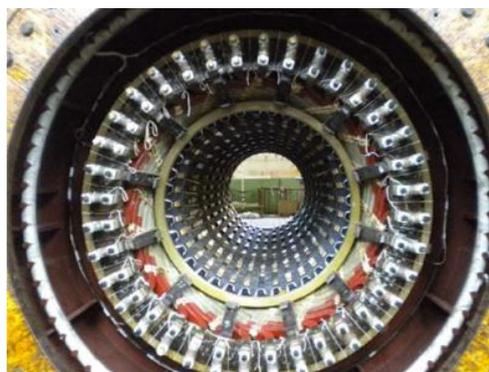


Figura 3.2 – Foto ilustrativa de um gerador [83]



Figura 3.3 – Foto de um dos PRs de 345 kV medidos em campo

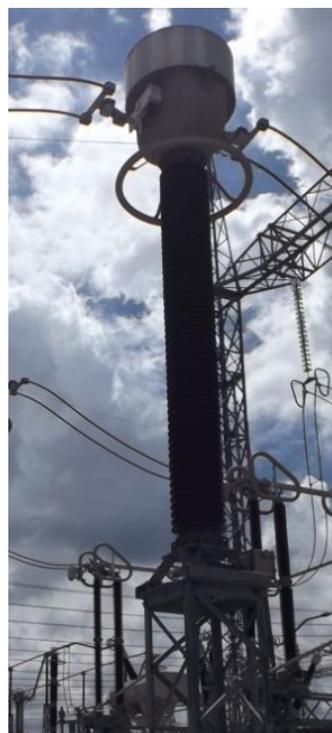


Figura 3.4 – Foto de um dos TCs de 345 kV medidos em campo

Os PRs, ou supressores de surtos de tensão, são os equipamentos instalados nas subestações responsáveis pelo controle de parte das sobretensões existentes nos sistemas elétricos de potência, contribuindo decisivamente para a sua confiabilidade, economia e continuidade de operação. Os equipamentos de uma subestação podem ser solicitados por sobretensões provenientes de ocorrências internas ao sistema ou associadas com a ocorrência de descargas atmosféricas. Com o objetivo de impedir que estes equipamentos sejam danificados, é necessária a instalação de dispositivos de proteção contra sobretensões, sendo os PRs os equipamentos adequados para esta finalidade. Atuam como limitadores de tensão, impedindo que valores acima de um determinado nível pré-estabelecido em projeto, adequados para fornecer uma margem de segurança, normalmente padronizada, possam alcançar os equipamentos para os quais fornecem proteção [11].

Nas subestações, os medidores, relés de proteção e dispositivos de controle do tipo de corrente alternada funcionam alimentados por tensões e correntes supridas por transformadores de potencial e de corrente (TPs e TCs), respectivamente. Estes transformadores proporcionam isolamento contra a alta tensão do circuito primário. São chamados de transformadores de instrumentos (TIs) e suprem os medidores, relés de proteção e dispositivos de controle com tensões e correntes proporcionais às que circulam pelos circuitos de potência, sendo estas suficientemente reduzidas para que estes instrumentos possam ser fabricados com um nível menor de isolamento. Os medidores, relés de proteção e outros dispositivos de controle ficam instalados no circuito secundário dos TIs. Os TCs têm o seu enrolamento primário ligado em série com o circuito de alta-tensão (ou de potência, também chamado de circuito primário) [11].

O gerador elétrico é o equipamento responsável por transformar energia mecânica em energia elétrica. Na grande maioria das centrais, tanto hidrelétricas quanto termelétricas, ele é do tipo síncrono, embora geradores de indução (assíncronos) sejam utilizados em pequeno número de projetos, em geral de pequeno porte (os dados utilizados neste estudo foram medidos em geradores síncronos). Esse tipo de equipamento normalmente apresenta elevadas dimensões físicas. Já o motor elétrico faz exatamente o contrário do gerador, ou seja, transforma energia elétrica em mecânica, podendo também ser síncrono (o qual foi utilizado neste estudo) ou assíncrono. Motores de grande porte são

encontrados principalmente em indústrias de papel e metalúrgicas, e servem, por exemplo, para bombeamento de água para refrigerar partes do sistema de produção.

Dos diversos tipos de isolamentos de alta tensão, pode-se dizer que existem aqueles mais frágeis, que suportam menos DP (como os encontrados nos PRs e TCs), e aqueles mais robustos, que suportam mais DP (como os encontrados nos geradores e motores). Nos mais frágeis, a detecção de DP já é suficiente para programar a retirada do equipamento de operação. Já nos mais robustos, é preciso verificar o número de fontes de DP presentes no isolamento e a classificação destas fontes, para saber se é preciso intervir na máquina (caso de defeitos com grau de risco de falha alto ou médio) ou apenas administrar a evolução das DP (caso de defeitos com grau de risco de falha baixo). A Figura 3.5 ilustra de forma resumida tudo isso que foi comentado.

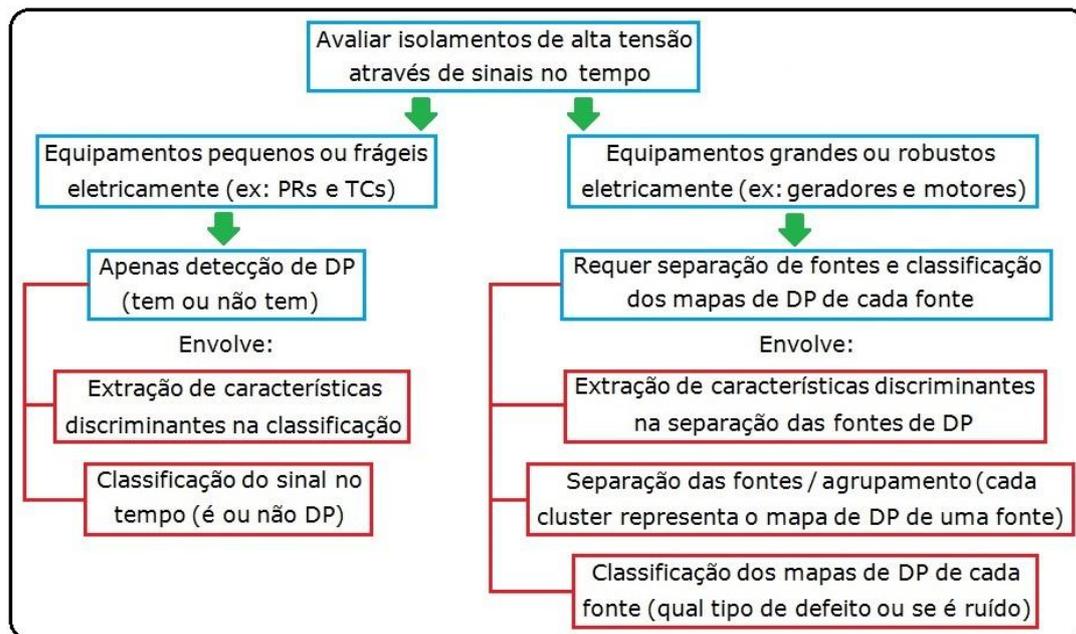


Figura 3.5 – Tratamento necessário em cada tipo de isolamento

3.2

Sistemas de Medição

O sistema de medição utilizado para obter os sinais das células de teste em laboratório, mostrado na Figura 3.6, é o mesmo utilizado para medir os sinais do gerador e motor no campo.

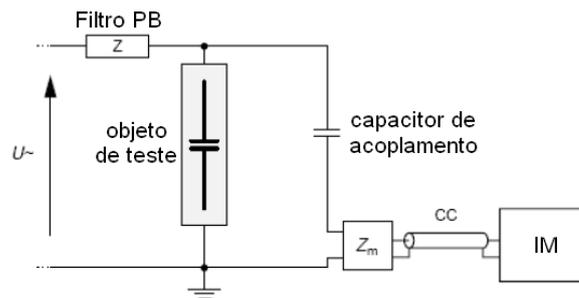


Figura 3.6 – Circuito de medição de DP utilizado nas células de teste em laboratório e no gerador em campo

Onde: Filtro PB → Filtro passa-baixas;
 U → Tensão de teste;
 CC → Cabo coaxial;
 Z_m → Impedância de medição;
 IM → Instrumento de medição.

Na Figura 3.6, o "filtro PB" é responsável por não deixar passar ao circuito de medição sinais de alta frequência oriundos da fonte de alta tensão, utilizada para aplicar a tensão de ensaio no objeto de teste (célula ou gerador), visto que tais sinais poderiam ser confundidos com sinais de DP vindos do objeto de teste durante a medição. Este tipo de ensaio é chamado de *off-line*, pois é feito com o objeto (no caso, o gerador) desligado e fora de operação. O "capacitor de acoplamento" funciona como um sensor de DP e também, juntamente com Z_m , como um divisor de tensão para diminuir a tensão que chega ao IM. Z_m também funciona como filtro passa-alta, deixando passar sinais de alta frequência ao IM (que representam as DP oriundas do objeto de teste) e reduzindo a próximo de zero a amplitude dos sinais de baixa frequência (como o 60 Hz, por exemplo). Através do cabo coaxial, o IM capta os sinais de DP para registro. O IM é

formado por um osciloscópio conectado a um computador com software de captura para registro dos sinais. O software utilizado neste caso é o IMA-DP (instrumentação para monitoramento e análise de descargas parciais), desenvolvido pelo Cepel [84]. No caso de alguns geradores e do motor referidos aqui, foram utilizados mapas de DP medidos com outro software comercial.

Para o caso de ensaios *on-line* com a utilização de vários sensores (capacitor de acoplamento), estes são instalados permanentemente fazendo alguns acessos em partes do objeto de teste (Figura 3.6) para a detecção ser mais sensível naquele ponto (caso contrário, os sinais seriam muito amortecidos). Nestes casos também não é utilizado o “Filtro PB” e a tensão de ensaio U (Figura 3.6) é a própria tensão gerada presente entre o ponto de acesso do sensor e o terra. No caso dos sinais de medição *on-line* deste estudo, o software utilizado também é o IMA-DP.

O circuito de medição utilizado para capturar os sinais dos PRs e TCs no campo é um pouco diferente, como mostrado na Figura 3.7. Neste circuito, o objeto de teste (PR ou TC) não é retirado de operação para a medição, e um sensor, denominado "transformador de corrente de alta frequência (TCAF)" (como ilustrado na Figura 3.8), é acoplado ao cabo de aterramento do equipamento (PR ou TC) para medição dos sinais drenados por este para o terra, estando o equipamento sob sua tensão de operação. Esta técnica de medição foi proposta primeiramente em [85] para aplicação em PRs e posteriormente aprimorada e estendida também a TCs pelo Cepel. Vem sendo amplamente aceita como solução para avaliação de PRs e TCs no campo sem necessidade de desligamento e com casos reais de sucesso na retirada de operação de equipamentos com DP internas pelo Cepel [86][87][88][10][79]. Isso permite a rápida avaliação destes equipamentos nas subestações e previne futuras falhas destes.

Na Figura 3.7, o TCAF funciona como sensor de DP e também como filtro passa-alta, reduzindo sinais de baixa frequência (como o de 60 Hz) que vêm do sistema e capturando somente os sinais de DP (que são de alta frequência) vindos do equipamento. Neste tipo de medição, o IM capta os sinais de DP para registro, sendo formado por um osciloscópio conectado a um computador com software de captura para gravação dos sinais. O software utilizado neste caso é um aquisitor do sinal mostrado na tela do osciloscópio, gravado em formato de texto.

No caso do PR e do TC retirados de operação e levados para análise em laboratório após medições do Cepel constatarem possibilidade de DP internas nestes, o circuito de medição utilizado é o mesmo da Figura 3.7, com a diferença de que a tensão U é aplicada pela fonte de tensão do laboratório, com baixo ruído de fundo (muito diferente do ambiente de subestação), e também de que neste caso é possível realizar a calibração da medição. Ou seja, antes de se aplicar tensão no objeto de teste, um sinal de carga conhecida é injetado em seus terminais e medido pela instrumentação, onde o valor lido é casado com o valor injetado, algo impossível de se fazer no campo sem desligar o equipamento.

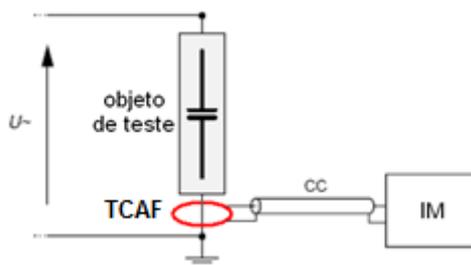


Figura 3.7 – Circuito de medição de DP utilizando TCAF



Figura 3.8 – Exemplo de um TCAF conectado ao cabo de aterramento do equipamento

Onde: $U \rightarrow$ Tensão de operação do equipamento;
 $CC \rightarrow$ Cabo coaxial;
 $IM \rightarrow$ Instrumento de medição;
 $TCAF \rightarrow$ Transformador de corrente de alta frequência.

3.3

Sinais Utilizados

É utilizado neste estudo um total de 28 sinais medidos nas duas células de teste e na barra de gerador com DP (tabela 3.1). São sinais no tempo com duração de 0,5 s cada, obtidos com a aplicação em laboratório de tensões de até 16 kV sob os terminais das células e da barra de gerador utilizando uma fonte de alta tensão CA de 60 Hz. Cada sinal contém vários pulsos no tempo com duração de 2 μ s

cada, em média, colhidos ensaiando os objetos individualmente e também em paralelo.

Tabela 3.1 - Sinais medidos em laboratório nas células de teste e na barra de gerador

Objeto de teste		Número de sinais	Detalhes
Individuais	Célula PPA	4	Sinais no tempo com duração de 0,5 s cada com objetos de teste sob tensão máxima de 16 kV. Cada sinal contém vários pulsos com duração de 2 μ s cada, em média.
	Célula EFA	4	
	Barra de Gerador	4	
Em paralelo	Célula PPA // Célula EFA	4	
	Célula PPA // Barra de gerador	4	
	Célula EFA // Barra de gerador	4	
	Célula PPA //Célula EFA // Barra de gerador	4	

Um total de 260 sinais medidos em PRs e 135 sinais medidos em TCs foi utilizado aqui (tabela 3.2). Esses sinais foram obtidos da avaliação em campo de 66 PRs e 41 TCs, com tensão de operação variando de 69 a 765 kV (tabela 3.3), colhidos ao longo de oito anos de medição em diversas subestações no Brasil, contendo equipamentos com e sem DP internas. São sinais no tempo com duração de 2 μ s cada medidos com os objetos de teste sob tensão de operação (sem necessidade de desligamento do equipamento), onde cada sinal representa um pulso no tempo.

Sinais medidos em laboratório em um PR e em um TC retirados de campo com suspeita de DP pelo Cepel (suspeita essa confirmada posteriormente em laboratório) também foram utilizados aqui, e são detalhados na tabela 3.4.

Um sinal medido de forma *off-line* em uma das fases de um gerador elétrico de grande porte, no qual o sinal de DP está em meio a sinais de ruído, também é utilizado aqui (tabela 3.5). O sinal tem duração de 0,5 s e foi obtido com a aplicação em campo de 8 kV de tensão na fase do gerador, utilizando uma fonte

de alta tensão CA de 60 Hz. Esse sinal é formado por vários pulsos no tempo com duração de 2 μ s, em média.

Tabela 3.2 - Sinais medidos em campo nos PRs e TCs

Tipo de equipamento	Número de sinais		Detalhes
	PRs	TCs	
Sem DP	121	75	Sinais no tempo com duração de 2 μ s cada com objetos de teste sob tensão de operação variando de 69 a 765 kV, onde cada sinal representa um pulso.
Com DP	139	60	
Total	260	135	

Tabela 3.3 - Distribuição dos PRs e TCs pela tensão de operação

Tensão de Operação	Número de equipamentos					
	765 kV	500 kV	345 kV	230 kV	69 kV	Total
PRs	7	27	0	21	11	66
TCs	7	11	4	19	0	41
Total	14	38	4	40	11	107

Tabela 3.4 - Sinais medidos em laboratório no PR e no TC retirados de operação

Tipo de equipamento	Número de sinais	Tensão de operação	Classe
PR	12	500 kV	Com DP
TC	21	500 kV	Com DP

Tabela 3.5 - Sinal no tempo medido em campo no gerador 1

Objeto de teste	Número de sinais	Detalhes
Uma fase do gerador 1	1	Sinal no tempo com duração de 0,5 s com objeto de teste sob tensão máxima de 8 kV. O sinal contém vários pulsos com duração de 2 μ s cada, em média.

São utilizados aqui também mapas de DP medidos de forma *on-line* em oito acopladores (sensores de DP) instalados ao longo de uma das fases do estator de um gerador hidrelétrico de uma grande usina, conforme detalhado na tabela 3.6.

Tabela 3.6 – Mapas de DP medidos em campo em monitoramento *on-line*

Objeto de teste	Número de sinais	Detalhes
Uma fase do gerador 2	320	Mapas de DP resultantes de monitoramento <i>on-line</i> no gerador 2 e medidos em 8 sensores diferentes espalhados pela fase (40 sinais de cada sensor). Objeto de teste sob tensão permanente de 8 kV.

Por fim, são utilizados também mapas de DP unipolares (sinais colhidos em módulo por limitações no hardware de medição do sistema comercial) medidos de forma *off-line* em 5 equipamentos de grande porte – 4 geradores e 1 motor –, conforme detalhado na tabela 3.7.

Tabela 3.7 – Mapas de DP medidos em campo em medições *off-line*

Objeto de teste	Número de sinais	Detalhes
Uma fase do gerador 3	1	Mapas de DP unipolares resultantes de medição <i>off-line</i> nos objetos de teste e medidos em sensor único instalado na saída de fase de cada um. Objetos de teste com tensão de ensaio variando de 8 a 15 kV (nominal de cada objeto).
Uma fase do gerador 4	1	
Uma fase do gerador 5	1	
Uma fase do gerador 6	1	
Uma fase do motor 1	1	

3.4

Extração de Características dos Sinais

Os dados estudados aqui foram caracterizados e analisados em relação a parâmetros extraídos do sinal no tempo, seja para realizar a separação das diferentes fontes de DP (caso das células e do gerador), seja para classificar os

sinais como sendo ou não de DP (caso dos PRs e TCs). Para isso, 23 diferentes variáveis (ou parâmetros) do sinal no tempo foram avaliadas, sendo algumas sugeridas na literatura e outras propostas aqui para testes. Vale ressaltar que os sinais no tempo utilizados aqui foram normalizados de forma que suas amplitudes assumissem valores entre 0 e 1. A clara distinção entre esta origem de cada uma das variáveis (se da literatura ou se proposta aqui) será apresentada na tabela 3.8, mostrada ao final da descrição destas. A seguir são apresentadas as variáveis:

Variável 1 - Disp_ACF_1: Maior pico de dispersão da Função de Autocorrelação (ACF) normalizada do sinal no banco de dados em questão. Equivale ao valor no eixo y da ACF do sinal, normalizada, na amostra em que ocorre a maior dispersão dessa função no banco de sinais disponível.

Para o cálculo da ACF, cada sinal de pulso é representado por um vetor S com N_c amostras de tempo digital, $\{s_1, s_2, \dots, s_{N_c}\}$. Dado um conjunto de N vetores, $\{S_1, S_2, \dots, S_N\}$, que representam a sequência de pulsos no banco de dados em análise, a ACF normalizada do sinal pode ser calculada da seguinte forma [80]:

$$R_S[k] = \frac{\sum_{i=1}^{N_c-k} s(i)s(i+k)}{N_c} \frac{1}{E_S} \quad \text{Eq.3.1}$$

onde k varia de 1 a N_c , sendo $R_S[k=N_c]=0$, formando um vetor ACF com N_c amostras. E_S é a energia total do sinal, dada por:

$$E_S = \frac{\sum_{i=1}^{N_c} s^2(i)}{N_c} \quad \text{Eq.3.2}$$

Cada ACF do conjunto de sinais é normalizada em relação à energia do sinal e as semelhanças entre diferentes ACFs depende apenas da sua forma. Para verificar as amostras (k) onde ocorrem os maiores picos de dispersão nas ACFs, é necessário calcular o desvio padrão, $\sigma[k]$, das amplitudes para cada componente k dos N vetores que representam os sinais medidos, da seguinte forma:

$$\sigma[k] = \frac{\sqrt{\sum_{i=1}^N (S_k(i) - S_{k\mu})^2}}{N - 1} \quad \text{Eq.3.3}$$

O valor a ser atribuído à variável 1 corresponde ao valor de $R_s[k_I]$, sendo k_I a amostra que apresenta a maior dispersão dentre os sinais em estudo ($\sigma[k_I]$).

Variável 2 - Disp_ACF_2: Segundo maior pico de dispersão da ACF normalizada do sinal no banco de dados em questão [80].

Variável 3 - Disp_ACF_3: Terceiro maior pico de dispersão da ACF normalizada do sinal no banco de dados em questão [80].

Vale ressaltar que os três maiores picos de dispersão da ACF normalizada vêm sendo utilizados com sucesso na separação de fontes de DP, segundo publicado por [80] em 2015.

Variável 4 - Disp_FFT_ACF_1: Maior pico de dispersão do valor absoluto da Transformada Discreta de Fourier (TDF ou FFT, do inglês *Fast Fourier Transform*) da ACF normalizada do sinal no banco de dados em questão.

Para um sinal discreto x (que seria, neste caso, a ACF do sinal normalizada) com N amostras, sua FFT é dada por:

$$X(k) = \sum_{n=1}^N x(n) \exp\left(\frac{-j2\pi(k-1)(n-1)}{N}\right), 1 \leq k \leq N \quad \text{Eq.3.4}$$

O vetor $X(k)$ é composto por números complexos, mas o cálculo da variável em questão está relacionado somente ao valor absoluto da FFT da ACF. Em termos de coordenadas polares, utiliza-se então somente a amplitude da FFT, deixando-se a fase de fora do cálculo.

O desvio padrão dos valores absolutos da FFT da ACF dos sinais é calculado para determinar as amostras onde ocorrem os maiores picos de dispersão, como realizado para a variável 1. Atribui-se à variável 4 o valor absoluto da FFT da ACF na amostra onde ocorre a maior dispersão dos sinais no banco em estudo.

Variável 5 - Disp_FFT_ACF_2: Segundo maior pico de dispersão do valor absoluto da FFT da ACF normalizada do sinal no banco de dados em questão.

Variável 6 - Disp_FFT_ACF_3: Terceiro maior pico de dispersão do valor absoluto da FFT da ACF normalizada do sinal no banco de dados em questão.

Variável 7 - DeltaTmax1max2: Intervalo de tempo (ΔT) entre o primeiro e o segundo máximos da ACF do sinal normalizada.

Variável 8 - DeltaTmax1min1: Intervalo de tempo (ΔT) entre o primeiro máximo e o primeiro mínimo da ACF do sinal normalizada.

Variável 9 - DeltaFmax1max2: Intervalo de frequência (ΔF) entre o primeiro e o segundo máximos do valor absoluto da FFT da ACF do sinal normalizada.

Variável 10 - Disp_FFT_Sinal_1: Maior pico de dispersão do valor absoluto da FFT do sinal normalizado no banco de dados em questão.

Variável 11 - Disp_FFT_Sinal_2: Segundo maior pico de dispersão do valor absoluto da FFT do sinal normalizado no banco de dados em questão.

Variável 12 - Disp_FFT_Sinal_3: Terceiro maior pico de dispersão do valor absoluto da FFT do sinal normalizado no banco de dados em questão.

Variável 13 - Pico_F_ACF: Valor da frequência onde ocorre o maior pico do valor absoluto da FFT da ACF normalizada do sinal no banco de dados em questão.

Variável 14 - Largura_Pulso_ACF: Intervalo de tempo (ΔT) entre o pico da ACF do sinal normalizada e o ponto onde esta cai abaixo de 30 % do valor de pico.

Variável 15 - Pico_F_Sinal: Valor da frequência onde ocorre o maior pico do valor absoluto da FFT do sinal normalizado no banco de dados em questão.

O trabalho [58], publicado em 2010, testa uma variável semelhante a esta e também às variáveis 10, 11 e 12, citadas anteriormente. Naquele caso, os autores realizam uma medição centrada em frequência (ainda na captura dos dados), escolhendo três diferentes frequências. A medição então é feita utilizando-se filtros passa-faixa, centrados nas frequências pré-selecionadas, com medição em três canais, um para cada filtro. Os picos de amplitude do valor absoluto da FFT do sinal de cada um dos três canais são então utilizados como parâmetros de caracterização dos sinais e plotados em um gráfico 3D (cada canal correspondendo a um eixo) para separação das diferentes fontes de DP. O método foi testado em um cabo de alta tensão e mostrou-se eficaz na separação. Uma desvantagem deste método está na necessidade de se escolherem a priori três frequências que sejam discriminantes entre os sinais, o que não ocorre no caso das quatro variáveis propostas aqui (variáveis 10, 11, 12 e 15).

Variável 16 - Largura_Pulso_Sinal: Intervalo de tempo (ΔT) entre o pico do sinal normalizado e o máximo local onde este cai abaixo de 30 % do valor de pico.

Variável 17 - Vetor_Tempo_em_TxF: Variável proposta em [21], com o objetivo de estabelecer uma métrica T no espaço tempo-frequência equivalente ao tempo, calculada da seguinte forma:

$$T = \sqrt{\frac{\int_{-\infty}^{\infty} (t - t_0)^2 S(t)^2 dt}{\int_{-\infty}^{\infty} S(t)^2 dt}} \quad \text{Eq.3.5}$$

onde:

$$t_0 = \frac{\int_{-\infty}^{\infty} t S(t)^2 dt}{\int_{-\infty}^{\infty} S(t)^2 dt} \quad \text{Eq.3.6}$$

$t = \text{Vetor tempo do sinal (eixo } x)$

$$S(t) = \text{Vetor amplitude do sinal (eixo } y)$$

Variável 18 - Vetor_Frequencia_em_TxF: Variável proposta em [21], com o objetivo de estabelecer uma métrica F no espaço tempo-frequência equivalente à frequência, calculada da seguinte forma:

$$F = \sqrt{\frac{\int_{-\infty}^{\infty} f^2 |S(f)|^2 df}{\int_{-\infty}^{\infty} |S(f)|^2 df}} \quad \text{Eq.3.7}$$

onde:

$$S(f) = \text{Transformada de Fourier de } S(t)$$

É importante ressaltar que o espaço de separação TxF, correspondente às variáveis 17 e 18 plotadas uma contra a outra, vem se mostrando como um dos melhores espaços de separação de fontes de DP, como pode ser visto em [81]. Neste trabalho os autores utilizam o espaço TxF para separar sinais de DP em barras e bobinas novas de máquinas rotativas submetidas a testes de aceitação, obtendo sucesso na separação.

A eficácia do espaço de separação TxF também pode ser comprovada em [82], onde os autores criam defeitos artificiais em barras de máquinas rotativas e as submetem a ensaios em laboratório para captura dos sinais de DP. O espaço TxF é utilizado para separar as diferentes fontes dos sinais medidos nas barras.

Variável 19 - Vetor_TempoACF_em_T_ACF_x_F_ACF: Variável semelhante à variável 17, mas a métrica correspondente ao tempo é calculada por meio da ACF do sinal, ao invés de se utilizar o sinal diretamente.

Variável 20 - Vetor_FrequenciaACF_em_T_ACF_x_F_ACF: Variável semelhante à variável 18, mas a métrica correspondente à frequência é calculada utilizando-se a transformada de Fourier da ACF do sinal ao invés de se utilizar a transformada de Fourier do sinal diretamente.

Variável 21 - Área_Média_do_Sinal: Resultado da multiplicação do intervalo de tempo (ΔT), calculado entre o início do sinal normalizado e o máximo local

onde este cai abaixo de 30 % do valor de pico, e o valor deste pico do sinal em torno de 30 %, utilizado como referência.

Variável 22 - Área_Total_do_Sinal: Resultado da soma de todas as amostras do sinal normalizado multiplicadas pelo intervalo de amostragem, ou seja, a área total da curva normalizada do sinal.

Variável 23 - Área_Total_da_FFT_do_Sinal: Resultado da soma de todas as amostras do valor absoluto da FFT do sinal normalizado multiplicadas pelo intervalo de frequência equivalente entre uma amostra e outra, ou seja, a área total da curva do valor absoluto da FFT do sinal normalizado.

As 23 variáveis acima devem ser avaliadas individualmente e em grupo com relação à sua capacidade de extrair informações dos sinais no tempo. Para isso, o banco de dados gerado pelo ensaio das células e da barra de gerador em laboratório, que possuem classe determinada (cada sinal pode ser associado a uma das duas células ou à barra), será utilizado para avaliar a capacidade das variáveis na separação das fontes de DP. Em paralelo, os sinais dos TCs e PRs, que também possuem classe determinada (sabe-se a priori se os sinais são ou não de DP), também serão avaliados em relação ao poder (discriminante) das 23 variáveis de separar sinais com DP e sem DP para os TCs e para os PRs. A tabela 3.8 mostra um resumo das 23 variáveis detalhadas anteriormente, indicando quais já vem sendo utilizadas na literatura (6 das 23 variáveis em questão) e quais foram propostas neste estudo para análise de desempenho em sinais relacionados a DP (17 das 23 variáveis em questão). Quando estas variáveis forem mencionadas no texto, o “número da variável” apontado na tabela 3.8 será utilizado como referência.

Tabela 3.8 – Resumo das variáveis utilizadas no estudo

Nº da variável	Nome resumido	Proposta neste estudo?
1	Disp_ACF_1	Não
2	Disp_ACF_2	Não
3	Disp_ACF_3	Não
4	Disp_FFT_ACF_1	Sim
5	Disp_FFT_ACF_2	Sim
6	Disp_FFT_ACF_3	Sim
7	DeltaT_max1_max2	Sim
8	DeltaT_max1_min1	Sim
9	DeltaF_max1_max2	Sim
10	Disp_FFT_Sinal_1	Sim
11	Disp_FFT_Sinal_2	Sim
12	Disp_FFT_Sinal_3	Sim
13	Pico_F_ACF	Sim
14	Largura_pulso_ACF	Sim
15	Pico_F_Sinal	Sim
16	Largura_pulso_Sinal	Sim
17	Vetor_Tempo_em_TxF	Não
18	Vetor_Frequência_em_TxF	Não
19	Vetor_Tempo-ACF_em_T-ACF_x_F-ACF	Sim
20	Vetor_Frequência-ACF_em_T-ACF_x_F-ACF	Sim
21	Área_média_do_Sinal	Sim
22	Área_total_do_Sinal	Não
23	Área_total_da_FFT_do_Sinal	Sim

Em resumo, o que se deseja é selecionar um conjunto de variáveis relevantes e não redundantes em cada aplicação que seja capaz de auxiliar com eficácia a determinação da classe de saída dos sinais categorizados. Para isso, as variáveis serão ordenadas utilizando-se o algoritmo de seleção “Relief” [89], que é baseado em medidas de distância. Este algoritmo foi o escolhido para ser

utilizado neste estudo principalmente devido à sua capacidade de operar tanto com atributos nominais (incluindo lógicos) quanto numéricos (inteiros ou reais).

O objetivo do método Relief é avaliar a importância de cada atributo na determinação da classe de saída das amostras. Ou seja, ele mede a relevância dos atributos em função da classe de saída a partir de medidas de distância (quanto maior o valor do Relief, maior a relevância). Os parâmetros do algoritmo especificam o número de instâncias a serem amostradas, o número de vizinhos a serem verificados, a ponderação de vizinhos por distância e a função exponencial que determina a velocidade com que os pesos diminuem com a distância.

O Relief é utilizado aqui também para avaliar a redundância entre variáveis. O critério adotado para eliminar variáveis redundantes é o seguinte: se o índice de relevância (valor do Relief) entre duas variáveis for maior do que o índice de relevância entre a melhor variável (primeira de todas na ordenação) e a saída (classe), elimina-se a variável de menor relevância com a saída dentre as duas analisadas. O processo é feito de forma que as variáveis com menor relevância na determinação da saída sejam eliminadas primeiro. Desta forma, restarão somente as variáveis relevantes e não redundantes no banco de dados a ser processado. Maiores detalhes sobre o Relief são apresentados no Anexo A.

3.5

Agrupamento dos Sinais

No caso dos sinais destinados à separação de fontes de DP (ou seja, excluídos os sinais dos PRs e TCs, que passarão direto à etapa de classificação), após a ordenação das variáveis por relevância na determinação das classes de saída, as mais bem avaliadas poderão ser plotadas umas contra as outras em espaços 2D e 3D (ou até superiores, mas não observáveis), onde será possível visualizar os *clusters* das diferentes fontes. A atribuição de cada amostra ao *cluster* (grupo de pontos) correspondente é chamada de agrupamento, sendo mais bem definida como uma técnica exploratória de dados com o objetivo de formar agrupamentos de objetos semelhantes em bancos de dados. No agrupamento, nenhuma suposição é feita a respeito dos grupos (classes), sendo estes obtidos por meio da aplicação dos conceitos de similaridade e distância, onde a medida de distância adotada é a Euclidiana. Sejam $X = (x_1, x_2, \dots, x_n)$ e $Y = (y_1, y_2, \dots, y_n)$

dois pontos de \mathbf{R}^n . A distância Euclidiana entre X e Y é dada pela seguinte equação:

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2} \quad \text{Eq.3.8}$$

Um dos algoritmos de agrupamento mais utilizados é o *k-means*. Supondo que cada uma das “ n ” amostras de um banco de dados tenha associada a si um vetor “ x_j ” contendo os atributos, os “ n ” vetores “ x_j ” serão agrupados em “ k ” *clusters*. O algoritmo *k-means* minimiza a seguinte função objetivo:

$$J(v_q) = \sum_{i=1}^k \sum_{j=1}^n \|x_j - v_i\|^2 \quad \text{Eq.3.9}$$

onde “ v_q ” é o centróide do q -ésimo *cluster*.

Outro algoritmo de agrupamento consagrado é o *Fuzzy c-means* (FCM). Neste caso, uma amostra poderá pertencer a todas as classes com diferentes graus de pertinência, sendo que a soma dos graus de pertinência deve ser igual a 1. Supondo que cada uma das “ n ” amostras de um banco de dados tenha associada a si um vetor “ x_j ” contendo os atributos, os “ n ” vetores “ x_j ” serão agrupados em “ c ” *clusters*. O algoritmo FCM minimiza a seguinte função objetivo:

$$J(u_{ij}, v_k) = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m \|x_j - v_i\|^2 \quad \text{Eq.3.10}$$

onde:

- u_{ij} : valor entre 0 e 1 que indica o grau de pertinência de cada indivíduo j a um *cluster* i ;
- m : coeficiente fuzzy (neste estudo utilizou-se sempre $m = 1,2$);
- v_k : centróide do k -ésimo *cluster*.

O uso do FCM permite a exclusão de amostras com baixo grau de pertinência às classes selecionadas.

Na equação 3.10, o número de *clusters* é um parâmetro de entrada e deve ser fornecido para cálculo. Propõe-se aqui que o número ideal de *clusters* seja baseado numa "matriz de dissimilaridade relativa" dos centróides, determinada da seguinte forma:

1) Será fornecido um valor elevado NC_E (número de *clusters* de entrada) para o parâmetro de entrada do FCM correspondente ao número de *clusters* (número acima do número de *clusters* esperado para o conjunto);

2) A distância Euclidiana (Eq.3.8) entre cada centróide determinado na saída do algoritmo FCM será calculada. Quanto maior a distância entre um par de centróides, maior a dissimilaridade entre eles;

3) A "matriz de dissimilaridade relativa" dos centróides (proposta neste estudo) será calculada dividindo-se a distância Euclidiana de cada par de centróides pela maior delas (referência), obtendo-se assim valores entre 0 e 1;

4) Considerando-se um "limiar de junção" L_J pré-determinado, pares de centróides próximos entre si (ou seja, com distância relativa $\leq L_J$) serão unidos (começando pelos mais próximos entre si), com o novo centróide sendo posicionado no centro da reta que liga os 2 centróides originais;

5) Uma nova matriz de dissimilaridade será calculada, agora com um centróide a menos, e o processo retorna ao "item 4" anterior até que a distância equivalente entre todos os centróides restantes seja superior a L_J ;

6) Ao final do processo, espera-se obter um número de *clusters* ideal $NC_I \leq NC_E$ que se aproxime melhor do número de fontes de DP diferentes presente no mapa de DP em análise.

Outra técnica conhecida na literatura, chamada "dendrograma", também foi utilizada para verificar o número de *clusters* ideal para o banco de dados [90][91]. O dendrograma é um diagrama de árvore que exhibe os grupos formados por agrupamento das amostras em cada passo e em seus níveis de dissimilaridade. Ao longo do eixo vertical do dendrograma é exibido o nível de dissimilaridade das amostras nos pontos de junção e as diferentes observações são listadas ao longo do eixo horizontal. Maiores detalhes sobre dendrogramas são apresentados no Anexo B.

3.6

Classificação dos Sinais

3.6.1

Isolamentos com Classificação Baseada na Assinatura do Sinal no Tempo

A classificação dos sinais medidos nos PRs e TCs, tanto no campo quanto em laboratório, com base na análise da assinatura do sinal no tempo, será efetuada por meio de duas técnicas conhecidas: "árvore de decisão" e "SVM" (*Support Vector Machine*). O que se pretende é avaliar o desempenho de cada uma dessas técnicas quando aplicadas ao banco de dados em questão.

A árvore de decisão é determinada através do algoritmo "J48" [93], que leva em consideração a capacidade das variáveis de entrada na distinção dos sinais com e sem características de DP. Há uma árvore de decisão para classificação dos PRs e outra para classificação dos TCs. De forma resumida, a estrutura de uma árvore de decisão é caracterizada conforme mostrado na Figura 3.9.

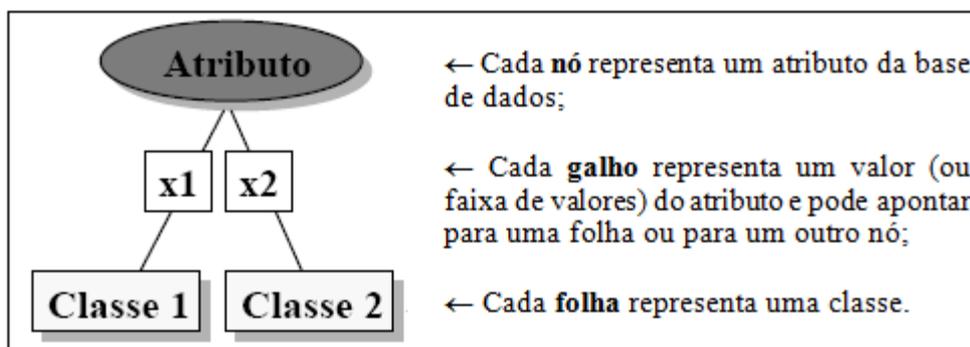


Figura 3.9 – Estrutura de uma árvore de decisão

A árvore vai sendo montada conforme a "taxa de ganho de informação" (*Gain Ratio*) das variáveis em relação à saída (classe). A raiz da árvore (primeiro atributo), por exemplo, corresponde à variável que fornece a maior taxa de ganho de informação com relação à classe de saída. De forma semelhante, os demais atributos da árvore são selecionados e nela posicionados, sempre com base na taxa de ganho de informação. Maiores detalhes sobre árvores de decisão são apresentados no Anexo C.

SVM é um algoritmo de aprendizado supervisionado, cujo objetivo é a separação de um conjunto de objetos de diferentes classes utilizando o conceito de planos de decisão [94]. Com princípios embasados na teoria de aprendizado estatístico, essa técnica de aprendizado de máquina se caracteriza por apresentar uma boa capacidade de generalização. A SVM também é robusta diante de dados de grande dimensão, sobre os quais outras técnicas de aprendizado comumente obtêm classificadores super ou sub ajustados. Outra característica atrativa é a convexidade do problema de otimização formulado em seu treinamento, que implica na existência de um único mínimo global. Essa é uma vantagem da SVM sobre, por exemplo, as Redes Neurais Artificiais (RNAs) Perceptron Multicamadas (*Multilayer Perceptron*) [64], em que há mínimos locais na função objetivo minimizada. Além disso, o uso de funções *Kernel* na não-linearização da SVM torna o algoritmo eficiente, pois permite a construção de simples hiperplanos em um espaço de alta dimensão de forma tratável do ponto de vista computacional [95][96]. Maiores detalhes sobre SVM são apresentados no Anexo D.

3.6.2

Isolamentos com Classificação Baseada na Assinatura do Mapa de Descargas Parciais

Após a separação das diferentes fontes de DP (agrupamento) nos sinais analisados neste estudo (caso das células PPA, EFA e da barra de gerador ensaiados em laboratório, citados na tabela 3.1, e do sinal de DP em meio a ruído no gerador 1, citado na tabela 3.5), é necessário realizar a classificação dos sinais (plotados nos mapas de DP para serem referenciados ao ângulo da senóide) com relação ao tipo de padrão de defeito que representam ou se são simplesmente ruído de medição.

No caso dos dados aqui analisados, que se mostram já na forma de mapa de DP (caso dos sinais do gerador 2, citado na tabela 3.6, e dos sinais dos geradores 3 a 6 e do motor 1, citados na tabela 3.7), a etapa de agrupamento não é necessária, passando-se direto à etapa de classificação.

A metodologia de classificação dos sinais mencionados é detalhada no tópico 3.6.2.1 apresentado a seguir.

3.6.2.1

Metodologia de Classificação

Sabe-se que a distribuição dos sinais no mapa de DP com relação à fase, onde ocorre a descarga, revela características do tipo de defeito em questão. Já a distribuição dos sinais no mapa de DP com relação à amplitude da descarga, revela características da severidade do defeito em questão. Desta forma, os histogramas de fase $H_n(\text{fase})$, $H_{q\max}(\text{fase})$ e $H_{q\text{med}}(\text{fase})$, tanto dos sinais positivos quanto dos negativos, afiguram-se como uma boa alternativa para apontar a classe de cada mapa de DP. Assim, os seis histogramas serão utilizados aqui para classificação dos mapas. São eles:

$H_n(\text{fase})+$: distribuição do número de sinais positivos em relação à fase;

$H_n(\text{fase})-$: distribuição do número de sinais negativos em relação à fase;

$H_{q\max}(\text{fase})+$: distribuição da carga máxima positiva em relação à fase;

$H_{q\max}(\text{fase})-$: distribuição da carga máxima negativa em relação à fase;

$H_{q\text{med}}(\text{fase})+$: distribuição da carga média positiva em relação à fase;

$H_{q\text{med}}(\text{fase})-$: distribuição da carga média negativa em relação à fase.

A metodologia proposta para a classificação é a seguinte:

- 1) Cálculo dos seis histogramas para cada mapa de DP do banco de dados: $H_n(\text{fase})+$, $H_n(\text{fase})-$, $H_{q\max}(\text{fase})+$, $H_{q\max}(\text{fase})-$, $H_{q\text{med}}(\text{fase})+$ e $H_{q\text{med}}(\text{fase})-$;
- 2) Normalização dos histogramas dividindo-se cada um por seu somatório;
- 3) Determinação do coeficiente de correlação entre os histogramas da amostra que se quer classificar e os histogramas de todos os N sinais presentes no banco de dados histórico (será detalhado no tópico 3.6.2.2. na sequência). Obtém-se assim uma matriz de coeficientes $6 \times N$, onde cada linha representa um tipo dos seis histogramas e cada coluna representa uma amostra padrão já classificada presente no banco de dados histórico;
- 4) A amostra teste será atribuída à mesma classe do padrão (do banco de dados histórico) com a qual ela mais se assemelhar, tomando-se a média da correlação dos seis histogramas como medida de semelhança (espaço de comparação em $1 \times N$).

Será utilizado aqui o "Coeficiente de Correlação de Pearson (r)", que mede o grau da correlação linear entre duas variáveis quantitativas. Ele é um índice adimensional com valores situados entre -1 e 1, que reflete a intensidade de uma relação linear entre dois conjuntos de dados (2 vetores, neste caso). Se $r = 1$, tem-se uma correlação perfeita positiva entre as duas variáveis. Para $r = -1$, a correlação é perfeita negativa entre as duas variáveis, ou seja, se uma aumenta, a outra sempre diminui. Já $r = 0$ significa que as duas variáveis não dependem linearmente uma da outra. Sejam $X = (x_1, x_2, \dots, x_n)$ e $Y = (y_1, y_2, \dots, y_n)$ os dois vetores para os quais se deseja avaliar a semelhança. O Coeficiente de Correlação de Pearson é dado por:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad \text{Eq.3.11}$$

onde:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad e \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \quad \text{Eq.3.12}$$

3.6.2.2

Banco de Dados Histórico

Propõe-se neste trabalho a criação de um banco de dados histórico para servir como base de comparação, conforme a metodologia proposta no tópico 3.6.2.1 anterior. Contudo, é importante salientar que o banco de dados histórico deve estar voltado para o tipo específico de equipamento. Como os sinais utilizados neste estudo são oriundos de geradores e de um motor, serão utilizados padrões de defeitos de máquinas rotativas apresentados na norma IEC 60034-27-02 [97]. Esses padrões serão capturados em forma de imagem, processados e então digitalizados em forma de matriz, de modo que a cor de cada pixel corresponda ao número de sinais de DP ocorridos com aquela amplitude e naquela fase correspondente, preenchendo cada loco da matriz com o número de sinais correspondente ao endereço. Desta forma, tem-se uma reprodução fiel da assinatura do mapa de DP apresentado na norma.

Outra forma de se obter os mapas de DP padrão seria ensaiando objetos de teste com os mesmos defeitos em laboratório, o que seria tarefa árdua dada a dificuldade de encontrar objetos para retratar os diferentes defeitos.

Nas Figuras 3.10 a 3.14, são apresentados os mapas de DP (com padrão de cores variando de cinza a amarelo intenso, utilizado por um sistema comercial de medição de DP) dos cinco diferentes tipos de defeito retirados da norma IEC 60034-27-02 [97] e os mapas de DP (com padrões de cores variando de azul a vermelho intenso, utilizado pelo sistema de medição IMA-DP, desenvolvido pelo Cepel) após o processamento das imagens originais e digitalização em forma de matriz. Na Tabela 3.9, são detalhados os cinco tipos de defeito retirados da norma e os respectivos graus de risco de falha associados a cada um [97].

Tabela 3.9 – Resumo dos defeitos encontrados em máquinas rotativas

Tipo de defeito	Grau de risco	Descrição resumida
DP1: Gap	Normal	DP que ocorrem entre duas barras na saliência do enrolamento ou entre uma barra e a prensa do núcleo do estator.
DP2: Ranhura	Alto	DP geradas devido a bolsas de ar ou gás dentro do núcleo do estator, ocorrendo entre a superfície de uma barra e o núcleo do estator.
DP3: Cavidades internas	Baixo	DP geradas em bolhas de ar ou gás embutidas no isolamento principal.
DP4: Delaminação interna	Alto	DP geradas em bolsas de ar ou gás alongadas longitudinalmente embutidas no isolamento principal.
DP5: Tracking de superfície	Normal	DP que ocorrem ao longo do braço de projeção do enrolamento devido à contaminação na interface ar/isolamento.

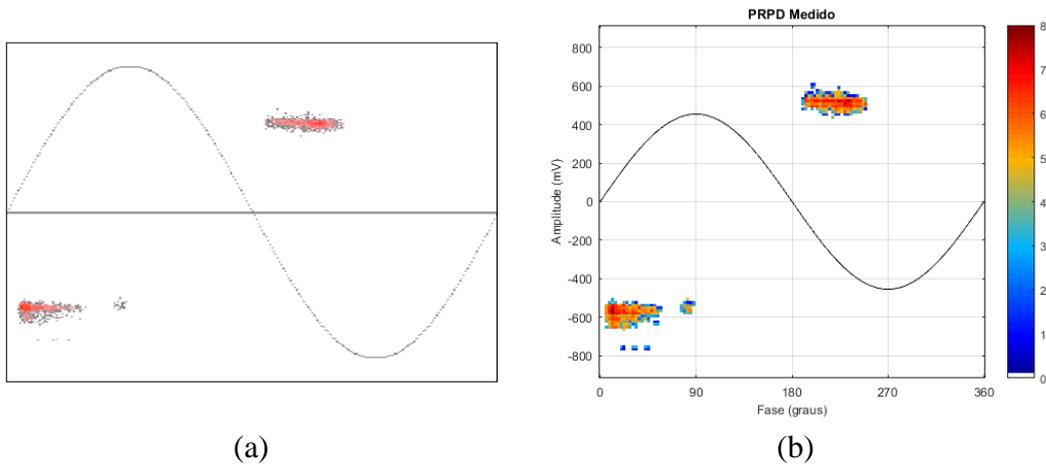


Figura 3.10 – Padrão DP1 (Gap) (a) Imagem retirada da norma e
(b) Mapa de DP após processamento

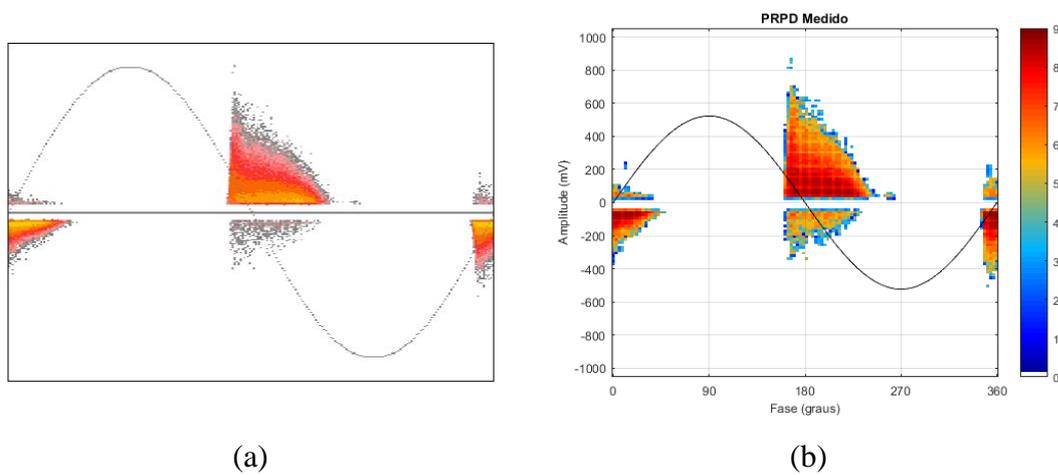


Figura 3.11 – Padrão DP2 (Ranhura) (a) Imagem retirada da norma e
(b) Mapa de DP após processamento

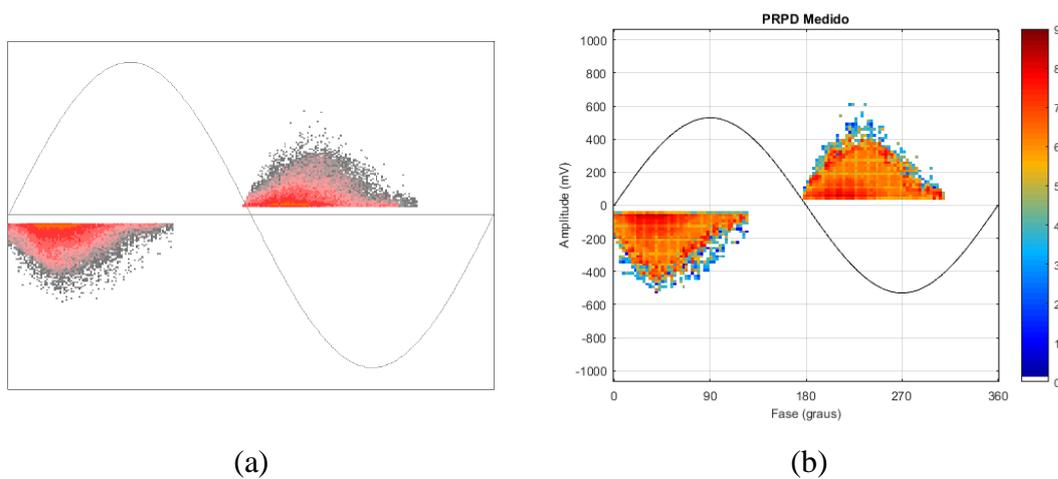


Figura 3.12 – Padrão DP3 (Cavidades internas) (a) Imagem retirada da norma e
(b) Mapa de DP após processamento

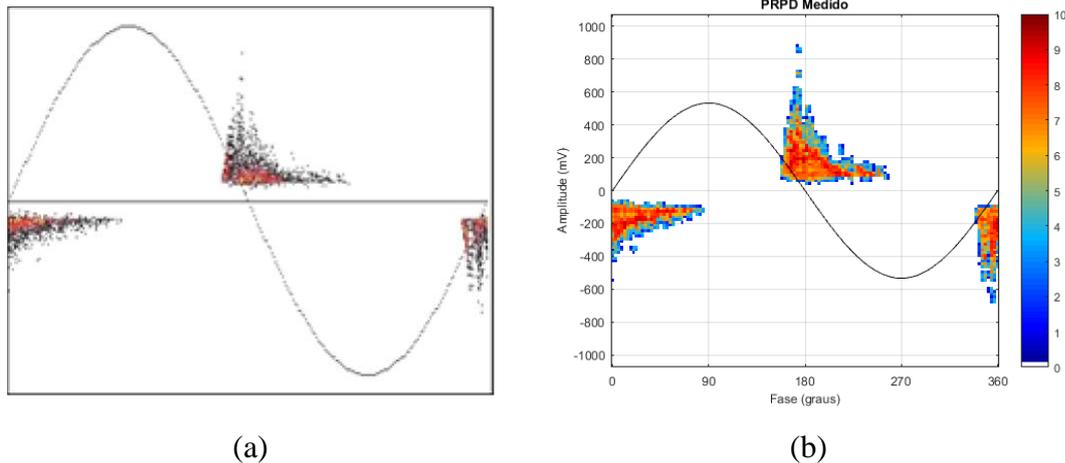


Figura 3.13 – Padrão DP4 (Delaminação interna) (a) Imagem retirada da norma e (b) Mapa de DP após processamento

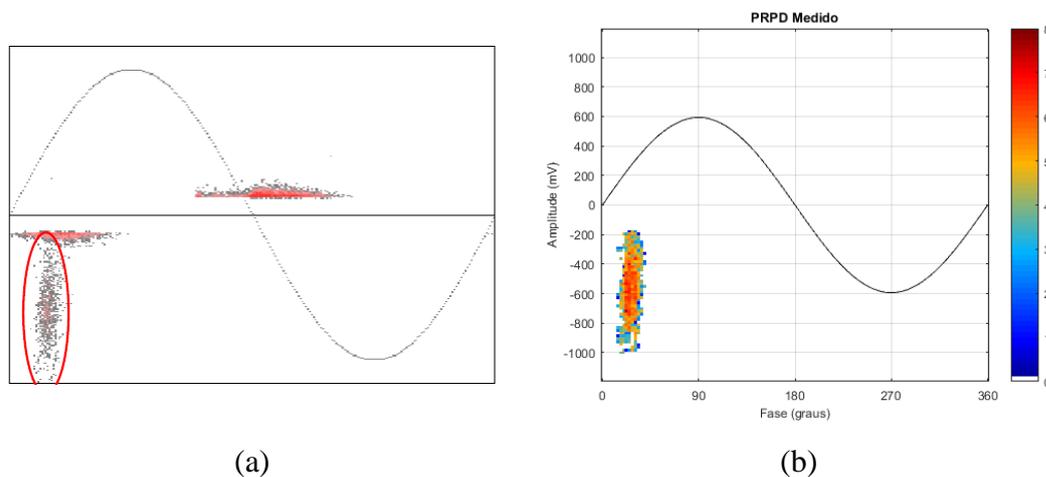


Figura 3.14 – Padrão DP5 (Tracking de superfície) (a) Imagem retirada da norma e (b) Mapa de DP após processamento

No caso deste estudo em particular, o banco histórico será criado também com os 12 sinais obtidos em laboratório das duas células (8 sinais) e da barra de gerador (4 sinais), que posteriormente serão tirados do banco para uso futuro em sinais de geradores e motores apenas. Além disso, também será incluído no banco um sinal de ruído, normalmente encontrado em campo, para identificação deste em meio aos sinais com informação útil. Os mapas de DP representando estes padrões (um de cada célula, um da barra de gerador e o único de ruído) são mostrados nas Figuras 3.15 a 3.18.

Nas Figuras 3.19 a 3.23, são mostradas imagens de alguns dos defeitos citados em barras de geradores para ilustração [97][13].

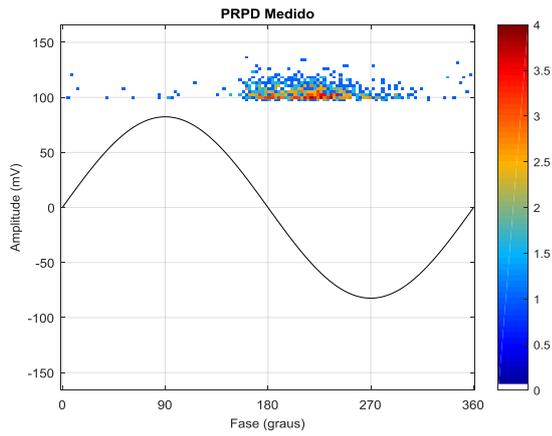


Figura 3.15 – Mapa padrão célula PPA

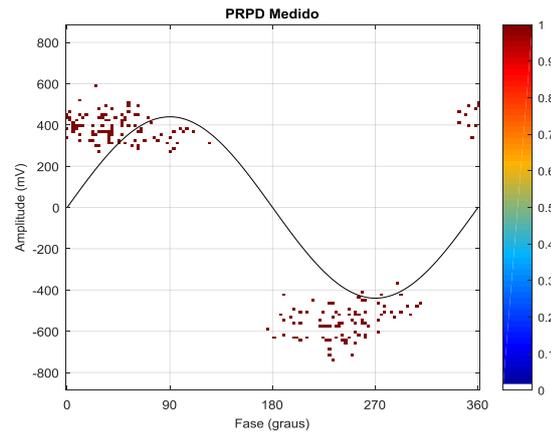


Figura 3.16 – Mapa padrão célula EFA

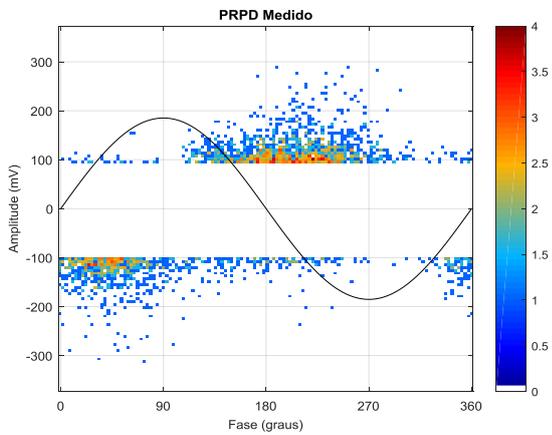


Figura 3.17 – Mapa padrão
barra de gerador

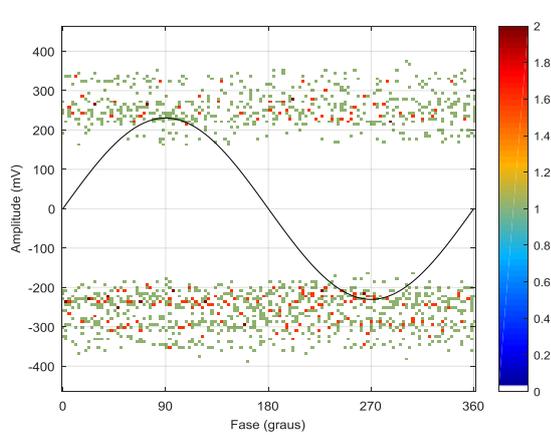


Figura 3.18 – Mapa padrão ruído



Figura 3.19 - DP de Gap entre
2 barras

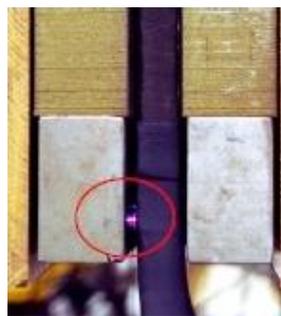


Figura 3.20 - DP de Gap entre a
barra e a prensa do núcleo do estator

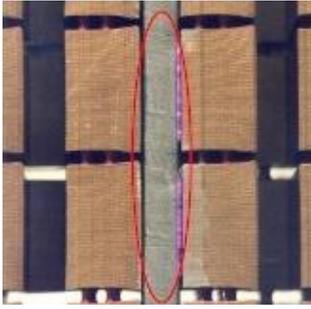


Figura 3.21 - DP de ranhura entre a barra e o núcleo do estator

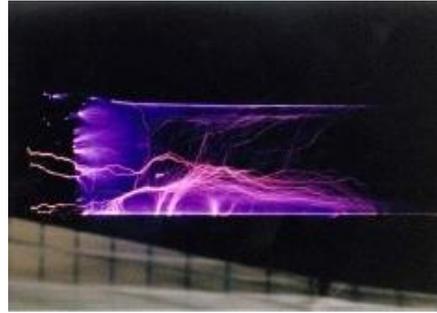


Figura 3.22 - DP por tracking de superfície no braço de projeção do enrolamento

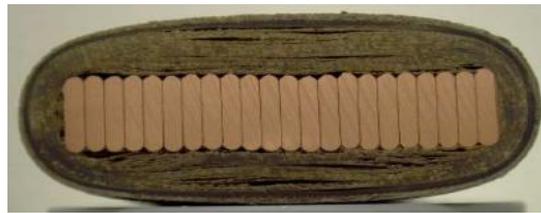


Figura 3.23 - Corte transversal em uma barra com delaminação

4

Resultados e Discussão

4.1

Utilizando os Sinais dos Para-raios e Transformadores de Corrente

A classificação dos sinais medidos nos PRs e TCs em campo é feita aqui utilizando-se os métodos de “árvore de decisão” e “SVM”, conforme mencionado no item 3.6.1, que leva em consideração as variáveis mais bem avaliadas (relevantes e não redundantes) para determinar se os sinais têm ou não características de DP. Nas tabelas 4.1 e 4.2 são apresentados os resultados da aplicação do Relief nos sinais dos PRs e TCs, respectivamente, para as 23 variáveis em análise, apresentadas já na ordem de importância para o tipo de equipamento (PR ou TC) e com as variáveis classificadas como redundantes (conforme critério detalhado no item 3.4) marcadas em vermelho.

Com o critério adotado para eliminação de variáveis redundantes da análise, restaram 19 variáveis para serem utilizadas nos PRs, e 22 para os TCs. Desta forma, pôde-se testar a classificação dos sinais por meio da árvore de decisão e do SVM variando-se o número de variáveis de entrada (na ordem de importância apontada pelo Relief) de dois até o máximo de variáveis selecionado em cada equipamento. Vale ressaltar que o critério para treino e teste adotado aqui foi o “*leave one out*”, no qual uma das N amostras do banco é separada para teste e as outras $N-1$ são utilizadas para treino. Repete-se o processo de forma que cada uma das N amostras tenha sido utilizada uma única vez como amostra de teste. Os resultados são apresentados nas tabelas 4.3 (árvore/PRs), 4.4 (SVM/PRs), 4.5 (árvore/TCs) e 4.6 (SVM/TCs) para os PRs e TCs utilizando as duas técnicas de classificação. É importante registrar que, nos resultados apresentados no SVM, a função *kernel* do tipo polinomial (de grau 3) foi a escolhida, visto que apresentou melhor resultado que as demais testadas (linear, quadrática, RBF – *Radial Basis Function* – e MLP – *Multilayer Perceptron*).

Tabela 4.1 - Resultado do Relief nos
PRs em relação às 23 variáveis

Variáveis	Relief
22	0,118
16	0,095
3	0,093
23	0,081
2	0,068
17	0,063
14	0,059
18	0,058
19	0,054
21	0,050
1	0,045
12	0,029
20	0,028
8	0,027
4	0,025
5	0,022
15	0,021
13	0,020
6	0,017
7	0,016
11	0,013
10	0,010
9	0,008

Tabela 4.2 - Resultado do Relief nos
TCs em relação às 23 variáveis

Variáveis	Relief
16	0,168
22	0,164
6	0,101
21	0,087
23	0,084
3	0,074
2	0,070
18	0,070
20	0,065
13	0,065
4	0,058
1	0,056
11	0,054
7	0,052
17	0,046
14	0,043
5	0,041
19	0,039
8	0,032
10	0,030
12	0,024
15	0,015
9	-0,002

Tabela 4.3 - Taxa de acerto na classificação dos PRs utilizando Árvore de Decisão

Nº de Variáveis de entrada	Taxa de acerto (%)		
	nos “Sem DP”	nos “Com DP”	Total
2	75,2	82,7	79,2
3	81,8	82,0	81,9
4	85,1	83,5	84,2
5	84,3	89,9	87,3
6	85,1	86,3	85,8
7	88,4	87,8	88,1
8	88,4	88,5	88,5
9	89,3	87,8	88,5
10	90,1	87,8	88,8
11	90,9	86,3	88,5
12	90,9	85,6	88,1
13	90,9	87,1	88,8
14	90,9	89,2	90,0
15	90,1	89,2	89,6
16	89,3	89,2	89,2
17	89,3	91,4	90,4
18	89,3	91,4	90,4
19	89,3	91,4	90,4

Tabela 4.4 - Taxa de acerto na classificação dos PRs utilizando SVM

Nº de Variáveis de entrada	Taxa de acerto (%)		
	nos “Sem DP”	nos “Com DP”	Total
2	86,8	78,4	82,3
3	81,0	88,5	85,0
4	81,0	87,8	84,6
5	83,5	86,3	85,0
6	85,1	90,6	88,1
7	83,5	91,4	87,7
8	87,6	94,2	91,2
9	90,1	94,2	92,3
10	87,6	94,2	91,2
11	89,3	94,2	91,9
12	90,9	92,8	91,9
13	93,4	92,8	93,1
14	91,7	92,1	91,9
15	89,3	93,5	91,5
16	91,7	92,1	91,9
17	92,6	92,1	92,3
18	91,7	92,1	91,9
19	93,4	92,1	92,7

Tabela 4.5 - Taxa de acerto na classificação dos TCs utilizando Árvore de Decisão

Nº de Variáveis de entrada	Taxa de acerto (%)		
	nos “Sem DP”	nos “Com DP”	Total
2	89,3	85,0	87,4
3	88,0	85,0	86,7
4	89,3	85,0	87,4
5	88,0	88,3	88,1
6	93,3	90,0	91,9
7	89,3	90,0	89,6
8	98,7	93,3	96,3
9	98,7	90,0	94,8
10	98,7	90,0	94,8
11	97,3	91,7	94,8
12	97,3	91,7	94,8
13	97,3	91,7	94,8
14	97,3	90,0	94,1
15	96,0	91,7	94,1
16	96,0	91,7	94,1
17	93,3	90,0	91,9
18	92,0	90,0	91,1
19	92,0	90,0	91,1
20	92,0	90,0	91,1
21	92,0	90,0	91,1
22	92,0	90,0	91,1

Tabela 4.6 - Taxa de acerto na classificação dos TCs utilizando SVM

Nº de Variáveis de entrada	Taxa de acerto (%)		
	nos “Sem DP”	nos “Com DP”	Total
2	84,0	83,3	83,7
3	90,7	88,3	89,6
4	92,0	86,7	89,6
5	92,0	83,3	88,1
6	98,7	85,0	92,6
7	96,0	81,7	89,6
8	93,3	86,7	90,4
9	93,3	86,7	90,4
10	93,3	91,7	92,6
11	96,0	90,0	93,3
12	97,3	91,7	94,8
13	97,3	91,7	94,8
14	97,3	90,0	94,1
15	96,0	91,7	94,1
16	100,0	91,7	96,3
17	98,7	91,7	95,6
18	96,0	90,0	93,3
19	98,7	88,3	94,1
20	96,0	91,7	94,1
21	96,0	90,0	93,3
22	98,7	90,0	94,8

Um resumo dos melhores resultados obtidos na classificação dos PRs e TCs é mostrado na tabela 4.7. Como pode ser observado, o melhor resultado nos PRs foi obtido para 13 variáveis de entrada e SVM, com 93,1 % de acerto total (a matriz de confusão desse resultado é mostrada na tabela 4.8). Já no caso dos TCs, o melhor resultado foi obtido para 8 variáveis de entrada e árvore de decisão (requer menos variáveis que o SVM para o mesmo percentual de acerto), com 96,3 % de acerto total (a matriz de confusão desse resultado é mostrada na tabela 4.9).

Tabela 4.7 – Melhores percentuais de acerto atingidos nos PRs e TCs

Tipo de equipamento	Método de classificação	Número de variáveis	Taxa de acerto na classificação (%)		
			nos "sem DP"	nos "com DP"	Total
PR	Árvore	17	89,3	91,4	90,4
	SVM	13	93,4	92,8	93,1
TC	Árvore	8	98,7	93,3	96,3
	SVM	16	100	91,7	96,3

Tabela 4.8 – Matriz de confusão da melhor classificação nos PRs (SVM/13 variáveis)

		Nº de sinais classificados como "Sem DP"	Nº de sinais classificados como "Com DP"	Taxa de acerto
Nº total de sinais "Sem DP"	121	113	8 (falsos positivos)	93,4 %
Nº total de sinais "Com DP"	139	10 (falsos negativos)	129	92,8 %
Nº total de sinais	260	123	137	93,1 %

Tabela 4.9 – Matriz de confusão da melhor classificação nos TCs (árvore de decisão/8 variáveis)

		Nº de sinais classificados como "Sem DP"	Nº de sinais classificados como "Com DP"	Taxa de acerto
Nº total de sinais "Sem DP"	75	74	1 (falso positivo)	98,7 %
Nº total de sinais "Com DP"	60	4 (falsos negativos)	56	93,3 %
Nº total de sinais	135	78	57	96,3 %

Para ilustrar, são apresentados na Figura 4.1 um sinal medido em um PR com DP e um sinal medido em um PR sem DP. Na Figura 4.2, faz-se o mesmo

para os TCs, com um sinal medido em um TC com DP e outro em um TC sem DP.

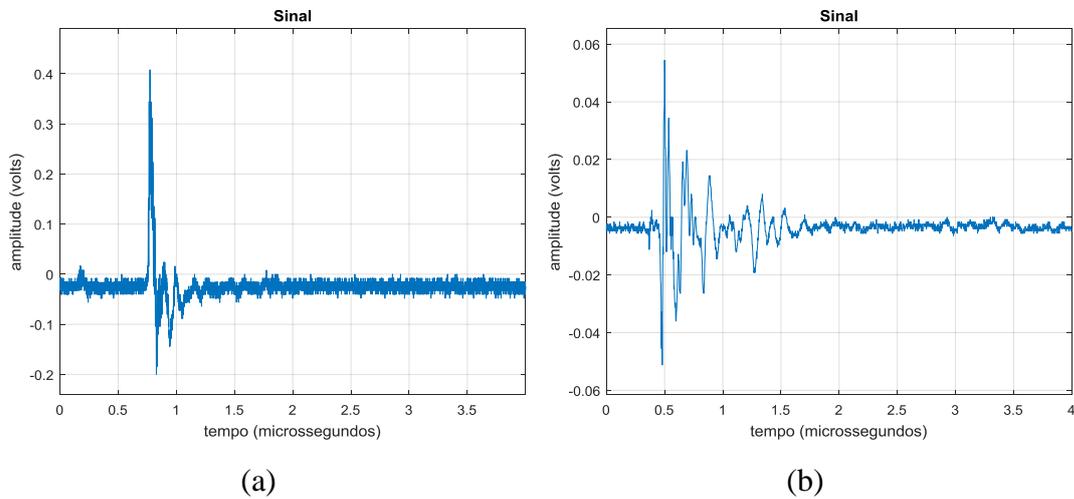


Figura 4.1 – Sinais medidos em PRs com e sem DP,
sendo: (a) com DP; e (b) sem DP

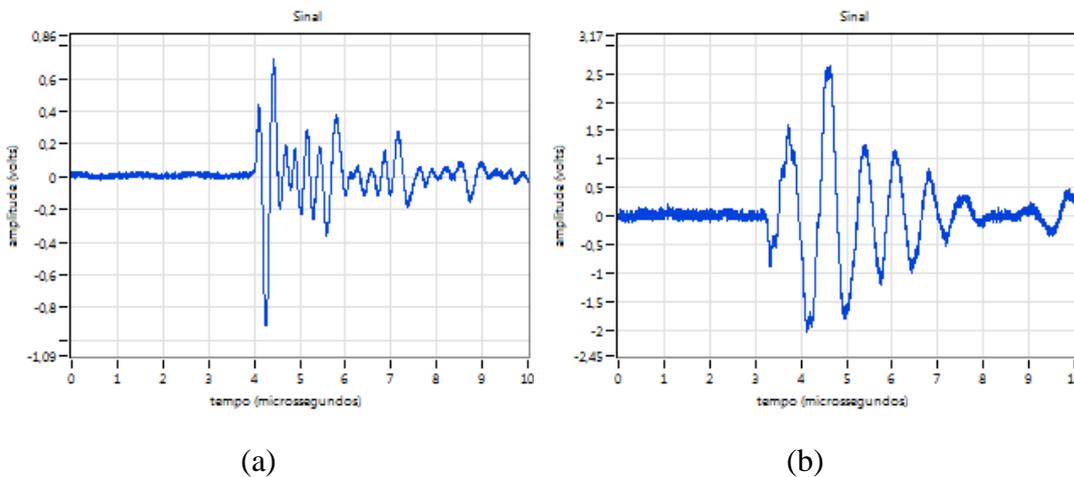


Figura 4.2 – Sinais medidos em TCs com e sem DP,
sendo: (a) com DP; e (b) sem DP

Analisando-se as melhores configurações de classificação e percentuais de acerto obtidos, percebe-se que a classificação automática de PRs e TCs é possível mediante o uso das metodologias propostas, aliado a uma técnica de medição nada invasiva desenvolvida pelo Cepel (o equipamento não necessita ser desligado). Através de inspeções regulares em períodos definidos de tempo, a ferramenta

desenvolvida mostra-se valiosa para a verificação da condição operativa destes equipamentos no campo. Pode-se destacar as seguintes vantagens da técnica:

- 1) Rapidez na execução dos ensaios e na avaliação dos equipamentos em operação, permitindo assim a sua retirada de serviço antes de uma eventual falha;
- 2) Realização do ensaio com o equipamento energizado e em funcionamento normal, sem a necessidade de intervenções nos equipamentos e nem no sistema elétrico de potência;
- 3) Capacidade de prevenir falhas dos equipamentos, evitando assim gastos desnecessários e avarias em equipamentos vizinhos ou protegidos por estes;
- 4) Menor dependência do olho clínico do especialista na determinação dos equipamentos com DP, por meio da classificação automática dos sinais;
- 5) Proporcionar maior segurança para as equipes de manutenção.

A partir dos resultados obtidos no estudo dos sinais medidos em campo nos PRs e TCs, foram identificadas as variáveis mais efetivas para separar sinais de equipamentos com DP e sem DP. Foi possível, então, desenvolver um classificador de sinais para ser utilizado nestes equipamentos, seja no campo ou em laboratório, mas utilizando-se a mesma metodologia de captura dos sinais através do TCAF acoplado ao cabo de terra do equipamento [98]. A interface deste classificador de sinais de PRs e TCs é apresentada na Figura 4.3.

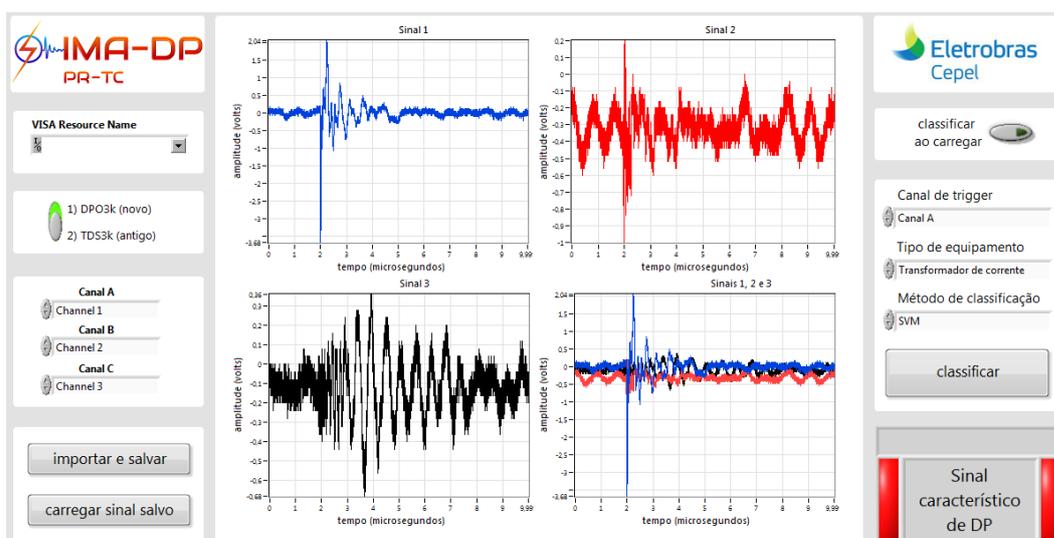


Figura 4.3 – Interface do classificador de sinais de PRs e TCs desenvolvido

Quando há suspeita de que o sinal medido em um PR ou TC no campo é característico de DP interna ao equipamento, é aconselhada a retirada de tal equipamento de operação, assim que possível, e a sua avaliação em laboratório (ambiente controlado e sem ruído externo elevado). Seguindo esta sistemática, dois equipamentos – um PR de 500 kV e um TC de 500 kV – foram considerados suspeitos (baseando-se somente no olho clínico do especialista na avaliação da assinatura do sinal à época), retirados de operação e levados para avaliação em laboratório. Feita a avaliação, as suspeitas de DP internas aos equipamentos foram confirmadas pela verificação do sinal no tempo medido pelo TCAF e analisado pelo “classificador automático” aqui desenvolvido, que indicou "sinal característico de DP" na maioria dos sinais, corroborando também com a classificação do especialista no laboratório, para os dois equipamentos. A tabela 4.10 mostra os percentuais de acerto atingidos nos 12 sinais do PR e nos 21 sinais do TC (número de sinais estes apontados na tabela 3.4) ensaiados em laboratório e testados no “classificador automático” desenvolvido.

Tabela 4.10 – Percentual de acerto do classificador automático nos equipamentos retirados de campo para ensaio no laboratório de AT

Equipamento	Método de classificação	Taxa de acerto (%)
PR	Árvore de decisão (com 17 variáveis)	91,7
	SVM (com 13 variáveis)	91,7
TC	Árvore de decisão (com 8 variáveis)	100
	SVM (com 16 variáveis)	95,2

Os sinais medidos em campo foram posteriormente (não no momento da captura dos sinais em campo, quando não havia ainda o classificador) submetidos ao classificador, que acusou "sinais característicos de DP" nos dois equipamentos, corroborando a classificação do especialista à época. Um exemplo de gráficos dos sinais no tempo medidos tanto no campo quanto em laboratório, para o PR e também para o TC, são apresentados nas Figuras 4.4 e 4.5, respectivamente.

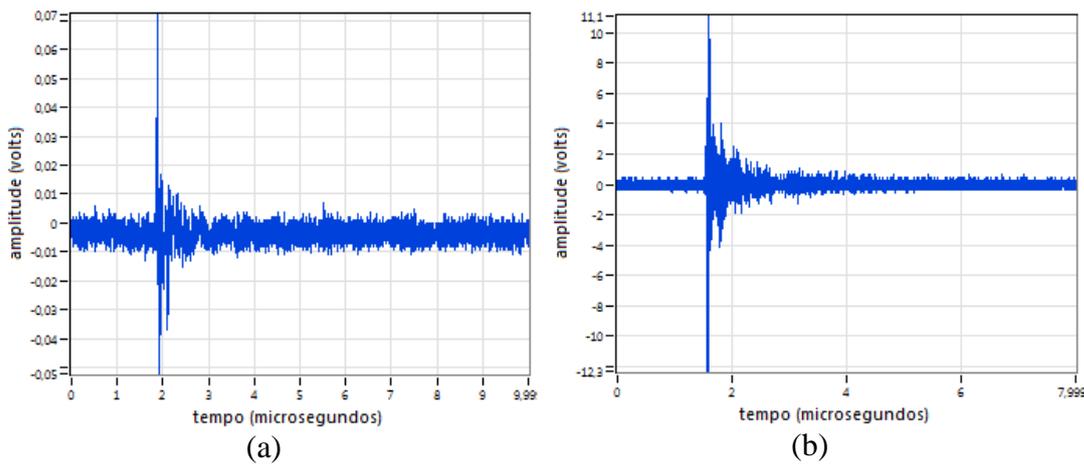


Figura 4.4 – Sinais medidos no PR identificado com DP, sendo: (a) Sinal medido no campo; e (b) sinal medido em laboratório

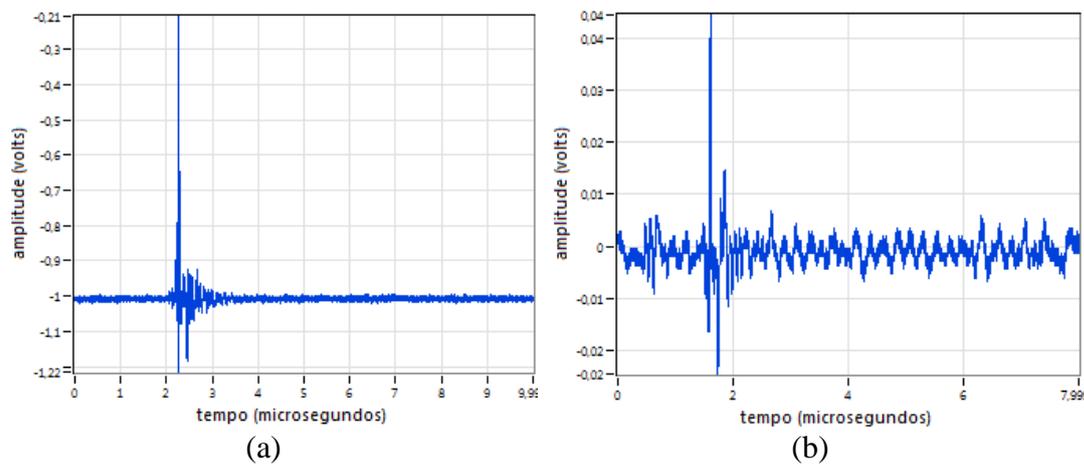


Figura 4.5 – Sinais medidos no TC identificado com DP, sendo: (a) Sinal medido no campo; e (b) sinal medido em laboratório

4.2

Utilização dos Sinais das Células e da Barra de Gerador

4.2.1

Seleção de Características e Agrupamento

Na tabela 4.11 é apresentado o resultado da aplicação do Relief nos sinais das células e da barra de gerador ensaiados em laboratório, para as 23 variáveis e já com a respectiva ordem de importância destas na determinação da classe de saída (que neste caso é apontar se o sinal no tempo pertence à célula PPA, à célula

EFA ou à barra de gerador). O critério de análise de redundância entre variáveis adotado aqui (conforme detalhado no capítulo 3), não apontou nenhuma variável redundante neste banco de dados. Conforme pode ser observado na tabela 4.11, das três primeiras variáveis mais bem avaliadas, duas foram propostas aqui neste estudo (4 e 5), o que demonstra o poder de discriminação entre os sinais de cada fonte segundo a informação extraída por estas variáveis dos sinais. Na Figura 4.6, é mostrado o espaço de separação 3D formado pela combinação das três melhores variáveis selecionadas (2, 4 e 5).

O objetivo principal na separação de fontes de DP é, dado um conjunto de sinais no tempo não categorizados (que dá origem ao mapa de DP quando associado à fase da senóide), utilizar variáveis relevantes e não redundantes que evidenciem as características destes sinais para que se possa fazer o agrupamento destes no espaço de separação formado por estas variáveis [99].

A ideia inicial aqui é avaliar a metodologia proposta (detalhada no capítulo 3) para a determinação do número ideal de *clusters* (NC_I) a ser utilizado no FCM. Para tal foram testados, para o conjunto com as 23 variáveis no espaço de separação, valores de Limiar de Junção (L_J) variando de 0 a 1 (com intervalo de 0,025). O resultado das simulações é apresentado na tabela 4.12, que mostra o NC_I obtido em cada configuração, com os três objetos (células PPA, EFA e barra de gerador) tomados 2 a 2 (3 combinações possíveis) e 3 a 3 (1 combinação possível). Vale ressaltar que o número de *clusters* de entrada (NC_E) definido, a ser minimizado, foi seis e que neste ponto do estudo foram utilizados os sinais de cada objeto ensaiado separadamente (não em paralelo) e unidos depois nas diferentes configurações testadas, de forma que a classe de cada sinal fosse conhecida a priori.

Pode-se observar pela tabela 4.12 que um valor bom para L_J deve estar entre 0,400 e 0,625. Nesta faixa, obtém-se sempre um NC_I igual ao número real de classes de saída do banco de sinais das células e da barra de gerador com as 23 variáveis, seja na configuração com 2 ou com 3 classes de saída (que neste caso é o tipo de fonte de DP: célula PPA, célula EFA ou barra de gerador).

Tabela 4.11 – Resultado do Relief para as 23 variáveis
nas células e na barra de gerador

Variáveis	Relief
2	0,331
4	0,263
5	0,255
1	0,183
16	0,169
3	0,161
17	0,150
23	0,143
18	0,134
10	0,133
12	0,116
11	0,109
14	0,108
22	0,099
19	0,096
20	0,093
13	0,079
15	0,078
21	0,074
7	0,058
8	0,053
9	0,042
6	0,040

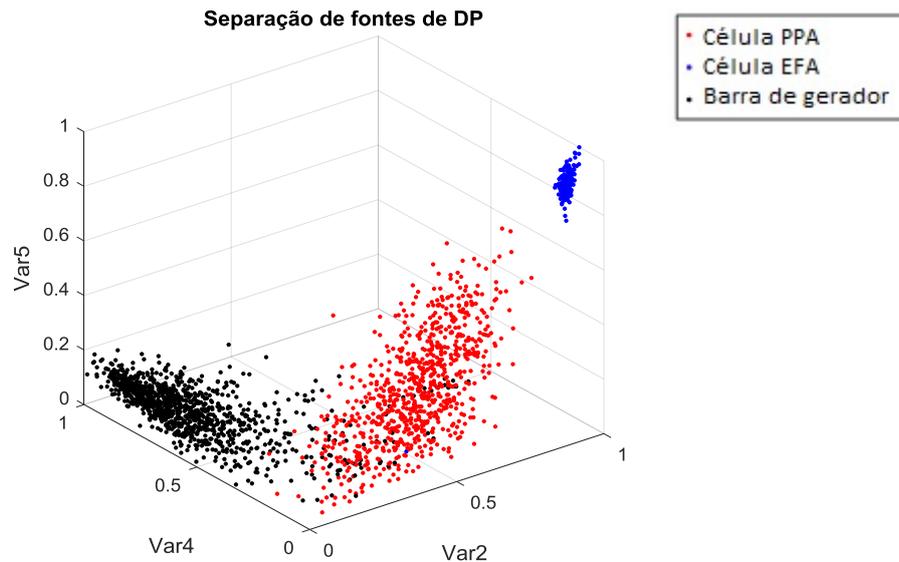


Figura 4.6 – Sinais das células plotados em relação às suas classes originais utilizando as variáveis 2, 4 e 5

Fixando-se agora o limiar de junção $L_J=0,500$ (que corresponde ao centro da faixa ideal) e utilizando-se as 4 diferentes configurações, seja com 2 ou 3 fontes de DP (PPA+EFA, PPA+barra de gerador, EFA+barra de gerador, e PPA+EFA+barra de gerador), foram feitas simulações com diferentes números de variáveis de entrada (na ordem de importância do Relief) para se determinar o NC_I pelo método proposto (descrito no capítulo 3). Em todos os casos, obteve-se o NC_I correto, seja com 2 ou 3 fontes de DP diferentes. A mesma simulação foi realizada utilizando-se o método dos dendrogramas (descrito no capítulo 3), também variando o número de variáveis de entrada. Neste caso, foi necessário determinar um ponto de corte no nível de similaridade normalizado (PCN) para a determinação do número ideal de *clusters*. Para analisar o melhor PCN, os dendrogramas obtidos com as 23 variáveis, mostrados nas figuras 4.7 a 4.10, foram utilizados como base. Nestes casos, pode-se perceber que dos níveis 0,40 a 0,91, nas 4 configurações diferentes (com 2 ou 3 classes), obtém-se o número correto de *clusters* do banco. Então, baseando-se na maior largura de faixa do dendrograma formado na simulação com 23 variáveis, pode-se utilizar $PCN = 0,65$ (centro da faixa ideal) como referência. Considerando-se este valor de PCN (0,65) e realizando simulações para diferentes números de variáveis de entrada, obteve-se sempre o valor correto de NC_I , seja ele 2 ou 3.

Tabela 4.12 - Influência do limiar de junção na determinação do número de *clusters* ideal utilizando as 23 variáveis

Limiar de Junção	Classes dos sinais a separar			
	PPA e EFA	PPA e Barra	EFA e Barra	PPA, EFA e Barra
0,000	6	6	6	6
0,025	3	4	6	4
0,050	3	4	6	4
0,075	3	4	6	4
0,100	3	4	4	4
0,125	3	4	4	4
0,150	3	4	4	4
0,175	3	4	4	4
0,200	2	3	3	3
0,225	2	3	3	3
0,250	2	3	3	3
0,275	2	3	3	3
0,300	2	3	3	3
0,325	2	2	3	3
0,350	2	2	3	3
0,375	2	2	3	3
0,400	2	2	2	3
0,425	2	2	2	3
0,450	2	2	2	3
0,475	2	2	2	3
0,500	2	2	2	3
0,525	2	2	2	3
0,550	2	2	2	3
0,575	2	2	2	3
0,600	2	2	2	3
0,625	2	2	2	3
0,650	2	2	2	2
0,675	2	2	2	2
0,700	2	2	2	2
0,725	2	2	2	2
0,750	2	2	2	2
0,775	2	2	2	2
0,800	2	2	2	2
0,825	2	2	2	2
0,850	2	2	2	2
0,875	2	2	2	2
0,900	2	2	2	2
0,925	2	2	2	2
0,950	2	2	2	2
0,975	2	2	2	2
1,000	1	1	1	1

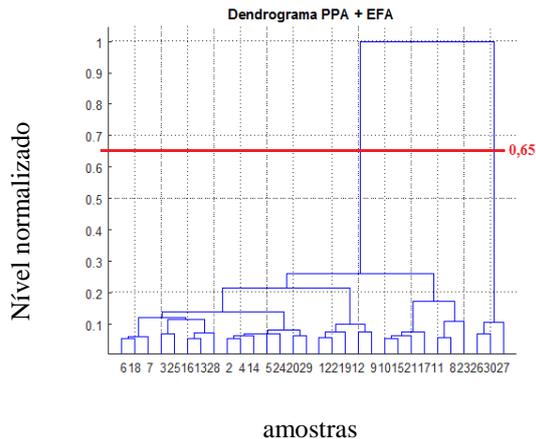


Figura 4.7 – Dendrograma PPA
+EFA com 23 variáveis

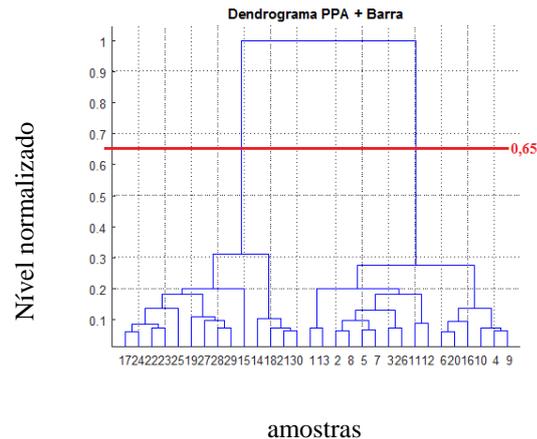


Figura 4.8 – Dendrograma PPA
+barra com 23 variáveis

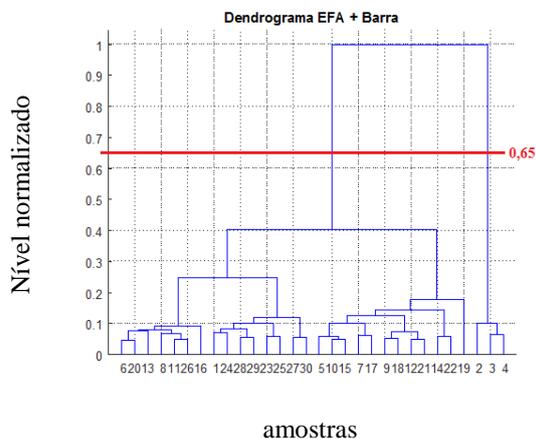


Figura 4.9 – Dendrograma EFA
+barra com 23 variáveis

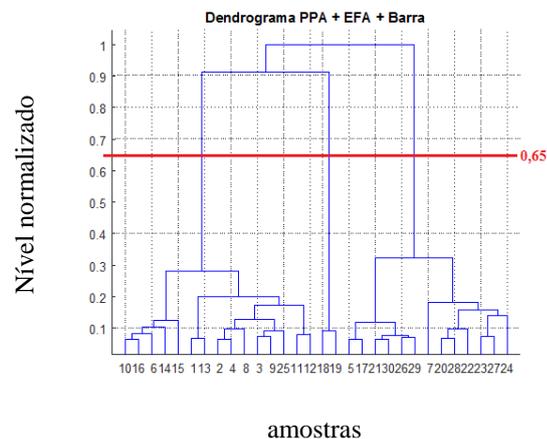


Figura 4.10 – Dendrograma PPA
+EFA+barra com 23 variáveis

O resultado aqui obtido demonstra que tanto o método proposto quanto o método dos dendrogramas, conhecido na literatura, podem ser utilizados na determinação do NC_1 com sucesso. É preciso agora avaliar o percentual de acerto do FCM na determinação da origem correta de cada sinal (não confundir com a classificação dos mapas de DP, que será tratada posteriormente). Utilizando-se a configuração com três fontes diferentes de DP (células PPA, EPA e barra de gerador), foram feitas simulações com diferentes números de variáveis de entrada (na ordem de importância do Relief) para se determinar a origem de cada sinal, já fornecendo ao FCM o NC_1 . O percentual de acerto obtido para cada conjunto de variáveis é mostrado na tabela 4.13.

Tabela 4.13 – Taxa de acerto do FCM para 3 *clusters* (classes conhecidas - objetos ensaiados separadamente e unidos depois)

Nº de Variáveis utilizadas (na ordem de relevância)	Taxa de Acerto (%)			
	Total	Na célula PPA	Na célula EFA	Na barra de gerador
2	91,9	93,1	99,6	89,6
3	93,7	95,9	99,5	90,9
4	94,8	97,0	99,5	92,2
5	96,0	99,6	99,5	92,5
6	95,6	99,5	99,5	91,9
7	95,6	99,8	99,6	91,6
8	95,2	99,8	99,6	91,0
9	95,1	99,8	99,5	90,7
10	95,1	99,8	99,5	90,7
11	94,9	99,8	99,5	90,4
12	94,9	99,8	99,5	90,4
13	94,8	99,8	99,5	90,1
14	94,8	99,8	99,5	90,0
15	94,8	99,8	99,5	90,1
16	94,6	99,8	99,5	89,6
17	94,7	99,8	99,5	89,9
18	94,9	99,8	99,5	90,4
19	94,9	99,8	99,5	90,4
20	94,9	99,8	99,5	90,3
21	94,9	99,8	99,5	90,3
22	95,0	99,8	99,5	90,5
23	95,0	99,8	99,5	90,5
Variáveis 17 e 18	88,4	99,0	99,5	78,1

A melhor taxa de acerto na determinação da origem de cada sinal foi obtida para 5 variáveis (96 % de acerto total). Foi elevada também a taxa de acerto obtida com todas as 23 variáveis (95 % de acerto total), que considera várias nuances dos sinais devido à diversidade de variáveis relevantes e não redundantes

consideradas. Como pode ser observado na tabela 4.13, foi realizada também uma simulação utilizando-se as variáveis 17 e 18 (referência na literatura pela eficácia na separação das fontes [81][82]). Observa-se que, com as variáveis 17 e 18 somente, o FCM tem um percentual de acerto mais baixo do que em todos os demais casos. Isto ressalta a quantidade e a qualidade de informação agregada entregue pelas variáveis propostas neste estudo.

4.2.2

Classificação dos Mapas de Descargas Parciais

Até agora foram avaliadas as técnicas aqui propostas relacionadas à etapa de agrupamento, utilizando-se para isso sinais categorizados (vindos dos objetos ensaiados individualmente). Visto que estas técnicas (seleção de variáveis com o Relief e determinação do NC_1 através da matriz de dissimilaridade) mostraram-se satisfatórias e que seus efeitos puderam ser medidos em virtude de os sinais serem categorizados, o objetivo agora é aplicar as mesmas técnicas para classificação dos mapas de DP utilizando-se os sinais medidos nos objetos ensaiados em paralelo, passando-se rapidamente pela etapa de agrupamento.

Serão utilizados neste tópico os 16 sinais medidos com os objetos ensaiados em paralelo nas 4 diferentes configurações (PPA//EFA, PPA//barra de gerador, EFA//barra de gerador, e PPA//EFA//barra de gerador). Na Figura 4.11 é mostrada uma foto do arranjo de ensaio montado em laboratório para ensaiar os 3 objetos em paralelo.

Utilizando-se o limiar de junção $L_J = 0,5$, foi obtido sempre o número de *clusters* ideal NC_1 (correto) nos 16 sinais em análise (seja ele 2 ou 3), independentemente do número de variáveis de entrada utilizado na simulação. O mesmo resultado foi obtido aplicando-se a técnica dos dendrogramas, ou seja, o NC_1 foi entregue corretamente em todas as configurações.



Figura 4.11 – Arranjo de ensaio montado em laboratório para ensaiar os 3 objetos em paralelo

O passo seguinte foi utilizar o NC_1 no FCM e realizar o agrupamento dos 16 sinais, obtendo-se os mapas de DP separados de cada fonte de DP. Neste ponto, um teste adicional foi realizado. No FCM, cada ponto é atribuído a um *cluster* com determinado grau de pertinência (GP) associado. A ideia é testar a classificação dos mapas de DP formados com todos os pontos (opção 1) e também testar a classificação dos mapas de DP formados somente com pontos de elevado grau de pertinência ao *cluster* (opção 2). Isso equivale a utilizar um grau de pertinência mínimo (GPMin) como ponto de corte para os pontos (sinais). Adotou-se 0 para a opção 1 (GPMin = 0) e 0,7 para a opção 2 (GPMin = 0,7). A tabela 4.14 apresenta as taxas de acerto obtidas para estas duas opções, com diferentes números de variáveis de entrada, segundo a metodologia de classificação aqui proposta, baseada na correlação dos histogramas de fase em relação ao banco histórico criado aqui (vide capítulo 3). A Figura 4.12 retrata os dados da tabela 4.14 em forma de gráfico.

Tabela 4.14 – Taxa de acerto na classificação dos mapas de DP separados dos sinais medidos em paralelo nos objetos em laboratório

Número de variáveis utilizado no agrupamento	Taxa de acerto (%) em relação ao GPMin adotado	
	GPMin = 0	GPMin = 0,7
2	77,7	75,0
3	77,7	80,6
4	80,6	83,3
5	80,6	80,6
6	77,7	80,6
7	77,7	75,0
8	80,6	80,6
9	80,6	80,6
10	80,6	75,0
11	80,6	75,0
12	80,6	80,6
13	86,1	80,6
14	83,3	72,2
15	86,1	75,0
16	88,9	69,4
17	86,1	69,4
18	86,1	75,0
19	86,1	75,0
20	88,9	75,0
21	86,1	80,6
22	86,1	75,0
23	86,1	86,1
Só variáveis 17 e 18	77,7	77,7

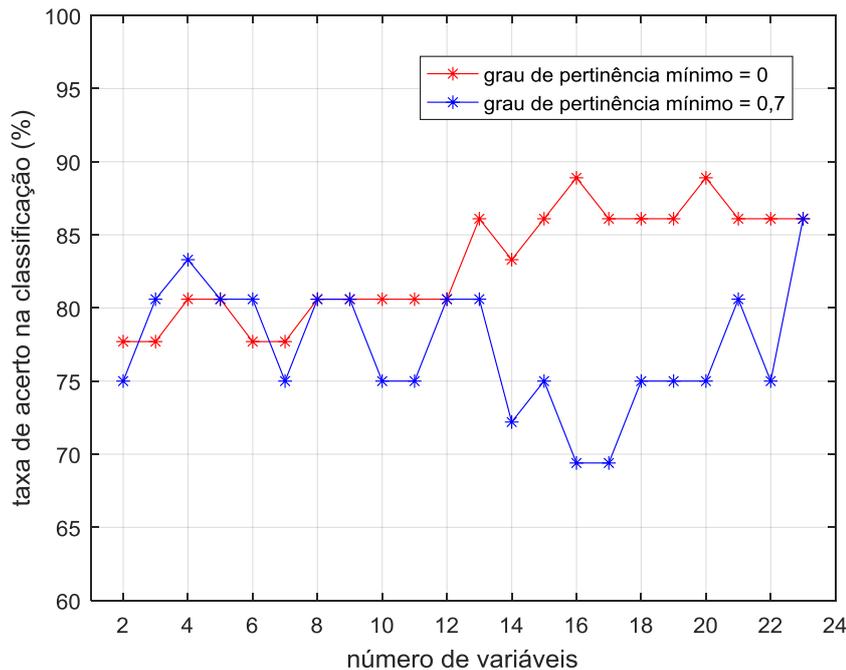


Figura 4.12 – Taxa de acerto na classificação dos mapas de DP separados em função do número de variáveis de entrada e analisando o grau de pertinência mínimo

A Tabela 4.14 leva em consideração a classificação de 36 mapas de DP (em cada opção de GPMín) originados dos 16 sinais (3 configurações com 2 mapas cada e 1 configuração com 3 mapas, multiplicados por 4, pois foram medidos 4 sinais em cada arranjo). Desta tabela e da Figura 4.12, pode-se extrair algumas informações:

- 1) O melhor percentual de acerto atingido na classificação dos mapas de DP foi 88,9 %, com GPMín = 0 e 16 variáveis de entrada;
- 2) Percebe-se que a utilização somente de pontos com elevado grau de pertinência (GPMín = 0,7) na formação dos mapas de DP a serem classificados não resultou, em geral, em uma maior taxa de acerto (quando comparado com GPMín = 0);
- 3) A taxa de acerto tende a aumentar com o número de variáveis de entrada utilizado no agrupamento quando se consideram todos os pontos (GPMín = 0), provavelmente devido à utilização de variáveis relevantes e não redundantes no agrupamento, o que pode encorajar a utilização das 23 variáveis em futuras classificações;

- 4) A taxa de acerto obtida utilizando-se somente as variáveis 17 e 18 (77,7 %), consagradas na literatura, foi consideravelmente baixa em relação ao melhor resultado obtido aqui com as variáveis propostas (88,9 %);
- 5) Pode-se concluir que a correlação de Pearson estabelece uma boa métrica de comparação dos histogramas de fase a classificar com os histogramas padrão do banco de dados histórico criado;
- 6) Comprovou-se que os histogramas de fase traduzem bem as características encontradas nos mapas de DP e podem ser usados na classificação destes;
- 7) A utilização de sinais de DP medidos em objetos ensaiados em paralelo no laboratório não comprometeu a classificação correta dos mapas de DP, visto o elevado percentual de acerto obtido na classificação destes.

Como exemplo, é apresentado na Figura 4.13 um dos 4 mapas de DP medidos no ensaio dos 3 objetos em paralelo (célula PPA//célula EFA//barra de gerador). Na Figura 4.14, são apresentados os 3 mapas obtidos após o agrupamento realizado no mapa original medido, considerando-se 23 variáveis de entrada, com a respectiva classificação de saída baseada no banco de dados histórico. Também a título de ilustração, são apresentados nas Figuras 4.15 a 4.17 os histogramas de fase de um mapa de DP medido na célula EFA, que é utilizado como referência no banco de dados histórico criado para determinar a classe de novos mapas a classificar.

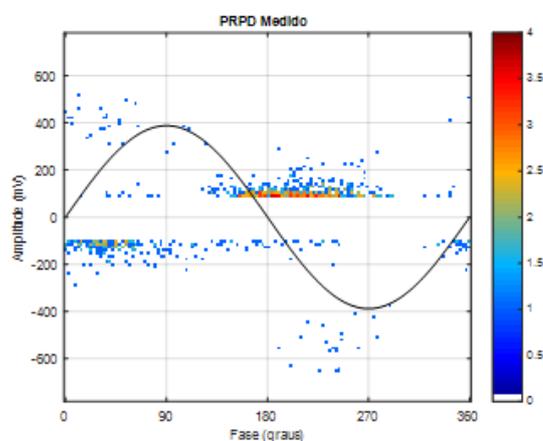


Figura 4.13 – Mapa de DP medido nos ensaio com as duas células e a barra de gerador em paralelo

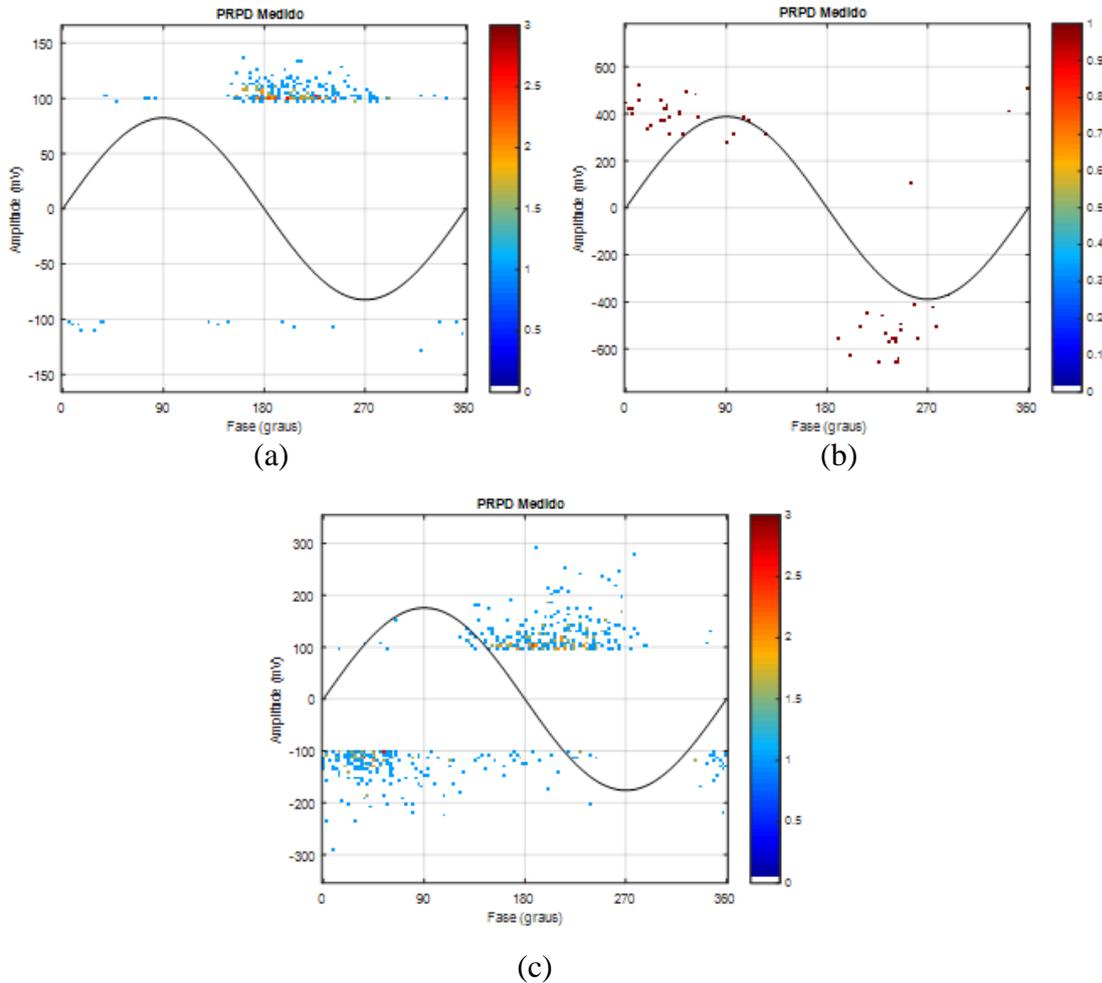


Figura 4.14 – Mapas de DP obtidos após agrupamento do sinal original, sendo a classe de saída: (a) célula PPA; (b) célula EFA; e (c) barra de gerador

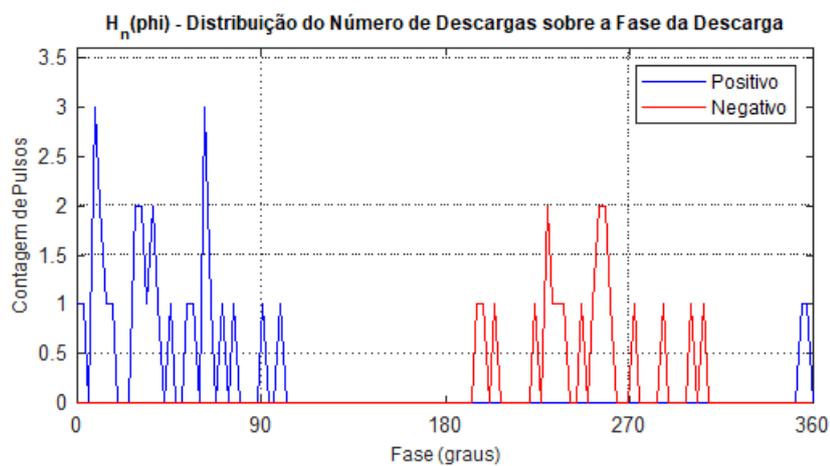


Figura 4.15 - Histogramas positivo e negativo do número de pontos em relação à fase – $H_n(\text{fase})$

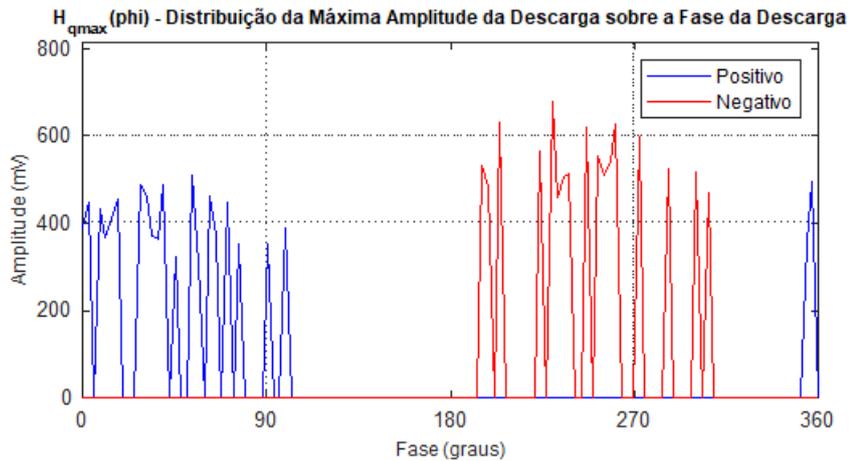


Figura 4.16 - Histogramas positivo e negativo da carga máxima em relação à fase – $H_{q_{max}}(fase)$

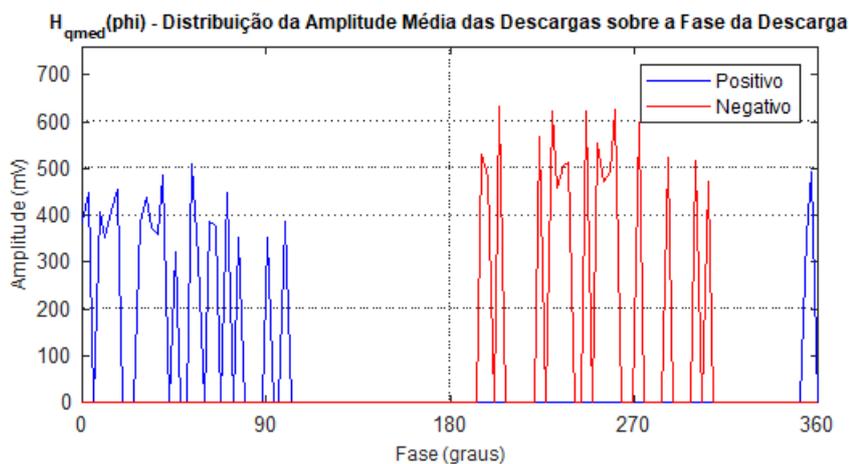


Figura 4.17 - Histogramas positivo e negativo da carga média em relação à fase – $H_{q_{med}}(fase)$

4.3

Aplicação das Técnicas Propostas em um Sinal Real com Ruído

Um problema enfrentado em grande parte das medições é a interferência de ruído externo ao equipamento que está sendo avaliado nos dados capturados, que pode ser proveniente de outras máquinas instaladas na planta. Se a faixa de frequência dos sinais de ruído estiver sintonizada com a dos sinais de DP que estão sendo capturados, o ruído pode mascarar a medição e prejudicar a interpretação dos sinais. Técnicas de separação de fontes de DP podem ser

empregadas para extrair o ruído da medição e deixar apenas os sinais de DP, que representam a informação útil desejada.

O sinal no tempo que forma o mapa de DP capturado num determinado gerador (referenciado no capítulo 3 como “gerador 1”), mostrado na Figura 4.18 (mapa original da medição), foi então caracterizado em relação às 23 variáveis de entrada. Utilizando-se um limiar de junção $L_J = 0,50$, chegou-se a um valor de número de *clusters* ideal $NC_1 = 3$. No dendrograma, com ponto de corte do nível em $PCN = 0,65$ (normalizado), também se obteve $NC_1 = 3$. Aplicou-se então o FCM (com coeficiente fuzzy $m = 1,2$ e $GPM_{in} = 0$ na seleção de pontos) para realizar-se o agrupamento dos pontos e a conseqüente separação dos três mapas de DP (fruto dos três *clusters* do agrupamento), mostrados na Figura 4.19, contidos no mapa original medido.

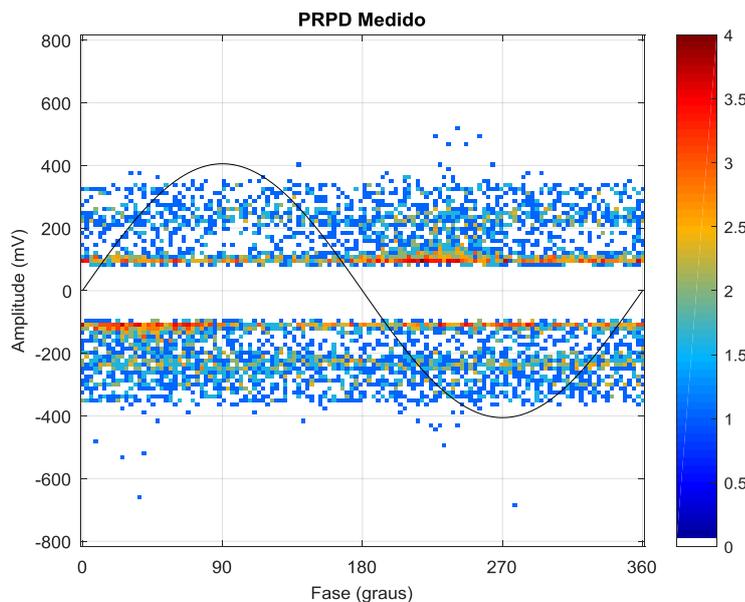


Figura 4.18 –Mapa de DP original do gerador 1 com ruído

O mapa da fonte 2 (Figura 4.19b) e o mapa da fonte 3 (Figura 4.19c) foram classificados dentro do banco de dados histórico criado (baseado na correlação de Pearson de seus histogramas de fase, conforme proposto aqui e detalhado no capítulo 3) como sendo sinais de ruído, cuja principal característica é a sua presença em praticamente todas as fases da senóide, o que pode ser constatado na observação dos mapas. Estes dois ruídos têm provavelmente duas origens diferentes. Percebe-se que o primeiro (fonte 2) apresenta sinais com valor

absoluto variando essencialmente entre 150 e 350 mV e baixa densidade de pontos. Já o segundo (fonte 3) apresenta sinais com valor absoluto concentrado entre 100 e 150 mV em média (sinais abaixo de 100 mV foram desconsiderados nessa medição) e densidade de pontos mais elevada. O mapa da fonte 1 (Figura 4.19a) diz respeito aos sinais realmente importantes (contém informação útil), pois representam sinais de DP oriundos do equipamento sob avaliação, concentrados principalmente na descida do semiciclo negativo e na subida do semiciclo positivo da senóide. A classificação deste mapa da fonte 1 perante o banco de dados histórico criado apontou para a classe DP3, cavidades internas, que correspondem a DP geradas em bolhas de ar ou gás embutidas no isolamento principal, apresentando baixo grau de risco de falha associado [97]. Neste caso, a indicação é somente continuar realizando as medições periódicas de DP na máquina para verificar se há evolução do defeito.

A análise deste exemplo de sinal real mostra a importância da separação de fontes de DP, visto que a mera observação do mapa da Figura 4.18 não permite identificar se há alguma assinatura característica de DP e, caso ela exista, se apresenta risco elevado de falha (requerendo neste caso uma intervenção corretiva na máquina).

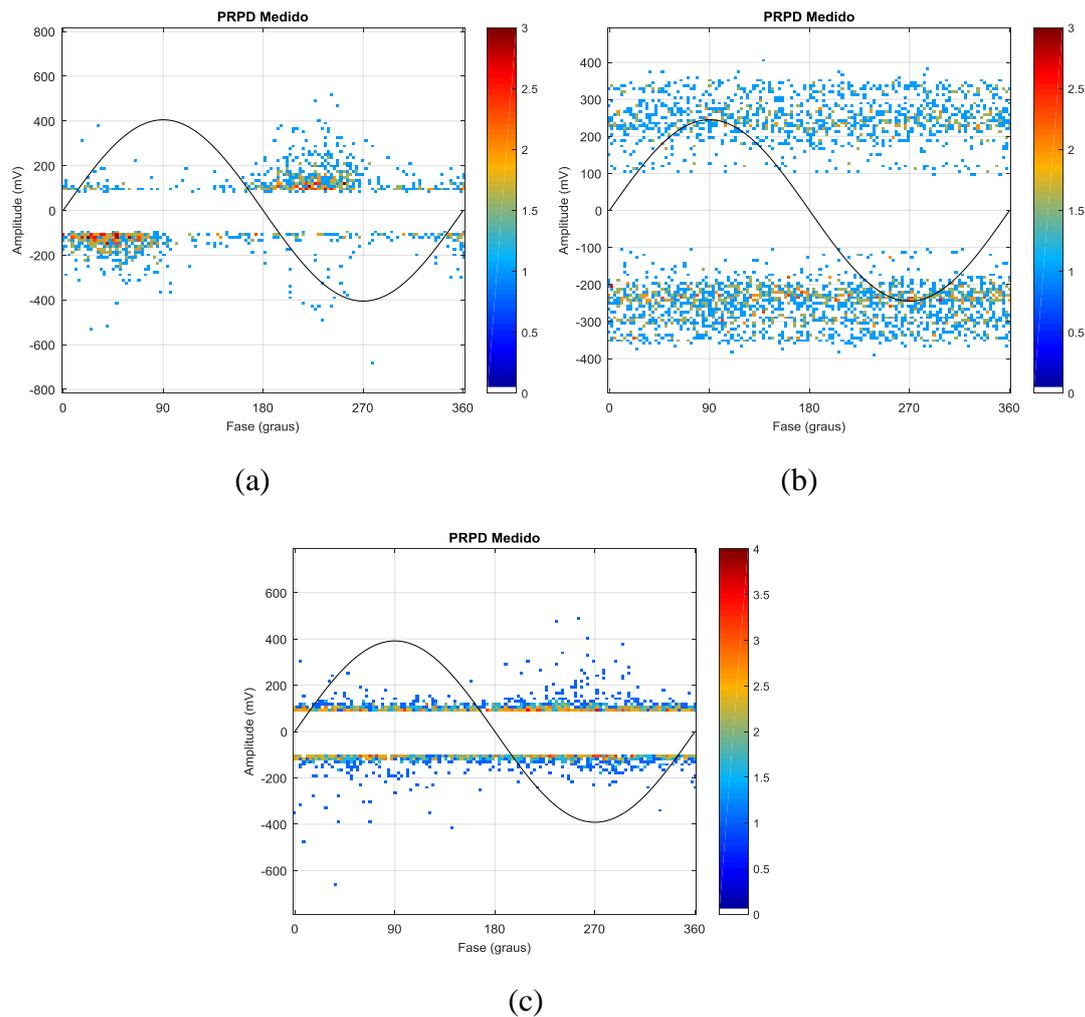


Figura 4.19 – Mapas derivados do mapa original medido em campo no gerador 1:

- (a) fonte 1 (DP3 – cavidades internas); (b) fonte 2 (ruído);
e (c) fonte 3 (ruído)

4.4

Testando a Eficácia da Classificação do Banco de Dados Histórico em Mapas de Descargas Parciais com Classe Conhecida

Para testar o desempenho da metodologia de classificação aqui proposta, baseada na correlação de Pearson dos histogramas de fase de mapas a classificar com os histogramas de fase dos mapas padrões que compõem o banco de dados histórico criado (baseado em padrões de norma de máquinas rotativas), é necessário um conjunto com vários mapas de teste com classes conhecidas. Será então utilizado um banco de dados com mapas de DP medidos em um gerador (referenciado no capítulo 3 como “gerador 2”) avaliado de forma *on-line*.

Em alguns casos, devido à grande dimensão dos geradores, é necessária a instalação de vários sensores (acopladores capacitivos) na máquina de forma que estes se aproximem fisicamente das possíveis fontes de DP para amenizar o efeito de atenuação dos sinais medidos. Assim, pode acontecer de os sinais capturados por meio de um mesmo acoplador estarem relacionados a uma única fonte de defeito e todos os mapas capturados serem então bastante semelhantes neste sensor. Isto ocorre com este banco. Ele é composto por 320 mapas de DP já classificados medidos em 8 acopladores instalados ao longo de uma das fases do estator de um gerador de grande porte. A tabela 4.15 apresenta o tipo de defeito detectado em cada acoplador e o grau de risco associado [97].

Tabela 4.15 – Tipo de defeito detectado em torno de cada acoplador e grau de risco associado

Acoplador	Tipo de defeito	Grau de risco	Nº de mapas
1	DP1 - Gap	normal	40
2	DP2 - ranhura	alto	40
3	DP1 - Gap	normal	40
4	DP3 - cavidades internas	baixo	40
5	DP1 - Gap	normal	40
6	DP4 - delaminação interna	alto	40
7	DP2 - ranhura	alto	40
8	DP4 - delaminação interna	alto	40

Nas Figuras 4.20 a 4.27, é apresentado um exemplo de mapa medido em cada acoplador. Nestes mapas, é possível observar a assinatura característica dos defeitos e compará-los com os padrões de norma (cujos mapas são mostrados no capítulo 3) para verificar a semelhança entre eles.

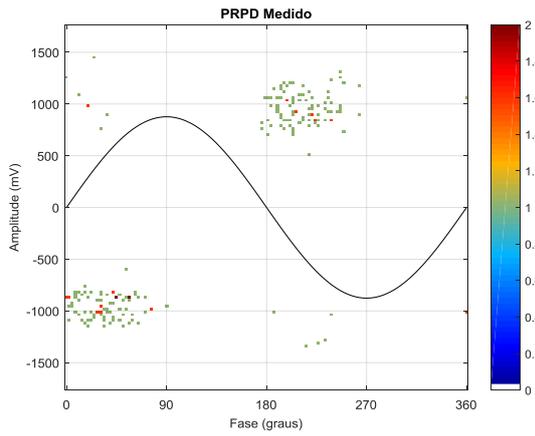


Figura 4.20 – Mapa medido no acoplador 1

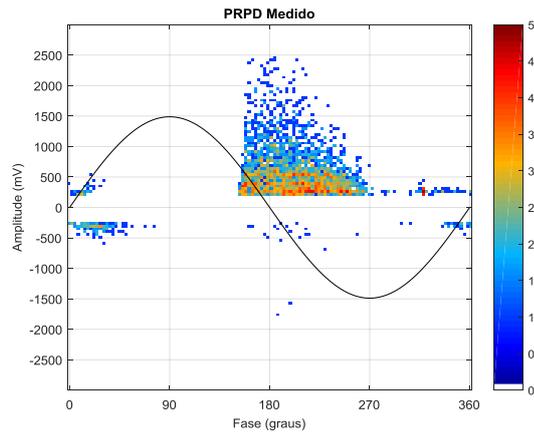


Figura 4.21 – Mapa medido no acoplador 2

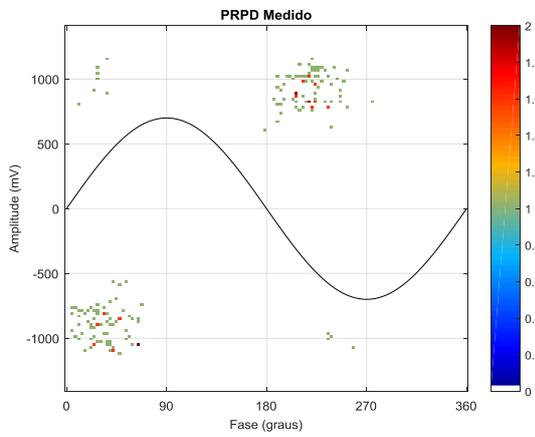


Figura 4.22 – Mapa medido no acoplador 3

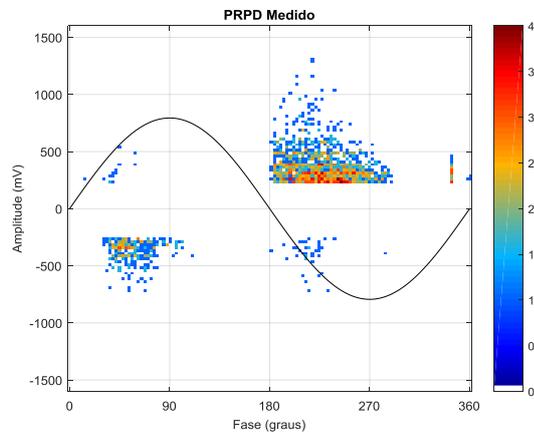


Figura 4.23 – Mapa medido no acoplador 4

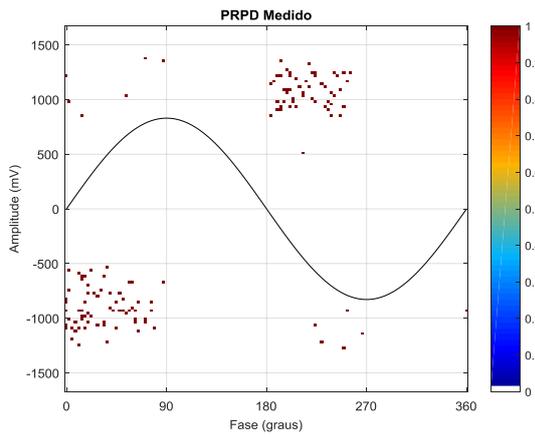


Figura 4.24 – Mapa medido no acoplador 5

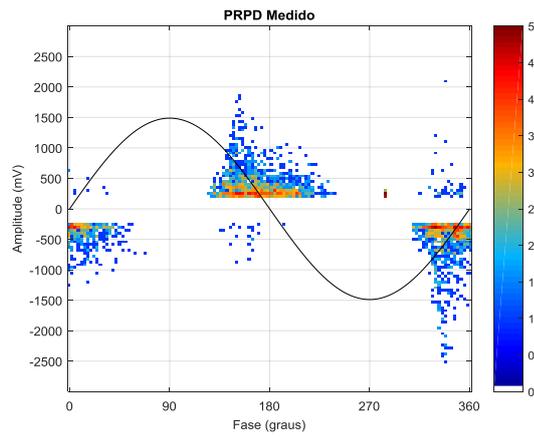


Figura 4.25 – Mapa medido no acoplador 6

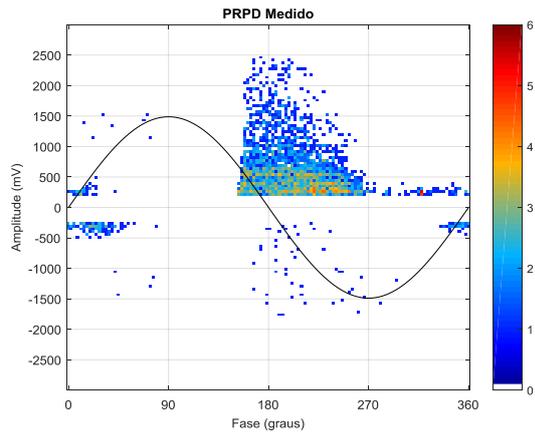


Figura 4.26 – Mapa medido no acoplador 7

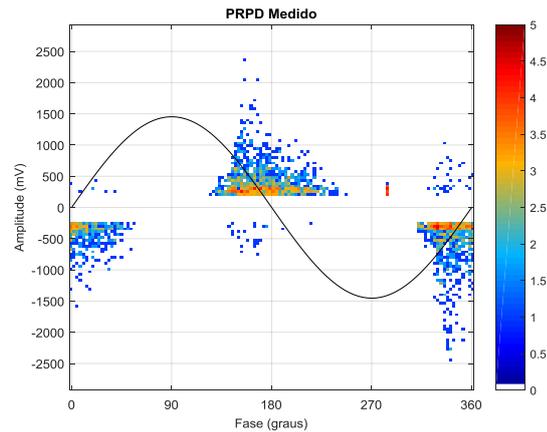


Figura 4.27 – Mapa medido no acoplador 8

Os 320 mapas foram então classificados em relação ao banco de dados criado utilizando a metodologia aqui proposta. A classe de defeito determinada para cada mapa foi então comparada à classe real pré-estabelecida pelo especialista, obtendo-se um elevado percentual de acerto em cada acoplador, conforme mostrado na tabela 4.16.

Tabela 4.16 – Taxa de Acerto em cada acoplador

Acoplador	Taxa de acerto (%)
1	95,0
2	100
3	97,5
4	100
5	100
6	100
7	100
8	100
Total	99,1

A elevada taxa de acerto de 99,1 % demonstra a eficácia da metodologia na classificação dos mapas, indicando que ela pode ser utilizada na classificação automática dos mapas medidos, sob supervisão controlada do especialista. Além disso, o banco de dados histórico criado pode ser utilizado como base de

comparação para máquinas rotativas (excetuando-se os sinais das células PPA, EFA e da barra de teste incluídos neste).

4.5

Adaptação do Banco de Dados Histórico para Mapas de Descargas Parciais Unipolares com Classe Conhecida

Nesta parte do estudo, são utilizados cinco mapas de DP unipolares (valores em módulo) capturados em medições *off-line* (com sensor único na saída de fase) em 4 geradores e 1 motor de grande porte (referenciados no capítulo 3 como “gerador 3”, “gerador 4”, “gerador 5”, “gerador 6”, e “motor 1”). São resultado de limitações do hardware comercial de medição em campo, necessário em algumas situações específicas para manter elevada a taxa de captura de pontos importantes (pulsos de DP no tempo) para formar o mapa de DP. Quando há muita interferência externa de ruído dos processos industriais do chão de fábrica na medição, por exemplo, o elevado número de pontos relacionados a ruído capturados no mapa perturba a captura de pontos importantes para caracterizar o defeito presente, forçando a utilização de uma espécie de filtro *on-line* (no momento da medição) destes ruídos que só permite capturar mapas unipolares.

O objetivo aqui é verificar se os mapas são classificados corretamente ao se adaptar o banco de dados histórico para funcionar também com dados unipolares medidos com o equipamento comercial em questão. Neste caso, os histogramas de fase $H_n(\text{fase})^+$, $H_n(\text{fase})^-$, $H_{q_{\max}}(\text{fase})^+$, $H_{q_{\max}}(\text{fase})^-$, $H_{q_{\text{med}}}(\text{fase})^+$, e $H_{q_{\text{med}}}(\text{fase})^-$, utilizados no banco histórico, são transformados em $H_n(\text{fase})$, $H_{q_{\max}}(\text{fase})$, e $H_{q_{\text{med}}}(\text{fase})$, da seguinte forma:

$$H_n(\text{fase}) = (H_n(\text{fase})^+) + (H_n(\text{fase})^-) \quad \text{Eq.4.1}$$

$$H_{q_{\max}}(\text{fase}) = \max [H_{q_{\max}}(\text{fase})^+, H_{q_{\max}}(\text{fase})^-] \quad \text{Eq.4.2}$$

$$H_{q_{\text{med}}}(\text{fase}) = \frac{ [H_n(\text{fase})^+ \cdot H_{q_{\text{med}}}(\text{fase})^+] + [H_n(\text{fase})^- \cdot H_{q_{\text{med}}}(\text{fase})^-] }{ (H_n(\text{fase})^+) + (H_n(\text{fase})^-) } \quad \text{Eq.4.3}$$

É importante ressaltar que os histogramas determinados nas equações 4.1 a 4.3 devem ser normalizados após o cálculo, ou seja, divididos pela área total de cada um. Assim, a correlação de Pearson é aplicada agora a três histogramas e não

mais a seis, para verificar quais mapas padrões, “transformados” em unipolares, apresentam maior semelhança com o mapa testado.

Na tabela 4.17, são apresentadas as classes reais dos mapas de DP medidos atribuídas pelo especialista pela comparação visual com os padrões de norma com o respectivo grau de risco de falha. Nas Figuras 4.28 a 4.32, são mostrados os mapas unipolares medidos nas máquinas, referidos aqui. Vale ressaltar que, apesar de não aparecerem no mapa, os valores no eixo vertical foram normalizados para se situarem entre 0 (mínimo no eixo) e 1 (máximo no eixo), visto que o importante, neste caso, é a assinatura do mapa de DP e não a sua amplitude. Já o eixo horizontal vai de 0 (mínimo no eixo) a 360 graus (máximo no eixo), como de costume.

Tabela 4.17 – Classes reais dos mapas avaliados

Máquina	Tipo de defeito	Grau de risco de falha
Gerador 3	DP2 - ranhura	alto
Gerador 4	DP3 - cavidades internas	baixo
Gerador 5	DP4 - Delaminação interna	alto
Gerador 6	DP2 - ranhura	alto
Motor 1	DP5 - Tracking de superfície	normal

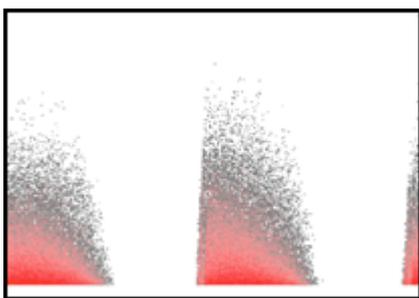


Figura 4.28 – Mapa unipolar medido no gerador 3

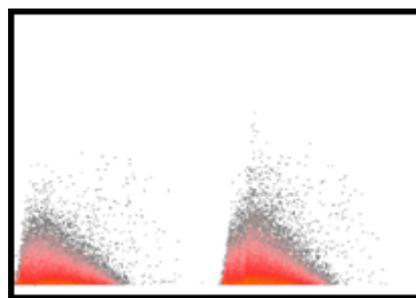


Figura 4.29 – Mapa unipolar medido no gerador 4

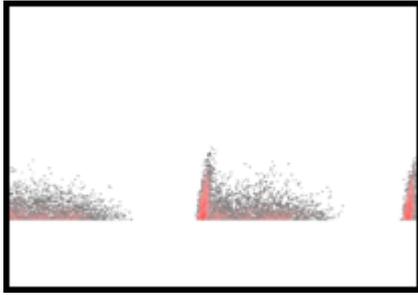


Figura 4.30 – Mapa unipolar
medido no gerador 5

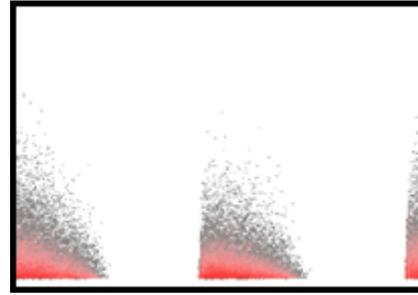


Figura 4.31 – Mapa unipolar
medido no gerador 6



Figura 4.32 – Mapa unipolar
medido no motor 1

Os 5 mapas unipolares em análise foram então classificados em relação ao banco de dados adaptado utilizando a metodologia proposta. A classe de defeito determinada para cada mapa foi compatível com a classe real de cada um em todos os casos, mostrando uma taxa de acerto de 100 %. Isso demonstra que é possível fazer também a classificação automática de mapas de DP unipolares com uma pequena adaptação (de fácil implementação) no banco de dados histórico criado.

5

Conclusões e Trabalhos Futuros

Uma nova sistemática foi proposta aqui para detectar descargas parciais em para-raios e transformadores de corrente de subestação a partir de sinais medidos em campo, considerando desde a extração de características discriminantes dos sinais no tempo medidos nestes até a determinação de técnicas utilizando árvores de decisão e máquina de vetor suporte (para cada tipo de equipamento separadamente), capazes de separar sinais de equipamentos com e sem descargas parciais internas com uma elevada taxa de acerto. A variável 16 (intervalo de tempo do pulso de sinal do pico até 30 % deste), proposta aqui para caracterização de sinais no tempo, mostrou-se como a mais bem avaliada para os transformadores de corrente e a segunda mais bem avaliada para os para-raios quanto à capacidade de discriminação de sinais com e sem descargas parciais nestes equipamentos.

As novas variáveis 4, 5 (primeiro e segundo maior pico de dispersão do valor absoluto da FFT da ACF do sinal normalizado no banco de dados em questão, respectivamente) e 16, propostas neste estudo, demonstram grande capacidade de discriminação das diferentes fontes de descargas parciais medidas nos sinais das duas células de teste de descargas parciais e da barra de gerador ensaiadas em laboratório. Estas variáveis podem também ser utilizadas em conjunto com as variáveis 17 (métrica T no espaço tempo-frequência equivalente ao tempo proposta por [21]) e 18 (métrica F no espaço tempo-frequência equivalente à frequência proposta por [21]), que vêm apresentando eficácia na separação de fontes de descargas parciais conforme a literatura [81][82].

Uma nova métrica, baseada no conceito de “matriz de dissimilaridade relativa”, foi proposta e se mostrou eficaz como critério de determinação do número de *clusters* ideal em espaços de separação de fontes de descargas parciais com duas ou três origens diferentes (número de fontes encontrado na maioria dos casos de máquinas rotativas). Neste processo, foi estabelecida uma faixa de valores para o parâmetro “Limiar de Junção”, proposto neste estudo, dentro da qual se obtém um número de *clusters* capaz de proporcionar uma discriminação

precisa e não excessiva dos sinais de descargas parciais presentes no mapa de descargas original medido.

Foi proposto também um critério de classificação de mapas de descargas parciais baseado na correlação de Pearson entre os histogramas de fase de mapas de descargas parciais a classificar e os histogramas de fase de mapas de descargas parciais já classificados, que compõem um banco de dados histórico para comparação. Esta técnica mostrou-se eficaz para o banco de sinais das duas células e da barra de gerador ensaiados em laboratório para este estudo, e também nos demais mapas medidos em máquinas rotativas e utilizados aqui também.

O processo de criação do banco de dados histórico com padrões de mapas de descargas parciais, com assinaturas de diferentes tipos de defeito em máquinas rotativas, a partir do processamento de imagens destes padrões retiradas de norma, mostrou-se eficaz quando testado na classificação de mapas com classe determinada. Isso se afigura como uma ferramenta de grande valor, haja vista a dificuldade de se obterem padrões característicos de defeito por meio de ensaios em laboratório, principalmente pela dificuldade em conseguir amostras de teste representativas dos diferentes tipos de defeito. O banco histórico pode ainda ser alimentado com mais padrões de defeito, visto que um determinado defeito pode ter mais de uma assinatura característica. Além disso, outros bancos podem ser criados, com base nos mesmos princípios, para uso em outros tipos de isolamento que não os de máquinas rotativas (transformadores, por exemplo).

Em relação à aplicação imediata, este estudo resultou uma metodologia completa para caracterização e classificação de sinais (quanto à presença de descargas parciais) em para-raios e transformadores de corrente de forma automática, que está pronta para ser incorporada às práticas de campo adotadas pelo Cepel e também disponibilizadas às empresas do Grupo Eletrobrás que tenham interesse no sistema para avaliar seus ativos.

As variáveis aqui propostas podem ser utilizadas para extrair informações de sinais no tempo em geral. Desta forma, elas podem ser aplicadas em diferentes bancos de dados oriundos de medições realizadas em diferentes equipamentos. Um caminho natural é a aplicação futura destas variáveis para detecção de descargas parciais em transformadores de potencial (TPs) de subestação, sejam eles indutivos ou capacitivos. O Cepel já vem realizando medições em equipamentos desta natureza, nos mesmos moldes das medições nos PRs e TCs

citadas aqui, porém o número de equipamentos avaliados em campo ainda é pequeno quando comparado aos PRs e TCs já avaliados. Desta forma, com a aquisição de mais dados oriundos de TPs, é natural que as técnicas utilizadas aqui em PRs e TCs também sejam estendidas a estes num futuro próximo.

Um estudo interessante poderia também ser realizado por meio do ensaio em paralelo de barras de gerador com diferentes tipos de defeito em laboratório para testar as técnicas propostas aqui para separação de fontes de DP nos dados obtidos. Mesmo que algumas barras tenham tipos repetidos de defeito (DP por delaminação, por exemplo, que é bastante comum), já seria um exercício interessante para avaliar as técnicas propostas, visto que a assinatura do sinal medido está também relacionada ao caminho que ele "percorre" até o sensor, além das características pela natureza do defeito.

Outro trabalho que pode ser feito de imediato é a ampliação do banco de dados histórico proposto aqui, composto por padrões de mapas de DP retirados de norma. Este banco pode ser ampliado com mapas de DP com assinaturas de diferentes tipos de defeito medidos em máquinas rotativas, dado que um mesmo tipo de defeito pode possuir diferentes variantes de assinatura, apesar de elas guardarem certa semelhança entre si. Quanto mais mapas de DP bem catalogados forem utilizados para compor o banco de dados histórico, maior a chance de se classificar corretamente de um novo mapa de DP em análise com a metodologia proposta neste estudo.

Todas as técnicas utilizadas aqui para máquinas rotativas poderiam também ser aplicadas para transformadores de potência, desde a determinação de variáveis relevantes e não redundantes para a separação de fontes de DP até a classificação dos mapas de DP pela correlação dos histogramas de fase, passando também pela criação de um banco de dados histórico para este tipo de equipamento.

As taxas de acerto obtidas neste estudo na classificação dos PRs e TCs utilizando SVM mostraram-se elevadas (93,1 % e 96,3 %, respectivamente), mas isso não impede que em trabalhos futuros nestes equipamentos obtenha-se melhores resultados explorando-se a variação de parâmetros do SVM. Já com relação às árvores de decisão, há espaço para explorar a estrutura formada por estas para a classificação dos equipamentos, visto que elas contam história (ressaltam nuances importantes) e podem trazer informações valiosas em determinadas aplicações.

Referências Bibliográficas

- [1] TIAN, Y., LEWIN, P.L., DAVIES, A.E., “Comparison of on-line partial discharge detection methods for HV cable joints”, *IEEE Trans. Dielectr. Electr. Insul.* 9 604–615, 2002.
- [2] SATISH, L., ZAENGL, W.S., “Artificial neural networks for recognition of 3-d partial discharge patterns”, *IEEE Trans. Dielectr. Electr. Insul.* 1, 265–275, 1994.
- [3] ANGRISANI, L., DAPONTE, P., LUPÒ, G., PETRARCA, C., VITELLI, M., “Analysis of ultrawide-band detected partial discharges by means of a multiresolution digital signal-processing method”, *Measurement* 27, 207–221, 2000.
- [4] SATISH, L., GURURAJ, B.I., “Partial discharge pattern classification using multilayer neural networks”, *Sci., Meas. Technol., IEEE Proc. A* 140, 323–330, 1993.
- [5] KARTHIKEYAN, B., GOPAL, S., VENKATESH, S., “ART 2 – An unsupervised neural network for PD pattern recognition and classification”, *Expert Syst. Appl.* 31, 345–350, 2006.
- [6] VENKATESH, S., GOPAL, S., “Robust heteroscedastic probabilistic neural network for multiple source partial discharge pattern recognition – significance of outliers on classification capability”, *Expert Syst. Appl.* 38, 11501–11514, 2011.
- [7] KARTHIKEYAN, B., GOPAL, S., VENKATESH, S., “Partial discharge pattern classification using composite versions of probabilistic neural network inference engine”, *Expert Syst. Appl.* 34, 1938–1947, 2008.
- [8] SHAFIQ, M., HUSSAIN, G.A., KÜTT, L., LEHTONEN, M., “Effect of geometrical parameters on high frequency performance of Rogowski coil for partial discharge measurements”, *Measurement* 49 (3), 126–137, 2014.
- [9] HOOF, M., FREISLEBEN, B., PATSCH, R., “PD source identification with novel discharge parameters using counter propagation neural networks”, *IEEE Trans. Dielectr. Electr. Insul.* 4, 17–32, 1997.

- [10] AMORIM JÚNIOR, H.P., CARVALHO, A.T., RODRIGUES, T.B., BATISTA, J.B., FLEMING, C.C., "Experience with On-Line Insulation Diagnostics of Surge Arresters by PD Measurement in the Field", 2016 IEEE International Conference on Dielectrics (ICD), v.1, p.p. 472-475, Montpellier - França, 2016.
- [11] FRONTIN, S.O. ET AL., "Equipamentos de Alta Tensão: Prospecção e Hierarquização de Inovações Tecnológicas", 1 ed. Brasília: Teixeira,. v. 2000, 934p, 2013.
- [12] CAVALLINI, A., MONTANAR, G.C., PULETTI, F., CONTIN, A., "A New Methodology for Identification of PD in Electrical Apparatus: Properties and Applications", IEEE Trans. Dielectr. Electr. Insul., Vol. 12, pp. 203-215, 2005.
- [13] AMORIM JÚNIOR, H.P., CARVALHO, A.T., SENS, M.A., LEVY, A.F.S., "Curso de Descargas Parciais em Equipamentos – Fundamentos, Técnicas de Medição e Diagnóstico", Curso ministrado pelo Cepel – material contido em apostila impressa, Rio de Janeiro, 2016.
- [14] INTERNATIONAL ELETROTECHNICAL COMMISSION, "IEC 60270 –High-voltage test techniques –Partial discharge measurements", Suíça, 2000.
- [15] ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, "ABNT NBR6940 –Técnicas de ensaios elétricos de alta tensão"– Medição de descargas parciais, Rio de Janeiro, 2015.
- [16] SAHOO, N.C., SALAMA, M.A., BARTNIKAS, R., "Trends in Partial Discharge Pattern Classification: A Survey", IEEE Transactions on Dielectrics and Electrical Insulation, v.12, n. 2, pp. 248-264, 2005.
- [17] MAZROUA, A.A., BARTNIKAS, R., SALAMA, M.M., "Discrimination Between PD Pulse Shapes Using Different Neural Network Paradigms", IEEE Transactions on Electrical Insulation, v. 1, n. 6, pp. 1119-1131, 1994.
- [18] JAMES, R.E., PHUNG, B.T., "Development of Computer-based Measurements and their Application to PD Pattern Analysis", IEEE Transactions on Electrical Insulation, v. 2, pp. 838-856, 1995.
- [19] CACCIARI, M., CONTIN, A., RABACH, G., MONANARI, C.G., "An

- Approach to Partial Discharge Investigation by Height Distribution Analysis”, IEEE Proc.: Sci., Meas. Technol., v.42, pp. 102-108, 1995.
- [20] SALAMA, M.M., BARTINIKAS, R., “Determination of Neural Network Topology for Partial Discharge Pulse Pattern Recognition”, IEEE Transactions Neural Networks, v. 13, pp. 446-456, 2002.
- [21] CONTIN, A., CAVALLINI, A., MONTANARI, G.C., PASINI, G., PULETTI, F., “Digital detection and fuzzy classification of partial discharge signals”, IEEE Trans. Dielectr. Electr. Insul. 9, 335–348, 2002.
- [22] VENKATESH, S., GOPAL, S., “Orthogonal least square center selection technique – A robust scheme for multiple source Partial Discharge pattern recognition using Radial Basis Probabilistic Neural Network”, Expert Syst. Appl. 38, 8978–8989, 2011.
- [23] CONTIN, A., "Diagnostics of Insulation Systems by Means of Partial Discharges", IEEE 10th International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives (SDEMPED), 191-197, 2015.
- [24] YU, H., SONG, Y.H., “Using improved self-organizing map for partial discharge diagnosis of large turbogenerators”, IEEE Trans. Energy Convers. 18, 392–399, 2003.
- [25] KRANZ, H.G., “Diagnosis of partial discharge signals using neural networks and minimum distance classification”, IEEE Trans. Electr. Insul. 28, 1016–1024, 1993.
- [26] GULSKI, E., “Computer-aided measurement of partial discharges in HV equipment”, IEEE Trans. Electr. Insul. 28, 969–983, 1993.
- [27] KREUGER, F.H., GULSKI, E., KRIVDA, A., “Classification of partial discharges”, IEEE Trans. Electr. Insul. 28, 917–931, 1993.
- [28] HUCKER, T., KRANZ, G.H., “Requirements of Automated PD Diagnosis System for Fault Identification in Noisy Conditions”, IEEE Transactions on Electrical Insulation, v. 2, pp. 544-556, 1995.
- [29] CHUI, C.K., “An Introduction to Wavelets”, Academic Press, Inc., 1992.
- [30] KIM, C.H., AGGARWAL, R., “Wavelet Transforms in Power Systems:

- Part 1 – General Introduction to the Wavelet Transforms”, *Power Engineering Journal*, 14, 2, 81-87, April, 2000.
- [31] CARMINATI, E., CRISTALDI, L. LAZZARONI, M., MONTI, A., “A Neurofuzzy Approach for the Detection of Partial Discharge”, *IEEE Trans. Instrum. Meas.*, v.50, pp. 1413-1417, 2001.
- [32] JANG, J.K., KIM, S.H. LEE, Y.S., “Classification of Partial Discharge Electrical Signals Using Wavelet Transforms”, *IEEE 13th Intern. Conf. Dielectr. Liquids*, Nara, Japan, pp. 552-555, 1999.
- [33] RAHMAN, M.K.A., ARORA, R. SRIVASTAVA, S.C., “Partial Discharge Classification Using Principal Component Transformation”, *IEEE Proc., Part A*, v. 147, pp. 7-13, 2000.
- [34] HARALIK, R.M., SHAMUGAN, K., DINSTEN, I., “Textural Features for Image Classification”, *IEEE Trans. Syst., Man and Cybernetics*, SMC-3, pp. 610-621, 1973.
- [35] CONNERS, R.W., HARLOW, C.A., “A Theoretical Comparison of Texture Algorithms”, *IEEE Trans. Pattern Analysis and Machine Intelligence*, PAMI-2, pp. 204- 222, 1980.
- [36] GALLOWAY, M.M., “Texture Analysis Using Gray Level Run Lengths”, *Computer Vision, Graphics and Image Processing*, v. 4, pp. 172-179, 1975.
- [37] PRATT, W.K., “Digital Signal Processing”, Wiley, 1978.
- [38] SATISH, L., ZAENGL, W.S., “Can Fractal Features be Used for Recognizing 3-D Partial Discharge Patterns?”, *IEEE Transactions on Electrical Insulation*, v. 2, pp. 352-359, 1995.
- [39] KRIVDA, A., GULSKI, E., SATISH, L., ZAENGL, W.S., “The use of fractal features for recognition of 3-D discharge patterns”, *IEEE Trans. Dielectr. Electr. Insul.* 2, 889–892, 1995.
- [40] KUNDU, P., KISHORE, N.K., SINHA, A.K., “Identification of two simultaneous partial discharge sources in an oil-pressboard insulation system using acoustic emission techniques”, *Appl. Acoust.* 73 (4), 395–401, 2012.
- [41] LALITHA, E.M., SATISH, L., “Fractal Image Compression for Classification of PD Sources”, *IEEE Transactions on Electrical*

- Insulation, v. 5, pp. 550-557, 1998.
- [42] ZHUORUI, J., YOUYUAN, W., JIAXIN, N., JIAN, L., XUESHONG, W., “Recognition of UHF PD signals in transformers based on wavelet and fractal theory”, in: *Electrical Insulation and Dielectric Phenomena, 2007. CEIDP 2007. Annual Report – Conference on, 2007*, pp. 192–195, 2007.
- [43] JIAN, L., CAIXIN, S., GRZYBOWSKI, S., “Partial discharge image recognition influenced by fractal image compression”, *IEEE Trans. Dielectr. Electr. Insul.* 15, 496–504, 2008.
- [44] CHEN, H.C., “Fractal features-based pattern recognition of partial discharge in XLPE power cables using extension method”, *Gener., Transm. Distrib., IET* 6, 1096–1103, 2012.
- [45] CHEN, H.C., GU, F.C., “Pattern recognition with cerebellar model articulation controller and fractal features on partial discharges”, *Expert Syst. Appl.* 39, 6575–6584, 2012.
- [46] GU, F.C., CHANG, H.C., CHEN, F.H., KUO, C.C., “Partial discharge pattern recognition of power cable joints using extension method with fractal feature enhancement”, *Expert Syst. Appl.* 39, 2804–2812, 2012.
- [47] JIAN, L., CAIXIN, S., GRZYBOWSKI, S., TAYLOR, C.D., “Partial discharge image recognition using a new group of features”, *IEEE Trans. Dielectr. Electr. Insul.* 13, 1245–1253, 2006.
- [48] LALITHA, E.M., SATISH, L., “Wavelet Analysis for Classification of Multisource PD Patterns”, *IEEE Transactions on Electrical Insulation*, v. 7, pp. 40-47, 2000.
- [49] YOUNG, R.K., “Wavelet Theory and Applications”, Kluwer Academic Publisher, 1993.
- [50] CONTIN, A., CAVALLINI, A., “Identification of Defects Supporting PD in Rotating Machines with Different Insulation Technologies”, 2018 IEEE 2nd International Conference on Dielectrics (ICD), Budapest, Hungria, Julho de 2018.
- [51] MORSHUIS, P., “Assessment of Dielectric Degradation by Ultrawide-Band PD Detection”, *IEEE Trans. Dielectr. Electr. Insul.*, Vol. 2, pp. 744-760, 1995.

- [52] HUDON, C., BARTNIKAS, R., WERTHEIMER, M.R., "Spark-to-Glow Discharge Transition due to Increased Surface Conductivity on Epoxy Resin Specimens", IEEE Trans. Electr. Insul., Vol. 28, pp. 1-8, 1993.
- [53] MEHDI, A., ASGHAR, A., "A method for discriminating original pulses in online partial discharge measurement", Measurement: Journal of the International Measurement Confederation, vol. 44, n°1, pp.148-158, 2011.
- [54] LIN, Y.H., "Using K-Means Clustering and Parameter Weighting for PD Noise Suppression", IEEE Trans. on Pow. Del., V.26, pp.2380-2390, 2011.
- [55] KOLTUNOWICZ, W., PLATH, R., "Synchronous Multi-Channel PD Measurements", IEEE Trans. Dielectr. Electr. Insul., Vol. 15, pp. 1715-1723, 2008.
- [56] BAKER, P. C., JUDD, M. D., MCARTHUR, S. D. J., "A Frequency-Based RF PD Detector for Low-Power Wireless Sensing", IEEE Trans. Dielectr. Electr. Insul., Vol. 17, pp. 133-140, 2010.
- [57] HAO, L., CONTIN, A., HUNTER, J.A., LEWIN, L., SWAFFIELD, D.J., WALTON, C., "Discrimination of Multiple PD Sources Using Wavelet Decomposition and Principal Component Analysis", IEEE Trans. Dielectr. Electr. Insul., Vol.18, pp. 1702-1711, October 2011.
- [58] KULJACA, N., MEREGALLI, S., CONTIN, A., UKOVICH, A., "Separation of Multiple Sources in PD Measurements Using an Amplitude-Frequency Relation Diagram", 10th IEEE ICSD, pp.480-483, Potsdam, Alemanha, July 2010.
- [59] CONTIN, A., PASTORE, S., "Classification and Separation of Partial Discharge Signals by Means of their Auto-Correlation Function Evaluation", IEEE Trans. Dielectr. Electr. Insul., Vol. 16, pp. 1609-1622, 2009.
- [60] MAZROUA, A.A., SALAMA, M.M.A., BARTNIKAS, R., "PD Patterns Recognition with Neural Networks Using the Multilayer Perceptron Technique", IEEE Transactions on Electrical Insulation, v. 28, pp. 1082-1089, 1993.

- [61] SU, M.S., CHEN, J.F., LIN, Y.H., “Phase determination of partial discharge source in three-phase transmission lines using discrete wavelet transform and probabilistic neural networks”, *Int. J. Electr. Power Energy Syst.* 51 (10), 27–34, 2013.
- [62] GULSKI, E., KRIVDA, A., “Neural Networks as a Tool for Recognition of Partial Discharges”, *IEEE Transactions on Electrical Insulation*, v. 28, pp. 984-1001, 100. 1993.
- [63] TOU, J.T., GONZALES, R.C., “Pattern Recognition Principles”, Addison-Wesley, 1974.
- [64] HAYKIN, S., “Neural Network: A Comprehensive Foundation”, 2da. Edition, ISBN 0-02-352761-7 Prentice Hall, 1999.
- [65] WANG, M.H, “Partial Discharge Pattern Recognition of Current Transformers Using an ENN”, *Transactions on Power Delivery*, v. 20, pp. 1984-1990, 2005.
- [66] CUENCA, W.M., SEIXAS, J.M., LEVY, A.F.S., “Sistema Caracterizador Híbrido para Identificar Descargas Parciais em Transformadores de Potência”, *Proceedings of the V Brazilian Conference on Neural Networks*, pp. 169-174, PUC, Rio de Janeiro – RJ, Brazil, April 2-5, 2001.
- [67] JIN, J., CHANG, C.S., CHANG, C., HOSHINO, T., HANAI, M., KOBAYASHI, N., “Classification of partial discharge events in gas-insulated substations using wavelet packet transform and neural network approaches”, *Sci., Meas. Technol., IEE Proc.* 153, 55–63, 2006.
- [68] CONTIN, A., CAVALLINI, A., MONTANARI, G.C., PASINI, G., “Digital Detection and Fuzzy Classification of Partial Discharge Signals”, *IEEE Transactions on Electrical Insulation*, v. 9, pp. 335-348, 1993.
- [69] SALAMA, M.M.A., BARTNIKAS, R., “Fuzzy logic applied to PD pattern classification”, *IEEE Trans. Dielectr. Electr. Insul.* 7, 118–123, 2000.
- [70] MAZZETTI, C., MASCIOLI, F.M.F., BALDINI, F., PANELLA, M., RISICA, R., BARTNIKAS, R., “Partial discharge pattern recognition by neuro-fuzzy networks in heat-shrinkable joints and terminations of

- XLPE insulated distribution cables”, IEEE Trans. Power Delivery 21, 1035–1044, 2006.
- [71] SIMPSON, P.K., “Fuzzy min–max neural networks – Part 2: Clustering”, IEEE Trans. Fuzzy Syst. 1, 32, 1993.
- [72] RIZZI, A., MASCIOLI, F.M.F., BALDINI, F., MAZZETTI, C., BARTNIKAS, R., “Genetic optimization of a PD diagnostic system for cable accessories”, IEEE Trans. Power Delivery 24, 1728–1738, 2009.
- [73] RIZZI, A., PANELLA, M., MASCIOLI, F.M.F., “Adaptive resolution min–max classifiers”, IEEE Trans. Neural Networks 13, 402–414, 2002.
- [74] MOTA, H.O., ROCHA, L.C.D., SALLES, T.C.M., VASCONCELOS, F.H., “Partial discharge signal denoising with spatially adaptive wavelet thresholding and support vector machines”, Electric Power Syst. Res. 81, 644–659, 2011.
- [75] HAO, L., LEWIN, P.L., “Partial discharge source discrimination using a support vector machine”, IEEE Trans. Dielectr. Electr. Insul. 17, 189–197, 2010.
- [76] SHARKAWY, R.M., MANGOUBI, R.S., ABDEL-GALIL, T., SALAMA, M.M.A., BARTNIKAS, R., “SVM classification of contaminating particles in liquid dielectrics using higher order statistics of electrical and acoustic PD measurements”, IEEE Trans. Dielectr. Electr. Insul. 14, 669–678, 2007.
- [77] IBRAHIM, K., SHARKAWY, R.M., SALAMA, M.M.A., BARTNIKAS, R., “Realization of partial discharge signals in transformer oils utilizing advanced computational techniques”, IEEE Trans. Dielectr. Electr. Insul. 19, 1971–1981, 2012.
- [78] HUNTER, J.A., LEWIN, P.L., HAO, L., WALTON, C., MICHEL, M., “Autonomous classification of PD sources within three-phase 11 kV PILC cables”, IEEE Trans. Dielectr. Electr. Insul. 20, 2117–2124, 2013.
- [79] RODRIGUES, T.B., AMORIM JÚNIOR, H.P., CARVALHO, A.T., FLEMING, C.C., BATISTA, J.B., "Experience with On-Line Insulation Diagnostics of Surge Arresters by Partial Discharges Measurement in the Field", Colloquium, Montreal - Canadá, 2016.
- [80] CONTIN, A., PASTORE, S., PAGANIN, R., "Evaluation of Spaces for

- the Separation of Signals Due to Multiple PD Sources", Electrical Insulation Conference (EIC), Seattle - USA, p.p. 209-213, 2015.
- [81] CONTIN, A., PICCOLO, A., "PD Acceptance Tests for New Bars and Coils", Electrical Insulation Conference (EIC), Montréal - Canada, 2016.
- [82] CONTIN, A., AIZZA, M., FANTIN, F., PICCOLO, A., "Identification of PD Sources and Insulation Technologies in Rotating Machines", Electrical Insulation Conference (EIC), Seattle - USA, 2015.
- [83] HAN, J., GE, B., LI, W., "Numerical calculation of CHTC on end metal parts and flow in end region of a turbogenerator", IET Science, Measurement & Technology, Vol. 12, Iss. 2, pp. 193-199, 2018.
- [84] CARVALHO, A.T., AMORIM JÚNIOR, H.P., FLEMING, C.C., RODRIGUES, T.B., BRASIL, F.S., VILHENA, P.R.M., CARVALHO, D.S., "Virtual Instrumentation for Partial Discharge Monitoring". 2017 IEEE Electrical Insulation Conference (EIC), Baltimore, MD, USA, Junho de 2017.
- [85] BACEGA, W.R., TATIZAWA, H., KANASHIRO, A.G., "Técnica de Identificação do Processo de Degradação de Para-raios de ZnO em Campo", XX SNPTEE - Seminário Nacional de Produção e Transmissão de Energia Elétrica, Recife, 2009.
- [86] AMORIM JÚNIOR, H.P., CARVALHO, A.T., RODRIGUES, T.B., RODRIGUES, J.A.P., OLIVEIRA, H.B., DIAS, J.B., "Avaliação de Para-raios no Campo e no Laboratório por meio das Descargas Parciais utilizando TC de Alta Frequência", IV Simpósio Brasileiro de Sistemas Elétricos (SBSE), Goiânia - Brasil, 2012.
- [87] AMORIM JÚNIOR, H.P., CARVALHO, A.T., RODRIGUES, T.B., RODRIGUES, J.A.P., OLIVEIRA, H.B., DIAS, J.B., "Evaluation of Surge Arresters in the field and in the laboratory by means of Partial Discharges using High Frequency Current Transformer", In: 2012 IEEE International Symposium on Electrical Insulation (ISEI), San Juan – Porto Rico, 2012.
- [88] AMORIM JÚNIOR, H.P., CARVALHO, A.T., RODRIGUES, T.B., BATISTA, J.B., FLEMING, C.C., "Evaluation of Surge Arresters in the

- Field and in the Laboratory by Means of Partial Discharges", In: ICPADM 2015, Sidney – Australia, 2015.
- [89] KIRA, K., RENDELL, L., "A practical approach to feature selection", In D. Sleeman, & P. Edwards (Eds.), *Proceedings of the Ninth International Workshop on Machine Learning* (pp. 249–258), Aberdeen - Scotland, 1992.
- [90] WARD, J.H., "Hierarchical grouping to optimize an objective function", *Journal of the American Statistical Association*, 58, 236–244, 1963.
- [91] MURTAGH, F., LEGENDRE, P., "Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion?" *Journal of Classification*, 31, 274–295, 2014.
- [92] BUSSAB, W., MIAZAKI, E., ANDRADE, D.F. "Introdução à análise de agrupamentos". Associação Brasileira de Estatística. ABE 9º Simpósio Nacional de Probabilidade e Estatística, São Paulo – SP, julho de 1990.
- [93] QUINLAN, R., "C4.5: Programs for Machine Learning", San Mateo - CA, Estados Unidos, 1993.
- [94] BOSER, B. E., GUYON, I.M., VAPNIK, V.N., "A Training Algorithm for Optimal Margin Classifiers". In: *Annual Workshop on Computational Learning*, 5, Pittsburgh. ACM Press. Pittsburgh: Haussler D, p.144-152, julho de 1992.
- [95] DRUCKER, H., BURGESS, C.J., KAUFMAN, L., SMOLA, A., VAPNIK, V., "Support vector regression machines. Advances in neural information processing systems", Morgan Kaufmann Publishers, p. 155–161, 1997.
- [96] SARADHI, V., KAMIK, H., MITRA, P., "A Decomposition Method for Support Vector Clustering". In *Proc. of the 2nd International Conference on Intelligent Sensing and Information Processing (ICISIP)*, p. 268-271, 2005.
- [97] INTERNATIONAL ELETROTECHNICAL COMMISSION, "IEC 60034-27-2, Rotating electrical machines – Part 27-2: On-line partial discharge measurements on the stator winding insulation of rotating electrical machines", Suíça, 2012.
- [98] RODRIGUES, T.B., TANSCHKEIT, R., VELLASCO, M.M.B.R.,

- AMORIM JUNIOR, H.P., "Avaliação Automática de Para-raios e Transformadores de Corrente no Campo Quanto à Presença de Descargas Parciais Internas". In: XVIII Eriac - Décimo Oitavo Encontro Regional Ibero-Americano de Cigré, Foz do Iguaçu - PR, Maio de 2019.
- [99] RAYMOND, W.J.K., ILLIAS, H.A., BAKAR, A.H.A., MOKHLIS, H., "Partial discharge classifications: Review of recent progress. Measurement Elsevier", 68 (164), p. 181, 2015.
- [100] AHA, D.W., KIBLER, D., ALBERT, M.K., "Instance-Based Learning Algorithms". Machine Learning, 6, 37-66. 1991.
- [101] CALLAN, J.P., FAWCETT, T.E., RISSLAND, E.L., "CABOT : An Adaptive Approach to Case-Based Search". Proceedings of the Twelfth International Joint Conference on Artificial Intelligence, 803-808. 1991.
- [102] VASCONCELLOS, E.C., "Árvores de Decisão Aplicadas ao Problema de Separação Estrela/Galáxia". Dissertação de Mestrado do Curso de Pós-Graduação em Computação Aplicada – INPE, São José dos Campos, SP, 2011.
- [103] LORENZETT, C.D.C., TELÖCKEN, A.V., "Estudo Comparativo entre os algoritmos de Mineração de Dados Random Forest e J48 na tomada de Decisão". In: II Simpósio de Pesquisa e Desenvolvimento em Computação - Universidade de Cruz Alta, Cruz Alta, RS. 2016.
- [104] LIBRELOTTO, S.R., "Análise dos Algoritmos de mineração J48 e Apriori Aplicados na Detecção de Indicadores da Qualidade de Vida e Saúde". Universidade de Cruz Alta, Cruz Alta, RS. 2014.
- [105] BRAGA, A., CARVALHO, A.C.P.L.F., LUDERMIR, T.B., "Redes Neurais Artificiais: Teoria e Aplicações". Editora LTC, 2000.
- [106] BURGESS, C.J.C., "A tutorial on support vector machines for pattern recognition. Knowledge Discovery and Data Mining", 2(2):1-43, 1998.
- [107] CAMPBELL, C., "An introduction to kernel methods". In R. J. Howlett and L. C. Jain, editors, Radial Basis Function Networks: Design and Applications, pages 155–192, Berlin, 2000.
- [108] CHANG, C.C., LIN, C.J. "LIBSVM: a library for support vector machines". Disponível em: <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>, 2004.

- [109] HEARST, M.A., SCHÖLKOPF, B., DUMAIS, S., OSUNA, E., PLATT, J., "Trends and controversies - support vector machines". IEEE Intelligent Systems, 13(4):18–28, 1998.
- [110] HERBRICH, R., "Learning Kernel Classifiers: Theory and Algorithms". MIT Press, 2001.
- [111] JOACHIMS, T., "Learning to classify texts using support vector machines: methods, theory and algorithms". Kluwer Academic Publishers, 2002.
- [112] KIM, K.I., JUNG, K., PARK, S.H., KIM, H.J., "Support vector machines for texture classification". IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(11):1542–1550, 2002.
- [113] LORENA, A.C., CARVALHO, A.C.P.L.F., "Classificadores de margens largas (Large Margin Classifiers)". Technical Report 195, Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2003.
- [114] LORENA, A.C., CARVALHO, A.C.P.L.F., "Introdução às máquinas de vetores suporte (Support Vector Machines)". Technical Report 192, Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos. Disponível em: ftp://ftp.icmc.usp.br/pub/BIBLIOTECA/rel_tec/RT_192.ps.zip, Abril 2003.
- [115] MERCER, J., "Functions of positive and negative type and their connection with the theory of integral equations". Philosophical Transactions of the Royal Society, A 209:415–446, 1909.
- [116] MITCHELL, T., "Machine Learning". McGraw Hill, 1997.
- [117] MÜLLER, K.R., MIKA, S., RÄTSCH, G., TSUDA, K., SCHÖLKOPF, B., "An introduction to kernel-based learning algorithms". IEEE Transactions on Neural Networks, 12(2):181-201, Março 2001.
- [118] NOBLE, W.S., "Support vector machine applications in computational biology". In B. Schölkopf, K. Tsuda, and J.-P. Vert, editors, Kernel Methods in computational biology, pages 71–92. MIT Press, 2004.
- [119] PASSERINI, A., "Kernel Methods, multiclass classification and applications to computational molecular biology". PhD thesis,

Università Degli Studi di Firenze, 2004.

- [120] PONTIL, M., VERRI, A., "Support vector machines for 3-D object recognition". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(6):637–646, 1998.
- [121] SCHÖLKOPF, B., GUYON, I., WESTON, J., "Statistical learning and kernel methods in bioinformatics". In P. Frasconi and R. Shamir, editors, *Artificial Intelligence and Heuristic Methods in Bioinformatics*, pages 1–21. IOS Press, 2003.
- [122] SMOLA, A.J., BARTLETT, P., SCHÖLKOPF, B., SCHUURMANS, D., "Introduction to large margin classifiers". In A. J. Smola, P. Barlett, B. Schölkopf, and D. Schuurmans, editors, *Advances in Large Margin Classifiers*, pages 1–28. MIT Press, 1999.
- [123] SMOLA, A.J., SCHÖLKOPF, B., "Learning with Kernels". The MIT Press, Cambridge, MA, 2002.
- [124] VAPNIK, V.N., "The nature of Statistical learning theory". Springer-Verlag, New York, 1995.
- [125] VAPNIK, V.N., CHERVONENKIS, A.Y., "On the uniform convergence of relative frequencies of events to their probabilities". *Theory of Probability and its Applications*, 16(2):283–305, 1971.

Anexo A - Relief

A representação de dados brutos geralmente utiliza muitos atributos, sendo que apenas alguns destes são relevantes na determinação da saída desejada. Como atributos relevantes são frequentemente desconhecidos em problemas do mundo real, é necessário avaliar vários atributos possíveis. Infelizmente, atributos redundantes degradam o desempenho da aprendizagem dos conceitos inerentes ao problema em análise, tanto em velocidade (devido à alta dimensionalidade) quanto em precisão preditiva (devido a informações irrelevantes). A seleção de atributos é o problema de escolher um pequeno subconjunto destes idealmente necessário e suficiente para descrever o conceito do problema em análise [89].

Para diversos problemas do mundo real, que possivelmente envolvem muita interação de atributos, é necessário um método confiável e eficiente para eliminar atributos irrelevantes. Foi com este objetivo que o algoritmo do Relief foi desenvolvido. Ele seleciona atributos relevantes utilizando um método estatístico (não depende de heurísticas), sendo preciso mesmo quando os atributos guardam relação entre si, e também tolerante a ruídos [89].

O Relief é um algoritmo que atribui pesos aos atributos levando em consideração o aprendizado baseado em instância [100][101]. Dado um conjunto de treinamento S , com tamanho de amostra m , e um limiar de relevância τ ($0 \leq \tau \leq 1$), o Relief detecta os atributos que são estatisticamente relevantes na determinação da saída. Ele permite que a escala de cada atributo seja nominal (incluindo lógica) ou numérica (inteira ou real). As diferenças de valores dos atributos entre duas instâncias X e Y são definidas pela seguinte função *diff*:

Quando x_k e y_k são nominais:

$$diff(x_k, y_k) = \begin{cases} 0, & \text{se } x_k \text{ e } y_k \text{ são o mesmo} \\ 1, & \text{se } x_k \text{ e } y_k \text{ são diferentes} \end{cases} \quad \text{Eq.A.1}$$

Quando x_k e y_k são numéricas:

$$diff(x_k, y_k) = (x_k - y_k) / nu_k \quad \text{Eq.A.2}$$

onde nu_k é uma unidade de normalização para colocar os valores de $diff$ no intervalo $[0,1]$.

O Relief (cujo algoritmo é apresentado na Figura A.1) seleciona uma amostra composta por m triplicatas de uma instância X , sendo estas a própria instância X e mais as suas instâncias "Near-hit" e "Near-miss" (chama-se uma instância de "Near-hit" de X se esta pertencer à vizinhança próxima de X e também à mesma categoria de X ; chama-se uma instância de "Near-miss" quando esta pertencer à vizinhança próxima de X , mas não à mesma categoria de X). O Relief usa a distância Euclidiana p -dimensional (onde " p " é o número total de atributos) para selecionar "Near-hit" e "Near-miss". Ele chama uma rotina para atualizar o vetor peso W do atributo para cada triplicata de amostra e determinar a relevância média do vetor peso do atributo (de todos os atributos em relação à saída) [89]. Por fim, o Relief seleciona os atributos cujos pesos médios (nível de relevância) estejam acima do limite τ .

Rotina: Relief (S, m, τ);
 Separar S em $S^+ = \{\text{instâncias positivas}\}$ e $S^- = \{\text{instâncias negativas}\}$;
 $W = (0, 0, 0, \dots, 0)$;
 Para $i = 1$ a m ;
 Escolher aleatoriamente uma instância $X \in S$;
 Escolher aleatoriamente uma das instâncias positivas mais próximas de X , $Z^+ \in S^+$;
 Escolher aleatoriamente uma das instâncias negativas mais próximas de X , $Z^- \in S^-$;
 Se (X é uma instância positiva);
 Então Near-hit = Z^+ ; Near-miss = Z^- ;
 Senão Near-hit = Z^- ; Near-miss = Z^+ ;
 Atualizar_Pesos (W , X , Near-hit, Near-miss);
 Relevância = $(1/m) * W$;
 Para $i = 1$ a p ;
 Se ($\text{Relevância}_i \geq \tau$)
 Então f_i é um atributo relevante;
 Senão f_i é um atributo irrelevante;

Rotina: Atualizar_Pesos (W, X, Near-hit, Near-miss);
 Para $i = 1$ a p ;
 $W_i = W_i - \text{diff}(x_i, \text{Near-hit}_i)^2 + \text{diff}(x_i, \text{Near-miss}_i)^2$

Figura A.1 – Algoritmo do Relief [89]

Anexo B - Dendrogramas

A análise dos agrupamentos para formação dos dendrogramas baseia-se em uma função de distância para avaliar a semelhança entre as amostras do banco de dados e em um algoritmo de cálculo para realizar a tarefa de construir a hierarquia de partições característica dos dendrogramas. A configuração que se mostrou mais adequada aqui, após alguns testes, (avaliando a qualidade pela maior largura de faixa de nível de dissimilaridade que apontava o número correto de *clusters*) foi a que utilizou a distância Euclidiana como métrica atrelada ao algoritmo de cálculo baseado em variância mínima proposto por Ward (método de Ward) [90].

Para melhor compreensão do processo de formação dos dendrogramas, pode-se utilizar o exemplo de agrupar indivíduos conforme altura, massa e idade, segundo dados apresentados na tabela B.1 [92]. Supõe-se que, utilizando o método de Ward e a métrica de distância Euclidiana, foram obtidos os níveis de junção apresentados na tabela B.2, com o dendrograma equivalente a esse processo sendo o mostrado na Figura B.1.

Tabela B.1 - Dados dos indivíduos

Indivíduo	Altura (cm)	Massa (kg)	Idade (anos)
A	180	79	15
B	175	75	28
C	170	70	50
D	167	63	25
E	180	71	80
F	165	60	31

Pelo exemplo apresentado, percebe-se que o dendrograma pode ser utilizado para visualizar como os agrupamentos são formados em cada passo e para avaliar os níveis de similaridade (ou distância) dos agrupamentos formados. O padrão de como a similaridade ou os valores de distância mudam de um passo para outro pode ajudar a escolher o agrupamento final dos dados. O passo onde os valores mudam abruptamente também pode identificar um bom ponto para definir o agrupamento final [92].

Tabela B.2 - Níveis de junção

Passo	Junção	Nível (valor relativo*)	Nível (valor relativo normalizado**)
1	D,F	0,37	0,23
2	A,B	0,67	0,41
3	AB,E	0,73	0,45
4	C,DF	0,95	0,58
5	ABE,CDF	1,64	1

Obs: * Imaginando-se uma medida de distância qualquer (cm, por exemplo) sendo ajustada para valores entre 0 e 1 antes do cálculo da distância Euclidiana, os valores aqui seriam adimensionais (valores relativos). **Dividindo-se o valor relativo dos níveis pelo maior deles (valor do último passo), tem-se os níveis com valores relativos normalizados.

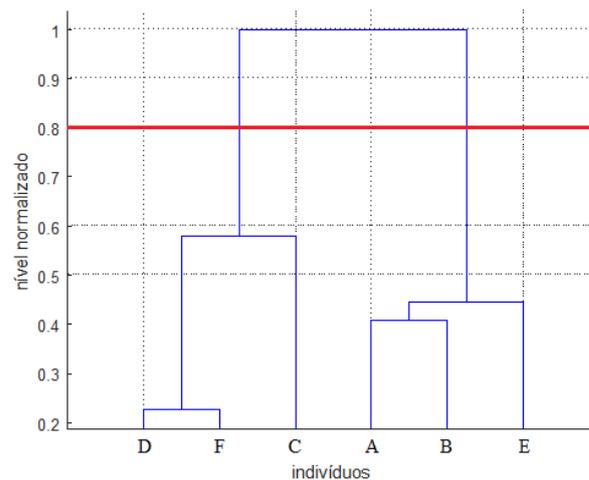


Figura B.1 - Dendrograma formado após processamento dos dados

A decisão sobre o agrupamento final também é chamada corte do dendrograma. O corte do dendrograma pode ser representado por uma linha horizontal, determinando o agrupamento final (quantidade de agrupamentos) pelo número de intersecções desta com as linhas verticais do dendrograma. Pode-se comparar os dendrogramas a agrupamentos finais diferentes para determinar qual agrupamento final faz mais sentido para os dados em análise. No caso do exemplo apresentado, pode-se escolher o nível normalizado 0,8 para se estabelecer o agrupamento final, originando-se assim dois *clusters*, um deles formado pelos indivíduos ABE e outro pelos indivíduos CDF [92].

Anexo C - Árvores de Decisão

Uma árvore de decisão consiste em um conjunto de regras geradas a partir de uma série de exemplos de treinamento para classificar dados que sejam definidos pelos mesmos atributos dos exemplos. Em uma árvore de decisão as regras de classificação são estruturadas na forma de nós e ramos. Cada nó abriga um determinado teste sobre um ou mais atributos, e cada ramo os possíveis resultados para o teste. O nó contendo o primeiro teste é denominado de raiz, e os nós finais, dos quais não saem ramos, são chamados folhas. As folhas, em geral, não contêm um teste, mas sim uma classificação. Cada caminho que começa na raiz e termina em uma folha pode ser definido como uma regra de classificação [102].

A ideia básica na qual se baseia a construção dessas árvores é resolver um problema complexo dividindo-o em uma série de subproblemas de complexidade menor. É um processo recursivo no qual o conjunto de exemplos é dividido até que seja fácil classificar os exemplos contidos nos subconjuntos gerados. A Tabela C.1 mostra um conjunto simples de exemplos de treinamento [102]. Os dados exibidos relacionam condições meteorológicas com duas possíveis decisões: sair ou não de casa. Neste contexto, as condições meteorológicas tempo, temperatura, umidade e vento são os atributos, ou seja, valores que devem ser analisados para se tomar a decisão. Já as decisões "sair" ou "não sair" são as classes possíveis.

O objetivo ao analisar os dados da Tabela C.1 é aprender como decidir sair ou não de casa com base nos atributos meteorológicos contidos na tabela. Em outras palavras, dada uma combinação desses atributos que não esteja representada na tabela, surge a dúvida de como decidir sobre sair ou não de casa. Para responder essa questão é que se constrói a árvore de decisão. Os algoritmos de construção irão gerar um conjunto de testes sobre os atributos em um formato de nós e ramos, e os resultados destes testes darão a resposta desejada.

O conjunto de dados mostrado na Tabela C.1 é muito interessante do ponto de vista acadêmico, pois ele contém atributos nominais e numéricos. Os atributos Tempo, Vento e a classe Decisão são chamados atributos nominais. Esses atributos possuem uma quantidade limitada e previamente conhecida de valores.

No caso do atributo Tempo, as possibilidades são sol, nublado ou chuva. Neste caso, um teste sobre o atributo Tempo pode ter no máximo três resultados, como mostrado na Figura C.1(a). Já os atributos ditos numéricos são assumidos como podendo apresentar qualquer valor real. Por exemplo, o atributo Temperatura pode ter qualquer valor x desde que $x \in \mathbb{R}$, sendo $x \geq 0$. Os testes sobre atributos numéricos geralmente fornecem resultados binários, como os da Figura C.1(b), mas podem também gerar resultados ternários como o mostrado na Figura C.1(c). A classe do exemplo é sempre um atributo nominal, ou seja, o número possível de classes deve sempre ser limitado e bem definido [102].

Tabela C.1 - Pequeno conjunto de exemplos de treinamento [93]

Tempo	Temperatura (°C)	Umidade (%)	Vento?	Decisão
sol	23,9	70	sim	sair
sol	26,7	90	sim	não sair
sol	29,4	85	não	não sair
sol	22,2	95	não	não sair
sol	20,6	70	não	sair
nublado	22,2	87	sim	sair
nublado	28,3	78	não	sair
nublado	17,8	65	sim	sair
nublado	27,2	75	não	sair
chuva	21,7	80	sim	não sair
chuva	18,3	70	sim	não sair
chuva	23,9	80	não	sair
chuva	20,0	80	não	sair
chuva	21,1	96	não	sair

A Tabela C.1 exibe um conjunto de dados relativamente pequeno, e mesmo assim não é fácil definir os critérios necessários para a tomada de decisão (classificação). Se já é difícil analisar um conjunto pequeno de dados como este, pode-se imaginar como seria praticamente impossível para um ser humano inferir critérios para um conjunto com um grande volume de dados. Os algoritmos de árvores de decisão foram criados para inferir critérios de classificação em grandes volumes de dados de forma que estes critérios fiquem claros para o usuário. Esta é

uma característica importante das árvores de decisão. Um desses algoritmos é o J48, criado inicialmente por Ross Quinlan em 1993 [93]. Trata-se de um algoritmo de mineração de dados voltado para classificação baseado em árvores de decisão. É caracterizado pela evolução de seus antecessores ID3, C4.5, C5.0, implementados na linguagem C, tendo o J48 a mesma lógica só que implementado em linguagem JAVA. Uma das vantagens da aplicação deste algoritmo na tomada de decisão é que ele se mostra adequado para procedimentos envolvendo variáveis contínuas e discretas presentes nas bases de dados, e permite a construção de árvores de decisão que classificam e apresentam em suas ramificações os atributos de maior relevância [103].

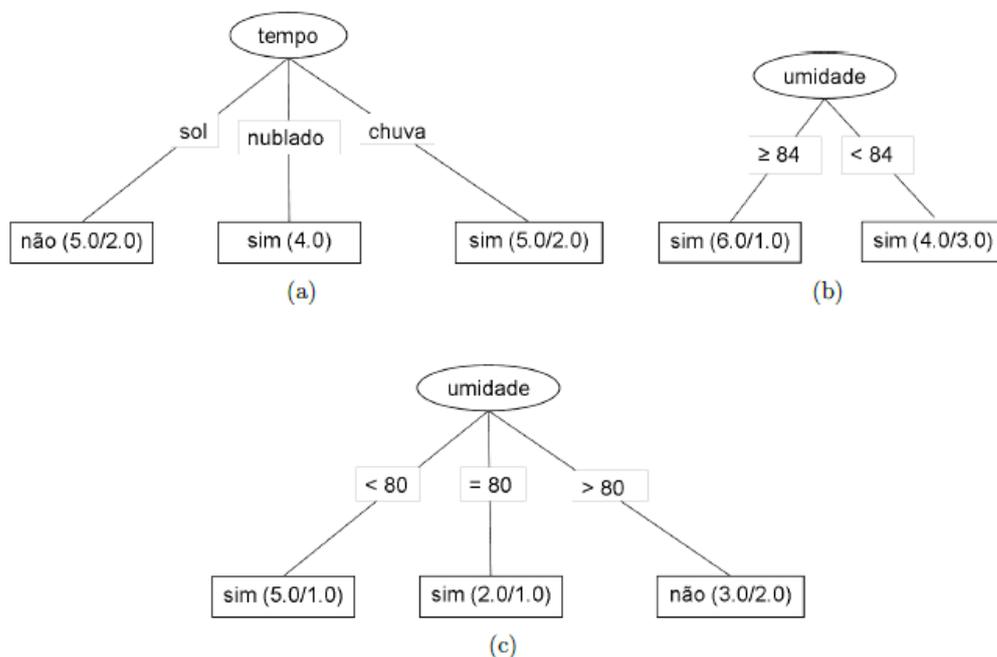


Figura C.1 – Exemplos de testes executados por uma árvore de decisão sobre um atributo nominal (a) e dois numéricos (b e c), sendo citado em cada folha: número de exemplos da classe indicada / número de exemplos da outra classe [102]

O J48 utiliza a taxa de ganho de informação (*information gain ratio*) para a construção da árvore de decisão. Esta taxa é caracterizada pelo uso de uma métrica para ranquear todos os atributos de uma base de dados, sendo calculada dividindo-se o ganho de informação (*Gain*) de um atributo pelo número de saídas que um teste com aquele atributo pode resultar (*SplitInfo*). A ideia é selecionar o teste que minimize a quantidade de informação necessária para a classificação de um objeto. Para a construção de um nó, a taxa de ganho de informação é calculada

para cada atributo e aquele que apresentar maior valor será utilizado no nó para dividir o conjunto de exemplos de treinamento. Se o subconjunto testado em um determinado nó só possuir exemplos da mesma classe ou se todos os exemplos nele contidos apresentam a mesma taxa de ganho de informação, então nenhum teste é proposto e uma folha é criada [102].

O algoritmo J48 tem a característica de testar um único atributo em cada nó, sendo por isso as árvores de decisão geradas por ele consideradas univariantes ou axis-parallel (eixo-paralelo). Este tipo de teste divide o espaço de atributos com um multiplano perpendicular ao eixo que representa o atributo em questão, ou seja, paralelo aos demais. Logo, a denominação axis-parallel refere-se a esta propriedade.

Após a criação da árvore, o J48 aplica técnicas de poda onde serão retiradas as “impurezas” da árvore deixando somente as informações importantes e que agregarão para a tomada de decisão. A poda deste método ocorre de maneira a comparar taxas de estimativa de erro de cada subárvore do nó folha gerado. São processados sucessivos testes a partir do nó raiz da árvore, de forma que, se a estimativa de erro indicar que a árvore será mais precisa se os nós descendentes (filhos) de um determinado nó “n” forem eliminados, então estes nós descendentes serão eliminados [104].

Anexo D - Máquinas de Vetor Suporte

As Máquinas de Vetor Suporte (*Support Vector Machines* - SVMs) constituem uma técnica de aprendizado que vem recebendo crescente atenção da comunidade de Aprendizado de Máquina (AM) [116]. Os resultados da aplicação dessa técnica são comparáveis e muitas vezes superiores aos obtidos por outros algoritmos de aprendizado, como as Redes Neurais Artificiais (RNAs) [105][64]. Exemplos de aplicações de sucesso podem ser encontrados em diversos domínios, como na categorização de textos [111], na análise de imagens [112][120] e em Bioinformática [118][121].

As SVMs são embasadas pela teoria de aprendizado estatístico, desenvolvida por Vapnik [124] a partir de estudos iniciados em [125]. Essa teoria estabelece uma série de princípios que devem ser seguidos na obtenção de classificadores com boa generalização, definida como a sua capacidade de prever corretamente a classe de novos dados do mesmo domínio em que o aprendizado ocorreu.

A seguir são descritos os conceitos básicos a respeito das SVMs para problemas de classificação e também o processo de determinação matemática do classificador, conforme pode ser encontrado em [114], apresentado aqui em uma versão mais resumida, os quais também podem ser consultados em [113].

D.1

A Teoria de Aprendizado Estatístico

Seja f um classificador e F o conjunto de todos os classificadores que um determinado algoritmo de AM pode gerar. Esse algoritmo, durante o processo de aprendizado, utiliza um conjunto de treinamento T , composto de n pares (x_i, y_i) , para gerar um classificador particular $f \in F$.

Considerando, por exemplo, o conjunto de treinamento da Figura D.1 [123]. O objetivo do processo de aprendizado é encontrar um classificador que separe os dados das classes “círculo” e “triângulo”. As funções ou hipóteses consideradas são ilustradas na figura por meio das bordas, também denominadas fronteiras de decisão, traçadas entre as classes.

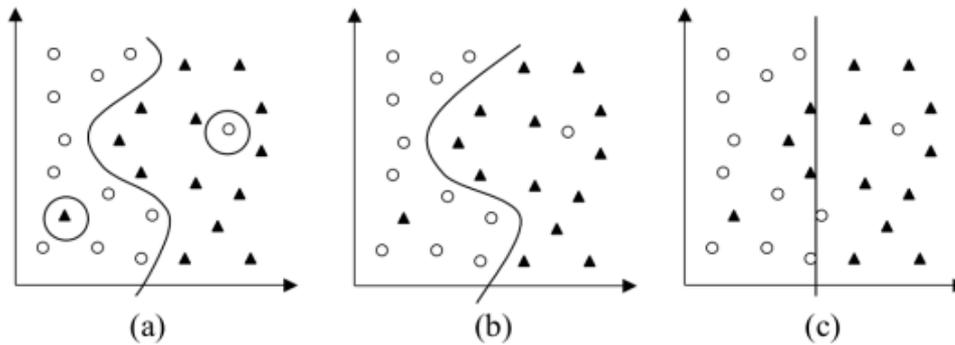


Figura D.1 – Conjunto de treinamento binário e três diferentes hipóteses [114]

Na imagem da Figura D.1(a), tem-se uma hipótese que classifica corretamente todos os exemplos do conjunto de treinamento, incluindo dois possíveis ruídos. Por ser muito específica para o conjunto de treinamento, essa função apresenta elevada suscetibilidade a cometer erros quando confrontada com novos dados. Esse caso representa a ocorrência de um superajustamento do modelo aos dados de treinamento.

Um outro classificador poderia desconsiderar pontos pertencentes a classes opostas que estejam muito próximos entre si. A ilustração da Figura D.1(c) representa essa alternativa. A nova hipótese considerada, porém, comete muitos erros, mesmo para casos que podem ser considerados simples. Tem-se assim a ocorrência de um sub-ajustamento, pois o classificador não é capaz de se ajustar mesmo aos exemplos de treinamento.

Um meio termo entre as duas funções descritas é representado na Figura D.1(b). Esse preditor tem complexidade intermediária e classifica corretamente grande parte dos dados, sem se fixar demasiadamente em qualquer ponto individual.

A Teoria de Aprendizado Estatístico (TAE) estabelece condições matemáticas que auxiliam na escolha de um classificador particular f a partir de um conjunto de dados de treinamento. Essas condições levam em conta o desempenho do classificador no conjunto de treinamento e a sua complexidade, com o objetivo de obter um bom desempenho também para novos dados do mesmo domínio.

D.2

SVMs Lineares

As SVMs surgiram pelo emprego direto dos resultados fornecidos pela teoria de aprendizado estatístico. É apresentado a seguir o uso de SVMs na obtenção de fronteiras lineares para a separação de dados pertencentes a duas classes, sendo a primeira formulação mais simples (SVMs com margens rígidas), para lidar com problemas linearmente separáveis, e a segunda formulação mais ampla (SVMs com margens suaves), para definir fronteiras lineares sobre conjuntos de dados mais gerais. Posteriormente descreve-se a obtenção de fronteiras não lineares com SVMs, por meio de uma extensão das SVMs lineares.

D.2.1

SVMs com Margens Rígidas

As SVMs lineares com margens rígidas definem fronteiras lineares a partir de dados linearmente separáveis. Seja T um conjunto de treinamento com n dados $x_i \in X$ e seus respectivos rótulos $y_i \in Y$, em que X constitui o espaço dos dados e $Y = \{-1, +1\}$. T é linearmente separável se é possível separar os dados das classes $+1$ e -1 por um hiperplano [123].

Classificadores que separam os dados por meio de um hiperplano são denominados lineares [107]. A equação de um hiperplano é apresentada na Equação D.1, em que $\mathbf{w} \cdot \mathbf{x}$ é o produto escalar entre os vetores \mathbf{w} e \mathbf{x} , $\mathbf{w} \in X$ é o vetor normal ao hiperplano descrito e b dividido por $\|\mathbf{w}\|$ corresponde à distância do hiperplano em relação à origem, com $b \in \mathbb{R}$.

$$f(\mathbf{x}) = \mathbf{w} \cdot \mathbf{x} + b = 0 \quad \text{Eq.D.1}$$

Essa equação divide o espaço dos dados X em duas regiões: $\mathbf{w} \cdot \mathbf{x} + b > 0$ e $\mathbf{w} \cdot \mathbf{x} + b < 0$. Uma função sinal $g(\mathbf{x}) = \text{sgn}(f(\mathbf{x}))$ pode então ser empregada na obtenção das classificações, conforme ilustrado na Equação D.2 [122].

$$g(\mathbf{x}) = \text{sgn}(f(\mathbf{x})) = \begin{cases} +1, & \text{se } \mathbf{w} \cdot \mathbf{x} + b > 0 \\ -1, & \text{se } \mathbf{w} \cdot \mathbf{x} + b < 0 \end{cases} \quad \text{Eq.D.2}$$

A partir de $f(\mathbf{x})$, é possível obter um número infinito de hiperplanos equivalentes, pela multiplicação de \mathbf{w} e b por uma mesma constante [119]. Define-se o hiperplano canônico em relação ao conjunto T como aquele em que \mathbf{w} e b são escalados de forma que os exemplos mais próximos ao hiperplano $\mathbf{w} \cdot \mathbf{x} + b = 0$ satisfaçam a Equação D.3 [117].

$$|\mathbf{w} \cdot \mathbf{x}_i + b| = 1 \quad \text{Eq.D.3}$$

Essa forma implica nas inequações apresentadas em Eq.D.4, resumidas na Expressão mostrada em Eq.11.

$$\begin{cases} \mathbf{w} \cdot \mathbf{x}_i + b \geq +1, \text{ se } y_i = +1 \\ \mathbf{w} \cdot \mathbf{x}_i + b \leq -1, \text{ se } y_i = -1 \end{cases} \quad \text{Eq.D.4}$$

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1 \geq 0, \forall (\mathbf{x}_i, y_i) \in T \quad \text{Eq.D.5}$$

Seja \mathbf{x}_1 um ponto no hiperplano $H_1: \mathbf{w} \cdot \mathbf{x} + b = +1$ e \mathbf{x}_2 um ponto no hiperplano $H_2: \mathbf{w} \cdot \mathbf{x} + b = -1$, conforme ilustrado na Figura D.2. Projetando $\mathbf{x}_1 - \mathbf{x}_2$ na direção de \mathbf{w} , perpendicular ao hiperplano separador $\mathbf{w} \cdot \mathbf{x} + b = 0$, é possível obter a distância entre os hiperplanos H_1 e H_2 [107]. Essa projeção é apresentada na Equação D.6.

$$(\mathbf{x}_1 - \mathbf{x}_2) \left(\frac{\mathbf{w}}{\|\mathbf{w}\|} \cdot \frac{(\mathbf{x}_1 - \mathbf{x}_2)}{\|\mathbf{x}_1 - \mathbf{x}_2\|} \right) \quad \text{Eq.D.6}$$

Tem-se que $\mathbf{w} \cdot \mathbf{x}_1 + b = +1$ e $\mathbf{w} \cdot \mathbf{x}_2 + b = -1$. A diferença entre essas equações fornece $\mathbf{w} \cdot (\mathbf{x}_1 - \mathbf{x}_2) = 2$ [109]. Substituindo esse resultado na Equação D.6, tem-se:

$$\frac{2 (\mathbf{x}_1 - \mathbf{x}_2)}{\|\mathbf{w}\| \|\mathbf{x}_1 - \mathbf{x}_2\|} \quad \text{Eq.D.7}$$

Como deseja-se obter o comprimento do vetor projetado, toma-se a norma da Equação D.7, obtendo-se:

$$\frac{2}{\|\mathbf{w}\|}$$

Eq.D.8

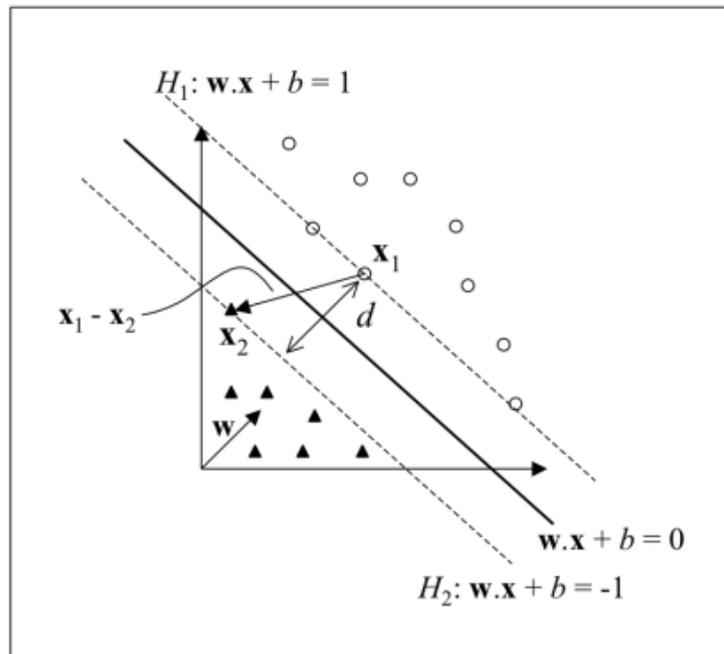


Figura D.2 - Cálculo da distância d entre os hiperplanos H_1 e H_2 [109]

Essa é a distância d , ilustrada na Figura D.2, entre os hiperplanos H_1 e H_2 , paralelos ao hiperplano separador. Como \mathbf{w} e b foram escalados de forma a não haver exemplos entre H_1 e H_2 , 1 dividido por $\|\mathbf{w}\|$ é a distância mínima entre o hiperplano separador e os dados de treinamento. Essa distância é definida como a margem geométrica do classificador linear [107].

A partir das considerações anteriores, verifica-se que a maximização da margem de separação dos dados em relação a $\mathbf{w} \cdot \mathbf{x} + b = 0$ pode ser obtida pela minimização de $\|\mathbf{w}\|$ [106]. Dessa forma, recorre-se ao seguinte problema de otimização [123]:

$$\text{Minimizar}_{\mathbf{w}, b} \frac{1}{2} \|\mathbf{w}\|^2 \quad \text{Eq.D.9}$$

$$\text{Com as restrições: } y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1 \geq 0, \forall i = 1, \dots, n \quad \text{Eq.D.10}$$

As restrições são impostas de maneira a assegurar que não haja dados de treinamento entre as margens de separação das classes. Por esse motivo, a SVM obtida possui também a nomenclatura de SVM com margens rígidas.

O problema de otimização obtido é quadrático, cuja solução possui uma ampla e estabelecida teoria matemática [123]. Como a função objetivo sendo minimizada é convexa e os pontos que satisfazem as restrições formam um conjunto convexo, esse problema possui um único mínimo global [119]. Problemas desse tipo podem ser solucionados com a introdução de uma função Lagrangiana, que engloba as restrições à função objetivo, associadas a parâmetros denominados multiplicadores de Lagrange α_i (Equação D.11) [123].

$$L(\mathbf{w}, b, \boldsymbol{\alpha}) = \frac{1}{2} \|\mathbf{w}\|^2 - \sum_{i=1}^n \alpha_i (y_i (\mathbf{w} \cdot \mathbf{x}_i + b) - 1) \quad \text{Eq.D.11}$$

A função Lagrangiana deve ser minimizada, o que implica em maximizar as variáveis α_i e minimizar \mathbf{w} e b [117]. Tem-se então um ponto de sela, no qual:

$$\frac{\partial L}{\partial b} = 0 \quad \text{e} \quad \frac{\partial L}{\partial \mathbf{w}} = 0 \quad \text{Eq.D.12}$$

A resolução dessas equações leva aos resultados representados nas equações D.13 e D.14.

$$\sum_{i=1}^n \alpha_i y_i = 0 \quad \text{Eq.D.13}$$

$$\mathbf{w} = \sum_{i=1}^n \alpha_i y_i \mathbf{x}_i \quad \text{Eq.D.14}$$

Substituindo as equações D.13 e D.14 na Equação D.11, obtém-se o seguinte problema de otimização:

$$\text{Maximizar}_{\alpha} \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j (\mathbf{x}_i \cdot \mathbf{x}_j) \quad \text{Eq.D.15}$$

$$\text{Com as restrições: } \begin{cases} \alpha_i \geq 0, \forall i = 1, \dots, n \\ \sum_{i=1}^n \alpha_i y_i = 0 \end{cases} \quad \text{Eq.D.16}$$

Essa formulação é denominada forma dual, enquanto o problema original é referenciado como forma primal. A forma dual possui os atrativos de apresentar restrições mais simples e permitir a representação do problema de otimização em termos de produtos internos entre dados, o que será útil na não-linearização das SVMs (que será tratada mais à frente). É interessante observar também que o problema dual é formulado utilizando apenas os dados de treinamento e os seus rótulos.

D.2.2

SVMs com Margens Suaves

Em situações reais, é difícil encontrar aplicações cujos dados sejam linearmente separáveis. Isso se deve a diversos fatores, entre eles a presença de ruídos e *outliers* nos dados ou à própria natureza do problema, que pode ser não linear. Neste caso, as SVMs lineares de margens rígidas podem ser estendidas para lidar com conjuntos de treinamento mais gerais. Para realizar essa tarefa, permite-se que alguns dados possam violar a restrição da Equação D.10. Isso é feito com a introdução de variáveis de folga ε_i , para todo $i = 1, \dots, n$ [122]. Essas variáveis relaxam as restrições impostas ao problema de otimização, que se tornam [123]:

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 - \varepsilon_i, \quad \varepsilon_i \geq 0, \quad \forall i = 1, \dots, n \quad \text{Eq.D.17}$$

A aplicação desse procedimento suaviza as margens do classificador linear, permitindo que alguns dados permaneçam entre os hiperplanos H_1 e H_2 e também

a ocorrência de alguns erros de classificação. Por esse motivo, as SVMs obtidas neste caso também podem ser referenciadas como SVMs com margens suaves.

Um erro no conjunto de treinamento é indicado por um valor de ε_i maior que 1. Logo, a soma dos ε_i representa um limite no número de erros de treinamento [106]. Para levar em consideração esse termo, minimizando assim o erro sobre os dados de treinamento, a função objetivo da Equação D.9 é reformulada como [5]:

$$\underset{w,b,\varepsilon}{\text{Minimizar}} \quad \frac{1}{2} \|\mathbf{w}\|^2 + C \left(\sum_{i=1}^n \varepsilon_i \right) \quad \text{Eq.D.18}$$

A constante C é um termo de regularização que impõe um peso à minimização dos erros no conjunto de treinamento em relação à minimização da complexidade do modelo [119]. A presença do termo $\sum_{i=1}^n \varepsilon_i$ no problema de otimização também pode ser vista como uma minimização de erros marginais, pois um valor de $\varepsilon_i \in (0, 1]$ indica um dado entre as margens. Tem-se então uma formulação de acordo com os princípios da TAE discutidos anteriormente.

Novamente o problema de otimização gerado é quadrático, com as restrições lineares apresentadas na Equação D.17. A sua solução envolve passos matemáticos semelhantes aos apresentados anteriormente, com a introdução de uma função Lagrangiana e tornando suas derivadas parciais nulas. Tem-se como resultado o seguinte problema dual:

$$\underset{\alpha}{\text{Maximizar}} \quad \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j (\mathbf{x}_i \cdot \mathbf{x}_j) \quad \text{Eq.D.19}$$

$$\text{Com as restrições:} \quad \begin{cases} 0 \leq \alpha_i \leq C, \forall i = 1, \dots, n \\ \sum_{i=1}^n \alpha_i y_i = 0 \end{cases} \quad \text{Eq.D.20}$$

Pode-se observar que essa formulação é igual à apresentada para as SVMs de margens rígidas, a não ser pela restrição nos α_i , que agora são limitados pelo valor de C .

D.3

SVMs Não Lineares

As SVMs lineares são eficazes na classificação de conjuntos de dados linearmente separáveis ou que possuam uma distribuição aproximadamente linear, sendo que a versão de margens suaves tolera a presença de alguns ruídos e *outliers*. Porém, há muitos casos em que não é possível dividir satisfatoriamente os dados de treinamento por um hiperplano. Um exemplo é apresentado na Figura D.3(a), em que o uso de uma fronteira curva seria mais adequada na separação das classes.

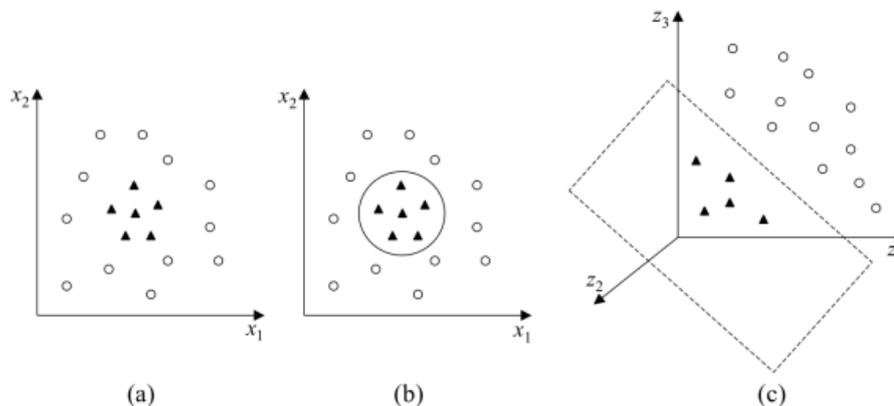


Figura D.3 – (a) Conjunto de dados não linear; (b) Fronteira não linear no espaço de entradas; (c) Fronteira linear no espaço de características [117]

As SVMs lidam com problemas não lineares mapeando o conjunto de treinamento de seu espaço original, referenciado como de entradas, para um novo espaço de maior dimensão, denominado espaço de características (*feature space*) [14]. Seja $\Phi: X \rightarrow \Psi$ um mapeamento, em que X é o espaço de entradas e Ψ denota o espaço de características. A escolha apropriada de Φ faz com que o conjunto de treinamento mapeado em Ψ possa ser separado por uma SVM linear.

O uso desse procedimento é motivado pelo teorema de Cover [64]. Dado um conjunto de dados não linear no espaço de entradas X , esse teorema afirma que X pode ser transformado em um espaço de características Ψ no qual com alta probabilidade os dados são linearmente separáveis. Para isso duas condições devem ser satisfeitas. A primeira é que a transformação seja não linear, enquanto a segunda é que a dimensão do espaço de características seja suficientemente alta.

Para ilustrar esses conceitos, considere o conjunto de dados apresentado na Figura D.3(a) [117]. Transformando os dados de \mathbb{R}^2 para \mathbb{R}^3 com o mapeamento representado na Equação D.21, o conjunto de dados não linear em \mathbb{R}^2 torna-se linearmente separável em \mathbb{R}^3 (Figura D.3(c)). É possível então encontrar um hiperplano capaz de separar esses dados, descrito na Equação D.22. Pode-se verificar que a função apresentada, embora linear em \mathbb{R}^3 (Figura D.3(c)), corresponde a uma fronteira não linear em \mathbb{R}^2 (Figura D.3(b)).

$$\Phi(\mathbf{x}) = \Phi(x_1, x_2) = (x_1^2, \sqrt{2}x_1x_2, x_2^2) \quad \text{Eq.D.21}$$

$$f(\mathbf{x}) = \mathbf{w} \cdot \Phi(\mathbf{x}) + b = w_1x_1^2 + w_2\sqrt{2}x_1x_2 + w_3x_2^2 + b = 0 \quad \text{Eq.D.22}$$

Logo, mapeia-se inicialmente os dados para um espaço de maior dimensão utilizando Φ e aplica-se a SVM linear sobre este espaço. Essa encontra o hiperplano com maior margem de separação, garantindo assim uma boa generalização. Utiliza-se a versão de SVM linear com margens suaves, que permite lidar com ruídos e *outliers* presentes nos dados. Para realizar o mapeamento, aplica-se Φ aos exemplos presentes no problema de otimização representado na Equação D.19, conforme ilustrado a seguir:

$$\text{Maximizar}_{\alpha} \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j (\Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j)) \quad \text{Eq.D.23}$$

De forma semelhante, a Equação D.23 está sob as mesmas restrições mostradas na Equação D.20.

Como Ψ pode ter dimensão muito alta (até mesmo infinita), a computação de Φ pode ser extremamente custosa ou inviável. Porém, percebe-se pela Equação D.23 que a única informação necessária sobre o mapeamento é de como realizar o cálculo de produtos escalares entre os dados no espaço de características, pois tem-se sempre $\Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j)$, para dois dados \mathbf{x}_i e \mathbf{x}_j , em conjunto. Isso é obtido com o uso de funções denominadas *Kernels*.

Um Kernel K é uma função que recebe dois pontos \mathbf{x}_i e \mathbf{x}_j do espaço de entradas e computa o produto escalar desses dados no espaço de características [110]. Tem-se então:

$$k(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j) \quad \text{Eq.D.24}$$

Para o mapeamento apresentado na Equação D.21 e dois dados $\mathbf{x}_i = (x_{1i}, x_{2i})$ e $\mathbf{x}_j = (x_{1j}, x_{2j})$ em \mathbb{R}^2 , por exemplo, o Kernel é dado por [14]:

$$k(x_i, x_j) = (x_{1i}^2, \sqrt{2}x_{1i}x_{2i}, x_{2i}^2) \cdot (x_{1j}^2, \sqrt{2}x_{1j}x_{2j}, x_{2j}^2) = (x_i \cdot x_j)^2 \quad \text{Eq.D.25}$$

É comum empregar a função Kernel sem conhecer o mapeamento Φ , que é gerado implicitamente. A utilidade dos Kernels está, portanto, na simplicidade de seu cálculo e em sua capacidade de representar espaços abstratos.

Para garantir a convexidade do problema de otimização formulado na Equação D.23 e também que o Kernel represente mapeamentos nos quais seja possível o cálculo de produtos escalares conforme a Equação D.24, utiliza-se funções Kernel que seguem as condições estabelecidas pelo teorema de Mercer [115][122]. De forma simplificada, um Kernel que satisfaz as condições de Mercer é caracterizado por dar origem a matrizes positivas semi-definidas \mathbf{K} , em que cada elemento K_{ij} é definido por $K_{ij} = K(\mathbf{x}_i, \mathbf{x}_j)$, para todo $i, j = 1, \dots, n$ [110].

Alguns dos Kernels mais utilizados na prática são os Polinomiais, os Gaussianos ou RBF (*Radial-Basis Function*) e os Sigmoidais, listados na Tabela D.1. Cada um deles apresenta parâmetros que devem ser determinados pelo usuário, indicados também na Tabela D.1. O Kernel Sigmoidal, em particular, satisfaz as condições de Mercer apenas para alguns valores de δ e κ . Os Kernels Polinomiais com $d = 1$ também são denominados lineares.

Tabela D.1 – Funções Kernel mais comuns [108]

Tipo de Kernel	Função $K(\mathbf{x}_i, \mathbf{x}_j)$	Parâmetros
Polinomial	$(\delta(\mathbf{x}_i \cdot \mathbf{x}_j) + \kappa)^d$	δ, κ e d
Gaussiano	$\exp(-\sigma \ \mathbf{x}_i - \mathbf{x}_j\ ^2)$	σ
Sigmoidal	$\tanh(\delta(\mathbf{x}_i \cdot \mathbf{x}_j) + \kappa)$	δ e κ