

INSTITUTO FEDERAL DE EDUCAÇÃO,
CIÊNCIA E TECNOLOGIA DO CEARÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM ENERGIAS RENOVÁVEIS

Navar de Medeiros Mendonça e Nascimento

Detecção e classificação de falhas de curto-circuito em geradores de indução
aplicados em geração eólica utilizando aprendizado de máquina

FORTALEZA, CEARÁ

2018

Navar de Medeiros Mendonça e Nascimento

Detecção e classificação de falhas de curto-circuito em geradores de indução aplicados em
geração eólica utilizando aprendizado de máquina

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Energias Renováveis do Instituto Federal de Educação, Ciência e Tecnologia do Ceará como um dos requisitos para a obtenção do título de mestre em energias renováveis. Área de concentração: Energias Renováveis.

Prof. Dr. Pedro Pedrosa Rebouças Filho
Prof. Dr. Cláudio Marques de Sá Medeiros

FORTALEZA, CEARÁ

2018

Dados Internacionais de Catalogação na Publicação (CIP)
Ficha catalográfica elaborada pelo Bibliotecário
Gláucio Barreto de Lima - CRB-3/995

N244 Nascimento, Navar de Medeiros Mendonça e.
Detecção e classificação de falhas de curto-
circuito em geradores de indução aplicados em
geração eólica utilizando aprendizado de máquina
/ Navar de Medeiros Mendonça e Nascimento --
2018.
136 f.

Dissertação (Mestrado em Energias Renováveis) -
- Instituto Federal de Educação Ciência e
Tecnologia do Ceará, *Campus* Maracanaú, 2018.
Orientador: Prof. Dr. Pedro Pedrosa
Rebouças Filho.

1.ENERGIA EÓLICA. 2.GERADOR DE INDUÇÃO.
3.APRENDIZADO DE MÁQUINA. I.Título.

621.312136

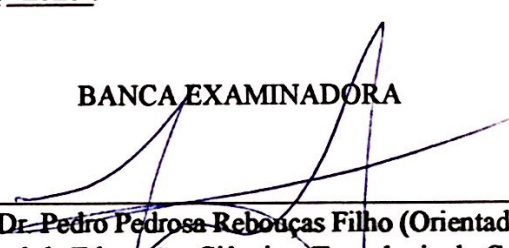
NAVAR DE MEDEIROS MENDONÇA E NASCIMENTO

DETECÇÃO E CLASSIFICAÇÃO DE FALHAS DE CURTO- CIRCUITO EM
GERADORES DE INDUÇÃO APLICADOS EM GERAÇÃO EÓLICA UTILIZANDO
APRENDIZADO DE MÁQUINA

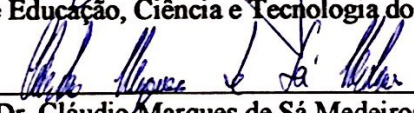
Dissertação submetida à Coordenação do Curso de Pós-graduação em Energias Renováveis do Instituto Federal de Educação, Ciência e Tecnologia do Ceará, como requisito parcial para a obtenção do título de Mestre em Energias Renováveis, área de concentração Energias Renováveis.

Aprovada em 05 / 04 / 2018.

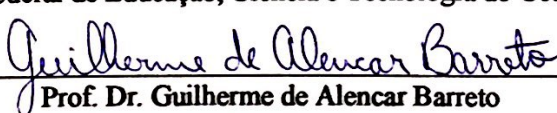
BANCA EXAMINADORA



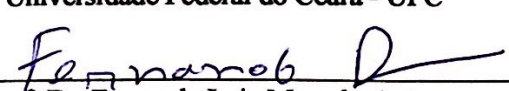
Prof. Dr. Pedro Pedrosa Rebouças Filho (Orientador)
Instituto Federal de Educação, Ciência e Tecnologia do Ceará - IFCE



Prof. Dr. Cláudio Marques de Sá Medeiros
Instituto Federal de Educação, Ciência e Tecnologia do Ceará - IFCE



Prof. Dr. Guilherme de Alencar Barreto
Universidade Federal do Ceará - UFC



Prof. Dr. Fernando Luiz Marcelo Antunes
Universidade Federal do Ceará - UFC

Dedico este trabalho ao meus pais.

Agradecimentos

Agradeço primeiramente a Deus, o grande arquiteto do universo.

A minha mãe, Jacqueline, por todo seu suporte dado em todas as minhas escolhas. Além disso, por seus incentivos diários e também, por acreditar tanto, às vezes mais que eu mesmo, no meu potencial.

Ao meu pai, Haroldo, por me ajudar racionalmente em minhas escolhas, além dos seus conselhos assertivos e engrandecedores.

Ao meu padastro, Nilo, pois sua calma diversas vezes me transmitiu paz. Sua simplicidade nas resolução de problemas me fez acreditar ser capaz.

Ao Instituto Federal do Ceará (IFCE), por me propiciar um ambiente saudável, servindo como uma segunda casa, na qual pude crescer e aprender desde os 17 anos de idade. Em especial, aos laboratórios de Processamento de Imagens, Sinais e Computação Aplicada (LAPISCO) e Processamento de Energias (LPE), por terem me provido suporte em tudo que precisei.

Ao meu orientador Prof. Pedro Pedrosa, por ter acreditado na minha capacidade e ter apostado nos meus interesses acadêmicos. Também, por me fazer imaginar grande e a acreditar que com trabalho duro somos capaz de conquistar tudo o que quisermos. Ao meu coorientador Prof. Cláudio Sá, pelo rigor imposto em seus questionamentos e sugestões, pois sei que sempre foram almejando evoluções.

Ao meus colegas de laboratório, todos sempre prestativos em me ajudar nos que lhe fossem tangível. Em especial, ao meu amigo Igor Sousa, nossas discussões acadêmicas me ajudaram bastante durante essa caminhada.

“It’s not who you are underneath, but what you do that defines you.”
(Rachel Dawes, The Dark Knight series)

Resumo

As fontes de energia renováveis vêm conquistando espaço nas matrizes energéticas à medida que a dependência de petróleo diminui. Dentre as fontes limpas, destaca-se a energia eólica, a qual apresentou crescimentos acelerados na última década. Contudo, pesquisas indicam que, para tornar a energia eólica uma solução efetiva, os custos operacionais com manutenção devem ser diminuídos. A manutenção voltou os olhos para essa necessidade e as técnicas para monitoramento das condições de funcionamento servem como um aliado às rotinas de gerenciamento em parques eólicos. Os geradores elétricos de aerogeradores são indicados como o equipamento de maior ônus financeiro a um parque eólico. As falhas de curto-circuito entre espiras do bobinamento estatórico são severas para essa máquina elétrica. Pesquisas buscam soluções para a detecção prematura de falhas. Baseado nesse contexto, objetivou-se construir uma bancada para emular um sistema de geração eólica que possibilite a inserção de falhas de curto-circuito em um gerador trifásico. Realizaram-se aquisições de dados de corrente elétrica para criar uma base para análise de falhas. Propõem-se o uso de quatro técnicas para extração de atributos, que são: a transformada de Fourier, o algoritmo de Goertzel, Estatística de Alta Ordem (HOS) e a Matriz de Coocorrência Estrutural (SCM). Os bancos de dados gerados serviram para o treinamento dos seguintes classificadores: mínimos quadrados ordinários (OLS), perceptron simples (PS), Gaussiano e perceptron multi-camadas (MLP). Propôs-se ainda um classificador binário, com opção de rejeição, que identificou 100% das condições normais com menos de 1% de taxa de falsos positivos e falsos negativos. Para a classificação dos tipos de falha, diferentes topologias da MLP foram utilizadas, o que tornou possível identificar um curto-circuito incipiente, de alta impedância, em 1,41% das espiras com 61,41% de acurácia, utilizando três sensores de corrente. Ao utilizar apenas um sensor de corrente a taxa de acerto dessa mesma falha cai para 38,06%. Porém, a acurácia da MLP para as condições de normais manteve-se em 99,25%. A combinação dos atributos gerador pelo extrator Fourier com a rede MLP demonstrou-se mais efetiva para a detecção de falhas, pois obteve 84,48% de acertos gerais na base de dados. O melhor resultado do classificador Gaussiano foi 76,72% de acerto na base formada pelos atributos gerados com o extrator HOS. Dentre os classificadores lineares o melhor resultado foi do OLS, com 64,91% de taxa de acerto considerando a base formada pelos atributos do extrator Fourier, superando o PS em todos os casos. Conclui-se que é possível identificar prematuramente as falhas do gerador de indução aplicado em um sistema de aerogeração, com garantia da confiabilidade na classificação da rede neural, utilizando apenas um sensor de corrente. Caso seja de interesse classificar o tipo da falha é necessário utilizar três sensores de corrente.

Palavras-chaves: Energia eólica; Gerador de indução; Análise da assinatura de corrente elétrica; Aprendizado de máquina.

Abstract

While the dependence on oil decreases as renewable energy sources gain space in the energy matrix. Among clean energy sources for electricity, there are highlights for wind power, which has experienced accelerated growth in the last decade. However, research indicates wind energy can be a more effective solution if problems like maintenance costs are mitigated, since it impacts on the final energy costs. Predictive maintenance meets this need and techniques for condition monitoring serves as an ally to management routines in wind farms, in particular to generators installed in wind turbines, which are indicated as the equipment of greater financial burden in a wind farm. Winding inter-turn Short-circuits might be destructive for the electric machine, and researches seek solutions to predict these types of problems, such as the use of electric current signature combined with machine learning methods. In this way, the objective was to build a test-bench for emulation of a wind turbine that allows the insertion of short-circuit faults in the three-phase induction generator. Electrical current data acquisition was performed to create a fault database. We propose the use of four techniques for extracting attributes, which are: the Fourier transform, the Goertzel algorithm, High Order Statistics (HOS) and the Structural Matrix of Co-Corruption (SCM). The constituted databases were used to train the classifiers: Ordinary Least Squares (OLS), Simple Perceptron (PS), Gaussian and Perceptron Multi Layer (MLP). A binary classifier with rejection option was proposed, which identified 100% of the generator Normal conditions with less than 1% false positive and negative rate. To classify the fault type different topologies of the MLP were used and it was possible to identify incipient faults, such as high impedance short circuit in 1.41% turns with 61.41% accuracy using three current sensors. When using only one current sensor the hit rate of that same fault drops to 38.06%. But, the MLP's accuracy for Normal's conditions remained at 99.25%. The combination of the features generated by Fourier extractor and the MLP network proved to be more effective for detection and fault in the wind turbine, since it obtained 84.48% of general hits in the database. The best result of the Gaussian classifier was 76.72% of hit rate in the database formed with features from HOS extractor. Among the linear classifiers, the best result was the OLS, with 64.91% of hit rate in the base formed with the Fourier Extractor, surpassing the PS in all cases. It is concluded that is possible to identify prematurely the faults of the induction generator applied in a wind turbine system, with guarantee of the reliability of the neural network classification, using only a current sensor. And, if the interest is to use classify fault by its type, it is necessary to use three current sensors.

Key-words: Wind Energy; Induction Generator; Current Signature Analysis; Machine Learning.

Lista de ilustrações

Figura 1 – <i>Ranking</i> mundial de potência eólica instalada.	21
Figura 2 – Panorama da aerogeração no Brasil em 2016.	22
Figura 3 – Impactos na operação de um parque eólico <i>versus</i> tipo de parada em máquinas.	23
Figura 4 – Extratificação das falhas em máquinas de indução.	26
Figura 5 – Sistema típico de aerogeração.	30
Figura 6 – Configurações de geradores elétricos aplicados na aerogeração.	31
Figura 7 – Partes do motor do indução.	35
Figura 8 – Rotor gaiola de esquilo.	36
Figura 9 – Diagrama esquemático da operação da máquina elétrica de indução como:	37
Figura 10 – Tipos de curto-circuito no bobinamento estatórico da máquina de indução.	39
Figura 11 – Topologia do neurônio de McCulloch e Pitts (1943).	43
Figura 12 – Arquitetura da rede neural do tipo MLP.	49
Figura 13 – Diagrama do sistema de emulação do aerogerador.	57
Figura 14 – Instalações da bancada de emulação do sistema de aerogeração.	58
Figura 15 – Ensaio preliminares para a caracterização do regime permanente do gerador de indução. Em 15a é o comportamento da corrente elétrica, em 15b da tensão no barramento CC e em 15c da potência trifásica gerada.	61
Figura 16 – Detalhe da inserção do curto-circuito entre espiras do bobinamento do estator.	62
Figura 17 – Esquema de ligações para os dois tipos de curto-circuito	63
Figura 18 – Foto da máquina elétrica preparada para inserção de curto-circuito entre espiras.	64
Figura 19 – Foto do módulo NI-USB-6009 em 19a e da interface no LabVIEW em 19b.	65
Figura 20 – Foto do sensor de corrente em 20a e o diagrama de instalação em 20b e 20c.	66
Figura 21 – Foto do sensor MEMs 21a o diagrama de instalação em 21b.	67
Figura 22 – Foto da bobina para captação de fluxo axial em 22a o diagrama de instalação em 22b.	68
Figura 23 – Comportamento do GP, em azul, e GM, em verde. Em 23a é o comportamento da corrente elétrica, em 23b da tensão no barramento CC e em 23c da potência trifásica gerada	69

Figura 24 – Curvas do SCIG modificado para inserção de curto-circuitos. Em 24a é o comportamento da corrente elétrica, em 24b da tensão no barramento CC e em 24c da potência trifásica gerada	70
Figura 25 – Corrente do gerador normal e com falha de BI-3, com $f_b = 60 Hz$. . .	72
Figura 26 – Corrente elétrica de linha do gerador normal e com falha, em regime de corrente nominal	73
Figura 26 – (Continuação) Corrente elétrica de linha do gerador normal e com falha, em regime de corrente nominal	74
Figura 27 – Comparativo da corrente elétrica das três linhas, à operação próxima a corrente nominal, em diferentes condições de falha. Na legenda (1) é o gerador em estado normal, (2) é falha de AI-1, (3) é BI-1, (4) AI-2, (5) é a BI-2, (6) é AI-3, (7) é a BI-3	75
Figura 28 – Curvas do SCIG modificado nas condições normais, em tons de azuis, e falha de BI-3, em tons de vermelhos. Em 28a é o comportamento da corrente elétrica, em 28b da tensão no barramento CC e em 28c da potência trifásica gerada	77
Figura 29 – Curvas do SCIG modificado nas nas condições normais, em azuis, e falha de CC AI-1, em amarelos. Em 29a é o comportamento da corrente elétrica, em 29b da tensão no barramento CC e em 29c da potência trifásica gerada	78
Figura 30 – Projeções bi-dimensionais dos bancos de dados para cada um dos extratores.	79
Figura 31 – <i>Boxplot</i> da acurácia da aplicação do OLS nas bases de dados formadas pelos extratores. Em 31a estão os resultados na base de treinamento e em 31b na base de teste.	81
Figura 32 – <i>Boxplot</i> da acurácia da aplicação do Perceptron simples nas bases de dados formadas pelos extratores. Em 32a estão os resultados na base de treinamento e em 32b na base de teste.	82
Figura 33 – <i>Boxplot</i> da acurácia da aplicação do Perceptron simples nas bases de dados formadas pelos extratores. Em 32a estão os resultados na base de treinamento e em 32b na base de teste.	83
Figura 34 – Histograma das harmônicas de Fourier utilizadas como atributos comparado com uma distribuição Gaussiana.	83
Figura 35 – <i>Boxplot</i> da acurácia da aplicação da MLP nas bases de dados formadas pelos extratores. Em 35a estão os resultados na base de treinamento e em 35b na base de teste.	85
Figura 36 – Impacto da variação do número de épocas na acurácia média após 50 treinamentos.	86

Figura 37 – Impacto da variação da taxa de aprendizagem na acurácia média após 50 treinamentos.	87
Figura 38 – Impacto da variação da taxa de momento na acurácia média após 50 treinamentos.	88
Figura 39 – Impacto da variação do número de neurônios ocultos, (Q), na acurácia média após 50 treinamentos.	88
Figura 40 – Impacto na acurácia da MLP mediante adição dos atributos para formação de novas bases de dados: tensão no barramento CC (BT-2), corrente eficaz (BT-3), f_g (BT-4) e com os três propostos (BT-5).	90
Figura 41 – Projeções bi-dimensionais das bases de dados, formadas com o extrator Fourier, após a adição dos atributos V_{cc} , f_g e I_{Rrms}	91
Figura 42 – Espectro de frequência em $f_g = 60\text{ Hz}$ da condição normal, com $f_g = 58,87\text{ Hz}$ e espectro da falha de CC BI NVL 3, com $f_g = 58,84\text{ Hz}$. . .	95
Figura 43 – Variação numérica das harmônicas normalizadas pela fundamental, da condição normal (azul) e falha de BI-3 (vermelho).	96
Figura 44 – Espectro de frequência da condição Normal para diferentes f_b	97
Figura 45 – <i>Boxplot</i> da acurácia da aplicação do MLP nas bases de teste e treinamento, formado por cada um dos sensores de corrente.	100
Figura 46 – Valores das harmônicas normalizadas pela fundamental, f_n , na condição de $f_b = 60\text{ Hz}$, da condição Normal e todos os tipos de falhas.	101
Figura 47 – Saída da MLP para a base de dados binária.	110
Figura 48 – Curvas de tensão e potência, em uma mesma condição de f_g para a Classe Normal e Falha BI-3.	111
Figura 49 – Saída da MLP para a base de dados binária, utilizando a frequência relativa como atributo.	112
Figura 50 – Etapas para a constituição da base de dados.	136

Lista de tabelas

Tabela 1 – Resumo dos estados da máquina considerando Alta Impedância (AI), Baixa Impedância (BI), com os níveis de intensidade equivalente aos percentuais de espiras em curto de 1.41%, 4.81% e 9.26%, totalizando 6 níveis de falha.	64
Tabela 2 – Passos para realização dos experimentos no gerador de indução	71
Tabela 3 – Comparativo numérico entre os valores de corrente elétrica de linha, tensão no barramento CC e potência para os estado normal e falha BI-3. . . .	72
Tabela 4 – Resumo dos hiperâmetros propostos para a combinação Fourier-MLP . . .	89
Tabela 5 – Resultados gerais dos extratores-classificadores para as seguintes métricas: Acurácia (Acc), Sensibilidade (Sen), Especificidade (Esp), F-score (Fsc) nas bases de treinamento e teste.	93
Tabela 6 – Taxa de acertos por classes de todos os classificadores utilizando o extrator Fourier.	94
Tabela 7 – Matriz de confusão percentual da combinação Fourier-MLP para a detecção em 6 níveis de falha.	99
Tabela 8 – Resultado médio da classificação da MLP nas bases de dados cujas correntes da fases S e T compõem novas amostras.	102
Tabela 9 – Matriz de confusão média proveniente da classificação da MLP na base de dados de testes cujas correntes da fases S e T compõem novas amostras. Em negrito estão as acurácias por classe.	103
Tabela 10 – Matriz de confusão média proveniente da classificação da MLP na base de dados de testes cujas correntes da fases S e T compõem novas amostras. A classificação levou em consideração a binarização da base em: classe Normal e Falha.	103
Tabela 11 – Resultado médio da classificação da MLP nas bases de dados cujas correntes da fases S e T compõem novos atributos.	104
Tabela 12 – Matriz de confusão média proveniente da classificação da MLP na base de dados de testes cujas correntes da fases S e T compõem novos atributos. Em negrito estão as acurácias por classe.	105
Tabela 13 – Resultado médio da classificação da MLP nas bases de dados, com as falhas agrupadas pelas quantidade de espiras em curto circuito de 1,41% (NVL-1), 4,81% (NVL-2) e 9,26% (NVL-3).	106
Tabela 14 – Matriz de confusão média proveniente da classificação da MLP na base de dados de testes, com as falhas agrupadas pelas quantidade de espiras em curto circuito de 1,41% (NVL-1), 4,81% (NVL-2) e 9,26% (NVL-3).	107

Tabela 15 – Resultado médio da classificação da MLP nas bases de dados, com as falhas agrupadas por tipo de curto (<i>i.e</i> Baixa impedância - BI e Alta impedância - AI).	107
Tabela 16 – Matriz de confusão média proveniente da classificação da MLP na base de dados de testes, com as falhas agrupadas por tipo de curto (<i>i.e</i> Baixa impedância - BI e Alta impedância - AI).	108
Tabela 17 – Resultado médio da classificação da MLP nas bases de dados, com as todas as falhas agrupadas em uma única classe.	108
Tabela 18 – Matriz de confusão média proveniente da classificação da MLP na base de dados de testes, com as todas as falhas agrupadas em uma única classe.	109

Lista de abreviaturas e siglas

IFCE	Instituto Federal de Educação, Ciência e Tecnologia do Ceará
SCIG	Gerador de Indução Gaiola Esquilo - <i>Squirrel Cage Induction Generator</i>
DIFG	Gerador de Indução Duplamente Alimentados - <i>Doubly Fed Induction Generator</i>
PMSG	Gerador Síncrono de Imã Permanente - <i>Permanent Magnet Synchronous Generator</i>
WRIG	Gerador de Indução de Rotor Bobinado - <i>Wound Rotor Induction Generator</i>
SA	Sistema de Aerogeração
CC	Corrente Contínua
CA	Corrente Alternada
MP	Máquina Primária
f_g	Frequência do gerador
f_b	Frequência de base
f_r	Frequência relativa
MCSA	Análise da Assinatura de Corrente do Motor - <i>Motor Current Signature Analysis</i>
OLS	Mínimos Quadrados Ordinários - <i>Ordinary Least Square</i>
PS	Perceptron Simples
MLP	Perceptron de Múltiplas Camadas - <i>Multilayer Perceptron</i>
HOS	Estatística de Alta Ordem - <i>Higher-Order Statistics</i>
SCM	Matriz Estrutural de Coocorrência - <i>Structural Co-occurrence Matrix</i>
QTCM	Quadro de Teste de Curto em Máquinas
AI	Alta Impedância
BI	Baixa Impedância

GP	Gerador Padrão
GM	Gerador Modificado
NVL-1	curto-circuito em 1,41% de espiras
NVL-2	curto-circuito em 4,81% de espiras
NVL-3	curto-circuito em 9,26 % de espiras

Lista de símbolos

$\mathbb{R}^p$	Espaço de atributos
p	atributo
$\mathbf{x}_n$	vetor de entrada
C_i	i-ésima Classe
d_i	i-ésima rotulo
b	viés, ou <i>bias</i>
$\mathbf{W}$	Matriz de pesos
$\mathbf{X}$	Matriz de dados
$\mathbf{D}$	Matriz dos rótulos
$\mathbf{I}$	Matriz identidade
$\mathbf{D}$	Matriz dos rótulos
$\mathbf{w}_i$	Vetor de pesos do i-ésimo neurônio
$P(c_i \mathbf{x})$	Probabilidade da amostra $\mathbf{x}$ em relação a classe i
$P(c_i)$	Probabilidade <i>a priori</i> a classe c_i
Σ_i	Matriz de covariância da classe i
$g_i(x)$	função discriminante da classe i
σ_{xy}	Covariância da variável x em relação à y
t	Iteração
$y_i^{(h)}(t)$	Saída do i-ésimo neurônio da camada oculta
$y_i^{(o)}(t)$	Saída do i-ésimo neurônio da camada de saída
$u_i^{(h)}(t)$	<i>Net</i> do i-ésimo neurônio da camada oculta
$u_i^{(o)}(t)$	<i>Net</i> do i-ésimo neurônio da camada de saída
w_{ij}	Peso que conecta a j-ésima entrada ao i-ésimo neurônio da camada oculta

m_{ki}	Peso que conecta o i-ésimo neurônio ao k-ésimo neurônio da camada de saída
φ_i	Função de ativação do i-ésimo neurônio
$\delta_k(t)$	gradiente do k-ésimo neurônio
η	taxa de aprendizagem
α	termo <i>momentum</i>
f_{cc}	frequência de curto-circuito
f_n	frequência fundamental
s	escorregamento
V_{cc}	Tensão elétrica no barramento CC

Sumário

1	INTRODUÇÃO	21
1.1	Motivação	23
1.2	Justificativa	25
1.3	Objetivos	25
1.3.1	Objetivos específicos	26
1.4	Produção científica	26
1.5	Organização do texto	27
2	FUNDAMENTAÇÃO TEÓRICA	29
2.1	Aerogeração e geradores elétricos	29
2.1.1	Geradores elétricos	30
2.1.1.1	Estado da arte para o SCIG aplicado em aerogeradores	33
2.1.1.2	Constituição do gerador de indução trifásico	34
2.1.1.3	Princípio de funcionamento	36
2.2	Monitoramento das condições de funcionamento	37
2.2.1	Falhas de curto-circuito em máquinas elétricas de indução	38
2.2.2	Estado da arte em detecção de falhas de curto-circuito entre espiras estatísticas	39
2.3	Técnicas de aprendizado de máquina	40
2.3.1	Classificadores lineares	42
2.3.1.1	Mínimos Quadrados Ordinário	42
2.3.1.2	Perceptron simples	43
2.3.2	Classificadores Gaussianos	44
2.3.2.1	Matrizes de covariância	45
2.3.3	Redes Neurais Artificiais	48
2.3.3.1	Perceptron de Múltiplas Camadas	48
2.3.3.1.1	Algoritmo do <i>back propagation</i>	50
2.3.4	Estado da arte de aplicações em máquinas rotativas	51
2.4	Técnicas de extração de atributos em sinais	53
2.4.1	A transformada de Fourier à luz da teoria de Penman et al. (1994)	53
2.4.2	O algoritmo de Goertzel	54
2.4.3	A Estatística de Alta Ordem (HOS)	54
2.4.4	Matriz de Coocorrência Estrutural	55
2.5	Contribuições da pesquisa	55
3	EXPERIMENTOS E AQUISIÇÕES DE DADOS NO SCIG	57

3.1	Descrição do processo de emulação do conjunto turbina eólica/gerador de indução	59
3.2	O protótipo do gerador	62
3.3	Sistema de monitoramento	64
3.3.1	Sensoriamento de corrente elétrica	65
3.3.2	Sensoriamento de vibração mecânica	66
3.3.3	Sensoriamento de fluxo magnético axial	67
3.4	Caracterização experimental do gerador	68
3.5	Extração de atributos e formação do banco para avaliação e validação de resultados	71
3.5.1	Análises preliminares	71
3.5.2	Formações das bases de teste de corrente elétrica para análise de falhas	79
4	INVESTIGAÇÃO PRELIMINAR SOBRE EXTRATORES DE ATRIBUTOS E CLASSIFICADORES	80
4.1	Procedimentos comuns a todos os classificadores	80
4.2	Classificadores lineares	81
4.3	Classificador Gaussiano	82
4.4	Classificador Neural MLP	84
4.4.1	Número de Épocas	86
4.4.2	Taxa de aprendizagem	86
4.4.3	Termo de <i>momentum</i>	87
4.4.4	Número de neurônios	88
4.5	Adição de novos atributos à base de dados	89
4.6	Resultados das aplicações extrator <i>versus</i> classificador	91
4.7	Conclusões parciais	97
5	PROJETO DO CLASSIFICADOR MLP	98
5.1	Classificador neural proposto	98
5.2	Avaliação do uso das correntes elétricas de todas as fases	100
5.2.1	Harmônicas das fases S e T como novos atributos	103
5.3	Aumento da confiabilidade na classificação da MLP	105
5.3.1	Agrupamento das falhas por quantidade de espiras em curto-circuito	105
5.3.2	Agrupamento das falhas entre baixa e alta impedância	107
5.3.3	Agrupamento binário	108
6	CONCLUSÕES	113
6.1	Conclusões à luz das hipóteses levantadas	114
6.2	Limitações	115
6.3	Considerações gerais	115

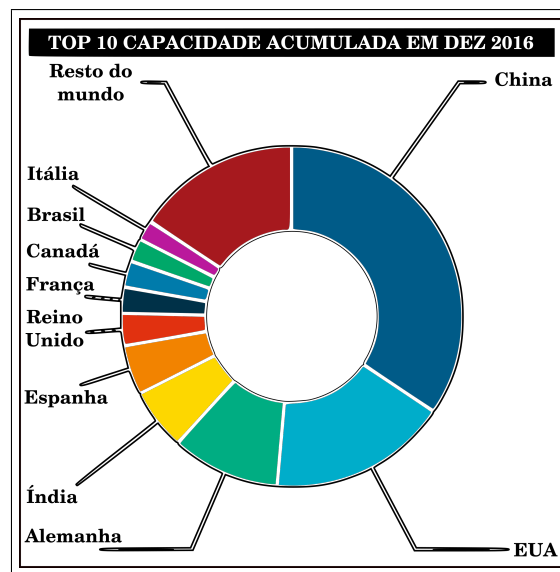
6.4	Sugestões para trabalhos futuros	116
	REFERÊNCIAS	117
	APÊNDICES	124
	APÊNDICE A – QUADRO DE TESTES DE CURTO-CIRCUITO EM MÁQUINAS (QTCM)	125
	APÊNDICE B – DETALHES SOBRE A FORMAÇÃO DA BASE DE DADOS	136

1 Introdução

Em virtude dos estudos e discussões sobre impactos ambientais causados pelas queimas de combustíveis fósseis, novas pesquisas emergiram sobre a utilização de energias renováveis para diversos fins. Nos últimos 40 anos, a utilização de fonte renováveis para geração de eletricidade cresceu 221%, enquanto o uso do petróleo e seus derivados decaiu 68% ([International Energy Agency, 2017](#)). Isso denota uma mobilização global para o uso de fontes limpas de energia. Dentre as fontes de energias renováveis, a energia eólica tornou-se mundialmente a solução mais efetiva e aceita para geração de energia elétrica, contribuindo com 486,7 GW para a demanda mundial ([PAO; JOHNSON, 2011](#); [GWEC, 2017](#)). Essa produção de energia representa apenas 3% da demanda energética mundial. Porém, estima-se que em 2030 a aerogeração será capaz de suprir de 17 a 19% da necessidade mundial ([Global Wind Energy Council, 2015](#)).

Esse crescimento também é perceptível no Brasil, pois a potência instalada atual é 45 vezes maior do que em 2006, fato esse que ajudou a colocar o país em 9º lugar do mundo em geração eólica ([GWEC, 2017](#)) e em 5º lugar com a maior capacidade instalada no ano de 2016. Na Figura 1 é exibida a representatividade dos principais países na matriz eólica mundial. Atualmente a China, EUA, Alemanha e Índia são os que detêm a maior capacidade instalada.

Figura 1 – *Ranking* mundial de potência eólica instalada.

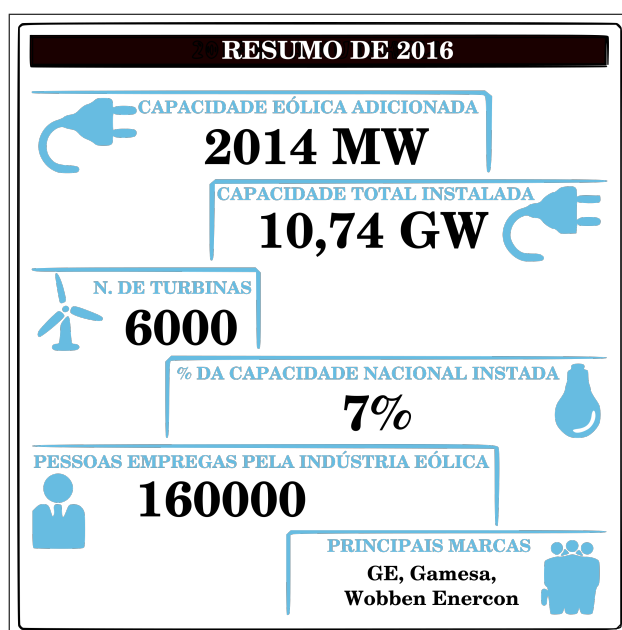


Fonte: Adaptado de [GWEC \(2017\)](#).

O crescimento no Brasil é notável. De acordo com [GWEC \(2017\)](#), o país é extremamente atrativo aos olhos dos investidores, pois a geografia nacional é propícia para a

geração eólica. Além disso, há políticas públicas para investimento nesse setor, como os créditos criados pelo Banco Nacional de Desenvolvimento Econômico e Social (BNDES) e o Programa de Incentivo às Fontes Alternativas de Energia (PROINFA). Os objetivos de ambos é diversificar a matriz energética brasileira e incentivar a geração distribuída, visando criar um sistema de geração descentralizado para aumentar a capacidade de produção e a confiabilidade do sistema elétrico (XAVIER et al., 2015). Na Figura 2, há indicativos da realidade atual do setor eólico, ficando claro que os investimentos e esforços tanto políticos quanto empresariais convergiram para o país atingir a capacidade de 10,74 GW, empregando ainda mais de 160 mil pessoas.

Figura 2 – Panorama da aerogeração no Brasil em 2016.



Fonte: Adaptado de GWEC (2017).

Apesar da crescente exploração desta fonte de energia, de 17% ao ano (Global Wind Energy Council, 2015), as tecnologias para tal não estão consolidadas. Ainda existem desafios de engenharia e ciência a serem resolvidos para suportar esta expansão (PAO; JOHNSON, 2011). A interdisciplinaridade deste campo proporciona pesquisas em diversas áreas. Alguns exemplos são: aspectos de natureza mecânica, que estão sendo estudados, como evidenciado em Barrett e Ning (2016), Ben Hassena et al. (2015), Mudi e Sinha (2014), Oh et al. (2015), Puleva (2015) e Wenlei et al. (2014); pesquisas sobre peculiaridades das construções *offshore* foram encontradas na literatura em Jeong, Park e Lee (2015) e Tahani et al. (2017); problemas de natureza elétrica são abordados em Alnasir et al. (2014), Matos et al. (2015) e Zhang et al. (2013); além de análises de viabilidade econômica, como em May, McMillan e Thöns (2015).

Dentre as áreas de estudo que cercam a aerogeração, o presente texto enfoca as pesquisas sobre os geradores elétricos, em especial o gerador de indução gaiola de esquilo,

do inglês *Squirrel Cage Induction Generator* (SCIG), propondo técnicas para a detecção incipiente de falhas no mesmo.

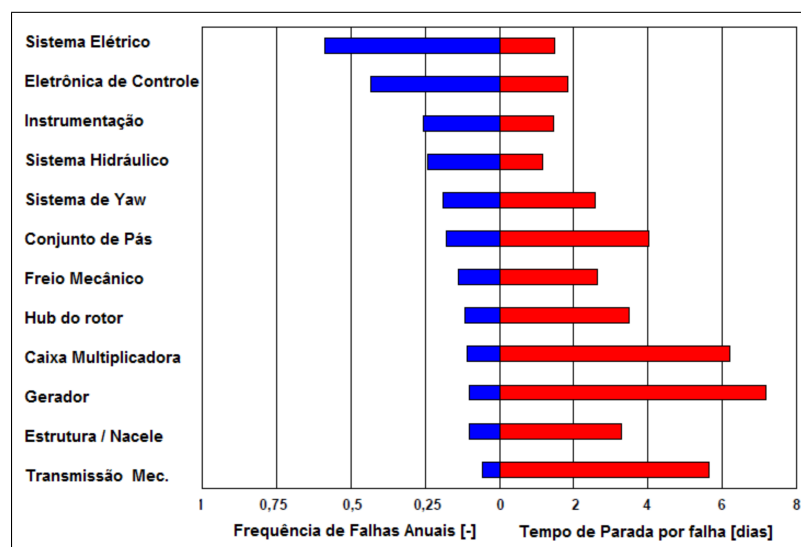
1.1 Motivação

Os dois principais motivadores para o desenvolvimento dessa pesquisa são: o uso dos geradores de indução em sistemas de aerogeração e o potencial eólico brasileiro, em especial, o Nordeste. Yaramasu et al. (2015) indicam SCIG como uma tendência futura, juntamente com grandes fabricantes de aerogeradores. As principais razões para isso serão melhores discutidas na Seção 2.1.1.1. Sabe-se que há uma busca alternativa que propicia a redução dos custos da energia.

Problemas operacionais impactam diretamente no custo de energia, segundo Polinder et al. (2013), somente com um sistema de aerogeração confiável e disponível é que o custo da energia pode ser mitigado. E, à luz da confiabilidade estão os custos operacionais de manutenção, os quais são responsáveis por até 30% do custo da energia (POLINDER et al., 2013).

A preocupação com manutenção em parques eólicos já foi evidenciada por Hahn, Durstewitz e Rohrig (2007). Em seu estudo são exibidos os principais componentes de um aerogerador e a sua onerosidade para os parques eólicos, caso haja paradas não programadas. Na Figura 3 é exibida a compilação dos dados de falhas de um conjunto de parques instalados na Europa.

Figura 3 – Impactos na operação de um parque eólico *versus* tipo de parada em máquinas.



Fonte: Adaptado de Hahn, Durstewitz e Rohrig (2007).

É notável que problemas relacionados ao gerador elétrico são extremamente onerosos para o parque eólico, causando a parada média de 7 dias no estudo de Hahn, Durstewitz e

Rohrig (2007). É importante frisar que o estudo foi feito na Europa, que é um ambiente estabelecido e experiente com a operação de usinas eólicas. Não há estudos dessa natureza no Brasil, mas, visto que o GWEC (2017) condena a cadeia de suprimentos para essa atividade no país, essa severidade pode ser ainda maior, talvez levando a meses de máquina parada.

Portanto, para a inserção eficiente da energia eólica na matriz energética é preciso garantir sua confiabilidade. Estudos como o de Kusiak e Li (2011) exibem que o sistema de monitoramento das condições de funcionamento de equipamentos de um parque eólico provê benefícios para as rotinas de manutenção. Especialmente para o componente central de um aerogerador, o gerador elétrico, que já tem sido alvo de pesquisas, como a de Toma, Capocchi e Capolino (2013), que exhibe métodos para diagnóstico de problemas nesse equipamento.

Além disso, os integrantes do grupo de pesquisa de Controle e Processamento de Energia do Instituto Federal do Ceará (IFCE), mesmo grupo no qual está inserido este trabalho, já realizaram detecção e classificação de falhas em motores elétricos, como exibido em Oliveira, Silva e Medeiros (2013), Coelho et al. (2014), Vieira (2016) e Sousa e Sá (2017). Os resultados obtidos nesses trabalhos indicam que as metodologias adotadas pelo grupo são promissoras. Sendo a abordagem do presente trabalho semelhante a elas no seu foco, diferindo pois o trabalho em questão está concentrado em geradores elétricos aplicados voltados para a energia eólica.

Outro principal motivo é o fato de a presente pesquisa estar alinhada com a realidade nacional, pois o potencial eólico do Brasil é tanto que o GWEC (2017) relata que há capacidade para atender três vezes a demanda por eletricidade. Isso já havia sido relatado, também, pelo Ministério de Minas Energia (2001), há 15 anos, quando apresentou um relatório com estudo sobre o potencial eólico do país. O estudo foi feito através de geoprocessamento e estimou a capacidade para instalar 143,47 GW de potência, admitindo apenas regiões com padrões de ventos médios e constantes (maiores que 7,0 m/s e menores que 7,5 m/s). Contudo, se fossem aproveitados ventos a partir de 6 m/s, a capacidade para instalar seria de 1334,78 GW. E, atualmente, existem menos do que 0,01 % dessa capacidade. Essa projeção seria capaz de suprir toda a demanda brasileira de 1624 TWh de energia em 2050, segundo Empresa de Pesquisa Energética (2014).

Além disso, o Nordeste brasileiro é um forte representante na matriz eólica nacional, pois é detentor de quase 82% da potência instalada no Brasil (GWEC, 2017). Isso é justificado pelo fato dessa região possuir as melhores condições climáticas e geográficas para a aplicação de energias renováveis (TÍBA; LEAL, 2012). O Ceará, em especial, detém a terceira posição no *ranking* nacional, possuindo atualmente 68 usinas eólicas instaladas em regiões costeiras (*onshore*). As perspectivas para o território cearense são boas, tanto que isso já foi percebido por pesquisas, como a de Lima et al. (2015), a qual avaliou o

potencial eólico para instalação de usinas *offshore* no estado, o que nunca foi pensando, até então.

Os benefícios da exploração eólica no Ceará vão além da matriz energética brasileira, pois Dutra e Szklo (2008) enfatizam que as comunidades locais também se beneficiarão, uma vez que parte dos investimentos serão redirecionados aquela mesma região, de acordo com as regulamentações nacionais. Em suma, o cerne motivador desta pesquisa permeia os estudos de viabilidade para implantação de usinas eólicas no Ceará, que visa garantir expansão da indústria local, do estado e da sistema energético brasileiro.

1.2 Justificativa

Visto os altos custos atribuídos a um parque eólico quando há falha em geradores elétricos, existe uma preocupação latente com a manutenção desses equipamentos. Esse fator é ainda mais crítico em parques eólicos *offshore*, especialmente pelos altos gastos operacionais. Isso foi constatado também por Lau, Ma e Pecht (2012) durante sua revisão de métodos para prognóstico de falhas em aerogeradores. Então, há um direcionamento de pesquisas que buscam maior confiabilidade em sistemas de aerogeração, através do prognóstico de falhas em geradores de indução.

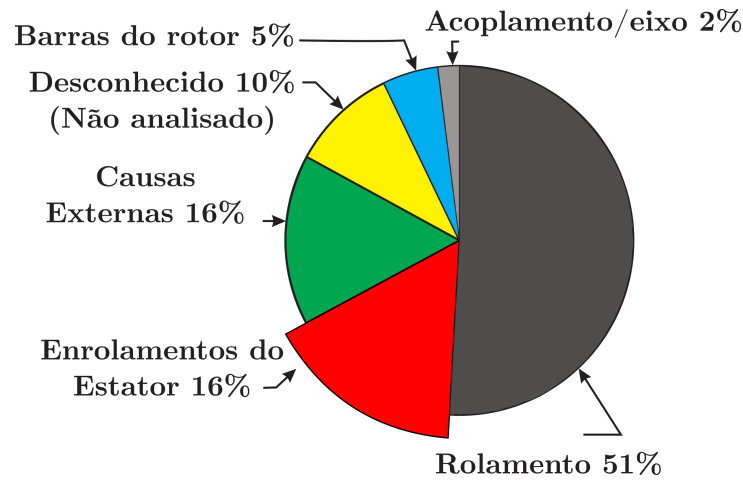
Uma compilação de dados de falhas em máquinas de indução na indústria química e do petróleo, em um período de 40 anos realizada por Bonnett (2010), exhibe quais as falhas mais expressivas nesse tipo de máquina. Em destaque estão as falhas em rolamentos, que representam 51% das falhas na máquina, seguida pelas falhas nos enrolamentos do estator, que dominam 16% do total, enquanto que falhas de barras no rotor representam 5%. Seus resultados são exibidos na Figura 4. Outros autores são mais enfáticos para as falhas de curto-circuito em bobinas do estator, como Kliman et al. (1996), os quais concluem que de 30% a 40% das falhas em motores de indução estão relacionadas a esse tipo de falha.

Baseado na importância do gerador, na representatividade do SCIG em aerogeradores e nas projeções futuras dessa tecnologia evidenciadas por Yaramasu et al. (2015), a presente pesquisa justifica-se pela necessidade de garantir maior confiabilidade a turbinas eólicas que utilizem o SCIG, embasado, também por Lima et al. (2015). Além de dar continuidade aos trabalhos já desenvolvidos pelo grupo de pesquisa do IFCE.

1.3 Objetivos

O objetivo deste trabalho é realizar a detecção de curto-circuito incipiente entre espiras do bobinamento estatórico do gerador de indução gaiola de esquilo, aplicado na geração eólica.

Figura 4 – Extratificação das falhas em máquinas de indução.



Fonte: Adaptado de Bonnett (2010).

1.3.1 Objetivos específicos

- Construir uma bancada para emulação de um sistema de aerogeração, que deve possuir um quadro de comandos para automatizar o processo de inserção do curto-circuito em um gerador elétrico e um sistema para sensoriamento e aquisição dos dados que vise replicar estudos anteriores e servir como gerador de dados para estudos futuros;
- Realizar aquisições dos sinais de caracterizam o funcionamento do gerador de indução, baseado na emulação de um sistema de aerogeração;
- Avaliar a aplicação de técnicas para extração de características em sinais;
- Avaliar métodos de aprendizado de máquina para detecção e classificação das falhas no gerador;

1.4 Produção científica

Durante o desenvolvimento dessa pesquisa os seguintes trabalhos científicos foram produzidos:

Artigos aceitos e aguardando publicação:

NASCIMENTO, N. M. M.; SOUSA, I. R.; REBOUÇAS FILHO, P. P.; MEDEIROS, C. M. S. A reliable approach for detection of incipient faults of short-circuits in induction generator using machine learning. **Computers and Electrical Engineering - Elsevier**, 2018.

SOUSA, P. H. F.; NASCIMENTO, N. M. M.; REBOUÇAS FILHO, P. P.; MEDEIROS, C. M. S. Detection and classification of faults in induction generator applied into wind

turbines through a machine learning approach. **International Joint Conference on Neural Networks (IJCNN)**, 2018.

NASCIMENTO, N. M. M.; REBOUÇAS FILHO, P. P.; MEDEIROS, C. M. S. A comparison of machine learning methods to identify broken bar failures in induction motors using statistical moments. In: **International Conference on Intelligent Systems Design and Applications (ISDA)**, 2017

NASCIMENTO, N. M. M.; REBOUÇAS FILHO, P. P.; MEDEIROS, C. M. S. Higher-Order Statistics applied to machine learning as an approach to identify broken rotor bars in induction motors. **IEEE Latin America**, 2017.

Artigos submetidos e/ou em processo de revisão:

REBOUÇAS FILHO, P. P.; GOMES, S. L.; NASCIMENTO, N. M. M.; MEDEIROS, C. M. S.; RODRIGUES, J. P. C.; ALBUQUERQUE, V. H. C. Machine Learning and Internet of Things to Predict Energy Production at a Wind Farm. **Applied Soft Computing**, 2017.

GONÇALVES, M. P.; NASCIMENTO, N. M. M.; ALMEIDA, J. S.; JÚNIRO, C. R. S.; FREITAS, F. N. C. F.; REBOUÇAS FILHO, P. P. Northeast Brazil as a keypoint in the development and popularization of modern renewable energies in Brazil. **Renewable & Sustainable Energy Reviews**, 2017.

Artigos em desenvolvimento:

RAMALHO, G. L. B.; REBOUÇAS FILHO, P. P.; MATOS, F.F.; ALMEIDA, J. F.; FREITAS, A. K. S.; PRATA, V. R.; NASCIMENTO, N. M. M. Detection of incipient cavitation in centrifugal pumps, using MEMS accelerometers and Machine Learning methods. **Sensors**, 2018.

1.5 Organização do texto

É importante enfatizar ao leitor que o texto dessa dissertação está organizado na ordem cronológica do desenvolvimento da pesquisa. Ao longo do trabalho foram elaboradas hipóteses e experimentos para suas validações. As confirmações ou refutações levaram à conclusões parciais, que incitaram novas hipóteses. Culminando, por fim, nas conclusões mediante os objetivos propostos.

Mediante isto, o texto dessa dissertação está organizado da seguinte forma: no Capítulo 2 são exibidas as fundamentações teóricas necessárias para o desenvolvimento da pesquisa; No Capítulo 3 explicam-se os passos para condução do experimento, que culminaram na formação dos quatro bancos de dados treinamento dos algoritmos de aprendizado de máquina; No Capítulo 4 são exibidos os passos para projeto do classificador de detecção incipiente de curto-circuito no SCIG, desde aplicação de ferramenta lineares

até as não lineares; As discussões e particularidades dos resultados são expostas no Capítulo 5 que culmina nas conclusões, exibidas no Capítulo 6.

2 Fundamentação teórica

Neste Capítulo há a fundamentação teórica necessária para o entendimento do trabalho. Os textos contém informações sucintas, porém importantes e são baseadas em referências clássicas das diversas áreas, como também é apresentado um estado da arte referente aos principais alicerces dessa pesquisa. O texto está organizado da seguinte forma: na Secção 2.1 é apresentada uma explicação sobre aerogeração, geradores elétricos e novidades que circunda essas áreas; na Secção 2.2 há um estado da arte sobre monitoramento das condições de funcionamento e as principais abordagens para detecção de falhas em máquinas elétricas; na Secção 2.3, apresenta-se uma explicação sucinta sobre os métodos de aprendizado de máquina utilizados nesse trabalho e, por fim, na Secção 2.4, as técnicas de extração de atributos utilizadas para formação dos bancos de dados.

2.1 Aerogeração e geradores elétricos

Há uma tendência natural no aumento de tamanho dos aerogeradores, pois as propriedades aerodinâmicas dos geradores eólicos demonstram que a potência elétrica gerada é proporcional ao quadrado do diâmetro do rotor e ao cubo da velocidade do vento (PAO; JOHNSON, 2011). Mediante um levantamento realizado por Yaramasu et al. (2015), a capacidade dos aerogeradores subiu de 50 kW para 7,5 MW e o diâmetro do seu rotor aumentou de 15 m para 126 m em 30 anos. Um impulsionador atual desta tendência são as tecnologias para geração mar adentro (*offshore*, do inglês) (BOOKMAN, 2005). Apesar de possuírem os mesmos objetivos das aplicações em terra (*onshore*, do inglês), possuem peculiaridades distintas, tais como: (i) tamanho e potências das turbinas; (ii) impactos socio-ambientais e (iii) fatores operacionais.

Pesquisas feitas em Global Wind Energy Council (2015) demonstraram que a potência média das turbinas em 2013 para aplicações *offshore* era de 3,613 MW contra 1,926 MW, em *onshore*, evidenciado também por Yaramasu et al. (2015), através de uma pesquisa de mercado, indicando que o tamanho dos aerogeradores instalados mar adentro são maiores que os instalados em terra. A justificativa para tal reside no fato de os ventos serem mais consistentes e menos turbulentos mar adentro, de acordo com Bookman (2005), então, espera-se melhor aproveitamento da capacidade de geração.

Outros fatores que permeiam os tipos de aerogeração são os impactos sociais associados, visto que a construção de parques eólicos *onshore* demanda grandes extensões terra, muitas vezes escassa, como em regiões desamente povoadas como a Europa e a China. Um levantamento feito por Global Wind Energy Council (2015) afirma que 91% da geração *offshore* está concentrada na Europa, o que indica a preferência europeia pela

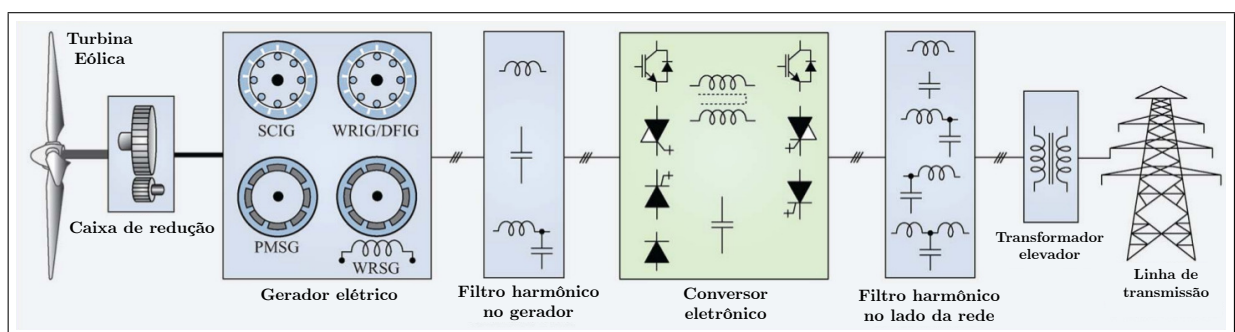
escolha dessa tecnologia. Outra vantagem das instalações de parques eólicos adentro do mar é a menor poluição visual causada, comparado aos parques em terra, especialmente quando situados na costa.

Entretando, há importantes particularidades que circundam as escolhas de construções entre parques *offshore* e *onshore*, entre as quais podem-se citar os elevados custos de instalação e manutenção de parques eólicos mar adentro. De acordo com [Polinder et al. \(2013\)](#), os custos de manutenção são responsáveis por 30% do valor da energia gerada. Outro fator preponderante está na interligação dos parques eólicos com a rede elétrica, pois a infraestrutura elétrica de conexão é mais complexa e cara. Porém, evidenciou-se pesquisas abordando novas técnicas e métodos, como em [Egea-alvarez et al. \(2014\)](#), [Kong e Jia \(2011\)](#) e [Yaramasu et al. \(2015\)](#), visando interligar o sistema de uma forma confiável, menos complexa e com menores custos.

2.1.1 Geradores elétricos

Um sistema típico de geração eólica é exibido na Figura 5. Dentre suas partes, há um eixo de pás responsáveis por transformar a velocidade do vento em movimento rotativo. Acoplado às pás está o gerador elétrico, que é o componente responsável por transformar a energia cinética rotativa em energia elétrica, tendo sendo assim, um papel crucial para o Sistema de Aerogeração (SA) ([YARAMASU et al., 2015](#)). Há também sistemas de filtros e conversores eletrônicos, os quais são responsáveis pelo gerenciamento e controle de potência, e filtros para condicionamento da tensão gerada para os padrões das linhas de transmissão.

Figura 5 – Sistema típico de aerogeração.

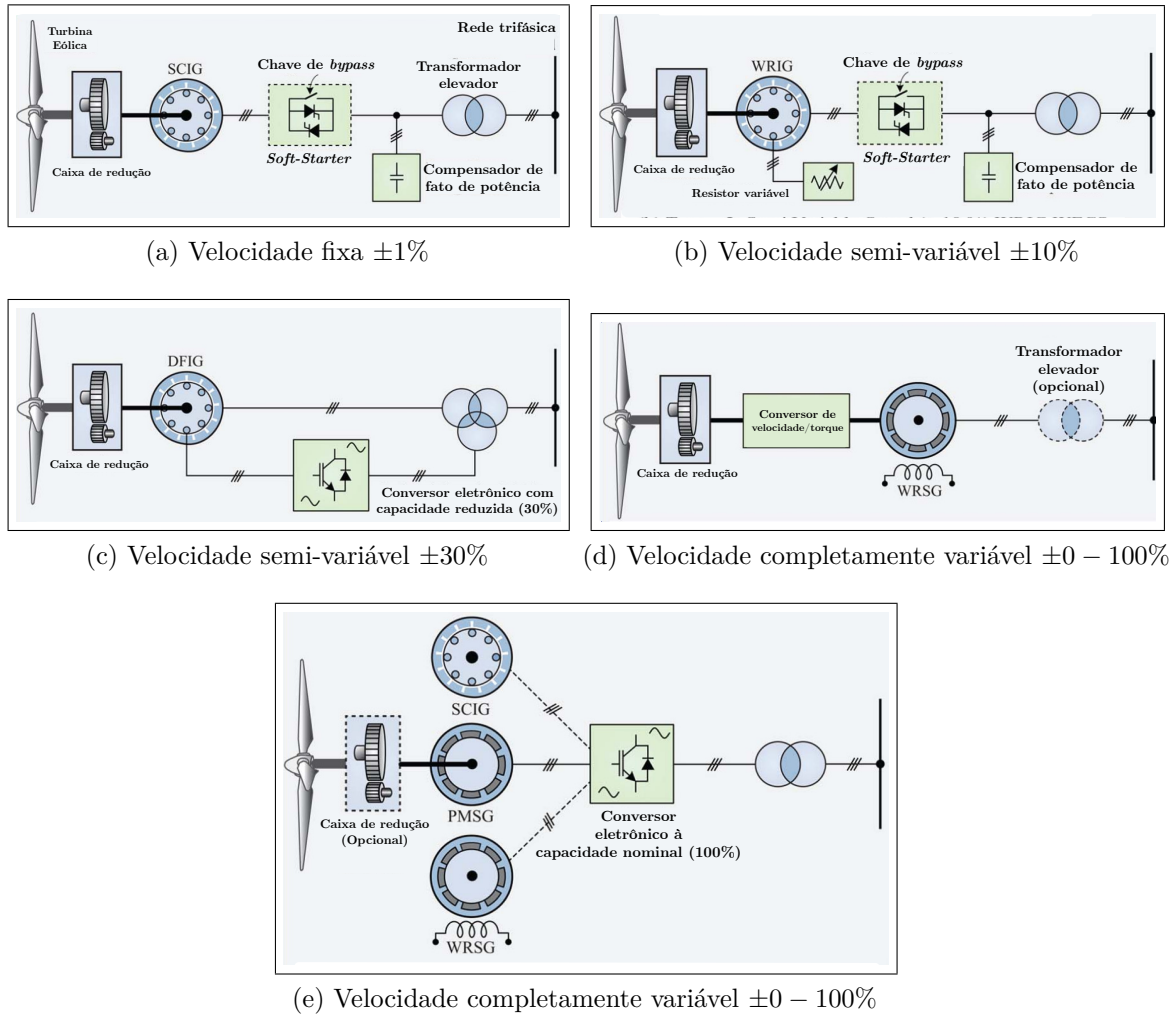


Fonte: Adaptado de [Yaramasu et al. \(2015\)](#).

Contudo, há diferentes tipos de geradores elétricos aplicados em SA. Dentre eles, o SCIG oferece vantagens sobre o Gerador de Indução Duplamente Alimentados (do inglês, *Doubly Fed Induction Generator* - DFIG), o Gerador Síncrono de Imã permanente (do inglês, *Permanent Magnet Synchronous Generator* - PMSG) e ao Gerador de Indução de

Rotor Bobinado (do inglês, *Wound Rotor Induction Generator* - WRIG) aoresentando menor custo de manutenção e instalação, tamanho reduzido, ausência de anéis ou escovas, controle mais simples e maior robustez (MESEMANOLIS; MADEMLIS; KIOSKERIDIS, 2013; ZOU; ELBULUK; SOZER, 2013; YARAMASU et al., 2015). Na Figura 6 são exibidas as diferentes configurações para os geradores supracitados.

Figura 6 – Configurações de geradores elétricos aplicados na aerogeração.



Fonte: Adaptado de Yaramasu et al. (2015).

Quanto as configurações de velocidades de rotação do gerador elétrico, pode-se classificar o SA em velocidades fixas (e.g. Figura 6a), semi-variáveis (e.g. Figuras 6b e 6c), e completamente variáveis (e.g. 6d e 6e) (YARAMASU et al., 2015), (MUDI; SINHA, 2014). Nestes cenários, o SCIG encaixa-se nas aplicações de: (i) velocidades fixas, estas configurações foram as primeiras e estão em desuso por possuírem baixa eficiência e serem susceptíveis às mudanças nos ventos, afetando diretamente na rede elétrica (YARAMASU et al., 2015); e (ii) velocidades completamente variáveis, que utilizam um conversor eletrônico de potência para conexão com a rede elétrica, sendo capaz de tornar o SA mais flexível e proporcionar menor estresse ao sistema, comparado ao de velocidade fixa (HOSSAIN;

ALI, 2015). Um Sistema de Aerogeração com SCIG e velocidade variável possui uma faixa de geração maior que a um DFIG, que utiliza velocidades semi-variáveis entre 25% e 30% (MÜLLER; DEICKE; W., & Rik Doncker, 2002; YARAMASU et al., 2015) a um WRIG, com velocidade semi-variáveis de aproximadamente 10% e não apresenta vantagens a um PMSG (ALNASIR et al., 2014) .

A necessidade de utilização de um conjunto de redução mecânico (do inglês *gearbox*) torna-se obrigatória em um SA com SCIG, em virtude da construção deste gerador. Essa obrigatoriedade pode ser eliminada, ao utilizar um PMSG, com um elevado número de pólos. Com o DFIG ou WRIG a obrigatoriedade é mantida (YARAMASU et al., 2015). Utilizar um PMSG parece vantajoso, pois conjuntos redutores mecânicos proporcionam perdas de rendimento, porém, a escolha de utilizá-lo é baseada em uma solução de compromisso que equipare peso, tamanho e necessidade da aplicação (DUAN; HARLEY, 2009).

A utilização de um conversor eletrônico de potência permite os sistemas de aerogeração operar em velocidade variável, eliminando o uso de uma *soft-starter* e de um compensador de potência reativa (YARAMASU et al., 2015). Além disso, os conversores garantem uma entrega de energia rápida, constante e livre de oscilações, possibilitando maior flexibilidade para estratégias de controle e melhoramento da qualidade energética (POLINDER, 2011). Sua utilização torna-se obrigatória em SA de velocidade variável (do inglês, *full speed range*), que utiliza um SCIG. Entretanto, a adição de mais um componente aumenta a complexidade do sistema e pode torná-lo mais susceptível a falhas. A faixa de potência na qual o conversor deverá operar é de suma importância, visto que o seu alto custo inicial gira em torno de 7 a 12% do custo total do aerogerador, de acordo com Yaramasu et al. (2015).

Quando o SCIG estiver operando em um sistema de velocidade completamente variável, há a necessidade de utilizar conversores eletrônicos na topologia *back-to-back* (YARAMASU et al., 2015). Essa configuração consiste em dois conversores em cascata, sendo um deles responsável pela conversão de corrente alternada (CA) para corrente contínua (CC), com tensão e frequência variáveis no lado do gerador e o outro pela conversão CC/CA, com tensão e frequências fixas ao lado da rede. Essa topologia permite fluxo bidirecional de potência e é necessário serem fabricados dois conversores por gerador, projetados para sua potência nominal. Contudo, no DFIG de velocidade semi-variável o conversor é limitado a 30% da potência do gerador (YARAMASU et al., 2015). Dessa forma, a escolha de utilizar um SCIG como gerador elétrico é pautada na localização, tipo e potência do parque eólico, devendo ser alinhado com a visão empresarial em termos de gestão de viabilidade e retornos. O movimento do mercado eólico tem-se demonstrado aceitador dessa tecnologia, indicando-o inclusive como tendências atuais e futuras.

2.1.1.1 Estado da arte para o SCIG aplicado em aerogeradores

Como evidenciado, a principal desvantagem da geração *offshore* está nos elevados custos de manutenção e transmissão. Portanto, é necessário que, além do SA ser confiável, também possua alta disponibilidade, a fim de minimizar os impactos da manutenção nos custos energéticos, como estimado por [Polinder et al. \(2013\)](#). Estas características são os fatores decisivos quanto a escolha de construção de um parque *offshore* e é em meio a esses aspectos que se encaixa o SCIG, pois a maturidade tecnológica desta máquina elétrica reflete em uma tecnologia robusta e barata.

Em um levantamento apresentado por [Zhang et al. \(2013\)](#), do total de aerogeradores instalados em parques eólicos *offshore*, 48,11% utilizam o SCIG, representando uma vantagem enorme aos PMSG e WRIG, e concorrendo apenas com o DFIG. O ambiente operacional hostil, no qual o gerador elétrico está inserido, torna praticamente inviável o uso de um WRIG mediante essas maiores necessidades de manutenções em bobinamentos e escovas, de acordo com [Yaramasu et al. \(2015\)](#).

Neste cenário, um SCIG com *gearbox* também possui vantagens sobre um PMSG, sem conjunto redutor mecânico (do inglês *gearless*), pois, apesar da inexistência de uma *gearbox*, esta solução torna o PMSG grande, pesado e de difícil instalação ([YARAMASU et al., 2015](#)). Além disso, um estudo realizado por [Tavner, Bussel e Spinato \(2006\)](#) concluiu que a taxa de falha média de um aerogerador não diminuiu ao retirar-se a *gearbox*, contradizendo a percepção de que esse componente mecânico causasse problemas ao sistema. Em geral, um PMSG é mais eficiente que um SCIG, entretanto, um estudo realizado por ([ALNASIR et al., 2014](#)) demonstrou que, em condições de ventos constantes e médios (entre 6-10 m/s), esses geradores apresentaram eficiências similares, e esse representa o cenário de aerogeração *offshore*. Portanto o uso de um PMSG não denota vantagens explícitas para essa aplicação. Além disso, há uma incerteza quanto ao futuro do uso do PMSG, visto que estes geradores utilizam ímãs permanentes, compostos por materiais raros na terra ([ALNASIR et al., 2014](#)). Mediante os fatos supracitados, um SCIG acoplado a um conjunto de redução mecânico, demonstra ser a opção ideal quanto a escolha do gerador elétrico, para o cenário de geração *offshore*.

Apesar de no cenário atual da aerogeração o gerador elétrico mais utilizado ser o DFIG - constituindo mais que 50% dos aerogeradores instalados, [Yaramasu et al. \(2015\)](#) afirmam que o SCIG, com *gearbox* e utilizando conversor eletrônico para operação em *full speed range*, dominarão o mercado nos próximos anos. Há uma tendência para escolha de sistemas de aerogeração que utilizem essa tecnologia em virtude da capacidade de operação em velocidade completamente variável, capaz de melhorar o rendimento e reduzir estresses mecânicos. Um levantamento feito por [Yaramasu et al. \(2015\)](#) exibiu que 90% dos fabricantes de turbinas eólicas do mundo anunciaram futuros projetos baseados nessa tecnologia. É neste cenário que se encaixa a utilização de um SCIG com um conversor

eletrônico *full scale*, capaz de extrair potência elétrica em diferentes regimes de operação dos ventos.

Os parques eólicos *offshore* convergem para esta realidade e estão cada vez mais ganhando a atenção das empresas, pelo fato de permitirem uma capacidade de produção maior e mais estável que *onshore*. Yaramasu et al. (2015) indicam que, em 2020, a capacidade de produção *offshore* deverá ser da ordem de 40 GW, o equivalente a praticamente 5 vezes a capacidade atual. Além disso, uma análise feita pelo Global Wind Energy Council (2015) mostrou que enquanto o mercado *onshore* decaiu 7,8% as instalações *offshore* mais que duplicaram em 2014.

Outro motivo que justifica as turbinas eólicas mais potentes estarem adentro do mar é o fato de o custo da geração de energia poder ser reduzido em 2,5% ao sair-se de aplicações em baixa tensão para média tensão (YARAMASU et al., 2015). Portanto preferem-se aerogeradores maiores do que um grupo de pequenos aerogeradores por serem mais capazes de gerar energia elétrica, tornando os custos de instalação e manutenção menores.

As expectativas futuras destas tecnologias também se refletem nas recentes pesquisas que estudam as problemáticas circundantes desta aplicação e propõem técnicas que permitam difundir a geração *offshore* e o uso do SCIG na aerogeração. Tais problemas são abordados de diversas óticas, tais como: (i) conversores eletrônicos, estudado por Egea-alvarez et al. (2014) e Matos et al. (2015); (ii) teorias de controle, como pesquisados por Cardoso et al. (2016) e Mesemanolis, Mademlis e Kioskeridis (2013); (iii) sistemas de transmissão para geração *offshore*, estudado por Kong e Jia (2011).

Em suma, o uso de SCIG em sistemas de aerogeração de alta potência, principalmente em parques *offshore*, é viável e vantajoso. A comprovação reside na tendência futura para a construção de parques *offshore* e em anúncios feitos pelas maiores empresas que circundam este mercado, as quais afirmam que haverá investimentos e desenvolvimentos nesta tecnologia voltados para aplicações de grande potência.

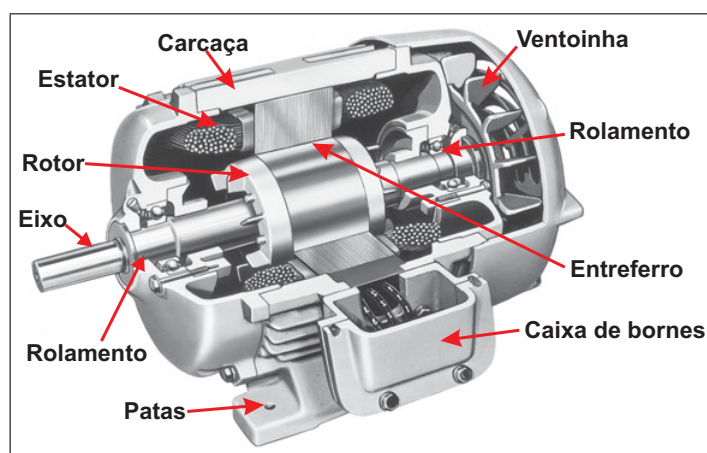
2.1.1.2 Constituição do gerador de indução trifásico

O princípio de funcionamento do gerador de indução trifásico é similar ao do motor de indução trifásico. Portanto, conhecer como a máquina opera na função de motor facilita o entendimento da sua operação.

Segundo FRANCISCO (2006), o motor de indução é constituído basicamente pelo estator e pelo rotor. O estator é fabricado com chapas ferromagnéticas prensadas e isoladas entre si que possuem cavas nas quais são colocados os enrolamentos estatóricos. O conjunto fica alojado no interior de uma carcaça de ferro, aço ou alumínio. Já o rotor, que é a parte móvel da máquina, é constituído por um núcleo feito com chapas ferromagnéticas

prensadas, nas quais se encontram enrolamentos (no caso do rotor bobinado) ou barras condutoras de alumínio ou cobre em paralelo (no caso do rotor gaiola de esquilo). O rotor fica apoiado no veio de rotação do motor, que possui rolamentos nos extremos, transmitindo a energia mecânica produzida à carga. Entre o rotor e o estator existe o entreferro, que deve ser o menor possível a fim de reduzir a relutância magnética total do circuito e, assim, aumentar o fluxo, já que o ar é muito menos permeável do que o ferro. Os elementos que constituem o motor de indução podem ser observados na [Figura 7](#).

Figura 7 – Partes do motor do indução.



Fonte: Adaptado de [CHAPMAN \(2013\)](#).

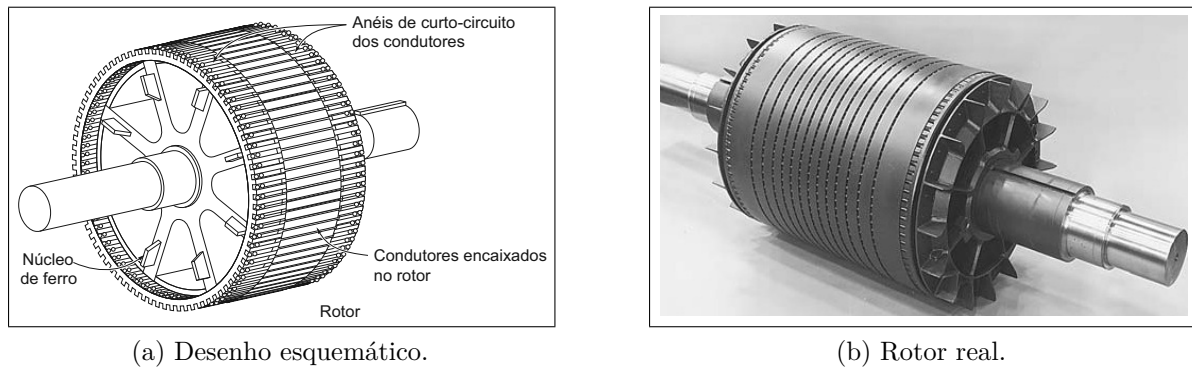
Há dois tipos diferentes de rotores na máquina de indução. Um deles é denominado rotor gaiola de esquilo e o outro é chamado de rotor bobinado ([CHAPMAN, 2013](#)). Este trabalho refere-se apenas ao rotor gaiola de esquilo. Logo, apenas este será abordado. Para mais informações sobre rotores bobinados, ver [FRANCISCO \(2006\)](#), [Fitzgerald \(2006\)](#), [Franchi \(2008\)](#) e [CHAPMAN \(2013\)](#).

O rotor gaiola de esquilo consiste em uma série de barras condutoras que estão encaixadas dentro de ranhuras próximas à superfície do rotor e postas em curto-circuito, em ambas as extremidades por grandes anéis condutores. Essa forma construtiva é conhecida como rotor gaiola de esquilo, porque os condutores, se fossem examinados isoladamente, seriam semelhantes àsquelas rodas nas quais os *hamsters* correm fazendo exercício ([CHAPMAN, 2013](#)). A extrema simplicidade e a robustez da construção em gaiola de esquilo representam vantagens notáveis para esse tipo de motor de indução e, de longe, fazem dele o tipo de motor mais comumente usado, indo desde motores fracionários até os de grande porte ([FITZGERALD, 2006](#)).

O motor com rotor gaiola de esquilo é um motor de uma única alimentação. Não necessita de coletor nem de escovas e, por isso, não possui contatos elétricos móveis. Este

fato tem como resultado um motor praticamente sem manutenção (FRANCISCO, 2006). Um desenho esquemático e uma ilustração de um rotor real podem ser observados nas Figuras 8a e 8b.

Figura 8 – Rotor gaiola de esquilo.



Fonte: CHAPMAN (2013).

2.1.1.3 Princípio de funcionamento

Pelos circuitos dos enrolamentos do estator e do rotor, da máquina de indução, circula apenas corrente alternada. Logo, pode-se compara-la com um transformador, em que o enrolamento secundário está no rotor e recebe energia por indução do primário, que está no estator (FRANCISCO, 2006; FITZGERALD, 2006).

Na máquina elétrica assíncrona trifásica, o bobinamento estatórico é formado por conjuntos de três enrolamentos, instalados de forma que suas forças magnetomotrizs estejam defasadas em um ângulo de 120° entre si. Pela lei de Faraday quando os enrolamentos são percorridos por correntes elétricas defasadas eletricamente de 120° , induz-se campos magnético em fase com a corrente elétrica, ditos como pulsantes. A combinação vetorial de três campos magnéticos pulsantes resulta em um campo magnético girante de frequência igual à da fonte de alimentação. Esse campo, ao atravessar o entreferro, provoca uma variação de fluxo magnético nos condutores do rotor, gerando uma força eletromotriz induzida nesses condutores, também de acordo com a lei de Faraday. Considerando que os condutores do rotor estão em um circuito fechado, os mesmos são percorridos por correntes induzidas. Estas correntes induzidas, de acordo com a lei de Lenz, têm um sentido tal que criam um campo magnético opositor ao campo magnético criado pelo estator. O rotor tenta acompanhar o campo girante produzido pelo estator, entrando, assim, em movimento (FRANCISCO, 2006), mas nunca alcançando a velocidade estatórica. Por isso, dá-se o nome de motor assíncrono (*i.e.* velocidade rotórica fora de sincronia com a estatórica).

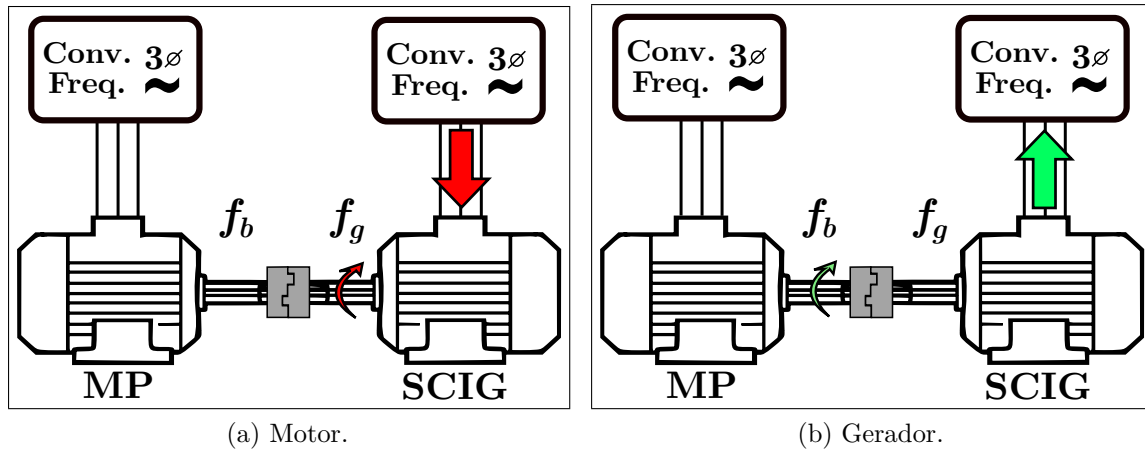
A diferença entre essas velocidade dá origem ao conceito de escorregamento, s , que é positivo quando a máquina funciona como motor, zero quando a máquina atinge a velocidade síncrona e negativo quando opera como gerador.

A frequência do estator é chamada nesse trabalho pela sigla f_g , fazendo alusão à frequência comandada pelo conversor de frequência que aciona o gerador de indução. Enquanto que a sigla f_b simboliza a frequência comandada pelo conversor que aciona a máquina primária, acoplada ao gerador, e faz alusão à rotação rotórica.

Na Figura 9 é exibido um esquemático contendo o gerador de indução gaiola de esquilo, SCIG, e uma Máquina Primária (MP), ambos alimentados por um conversor de frequência trifásica. O conjunto pode funcionar como: (i) motor, Figura 9a, ao f_g ser maior que f_b , nesse momento, o sentido do fluxo de potência é da fonte de alimentação para a máquina de indução, propiciando potência mecânica na ponta do eixo; (ii) gerador, Figura 9b, desde que f_b seja maior que f_g , nessa operação, o fluxo de potência inverte-se e passa a ser no sentido da máquina para fonte de alimentação. É importante salientar que, mesmo operando como gerador, a máquina de indução precisa de uma alimentação estatórica para fins de magnetização da mesma. Quando $f_b = f_g$, não há fluxo de potência e, por consequência, não há geração de energia elétrica.

Não é qualquer diferença entre f_b e f_g que propicia a geração de energia. A diferença deve ser suficiente para que não haja circulação de corrente da rede elétrica para o conversor de frequência do gerador. Esse momento é chamado, neste trabalho, de **primeiro ponto de geração**, pois o SCIG está gerando tanto energia elétrica para o barramento CC do conversor de frequência, quanto reativos necessários para a magnetização do seu estator.

Figura 9 – Diagrama esquemático da operação da máquina elétrica de indução como:



Fonte: Autor.

2.2 Monitoramento das condições de funcionamento

Há uma constante necessidade por redução de custos operacionais e de manutenção em parque eólicos (ASFANI; PURNOMO; SAWITRI, 2013). Segundo Asfani, Purnomo e Sawitri (2013), a forma mais eficiente de se reduzir tais custos é continuamente monitorar

esses sistemas, de forma a prever a degradação dos componentes para otimizar rotinas de manutenção e planejamento do parque eólico, como também prover maior disponibilidade do sistema.

Kusiak e Li (2011) relatam que para tirar proveitos de um sistema de monitoramento das condições de funcionamento de um parque eólico é necessário, além de indicar condições falha e não falha, também indicar severidade e ainda prever futuras falhas. O autor utiliza técnicas baseadas em dados, como redes neurais integradas com informações de um sistema de controle e aquisição de dados (SCADA).

2.2.1 Falhas de curto-circuito em máquinas elétricas de indução

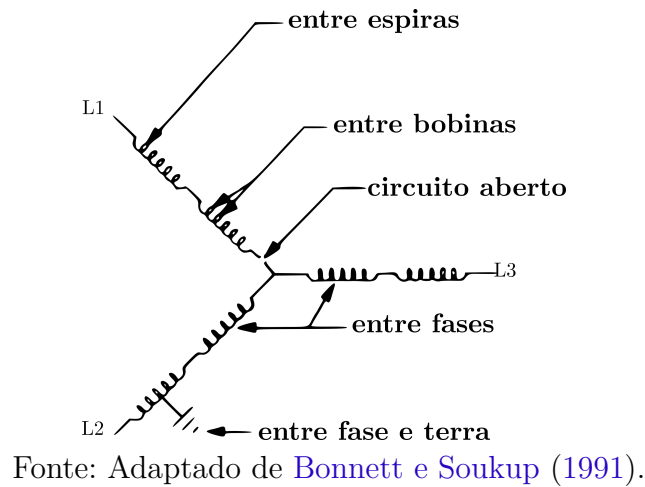
Apesar da versatilidade do SCIG, ele não é imune a falhas, e possui limitações. Em termos gerais, as falhas neste tipo de equipamento estão associadas a diversos fatores em conjunto, dentre eles: sobreaquecimento, efeitos elétricos, dinâmicos e mecânicos (BONNETT; SOUKUP, 1991). A literatura enfatiza que, de uma maneira geral, a combinação de fatores induz a ocorrência, prematura ou não, de falhas na máquina elétrica.

Segundo Bonnett e Soukup (1991), as falhas relacionadas ao bobinamento do estator são de 5 tipos: (i) entre espiras, (ii) entre bobinas, (iii) circuito aberto, (iv) fase-fase e (v) fase-terra, exibidas na Figura 10. Todos os tipos de falhas iniciam-se de uma condição de curto-circuito entre espiras em uma mesma fase. Nessa condição inicia-se a circulação de elevadas correntes na região em curto-circuito que ocasiona uma degradação do isolamento por sobreaquecimento (BONNETT; SOUKUP, 1991). A esse processo é dado o nome de curto-circuito incipiente. Entretanto, esse tipo de falha é geralmente não perceptível, sendo por isso, negligenciada por métodos de proteção. E assim, o curto-circuito continua evoluindo até que o isolamento de área maior seja danificado e cause falhas mais severas, como o curto-circuito entre fase-fase ou entre fase-terra, levando a degradação total da máquina (BONNETT; SOUKUP, 1991; SHARIFI; EBRAHIMI, 2011).

Há, ainda, uma característica determinante para a ocorrência de falhas em bobinamentos de máquina de indução: demonstrou-se que utilizar conversores de frequência para acionamento da máquina pode causar dez vezes mais estresses no bobinamento da máquina (KAUFHOLD et al., 2002). Isso porquê o chaveamento da ordem de microsegundos causa distribuições não lineares de tensão que acarretam em estresses elétricos entre espiras ou na terminação de uma bobina (KAUFHOLD et al., 2002). Esse fato vem de encontro com os sistemas de aerogeração que utilizam SCIG com conversores eletrônicos, como os exibidos na Seção 2.1.

Então, há a necessidade de haver sistemas para detecção prematura de falhas de curto-circuito entre espiras do estator no gerador de indução, visto que Bonnett e Soukup (1991) enfatiza que o curto entre espiras é a falha incipiente e a mais difícil de ser detectada.

Figura 10 – Tipos de curto-circuito no bobinamento estatórico da máquina de indução.



2.2.2 Estado da arte em detecção de falhas de curto-circuito entre espiras estatóricas

Quando se trata de análise de falhas em máquinas elétricas, há métodos relevantes como a análise de vibração ([TORABIZADEH; NOSHADI, 2011](#)). Contudo, a busca na literatura evidenciou a técnica de análise da assinatura de corrente (do inglês, *Motor Current Signature Analysis* - MCSA) como uma técnica estabelecida. Essa consiste em um conjunto de métodos consolidados, que desenvolveram-se principalmente entre os anos de 1975 e 1985 em pesquisas ao redor do mundo, que tinham em comum a análise da corrente e do seu espectro de frequência para caracterizar máquinas de indução ([THOMSON, 2017](#)).

Em base à manutenção centrada em confiabilidade, [Gradin e Burstein \(1993\)](#) propuseram a utilização MCSA como uma técnica não invasiva para identificação de problemas em rolamentos de motores de ventiladores de usinas nucleares. Os autores ratificam que havendo a possibilidade de realizar a leitura do sinal de corrente elétrica de um ambiente controlado, sem necessidade de exposição da planta industrial, pode-se aplicar a técnica e identificar, de maneira operacionalmente segura, problemas no sistema de ventilação.

[Albizu et al. \(2004\)](#) realizam uma revisão sobre métodos para detecção *online* de falhas no estator de geradores de indução aplicados em energias renováveis. O estudo elaborado pelos autores indica que há métodos confiáveis para detecção de falhas em MI de alta tensão. Contudo, as técnicas para baixa tensão ainda estão sendo padronizadas, o que também foi ratificado posteriormente por [Nandi, Toliyat e Li \(2005\)](#). Para máquinas de baixa tensão (abaixo de 600 V), métodos como assinatura de corrente são válidos, porém, a principal desvantagem é a dificuldade em relacionar a assinatura de corrente com a severidade da falha ([ALBIZU et al., 2004](#)).

Royo e Arcega (2007) estuda a aplicação da Transformada de Fourier com o método MCSA em um SCIG submetido à falhas de curto-circuito entre espiras, quebra de barras e com rolamento defeituoso. Os autores utilizam modelos matemáticos para encontrar as componentes de frequência em máquinas submetidas à essas condições. E analisam os espectros em frequência obtidos de ensaios. Utilizam sensores de corrente e velocidade para validar a aplicação da técnica matemática e, mediante os critérios estabelecidos, concluem que é possível discernir características normais dos três tipos de falhas. Porém, a variação de carga no gerador, algumas vezes, induz componentes no espectro de Fourier que podem mascarar as frequências de falhas.

A necessidade por detecção de curtos-circuitos incipientes continuou e atualmente empregam-se técnicas de inteligência computacional para as detecções, como exibido nas pesquisas de Asfani, Purnomo e Sawitri (2013). Os autores obtêm taxas de acertos de 100% para a detecção temporária com 75% de espiras em curto-circuito. As desvantagens dessa metodologia estão na necessidade de utilizar três sensores de corrente elétrica, visto que o autor assume que haverá uma variação na corrente das fases do gerador e, segundo Umans et al. (2014), o circuito elétrico de uma máquina de indução é teoricamente equilibrado, portanto, a inserção de um curto-circuito entre espiras proporciona um desequilíbrio no estator e a percepção desse problema só é conclusiva ao analisar-se as três fases da máquina elétrica. Além disso, o curto incipiente pode ocorrer em uma quantidade pequena de espiras, o que não é contemplado na pesquisa do autor.

Oliveira, Silva e Medeiros (2013) utilizam um único sensor de corrente, sem sensor de velocidade, para identificar falhas de curto-circuito entre espiras a partir de 1.4% em uma máquina elétrica de indução operando como motor. Os autores conseguem identificar 67% das condições de curto. Os autores propuseram o uso da técnica MCSA, combinada com Fourier, a partir da teoria do espectro de frequência de Penman et al. (1994), e Thomson e Fenger (2001). Vieira (2016) continuaram esses estudos, e conseguiram desenvolver um sistema embarcado para detecção dessas falhas.

2.3 Técnicas de aprendizado de máquina

O termo aprendizado de máquina (do inglês, *machine learning*) é especialmente utilizado na âmbito da ciência da computação. Esse termo faz alusão à capacidade computacional de resolver problemas a partir de um conhecimento prévio, ou seja, a partir de dados. É um campo de estudo vasto, que compreende diferentes estudos e aplicações, dentre eles as tarefas de reconhecimento de padrões.

Padrões podem ser interpretados de acordo com a subjetividade do projetista, por exemplo: padrões visuais, temporais ou lógicos. A identificação do sexo de uma pessoa, a partir de uma foto, feita por um computador é um exemplo, nesse caso em específico, é

uma tarefa denominada classificação. Em uma maneira filosófica pode-se dizer que tarefas de reconhecimento de padrões são encontradas em toda atividade inteligente.

A classificação é uma tarefa que consiste em associar dados a uma determinada categoria, como Falha e não Falha de um gerador, por exemplo. Existem diversas técnicas para tratar esse tipo de problema. Uma outra forma de entender no que consiste essa área é fazer uma alusão a problemas matemáticos. Classicamente é comum fazer equacionamentos do problema que se almeja resolver, isso consiste em uma abordagem tradicionalista, dita como determinística. A ideia por detrás do aprendizado de máquina é estocástica, pois ao invés de realizar-se equacionamentos específicos para o problema, pode-se deixar a "máquina" aprender o comportamento a partir dos dados. A versatilidade dessa abordagem é a possibilidade da aplicação de técnicas não lineares para problemas deterministicamente complexos. Entretanto, enfatiza-se, que esse conjunto de técnicas dependem exclusivamente dos dados.

Os problemas de reconhecimento de padrões abordados nessa dissertação serão tratados como problemas de classificação. Antes de se iniciar as fundamentações essa seção, é importante fazer algumas definições preliminares:

- **Atributo:** é uma característica que possa descrever o problema a ser tratado. Por exemplo, altura, peso, quantidade de habitantes, ou ainda *amplitude harmônica da corrente elétrica*. Convenciona-se atributo como sendo a variável p e que o espaço de atributos é chamado de $\mathbb{R}^p$;
- **Amostra:** é a composição de um ou mais atributos que possam descrever aquele tipo de dado. Também chamado de vetor de entrada para as notações matemática, $\mathbf{x}_n$;
- **Rótulo:** valor que define a que grupo (*i.e.* classe C_i) pertence aquela amostra. Por exemplo, Normal e Falha de curto-circuito são duas classes distintas que rotulam todas as amostras do banco de dados. É comum quantizar-se esses rótulo, d_i , durante a formação do banco de dados, *e.g.*, atribuir valor numérico 1 à classe Normal e 2 à classe de Falha;
- **Treinamento e parametrização:** Esses termos fazem alusões ao processo de identificação automática do modelo por parte do método de aprendizado de máquina. Pode ser recursivo, aonde cada amostra produz um sinal de erro que servirá para ajuste ou em lote (do inglês *batch*), na qual o modelo é estimado com todos os dados de uma vez só. Há diferenças operacionais, vantagens e desvantagens entre esses tipos, e serão discutidas adiante;
- **Teste:** é o momento que se apresentam dados não utilizados na etapa de treinamento para avaliar a capacidade do método em generalizar o "conhecimento" obtido.

- Classificador: método ou algoritmo que realiza a classificação da amostra, $\mathbf{x}_j$ como sendo da classe C_i .
- Função discriminante: é a função mapeada pelo classificador para realizar a classificação do vetor de entrada.
- Hiperespaço: é o espaço n -dimensional de projeção dos dados de entrada.
- Hiperparâmetros: são parâmetros inerentes a cada um dos classificadores que precisam ser definidos *a priori* do treinamento. É comum nessa área serem adotadas passos, *i.e.*, heurísticas, para a seleção desses valores.

2.3.1 Classificadores lineares

Nesse tipo de classificadores tenta-se obter um mapeamento linear entre os dados de entrada (*i.e.* vetores de atributos $\mathbf{x}_n$), e os dados de saída correspondentes (*i.e.* vetores de rótulos $\mathbf{d}_n$). É comum adicionar um termo de viés b (*bias*, do inglês) aos vetores de entrada do problema. Tal mapeamento é representado pela equação

$$\mathbf{d}_n = \mathbf{W}\mathbf{x}_n \quad (2.1)$$

na qual a matriz $\mathbf{W}$ é obtida através dos dados por técnicas mostradas posteriormente.

A identificação da classe, C_j , é feita ao selecionar a posição do máximo valor da multiplicação expressa na Equação 2.2. Essa é a operação da função discriminante, traduzida para equação

$$C_j = \operatorname{argmax}_{n=1,\dots,N} (\mathbf{d}_n). \quad (2.2)$$

Diz-se que a matriz $\mathbf{W}$ contém o conhecimento do comportamento dos dados pois, utilizam critérios de otimização para minimizar os erros durante processos de treinamento ou parametrização dos classificadores. Alguma das formas de encontrar a matriz $\mathbf{W}$ é através do algoritmo dos mínimos quadrados ordinário ou do perceptron simples.

2.3.1.1 Mínimos Quadrados Ordinário

Primeiramente, admite-se que o agrupamento dos vetores de atributos x_n irão constituir a matriz dos dados $\mathbf{X}$ e o agrupamento dos rótulos d_n a matriz $\mathbf{D}$. Se a matriz $\mathbf{X}$ for quadrada (*i.e.* o número de amostrar é igual a $p+1$, quantidade de atributos junto do intercepto ou *bias*) pode-se simplesmente realizar a seguinte operação:

$$\mathbf{W} = \mathbf{D}\mathbf{X}^{-1}. \quad (2.3)$$

Entretanto, não é comum haver tão poucos dados quanto atributos. Logo, uma contorno para isso é utilizar a matriz pseudo-inversa, que dá origem ao mínimos quadrados ordinários (do inglês *Ordinary Least Square* - OLS). A Equação 2.3 passa a ser,

$$\mathbf{W} = \mathbf{D}\mathbf{X}^T(\mathbf{X}\mathbf{X}^T)^{-1}, \quad (2.4)$$

e $\mathbf{X}^T(\mathbf{X}\mathbf{X}^T)^{-1}$ é chamada de pseudo-inversa de $\mathbf{X}$.

É comum ter-se problemas de mal condicionamento com a inversão de matrizes, para isso contornar essa problemática há técnicas como a regularização de Tikhonov que leva a Equação 2.3 à

$$\mathbf{W} = \mathbf{D}\mathbf{X}^T(\mathbf{X}\mathbf{X}^T + l\mathbf{I})^{-1}, \quad (2.5)$$

em que l é a constante de regularização e $\mathbf{I}$ a matriz identidade de dimensões $p + 1 \times p + 1$.

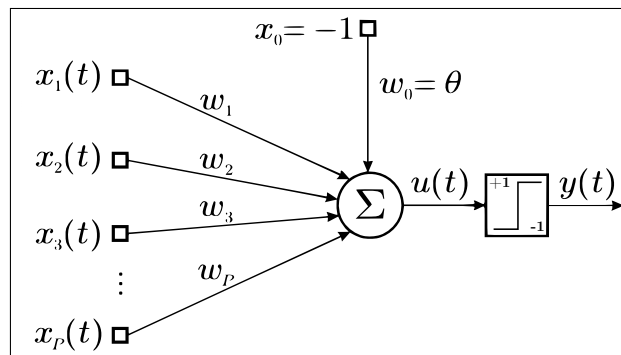
É importante salientar que a estimativa da matriz $\mathbf{W}$ é feita em modo *batch*, o que torna a estimação dos parâmetros desse classificador rápida.

2.3.1.2 Perceptron simples

Uma das diferenças entre esse classificador e o OLS é apenas a regra de aprendizagem que torna-o recursivo. Uma interpretação para esse classificador é vê-lo como um modelo de neurônio artificial. As linhas da matriz $\mathbf{W}$ são os pesos sinápticos $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_p$, que conecta cada atributo, $x_1, x_2, \dots, x_p$, do vetor de dados $\mathbf{x}_n$. Há novamente um termo de bias, w_0 , e outra diferença é função de ativação do neurônio, $y(t)$, expressa por

$$y(t) = \text{sign}(u(t)) = \text{sign}(\mathbf{w}_n \mathbf{x}_n) \quad (2.6)$$

Figura 11 – Topologia do neurônio de McCulloch e Pitts (1943).



Fonte: Autor.

Esse modelo é baseado no neurônio de McCulloch e Pitts (1943) e proposto inicialmente por Rosenblatt (1958). Diferente do OLS a estimativa da matriz de pesos é

feita recursivamente e baseada no erro da apresentação de amostras ao Perceptron. Isso dar-se por meio da **regra de aprendizagem**.

Baseada em critérios de otimizações a equação de aprendizagem do Perceptron é expressa por

$$\mathbf{w}(t+1) = \mathbf{w}(t) + \Delta\mathbf{w}(t) = \mathbf{w}(t) + \alpha\mathbf{e}(t)\mathbf{w}(t), \quad (2.7)$$

em que $1 > \alpha > 0$ é chamada de taxa de aprendizagem e $\mathbf{e}(t) = \mathbf{d}(t) - \mathbf{y}(t)$.

É importante salientar que um perceptron é capaz de separar o hiperespaço em dois, portanto para um problema de classificação com o número de classes acima de dois pode-se fazer combinações com perceptrons. Dos quais, cada unidade será responsável por criar a função discriminantes daquela determinada classe.

2.3.2 Classificadores Gaussianos

Sabe-se que uso desses classificadores foram proposto por [Fisher \(1922a\)](#), baseado na regra de Bayes. Apesar da proposição, esses são principalmente conhecido como classificadores Bayesianos, e as alusões feitas à esse classificador neste trabalho serão feitas dessa forma.

São ditos como métodos paramétricos, ou seja, não realizam aprendizado recursivo baseado em erro. Os classificadores Gaussianos utilizam cálculos probabilísticos entre as classes presente nos dados para classificar novas amostras como pertencentes à classe mais provável ([HAYKIN, 2001](#)). Levando em consideração as diferentes quantidades de amostras por classe e diferentes valores de variância das mesmas, a teoria de decisão estatística proposta por [Fisher \(1922b\)](#) maximiza a probabilidade a posteriori $P(c|x)$ e classifica x como pertencente à classe c_i conforme

$$\mathbf{x} \in c_i \text{ se } P(c_i|\mathbf{x}) > P(c_j|\mathbf{x}), \text{ para todo } j \neq i. \quad (2.8)$$

A matemática pode detrás dos classificadores Bayesianos assumem que os dados são variáveis aleatórias, ou seja, se comportam mediante uma função de densidade de probabilidade Gaussiana. Porém, este classificador pode ser empregado indiscriminadamente em relação a quaisquer dados.

Usando a regra de Bayes pode-se calcular a probabilidade a posteriori a partir da probabilidade a priori $P(c_i)$ e da função densidade de probabilidade $p(\mathbf{x}|c_i)$ das classes, e normalizada pela probabilidade $P(\mathbf{x})$ como sendo

$$P(c_i|\mathbf{x}) = \frac{p(\mathbf{x}|c_i)P(c_i)}{P(\mathbf{x})}. \quad (2.9)$$

Considerando a função densidade de probabilidade $p(\mathbf{x}|c_i)$ como Gaussiana, a Equação 2.9 é escrita como

$$P(c_i|\mathbf{x}) = \frac{P(c_i)}{P(\mathbf{x})} \left[\frac{1}{\sigma_i \sqrt{2\pi}} \exp \left(-\frac{(\mathbf{x} - \mu_i^2)}{2\sigma_i^2} \right) \right]. \quad (2.10)$$

Para o caso onde os dados são D dimensionais,

$$P(c_i|\mathbf{x}) = \frac{P(c_i)}{P(\mathbf{x})} \left[\frac{1}{|\Sigma_i|^{\frac{1}{2}} (2\pi)^{\frac{D}{2}}} \exp \left(-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}_i)^T \Sigma_i^{-1} (\mathbf{x} - \boldsymbol{\mu}_i) \right) \right]. \quad (2.11)$$

onde Σ_i é a matriz de covariância da classe c_i . Manipulando a equação 2.11 encontre-se a equação das funções discriminantes $g_i(\mathbf{x})$ dada por

$$g_i(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \Sigma_i^{-1} \mathbf{x} + \boldsymbol{\mu}_i^T \Sigma_i^{-1} \mathbf{x} + \left[\ln(P(c_i)) - \frac{1}{2} \ln(|\Sigma_i|) - \frac{1}{2} \boldsymbol{\mu}_i^T \Sigma_i^{-1} \boldsymbol{\mu}_i \right]. \quad (2.12)$$

em que $\mathbf{x}$ pertence à classe c_i se $g_i(\mathbf{x}) > g_j(\mathbf{x})$.

Pode-se observar que a Equação 2.12 se trata de um discriminante quadrático da forma $A\mathbf{x}^2 + B\mathbf{x} + C$, em que $A = \frac{1}{2}\Sigma^{-1}$, $B = \boldsymbol{\mu}_i^T \Sigma_i^{-1} \mathbf{x}$ e $C = \left[\ln(P(c_i)) - \frac{1}{2} \ln(|\Sigma_i|) - \frac{1}{2} \boldsymbol{\mu}_i^T \Sigma_i^{-1} \boldsymbol{\mu}_i \right]$. Assim, percebe-se que a matriz de covariância dos atributos é de extrema importância para o classificador. Diversas são as formas que a matriz de covariância pode assumir, e algumas destas são apresentadas a seguir.

2.3.2.1 Matrizes de covariância

A covariância é um momento conjunto de segunda ordem de variáveis aleatórias, centrados nas respectivas médias e mede o grau de interdependência ou interrelação numérica linear entre elas. A covariância σ_{xy} pode ser calculada como

$$\sigma_{xy} = cov(x, y) = \frac{1}{N} \left[\sum_{i=1}^N x_i y_i - \frac{1}{N} \left(\sum_{i=1}^N x_i \right) \left(\sum_{i=1}^N y_i \right) \right]. \quad (2.13)$$

A matriz de covariância Σ_i é obtida de acordo com

$$\Sigma_i = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1m} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{m1} & \sigma_{m2} & \cdots & \sigma_m^2 \end{pmatrix}, \quad (2.14)$$

em que m é o número de atributos pertencentes à classe c_i . Alguns detalhes sobre a matriz de covariância: é quadrada e simétrica, isso significa que $\Sigma_i^T = \Sigma_i$; na diagonal principal, posição i, j , para $i = j$, estão as variâncias individuais, σ^2 , de cada atributo no conjunto de dados; nos elementos de posição i, j , para $i \neq j$, estão as covariância entre os atributos x_i e x_j , calculados pela Equação 2.13 e $\sigma_{ij} = \sigma_{ji}$.

- Matriz de covariância independente por classes

Assumindo que as classes possuam matrizes de covariância diferentes, o discriminante é dito quadrático, como já apresentado na [Equação 2.12](#). Porém, esta abordagem pode reduzir o desempenho do classificador quando há poucas amostras por classes ([MARKS; DUNN, 1974](#)). Visando reduzir este efeito, algumas outras imposições podem ser feitas à matriz de covariância, como seguem.

- Matriz de covariância geral para o conjunto de dados

Assume-se que todos os dados pertençam à mesma classe, logo a matriz de covariância Σ_{geral} será a mesma para todas as classes e a [Equação 2.12](#) é reduzida para

$$g_i(x) = \boldsymbol{\mu}_i^T \Sigma_{geral}^{-1} \mathbf{x} + \left[\ln(P(c_i)) - \frac{1}{2} \boldsymbol{\mu}_i^T \Sigma_{geral}^{-1} \boldsymbol{\mu}_i \right], \quad (2.15)$$

uma vez que ao se comparar $g_i(x)$ e $g_j(x)$, os termos $\frac{1}{2} \mathbf{x}^T \Sigma_{geral}^{-1} \mathbf{x}$ e $\frac{1}{2} \ln(|\Sigma_{geral}|)$ são os mesmos em ambos os lados da inequação. Assim, a matriz de covariância geral torna o discriminante linear, na forma $Ax + b$.

- Matriz de covariância agregada

A matriz de covariância agregada (*pooled covariance matrix*) promove a junção de todas as amostras na mesma classe com o intuito de se obter uma matriz única diferente da matriz geral, pois a matriz agregada Σ_{pool} combina as classes de forma ponderada da seguinte forma,

$$\Sigma_{pool} = \sum_{i=0}^C \frac{n_i}{n} \Sigma_i, \quad (2.16)$$

em que n_i é o número de amostras da classe c_i , $n = \sum_{i=1}^C n_i$ é o número total de amostras e C é o número de classes. Assim, a [Equação 2.12](#) também se torna um discriminante linear e pode ser escrita como

$$g_i(\mathbf{x}) = \boldsymbol{\mu}_i^T \Sigma_{pool}^{-1} \mathbf{x} + \left[\ln(P(c_i)) - \frac{1}{2} \boldsymbol{\mu}_i^T \Sigma_{pool}^{-1} \boldsymbol{\mu}_i \right]. \quad (2.17)$$

- Matriz de covariância diagonal

Quando utiliza-se essa matriz assume-se que todos os atributos são estatisticamente independentes, ou seja não há covariância entre si, $\sigma_{ij} = 0$ para todo $i \neq j$, como em

$$\Sigma_{diagonal} = \begin{pmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_m^2 \end{pmatrix}. \quad (2.18)$$

Esta abordagem é utilizada pelo classificador Naive-Bayes, e há aplicações na literatura desse classificador para diversos fins. Inclusive, para detecção de falha em motores de indução como em [Asfani, Purnomo e Sawitri \(2013\)](#). Assim a Equação 2.12 é dada por

$$g_i(\mathbf{x}) = \boldsymbol{\mu}_i^T \Sigma_{diagonal}^{-1} \mathbf{x} + \left[\ln(P(c_i)) - \frac{1}{2} \boldsymbol{\mu}_i^T \Sigma_{diagonal}^{-1} \boldsymbol{\mu}_i \right]. \quad (2.19)$$

Deve-se ter cuidado com essa abordagem, pois assumir que não há dependência estatísticas entre os atributos é uma premissa forte. Ressalta-se, também, que a operação desse classificar continua sendo linear, de acordo com seu discriminante, na Equação 2.19.

- Matriz de covariância igual a identidade

Um caso particular da matriz de covariância diagonal acontece quando a variância dos atributos são iguais, como em

$$\Sigma_{identidade} = \sigma^2 \mathbf{I}_D = \begin{pmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{pmatrix}. \quad (2.20)$$

em que $\mathbf{I}_D$ é a matriz identidade de dimensão $D \times D$. Assim, a Equação 2.12 pode ser escrita como

$$g_i(\mathbf{x}) = \frac{1}{\sigma^2} \boldsymbol{\mu}_i^T \mathbf{x} + \left[\ln(P(c_i)) - \frac{1}{2\sigma^2} \boldsymbol{\mu}_i^T \boldsymbol{\mu}_i \right]. \quad (2.21)$$

- Matriz de covariância proposta por Friedman

[Friedman \(1988\)](#) propôs um método de evitar a redução do desempenho do discriminante quadrático quando o mesmo tenta classificar classes com poucas amostras. Seu método fornece uma matriz de covariância Σ_i^λ composta a partir da combinação linear de Σ_i com a matriz agregada Σ_{pool} , como sendo

$$\Sigma_i^\lambda = \frac{(1 - \lambda)n_i \Sigma_i + \lambda n \Sigma_{pool}}{(1 - \lambda)n_i + \lambda n} \quad (2.22)$$

em que $0 \leq \lambda \leq 1$ é o coeficiente de Friedman. Assim, a [Equação 2.12](#) é escrita como

$$g_i(\mathbf{x}) = \boldsymbol{\mu}_i^T (\Sigma_i^\lambda)^{-1} \mathbf{x} + \left[\ln(P(c_i)) - \frac{1}{2} \boldsymbol{\mu}_i^T (\Sigma_i^\lambda)^{-1} \boldsymbol{\mu}_i \right]. \quad (2.23)$$

Percebe-se então que para $\lambda = 0$, a matriz Σ_i^λ iguala-se à Σ_i , enquanto que para $\lambda = 1$, a mesma se iguala à Σ_{pool} . Assim, tem-se que a matriz de covariância proposta por Friedman confere ao classificador um discriminante presente entre o discriminante linear utilizando a matriz Σ_{pool} e o discriminante quadrático.

2.3.3 Redes Neurais Artificiais

Uma rede neural é um processador maciçamente paralelamente distribuído constituído de unidades de processamento simples, que têm a propensão natural para armazenar conhecimento experimental e torná-lo disponível para uso. Ela se assemelha ao cérebro pelo fato de que o conhecimento é adquirido pela rede a partir de um processo de aprendizagem e de que esse conhecimento é armazenado em conexões sinápticas ([HAYKIN, 2001](#)).

De acordo com [Zhang \(2000\)](#), as redes neurais têm emergido como uma importante ferramenta para a classificação, revelando-se como uma alternativa promissora aos vários métodos de classificação convencionais. Por serem modelos não lineares, elas podem ajustar-se aos dados sem qualquer especificação explícita, modelando relações complexas do mundo real.

Uma rede neural supervisionada começa ignorante. Seu professor, que pode ser um humano, um programa de computador ou outra rede neural mais madura, a ensina. Esse *feedback* é usado pela rede aprendiz para ajustar seus pesos sinápticos para, com o tempo, a rede apresentar respostas corretas sem supervisão.

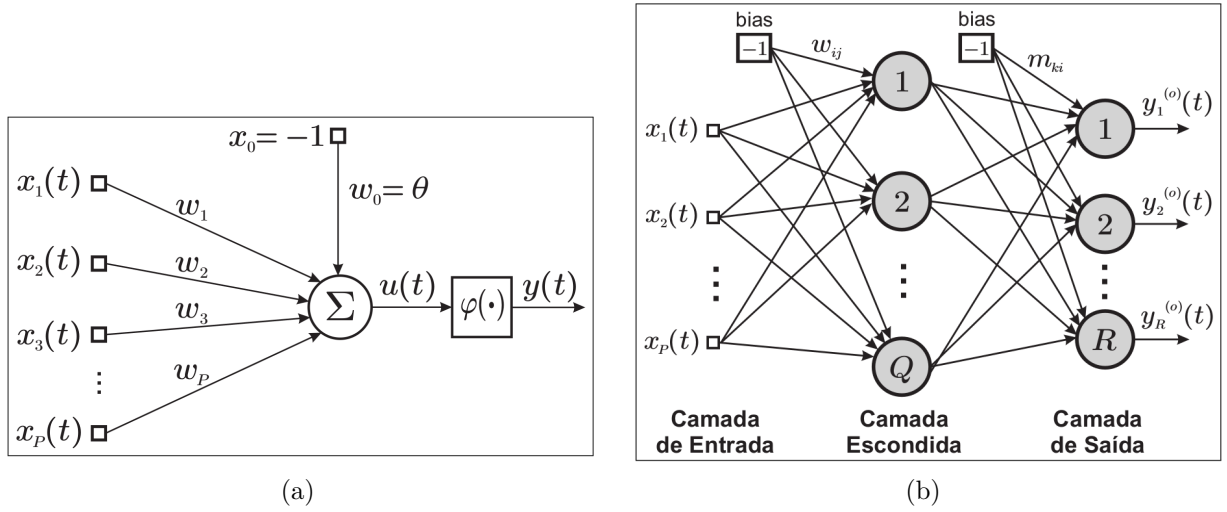
Entre os vários tipos de redes neurais, a rede Perceptron de Múltiplas Camadas (*Multilayer Perceptron* - MLP) tem se mostrado uma ferramenta poderosa na modelagem de mapeamentos entrada-saída tipicamente encontrados em problemas de aproximação de funções (regressão) e classificação de padrões ([MEDEIROS; BARRETO, 2015](#)).

2.3.3.1 Perceptron de Múltiplas Camadas

O perceptron multicamadas é constituído por um sistema de neurônios simples interconectados, representando um mapeamento não linear entre um vetor de entrada e um vetor de saída. Os neurônios estão ligados por pesos (sinapses) e sinais de saída, que são a soma das entradas ponderadas pelos pesos e modificadas por uma função de ativação ([HAYKIN, 2001](#)). A rede MLP é ilustrada na [12](#), e um de seus neurônios na [12a](#), chamado de Perceptron.

O vetor $x(t)$ é o vetor de entradas da rede neural, enquanto $y^{(o)}(t)$ é o vetor que representa a saída do mesmo. A função de ativação do neurônio é representada por $\varphi(\cdot)$

Figura 12 – Arquitetura da rede neural do tipo MLP.



Fonte: 12a: Autor; 12b: Adaptado de Sousa (2017).

que geralmente assume, dentre outras funções possíveis, a função tangente hiperbólica ou a função sigmóide, por suas altas não linearidades. Os pesos da camada oculta são dados por w_{ij} e conectam a j -ésima entrada ao i -ésimo neurônio da camada escondida, enquanto os pesos da camada de saída são dados por m_{ki} . O elemento *bias* é um elemento de controle do neurônio. As representações matemáticas das saídas dos neurônios da camada escondida e de saída no passo t são apresentadas nas Equações 2.24 e 2.25, respectivamente,

$$y_i^{(h)}(t) = \varphi \left[u_i^{(h)}(t) \right] = \varphi_i \left[\sum_{j=0}^P w_{ij}(t) x_j(t) \right], \quad i = 1, \dots, Q \quad (2.24)$$

e

$$y_k^{(o)}(t) = \varphi \left[u_k^{(o)}(t) \right] = \varphi_k \left[\sum_{i=0}^Q m_{ki}(t) y_i^{(h)}(t) \right], \quad k = 1, \dots, R. \quad (2.25)$$

A saída da rede MLP é representada na Equação 2.26 a seguir

$$y_k^{(o)}(t) = \varphi_k \left\{ \sum_{i=0}^Q m_{ki}(t) \varphi_i \left[\sum_{j=0}^P w_{ij}(t) x_j(t) \right] \right\}, \quad k = 1, \dots, R. \quad (2.26)$$

O treinamento da MLP é supervisionado, ou seja, cada entrada vem acompanhada de uma resposta desejada e então, os pesos sinápticos da rede são ajustados de forma que a saída seja a mais próxima possível da desejada (HAYKIN, 2001).

O aprendizado a partir de exemplos, embora seja uma ideia intuitiva, não pode ser tratada como uma questão banal. Existem algoritmos bastante sofisticados sendo utilizados para realizar a adaptação dos pesos do modelo neural. A implementação de alguns destes algoritmos requer conhecimentos sólidos na área de otimização não-linear (MEDEIROS, 2014). Um desses algoritmos, o *back propagation* ou gradiente descendente, é brevemente

descrito a seguir para o treinamento de uma MLP com apenas uma camada escondida. Maiores detalhes são encontrados em Haykin (2001).

2.3.3.1.1 Algoritmo do *back propagation*

Após a apresentação de um dado $x(t)$ à rede, é necessário calcular a(s) saída(s) $y_k^{(o)}(t)$ segundo a Equação 2.26, fazendo o que se chama de sentido direto da informação, no qual os pesos sinápticos permanecem fixos. Em seguida, é feito o sentido reverso onde os pesos são ajustados pela retro propagação dos sinais de erro de saída através da camada de saída e escondida, até atingir a camada de entrada. Para isso, é necessário calcular o valor do erro $e_k^{(o)}(t)$ gerado por cada neurônio de saída como

$$e_k^{(o)}(t) = d_k(t) - y_k^{(o)}(t), \quad k = 1, \dots, R, \quad (2.27)$$

em que $d_k(t)$ é o valor desejado para a saída do k -ésimo neurônio da camada de saída (MEDEIROS; BARRETO, 2015). Para o ajustes dos pesos da camada de saída (m_{ki}), é necessário que a informação de erro calculada na Equação 2.27 seja retro propagada na função de ativação φ_k . A isso dá-se o nome de gradiente local (δ) e pode ser calculado da seguinte forma,

$$\delta_k^{(o)}(t) = \varphi'_k \left[u_k^{(o)}(t) \right] e_k^{(o)}(t), \quad (2.28)$$

em que φ'_k é a derivada da função de ativação. De uma forma semelhante, pode-se encontrar o gradiente local para a camada escondida $\delta_i^{(h)}$ como a seguir,

$$\delta_i^{(h)}(t) = \varphi'_i \left[u_i^{(h)}(t) \right] e_i^{(h)}(t) = \varphi'_i \left[u_i^{(h)}(t) \right] \sum_{k=1}^R m_{ki}(t) \delta_k^{(o)}(t), \quad (2.29)$$

em que $e_i^{(h)}(t)$ representa o erro projetado do i -ésimo neurônio da camada escondida.

Assim, os pesos da camada oculta e da camada de saída podem ser ajustados de acordo com as Equações 2.30 e 2.31, respectivamente,

$$w_{ij}(t+1) = w_{ij}(t) + \eta \delta_i^{(h)}(t) x_j(t) \quad (2.30)$$

e

$$m_{ki}(t+1) = m_{ki}(t) + \eta \delta_k^{(o)}(t) y_i^{(h)}(t), \quad (2.31)$$

em que η é a taxa de aprendizagem, a qual deve estar dentro do intervalo de 0 a 1.

Uma variação do algoritmo *back propagation* é obtida com a simples adição do termo de *momentum*, pois basta inserir um único parâmetro nas Equações 2.30 e 2.31 visando ponderar o quão os pesos sinápticos foram alterados entre duas iterações anteriores e sucessivas. Desse modo, os pesos da camada escondida e de saída são alterados como sendo

$$w_{ij}(t+1) = w_{ij}(t) + \eta \delta_i^{(h)}(t) x_j(t) + \alpha (w_{ij}(t) - w_{ij}(t-1)) \quad (2.32)$$

e

$$m_{ki}(t+1) = m_{ki}(t) + \eta \delta_k^{(o)}(t) y_i^{(h)}(t) + \alpha (m_{ki}(t) - m_{ki}(t-1)) , \quad (2.33)$$

em que α é definida como a taxa de *momentum* e seu valor está compreendido entre 0 e 1.

Cada apresentação das N amostras de treinamento é chamada de época. No entanto, se o número de amostras não for suficiente para uma classificação satisfatória, os dados são novamente apresentados à rede, época por época, de forma aleatória para que a busca da solução ótima, no espaço de pesos, seja aproximadamente estocástica.

Uma maneira simples de avaliar a convergência do algoritmo é através do erro quadrático médio, dado por

$$\varepsilon_{QM} = \frac{1}{2N} \sum_{t=1}^N \sum_{k=1}^R [e_k^{(o)}(t)]^2 , \quad (2.34)$$

calculado após cada época. Se o valor obtido está abaixo de um patamar especificado, então a convergência é atingida.

Medeiros (2014) recomenda que antes da rede ser aplicada para solucionar um determinado problema, a mesma deve ser submetida a uma avaliação do seu desempenho sobre um conjunto de dados contendo amostras jamais apresentadas anteriormente à rede, chamado conjunto de teste, avaliando assim a capacidade de generalização da rede.

2.3.4 Estado da arte de aplicações em máquinas rotativas

Em termos de aprendizado de máquina o campo de reconhecimento de padrões visa a identificação automática de regularidades em dados, a partir do uso de algoritmos que possam descrever o comportamento dos dados e realizar tarefas de classificação (BISHOP, 2013). Há diversos métodos para diferentes aplicações na literatura e a vantagem da utilização de métodos de aprendizado de máquina reside na possibilidade de adaptação à necessidade do problema em questão.

As redes neurais artificiais tem sido utilizadas para a detecção de problemas em caixas de engrenagens de aerogeradores por Yang, Li e Wang (2008). A utilização de 9 sensores de vibração instalados no gerador, em conjunto com decomposições de Wavelet provê capacidade de auto-adaptação e identificação de problemas no sistema mecânico de engrenagens (YANG; LI; WANG, 2008).

Ju et al. (2011) realiza diagnóstico do *gearbox*, geradores e sistema hidráulico de duas turbinas instaladas em um parque eólico, combinando informações provenientes do vento, potência gerada, temperatura de diversas partes do equipamento e parâmetros de operação do equipamento e foi capaz de identificar 80% das condições defeituosas. Contudo, Ju et al. (2011) expressa que para poder generalizar a utilização da técnica é necessário maior quantidade de dados. Outras metodologias para análises generalistas em turbinas

eólicas foram apresentadas em [An, Jiang e Li \(2010\)](#), [Manrique, Giraldo e Esmeral \(2012\)](#), levando em consideração problemas gerais na turbina, sem forçar-se em um componente específico, como o gerador elétrico.

Já a pesquisa de [Toma, Capocchi e Capolino \(2013\)](#) realiza um estudo para implementação de uma rede neural embarcada para identificação das falhas de curto-circuito entre espiras de um WRIG. O autor exhibe um estudo preliminar sobre a implementação em um processador do tipo FPGA. Pesquisas como essa denotam que há uma mobilidade na comunidade científica em desenvolver-se sistemas de prognóstico de falhas para geradores de indução.

Por outro lado, [Asfani, Purnomo e Sawitri \(2013\)](#) utiliza outras técnicas bastante populares na literatura, os métodos Bayesianos. Nesse caso a ideia é a identificação temporário de curto-circuito, exibindo ser capaz de identificar desde que se utilizem um sensor de corrente para cada fase do gerador e um sensor de velocidade na ponta do eixo da máquina.

Tomando de mãos da técnica MCSA, [Oliveira, Silva e Medeiros \(2013\)](#) utiliza as harmônicas de curto-circuito segundo a equação de [Penman et al. \(1994\)](#) para treinar uma rede neural do tipo perceptron de multi camadas para identificar falhas no motor de indução gaiola de esquilo. O autor realiza, também, um estudo de relevância das harmônicas exibindo que nem todas as retiradas da equação de [Penman et al. \(1994\)](#) foram relevantes para a rede neural. A metodologia para o treinamento e validação da rede é exibida em [Oliveira \(2014\)](#) e atingiu-se acerto médio de falhas acima de 80%, identificando 100% dos caso normais. Os erros cometidos pela rede neural proposta por [Oliveira \(2014\)](#) foram entre níveis de curto-circuito, mas se pensar que o objetivo é a detecção do curto incipiente, ainda em formação, a rede obteve êxito de 100% em sua tarefa.

O trabalho de [Oliveira \(2014\)](#) motivou a pesquisa de [Coelho et al. \(2014\)](#), que propõe o uso de outras técnicas, como variantes das rede neural MLP, a ELM e MLM, e ainda técnicas baseadas na teoria do aprendizado estatístico como as máquinas de vetores suporte e suas variantes. Sua metodologia é exposta em [Coelho \(2015\)](#) e o autor consegue resultados da ordem de 98% de acerto. [Vieira \(2016\)](#), também, utilizou a pesquisa de [Oliveira \(2014\)](#) para projetar um sistema embarcado. Esse sistema foi baseado em processadores digitais de sinais, para identificar as falhas de curto-circuito entre espiras. Os autores obtiveram resultados que consideraram satisfatórios.

Em suma, a técnica de MCSA, combinada com as harmônicas de [Penman et al. \(1994\)](#), e com os métodos de aprendizado de máquina proporcionam uma abordagem promissora para detecção prematura de falhas de curto-circuito em máquinas de indução acionadas por conversores de frequência. A seguir são apresentadas explicações sucinta das técnicas de processamento de sinais utilizadas como extrator de atributos.

2.4 Técnicas de extração de atributos em sinais

A etapa de extração de atributos consiste em identificar informações presentes em um sinal. Essas informações, chamadas de atributos, podem ser quaisquer uma que descrevam quantitativamente ou qualitativamente um problema. Por exemplo, a média salarial dos parlamentares, ou ainda, variância estatística de uma sinal de vibração mecânica, ou até, amplitude da harmônica fundamental em um sinal periódico. É uma etapa extremamente importante na área de aprendizado de máquina, pois é uma das etapas de pré-processamento dos dados. A seguir, estão uma breve explicação da técnicas utilizadas nesse trabalho.

2.4.1 A transformada de Fourier à luz da teoria de [Penman et al. \(1994\)](#)

[Penman et al. \(1994\)](#) estudou a ocorrência de curtos-circuitos entre espiras de bobinamentos estatóricos de motores e desenvolveu uma modelagem matemática que relaciona a dispersão de fluxo magnético e as harmônicas mais importantes nos sinais de corrente. O modelo proposto pelos autores baseia-se na equação:

$$f_{cc} = [k \pm n(1 - s)/p]f_0, \quad (2.35)$$

na qual f_{cc} são as harmônicas em função do curto-circuito entre espiras; $k = 1, 2, 3, \dots$, é a ordem das harmônicas temporais; $n = 1, 2, 3, \dots$ é ordem das harmônicas espaciais; s é o escorregamento do motor; p é o número de pares de pólos e f_0 é a frequência fundamental da tensão de alimentação.

A Equação 2.35 é dependente do escorregamento, que por sua vez é dependente da rotação do motor pela equação,

$$s = \frac{n_s - n_r}{n_s}, \quad (2.36)$$

sendo n_s a rotação síncrona e n_r a rotação do rotor da máquina. O uso dessas equações acarreta a necessidade de medição de velocidade no eixo da máquina para estimar as f_{cc} . Isso pode não ser usual para aplicações em motores de baixa potência, visto que os sensores de velocidade podem ser tão caros quanto essa máquina. Porém, [Oliveira \(2014\)](#) contornou essa problemática. O autor admite que o motor esteja funcionando com velocidade nominal e sem carga e, a partir dos dados fornecidos pelo fabricante da máquina, ele estima as componentes de frequência por meio da Equação 2.35. [Oliveira \(2014\)](#) evidenciou que que as componentes relativas à falha de sua máquina aparecem a cada $0,5f_0$. O autor obteve êxito em identificar as falhas utilizando as harmônicas propostas por [Penman et al. \(1994\)](#), validando a aplicação do método para detecção de curto-circuito entre espiras do bobinamento estatórico.

2.4.2 O algoritmo de Goertzel

O algoritmo de [Goertzel \(1958\)](#) foi proposto como uma forma para calcular os coeficientes da transformada discreta de Fourier. É um algoritmo recursivo que realiza operações aritméticas com um único valor a cada interação, facilitando o cálculo de coeficientes (*i.e.* harmônicas) singulares do sinal. A vantagem da utilização do algoritmo é que ele é numericamente mais eficiente para computar valores individuais das harmônicas, o que pode tornar seu uso viável em aplicações embarcadas.

2.4.3 A Estatística de Alta Ordem (HOS)

O uso de Curtose (*Kurtosis*, do inglês) foi primeiramente proposto por [Dwyer \(1983\)](#) como uma ferramenta estatística para indicar componentes não-Gaussianas em um sinal. Porém, em anos posteriores, a teoria foi reformulada por [V. Capdevielle e Lacoume \(1996\)](#) à luz da Estatística de Alta Ordem (HOS, do inglês *Higher-Order Statistics*) e descrita a partir do momento estatístico normalizado de quarta ordem, aplicáveis apenas a sinais estacionários. Entretanto, proposições feitas por [Antoni \(2006\)](#) já preconizam eficácia do método e capacidade de caracterizar sinais não estacionários.

A busca na literatura evidenciou, em combinação com a Curtose, o uso do momento estatístico normalizado de terceira ordem, chamado de Assimetria (*Skewness*, do inglês). Evidências da utilização da assimetria são encontradas em [Martin e Honarvar \(1995\)](#), para auxílio na análise vibracional de rolamentos, e por [B e KR \(2003\)](#) em conjunto com técnicas de aprendizado de máquina. A vantagem de utilização de HOS em sinais reside na insensibilidade a ruídos do tipo gaussiano ao utilizar momentos de ordem superior a dois ([MENDEL, 1991](#)). Tal afirmativa foi posteriormente confirmada por [Fucai Li et al. \(2008\)](#).

As fórmulas utilizadas para cálculo da Curtose, K , e Assimetria, γ , são respectivamente,

$$K(\mathbf{X}) = E \left[\left(\frac{\mathbf{X} - \mu}{\sigma} \right)^4 \right] e \quad (2.37)$$

$$\gamma(\mathbf{X}) = E \left[\left(\frac{\mathbf{X} - \mu}{\sigma} \right)^3 \right], \quad (2.38)$$

nas quais μ é a média e σ é o desvio padrão.

Além desses dois valores a aplicação de HOS nos sinais utilizados nesse trabalho extraiu, também, variância e valor quadrático médio (*rms* - do inglês, *root mean square*).

2.4.4 Matriz de Coocorrência Estrutural

A Matriz de Coocorrência Estrutural (SCM, do inglês *Structural Co-occurrence Matrix*) consiste em um método e análises estruturais de sinais discretos, baseado em estatísticas de coocorrência. A premissa desse método é de que há conexões entre estruturas de baixo nível de dois sinais discretos em n -dimensionais (RAMALHO et al., 2016). Sua principal característica é introduzir um conhecimento prévio sobre os sinais analisados, enfatizando a detecção de detalhes. Sua saída é um histograma bidimensional, em que SCM fornece as co-ocorrências entre as estruturas dos sinais de entrada (RAMALHO et al., 2016). O sinal de entrada representado será comparado com outro sinal, que corresponde ao sinal de entrada após a aplicação de um filtro pré-selecionado. Após esta comparação, é gerada uma matriz com as diferenças estruturais entre os sinais. Então, do histograma bidimensional são retirados atributos, propostos por Ramalho et al. (2016). Na literatura, o uso de SCM foi relatado apenas em tarefas de classificação de imagens, como em Marinho et al. (2017). Nos sinais do domínio do tempo, não há relatórios sobre o uso do SCM como técnica de extração de recursos. (RAMALHO et al., 2016) propôs 6 atributos a serem extraídos de qualquer sinal e os cálculos são mostrados em seu artigo.

2.5 Contribuições da pesquisa

A revisão bibliográfica realizada possibilitou identificar o espaço no qual a presente pesquisa está inserida, pois contribui com embasamentos teóricos e tendências de estudos atuais. Mediante isso, as diretrizes desse trabalho são listadas a seguir.

- O trabalho está alinhado com os planos mundiais para o desenvolvimento da energia eólica, como também para o Nordeste brasileiro, em especial para o Ceará que é extremamente representativo para a matriz energética brasileira em termos de energia eólica.
- Além disso, a escolha do gerador de indução gaiola de esquilo está pautada nas projeções futuras feitas por Yaramasu et al. (2015), e também está alinhada com as pesquisas de Lima et al. (2015). Então, isso torna o presente trabalho inédito para região, pois proverá um estudo que poderá ser utilizado pelos gestores dos parques eólicos locais para otimizar rotinas de manutenção e reduzir custos operacionais em geradores elétricos.
- Será possível também detectar as falhas utilizando apenas um único sensor de corrente e sem empregar o sensor de velocidade do gerador, em contrapartida aos trabalhos evidenciados na literatura. Isso tornará nosso sistema apto a trabalhar em geradores alimentados por conversores de frequência, que já possuem embarcado em sua eletrônica um sensor de corrente elétrica. A solução proposta poderá ser

disponibilizada para empresas que desenvolvem conversores de frequência para aerogeradores.

- A análise dos dados proposta possibilitará a detecção prematura de falhas no SCIG, indicando ainda a severidade do defeito na máquina elétrica, refutando a afirmativa de Albizu et al. (2004) de que a utilização da técnica MCSA é válida para identificar anormalidades, porém, carente para indicar intensidades.

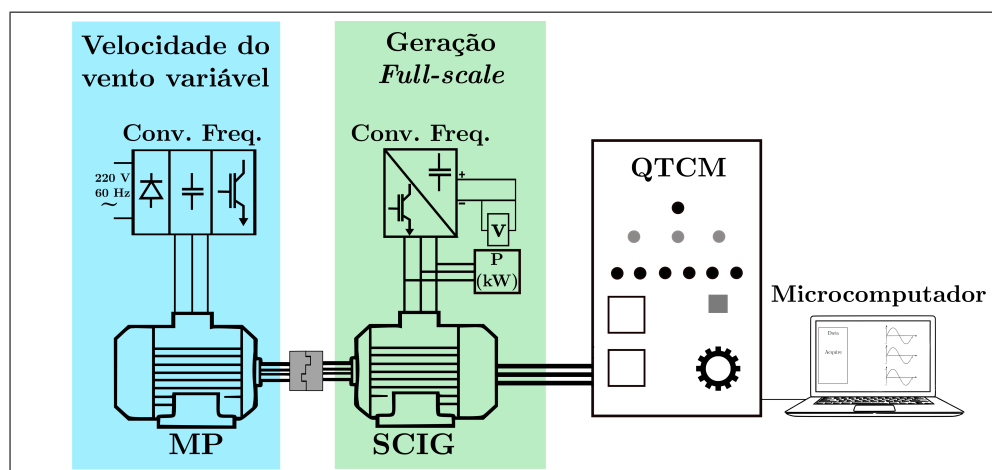
3 Experimentos e aquisições de dados no SCIG

O aerogerador simulado neste trabalho é baseado na configuração descrita por [Yaramasu et al. \(2015\)](#), que consiste em um gerador de indução gaiola de esquilo, do tipo *Full-scale* e *Full-Variable-Speed*, ou seja, a máquina elétrica está apta a gerar energia elétrica em toda a faixa de operação de velocidade. Na Figura 13 é exibido o diagrama do sistema de emulação do aerogerador.

Para realizar experimentos, utilizou-se um gerador de indução do tipo gaiola de esquilo (SCIG) com as seguintes características: 4 pólos, potência mecânica de 1 CV, eletricamente ligado em configuração de delta para tensão de alimentação de 220 V, com corrente nominal de 3 A. As bobinas estatóricas são alimentadas por um conversor de frequência trifásico com alimentação monofásica, CFW-08 da marca WEG. Como Máquina Primária (MP), acoplada mecanicamente ao gerador, utilizou-se um motor de indução com as mesmas características e ligações, e acionado por conversor de frequência.

O conversor de frequência que aciona a MP faz o papel do vento e irá propiciar velocidade variáveis a MP, lado marcado em azul na Figura 13. A máquina primária faz o papel das pás da turbina eólica, provendo movimento rotativo ao eixo do gerador, que é marcado pelo lado verde na mesma figura.

Figura 13 – Diagrama do sistema de emulação do aerogerador.



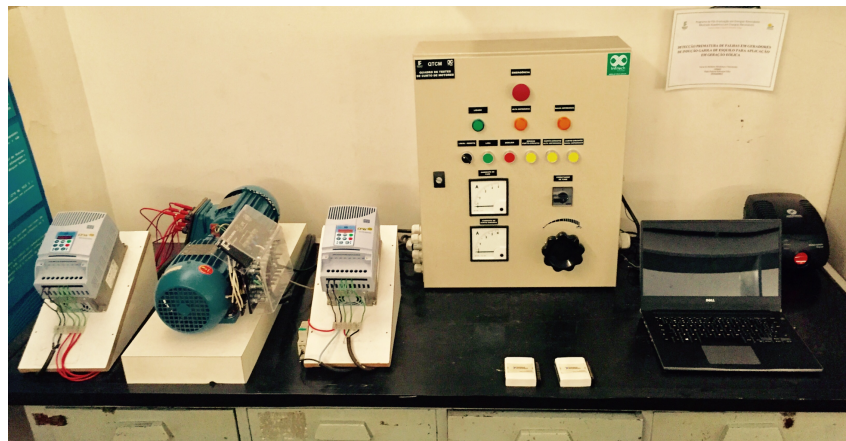
Fonte: Autor.

O SCIG utilizado no sistema de aerogeração emulado foi devidamente modificado para facilitar a inserção de curto-circuito entre as espiras. Os detalhes sobre esse procedimento serão expostos na Seção 3.2. Os experimentos são comandados por um Quadro de Teste de Curto em Máquinas (QTCM), cuja a função é realizar ligações elétricas que

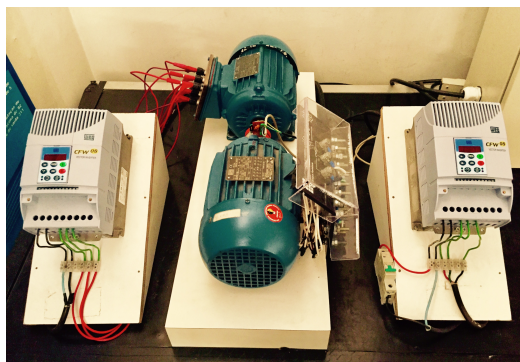
emulem o processo de curto-circuito entre espiras do bobinamento estatórico. Há também módulos para aquisição e envio de dados a um microcomputador, que possui uma interface gráfica para gerenciamento dos teste.

Na Figura 14 são exibidas as instalações aonde foram realizados os experimentos. Na Figura 14a exibe-se uma vista completada da bancada de experimentos. Na Figura 14b estão as instalação das máquinas e conversores de frequência, e em 14c o QTCM. Esse quadro foi projetado para provê energização ao circuito dos motores e conversores, comandar as emulações de curto-circuito, descritas na Subseção 3.2, e realizar o sensoriamento de sinais do gerador. Os detalhes construtivos são exibidos no Apêndice A. Nas instalações estão dispostas também os módulos de aquisição de dados NI-USB-6009 da *National Instruments* e o microcomputador, onde roda a interface do LabVIEW para aquisição de dados.

Figura 14 – Instalações da bancada de emulação do sistema de aerogeração.



(a) Bancada de experimentos.



(b) emulação do aerogerador.



(c) QTCM.

Fonte: Autor.

Os dados coletados com o sistema proposto conterão informações do SCIG operando em diferentes regimes de geração, submetidos a falhas ou não. Isso possibilitará criar uma base de dados em que será aplicada posteriormente técnicas de aprendizado de máquina

para detecção prematura das falhas.

3.1 Descrição do processo de emulação do conjunto turbina eólica/gerador de indução

Nesta seção são descritos os passos para a realização dos experimentos no sistema de aerogeração emulado. As explicações sobre o sistema são baseadas no mesmo descrito na Seção 2.1.1.3. Deve-se lembrar que f_b é valor de frequência comandado pelo conversor da MP e f_g do SCIG. Os passos para realização respeitam a seguinte ordem:

1. Ajustar as rampas de aceleração e desaceleração dos dois conversores para o mesmo valor. Isso evitará que no momento da partida hajam sobrecorrentes em algumas das máquinas elétrica;
2. Ajusta-se $f_b = f_g$ para garantir que, após a rampa de partida, ambas as máquinas estejam na mesma velocidade;
3. Dar o comando de partida, simultaneamente, em ambos os conversores de frequência;
4. Quando o conjunto estiver em regime permanente de funcionamento reduz gradativamente f_g até atingir o primeiro ponto de geração (*i.e.*, ver Seção 2.1.1.3).
5. Desliga-se a alimentação da rede elétrica no conversor do SCIG;
6. Após o desligamento da alimentação da rede elétrica a tensão no barramento de corrente contínua (CC) do conversor de frequência do gerador deverá cair. Ajustar f_g até atingir 210 V.
7. Realizar aquisição dos dados, através da interface LabVIEW;
8. Reduz f_g até f_{max} , até que uma das seguintes situações aconteça:
 - O gerador atinga a corrente nominal;
 - A tensão do barramento CC do conversor, V_{cc} , atinja o valor máximo permitido, pelo equipamento, de 380 V;
 - Seja atingida a corrente nominal na máquina primária, I_{mp} ;
9. Repetir o passo 7;
10. Dar o comando de desligamento, simultaneamente, em ambos os conversores de frequência e esperar o conjunto mecânico ficar totalmente parado;
11. Definir um novos valores pra f_b e f_g ;

12. Repetir o passo 1.

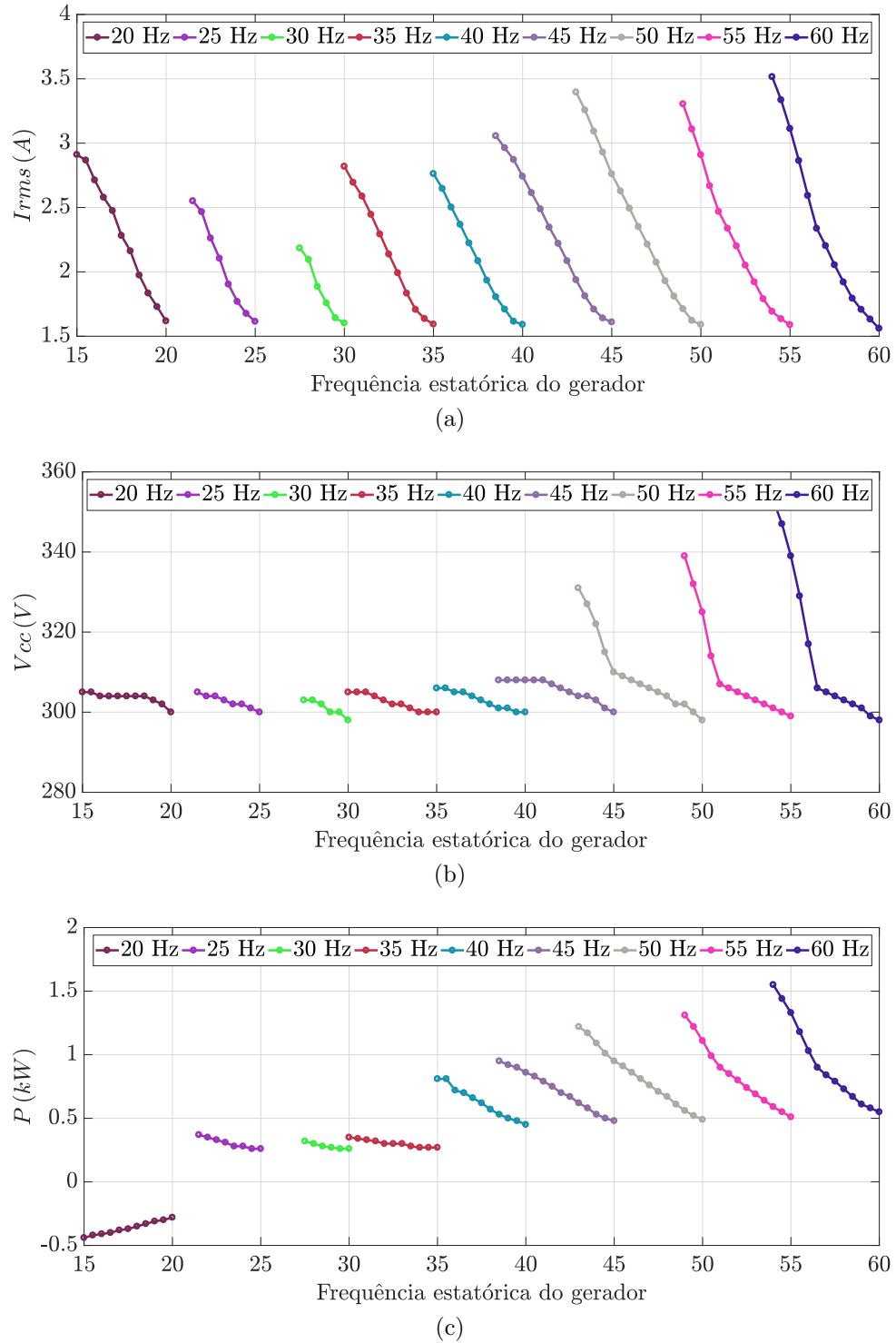
Como a frequência nominal de placa do gerador de indução é 60 Hz , esse é o primeiro valor de f_g e f_b utilizado nos experimentos. Nos experimentos realizados nesse trabalho, evidenciou-se que a geração ocorre a partir de uma diferença positiva de aproximadamente $0,5\text{ Hz}$ entre f_b e f_g . As aquisições e medições dos dados são feitas a cada intervalo do passo número 7.

Com o intuito de familiarizar-se com o problema foram realizados ensaios preliminares no gerador, adotando intervalos de 5 Hz para f_b e para cada um destes, intervalos de $0,5\text{ Hz}$ a partir do primeiro ponto de geração. Na Figura 15 exibem-se as curvas de corrente elétrica na fase R, tensão no barramento CC, e potência medida no gerador. A análise da tensão do barramento CC na Figura 15b exibe que a geração máxima ocorre com $f_b = 60\text{ Hz}$ e f_g entre 55 e 60 Hz , porém, até 45 Hz ainda há geração elétrica efetiva, desde que haja diferença de frequência entre f_b e f_g . Ainda há geração após 45 Hz , mas a um custo de diferenças de frequência elevadas o suficiente para ultrapassar a corrente nominal da máquina primária.

A potência elétrica do gerador é exibida na Figura 15c e identificou-se que a partir de 35 Hz a geração torna-se pequena em comparação com a região nominal, e mesmo com uma diferença de 5 Hz entre f_b e f_g o gerador não atingiu a corrente nominal, mas o ensaio foi interrompido pois a corrente nominal havia sido atingida na máquina primária. A partir de $f_b = 25\text{ Hz}$ não há mais fluxo de potência do SCIG para o barramento CC e a potência gerada é suficiente apenas para manter a energia reativa para magnetização da máquina elétrica.

Estes ensaios preliminares possibilitaram o conhecimento do funcionamento em regime permanente do gerador de indução. Portanto, na próxima seção são explicadas as modificações feitas no SCIG, para, então, realizar-se aquisições das condições de funcionamento da máquina.

Figura 15 – Ensaios preliminares para a caracterização do regime permanente do gerador de indução. Em 15a é o comportamento da corrente elétrica, em 15b da tensão no barramento CC e em 15c da potência trifásica gerada.

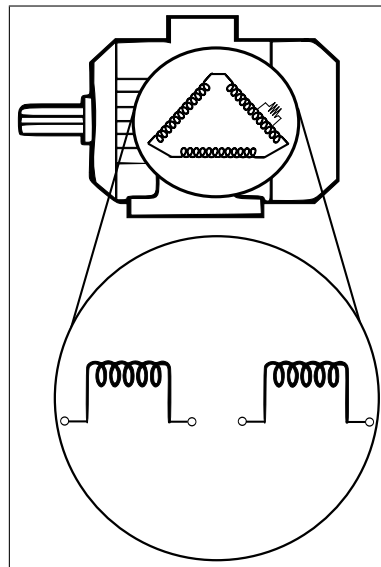


Fonte: Autor.

3.2 O protótipo do gerador

Para realizar os ensaios de curto-circuito, a máquina elétrica utilizada nos experimentos foi devidamente preparada por uma empresa especializada, pois sofreu processo de rebobinamento do seu circuito estatórico, tornando acessíveis as ramificações das bobinas por meio de terminais elétricos que permitiram a injeções de curto-circuito entre espiras. A preparação da máquina foi realizada de uma forma que pudessem serem feitas emulações da evolução do curto-circuito, desde sua ocorrência incipiente, quase imperceptível, até situações severas de pré-degradação do bobinamento estatórico. A mesma máquina foi utilizada também por [Oliveira, Silva e Medeiros \(2013\)](#) em seus experimentos para detecção de falhas em motores elétricos, seguindo os mesmos critérios para emulação das falhas. Um visão geral das alterações feitas são exibidas na Figura 16.

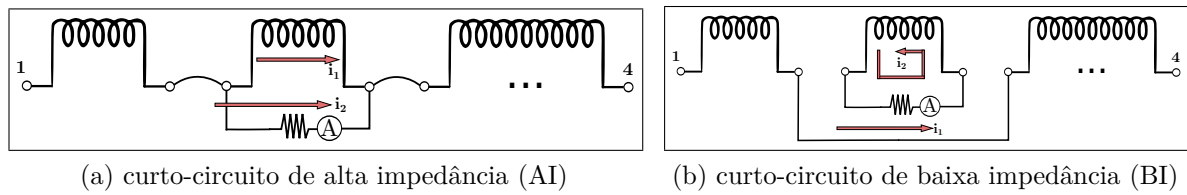
Figura 16 – Detalhe da inserção do curto-circuito entre espiras do bobinamento do estator.



Fonte: Autor.

O esquema de ligações das bobinas para inserção do curto-circuito é exibido na Figura 17. Quanto ao tipo, os curto-circuitos se classificam como (i) Alta Impedância (AI), pois há um caminho paralelo criado para a circulação de corrente elétrica, Figura 17a. Esse estado indica o curto incipiente, em momentos iniciais, pois parte do isolante elétrico entre espiras está se deteriorando, criando um ramo secundário para passagem da corrente elétrica e sofrendo sobreaquecimento até que haja a ruptura total do isolante. A ruptura total do isolamento é emulada como exibido na Figura 17b. Uma quantidade de bobinas é retirada do circuito elétrico da máquina e colocadas em curto-circuito pleno por meio de comandos elétricos. Pela lei de Faraday, uma tensão é induzida nos terminais dessa bobina, proveniente no campo magnético do restante das bobinas que permaneceram alimentadas. Esse comportamento é análogo à um ramo secundário de um transformador, e dar-se o nome a esse processo de curto-circuito de (ii) Baixa Impedância (BI).

Figura 17 – Esquema de ligações para os dois tipos de curto-circuito



Fonte: Autor.

É importante enfatizar que em todos os ensaios a corrente de curto é limitada ao valor da corrente nominal da máquina, por meio de uma resistência variável, cujo ajuste é feito mediante aferições em um amperímetro analógico. Isto se dá porque em situações reais de curto-circuito a corrente nas bobinas pode subir excessivamente ao ponto de danificar a máquina. Isso tornaria o ensaio destrutivo. Portanto, a emulação de BI e AI feita na máquina são, na verdade, um curto-circuito de baixa impedância, mas visto que um dos objetivos desse trabalho é a detecção incipiente de curto-circuitos, espera-se que se for possível identificar o curto BI emulado também será possível detectá-lo em condições reais.

Pode-se também escolher o curto-circuito que deseja-se emular em termos de intensidades, que representa quantidade de bobinas submetida à falha. Isso pode ser feito em três níveis, que consistem em 1,41% (1), 4,81% (2) e 9,26% (3) do total de espiras, respectivamente. Os textos seguintes deste trabalho farão referências às falhas de curto-circuitos pelas siglas AI-1, AI-2, AI-3, BI-1, BI-2 e BI-3. A primeira parte da sigla representa o tipo, baixa ou alta impedância, e a segunda parte a quantidade de espiras em falha. Por fim, há a possibilidade de inserção de 6 níveis diferentes de falhas no SCIG, resumidos na Tabela 1. Há também uma definição de uma escala evolutiva de severidade.

O grau de severidade do curto-circuito, estabelecido na Tabela 1, é uma premissa para detecção do curto incipiente: identifica-lo em situações iniciais. Apesar de Albizu et al. (2004) indicar que com a técnica MCSA não é possível identificar a severidade de curto-circuito entre espiras, acredita-se que com a metodologia proposta nesse trabalho seja possível refutar essa afirmativa, exibindo que é possível identificar curtos incipientes.

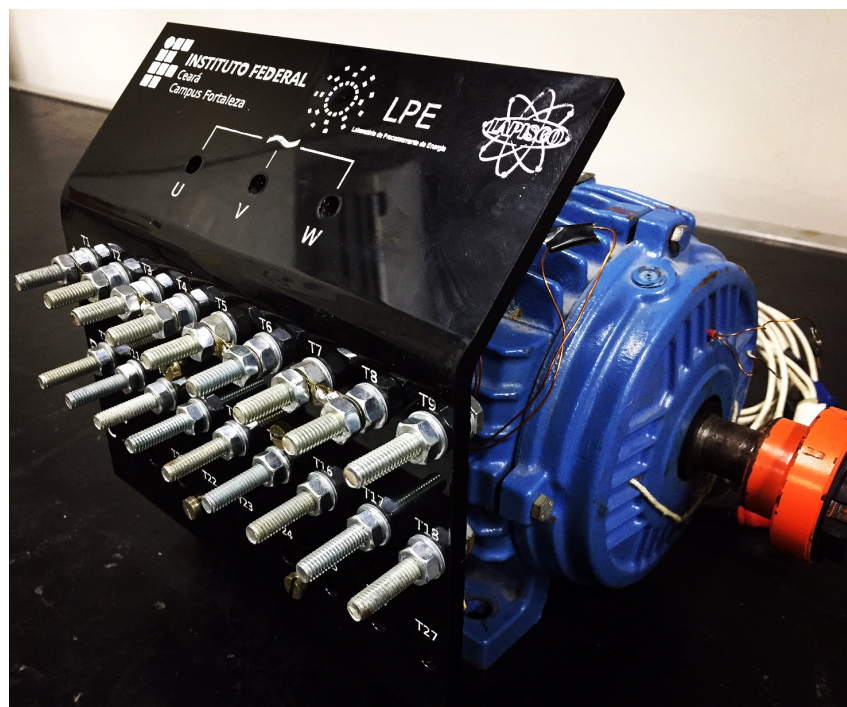
A emulação da falha quanto a número de espiras é feita por meio dos terminais da máquina, exibidos na Figura 18. A escolha do tipo de curto é feita com o auxílio do QTCM e os detalhes do esquema elétrico estão no Apêndice A.

Tabela 1 – Resumo dos estados da máquina considerando Alta Impedância (AI), Baixa Impedância (BI), com os níveis de intensidade equivalente aos percentuais de espiras em curto de 1.41%, 4.81% e 9.26%, totalizando 6 níveis de falha.

Tipos de curto	% de espiras	Graus de severidade
AI	1.41	1
	4.81	2
	9.26	3
BI	1.41	4
	4.81	5
	9.26	6

Fonte: Autor.

Figura 18 – Foto da máquina elétrica preparada para inserção de curto-circuito entre espiras.



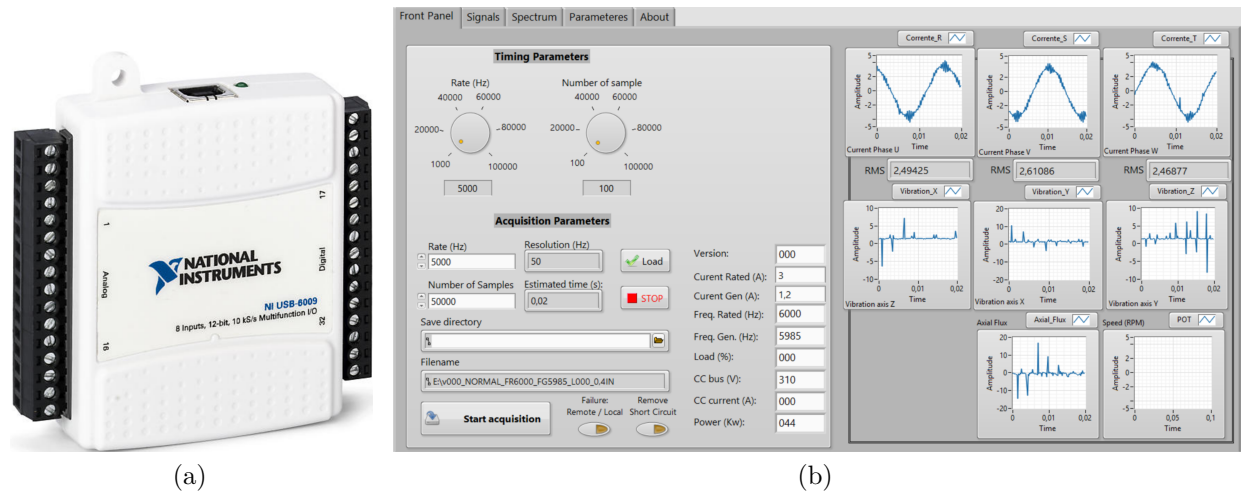
Fonte: Autor.

3.3 Sistema de monitoramento

Para realizar as aquisições utilizou-se uma interface gráfica desenvolvida no *software* LabVIEW® da *National Instruments*, especificamente para esta aplicação. O dois módulos de aquisições de dados utilizado foi o NI-USB-6009, de mesma marca. Esse módulo possui resolução de 14 bits ([National Instruments, 2015](#)), 8 entradas analógicas, sendo 4 diferenciais. O módulo foi programado para realizar 10 segundos de aquisição a uma taxa de amostragem

de 5 kHz . Na Figura 19b é exibida a tela da interface gráfica de aquisição de dados, que possui opções para visualização em tempo real dos sinais provenientes do gerador, por meio do módulo exposto na Figura 19a.

Figura 19 – Foto do módulo NI-USB-6009 em 19a e da interface no LabVIEW em 19b.



Fonte: 19a: Adaptado de (National Instruments, 2015); 19b: Autor.

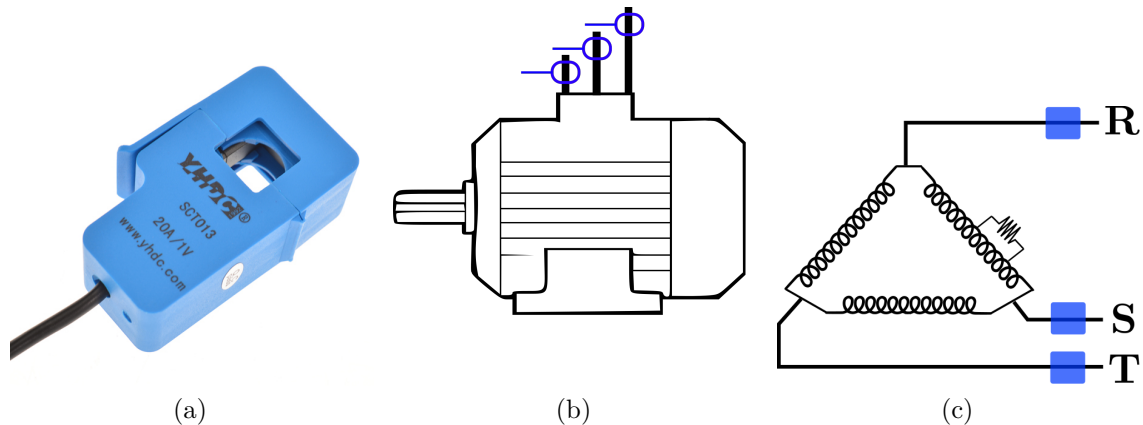
Para realizar o monitoramento instalaram-se 3 sensores de corrente elétrica, 1 de vibração e 1 de fluxo magnético axial no gerador de indução. Suas características são descritas a seguir.

3.3.1 Sensoriamento de corrente elétrica

O sensoriamento de corrente em máquina elétrica é uma técnica efetiva para análise de monitoramento da condição de funcionamento, como apontado por Sharifi e Ebrahimi (2011). O uso dos sensores de corrente é justificado pois a ocorrência do curto-circuito entre espiras causa assimetrias no circuito elétrico da máquina (SHARIFI; EBRAHIMI, 2011). Morsalin et al. (2014) também exibem um estudo que utiliza a corrente elétrica de um motor de indução, junto com informações de tensão e velocidade, para classificar condições normais e falhas de curto-circuito entre espiras. Dessa forma, escolheu-se por realizar aquisição dos sinais da corrente elétrica das três fases do gerador de indução. Utilizou-se sensores de corrente do tipo Transformador de Corrente (TC). Na Figura 20 é exibido o sensor de corrente utilizado e o diagrama de instalação no gerador.

O sensor possui não linearidade de 3%, o que garante confiabilidade na medição entre 10% e 120% da corrente nominal do instrumento (0 a 20 A), e fornece uma saída alternada entre 0 V e 1 V (YDHC, 2010), que é conectada a entrada diferencial do módulo de aquisição de dados.

Figura 20 – Foto do sensor de corrente em 20a e o diagrama de instalação em 20b e 20c.



Fonte: 20a: Adaptada de (YDHC, 2010); 20b: Autor.

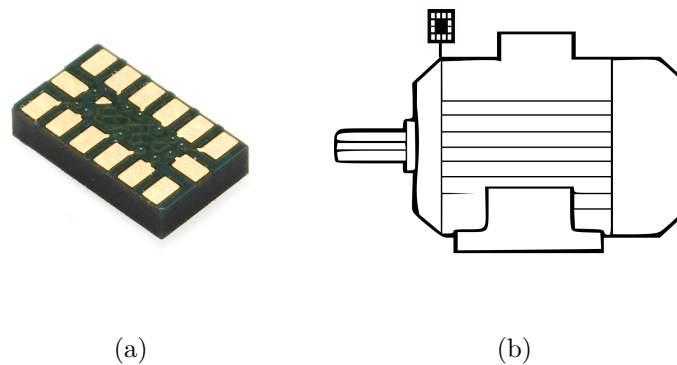
3.3.2 Sensoriamento de vibração mecânica

Monitoramento por vibração é uma das técnicas mais difundidas e mais apropriadas para análises em máquinas rotativas (BEEBE, 2004), tanto que tem sido utilizada em trabalhos recentes como em Azizi et al. (2017), que empregou análise vibracional para detecção de cavitação em bombas centrífugas. Stopa, Cardoso Filho e Martinez (2014) apresentam um estudo do espectro de frequência em sinais de vibração em bombas rotativas submetidas a falhas.

Ramalho et al. (2013) exhibe os padrões de vibração de uma motor elétrico submetido à operação com diferentes tipos de carga e com desalinhamento que podem ser detectados junto com a decomposição de Wavelet e técnicas de inteligência computacional. Anos posteriores, Ramalho et al. (2014) também utiliza uma metodologia semelhante, porém para detecção de curto circuito entre espiras de bobinamentos estatórico em motores de indução.

Para aquisição de vibração utilizou-se um sensor acelerômetro do tipo MEMs (do inglês *Micro Electro Mechanical system*) de três eixos, como proposto por Ramalho et al. (2014), exposto na Figura 21a e instalado no gerador de acordo como o exposto na Figura 21b. O sensor possui as seguintes características: não linearidade de 1%, sensibilidade de 800 mV/g e alimentação de 3,3 V.

Figura 21 – Foto do sensor MEMs [21a](#) o diagrama de instalação em [21b](#).



Fonte: [21a](#): Adaptada de ([SEMICONDUCTOR, 2008](#)) [20b](#): Autor.

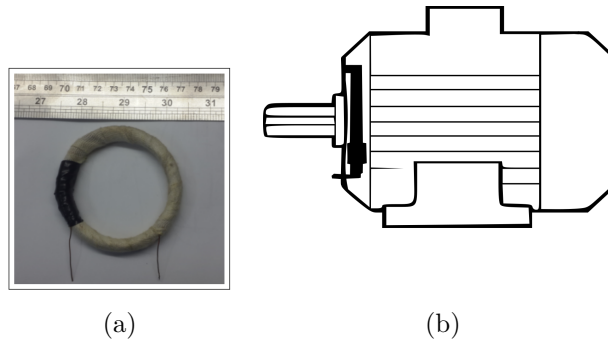
3.3.3 Sensoriamento de fluxo magnético axial

O fluxo axial que percorre uma máquina elétrica é uma característica natural do equipamento, gerado pelo acoplamento magnético entre o circuito do estator e do rotor ([JARZYNA, 1995](#)), devido a imprecisões física durante a construção do gerador. Contudo, [Jarzyna \(1995\)](#) relata que esse efeito pode ser influenciado por assimetrias entre o eixo do rotor e do campo gerado no estator.

[Henao, Demian e Capolino \(2003\)](#) realiza uma análise no domínio da frequência para detecção de falhas no bobinamento estatórico de um motor submetido a curto-circuito entre espiras. O autor afirma que a utilização de um sensor de fluxo é mais eficiente que a abordagem clássica pela análise da assinatura de corrente do motor. [Henao, Demian e Capolino \(2003\)](#) exhibe que o espectro de frequência de fluxo magnético é tão conclusivo quanto a corrente elétrica para a detecção de falhas no motor elétrico.

[Sousa e Sá \(2017\)](#) utilizou uma bobina constituída de 100 espiras de fio de cobre 24 AWG, inserida no eixo da máquina, como exibido na [Figura 22](#). Pelo fato da bobina estar imersa no fluxo magnético produzido pelo gerador, haverá uma tensão induzida de acordo com a lei de Faraday, porém não haverá corrente elétrica circulando pela mesma, visto que seus terminais permanecerão em aberto.

Figura 22 – Foto da bobina para captação de fluxo axial em 22a o diagrama de instalação em 22b.

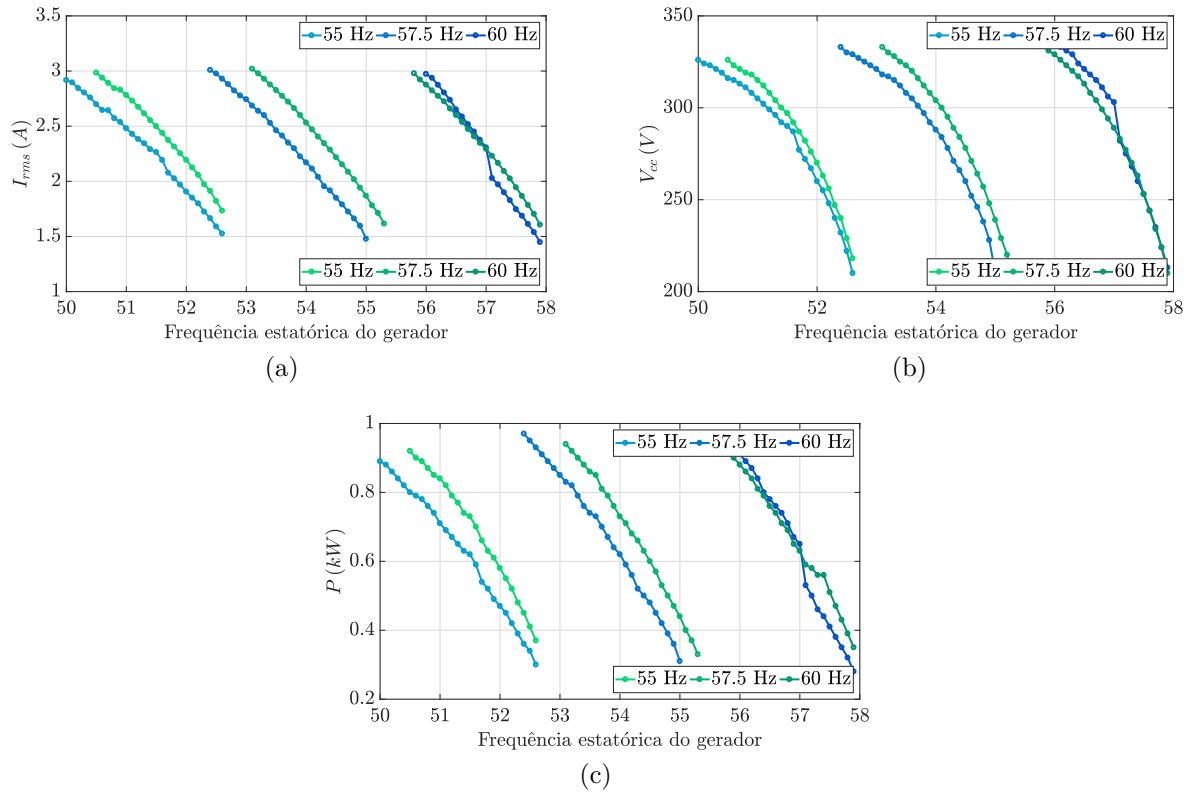


Fonte: 22a: Adaptada de (SOUSA, 2017); 20b: Autor.

3.4 Caracterização experimental do gerador

Objetivando identificar o quanto o processo de preparação do gerador modificou as características da máquina, realizaram-se experimentos nesse gerador e os dados foram comparados com os obtidos pelo gerador padrão. Na Figura 23 é exibido o comportamento da corrente elétrica, da tensão no barramento CC do conversor e da potência pela variação da frequência da máquina primária e do gerador. As curvas em azul são as medições realizadas no gerador padrão, enquanto que as curvas em verde são as curvas do gerador modificado. A diferença de frequência entre a GP e o GM para um mesmo ponto de geração é de $0,5 \text{ Hz}$.

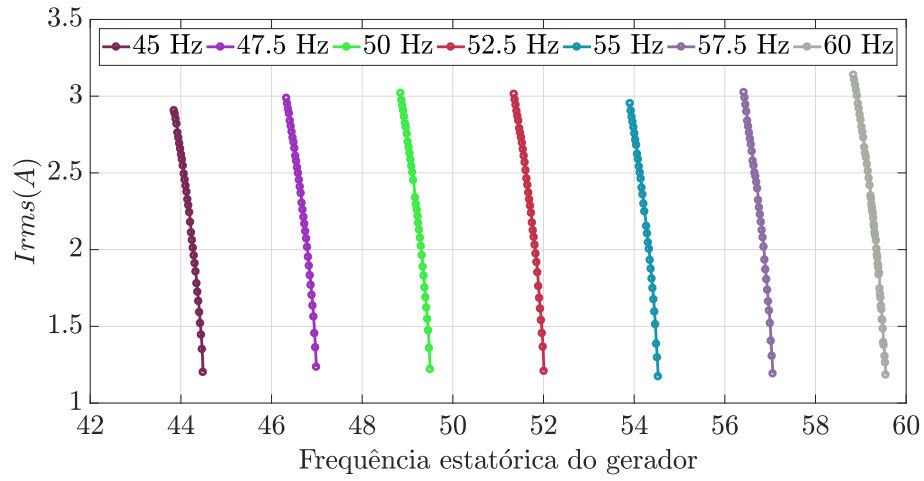
Figura 23 – Comportamento do GP, em azul, e GM, em verde. Em 23a é o comportamento da corrente elétrica, em 23b da tensão no barramento CC e em 23c da potência trifásica gerada



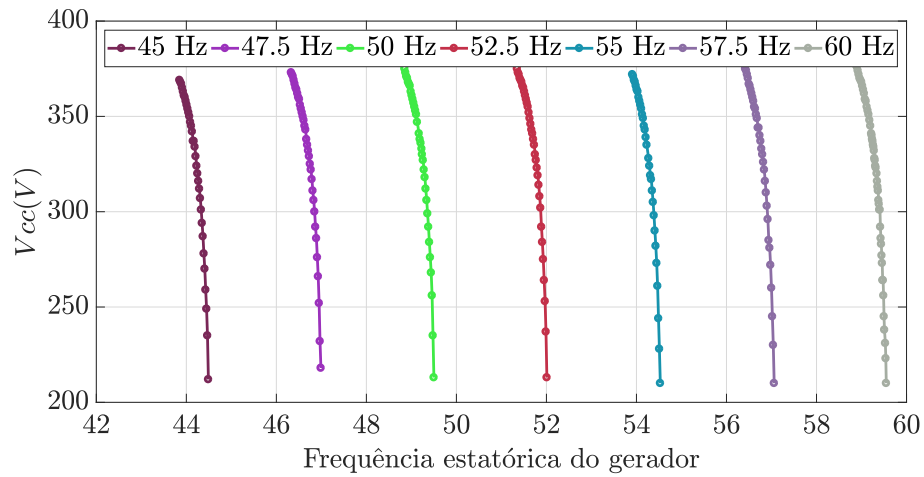
Fonte: Autor.

Essas evidências contribuíram para guiar a realização dos experimentos para aquisições dos dados do gerador. Os passos para aquisição dos dados são os mesmos descritos na Seção 3.1. Define-se para f_b valores entre 45 e 60 Hz espaçados de 2,5 Hz e para f_g passos de 0,02 Hz . As curvas exibidas na Figura 24 foram coletadas do protótipo do gerador. Na Tabela 2 são resumidos os procedimentos descritos.

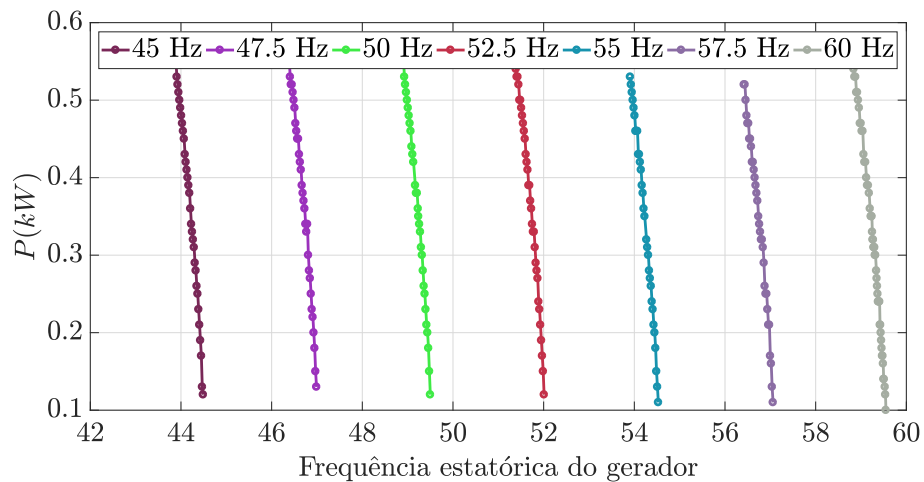
Figura 24 – Curvas do SCIG modificado para inserção de curto-circuitos. Em 24a é o comportamento da corrente elétrica, em 24b da tensão no barramento CC e em 24c da potência trifásica gerada



(a)



(b)



(c)

Fonte: Autor.

Considerando as seis condições de falhas estabelecidas na Seção 3.2, junto com a condição de funcionamento Normal, os experimentos foram divididos em 7 grupos. Foi feito um total de 1356 aquisições de dados, sendo 248 de condições normais do gerador e 1108 de condições de falha. Na Tabela 2, há um resumo dos experimentos realizados.

Tabela 2 – Passos para realização dos experimentos no gerador de indução

Estados	f_b	f_g
Normal	45 a 60 Hz , passos de 2,5 Hz	$f_1(V_{cc} = 210 V)$ até $f_{max}(V_{cc} = 380 V$ ou $I_g = 3 A$ ou $I_{mp} = 3 A)$, com passos de 0,02 Hz
Falha AI-1		
Falha AI-2		
Falha AI-3		
Falha BI-1		
Falha BI-2		
Falha BI-3		

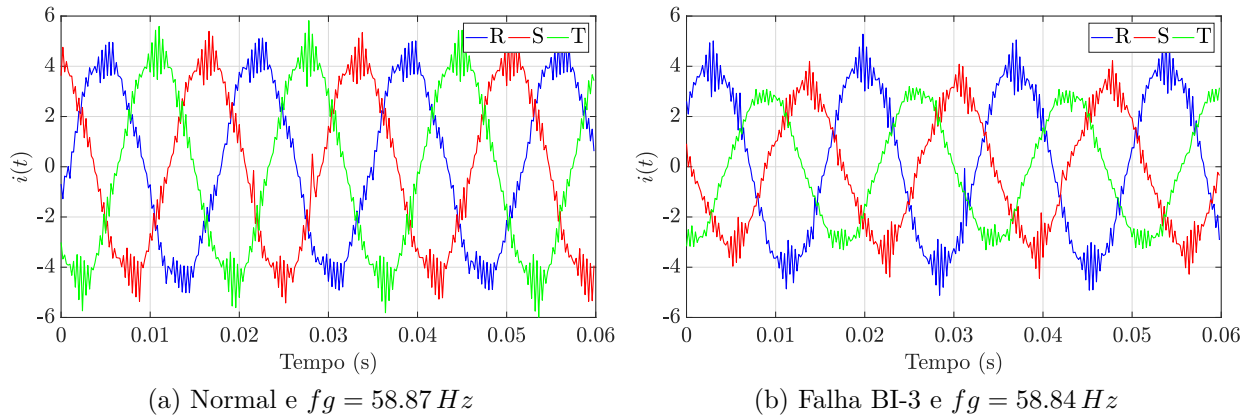
Fonte: Autor.

3.5 Extração de atributos e formação do banco para avaliação e validação de resultados

De acordo com a busca na literatura, foi escolhido utilizar os sinais de corrente elétrica do gerador para formação da base de dados, pois a técnica MCSA é promissora para detecção de falhas de curto-circuito entre espiras de máquinas elétrica de indução. Uma análise preliminar dos sinais feita na Seção 3.5.1 para, então, explicar como é gerado os bancos de dados na Seção 3.5.2.

3.5.1 Análises preliminares

Na Figura 25 é exibido um comparativo entre as aquisições dos sinais das correntes no SCIG operando em regime próximo a corrente nominal nas condições Normal e com Falha BI-3. A frequência de base é 60 Hz , a frequência do gerador é 58,87 Hz , considerando-o sem falha, e 58,84 Hz para o gerador com falha. Após a inserção do curto-circuito entre as fase R e T, percebeu-se que houve um desequilíbrio entre as correntes elétricas das três fases do gerador, como pode-se ver nas Figuras 25a e 25b. Na Tabela 3 é exibido um comparativo entre os valores da corrente elétrica medidos na linha do conversor para o gerador, e constatou-se que a ocorrência do curto, seja de alta ou baixa impedância, desbalanceia o sistema, e isso acarreta em uma subcorrente e sobrecorrente nas outras fases. Tanto a potência quanto a tensão do barramento CC sofreram diminuições em seus valores nominais, indicando também a ocorrência de uma anomalia.

Figura 25 – Corrente do gerador normal e com falha de BI-3, com $f_b = 60\text{ Hz}$ 

Fonte: Autor.

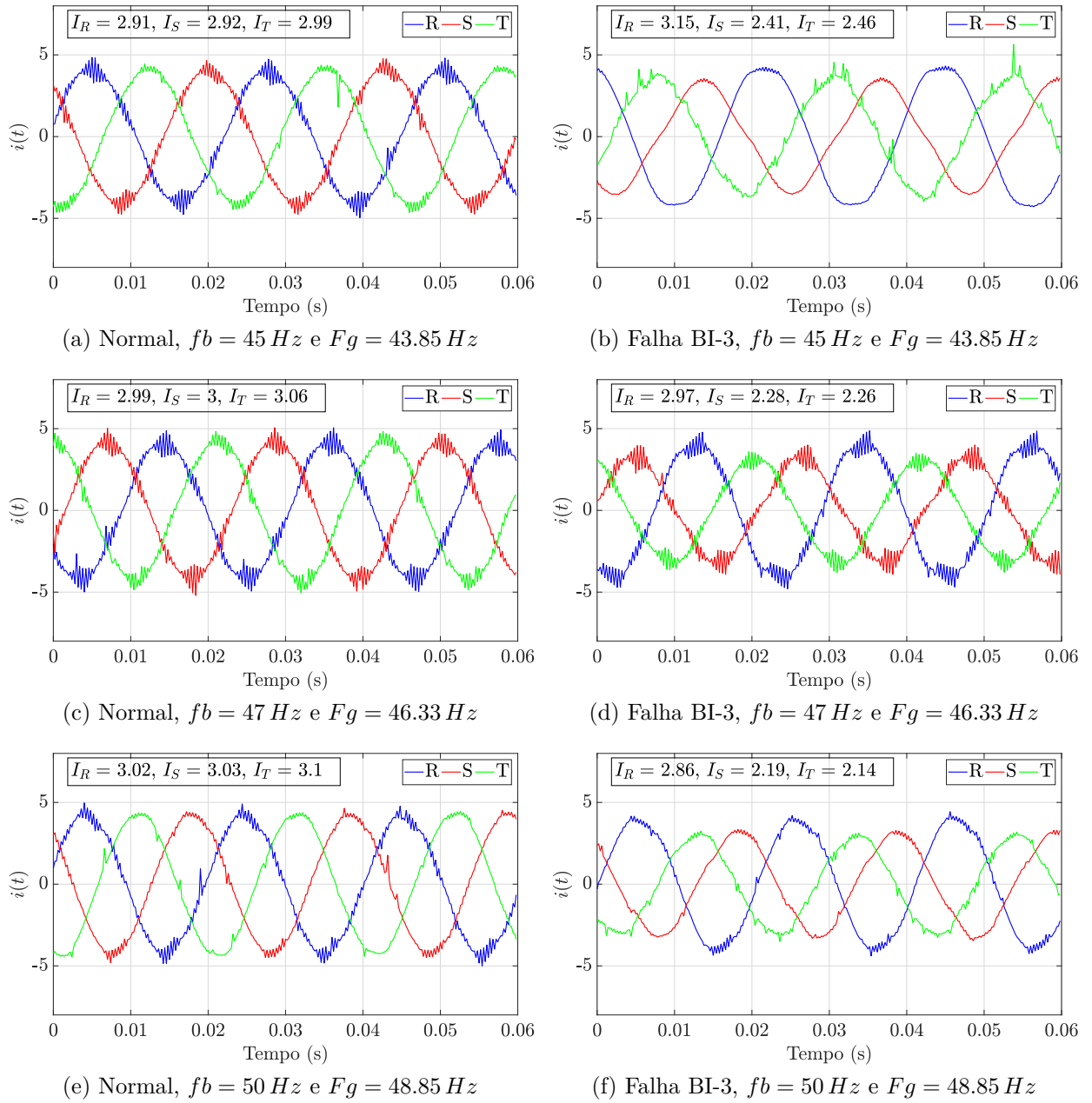
Tabela 3 – Comparativo numérico entre os valores de corrente elétrica de linha, tensão no barramento CC e potência para os estado normal e falha BI-3.

Medição	NORMAL	FALHA BI-3
$I_R\text{ (A)}$	3.1383	2.9376
$I_S\text{ (A)}$	3.1153	2.2624
$I_T\text{ (A)}$	3.2292	2.1119
$V_{cc}\text{ (V)}$	380	327
$P\text{ (kW)}$	0.54	0.42

O mesmo fenômeno se repete nas frequências de base de 45, 47,5, 50, 52,5, 55 e 57,5 Hz ao inserir o curto-circuito na fase R, que é mostrado na Figura 26. E de fato, a presença do curto na fase R impactou nas fases S e T, como se vê nos sinais de corrente elétrica entre as Figuras 26a e 26l.

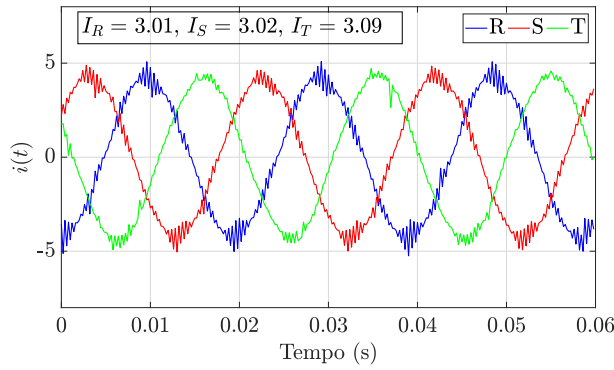
Mediante as análises na Tabela 3 e Figura 25 evidenciou-se que as situações extremas do experimento, funcionamento normal e curto-circuito BI-3, são facilmente descritas no domínio do tempo como visto anteriormente, pois um curto-circuito de baixa impedância entre 9.26% de espiras é perceptível. Mesmo que sua ocorrência tenha sido controlada durante os experimentos, o seu potencial de severidade para destruir a máquina é manifestado no desbalanceamento do sistema elétrico do gerador. Entretanto, nas ocorrências dos curtos de alta impedância, o desequilíbrio no gerador é latente, sendo difícil de ser detectado em análises no domínio do tempo. Essas condições são exibidas nas Figuras 27a até 27c.

Figura 26 – Corrente elétrica de linha do gerador normal e com falha, em regime de corrente nominal

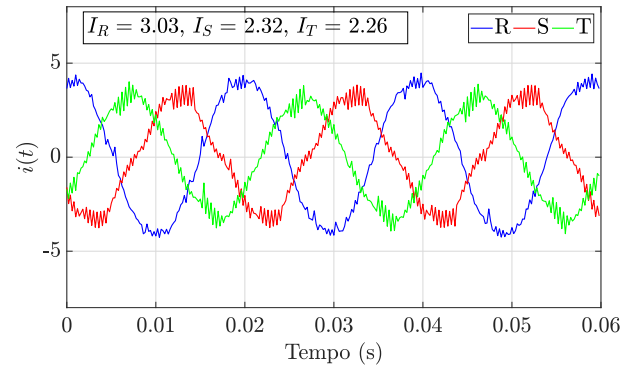


Fonte: Autor.

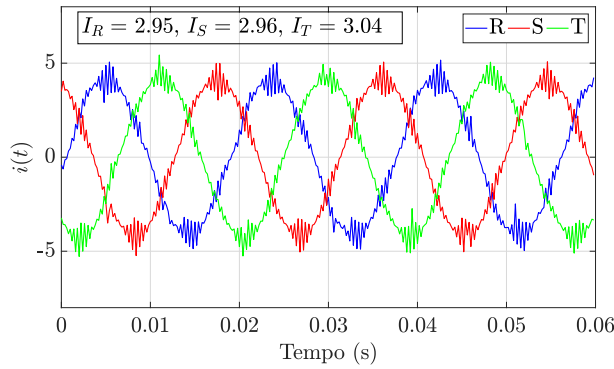
Figura 26 – (Continuação) Corrente elétrica de linha do gerador normal e com falha, em regime de corrente nominal



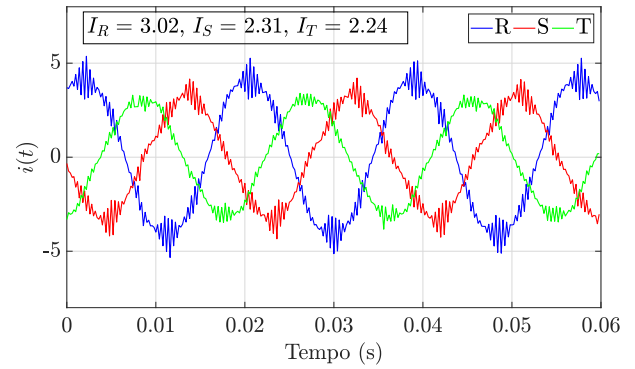
(g) Normal, $fb = 52 \text{ Hz}$ e $Fg = 51.35 \text{ Hz}$



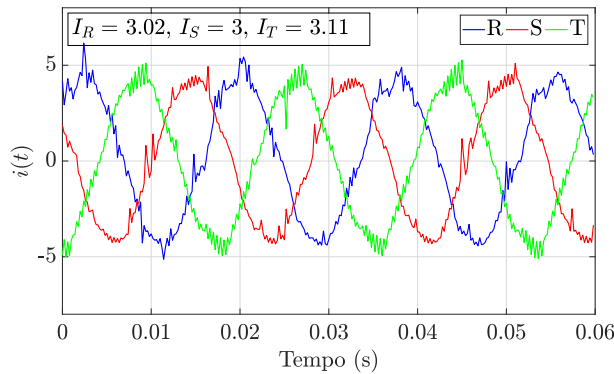
(h) Falha BI-3, $fb = 52 \text{ Hz}$ e $Fg = 51.35 \text{ Hz}$



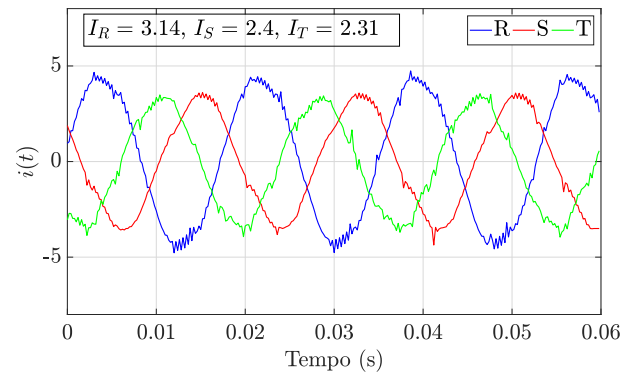
(i) Normal, $fb = 45 \text{ Hz}$ e $Fg = 53.91 \text{ Hz}$



(j) Falha BI-3, $fb = 55 \text{ Hz}$ e $Fg = 53.91 \text{ Hz}$



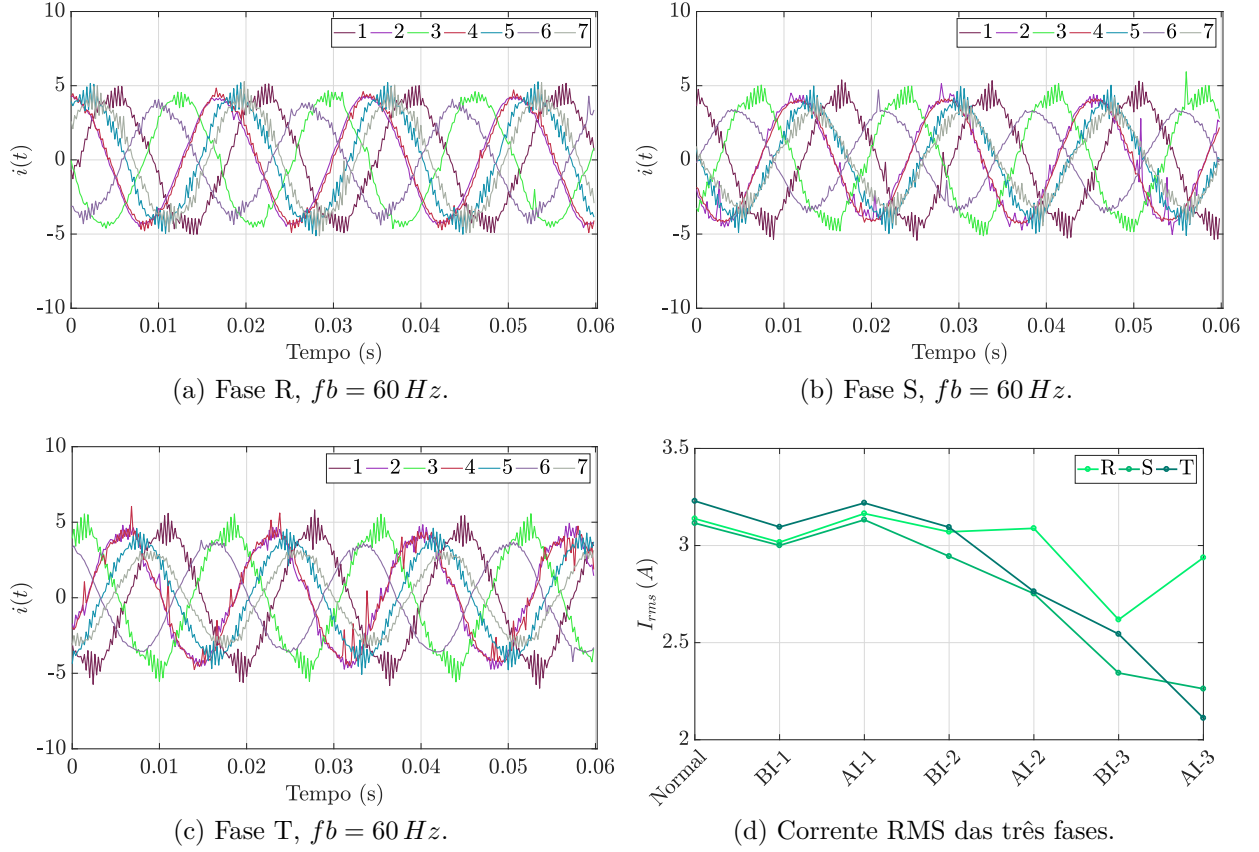
(k) Normal, $fb = 57.5 \text{ Hz}$ e $Fg = 56.42 \text{ Hz}$



(l) Falha BI-3, $fb = 57.5 \text{ Hz}$ e $Fg = 56.3 \text{ Hz}$

Fonte: Autor.

Figura 27 – Comparativo da corrente elétrica das três linhas, à operação próxima a corrente nominal, em diferentes condições de falha. Na legenda (1) é o gerador em estado normal, (2) é falha de AI-1, (3) é BI-1, (4) AI-2, (5) é a BI-2, (6) é AI-3, (7) é a BI-3



Fonte: Autor.

Em uma mesma condição de f_b e f_g , o curto inserido na fase R não é diretamente visível em nenhuma das 7 condições de operação mostradas nas Figuras 27a, 27b e 27c, tanto que a olho nu é fácil confundir a operação normal com quaisquer uma das falhas. A evolução do comportamento da corrente das três fases é exibido na Figura 27d e nota-se que, para um mesmo número de espiras há um desequilíbrio na corrente ao passar do regime de alta para baixa impedância, o que já era esperado. Apesar da ocorrência do curto-circuito, em todos os momentos a corrente esteve abaixo da nominal do gerador, portanto, a utilização apenas da diferença numérica do valor *rms* de corrente não é conclusiva para identificação da falha. Comportamentos similares são equivalentes para as fases S e T.

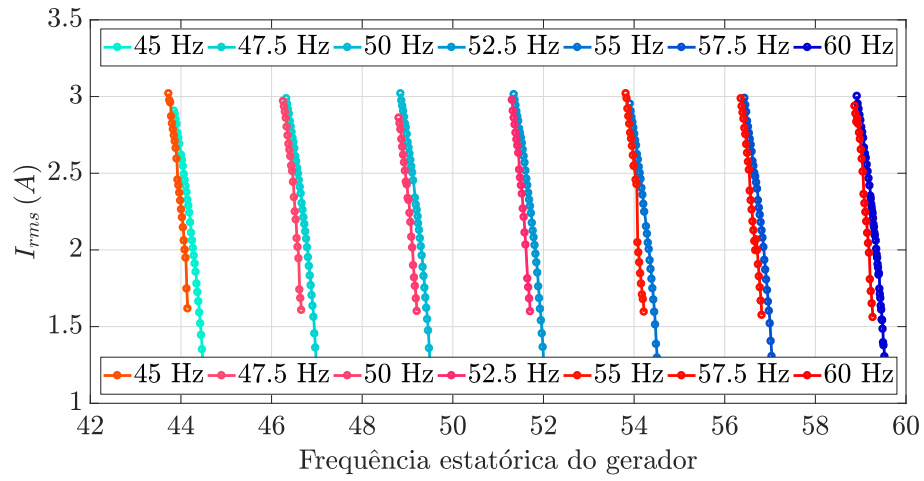
As evidências do curto-circuito durante a ocorrência da falha crítica são perceptíveis nas curvas do gerador apresentadas na Figura 28. As curvas de tensão no barramento CC e potências ficam distantes dos valores nominais, em comparação às condições normais. Entretanto, na ocorrência de uma falha incipiente (i.e. AI-1), praticamente não há diferenças

nas curvas de tensão e potência do aerogerador, como é exibido na Figura 29. As informações de potência e tensão no barramento CC proveem indicativos da ocorrência de problemas no gerador, dessa forma utilizar essas informações em conjunto com outras técnicas pode auxiliar na identificação das falhas do gerador de indução.

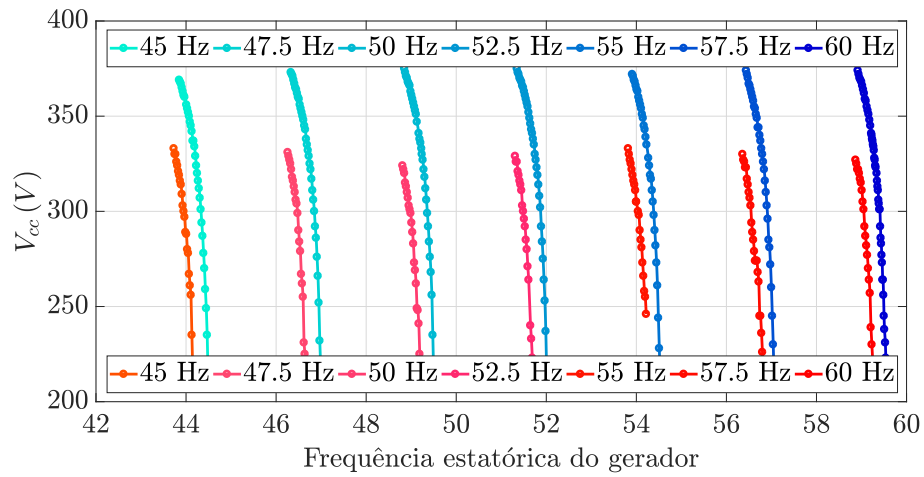
Junto com as evidências encontradas nos experimentos e com as premissas identificadas na Seção 2.2.1, para aplicar o método MCSA é necessário que: (i) utilize-se técnicas de extração de atributos, visto que a análise simples no domínio do tempo não é muito descritiva e ainda propõe-se (ii) a utilização do sinal de apenas um sensor de corrente, pois o estudo de detecção prematura de falhas de curto-circuito entre espiras do estator em motores elétricos, feito por Oliveira (2014), baseou-se na utilização do sinal de corrente de apenas uma das fase dos motor. Coelho (2015) também ressalta que o projeto de um sistema para detecção de falhas em motores acionados por conversores de frequência deve ser feito utilizando apenas um sensor de corrente elétrica, pois o autor ratifica que todas as fases do motor são penalizada mediante a presença de um curto circuito e, portanto, a utilização de apenas um sensor eliminaria informações redundantes.

A emulação do curto-circuito é inserida no ramo entre as fases R e S e, apesar de a fase T não está no ramo do curto-circuito, acredita-se que a aparição de uma falha é percebida em todas as três correntes do gerador. A extração de atributos será feita na fase R. A formação da base de dados é descrita na seção a seguir.

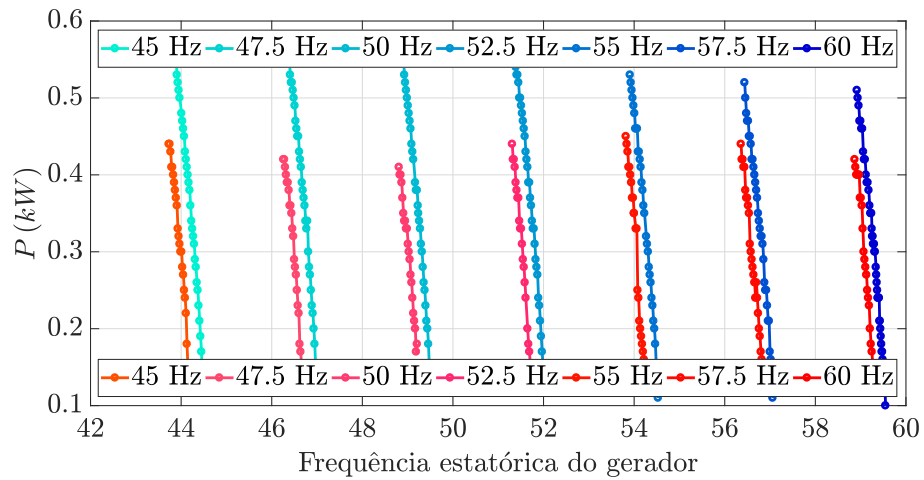
Figura 28 – Curvas do SCIG modificado nas condições normais, em tons de azuis, e falha de BI-3, em tons de vermelhos. Em 28a é o comportamento da corrente elétrica, em 28b da tensão no barramento CC e em 28c da potência trifásica gerada



(a)



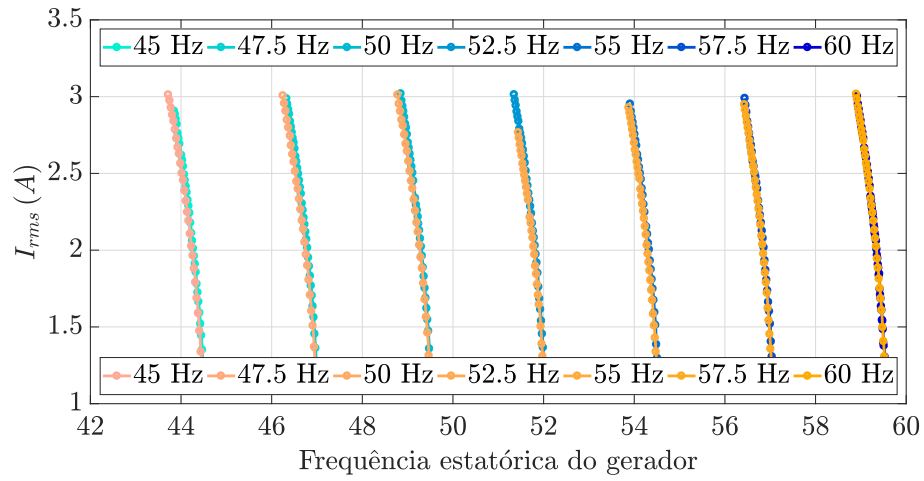
(b)



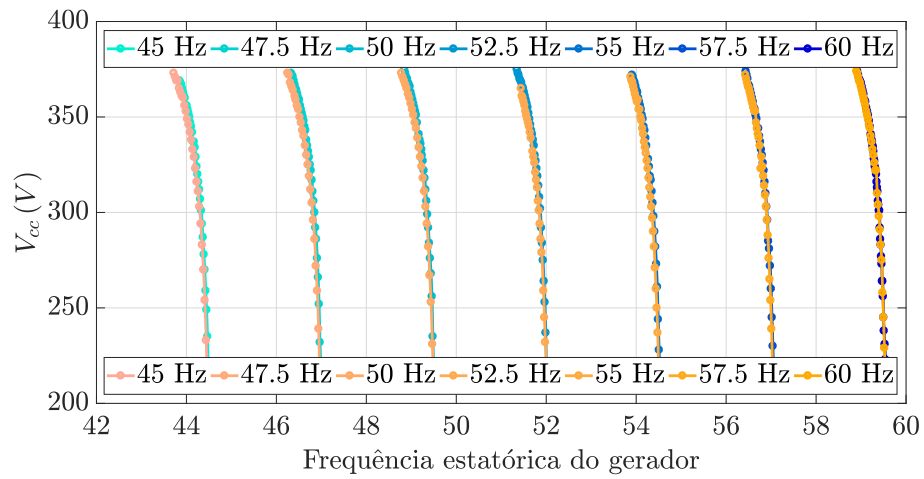
(c)

Fonte: Autor.

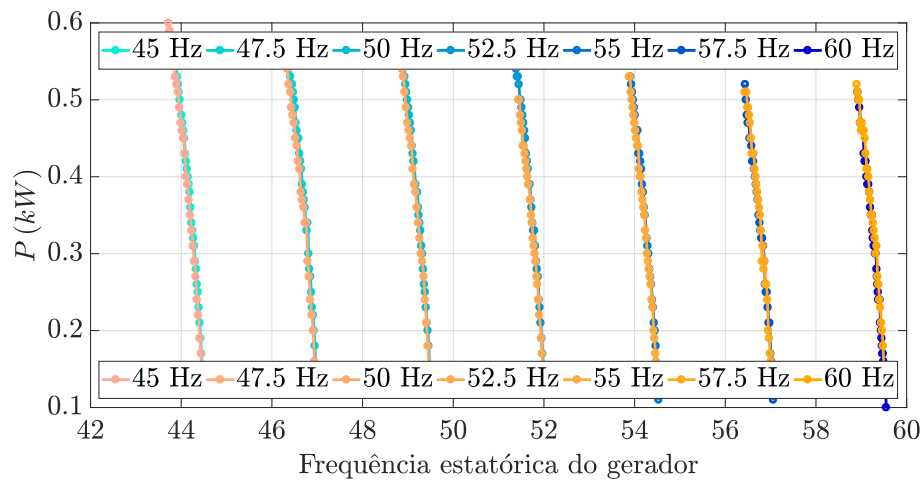
Figura 29 – Curvas do SCIG modificado nas condições normais, em azuis, e falha de CC AI-1, em amarelos. Em 29a é o comportamento da corrente elétrica, em 29b da tensão no barramento CC e em 29c da potência trifásica gerada



(a)



(b)



(c)

Fonte: Autor.

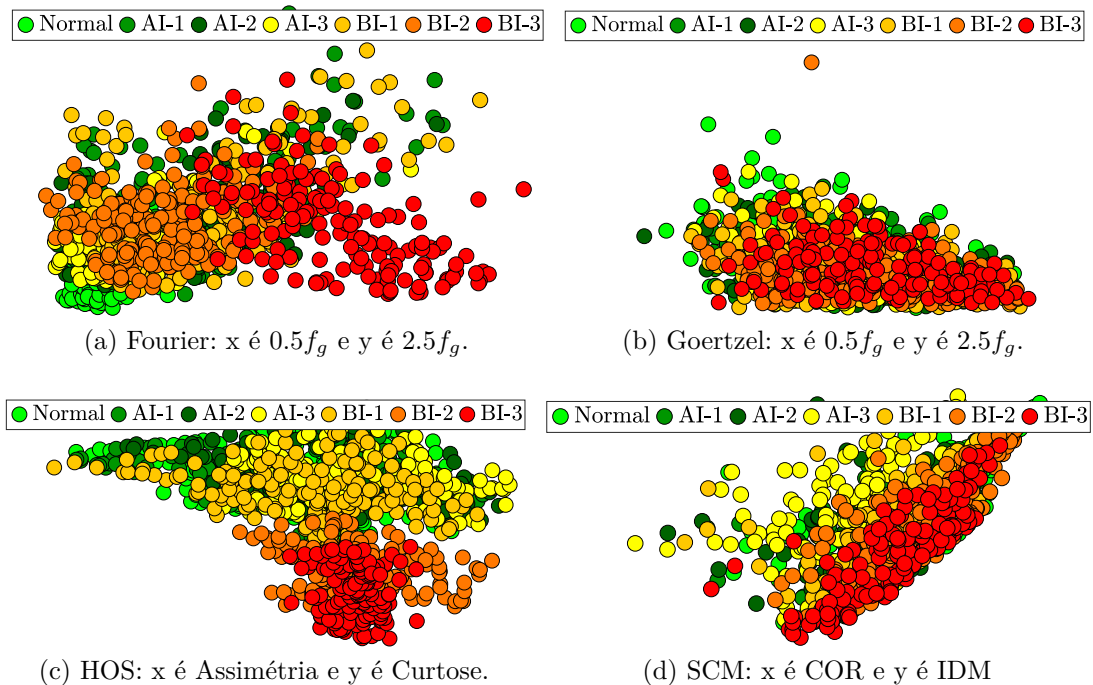
3.5.2 Formações das bases de teste de corrente elétrica para análise de falhas

Utilizam-se as técnicas de Fourier, Goertzel, HOS e SCM para extração de atributos. Todas as bases de dados¹ formadas possuem 1356 amostras e 7 classes, dispostas da seguinte forma: 248 amostras na classe Normal, 203 da falha AI-1, 179 da falha AI-2, 183 da falha AI-3, 177 da falha BI-1, 208 da falha BI-2, 158 da falha BI-3.

A base de dados formada pelo extrator Fourier possui 6 atributos, que são: harmônica $0,5f_g$, $2,5f_g$, $3f_g$, $5f_g$ e $7f_g$ normalizadas pelas fundamentais. A base formada pelo algoritmo de Goertzel possui os mesmos atributos e as mesmas harmônicas. A base formada por HOS possui 4 atributos, que são: assimetria, curtose, variância e *rms*. A base da SCM possui 6 atributos, que são: correlação, momento inverso diferencial, entropia, distância *Chi-square*, razão da distância *Chi-square* e a taxa média absoluta diferencial.

Na Figura 30 são exibidas as projeções bi-dimensionais das bases de dados formadas pelos quatro extratores. Há sobreposições entre as classes e isso incita que talvez não exista uma superfície de decisão linear para classificação. Porém, dado a dimensionalidade do problema, pode ser que haja um hiperplano capaz de separar linearmente os bancos de dados, mesmo que as projeções bi-dimensionais não acusam.

Figura 30 – Projeções bi-dimensionais dos bancos de dados para cada um dos extratores.



Fonte: Autor.

¹ Os detalhes referentes a formação da base de dados, como tipo de arquivo utilizado, padrão adotado para montagem do banco de dados e outros, estão detalhados no Apêndice B, bem como o *link* para *download* do banco de dados.

4 Investigação preliminar sobre extratores de atributos e classificadores

Nesse Capítulo são exibidos os passos relativos ao projeto do classificador de falhas incipiente de curto-circuito. A ordem do texto é a seguinte: na Seção 4.1 são discutidos os procedimentos experimentais comuns a todos os classificadores. Na Seção 4.2 são exibidos os resultados dos classificadores lineares e na Seção 4.3 dos classificadores paramétricos. O treinamento da rede neural MLP é exibida na Seção 4.4. Na Seção 4.6 é feito o comparativo entre as técnicas de extração de atributos e os resultados dos classificadores. O final desse capítulo discute as melhores combinações.

4.1 Procedimentos comuns a todos os classificadores

Para comparar o desempenho dos classificadores, todos os experimentos seguem os mesmos procedimentos em comum:

- A normalização utilizada em todos as bases de dados é a estatística. A qual é feita a subtração da média e divisão pelo desvio padrão em cada um dos vetores de atributos, x_{nj} , da seguinte forma:

$$x_{nj}^* = \frac{x_{nj} - \mu_j}{\sigma_j}, \quad (4.1)$$

na qual, μ_j é a média do atributo j e σ_j é o seu desvio padrão. Isso garante variância unitária para cada uma das dimensões de atributos;

- A separação dos dados foi feita utilizando a técnica *hold out*, que consistiu em utilizar 80% das amostras para criar uma base de treinamento e 20% para uma base de teste, separadas aleatoriamente. O base de dados separado para treinamento possui quantidades de amostras balanceada por classes, ou seja, a quantidade de amostras é a mesma para todas as classes;
- Para todos os experimentos exidos ao longo desse trabalho foram realizados 50 treinamentos independentes. A partir destes são gerados valores de acurácia geral, por classes e matrizes de confusão média. Outras métricas como sensibilidade, especificidade e *f-score* também são apresentadas e ajudando a compor a discussão dos classificadores;
- As definições dos classificadores serão feitas com bases nos resultados obtidos nas bases de teste, entretanto o resultados na base de treinamento também ajudarão nas

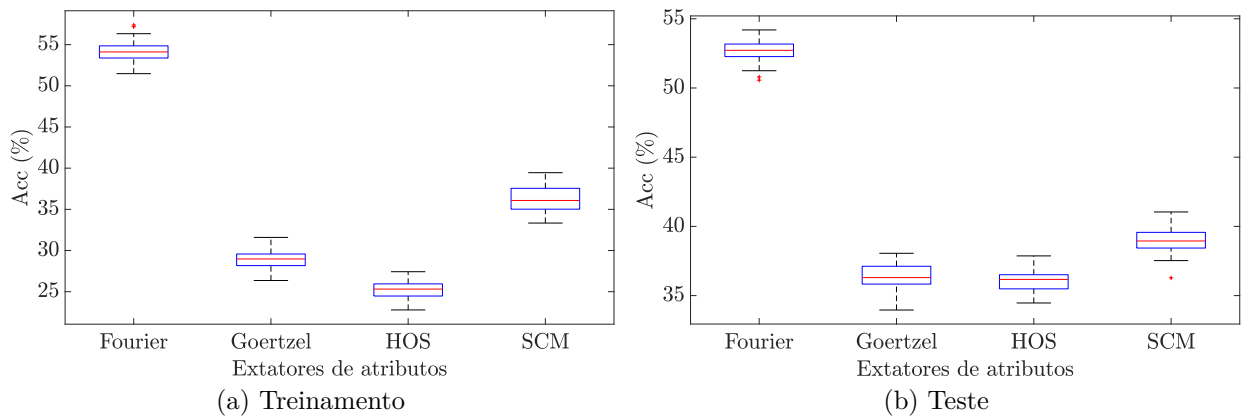
análises. Pois, acredita-se que avaliação em pares dessas duas bases provê robustez ao processo de projeto do classificador, pois possibilita identificação de problemas como sobreajustamento, por exemplo.

4.2 Classificadores lineares

Visto que não há um conhecimento prévio dos bancos de dados constituídos, primeiramente é utilizada uma ferramenta linear com o intuito de verificar a separabilidade da base de dados. Para tal, utilizou-se, na seguinte ordem, os classificadores OLS e Perceptron Simples. Em ambos os classificadores utilizou o rótulo $d_i = 1$ para a classe de interesse na classificação e $d_i = -1$ para as classes opostas à classificação.

O classificador OLS não possui hiperparâmetros, seu treinamento é realizado em lote (*batch*) e a variação das taxas de acerto gerais do classificador OLS na base de treinamento e de teste são exibidas na Figura 31. O melhor resultado foi obtido na base de dados formada pelo extrator de atributos Fourier, entretanto em seu melhor resultado atingiu-se menos que 60% de classificações corretas. Visto que o classificador dos mínimos quadrados ordinário utiliza todas os vetores da base de dados de uma única vez para estimação da matriz de pesos $\mathbf{W}$, acredita-se que usar a atualização iterativa por amostra, baseada em uma regra de aprendizagem, possa provê melhores resultados. Portanto, o próximo teste é feito utilizando o Perceptron Simples como classificador.

Figura 31 – *Boxplot* da acurácia da aplicação do OLS nas bases de dados formadas pelos extratores. Em 31a estão os resultados na base de treinamento e em 31b na base de teste.

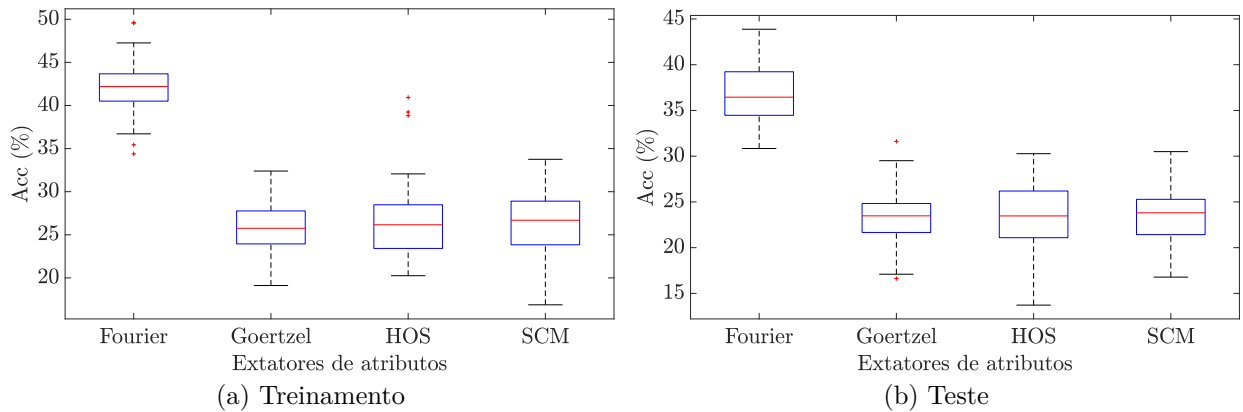


Fonte: Autor.

Para treinamento do Perceptron Simples utilizam-se o número de épocas, $N_e = [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000]$ e a taxa de aprendizagem com decaimento exponencial, partindo de $\eta = [0,5, 0,4, 0,3, 0,2, 0,1, 0,01]$ até em 0.0001. Baseado na acurácia

geral, o melhor dos casos foi com $N_e = 1000$ e $\eta = 0,1$ e a distribuição da acurácia é exibida na Figura 32. Mediante os resultados obtidos com esse classificador identificou-se que o OLS, mesmo sendo treinado em modo *batch*, obteve taxas superiores ao Perceptron Simples. Em ambos os classificadores, o uso de Fourier como extrator de característica demonstrou-se superior aos outros extratores.

Figura 32 – *Boxplot* da acurácia da aplicação do Perceptron simples nas bases de dados formadas pelos extratores. Em 32a estão os resultados na base de treinamento e em 32b na base de teste.



Fonte: Autor.

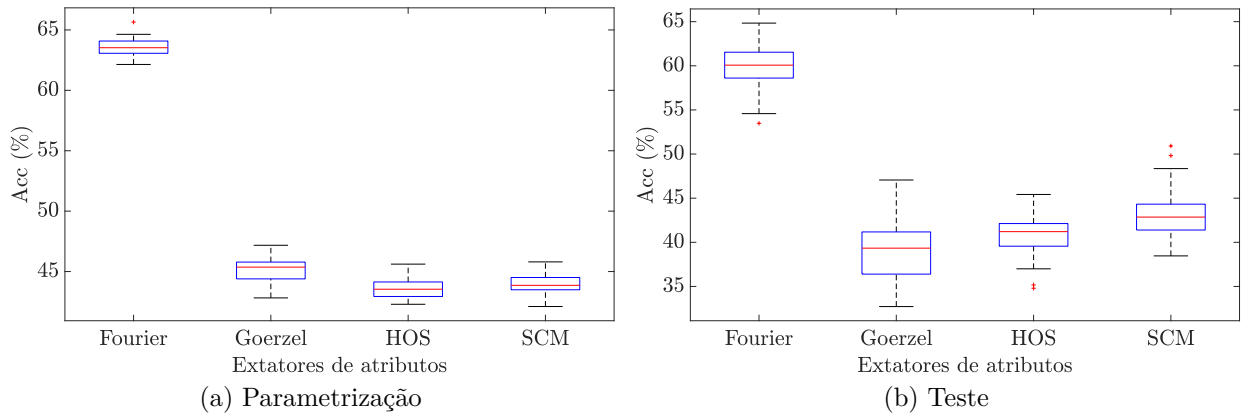
Os resultados médios, de ambos os classificadores foram próximos a 50% e, para justificar a baixa eficiência, é importante ressaltar que as projeções bidimensionais dos dados não demonstravam uma separabilidade linear entre as classes, e isso foi afirmado pela operação desses classificadores. Entretanto, para a identificação incipiente de falhas de curto-circuito no gerador espera-se um classificador com maior capacidade em detectar falhas. Dessa forma, o próximo passo é testar o classificador Gaussiano.

4.3 Classificador Gaussiano

Os classificadores Gaussianos partem da premissa de que os dados se comportam como variáveis aleatórias regidas por funções densidades de probabilidade gaussianas. As variações desse classificador são decorrentes das diferentes formas de estimar a matriz de covariância exibida na Subseção 2.3.2. Esse classificador não possui hiperparâmetros, com exceção da estimativa de Friedman (1988), que possui o fator λ , escolhido em 0,5, objetivando permitir uma operação com funções discriminantes entre quadráticas e lineares. De todas as configurações exibidas na Subseção 2.3.2, a que apresentou melhores resultados foi a estimativa das matrizes de covariância independentes por classe, operando como discriminante quadrático. Seus resultados são exibidos na Figura 33. Percebe-se que as taxas de acertos foram superiores aos classificadores OLS e Perceptron Simples. A melhoria

a frente dos classificadores lineares é esperada, pois um discriminante quadrático possui maior margem de ajustes em contrapartida a um linear. O uso do extrator Fourier também demonstrou ser melhor nesse classificador.

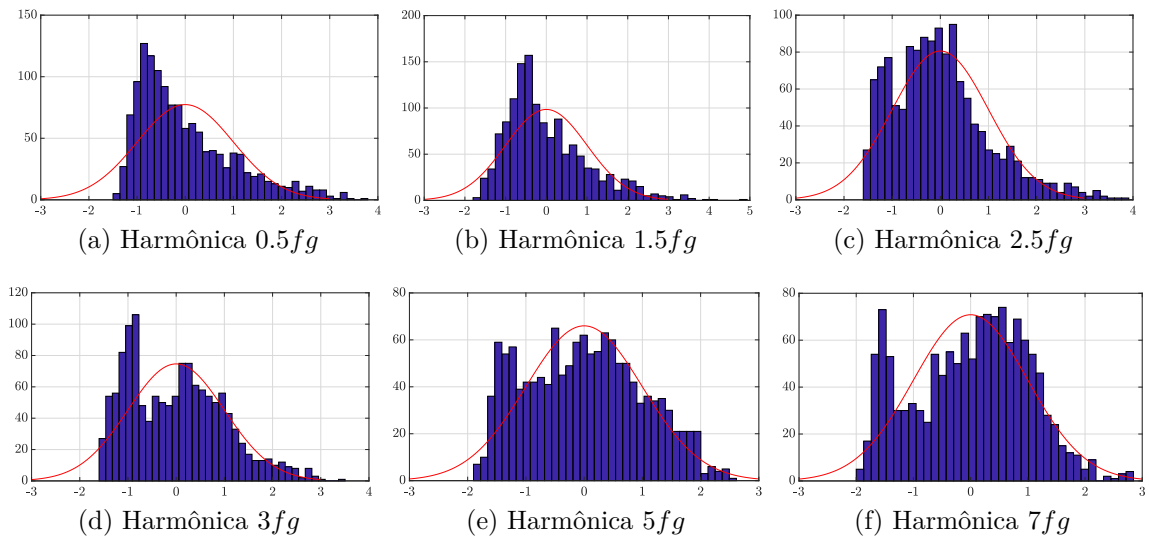
Figura 33 – *Boxplot* da acurácia da aplicação do Perceptron simples nas bases de dados formadas pelos extratores. Em 32a estão os resultados na base de treinamento e em 32b na base de teste.



Fonte: Autor.

Apesar das melhorias, os acertos do classificador Gaussiano quadrático não superaram 65% de acertos, considerando tanto os conjuntos de parametrização quanto os de teste. A premissa de que o comportamento dos dados é regido por um processo gaussiano é refutada ao analisar as distribuição dos atributos, expostas na Figura 34.

Figura 34 – Histograma das harmônicas de Fourier utilizadas como atributos comparado com uma distribuição Gaussiana.



Fonte: Autor.

Mediante as evidências supracitas identificou-se a necessidade da aplicação de um classificador não-linear e escolheu-se o Perceptron multi-camadas - MLP, cujos treinamentos são expostos a seguir.

4.4 Classificador Neural MLP

A arquitetura deste classificador, exibida na Subseção 2.3.3.1, utiliza como unidade básica o mesmo neurônio do classificador Perceptron Simples, porém, utilizando funções de ativação não-lineares em suas saídas, além de estrutura-los em cascata com outras camadas de neurônios. A necessidade de utilização de funções não-lineares e múltiplas camadas garante a operação não-linear da rede (BOSE; GARGA, 1993) e (MINSKY; PAPERT; BOTTOU, 2017).

O custo da versatilidade da operação não-linear da rede MLP reside respectivamente na quantidade de hiperparâmetros e na característica estocástica da aprendizagem (*i.e.* treinamento) desse classificador. Para realizar um ajuste apropriado dos hiperparâmetros, as heurísticas iniciais para os primeiros experimentos nesse classificador são feitas:

- Ajuste dos pesos (*i.e.* aprendizagem) pelo gradiente descendente e retropropagação do erro (*backpropagation*);
- Função de ativação tangente hiperbólica, $\tanh(x)$, na saída de todos os neurônios;
- Critério de parada pelo número de épocas de treinamento;
- Termo de viés, ou *bias*, ajustado em -1;
- Número inicial de neurônios da camada oculta, Q , de acordo com a regra do valor médio de Blum (1992) e Masters (1993):

$$Q_0 = \frac{M + P}{2}, \quad (4.2)$$

na qual M é o número de neurônios de saída (*i.e.* número de classes para esse problema) e P o número de unidades de entrada (*i.e.* atributos).

- Hiperparâmetro relativo ao número de épocas de treinamento, N_e , inicialmente ajustado em 1000;
- Valor inicial do hiperparâmetro relativo à taxa de aprendizado, η_0 , com decaimento exponencial, começando em 0.5 e terminando em 0.0001;
- Hiperparâmetro relativo ao termo de *momentum*, α , inicialmente ajustado em 0.5;

- Por critérios simplistas foi escolhido, primeiramente, avaliar uma camada oculta. E de acordo com Haykin (2001) o número de amostras, N , do conjunto de dados deve seguir a relação,

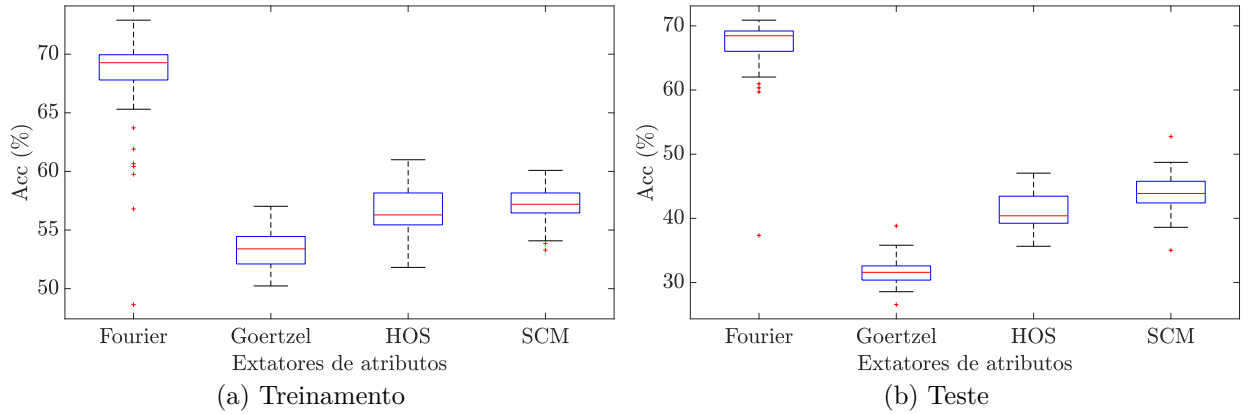
$$N > \frac{W}{\epsilon}, \quad (4.3)$$

na qual W é o número total de pesos da rede e ϵ é o erro que deseja-se atingir no conjunto de treinamento. As heurísticas indicam que uma boa aproximação é $N \approx 10W$. Entretanto, caso seja necessário, utilizar-se-á outras configurações desse hiperparâmetro;

- Rotulação $d_i = +0,97$ para a classe de interesse e $-0,97$ para as demais, com o intuito de evitar a região de saturação da função $\tanh(x)$.

Os resultados dos primeiros experimentos desse classificador nos quatro diferentes bancos de dados são exibidos na Figura 35. Novamente, o extrator Fourier destaca-se dos demais e, sem haver sequer seleções criteriosas dos hiperparâmetros a rede MLP já apresentou taxas de acerto superiores ao classificador Gaussiano quadrático, indicando maior aptidão para a detecção prematura de falhas de curto-circuito.

Figura 35 – *Boxplot* da acurácia da aplicação da MLP nas bases de dados formadas pelos extratores. Em 35a estão os resultados na base de treinamento e em 35b na base de teste.



Fonte: Autor.

O próximo passo é estimar os melhores hiperparâmetros. Portanto, nas próximas subseções será exibida a sequência de experimentos realizados para seleção dos seguinte hiperparâmetros: (i) Número de épocas - N_e ; (ii) Taxa de aprendizagem - η ; (iii) Termo de *momentum* - α e (iv) Número de neurônios da camada oculta - Q , nesta mesma ordem.

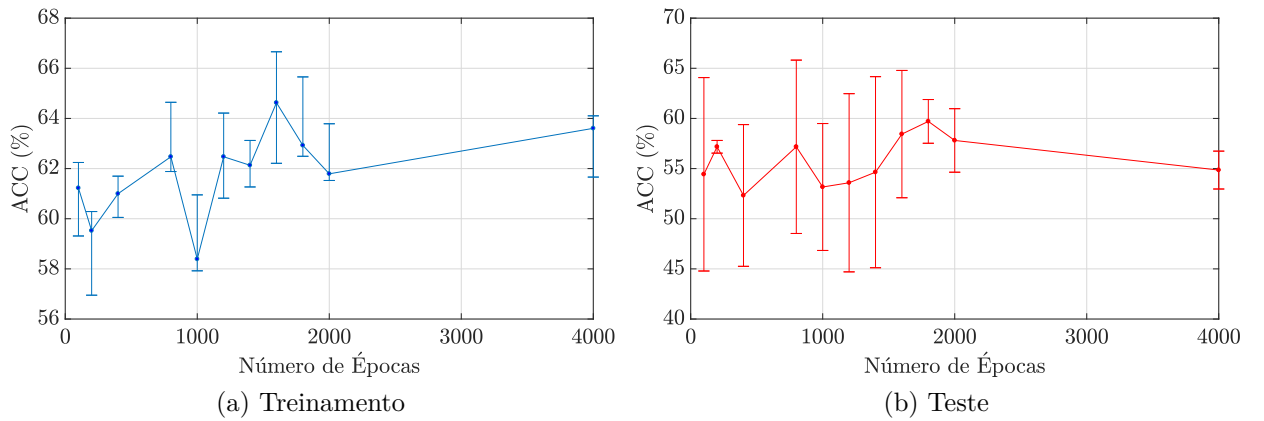
Antes de dar continuidade aos experimentos é importante salientar que, dentre as aplicações de todos os classificadores, utilizados até então, nos bancos de dados formados

pelos extratores propostos, o que apresentou melhores resultados foi o extrator Fourier. Por isso, a exibição dos resultados dos próximos experimentos focam apenas no bando de dados formado por esse extrator, porém a mesma metodologia foi seguida para treinar distintas redes MLP para os bancos formados pelos extratores Goertzel, HOS e SCM. Seus resultados serão analisados no capítulo posterior.

4.4.1 Número de Épocas

Realizaram-se 50 treinamentos independentes utilizando distintos números de épocas de treinamento, $N_e = [100, 200, 400, 800, 1000, 1200, 1400, 1800, 2000, 4000]$. O impacto da variação do número de épocas na acurácia é exibido na Figura 38. Percebe-se, que tanto nos conjuntos de treinamento quanto nos de teste, os melhores resultados foram obtidos com 1800 épocas de treinamento.

Figura 36 – Impacto da variação do número de épocas na acurácia média após 50 treinamentos.



Fonte: Autor.

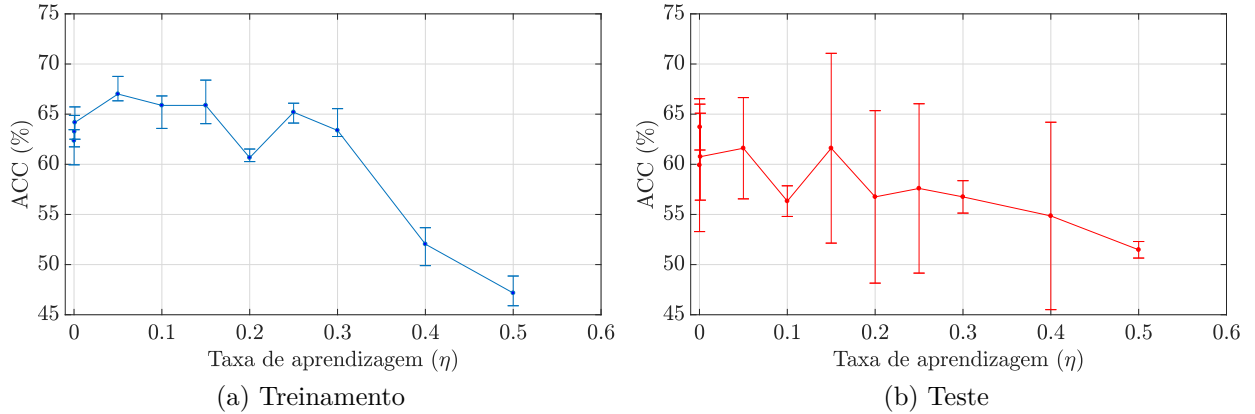
A maior taxa de acerto média, nos conjuntos de treinamento e de teste, obteve-se com 1800 épocas. Entretanto, a análise da distribuição da taxa de acerto na base de treinamento, mostrada na Figura 36a, pode induzir que após 2000 épocas a taxa de acerto volta a aumentar, podendo chegar a um valor superior a 1800 épocas. Porém, analisando a distribuição de acertos na base de testes, apresentado na Figura 36b, após 1800 épocas a taxa de acerto média cai. Isso é indício de que nesses treinamentos possivelmente houver sobreajustamento (i.e. *overfitting*). Portanto, o valor do hiperparâmetro N_e adotado para os próximos experimentos é 1800.

4.4.2 Taxa de aprendizagem

Para estimar a taxa de aprendizagem inicial η_0 que melhor se adapta à base de dados, testaram-se os seguintes valores: $\eta = [0,0001, 0,0005, 0,001, 0,05, 0,1, 0,15, 0,2, 0,25, 0,3,$

0,4, 0,5] com decaimento exponencial até 0,00001. Na Figura 37 exibe-se a evolução da acurácia para os diferentes valores de η e percebe-se que, acima de 0,1, há uma instabilidade crescente na classificação, que pode ser visto Figura 37b.

Figura 37 – Impacto da variação da taxa de aprendizagem na acurácia média após 50 treinamentos.



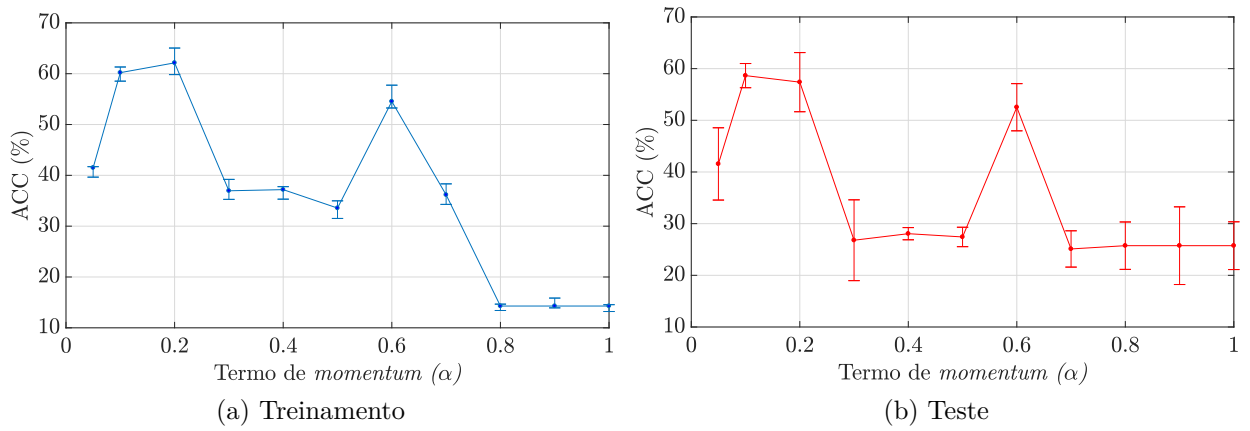
Fonte: Autor.

Além disso, heurísticas indicam que esse hiperparâmetro deve ser ajustado em valores abaixo de 0,5 com o objetivo de facilitar a convergência do treinamento. Abaixo de 0,05 a acurácia manteve-se pouco dispersa para a base de treinamento. Portanto, o valor selecionado foi $\eta = 0,1$ para, então, partir-se para ajuste do termo de *momentum*.

4.4.3 Termo de *momentum*

Para identificar qual o melhor valor do hiperparâmetro de *momentum*, realizam-se os experimentos utilizando $\alpha = [0,05, 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9, 1]$. Com $\alpha = 0,2$ obteve-se os melhores resultados. Portanto, esse foi o valor selecionado. É importante ressaltar que a partir de 0,6 a taxa de acerto cai consideravelmente, pois, como esse termo confere uma característica de conjugado ao método do gradiente descendente, valores elevados podem distanciar o treinamento dos mínimos locais da curva de erro.

Figura 38 – Impacto da variação da taxa de momento na acurácia média após 50 treinamentos.

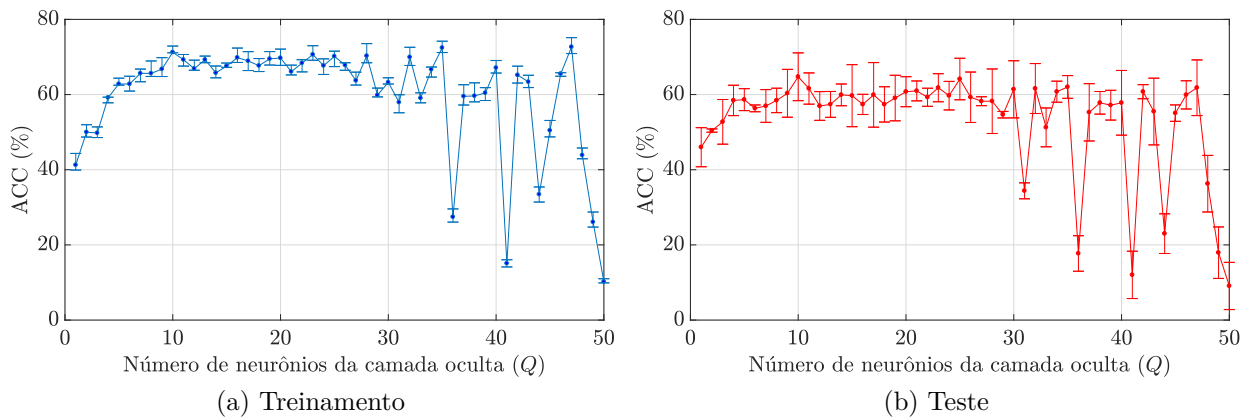


Fonte: Autor.

4.4.4 Número de neurônios

Para definição do número de neurônios utilizados na camada oculta, realizam-se experimentos variando o Q de 1 até 50 e os resultados são exibidos na Figura 39. Percebe-se que a taxa de acerto cresce a medida que aumenta-se o número de unidades ocultas, chegando a um valor quase que constante a partir de $Q = 10$. Esse comportamento mantém-se até $Q = 25$, pois a partir desse ponto a acurácia média passa a ter uma característica oscilatória, indicando que há dificuldades para convergência no treinamento. Dessa forma, o número de neurônios que melhor adapta-se a este problema deve estar entre 10 e 25.

Figura 39 – Impacto da variação do número de neurônios ocultos, (Q), na acurácia média após 50 treinamentos.



Fonte: Autor.

Isso vai de encontro com a afirmativa de Swingler (1996), na qual o número de

neurônios na camada oculta deve ser menor que duas vezes o número de atributos. Portanto, escolheu-se $Q = 10$. Um resumo dos hiperparâmetro utilizados é exibido na Tabela 4.

Tabela 4 – Resumo dos hiperâmetros propostos para a combinação Fourier-MLP

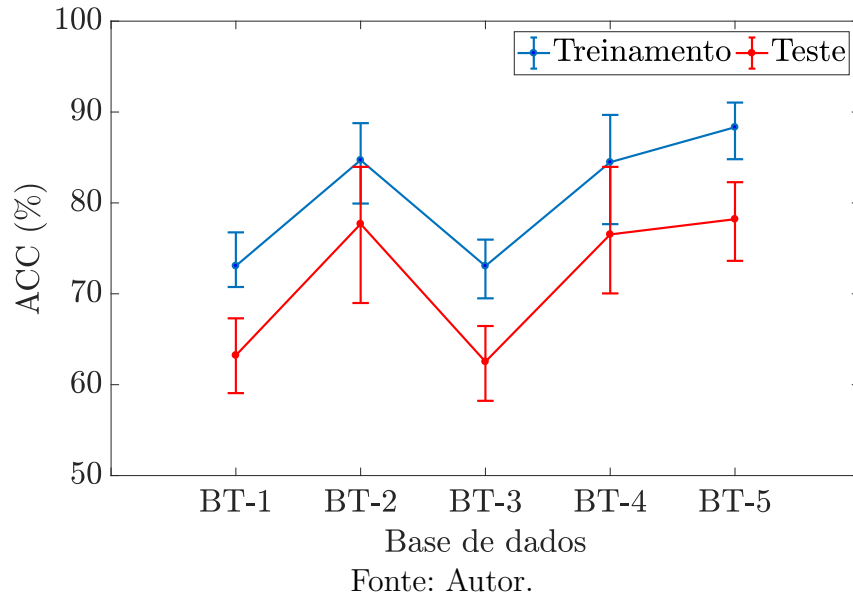
Hiperparâmetro	Valor
N_e	1800
η	0,1
α	0,2
Q	10
d_i	+0,97
b	-1
φ	$\tanh(x)$
Camadas ocultas	1

Após a definição dos hiperparâmetros, identificou-se que a taxa de acerto nas base de treinamento subiu de 67,90% para 71,25% e na base de teste de 63,93% para 64,74%. Esses aumentos não representam parcelas significativas para apontar a rede neural como um método para detecção prematura de falhas de curto-circuito entre espiras do gerador de indução. Porém, há outros tipos de informações relevantes à característica de operação da máquina, que podem indicar diferenças na curva de geração da máquina, como visto na Seção 3. Logo, levanta-se a hipótese que a adição da leitura de tensão do barramento CC, do valor da corrente eficaz e da frequência comandada pelo inversor no gerador (f_g) na base de dados, considerando-os como novos atributos, possa melhorar a classificação das falhas no gerador.

4.5 Adição de novos atributos à base de dados

Os experimentos feitos, a partir de agora, utilizam a mesma configuração dos hiperparâmetros da MLP definidos anteriormente. Para avaliar individualmente o impacto da adição de cada um dos atributos supracitados, realizam-se 50 treinamentos para cada um dos três. É chamada de BT-1 a base de dados original utilizada até então. BT-2 refere-se à base formada pela adição da tensão no barramento CC, enquanto que BT-3 é proveniente da adição da corrente eficaz na mesma fase onde são extraídas as harmônicas. A base BT-4 é formado pela adição de f_g . Finalmente, BT-5 a base formada ao serem adicionados os três atributos juntos. Na Figura 40 é exibida a distribuição da taxa de acerto, em todos os bancos de dados, e na Figura 41 é apresentada a dispersão dos dados em cada uma das bases.

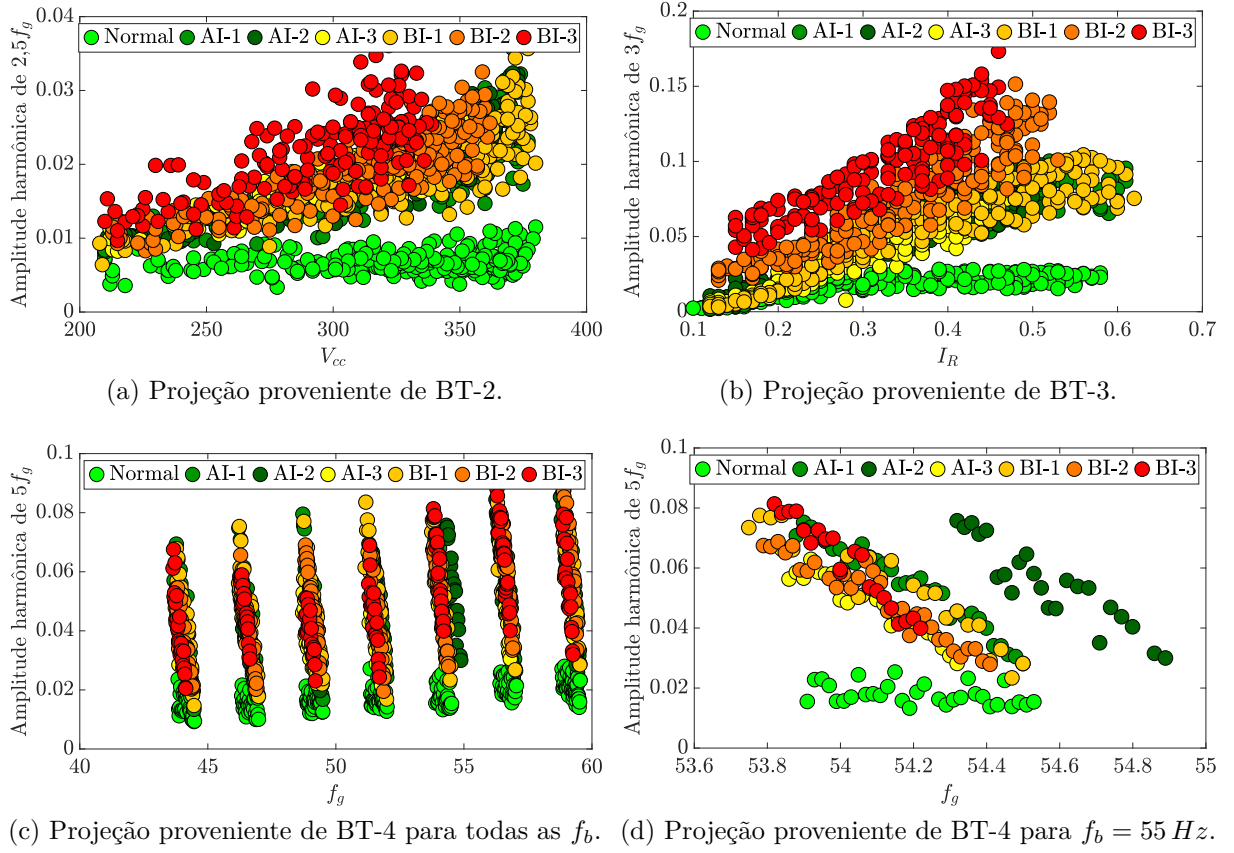
Figura 40 – Impacto na acurácia da MLP mediante adição dos atributos para formação de novas bases de dados: tensão no barramento CC (BT-2), corrente eficaz (BT-3), f_g (BT-4) e com os três propostos (BT-5).



O primeiro teste, feito em BT-2, apresentou resultados expressivos, pois a taxa de acerto subiu praticamente 10%, nos conjuntos de treinamento e de teste. Isso indica que a tensão do barramento CC contém informações relevantes para a classificação o que pode ser ratificado pela Figura 41a, pois a adição dessa informação melhorou a separabilidade dos dados Normais. No teste seguinte, feito em BT-3, a taxa de acerto desceu as condições próximas à BT-01, indicando que a valor eficaz da corrente elétrica não provê tantas informações conclusivas sobre o curto-circuito, pois o aumento de valor pode indicar também um ponto de geração distinto, como visto na Seção 3. A projeção exibida na Figura 41b ratifica essa análise.

No teste seguinte, feito em BT-4, avalia-se a influência da adição de f_g . A taxa de acertos da MLP tornou a subir, como em BT-2, indicando que esta informação é importante para a detecção da falha. A informação de f_g é pautada na dispersão dos dados, exibida na Figura 41c, pois, ao adicionar-se esse atributo, é possível separar os pontos de operação do gerador em agrupamentos distintos. Por exemplo, ao analisar-se apenas a condição quando f_b é 55 Hz, Figura 41d, a separabilidade dos dados normais é quase que linear.

Figura 41 – Projeções bi-dimensionais das bases de dados, formadas com o extrator Fourier, após a adição dos atributos V_{cc} , f_g e I_{Rrms} .



Fonte: Autor.

Por fim, adicionou-se todos os três atributos, chegando à base BT-05, e a acurácia atingida pelo classificador tornou a subir novamente, superando os testes feitos em BT-2 e BT-4. As informações da tensão do barramento CC e f_g denotam corroboração para o processo de detecção de falhas feito pela MLP. A corrente eficaz também contribuiu, porém, em uma parcela menor. A informação da corrente eficaz poderá ajudar a detecção da falha por parte do classificador, pois, em uma falha real a corrente será mais elevada que limitação imposta nos ensaios.

Visto as evidências supracitadas, é importante realizar avaliações da inserção dos mesmo atributos para os bancos de dados formados pelos extratores Goertzel, HOS e SCM. Foram feitos novos treinamentos nos classificadores OLS, Perceptron Simples, Gaussiano e MLP. Desse modo, a discussão sobre os resultados obtidos são realizadas a seguir.

4.6 Resultados das aplicações extrator *versus* classificador

Na Tabela 5 são exibidos os resultados gerais após 50 treinamentos de todas as combinações extrator-classificador, nas bases de treinamento e de teste. Usando Fourier

como extrator de atributos, a MLP alcançou em média uma taxa de acerto, entre todas as classes, de 84,48% e 76.53% nas bases de treinamento e teste, respectivamente. Isso foi melhor do que Gaussiano e, como esperado, melhor que os classificadores lineares.

A adição dos novos atributos à base de dados proporcionou melhorias em todas as combinações extrator-classificador, porém, não tornou o problema linearmente separável, o que justifica os resultados dos classificadores OLS e Perceptron Simples atingir resultados abaixo da MLP e do classificador Gaussiano. A sensibilidade indica que a MLP foi melhor do que outros métodos para identificar as condições de operação Normal do gerador. A especificidade também indica a MLP como melhor do que os outros classificadores ao fazer previsões de condições defeituosas.

O algoritmo de Goertzel, apesar de implementar a transformada de Fourier, proveu resultados abaixo da transformação clássica de Fourier para o classificador MLP. Isso é visto especialmente na base de teste, cuja as diferenças entre os resultados nas bases de treinamento e de teste é maior do que no extrator Fourier. Evidenciou-se que, com esse extrator, os classificadores Perceptron Simples e OLS obtiveram resultados muito piores que no extrator Fourier, isso indica que o uso desse extrator piorou a separabilidade da base de dados.

Ainda na Tabela 5, a combinação de HOS-MLP fica em segundo lugar na acurácia geral. Há um destaque para a combinação desse extrator com o classificador Gaussiano. Isso ocorre, pois tanto o extrator HOS como o classificador de Gaussiano baseiam suas teorias em processos Gaussianos. Os resultados desse extrator nos classificadores lineares foram melhores do que no extrator Goertzel, indicando melhor separabilidade. O extrator SCM não demonstrou-se efetivo para a classificação das falhas do gerador de indução. Mediante essas análises, identificou-se que o extrator Fourier é mais efetivo para a detecção de curto-circuitos no gerador de indução.

Tabela 5 – Resultados gerais dos extratores-classificadores para as seguintes métricas: Acurácia (Acc), Sensibilidade (Sen), Especificidade (Esp), F-score (Fsc) nas bases de treinamento e teste.

Classifier	Acc (%)	Sen (%)	Esp (%)	Fsc (%)
Treinamento				
Fourier				
MLP	84,48±2,65	84,48±2,65	97,01±0,59	84,48±2,65
Gaussiano quadrático	77,30±0,63	77,47±1,13	95,33±0,16	77,39±0,81
Perceptron simples	66,56±1,17	67,19±1,71	92,43±0,43	67,13±1,37
OLS	64,91±0,88	65,10±1,03	91,73±0,29	65,00±0,87
Goertzel				
MLP	81,42±1,46	81,42±1,46	96,33±0,34	81,42±1,46
Gaussiano quadrático	75,40±0,58	73,59±1,56	91,26±0,98	75,54±0,87
Perceptron simples	32,80±2,53	33,01±2,71	74,66±2,40	33,04±2,73
OLS	42,48±0,99	42,86±1,19	81,64±0,60	42,72±1,03
HOS				
MLP	83,31±1,69	85,31±3,57	98,82±0,74	93,31±2,03
Gaussiano quadrático	78,96±0,64	78,96±0,64	95,75±0,16	78,96±0,64
Perceptron simples	49,73±2,61	49,75±2,59	85,69±1,54	49,91±2,80
OLS	59,31±1,26	59,67±1,72	89,80±0,51	59,58±1,40
SCM				
MLP	40,93±5,58	40,93±7,97	80,61±3,56	40,93±4,75
Gaussiano quadrático	54,97±1,00	55,03±1,05	88,00±0,41	55,03±1,00
Perceptron simples	30,75±1,72	31,03±1,71	73,21±1,71	31,18±1,64
OLS	43,62±0,97	43,93±1,36	82,34±0,54	43,83±1,08
Teste				
Fourier				
MLP	76,53±3,45	76,98±3,59	95,13±0,87	76,79±3,40
Gaussiano quadrático	72,32±2,29	72,61±2,42	94,09±0,68	72,65±2,36
Perceptron simples	66,56±1,17	67,19±1,71	92,43±0,43	67,13±1,37
OLS	64,91±0,88	65,10±1,03	91,73±0,29	65,00±0,87
Goertzel				
MLP	65,59±3,16	65,59±3,16	92,16±0,94	65,92±2,99
Gaussiano quadrático	71,30±0,78	70,32±1,21	88,57±0,78	71,53±0,54
Perceptron simples	32,80±2,53	33,01±2,71	74,66±2,40	33,04±2,73
OLS	42,48±0,99	42,86±1,19	81,64±0,60	42,72±1,03
HOS				
MLP	73,54±1,42	73,54±2,31	86,82±0,34	73,54±1,57
Gaussiano quadrático	76,72±2,13	76,94±2,25	95,25±0,59	76,97±2,10
Perceptron simples	49,73±2,61	49,75±2,59	85,69±1,54	49,91±2,80
OLS	59,31±1,26	59,67±1,72	89,80±0,51	59,58±1,40
SCM				
MLP	57,15±1,52	57,24±1,52	88,88±0,62	57,20±1,51
Gaussiano quadrático	51,49±2,37	52,25±2,87	86,73±1,09	52,21±2,42
Perceptron simples	30,75±1,72	31,03±1,71	73,21±1,71	31,18±1,64
OLS	43,62±0,97	43,93±1,36	82,34±0,54	43,83±1,08

Na Tabela 6 exibem-se as taxas de acertos por classe dos classificadores propostos. Percebe-se que em todos os classificadores, a classe Normal obteve mais que 98%. A classe BI-3 foi classificada pela MLP com 100% e 99,81% de acerto nas bases de treinamento e de teste, respectivamente. E isso já era esperado pois essa é a condição de falha mais crítica da base dados. Os classificadores lineares obtiveram taxas de acerto menores que MLP para as outras classes, portanto, a MLP e o classificador Gaussiano quadrático serão analisados a partir de então.

Tabela 6 – Taxa de acertos por classes de todos os classificadores utilizando o extrator Fourier.

Class	Treinamento			
	MLP	Gaussiano quadrático	Perceptron Simples	OLS
Normal	99,98±0,11	99,14±0,33	100,00±0,00	99,81±0,34
AI-1	73,51±12,26	62,94±3,27	54,05±5,85	58,51±4,10
AI-2	66,00±8,74	54,11±3,28	24,68±13,85	27,86±5,60
AI-3	94,05±3,82	86,71±1,31	61,57±11,78	89,68±1,76
BI-1	58,94±12,36	39,49±5,91	9,51±9,82	26,30±0,68
BI-2	98,89±1,12	92,10±1,21	33,29±17,12	54,08±3,61
BI-3	100±0,00	100±0,00	77,06±18,53	98,13±0,57
Class	Teste			
	MLP	Gaussiano quadrático	Perceptron Simples	OLS
Normal	99,25±1,18	98,76±1,61	99,98±0,12	99,70±0,46
AI-1	52,16±14,23	53,80±7,01	53,32±7,67	55,53±5,09
AI-2	49,96±11,07	47,78±8,02	22,00±13,46	26,15±6,92
AI-3	85,58±6,92	82,70±6,80	61,30±13,17	89,16±3,70
BI-1	36,94±14,46	25,03±9,91	7,88±9,18	23,37±4,18
BI-2	92,44±4,09	89,10±4,46	32,93±18,76	52,71±4,38
BI-3	98,81±1,99	100±0,00	78,31±18,07	97,69±2,35

Em ordem decrescente de taxas de acertos estão as falhas BI-3, BI-2, AI-3, AI-1, AI-2 e BI-1, para os dois classificadores. Analisando apenas as falhas de baixa impedância (*i.e.* BI-1, BI-2 e BI-3), há um indício de correlação positiva entre a taxa de acerto e os níveis de espiras, isto é, a medida que aumenta-se a quantidade de espiras em curto-circuito o classificador tende a acertar mais. Por outro lado, ao olhar para as falhas de alta impedância (*i.e.* AI-1, AI-2 e AI-3), não ocorre o mesmo fenômeno, pois a falha mais incipiente, AI-1, foi melhor detectada do que a falha AI-2, nos conjuntos de treinamento e de teste. Em todos os casos a MLP obteve acurácia maior que o Gaussiano, portanto, será o classificador escolhido para o projeto do classificador no Capítulo posterior.

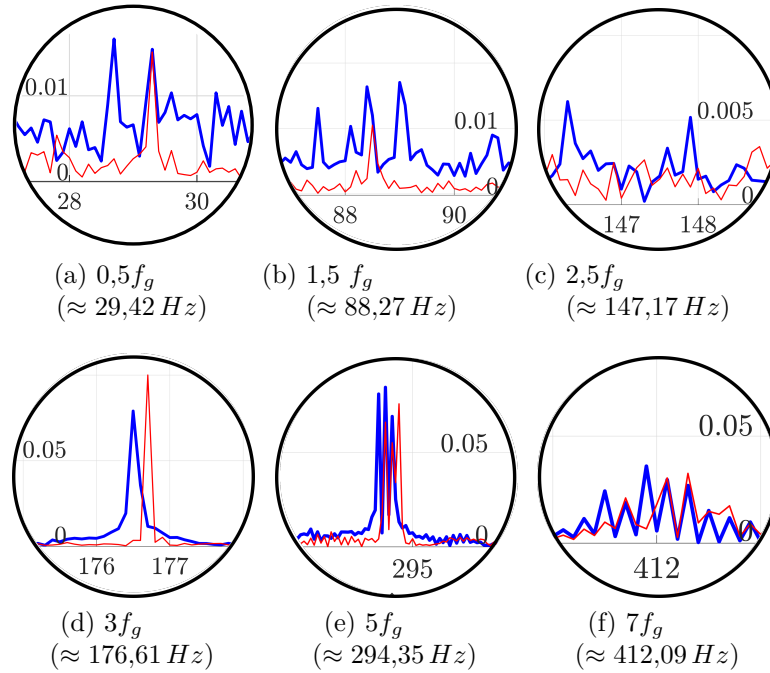
As análises supracitadas podem levar ao seguinte questionamento:

- Por que o extrator de atributos Fourier possibilitou melhores resultados que os outros?

A interpretação do espectro de frequência de Fourier explica o funcionamento, como será feito adiante.

São realizadas comparações entre as condições extremas de operação do gerador e, na Figura 42, é exibido um exemplo que possui um comparativo do espectro de frequência para $f_b = 60\text{ Hz}$ na condição de corrente nominal e de geração máxima. Para este caso, a harmônica $0,5f_g$, tanto para o caso normal quanto para o caso de falha, apresentou valores numericamente aproximados, que podem ser vistos na Figura 42a. O mesmo ocorre para as harmônicas $5f_g$ e $7f_g$, mostradas nas Figuras 42e e 42f, respectivamente. Por outro lado, as harmônicas $1,5f_g$, $2,5f_g$ e $3f_g$, Figuras 42b, 42c e 42d, apresentaram diferenças numéricas para os ambas condições. O mesmo vale para as frequência fundamentais.

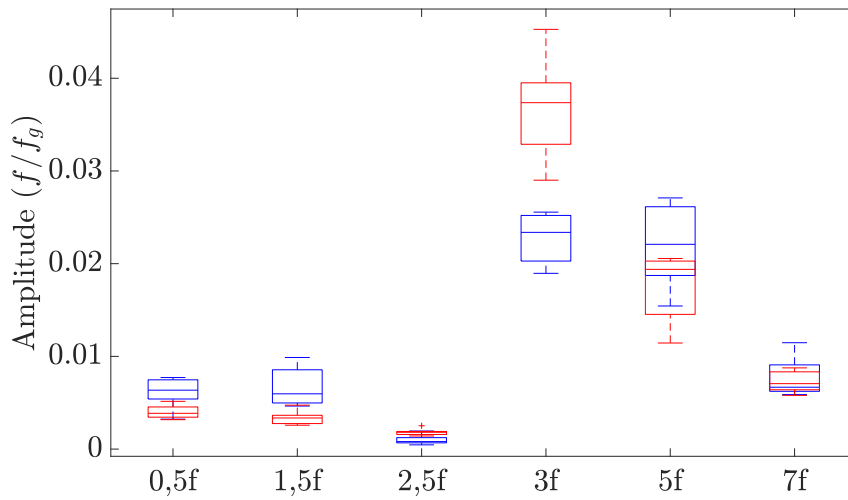
Figura 42 – Espectro de frequência em $f_g = 60\text{ Hz}$ da condição normal, com $f_g = 58,87\text{ Hz}$ e espectro da falha de CC BI NVL 3, com $f_g = 58,84\text{ Hz}$



Fonte: Autor.

Na Figura 43 exibe-se a variação dos valores harmônicos das condições normais e com falha mais críticas para todas as frequências f_b e f_g , normalizadas individualmente por suas fundamentais.

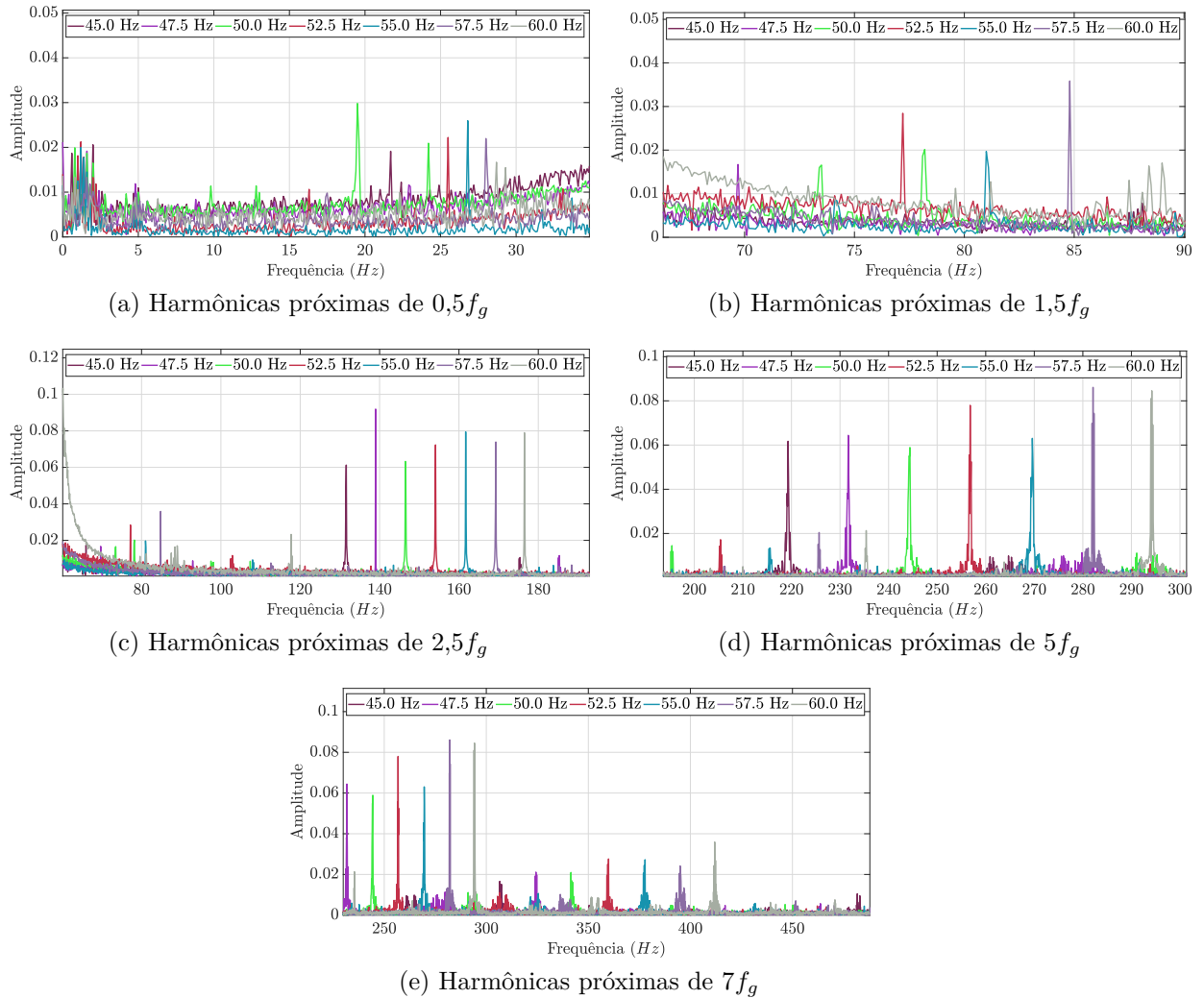
Figura 43 – Variação numérica das harmônicas normalizadas pela fundamental, da condição normal (azul) e falha de BI-3 (vermelho).



Fonte: Autor.

As variâncias dos dados mostram que: (i) a harmônica $0,5f_g$ da condição normal possui menor variabilidade que para a condição de falha e (ii) o contrário ocorre para a harmônica $1,5f_g$. As harmônicas (iii) $2,5f_g$ e (iv) $3f_g$ possuem faixa de variação distintas na base dados; 75% dos valores das harmônica $5f_g$ da classe Normal coincidem quase que 50% com os valores das Falhas BI-3. As harmônicas $7f_g$ são similares para ambas as falhas.

Na classe normal obteve-se quase que 100% de acerto, pois as harmônicas dessa classe são caracterizadas entre si, desde que saiba-se os valores da frequência de operação do conversor, f_g . As harmônicas da classe Normal, para diferentes condições de f_b , são exibidas na Figura 44. A utilização do valor de f_g na composição da base de dados é importante, pois se não houvesse essa informação, todas essas harmônicas estariam sobrepostas, podendo causar confusões com as harmônicas das classes de Falha.

Figura 44 – Espectro de frequência da condição Normal para diferentes f_b 

Fonte: Autor.

4.7 Conclusões parciais

As análises realizadas nos quatro extratores de atributos indicou que, em todos os classificadores, os resultados da utilização de Fourier para composição das bases de dados é mais efetiva. A adição dos atributos f_g , V_{cc} e I_{rms} proporcionou melhores classificações, em todos os casos. Portanto, essas informações são relevantes para realizar a detecção de curto-circuitos no SCIG.

Diante das parametrizações e treinamentos realizados nos classificadores, a rede neural MLP apresentou resultados acima dos outros classificadores. Logo, esse é o classificador escolhido para o projeto do detector de falhas incipientes de curto-circuito. O próximo capítulo trará detalhes dos resultados obtidos por esse classificador.

5 Projeto do Classificador MLP

Neste capítulo são discutidas as etapas para o projeto do detector de falhas de curto-circuito no gerador de indução. A ordem em que os resultados são apresentados segue uma cronologia do desenvolvimento da pesquisa. São levantadas hipóteses e experimentos para testá-las. À medida que os resultados são analisados, as hipóteses são refutadas ou confirmadas convergindo até os resultados finais.

O texto está organizado da seguinte forma: na Seção 5.1 são exibidos os resultados do classificador neural proposto nas bases de dados utilizadas até então. Na Seção 5.2, são impostas outras informações para o classificador, a fim de verificar sua confiabilidade de classificação. As proposições para prover robustez ao detector são discutidas na Seção 5.3.

5.1 Classificador neural proposto

Mediante os treinamentos realizados na Seção 4.4, a topologia da MLP proposta possui 10 neurônios na camada oculta, todos com função de ativação tangente hiperbólica.

Os resultados desse classificador para as bases de treinamento e teste são exibidas na matriz de confusão exposta na Tabela 7. Essa matriz apresenta percentuais de acertos das amostras por cada classe, isto é, a soma do percentual de 100% é a cada linha. Esse é o único caso em que da matriz de confusão na base de treinamento é exibida. As análises posteriores serão feitas em cima na matriz de confusão da base de testes. É importante ressaltar que apesar de nas análises posteriores não ser exibida a matriz de confusão na base de treinamento, houve o mesmo cuidado para avaliação do desempenho do classificador.

Identificou-se que a grande parte dos erros da MLP estão entre as falhas de severidades intermediárias, as quais são BI-1, BI-2, AI-1, AI-2. Isso é visto na matriz de confusão da Tabela 7. É possível ver que a falha AI-1 é, na maior parte das vezes, incorretamente classificada como BI-1. Isso acontece porque a baixa quantidade de espiras no enrolamento do estator sob curto-circuito de 1,41% não é suficiente para fornecer diferenças significativas nos espectros de frequência entre tipos de falha de alta e baixa impedância. Algumas amostras de AI-1 são classificadas como AI-2 devido à similaridade tanto na quantidade de espiras como no nível de impedância de curto-circuito.

Havia uma preocupação com os falsos negativos oriundos da falha AI-1, pois esta condição tende a ser mais semelhante à condição operacional Normal do que as outras falhas. Porém, a robustez do classificador vem com 0,93% e 1,69% dessas amostras classificadas como Normal, nas bases de treinamento e teste, respectivamente. Por outro lado, a falha AI-2 apresenta maior taxa de falsos negativos do que as outras.

Ainda na Tabela 7, é importante salientar a coerência na classificação da MLP, pois, na base de testes, os erros decrescem à medida que se distanciam do alvo (*i.e.* marcado em negrito). O mesmo acontece na aplicação da rede neural na base de treinamento, indicando que não houve *overfitting* da MLP.

Tabela 7 – Matriz de confusão percentual da combinação Fourier-MLP para a detecção em 6 níveis de falha.

Treinamento							
Rótulos	Predições						
	Normal	AI-1	AI-2	AI-3	BI-1	BI-2	BI-3
Normal	99,98%	0,00%	0,01%	0,00%	0,00%	0,00%	0,00%
AI-1	0,93%	73,50%	4,41%	0,49%	19,25%	0,85%	0,53%
AI-2	1,84%	7,44%	66%	6,38%	14,12%	3,25%	0,95%
AI-3	0,12%	0,82%	3,26%	94,04%	1,00%	0,68%	0,04%
BI-1	1,09%	28,11%	8,84%	1,26%	58,93%	1,04%	0,69%
BI-2	0,02%	0,05%	0,25%	0,49%	0,14%	98,88%	0,15%
BI-3	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	100%
Teste							
Rótulos	Predições						
	Normal	AI-1	AI-2	AI-3	BI-1	BI-2	BI-3
Normal	99,25%	0,10%	0,13%	0,28%	0,13%	0,08%	0,03%
AI-1	1,69%	52,16%	8,31%	0,91%	34,57%	1,43%	0,94%
AI-2	2,53%	10,19%	49,96%	11,36%	19,77%	5,25%	0,94%
AI-3	0,56%	0,98%	8,49%	85,58%	1,68%	2,60%	0,11%
BI-1	1,02%	43,53%	14,39%	2,08%	36,94%	1,33%	0,71%
BI-2	0,12%	0,29%	1,78%	3,00%	0,27%	92,44%	2,10%
BI-3	0,00%	0,00%	0,00%	0,06 %	0,06%	1,06%	98,81%

As falhas AI-2 e BI-1 possuem menores taxas de acerto gerais, pois, por serem falhas intermediárias, as classificações errôneas ocorrem principalmente com AI-3 e AI-1. A falha AI-3 é confundida principalmente com AI-2. Já a falha BI-1 é classificado erroneamente, em sua maioria, como AI-1, visto que a pouca quantidade de espiras não proporciona diferenças significativas entre o tipo de impedância. Algumas falhas BI-1 são classificadas como AI-2 e AI-3. Mas, as falhas AI-3, BI-2 e BI-3 têm uma taxa de acerto maior, porque a quantidade de voltas sob curto-circuito cresce ao lado de sua discrepância de condições normais. Apenas 4% e 5,92% das amostras em condições defeituosas, das bases de treinamento e de teste, respectivamente, foram classificadas erroneamente como normais, o que afirma novamente a hipótese de que as classificações erradas por MLP tendem a estar entre classes defeituosas e não com condições normais.

Os resultados dos experimentos exibidos até então foram feitos utilizando o sensor de corrente da fase R, porém, sabe-se que, na ocorrência da falha, há um desbalanceio na corrente das outras fases, causado pela distorção do campo magnético girante e que atinge

todas as fases com maior ou menor intensidade. Por esse motivos, avaliou-se a utilização das correntes das outras fases, que será exibido a seguir.

5.2 Avaliação do uso das correntes elétricas de todas as fases

Primeiramente, cria-se uma base para cada um dos sensores de corrente das fases do gerador. Manteve-se a mesma topologia e hiperparâmetros propostos para MLP com o extrator Fourier definido anteriormente. Realizam-se 50 treinamentos independentes. Os resultados da classificação nas bases de treinamento e de teste são exibidos na Figura 45.

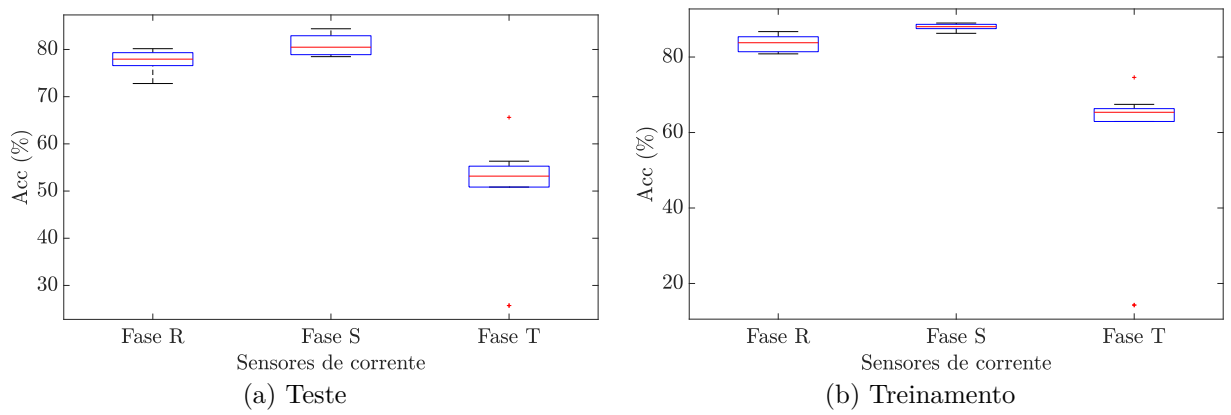
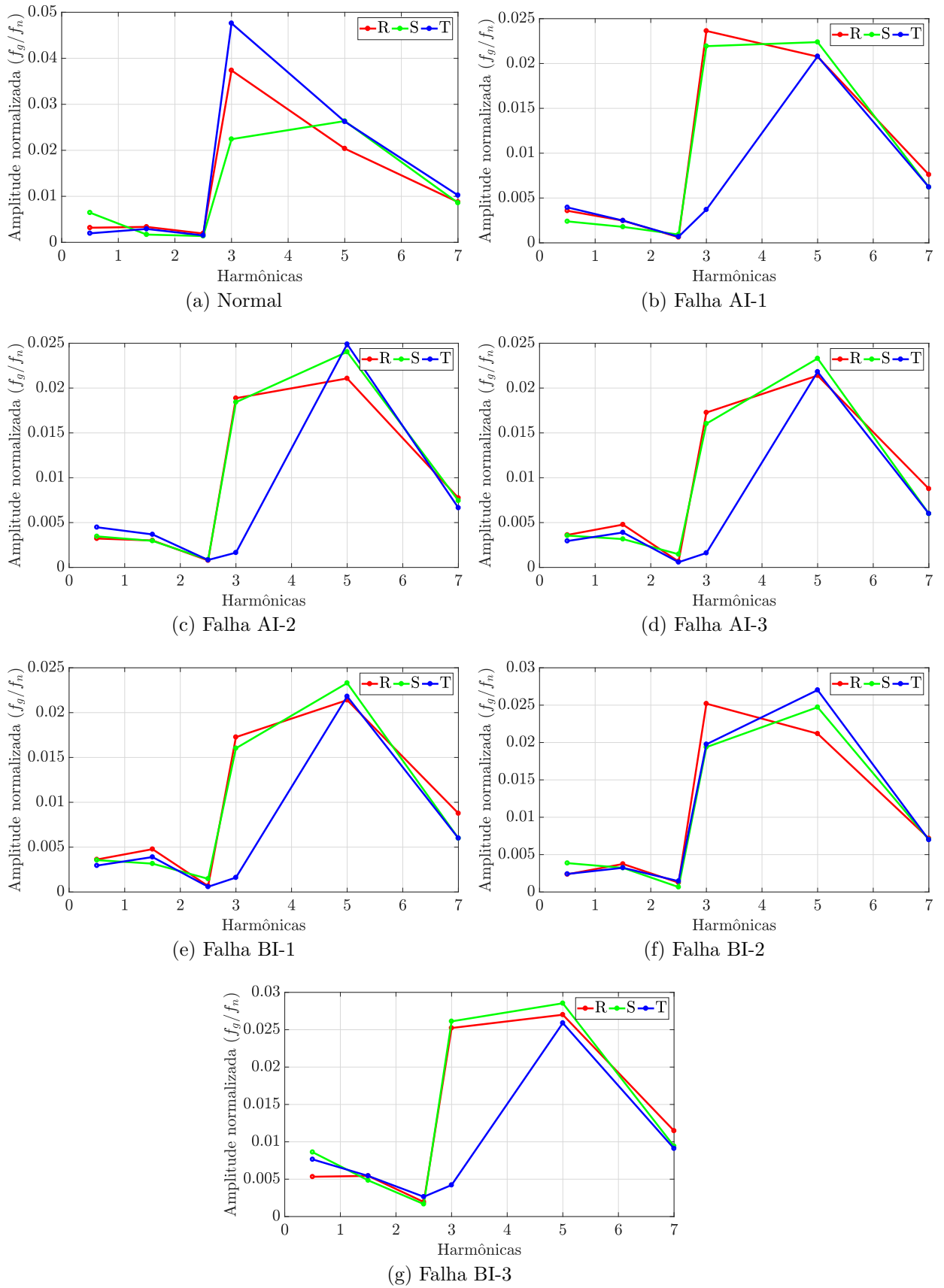


Figura 45 – *Boxplot* da acurácia da aplicação do MLP nas bases de teste e treinamento, formado por cada um dos sensores de corrente.

A base de dados formada a partir do sensor da fase S obteve resultados acima do resultados gerados pela base usado testado até então (*i.e.* fase R). Por outro lado, a base de dados formada a partir do sensor da fase T obteve resultados inferiores ao da fase R. Apesar de haver um desequilíbrio nas correntes do gerador durante a ocorrência de uma falha, a fase T é menos afetada por estar distante da inserção do curto-circuito, ao contrário das fases R e S, cuja a falha é inserida entre as mesmas.

Na Figura 46 exibem-se os valores das componentes harmônicas normalizadas por suas respectivas fundamentais, para todas as classes. E, de fato, a fase T possui uma resposta diferente nas suas harmônicas durante qualquer uma das condições de falha.

Figura 46 – Valores das harmônicas normalizadas pela fundamental, f_n , na condição de $f_b = 60\text{ Hz}$, da condição Normal e todos os tipos de falhas.



Até o momento, a base de dados utilizada para treinar o classificador foi formada pela corrente fase R e possui 1357 amostras. É importante ressaltar que na ocorrência real de um curto-circuito não se sabe em qual das fases ocorrerá o problema, portanto, treinar o classificador com apenas uma das correntes não é uma abordagem assertiva. Logo, as correntes das fases R e T são utilizadas também, mas para comporem novas amostras à base de dados, levando a um total de 4071 amostras. Espera-se que, com essa abordagem, se possa utilizar apenas um sensor de corrente, instalado em qualquer uma das fases, para detecção *a posteriori* e em tempo real. O resultados desses experimentos são expostos nas Tabelas 8.

Tabela 8 – Resultado médio da classificação da MLP nas bases de dados cujas correntes da fases S e T compõem novas amostras.

Base de dados	Teste	Treino
Classe	Acc (%)	Acc (%)
Normal	99,89±0,25	100±0,00
AI-1	38,09±13,73	41,95±10,10
AI-2	34,56±7,87	40,42±7,06
AI-3	62,35±7,92	64,33±5,44
BI-1	35,92±0,88	41,74±8,14
BI-2	64,65±3,01	71,29±4,22
BI-3	95,58±0,88	97,26±1,08
Geral	64,78±3,62	65,28±1,65

Apesar de acreditar-se que a adição das corrente das fases S e T, como novas amostras, fosse enriquecer o banco de dados para o classificador, os resultados indicaram o contrário. A acurácia geral na base de teste foi 64,78%. Contudo, a análise da acurácia por classes indica o mesmo comportamento anterior: a condição de funcionamento Normal do gerador de indução foi identificada 99,89% da vezes na base de teste. Na verdade, para essa mesma classe, houve um pequeno acréscimo da taxa de acerto comparado à base de teste composta apenas com o sensor da fase R. Isso é um indício de que a adição na outras correntes ajudaram a identificação dos dados normais perante às falhas.

Ao analisar a matriz de confusão na Tabela 9, nota-se que, em comparação à classificação da base de dados da fase R, houve uma redução drástica na taxa de falsos negativos em todos os níveis de falhas. Na verdade, a única falha que apresentou falsos negativos foi a falha BI-2, com 0,16%. A maior parte dos erros permanecem entre classes.

Tabela 9 – Matriz de confusão média proveniente da classificação da MLP na base de dados de testes cujas correntes das fases S e T compõem novas amostras. Em negrito estão as acurácias por classe.

Rótulos	Predições						
	Normal	AI-1	AI-2	AI-3	BI-1	BI-2	BI-3
Normal	99,89%	0%	0,11%	0%	0%	0%	0%
AI-1	0%	38,09%	14,00%	10,09%	26,52%	10,70%	0,61%
AI-2	0%	13,42%	34,56%	17,22%	22,03%	11,65%	1,14%
AI-3	0%	8,24%	13,18%	62,35%	3,18%	11,18%	1,88%
BI-1	0%	26,45%	19,47%	10,13%	35,92%	7,11%	0,92%
BI-2	0,16%	9,80%	9,96%	6,45%	5,22%	64,65%	3,76%
BI-3	0%	0%	0,42%	0%	0%	4,00%	95,58%

Se as todas as falhas forem agrupadas em uma única classe, isso resultará em uma classificação binária, que é exibida na Tabela 10. As taxas de falsos positivos e negativos se mantiveram abaixo de 0,5%, o que é potencialmente bom para o classificador, pois os *falsos positivos ocasionam paradas desnecessárias do aerogerador*, levando a descrédito do classificador. Os *falsos negativos* são condições em que o classificador permite o *gerador continuar funcionando, mesmo com um curto incipiente*.

Tabela 10 – Matriz de confusão média proveniente da classificação da MLP na base de dados de testes cujas correntes das fases S e T compõem novas amostras. A classificação levou em consideração a binarização da base em: classe Normal e Falha.

Rótulos	Predições	
	Normal	Falha
Normal	99,89%	0,11%
Falha	0,04%	99,96%

Os resultados anteriormente exibidos indicam que utilizar apenas um sensor de corrente para detecção de falhas de curto-circuito na máquina é válida, visto que há conversores de frequência comerciais que possui apenas um sensor de corrente embarcado. Porém, visto a importância de detectar prematuramente falhas em geradores de induções aplicados na aerogeração, acredita-se que utilizar os três sensores de corrente simultaneamente melhorará os resultados da classificação da MLP. Portanto, na próxima subseção, serão discutidos os resultados desses experimentos.

5.2.1 Harmônicas das fases S e T como novos atributos

Nesse experimento, uma nova base de dados foi composta, a partir dos sensores de correntes das fases R, S e T, porém, as harmônicas extraídas de cada uma das fase

compuseram novos atributos a mesma. A base ficou com 1356 amostras e 23 atributos, dos quais 18 são as harmônicas das três fases, o valor de tensão no barramento CC (V_{cc}), a frequência comandada no gerador (f_g) e o valor eficaz da corrente de cada uma das fases (I_{Rrms} , I_{Srms} , I_{Trms}). Os resultados são exibidos na Tabela 11.

Houve um significativo aumento na acurácia geral do classificador, ao comparar-se com os resultados da utilização apenas do sensor da fase R, exibido na Tabela 5. A taxa de acerto geral passou de 84,48% para 94,01% na base de treinamento e de 76,53% para 86,35% na base de testes. A classe Normal continuou sendo detectada em mais que 99% dos casos. Os erros continuam ocorrendo entre as classes de Falhas, contudo houve uma significativa melhora na classificação da falha AI-2 que passou de 49,96% de acertos para 92,68% na base de testes. As menores taxas de acertos são das classes AI-1 e BI-1, porém, houve um aumento em relação ao exibido na Tabela 5. De acordo com a matriz de confusão da base de testes, exibida a Tabela 12, a falha AI-3 quase não é mais confundida com as outras classes, como ocorria ao usar apenas o sensor da fase R, e o mesmo vale para as falhas BI-2 e BI-3.

Entretanto, essa abordagem tornou a subir a taxa de falsos positivos nas classes de falhas para aproximadamente 4% de todas as condições e isso pode ser potencialmente ruim, pois erros falsos negativos implicam em manter o funcionamento da máquina mesmo havendo uma condição de curto-circuito, permitindo a evolução desse processo de falha.

Tabela 11 – Resultado médio da classificação da MLP nas bases de dados cujas correntes da fases S e T compõem novos atributos.

Base de dados	Teste	Treino
Classe	Acc (%)	Acc (%)
Normal	99,70±0,70	100±0,00
AI-1	61,48±18,23	87,38±19,78
AI-2	92,68±19,28	96,00±19,79
AI-3	95,02±19,66	96,00±19,79
BI-1	58,86±17,50	86,97±19,45
BI-2	93,44±19,36	95,73±19,74
BI-3	94,69±19,66	96,00±19,79
Geral	86,35±12,85	94,01±16,52

Tabela 12 – Matriz de confusão média proveniente da classificação da MLP na base de dados de testes cujas correntes da fases S e T compõem novos atributos. Em negrito estão as acurácias por classe.

Rótulos	Predições						
	Normal	AI-1	AI-2	AI-3	BI-1	BI-2	BI-3
Normal	99,79%	0,08%	0%	0%	0,08%	0,03%	0,02%
AI-1	4,13%	61,48%	0,10%	0%	34,26%	0,03%	0%
AI-2	4,23%	0,30%	92,68%	2,15%	0,53%	0,11%	0%
AI-3	4,00%	0%	0,91%	95,02%	0,04%	0,04%	0%
BI-1	4,00%	35,96%	0,98%	0%	58,86%	0,20%	0%
BI-2	4,05%	0,27%	0,24%	0,39%	0,24%	93,44%	1,37%
BI-3	4,00%	0,06%	0%	0%	0%	1,25%	94,6%

Acredita-se que a robustez do classificador baseia-se não apenas na sua taxa de acerto, mas na sua capacidade de identificar falsos positivos e falsos negativos, além de rejeitar amostras que proporcionam confusão. Assim, para aumentar a confiabilidade do classificador, propõe-se o agrupamento das falhas por tipo e/ou intensidades. Então os resultados dos experimentos posteriores são exibidos após treinar-se a MLP com esses agrupamentos.

5.3 Aumento da confiabilidade na classificação da MLP

Novos experimentos com a MLP foram realizados, mantendo a mesma topologia e configurações adotadas nos experimentos anteriores. Esses experimentos foram realizados na seguinte ordem: (i) agrupamento por quantidade de espiras em curto; (ii) agrupamento dos dados de falha por tipo de impedância (*i.e* baixa ou alta) e (iii) agrupamento binário, rotulando todas as falhas como apenas uma.

5.3.1 Agrupamento das falhas por quantidade de espiras em curto-circuito

Os experimentos realizados nessa etapa levaram em consideração as classes de falha agrupadas em três classes, de curto-circuito em 1,41% (NVL-1), em 4,81% (NVL-2) e em 9,26% (NVL-3) de espiras. Os resultados da aplicação da MLP nessa base de teste são mostrados na Tabela 13.

Tabela 13 – Resultado médio da classificação da MLP nas bases de dados, com as falhas agrupadas pela quantidade de espiras em curto circuito de 1,41% (NVL-1), 4,81% (NVL-2) e 9,26% (NVL-3).

Base de dados	Teste	Treino
Classe	Acc (%)	Acc (%)
Normal	100±0,00	100±0,00
NVL-1	96,92±2,28	99,85±0,34
NVL-2	91,11±4,92	99,55±0,44
NVL-3	96,78±2,56	100±0,03
Geral	95,21±2,27	99,85±0,14

A taxa de acerto geral aumentou para 99,8% na base de treinamento e para 95,21% na base de teste. Mais uma vez a classe Normal é classificada corretamente em quase 100% das vezes, em ambas as bases. Um comportamento similar ao classificador treinado com as 6 falhas ocorre. A taxa de acerto do NVL-2 é a menor de todas, e, de acordo com as matrizes de confusão nos erros, estão principalmente em relação às falhas NVL-1. Isso ocorre porque as condições AI-2 são muito confundidas com AI-1, como evidenciado no classificador treinado com 7 classes.

Visto que a preocupação do trabalho é detectar curto incipiente no gerador, detectar as falhas de NVL-1 é mais vantajoso, uma vez que são as falhas mais difíceis de detectar, pois uma pequena quantidade de espiras está em curto-circuito. Isso foi obtido pelo classificador o que é possível concluir ao analisar a matriz de confusão exposta na Tabela 14. Os erros das falhas NVL-2 não são tão graves, pois foram em sua maioria confundidos com a falha incipiente de NVL-1 e com a falha de NVL-3 e apenas 0,26% são falsos negativos. Há um destaque para as falhas de NVL-1, pois a MLP as classificou corretamente 96,82% das vezes na base de treinamento. Esta é uma grande melhoria, uma vez que, na classificação de 7 classes, as falhas AI-1 e BI-1 apresentaram taxas na base de treinamento de 52,16% e 36,94%, respectivamente, e na base de teste de 61,48% e 58,86%, respectivamente, como visto na Tabela 7.

Com essa agrupamento, foi possível reduzir a taxa de falsos negativos, mas ainda buscaram-se melhorias. Portanto, experimentos utilizando um novo agrupamento das falhas foi feito e os resultados são exibidos a seguir.

Tabela 14 – Matriz de confusão média proveniente da classificação da MLP na base de dados de testes, com as falhas agrupadas pela quantidade de espiras em curto circuito de 1,41% (NVL-1), 4,81% (NVL-2) e 9,26% (NVL-3).

Rótulos	Predições			
	Normal	NVL-1	NVL-2	NVL-3
Normal	100%	0%	0%	0%
NVL-1	0,22%	96,92%	2,80%	0,05%
NVL-2	0,26%	5,13%	91,11%	3,49%
NVL-3	0,07%	0%	3,15%	96,78%

5.3.2 Agrupamento das falhas entre baixa e alta impedância

Na Tabela 15 são exibidos os resultados do experimento na base de dados com as três classes: motor normal, falha de alta impedância (AI) e falha de baixa impedância (BI). Nesse experimento, a MLP alcançou 99,97% e 99,67% de acertos nas bases de treinamento e teste, respectivamente. É importante ressaltar que a classe Normal obteve taxa de acerto de 100%, não havendo falsos positivos.

Tabela 15 – Resultado médio da classificação da MLP nas bases de dados, com as falhas agrupadas por tipo de curto (*i.e* Baixa impedância - BI e Alta impedância - AI).

Base de dados	Teste	Treino
Classe	Acc (%)	Acc (%)
Normal	100±0,00	100±0,00
AI	99,74±0,30	100±0,00
BI	99,74±0,30	99,90±0,21
Geral	99,67±0,21	99,97±0,07

Como mostra é mostrado na matriz de confusão da Tabela 16, a classe BI é confundida menos que 1% das vezes com a classe AI, levando em consideração o treinamento e o teste. Tais confusões são referentes ao nível de 1,41% de espiras em curto-circuito, que representa uma pequena quantidade de espiras, levando a classificações errôneas entre as classes AI e BI. Além disso, é importante lembrar que as falhas emuladas no gerador não são um curto-circuito real, uma vez que a corrente de curto-circuito é limitada na corrente nominal do gerador para não danificar a máquina. Então, todos os curto-circuitos emulados no gerador são na verdade uma condição de alta impedância, apesar da distinção utilizada nesse trabalho, e isso incorpora uma dificuldade a mais no processo de classificação da rede neural.

Contudo, a rede neural treinada nesse experimento apresentou robustez superior ao experimento anterior, pois diminui a ocorrência de falsos negativos para menos que

0,2% e também demonstrou ser capaz de detectar curtos incipientes de baixa impedância em uma pequena quantidade de espiras.

Tabela 16 – Matriz de confusão média proveniente da classificação da MLP na base de dados de testes, com as falhas agrupadas por tipo de curto (*i.e* Baixa impedância - BI e Alta impedância - AI).

Rótulos	Predições		
	Normal	AI	BI
Normal	100%	0%	0%
AI	0,20%	99,74%	0,06%
BI	0%	0,65%	99,35%

Visando reduzir ainda mais a taxa de falsos negativos, agrupam-se todas falhas em uma única classe, criando uma banco de dados de classes binárias. Esses resultados são exibidos a seguir.

5.3.3 Agrupamento binário

A diferença para esse experimento é que utilizou-se um conjunto desbalanceado, pois há muito mais amostras no conjunto de falhas do que no conjunto normal, e utilizar uma base de treinamento equilibrado entre classe implicaria em retirar a maior parte do conhecimento das falhas da rede. Manteve-se a proporção de 80% dos dados de cada classe separados para o treinamento e 20% para o teste. Os resultados exibidos na Tabela 17 demonstram que não há presença de falsos positivos ou falsos negativos, pois atingiu-se 100% de taxa de acerto nas bases de treinamento e de teste. Este resultado era esperado, visto que os resultados exibidos desde a Tabela 7 até a Tabela 14 já mostraram valores semelhantes à acurácia média para a classe Normal.

Tabela 17 – Resultado médio da classificação da MLP nas bases de dados, com as todas as falhas agrupadas em uma única classe.

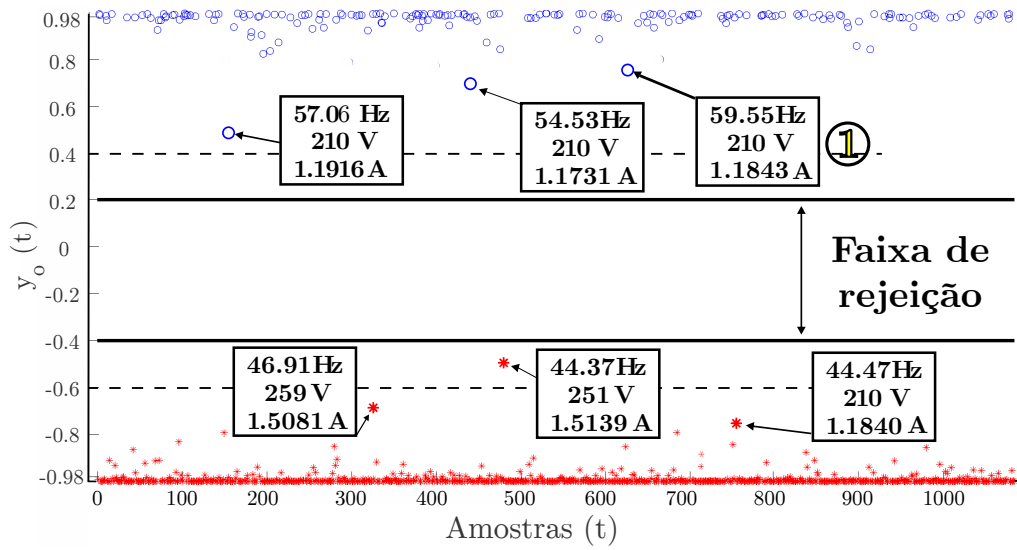
Base de dados	Teste	Treino
Classe	Acc (%)	Acc (%)
Normal	100±0,00	100±0,00
Falha	100±0,00	100±0,00
Geral	100±0,00	100±0,00

Tabela 18 – Matriz de confusão média proveniente da classificação da MLP na base de dados de testes, com as todas as falhas agrupadas em uma única classe.

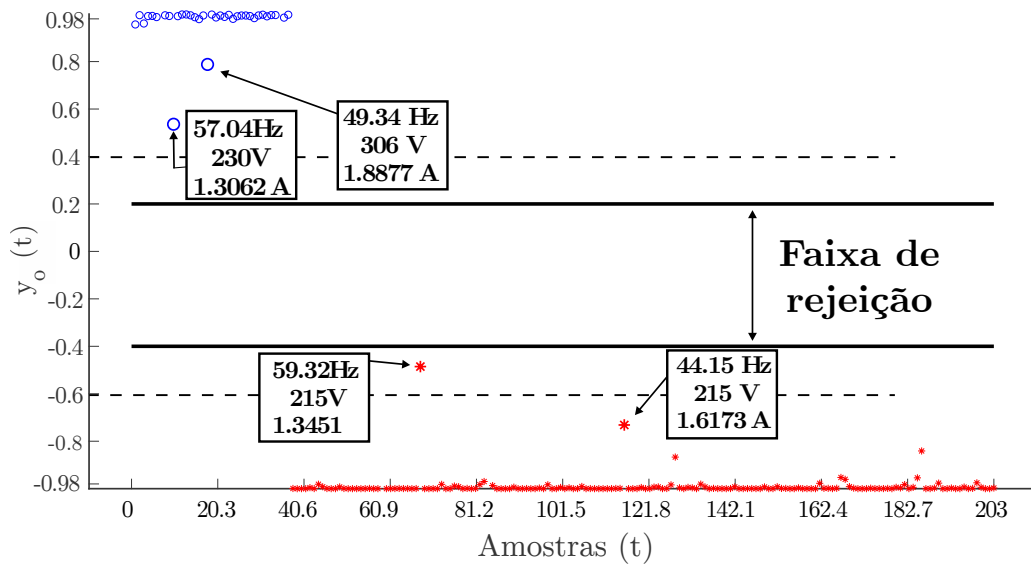
Rótulos	Predições	
	Normal	Falha
Normal	100%	0%
Falha	0%	100%

Nas Figuras 47a e 47b é apresentada a saída $y_o(t)$ da MLP para as bases de treinamento e de teste, respectivamente. É possível observar que, apesar de 100% de acurácia ter sido obtida em ambas as bases, algumas amostras são propensas a uma possível classificação errônea, como as marcadas nos gráficos. Para evitar que amostras deste tipo sejam erroneamente classificadas, foi introduzida uma faixa de rejeição na saída do MLP. Na prática, se alguma amostra estiver dentro desse intervalo, ela será rejeitada pela rede e uma nova amostra será analisada. Este procedimento garante ainda mais robustez a rede neural pois reduz-se a ocorrência de casos falsos positivos e falsos negativos. No caso específico deste trabalho, foi implementado um intervalo de rejeição assimétrica entre +0,2 e -0,4 para priorizar a diminuição de falsos positivos, uma vez que a classificação errônea da condição Normal como Falha ocasiona a parada desnecessária da máquina, implicando em indisponibilidade do aerogerador e levando a descrédito no sistema de detecção de falhas.

Figura 47 – Saída da MLP para a base de dados binária.



(a) Saída para a base de treinamento.



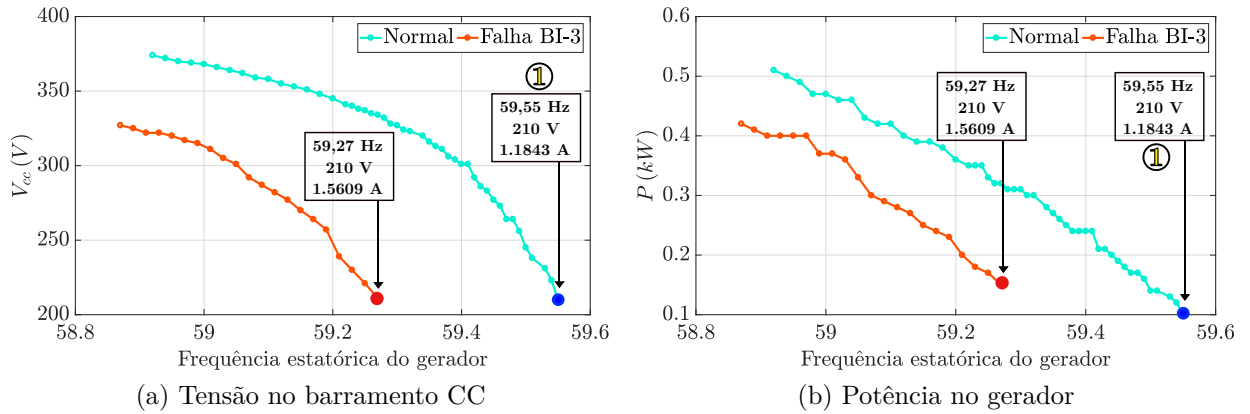
(b) Saída para a base de teste.

A análise da base de dados de treinamento na Figura 47a indica que amostras de condições normais que estão mais próximas do limiar têm baixa tensão no barramento CC, em torno de 210 V. Isso ocorre em diferentes frequências do gerador. Nas condições defeituosas ocorre o mesmo, em amostras próximas ao limite, a tensão no barramento CC está abaixo do valor nominal de 311 V. A saída da MLP para a base de dados de teste, apresentado na Figura 47b, mostra o mesmo padrão, com exceção da amostra de 49,34 Hz, onde o tensão no barramento CC atingiu 306 V, sendo esse exemplo um possível *outlier* gerado durante os ensaios.

Esse fenômeno pode ocorrer porque o desequilíbrio produzido pelo curto-circuito deprecia a capacidade de geração. Esta hipótese é afirmada pela tensão medida no

barramento CC do gerador modificado durante os ensaios. Na Figura 48 é exibido um exemplo, nos quais as curvas de tensão e de potência por frequência do gerador nas condições Normal e Falha BI-3 são mostradas. O ponto marcado pelo número 1 nas Figuras 47a, 48a e 48b são a mesma amostra. Por exemplo, quando f_g é 59,55 Hz e $f_b = 60$ Hz, iniciou-se a geração de energia. Esse momento pode ser interpretado como uma condição de baixa velocidade do vento, quando a geração está reduzida, e por consequência a potência gerada está muito abaixo de seu valor nominal, o que é visto nas Figuras 48b e 48a, respectivamente. Pelo mesmo motivo, a tensão no barramento CC do conversor de frequência é inferior ao valor nominal. Nas curvas em vermelho estão representadas a Falha BI-3. É possível observar que a mesma tensão no barramento CC e potência ocorrem em $f_g = 59,55$ Hz, para a condição Normal e em $f_g = 59,27$ Hz, para a condição de Falha. Isso pode levar a uma classificação errada. O mesmo pode ser interpretado para os outros pontos marcados na Figura 47.

Figura 48 – Curvas de tensão e potência, em uma mesma condição de f_g para a Classe Normal e Falha BI-3.

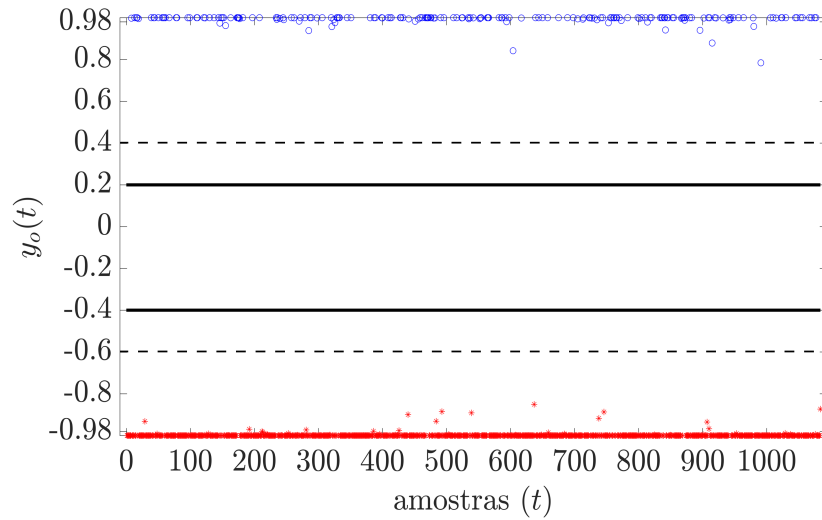


Essas evidências indicam que a diferença de frequência da base (f_b) e da frequência do gerador (f_g), chamada de frequência relativa, f_r , é diretamente proporcional a tensão no barramento CC e à potência gerada. Isso incitou a hipótese de que utilizar f_r como atributo ao invés de f_g pode melhorar a classificação da MLP. Foi realizado um novo experimento, mantendo a mesma topologia da MLP e um agrupamento binário da base de dados, apenas alterando esse atributo. Os resultados de classificação mantiveram-se os mesmos, porém, a separabilidade das bases de dados melhorou, como pode ser visto nas Figuras 49a e 49b. Nas Figuras 47 e 49, a faixa tracejada foi colocada para feitos comparativos. A margem de separação dos dados aumentou, tornando o classificador ainda mais robusto a falsos positivos.

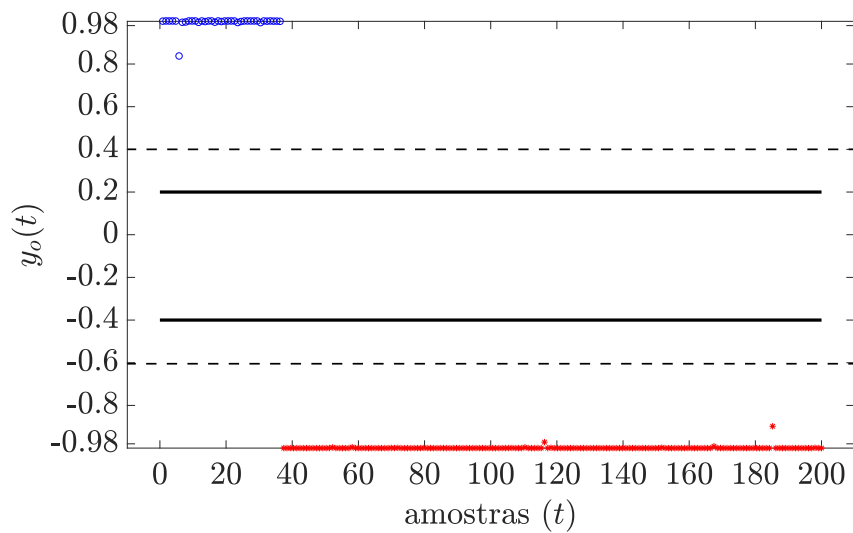
Logo, as evidências indicaram que o ideal é fornecer ao classificador informação de f_r , porém, não fazia parte do escopo dessa pesquisa utilizar um sensor de velocidade acoplado ao gerador. Nesses casos, pode-se utilizar a informação de f_g para a rede neural,

pois apenas a informação de tensão no barramento CC poderia ser tendenciosa a uma má classificação. Mesmo em condições confusas, a confiabilidade dos classificadores é mantida, pois nenhuma amostra foi classificada erroneamente e a faixa de rejeição evita falsas classificações positivas.

Figura 49 – Saída da MLP para a base de dados binária, utilizando a frequência relativa como atributo.



(a) Saída para a base de treinamento.



(b) Saída para a base de teste.

6 Conclusões

Neste trabalho, buscou-se realizar a detecção e classificação das falhas de curto-circuito entre espiras do bobinamento estatórico do gerador de indução gaiola de esquilo, aplicado na geração eólica. Para tal, foi necessário construir um sistema para emulação e aquisição de dados do aerogerador. Após, utilizou-se quatro técnicas de extração de atributos e algoritmos de aprendizado de máquina para detecção incipiente das falhas de curto-circuito. Os objetivos traçados foram alcançados com êxito e suas discussões são expostas a seguir.

O projeto da bancada de experimentos foi executado com sucesso. As etapas inerentes a esse projeto foram: (i) construir um quadro de comando, projetado especificamente para esse propósito, que contivesse um sistema para emulação das falhas de curto-circuito; (ii) adequar a máquina de indução trifásica, modificada, para emular um sistema de aerogeração do tipo *Full-scale* e *Full-variable-speed*; (iii) criar uma interface para aquisição de dados utilizando *Lab View* e (iv) especificar transdutores para aquisição dos dados de corrente elétrica, fluxo magnético axial e vibração mecânica. Os experimentos para aquisições dos dados seguiram padrão metodológicos adequados que possibilitaram, inclusive, as aquisições de novos dados para futuras pesquisas. É relevante destacar o alcance do êxito desses objetivos, pois somente quando esta etapa foi concluída e validada é que foi possível avançar a abordagem proposta.

Dentro os quatro extratores de atributos escolhidos, três são técnicas para análises no domínio da frequência e uma para o domínio do tempo do sinal de corrente elétrica do gerador. Formaram-se quatro bases de dados distintas, que serviram para o treinamento dos classificadores propostos. Os resultados indicaram que nenhuma das bases de dados é linearmente separável, pois os classificadores OLS e Perceptron simples não obtiveram sucesso na classificação das falhas. A melhor taxa de acerto obtida foi de 64,91% e 66,56% na base formada pelo extrator Fourier. Dentre as configurações do classificador Gaussiano, o discriminante quadrático foi melhor, pois obteve 76,72% de acertos gerais na base formada pelo extrator HOS. A combinação MLP-Fourier demonstrou-se a mais efetiva para o problema exposto nessa pesquisa.

Ao compor as informações de corrente elétrica como novas amostras à base de dados a classificação geral da MLP piorou (*i.e* 64,78%), porém, a classificação binária dessa mesma rede atingiu 99,89% de acerto na classe Normal e 99,96% nas classes de Falha, e a taxa de falsos negativos diminuiu para menos de 0,5%. Com isso, conclui-se que é possível identificar condições de falha e não-falha do gerador de indução utilizando apenas uma corrente elétrica.

A detecção prematura das falhas de curto-circuito foi possível ao utilizar as informações dos três sensores de corrente elétrica como atributos. A taxa de acerto das falhas mais incipientes (*i.e.* em 1,41% de espiras - AI-1 e BI-1) subiram de 38,09% e 35,92% para 61,48% e 58,86% respectivamente. Nessa proposta, manteve-se também a taxa de detecção de condições normais acima de 99%, porém, a taxa de falsos negativos subiu para ordem de 4%.

Vislumbrando projetar um classificador para operação em tempo real, foi explorada a possibilidade de aumentar a confiabilidade do classificador neural ao propor diferentes topologias. A rede utilizada para classificar as falhas quanto a quantidade de espiras obteve 95,21% de acertos com 0,26% de taxa de falsos positivos. A rede neural proposta para identificar falhas de curto-circuito por tipo de impedância foi capaz de reduzir a taxa de falsos positivos para 0,2% e aumentar a taxa de acerto para 99,67%. Conclui-se que isso provê robustez à classificação da MLP, pois credita confiança no sistema de detecção de falhas, evitando paradas desnecessárias no parque eólico. Por fim, a proposição da rede binária garantiu taxa de acerto de 100% em ambas as condições de funcionamento da máquina, acima do que quando utiliza-se apenas um sensor de corrente. Em todas as topologias propostas a taxa de falsos positivos manteve-se em 0% e conclui-se que isso garante a parada prematura da máquina elétrica na ocorrência de um curto incipiente.

A evolução dos experimentos exibiu que a condição de funcionamento Normal do gerador foi identificada em quase que 100% das ocorrências. Os erros da MLP estavam entre os níveis de falhas de curto-circuito, especialmente com 1,41% de espiras em curto-circuito. Dessa forma, não houveram erros falso positivos, que, como discutidos, implicam na disponibilidade energética do parque eólico.

6.1 Conclusões à luz das hipóteses levantadas

Ao longo do trabalho, foram levantadas hipóteses e projetados experimentos para seus testes, feitos na ordem cronológica da aparição das evidências encontradas na pesquisa. Dessa forma, as conclusões à luz das hipóteses são exibidas a seguir.

Havia uma percepção de que a transformada de Fourier poderia ser uma ferramenta promissora para esse problema, por possuir ser uma técnica consolidada e vastamente utilizada na literatura, e, de fato, foi afirmado pelos resultados dos experimentos exibidos no Capítulo 5.

Entretanto, a busca para melhorar os resultados dos classificadores instigou a hipótese que a adição de novos atributos no banco de dados melhoraria a classificação da MLP. Os atributos escolhidos foram: a tensão no barramento CC; corrente eficaz e frequência comandada no conversor de frequência. Essa hipótese foi confirmada pois a classificação feita pela MLP obteve um acréscimo de 13,23%.

A base de dados dos testes iniciais para os classificadores foi montada com a corrente da fase R, que estava no ramo do curto-circuito. Acreditava-se que poderia utilizar apenas uma corrente elétrica para identificação de problemas no gerador, pois pesquisas anteriores haviam proposto tal abordagem para classificação de falhas em motores de indução. Porém, essa hipótese foi parcialmente confirmada, visto que os resultados dos experimentos exibidos na Seção 5.2 promoveram um decréscimo na classificação da MLP, diminuindo principalmente a classificação das falhas incipientes. Seguido desse experimento, formulou-se que utilizar as harmônicas provenientes dos três sensores de corrente elétrica simultaneamente poderia provê melhores resultados. De fato, isso foi confirmado, pois a classificação da MLP atingiu 94,01%, com 100% de acerto da classe Normal em ambas as bases de treinamento e de teste.

Utilizar a informação de frequência relativa é o ideal para a tornar o classificador mais robusto, entretanto, isso implica em utilizar um sensor de velocidade acoplado ao rotor do gerador.

6.2 Limitações

Embora o classificador proposto nesse trabalho tenha atingido 100% de acerto na classificação binária e menos de 1% de falsos negativos para a classificação em grupos de impedância de curto-circuito, há limitações para a sua utilização. Primeiramente, os ensaios foram feitos em uma máquina elétrica modificada, cujas emulações dos curto-circuitos foram limitadas a corrente nominal por questões de repetibilidade do ensaio. Seria válido realizar testes da operação do classificador em uma situação de curto-circuito real, porém, por falta de recursos não havia uma máquina elétrica disposta para um ensaio destrutivo. Entretanto, visto que o classificador foi apto a identificar as falhas abordadas nesse trabalho, que caracterizam condições extremamente incipientes, espera-se que em uma condição real o mesmo ocorra.

Apesar da análise dos resultados da aplicação de diversos extratores exibida na Seção 4.6 indicar a transformada de Fourier como uma ferramenta promissora, é importante salientar que o extrator HOS exibiu resultados que merecem investigações que fugiam o escopo dessa pesquisa. A sua utilização pode ser vantajosa para sistemas embarcados, pois visto que as operações matemáticas limitam-se ao domínio do tempo, há uma percepção de ser uma abordagem mais simplista para esse tipo de aplicação.

6.3 Considerações gerais

A importância desse trabalho permeia sua natureza multidisciplinar, por apresentar conceitos de instrumentação eletrônica, processamento de sinais e inteligência

computacional aplicada ao campo de estudo da máquina elétrica de indução.

Uma importante contribuição reside na metodologia desde a aquisições dos dados até a detecção e classificação das falhas de curto-circuito. É importante ter em mente que essa abordagem pode ser facilmente estendida à problemas similares. Por exemplo, outros sistemas de geração eólica que utilizem geradores de indução podem utilizar a mesma abordagem. Por exemplo, máquinas como o DFIG e o WRIG também se beneficiariam do sistema, visto que possuem características eletromagnéticas similares ao SCIG.

O sistema proposto nesse trabalho foi desenvolvido com sensores de corrente elétrica instalados na linha de alimentação do gerador. Porém, o mesmo trabalho pode ser feito com os sensores de corrente elétrica já existentes no conversor de frequência. Dessa forma, o sistema para detecção de falha estará embarcado junto com o módulo conversor. Inclusive, o estudo levou em consideração conversores que utilizem um ou três sensores de corrente elétrica.

6.4 Sugestões para trabalhos futuros

Esse trabalho faz parte das pesquisas desenvolvidas pelo grupo de pesquisa de controle e processamento de energia do Instituto Federal de Educação, Ciência e Tecnologia do Ceará. Tal grupo já possui outros trabalhos que relaciona máquinas de indução e aprendizado de máquina. Então, os desenvolvimentos realizados nessa pesquisa poderão ajudar os seguintes estudos:

- Embarcar o classificador neural em um sistema microcontrolado;
- Avaliar a utilização dos outros sinais, coletados durante a realização dos ensaios;
- Aplicar outras técnicas de aprendizado de máquina, como SVM (do inglês *Support Vector Machine*) (CORTES; VAPNIK, 1995);

Referências

- ALBIZU, I. et al. On-line stator winding fault diagnosis in induction generators for renewable generation. **Proceedings of the 12th IEEE Mediterranean Electrotechnical Conference (IEEE Cat. No.04CH37521)**, v. 3, n. 690 V, p. 1017–1020, 2004.
- ALNASIR, Z. et al. CCECE 2014 1569877045 Performance Comparison of Standalone SCIG and PMSG-Based Wind Energy Conversion Systems. p. 1–8, 2014.
- AN, X.; JIANG, D.; LI, S. Application of back propagation neural network to fault diagnosis of direct-drive wind turbine. **2010 World Non-Grid-Connected Wind Power and Energy Conference**, p. 1–5, 2010.
- ANTONI, J. The spectral kurtosis: A useful tool for characterising non-stationary signals. **Mechanical Systems and Signal Processing**, v. 20, n. 2, p. 282–307, 2006. ISSN 08883270.
- ASFANI, D. A.; PURNOMO, M. H.; SAWITRI, D. R. Naïve Bayes classifier for Temporary short circuit fault detection in Stator Winding. p. 288–294, 2013.
- AZIZI, R. et al. Improving accuracy of cavitation severity detection in centrifugal pumps using a hybrid feature selection technique. **Measurement**, v. 108, p. 9–17, 2017.
- B, S.; KR, A.-B. Artificial neural network based fault diagnostics of rolling element bearings using time-domain features. **Mechanical Systems and Signal Processing**, v. 17, n. 2, p. 317–328, 2003.
- BARRETT, R.; NING, A. Comparison of Airfoil Precomputational Analysis Methods for Optimization of Wind Turbine Blades. p. 1–8, 2016. ISSN 19493029.
- BEEBE, R. S. **Predictive maintenance of pumps using condition monitoring**. [S.l.: s.n.], 2004. ISBN 1856174085.
- Ben Hassena, M. A. et al. Validation of a New Structural Health Monitoring Technique of a Wind Turbine Prototype. **12th International Multi-Conference on Systems, Signals and Devices (SSD'15)**, p. 1–5, 2015.
- BISHOP, C. M. **Pattern Recognition and Machine Learning**. [S.l.: s.n.], 2013. v. 53. 1689–1699 p. ISSN 1098-6596. ISBN 9788578110796.
- BLUM, A. **Neural networks in C++: an object-oriented framework for building connectionist systems**. [S.l.]: John Wiley & Sons, Inc., 1992.
- BONNETT, A. H. Root cause failure analysis for AC Induction Motors in the petroleum and chemical industry. **Petroleum and Chemical Industry Conference (PCIC), 2010 Record of Conference Papers Industry Applications Society 57th Annual**, p. 1–13, 2010.

- BONNETT, A. H.; SOUKUP, G. C. Cause and analysis of stator and rotor failures in 3-phase squirrel cage induction motors. **Pulp and Paper Industry Technical Conference, 1991., Conference Record of 1991 Annual**, p. 22 – 42, 1991.
- BOOKMAN, T. T. S focus: Wind energy's promise, offshore. **IEEE Technology and Society Magazine**, v. 24, n. 2, p. 9–15, 2005.
- BOSE, N. K.; GARGA, A. K. Neural network design using Voronoi diagrams. **IEEE Transactions on Neural Networks**, IEEE, v. 4, n. 5, p. 778–787, 1993.
- CARDOSO, J. G. et al. SCIG wind turbine wireless controlled using morphological filtering for power quality enhancement. **Renewable Energy**, v. 92, p. 303–311, 2016.
- CHAPMAN, S. J. Fundamentos de Máquinas Elétricas. **Bookman**, 2013.
- COELHO, D. N. **Técnicas de Aprendizado de Máquina para Detecção de Falhas em Motores de Indução Trifásicos**. Tese (Doutorado), 2015.
- COELHO, D. N. et al. detection of Short Circuit Incipient Fault in a Three- Phase Induction Motor. p. 1–7, 2014.
- CORTES, C.; VAPNIK, V. Support Vector Networks. **Machine Learning**, v. 20, n. 3, p. 273–297, 1995.
- DUAN, Y.; HARLEY, R. G. Present and Future Trends in Wind Turbine Generator Designs. n. 11, 2009.
- DUTRA, R. M.; SZKLO, A. S. Incentive policies for promoting wind power production in Brazil: Scenarios for the Alternative Energy Sources Incentive Program (PROINFA) under the New Brazilian electric power sector regulation. **Renewable Energy**, v. 33, n. 1, p. 65–76, jan 2008.
- DWYER, R. Detection of non-Gaussian signals by frequency domain Kurtosis estimation. **ICASSP '83. IEEE International Conference on Acoustics, Speech, and Signal Processing**, v. 8, p. 607–610, 1983.
- EGEA-ALVAREZ, A. et al. Control of a wind turbine cluster based on squirrel cage induction generators connected to a single VSC power converter. v. 61, p. 523–530, 2014.
- Empresa de Pesquisa Energética. **Demanda de Energia 2050**. [S.l.], 2014. Disponível em: <<http://www.pac.gov.br/noticia/13554306>>.
- FISHER, R. A. On the mathematical foundations of theoretical statistics. **Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character**, JSTOR, v. 222, p. 309–368, 1922.
- FISHER, R. A. On the Mathematical Foundations of Theoretical Statistics. **Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences**, v. 222, n. 594-604, p. 309–368, 1922. ISSN 1364-503X.
- FITZGERALD, A. E. **Máquinas Elétricas**. [S.l.]: Bookman, 2006.
- FRANCHI, C. M. **Acionamentos Elétricos**. [S.l.]: Érica, 2008.
- FRANCISCO, A. M. S. Motores de Indução Trifásicos. 2006.

- FRIEDMAN, J. H. Regularized Discriminant. **Journal of the American Statistical Association**, n. July, 1988.
- Fucaí Li et al. Wavelet Transform-based Higher-order Statistics for Fault Diagnosis in Rolling Element Bearings. **Journal of Vibration and Control**, v. 14, n. 11, p. 1691–1709, 2008. ISSN 1077-5463. Disponível em: <<http://jvc.sagepub.com/content/14/11/1691.abstract>>. <<http://jvc.sagepub.com/cgi/doi/10.1177/1077546308091>>.
- Global Wind Energy Council. Global Wind Report. 2015.
- GOERTZEL, G. An Algorithm for the Evaluation of Finite Trigonometric Series. **The American Mathematical Monthly**, v. 65, n. 1, p. 34–35, 1958.
- GRADIN, L.; BURSTEIN, N. Assessment Of Containment Fan Cooler Degradation By Motor Current Signature Analysis. **1993 IEEE Conference Record Nuclear Science Symposium and Medical Imaging Conference**, p. 990–994, 1993.
- GWEC. Global Wind Report 2016. p. 76, 2017. Disponível em: <<http://files.gwec.net/files/GWR2016.pdf>>.
- HAHN, B.; DURSTEWITZ, M.; ROHRIG, K. Reliability of Wind Turbines. **Wind Energy**, p. 1–4, 2007.
- HAYKIN, S. Neural Networks: a comprehensive foundation. **Pearson Education**, v. 2. ed, 2001.
- HENAO, H.; DEMIAN, C.; CAPOLINO, G. A. A frequency-domain detection of stator winding faults in induction machines using an external flux sensor. **IEEE Transactions on Industry Applications**, v. 39, n. 5, p. 1272–1279, 2003.
- HOSSAIN, M. M.; ALI, M. H. Future research directions for the wind turbine generator system. **Renewable and Sustainable Energy Reviews**, Elsevier, v. 49, p. 481–489, 2015.
- International Energy Agency. **Key world energy statistics**. [S.l.], 2017. Disponível em: <<http://www.iea.org/publications/freepublications/publication/KeyWorld2017.pdf>>.
- JARZYNA, W. Diagnostic characteristics of axial flux in an induction machine. n. 4, p. 141–146, 1995. ISSN 05379989.
- JEONG, Y.-j.; PARK, M.-s.; LEE, D.-h. Shape Dependent Wave Force and Bending Moment of Offshore Wind Substructure System. p. 0–3, 2015.
- JU, L. et al. Fault predictive diagnosis of wind turbine based on LM arithmetic of artificial neural network theory. **Proceedings - 2011 7th International Conference on Natural Computation, ICNC 2011**, v. 1, p. 575–579, 2011. ISSN 2157-9555.
- KAUFHOLD, M. et al. Interface phenomena in stator winding insulation-Challenges in design, diagnosis, and service experience. **IEEE Electrical Insulation Magazine**, v. 18, n. 2, p. 27–36, 2002. ISSN 08837554.
- KLIMAN, G. et al. A new approach to on-line turn fault detection in AC motors. **IAS '96. Conference Record of the 1996 IEEE Industry Applications Conference Thirty-First IAS Annual Meeting**, v. 1, p. 687–693, 1996. ISSN 0197-2618.

- KONG, X.; JIA, H. Techno-economic analysis of SVC-HVDC transmission system for offshore wind. **Asia-Pacific Power and Energy Engineering Conference, APPEEC**, 2011. ISSN 21574839.
- KUSIAK, A.; LI, W. The prediction and diagnosis of wind turbine faults. **Renewable Energy**, Elsevier Ltd, v. 36, n. 1, p. 16–23, 2011.
- LAU, B. C. P.; MA, E. W. M.; PECHT, M. Review of offshore wind turbine failures and fault prognostic methods. **Proceedings of IEEE 2012 Prognostics and System Health Management Conference, PHM-2012**, 2012. ISSN 2166-563X.
- LIMA, D. K. S. et al. Estimating the offshore wind resources of the State of Ceará in Brazil. **Renewable Energy**, v. 83, p. 203–221, 2015.
- MANRIQUE, R. F.; GIRALDO, F. A.; ESMERAL, J. S. Fault detection and diagnosis for wind turbines using data-driven approach. **2012 7th Colombian Computing Congress (CCC)**, p. 1–6, 2012.
- MARINHO, L. B. et al. A novel mobile robot localization approach based on topological maps using classification with reject option in omnidirectional images. **Expert Systems with Applications**, Elsevier Ltd, v. 72, p. 1–17, 2017.
- MARKS, S.; DUNN, O. J. American Statistical Association. v. 51, n. 273, p. 111–121, 1974.
- MARTIN, H. R.; HONARVAR, F. Application of statistical moments to bearing failure detection. **Applied Acoustics**, v. 44, n. 1, p. 67–77, 1995. ISSN 0003682X.
- MASTERS, T. **Practical neural network recipes in C++**. [S.l.]: Morgan Kaufmann, 1993.
- MATOS, F. F. V. et al. a Multilevel Wind Power Conversion System With an Open Winding Squirrel Cage Induction Generator. p. 1–6, 2015.
- MAY, A.; MCMILLAN, D.; THÖNS, S. Economic analysis of condition monitoring systems for offshore wind turbine sub-systems. **IET Renewable Power Generation**, v. 9, n. 8, p. 900–907, 2015.
- MCCULLOCH, W. S.; PITTS, W. A Logical Calculus of the Idea Immanent in Nervous Activity. **Bulletin of Mathematical Biophysics**, v. 5, p. 115–133, 1943. ISSN 0007-4985. Disponível em: <<http://www.cse.chalmers.se/{~}coquand/AUTOMATA/mcp.p>>.
- MEDEIROS, C. M. d. S. Notas de aula: Sistemas Lineares. 2014.
- MEDEIROS, C. M. d. S.; BARRETO, G. A. Notas de aula: Inteligência computacional aplicada. 2015.
- MENDEL, J. M. Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications. **Proceedings of the IEEE**, v. 79, n. 3, p. 278–305, 1991.
- MESEMANOLIS, A.; MADEMLIS, C.; KIOSKERIDIS, I. Optimal Efficiency Control Strategy in Wind Energy Conversion System With Induction Generator. **IEEE Journal of Emerging and Selected Topics in Power Electronics**, v. 1, n. 4, p. 238–246, 2013.

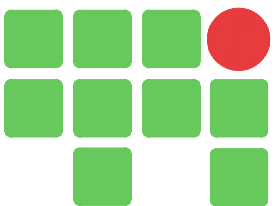
- Ministério de Minas Energia. **Atlas do Potencial Eólico Brasileiro**. [S.l.], 2001. 44 p.
- MINSKY, M.; PAPERT, S. A.; BOTTOU, L. **Perceptrons: An introduction to computational geometry**. [S.l.]: MIT press, 2017.
- MORSALIN, S. et al. Induction motor inter-turn fault detection using heuristic noninvasive approach by artificial neural network with Levenberg Marquardt algorithm. **IEEE International Conference on Informatics, Electronics & Vision (ICIEV)**, p. 1–6, 2014.
- MUDI, J.; SINHA, D. Comparative study among different wind turbines used for wind energy system. **Conventional Energy (ICONCE)**, 2014 1st, n. Iconce, p. 175–179, 2014.
- MÜLLER, S.; DEICKE, M.; W., & Rik Doncker, D. Doubly fed induction generator systems for wind turbines. **Industry Applications Magazine, IEEE**, v. 8, n. 3, p. 26–33, 2002.
- NANDI, S.; TOLYAT, H. A.; LI, X. Condition monitoring and fault diagnosis of electrical motors - A review. **IEEE Transactions on Energy Conversion**, v. 20, n. 4, p. 719–729, 2005.
- National Instruments. **NI USB-6009 Device Specifications**. [S.l.]: IEEE, 2015. 1–10 p.
- OH, K. Y. et al. A Novel Method and its Field Tests for Monitoring and Diagnosing Blade Health for Wind Turbines. **IEEE Transactions on Instrumentation and Measurement**, v. 64, n. 6, p. 1726–1733, 2015. ISSN 00189456.
- OLIVEIRA, Á. G.; SILVA, R.; MEDEIROS, C. M. S. Neural network used to stator winding interturn short-circuit fault detection in an induction motor driven by frequency converter. **Proceedings - 1st BRICS Countries Congress on Computational Intelligence, BRICS-CCI 2013**, p. 459–464, 2013.
- OLIVEIRA, A. G. de. **Classificadores neurais aplicados na detecção de curto-circuito entre espiras estatóricas em motores de indução trifásicos acionados por conversor de frequência**. Tese (Doutorado), 2014.
- PAO, L. Y.; JOHNSON, K. E. Control of Wind Turbines. **IEEE Control Systems**, v. 31, n. 2, p. 44–62, 2011.
- PENMAN, J. et al. Detection and location of interturn short circuits in the stator windings of operating motors. **IEEE Transactions on Energy Conversion**, v. 9, n. 4, p. 652—658, 1994.
- POLINDER, H. Overview of and trends in wind turbine generator systems. **2011 IEEE Power and Energy Society General Meeting**, p. 1–8, 2011. ISSN 1944-9925.
- POLINDER, H. et al. Trends in Wind Turbine Generator Systems. **IEEE Journal of Emerging and Selected Topics in Power Electronics**, v. 1, n. 3, p. 174–185, 2013. ISSN 2168-6777.
- PULEVA, T. Wind Turbine Power Control Based on Aero Dynamical Torque Estimation. 2015.

- RAMALHO, G. L. B. et al. Rotation-invariant feature extraction using a structural co-occurrence matrix. **Measurement**, v. 94, p. 406–415, 2016.
- RAMALHO, G. L. B. et al. Detecção De Falhas Em Motores Elétricos Através Da Classificação De Padrões De Vibração Utilizando Uma Rede Neural Elm. **Holos**, v. 4, n. 0, p. 185, 2014.
- RAMALHO, G. L. B. et al. Detecção de falhas através de características do sinal de vibração e rede SOFM. **XI Simpósio Brasileiro de Automação Inteligente, 2013, Fortaleza-CE. Simpósio Brasileiro de Automação Inteligente 2013 (SBAI 2013)**, 2013. ISSN 1098-6596.
- ROSENBLATT, F. The perceptron: a probabilistic model for information storage and organization in the brain. **Psychological review**, American Psychological Association, v. 65, n. 6, p. 386, 1958.
- ROYO, J.; ARCEGA, F. J. Machine current signature analysis as a way for fault detection in squirrel cage wind generators. **2007 IEEE International Symposium on Diagnostics for Electric Machines, Power Electronics and Drives, SDEMPED**, p. 383–387, 2007.
- SEMICONDUCTOR, F. **Three Axis Low-g Micromachined Accelerometer**. 2008. 1–11 p.
- SHARIFI, R.; EBRAHIMI, M. Detection of stator winding faults in induction motors using three-phase current monitoring. **ISA Transactions**, v. 50, n. 1, p. 14–20, 2011.
- SOUSA, I. R. Detecção prematura de falhas por abertura de barras rotóricas em motor de indução trifásico aplicando redes neurais artificiais. 2017.
- SOUSA, I. R. de; SÁ, C. M. d. M. Aplicação de redes neurais artificiais para detecção de falhas em barras rotóricas de motores de indução. In: **XIII Simpósio Brasileiro de Automação Inteligente**. [S.l.: s.n.], 2017.
- STOPA, M. M.; Cardoso Filho, B. J.; MARTINEZ, C. B. Incipient detection of cavitation phenomenon in centrifugal pumps. **IEEE Transactions on Industry Applications**, v. 50, n. 1, p. 120–126, 2014.
- SWINGLER, K. **Applying neural networks: a practical guide**. [S.l.]: Morgan Kaufmann, 1996.
- TAHANI, M. et al. Vibrational Simulation of Offshore Floating Wind Turbine and its Directional Movement Control by Fuzzy Logic. 2017.
- TAVNER, P. J.; BUSSEL, G. J. W. V.; SPINATO, F. Machine and Converter Reliabilities in Wind Turbines. **International Conference on Power Electronics, Machines and Drives**, n. April, p. 1–4, 2006.
- THOMSON, W. T. Motor Current Signature Analysis for Induction Motors. **Industrial Application and Case Histories**, p. 359–373, 2017.
- THOMSON, W. T.; FENGER, M. Current signature analysis to detect induction motor faults. **IEEE Industry Applications Magazine**, v. 7, n. 4, p. 26–34, 2001.

- TÍBA, C.; LEAL, S. S. Measuring and modelling illuminance in the semi-arid Northeast of Brazil. **Renewable Energy**, v. 48, p. 464–472, 2012.
- TOMA, S.; CAPOCCHI, L.; CAPOLINO, G.-A. Wound Rotor Induction Generator Inter-Turn Short-Circuits Diagnosis Using a New Digital Neural Network. **IEEE Transactions on Industrial Electronics**, v. 60, n. c, p. 1–1, 2013.
- TORABIZADEH, M.; NOSHADI, A. Artificial Neural Network-based fault diagnostics of an electric motor using vibration monitoring. **Proceedings 2011 International Conference on Transportation, Mechanical, and Electrical Engineering (TMEE)**, p. 1512–1516, 2011.
- UMANS, S. D. et al. **Máquinas Elétricas de Fitzgerald e Kingley**. [S.l.: s.n.], 2014. 724 p.
- V. Capdevielle, C. S.; LACOUME, J.-L. Blind separation of wide-band sources: application to rotating machine signals. **Proceedings of the Eighth European Signal Processing Conference**, v. 3, p. 2085–2088, 1996.
- VIEIRA, R. G. **Processo de Prototipação de um Classificador Neural Embarcado para Detecção de Falhas em Motores de Indução Trifásicos**. Tese (Doutorado), 2016.
- WENLEI, S. U. N. et al. The Development of Large-scale Wind Turbine Blade Redesign System. n. September, p. 12–13, 2014.
- XAVIER, G. A. et al. Simulation of distributed generation with photovoltaic microgrids-Case study in Brazil. **Energies**, v. 8, n. 5, p. 4003–4023, 2015.
- YANG, S.; LI, W.; WANG, C. The intelligent fault diagnosis of wind turbine gearbox based on artificial neural network. **2008 International Conference on Condition Monitoring and Diagnosis**, p. 1327–1330, 2008.
- YARAMASU, V. et al. High-Power Wind Energy Conversion Systems: State-of-the-Art and Emerging Technologies. **Proceedings of the IEEE**, v. 103, n. 5, 2015.
- YDHC. **Split core current transformer**. [S.l.], 2010.
- ZHANG, G. Neural networks for classification: a survey. **IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)**, v. 30, n. 4, p. 451–462, 2000.
- ZHANG, Z. et al. High-power generators for offshore wind turbines. **Energy Procedia**, v. 35, n. 1876, p. 52–61, 2013.
- ZOU, Y.; ELBULUK, M. E.; SOZER, Y. Simulation comparisons and implementation of induction generator wind power systems. **IEEE Transactions on Industry Applications**, v. 49, n. 3, p. 1119–1128, 2013. ISSN 00939994.

Apêndices

APÊNDICE A – Quadro de Testes de Curto-circuito em Máquinas (QTCM)



INSTITUTO
FEDERAL
Ceará

CONTRATANTE:

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA
E TECNOLOGIA DO CEARÁ – IFCE

PROJETO:

QUADRO DE TESTE DE CURTO DE
MOTORES

DESCRIÇÃO:

QTCM

CONTRATADA:

INFITECH AUTOMAÇÃO

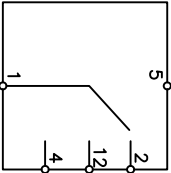
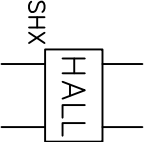
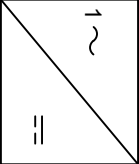
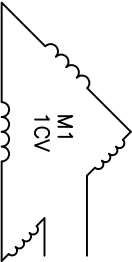
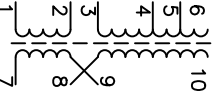
ENDEREÇO:

AV. MINISTRO JOSÉ AMÉRICO, 275
CAMBEBA – FORTALEZA – CEARÁ
TEL: (85)3276–4074
<http://www.infitech.com.br>



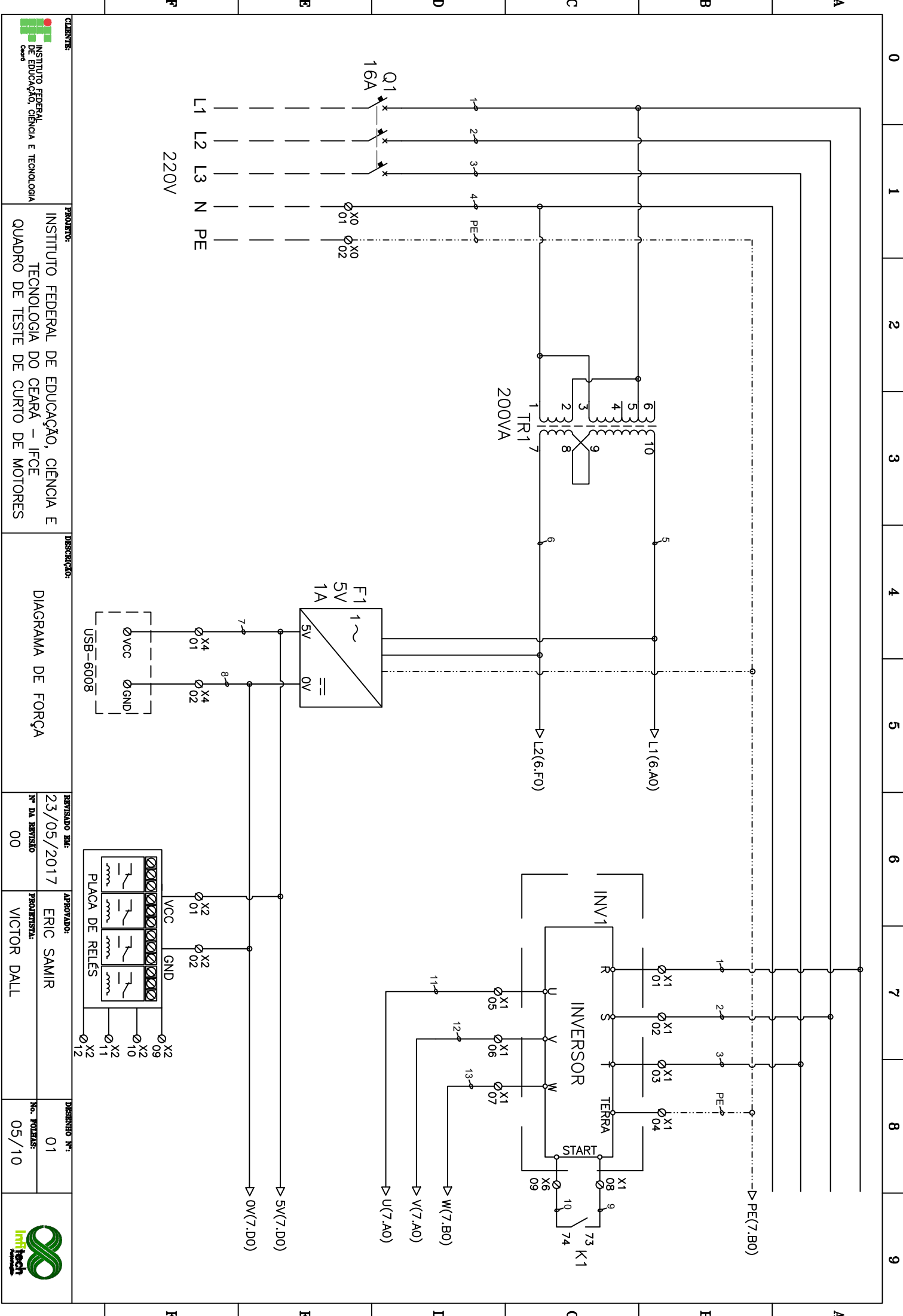
REVISÃO	DESCRIÇÃO
01	[26/02] Correção na régua de bornes e indicação de cor de condutores.
02	
03	
04	
05	
06	

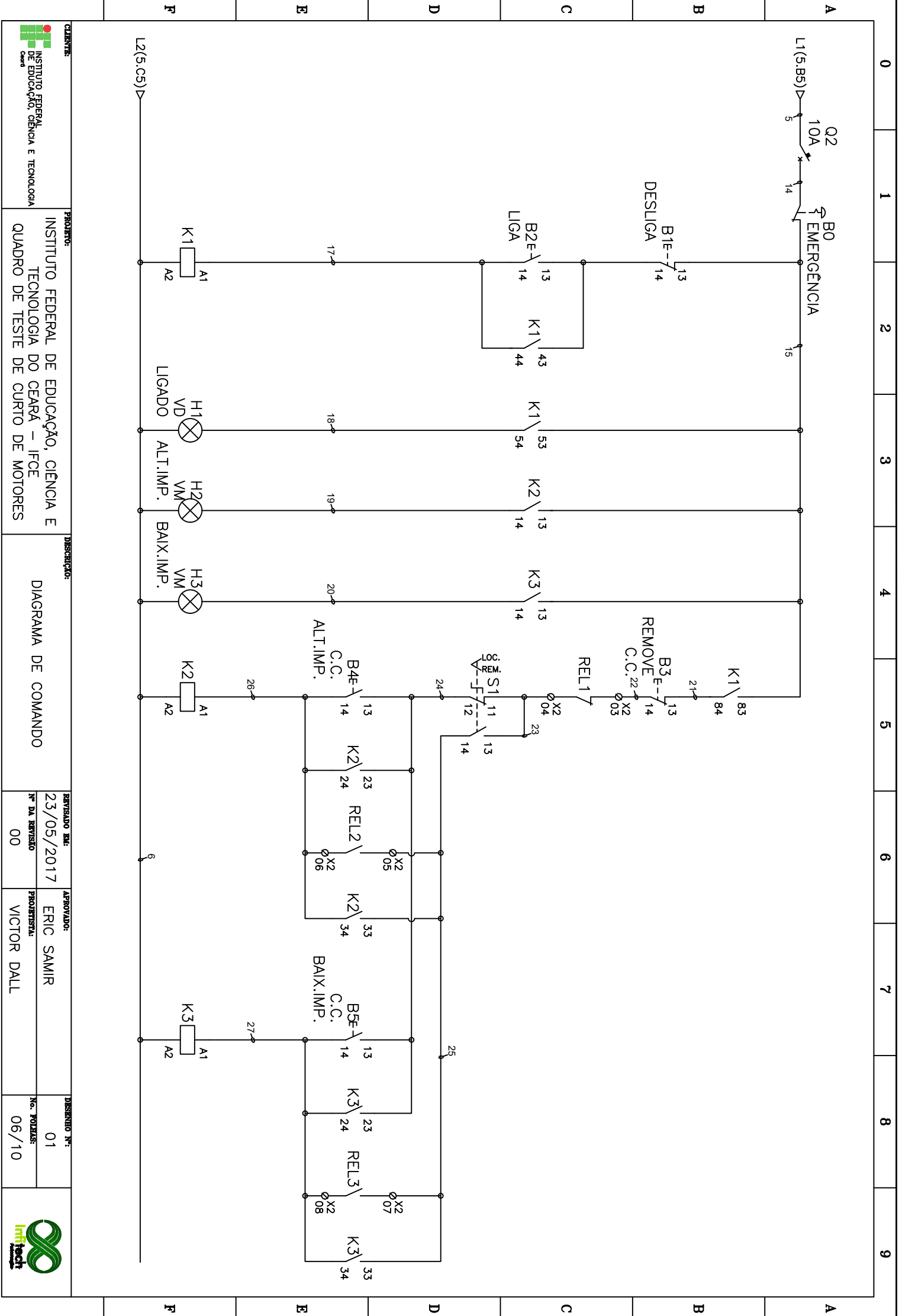
LEGENDA DE SÍMBOLOS

	Chave comutadora de amperímetro		Sensor de efeito Hall		Fonte CC 5A
	Motor trifásico com fechamento em triângulo		Transformador isolador		
	REOSTATO		Disjuntor tripolar		
	Contatos NA/NF		Disjuntor monopolar		
	Botoeira sem retenção		Borne		
	Amperímetro		Indicador de localização externa		
	Botoeira de emergência com retenção		Tomada 2P+T		

	0	1	2	3	4	5	6	7	8	9														
A	<u>CARACTERÍSTICAS TÉCNICAS</u> TENSÃO DE ALIMENTAÇÃO 220V 3n2,5(2,5)+Tn2,5 FREQUÊNCIA 60HZ <u>FIÇÃO (SOMENTE ONDE NÃO INDICADO)</u> <u>CIRCUITO CA</u> FASE : BRANCO – SEÇÃO 1,5mm2 NEUTRO : AZUL – SEÇÃO 1,5mm2 TERRA : VERDE – SEÇÃO 1,5mm2 COMANDO CA : BRANCO – SEÇÃO 1,5mm2 <u>CIRCUITO CC</u> SINAL CC: CINZA – SEÇÃO 1,0 mm2 ALIMENTAÇÃO CC + : VERMELHO – SEÇÃO 1,0 mm2 ALIMENTAÇÃO CC – : PRETO – SEÇÃO 1,0 mm2					LEGENDA DE SÍMBOLOS CARACTERÍSTICAS TÉCNICAS LEIAUTE DIAGRAMA DE FORÇA DIAGRAMA DE COMANDO DIAGRAMA DE ALIMENTAÇÃO DO MOTOR LISTA DE MATERIAIS IDENTIFICAÇÕES EXTERNAS					02	03	04	05	06	07	08							
B										09 e 10														
C																								
D																								
E																								
F																								
<table><tr><td>CLIENTE:</td><td>PROJETO:</td><td>DESCRIÇÃO:</td><td>REVISÃO EM:</td><td>APROVADO:</td><td>DESENHO Nº:</td><td></td></tr><tr><td> INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA Ceará</td><td>INSTITUTO FEDERAL DE EDUCAÇÃO, CIENCIA E TECNOLOGIA DO CEARÁ – IFCE QUADRO DE TESTE DE CURTO DE MOTORES</td><td>CARACTERÍSTICAS TÉCNICAS</td><td>23/05/2017 Nº DA REVISÃO 00</td><td>ERIC SAMIR PROJETISTA VICTOR DALL</td><td>01 Mo. FORNAR 03/10</td><td></td></tr></table>											CLIENTE:	PROJETO:	DESCRIÇÃO:	REVISÃO EM:	APROVADO:	DESENHO Nº:		 INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA Ceará	INSTITUTO FEDERAL DE EDUCAÇÃO, CIENCIA E TECNOLOGIA DO CEARÁ – IFCE QUADRO DE TESTE DE CURTO DE MOTORES	CARACTERÍSTICAS TÉCNICAS	23/05/2017 Nº DA REVISÃO 00	ERIC SAMIR PROJETISTA VICTOR DALL	01 Mo. FORNAR 03/10	
CLIENTE:	PROJETO:	DESCRIÇÃO:	REVISÃO EM:	APROVADO:	DESENHO Nº:																			
 INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA Ceará	INSTITUTO FEDERAL DE EDUCAÇÃO, CIENCIA E TECNOLOGIA DO CEARÁ – IFCE QUADRO DE TESTE DE CURTO DE MOTORES	CARACTERÍSTICAS TÉCNICAS	23/05/2017 Nº DA REVISÃO 00	ERIC SAMIR PROJETISTA VICTOR DALL	01 Mo. FORNAR 03/10																			

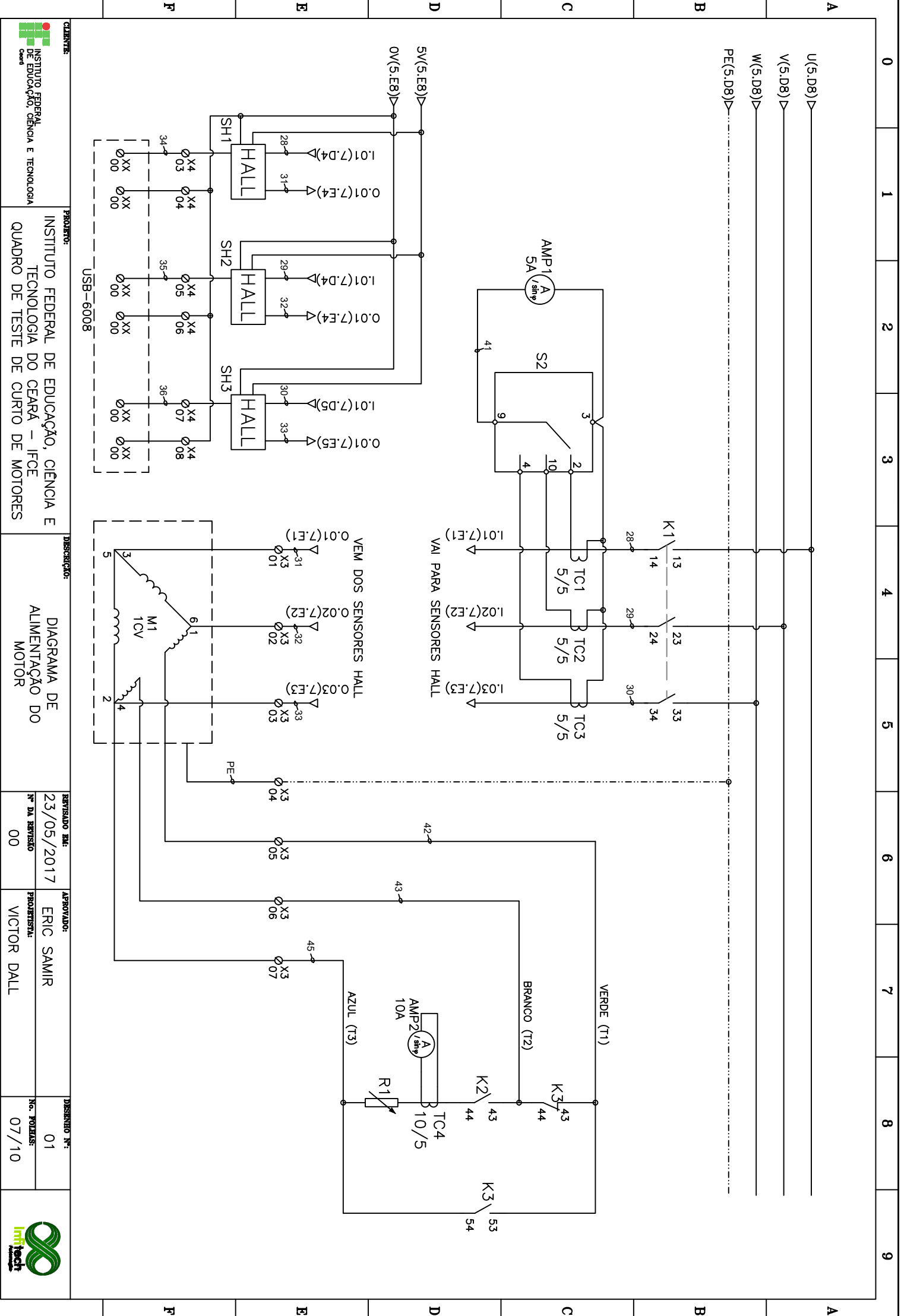
	A	0	1	2	3	4	5	6	7	8	9
	B										
	C										
	D										
	E										
	F										
CLIENTE:	INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA do Ceará - IFCE										
PROJETO:	INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO CEARÁ – IFCE QUADRO DE TESTE DE CURTO DE MOTORES										
DESCRIÇÃO:	LAYOUT DO QUADRO										
REVISÃO EM:	23/05/2017										
Nº DA REVISÃO	00										
APROVADO:	ERIC SAMIR										
PROJETISTA:	VICTOR DALL										
DESSENHO Nº:	01										
Nº. FOLHAS:	04/10										
	ImTech										





CLIENTE:	PROJETO:	DESCRIÇÃO:	REVISÃO EM:	APROVADO:	DESENHO N.º:
INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA Campus Ceará	INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO CEARÁ – IFCE QUADRO DE TESTE DE CURTO DE MOTORES	DIAGRAMA DE COMANDO	23/05/2017 N.º DA REVISÃO 00	ERIC SAMIR PROJETISTA: VICTOR DALL	01 N.º. FOLHAS: 06/10





ETIQUETAS

A	B	C	D	E	F	LIGA 1. MEDIDAS: 45mm X 15mm X 3mm 2. PLAQUETAS EM ACRILICO PRETO BRILHOSO 3. LETRAS BRANCAS EM FONTE ARIAL BLACK	DESLIGA 1. MEDIDAS: 45mm X 15mm X 3mm 2. PLAQUETAS EM ACRILICO PRETO BRILHOSO 3. LETRAS BRANCAS EM FONTE ARIAL BLACK
						LIGADO 1. MEDIDAS: 45mm X 15mm X 3mm 2. PLAQUETAS EM ACRILICO PRETO BRILHOSO 3. LETRAS BRANCAS EM FONTE ARIAL BLACK	EMERGÊNCIA 1. MEDIDAS: 45mm X 15mm X 3mm 2. PLAQUETAS EM ACRILICO PRETO BRILHOSO 3. LETRAS BRANCAS EM FONTE ARIAL BLACK
						REMOVE CURTO-CIRCUITO 1. MEDIDAS: 45mm X 15mm X 3mm 2. PLAQUETAS EM ACRILICO PRETO BRILHOSO 3. LETRAS BRANCAS EM FONTE ARIAL BLACK	LOCAL / REMOTO 1. MEDIDAS: 45mm X 15mm X 3mm 2. PLAQUETAS EM ACRILICO PRETO BRILHOSO 3. LETRAS BRANCAS EM FONTE ARIAL BLACK
						CURTO-CIRCUITO ALTA IMPEDÂNCIA 1. MEDIDAS: 45mm X 15mm X 3mm 2. PLAQUETAS EM ACRILICO PRETO BRILHOSO 3. LETRAS BRANCAS EM FONTE ARIAL BLACK	CURTO-CIRCUITO BAIXA IMPEDÂNCIA 1. MEDIDAS: 45mm X 15mm X 3mm 2. PLAQUETAS EM ACRILICO PRETO BRILHOSO 3. LETRAS BRANCAS EM FONTE ARIAL BLACK
A	B	C	D	E	F	ALTA IMPEDÂNCIA 1. MEDIDAS: 45mm X 15mm X 3mm 2. PLAQUETAS EM ACRILICO PRETO BRILHOSO 3. LETRAS BRANCAS EM FONTE ARIAL BLACK	BAIXA IMPEDÂNCIA 1. MEDIDAS: 45mm X 15mm X 3mm 2. PLAQUETAS EM ACRILICO PRETO BRILHOSO 3. LETRAS BRANCAS EM FONTE ARIAL BLACK

0	1	2	3	4	5	6	7	8	9		
ETIQUETAS											
INSTITUTO FEDERAL QTCM Infotech QUADRO DE TESTES DE CURTO DE MOTORES			1. MEDIDAS: 57,6mm X 28,8mm X 3mm 2. PLAQUETAS EM ACRÍLICO PRETO BRILHOSO 3. LETRAS BRANCAS EM FONTE ARIAL BLACK			Infotech Automação			1. MEDIDAS: 50mm X 50mm 2. ETIQUETA EM PAPEL VERDE BRILHOSO 3. LETRAS BRANCAS EM FONTE ARIAL BLACK		

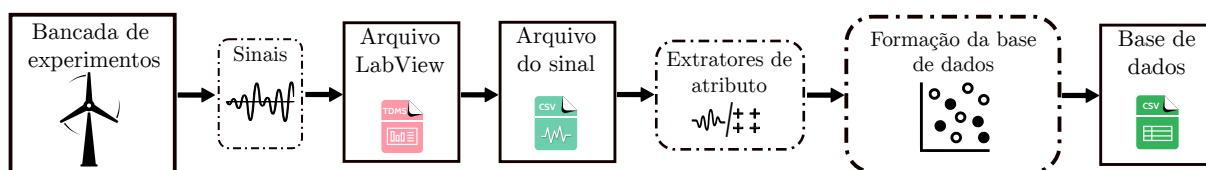
INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA Ceará	PROJETO: INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO CEARÁ – IFCE QUADRO DE TESTE DE CURTO DE MOTORES	DESCRIÇÃO: IDENTIFICAÇÕES EXTERNAS	REVISÃO EM: 23/05/2017 Nº DA REVISÃO 00	APROVADO: ERIC SAMIR PROJETISTA: VICTOR DALL	DESENHO Nº: 01 Nº. FOLHAS: 10/10	
--	--	---	--	---	---	--------------

APÊNDICE B – Detalhes sobre a formação da base de dados

Uma visão geral sobre as etapas inerentes a constituição da base de dados é exibida na Figura 50. Os sinais de funcionamento exibidos na Seção 3.3 são adquiridos na bancada de experimentos pelo *software* desenvolvido no LabVIEW. As aquisição são feitas em diferentes pontos de funcionamento do gerador, de acordo com o exibido na Tabela 2.

Ao todo, sete sinais, de 50 mil pontos, são salvos em cada aquisição. Um arquivo de extensão *.tdms* é gerado pelo LabVIEW. Esses sinais são as correntes elétricas das três fases do gerador, dos três eixos de vibrações e do fluxo magnético axial. Além disso, são salvos registros adicionais feitos nos ensaios, que são referentes a tensão no barramento CC (V_{cc}), a frequência no gerador (f_g), a frequência de base (f_b) e a potência gerada (P).

Figura 50 – Etapas para a constituição da base de dados.



Fonte: Autor.

Um *script* feito no MATLAB converte os arquivos *.tdms*, junto com os registros adicionais, em arquivos de formato *.csv*, separados pelos seguintes tipos de sinais: corrente elétrica, vibração e fluxo magnético axial. Um outro *script* é responsável por gerar os atributos dos extratores propostos na Seção 3.5. A formação da base de dados é feita ao concatenar os atributos de todas as aquisições e rotular essas amostras, de acordo com o tipo de operação do gerador (*i.e.* classes).

Nesse trabalho, foram utilizados apenas os sinais de corrente elétrica para compor a base de testes para análise de falhas. A base utilizada para gerar os resultados está disponível para *download* e pode ser encontrada no [repositório online](https://github.com/navarmn/Wind_turbine_failure_prediction)¹.

¹ Endereço para *download*: <https://github.com/navarmn/Wind_turbine_failure_prediction>. Caso encontre problemas com o *link* disponibilizado, contate navarmedeiros@gmail.com