

UNIVERSIDADE FEDERAL DO PARÁ – UFPA
INSTITUTO DE TECNOLOGIA – ITEC
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

CARACTERIZAÇÃO DE PADRÕES DE DESCARGAS PARCIAIS EM
HIDROGERADORES UTILIZANDO TÉCNICAS DE INTELIGÊNCIA
COMPUTACIONAL

MEDILLIN PEREIRA ALVES

DM 32 / 2015

UFPA / ITEC / PPGEE
Campus Universitário do Guamá
Belém-Pará-Brasil
2015

UNIVERSIDADE FEDERAL DO PARÁ – UFPA
INSTITUTO DE TECNOLOGIA – ITEC
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

MEDILLIN PEREIRA ALVES

CARACTERIZAÇÃO DE PADRÕES DE DESCARGAS PARCIAIS EM
HIDROGERADORES UTILIZANDO TÉCNICAS DE INTELIGÊNCIA
COMPUTACIONAL

Dissertação de Mestrado
submetida à Banca
Examinadora do Programa de
Pós-Graduação em Engenharia
Elétrica da UFPA para obtenção
do Grau de Mestre em
Engenharia Elétrica na área de
Sistemas de Energia Elétrica.

UFPA / ITEC / PPGEE
Campus Universitário do Guamá
Belém-Pará-Brasil
2015

UNIVERSIDADE FEDERAL DO PARÁ – UFPA

INSTITUTO DE TECNOLOGIA – ITEC

PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

AUTOR: MEDILLIN PEREIRA ALVES

Título: Caracterização de Padrões de Descargas Parciais em Hidrogeradores Utilizando Técnicas de Inteligência Computacional

DISSERTAÇÃO DE MESTRADO SUBMETIDA À BANCA EXAMINADORA APROVADA PELO COLEGIADO DO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA, SENDO JULGADA ADEQUADA PARA A OBTENÇÃO DO GRAU DE MESTRE EM ENGENHARIA ELÉTRICA NA ÁREA DE SISTEMA DE ENERGIA ELÉTRICA.

APROVADA EM: **24/09/2015**

BANCA EXAMINADORA:

Prof. Dr. Marcus Vinícius Alves Nunes
(Orientador – PPGEE/UFPA)

Prof. Dr. Ubiratan Holanda Bezerra
(Membro – PPGEE/UFPA)

Prof. Dr. João Paulo Abreu Vieira
(Membro – PPGEE/UFPA)

Prof. Dr. André Cavalcante do Nascimento
(Membro – IFPA)

VISTO:

Prof. Dr. Evaldo Gonçalves Pelaes
(Coordenador do PPGEE/ITEC/UFPA)

O potencial da humanidade é infinito e todo ser tem uma contribuição a fazer por um mundo mais grandioso. Estamos nele juntos. Somos UM.

Helena Petrovna Blavatsky

DEDICATÓRIA

Aos Sebastões da minha vida, minha mãe e meu marido, que tanto me incentivaram e apoiaram nesta empreitada. Aos meus queridos filhos de estimação, Minnie e Negão, que estiveram ao meu lado em tempo integral durante a produção deste trabalho.

Medillin Pereira Alves

AGRADECIMENTOS

Primeiramente, agradeço a Deus por proporcionar à minha vida grandes ensinamentos, agraciar-me com opções de escolhas, conceder-me luz para meus caminhos e me oportunizar a gratidão.

À minha mãe, Sebastiana Veracy de Oliveira Pereira, cujo amor imensurável sempre me orienta ao caminho do bem, incentivando incansavelmente, e sendo minha plateia mais calorosa.

Ao meu pai, Jesse Luis Teixeira da Silva, que foi empenhado a proporcionar todos os meios que fazem de mim o que sou hoje.

Ao meu marido, Sebastião Borges Fonseca, por apoiar e respeitar todos os meus desejos e ideais, tornando-se parceiro. Agradeço pelo apoio direto que exerceu, apontando caminhos fundamentais que ajudaram a decidir alguns rumos desta pesquisa.

À minha prima, Cecília Santos de Oliveira Cardoso, cuja alegria é sincera por nossos progressos.

Aos meus filhos de estimação, Minnie e Negão, por estarem ao meu lado integralmente durante a realização deste trabalho.

Ao meu orientador, Prof^o Dr^o Marcus Vinícius Alves Nunes, por me aceitar como sua orientanda de mestrado, por sempre se dispor às minhas recorrências, e por me proporcionar todos os caminhos que me levaram à conclusão deste trabalho.

Ao Engenheiro de Manutenção Elétrica da ELETRONORTE, Fernando de Souza Brasil e, ELETRONORTE Centrais Elétricas do Brasil S/A, por fazer parceria nesta empreitada, concedendo-me os dados, materiais e orientações necessárias para a realização deste trabalho.

Medillin Pereira Alves

RESUMO

ALVES, M. P., **Caracterização de Padrões de Descargas Parciais em Hidrogeradores Utilizando Técnicas de Inteligência Computacional**. Dissertação de Mestrado. Programa de Pós-graduação em Engenharia Elétrica da Universidade Federal do Pará, Belém, Pará, 2015. 121p.

Esta dissertação apresenta os experimentos com aplicações de técnicas de Inteligência computacional para caracterização de descargas parciais em hidrogeradores. A classificação das descargas parciais contribui para uma análise prévia de problemas e permite a manutenção preditiva nas máquinas, reduzindo a possibilidade de falhas nas mesmas. Os dados foram coletados de modo online (máquina em operação) na Usina Hidrelétrica de Tucuruí, sendo observados os padrões de descarga interna, de laminação e entre barras. O software IMA-DP, desenvolvido em parceria entre Eletronorte e Cepel, permitiu que esses dados fossem medidos e registrados de maneira rápida, e organizados através dos mapas PRPD (Phase Resolved Partial Discharges). As técnicas de binarização, ANOVA (Analysis of Variance), ACI (Análise de Componentes Independentes) e ACP (Análise de Componentes Principais) foram aplicadas aos sinais para adequar os mesmos ao uso das técnicas de inteligência computacional. O trabalho foi desenvolvido no ambiente IPython usando a biblioteca scikit-learn, a qual possui eficientes algoritmos de inteligência. Os experimentos foram realizados fazendo-se uso das técnicas: KNN (K-Nearest Neighbors.), Floresta Randômica e MVS (Máquinas de Vetores de Suporte). Tais técnicas apresentaram bons resultados com os experimentos realizados, destacando-se aqueles obtidos para MVS que apresentaram os melhores resultados, atingindo uma acurácia de 96.07%, devido possuir mecanismos de seleção das principais variáveis durante o processo de treinamento.

Palavras Chaves: Caracterização de Descargas Parciais, Inteligência Computacional, Hidrogeradores, Isolamentos, Bobinas Estatóricas.

ABSTRACT

ALVES, M. P., **Partial Discharge Standards Characterization in Hydrogenerators Using Techniques of Computational Intelligence**. Master Degree Dissertation. Programa de Pós-graduação em Engenharia Elétrica da Universidade Federal do Pará, Belém, Pará, 2015. 121p.

This master's thesis presents the experiments with applications of computational intelligence techniques for the characterization of partial discharges in hydrogenerators. The classification of the partial discharge contributes to a prior analysis problems and allows predictive maintenance on machinery, reducing the possibility of failures in them. Data were collected online mode (operation machine) in the Tucuruí Hydroelectric Power Plant, observed the internal discharge standards, delamination and between bars. The IMA-DP software, developed in partnership with Eletronorte and Cepel, allowed these data were measured and recorded quickly, and organized through PRPD maps (Phase resolved Partial Discharges). Binarization techniques, ANOVA (Analisis of Variance), ICA (Independent Component Analysis) and PCA (Principal Component Analysis) were applied to the signals to adapt them to the use of computational intelligence techniques. The study was developed in IPython environment using scikit-learn library, which has efficient intelligence algorithms. The experiments were performed making use of techniques: KNN (K-Nearest Neighbors), Random Forest and MVS (Support Vector Machines). Such techniques showed good results with the experiments, highlighting those obtained for MVS that showed the best results, achieving an accuracy of 96.07%, due possess selection mechanisms of the main variables during the training process.

Key Words: Characterization of Partial Discharge, Computational Intelligence, Hydrogenerators, Insulations, Stator Coils.

SUMÁRIO

1.0 INTRODUÇÃO.....	1
1.1 MOTIVAÇÃO	1
1.2 REVISÃO BIBLIOGRÁFICA	2
1.3 OBJETIVOS	4
1.3.1 Objetivo geral	4
1.3.2 Objetivos específicos	4
1.4 ESCOPO DO TRABALHO	5
2.0 DESCARGAS PARCIAIS.....	7
2.1 INTRODUÇÃO.....	7
2.2 O CONCEITO DE DESCARGAS PARCIAIS	7
2.3 FÍSICA DAS DESCARGAS PARCIAIS	9
2.3.1 Da ionização dos átomos à avalanche de elétrons.....	9
2.3.2 Forma do pulso das DPs	11
2.4 CIRCUITO EQUIVALENTE DE UMA CAVIDADE	13
2.5 TIPOS DE DP.....	15
2.5.1 Classificação geral das DPs.....	15
2.5.2 Classificação das DPs em máquinas elétricas rotativas.....	16
2.5.2.1 Descargas internas	16
2.5.2.1.1 Cavidades internas	16
2.5.2.1.2 Delaminação interna	17
2.5.2.1.3 Delaminação entre condutores e a isolação	18
2.5.2.1.4 Arboreescência elétrica.....	19
2.5.2.2 Descargas de ranhura.....	19
2.5.2.3 Descargas na cabeça da bobina	20
2.5.2.3.1 Descargas de superfície	20
2.5.2.3.2 Descargas entre fases	20
2.5.2.4 Partículas condutoras.....	20
2.6 GRANDEZAS ASSOCIADAS ÀS DPS.....	21
2.7 GRANDEZAS ASSOCIADAS À FASE – OS MAPAS ESTATÍSTICOS	22
2.8 MÉTODOS DE MEDIÇÃO DAS DPS.....	24
2.8.1 Métodos não elétricos.....	24
2.8.1.1 Método acústico e ultrassônico	24
2.8.1.2 Detecção de ozônio.....	25

2.8.1.3	Teste de black-out	26
2.8.1.4	Teste de corona	26
2.8.2	Métodos elétricos.....	26
2.8.2.1	Medição do pulso elétrico	26
2.8.2.2	Medição da radiação de rádio frequência	28
2.9	DIFICULDADES NOS PROCESSOS DE MEDIÇÃO DAS DPS	28
2.10	CONCLUSÕES DO CAPÍTULO	28
3.0	OS ENROLAMENTOS ESTATÓRICOS DO HIDROGERADOR	30
3.1	INTRODUÇÃO.....	30
3.2	ESTRUTURA DOS ENROLAMENTOS DO ESTATOR	30
3.2.1	Estrutura geral do hidrogerador	30
3.2.2	Os Enrolamentos Estatóricos do Hidrogerador	32
3.2.2.1	Modelo barra	34
3.2.2.2	Modelo de Múltiplas Espiras.....	37
3.3	OS ISOLAMENTOS DAS BOBINAS ESTATÓRICAS	39
3.3.1	Breve histórico sobre a constituição dos isolamentos das bobinas estatóricas	39
3.3.2	Tipos de isolamentos	39
3.4	FALHAS NO ISOLAMENTO DO ESTATOR	40
3.5	CONCLUSÕES DO CAPÍTULO	44
4.0	CLASSIFICADORES BASEADOS EM TÉCNICAS DE INTELIGÊNCIA E TRATAMENTO DOS DADOS.....	45
4.1	INTRODUÇÃO.....	45
4.2	TRATAMENTO DE DADOS	45
4.2.1	Análise de Variância	45
4.2.1.1	Anova de um fator	46
4.2.2	Análise de Componentes Principais	48
4.2.2.1	Decomposições em Valores Singulares	49
4.2.2.2	Aplicação da Decomposição em Valores Singulares na Análise de Componentes Principais	52
4.2.3	Análise de Componentes Independentes	56
4.2.3.1	Definição	57
4.2.3.2	Restrições do Método ACI	58
4.2.3.3	Técnicas de Pré-Processamento.....	59
4.2.3.4	Estimativa dos Componentes Independentes	61

4.2.3.5 Medida de Não Gaussianidade por Kurtosis	61
4.2.3.6 Negentropia	62
4.2.3.7 Algoritmo Rápido de Ponto Fixo (FastICA)	63
4.3 TÉCNICAS DE CLASSIFICAÇÃO	65
4.3.1 K-Vizinhos Mais Próximos	65
4.3.2 Floresta Randômica	67
4.3.3 Máquinas de Vetores de Suporte	69
4.3.3.1 Teoria do Aprendizado Estatístico	70
4.3.3.2 Classificação Linear	72
4.3.3.3 Hiperplano Ótimo	73
4.3.3.4 Classificação Multiclasses	76
4.4 CONCLUSÕES DO CAPÍTULO	77
5.0 METODOLOGIA, EXPERIMENTOS E RESULTADOS	79
5.1 INTRODUÇÃO	79
5.2 PROCEDIMENTOS METODOLÓGICOS	79
5.3 COLETA DE DADOS	81
5.4 FERRAMENTAS COMPUTACIONAIS	83
5.5 TRATAMENTO DOS DADOS	83
5.6 RESULTADOS DE EXPERIMENTOS	85
5.6.1 Configurações no algoritmo de KNN	86
5.6.2 Configurações no algoritmo de Floresta Randômica	86
5.6.3 Configurações no algoritmo de MVS	87
5.6.4 Configurações no algoritmo de ANOVA aplicado às técnicas classificadoras	88
5.6.5 Configurações no algoritmo de FastICA aplicado às técnicas classificadoras	89
5.6.6 Configurações no algoritmo de ACP aplicado às técnicas classificadoras	89
5.7 RESULTADOS COMPLEMENTARES	91
5.8 CONCLUSÕES DO CAPÍTULO	97
6.0 CONCLUSÕES	99
6.1 CONSIDERAÇÕES FINAIS	99
6.2 SUGESTÕES PARA TRABALHOS FUTUROS	101
REFERÊNCIAS BIBLIOGRÁFICAS	102

LISTA DE FIGURAS

Figura 2.1 - Representação esquemática de uma cavidade no interior do material dielétrico...	8
Figura 2.2 - Representação esquemática do fenômeno de arborescência elétrica no interior do material dielétrico.....	8
Figura 2.3 - Formação de dipolos. a) Átomo não polarizado. b) Átomo polarizado comportando-se como dipolo sob a ação de um campo elétrico.....	9
Figura 2.4 - Processo de avalanche de elétrons. a) Elétron ganha energia tornando-se elétron livre e átomo torna-se íon positivo. b) Elétron livre colide com átomo e libera novo elétron livre e íon positivo. c) A sucessão desse processo resulta em várias cargas pos positivas e negativas que resultam na avalanche de elétrons.....	10
Figura 2.5 - Processo de centelhamento. a) Elétron livre é reincorporado no íon positivo formando novamente o átomo. b) Na formação do átomo um quantum de energia liberado gera luz visível na forma de um centelhamento.....	11
Figura 2.6 - Pulso de DP medido através da variação de corrente pelo tempo.....	12
Figura 2.7 - Circuito equivalente de uma cavidade no interior do dielétrico.....	13
Figura 2.8 - Pulsos de DPs em uma cavidade.....	14
Figura 2.9 - Corona no isolador.....	15
Figura 2.10 - Imagem por microscopia óptica de um corte da seção transversal do isolamento epóxi-mica de uma barra estatórica de hidrogerador com um vazio na região do epóxi.....	17
Figura 2.11 - Imagem por microscopia óptica de um corte da seção transversal do isolamento de uma barra estatórica de hidrogerador com grande delaminação do compósito epóxi-mica	18
Figura 2.12 - Fotografia da seção transversal de uma barra estatórica, onde se podem observar delaminações internas no isolamento.....	18
Figura 2.13 - Fontes de descargas internas na seção transversal de uma bobina estatórica. a) Cavidade interna. b) Delaminação interna. c) Delaminação entre condutores e isolamento. d) Arboreamento elétrico.....	19
Figura 2.14 - Fontes de DPs no isolamento do estator.....	21
Figura. 2.15 - Sensores acústicos instalados em um transformador de potência.....	25

Figura 2.16 - Capacitores de acoplamento. a) Instalado no interior do enrolamento estatórico. b) Instalado no terminal fase.....	27
Figura 3.1 - Esquema de uma usina hidrelétrica.....	31
Figura 3.2 - Estrutura do hidrogerador. a) Seção do estator: enrolamentos fixados no núcleo. b) Rotor. c) Seção da máquina completa: estator e rotor.....	32
Figura 3.3 - Estator de hidrogerador.....	33
Figura 3.4 - Região das cabeças de bobinas interligadas.....	33
Figura 3.5 - Figura esquemática do corte transversal das barras inseridas na ranhura.....	34
Figura 3.6 - Barras. a) Modelo de barra imbricado; b) Modelo de barra ondulado.....	34
Figura 3.7 - Conexões dos enrolamentos. a) Enrolamentos ondulados; b) Enrolamentos imbricados.....	35
Figura 3.8 - Figura esquemática da transposição Roebel.....	36
Figura 3.9 - Múltiplas espiras. a) Modelo múltipla espira diamante; b) Modelo múltipla espira concêntrico.....	37
Figura 3.10 - Figura esquemática do corte de uma ranhura com bobinas de múltiplas espiras.....	38
Figura 3.11 - Figura esquemática do corte de uma ranhura com bobinas de múltiplas espiras com isolamento complementar entre espiras.....	38
Figura 3.12 - Os tipos de isolamento das bobinas. a) Isolamento das bobinas Roebel. b) Isolamento das bobinas de múltipla espira.....	40
Figura 3.13 - Causas raízes de falhas nos hidrogeradores.....	41
Figura 3.14 - Causas de falhas no sistema de isolamento elétrica.....	42
Figura 4.1 - Classificação de objeto usando K-NN. Para $k=1$, x é assumido positivo; $k=3$, x é negativo; $k=5$, x é positivo.....	66
Figura 4.2 - Classe de hiperplanos com um hiperplano ótimo.....	71
Figura 4.3 - A margem ρ e os vetores suporte sobre a linha pontilhada.....	72
Figura 4.4 - Distância entre hiperplanos e vetores suporte.....	74
Figura 5.1 – Definição da trajetória da pesquisa.....	80

Figura 5.2 - (a) Caixa terminal com as impedâncias de medição; (b) IMA-DP autônomo.....	82
Figura 5.3 - Exemplo de uma medição sob a forma de mapa PRPD.....	82
Figura 5.4 - Sinal de descarga interna com visualização no IMA-DP.....	85
Figura 5.5 - Sinal de descarga interna com visualização no IPython utilizando a biblioteca matplotlib. a) Sinal sem tratamento; b) Sinal binarizado.....	85
Figura 5.6 - 30000 atributos selecionados por ANOVA.....	89
Figura 5.7 - Acurácia usando ACP para cada n componentes para a técnica MVS.....	90
Figura 5.8 - Acurácia usando ACP para cada n componentes para a técnica KNN.....	91
Figura 5.9 - Acurácia usando ACP para cada n componentes para a técnica Floresta Randômica.....	91

LISTA DE TABELAS

Tabela 3.1 - Tipos de falhas nos enrolamentos estatóricos.....	43
Tabela 4.1 - Experimento de um fator.....	46
Tabela 4.2 - Equações de SQG, SQR e SQT.....	47
Tabela 4.3 - Equações para cálculo das variâncias.....	48
Tabela 5.1 - Acurácias para cada número de vizinho.....	86
Tabela 5.2 - Acurácias para cada número de árvores.....	87
Tabela 5.3 - Acurácias para cada kernel.....	87
Tabela 5.4 - Acurácias para cada experimento.....	92
Tabela 5.5 - Matriz de confusão para predição com MVS.....	93
Tabela 5.6 - Matriz de confusão para predição com Floresta Randômica.....	94
Tabela 5.7 - Matriz de confusão para predição de KNN.....	94
Tabela 5.8 - Matriz de confusão para predição de MVS com PCA.....	94
Tabela 5.9 - Matriz de confusão para predição de MVS com ANOVA.....	95
Tabela 5.10 - Matriz de confusão para predição de MVS com ACI.....	95
Tabela 5.11 - Matriz de confusão para predição de KNN com PCA.....	95
Tabela 5.12 - Matriz de confusão para predição de KNN com ANOVA.....	96
Tabela 5.13 - Matriz de confusão para predição de KNN com ACI.....	96
Tabela 5.14 - Matriz de confusão para predição de Floresta Randômica com PCA.....	96
Tabela 5.15 - Matriz de confusão para predição de Floresta Randômica com ANOVA.....	97
Tabela 5.16 - Matriz de confusão para predição de Floresta Randômica com ACI.....	97

LISTA DE QUADROS

Quadro 2.1 - Padrões de mapas estatísticos de DPs.....	23
--	----

LISTA DE SIGLAS

ACI - Análise de Componentes Independentes

ACP - Análise de Componentes Principais.

ANOVA - Analysis of Variance.

DP – Descarga Parcial.

DVS - Decomposição por Valores Singulares.

IMA-DP – Instrumento de Monitoração e Análise de Descargas Parciais.

KKT - Karush-Kuhn-Tucker

KNN - K-Nearest Neighbors.

MVS - Máquinas de Vetores de Suporte.

PC – Picocoulombs.

PRPD - Phase Resolved Partial Discharges.

SNPTEE - Seminário Nacional de Produção e Transmissão de Energia Elétrica

SQG - Soma dos Quadrados dos Grupos

SQR - Soma dos Quadrados dos Resíduos.

SQT - Soma dos Quadrados Totais

TCT - Todos-Contra-Todos.

UCT - Um-Contra-Todos

UHT – Usina Hidrelétrica de Tucuruí.

VC - Vapnik-Chervonenkis.

1.0 INTRODUÇÃO

1.1 MOTIVAÇÃO

As transformações no cenário energético brasileiro, que vêm ocorrendo nos últimos anos, impõem novas condições às empresas de geração, transmissão e distribuição de energia elétrica. Estas mudanças incluem o aumento da demanda por energia, que pode conduzir as máquinas a estados críticos de operação, por tempo prolongado (1). Além disso, as empresas precisam atender os procedimentos do Operador Nacional do Sistema (ONS) e os regulamentos da Agência Nacional de Energia Elétrica (ANEEL), que estabelecem as regras para garantir o fornecimento de energia.

Para atender às exigências do mercado e diminuir os elevados custos envolvidos para atender estas demandas, as companhias do setor elétrico estão investindo, cada vez mais, em estudos para buscar uma maior eficiência na operação dos equipamentos. Para isso, é fundamental elaborar técnicas adequadas para a medição, monitoração e diagnóstico dos equipamentos e do sistema elétrico como um todo, para evitar o desligamento não previsto de seus componentes, principalmente daqueles mais importantes e de maior custo, como os transformadores de potência e geradores.

Em vista disso, a manutenção preventiva, que se baseia no tempo de vida útil e de uso do equipamento, vem perdendo espaço para manutenção preditiva desde a última década, que se baseia nas condições de operação do equipamento. A manutenção preditiva permite que seja realizado o diagnóstico de possíveis falhas, que proporcionam o acompanhamento do estado operativo do equipamento ao longo do seu tempo de uso. Neste sistema, as paradas de máquinas para manutenção são melhores planejadas de forma que minimize ou elimine as perdas, evitando desligamentos não previstos que comprometem o fornecimento de energia elétrica. Com isso, a expectativa das empresas é aumentar, o tanto quanto possível, o tempo de vida útil e o intervalo entre manutenções dos equipamentos.

Pesquisas engendradas, a fim de compreender os problemas que afetam os hidrogeradores, apontaram que a maioria das falhas está relacionada com o aumento do nível de descargas parciais (DPs) na isolação elétrica da máquina. DPs são descargas elétricas intermitentes que atingem parcialmente o isolamento, ocorrendo em cavidades ou adjacente a um condutor. Sua presença tende a aumentar a degradação da isolação, pois provoca o rompimento dielétrico de outras cavidades, causando a perda das propriedades elétricas e

mecânicas do material isolante, levando o equipamento a graves danos. Assim, a presença das DPs constitui um excelente indicador da saúde do material isolante, possibilitando prever seu tempo de vida útil e falhas futuras. Sua detecção e medição é uma das principais estratégias utilizadas para detectar defeitos em máquinas rotativas, que podem ter sido originados durante a fabricação do equipamento, como também durante sua operação.

Apesar de constituir uma boa ferramenta de diagnóstico para os hidrogeradores, a medição de DPs faz-se acompanhada por diversos sinais ruidosos que se misturam aos sinais das DPs. Na maioria dos casos, as descargas ruidosas ocorrem com amplitudes e taxa de ocorrências similares, ou até mais elevadas, dificultando a caracterização dos dados de DPs e, conseqüentemente, inviabilizando um diagnóstico do estado operativo do equipamento.

Uma vez que as empresas do setor elétrico possuem a necessidade de ter confiabilidade na interpretação dos dados de DPs coletados, esta dissertação tem como objetivo realizar a classificação automática de padrões de DPs, utilizando técnicas de inteligência computacional para tratar os dados e, posteriormente, classificá-los. Os dados utilizados neste trabalho foram coletados e armazenados pelo software Instrumento de Monitoração e Análise de Descargas Parciais Autônomo (IMA-DP autônomo), desenvolvido em parceria entre CEPEL e ELETRONORTE, que possibilitou a realização desta empreitada.

Este trabalho propiciou a publicação de um artigo, com o título Caracterização de Padrões de Descargas Parciais em Hidrogeradores Utilizando Técnicas de Inteligência Computacional, no XXIII Seminário Nacional de Produção e Transmissão de Energia Elétrica (SNPTEE) (2).

1.2 REVISÃO BIBLIOGRÁFICA

Esta seção apresenta um resumo dos trabalhos relacionados aos métodos empregados para caracterização de DPs, abrangendo também, investigações na área das técnicas mais utilizadas de inteligência computacional para problemas que envolvem classificação de dados.

Os resultados obtidos com a pesquisa realizada por (3) possibilitaram a determinação dos padrões de vários tipos de fontes de DPs, através dos mapas de descargas parciais associadas à fase. Os padrões encontrados foram: descargas internas, descargas de ranhura, descargas na cabeça de bobina e descargas de delaminação. Tais resultados permitem um melhor diagnóstico do estado do gerador, com um impacto mínimo na operação, pois as medições podem ser levadas a cabo de modo online (máquina em operação). No ano seguinte,

na pesquisa de (4) procurou-se reduzir o risco de falsas interpretações de DPs, determinando a condição de isolamento do estator em relação a outras máquinas semelhantes.

No trabalho de (5), foi utilizado o método das curvas principais com o algoritmo k-segmentos não suaves para classificação de padrões de DPs, a partir de seus modelos experimentais. A técnica de curvas principais foi associada com redes neurais, distância euclidiana e método da votação, obtendo resultados significativos.

Em (6) foi realizada uma extensa pesquisa com diversos métodos de classificação, com experimentos com redes neurais e métodos de clusterização. O trabalho de (7) utilizou redes neurais para classificação de sinais de DPs em transformadores de potência. A técnica de classificação foi usada em conjunto com análise de componentes principais e análise de componentes principais não-lineares. Os melhores resultados foram obtidos com o uso de redes neurais e análise de componentes principais não-lineares. No trabalho de (1) foi desenvolvida uma pesquisa para caracterização de DPs em equipamentos de alta tensão. Utilizou-se análise de componentes principais para classificar sinais de DPs, análise de componentes independentes para extração de ruído e Discriminador híbrido com Redes Neurais para classificação.

Na pesquisa realizada a respeito das técnicas inteligentes, mais recentemente utilizadas, apontou-se que o uso de máquinas de vetores de suporte é um bom método para problemas que envolvem dados multidimensionais, com bons resultados para casos de classificação. No trabalho de (8) foram testados diferentes configurações da função kernel para máquinas de vetores de suporte, com aplicações em grandes massas de dados e várias classes, atingindo bons resultados de acurácia.

Floresta randômica e k-vizinhos mais próximos também são técnicas que mostraram-se eficientes em casos de classificação. Em (9) foi exposto como as florestas randômicas são capazes de melhorar a precisão das estimativas dos classificadores resultantes. O trabalho de dissertação de (10) abordou a comparação de variações do método de k-vizinhos mais próximos para aumentar a eficácia da classificação automática de textos. No trabalho de (11), em problemas envolvendo modelagem de distribuição potencial, foram testadas nove técnicas de classificação de dados, onde floresta randômica foi a que apresentou os melhores resultados de acurácia na grande maioria dos testes realizados.

No Guia IEEE 1434 (12) é apresentada uma revisão da natureza das DPs nos enrolamentos do estator, técnicas de medição nas condições da máquina em operação e fora de operação, assim como em bobinas de múltiplas espiras e meias bobinas. Na norma internacional IEC 60270 (13), são ditadas regras para medições de DPs em ensaios com alta

tensão, as principais grandezas relacionadas às DPs, os tipos de circuitos utilizados para medição, os sistemas de medição e os procedimentos utilizados para a calibração do circuito. O Cigrè (14) realizou um estudo com hidrogeradores em diferentes países do mundo. Nesta pesquisa foram obtidos dados estatísticos que ajudam a avaliar os riscos das principais falhas dos hidrogeradores, sendo as DPs apontadas como a principal causa de falhas elétricas.

Na norma IEC 60034-27-2 (15), sobre medição *on-line* de DPs em máquinas rotativas, são apresentados os assuntos já abordados nas normas já citadas, e também técnicas de supressão de ruído e os padrões de mapas estatísticos levantados por (3).

1.3 OBJETIVOS

1.3.1 Objetivo geral

Esta dissertação possui como objetivo geral apresentar uma metodologia de classificação automática dos padrões de DPs, através do uso de técnicas de inteligência computacional e, tratar os dados de DPs na tentativa de minimizar os ruídos acoplados aos sinais.

1.3.2 Objetivos específicos

O objetivo geral desta dissertação proporciona a realização dos seguintes objetivos específicos:

- Propiciar conhecimento acerca dos fenômenos que envolvem as DPs, conceitos, natureza física, padrões e métodos de medição;
- Apresentar os aspectos construtivos de um hidrogerador, dando destaque para a isolamento elétrica dos enrolamentos estatóricos;
- Investigar técnicas de inteligência computacional mais recentemente empregadas para a classificação de dados;
- Tratar os dados de DPs, de forma que, os ruídos sejam minimizados o tanto quanto possível, para facilitar o processo de classificação dos padrões de DPs;
- Investigar uma linguagem de programação que ofereça amplas vantagens para a

realização dos presentes objetivos;

- Comparar e avaliar as respostas obtidas das classificações realizadas por cada técnica empregada.

1.4 ESCOPO DO TRABALHO

Esta dissertação foi organizada em sete capítulos, conforme explanado a seguir.

No primeiro capítulo, com o tema **Introdução**, são apresentadas as motivações que levaram à produção deste trabalho. Também é abordada, de modo geral, a importância de se compreender e desenvolver novas técnicas que venham aprimorar os métodos de detecção, medição e caracterização de DPs, para as empresas do setor elétrico manterem uma boa qualidade no fornecimento de energia aos seus consumidores. Além disso, é retratada a dificuldade no processo de caracterização das DPs devido às interferências dos sinais ruidosos que se misturam aos sinais de DPs.

Neste capítulo também é apresentada uma breve revisão bibliográfica acerca de trabalhos desenvolvidos que envolvem a caracterização de DPs, como também, da aplicação das técnicas de inteligência computacional escolhidas para este trabalho. Em seguida, no capítulo são descritos os objetivos, geral e específico, que envolvem a realização deste trabalho.

No segundo capítulo, com o tema **Descargas Parciais**, é realizado um estudo geral, onde é apresentado o conceito de DPs, abordando a estrutura da máquina em quem ocorrem e quais os fatores que desencadeiam o surgimento das descargas. A natureza física e o circuito equivalente de uma DP são apresentados para proporcionar melhor esclarecimento do que acontece, a nível estrutural, quando o isolamento é afetado por uma DP. Em seguida são apresentados os tipos de DPs, conforme catalogado em norma, segundo suas fontes de ocorrência.

Neste capítulo também são abordados os mapas PRPD onde os dados de DPs são condensados e armazenados, registrando o número de pulsos de DP com suas amplitudes e fases. Para finalizar, são apresentados os métodos utilizados para detecção e medição de DPs e as dificuldades inerentes a este processo.

No terceiro capítulo, com o tema **Os Enrolamentos Estatóricos do Hidrogerador**, são apresentadas, de modo sucinto, as estruturas básicas destas máquinas dando enfoque nos

enrolamentos estáticos e os isolamentos, que fazem parte de suas estruturas, apresentando os modelos de barras, os tipos de isolamentos que as envolvem e um breve histórico do material constituinte dos isolamentos. Em seguida são apresentados resultados estatísticos, de uma pesquisa do CIGRÈ, realizada para investigar as causas de falhas que afetam os hidrogeradores.

No quarto capítulo, com o tema **Classificadores Baseados em Técnicas de Inteligência e Tratamento de Ruídos**, são abordados os fundamentos das técnicas Análise de Variância e Análise de Componentes Principais utilizadas na tentativa de reduzir a dimensão dos dados de DPs. A técnica Análise de Componentes Independentes também foi utilizada na tentativa de separar dos dados de DPs os sinais ruidosos. Em seguida são abordadas as técnicas utilizadas para a classificação dos dados de DPs, quais foram: K-Vizinhos Mais Próximos, Floresta Randômica e Máquinas de Vetores de Suporte.

No quinto capítulo, com o tema **Metodologia**, são apresentados o cronograma e os procedimentos metodológicos que envolveram a confecção desta dissertação. Em seguida, é explanada a origem dos dados DPs coletados, os tipos de DPs encontrados e o instrumento de medição utilizado em conjunto com o sistema IMA-DP.

Neste capítulo também são retratados o software e a linguagem de programação escolhidas para este trabalho e o tratamento prévio realizado nos dados de DPs antes de serem processados pelas técnicas citadas no capítulo anterior.

No capítulo 6, com o tema **Experimentos e Resultados**, estão expostos quais experimentos e resultados atingidos de acordo com as técnicas descritas no capítulo anterior.

Para finalizar, no capítulo 7, com o tema **Considerações Finais**, são relatadas as principais conclusões obtidas ao longo desta dissertação e algumas propostas para trabalhos futuros.

2.0 DESCARGAS PARCIAIS

2.1 INTRODUÇÃO

A presença de DPs é indicativo da ocorrência de falha de isolamento no sistema. Esse fenômeno contribui para a degradação do material afetado, acelerando o aparecimento de novas DPs e diminuindo o tempo de vida útil do dispositivo. Assim, o estudo das DPs constitui uma importante ferramenta para a análise da qualidade dos equipamentos de alta tensão, como também, para acompanhar seu desempenho operativo.

Este capítulo apresenta, em visão geral, os conceitos e características associadas às DPs, mostrando sua natureza física. Devido ao tema deste trabalho, uma atenção especial será dada para as DPs que ocorrem nos isolamentos dos estatores de hidrogeradores, no que concerne aos tipos de padrões existentes catalogados em norma e suas origens.

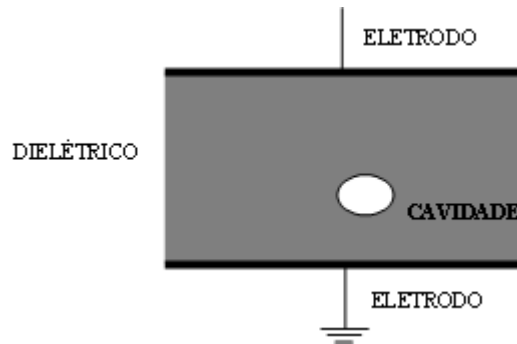
2.2 O CONCEITO DE DESCARGAS PARCIAIS

Materiais isolantes submetidos a valores de tensão acima do limite de sua suportabilidade dielétrica ficam susceptíveis à ocorrência de DPs. Quando essas descargas ocorrem em torno de um elemento condutor, no ar, as descargas caracterizam o chamado efeito corona, sendo visíveis à percepção humana (7).

Estas descargas também podem ocorrer no interior dos equipamentos, sendo nas superfícies condutoras, entre dois diferentes tipos de materiais isolantes ou nas falhas dos materiais isolantes, compreendendo estas falhas pequenas bolhas ou cavidades gasosas. Essas descargas são conhecidas como DPs (7).

De acordo com (12) e (13), DPs são descargas elétricas ocasionadas pela ionização do meio gasoso nas pequenas cavidades dos materiais dielétricos ou em suas superfícies, através do aumento da tensão elétrica até o limiar da suportabilidade dielétrica do meio gasoso. São chamadas DPs por ocorrerem apenas no interior da cavidade, não percorrendo inteiramente o caminho entre o isolamento e o condutor. Na Figura 2.1 é ilustrada uma cavidade no interior do dielétrico.

Figura 2.1 - Representação esquemática de uma cavidade no interior do material dielétrico.

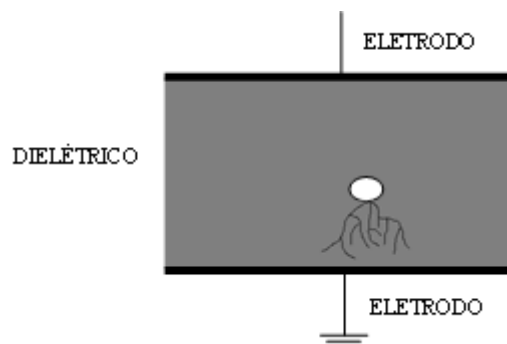


Fonte: Elaborado pelo autor.

As DPs, dependendo de sua intensidade e, ao longo do tempo, alastram-se por toda a estrutura do material isolante comprometendo sua função, como também, diminuindo o tempo de vida útil do equipamento de alta tensão.

As pequenas cavidades no material isolante são defeitos originados em deficiências nas suas tecnologias de fabricação. Com o decorrer do tempo, as DPs nessas cavidades aceleram o processo de envelhecimento do material, que segundo (16), “[...] causa alterações nas propriedades das superfícies internas dos vazios, e a quebra das moléculas do polímero inicia a formação dos caminhos condutores”. Esses caminhos condutores são finos canais erosivos que constituem a chamada arborescência elétrica, altamente prejudicial para o equipamento, pois sua proliferação provoca a ruptura das propriedades dielétricas do material isolante, danificando-o. Na Figura 2.2 está ilustrado o fenômeno da arborescência elétrica:

Figura 2.2 - Representação esquemática do fenômeno de arborescência elétrica no interior do material dielétrico.



Fonte: Elaborado pelo autor.

2.3 FÍSICA DAS DESCARGAS PARCIAIS

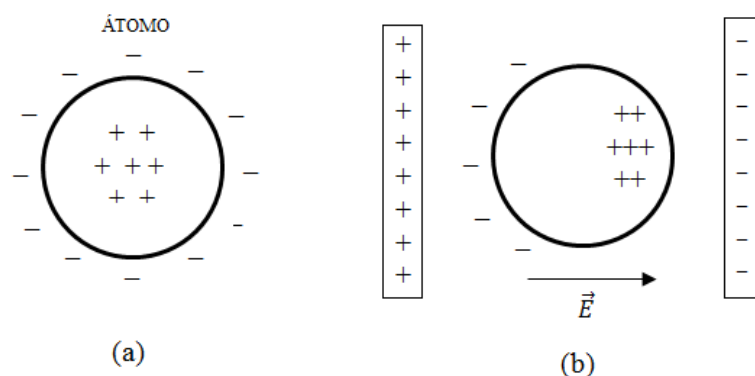
2.3.1 Da ionização dos átomos à avalanche de elétrons

Para compreender a natureza física do fenômeno das DPs é necessário abranger como ocorre a ionização das partículas gasosas até a avalanche de elétrons.

Segundo (16), dois requisitos são necessários para a ocorrência de DPs nos materiais que são fabricados os isolantes (poliméricos): o campo elétrico na cavidade gasosa deve ultrapassar o valor crítico e a existência de elétrons livres.

O processo de ionização começa com o aumento da tensão elétrica que provoca o aumento do campo elétrico no interior das cavidades gasosas. Além disso, a ionização também depende dos gases que estão preenchendo a cavidade, da pressão exercida por eles, e do produto da pressão do gás pela distância entre os materiais condutores que estão envolvidos pelo isolamento, funcionando como eletrodos. A partir disso, o campo elétrico exerce uma força sobre os elétrons da camada de valência mais externa, provocando um deslocamento das cargas positivas do átomo, para o mesmo sentido do campo, e deslocando as cargas negativas no sentido contrário ao do campo elétrico. Então, os átomos ficam polarizados, comportando-se como dipolos, conforme observa-se esquematicamente na Figura 2.3, a seguir:

Figura.2.3 - Formação de dipolos. a) Átomo não polarizado. b) Átomo polarizado comportando-se como dipolo sob a ação de um campo elétrico.



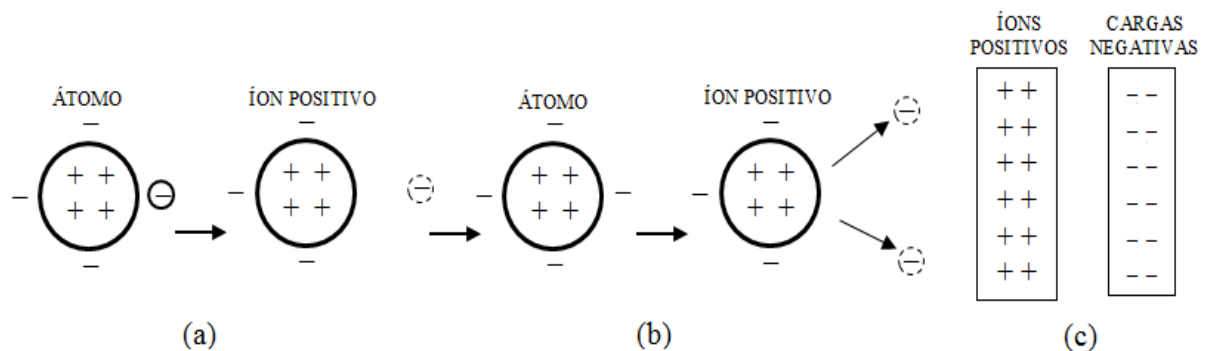
Fonte: Elaborado pelo autor.

Quando o campo elétrico excede o limite da suportabilidade dielétrica dos gases nas cavidades, os elétrons das últimas camadas de valência ganham energia para se desprender da sua órbita e, o átomo fica positivamente ionizado. O elétron desprendido é acelerado sob a

ação do campo elétrico. Quanto maior a intensidade do campo maior será a aceleração do elétron.

O elétron acelerado irá iniciar o processo da avalanche de elétrons. De acordo com (1), esta consiste na colisão do elétron em alta velocidade com um átomo ou molécula. Durante a colisão, se a transferência de energia do elétron para o átomo ou molécula não for suficiente, não haverá desprendimento de outro elétron e nem a formação de outro íon. Essa colisão é dita elástica. Entretanto, se houver transferência de energia suficiente, outro elétron livre e íon positivo serão gerados, caracterizando a colisão inelástica. O elétron inicial, após a colisão, terá sua velocidade significativamente diminuída e o novo elétron livre irá colidir com outros átomos. Na Figura 2.4 está ilustrado o processo de avalanche de elétrons:

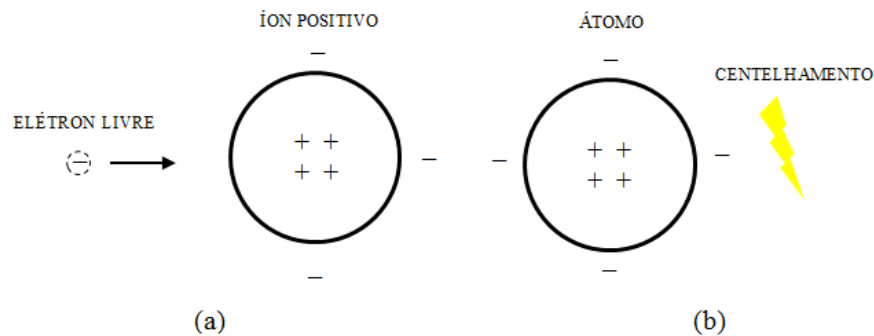
Figura 2.4 - Processo de avalanche de elétrons. a) Elétron ganha energia tornando-se elétron livre e átomo torna-se íon positivo. b) Elétron livre colide com átomo e libera novo elétron livre e íon positivo. c) A sucessão desse processo resulta em várias cargas pos positivas e negativas que resultam na avalanche de elétrons.



Fonte: Adaptado de (1).

Os íons positivos possuem massa muito maior que a massa do elétron, por conta disso, movem-se numa velocidade bem menor que os elétrons, no mesmo sentido que o campo elétrico. Quando o íon consegue atrair e reincorporar o elétron na sua última camada de valência, um quantum de energia, com a mesma magnitude que gerou o elétron livre, é liberado. Esse quantum corresponde a uma radiação na faixa de luz perceptível ao ser humano, na forma de um centelhamento, conforme observa-se na Figura 2.5. Esse retorno dos íons positivos aos átomos neutros continuará e cessará o processo da DP (1).

Figura 2.5 - Processo de centelhamento. a) Elétron livre é reincorporado no íon positivo formando novamente o átomo. b) Na formação do átomo um quantum de energia liberado gera luz visível na forma de um centelhamento.



Fonte: Elaborado pelo autor.

Em (17), a quantidade de elétrons liberados na avalanche de elétrons, num período estimado de 100 nanosegundos, pode variar de centenas de elétrons por segundo até 10^{22} elétrons por segundo. A corrente elétrica gerada durante a DP pode variar de 10^{-17} A até alguns milhares de Ampères (1).

2.3.2 Forma do pulso das DPs

As formas de onda dos sinais medidos de DPs dependem dos equipamentos utilizados para o sistema de medição e pelas características físicas das descargas.

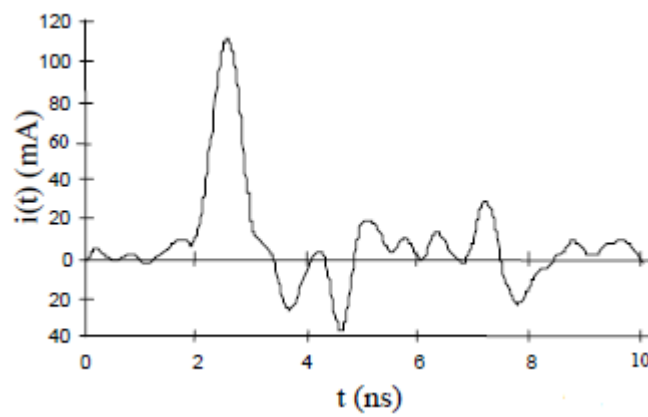
Cada sinal de DP possui diferentes valores de amplitude, largura (tempo de duração), frequência, tempo de subida e tempo de descida. Esses parâmetros dependem de vários fatores, dentre eles (12) e (17):

- O tamanho, a geometria e a localização das cavidades gasosas, que podem ser encontradas tanto no interior do material isolante como nos seus revestimentos semi-condutores;
- O gás que preenche as cavidades e sua pressão;
- O valor da sobretensão exercida sobre a cavidade;
- O material dos isolamentos;
- O valor de impedância da fonte de tensão aplicada;
- A geometria da máquina;

- A propagação do sinal no instrumento de medição, que possui circuito de medição, partes condutoras e linhas de transmissão.

De modo geral, o tempo de subida das DPs é da ordem de nanossegundos e o tempo de descida é da ordem de microssegundos, logo, caracterizando uma função do tipo impulso. Os pulsos de DPs apresentam também oscilações, onde considera-se o primeiro pico para cálculo dos parâmetros do sinal. Na Figura 2.6 o pulso de DP apresenta pico em 2,5 ns, com duração de descarga de aproximadamente 10 ns:

Figura 2.6 - Pulso de DP medido através da variação de corrente pelo tempo.



Fonte: Adaptado de (1).

Em (17), o tempo de subida do pulso corresponde ao movimento dos elétrons livres em direção contrária ao do campo, ou seja, na direção que atua como ânodo. No tempo de descida, os íons positivos, maiores e mais pesados que os elétrons livres, movimentam-se a favor do campo elétrico, ou seja, na direção que atua como cátodo. Em gases eletronegativos, como o CO_2 , alguns íons negativos formados movimentam-se na direção do ânodo, no tempo de descida do pulso.

Os sinais de DPs, numa mesma avalanche de elétrons, podem apresentar diversas bandas de frequência, de modo que, conforme o objetivo e instrumento de medição disponível, pode-se analisar somente uma faixa da banda de frequência. Em (17) e (18), a maioria dos instrumentos de medição detectam o início do pulso de DP, que tem um tempo de subida de 1 a 5 ns, através da corrente elétrica nos condutores provocada pelas DPs nos isolamentos.

2.4 CIRCUITO EQUIVALENTE DE UMA CAVIDADE

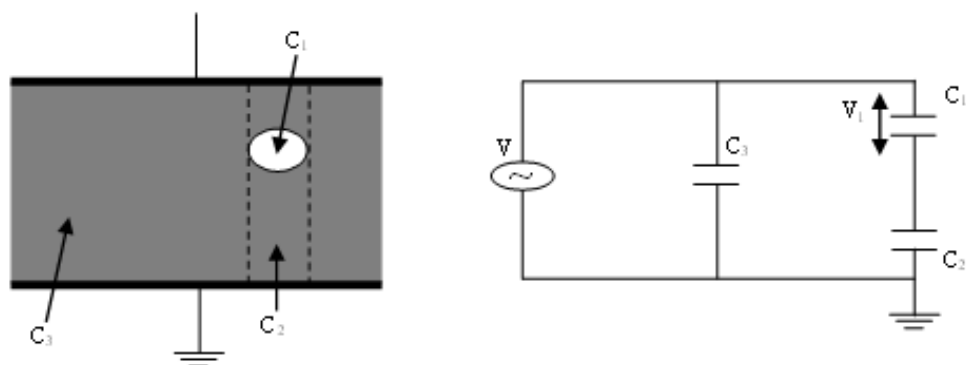
As cavidades presentes nos materiais dielétricos possuem constante dielétrica menor que os materiais de que são fabricados os isolantes. Quando uma tensão é aplicada nos terminais dos isolamentos, o campo elétrico e a diferença de potencial elétrico gerada no interior da cavidade são maiores do que no material dielétrico. Em (18), a Lei de Paschen determina que a tensão elétrica na cavidade é função do produto entre a pressão do gás (p) e o diâmetro da cavidade (d), de acordo com a Equação 2.1.

$$V = f(pd) \quad (2.1)$$

Ao ultrapassar o limite de suportabilidade dielétrica do gás, a tensão na cavidade gera a descarga elétrica, conhecida como DP. Segundo (15), a tensão nas cavidades pode exceder o limite de suportabilidade do gás mesmo que a máquina esteja operando com sua tensão de trabalho nominal, pois a magnitude de tensão no gás que preenche as cavidades sempre será maior do que nos isolamentos.

Conforme a Figura 2.7, considerando um isolamento entre dois condutores, pode-se dividi-lo em três partes: a capacitância da cavidade C_1 , a capacitância do dielétrico que está em série com a cavidade C_2 , e a capacitância do restante do dielétrico C_3 .

Figura 2.7 - Circuito equivalente de uma cavidade no interior do dielétrico.



Fonte: Elaborado pelo autor.

Quando a tensão V no condutor é aplicada, a tensão na cavidade V_1 é dada pela Equação 2.2:

$$V_1 = \frac{V d_1}{d_1 + \left(\frac{\epsilon_0}{\epsilon_1}\right) d_2} \quad (2.2)$$

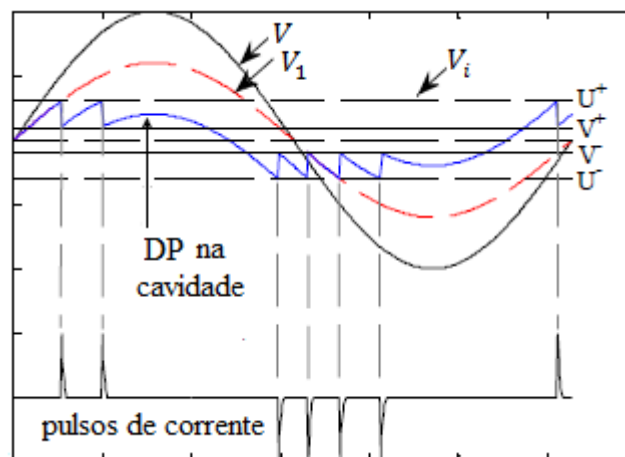
onde d_1 e d_2 são, respectivamente, a espessura da cavidade e do dielétrico, e ϵ_0 e ϵ_1 são suas respectivas permissividades. Comumente, $d_1 \gg d_2$, de tal forma que, assume-se que a cavidade é totalmente preenchida por um gás, transformando a Equação (2.2) na Equação (2.3):

$$V_1 = V \epsilon_r \frac{d_1}{d_2} \quad (2.3)$$

onde, ϵ_r é a permissividade relativa do dielétrico.

Quando aplica-se V aos condutores, V_1 aumenta até atingir o valor de tensão de ruptura, V_i , da resistência dielétrica do gás. Nesse momento, a descarga na cavidade acontece, V_1 torna-se nulo e a centelha da descarga extingue-se. Então, novamente a tensão volta a subir e todo o processo repete-se. Se a tensão V aplicada for alternada, as descargas ocorrem em ambos os meios ciclos. Na Figura 2.8 é ilustrada a forma de onda das descargas de acordo com a curva de DP na cavidade.

Figura 2.8 - Pulsos de DPs em uma cavidade.



Fonte: Adaptado de (19).

2.5 TIPOS DE DP

Os equipamentos que compõem um sistema de alta tensão possuem em seus sistemas físicos os mais diversos designs, materiais, tipos de manutenção e condições operacionais que influenciam a localização, características, quantificação, nível de periculosidade e a forma de evolução da DP (12).

2.5.1 Classificação geral das DPs

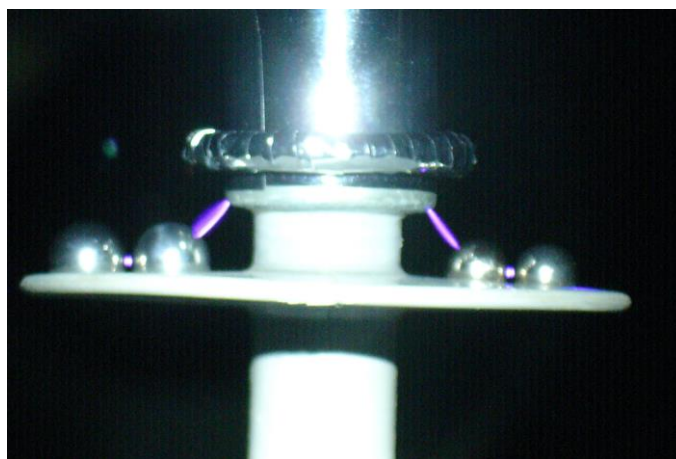
De um modo geral, as DPs podem ser classificadas em três grupos: descargas parciais internas, descargas parciais superficiais e descargas corona.

As descargas parciais internas ocorrem em cavidades no interior de dielétricos sólidos ou líquidos, ou na região de interface entre o dielétrico e o condutor. A constante ocorrência das DPs provoca desgaste do material dielétrico na superfície das cavidades, formando caminhos condutores conhecidos como arborescência elétrica.

As descargas parciais superficiais localizam-se na superfície do material dielétrico. Geralmente ocorrem entre a superfície do dielétrico e um condutor, causado pelo excessivo aumento do campo elétrico nessa região.

As descargas corona ocorrem no ar, em torno de pontas agudas de eletrodos metálicos. O elevado campo elétrico nessas regiões é explicado pelo poder das pontas. Em (16), a descarga pode ser facilmente visualizada quando a tensão aplicada é alternada e senoidal, pois na máxima magnitude do semi-ciclo negativo tem início as descargas corona, sendo seu efeito visual conhecido como efeito corona.

Figura 2.9 - Corona no isolador.



Fonte: (20)

2.5.2 Classificação das DPs em máquinas elétricas rotativas

Nas máquinas elétricas rotativas - em especial, os hidrogeradores - a análise dos diferentes padrões de DPs, no sistema de isolamento do estator, é de extrema importância para o acompanhamento do seu estado operativo e para o planejamento adequado de manutenções. Através dos padrões de DP é possível saber o nível de periculosidade das descargas que estão afetando o equipamento, estimar o tempo de vida da máquina, a localização das DPs, entre outros. As descargas coronas não ocorrem nas máquinas rotativas, pois seu design não propicia ambiente ao aparecimento destas.

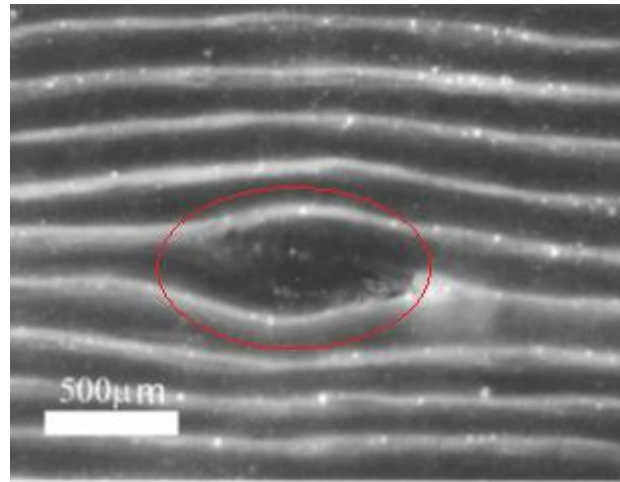
De acordo com (15), existem quatro grupos onde as fontes de DPs são classificadas: descargas internas, descargas de ranhura, descargas na cabeça da bobina e descarga por partículas condutivas.

2.5.2.1 Descargas internas

2.5.2.1.1 Cavidades internas

Pequenas cavidades são formadas durante o processo de fabricação dos materiais isolantes de alta tensão, compostos por fitas de mica impregnadas em resina. Nas metodologias de fabricação empregadas busca-se reduzir o número de cavidades presentes no material, porém não consegue-se ainda o intento de eliminá-las por completo. Entretanto, a quantidade de cavidades presentes após a fabricação é pequena, de modo que, uma determinada presença de descargas parciais em isolantes novos seja aceitável. Além disso, o material dos isolantes é projetado para que suporte as DPs por um longo período sem que estas atravessem completamente o isolamento, destruindo-o. Na Figura 2.10 está ilustrado um exemplo de cavidade interna:

Figura 2.10 - Imagem por microscopia óptica de um corte da seção transversal do isolamento epóxi-mica de uma barra estatórica de hidrogerador com um vazio na região do epóxi.

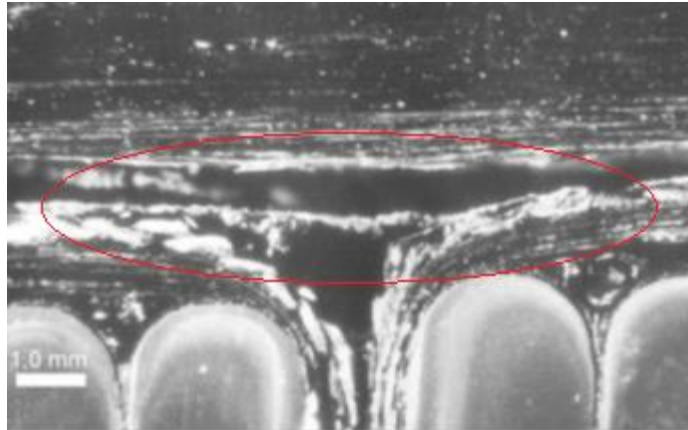


Fonte: (16).

2.5.2.1.2 Delaminação interna

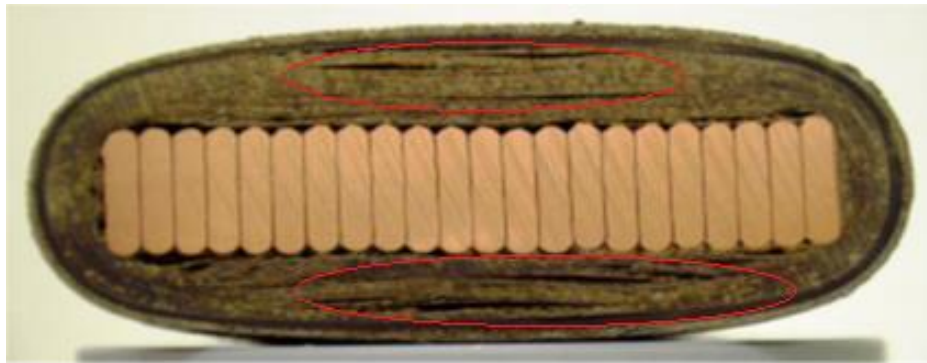
Delaminações internas são grandes cavidades que podem originar-se no processo de fabricação ou através de sobretensões e perturbações mecânicas ou térmicas durante a operação da máquina. DPs em delaminações possuem uma quantidade de energia relativamente alta que podem corroer o isolamento comprometendo a sua condutividade térmica. Em (18), “quando ocorre a delaminação das camadas de isolação, os condutores internos de cobre ficam livres para vibrar e então ocorrem descargas parciais de alta energia [...]”. Análises cuidadosas são necessárias na ocorrência deste tipo de descarga, pois sua evolução pode destruir o isolamento. Nas Figuras 2.11 e 2.12 estão ilustrados exemplos de delaminações internas:

Figura 2.11 - Imagem por microscopia óptica de um corte da seção transversal do isolamento de uma barra estatórica de hidrogerador com grande delaminação do compósito epóxi-mica.



Fonte: (16).

Figura 2.12 - Fotografia da seção transversal de uma barra estatórica, onde se podem observar delaminações internas no isolamento.



Fonte: (17).

2.5.2.1.3 Delaminação entre condutores e a isolação

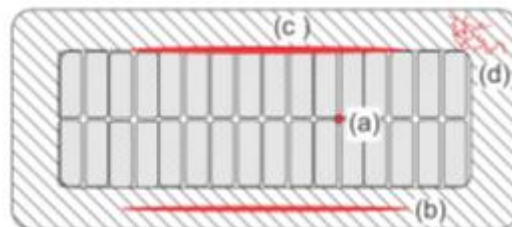
Delaminações podem ocorrer na interface do isolamento principal com o condutor, tendo como causa os ciclos térmicos de operação nesta área de contato. Este tipo de DP pode evoluir rapidamente destruindo o isolamento, e segundo (18), pode levar às falhas principalmente em bobinas de múltiplas espiras.

2.5.2.1.4 Arborescência elétrica

A arborescência elétrica pode ter origem onde o campo elétrico esteja concentrado no interior do isolamento, como em cavidades, delaminações, estruturas pontiagudas ou mesmo por impurezas (17). Constitui-se num processo de erosão das DPs internas, em que sucessivas descargas provocam a quebra das moléculas do polímero que constituem o dielétrico, formando finos canais erosivos que proliferam-se ao longo do isolante. Esse processo ocorre ao longo do tempo levando o isolante à perda de suas propriedades elétricas, possibilitando a ocorrência de descargas completas entre os condutores em torno do isolamento e danificando o equipamento.

A Figura 2.13 ilustra as descargas internas acima explanadas, numa seção transversal de uma barra estatórica:

Figura 2.13 - Fontes de descargas internas na seção transversal de uma bobina estatórica. a) Cavidade interna. b) Delaminação interna. c) Delaminação entre condutores e isolamento. d) Arboreamento elétrico.



Fonte: (17).

2.5.2.2 Descargas de ranhura

Descargas de ranhura ocorrem quando o revestimento da parte condutora na ranhura danifica-se, causado pelo movimento da barra ou da bobina na ranhura em virtude de defeito de fabricação ou montagem, erosão do material, desgaste químico ou afrouxamento na fixação.

Descargas de ranhura são descargas de alta energia, que evoluem num tempo bem mais curto comparado a outros tipos de DPs, e geralmente aparecem junto de sérios danos mecânicos. Ocorrem nos terminais de alta tensão de cada fase da máquina, principalmente se houver vibração das barras ou bobinas. É de suma importância a detecção desse tipo de DP

ainda nos seus estágios iniciais, para que medidas adequadas de manutenção possam ser planejadas.

2.5.2.3 Descargas na cabeça da bobina

Descargas na cabeça da bobina ocorrem nas interfaces entre diferentes elementos da cabeça da bobina dos enrolamentos do estator (19).

2.5.2.3.1 Descargas de superfície

Descargas de superfície ocorrem quando o valor do campo elétrico em uma superfície rompe o limite de rigidez dielétrica do gás ao seu redor. Originam-se na ausência, desgaste ou contaminação por poeira, umidade, etc., do revestimento semicondutor em uma barra estatórica na região da cabeça da bobina. Esse tipo de DP geralmente provoca falhas entre fase e terra, tendo um processo erosivo lento, mesmo que os efeitos de superfície provoquem alterações relativamente rápidas na dinâmica das DPs.

2.5.2.3.2 Descargas entre fases

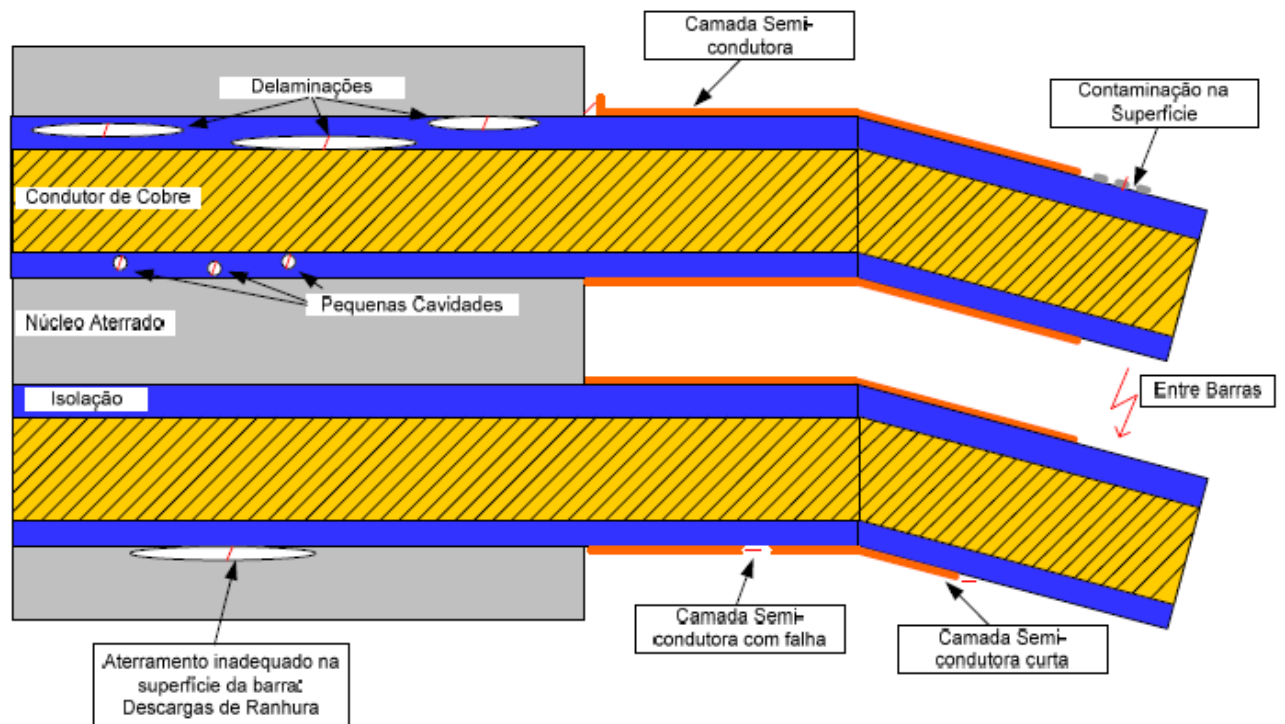
DPs entre fases ocorrem em virtude de uma distância inadequada de isolamento entre fases ou devido ao uso incorreto dos sistemas de suporte da cabeça da bobina. Esse tipo de DP pode ter amplitudes elevadas e resultar em falhas entre fase e fase.

2.5.2.4 Partículas condutoras

Contaminações nos enrolamentos podem resultar em uma forte concentração de DPs, devido à presença de pequenas partículas condutoras. Essas DPs podem provocar pequenas cavidades no isolamento, através do processo erosivo.

Na figura 2.14 estão ilustradas algumas das fontes de DPs no isolamento do estator mencionadas anteriormente.

Figura 2.14 - Fontes de DPs no isolamento do estator.



Fonte: Adaptado de (21).

De acordo com o nível de periculosidade para o bom funcionamento da máquina, associado às principais fontes de DP, a norma (15) classificou as DPs em três níveis:

- Baixo risco: Cavidades internas, normalmente, não reduzem aceleradamente o tempo de vida útil da máquina;
- Médio risco: Descargas na cabeça na bobina e por partículas condutoras;
- Alto risco: Delaminações internas e descargas de ranhura reduzem drasticamente o tempo de vida útil da máquina, de modo que, nos primeiros estágios dessas descargas é necessário algum tipo de intervenção para manutenção.

2.6 GRANDEZAS ASSOCIADAS ÀS DPS

Alguns conceitos relacionados às grandezas acerca das DPs serão relacionados a seguir:

Carga Aparente (q):

O valor absoluto da carga aparente (q) de um pulso DP é a carga que se injeta em um tempo curtíssimo entre os terminais do objeto de teste de um circuito, variando a tensão entre seus terminais com valor igual ao da DP. A carga aparente é normalmente expressa em

picocoulombs (PC).

Taxa de Repetição de Pulso (n):

É a relação entre o número total de pulsos de DPs medidos em um dado intervalo de tempo, e a duração deste intervalo de tempo. Em (18), geralmente são considerados pulsos de DP acima de uma determinada intensidade.

Ângulo de Fase das DPs (ϕ_i):

É a posição em que ocorre o pulso de DP em relação ao período T da senóide de tensão aplicada, de acordo com a Equação x. O ângulo de fase é expresso, em graus ($^\circ$), $\phi_i = 360(t_i/T)$, onde t_i é o tempo de ocorrência dos pulsos de DP e T é o período da tensão aplicada.

Amplitude (mV):

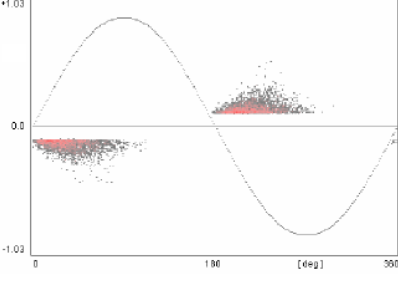
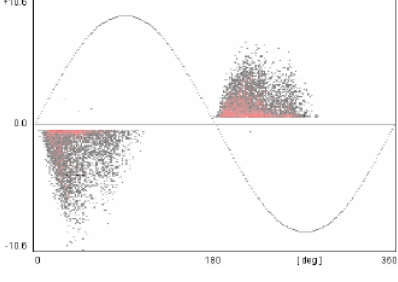
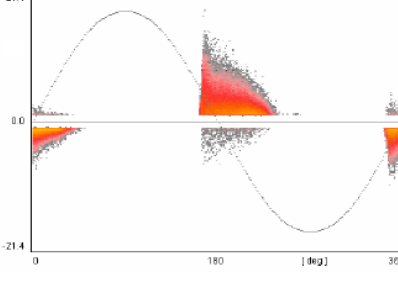
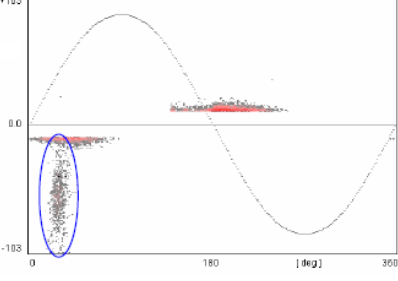
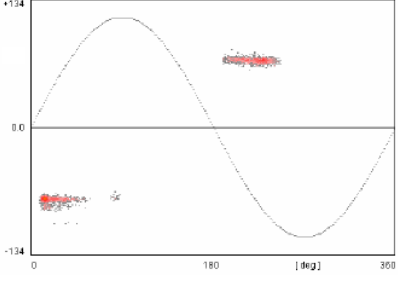
É o valor medido, geralmente em milivolts, do pulso de tensão elétrica gerado pelas DPs.

2.7 GRANDEZAS ASSOCIADAS À FASE – OS MAPAS ESTATÍSTICOS

Uma medição de DPs contém uma quantidade muito grande de informações a serem processadas e armazenadas. O método utilizado para condensar e armazenar esses dados, conforme (15), registra o número de pulsos de DP com suas amplitudes e fases. É conhecido como Mapas Estatísticos de DPs ou Mapas PRPD (*Phase Resolved Partial Discharge – PRPD*).

Nesse tipo de análise, a tensão de ensaio é considerada constante e os ângulos de fase são divididos em faixas, considerando um tempo previamente determinado para a medição. O modo de visualização é comumente bidimensional, onde os valores de amplitude estão representados no eixo das ordenadas, os ângulos de fase estão representados no eixo das abcissas e o número de pulsos de DP para cada combinação de amplitude e fase é representado por variação de cores. Na Tabela 2.1 observa-se os mapas estatísticos dos padrões de DPs, de acordo com (15):

Quadro 2.1 - Padrões de mapas estatísticos de DPs.

Tipos de descargas	Mapas estatísticos
Descargas Internas	
Delaminação Interna	
Descargas de Ranhura	
Descargas de Superfície	
Descargas entre fases	

Fonte: Adaptado de (3).

Durante a medição das DPs sinais ruidosos são também captados interferindo na leitura dos sinais de DP provenientes dos enrolamentos estáticos. Assim, faz-se necessário adotar medidas para suprimir os ruídos, pois estes confundem-se aos sinais de DP, proporcionando

leitura de falsos positivos e comprometendo a avaliação do estado operacional da máquina. Ou então, os ruídos podem inviabilizar qualquer tentativa de diagnóstico da máquina. Podem ter como origem diversas fontes, dentre elas, ruídos térmicos, chaveamento de dispositivos, interferências do sistema de medição.

2.8 MÉTODOS DE MEDIÇÃO DAS DPS

As sucessivas e intermitentes DP's produzem diversos efeitos que proporcionam sua detecção e medição, através de técnicas que identificam e mensuram esses efeitos, a fim de obter os valores intrínsecos aos pulsos de DP's. Em (12), esses efeitos incluem fenômenos luminescentes, reações químicas, ondas acústicas, pulsos elétricos e pulsos de rádio frequência. Além desses, as DP's também podem apresentar outras manifestações como vibrações mecânicas e efeitos térmicos.

As técnicas de detecção e medição podem ser divididas em dois grupos: os métodos elétricos e os não elétricos. O método elétrico funciona com a inserção de um circuito de detecção e medição nos enrolamentos estatóricos. Os métodos não elétricos, comumente, servem de complemento ou suporte aos métodos elétricos (5).

2.8.1 Métodos não elétricos

A seguir serão brevemente explanados alguns métodos não elétricos, conforme (12):

2.8.1.1 Método acústico e ultrassônico

As perturbações acústicas originam-se das pequenas ondas de choque causadas pelo súbito aumento da temperatura do gás em volta do pulso de DP. Quando muitos pulsos ocorrem na superfície das bobinas, ruídos audíveis podem ser detectados. O ruído ocorre no intervalo de várias centenas de hertz e até acima de 150 kHz, com grande parcela da energia acústica variando em torno de 40 kHz.

Para a captação dos sinais audíveis são utilizados microfones direcionais que, também, podem localizar as DP's superficiais. Entretanto, DP's no interior da parede de aterramento não são detectadas, a menos que o nível de energia nas DP's esteja em nível elevado.

Hastes de fibra de vidro podem ser utilizadas em conjunto com os microfones direcionais, atuando como guias dos sinais acústicos. Proporcionam isolamento elétrico entre o sensor e o equipamento de teste. Na figura 2.15 estão ilustrados alguns sensores acústicos instalados em um transformador de potência:

Figura 2.15 – Sensores acústicos instalados em um transformador de potência.



Fonte: Adaptado de (22).

2.8.1.2 Detecção de ozônio

Em máquinas com arrefecimento a ar, a presença de DPs na superfície das bobinas ou barras provoca reações químicas no ar no entorno imediato, produzindo gás ozônio. O aumento de DPs superficiais contribui para o aumento da concentração de ozônio, no entanto, DPs adjacentes aos condutores ou no interior da parede de aterramento não geram uma quantidade de ozônio que seja mensurável.

Diversos fatores influenciam na concentração do ozônio como a temperatura, fluxo do ar, umidade do ambiente, e até mesmo, pode estar relacionado à carga da máquina e ao fator de potência.

Para detecção e medição da concentração de ozônio podem ser utilizadas técnicas como tubos químicos até sensores. Essa técnica é mais eficiente para monitoramentos on-line.

2.8.1.3 Teste de black-out

O teste de black-out é um método empregado que consiste em energizar a bobina e, através de inspeção visual à distância segura, verificar em completa escuridão a atividade das DPs. Essa técnica é útil para verificar a presença e localização de DPs superficiais em revestimentos do controle de estresse, aberturas de ar em cabeças de bobinas, como também, para DPs externas com tensão fase-fase e para DPs de ranhura em bobinas individuais. Esse método também é útil para localizar rachaduras.

2.8.1.4 Teste de corona

É um método similar ao teste de black-out, com a vantagem de que pode ser realizado em condições normais de iluminação e o instrumento utilizado funciona como um visor de raios ultravioletas, proporcionando maior segurança e confiabilidade.

2.8.2 Métodos elétricos

2.8.2.1 Medição do pulso elétrico

Cada vez que ocorre uma DP uma pequena corrente elétrica também flui nos enrolamentos da máquina, compreendendo um determinado número de cargas, em pico Coulomb (pC), mensurável pelo pulso de tensão elétrica gerado. Esse pulso de tensão constitui-se num dos principais meios de detecção e medição das DPs, também podendo-se realizar através do pulso de corrente resultante, em circuito remotos. Nos enrolamentos ou barras do hidrogerador centenas de DP, que ocorrem a cada segundo, podem ser detectadas.

Os capacitores de acoplamento de alta tensão são os equipamentos utilizados para detectar os pulsos de tensão. Normalmente são conectados nos enrolamentos ou no terminal fase, conforme observa-se na Figura 2.15. Funciona com baixa impedância para pulsos de tensão de alta frequência das DPs, porém apresentam alta impedância para a tensão de frequência industrial. Os capacitores transmitem na saída pulsos de tensão que podem ser medidos por osciloscópio, analisador de espectro ou um medidor de DPs.

As detecções e medições de DPs utilizando este método podem ser realizadas tanto para testes on-line como off-line.

Figura. 2.16 - Capacitores de acoplamento. a) Instalado no interior do enrolamento estatórico.
b) Instalado no terminal fase.



(a)



(b)

Fonte: (18).

A magnitude de uma DP além de poder ser mensurada em termos de tensão (mV), pode também ser considerada em termos de carga aparente (pC), dependendo da faixa de frequência do sistema de medição de DP. No entanto, a conversão de magnitudes de tensão de DP em nível de carga aparente, e vice-versa, geralmente não é possível, principalmente em faixas de frequências mais altas. Isto se deve ao fato que a maioria dos instrumentos digitais de medição de DP, ao detectar um impulso de DP, precisa de um “tempo morto” para processar a informação recebida, sendo que, alguns pulsos de alta magnitude podem não ser registrados, comprometendo a medição em termos de carga. Além disso, em instrumentos de medição que operam na faixa de alta e baixa frequência, pode ocorrer superposição de dois impulsos de DP, resultando na soma de magnitudes dos impulsos. Mudanças na escala de magnitude podem, também, alterar o número de pulsos de DP registrados, afetando a medição em carga aparente. E, na maioria dos instrumentos de medição de DP, existe um limiar abaixo do qual, pulsos de DP não são registrados, para evitar que o tempo morto do equipamento não seja continuamente acionado, impedindo o registro de dados de DPs (15).

2.8.2.2 Medição da radiação de rádio frequência

Os centelhamentos gerados pelas DPs produzem ondas eletromagnéticas de rádio frequência que se propagam longe do local de ocorrência das DPs. Esses distúrbios têm frequência de rádio frequência 100 kHz até centenas de MHz.

Antenas apropriadas podem ser utilizadas para detectar a atividade das DPs, sendo possível até localizar as DPs que estão ocorrendo no interior do hidrogerador.

2.9 DIFICULDADES NOS PROCESSOS DE MEDIÇÃO DAS DPS

As interferências nos processos de detecção e medição das DPs dificultam a interpretação e classificação das DPs, pois os dados apresentam-se com vários sinais de ruídos que se confundem às DPs. De acordo com (7), vários fatores interferem no processo de detecção e medição das DPs, dividindo-se em duas categorias: perturbações que ocorrem quando o circuito de ensaio está energizado e quando o circuito não está energizado.

Os distúrbios que ocorrem quando o circuito de ensaio está energizado aumentam segundo o aumento do nível de tensão. Esses distúrbios originam-se nos condutores, em falhas nos aterramentos e em harmônicos da fonte de tensão próximos à banda de frequência da medição (7).

Nos circuitos de ensaio não energizados, as perturbações ocorrem por chaveamentos em outros circuitos, ensaios de alta tensão realizados nas proximidades, interferências intrínsecas aos sistemas de medição (7). Diferentes irradiações de ondas eletromagnéticas também podem interferir no processo de medição das DPs, como os ruídos pulsantes originados de comutadores, ruídos aleatórios e ruídos senoidais oriundos de serviços de comunicação.

Além disso, pulsos de DPs podem ocorrer em uma velocidade acima da capacidade do equipamento de medição, a ponto deste não conseguir detectá-los.

2.10 CONCLUSÕES DO CAPÍTULO

Os materiais isolantes das barras estatóricas constituem um dos componentes mais delicados da máquina e que possuem grande importância para o funcionamento desta. Em

face disto, é indispensável a realização de monitoramentos periódicos para avaliar o estado operacional da máquina, a fim de evitar interrupções não planejadas, comprometendo o fornecimento de energia elétrica.

A presença de DPs, nos seus mais variados tipos, indicam falhas que contribuem para a degradação do material isolante, gerando novas regiões com ocorrência das descargas. Dependendo das regiões onde ocorrem as DPs, estas terão um padrão específico de ocorrência catalogado e um nível de periculosidade associado.

Assim, é importante monitorar o comportamento das DPs nos isolamentos das máquinas para planejamento de reparos, como também, formar uma base de dados para o desenvolvimento de estudos de DPs, que possam proporcionar novos métodos de acompanhamento, controle e prevenção das DPs.

3.0 OS ENROLAMENTOS ESTATÓRICOS DO HIDROGERADOR

3.1 INTRODUÇÃO

Hidrogeradores são máquinas rotativas projetadas para operar durante longo tempo. No entanto, seu tempo de vida útil pode ficar comprometido devido aos diversos tipos de falhas que podem acometer a máquina.

A estrutura mais frágil destas máquinas são os isolamentos das bobinas estatóricas. Nessas estruturas ocorre a maioria dos problemas que comprometem gradualmente o funcionamento das máquinas.

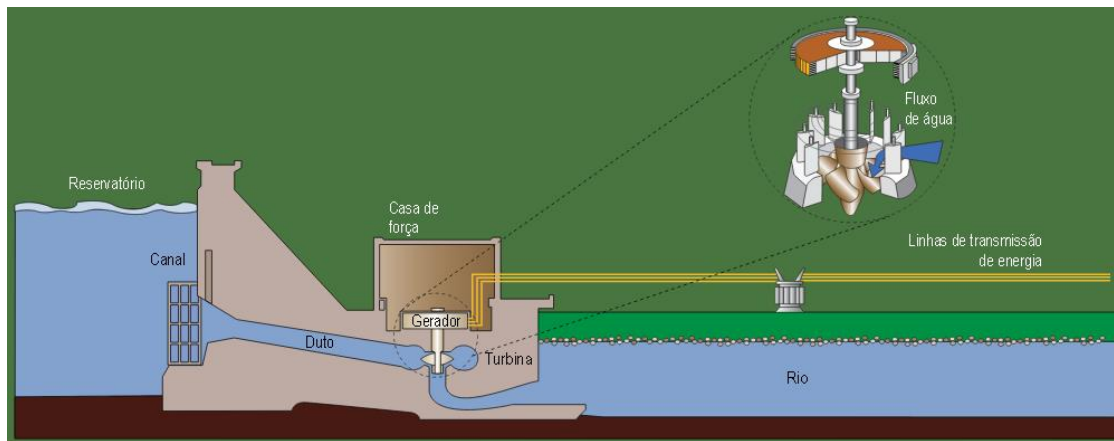
Assim, este capítulo apresentará a estrutura geral dos hidrogeradores, com ênfase nos isolamentos das bobinas do estator. Além disso, será explanado os tipos e as causas das falhas a que estão sujeitos estes equipamentos.

3.2 ESTRUTURA DOS ENROLAMENTOS DO ESTATOR

3.2.1 Estrutura geral do hidrogerador

Hidrogeradores são máquinas rotativas que convertem a energia mecânica, fornecida pela força da água no eixo da turbina hidráulica, em energia elétrica. São normalmente empregados em usinas hidrelétricas construídas em um rio. Na Figura 3.1 é ilustrada esquematicamente uma usina hidrelétrica:

Figura 3.1 - Esquema de uma usina hidrelétrica.

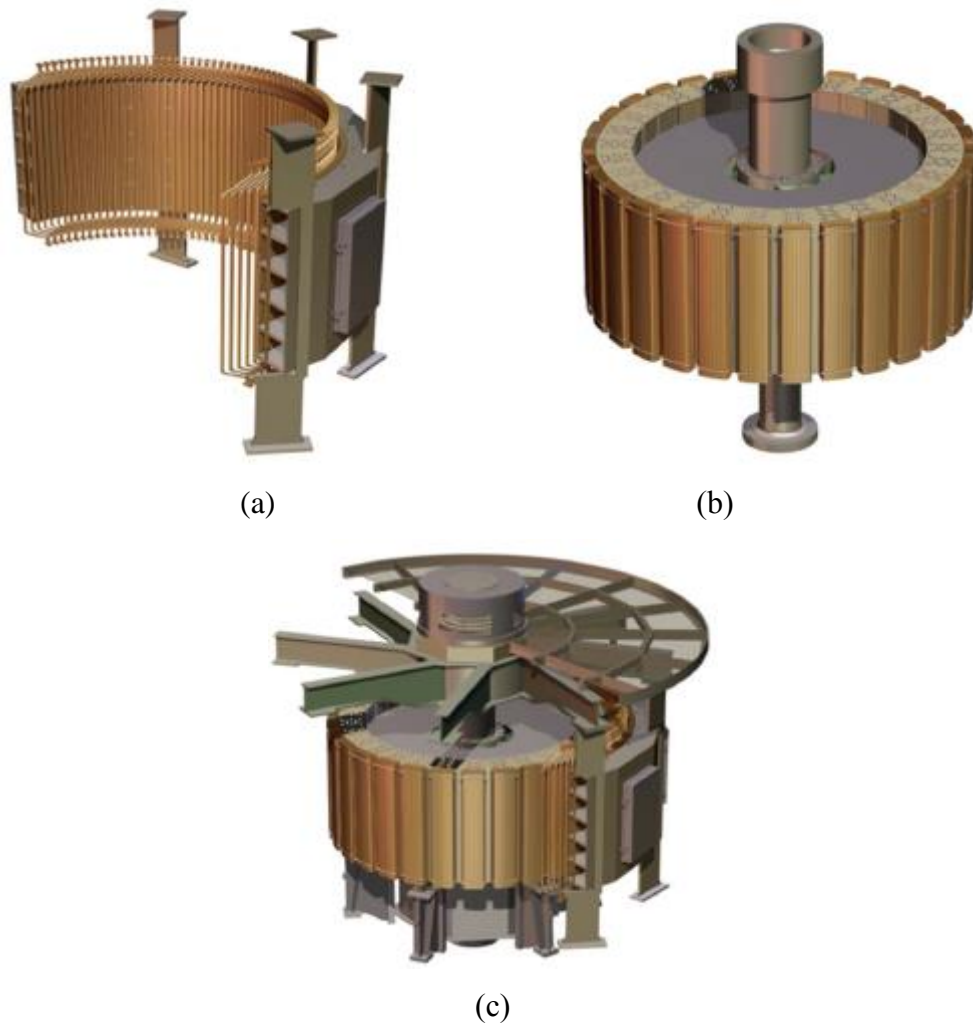


Fonte: (23).

Basicamente, os hidrogeradores são compostos por duas estruturas principais: o rotor e o estator. O rotor constitui-se na parte móvel, em torno do próprio eixo, impulsionado por uma fonte primária de energia. É construído com material ferromagnético e possui bobinas que formam um enrolamento, chamado enrolamento de campo, responsável pela geração do campo magnético girante. É responsável pela criação de um fluxo magnético, à semelhança de um eletroímã, com tantos pólos quantos sejam os do enrolamento do estator ou armadura (24). O rotor empregado em hidrogeradores é do tipo pólos salientes.

O estator, também conhecido como armadura, é a parte fixa composta por um conjunto de bobinas, que formam os enrolamentos, simetricamente alocadas e instaladas em ranhuras em um núcleo fixo e aterrado. A ação do fluxo magnético gerado pelo rotor induz uma força eletromotriz nos enrolamentos, e estes funcionam como condutor proporcionando um caminho à corrente elétrica, enquanto que o núcleo concentra o campo magnético, fixa os enrolamentos e dissipa o calor (18). O comprimento axial do núcleo do estator é relativamente curto para baixa velocidade, e os geradores são de grande diâmetro (25). No caso dos hidrogeradores, a quantidade de polos é maior que em geradores à gás ou vapor. Na Figura 3.2 são ilustrados o rotor e o gerador e como estes são estruturados juntos:

Figura 3.2 - Estrutura do hidrogerador. a) Seção do estator: enrolamentos fixados no núcleo.
b) Rotor. c) Seção da máquina completa: estator e rotor.



Fonte: (26).

Quando o hidrogerador opera desconectado da carga e a corrente é nula nos enrolamentos, não há queda de tensão nas reatâncias internas da máquina e a força eletromotriz é igual à tensão nos terminais do gerador. Já quando o hidrogerador opera conectado a uma carga, a corrente nos enrolamentos estáticos não é nula, e ocorre uma queda de tensão nas reatâncias internas. Assim, a força eletromotriz será diferente da tensão dos terminais da máquina.

3.2.2 Os Enrolamentos Estáticos do Hidrogerador

Os enrolamentos, bobinas ou barras estáticas tem a função primordial de conduzir as

correntes induzidas pelo fluxo magnético gerado pelo rotor. São constituídas por barras de cobre isoladas, inseridas nas ranhuras do núcleo. As cabeças de bobina fazem o interligamento das barras em suas extremidades, a fim de tornar o enrolamento trifásico, conforme as Figuras 3.3, 3.4 e 3.5, apresentadas a seguir:

Figura 3.3 - Estator de hidrogerador.



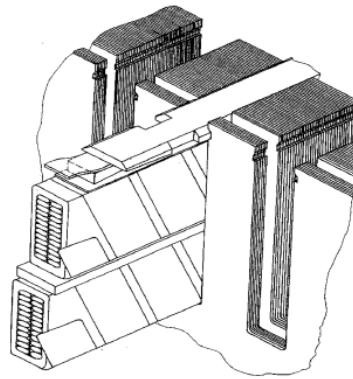
Fonte: (26).

Figura 3.4 - Região das cabeças de bobinas interligadas.



Fonte: (21).

Figura 3.5 - Figura esquemática do corte transversal das barras inseridas na ranhura.



Fonte: (26).

De acordo com (27), dois modelos de barras são comumente utilizados nos projetos dos enrolamentos estatóricos: o modelo de barra ou meia bobina ou barras Roebel e o modelo de múltiplas espiras. A seleção do modelo da barra depende da potência de saída e do fator de potência da máquina, como também fatores como capacidade de fabricação, facilidade de instalação e reparo são considerados. Máquinas abaixo de 75 MW comumente serão construídas com bobinas de múltiplas espiras e, acima de 75MW, são normalmente utilizadas bobinas do modelo de meia bobina.

3.2.2.1 Modelo barra

Esse modelo é construído com dois tipos de design para os terminais da barra: ondulado e imbricado. Na Figura 3.6 estão ilustrados os dois designs:

Figura 3.6 - Barras. a) Modelo de barra imbricado; b) Modelo de barra ondulado.



(a)

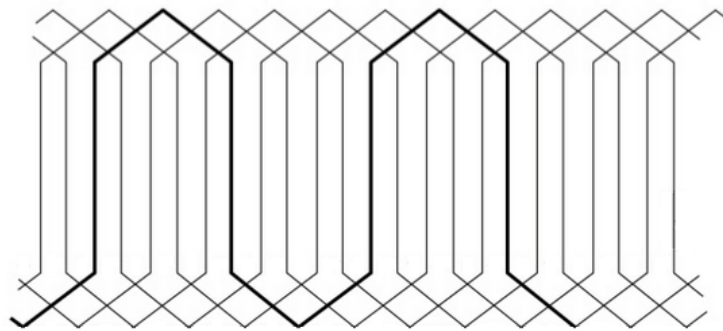


(b)

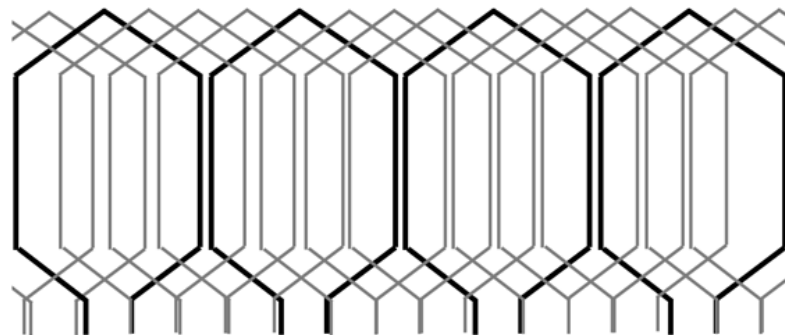
Fonte: Elaborado pelo autor.

São empregados de acordo com o desenho do enrolamento para o encaixe no núcleo, ou seja, a maneira com que as bobinas são colocas na barra. A barra imbricada tem seus terminais conectados às barras vizinhas. Os terminais das barras onduladas são conectados às barras deslocados 360° elétrico entre si. Na Figura 3.7 são ilustradas esquematicamente as conexões das barras citadas:

Figura 3.7 - Conexões dos enrolamentos. a) Enrolamentos ondulados; b) Enrolamentos imbricados.



(a)



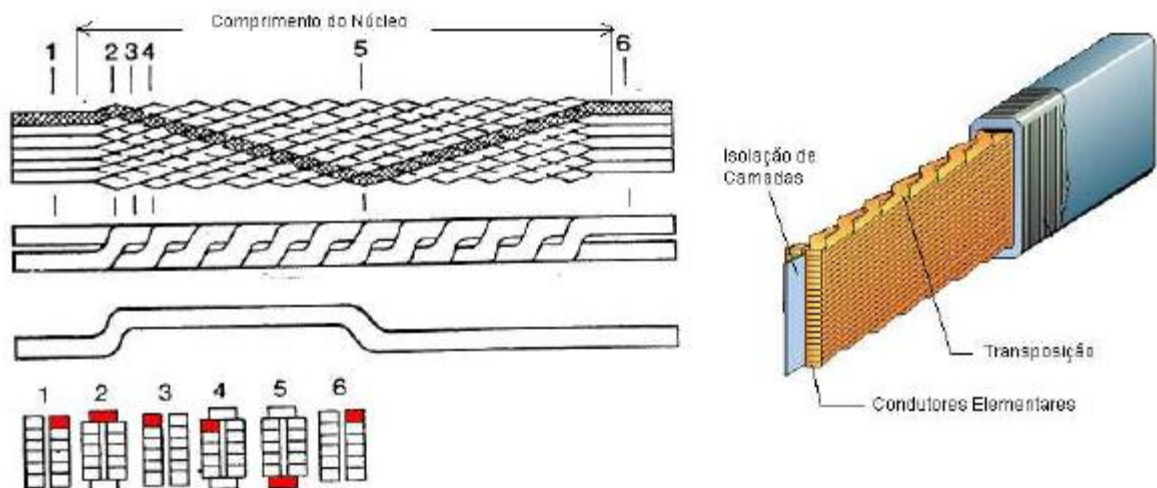
(b)

Fonte: Elaborado pelo autor.

Esses tipos de barras são comumente empregados em máquinas que possuem núcleo longo ou com diâmetro pequeno. Para facilitar a montagem das barras em estatores com este perfil, estas são divididas em duas seções, por isso o nome do modelo também conhecido como meia bobina ou barras Roebel.

A configuração da barra Roebel, conhecida como transposição Roebel, determina que esse tipo de barra é formado por condutores em seção retangular e isolados entre si. Além disso, os condutores são transpostos de 360° , a fim de distribuir a corrente de modo homogêneo pela seção da barra e facilitar a dissipação de calor (28). Na Figura 3.8 é ilustrada esquematicamente a transposição Roebel:

Figura 3.8 - Figura esquemática da transposição Roebel.



Fonte: (28).

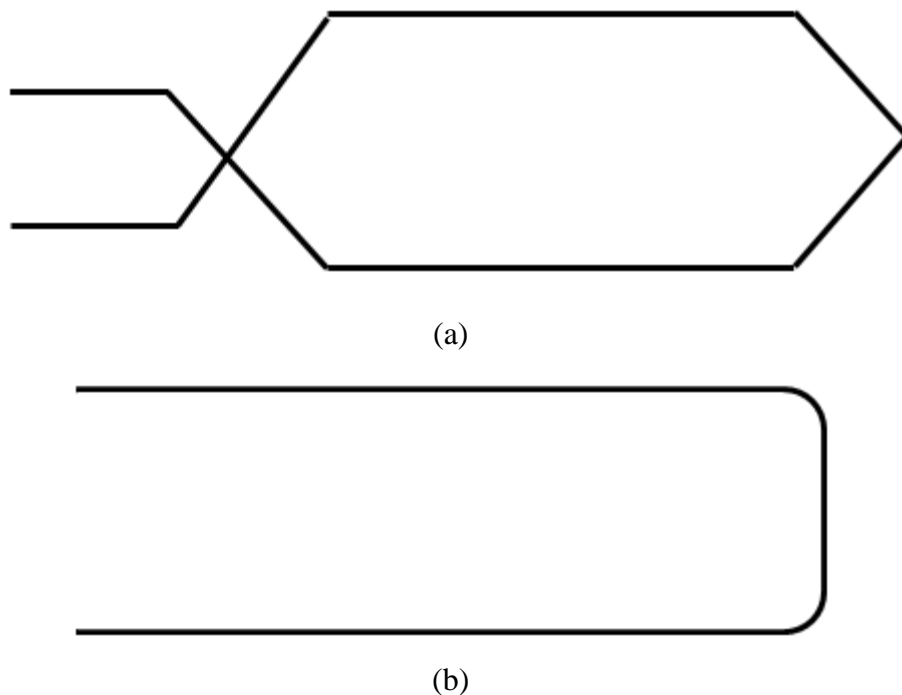
De acordo com (28), “existe uma limitação construtiva na confecção de barras Roebel imposta pelo comprimento do núcleo estatórico. Uma vez que para completar a transposição de 360° se faz necessário um comprimento mínimo dependente do número e dimensões do fio.”.

Uma grande vantagem neste modelo de barra é a versatilidade para a montagem, já que cada meia bobina é inserida individualmente e soldada à outra parte. A mesma vantagem também se aplica para manutenções, sendo necessária apenas a retirada da soldagem para atender aos reparos de uma região específica da bobina (28).

3.2.2.2 Modelo de Múltiplas Espiras

Assim como no modelo anterior, as barras de múltiplas espiras podem ser construídas com dois tipos de design: concêntrico e diamante. Na Figura 3.9 estão ilustrados os modelos mencionados:

Figura 3.9 - Múltiplas espiras. a) Modelo múltipla espira diamante; b) Modelo múltipla espira concêntrico.



Fonte: Elaborado pelo autor.

Barras de múltiplas espiras diamante são bastante versáteis para as mais diversas máquinas e são largamente utilizadas em projetos modernos. Já as barras de múltiplas espiras concêntricas são mais encontradas em máquinas antigas (27).

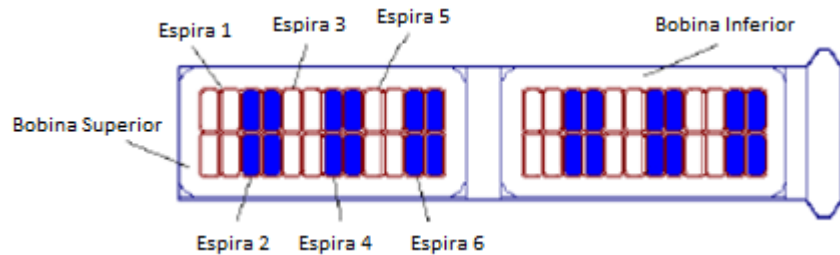
A bobina concêntrica é constituída por grupos de fios de cobre isolados que formam o condutor com várias camadas de isolamento em torno. O terminal da bobina, em forma de “U” é isolado com várias camadas de fita isolante. Diferentes tamanhos de bobina no mesmo estator são necessários para esta configuração. No núcleo com ranhuras semi-fechadas, bobinas concêntricas são inseridas numa extremidade e conectadas na outra extremidade em outra bobina (27).

A bobina de múltipla espira diamante é constituída por um grupo de condutores de

cobre isolados em cada espira. Conforme a necessidade, a bobina poderá ter mais de uma espira, tornando-se mais sensível a surtos de tensão oriundos de operações de chaveamento ou descargas atmosféricas. Por isso, a isolamento entre as espiras deve ser bem projetada (28).

Na figura (3.10) observa-se o corte da ranhura de um enrolamento de dupla camada formada por bobinas de múltiplas espiras, constituída por seis espiras sendo cada espira formada por quatro fios:

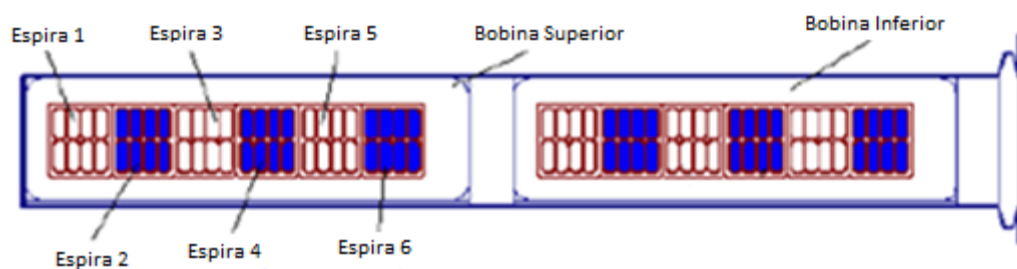
Figura 3.10 - Figura esquemática do corte de uma ranhura com bobinas de múltiplas espiras.



Fonte: Adaptado de (28).

Em (28), quando o número de fios por espira for grande indica-se o uso de isolamento entre as espiras, mantendo a isolamento de cada fio no mínimo necessário, conforme Figura (3.11):

Figura 3.11 - Figura esquemática do corte de uma ranhura com bobinas de múltiplas espiras com isolamento complementar entre espiras.



Fonte: Adaptado de (28).

3.3 OS ISOLAMENTOS DAS BOBINAS ESTATÓRICAS

3.3.1 Breve histórico sobre a constituição dos isolamentos das bobinas estatóricas

A isolação tem a principal função de separar os condutores de cobre do núcleo do estator, permitindo que haja diferença de potencial entre esses dois componentes. No entanto, as bobinas estatóricas são submetidas a esforços elétricos, mecânicos, térmicos e ambientais, que aceleram o desgaste da isolação.

Em (18), para empregar um material isolante cada vez mais resistente os componentes frequentemente adotados são o vidro, mica e papel, com boas propriedades de isolação elétrica. Entretanto, o vidro e a mica tendem ao estilhaçamento quando submetidos a vibrações com frequências de 100 a 120 Hz. Por conta disso, compostos orgânicos como o asfalto, poliéster e epóxi são também utilizados com a finalidade de fortalecer mecanicamente os isoladores.

Segundo (29), os isolamentos das bobinas estatóricas eram constituídos de fitas de papel e lâminas de mica impregnadas com betume. A este arranjo, utilizado até 1960, denominava-se *Micafolium asfáltico*. Também havia o *Micafolium gomalaca*, que diferenciava-se do arranjo anterior por empregar resina natural no lugar do betume asfáltico.

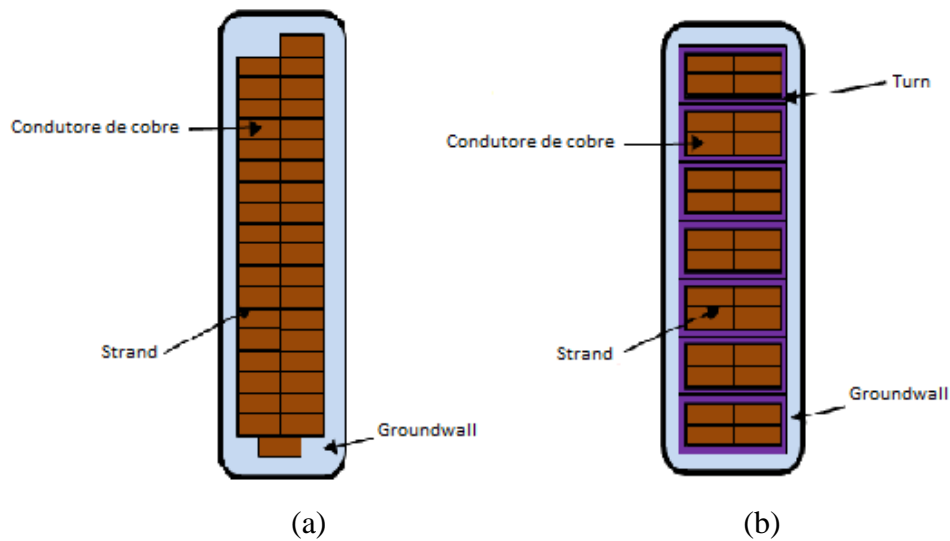
Após 1960, uma nova composição de isoladores surgiu com propriedades isolantes e térmicas superiores. Denominava-se *Micafolium poliéster*. Sua composição era de folhas de mica e tecido de fibra de vidro impregnados com a resina sintética poliéster (29).

Atualmente, a composição empregada nos isolamentos constitui-se de fitas de mica e resina epóxi, que possui desempenho superior às composições utilizadas algumas décadas atrás.

3.3.2 Tipos de isolamentos

Existem três tipos de isolamentos em uma bobina estatórica: a isolação individual do condutor de cobre (*strand*), a isolação de um grupo de condutores formando uma espira (*turn*) a isolação externa da bobina ou barra (*groundwall*). Nas bobinas de múltiplas espiras são encontrados os três tipos de isolamentos. As barras Roebel não possuem isolamento do tipo *turn* por não possuírem espiras. Na Figura 3.12 estão ilustrados os tipos de isolação das bobinas:

Figura 3.12 - Os tipos de isolamento das bobinas. a) Isolação das bobinas Roebel. b) Isolação das bobinas de múltipla espira.



Fonte: Elaborado pelo autor.

A falha entre os isolamentos dos condutores individuais geralmente acarreta apenas um aumento das perdas pelo fluxo magnético. Mas, se houver centelhamento, a barra pode ser destruída. A isolamento *turn* evita curto-circuito entre espiras e protege de altas tensões transitórias sob o enrolamento. A isolamento externa (*groundwall*) previne contra curtos entre os condutores de cobre e o núcleo aterrado. Sua espessura depende da tensão da máquina e do estresse volts/mm definido pelo fabricante. Quanto maior a espessura do isolamento, menos cobre e maior será o estresse térmico. Por outro lado, quanto mais fina a espessura do isolamento maior será o estresse de tensão e risco de falha.

3.4 FALHAS NO ISOLAMENTO DO ESTATOR

As máquinas rotativas de alta tensão, em sistemas de produção de energia elétrica, apresentam um dos principais custos operacionais os quais estão relacionados principalmente às falhas. O modo de manutenção preditivo que, atualmente, condiz melhor com as necessidades de manutenção que as concessionárias de energia precisam atender, possibilita que paradas previamente programadas para fins de reparos sejam realizadas. Para isso, é fundamental conhecer a origem dessas falhas, onde são mais recorrentes, e diagnosticá-las a tempo de evitar a perda total do equipamento. A distribuição das falhas, causas, tempo de

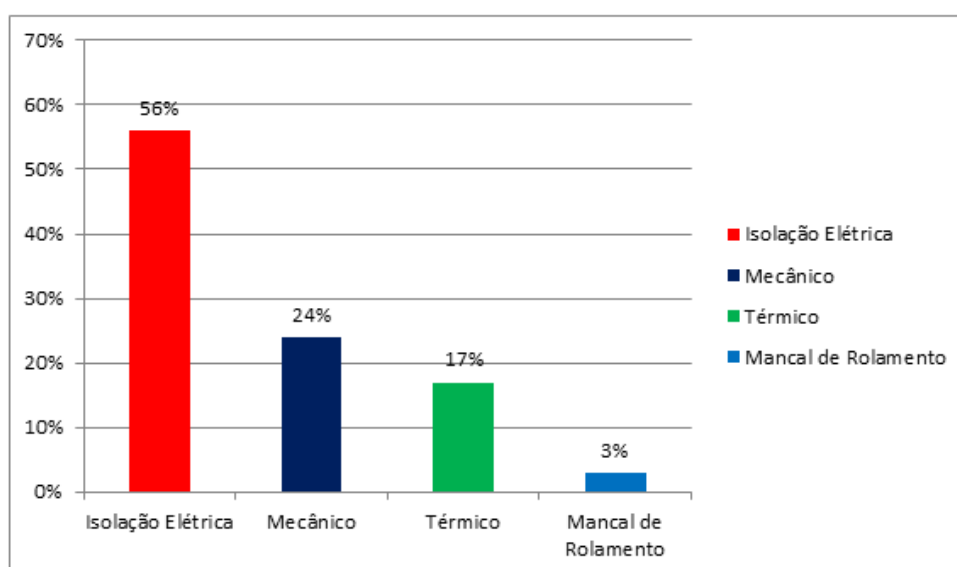
evolução, dentre outros fatores, variam entre os diferentes tipos de máquinas dadas as especificações de cada uma, de acordo com a geração de energia, disponibilidade dos recursos fontes de produção, entre outros. Mesmo diante este cenário, é possível realizar uma análise aproximada que atenda os principais requisitos para evitar o dano destes equipamentos.

Outro aspecto positivo em manutenções preditivas é que além das interrupções poderem ser programadas, o seu número pode ser reduzido. Os isolamentos da máquina constituem e uma das partes mais delicadas, que envelhecem mais rápido e que, conseqüentemente, são susceptíveis à temperatura. Em (30), a quantidade de paradas e partidas das máquinas, afeta o ciclo de temperatura dos enrolamentos e da máquina em geral, causando expansão e contração, afrouxamento e vibração dos enrolamentos nas ranhuras. Assim, reduzir ao máximo o número de manutenções é importante para a vida útil dos enrolamentos estatóricos.

Com o intuito de reunir informações úteis para fabricantes e usuários, o CIGRÉ (14) realizou um estudo com hidrogeradores em diferentes países do mundo. A pesquisa obteve dados estatísticos que ajudam a avaliar os riscos das principais falhas dos hidrogeradores. Foram investigados 69 incidentes.

Conclui-se que as principais causas raízes de falhas nos hidrogeradores são: falhas nos isolamentos, defeitos mecânicos, influências térmicas e danos nos mancais de rolamento. Os resultados estatísticos estão ilustrados na Figura (3.13):

Figura 3.13 - Causas raízes de falhas nos hidrogeradores.

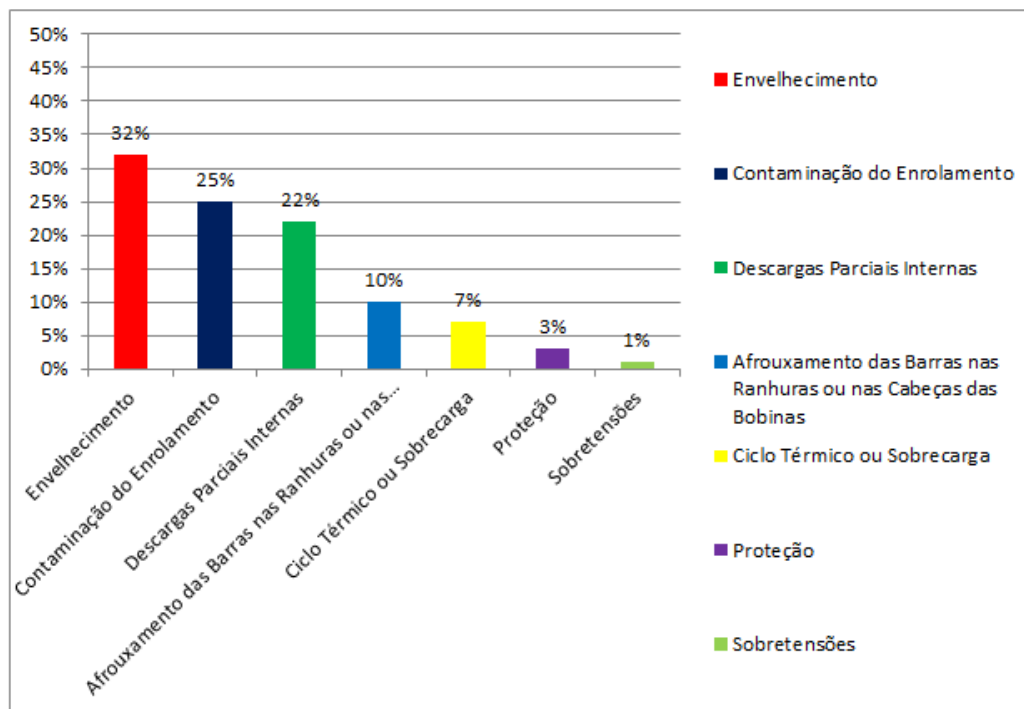


Fonte: Adaptado de (14).

A quantidade de falhas no isolamento corresponde a mais da metade da amostra. Assim,

investigando as causas das falhas dos isolamentos obteve-se os resultados ilustrados na Figura (3.14), a seguir:

Figura 3.14 - Causas de falhas no sistema de isolamento elétrica.



Fonte: Adaptado de (14).

As falhas mais frequentes nos isolamentos estão relacionadas ao envelhecimento do material constituinte da isolação, e por contaminação dos enrolamentos por poeira e umidade. As falhas elétricas são causadas por descargas parciais internas, sobretensões e proteção. As falhas devido ao afrouxamento das barras nas ranhuras ou na cabeça da bobina acontecem devido à vibração da máquina.

No trabalho (31) relacionou-se os mecanismos de falhas que ocorrem em diferentes tipos de máquinas, seus sintomas e os modos de detecção que podem ser empregados, como pode ser visualizado na Tabela (3.1):

Tabela 3.1 - Tipos de falhas nos enrolamentos estatóricos.

Mecanismo de Falha	Sintomas	Teste de Detecção	Tipo de Máquina
Impregnação Inadequada	DPs	Medição de DPs, tan δ , fator de Potência	Impregnação a vácuo e fita com resina
Falhas na camada semicondutora	DPs, descarga na ranhura, ozônio	Medição de DPs, inspeção visual e monitorar ozônio	Refrigeradas a ar
Enrolamento frouxo	DPs, ozônio, cunhas soltas	Medição de DPs, inspeção visual, monitorar ozônio e testes da cunhagem	Sistemas com isolamento termorrígida (epóxi e poliéster)
Centelhamento por Vibração	DPs, ozônio, cunhas soltas	Medição de DPs, inspeção visual, monitorar ozônio e testes da cunhagem	Com isolações à base de fita com resina ou Impregnação a vácuo
Interface de alívio de Tensão	DPs, pó branco (ácido nítrico), ozônio	Medição de DPs, tan δ , fator de potência, inspeção visual e monitorar ozônio	Refrigeradas a ar e com pintura na interface de alívio de tensão
Espaçamento Inadequado	DPs, pó branco, descoloração da isolação e ozônio	Medição de DPs, inspeção visual, monitorar ozônio	Refrigeradas a ar, seccionadoras e conexões de motores
Vibração das cabeças de bobina	Perda do sistema de bloqueio e amarração, pó branco disperso ou graxa	Inspeção visual, acelerômetro de fibra ótica	Alta tensão, alta velocidade com grandes cabeças de bobinas
Má conexão elétrica	DPs, pó branco e	Medição de DPs,	Qualquer conexão

	descoloração das conexões	inspeção visual, temperatura com infravermelho	
Deterioração térmica	DPs, descoloração da isolação	Medição de DPs, $\tan \delta$, fator de potência, inspeção visual	Todos os tipos de máquina
Ciclos de carga	DPs, ruptura da isolação na saída da barra	Medição de DPs, $\tan \delta$, fator de potência, inspeção visual, Surge Test, Hipot	Todos os tipos de máquina (com barras longas)

Fonte: Adaptado de (31).

3.5 CONCLUSÕES DO CAPÍTULO

Neste capítulo foi descrita a estrutura geral de um hidrogerador, dando ênfase para as barras estáticas e seus designs. Foi abordado também a constituição do material dos isolamentos das barras e seus tipos.

No estudo do CIGRÉ, a respeito das principais falhas que acometem os hidrogeradores, apontou-se que a maioria destas ocorre nas isolações elétricas. Também revelou-se que dentre as falhas elétricas, as DPs possuem maior taxa de ocorrência. Com esses dados, é notável o quanto se faz necessário monitorar o comportamento das DPs nas máquinas e realizar estudos para estes tipos de falhas. Além disso, é de extrema importância que os isolantes sejam submetidos à avaliação antes de entrarem em operação na máquina, para que esta não tenha um tempo de vida útil afetado.

4.0 CLASSIFICADORES BASEADOS EM TÉCNICAS DE INTELIGÊNCIA E TRATAMENTO DOS DADOS

4.1 INTRODUÇÃO

Os dados de DPs coletados possuem uma grande dimensão de informações acompanhadas de ruídos que dificultam a eficácia de técnicas inteligentes para o reconhecimento de padrões de DPs. Em muitos casos, antes de executar a classificação dos dados é necessário uma etapa de pré-processamento, objetivando reduzir as interferências dos ruídos e diminuir o esforço computacional realizado pelas técnicas de classificação aplicadas.

Nesta dissertação foram utilizadas técnicas de compactação dos dados, a fim de minimizar os ruídos. Para tal procedimento foram aplicadas a Análise de Componentes Principais (ACP) e Análise de Variância (ANOVA). Ambas baseiam-se na redução da dimensionalidade dos dados fornecidos. Outro recurso utilizado com o mesmo propósito foi Análise de Componentes Independentes (ACI) que separa fontes de sinais independentes entre si. Cada uma dessas técnicas foi aplicada separadamente.

Para o reconhecimento e classificação dos padrões de DPs foram empregadas técnicas mais recentes na literatura. Uma maneira de classificar os dados é considerando a similaridade entre as características dos sinais mais próximos, através da técnica K-Vizinhos Mais Próximos. Outra técnica, Floresta Randômica, baseia-se em um conjunto de árvores de decisão para realizar a melhor classificação. E ainda, a técnica de Máquinas de Vetores de Suporte, do inglês *support vector machines*, que utiliza diversos hiperplanos para delimitar as regiões entre as classes.

No presente capítulo são descritas as técnicas citadas que foram utilizadas para a classificação dos dados de DPs deste trabalho.

4.2 TRATAMENTO DE DADOS

4.2.1 Análise de Variância

A Análise de Variância (ANOVA, do inglês *ANalysis Of Variance*) é uma técnica de seleção de características que pode ser usada para identificar as variáveis mais relevantes para a classificação de um conjunto de informações, reduzindo a dimensionalidade para aumentar a

eficiência e acurácia na classificação. Trata-se de um método estatístico que analisa a existência de diferença significativa entre as médias e, também, se os fatores em questão influenciam em alguma variável dependente. Tais fatores podem configurar-se em variáveis qualitativas ou quantitativas, enquanto que, as variáveis dependentes devem ser quantitativas (32).

Existem condições para o método poder trabalhar de modo eficaz: as amostras devem possuir caráter aleatório, independente, e ser de populações normais desde que estas tenham variâncias iguais. Satisfeitas estas condições, as médias são comparadas para então verificar qual delas é mais significativa (32).

Matematicamente, o processo consiste em definir um grupo de fatores de tamanho N para cada classe e calcular a variância (v) para todos os elementos dos grupos (número de grupos $\times N$), através da Equação (4.1):

$$v = \left(\frac{\sigma}{N}\right)^{1/2} \quad (4.1)$$

Onde, σ é o desvio padrão e N é a quantidade total de elementos.

4.2.1.1 Anova de um fator

De acordo com (32), em ANOVA de um fator, uma única variável independente é subdividida em tratamentos (níveis do fator). A variável de estudo, que constitui-se na variável dependente, será medida através das amostras de cada tratamento. Na Tabela 4.1 são representados os elementos das amostras com seus tratamentos:

Tabela 4.1 - Experimento de um fator.

		Tratamentos					
		1	2	...	K		
Elemento das amostras	1	x_{11}	x_{21}	...	x_{k1}		
	2	x_{12}	x_{22}	...	x_{k2}		
		
		
	n_j	x_{n1}	x_{n2}	...	x_{nk}		
	Σ						Total
	Médias	\bar{x}_1	\bar{x}_2	...	\bar{x}_k	\bar{X}	

Fonte: (24)

O resultado de cada amostra pode ser representado pelo modelo estatístico de (33):

$$X_{ij} = \mu + \tau_i + \varepsilon_{ij} \quad (4.2)$$

onde, μ é a média geral de todas as amostras, τ_i é o efeito do nível i , ε_{ij} é o erro aleatório, $i = 1, \dots, I$ corresponde ao número de amostras de cada grupo.

As variâncias são calculadas a partir das somas dos quadrados dos resíduos. A Soma dos Quadrados dos Grupos (SQG) está associada ao desvio da média do grupo i em relação à média geral. A Soma dos Quadrados dos Resíduos (SQR) está associada ao desvio da observação individual em relação à média do grupo. A Soma dos Quadrados Totais (SQT) é a soma entre SQG e SQR (33). Na Tabela 4.2 são exibidas as equações de cada soma dos quadrados citada:

Tabela 4.2 - Equações de SQG, SQR e SQT.

SQG	SQR	SQT
$SQG = \sum_i n_i (\bar{X}_i - \bar{X}_{..})^2$	$SQR = \sum (X_{ij} - \bar{X}_{.j})^2$	$SQT = \sum_{ij} (X_{ij} - \bar{X}_{..})^2$ $SQT = SQG + SQR$

Sendo, $\bar{X}_{..}$ a média geral correspondente a todas as amostras e \bar{X}_i a média das observações do grupo i .

Em (33), a estimativa das variâncias é obtida através da divisão das somas quadráticas pelos seus respectivos graus de liberdade. A distribuição dos resíduos deve ser aproximadamente normal e a variância seja idêntica e independentemente distribuída. Se não houver efeito entre os grupos, as estimativas das variâncias dos grupos e dos resíduos devem ser aproximadamente iguais. A Tabela 4.3 exibe as equações para cálculo das variâncias:

O valor de F indica a diferença entre os grupos em função da variância dentro de cada grupo. Se F for aproximadamente 1, a variância não é considerável. F é calculado pela razão entre MQG e MQR.

Tabela 4.3 - Equações para cálculo das variâncias.

Fonte de Variação	Soma dos Quadrados	Graus de Liberdade	Média Quadrática	F_{cal}
Entre Grupos	SQG	$J - 1$	$MQG = \frac{SQG}{J - 1}$	$\frac{MQG}{MQR}$
Dentro do Grupo	SQR	$N - J$	$MQR = \frac{SQR}{N - J}$	
Total	SQT	$N - 1$		

4.2.2 Análise de Componentes Principais

A Análise de Componentes Principais (ACP) é uma das técnicas mais utilizadas para redução de dimensionalidade de dados multidimensionais. Os dados multidimensionais demandam maior esforço computacional para serem processados e armazenados, além disso, apresentam uma maior dificuldade na identificação de características semelhantes.

A ACP, também conhecida como *Karhunen-Loeve Transform*, reduz a dimensionalidade de dados multidimensionais, elimina sobreposições, e procura a existência de padrões que podem estar ocultos, através de características semelhantes, chamadas componentes principais. A técnica encontra formas mais representativas de dados através de combinações lineares das variáveis originais. Esse processo facilita a compactação dos dados antes que estes sejam analisados por outro sistema, diminuindo o esforço computacional realizado pelas técnicas posteriormente aplicadas e possibilitando a classificação de padrões encontrados com uma quantidade menor de dados ruidosos ou com pouca influência (34), (35).

A técnica correlaciona estatisticamente os dados decompondo-os em um conjunto de dados multivariados em ordem decrescente de importância, preservando os componentes principais, como um conjunto de variáveis não correlacionadas, e mantendo a máxima variância possível dos dados originais. O componente principal posterior deve ser ortogonal em relação ao primeiro componente principal, sendo assim, estatisticamente independentes entre si, baseando-se nos autovetores e autovalores da matriz de covariância dos dados. ACP trabalha nos dados encontrando um conjunto de vetores que se estendem por um subespaço (34). Como exemplo, imagens tridimensionais de um objeto podem ser reduzidas em duas dimensões, onde a nova imagem irá ilustrar o objeto com o máximo possível de suas características.

Outra solução algébrica que pode ser utilizada com ACP é a Decomposição por Valores Singulares (DVS). A DVS possibilita que qualquer matriz de dados possa ser decomposta em uma soma ponderada finita de matrizes de posto unitário, encontrando o posto ou grau de singularidade da matriz (35). É uma técnica que pode ser aplicada à compressão de dados, de modo que preserve uma quantidade mínima de dados para recuperá-los. DVS e PCA são técnicas intimamente relacionadas e comumente aplicadas juntas.

4.2.2.1 Decomposições em Valores Singulares

Em (36), para uma matriz $(A_{ij})_{m \times n}$, e considerando que a matriz $(A^T A_{ij})_{n \times n}$ é simétrica, a mesma é possível de ser diagonalizada ortogonalmente. Os autovalores de $A^T A$ são todos reais e não negativos. Para demonstrar, supondo que v_i seja o autovetor de $A^T A$ correspondente ao autovalor λ_i , então se $1 \leq i \leq n$:

$$\|Av_i\|^2 = (Av_i)^T Av_i = v_i^T A^T Av_i = v_i^T (\lambda_i v_i) \quad (4.3)$$

Os Valores Singulares de uma matriz $A_{m \times n}$ são as raízes quadradas dos autovalores de $A^T A$, e são denotados por $\sigma_1, \sigma_2, \sigma_3, \dots, \sigma_n$. Comumente, ordena-se os valores singulares de modo decrescente, em que $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \dots \geq \sigma_n$.

Considerando os autovetores de $A^T A$. Como $A^T A$ é simétrica, existe uma base ortonormal para \mathbb{R}^n formada por autovetores de $A^T A$. Seja $\{v_1, v_2, \dots, v_n\}$ a base correspondente aos autovalores de $A^T A$, ordenados de modo que $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_n$:

$$\lambda_i = \|Av_i\|^2 \quad (4.4)$$

Portanto:

$$\sigma_i = \sqrt{\lambda_i} = \|Av_i\| \quad (4.5)$$

Assim, os valores singulares de A são os comprimentos dos vetores Av_1, Av_2, \dots, Av_n .

Supondo que v_1, v_2, \dots, v_n é uma base ortonormal consistindo em autovetores de $A^T A$ ordenados de tal forma que os autovalores associados satisfaçam $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$. E ainda considerando que A tem r valores singulares não nulos, então $\{Av_1, Av_2, \dots, Av_r\}$ é uma base ortogonal para o espaço gerado pelas colunas de A ($ColA$) e $posto(A) = r$.

Para demonstrar, como v_i e $\lambda_j v_j$ são ortogonais para $i \neq j$: $\langle Av_i, Av_j \rangle = \langle v_i, A^T Av_j \rangle = \langle v_i, \lambda_j v_j \rangle = \lambda_j \langle v_i, v_j \rangle = 0$. Logo, Av_1, Av_2, \dots, Av_n é um conjunto ortogonal. Além disso, como os comprimentos dos vetores Av_1, Av_2, \dots, Av_n são os valores singulares de A e como existem r valores singulares não-nulos, $Av_i \neq 0$ se, e somente se $1 \leq i \leq r$. Portanto, são vetores linearmente independentes e pertencem ao espaço gerado pelas colunas de A , qualquer que seja y em $ColA$:

$$y = Ax \quad (4.6)$$

$$y = Ac_1 v_1 + Ac_2 v_2 + \dots + Ac_r v_r + Ac_{r+1} v_{r+1} + \dots + Ac_n v_n \quad (4.7)$$

$$y = Ac_1 v_1 + Ac_2 v_2 + \dots + c_r Av_r + c_{r+1} Av_{r+1} + \dots + c_n Av_n \quad (4.8)$$

$$y = Ac_1 v_1 + Ac_2 v_2 + \dots + c_r Av_r + 0 + \dots + 0 \quad (4.9)$$

Então, y pertence a $\{Av_1, Av_2, \dots, Av_r\}$, o que mostra que é uma base ortogonal para o espaço gerado pelas colunas de A e que $Posto(A) = r$.

Para realizar a Decomposição em Valores Singulares considera-se a matriz $(A_{ij})_{m \times n}$ com valores singulares $\sigma_1 > \sigma_2 > \dots > \sigma_r > 0$ e $\sigma_{r+1} = \sigma_{r+2} = \dots = \sigma_n = 0$. Então, existem duas matrizes ortogonais, $(U_{ij})_{m \times n}$ e $(V_{ij})_{n \times n}$, e uma matriz $(S_{ij})_{m \times n}$ tais que $A = USV^T$:

$$D = \begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \sigma_r \end{bmatrix} \quad (4.10)$$

$$S = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \quad (4.11)$$

Para construir a matriz ortogonal V , deve-se encontrar uma base ortonormal v_1, v_2, \dots, v_n de \mathbb{R}^n formada por autovetores da matriz simétrica. Então, $A^T A \in M_{(n \times n)}$. E também, $V = [v_1 v_2 v_3 \dots v_n]$ é uma matriz ortogonal $n \times n$.

Para encontrar a matriz ortogonal U , Av_1, Av_2, \dots, Av_r é um conjunto ortogonal de vetores de \mathbb{R}^m . Supondo que v_i seja o autovetor de $A^T A$ correspondente ao autovalor λ_i . Então, para $i \neq j$: $\langle Av_i, Av_j \rangle = \langle v_i, A^T Av_j \rangle = \langle v_i, \lambda_j v_j \rangle = \lambda_j \langle v_i, v_j \rangle = 0$, visto que, os autovetores v_i são ortogonais.

Como o valor singular satisfaz $\sigma_i = \|Av_i\|$, e os primeiros r são não nulos. Portanto, Av_1, Av_2, \dots, Av_r pode ser normalizado considerando $u_i = 1/\sigma_i \times Av_i$ para $1 \leq i \leq r$.

Isto garante que o conjunto $u_1, u_2, u_3, \dots, u_r$, chamado Processador de Gram-Schmidt e desenvolvido no trabalho de (36), seja um conjunto ortonormal de \mathbb{R}^m . Mas, se $r < m$, ele não será uma base para \mathbb{R}^m . Neste caso, estende-se o conjunto $u_1, u_2, u_3, \dots, u_r$ para uma base ortonormal $u_1, u_2, u_3, \dots, u_m$ de \mathbb{R}^m . Então, considerando $U = [u_1, u_2, u_3, \dots, u_m]$:

$$V^T = V^{-1} \quad (4.12)$$

$$Av_i > 0, \quad \forall 1 \leq i \leq r \quad (4.13)$$

$$Av_i = \sigma_i u_i, \quad \forall 1 \leq i \leq r \quad (4.14)$$

$$\|Av_i\| = \sigma_i = 0, \quad \forall r + 1 \leq i \leq n \quad (4.15)$$

Assim, $Av = A[v_1 v_2 \dots v_n]$:

$$Av = [Av_1 Av_2 \dots Av_n] \quad (4.16)$$

$$Av = [\sigma_1 u_1 \sigma_2 u_2 \dots \sigma_r u_r] \quad (4.17)$$

$$Av = [u_1 u_2 \dots u_m] \begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \sigma_n \end{bmatrix} \quad (4.18)$$

$$Av = US \quad (4.19)$$

$$A = USV^T \quad (4.20)$$

Onde, a DVS de A também pode ser escrita como o somatório de r matrizes de posto 1:

$$Av = [Av_1 \ Av_2 \ \dots \ Av_r \ Av_{r+1} \ \dots \ Av_n] \quad (4.21)$$

$$Av = [Av_1 \ Av_2 \ \dots \ Av_r \ 0 \ \dots \ 0] \quad (4.22)$$

$$Av = [\sigma_1 u_1 \ \sigma_2 u_2 \ \dots \ \sigma_r u_r \ 0 \ \dots \ 0] \quad (4.23)$$

$$Av = [u_1 \ u_2 \ \dots \ u_r] \begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \sigma_r \end{bmatrix} \quad (4.24)$$

$$Av = US \quad (4.25)$$

$$A = USV^T \quad (4.26)$$

$$A = \sum_{i=1}^r \sigma_i u_i v_i^T \quad (4.27)$$

4.2.2.2 Aplicação da Decomposição em Valores Singulares na Análise de Componentes Principais

Em (36), considerando uma matriz de dados A de m indivíduos caracterizados por n variáveis quantitativas:

$$A = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \vdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \vdots & x_{mn} \end{bmatrix} \quad (4.28)$$

Sejam x , y e z vetores de \mathbb{R}^m . Uma distância em \mathbb{R}^m é uma aplicação $d: \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$ satisfazendo as seguintes condições:

$$\text{i.} \quad d(x, y) \geq 0, d(x, y) = 0 \leftrightarrow x = y \quad (4.29)$$

$$\text{ii.} \quad d(x, y) = d(y, x) \quad (4.30)$$

$$\text{iii.} \quad d(x, y) \leq d(x, z) + d(z, y) \quad (4.31)$$

Para estabelecer distâncias entre os indivíduos da matriz A , não é aconselhável utilizar a distância usual, pois as variáveis podem apresentar unidades de medida diferentes ou até unidades de medidas não comparáveis. Então, para estabelecer a distância sem o problema das diferenças entre as unidades de medida, divide-se a coluna j da matriz X por $\sqrt{S_{jj}}$. A distância entre a primeira e segunda coluna, por exemplo, é dada por:

$$(x_{11} - x_{12})^2/S_{11} + \cdots + (x_{m1} - x_{m2})^2/S_{mm} \quad (4.31)$$

Tomando a matriz:

$$D_s = \begin{bmatrix} 1/S_{11} & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 1/S_{nn} \end{bmatrix} \quad (4.32)$$

Assim, (4.31) pode ser escrito como sendo: $(x_1 - x_2)^T D_s (x_1 - x_2)$.

Para encontrar os subespaços ótimos, dada as observações dos m , objetiva-se encontrar um subespaço vetorial de \mathbb{R}^n com dimensão menor que n , que melhor represente os dados.

Seja x_j um indivíduo e $p(x_j)$ a projeção sobre o subespaço, a distância entre x_j e $p(x_j)$ é dada por:

$$d_j^2 = (x_j - p(x_j))^T D_s (x_j - p(x_j)) \quad (4.33)$$

Para minimizar:

$$\sum_j w_j d_j^2 \quad (4.34)$$

Ou seja, encontrar um subespaço vetorial de R^n que minimiza (4.34). "Pesos" são atribuídos para as j observações, e conseqüentemente para as j distâncias.

Um subespaço p -dimensional em R^n ($k \leq n$) é um subconjunto de vetores da forma:

$$u + \lambda_1 v_1 + \dots + \lambda_k v_p \quad (4.35)$$

onde $u \in \mathbb{R}^n$ é um vetor fixo. Assim:

$$\sum_j w_j d_j^2 = \quad (4.36)$$

$$\sum_j w_j (x_j - b_j)^T D_s (x_j - b_j) = \quad (4.37)$$

$$\sum_j w_j (a_j - y_j)^T D_s (a_j - y_j), \text{ onde } a_j = x_j - \tilde{x} \text{ e } y_j = f_{j1} u_1 + \dots + f_{jp^*} u_{p^*}. \quad (4.38)$$

Seja $A = [a_1 \dots a_n]$ e $Y = [y_1 \dots y_n]$. Objetiva-se encontrar uma matriz Y que minimize:

$$\|A - Y\|^2 = \sum_j w_j (a_j - y_j)^T D_s (a_j - y_j) \quad (4.39)$$

entre todas as matrizes Y de posto p^* .

Dada uma matriz A de posto r , existe uma matriz A de posto p onde:

$$\|A - A_p\|_2^2 = \sigma_{p+1}^2 \quad (4.40)$$

Seja $A = \sum_{i=1}^r \sigma_i u_i v_i^T$ a SVD de A . Então $A_p = \sum_{i=1}^p \sigma_i u_i v_i^T$ e $A - A_p = \sum_{i=p+1}^r \sigma_i u_i v_i^T$. Seja v o vetor singular $A - A_p$. Expressa-se v como combinação linear de $v_1, v_2, v_3, \dots, v_r$. Assim, $v = \sum_{i=1}^r \alpha_i v_i$. Logo:

$$|(A - A_p)v| = |\sum_{i=p+1}^r \sigma_i u_i v_i^T \sum_{i=1}^r \alpha_i v_i| \quad (4.41)$$

$$= |\sum_{i=p+1}^r \alpha_i \sigma_i u_i v_i^T v_i| \quad (4.42)$$

$$= |\sum_{i=p+1}^r \alpha_i \sigma_i u_i| \quad (4.43)$$

$$= \sqrt{\sum_{i=p+1}^r \alpha_i^2 \sigma_i^2} \quad (4.44)$$

Onde v atinge seu máximo quando está sujeito a restrição de que $|v|^2 = \sum_{i=1}^r \alpha_i^2 = 1$, que ocorre quando $\alpha_{p+1} = 1$ e o restante dos α_i são iguais a 0. Assim,

$$\|A - A_p\|_2^2 = \sigma_{p+1}^2 \quad (4.45)$$

Ou seja, $Y = A_p$ é a melhor aproximação para a matriz A pela norma-2.

Seja A uma matriz $m \times n$. Para todas matrizes B de posto no máximo p é possível provar que:

$$\|A - A_p\|_2 \leq \|A - B\|_2 \quad (4.46)$$

Se A é uma matriz de posto menor ou igual a p , o teorema é obviamente verdadeiro, pois $\|A - A_p\|_2 = 0$. Supondo que posto de A é maior que p , então:

$$\|A - A_p\|_2^2 = \sigma_{p+1}^2 \quad (4.47)$$

Supondo agora que uma matriz B de posto no máximo p , de tal modo que B é uma melhor aproximação para A pela norma-2 do que A_p . Isto é, $\|A - B\|_2 < \sigma_{p+1}$.

O nulo de B , $Null(B)$ (conjunto de vetores v tais que $Bv = 0$), tem dimensão de pelo menos $n-p$.

Seja v_1, v_2, \dots, v_{p+1} , onde $p+1$ é o primeiro dos vetores singulares de A . Por dimensão, segue-se que existe um $z = 0$ em $Null(B) \cap Span \{v_1, v_2, \dots, v_{p+1}\}$. De forma que $|z| = 1$. O vetor z , que se encontra no espaço do primeiro vetor singular (v_{p+1}) de A , faz com que $(A - B)z \geq \sigma_{p+1}$. Assim a norma-2 de $A - B$ é pelo menos σ_{p+1} contradizendo que $\|A - B\|_2 < \sigma_{p+1}$.

Primeiro:

$$\|A - B\|_2^2 \geq |(A - B)z|^2 \quad (4.48)$$

Como $Bz = 0$

$$\|A - B\|_2^2 \geq |Az|^2 \quad (4.49)$$

Como z está em $\text{Span}\{v_1, v_2, \dots, v_{p+1}\}$:

$$|Az|^2 = \left| \sum_{i=1}^n \sigma_i u_i v_i^T z \right|^2 = \sum_{i=1}^n \sigma_i^2 (v_i^T z)^2 = \sum_{i=1}^{p+1} \sigma_i^2 (v_i^T z)^2 \geq \sigma_{p+1}^2 = \sum_{i=1}^{p+1} (v_i^T z)^2 = \sigma_{p+1}^2 \quad (4.50)$$

Segue-se que $\|A - B\|_2^2 \geq \sigma_{p+1}^2$ contradizendo a suposição que $\|A - B\|_2 < \sigma_{p+1}$. Ou seja, o mínimo é atingido quando:

$$Y = A_{[p^*]} = \sum_{p=1}^{p^*} \lambda_p u_p v_p^T, \quad (4.51)$$

assim $A_{[p^*]}$ é a matriz de ordem $m \times n$ com posto p^* que melhor aproxima a matriz A no sentido de (4.39).

Da SVD da matriz A , $A = U D_{\lambda} v^T$, ou seja $A = U F^T$, onde $F = V D_{\lambda}$. Seja:

$$F^T = \begin{bmatrix} F_{11} & \cdots & F_{n1} \\ \vdots & \vdots & \vdots \\ F_{1p} & \cdots & F_{np} \end{bmatrix} \quad (4.52)$$

Da relação $|a_1 a_2 \dots a_n| = |u_1 u_2 \dots u_p| F^T$ vem que:

$$a_1 = f_{11} u_1 + \cdots + f_{1p} u_p \quad (4.53)$$

$$a_n = f_{n1} u_1 + \cdots + f_{np} u_p \quad (4.54)$$

De modo que os números f_{ii} fornecem as coordenadas do vetor a_j no espaço gerado pelas colunas de U . $u_1 u_2 \dots u_{p^*}$ formam uma base para o subespaço ótimo. Desta forma, os

f_{ii} , $1 \leq i \leq p^*$ são as coordenadas da projeção de a_j no subespaço ótimo.

A projeção de a_1 sobre o espaço gerado por u_1, \dots, u_{p^*} é dada por:

$$p(a_1) = (a_{11})u_1 + \dots + (a_{1p^*})u_{p^*} \quad (4.55)$$

Então:

$$f_{ji} = a_j \cdot u_i \quad (4.56)$$

Como $a_j \cdot m_i = a_j^T D_\lambda u_i$:

$$f(i) = \begin{bmatrix} f_{li} \\ f_{ni} \end{bmatrix} \quad (4.57)$$

é denominado a i -ésima componente principal. E, assim a relação (4.56) é uma fórmula para obtermos o vetor $f(i)$.

4.2.3 Análise de Componentes Independentes

A Análise de Componentes Independentes (ACI) é uma técnica que identifica e recupera sinais estatisticamente independentes entre si, a partir de um conjunto de dados misturados. Funciona encontrando uma representação linear de dados não gaussianos, minimizando a dependência estatística entre eles, de forma que os componentes resultantes sejam estatisticamente independentes entre si o tanto quanto possível. Tem aplicação em diversas áreas de conhecimento, podendo ser utilizada em separação de sinais, como também em extração de características.

Existem duas condições para que a aplicação da ACI em um conjunto de dados seja eficiente. A primeira condição é que os sinais do conjunto de dados devem ser independentes. A segunda condição é que a quantidade de componentes independentes a ser identificado deve ser igual ao número de sinais no conjunto de dados misturados.

Um exemplo bem conhecido da aplicação de ACI é em sinais acústicos. Considerando que um grupo de pessoas falando ao mesmo tempo, o sinal acústico de cada pessoa será misturado ao das outras. Para duas pessoas poderem estabelecer a comunicação numa

conversa, é necessário que cada uma possa identificar as palavras que a outra emite. Ao passo que, cada pessoa escuta o sinal acústico de todos os participantes, o cérebro realiza o trabalho de reconhecer e separar o sinal acústico da pessoa emissora a que está interessado em se comunicar. O cérebro identifica o sinal acústico específico de uma pessoa pelo fato deste ser estatisticamente independente das outras pessoas, onde cada uma destas constitui uma fonte de sinal independente (37).

Vários algoritmos foram elaborados para automatizar a busca de fontes independentes utilizando ACI. Dentre eles, os mais utilizados são JADE, infomax e FastIca. Segundo o trabalho de (38), dos três algoritmos citados, o FastIca foi o que obteve melhor desempenho. Em outros trabalhos, (39), (40) e (41), o algoritmo FastIca é o mais aplicado em dados médicos, agropecuários, financeiros, elétricos, e outros.

4.2.3.1 Definição

Para estimar quantidades desconhecidas, considerando medidas conhecidas e restrições estatísticas sobre os dados originais, utiliza-se um modelo que expresse a relação entre as grandezas desconhecidas. O modelo mais utilizado é o modelo linear:

$$x(k) = As(k) \quad (4.58)$$

onde $x(k)$ é o vetor das medidas realizadas pelos sensores e $s(k)$ contém os dados originais. A matriz A é a matriz de mistura, que corresponde aos parâmetros de combinação dos sinais. Os coeficientes desta matriz são considerados constantes e desconhecidos.

A partir do modelo (4.58) e da condição de que os sinais devem ter fontes originais estatisticamente independentes, a ACI estima um conjunto de fatores $y(k)$ independentes ou independentes o quanto possível, através de um sistema de separação W , quando as medidas de $x(k)$ são conhecidas. A restrição estatística utilizada é que a função densidade de probabilidade das fontes sejam não-gaussianas, ou que, no máximo uma das fontes tenha função densidade de probabilidade gaussiana. O modelo de separação é descrito por:

$$y(k) = \hat{s}(k) = Wx(k) \quad (4.59)$$

W pode ser também considerada a matriz inversa de A . A determinação da matriz W passa a

ser o próximo problema. Este será resolvido se as fontes originais forem consideradas não-gaussianas, pois assim a condição de independência estatística será suficiente para a determinação dos coeficientes de W e dos componentes de $y(k)$.

4.2.3.2 Restrições do Método ACI

Considerando a restrição de que os sinais originais precisam ter fontes estatisticamente independentes entre si, a matriz de separação W é determinada de forma que os componentes de $y(k)$ também sejam estatisticamente independentes. Isto é, as variáveis são independentes se a função densidade de probabilidade conjunta de $y(k)$ puder ser fatorada no produto de suas funções densidade de probabilidade marginais, de acordo com a Equação (4.60):

$$p(y_1, \dots, y_n) = \prod_{i=1}^n p_{y_i}(y_i) \quad (4.60)$$

A Equação (4.60) pode ser utilizada para definir uma importante propriedade das variáveis aleatórias independentes, descrita em (34), conforme Equação (4.61):

$$E[g(y_1), \dots, g(y_n)] = E[g(y_1)]E[g(y_2)] \dots E[g(y_n)] \quad (4.61)$$

onde, $g(\cdot)$ são funções não lineares.

Na Equação (4.61) pode-se observar que se uma variável aleatória independente sofrer uma transformação não linear, a condição de independência não será afetada. Uma forma suave de independência é a não correlação. Um conjunto de n variáveis aleatórias (y_1, \dots, y_n) é não correlacionado se sua covariância é igual a zero, de acordo com a Equação (4.62):

$$E[y_1, \dots, y_n] - E[y_1]E[y_2] \dots E[y_n] = 0 \text{ para } y_1 \neq y_2 \neq \dots \neq y_n \quad (4.62)$$

As Equações (4.60) e (4.62) mostram que as variáveis aleatórias estatisticamente independentes são não correlacionadas, contudo, a não correlação não implica em independência. Exceto quando as variáveis aleatórias são gaussianas, a não correlação implica em independência.

Além de não poder garantir a independência, a não correlação também não garante a separação em componentes independentes. Para garantir a separação é necessário fazer uso de

estatísticas de ordem superior, que fazem com que a separação seja garantida para no máximo uma fonte gaussiana. Como a independência estatística implica na não correlação, a maioria dos algoritmos de ACI trabalham de forma que os componentes independentes sejam não correlacionados.

Outra restrição para a aplicação da ACI, é que os componentes independentes sejam não gaussianos. Admite-se que, no máximo, um desses componentes possa ser gaussiano.

A exemplo, considerando um sistema de sinais misturados (2 x 2) ortogonal e duas fontes gaussianas (s_1 e s_2). Os sinais resultantes (x_1 e x_2) são gaussianos, não correlacionados e tem variância unitária. Como a função densidade de probabilidade conjunta de (x_1 e x_2) é simétrica, esta não contém informação sobre a matriz de mistura A . Assim a matriz A não pode ser estimada, se mais de uma das fontes originais for gaussiana. A distribuição das variáveis gaussianas não é afetada por qualquer transformação ortogonal e as variáveis são independentes. Para variáveis não gaussianas a independência estatística é a principal garantia de que os sinais estimados serão distintos.

A Equação (4.63) descreve a função densidade de probabilidade para a distribuição gaussiana:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (4.63)$$

onde, σ é o desvio padrão e μ é a média.

4.2.3.3 Técnicas de Pré-Processamento

Existem dois principais métodos de pré-processamento empregados em ACI, que são a Centralização e o Branqueamento.

A centralização tem a função de simplificar os dados de entrada para o algoritmo. Trabalha subtraindo a média $E[x(k)]$ de cada elemento de $x(k)$, tornando $x(k)$ uma variável de média zero. Assim a matriz das componentes $s(k)$ também terá média zero. Após a estimativa da matriz de mistura W , com os dados centralizados, pode-se completar a estimativa somando vetor média de $s(k)$ ao vetor $s(k)$ centralizado.

No branqueamento, o vetor de medidas $x(k)$ é linearmente transformado em um novo vetor $z(k)$ branco e definido como $z(k) = Qx(k)$, onde Q é uma matriz de transformação linear. O branqueamento faz com que os componentes de $z(k)$ sejam não correlacionados e

que sua matriz de covariância seja unitária. Portanto, se Q é uma matriz de branqueamento para $x(k)$, então $z(k) = Qx(k)$, é branco, seus componentes são não correlacionados e sua variância é unitária.

Informações estatísticas de segunda ordem não permitem a separação em componentes independentes. Entretanto, a independência de segunda ordem ou não correlação espacial permite expressar o espaço dos sinais medidos em um sistema ortogonal, ou branco, que não implica na independência estatística dos sinais, mas que simplifica o problema. O branqueamento é aplicado após a centralização e antes do algoritmo principal.

O branqueamento pode ser aplicado através da decomposição em autovetores da matriz de covariância dos sinais medidos, ou seja:

$$E[x(k)x(k)^T] = MDM^T \quad (4.64)$$

onde M é a matriz ortogonal de autovetores de $E[x(k)x(k)^T]$ e D é a matriz diagonal de seus autovalores. Os dados branqueados são obtidos:

$$z(k) = MD^{-1/2}Mx(k) \quad (4.65)$$

onde a matriz $D^{-1/2}$, é calculada como $D^{-1/2} = (d_1^{-1/2}, \dots, d_n^{-1/2})$. Então, $E[x(k)x(k)^T] = I$, onde I é a matriz identidade.

O branqueamento transforma a matriz de mistura em uma nova matriz A_b :

$$z(k) = MD^{-1/2}Mx(k) \quad (4.66)$$

$$z(k) = MD^{-1/2}MA_s(k) \quad (4.67)$$

$$z(k) = A_b s(k) \quad (4.68)$$

A matriz branqueada A_b é uma matriz ortogonal, o que reduz o número de variáveis desconhecidas na matriz de mistura A de n^2 para $\frac{n(n-1)}{2}$.

Em sistemas de grandes dimensões, uma matriz ortogonal contém aproximadamente metade do número de parâmetros de uma matriz arbitrária. Para grandes dimensões, pode-se considerar que o branqueamento reduz metade do problema da ACI.

4.2.3.4 Estimativa dos Componentes Independentes

Para separar componentes utilizando ACI é necessário aplicar um algoritmo de otimização e uma função custo, ambos independentes. A função custo é qualquer função cuja otimização permita a estimativa de componentes independentes.

Pode-se estimar os componentes independentes através da medida da não gaussianidade dos componentes, sendo que, estes devem ser centralizados e ter variância unitária. A medida de não gaussianidade pode ser feita por meio do Kurtosis.

4.2.3.5 Medida de Não Gaussianidade por Kurtosis

O Kurtosis de uma variável aleatória $y(k)$ é definido como:

$$Kurt(y(k)) = E[y(k)^4] - 3(E[y(k)^2])^2 \quad (4.69)$$

como $y(k)$ tem variância unitária:

$$Kurt(y(k)) = E[y(k)^4] - 3 \quad (4.70)$$

Assim, o kurtosis é simplesmente uma normalização do momento de quarta ordem, ou seja, $E[y(k)^4]$.

Uma propriedade importante do kurtosis para uma variável gaussiana, é que o momento de quarta ordem é dado por $3E[[S^2(k)]^2]$. Como a variância $E[S^2(k)] = 1$, o kurtosis é nulo. Por isso, variáveis gaussianas não são aconselháveis para ACI.

O kurtosis de uma variável aleatória pode ser positivo, negativo ou nulo, sendo que:

- para $k > 0$ a variável é sub-gaussiana. Um exemplo de função densidade de probabilidade é a uniforme;
- para $k < 0$ a variável é super-gaussiana. Um exemplo de função densidade de probabilidade é a Laplaciana;
- para $k = 0$ a variável é gaussiana. Um exemplo de função densidade de probabilidade gaussiana ou normal.

A não gaussianidade pode ser medida através do valor absoluto do kurtosis, sendo que:

- $|Kurt| = 0$, as variáveis são gaussianas;
- $|Kurt| > 0$, as variáveis são não gaussianas.

Os componentes independentes são encontrados maximizando ou minimizando o kurtosis, de acordo com o algoritmo de otimização aplicado. Para encontrar a matriz de separação W que fornecerá os componentes independentes $y(k)$ com base na medida dos kurtosis das fontes, deve-se:

- Utilizar o branqueamento para obter $y(k)$;
- Inicializar W com um valor qualquer;
- Obter a direção em que o kurtosis de $y(k)$ cresce ou decresce mais acentuadamente;
- Encontrar um novo vetor W ;
- Repetir o processo descrito até que todas as fontes sejam estimadas.

4.2.3.6 Negentropia

A entropia H para uma dada variável aleatória define o grau de informação dessa variável aleatória. Quanto mais imprevisível for a variável observada, maior será sua entropia. Se há uma grande probabilidade de um evento específico ocorrer, o ganho informacional não é tão grande em cada observação, pois já se espera que este evento seja bastante provável. As variáveis gaussianas possuem maior entropia entre variáveis de mesma variância. Assim, a entropia pode ser usada como medida de não-gaussianidade:

$$H(y) = -\sum P(y = a_i) \log(P(y = a_i)) \quad (4.71)$$

onde $P(y = a_i)$ é a probabilidade de $y(k) = y$ ser igual a a_i . A negentropia (J) é dada por:

$$J(y) = H(y_{gauss}) - H(y) \quad (4.72)$$

onde y_{gauss} é uma variável aleatória gaussiana com a mesma matriz de covariância de y . Deste modo, obtém-se uma medida que é sempre positiva e mede a diferença de entropia

existente entre a variável gaussiana com os mesmos parâmetros. Esta medida recebe o nome de negentropia, constituindo-se numa versão normalizada da entropia e pode ser utilizada pelo algoritmo de ICA para encontrar os componentes independentes baseando-se na não gaussianidade da distribuição. A negentropia aproximada é calculada através de momentos de ordem superior, dado por:

$$J(y) = \frac{1}{12}E[y^3]^2 + \frac{1}{48}Kurtosis(y)^2 \quad (4.73)$$

Como y tem média zero e variância unitária e a validade desta aproximação é limitada, devido utilizar a kurtosis, acaba sofrendo a não robustez dela, chegando algumas vezes a ser imprecisa e em muitos casos sensíveis a *outliers*. Assim, outras técnicas foram propostas. A técnica mais simples, utiliza uma função não quadrática $G(\cdot)$ aplicável em função não quadrática:

$$J_G(y) = \alpha[E[G(y)] - E[G(b)]]^2 \quad (4.74)$$

onde b é uma variável aleatória gaussiana com média zero e variância unitária. As funções normalizadas usadas são:

$$G_1(u) = u^3, G_2(u) = \tanh(a_1 \cdot u), G_3(u) = u \cdot \exp\left(-a_2 \frac{u^2}{2}\right), G_4(u) = u^2 \quad (4.75)$$

Um dos principais algoritmos que utilizam a aproximação da negentropia para medida de não gaussianidade é o algoritmo rápido de ponto fixo (FastICA), um dos mais utilizados e que apresentam uma melhor estimativa das componentes independentes.

4.2.3.7 Algoritmo Rápido de Ponto Fixo (FastICA)

O algoritmo rápido de ponto fixo denominado de FastICA visa encontrar componentes independentes através da maximização da negentropia. Procura encontrar um vetor cuja projeção $W^T x(k)$ maximiza a função $J_G(W)$, ou seja, uma aproximação da negentropia dada por:

$$J_G(W) = [E[G(W^T X(k))] - E[G(b)]]^2 \quad (4.76)$$

onde $G(\cdot)$ é qualquer função não quadrática e b é uma variável gaussiana de média zero e variância unitária, $E[\cdot]$ são as esperanças.

A regra de treinamento do algoritmo consiste de um método que estima os componentes um a um. Em dados branqueados, o algoritmo FastICA pode ser equacionado da seguinte forma:

1. Inicializar W com um valor qualquer;
2. $W_{n+1} = E[X \cdot G(W^T \cdot X)] - E[G'(W^T \cdot X)] \cdot W$;
3. $W = \frac{W_{n+1}}{\|W_{n+1}\|}$.
4. Repetir o passo 2 até que haja convergência.

As esperanças $E[\cdot]$ são estimadas através da média amostral de um determinado número de amostras dos dados de entrada para garantir a precisão.

O algoritmo FastICA possui vários atributos importantes citados abaixo:

- O FastICA não necessita de ajuste no passo de adaptação comparado com algoritmos baseados em gradiente;
- Por ter uma estrutura paralela e distribuída, mas não adaptativa, pode ser considerado um algoritmo neural;
- A baixa quantidade de interações, de 5 a 10, que são necessárias na maioria dos casos faz com que o método tenha uma máxima precisão com um pequeno número de interações;
- O FastICA utiliza as médias amostrais calculadas sobre uma amostra grande de dados, ao invés de utilizar todos os dados imediatamente para treinamento;
- O algoritmo encontra diretamente as fontes para praticamente qualquer função não gaussiana, ou seja, não é necessário que seja feita uma distinção entre as fontes com kurtosis positivas ou negativas;
- O desempenho do algoritmo pode ser melhorado de acordo com a escolha da função não linear utilizada;
- Identifica os componentes independentes um a um.

4.3 TÉCNICAS DE CLASSIFICAÇÃO

4.3.1 K-Vizinhos Mais Próximos

No método K-Vizinhos Mais Próximos, (K-NN, do inglês *K-Nearest Neighbors*) é considerada a proximidade entre os dados para a realização de problemas envolvendo classificação. As variações dos algoritmos são definidas pelo número escolhido de vizinhos mais próximos. A técnica parte do pressuposto de que dados que apresentam similaridades concentram-se em uma mesma região do espaço de entrada (42). A classificação de um objeto baseia-se no conjunto de dados de treinamento para verificar quais são os objetos nesta base de dados que mais se assemelham ao novo objeto. Este será classificado dentro da classe mais comum a que pertencem os objetos mais similares a ele, ou seja, a classificação é realizada por analogia.

O algoritmo do 1-Vizinho Mais Próximo (1-NN, do inglês *1-Nearest Neighbors*) constitui-se na variação mais simples do algoritmo KNN. Cada objeto representa um ponto no espaço de entrada, onde são calculadas as distâncias entre cada dois pontos. Normalmente, para calcular essas distâncias, utiliza-se como métrica a distância euclidiana, dada pela Equação (4.77):

$$d(x_i, x_j) = \sqrt{\sum_{l=1}^d (x_i^l - x_j^l)^2} \quad (4.77)$$

onde x_i e x_j são dois objetos representados por vetores no espaço \mathfrak{R}^d , e x_i^l e x_j^l são elementos desses vetores, que correspondem aos valores da coordenada l .

Para classificar um exemplo não rotulado o algoritmo precisa de antemão memorizar os exemplos rotulados do conjunto de treinamento. Feito isso, o algoritmo calcula a distância entre o vetor de valores de atributos e cada exemplo rotulado. O rótulo da classe associado ao exemplo de treinamento mais próximo do exemplo de teste é utilizado para classificar o novo exemplo.

Comumente, é necessário normalizar os valores de cada atributo, para que todos tenham um mesmo intervalo de variação, não havendo muita discrepância entre os valores dos diferentes atributos. O processo de normalização consiste em v sendo um valor do atributo, considerando que o valor $v' \in [0, 1]$ é calculado como:

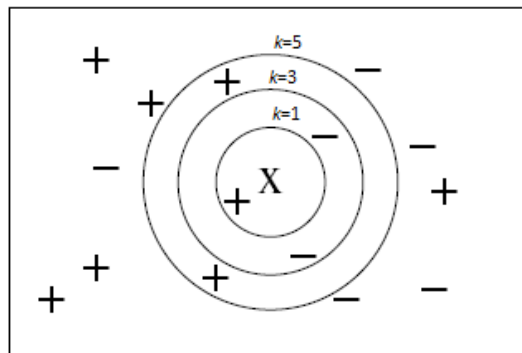
$$v' = \frac{v - \min_l}{\max_l - \min_l} \quad (4.78)$$

onde \min_l e \max_l são os valores mínimos e máximos que pode assumir o atributo l .

Considerando que, ao invés de 1 vizinho, são considerados os k vizinhos mais próximos do conjunto de treinamento do objeto de teste x_t . O valor de k é um parâmetro do algoritmo definido pelo usuário. Em (42), quando o valor de k é maior que 1, para cada objeto de teste, são obtidos k vizinhos. Cada vizinho soma voto em uma classe de acordo com a sua, para a classe futura do objeto de teste. As previsões dos diferentes vizinhos são agregadas de forma a classificar o ponto de teste. O objeto de teste é classificado na classe mais votada. Matematicamente, esse processo é equivalente a: $\hat{f}(x_t) \leftarrow \text{moda}(f(x_1), f(x_2), \dots, f(x_k))$, pois a constante que minimiza a função de custo 0-1 é a moda.

Como exemplo, na Figura (4.1), em um problema com duas classes o objeto de teste é representado por “x”, e k é definido com os valores 1, 3 e 5. No espaço definido pelos atributos, e usando a distância euclidiana, o objeto de teste será classificado com o mesmo rótulo da maior quantidade de objetos de treinamento mais próximos de x .

Figura 4.1 - Classificação de objeto usando K-NN. Para $k=1$, x é assumido positivo; $k=3$, x é negativo; $k=5$, x é positivo.



Fonte: Elaborado pelo autor.

Frequentemente, o valor de k escolhido é pequeno e ímpar, como $k=3, 5, \dots$. Valores pares não são recomendados para evitar empates. O valor de k também pode ser definido por validação cruzada e através da associação de um peso à contribuição de cada vizinho. Nesse último caso, a contribuição de cada um dos k vizinhos é pesada de forma inversamente proporcional à distância ao ponto de teste. Assim, é possível utilizar k igual à quantidade de todos os objetos de treinamento.

Nos problemas de classificação a moda ponderada: $y_t = \operatorname{argmax}_{c \in Y} \sum_{i=1}^k w_i I(c, y_i)$,

com $w_i = \frac{1}{d(x_t, x_i)}$ e $I(a, b)$ é uma função que retorna 1 se e só se $a = b$, em que y_i é a classe do exemplo x_i , w_i é o peso associado ao exemplo x_i e c é a classe com maior moda ponderada.

O método K-NN tem a vantagem de possuir um algoritmo de treinamento simples, ser aplicável mesmo em problemas de classificação mais complexos e ter característica incremental. No entanto, o custo computacional para o processamento do algoritmo é considerável.

4.3.2 Floresta Randômica

Em (43), Floresta Randômica é uma técnica de aprendizado de máquina onde diversas árvores de decisão são combinadas para resolver problemas, de modo que cada árvore depende dos valores dos vetores aleatórios amostrados independentes. Com a mesma distribuição de todas as árvores na floresta, os resultados são combinados para determinar a classe de cada objeto. Tem vantagem em relação aos modelos com uma única árvore de decisão por esta ser sensível a determinados tipos de dados.

A técnica combina conceitos de *bagging*, onde modelos individuais são desenvolvidos através de amostragem com reposição de dados do conjunto de treinamento, com o método de subespaço randômico (*randomizing*), onde cada árvore em um conjunto é construída de uma amostra randômica de atributos (43).

De modo sucinto, funciona com a construção de cada árvore com uma amostra de dados do conjunto de treinamento. Em seguida, para cada nó de cada árvore, apenas uma parte dos atributos selecionada aleatoriamente é considerada. Os parâmetros a serem definidos pelo usuário são o número total de árvores a serem construídas e a quantidade de atributos usada para cada nó. A classificação de um objeto funciona através de votação, onde cada objeto é classificado por cada árvore individualmente, sendo o critério da maioria dos votos utilizado na maioria das aplicações.

Matematicamente, sendo $D = \langle x_i, c_i \rangle \mid x_i = (y_{i,1}, y_{i,2}, \dots, y_{i,m})$ um vetor de características e $c_i \in \{-1, 1\}, 1 \leq i \leq n$. O método floresta randômica permite a utilização de diversos tipos de atributos, isto é, o conjunto-domínio de cada característica $y_{i,j}$ de x_i , podendo ser um conjunto de dados categóricos ou um conjunto numérico.

Considera-se T o número total de árvores inseridas pelo usuário e $M \leq m$ é o número de atributos do conjunto de m atributos considerados para cada nó de cada árvore.

Antes de gerar cada árvore t , onde $1 \leq t \leq T$, é gerado uma amostra D_{At} de D , obtendo de modo randômico, com reposição, $|D|$ elementos de D . A amostra D_{At} é utilizada como conjunto de treino para construção de cada árvore t . Então, inicia-se a construção da árvore.

Cada nó da árvore representa um subconjunto S de D_{At} . O nó raiz representa o próprio conjunto D_{At} . Em cada nó utiliza-se a característica y_j , entre M atributos, que particione o conjunto S correspondente em dois outros subconjuntos disjuntos S_1 e S_2 que correspondem, respectivamente, às subárvores esquerda e direita do nó correspondente a S . S_1 e S_2 devem ser os mais puros possíveis em relação às classes -1 e 1, ou seja, devem ser tais que o atributo y_j dentre os M atributos é o que melhor divide S , e também, o valor de v_{y_j} de y_j que gera a melhor divisão, em termos de pureza.

O critério mais utilizado para determinar a pureza de uma subárvore é o ganho de informação, que resulta da bisseção do conjunto S nos conjuntos S_1 e S_2 , definida por um valor v_{y_j} de y_j . O ganho de informação $GI(S, y_j, v_{y_j})$ devido a um valor v_{y_j} de uma característica y_j , em um conjunto S , é dado pela Equação (4.79):

$$GI(S, y_j, v_{y_j}) = E(S) - (p(S_1)E(S_1) + p(S_2)E(S_2)) \quad (4.79)$$

Na Equação (4.79), $E(X)$ corresponde à entropia do conjunto X e $p(X)$ corresponde à probabilidade de ocorrência de X . A entropia de um conjunto X em relação às possíveis classes C que ocorrem em X é dada pela Equação (4.80):

$$E(x) = - \sum_{c=1}^{|C|} p(c_i) \log_2 p(c_i) \quad (4.80)$$

Ao término da construção da floresta, cada árvore emite um voto -1 ou 1, em que a maioria dos votos computados ditará a classe do objeto.

Para a construção da técnica floresta randômica utiliza-se as seguintes etapas:

1. Selecionar randomicamente M das m características possíveis;
2. Se todos os elementos de S possuírem a mesma classe ou se todos os elementos de S possuírem os mesmos valores para cada característica y_j escolhida, não dividir o conjunto S e criar um nó folha para representar S na árvore. Caso contrário, a próxima etapa será seguida.

3. Para cada característica y_j escolhida, determinar o valor v_{y_j} de y_j que divida S em S_1 e S_2 , sendo que estes correspondam à partição mais pura possível entre todas as partições possíveis com os diferentes valores de y_j em S . Se y_j for uma característica contínua ou ordinal, a classificação é feita em ordem crescente dos seus valores. Sejam x um elemento de S e v_{xj} o valor da característica y_j correspondente ao elemento x . Incluir x em S_1 se $v_{xj} \leq v_{y_j}$, ou incluir x em S_2 se $v_{xj} > v_{y_j}$. Mas, se y_j for uma característica nominal, para cada valor v_{y_j} de y_j percorrer todos os elementos x de S . Seja v_{xj} o valor da característica y_j correspondente ao elemento x . Se $v_{xj} = v_{y_j}$ incluir x em S_1 , mas se $v_{xj} \neq v_{y_j}$ incluir x em S_2 ;
4. Determinar o par $\langle y_j, v_{y_j} \rangle$ que obtém a melhor partição de S ;
5. Gerar a bipartição de S , criando as subárvores S_1 e S_2 ;
6. Repetir todo o processo desde I, para S_1 e S_2 .

4.3.3 Máquinas de Vetores de Suporte

Máquinas de Vetores de Suporte (MVS) é uma técnica estatística desenvolvida para resolver problemas de predição. MVS classifica padrões através da construção de uma superfície de decisão, chamada hiperplano, que procura separar padrões linearmente separáveis. Para padrões não linearmente separáveis, é utilizada uma função de mapeamento mais adequada para separar os padrões de maneira mais eficiente (44).

A técnica é eficiente para trabalhar com dados multidimensionais, através do hiperplano de margem máxima, cujos dados mais próximos deste hiperplano serão os vetores de suporte (44). Em cada classe haverá pelo menos um vetor de suporte, geralmente existindo mais de um.

Em (45), as MVS possuem algumas características atrativas:

- Boa capacidade de generalização: A capacidade de generalização de um classificador é dada pela sua eficiência na classificação de dados que não pertençam ao conjunto utilizado em seu treinamento. Ou seja, na geração de preditores é evitado o *overfitting*, caso em que o preditor se torna muito especializado no conjunto de treinamento, obtendo baixo desempenho ao classificar novos padrões.
- Robustez em grandes dimensões: MVS são robustas diante de objetos de grandes

dimensões. Em outras técnicas inteligentes é comum a ocorrência de *overfitting* sobre esses tipos de dados.

- Convexidade da função objetivo: A aplicação das MVS implica na otimização de uma função quadrática, que possui apenas um mínimo global.

Neste trabalho será apenas explanado MVS para padrões linearmente separáveis.

4.3.3.1 Teoria do Aprendizado Estatístico

A Teoria do Aprendizado Estatístico objetiva encontrar um bom classificador, considerando todo o conjunto de dados. O desempenho esperado de um classificador f é que este tenha o menor erro de predições durante o treinamento. A medida de perda entre a resposta desejada e a resposta real é o risco empírico $R_{emp}(f)$, dado pela Equação (4.81):

$$R_{emp}(f) = \frac{1}{n} \sum_{i=1}^n c(f(x_i), y_i) \quad (4.81)$$

onde $c(\cdot)$ é a função de custo relacionada a previsão $f(x_i)$ com a saída desejada y_i , onde um tipo de função de custo é a perda 0/1 definida pela Equação (4.82).

$$c(f(x_i), y_i) = \begin{cases} 1, & \text{se } y_i f(x_i) < 0 \\ 0, & \text{caso contrário} \end{cases} \quad (4.82)$$

O processo de busca por uma função f' que represente um menor valor de R_{emp} é denominado de Minimização do Risco Empírico.

Os padrões de treinamento (x_i, y_i) são gerados por uma distribuição de probabilidade $P(x, y)$ em $\mathbb{R}^N \times \{-1, 1\}$ sendo P desconhecida. A probabilidade de classificação incorreta do classificador f é denominada de Risco Funcional, o qual quantifica a capacidade de generalização, conforme dado pela Equação (4.83):

$$R(f) = \int c(f(x_i), y_i) dP(x_i, y_i) \quad (4.83)$$

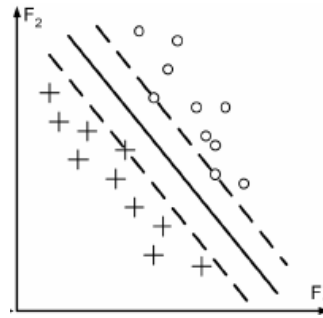
Durante o treinamento, $R_{emp}(f)$ é mais fácil de ser obtido que $R(f)$, pois, geralmente a distribuição de probabilidades P é desconhecida. A partir disto, dado um conjunto de dados de treinamento (x_i, y_i) com $x_i \in \mathbb{R}^N$ e $y_i \in \{-1, +1\}$, $i = \{1, 2, \dots, n\}$, sendo x_i o vetor de entrada

e y_i o rótulo da classe.

O objetivo é estimar uma função $f: \mathbb{R}^N \rightarrow \{-1, +1\}$. Se não houver restrição imposta na classe de funções em que se escolhe a estimativa f , a função pode ter um bom desempenho no conjunto de treinamento, porém não tendo o mesmo desempenho em padrões desconhecidos. Ou seja, a minimização apenas do risco empírico $R_{emp}(f)$ não garante uma boa capacidade de generalização, sendo desejado um classificador f^* tal que $R(f^*) = \min_{f \in F} R(f)$, onde F é o conjunto de funções f possíveis.

A Figura (4.2) ilustra uma classe de funções que pode ser utilizada para separar padrões linearmente separáveis. É necessário determinar uma função que minimize o R_{emp} .

Figura 4.2 - Classe de hiperplanos com um hiperplano ótimo.



Fonte: Elaborado pelo autor.

A teoria do aprendizado estatístico proporciona formas de limitar os hiperplanos, para prevenir modelos ruins, implementando uma função com a capacidade adequada para o conjunto de dados de treinamento. Estas limitações são impostas ao risco funcional da função. Os limites utilizam o conceito de dimensão Vapnik-Chervonenkis (VC).

Dado um conjunto de funções sinal G , sua dimensão VC é definida como o tamanho do maior conjunto de pontos que pode ser particionado aleatoriamente pelas funções contidas em G . A dimensão VC do conjunto de funções de classificação G é o número máximo de exemplos de treinamento que pode ser aprendido pela máquina sem erro. Para funções lineares no \mathbb{R}^N para $n \geq 2$ a dimensão VC é dada pela Equação (4.84):

$$VC(n) = n + 1 \quad (4.84)$$

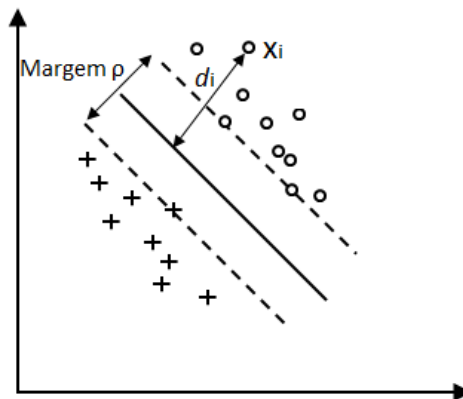
A margem separa duas classes da melhor maneira. É a menor distância entre os exemplos do conjunto de treinamento e o hiperplano utilizado para separação destas classes. O hiperplano é

dado por $f(x) = (w \cdot x) + b$. A margem máxima é obtida com a utilização do hiperplano ótimo. A margem ρ de um classificador f é definida por:

$$\rho = \min_i y_i f(x_i) \quad (4.85)$$

A margem é obtida pela distância entre o hiperplano e os vetores que estão mais próximos a ele, sendo estes os vetores suporte. Os vetores suporte são padrões críticos, que determinam o hiperplano ótimo. Os outros padrões irrelevantes podem ser removidos do conjunto de treinamento. Na Figura (4.3), são ilustrados os vetores suporte sobre as linhas pontilhadas e a margem.

Figura 4.3 - A margem ρ e o vetores suporte sobre a linha pontilhada.



Fonte: Elaborado pelo autor.

4.3.3.2 Classificação Linear

Classificação linear consiste em determinar uma função $f: X \subseteq \mathbb{R}^N \rightarrow \mathbb{R}^N$, que atribui um rótulo (+1) se $f(x) \geq 0$ e (-1), caso contrário. Considerando uma função linear, dada pela Equação (4.86):

$$f(x) = \langle w \cdot x \rangle + b \quad (4.86)$$

$$f(x) = \sum_{i=1}^n w_i x_i + b \quad (4.87)$$

onde w e $b \in \mathbb{R}^N \times \mathbb{R}^N$, são o vetor peso e bias, respectivamente, responsáveis por controlar

a função e a regra de decisão. Os valores de w e b são obtidos pelo processo de aprendizagem a partir dos dados de entrada.

O vetor peso e bias podem ser interpretados geometricamente sobre um hiperplano. Um hiperplano é um subespaço afim, que divide um espaço em duas partes, correspondendo à dados de duas classes distintas. O vetor peso define uma direção perpendicular ao hiperplano, e com a variação do bias o hiperplano é movido paralelamente a ele mesmo.

4.3.3.3 Hiperplano Ótimo

O hiperplano ótimo separa perfeitamente os dados de cada classe com uma margem de separação máxima. Considerando o conjunto de treinamento linearmente separável, o hiperplano ótimo é o hiperplano de separação com maior margem. O hiperplano ótimo é dado por:

$$\langle w \cdot x \rangle + b = 0 \quad (4.88)$$

Considerando a restrição imposta pela Equação (4.89), os classificadores lineares que separam um conjunto de treinamento possuem margem positiva. Ou seja, esta restrição afirma que não há nenhum dado entre $\langle w \cdot x \rangle + b = 0$ e $\langle w \cdot x \rangle + b = \pm 1$, sendo a margem sempre maior que a distância entre os hiperplanos $\langle w \cdot x \rangle + b = 0$ e $|\langle w \cdot x \rangle + b = 1|$. Devido a estas suposições as MVS obtidas são normalmente chamadas MVS com margens rígidas.

$$\begin{cases} \langle w \cdot x_i \rangle + b \geq +1, & \text{para } y_i = +1 \\ \langle w \cdot x_i \rangle + b \leq -1, & \text{para } y_i = -1 \end{cases} \quad (4.89)$$

Estas equações podem ser combinadas em:

$$y_i(\langle w \cdot x_i \rangle + b) \geq 1, \quad i = \{1, 2, \dots, n\} \quad (4.90)$$

Seja d_+ (d_-) a distância euclidiana entre os vetores suporte positivos (negativos) e o hiperplano, a margem ρ de um hiperplano de separação sendo a maior margem geométrica entre todos os hiperplanos, $\rho = (d_+ + d_-)$. $d_i(w, b; x_i)$, é a distância de um dado x_i ao hiperplano (w, b) , sendo calculado pela Equação (4.91):

$$d_i(w, b; x_i) = \frac{|\langle w, x_i \rangle + b|}{\|w\|} = \frac{y_i(\langle w, x_i \rangle + b)}{\|w\|} \quad (4.91)$$

Considerando a restrição imposta pela Equação (4.90), pode-se escrever:

$$d_i(w, b; x_i) \geq \frac{1}{\|w\|} \quad (4.92)$$

Assim, identifica-se $\frac{1}{\|w\|}$ como o limite inferior da distância entre os vetores suporte x_i e o hiperplano de separação (w, b) , as distâncias d_+ e d_- :

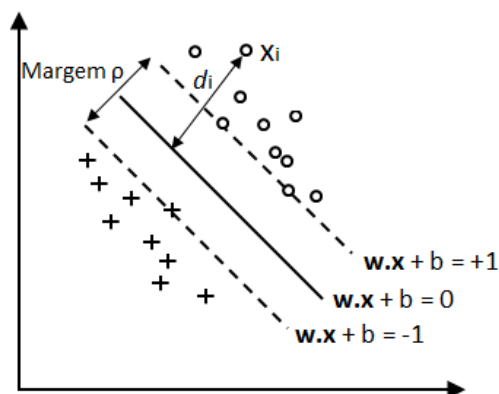
$$d_+ = d_- = \frac{1}{\|w\|} \quad (4.93)$$

A margem é sempre maior que a última instância, a minimização de $\|w\|$ leva a maximização da margem. A partir disto pode-se definir a margem ρ :

$$\rho = (d_+ + d_-) = \frac{2}{\|w\|} \quad (4.94)$$

Na Figura (4.4) é ilustrada a distância entre hiperplanos e os vetores suporte.

Figura 4.4 - Distância entre hiperplanos e vetores suporte.



Fonte: Elaborado pelo autor.

O hiperplano ótimo é dado pela minimização da norma $\|w\|$, considerando a restrição

da Equação (4.90). Pode-se assim reescrever:

$$\begin{aligned} &\text{Problema } P1 \\ &\text{Minimizar } \|w\| \qquad (4.95) \\ &\text{Sujeito a } y_i(\langle w \cdot x_i \rangle + b) \geq 1, \text{ para } i = \{1, 2, \dots, n\} \end{aligned}$$

O problema clássico de otimização acima, denominado programação quadrática pode ser resolvido com o método clássico de multiplicadores de Lagrange. Utilizando a teoria dos multiplicadores de Lagrange:

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^n \alpha_i (y_i (\langle w \cdot x_i \rangle + b) - 1) \quad (4.96)$$

onde α_i são os multiplicadores de Lagrange. O próximo passo é a minimização de $L(w, b, \alpha)$, em relação a w e b , e a maximização dos α_i . O método dos multiplicadores encontra os pontos ótimos igualando as derivadas parciais a zero. Sendo assim os pontos ótimos da Equação (4.96) são obtidos por meio da resolução das igualdades:

$$\frac{\partial L}{\partial b} = 0 \quad (4.97)$$

$$\frac{\partial L}{\partial w} = 0 \quad (4.98)$$

Sendo:

$$\frac{\partial L}{\partial w} = \left(\frac{\partial L}{\partial w_1}, \frac{\partial L}{\partial w_2}, \dots, \frac{\partial L}{\partial w_n} \right) \quad (4.99)$$

A partir da Equação (4.97) e (4.98) obtém-se:

$$\sum_{i=1}^n \alpha_i y_i = 0 \quad (4.100)$$

$$w = \sum_{i=1}^n \alpha_i y_i x_i \quad (4.101)$$

Substituindo as equações (4.100) e (4.101) no lado direito da Equação (4.96), chegamos ao seguinte problema de otimização:

Problema P2

$$\text{Maximizar } \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j \langle x_i \cdot x_j \rangle \quad (4.102)$$

$$\text{Sujeito a } \begin{cases} \alpha_i \geq 0, & i = \{1, \dots, n\} \\ \sum_{i=1}^n \alpha_i y_i = 0 \end{cases}$$

O objetivo então é determinar os valores ótimos de (w, b) , que representaremos por (w^*, b^*) . Através da Equação (4.101) podemos calcular w^* e b^* como segue:

$$w^* = \sum_{i=1}^n \alpha_i^* y_i x_i \quad (4.103)$$

o valor de b^* pode ser obtido utilizando as equações de Karush-Kuhn-Tucker (KKT):

$$\alpha_i^* (y_i (\langle w^* \cdot x_i \rangle + b^*) - 1) = 0, \quad i = \{1, \dots, n\} \quad (4.104)$$

Os valores de α_i^* assumem valores positivos para os exemplos de treinamento que estão a uma distância do hiperplano ótimo igual a largura da margem, ou seja, os vetores suporte. Para o restante dos exemplos, α_i^* é nulo. Portanto, o hiperplano ótimo é obtido pelos vetores suporte. Dado um vetor suporte x_j , obtém-se b^* por meio da condição de KKT:

$$b^* = y_j - \langle w^* \cdot x_j \rangle \quad (4.105)$$

Com os valores dos parâmetros w^* e b^* calculados, pode-se classificar de um novo padrão z apenas calculando:

$$\text{sgn}(\langle w^* \cdot z \rangle + b^*) \quad (4.106)$$

A classificação é dada pelo cálculo do produto interno entre o novo padrão e todos os vetores suporte.

4.3.3.4 Classificação Multiclasses

As MVS foram propostas inicialmente como ferramenta de classificação binária.

Entretanto, a maioria das aplicações reais possuem características multiclases. Para estes casos, alguns procedimentos para MVS multiclases foram elaborados como a seguir:

Num sistema multiclases o conjunto de treinamento é composto por pares (x_i, y_i) , tal que $y_i \in \{1, \dots, k\}$, com $k > 2$, sendo k o número de classes.

Os principais métodos utilizam como base a decomposição de um problema multiclasse com $k > 2$, em k problemas binários são: decomposição Um-Contra-Todos (UCT) e decomposição Todos-Contra-Todos (TCT).

O método UCT baseia-se na construção de k classificadores binários. Cada classificador f_i classifica uma classe i das demais. Em um novo padrão x , a classe a qual este novo padrão pertence é a classe representada pelo classificador que obteve o valor máximo entre o k classificadores. É definido por:

$$f(x) = \arg \max_{1 \leq i \leq k} (f_i(x)) \quad (4.107)$$

O método (TCT) compara as classes duas a duas, sendo necessárias $k \cdot (k - 1)/2$ MVS. Para decidir a qual classe pertence um novo padrão x , utiliza-se um esquema de votação por maioria.

4.4 CONCLUSÕES DO CAPÍTULO

Este capítulo apresentou as técnicas de inteligência computacional utilizadas na presente dissertação. As técnicas ANOVA, ACP e ACI, foram utilizadas para tratar os dados de DPs para diminuir a quantidade de informações ruidosas que interferem na leitura dos dados e aumentam o esforço computacional realizado pelas técnicas de classificação.

ANOVA e ACP atuam reduzindo a dimensionalidade dos dados. ANOVA faz essa redução analisando a diferença significativa entre as médias e se as variáveis influenciam em alguma variável independente. ACP trabalha procurando a existência de padrões ocultos, através dos componentes principais, encontrando formas mais representativas de dados através de combinações lineares das variáveis originais. Assim, essas técnicas apresentam potencial em tratar os dados para classificação.

A técnica ACI identifica e separa os sinais estatisticamente independentes entre si, encontrando uma representação linear de dados não gaussianos, minimizando a dependência estatística entre eles, de forma que os componentes resultantes sejam estatisticamente

independentes. Esta técnica apresentou um bom potencial para extrair os sinais ruidosos dos dados de DPs, em condições favoráveis, no entanto, os resultados são comprometidos quando os sinais são de fontes diferentes.

Para realizar a classificação dos dados de DPs, foram utilizadas três ferramentas: KNN, Floresta Randômica e MVS. KNN realiza uma classificação por analogia, considerando que dados semelhantes concentram-se em uma mesma região do espaço de entrada. Floresta Randômica combina diversas árvores de decisão que dependem dos valores dos vetores aleatórios independentes e com a mesma distribuição de todas as árvores na floresta. MVS baseia-se na construção do hiperplano para separar padrões linearmente separáveis.

As técnicas abordadas neste capítulo apresentam boas perspectivas quanto ao tratamento e classificação dos dados de DPs.

5.0 METODOLOGIA, EXPERIMENTOS E RESULTADOS

5.1 INTRODUÇÃO

Este capítulo apresenta a metodologia aplicada para a classificação dos dados de DPs coletadas da Usina Hidrelétrica de Tucuruí (UHT), utilizando para isso as técnicas de inteligência artificial empregadas.

São apresentados os procedimentos metodológicos, a origem dos dados coletados e quais instrumentos foram utilizados para sua captação, o software utilizado para tratar e classificar os dados de DPs e seu processo de tratamento e classificação.

Os sinais de DPs encontrados nos dados coletados da UHT, do tipo descarga interna, delaminação e entre barras, foram submetidos à experimentos com fase de treino e teste nas técnicas inteligentes descritas no Capítulo 4. Todos os experimentos foram realizados no ambiente IPython.

Os sinais de DPs utilizados foram um a um analisados e rotulados manualmente, quando apresentavam padrões mais próximos possíveis dos padrões que constam na norma (15). Dessa forma foi possível realizar o aprendizado supervisionado, utilizando parte dos dados para o treino e a outra parte para ser avaliada durante o teste. O objetivo do treinamento foi identificar os padrões de DPs, e o da fase de teste foi medir a acurácia das técnicas de classificação destes padrões.

Assim, neste capítulo são apresentados os experimentos realizados com os dados de DPs utilizados e, por fim, os resultados obtidos em cada procedimento.

5.2 PROCEDIMENTOS METODOLÓGICOS

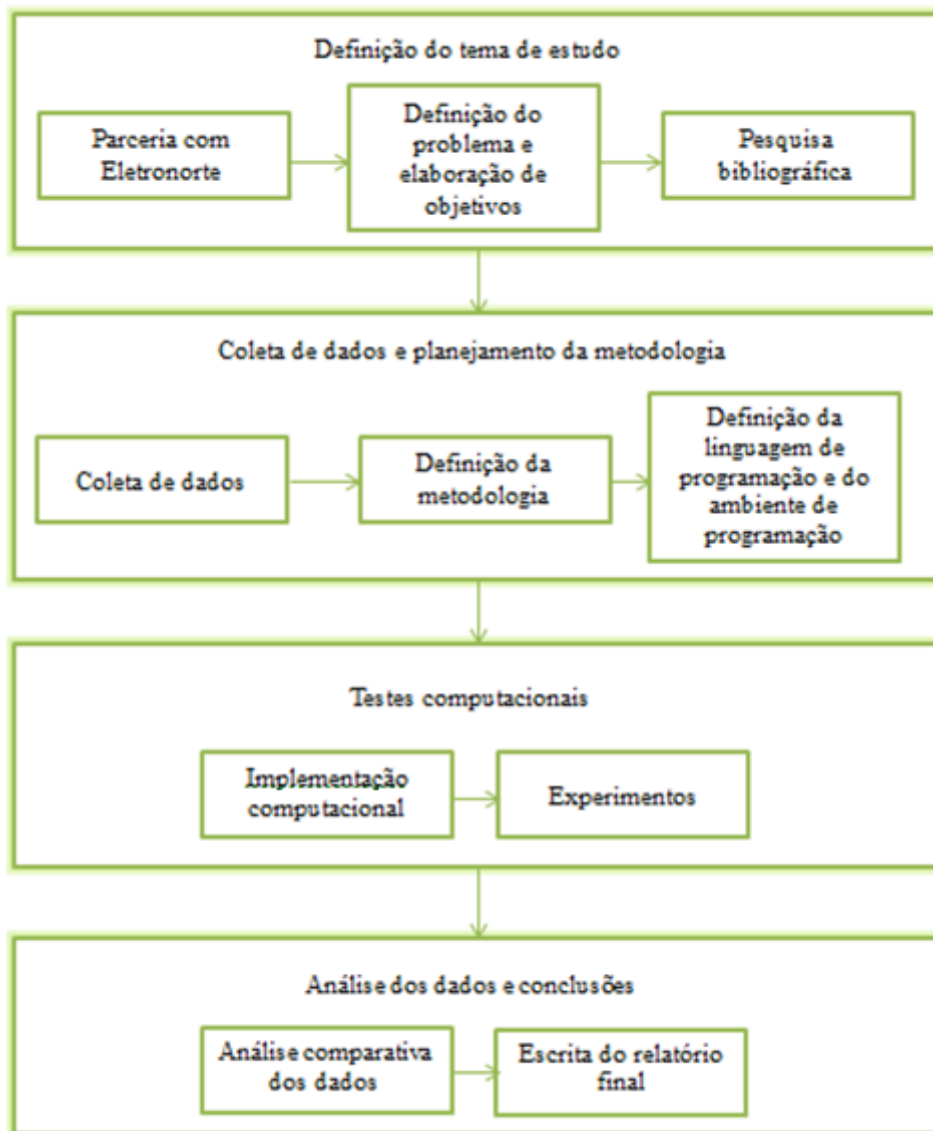
Para o desenvolvimento deste trabalho, primeiramente foi realizada uma pesquisa bibliográfica acerca do tema de DPs abordando suas causas e os diferentes tipos de padrões associados aos dados de DPs nos enrolamentos estatóricos de hidrogeradores.

Em seguida, foi realizada uma pesquisa bibliográfica sobre as principais técnicas de inteligência computacional, utilizadas para classificar padrões de DPs, e sobre as técnicas inteligentes mais atuais adotadas na classificação de outros tipos de padrões. Após isso, foi realizada uma pesquisa exploratória acerca de qual linguagem de programação e software

proporcionariam maiores recursos e flexibilidade para elaboração dos códigos necessários.

Após essas etapas, deu-se início à fase de implementação computacional, onde os dados de DPs coletados foram previamente classificados de forma manual, tratados e classificados pelas técnicas computacionais. Os resultados da acurácia das classificações realizados por cada técnica foram comparados para verificar qual das técnicas foi mais eficiente na realização de tais classificações. Para uma melhor compreensão da proposta metodológica do trabalho, a Figura 5.1 apresenta uma síntese da trajetória da pesquisa:

Figura 5.1 – Definição da trajetória da pesquisa.



Fonte: Elaborado pelo autor.

5.3 COLETA DE DADOS

As medições de DPs utilizadas nesta pesquisa foram fornecidas pelo Centro de Tecnologia da Eletronorte. Tais medições foram efetuadas nas máquinas em operação, em 23 unidades geradoras da Usina Hidrelétrica de Tucuruí (UHT), nos meses de março e setembro de 2014, totalizando 1960 medições. Cada medição de DPs foi analisada visualmente, a fim de identificar padrões que pudessem ser classificados conforme a norma (15). Foram identificados 159 medições do tipo descarga interna, 72 do tipo delaminação e 66 do tipo entre fases. Então, do conjunto de medições do tipo descarga interna e delaminação foram utilizados 66 medições de cada tipo, com o objetivo de equiparar as quantidades de cada tipo estudado, para que as técnicas inteligentes tivessem o mesmo nível de análise entre cada tipo.

Os sensores utilizados para captação dos dados de DPs foram os acopladores capacitivos, também chamados capacitores de acoplamento. Tais geradores possuem, devido às suas dimensões, uma quantidade razoável de acopladores capacitivos: alguns com 24, outros com 16 e poucos com 12 acopladores. Os acopladores capacitivos utilizados possuem 80 pF com uma resistência de 690 Ω . Esta configuração apresenta uma frequência de corte inferior aproximada de 2,88 MHz, o que significa que sinais abaixo desta frequência serão significativamente atenuados.

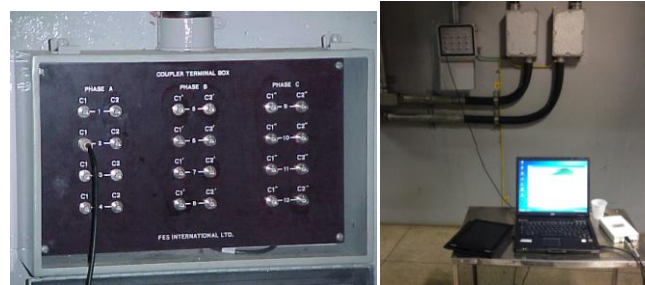
Num total de 23 máquinas a serem avaliadas, há um esforço considerável para o registro das medições, sendo necessário um mecanismo mais automatizado possível, com facilidades de navegação e rapidez para tornar o processo mais adequado no que tange a sua realização.

Para isso, foi utilizado o sistema IMA-DP Autônomo. Este foi desenvolvido em parceria entre Eletronorte e Cepel, para atender à demanda de medições na máquina em operação realizadas nos mais variados tipos de equipamentos de alta tensão. Este realiza medições esporadicamente em máquinas rotativas em operação com sensores permanentemente instalados, constituindo numa alternativa de baixo custo à substituição de sistemas de medição obsoletos adquiridos pelas empresas. Compreende a especificação de um hardware de medição compatível e de um software de processamento digital de sinais que registra os sinais de DPs medidos. Ele baseia-se na utilização de uma placa de osciloscópio de conexão USB conectado ao Notebook. Através de cabos coaxiais interligando os terminais da caixa de conexão ao osciloscópio, pode-se registrar de maneira rápida e eficiente os sinais oriundos de cada terminal.

Na Figura 5.2 estão ilustrados a caixa de acesso aos terminais dos sinais de DP, onde são realizadas as medições periódicas, e os resultados da avaliação dos níveis de DP

mostrados por ponto de medição em uma máquina.

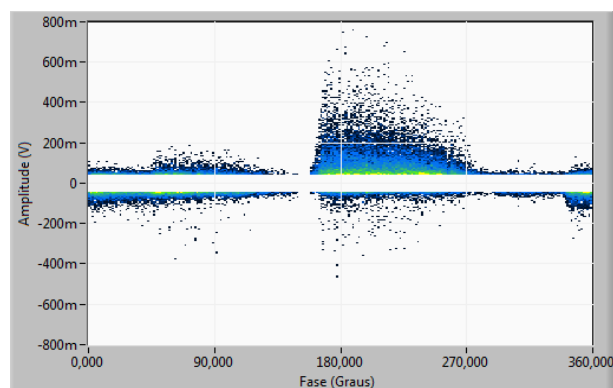
Figura 5.2 - (a) Caixa terminal com as impedâncias de medição; (b) IMA-DP autônomo.



Fonte: Elaborado pelo autor.

As informações que chegam dos sistemas de medição de DPs são tratadas pelo sistema IMA-DP, utilizando os mapas de descargas parciais associadas à fase (PRPD, do inglês *Phase Resolved Partial Discharges*) para organizar os pulsos elétricos das DPs, que serão tratados e armazenados em matrizes onde: cada linha corresponde aos valores de amplitude de tensão do pulso de DP; cada coluna, aos valores de ângulo de fase da tensão aplicada; cada elemento da matriz carrega a informação da quantidade de pulso de DP registrado com os mesmos valores de amplitude e ângulo de fase. Nesta dissertação, as medições de DPs analisadas possuem matrizes com resolução de 256 x 256. A Figura 5.3, a seguir, apresenta um exemplo das DPs associadas à fase.

Figura 5.3 - Exemplo de uma medição sob a forma de mapa PRPD.



Fonte: Elaborado pelo autor.

5.4 FERRAMENTAS COMPUTACIONAIS

Neste trabalho foram usadas diversas bibliotecas e ferramentas relacionadas à linguagem Python. Python é uma linguagem de propósito geral, que possui diversas bibliotecas e ambientes para computação científica. É uma das principais linguagens adotadas em projetos de inteligência computacional que vêm sendo utilizadas atualmente.

O ambiente de desenvolvimento utilizado foi o IPython. Este ambiente é apropriado para experimentos computacionais que utilizam a linguagem Python, pois permite a análise e visualização de dados de forma interativa, computação numérica e programação paralela.

A ferramenta se integra com bibliotecas robustas de computação científica, como NumPy, SciPy, pandas, matplotlib, etc. NumPy fornece um objeto de matriz multidimensional que suporta operações vetoriais otimizadas. SciPy oferece uma ampla variedade de algoritmos, de processamento de sinais, otimização, entre outros, que são baseados em NumPy. Pandas propõe estruturas de dados convenientes para tabulação de dados reais. Matplotlib permite traçar gráficos facilmente para a visualização interativa de vários tipos de dados. IPython é um ambiente adequado para usar todas essas ferramentas de uma forma simplificada (46).

A biblioteca Scikit-Learn, integrada ao IPython, foi utilizada para realizar os experimentos de aprendizado supervisionado. Scikit-Learn é uma biblioteca em Python com diversas implementações de algoritmos de inteligência computacional, possui uma sintaxe simples e muito útil para análise de dados, além de ser *open-source* e multiplataforma (47).

5.5 TRATAMENTO DOS DADOS

As 1960 medições de DPs utilizados neste trabalho passaram, primeiramente, por uma análise visual para identificar quais apresentavam padrões que poderiam ser classificados conforme a norma (15), para que estes fossem utilizados para treinar e testar as técnicas inteligentes que foram aplicadas. Através dos mapas PRPD foram identificados 159 medições apresentando padrões do tipo descarga interna, 72 do tipo delaminação e 66 do tipo entre fases. Então, do conjunto de dados do tipo descarga interna e delaminação foram utilizados 66 medições de cada tipo, para equiparar as quantidades de cada tipo, para que as técnicas inteligentes tivessem o mesmo nível de análise entre as classes.

O procedimento computacional para tratamento dos dados e classificação pode ser

orientado pelas etapas abaixo:

- Repetir para cada tipo de DP:
 - ↳ Inicializar lista de matrizes vazia;
 - ↳ Inicializar lista de vetores vazia;
 - ↳ Inicializar lista de tipos de DPs vazia;
 - ↳ Repetir para cada arquivo:
 - ✦ Carregar arquivo .dp;
 - ✦ Extrair lista com valores de tensão;
 - ✦ Extrair lista com valores de fase;
 - ✦ Extrair lista com número de ocorrências;
 - ✦ Gerar matriz de ocorrências, onde as linhas são as amplitudes e as colunas são as fases;
 - ✦ Binarizar os valores da matriz;
 - ✦ Adicionar matriz de ocorrências na lista de matrizes;
 - ✦ Converter matriz em vetor;
 - ✦ Adicionar vetor na lista de vetores;
 - ✦ Adicionar tipo de DP na lista de tipos de DPs.
 - ↳ Separar lista de vetores em lista de treino e lista de teste;
 - ↳ Separar lista de tipos DPs em lista de treino e lista de teste.
- Unificar as listas de treino de todos os tipos;
- Unificar as listas de teste de todos os tipos;
- Realizar o treinamento usando a lista de treino com todos os tipos;
- Realizar a predição usando a lista de teste com todos os tipos.

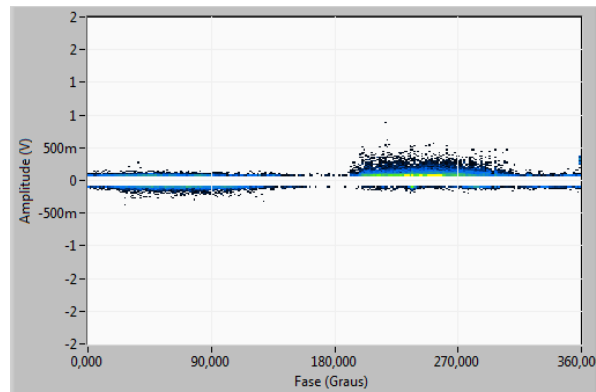
Os arquivos de DPs gerados pelo IMA-DP, no formato “.dp”, foram transformados em matrizes de dados, onde as informações de fase, tensão e número de ocorrência, representam as posições das colunas, linhas e número de ocorrência de cada DP de acordo com seu valor de tensão e fase, respectivamente.

As matrizes foram geradas com ordem de 256 x 256 elementos. Os valores dos números de ocorrência foram binarizados, onde o limiar utilizado foi à média das frequências, com isso, eliminando-se as descargas de baixa ocorrência. Na Figura 5.4 abaixo está ilustrado um sinal de DPs do tipo descarga interna, com visualização no IMA-DP. Na Figura 5.5 a) e Figura 5.5 b) , o mesmo sinal de DP é ilustrado sem qualquer tratamento e com a binarização, com visualização no IPython utilizando a biblioteca matplotlib. Por fim, os dados binarizados na

estrutura matricial foram transformados em vetores unidimensionais, ou seja, as matrizes de ordem 256 x 256 foram redimensionadas para vetores de ordem 1 x 65536.

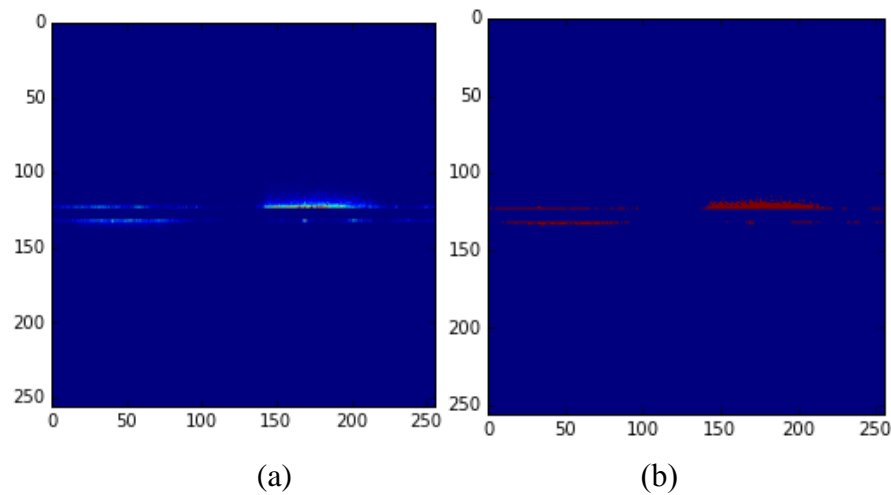
O pré-processamento contribuiu com a preparação dos dados para serem treinados e caracterizados com as técnicas de inteligência computacional.

Figura 5.4 - Sinal de descarga interna com visualização no IMA-DP.



Fonte: Elaborado pelo autor.

Figura 5.5 - Sinal de descarga interna com visualização no IPython utilizando a biblioteca matplotlib. a) Sinal sem tratamento; b) Sinal binarizado.



Fonte: Elaborado pelo autor.

5.6 RESULTADOS DE EXPERIMENTOS

Para realizar a caracterização de descargas parciais foi necessário avaliar a acurácia

das técnicas de inteligência computacional selecionadas. O método experimental foi escolhido para identificar a técnica que melhor caracterizasse os tipos de descargas parciais corretamente, tal como utilizado em (2).

Dos dados binarizados, conforme exposto no item 5.5 do capítulo anterior, foram selecionados para o experimento 66 sinais de cada tipo de DP (cavidade interna, delaminação e entre barras). O conjunto com 198 sinais foram submetidos ao experimento de caracterização com os algoritmos MVS, Floresta Randômica e KNN.

Considerando a quantidade reduzida de dados usados no experimento, optou-se em utilizar 75% dos dados para treinamento dos algoritmos de aprendizagem e 25% dos dados para teste. O conjunto de dados no teste foi o mesmo usado para validação.

5.6.1 Configurações no algoritmo de KNN

A técnica KNN possui apenas uma configuração a ser escolhida pelo usuário, que constitui-se no número de vizinhos, a ser considerado, em relação ao elemento a ser rotulado. Foram realizados quatro ensaios com o número n de vizinhos iguais a 1, 3, 5 e 7. Tal como exposto anteriormente é aconselhável que o número de vizinhos seja ímpar para evitar casos de empate durante a classificação.

Na Tabela 6.1 a seguir são expostos os valores de acurácia para cada número de vizinhos utilizados durante o experimento. Observa-se que o número 1 de vizinho foi a configuração que obteve melhor resultado. Isto indica que os dados os quais foram previstos obtiveram maior acurácia quando os estes foram rotulados com a mesma classe do dado de maior similaridade, ou seja, com o vizinho imediatamente mais próximo que continha características mais semelhantes. Para este caso, quanto maior a quantidade de vizinhos, maior a chance de rotular o dado com uma classe indevida.

Tabela 5.1 - Acurácias para cada número de vizinho.

Número de vizinhos	Acurácia
1	82,35 %
3	76,47 %
5	72,54 %
7	70,58%

5.6.2 Configurações no algoritmo de Floresta Randômica

As principais configurações da técnica Floresta Randômica são o número total de

árvores a serem construídas e a quantidade de atributos usados para cada nó.

A biblioteca Scikit-Learn define a raiz quadrada do número de atributos como o valor padrão para a quantidade atributos por nó. Para este experimento, cada elemento possuía 65536 atributos, portanto, esta configuração foi definida com o valor de 256.

O número total de árvores assumiu valores diferentes para cada ensaio. No experimento a configuração foi variada em 10, 50 e 100 árvores. Os valores para o mesmo número de árvores variaram durante os ensaios, devido à característica randômica da técnica. A melhor acurácia foi obtida com o número de 100 árvores, conforme pode ser visto na Tabela 6.2. Os ensaios com números superiores a 100 árvores retornaram no máximo a mesma acurácia que 100 árvores, com um custo computacional muito maior durante a execução.

Tabela 5.2 - Acurácias para cada número de árvores.

Número de árvores	Acurácia
10	88,23% a 86,27%
50	90,19 % a 92,15%
100	94,11 %

5.6.3 Configurações no algoritmo de MVS

O aprendizado supervisionado com MVS possui diversas funções *kernels*. Existem problemas que exigem o uso das funções *kernel* lineares ou não-lineares.

A biblioteca Scikit-Learn possui as implementações do MVS com kernel linear ou com funções polinomiais e gaussianas, as quais são não-lineares. O uso do MVS com os kernels não-lineares não atingiram resultados tão satisfatórios quanto o linear. Os resultados com a acurácia dos kernels linear, polinomial e gaussiano, são apresentados na Tabela 6.3:

Tabela 5.3 - Acurácias para cada kernel.

Kernel	Acurácia
Linear	96,07 %
Polinomial	78,43 %
Gaussiano	88,23 %

Conforme exposto no item 2.7 do Capítulo 2, os sinais de DPs são acompanhados de sinais ruidosos que se misturam aos sinais de DPs, podendo inviabilizar a leitura destes dados.

Com o objetivo de melhorar o resultado da classificação dos sinais de DPs, realizados pelo MVS Linear, os sinais foram, antes de serem classificados, submetidos às seguintes técnicas de pré-processamento: ANOVA, ACP e ACI, descritas na presente dissertação.

A binarização, explanada no item 5.5 deste trabalho, não foi aplicada nos dados utilizados para os experimentos com as técnicas ANOVA, ACP e ACI, pois o uso da binarização combinado às técnicas citadas para diminuir os sinais ruidosos de cada medição, diminui a eficiência das técnicas classificadoras por extrair uma quantidade significativa de sinais que, além de não constituírem interferências, poderiam ser importantes no momento da classificação, por localizarem-se, por exemplo, em uma região de fronteira ou possuir valores de tensão característicos de um determinado tipo de DP.

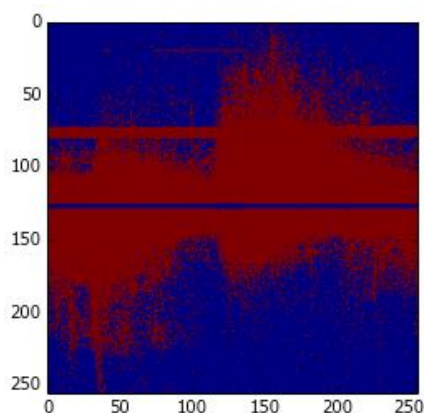
5.6.4 Configurações no algoritmo de ANOVA aplicado às técnicas classificadoras

O uso do método estatístico ANOVA possibilitou a seleção dos principais atributos dos sinais. Os sinais coletados de DPs foram mapeados em matrizes de dimensão 256x256. A técnica identificou os pontos mais relevantes para a caracterização.

As principais características foram selecionadas através de um algoritmo com a biblioteca Scikit-Learn no ambiente IPython. Foram definidos 30000 dos 65536 componentes das matrizes, ou seja, foram utilizados menos que a metade do total de atributos.

A Figura 6.1 representa a matriz de DPs, onde a área vermelha corresponde aos atributos selecionados pela técnica de seleção de características com ANOVA, e a parte azul são os atributos não selecionados.

Figura 5.6 - 30000 atributos selecionados por ANOVA.



Fonte: Elaborado pelo autor.

Em cada medição de DPs foram reduzidos alguns dados pela técnica e, em seguida, foram submetidos às etapas de treino e teste com MVS de *kernel* linear, KNN para K=1 vizinho mais próximo e Floresta Randômica.

Os testes realizados pelas técnicas MVS linear com a ANOVA e Floresta Randômica com ANOVA apresentaram resultados próximos, no entanto, inferior ao MVS linear. Com o uso de tais configurações propostas neste trabalho, os ensaios obtiveram a acurácia de 94,11 %. No ensaio realizado com a técnica KNN com ANOVA, o resultado atingido foi de 86,27%.

5.6.5 Configurações no algoritmo de FastICA aplicado às técnicas classificadoras

A técnica FastICA é comumente usada para separação cega em sinais ruidosos. Os dados de DPs foram submetidos à técnica, na tentativa de redução de ruídos, com o objetivo de melhorar os resultados na caracterização das descargas parciais.

Foram definidos 198 componentes como configuração do FastICA, ou seja, o mesmo número de amostras de DPs utilizadas no experimento. O algoritmo extrai as informações da fonte ruidosa, comum a todos os sinais, em seguida, aplica a transformação de matrizes para separar os ruídos dos dados submetidos à técnica.

As técnicas classificadoras MVS linear, KNN para K=1 vizinho e Floresta Randômica foram aplicada nos dados, após o pré-processamento com o FastICA, para avaliar a caracterização com o uso das técnicas combinadas, apresentando como resultados acurácias de 23,52 %, 23,52% e 25,49%, respectivamente. Os devidos valores representaram o pior resultado de todos os experimentos realizados durante a pesquisa.

5.6.6 Configurações no algoritmo de ACP aplicado às técnicas classificadoras

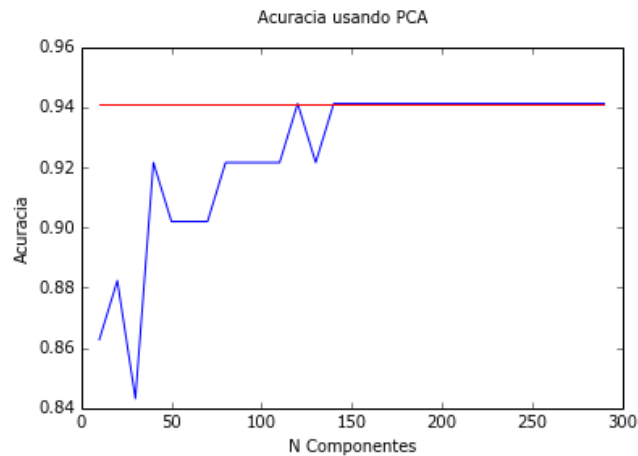
A técnica ACP é largamente utilizada na redução de dimensionalidade de dados multidimensionais, pois esta se propõe a procurar padrões existentes entre os sinais. Isto proporcionou a redução de dimensões da matriz de 256 x 256 dos sinais de DPs, facilitando a análise dos dados durante a sua classificação, pois além de diminuir o esforço computacional realizado pelas técnicas classificadoras, proporciona uma redução de sinais que podem ser ruidosos, considerando que a técnica encontra padrões.

O número de componentes é uma das principais configurações de ACP, onde a escolha pode ser feita considerando diversos aspectos, como esforço computacional ou acurácia. Para

selecionar o número de componentes adequado para o problema foram realizados diversos ensaios.

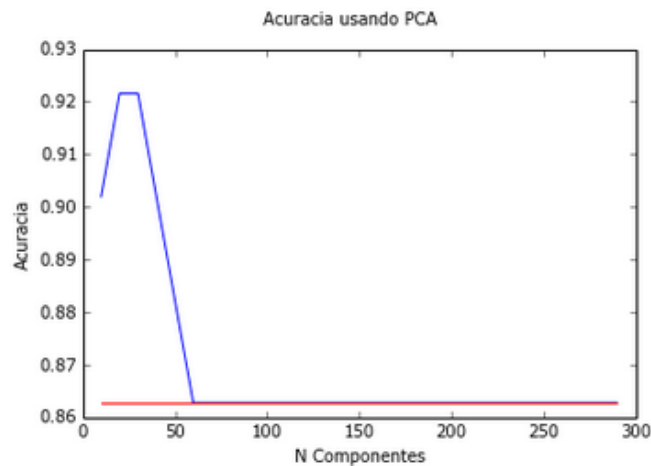
Nos ensaios foi avaliada a acurácia da utilização de cada uma das técnicas classificadoras com diferentes números de componentes do ACP. O número de componentes variou de 10 a 300, incrementando de 10 em 10. Tal como pode ser visto nas Figura 5.7, Figura 5.8 e Figura 5.9:

Figura 5.7 - Acurácia usando ACP para cada n componentes para a técnica MVS.



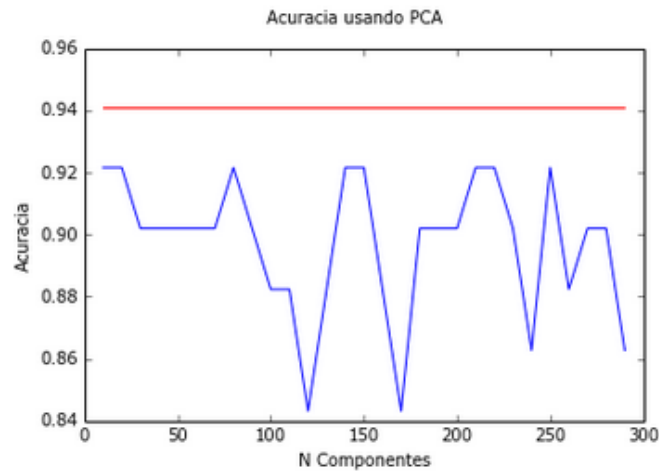
Fonte: Elaborado pelo autor.

Figura 5.8 - Acurácia usando ACP para cada n componentes para a técnica KNN.



Fonte: Elaborado pelo autor.

Figura 5.9 - Acurácia usando ACP para cada n componentes para a técnica Floresta Randômica.



Fonte: Elaborado pelo autor.

O menor número de componentes que obteve o melhor resultado foi o de 120 componentes para MVS, 20 componentes para KNN e 10 componentes para Floresta Randômica. A melhor acurácia do uso de MVS, KNN e Floresta Randômica após o tratamento de ACP foi de 94,11%, 92,15% e 92,15%, respectivamente.

5.7 RESULTADOS COMPLEMENTARES

Na Tabela 5.4 a seguir constam os resultados dos experimentos acima relatados. Observa-se que a técnica MVS foi a que obteve melhor resultado.

Tabela 5.4 - Acurácias para cada experimento.

	Acurácia
MVS	96,07%
KNN	82,35%
Floresta Randômica	94,11%
MVS com ANOVA	94,11%
MVS com FastICA	23,52%
MVS com ACP	94,11%
KNN com ANOVA	86,27%
KNN com FastICA	23,52%
KNN com ACP	92,15%
Floresta Randômica com ANOVA	94,11%
Floresta Randômica com FastICA	25,49%
Floresta Randômica com ACP	92,15%

Os resultados das classificações realizadas com os dados tratados pelas técnicas ANOVA e ACP, foram ligeiramente inferiores ao resultado da classificação de MVS sem tratamento, pois alguns sinais que foram retirados não constituíam propriamente ruídos. Apesar de aparentarem menos significância, estes sinais eram significativos para a classificação dos padrões. Isto pode ser verificado na pequena diferença de percentual na classificação com MVS sem o tratamento de ANOVA e ACP, de 96,07%, e com o tratamento destas, de 94,11% e 92,15%, com exceção de KNN com ANOVA, cujo resultado de 86,27% representa uma significativa diferença.

O tratamento das técnicas ANOVA e ACP com MVS, KNN e Floresta Randômica reduziu a matriz dos sinais de DPs de 65536 componentes para 30000, 120, 20 e 10 componentes, respectivamente. Essa redução de dimensionalidade, com pouca perda de acurácia, constitui-se numa vantagem, pois o esforço computacional realizado pelas técnicas classificadoras é consideravelmente reduzido, aumentando a velocidade na classificação. Isto

é um fator importante para os casos de classificações realizadas com a máquina em operação.

ACI foi a única técnica de tratamento de dados que atingiu resultado muito abaixo do estimado. O motivo para este fato atribui-se aos dados não terem a mesma fonte de descargas, ocorrendo falhas no processo de recuperação de sinais. Ou seja, para este caso, esta técnica não foi eficiente pelas medições terem sido realizadas em 23 máquinas.

Para analisar o desempenho das técnicas classificadoras foi utilizado o método da matriz de confusão. A matriz de confusão proporciona a visualização de desempenho de uma técnica classificadora onde, através do seu exame, pode-se avaliar quantitativamente as classes que o algoritmo classificador tem maior dificuldade para classificar.

Na matriz é ilustrado o número de predições corretas e incorretas para cada classe, e com quais classes houve erro durante a predição. Em um conjunto de dados, cada elemento m_{ij} da matriz de confusão possui i número de classes associados a uma classe j . Para k classes, a matriz possui dimensão $k \times k$. As linhas da matriz correspondem às classes originais e, as colunas, as classes preditas. A diagonal apresenta os elementos classificados corretamente, enquanto que, os demais elementos da matriz apresentam a quantidade de erros cometidos de acordo com a classe original e a predição realizada (42).

Assim, foi obtida a matriz de confusão da classificação realizada para cada experimento da Tabela 5.4, conforme as Tabelas 5.5 à Tabela 5.16, mostradas a seguir:

Tabela 5.5 - Matriz de confusão para predição com MVS.

		Classes Preditas		
		Cavidade Interna	Delaminação	Entre Barras
Classes Originais	Cavidade Interna	15	0	2
	Delaminação	0	17	0
	Entre Barras	0	0	17

Tabela 5.6 - Matriz de confusão para predição com Floresta Randômica.

		Classes Preditas		
		Cavidade Interna	Delaminação	Entre Barras
Classes Originais	Cavidade Interna	14	0	3
	Delaminação	0	17	0
	Entre Barras	0	0	17

Tabela 5.7 - Matriz de confusão para predição de KNN.

		Classes Preditas		
		Cavidade Interna	Delaminação	Entre Barras
Classes Originais	Cavidade Interna	11	0	6
	Delaminação	3	14	0
	Entre Barras	0	0	17

Tabela 5.8 - Matriz de confusão para predição de MVS com PCA.

		Classes Preditas		
		Cavidade Interna	Delaminação	Entre Barras
Classes Originais	Cavidade Interna	15	0	2
	Delaminação	0	17	0
	Entre Barras	1	0	16

Tabela 5.9 - Matriz de confusão para predição de MVS com ANOVA.

		Classes Preditas		
		Cavidade Interna	Delaminação	Entre Barras
Classes Originais	Cavidade Interna	14	0	3
	Delaminação	0	17	0
	Entre Barras	0	0	17

Tabela 5.10 - Matriz de confusão para predição de MVS com ACI.

		Classes Preditas		
		Cavidade Interna	Delaminação	Entre Barras
Classes Originais	Cavidade Interna	6	4	7
	Delaminação	7	0	10
	Entre Barras	10	1	6

Tabela 5.11 - Matriz de confusão para predição de KNN com PCA.

		Classes Preditas		
		Cavidade Interna	Delaminação	Entre Barras
Classes Originais	Cavidade Interna	13	0	4
	Delaminação	0	17	0
	Entre Barras	0	0	17

Tabela 5.12 - Matriz de confusão para predição de KNN com ANOVA.

		Classes Preditas		
		Cavidade Interna	Delaminação	Entre Barras
Classes Originais	Cavidade Interna	11	0	6
	Delaminação	1	16	0
	Entre Barras	0	0	17

Tabela 5.13 - Matriz de confusão para predição de KNN com ACI.

		Classes Preditas		
		Cavidade Interna	Delaminação	Entre Barras
Classes Originais	Cavidade Interna	0	17	0
	Delaminação	2	10	5
	Entre Barras	0	15	2

Tabela 5.14 - Matriz de confusão para predição de Floresta Randômica com PCA.

		Classes Preditas		
		Cavidade Interna	Delaminação	Entre Barras
Classes Originais	Cavidade Interna	14	0	3
	Delaminação	0	17	0
	Entre Barras	1	0	16

Tabela 5.15 - Matriz de confusão para predição de Floresta Randômica com ANOVA.

		Classes Preditas		
		Cavidade Interna	Delaminação	Entre Barras
Classes Originais	Cavidade Interna	14	0	3
	Delaminação	0	17	0
	Entre Barras	0	0	17

Tabela 5.16 - Matriz de confusão para predição de Floresta Randômica com ACI.

		Classes Preditas		
		Cavidade Interna	Delaminação	Entre Barras
Classes Originais	Cavidade Interna	6	6	5
	Delaminação	6	4	7
	Entre Barras	6	8	3

Conforme observado nas tabelas anteriores, a classificação com MVS foi a que menos apresentou erro, apenas na classificação de dois sinais de DPs do tipo cavidade interna que foram classificados como do tipo entre barras. Isto ocorreu em função da grande capacidade de generalização desta técnica e também da mesma possuir mecanismos de seleção das principais variáveis durante o processo.

5.8 CONCLUSÕES DO CAPÍTULO

Foram encontrados três padrões de DPs nas medições coletadas na máquina em operação de UHT: DPs internas, delaminação e entre barras. O sistema IMA-DP Autônomo possibilitou que esses dados fossem medidos e registrados de maneira rápida e eficiente, em formato de matrizes, sendo organizados através dos mapas PRPD. As matrizes possuem

resolução de 256 x 256, sendo que as linhas correspondem aos valores de tensão, as colunas, aos valores de ângulo de fase, e cada elemento da matriz representa a quantidade de pulso de DP que foi registrado com os valores de tensão e ângulo de fases especificadas. Os mapas PRPD foram importantes por possibilitar a visualização dos padrões formados pelos pulsos de descargas e sua classificação.

A linguagem Python e o software IPython, direcionado para computação científica, proporcionaram grande flexibilidade, facilidade e clareza durante a elaboração do código que atendeu aos propósitos desta dissertação. As bibliotecas fornecidas em Python oferecem variadas ferramentas para aplicações diversas, com a vantagem de possuir caráter *open-source*.

O processo de binarização utilizado conferiu aos dados de DPs uma maior limpeza com a eliminação de descargas de baixa ocorrência, em relação à média da quantidade de pulsos de DP. Esse procedimento facilitou a classificação dos dados de DPs pelas técnicas de classificação.

As técnicas de classificação utilizadas neste trabalho apresentaram, de modo geral, bons resultados com os experimentos realizados, com destaque para MVS que apresentou os melhores resultados para esta aplicação devido aos motivos expostos anteriormente, lembrando que esta técnica trabalha bem com problemas de grandes dimensões.

As classificações realizadas após os dados terem sido submetidos às técnicas de tratamento dos dados, ANOVA e ACP, embora não tenham atingido o resultado da classificação de MVS sem tratamento, apresentaram resultados bem próximos a este, com exceção da classificação realizada por KNN com ANOVA.

ACI foi a única técnica de tratamento de dados que não atingiu o resultado esperado. Este fato atribui-se aos dados não terem a mesma fonte de descargas. Com isso, ocorreram falhas no processo de recuperação de sinais. Em vez de a técnica extrair os ruídos, houve extração de sinais relevantes para a caracterização.

6.0 CONCLUSÕES

6.1 CONSIDERAÇÕES FINAIS

Esta dissertação de mestrado teve como objetivo a caracterização automática de padrões de DPs utilizando técnicas de inteligência computacional. Os dados foram provenientes da UHT, coletados no ano de 2014, totalizando 1960 medições de DPs. Destas foram identificados visualmente 159 do tipo descarga interna, 72 dados do tipo delaminação e 66 dados do tipo entre fases. Então, do conjunto de dados do tipo descarga interna e delaminação foram utilizados 66 dados de cada, com o objetivo de equiparar as quantidades de cada tipo estudado, para que as técnicas inteligentes tivessem o mesmo nível de análise entre cada um.

As técnicas inteligentes utilizadas neste trabalho foram escolhidas por serem de uso relativamente recente no meio acadêmico/científico e, por apresentarem bons resultados em pesquisas de diversas áreas de conhecimento.

Das técnicas utilizadas para classificar os padrões de DPs, MVS foi a que obteve maior percentual de acerto (96,07%).

Os sinais de DPs coletados contêm ruídos – oriundos dos próprios sistemas de medições de DPs, de atividades térmicas ou de chaveamentos, etc. - que se somam às descargas que são medidas e comprometem a avaliação do estado operacional da máquina. Neste sentido, foram utilizadas as técnicas ANOVA, ACI E ACP com o objetivo de diminuir a quantidade de sinais que não fossem significativos para a classificação dos padrões. Os sinais de DPs foram submetidos a cada uma destas ferramentas, sendo logo após, submetidos também à classificação pelo MVS, técnica que se mostrou mais eficiente para este caso. No entanto, os resultados obtidos foram inferiores à classificação realizada pelo MVS sem o tratamento prévio destas técnicas.

A ANOVA atuou nos dados analisando todos os sinais de DPs submetidos, comparando-os e subtraindo as variáveis menos significativas. Uma explicação para a diminuição da acurácia obtida após o tratamento da ANOVA é que esta pode ter retirado alguns sinais que não fossem propriamente ruídos, apesar de aparentarem menos significância, eram sinais que faziam a diferença para a classificação dos padrões. Isto pode ser verificado na pequena diferença de percentual nas classificações com MVS sem o tratamento da ANOVA, de 96,07%, e com o tratamento desta juntamente às técnicas classificadoras, de 94,11%, excetuando-se o tratamento a classificação de KNN após o tratamento de ANOVA.

A utilização da ACP teve o mesmo intuito da técnica acima citada. No entanto, atuou reduzindo a dimensionalidade da matriz dos sinais de DPs. Após este tratamento, a classificação realizada pelo MVS teve resultado de 94,11%, usando apenas 120 componentes, e a classificação realizada por KNN e Floresta Randômica teve resultado de 92,15%, utilizando 20 e 10 componentes, respectivamente. Ou seja, analisando um número muito menor de dados em cada medição, foi possível atingir uma acurácia muito próxima ao melhor resultado obtido.

O tratamento realizado por FastICA teve como objetivo analisar todos os sinais de DPs e separar os sinais que eram ruídos dos sinais que eram descargas. Após o tratamento desta técnica, a classificação realizada pelas técnicas obtiveram resultados em torno de 23% à 25%, representando o pior resultado com uma significativa diferença em relação às outras técnicas utilizadas para tratar os sinais com ruídos. Isto se deve ao fato que, para separar os sinais de descarga dos sinais ruidosos seria necessário que todas as medições tivessem sido realizadas na mesma máquina para garantir que a fonte das descargas fosse a mesma.

Assim, a técnica MVS apresentou o melhor resultado para a classificação dos três tipos de descargas utilizados, mesmo sem a utilização prévia das técnicas ANOVA, ACP e ACI. Isto se deve ao fato que a técnica possui uma grande capacidade de generalização e também mecanismos de seleção das principais variáveis durante o treinamento. O MVS possui a característica de apresentar bons resultados em problemas de classificação de dados de alta dimensionalidade.

O tratamento das técnicas ANOVA e ACP com MVS, KNN e Floresta Randômica reduziu a matriz dos sinais de DPs de 65536 componentes para 30000, 120, 20 e 10 componentes, respectivamente.

Apesar dos resultados com o tratamento das técnicas ANOVA e ACP com MVS, KNN e Floresta Randômica terem sido um pouco mais baixos, vale considerar que essas técnicas reduziram a matriz dos sinais de DPs de 65536 componentes para 30000, 120, 20 e 10 componentes, respectivamente. Isto é importante considerar no caso de identificações realizadas na máquina em operação, pois o esforço computacional realizado pelas técnicas classificadoras é reduzido, propiciando maior velocidade na classificação e mantendo o satisfatório resultado próximo aos 96,07%.

A binarização foi aplicada nos sinais de DPs que foram diretamente classificados pelas técnicas MVS Linear, Floresta Randômica e KNN, para eliminar as descargas de baixa ocorrência facilitando a classificação das medições de DPs. Em contrapartida, a binarização não foi utilizada quando os sinais de DPs foram anteriormente tratados pelas técnicas de pré-

processamento, pois sua aplicação diminuiria a eficiência das técnicas classificadoras por extrair uma quantidade significativa de sinais que poderiam ser importantes durante a classificação.

Ressalta-se ainda como resultados da presente dissertação a produção de um artigo, na área da presente dissertação para apresentação oral em congresso, no XXIII SNPTEE, Seminário Nacional de Produção e Transmissão de Energia Elétrica, a ser realizado este ano de 2015 em Foz do Iguaçu, artigo este desenvolvido em colaboração com a ELETRONORTE-ELETROBRÁS.

6.2 SUGESTÕES PARA TRABALHOS FUTUROS

Para trabalhos futuros sugere-se realizar as classificações de DPs com um conjunto maior de medições que contenham sinais de DPs dos outros tipos de padrões que não foram tratados neste trabalho.

Outras técnicas de inteligência computacional e processamento de imagem podem ser utilizadas para efeitos de comparação com os resultados aqui apresentados.

Sugere-se também, após a classificação de todos os padrões de DPs, a elaboração de um sistema inteligente que possa indicar a localização da ocorrência da descarga, a fim de que as manutenções possam ser cada vez mais controladas e melhor planejadas, com consequentes menores custos.

Além disso, podem ser estudadas formas de avaliar os impactos das DPs utilizando outros métodos de medição e outras variáveis do sistema, trabalhando-se, por exemplo, com medidas indiretas como tensão e corrente.

REFERÊNCIAS BIBLIOGRÁFICAS

- (1) CUENCA, W. M. H. **Caracterização dos Sinais de Descargas Parciais em Equipamentos de Alta Tensão a Partir de Modelos Experimentais**. Tese de Doutorado, Programa de Pós-Graduação de Engenharias, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil: 2005.
- (2) ALVES, M. NUNES, M. V. A. BRASIL, F. CARVALHO, D. CONCEIÇÃO, F. VILHENA, D. **Caracterização de Padrões de Descargas Parciais em Hidrogeradores Utilizando Técnicas de Inteligência Computacional**. Seminário Nacional de Produção e Transmissão de Energia Elétrica. Foz do Iguaçu, Brasil: 2015.
- (3) HUDON, C., BÉLEC, M. **Partial Discharges Signal Interpretation for Generator Diagnostics**. IEEE Transaction on Dielectrics and Electrical Insulation. No 2, Vol. 12, 2005.
- (4) STONE, G. C., WARREN, V. **Objective Methods to Interpret Partial-Discharge Data on Rotating Machine Stator Windings**. IEEE Transactions on Industry Applications. No 1, Vol. 42, 2006.
- (5) FAIER, J. M. **Curvas Principais Aplicadas na Identificação de Descargas Parciais em Equipamentos de Potência**. Dissertação de Mestrado, Programa Pós-Graduação de Engenharias, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil: 2006.
- (6) KRIVDA, A. **Automated Recognition of Partial Discharges**. IEEE Transactions on Dielectrics and Electrical Insulation. No 5, Vol. 2, 1995.
- (7) DINIZ, F. C. da. C. B. **Supressão de Ruído, Detecção e Classificação de Sinais de Descargas Parciais em Transformadores de Potência**. Tese de Doutorado, Programa Pós-Graduação de Engenharias, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil : 2005.
- (8) SRIVASTAVA, D. K., BHAMBHU, L. **Data Classification Using Support Vector**. **Journal of Theoretical and Applied Information Technology**. No 1, Vol. 12. Bahal, India, 2009.
- (9) LUCAS, L. C. de S. **Árvores, Florestas e Sua Função Como Preditores: Uma Aplicação na Avaliação do Grau de Maturidade de Empresas**. Revista Brasileira de Pesquisas de Marketing, Opinião e Mídia. ISSN 1983-9456. No 6. 2011.
- (10) SANTOS, F. C. **Variações do Método kNN e suas Aplicações na Classificação Automática de Textos**. Dissertação de Mestrado, Programa de Pós-Graduação do Instituto de Informática da Universidade Federal de Goiás, Universidade Federal de Goiás, Goiânia, Brasil: 2009.
- (11) LORENA, A. C. JACINTHO, L. F. O. **Comparing Machine Learning Classifiers in Potential Distribution Modelling**. Expert Systems with Applications. No 5, Vol. 38, 2010, p 5268-5275.
- (12) INSTITUTE OF ELECTRICAL AND ELECTRONIC ENGINEERS. **IEEE P1434 Guide for the Measurement of Partial Discharge in AC Electric Machinery**. USA: 2014.

- (13) INTERNACIONAL ELETROTECHNICAL COMMISSION. **IEC 60270 High Voltage Test Techniques – Partial Discharges Measurements**. 2000.
- (14) CONSEIL INTERNATIONAL DES GRANDS RÉSEAUX ÉLECTRIQUES. **CIGRÉ 392 Survey of Hydrogenerator Failures**. Working Group A1.10. 2009.
- (15) INTERNACIONAL ELETROTECHNICAL COMMISSION. **IEC 60034-27-2 Rotating Electrical Machines – Part 27-2: On-line Partial Discharges Measurements on the Stator Winding Insulation of Rotating Electrical Machines**. Suíça: 2012.
- (16) FILHO, V. S. **Imagens de Defeitos por Descarga Parcial Estimulada por Raio X Pulsado em Materiais Dielétricos Poliméricos**. Tese de Doutorado, Programa Interdisciplinar de Pós-Graduação em Engenharia, Universidade Federal do Paraná, Curitiba, Brasil : 2000.
- (17) CARVALHO, A. T. de. **Seleção de Bases Wavelet para Redução do Ruído de Banda Larga em Medições de Descargas Parciais em Máquinas Síncronas**. Tese de Doutorado, Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal do Rio de Janeiro. Rio de Janeiro, Brasil: 2014.
- (18) BRASIL, F. de S. **Estudo Experimental de Técnicas de Medição de Descargas Parciais nos Enrolamentos Estatóricos de Hidrogeradores**. Dissertação de Mestrado, Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal do Pará, Belém, Pará, Brasil: 2013.
- (19) NIASAR, M. G. **Partial Discharge Signatures of Defects in Insulation Systems Consisting of Oil and Oil-impregnated Paper**. Tese de Doutorado, Division of Electromagnetic Engineering, KTH School of Electrical Engineering. Estocolmo, Suécia: 2012.
- (20) SOUZA, A. L. **Estudo de Modelos para Definição de um Critério de Início de Corona Através da Análise de Experimentos e Simulações Computacionais**. Dissertação de Mestrado. Programa de Pós-Graduação em Engenharia Elétrica da Escola de Engenharia da Universidade Federal de Minas Gerais. Universidade Federal de Minas Gerais. Belo Horizonte, Minas Gerais: 2009.
- (21) TAYLOR, N. **Dielectric Response and Partial Discharge Measurements on Stator Insulation at Varied Low Frequency**. Tese de Doutorado, School of Electrical Engineering. Suécia: 2010.
- (22) AZEVEDO, C. H. Universidade Federal de Goiás, Goiânia, Brasil: 2009.. **B. Metodologia para a eficácia da detecção de descargas parciais por emissão acústica como técnica preditiva de manutenção em transformadores de potência imersos em óleo isolante**. Dissertação de Mestrado. Programa de Pós-Graduação da Escola de Engenharia Elétrica e de Computação da Universidade Federal de Goiás.
- (23) PANZERA, A. C., GOMES A. E. Q., MOURA, D. G. **Impactos Ambientais da Produção de Energia Elétrica**. Minas Gerais, Brasil : Centro de Referência Virtual do Professor, SEE, 2010.

- (24) SIMONE, Gilio Aluisio. **Centrais e aproveitamentos hidrelétricos : Teoria e exercícios.** São Paulo: Érica, 2000.
- (25) GURU, B. S., HIZIROĞLU, H. R. **Electric machinery and transformers.** New York: Oxford, 3 ed. 2001.
- (26) UEMORI, M., ROCHA, J. **Fundamentos de Projetos e Reabilitação de Geradores para Centrais Hidrelétricas, Princípios de Manutenção e Reabilitação de Geradores Síncronos.** Brasil: 2012.
- (27) IRIS POWER ENGINEERING. **PD Seminar – Volume 1.** version 4.2. Toronto, Canada : 2004.
- (28) OLIVEIRA, W., UEMORI, M., ROCHA, J., DIAS, C. **Barras Roebel Versus Bobinas de Múltiplas Espiras: Impactos na Reabilitação de Hidrogeradores.** XXI SNPTEE – Seminário Nacional de Produção e Transmissão de Energia Elétrica. Grupo de Estudo de Geração Hidráulica – GGH, Florianópolis, Brasil: 2011.
- (29) MARQUES, A.P., SÁ, L.C.R. **Avaliação de Envelhecimento em Barras de Estator de Hidrogeradores.** XV SNPTEE – Seminário Nacional de Produção e Transmissão de Energia Elétrica. Grupo de Estudo de Geração Hidráulica – GGH, Foz do Iguaçu, Brasil: 1999.
- (30) SANTOS, R.L.S. **Estudo da Influência do Ciclo de Partidas e Paradas no Estator de Geradores Síncronos.** Dissertação de Mestrado, Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal do Pará, Belém, Brasil: 2011.
- (31) STONE, G.C., SASIC, M., DUNN, D., CULBERT, I. **Recent Problems Experienced With Motor and Generation Windings.** Copyright Material IEEE, No. PCIC-2009-6. 2009.
- (32) DANTAS, R.D. **Método SVD: Análise Wavelets-Mãe na Extração de Textura Aplicadas em Mamogramas.** Conferência Brasileira de Dinâmica, Controle e Aplicações. Águas de Lindóia, São Paulo: 2011.
- (33) MONTGOMERY, D.C. **Design and Analysis of Experiments.** New York : John Wiley and Sons, 2008.
- (34) HACKELING, G. **Mastering Machine Learning with Scikit-Learn.** Birmingham, UK : Packt Publishing Ltd, 2014.
- (35) CARDOSO, C. A. R. **Atenuação do Ground Roll Utilizando Filtragem SVD.** Trabalho de Graduação, Universidade Federal da Bahia, Salvador, Bahia, Brasil: 2010.
- (36) GANDRA, S. R. **Aplicação da Decomposição em Valores Singulares à Análise de Dados Multivariados.** Relatório Final – Programa Institucional de Iniciação Científica. Universidade Federal do Amazonas. Manaus, Amazonas, Brasil: 2013.
- (37) BENÍTEZ, R. ESCUDERO, G. KANAAN, S. **Inteligencia Artificial Avanzada.** Universitat Oberta de Catalunya. ISBN:978-84-9029-887-9, 2013.

- (38) BARAGONA, R. BATTAGLIA, F. **Outliers Detection in Multivariate Time Series by Independent Component Analysis.** Neural Computation. No 7, Vol. 19, 2007, p. 1962-1984.
- (39) LEITE, L. **Análise de Componentes Independentes Aplicada à Identificação de Regiões Lesionadas em Mamogramas.** Tese de Doutorado. Programa de Pós-Graduação de Engenharia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil: 2005.
- (40) LEITE, I.C.C. **Análise de Componentes Independentes Aplicada à Avaliação de Imagem Radiográfica de Sementes.** Tese de Doutorado. Programa de Pós-Graduação em Estatística e Experimentação Agropecuária. Universidade Federal de Lavras, Lavras, Minas Gerais, Brasil: 2013.
- (41) OJA, E. HYVÄRINEN, A. **Independent Component Analysis: Algorithms and Applications.** Neural Networks, Helsinki, Finland: 2000.
- (42) FACELI, K. LORENA, A.C. GAMA, J. CARVALHO, A.C.P.L.F. **Inteligência Artificial: Uma Abordagem de Aprendizado de Máquina.** Rio de Janeiro, Brasil: LTC, 2011.
- (43) BREIMAN, L. **Random Forests.** Kluwer Academic Publishers. Manufactured in The Netherlands. Machine Learning. 2001, p 5-21.
- (44) VAPNIK, V. N. **The nature of statistical learning theory.** Springer-Verlag. ISBN 0387945598, New York, USA: 1995.
- (45) SMOLA, A. J. SCHÖLKOPF, B. **Learning with Kernels.** The MIT Press. Cambridge, 1999.
- (46) ROSSANT, C. **Learning IPython for Interactive Computing, and Data Visualization. Birmingham.** UK: Packt Publishing, 2013.
- (47) PEDREGOSA, F. VAROQUAUX, G. GRAMFORT, A. MICHEL, V. **Scikit-Learn: Machine Learning in Python.** Journal of Machine Learning. 2011.