

Karine Midori Nacano

Modelagem e Identificação de Anomalias em Buchas Capacitivas

Curitiba - PR, Brasil

2021

Karine Midori Nacano

Modelagem e Identificação de Anomalias em Buchas Capacitivas

Dissertação apresentada ao Programa de Pós-Graduação em Informática da Pontifícia Universidade Católica do Paraná como requisito parcial para obtenção do título de mestre em Informática.

Pontifícia Universidade Católica do Paraná - PUCPR

Programa de Pós-Graduação em Informática - PPGIa

Orientador: Marcelo Eduardo Pellenz

Coorientador: Edgard Jamhour

Curitiba - PR, Brasil

2021

Dados da Catalogação na Publicação
Pontifícia Universidade Católica do Paraná
Sistema Integrado de Bibliotecas – SIBI/PUCPR
Biblioteca Central
Margareth Jansson Zanetti – CRB 9/1117

N118m Nacano, Karine Midori
2021 Modelagem e identificação de anomalias em buchas capacitivas / Karine Midori Nacano ; orientador: Marcelo Eduardo Pellenz ; coorientador Edgard Jamhour. -- 2021.
81 f. : il. ; 30 cm

Dissertação (mestrado) – Pontifícia Universidade Católica do Paraná,
Curitiba, 2021
Bibliografia: f. 78-81

1. Máquinas - Monitoração. 2. Modelagem da informação.
3. Aprendizado do computador. 4. Análise de sistemas. 5. Monitoramento eletrônico no ambiente de trabalho 6. Buchas. I. Pellenz, Marcelo Eduardo. II. Jamhour, Edgard. III. Pontifícia Universidade Católica do Paraná. Programa de Pós-Graduação em Informática. IV. Título.

CDD 20. ed. – 004



Pontifícia Universidade Católica do Paraná
Escola Politécnica
Programa de Pós-Graduação em Informática

Curitiba, 19 de janeiro de 2022.

02-2023

DECLARAÇÃO

Declaro para os devidos fins, que **KARINE MIDORI NACANO** defendeu sua dissertação de Mestrado intitulada “**Modelagem e Identificação de Anomalias em Buchas Capacitivas**”, na área de concentração Ciência da Computação no dia 28 de outubro de 2021, no qual foi aprovada.

Declaro ainda, que foram feitas todas as alterações solicitadas pela Banca Examinadora, cumprindo todas as normas de formatação definidas pelo Programa.

Por ser verdade firmo a presente declaração.

Prof. Dr. Emerson Cabrera Paraiso
Coordenador do Programa de Pós-Graduação em Informática

Agradecimentos

Primeiramente agradeço a Deus pela vida, força e inspiração para a realização de mais essa jornada na minha vida acadêmica. Estando ao meu lado nas horas boas e nos momentos difíceis me confortando.

A PUCPR, P&D ANEEL e CISEI, por acreditarem em meu projeto de mestrado, concedendo a mim uma bolsa para melhor me dedicar aos estudos do tema, por me auxiliar nas pesquisas e recursos quando fora necessário. Ao orientador, Prof. Marcelo Eduardo Pellenz, e ao Co-orientador, Prof. Edgard Jamhour, por todo o tempo e dedicação investidos, por suas contribuições, sugestões e ideias que tornaram esse trabalho melhor. Sou eternamente grata a toda essa atenção e paciência nos últimos anos.

Agradeço a Copel pela possibilidade de obter os dados relevantes, pelo financiamento da pesquisa, assim como pela visitação às subestações de Bateias, que foi necessária para realização deste projeto. Agradeço também os engenheiros Marcos Vinicio Haas Rambo e Daniel Benetti que compartilharam os seus conhecimentos, tempo e auxiliaram com os dados para que este trabalho fosse realizado.

Não podendo esquecer de agradecer minha família e amigos pela colaboração, compreensão e apoio, por todo o carinho e amor que me mantiveram focados nos estudos. Aos meus pais, por me educarem e me tornarem na pessoa que sou hoje. Aos amigos que estiveram do meu lado sempre, dos quais não irei mencionar para não cometer a injustiça de esquecer de mencionar alguém. Gostaria de apenas adicionar que todos têm um lugar especial em meu coração.

Resumo

Este trabalho é parte dos resultados obtidos durante a execução do projeto de P&D de código PD-06491-0373/2012, executado pela PUCPR para Copel GeT no âmbito do Programa de Pesquisa e Desenvolvimento Tecnológico do Setor Elétrico, regulamentado pela Agência Nacional de Energia Elétrica (ANEEL). No atual contexto das redes elétricas inteligentes (Smart Grids), os sistemas de monitoramento em tempo real das condições de operação dos sistemas de geração, transmissão e distribuição de energia elétrica, se tornam uma ferramenta essencial para a operação eficiente da rede. Neste cenário, um dos principais equipamentos das subestações é o transformador de potência. A operação adequada desse equipamento é de grande importância para o fornecimento de energia. Um fator essencial para o funcionamento adequado do transformador é o sistema de isolamento dos terminais. No transformador, o componente responsável pela isolação é a bucha. Como o tempo de vida da bucha pode alterar de acordo com as demandas do sistema, forma de fabricação e os estresses recebidos, cada vez mais são utilizados sistemas para o monitoramento *on-line* destes componentes. Um sistema eficiente para identificação de possíveis problemas na bucha permite reduzir o número de desligamentos do transformador na subestação, diminuir os acidentes causados por explosões do transformador e amenizar os custos de manutenção. Este trabalho propõe estratégias para detectar o estado da bucha, sendo a ideia principal, utilizar os sinais coletados nas subestações e aplicar algoritmos de aprendizagem de máquina para a identificação de possíveis anomalias, causadas pelo envelhecimento. Outro ponto analisado, é a correlação da capacitância e do fator de potência com a temperatura do óleo para gerar um modelo elétrico da bucha mais próximo do comportamento real. Sendo assim, este trabalho analisa as formas de avaliar uma bucha mais usadas pelas concessionárias, verificando se é a melhor escolha a ser utilizada. Além de usar os resultados para gerar um modelo elétrico que seja mais completo.

Palavras-chave: Monitoramento *On-line*, Transformador, Bucha Capacitiva, Modelagem, Monitoramento, Aprendizagem de Máquina.

Lista de Ilustrações

Figura 1 – Arquitetura Geral do Sistema Elétrico de Potência (SEP).	12
Figura 2 – Transformador.	13
Figura 3 – Incêndio em subestação de energia na Zona Norte de Macapá (03/11/2020).	15
Figura 4 – Exemplo de bucha não-capacitiva.	20
Figura 5 – Distribuição do campo elétrico nas buchas.	21
Figura 6 – Exemplo de construção de uma bucha capacitiva.	22
Figura 7 – (a) Vista lateral. (b) Representação elétrica. (c) Vista superior.	23
Figura 8 – (a) Detalhes construtivos de uma bucha capacitiva de alta tensão. (b) Circuito equivalente das perdas dielétricas. (c) Diagrama vetorial das correntes.	24
Figura 9 – Circuito equivalente do isolamento da bucha de alta tensão e do sensor.	25
Figura 10 – Tipos de aprendizagem de máquina.	28
Figura 11 – Estrutura de uma Árvore de Decisão.	30
Figura 12 – Aplicação de uma KNN.	30
Figura 13 – Modelo de monitoramento de capacitância.	35
Figura 14 – Modelo de monitoramento pelas correntes balanceadas.	36
Figura 15 – Modelo com sinal de referência.	37
Figura 16 – Diagrama de blocos do método de comparação fase a fase.	38
Figura 17 – Etapas da Proposta.	43
Figura 18 – Características de Caxias.	44
Figura 19 – Gráfico de conversão da temperatura ambiente em temperatura do óleo.	50
Figura 20 – Gráfico de correção do efeito da temperatura ambiente na capacitância.	51
Figura 21 – Diagrama do modelo elétrico.	52
Figura 22 – Características de Bateias.	53
Figura 23 – Dados do transformador da Unidade 1 de Caxias.	56
Figura 24 – Dados do transformador da Unidade 2 de Caxias.	57
Figura 25 – Dados do transformador da Unidade 3 de Caxias.	58
Figura 26 – Dados do transformador da Unidade 4 de Caxias.	59
Figura 27 – Correlação dos dados para buchas em bom estado.	66
Figura 28 – Correlação dos dados para buchas em estado de alerta.	67
Figura 29 – Correlação dos dados para buchas que devem ser trocadas.	67
Figura 30 – Comparação da capacitância e fator de potência com a temperatura do óleo	69
Figura 31 – Dados simulados.	70
Figura 32 – Comparação dos dados simulados com dados reais de buchas antes de trocar e depois de trocar.	71

Figura 33 – Comparação dos dados simulados com dados reais de uma bucha em estado normal de operação.	72
Figura 34 – Características obtidas em Bateias.	73
Figura 35 – Dados do transformador 1 de Bateias para a tensão de 138kV.	74
Figura 36 – Dados do transformador 1 de Bateias para a tensão de 230kV.	75

Lista de Tabelas

Tabela 1 – Diagnóstico das buchas.	37
Tabela 2 – Tabela de comparação dos métodos de monitoramento.	39
Tabela 3 – Tabela de comparação dos métodos de modelagem.	41
Tabela 4 – Tabela de falhas em buchas.	41
Tabela 5 – Tabela de falhas em buchas.	42
Tabela 6 – Tabela de limiares de decisão usando o fator de dissipação.	46
Tabela 7 – Tabela de fator de correção (K) de temperatura do óleo.	49
Tabela 8 – Tabela de limiares de decisão usando o fator de potência e capacitância.	53
Tabela 9 – Características selecionadas.	60
Tabela 10 – Instâncias de Caxias.	61
Tabela 11 – Matrizes de confusão da árvore de decisão.	62
Tabela 12 – Matrizes de confusão do KNN.	63
Tabela 13 – Matrizes de confusão do <i>Naive Bayes</i>	63
Tabela 14 – Matrizes de confusão do <i>Random Forest</i>	63
Tabela 15 – Matrizes de confusão do <i>Bagging</i>	64
Tabela 16 – Matrizes de confusão do <i>AdaBoost</i>	64
Tabela 17 – Matrizes de confusão do <i>Random Subspaces</i>	65
Tabela 18 – Tabela de comparação dos classificadores	65

Lista de Abreviaturas e Siglas

SEP	Sistema Elétrico de Potência
RBP	<i>Resin Bonded Paper</i>
OIP	<i>Oil Impregnated Paper</i>
RIP	<i>Resin Impregnated Paper</i>
PF	<i>Power Factor</i>
DF	<i>Dissipation Factor</i>
ML	<i>Machine Learning</i>
IA	Inteligência Artificial
FP	Falso Positivo
VP	Verdadeiro Positivo
FN	Falso Negativo
VN	Verdadeiro Negativo
WEKA	<i>Waikato Environment for Knowledge Analysis</i>
KNN	<i>K Nearest Neighbor</i>
SMC	Sistemas de Múltiplos Classificadores
ESA	<i>Electrical Signature Analysis</i>
CNN	<i>Convolutional Neural Network</i>
AdaBoost	<i>Adaptive Boosting</i>
SMO	<i>Sequential Minimal Optimization</i>

Lista de Símbolos

C_1	Capacitância Principal
C_2	Tape Capacitivo
$\tan(\delta)$	Fator de Dissipação
I_R	Corrente Resistiva
I_C	Corrente Capacitiva
$\cos(\phi)$	Fator de Potência
I_B	Corrente Total da Bucha
I'_B	Corrente do Sensor
R_1	Resistência Paralela a Capacitância Principal
R_2	Resistência Paralela ao Tape Capacitivo
R_S	Resistência do Sensor
V_B	Resistência do Sensor
I_C	Corrente Capacitiva
I_R	Corrente Resistiva
f	Frequência da Rede
w	Frequência Angular da Rede
x_{C1}	Reatância Capacitiva Relativo à Capacitância Principal
x_{C2}	Reatância Capacitiva Relativo ao Tape Capacitivo
z_2	Impedância Relacionada ao Tape Capacitivo
δ	Ângulo do Fator de Dissipação
ϕ	Ângulo do Fator de Potência
k	Quantidade de Amostras
δ_{20}	Ângulo do Fator de Dissipação na Temperatura 20°

δ_m	Ângulo do Fator de Dissipação na Temperatura Medida
K	Fator de Correção da Temperatura do Óleo
K_C	Fator de Correção da Temperatura Ambiente para a Capacitância
C_{var}	Variação da Capacitância
$C_{1,Placa}$	Capacitância Principal da Placa
FP_{var}	Variação do Fator de Potência
FP_{Placa}	Fator de Potência da Placa
T_M	Temperatura do Óleo
T_a	Temperatura Ambiente
P	<i>precision</i>
R	<i>recall</i>

Sumário

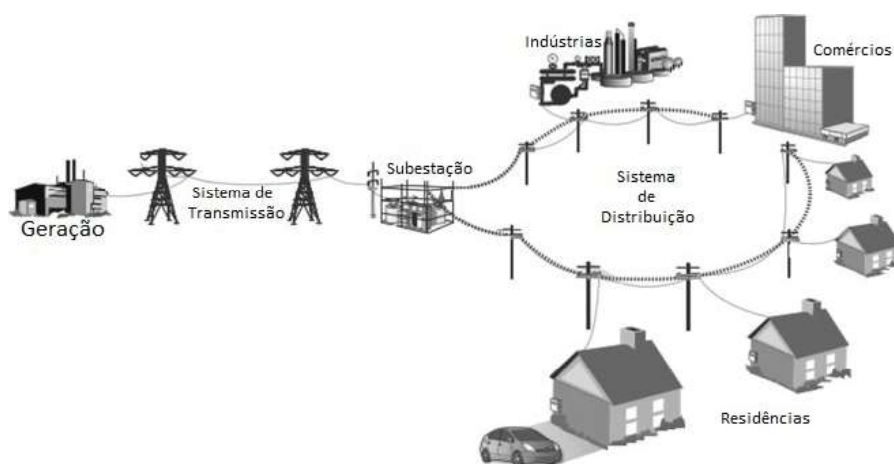
1	Introdução	12
1.1	Motivação	14
1.2	Objetivos	16
1.3	Hipóteses de Pesquisa	16
1.4	Contribuições	17
1.5	Organização do Trabalho	17
2	Fundamentação Teórica	19
2.1	Buchas	19
2.1.1	Buchas Não-Capacitivas	19
2.1.2	Buchas Capacitivas	20
2.1.3	Modelo Elétrico da Bucha Capacitiva	21
2.1.4	Capacitância, Fator de Dissipação e Fator de Potência	22
2.2	Aprendizagem de Máquina	26
2.2.1	Árvore de Decisão	29
2.2.2	KNN	29
2.2.3	<i>Naive Bayes</i>	31
2.2.4	Combinação de Classificadores	31
2.3	Correlação de Pearson	32
3	Trabalhos Relacionados	34
3.1	Monitoramento de Buchas	34
3.2	Modelagem do Sistema Trifásico	39
3.3	Tipos de Falhas em Buchas Capacitivas	40
4	Proposta	43
4.1	Coleta de Dados	43
4.2	Pré-Processamento	44
4.2.1	Remoção das Instâncias Incompletas	45
4.2.2	Seleção de Características	45
4.2.3	Rotulação dos Dados	45
4.3	Identificação e Detecção de Anomalias	46
4.3.1	Algoritmos de Aprendizagem de Máquina	46
4.3.2	Avaliação dos Algoritmos	47
4.4	Estudo dos Efeitos da Temperatura	47
4.5	Dados Adicionais	53
5	Resultados	55
5.1	Coleta de Dados	55
5.1.1	Base de Dados de Caxias	55

5.2	Etapa de Pré-Processamento	59
5.2.1	Remoção das Instâncias Incompletas	60
5.2.2	Seleção de Características	60
5.2.3	Rotulação de Dados	61
5.3	Identificação de Anomalias	61
5.4	Estudo dos Efeitos da Temperatura	65
5.4.1	Correlação entre as Características	66
5.4.2	Modelagem do Efeito da Temperatura	68
5.5	Dados Adicionais	71
6	Conclusão	76
6.1	Trabalhos Futuros	77
	Referências	78

1 Introdução

Como parte dos resultados obtidos durante a execução do projeto de P&D de código PD-06491-0373/2012, executado pela PUCPR para Copel GeT no âmbito do Programa de Pesquisa e Desenvolvimento Tecnológico do Setor Elétrico, regulamentado pela Agência Nacional de Energia Elétrica (ANEEL), este trabalho realiza um estudo da influência da variação da temperatura do óleo na avaliação do estado de uma bucha. Sendo assim, sabendo que a energia elétrica é uma das principais formas de uso da energia e está cada vez mais presente na vida da população. Com os avanços da tecnologia, há também o aumento da demanda, além de consumidores mais exigentes. A suspensão do fornecimento da energia elétrica pode provocar perdas financeiras às concessionárias de energia, transtornos à população e gerar insatisfações aos consumidores. Além disso, ao cessar o fornecimento de energia, outros serviços também são afetados. Nesse contexto, é preciso ter em mente que o Sistema Elétrico de Potência (SEP) consiste em todas as instalações e equipamentos, desde os geradores, transformadores elevadores e redutores, linhas de transmissão e os alimentadores de distribuição. Portanto, todo o conjunto deve ser o mais eficiente possível para prover um serviço de qualidade e com menos interrupções. Nesse contexto, analisando a arquitetura do sistema elétrico é possível encontrar buchas capacitivas que são um dos componentes mais críticos e uma das principais fontes de falhas em grandes transformadores.

Figura 1 – Arquitetura Geral do Sistema Elétrico de Potência (SEP).

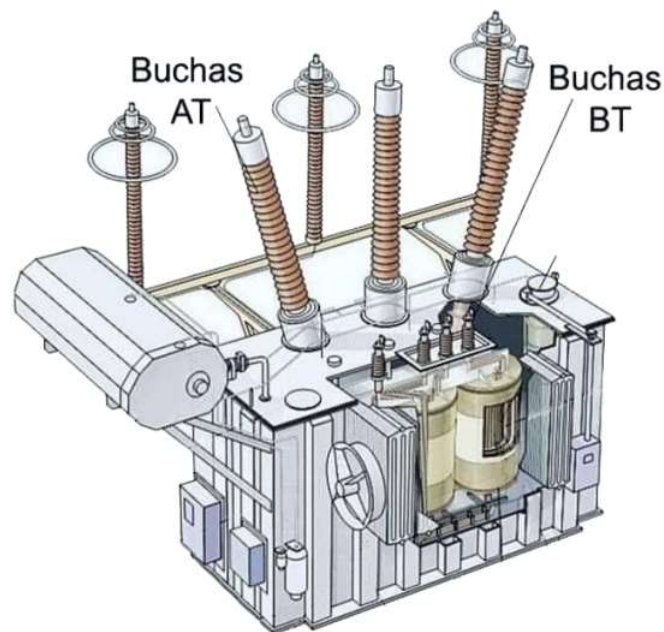


Fonte: Adaptado de KHOUSSEI; MATTAS, 2017

A Figura 1 mostra um exemplo de SEP e a aplicabilidade dos transformadores nesse contexto. Nota-se a importância de tais equipamentos para um sistema elétrico interligado,

por isso é essencial que haja um alto nível de confiabilidade na sua operação para o bom funcionamento de uma malha energética. Como os transformadores são componentes por onde passa toda a energia proveniente das unidades geradoras até os consumidores finais, são necessários dispositivos que interliguem os cabos e os transformadores de forma a permitir a passagem isolada de um condutor energizado. Estes dispositivos são as buchas de conexão, conforme ilustrado na Figura 2. Dentre os elementos que compõem a estrutura de um transformador, a bucha é um dos componentes que se destacam. Segundo MEHTA et al., os problemas de isolamento podem ser causa de aproximadamente 26% dos danos em transformadores e segundo RAMBO et al. as buchas podem ser responsáveis por aproximadamente 30% dos problemas em transformadores de potência. Apesar do seu custo ser relativamente baixo em comparação ao transformador, a manutenção de uma bucha leva ao desligamento do fornecimento de energia e uma falha nesse componente pode gerar danos a toda estrutura do transformador. Além disso, com o desbalanceamento natural de cargas que ocorre no sistema trifásico, torna-se ainda mais complicado a previsão de degradação do isolamento da bucha.

Figura 2 – Transformador.



Fonte: Adaptado de FRONTIN et al., 2013

O monitoramento de uma bucha pode ser realizado de duas maneiras *off-line* e *on-line*. O método de medição *off-line* é mais preciso e seu procedimento é mais eficiente devido a sua capacidade de detecção precoce de contaminações (RAMBO et al., 2007) (MEIRA et al., 2018). Entretanto, este método se tornou pouco viável devido a necessidade de se interromper o fornecimento da energia periodicamente para realizar ensaios nas buchas

e verificar a qualidade do equipamento, além de trazer mais risco aos colaboradores que operam diretamente nos equipamentos. Já os métodos de medição *on-line* proporcionam um acompanhamento diário e auxiliam no planejamento dos desligamentos para a manutenção, trazendo mais segurança. Dessa forma, é preciso que o monitoramento *on-line* tenha uma precisão similar à do monitoramento *off-line*.

Atualmente, com as alterações constantes no setor elétrico, como implantações de novas tecnologias e alterações nas plantas elétricas, o monitoramento de cada equipamento está cada vez mais presente. Além disso, há também uma preocupação em aperfeiçoar os métodos para a segurança dos profissionais e para a satisfação dos consumidores. Dessa forma, a confiabilidade na detecção de degradações e de anomalias deve ser pouco suscetível a ruídos ou outros distúrbios que afetem o seu comportamento. Neste trabalho focamos nas medidas de *capacitância* e *tangente delta* das buchas capacitivas para verificar a sua condição de operação.

1.1 Motivação

As faltas de energia elétrica são geralmente associadas apenas às linhas de transmissão. Entretanto, sem o funcionamento adequado das subestações e de seus equipamentos também não é possível entregar a energia com qualidade. No caso das buchas de alta tensão, a maioria dos danos requer a troca do equipamento para a segurança do seu funcionamento, caso contrário podem ocorrer destruições catastróficas. A manutenção desse componente causa perturbações na rede e perdas tanto energéticas quanto monetárias. Portanto, quanto menos desligamentos ocorrerem, menos distúrbios e menos transtornos serão gerados aos consumidores. Dessa forma, o monitoramento e a avaliação do estado da isolação de uma bucha tornam-se de extrema importância para amenizar estes transtornos.

As buchas capacitivas nas subestações são uma das principais fontes de falhas e podem ser responsáveis por aproximadamente 26% dos problemas em transformadores de potência (MEIRA et al., 2018). Além disso, qualquer falha na isolação de uma bucha pode acarretar a interrupção do serviço devido às explosões, que podem expor funcionários ao risco, danificar outros equipamentos da subestação e causar a interrupção do fornecimento de energia (RAMBO et al., 2007). Os autores MEIRA et al. comentam que o custo desse equipamento equivale a aproximadamente 0,05% do valor do transformador, portanto mesmo tendo um custo relativamente baixo, caso ocorra algum defeito na bucha, as perdas são enormes. Além disso, todo o processo de troca de uma bucha capacitiva (compra, transporte e instalação) pode ser demorado, visto que as fabricantes são em geral de fora do Brasil. Nesse contexto, o diagnóstico precoce pode evitar a expansão dessas falhas para outros equipamentos, reduzindo gastos, evitando transtornos à população e melhorando a qualidade dos serviços prestados pelas concessionárias de energia.

Figura 3 – Incêndio em subestação de energia na Zona Norte de Macapá (03/11/2020).



Fonte: G1-AP, 2020

Um exemplo recente da importância de uma bucha foi o apagão que ocorreu no Amapá no dia 3 de novembro de 2020. Segundo o Ministério de Minas e Energia, uma explosão seguida de um incêndio atingiu um transformador da subestação. Inicialmente, acreditou-se que o incidente foi causado por um raio. Entretanto, as investigações comprovaram que o início do problema ocorreu em uma das buchas do transformador, causando um superaquecimento e, conseqüentemente, gerando a explosão que levou ao incêndio, inutilizando um transformador e danificando outro transformador (BRANDÃO; LEAL, 2020), mostrado na Figura 3. Nesse contexto, nota-se que um equipamento relativamente barato pode gerar grandes perdas, pois o desligamento provocou um apagão em 13 das 16 cidades do estado, ou seja, cerca de 700 mil pessoas foram afetadas nesse incidente (BRANDÃO; LEAL, 2020). Os prejuízos socioeconômicos foram enormes, visto que a falta de energia elétrica afeta serviços básicos como o abastecimento de água da cidade, comprometendo a saúde e o cotidiano da população.

Em (MORALES, 2012) os autores mencionam que a expectativa de vida de um transformador é de aproximadamente 50 anos, enquanto que a vida útil de uma bucha está em torno de 25 anos. Entretanto, a bucha apresenta dois pontos críticos durante a sua operação. A primeira, entre 10 a 13 anos, devido a problemas na qualidade da produção, e a segunda entre 20 e 30 anos, que é o tempo de vida esperado. Por esse motivo o diagnóstico confiável do estado de uma bucha capacitiva deve ser realizado de forma rápida e com o menor custo possível. Portanto, dentre as opções de monitoramento das buchas (*off-line* ou *on-line*) é preciso ter a precisão de uma medição *off-line* com a rapidez da medição *on-line*. Além disso, segundo um estudo de avaliação realizado em 2012 (PAULINO; ALMEIDA, 2012), cerca de 1/3 dos transformadores apresentavam um tempo de operação acima de 30

anos. Diante desse cenário de tempo de operação dos transformadores, é difícil se obter dados de buchas em estados mais avançados de operação, visto que a maioria já teria sido trocada e os sensores seriam colocados na troca desses equipamentos, caso a concessionária responsável desejasse que esse monitoramento fosse realizado.

Neste contexto, para este trabalho foram obtidos duas bases de dados, a de SE-Caxias e a de SE-Bateias. Os dados obtidos em Bateias foram apenas dados preliminares e com condição de operação normal. Já a base de dados de Caxias contém um histórico com períodos de operação normal e com anomalia.

1.2 Objetivos

O objetivo geral desta pesquisa é propor uma estratégia para melhorar a identificação de anomalias em buchas capacitivas, além de avaliar a modelagem das buchas em função de outros fatores que possam alterar o seu comportamento, como por exemplo efeitos de temperatura. Neste trabalho utilizamos dados reais coletados por sistemas de monitoramento *online* instalados em transformadores em subestações. Para se alcançar o objetivo geral deste trabalho, foram definidos os seguintes objetivos específicos:

- Estudo dos métodos utilizados para detecção de falhas em buchas capacitivas.
- Aquisição de dados de buchas capacitivas em operação e sob diferentes condições de operação.
- Análise dos dados e seleção de características para identificação de anomalias.
- Estruturação e rotulação da base de dados.
- Avaliação de diferentes estratégias de aprendizagem de máquina para detectar anomalias a partir das características selecionadas.
- Estudo e modelagem dos efeitos de temperatura nas buchas capacitivas.

1.3 Hipóteses de Pesquisa

As hipóteses deste projeto são:

- $H_{1.0}$ O estado de operação da bucha capacitiva pode ser avaliado melhor com as técnicas de aprendizagem de máquina;
- $H_{2.0}$ A adição do efeito da temperatura no modelo elétrico da bucha capacitiva gera dados com características próximas ao comportamento real.

1.4 Contribuições

A principal contribuição deste trabalho é uma estratégia para identificação de anomalias em buchas capacitivas. Claramente existe uma forte limitação de acesso a dados nesta área, devido ao tempo de operação dos equipamentos e ao custo do monitoramento, visto que quanto mais informações deseja-se obter, mais sensores e investimento é necessário por parte das concessionárias. Neste contexto, com o desenvolvimento da proposta, também foram geradas outras contribuições secundárias:

- Uma nova modelagem elétrica das buchas capacitivas com variações térmicas, gerado a partir dos estudos de trabalhos relacionados com o tema e dos dados adquiridos nas subestações.
- Um *dataset* para detecção do estado de uma bucha, gerado a partir dos dados de quatro unidades da subestação, com a rotulação de acordo com o estágio de degradação.
- Uma estratégia para estudar e representar o efeito da temperatura nas buchas capacitivas.

1.5 Organização do Trabalho

Este documento foi organizado da seguinte forma:

- Capítulo 1: Este capítulo apresenta a introdução e a motivação para o estudo sobre as buchas capacitivas. Também descreve o cenário em que a pesquisa está inserida e contextualiza o problema abordado no trabalho.
- Capítulo 2: Apresenta a fundamentação teórica, com os conceitos básicos sobre as buchas capacitivas, incluindo seu modelo elétrico e as principais métricas que podem ser monitoradas. Além disso, é feita uma breve descrição sobre a geração de dados e as técnicas utilizadas neste trabalho.
- Capítulo 3: Neste capítulo são apresentados os trabalhos relacionados, incluindo aspectos de monitoramento das buchas, modelagem do sistema trifásico e tipos de falhas em buchas capacitivas. Também são descritas as principais estratégias utilizadas para o monitoramento das buchas capacitivas.
- Capítulo 4: Este capítulo apresenta a metodologia utilizada para a realização deste projeto de pesquisa, incluindo a estratégia de identificação das anomalias e a modelagem da bucha capacitiva.

- Capítulo 5: Neste capítulo são apresentadas as análises e resultados obtidos, bem como uma discussão sobre estes resultados.
- Capítulo 6: Este capítulo apresenta a conclusão, também discute sobre possíveis trabalhos futuros e as dificuldades enfrentadas durante a pesquisa.

2 Fundamentação Teórica

Neste trabalho, focamos em sistemas de detecção de anomalias para buchas capacitivas. Estes sistemas visam possibilitar a identificação precoce dos defeitos nas buchas pelas concessionárias. Com a melhora na precisão destes sistemas é possível realizar ações de manutenção mais eficientes e, conseqüentemente, menos danos serão gerados ao SEP. Neste trabalho, em uma primeira etapa, serão estudadas as características das buchas capacitivas, o seu funcionamento e como são modelados os sinais monitorados, como a capacitância e o fator de potência. Além disso, o trabalho também analisa o desempenho de algoritmos de aprendizagem de máquina para avaliar os dados das buchas obtidos em subestações e a correlação entre as suas características.

2.1 Buchas

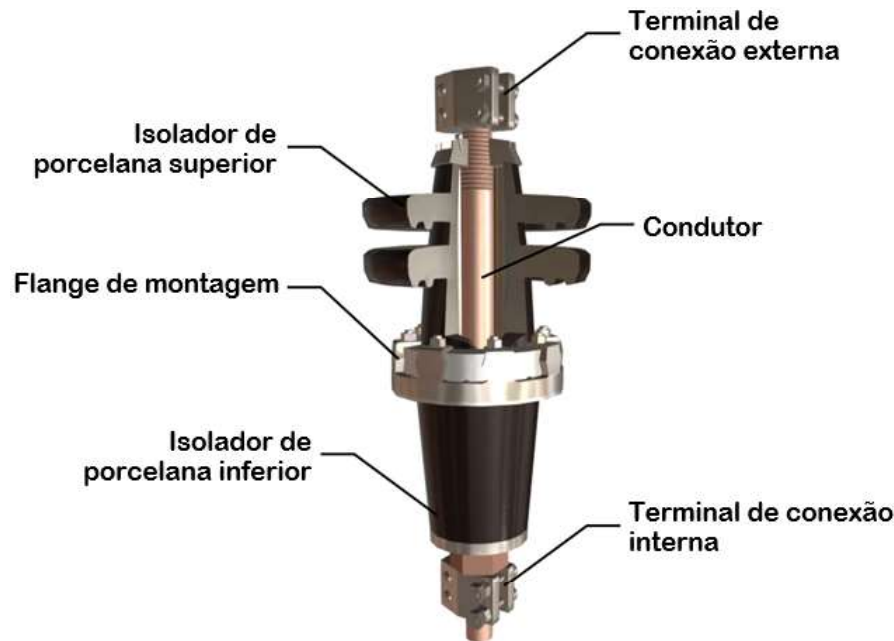
As buchas são um dos componentes mais importantes em um transformador de potência, visto que elas precisam fornecer tanto o suporte elétrico quanto mecânico (AHMED, 2011). Uma bucha de alta tensão deve fornecer o isolamento elétrico para a tensão nominal e eventuais sobretensões do sistema, além de servir como suporte mecânico para os condutores e outras conexões externas. A norma NBR-5034 (ABNT, 2014) define a bucha como "peça ou uma estrutura de material isolante, que assegura a passagem isolada de um condutor através de uma parede não isolante". Portanto, para verificar se a bucha está em condições de operação é importante que o isolamento seja avaliado. Para isso, é necessário avaliar de acordo com o tipo de bucha existente.

As buchas podem ser classificadas de acordo com seu isolamento das extremidades, isolamento interno ou sua construção (WAGENAAR, 2012). A escolha de uma bucha é determinada por algumas características como a tensão nominal, o tipo de material isolante e o ambiente em que está inserida (BOTELHO, 2011).

2.1.1 Buchas Não-Capacitivas

As buchas do tipo não-capacitivas ou não-condensivas são compostas por um condutor central revestido por um material isolante, como por exemplo a porcelana e a resina. Seu dimensionamento é relativo ao campo elétrico. Entretanto, o campo elétrico de buchas não capacitivas não é linear, tornando a sua fabricação inviável para tensões mais altas (acima de 52kV) (FRONTIN et al., 2013). Portanto, normalmente são buchas utilizadas em transformadores de distribuição e comutação de circuitos (AHMED, 2011). Na Figura 4 é mostrado um exemplo de bucha não-capacitiva.

Figura 4 – Exemplo de bucha não-capacitiva.



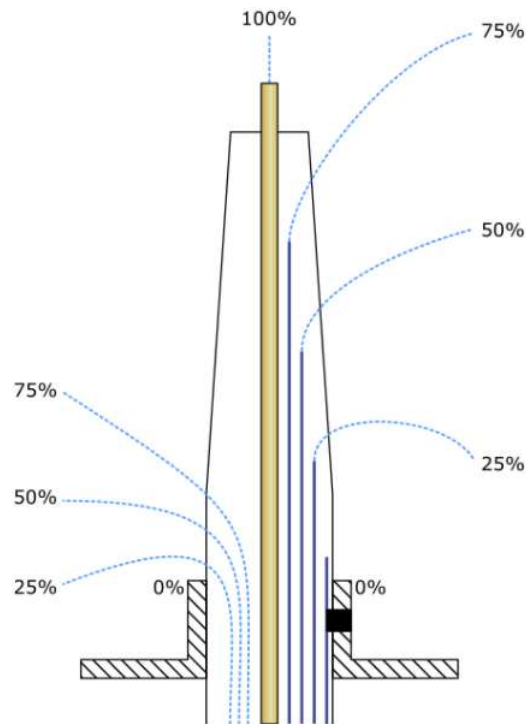
Fonte: Adaptado de SAVREE

2.1.2 Buchas Capacitivas

As buchas capacitivas são normalmente utilizadas em tensões acima de 52kV e têm a isolamento composta de material isolante e camada condutora ou semicondutora (FRONTIN et al., 2013). Um campo elétrico mais uniforme possibilita a redução das dimensões de uma bucha. A principal diferença entre buchas capacitivas e não-capacitivas é a formação de seu campo elétrico, como mostra a Figura 5. O lado esquerdo mostra uma bucha sem graduação capacitiva, ou seja, uma bucha não-capacitiva, que são constituídas de um corpo isolante sólido. Já o lado direito mostra a graduação capacitiva. Essa graduação em buchas capacitivas ocorre devido a construção do seu corpo em camadas, que causa uma melhor distribuição da tensão, criando um efeito capacitivo e um campo elétrico mais consistente. Conseqüentemente, é possível reduzir as dimensões, visto que há um maior controle do seu campo, possibilitando a sua utilização em tensões elevadas (FRONTIN et al., 2013). Entretanto, o processo de fabricação se torna um pouco mais complexo e, conseqüentemente, com custo mais elevado.

Para a construção do modelo elétrico de uma bucha capacitiva é preciso compreender como é a sua estrutura. A bucha capacitiva é feita utilizando um material isolante e chapas condutoras ou semicondutoras com espessura fina que são dispostos em camadas alternadas em volta do condutor central da bucha (FRONTIN et al., 2013). Segundo AHMED, as buchas capacitivas têm em sua estrutura cilindros metálicos concêntricos para equalizar a distribuição potencial. A sua construção consiste em folhas metálicas combinadas com

Figura 5 – Distribuição do campo elétrico nas buchas.



Fonte: SAVREE

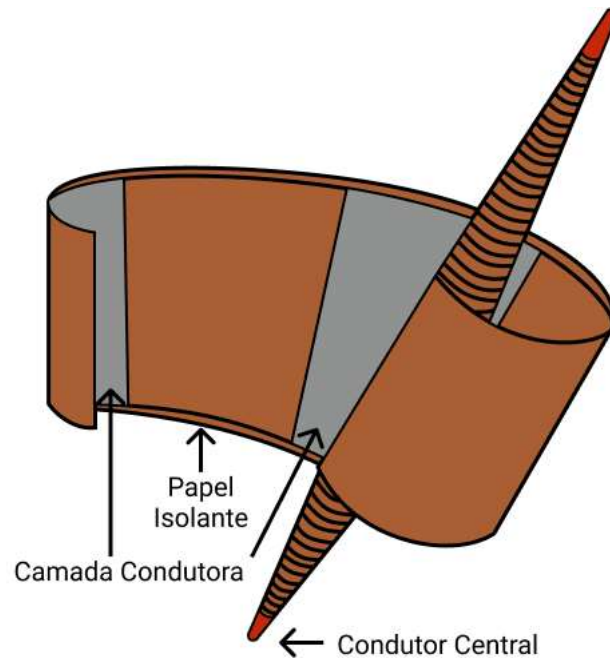
folhas impregnadas de óleo, como mostrado na Figura 6. As buchas são feitas de papel isolante enroladas em torno do condutor até o flange metálico com folhas metálicas inseridas entre as camadas de papel para gerar uma série de capacitores entre o condutor central e a flange. A distância entre cada folha é calculada de tal forma a uniformizar a distribuição da tensão. Existem três tipos de buchas com essa estrutura (FRONTIN et al., 2013):

- Papel Aglutinado com Resina (*Resin Bonded Paper - RBP*)
- Papel Impregnado com Óleo (*Oil Impregnated Paper - OIP*)
- Papel Impregnado com Resina (*Resin Impregnated Paper - RIP*)

2.1.3 Modelo Elétrico da Bucha Capacitiva

Devido a sua estrutura em camadas, a bucha capacitiva pode ser representada por um conjunto de capacitores iguais em série, dividido em duas partes. A primeira parte é a mais interna e conhecida como C_1 e a segunda é a mais externa e conhecida como C_2 . A Figura 7 ilustra as vistas superior e lateral da bucha capacitiva, além do seu modelo elétrico equivalente elétrico. Nessa modelagem, a capacitância C_1 representa todos os capacitores que estão entre o condutor central e o tape de teste, já a capacitância C_2 representa a

Figura 6 – Exemplo de construção de uma bucha capacitiva.



Fonte: Adaptada de BHUTADA et al., 2017

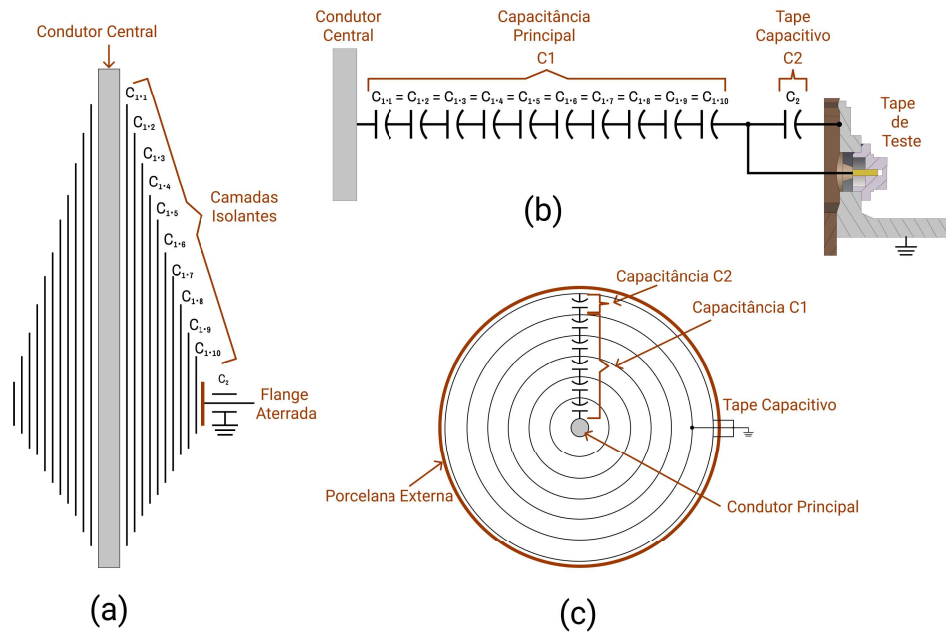
capacitância do tape. Nas Figuras 7(a) e 7(c) é possível visualizar as capacitâncias de acordo com as camadas da estrutura (C_1) e com o tape capacitivo (C_2). O tape capacitivo é o local acessível no lado externo da bucha, que possibilita as medições. Ele fica localizado próximo ao flange e deve ser aterrado quando não está sendo utilizado, como é mostrado na Figura 7(b).

2.1.4 Capacitância, Fator de Dissipação e Fator de Potência

Neste contexto, a capacitância e a tangente delta são as principais medidas para diagnosticar o estado de uma bucha. Essas medidas são sensíveis às mudanças na bucha, sendo assim o ingresso de umidade no equipamento aumenta o fator de dissipação e um curto-circuito de uma camada de isolamento aumenta o valor da capacitância (ARAÚJO et al., 2013).

Para tensões iguais ou acima de 52kV, a norma IEC 60137-2008, "*Insulated bushings for alternating voltages above 1 000 V*" (COMMISSION, 2008), define que as medições do fator de potência e da capacitância devem ser realizadas rotineiramente pela fabricante. Para esses testes a temperatura ambiente está entre 10 °C e 40 °C e a tensão é de 1 a 1,05 vezes a tensão máxima de operação da bucha. Nessas condições, a tangente delta não deve ultrapassar o valor de 0,7% e para a medição de capacitância não são definidos valores limites.

Figura 7 – (a) Vista lateral. (b) Representação elétrica. (c) Vista superior.



Fonte: Elaborada pela autora.

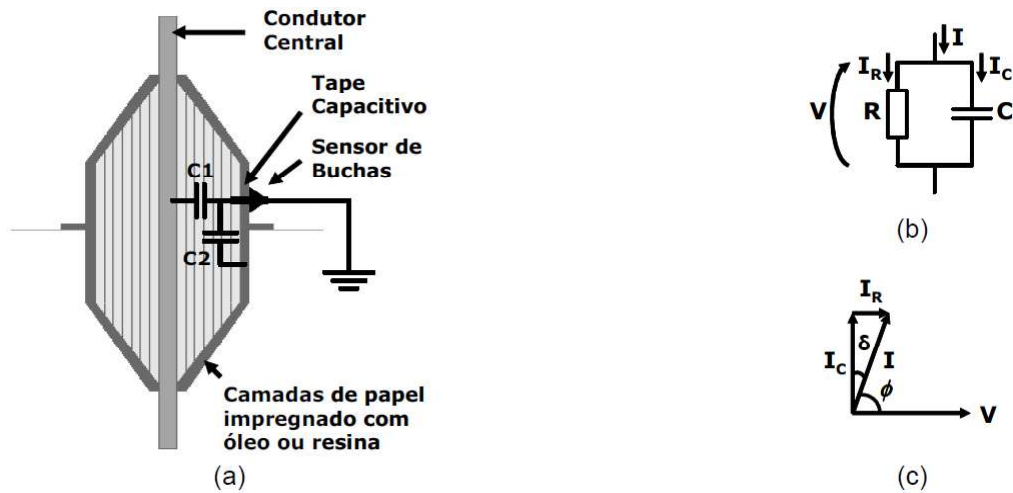
A norma ABNT NBR 5034-2014 ((ABNT), 2014) apresenta o limiar para a tangente delta para buchas do tipo RIP de 1,5% ao invés de 0,7% e um limiar de 1% para a variação de capacitância antes e depois da série de ensaios rotineiros.

Já a norma ANSI/IEEE C57.19.100-2012, "*IEEE Guide for Application of Power Apparatus Bushings*" (ELECTRICAL; ENGINEERS; C57.19.100-2012, 2013), sugere um aumento na frequência dos testes, caso a bucha apresente um aumento de tangente delta entre 1,5 e 2 vezes do seu valor inicial ou um aumento de 5% da capacitância em relação ao seu valor inicial.

As principais medidas para avaliar uma bucha capacitiva são a medida da capacitância e o fator de dissipação, primeiramente publicadas por Schering (SCHERING, 1919). As buchas dos transformadores de alta tensão são construídas de forma que a capacitância e o fator de dissipação sejam constantes ao longo de sua vida útil. Sendo assim, a variação desses parâmetros indica a deterioração do equipamento. Essas medidas são calculadas com base no modelo elétrico. A Figura 8(a) mostra uma simplificação do modelo, ou seja, a bucha modelada por apenas duas capacitâncias principais (C_1 e C_2). Entretanto, para abranger o efeito de envelhecimento é preciso um modelo mais completo que demonstre as perdas dielétricas.

Idealmente, considerando um circuito de corrente alternada, o diagrama vetorial de uma bucha capacitiva é uma corrente defasada de 90° da tensão, visto que é um circuito puramente capacitivo. Entretanto, para que o modelo seja mais real é necessário representar

Figura 8 – (a) Detalhes construtivos de uma bucha capacitiva de alta tensão. (b) Circuito equivalente das perdas dielétricas. (c) Diagrama vetorial das correntes.



Fonte: RAMBO et al., 2007

as imperfeições como as perdas dielétricas do sistema. Essas perdas de potência ativa são causadas pela passagem das correntes pelo isolamento da bucha e é representada pelo resistor em paralelo à capacitância como ilustrado na Figura 8(b). O vetor de corrente resultante desse modelo é demonstrado na Figura 8(c), sendo possível visualizar que a corrente resultante é composta por duas partes: uma resistiva e uma capacitiva. Dessa forma, é possível reproduzir a variação do fator de dissipação.

O valor da $\tan(\delta)$, também conhecido como fator de dissipação (*Dissipation Factor* – DF), é a medida do condutor central até o tape de teste que representa a qualidade da isolamento do equipamento. Esse fator é determinado principalmente pelas contaminações no sistema de isolamento, além de depender da umidade no papel e da temperatura. O conhecimento do fator de dissipação é importante para averiguar as perdas capacitivas, que podem indicar o envelhecimento prematuro do isolamento. Outro motivo é a verificação da qualidade da fabricação do equipamento. Portanto, este fator é uma das principais medidas para o monitoramento de buchas. O valor é calculado como a razão entre a corrente resistiva (I_R) e a corrente capacitiva (I_C), conforme definido pela Equação 2.1. Nesse contexto, o aumento do fator indica uma deterioração da bucha, ou seja, a qualidade do seu isolamento está comprometida.

$$DF = \tan(\delta) = \frac{I_R}{I_C} \quad (2.1)$$

- DF: Fator de Dissipação;
- δ : Ângulo entre a corrente capacitiva e a corrente resultante;

- I_R : Corrente Resistiva;
- I_C : Corrente Capacitiva.

$$PF = \cos(\phi) = \frac{I_R}{I} \quad (2.2)$$

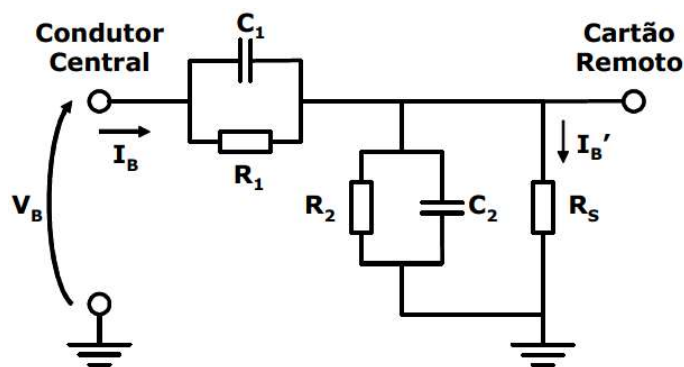
- PF: Fator de Potência;
- ϕ : Ângulo entre a tensão e a corrente resultante;
- I_R : Corrente Resistiva;
- I_C : Corrente Resultante.

Já o fator de potência (Power Factor – PF) é obtido calculando o $\cos(\phi)$ que é a corrente do resistor dividida pela corrente total ou, como mostrado na Equação 2.2. O ângulo δ representa o ângulo de perda e é o complemento do ângulo ϕ (Figura 8(c)). Esse fator também especifica as perdas do isolamento, porém essa medida é mais utilizada nos Estados Unidos, Canadá e alguns outros países (MURTY, 2017a). A relação entre os fatores de potência e de dissipação estão definidas nas Equação 2.3 e Equação 2.4.

$$\cos(\phi) = \frac{\tan(\delta)}{\sqrt{1 + \tan^2(\delta)}} \quad (2.3)$$

$$\tan(\delta) = \frac{\cos(\phi)}{\sqrt{1 + \cos^2(\phi)}} \quad (2.4)$$

Figura 9 – Circuito equivalente do isolamento da bucha de alta tensão e do sensor.



Fonte: Adaptado de RAMBO et al., 2007

A Figura 9 ilustra um circuito equivalente com o sensor sendo representado por uma resistência R_S . Os autores RAMBO et al. simplificam o sensor como sendo um

resistor R_S de baixa impedância se comparado com a do isolamento da bucha. Utilizando o circuito equivalente apresentado, é preciso fazer algumas considerações para o cálculo da capacitância C_1 . Primeiramente, as perdas são causadas pela passagem das correntes pelo isolamento da bucha. Normalmente, o tape capacitivo está aterrado e não recebe esses estresses elétricos (RAMBO et al., 2007), portanto é considerado que a capacitância C_2 não altera seu valor. Além disso, assume-se que $|X_{C_2}| \gg R_S$, portanto a corrente que passa pela capacitância C_2 e pela resistência R_2 é praticamente nula, ou seja, $I_B \approx I'_B$. Em buchas novas, há pouca perda dielétrica, conseqüentemente o fator de potência é pequeno, ou seja, baseando-se no diagrama vetorial da Figura 8(c) implica que $R_1 \gg X_{C_1}$ e o circuito resultante é R_S em série com C_1 . Assim, a capacitância pode ser determinada pela divisão de tensão e calculada pela Equação 2.6.

$$X_{C_2} = \frac{1}{j\omega C_2} \quad (2.5)$$

- X_{C_2} : Reatância Capacitiva do Capacitor C_2 ;
- ω : Frequência angular;
- C_2 : Capacitância do Tape Capacitivo.

$$C_1 = \frac{|I'_B|}{2\pi f |V_B|} \quad (2.6)$$

- C_1 : Capacitância Principal;
- I'_B : Corrente no Sensor;
- f : Frequência da Rede;
- V_B : Tensão da Bucha.

2.2 Aprendizagem de Máquina

O crescente número de dados, desenvolvimento de novas tecnologias e a necessidade de avaliações mais complexas, faz com que novos métodos sejam necessários para que haja uma solução mais condizente com o desejável. Nesse contexto, a Aprendizagem de Máquina (*Machine Learning - ML*), que é um ramo da Inteligência Artificial (IA), está cada vez mais se destacando no campo de análise de dados. Esse campo de estudo é utilizado para realizar análises mais complexas onde algoritmos mais clássicos não conseguem resolver, visto que utilizam técnicas de forma a encontrar soluções a partir da identificação de padrões de dados. Essa abordagem pode ser aplicada em combate a fraudes, classificação de *e-mails*,

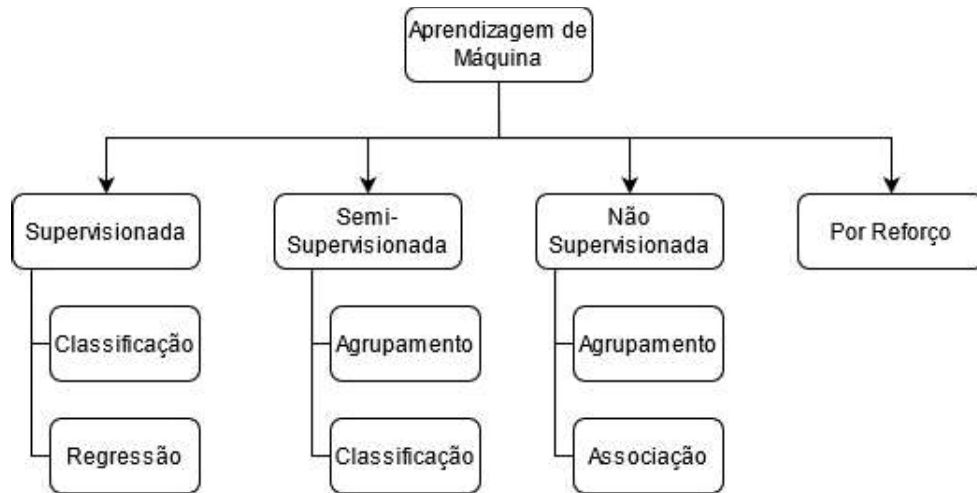
análise textual, recomendação de conteúdos, *smart grids*, detecção de anomalias, entre outros. Dessa forma, essas técnicas tornaram-se um dos componentes mais importantes em diagnósticos e que auxiliam em diversas aplicações do cotidiano.

O pioneiro da inteligência artificial foi Arthur Samuel em 1959. Ele definiu a aprendizagem de máquina como: "campo de estudo que dá aos computadores a habilidade de aprender sem serem explicitamente programados"(utilizando a tradução livre) (SIMON, 2013). Nesse sentido, é um método de análise de dados que simula a aprendizagem de um humano em uma máquina através de treinos. Nesse contexto, cada conjunto de dados, também conhecidas como instâncias, são formados por atributos de entrada (*features*), que podem ter um atributo alvo (BARANAUSKAS; MONARD, 2000a).

Os tipos de aprendizagem de máquina estão mostrados na Figura 10 e são basicamente: supervisionada, não supervisionada, semi-supervisionada e por reforço (GÉRON, 2019). Na aprendizagem supervisionada o modelo é gerado a partir de dados rotulados, sendo a rotulação com números contínuos chamada de regressão e a rotulação categórica de classificação. De forma oposta, a aprendizagem não supervisionada utiliza dados não rotulados. Nesse contexto, o objetivo é identificar similaridades entre as instâncias e agrupá-las, caso haja alguma instância muito fora de qualquer grupo, pode ser considerado como uma anomalia do sistema. Já a aprendizagem semi-supervisionada utiliza tanto instâncias rotuladas quanto instâncias não rotuladas. Essa é uma alternativa para diminuir o custo da rotulação e ainda obter resultados satisfatórios (GÉRON, 2019). Por último, a aprendizagem por reforço é a que difere mais dentre os tipos de aprendizagem, pois não utiliza uma base de dados e a aprendizagem ocorre por meio de estímulos como a recompensa por ações desejadas, como acontece em jogos digitais, carros autônomos e na robótica em geral.

Há vários algoritmos de aprendizagem de máquina e cada um é apropriado para um tipo de problema específico. Algumas das principais tarefas são: regressão, classificação, árvore de decisão e redes neurais. Para o presente trabalho foi dado um foco maior para as técnicas de classificação. Como já mencionado antes, a classificação é uma tarefa supervisionada e, portanto, necessita de uma base de dados rotulada. Nesse caso, algumas das métricas de avaliação dos modelos de classificação são a taxa de acerto, a precisão (*precision* - P), revocação (*recall* - R) e a *f-measure* ou *f1-score* (GÉRON, 2019). Sendo assim, em um problema binário (Positivo e Negativo), o Verdadeiro Positivo representado por VP e Verdadeiro Negativo por VN são dados que foram corretamente classificados. Por outro lado, o Falso Positivo (FP) e o Falso Negativo (FN) representam a parte dos dados que foram classificados incorretamente. Em problemas como a desse trabalho que não apresentam apenas duas classes, para estimar esses valores (VP, FP, FN e FP) é possível analisar a classe desejada como positiva e as demais como negativas (BARANAUSKAS; MONARD, 2000b).

Figura 10 – Tipos de aprendizagem de máquina.



Fonte: Elaborada pela autora.

A taxa de acerto é calculada pela Equação 2.7. Essa taxa avalia a porcentagem de dados classificados corretamente, mas não é a melhor métrica utilizada para avaliar um problema desbalanceado, ou seja, problemas onde a quantidade de exemplos de cada classe é desequilibrada. Já a precisão, que é calculada na Equação 2.8, verifica a proporção das instâncias classificadas como positivas que foram corretas. Nesse sentido, caso não haja FP, a precisão é igual a 1 (BARANAUSKAS; MONARD, 2000b). Já a revocação analisa a proporção das amostras positivas que foram classificadas corretamente, por isso, utilizado Equação 2.9, caso o FN seja nulo, a *recall* é igual a um. Por fim, a *f-measure* consiste na média harmônica entre a precisão e a revocação (BARANAUSKAS; MONARD, 2000b), sendo que o valor mínimo é zero e o máximo é um. O resultado ser um significa que a média entre a *recall* e a precisão é perfeita. A utilização dessas métricas depende da base ser balanceada ou desbalanceada e o tipo de algoritmo utilizado. Como será utilizado apenas classificadores, essas métricas são as principais formas de avaliar os algoritmos.

$$\text{Taxa de acerto} = \frac{VP + VN}{VP + VN + FP + FN} \quad (2.7)$$

$$\text{Precisão (P)} = \frac{VP}{VP + FP} \quad (2.8)$$

$$\text{Revocação (R)} = \frac{VP}{VP + FN} \quad (2.9)$$

$$\text{F-measure} = 2 \cdot \frac{P \cdot R}{P + R} \quad (2.10)$$

- VP: Verdadeiro Positivo;
- VN: Verdadeiro Negativo;
- FP: Falso Positivo;
- FN: Falso Negativo;
- P: Precisão;
- R: Revocação.

Existem diversas abordagens voltadas para classificação de dados presentes na literatura, as quais apresentam diferenças relacionadas ao custo computacional, dificuldade de implementação, facilidade de interpretação dos resultados, dentre outros fatores que influenciam na escolha da melhor técnica considerando o problema a ser resolvido. As principais abordagens de classificação utilizadas consistem na utilização de algoritmos monolíticos, combinação de classificadores e aprendizagem profunda (*Deep Learning*). Dentre os vários algoritmos existentes de aprendizagem de máquina, para este trabalho foram selecionado alguns dos mais utilizados, como por exemplo: árvores de decisão (algoritmo J48), K Vizinhos Mais Próximos (*K Nearest Neighbor - KNN*) e *Naive Bayes*. Além dos sistemas de múltiplos classificadores como: *Random Forest*, *Bagging*, *Boosting* e *Random Subspaces*

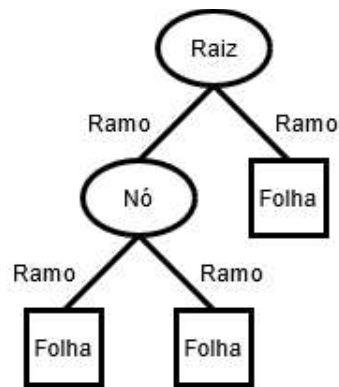
2.2.1 Árvore de Decisão

A árvore de decisão é um algoritmo de classificação que é fácil de compreender devido à sua forma simples e intuitiva de interpretação. O algoritmo de Hunt é a base de vários algoritmos de árvore de decisão, como a ID3, tendo como sucessor o C4.5 (TAN; STEINBACH, 2004). Para resolver um problema de classificação com uma árvore de decisão, são realizadas séries de comparações até obter uma resposta. Este algoritmo mapeia os possíveis resultados ordenando as classificações das instâncias abaixo da árvore iniciando pela raiz. Segundo a Figura 11, o nó raiz gera a primeira avaliação e seus ramos, que representam os possíveis valores dos atributos, levam aos prováveis resultados: um novo nó ou uma folha, que representa as possíveis classes do problema. Geralmente o atributo na raiz é a que melhor classifica sozinho os dados. Portanto, é utilizado testes estatísticos para a escolha de qual deve ser o atributo testado na raiz.

2.2.2 KNN

O algoritmo k vizinhos mais próximos, mais conhecido como *K Nearest Neighbors* (KNN) é um classificador mais básico baseado em instâncias. O algoritmo assume que

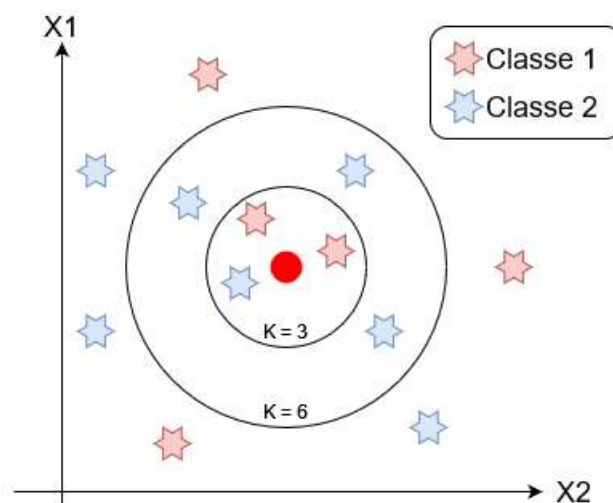
Figura 11 – Estrutura de uma Árvore de Decisão.



Fonte: Elaborada pela autora.

as instâncias correspondem a um ponto em um espaço de dimensão n para classificá-los de acordo com a distância entre cada ponto. O KNN tem como objetivo classificar cada instâncias considerando as k amostras mais próximas, ou seja, há uma avaliação de qual classe aparece mais em sua vizinhança para ser rotulado de acordo com a maioria. Normalmente, essa distância é definida de acordo com a distância Euclidiana. Como demonstrado na Figura 12, dependendo do número k , a classe em que o ponto vermelho pertence é alterado. Neste contexto, é visível a facilidade de compreensão e aplicação. Entretanto, dependendo do problema, pode virar uma solução muito demorada e não ser muito eficiente.

Figura 12 – Aplicação de uma KNN.



Fonte: Elaborada pela autora.

2.2.3 Naive Bayes

O algoritmo *Naive Bayes* é um classificador probabilístico que foi baseado no “teorema de Bayes”, criado por Thomas Bayes. Esse algoritmo é simples, rápido e escalonável. O teorema descreve a probabilidade de ocorrer um evento a partir do conhecimento obtido anteriormente. Além disso, um de seus princípios é a forte independência entre as variáveis de uma instância, ou seja, o algoritmo desconsidera a correlação entre as variáveis para a rotulação, por esse motivo o *Naive Bayes* é chamado de ingênuo (*naive*). O *Naive Bayes* usa a teoria bayesiana de forma que as densidades de probabilidade condicional da classe e as probabilidades anteriores são estimadas a partir de um conjunto de treinamento (DUDA; HART; G.STORK, 2001). Portanto, a classificação é obtida através da atribuição de cada amostra à classe de uma probabilidade máxima posteriori calculada pelas Equação 2.11 e Equação 2.12, onde y especifica os rótulos da classe e x é uma dada amostra.

$$Y_{MAP} = \operatorname{argmax}_i P(y_i|x) \quad (2.11)$$

- Y_{MAP} : Classe Máxima a *Posteriori*;
- $\operatorname{argmax}_i P(y_i|x)$: Maior probabilidade condicional da amostra x de ser a classe y_i .

$$P(y_i|x) = \frac{P(y_i)P(x|y_i)}{P(x)} \quad (2.12)$$

- $P(y_i|x)$: Probabilidade da classe y_i dado a amostra x ;
- $P(y_i)$: Probabilidade a *priori* da classe y_i ;
- $P(x|y_i)$: Probabilidade da amostra x dado a classe y_i ;
- $P(x)$: Probabilidade a *priori* do evento x .

2.2.4 Combinação de Classificadores

A abordagem, conhecida também como Sistemas de Múltiplos Classificadores (SMC) ou *pools* de classificadores são uma alternativa para a melhora da eficácia dos métodos monolíticos, adotando vários classificadores no processo classificatório (KO; SABOURIN; Britto, Jr., 2008). Quanto mais diversos a combinação de classificadores, os resultados tendem a ser melhores. Nesse contexto, a combinação tem como objetivo aumentar a precisão em troca de uma maior complexidade. Segundo KO; SABOURIN; Britto, Jr., esse sistema pode ser dividido em três fases: geração, seleção e integração.

Na primeira etapa, é feita a geração de *pools* de classificadores. Os classificadores podem ser homogêneos ou heterogêneos. Para KO; SABOURIN; Britto, Jr., independentemente do tipo utilizado é preferível que os classificadores tenham características complementares, ou seja, se um classificador tem uma alta taxa de acerto em apenas uma classe, o outro classificador ou o resto do conjunto de classificadores necessitariam apresentar uma taxa de acerto elevada nas outras classes para que a combinação gerasse um resultado melhor. Já a segunda fase seleciona tendo como critério o tipo (estático ou dinâmico) e a competência do classificador. Essa competência pode ser avaliada de forma estatística, computacional ou representacional. Por último, é feita a integração dos classificadores selecionados utilizando estratégias como a fusão ou seleção dos classificadores. Sendo que na fusão cada classificador tem conhecimento completo das características e na seleção cada classificador conhece bem e é responsável por parte do espaço.

Além das três fases, a abordagem também pode ser dividido basicamente em dois subconjuntos, homogêneos e heterogêneos. Na primeira abordagem, durante o processo de geração são adotados os métodos semelhantes de construção. Já na segunda, diferentes algoritmos são aplicados ao longo do processo. Para este trabalho, foram utilizadas apenas as abordagens homogêneas. Dentre as abordagens homogêneas mais aplicadas tem-se *Random Forest* (HO, 1995), *Bagging* (BREIMAN, 1996), *Boosting* (FREUND; SCHAPIRE et al., 1996) e *Random Subspaces* (HO, 1998).

2.3 Correlação de Pearson

A correção de Pearson (ρ) calcula o grau de correlação e a direção dela entre duas variáveis. Essa correlação é utilizada para indicar como duas variáveis são associadas (HOLLANDER; WOLFE; CHICKEN, 2013). O valor do coeficiente varia entre -1 e 1. Sendo que o valor 0 indica que não há uma associação. Um resultado maior que 0 indica uma associação positiva e menor que 0 uma associação negativa. Entretanto, o coeficiente só representará adequadamente se tiverem uma relação linear, as variáveis precisam ter uma distribuição aproximada e os dados não podem conter dados atípicos. Dessa forma, na Equação 2.13 é mostrado cálculo da correlação, onde X e Y representam os vetores das variáveis que serão analisadas a correlação, portanto \bar{X} e \bar{Y} são as suas médias e X_i e Y_i os elementos dos respectivos vetores.

$$\rho = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X}) \sum_{j=1}^n (Y_i - \bar{Y})}} \quad (2.13)$$

- ρ : Probabilidade da classe y_i dado a amostra x ;
- X_i : Elemento i do vetor X ;

- \bar{X} : Média dos elementos do vetor X ;
- Y_i : Elemento i do vetor Y ;
- \bar{Y} : Média dos elementos do vetor Y .

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} \quad (2.14)$$

- X_i : Elemento i do vetor X ;
- \bar{X} : Média dos elementos do vetor X .
- n : Número de elementos.

$$\bar{Y} = \frac{\sum_{i=1}^n Y_i}{n} \quad (2.15)$$

- Y_i : Elemento i do vetor Y ;
- \bar{Y} : Média dos elementos do vetor Y .
- n : Número de elementos.

Para interpretar os valores, o autor COHEN (2013) define que entre 0,10 e 0,29 os resultados podem ser considerados com correlação pequena, entre 0,30 e 0,49 como média correlação e entre 0,50 e 1 podem ser interpretados como alta correlação entre as variáveis. Entretanto, é preciso ressaltar que a avaliação da correlação é sensível a anomalias, ou seja, é sugerido que essa avaliação seja realizada por meios visuais, como gráficos, para ter mais certeza da correlação calculada.

3 Trabalhos Relacionados

A identificação precoce de falhas nas buchas capacitivas pode ser feita através de várias medidas. Para este trabalho, o foco foi dado ao fator de potência e o valor da capacitância da bucha. Este capítulo apresenta os trabalhos relacionados à pesquisa realizada dividida em três categorias: monitoramento das buchas, modelagem do sistema trifásico e tipos de falhas em buchas capacitivas.

3.1 Monitoramento de Buchas

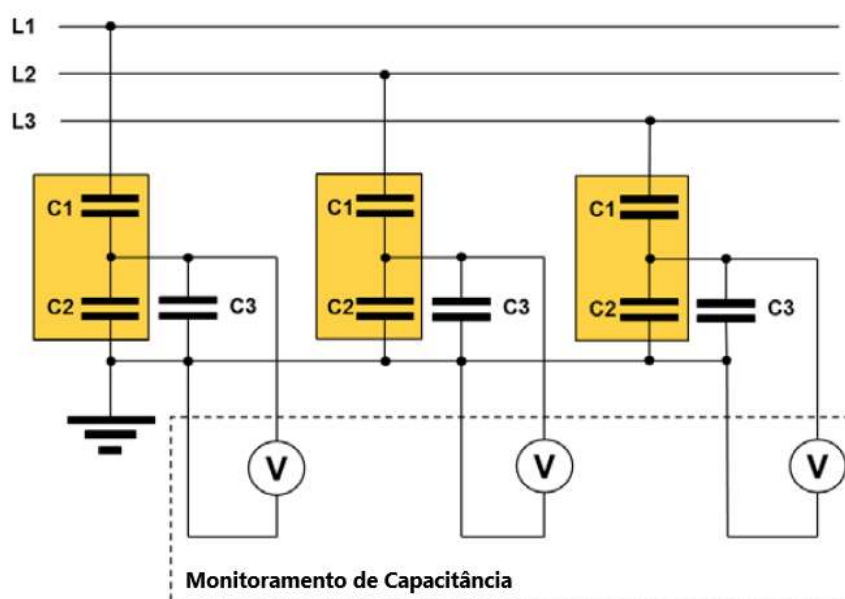
O monitoramento de buchas de alta tensão dos equipamentos tem como objetivo detectar a degradação do isolamento e auxiliar nas tomadas de decisão. A vida útil de uma bucha capacitiva é praticamente a metade da vida útil de um transformador, segundo o autor MORALES. Entretanto, mesmo com essa estimativa de vida, esse valor pode variar dependendo do estresse recebido. Sendo o envelhecimento, a umidade e qualidade da fabricação as principais causas da diminuição do tempo de degradação. Conforme é disposto na literatura, existem diversas formas de monitoramento de buchas. As principais métricas são a tangente delta ou a variação de sua capacitância. Inicialmente, essas métricas eram apenas medidas de forma *off-line*. Mas diversas abordagens já foram propostas para que monitoramento *on-line* dessas métricas fosse mais precisa. Para esta pesquisa, os trabalhos relacionados referentes ao monitoramento de buchas são utilizados como referência para compreender como são medido a capacitância e/ou a tangente delta.

Segundo o autor AJ et al., uma das técnicas que se destacou na indústria é a Análise de Assinatura Elétrica (*Electrical Signature Analysis-ESA*). A ESA é um termo genérico para o conjunto de medições e técnicas aplicadas para o monitoramento da condição das máquina elétricas. Essa análise é realizada por meio de medições de sinais elétricos como a tensão e a corrente. Essa técnica é cada vez mais conhecida devido a sua precisão nas medidas e a otimização das previsões relacionadas às condições dos equipamentos elétricos.

A avaliação das buchas capacitivas pode ser tanto *on-line* quanto *off-line*. Os principais métodos *off-line* encontrados foram: as medidas de capacitância e da tangente de delta, ensaios de suportabilidade e de descargas parciais, ensaios de impulso de tensão e ensaios de estabilidade térmica. Além disso, há autores como ATALLAH et al. que defendem a importância da verificação das condições de armazenamento para comparar com os resultados dos ensaios. Os métodos *off-line* ainda são utilizadas como referência de medição para a capacitância e o fator de dissipação devido a sua precisão, portanto o monitoramento de uma bucha será, em sua maioria, com referência ao seu valor medido enquanto desligada. Portanto, as medições *off-line* são utilizadas junto com as *on-line*.

Uma das formas mais fáceis de monitoramento *on-line* é o monitoramento da capacitância (MORALES, 2012). Esse método é utilizado para detectar falhas parciais entre uma ou mais camadas capacitivas com o auxílio de um capacitor C_3 conectado no tape e aterrado, como mostra a Figura 13. As medidas são baseadas em um divisor capacitivo em que o aumento permanente da tensão de 5% a 10% equivale à uma falha no equipamento, segundo o autor MORALES. Dependendo do ruído existente no sistema, essa alteração pode ser imperceptível. Por esse motivo, o sistema de aquisição de medidas deve ser construído de forma a amenizar o máximo de ruídos possível.

Figura 13 – Modelo de monitoramento de capacitância.



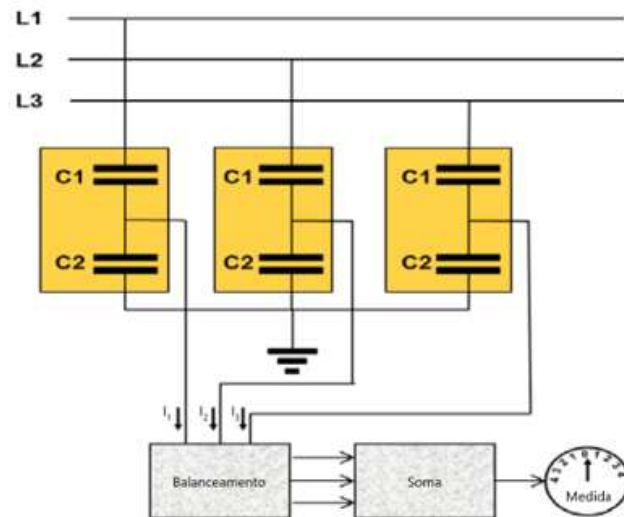
Fonte: Adaptado de MORALES, 2012

Como já foi possível deduzir, a desvantagem deste método de monitoramento de capacitância é a suscetibilidade a variações de tensão, ou seja, as flutuações da rede e variações de quantidade de camada de uma bucha. Outro fator, é que é possível medir apenas perdas parciais da corrente e problemas de contato. Dependendo do nível de tensão, a diferença de uma camada pode alterar em aproximadamente 1% o resultado. Além disso, em tensões altas, uma variação de 2% na magnitude poderá comprometer toda a análise (MORALES, 2012). Para ajustar os valores seria preciso fazer comparações entre a tensão primária e a secundária do transformador.

A medição das correntes balanceadas ou soma das correntes é um dos métodos mais utilizados. Como o nome sugere, trata-se de somar as correntes de cada fase de um mesmo transformador. Idealmente, quando a bucha é nova o resultado da soma é nulo. Isso acontece se considerarmos que os valores são de mesma magnitude e com

defasagem de 120° , adquiridas do mesmo transformador e com mesmo tipo de bucha nas três fases. Entretanto, sabe-se que há desbalanceamento de cargas na rede elétrica, que geram resultados diferentes de zero.

Figura 14 – Modelo de monitoramento pelas correntes balanceadas.



Fonte: Adaptado de MORALES, 2012

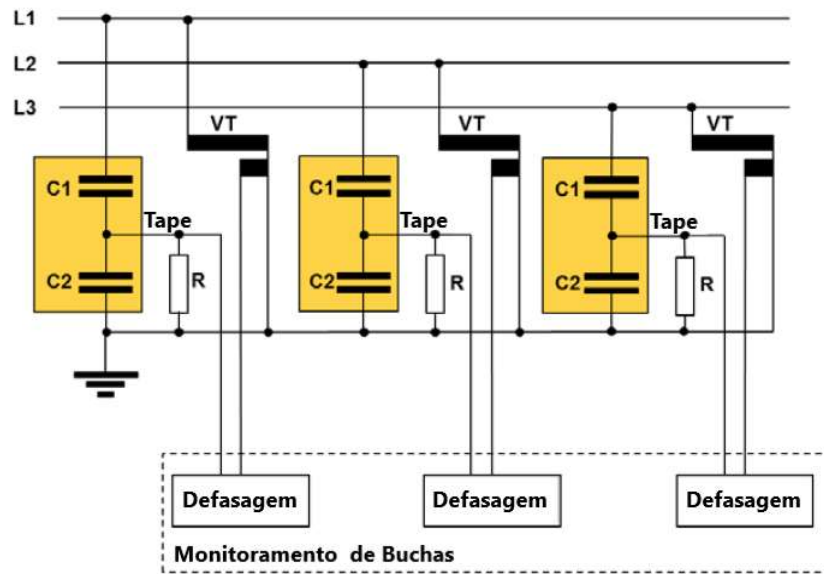
As capacitâncias C_1 e C_2 apresentadas na Figura 14 são as representações da capacitância principal e do tape capacitivo, respectivamente. Como nas subestações não é possível obter as condições perfeitas, é preciso uma etapa de pré-processamento dos sinais. As variações presentes nos sinais são tipicamente de 1% a 1,5% na magnitude e de 0,2% de defasagem do ângulo (MORALES, 2012). Entretanto, em redes mais instáveis e com cargas bastantes desbalanceadas pode-se obter porcentagens maiores. Neste cenário, os valores de cada sensor serão ajustados antes de serem somados ou terão algum sistema de auto aprendizagem para compensar o desbalanceamento causado pela rede.

Nota-se que esta técnica presume que a degradação ocorrerá de forma diferente nas três buchas, pois a soma das correntes é diferente de zero quando há degradação. Outro ponto importante é que as flutuações na magnitude das três fases têm uma compensação melhor que no método de monitoramento de capacitância, pois é considerado a flutuação geral e os sinais podem ser comparados de forma pareada com os outros transformadores da subestação.

Para aumentar a precisão no monitoramento e superar as desvantagens dos outros métodos, o autor MORALES sugere uma fonte de sinal que seja independente da carga como referência. Dessa forma, os ruídos em uma mesma fase e os efeitos de cargas podem ser descartados para detectar o envelhecimento da bucha.

A Figura 15 mostra um exemplo de referência, a tensão (VT) de um transformador

Figura 15 – Modelo com sinal de referência.



Fonte: Adaptado de MORALES, 2012

sugerida pelo autor MORALES. A comparação entre a medida da referência e a corrente de fuga reduz a utilização do *hardware* ao mínimo. Como é utilizada a mesma fase, conseqüentemente, espera-se que haja uma maior precisão devido à ausência da assimetria.

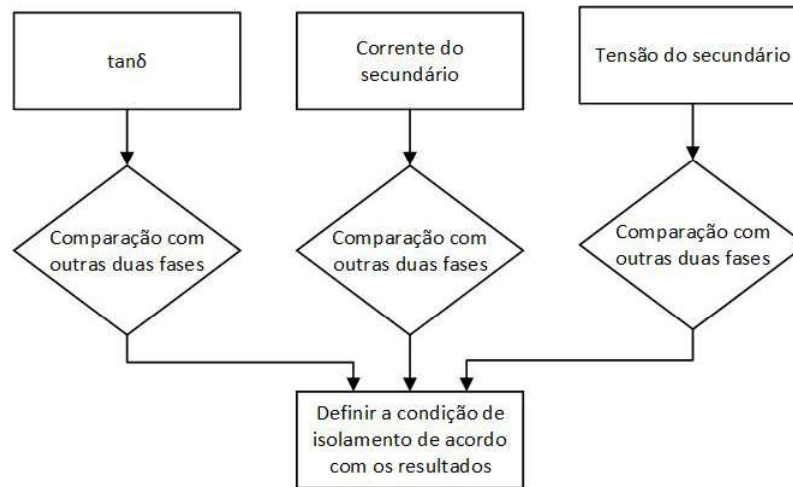
XIA et al. realizaram uma pesquisa sobre um método de monitoramento *on-line* para analisar a condição do isolamento das buchas de transformadores. No artigo é discutido os fatores que influenciam na precisão da medição do ângulo da perda dielétrica, como: a interferência entre fases, a flutuação de frequência da rede, a temperatura/umidade do ambiente e a interferência das harmônicas do sistema. Para amenizar esses efeitos indesejáveis foi proposto a combinação do método de comparação fase a fase e do método de análise de harmônicas (Figura 16). Sendo assim, a Tabela 1 mostra como são feitas as decisões utilizando o método proposto. Dessa forma, o método proposto pelos autores XIA et al. visa amenizar esses efeitos para melhorar a precisão das medidas.

Tabela 1 – Diagnóstico das buchas.

δ_{12}	δ_{13}	δ_{23}	Resultado
Houve mudanças	Houve mudanças	Quase invariante	Falha na bucha 1
Houve mudanças	Quase invariante	Houve mudanças	Falha na bucha 2
Quase invariante	Houve mudanças	Houve mudanças	Falha na bucha 3

Fonte: Adaptado de XIA et al., 2016

Figura 16 – Diagrama de blocos do método de comparação fase a fase.



Fonte: Adaptado de XIA et al., 2016

Como o monitoramento *on-line* apresenta algumas incertezas devido a ruídos do sistema, ANDRZEJEWSKI; GIL; MIKRONIKA propõem um método mais complexo, visto que acreditam que a aquisição de dados com maior precisão necessita de um algoritmo que seja mais elaborado. Devido às aquisições ocasionais, o monitoramento *off-line* não é suficiente para a tomada de decisões. Tendo em vista que os métodos mais tradicionais focam em falhas que se desenvolvem mais lentamente, o autor entende que essa análise não é muito condizente com o real estado da bucha, já que é medido quando o equipamento está desligado e fora das condições normais de desempenho. Portanto, o sistema aplicado por ANDRZEJEWSKI; GIL; MIKRONIKA realiza o monitoramento *on-line* com integração de dados.

O transformador monitorado é analisado a partir de cinco informações principais: o monitoramento da bucha, a análise da unidade, o monitoramento do comutador, as descargas parciais e a refrigeração. Com essa combinação de informação é possível obter uma resposta mais apropriada do sistema e, conseqüentemente, uma maior aceitação pela comunidade. Entretanto, o custo para aplicar e monitorar todos esses dados pode inviabilizar o projeto. Portanto, é preciso encontrar uma solução que não exija tantos recursos, seja de *hardware*, *software* ou até mesmo recursos financeiro.

A seguir é apresentado uma tabela comparativa dos métodos citadas nessa seção. A Tabela 2 mostra os métodos, as vantagens e as desvantagens. Nota-se que para essa comparação foram utilizadas apenas as técnicas de monitoramento *on-line* tendo como objetivo o acompanhamento de toda a vida útil de uma bucha.

Tabela 2 – Tabela de comparação dos métodos de monitoramento.

Trabalho	Método	Desvantagem	Vantagens
AJ et al.	Análise de Assinatura Elétrica	Alto custo	Precisão das medidas
MORALES	Monitoramento de capacitância a partir de uma capacitância C_3	Suscetível a flutuações de tensões	Simple e usa uma fase
MORALES	Correntes balanceadas ou soma das correntes	Desbalanceamento das cargas	Medição mais precisa e melhor compensação das flutuações da rede
MORALES	Referência de sinal independente da assimetria das fases	Dificuldade na eliminação e limitação de ruídos	Medida de um referencial independente de flutuações
XIA et al.	Comparação fase a fase	Custo para a aquisição dos dados	Reduz influência das interferências
ANDRZEJEWSKI; GIL; MIKRO- NIKA	Integração de dados	Alto custo	Precisão do estado do transformador

Fonte: Elaborada pela autora

3.2 Modelagem do Sistema Trifásico

A modelagem de um sistema elétrico pode ser realizada por diversas maneiras. A primeira forma de modelagem já foi citada na fundamentação teórica (Seção 2.1.1). A bucha capacitiva é modelada através de um resistor em paralelo com um capacitor em série com outra combinação de resistor em paralelo com o capacitor. Esse modelo pode ser simplificado: utilizando apenas dois capacitores em série. Nessa configuração não é possível analisar a variação do fator de potência, por isso atualmente não é mais tão interessante essa modelagem. Por outro lado, ao deixar mais complexo o modelo, é possível introduzir alterações de temperatura e ruídos do sistema. Os trabalhos citados nesta seção fazem algum tipo de modelagem para análises elétricas que podem ser alternativas para a forma de implementação do modelo. Essas modelagens são realizadas por meio de diagramas, equações ou a combinação das duas técnicas.

PITHWA realiza a simulação de um transformador utilizando a ferramenta *Simulink* do *Matlab*. O trabalho visa analisar os modelos de proteção diferencial para transformadores e algumas falhas que podem ocorrer nesse sistema. Como o Simulink contém diagramações gráficas por blocos da parte elétrica, as modelagens se tornam mais compactas. Além

disso, é possível analisar sistemas dinâmicos e modelar as operações de acordo com as características desejadas.

GUO et al., que também utiliza o *Simulink*, realizam simulações em um sistema trifásico desbalanceado. O autor estuda sobre bloqueios de fase e o modelo verifica a capacidade do filtro de Kalman de bloquear a fase e rastrear a frequência no sinal de tensão. No trabalho é verificado o uso do filtro de Kalman, entretanto usam um comparador de histerese para que a resposta do sistema seja mais rápida. Nesse contexto, nota-se que o Simulink também é capaz de adicionar funções em seus modelos, ou seja, a ferramenta é utilizada para implementar o comportamento do filtro no sistema. Com isso, é possível compreender tanto o sistema sem a implementação quanto o comportamento do método aplicado.

Já o WANG utiliza variáveis aleatórias para a simulação da tensão trifásica. O artigo combina técnicas numéricas de carga, simulação de Monte Carlo e a proposta de um novo modelo para simular as flutuações causadas por cargas. As flutuações da rede causadas por cargas são representadas por seis variáveis gaussianas correlacionadas. Essa quantidade se deve ao fato da presença de potência ativa e reativa em um sistema elétrico. Inicialmente, o autor simula essas variáveis como variáveis aleatórias com correlação Gaussiana. Além disso, o autor utiliza o fator de desbalanceamento complexo para avaliar o comportamento da rede e obter simulações condizentes com a realidade.

Outra forma de modelar o sistema para WANG é utilizando a tensão de linha. Nesse trabalho, são utilizados as mesmas técnicas que o trabalho anterior, porém o autor modifica a forma de obter o fator de desbalanceamento de forma a facilitar a aplicação. Mesmo usando o fator de desbalanceamento complexo, o método necessita apenas das magnitudes das tensões, simplificando os cálculos e medidas. Dessa forma, os dados simulados são comparados com dados reais para validar o modelo.

A Tabela 3 mostra uma comparação das modelagens e o que foram avaliadas pelo autor utilizando essas modelagens. Dessa forma, é apresentada alternativas de modelar um sistema elétrico. Foram selecionados os trabalhos que incluíram o modelo para aquisição de tensão ou corrente no sistema elétrico. Portanto, não há citações de modelagem estrutural da bucha. Além disso, os trabalhos relacionados dessa seção são algumas possibilidades de aplicações do modelo.

3.3 Tipos de Falhas em Buchas Capacitivas

As falhas em buchas capacitivas podem gerar diferentes tipos de problemas. Como uma falha pode implicar em um prejuízo econômico muito grande, a cada falha na rede elétrica são implantadas novas tecnologias com o intuito de averiguar o desempenho dos equipamentos. Para isso é necessário conhecer os principais tipos de falhas encontradas

Tabela 3 – Tabela de comparação dos métodos de modelagem.

Trabalho	Forma de modelagem	Avaliação
PITHWA	Software de simulação de circuito elétrico	Proteção para transformadores diferenciais
GUO et al.	Software de simulação de circuito elétrico + Equações	Bloqueios de fase
WANG	Equacionamento	Comportamento da rede
WANG	Equacionamento	Comportamento da rede

Fonte: Elaborada pela autora

nas buchas capacitivas. Neste trabalho o foco é dado no envelhecimento das buchas, mas para a avaliação dos dados é preciso considerar dados anômalos e, conseqüentemente, a causa dessas anomalias.

Uma bucha é formada de 4 partes principais: a isolamento, o condutor, a braçadeira de conexão e os acessórios (SEPTYANI; ARIFIANTO; PURNOMOADI, 2011). Segundo a Tabela 10, que mostra a estatística realizada nos anos 2004 a 2009 em PLN P3B Jawa Bali pelos autores SEPTYANI; ARIFIANTO; PURNOMOADI, o superaquecimento é um dos maiores problemas da bucha. Dessa forma, avaliar a temperatura de uma bucha pode ajudar na avaliação das condições de operação desse equipamento.

Tabela 4 – Tabela de falhas em buchas.

Problema	Porcentagem
Superaquecimento/ <i>Hotspot</i>	32%
Rachaduras/Rupturas	26%
Vazamento de óleo	21%
Dano na vedação	11%
Terminal quebrado	5%
Sobrecarga de corrente	5%

Fonte: Adaptado de SEPTYANI; ARIFIANTO; PURNOMOADI, 2011

Além disso, na revisão feita pelos autores AJ et al., é possível analisar ainda mais problemas e medidas preventivas nas buchas. A Tabela 5 mostra algumas das principais falhas que foram resumidas pelos autores AJ et al. a partir das pesquisas realizadas.

Tabela 5 – Tabela de falhas em buchas.

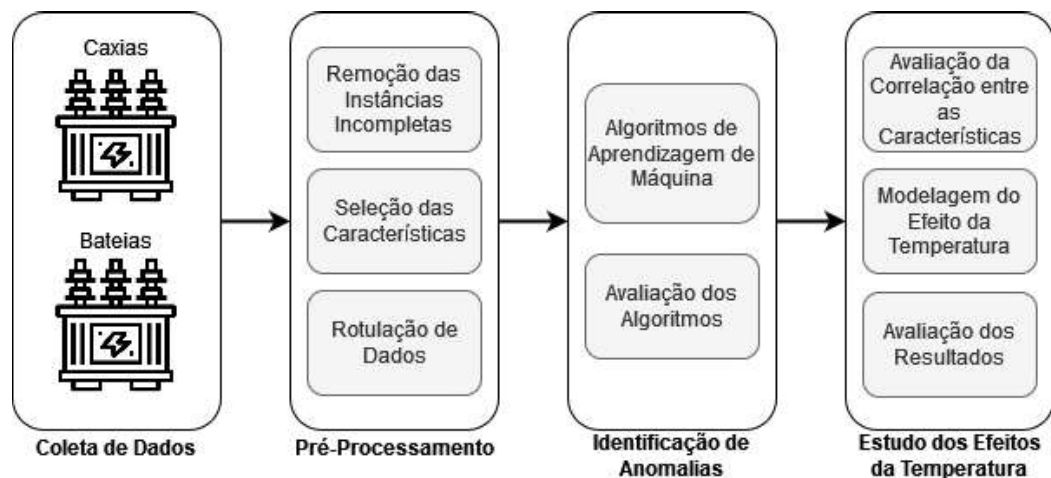
Local do Problema	Impacto
Rachaduras na porcelana	Entrada de umidade, vazamento de óleo e/ou gás, o preenchimento interno vaza
Determinação de encaixes cimentadas	Entrada de umidade, vazamento de óleo e/ou gás, o preenchimento interno vaza
Vazamento na vedação	Entrada de umidade, vazamento de óleo e/ou gás, o preenchimento interno vaza
Isolamento da umidade	Entrada de umidade
Vazamento na solda	Entrada de umidade, vazamento do preenchimento interno
Conexão quebrada entre a bucha aterrada e flange	Faíscas no tanque (aparelho) ou dentro bucha, óleo descolorido
Vazios no composto	Coroa Interna
Diminuição do óleo	Contaminação do preenchimento interno
Sem óleo	Entrada de umidade
Deslocamento do anteparo do gradiente	Fagulhas internas, descoloração do óleo
<i>Flashover</i> elétrico	Rachadura em na porcelana quebrada
Raio	Rachadura em na porcelana quebrada
Descarga de corona	Falha interna, interferência eletromagnética, figuras de Lichtenberg ao longo da superfície do papel ou da superfície interna
Curto-circuito das seções do condensador	Aumento da capacitância, redução da tensão no terminal do capacitor, aumento do estresse interno no isolamento
Escurecimento do óleo	Interferência eletromagnética e resultados de testes ruins

Fonte: Adaptado de AJ et al., 2018

4 Proposta

A proposta deste trabalho foi dividida em quatro etapas principais: a coleta dos dados, o pré-processamento, a identificação de anomalias e o estudo dos efeitos da temperatura. Conforme apresentado na Figura 17, os procedimentos iniciam com a coleta de dados nas subestações. Em seguida é preciso pré-processar os dados, selecionando as características que serão utilizados para a aplicação das técnicas de aprendizagem de máquina. Além disso, nesta etapa também é necessário retirar as instâncias com dados incompletos ou inconsistentes da base e rotular os dados em função das condições de operação das buchas. Após a rotulação, a base estará estruturada para estudos de identificação de anomalias usando os algoritmos de aprendizagem de máquina e avaliação de desempenho destes algoritmos. Por fim, combinando os resultados dos algoritmos de aprendizagem de máquina com as avaliações das correlações das características, são feitos estudos dos efeitos da temperatura nas buchas capacitivas.

Figura 17 – Etapas da Proposta.



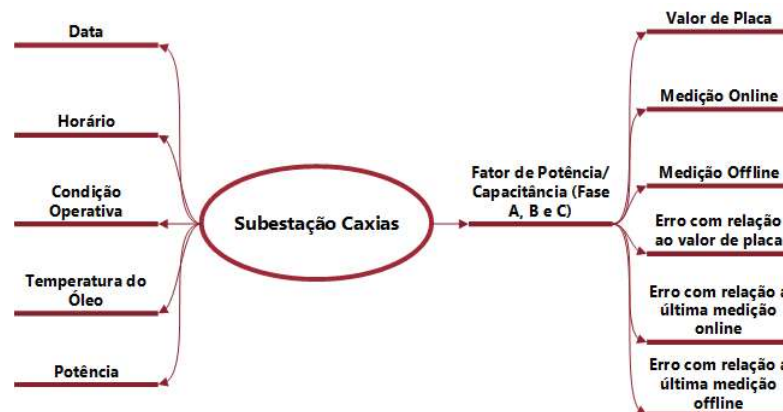
Fonte: Elaborada pela autora

4.1 Coleta de Dados

Inicialmente para este estudo estavam disponíveis apenas os dados da subestação de Salto Caxias que fica localizada no trecho final do Rio Iguaçu, cerca de 400 km de Curitiba. Os dados coletados dessa subestação são referentes aos meses de maio, junho e julho, obtidos com uma frequência de amostragem de aproximadamente uma coleta por dia. A Figura 18 mostra as características obtidas nessa subestação, que são de transformadores

ligados a quatro unidades geradoras U_1 , U_2 , U_3 e U_4 . Esses dados incluem tanto períodos com a bucha operando normalmente, quanto a bucha com problemas de operação. Isso possibilitou uma análise mais completa do sistema.

Figura 18 – Características de Caxias.



Fonte: Elaborada pela autora

A coleta de dados contém valores da *Data*, *Horário*, *Condição Operativa*, *Temperatura do Óleo* do transformador, *Potência*, além da *Capacitância* e do *Fator de Potência* de todas as três fases (A, B e C). Esses dados têm os dados da placa, além de conter também tanto as últimas medições *off-line* quanto as medições realizadas *on-line*. Com esses valores são calculados os erros da medição *on-line* em relação ao valor da placa, à última medição *off-line* e à última medição *on-line*. Sendo assim, antes de avaliar esses dados com algoritmos de aprendizagem de máquina é necessário realizar um pré-processamento.

4.2 Pré-Processamento

As etapas descritas nesta seção têm como objetivo preparar os dados a serem estudados. Os dados obtidos inicialmente são da subestação de Caxias com buchas em diferentes estágios da vida. Considerando o cenário atual das buchas capacitivas foram selecionadas as características mais comuns para analisar o estado da bucha. Em seguida, como a base de dados contém valores faltantes em algumas instâncias, algumas instâncias foram retiradas. Esse processamento auxilia na obtenção de resultados mais coerentes, visto que uma instância com características faltantes pode alterar o resultado da classificação. Por último, para utilizar os algoritmos de aprendizagem de máquina supervisionados é preciso que os dados sejam rotulados. Nessa etapa, a rotulação ocorreu com auxílio de alguns engenheiros da subestação e de uma pré-rotulação feita durante a aquisição dos dados.

4.2.1 Remoção das Instâncias Incompletas

Os dados obtidos na subestação de Caxias contém dados sem alguns valores e dados obtidos durante o desligamento para a troca de equipamento. Portanto, é necessário retirar instâncias incompletas ou dados em que a unidade estava fora de operação para que não houvesse inconsistências nos resultados da detecção de anomalias através da aprendizagem de máquina. Já para a análise da correlação, é preciso tirar dados atípicos, para que o resultado seja o mais adequado para a situação.

4.2.2 Seleção de Características

Para a subestação de Caxias foram obtidas várias informações referentes a cada uma das unidades em operação. Nesse contexto, foi considerado que a concessionária segue os intervalos sugeridos na norma IEEE C57.19.100-1995, ou seja, são realizados ensaios *off-line* a cada 3 a 6 anos para medir os valores de capacitância e fator de potência. Portanto, os dados *off-line* e os dados dos valores de placa podem ser bem discrepantes (IEEE..., 1995).

Para essa etapa, é possível utilizar a ferramenta do *software* WEKA que seleciona características. De acordo com a literatura, as principais formas de avaliação da bucha utilizam a capacitância e o fator de potência. Portanto, utilizando os avaliadores disponíveis no WEKA será possível verificar se os valores de capacitância e fator de potência são as melhores opções para decidir o estado do equipamento, visto que essas métricas estão diretamente relacionadas com a deterioração das buchas capacitivas.

4.2.3 Rotulação dos Dados

As medições da subestação de Caxias foram obtidas das três fases de quatro unidades em operação. Esses dados já estavam pré-rotulados pelos operadores da subestação. Sendo assim, o processo de rotulação foi mais fácil de ser concluído. Os dados da subestação de Caxias foram rotulados em: *bom estado*, *alerta* e *substituição*.

Essas classes foram definidas a partir dos limiares de decisão utilizados em concessionárias (RAMBO et al., 2007) e relatórios da subestação. Neste cenário, uma das métricas utilizadas para traçar um limiar é o fator de dissipação. Esse limiar é aplicado de acordo com a variação do fator de dissipação, que é definido em quatro classes, conforme apresentado na Tabela 6. Entretanto, para a classificação foram utilizadas apenas três classes. A primeira classe é a bucha em *Bom Estado*, que engloba os equipamentos novos e velhos. A segunda classe é a de *Alerta*, que equivale a classe *Investigação* da tabela. Por último, a classe *Substituição* que manteve a mesma nomenclatura. Nota-se que entre 0,3% e 0,4% a bucha está na faixa de transição de bucha nova para uma bucha antiga, portanto não há uma separação exata. Por esse motivo, essa parte inicial foi classificada

como apenas uma classe. Já o estado de alerta e substituição são estágios da vida da bucha que requerem mais atenção e foram mantidas estas separações.

Tabela 6 – Tabela de limiares de decisão usando o fator de dissipação.

Buchas em Bom Estado		Alerta	Substituição
Equipamentos Novos	Equipamentos Antigos	Investigação	Substituição
0,1 a 0,4%	0,3 a 0,6%	0,6 a 1%	>1%

Fonte: Elaborada pela autora.

4.3 Identificação e Detecção de Anomalias

Para a etapa de identificação e detecção de anomalias foram utilizadas informações coletadas em alguns trabalhos relacionados, técnicas de aprendizagem de máquinas e a avaliação realizada pelos operadores da subestação de Caxias. Dado o contexto do trabalho, as classificações utilizadas foram: a bucha em bom estado, alerta e retirar a bucha de operação. As buchas novas ou em bom estado é a classe com maior número de instâncias adquiridas. Esta classe é o estado de referência de uma bucha. Como cada bucha pode ter um valor inicial de capacitância diferente, o estudo é feito a partir da variação dos valores de capacitância e fator de potência. Já a classe alerta, é o estado da bucha que necessita de mais atenção. Como já foi citado, a bucha capacitiva tem dois valores de expectativa de vida: em torno dos 10 a 13 anos e em torno 20 a 30 anos (QUALITROL, 2017). Portanto, quando a bucha entra em estado de alerta, as concessionárias precisam estar atentas para realizar manutenções. Por fim, quando uma bucha chega ao último estágio, é quando não é mais recomendada a sua operação, ou seja, é necessário que a bucha seja trocada. Mesmo chegando a esse estágio, a bucha ainda se mantém em operação por um período. Isso ocorre por causa do tempo para fazer o pedido da bucha nova para a fabricante, somado com o tempo da entrega, do agendamento de desligamento do transformador e da troca da bucha.

4.3.1 Algoritmos de Aprendizagem de Máquina

Para este trabalho são utilizados os valores de capacitância e fator de potência para a análise do estado de uma bucha de cada fase. O estudo realizado não engloba as dependências entre fases para a avaliação do estado da bucha. Portanto, com os dados rotulados são aplicados os algoritmos de aprendizagem de máquina, ou seja, são realizados treinamentos dos modelos de aprendizagem de máquina (Seção 2.2) para a detecção de anomalias (estado alerta e retirar de operação) utilizando *software* livre WEKA (*Waikato Environment for Knowledge Analysis*) (WITTEN et al., 2005). Essa etapa pode ser feita também utilizando o *Scikit Learn* (PEDREGOSA et al., 2011). Entretanto, como a base é

pequena e o objetivo principal é realizar testes com algoritmos conhecidos de forma mais rápida e sem a necessidade de integração no momento, foi decidido utilizar o *software* WEKA.

Neste trabalho, o uso da aprendizagem de máquina tem como objetivo auxiliar no estudo das buchas e testar vários algoritmos para identificar a degradação das buchas. Devido a alteração do estado da bucha ser suave e sofrer influências da temperatura e dos ruídos da rede, é desejável compreender como a aprendizagem de máquina avalia os dados com essas interferências. Além disso, os modelos são testados para que fosse possível analisar qual é o comportamento do algoritmo para detectar as anomalias quando há informação de potência e temperatura do óleo e o quanto essas características afetam no resultado da análise. Por último, são extraídos dados estatísticos para auxiliar posteriormente na avaliação do modelo de bucha. Para isso, são obtidos a correlação entre as características utilizadas para o treinamento e os dados estatísticos (média, mediana, desvio padrão, entre outros).

4.3.2 Avaliação dos Algoritmos

A análise de dados tem como objetivo observar os principais comportamentos visando a modelagem mais completa de uma bucha capacitiva. Para isso, na análise é verificado qual é a correlação entre as características para saber a influência que um parâmetro têm em outro. Além disso, o uso de aprendizagem de máquina pode gerar informações diferentes, pois cada algoritmo de aprendizagem de máquina tem sua peculiaridade, que pode demonstrar alguma particularidade dos dados obtidos nas subestações. Neste trabalho, algumas métricas padrões foram utilizadas para avaliar a eficácia dos algoritmos como: os valores de acurácia, a matriz de confusão e o *f1-score*. O valor da acurácia é calculado para saber os acertos totais, a matriz de confusão é gerado para avaliar e visualizar melhor onde estão ocorrendo os erros e o *f1-score* para verificar a performance dos algoritmos. Essas métricas, são importantes para entender o funcionamento dos algoritmos ao classificar os dados, mas também trazem informações de quão difícil é classificar os dados obtidos em subestações.

4.4 Estudo dos Efeitos da Temperatura

O estudo do efeito da temperatura é realizado para gerar um modelo mais completo das buchas capacitivas. Nesta pesquisa, tanto a análise após a detecção do estado da bucha utilizando a aprendizagem de máquina quanto a análise após a geração do modelo de bucha tem como objetivo compreender melhor o funcionamento da bucha. A análise do modelo verifica a influência da temperatura do óleo no comportamento da bucha e será utilizado para comparar com os dados obtidos na subestação. Neste contexto, um modelo

mais completo pode ajudar em futuros estudos das concessionárias e melhorar a avaliação da bucha para que as decisões tomadas pelos operadores sejam cada vez mais precisas.

Para iniciar os estudos dos efeitos da bucha é feito a avaliação das correlações entre as características. Essa correlação permite verificar se há alguma dependência entre as características obtidas das buchas capacitivas. Sendo assim, a avaliação da correlação será importante para verificar se há algum padrão de acordo com o estado de degradação, com a fase ou com o transformador. Dessa forma, é possível estudar outros comportamentos das buchas capacitivas para que seja considerado na avaliação dos resultados finais.

No contexto das buchas capacitivas, ainda é difícil obter dados de estados de buchas mais degradadas. Isso ocorre devido ao monitoramento *on-line* ser relativamente recente e, por isso, os sensores são normalmente adicionados após a troca das buchas. Por esse motivo, a maioria dos dados disponíveis são de buchas em bom estado. Além disso, os sistemas de cada subestação podem ser diferentes, pois cada concessionária utiliza escalas e parâmetros diferentes. É necessário ressaltar que pode haver falhas nas aquisições de dados, dificuldade em armazenar e manter as coletas de dados atualizadas problemas relacionados com o sigilo de informações. Dessa forma, a modelagem de bucha auxilia na simulação de seu comportamento para que os seus estudos não dependam tanto dos dados de subestações. Neste trabalho, a principal dificuldade na análise com aprendizagem de máquina é que a base é muito desbalanceada, ou seja, a maioria dos dados obtidos são referentes a apenas uma classe: buchas em bom estado. Portanto, pensando em um modelo mais completo para compreender melhor as buchas capacitivas, o modelo proposto da bucha adiciona a variação com a temperatura para gerar esses dados sintéticos.

Conforme apresentado na Seção 2.1.3, o modelo da bucha pode ser representado por um resistor em paralelo com uma capacitância. A modelagem visa estudar o efeito da variação da temperatura nas medidas de capacitância e do fator de potência. A principal informação utilizada inicialmente é a correlação entre os parâmetros. Dessa forma, como o objetivo final é estruturar um modelo que seja mais próximo do real, o modelo acrescenta a variação da temperatura. Para isso é utilizado a Tabela 7 para adicionar o efeito da temperatura do óleo do transformador na modelagem. Além disso, é possível acrescentar modelos probabilísticos para simular os ruídos da rede ou até mesmo um desbalanceamento de carga entre as fases, como demonstrado no estudo realizado em (WANG, 2001). Entretanto, como nesse trabalho o foco é na análise por fase, não será estudado o efeito do desbalanceamento de carga. Para isso a Equação 4.1 é utilizada para fazer a correção dos valores medidos de forma *on-line*.

$$\tan(\delta_{20}) = \frac{\tan(\delta_m)}{K}. \quad (4.1)$$

- $\tan(\delta_{20})$: Ângulo de Dissipação em 20°C;

- $\tan(\delta_m)$: Ângulo de Dissipação na Temperatura Medida;
- K : Fator de Correção.

Tabela 7 – Tabela de fator de correção (K) de temperatura do óleo.

Intervalos de Temperatura em °C	Fator de Correção	
	OIP	RIP
0-2	0.8	0.76
3-7	0.85	0.81
8-12	0.9	0.87
13-17	0.95	0.93
18-22	1	1
23-27	1.05	1.07
28-32	1.1	1.14
33-37	1.15	1.21
38-42	1.2	1.27
43-47	1.25	1.33
48-52	1.3	1.37
53-57	1.34	1.41
58-62	1.35	1.43
63-67	1.35	1.43
68-72	1.3	1.42
73-77	1.25	1.39
78-82	1.2	1.35
83-87	1.1	1.29

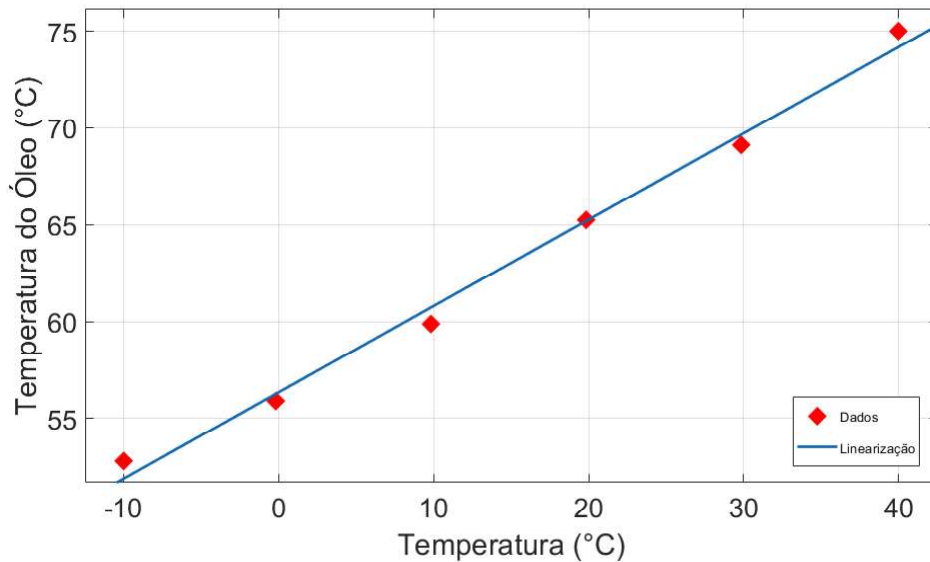
Fonte: Adaptado de MURTY, 2017b

A Tabela 7 mostra o fator de correção (K) para determinadas temperaturas de óleo tendo 20° como temperatura de referência. Essa tabela é utilizada como forma de análise do fator de potência em buchas capacitivas novas pelas fabricantes. Portanto, para a modelagem é possível realizar o cálculo inverso, ou seja, utilizar valores da temperatura do óleo para gerar a variação da simulação. Entretanto, os dados de temperaturas disponíveis são de temperatura ambiente, sendo assim é necessário realizar outra conversão antes de aplicar a variação de temperatura no modelo.

Segundo CHEN et al., é possível assumir que a temperatura da bucha é equivalente a maior temperatura do óleo do transformador, também denominado de *Hotspot*. Por esse motivo, a modelagem utiliza a conversão apresentada pela Figura 19 e a sua linearização é representada pela Equação 4.2. Essa linearização foi utilizada para simplificar a conversão e por gerar uma boa aproximação para estudos de modelagem. A Figura 19 mostra os pontos

adquiridos pelos autores CHEN et al., entretanto, no artigo é utilizada uma aproximação polinomial.

Figura 19 – Gráfico de conversão da temperatura ambiente em temperatura do óleo.



Fonte: Adaptado de CHEN et al., 2012

$$T_M = 0,4461 \cdot T_a + 56,34 \quad (4.2)$$

- T_M : Temperatura do Óleo;
- T_a : Temperatura Ambiente.

O modelo de bucha capacitiva utiliza como base o modelo elétrico ideal apresentado na Subseção 2.1.3. Portanto, a criação do novo modelo utiliza o modelo capacitivo e também a adição da resistência em série para representar as perdas dielétricas. O modelo utiliza duas capacitâncias C_1 e C_2 , que são a capacitância principal e a capacitância do *tape*, respectivamente. A partir dessas capacitâncias, foi adicionada em paralelo a cada uma das capacitâncias, as resistências R_1 e R_2 . Sendo R_1 em paralelo com a C_1 e o R_2 em paralelo com a C_2 .

Para gerar o modelo combina-se os limiares de decisão do fator de potência utilizados por concessionárias com a modelagem ideal da bucha. A Equação 4.3 mostra a equação utilizada para determinar o valor do R_1 variando com a temperatura do óleo. Na equação PF é o fator de potência, que será alterado de acordo com a porcentagem da variação do

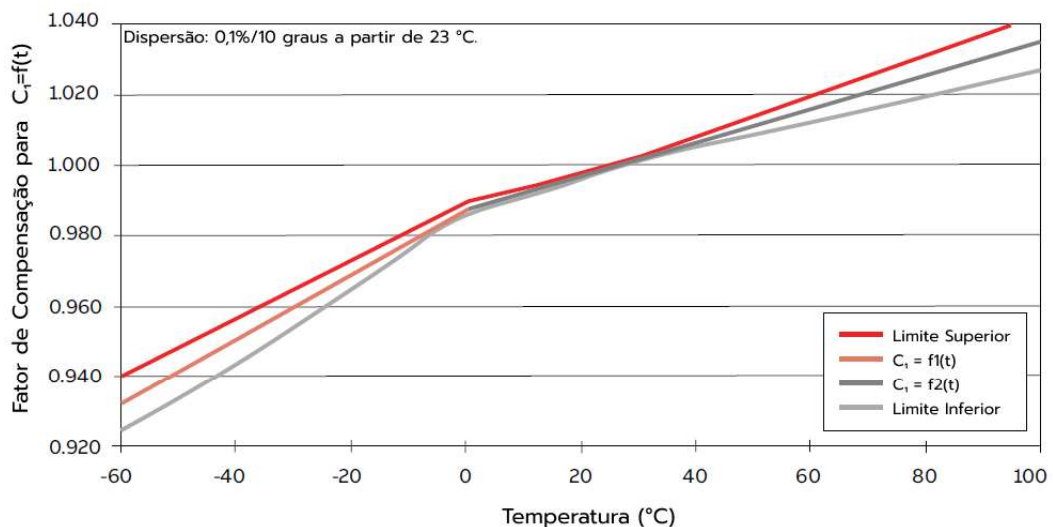
fator de dissipação. O f é a frequência da rede, o K é o fator de conversão da temperatura e o C_1 a capacitância principal da bucha.

$$R_1 = \frac{\tan(\arccos(FP_{var} \cdot FP_{Placa} \cdot K))}{2 \pi f C_{var} C_{1,Placa}} \quad (4.3)$$

- R_1 : Resistência R_1 ;
- FP_{var} : Variação do Fator de Potência;
- FP_{Placa} : Fator de Potência da Placa;
- K : Fator de Correção;
- f : Frequência da Rede;
- $C_{1,var}$: Variação da Capacitância Principal;
- $C_{1,Placa}$: Capacitância Principal da Placa.

Além da variação da temperatura no fator de potência é adicionado o efeito da temperatura na capacitância por um fator K_C . Essa variação é gerada utilizando o gráfico apresentado na Figura 20. Os limites superiores e inferiores não são utilizados neste trabalho, mas as fabricantes relatam que há uma dispersão de 0,1% a cada 10°C a partir de 23°C. Sendo assim, para facilitar o modelo são utilizadas as Equação 4.4 e Equação 4.5 para variar a temperatura na capacitância.

Figura 20 – Gráfico de correção do efeito da temperatura ambiente na capacitância.



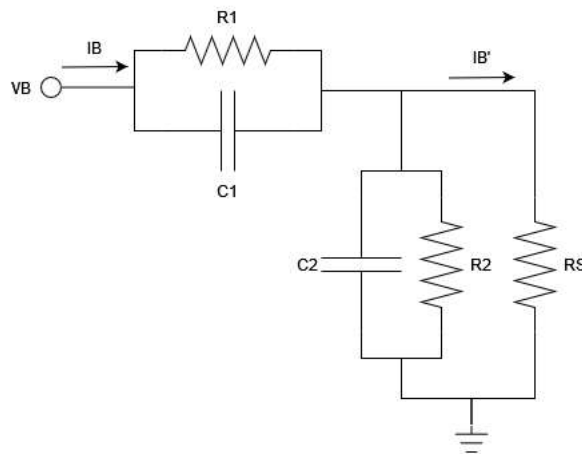
Fonte: Adaptado de ABB, 2021

$$K_C = 0.00048 Ta + 0.9 \quad (4.4)$$

$$K_C = 0,00092 Ta + 0,9 \quad (4.5)$$

- K_C : Fator de Correção do efeito da temperatura na Capacitância;
- Ta : Temperatura Ambiente.

Figura 21 – Diagrama do modelo elétrico.



Fonte: Elaborada pela autora

As buchas em operação sempre estão com os tapes capacitivos aterrados, como mostrado na Figura 21, por isso não há circulação de corrente por C_2 . Portanto, neste modelo, os valores de R_2 e C_2 são mantidos inalterados ao longo das simulações, pois não há uma degradação significativa. Por isso a impedância Z_2 ($R_2 || C_2$) é mantida com um valor alto. Já a capacitância C_1 é definida de acordo com os valores da última medição *off-line* dos dados de Caxias, que foi o valor de referência utilizado para gerar as variações. Além disso, um dos exemplos de limiar usado para decisão em concessionárias, que é mostrada na Tabela 8, pode ser uma das formas de gerar a variação do fator de potência e da capacitância para simular o envelhecimento da bucha. Os valores C_{Placa} e FP_{Placa} representam os valores iniciais de capacitância e fator de potência respectivamente, ou seja, as primeiras medições realizadas quando a bucha é fabricada. Portanto, para gerar as variações nos dados sintéticos, os valores apresentados na Equação 4.3 são variados de acordo com os limites e com o valor inicial da capacitância da bucha e do fator de potência. Por fim, é validado a influência da temperatura, utilizando os dados da subestação de Caxias.

Tabela 8 – Tabela de limiares de decisão usando o fator de potência e capacitância.

Parâmetro	Bom Estado	Alerta	Substituição
Capacitância	1 a 1,05 C_{Placa}	1,05 a 1,1 C_{Placa}	>1,1 C_{Placa}
Fator de Potência	1 a 2 FP_{Placa}	2 a 3 FP_{Placa}	>3 FP_{Placa}

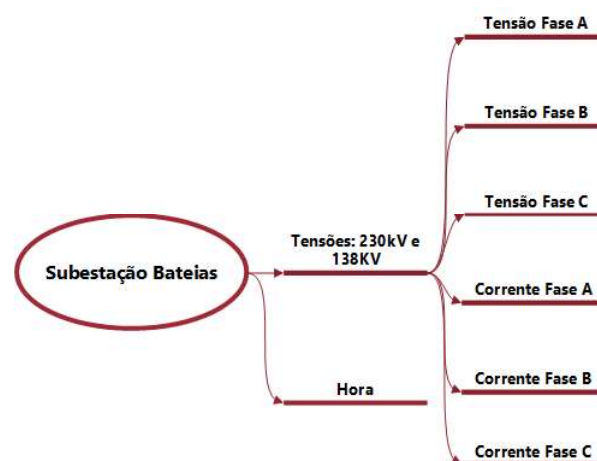
Fonte: Adaptado de RAMBO et al., 2007

Para finalizar, é feito a avaliação dos resultados. Como os dados gerados podem utilizar os valores das medidas *off-line* da subestação de Caxias e os valores da temperatura do óleo, pode-se usar como referência para avaliação as medições *on-line* para comparar com os dados gerados. Além disso, é possível analisar a diferença entre a variação dos dados sintéticos e dos dados reais.

4.5 Dados Adicionais

Atualmente, o sistema de coleta de Bateias já está operacional e com os novos dados que estão sendo coletados pretende-se ampliar os estudos de identificação de anomalia. A subestação Bateias fica localizada em Campo Largo (PR). Os dados são obtidos nesta subestação com uma taxa média de 21 amostras por dia e os dados coletados são referentes a julho, agosto e setembro.

Figura 22 – Características de Bateias.



Fonte: Elaborada pela autora

Para a etapa do processamento dos dados é preciso ressaltar algumas diferenças entre as informações adquiridas em cada subestação. A Figura 22 apresenta as medidas obtidas na subestação de Bateias. Observa-se que os dados de Caxias já tem os valores de capacitância e fator de potência de cada fase, enquanto em Bateias foram adquiridos os

valores de tensão e corrente de cada fase. Portanto, para os dados de Bateias é preciso calcular os valores de capacitância e fator de potência a partir dos valores de tensão e corrente medidos na subestação. Além disso, os dados de Caxias apresentam os valores iniciais da capacitância e do fator de potência. Como a estrutura de dados obtida em Bateias é a mais simples optou-se que a pesquisa fosse mantida utilizando a mesma quantidade de informações disponível dessa subestação. As medições dessa subestação são referentes à dois transformadores e foram realizadas nas três fases de cada um dos lados do transformador.

5 Resultados

Neste capítulo são apresentadas as análises e resultados das etapas abordadas na proposta. Inicialmente, são apresentados os dados obtidos nas subestações. Em seguida, é realizada a primeira avaliação utilizando as técnicas de aprendizagem de máquina. A partir das informações coletadas, é feito o estudo do efeito da temperatura nos valores da capacitância e fator de potência das buchas capacitivas. Por fim, os resultados são avaliados, verificando-se as similaridades e diferenças entre as informações.

5.1 Coleta de Dados

Neste trabalho foram utilizadas duas bases de dados distintas. A primeira é referente a subestação de Caxias, que possui as três classes (estados) de operação das buchas capacitivas, conforme discutido no capítulo anterior. A segunda base é referente a subestação de Bateias, com dados de apenas um estado de operação, ou seja, a bucha em bom estado. Ambas são bases de dados bem desbalanceadas, mas que ilustram bem o contexto atual das informações disponíveis das buchas de transformadores.

Os dados de cada uma das subestações tem tratamentos diferentes. Nesse contexto, os estudos foram iniciados com a base de Caxias, visto que é possível visualizar dados de cada estágio da vida da bucha, portanto, uma análise mais completa. Já com relação a base de Bateias, é possível verificar se há alguma semelhança entre as medições de buchas em boas condições e comparar os resultados com as buchas em bom estado da subestação de Caxias.

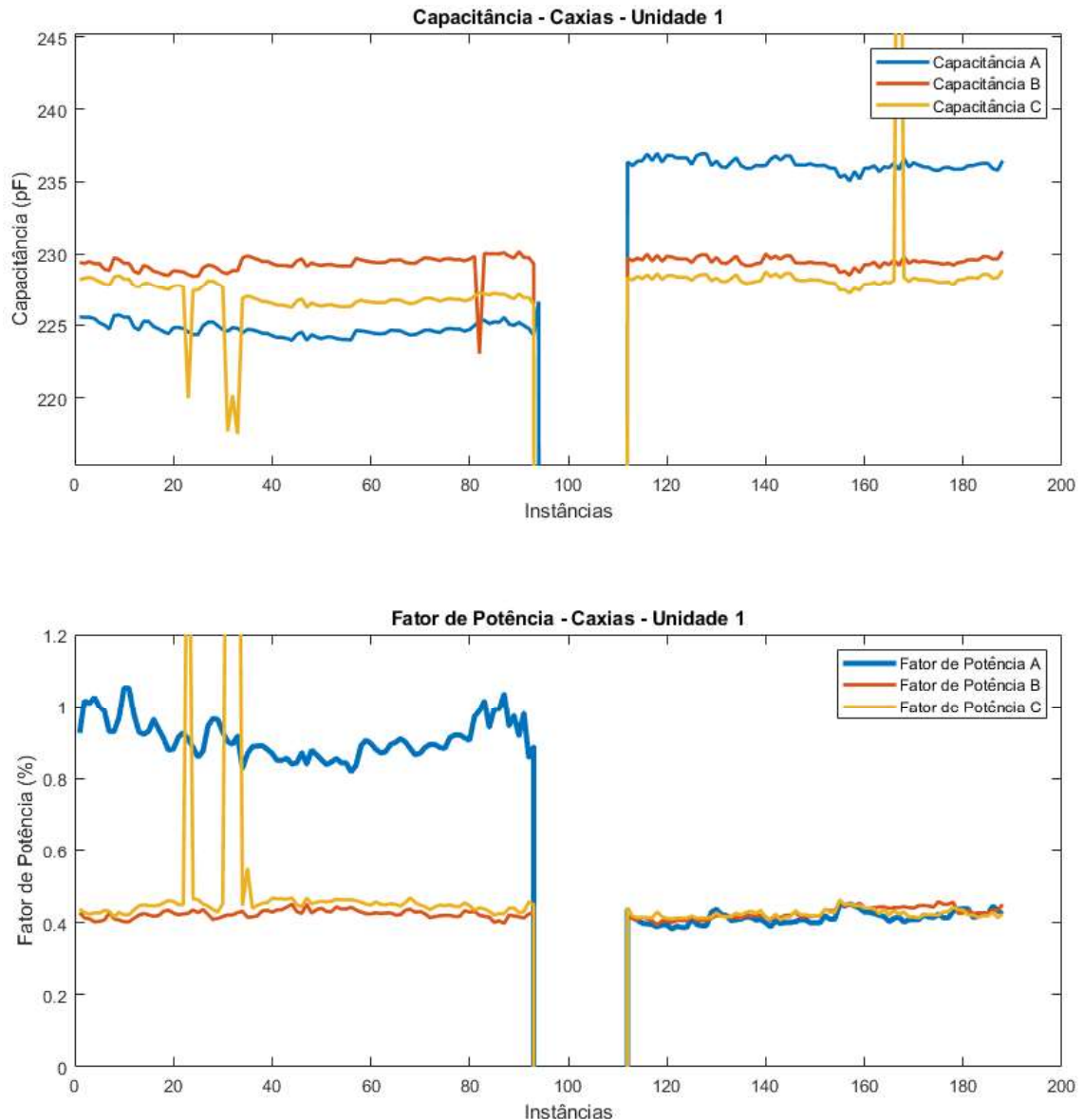
5.1.1 Base de Dados de Caxias

Os dados obtidos na subestação de Caxias foram coletados de transformadores conectados a quatro unidades geradoras. Essas unidades geradoras operam nas seguintes condições: compensador, gerador ou estão fora de operação. As instâncias que foram obtidas com as unidades fora de operação foram retiradas posteriormente, pois são dados atípicos ou inconsistentes, e que não devem ser analisados. A série temporal obtida nessa subestação equivalente a um conjunto de 2283 instâncias, que representam características de cada fase e equivalem a 3 meses de coleta. A média é de uma coleta/medida por dia. Portanto, são 200 instâncias caso a análise fosse feita por transformador, ou seja, com as características das três fases.

As Figuras 23, 24, 25 e 26 apresentam os gráficos das séries temporais de capacitância e fator de potência dos transformadores das unidades geradoras U_1 , U_2 , U_3 e U_4 ,

respectivamente. A quantidade de instâncias apresentadas nas figuras equivale aos dados de cada transformador, portanto cada figura contém em torno de 200 instâncias cada uma, visto que as imagens estão mostrando cada uma das fases de cada unidade.

Figura 23 – Dados do transformador da Unidade 1 de Caxias.



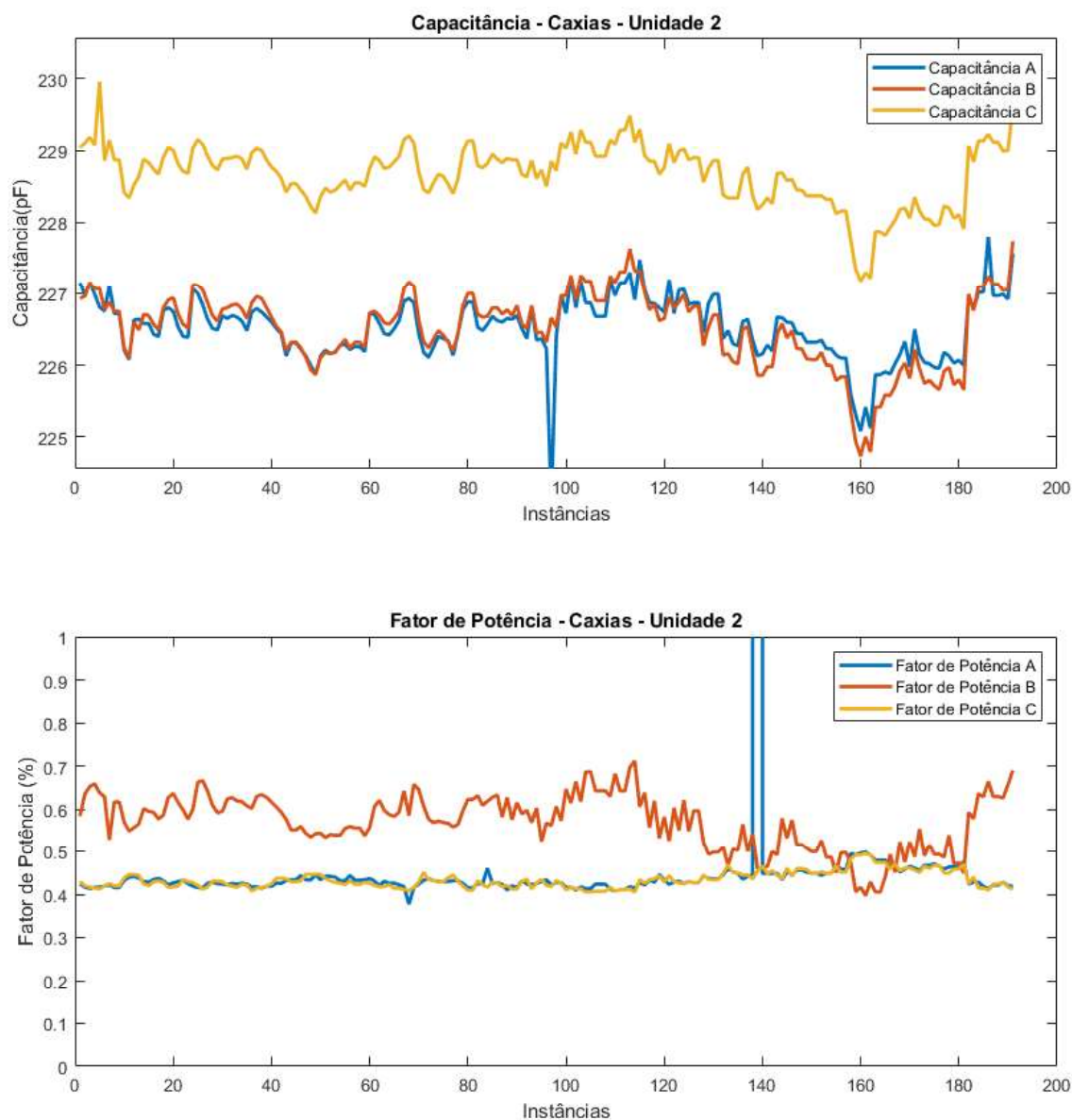
Fonte: Elaborada pela autora

Na Figura 23 (Unidade 1) é visível um intervalo com medições irregulares, esse intervalo é referente ao período em que a unidade ficou desativada devido a troca da bucha da fase A que estava em estado crítico de operação. Também é possível visualizar a alteração do comportamento dos valores de capacitância e fator de potência da bucha da fase A após o intervalo de troca. Os valores da capacitância aumentam, pois é um novo equipamento e cada bucha tem sua capacitância inicial diferente. Além disso, os valores

do fator de potência diminuem, que implicam no estado de operação da bucha.

Os intervalos da Figura 23 (Unidade 1) são medições com a unidade fora de operação, que retirados para realizar as análises. Outros pontos fora do padrão como alguns valores da fase C são tratados como anomalias locais, causados por descargas atmosféricas, manobras, algum problema na aquisição ou na própria rede.

Figura 24 – Dados do transformador da Unidade 2 de Caxias.

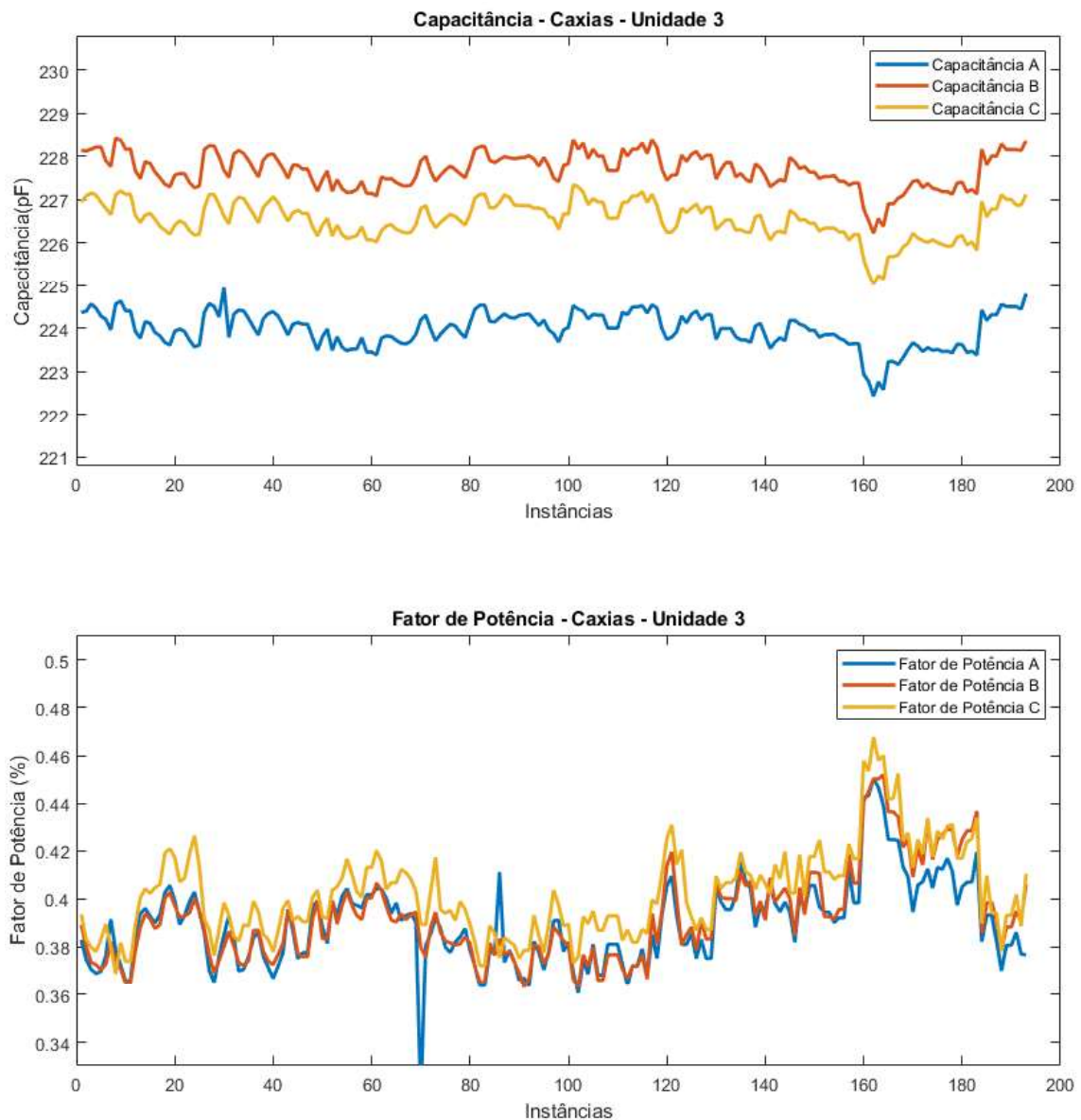


Fonte: Elaborada pela autora

Nas Figuras 24 e 25, referentes aos transformadores das Unidades 2 e 3, é possível visualizar um comportamento mais estável nas medições adquiridas, ou seja, não houve praticamente nenhuma perda de informação ao longo do período de coleta. Entretanto, na Unidade 4, assim como na Unidade 1, há um intervalo com aquisições que estão fora do

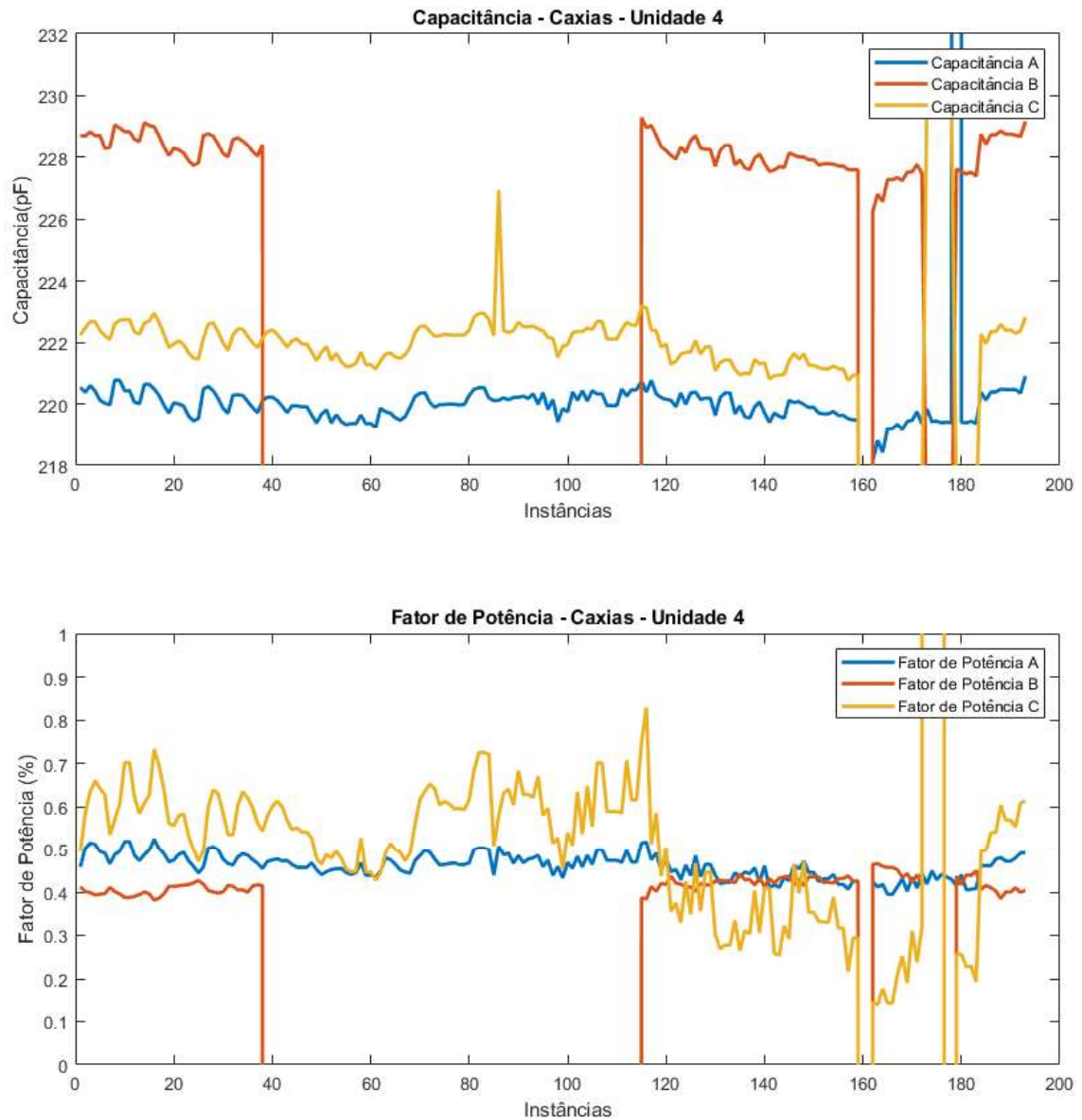
padrão. Porém, na Unidade 4 apenas a fase B está sem as medições por alguns instantes. Neste contexto, esse intervalo também é retirado na etapa de processamento dos dados mesmo que a falha na fase B não seja de origem de desligamento do sistema, pois as informações contidas nesse trecho não são relevantes para a análise. Enfatizando, são retirados apenas dados da fase B, mantendo as séries temporais da fase A e fase C como ilustrado na Figura 26.

Figura 25 – Dados do transformador da Unidade 3 de Caxias.



Fonte: Elaborada pela autora

Figura 26 – Dados do transformador da Unidade 4 de Caxias.



Fonte: Elaborada pela autora

5.2 Etapa de Pré-Processamento

O pré-processamento consiste em selecionar as características, retirar as instâncias incompletas e rotular os dados. Como a base de bateias é a mais simples, não foi necessário realizar nenhum pré-processamento. Sendo assim, essa seção é referente apenas à base de dados de Caxias.

5.2.1 Remoção das Instâncias Incompletas

A base de Caxias contém instâncias com dados faltando devido a problemas no momento da aquisição, como na fase B da Unidade 4. Além disso, houve também dados obtidos durante a troca de equipamentos, como é o caso da Unidade 1. Como esses dados não agregam para a avaliação do estado da bucha foram retirados. Outro detalhe importante é que algumas instâncias estavam sem a informação da temperatura. Estas instâncias também foram removidas para as análises de correlação dos valores da capacitância e do fator de potência com a temperatura. Portanto, após retirar os dados que não agregariam na avaliação do estado da bucha, a base de dados ficou com 2.058 instâncias no total.

5.2.2 Seleção de Características

Tabela 9 – Características selecionadas.

Atributos	Manual	Algoritmo de Seleção
Data		
Horário		
Valor de placa (C)		X
Medição offline (C)		X
Medição online (C)		
Erro com relação ao valor de placa (C)		
Erro com relação a última medição offline (C)	X	
Erro com relação a última medição online (C)		
Valor de placa (FP)		
Medição offline (FP)		X
Medição online (FP)		
Erro com relação ao valor de placa (FP)		
Erro com relação a última medição offline (FP)	X	
Erro com relação a última medição online (FP)		
Potência (Tr)	X	
Temperatura do Óleo (Tr)	X	X

C - Capacitância FP - Fator de Potência Tr - Transformador

Fonte: Elaborada pela autora.

Para a etapa de seleção de características foram feitas duas seleções. A primeira seleção de dados foi utilizado as características mais comuns na avaliação de buchas (capacitância e fator de potência em relação ao valor inicial, ou seja, foi utilizado o erro da medição *on-line* em relação o valor da placa), pois tem a medida *on-line* com referência ao seu valor inicial. A segunda forma, com a finalidade de obter um conjunto das

características que fossem mais relevantes para detectar anomalias com a aprendizagem de máquina, foi utilizado o algoritmo *Correlation Attribute Eval* do *software* WEKA, que analisa a correlação das características.

O resumo das características selecionadas são apresentadas na Tabela 9. Nessa tabela as características da coluna Manual são as características normalmente utilizadas no cenário atual do monitoramento das buchas capacitivas e foram manualmente selecionadas. Já as características da coluna do algoritmo de seleção foram selecionadas as quatro primeiras características mais relevantes obtidas a partir do avaliador de características do WEKA. Com essa seleção é possível avaliar se as características normalmente usadas são as melhores para tomar uma decisão sobre o estado da bucha.

5.2.3 Rotulação de Dados

Para esta etapa, os dados utilizados já haviam sido pré-rotulados na subestação. Nesta base, o transformador da Unidade 1 é o que contém a bucha com defeito. A quantidade que representa esse estado é 81 instâncias. As instâncias que representam buchas com necessidade de observações mais periódicas ou em um estado de alerta são 350, que são da fase B da Unidade 2 e da fase C da Unidade 4. O restante das instâncias pertencem a classe de buchas em bons estados ou novas, o que equivale a 1.545 instâncias. A Tabela 10 mostra a quantidade de instâncias em cada uma das classes.

Tabela 10 – Instâncias de Caxias.

Classe	Instâncias	Porcentagem
Bom Estado	1545	78%
Alerta	350	18%
Substituição	81	4%
Total	1976	100%

Fonte: Elaborada pela autora

5.3 Identificação de Anomalias

Para a detecção de anomalias foram utilizados alguns algoritmos de aprendizagem de máquina, como a árvore de decisão, KNN, *Naive Bayes* e *pools* de classificadores. Nesse trabalho foi utilizado o *software* livre WEKA (WITTEN et al., 2005) para realizar essa tarefa. Dentro do *software* é possível alterar alguns parâmetros, entretanto em sua maioria foi mantido os parâmetros padrões e caso haja alguma alteração será comentado durante a apresentação dos resultados. Para facilitar a análise, os resultados da base com as características selecionadas manualmente é definida como sendo a base A e os dados

selecionados com o algoritmo de seleção de características é a base B. Por fim, o protocolo utilizado foi a validação cruzada com 10 subconjuntos.

Inicialmente, foi utilizado a árvore de decisão que pelo *software* a implementação equivalente do C4.5 é a J48 (TAN; STEINBACH, 2004). As Tabelas 11a e 11b mostram as matrizes de confusão geradas ao aplicar o algoritmo nos dados da subestação de Caxias. Pela matriz é possível visualizar a diferença de acerto de acordo com as classes. A árvore obteve uma taxa de acerto de 89% (base A) e 99% (base B). Nesse algoritmo, a sua precisão foi de 88%, o valor de *recall* foi 89% e a *F-measure* foi 88% utilizando os dados selecionados manualmente e a precisão, *recall* e a *F-measure* foram de 99% utilizando a base com a seleção realizada pelo algoritmo. Nota-se que o resultado já é quase perfeito utilizando as características selecionadas pelo algoritmo. Nesse contexto, isso pode ocorrer por causa do algoritmo utilizado para selecionar as características usar uma árvore de decisão como base.

Tabela 11 – Matrizes de confusão da árvore de decisão.

(a) Base A				(b) Base B			
a	b	c	<-Classificado como	a	b	c	<-Classificado como
27	44	10	a = Substituição	80	1	0	a = Substituição
15	1491	39	b = Bom Estado	0	1537	8	b = Bom Estado
7	106	237	c = Alerta	1	1	348	c = Alerta

Fonte: Elaborada pela autora.

O algoritmo testado em seguida foi o KNN e o seu equivalente no *software* WEKA é o IBK. Os resultados das matrizes de confusão deste algoritmo estão mostrados nas Tabelas 12a e 12b. Como a base tem uma seleção realizada a partir de modelos de árvores é esperado que o algoritmos que usem estratégias de classificação mais próximas do algoritmo da árvore de decisão tenham resultados melhores. Portanto, nota-se que os resultados demonstrados nas matrizes estão piores que as matrizes da árvore de decisão. A sua taxa de acerto é em torno de 62% para a base A e em torno de 92% para a fase B. Utilizando o KNN obteve-se resultados parecidos em suas métricas. Portanto, a média de suas medidas de precisão, valor de *recall* e de *F-measure* são de aproximadamente 62% para a base A e 92% para a base B.

Outro algoritmo testado foi o *Naive Bayes*. As suas matrizes de confusão estão mostradas nas Tabelas 13a e 13b. As instâncias rotuladas como buchas em bom estado são confundidas em sua maioria como buchas em estado de substituição em ambas as bases. Esse algoritmo classifica bem as buchas que necessitam ser trocadas nas duas bases. Considerando a base de dados desbalanceada, é visível a diferença da taxa de acerto, visto que a taxa de acerto do *Naive Bayes* foi de aproximadamente 7% (base A) e 23% (base B). Os valores da sua precisão foi de 30%, do valor de *recall* foi 7% e a *F-measure* foi 5% para

Tabela 12 – Matrizes de confusão do KNN.

(a) Base A				(b) Base B			
a	b	c	<-Classificado como	a	b	c	<-Classificado como
15	65	1	a = Substituição	67	13	1	a = Substituição
98	1156	291	b = Bom Estado	7	1467	71	b = Bom Estado
1	301	48	c = Alerta	0	73	277	c = Alerta

Fonte: Elaborada pela autora.

a base A e o valor da sua precisão foi de 70%, do valor de *recall* foi 23% e a *F-measure* foi 25% para base B.

Tabela 13 – Matrizes de confusão do *Naive Bayes*.

(a) Base A				(b) Base B			
a	b	c	<-Classificado como	a	b	c	<-Classificado como
73	0	8	a = Substituição	81	0	0	a = Substituição
1255	12	278	b = Bom Estado	753	230	562	b = Bom Estado
269	22	59	c = Alerta	169	41	140	c = Alerta

Fonte: Elaborada pela autora.

Além dos algoritmos monolítico foi aplicado alguns Sistemas de Múltiplos Classificadores. O primeiro teste foi realizado utilizando o *Random Forest*, visto que o resultado da árvore de decisão foi o melhor e este algoritmo é um conjunto de árvores. As Tabelas 14a e 14b mostram as matrizes de confusão obtidas. Os resultados de todas as medidas foram aproximadamente 90% e 99% para a base A e B respectivamente. Portanto, não houve uma melhora significativa com esse sistema, sendo possível utilizar apenas a árvore de decisão que é mais simples.

Tabela 14 – Matrizes de confusão do *Random Forest*.

(a) Base A				(b) Base B			
a	b	c	<-Classificado como	a	b	c	<-Classificado como
37	39	5	a = Substituição	80	0	1	a = Substituição
9	1501	35	b = Bom Estado	0	1540	5	b = Bom Estado
6	89	255	c = Alerta	1	0	349	c = Alerta

Fonte: Elaborada pela autora.

O próximo algoritmo testado foi o *Bagging*. Este algoritmo gera diversos classificadores e seleciona aleatoriamente os subconjuntos de amostras para o treinamento (KO; SABOURIN; Britto, Jr., 2008). Os classificadores irão se comportar dependendo do subconjunto amostrado. A opção utilizada para essa classificação foi o *REPTree*, que é uma árvore de decisão de aprendizagem rápida. Analisando as Tabelas 15a e 15b, conclui-se que o resultado foi praticamente o mesmo gerado pelo *Random Forest* em ambas as situações.

Este algoritmo teve as métricas com o valor de aproximadamente 90% (base A) e 99% (base B).

Tabela 15 – Matrizes de confusão do *Bagging*.

(a) Base A				(b) Base B			
a	b	c	<-Classificado como	a	b	c	<-Classificado como
27	50	4	a = Substituição	80	0	1	a = Substituição
12	1498	35	b = Bom Estado	0	1537	8	b = Bom Estado
6	99	245	c = Alerta	1	1	348	c = Alerta

Fonte: Elaborada pela autora.

Ao contrário do *Bagging*, o *Boosting* usa uma parte das amostras para treinar classificadores sem ser de forma aleatória. Nesse contexto, as amostras consideradas mais difíceis têm maior probabilidade de serem selecionadas, e as amostras mais fáceis têm menos chance de serem usadas para treinamento (KO; SABOURIN; Britto, Jr., 2008). Portanto, a maioria dos classificadores concentrarão em amostras difíceis, tornando o algoritmo mais eficaz dependendo da situação. Entretanto, pelo *software* WEKA, o *Boosting* é alterado para *AdaBoost*, ou seja, é um *Adaptive Boosting*. Sendo assim, este algoritmo é adaptável, ou seja, as classificações subsequentes serão ajustadas de tal forma que as instâncias classificadas negativamente pelas classificações anteriores sejam priorizadas. As Tabelas 16a e 16b mostram os resultados do *AdaBoost*. Assim como no *Bagging*, foi utilizado a opção de classificador *REPTree*. O resultado da base B foi novamente melhor. Para este algoritmo, as métricas foram todas de aproximadamente 82% e 94%.

Tabela 16 – Matrizes de confusão do *AdaBoost*.

(a) Base A				(b) Base B			
a	b	c	<-Classificado como	a	b	c	<-Classificado como
0	81	0	a = Substituição	80	0	1	a = Substituição
0	1538	7	b = Bom Estado	9	1524	12	b = Bom Estado
0	261	89	c = Alerta	5	100	245	c = Alerta

Fonte: Elaborada pela autora.

Por fim, foi testado o algoritmo *Random Subspaces*. Este método utiliza diversos subconjuntos e cria diversos classificadores para realizar o treinamento (KO; SABOURIN; Britto, Jr., 2008). Nesse cenário, diferentes classificadores são utilizados para classificação de subespaços diferentes. Os resultados são mostrados nas Tabelas 17a e 17b. Novamente, o classificador utilizado foi o classificador *REPTree*. Os resultados da taxa de acerto obtidos foram de aproximadamente 85% e 98,8%, sendo a maior taxa referente a base B. Para a base A, a precisão foi de 86,7%, do valor de *recall* foi 85,1% e a *F-measure* foi 81,7%. Enquanto para a base B, todas as métricas resultaram em 98,8%.

Tabela 17 – Matrizes de confusão do *Random Subspaces*.

(a) Base A				(b) Base B			
a	b	c	<-Classificado como	a	b	c	<-Classificado como
5	73	3	a = Substituição	80	0	1	a = Substituição
0	1539	6	b = Bom Estado	0	1534	11	b = Bom Estado
0	212	138	c = Alerta	1	10	339	c = Alerta

Fonte: Elaborada pela autora.

Nas Tabelas 18a e 18b é visível a diferença de desempenho entre os classificadores utilizando as diferentes combinações de características. No cenário das buchas, nota-se que um algoritmo robusto como a árvore de decisão é solução para auxiliar nas decisões e sem necessitar de uma normalização prévia dos dados, o que facilita na implementação. Além disso, são robustas à ruídos, tornando-a uma boa alternativa para o cenário da redes elétricas que sofrem constantemente a influência de alteração de carga e da variação de temperatura.

Tabela 18 – Tabela de comparação dos classificadores

(a) Base A				
Algoritmo	Taxa de Acerto	Precisão	<i>Recall</i>	<i>F-Measure</i>
C4.5	88,82%	88%	88,8%	88,1%
KNN	61,69%	62,4%	61,7%	62%
<i>Naive Bayes</i>	7,29%	30,8%	27,3%	4,6%
<i>Random Forest</i>	90,74%	90,3%	90,7%	90,2%
<i>Bagging</i>	89,57%	88,9%	89,6%	88,8%
<i>AdaBoost</i>	82,34%	-	82,3%	-
<i>Random Subspaces</i>	85,12%	86,7%	85,1%	81,7%

(b) Base B				
Algoritmo	Taxa de Acerto	Precisão	<i>Recall</i>	<i>F-Measure</i>
C4.5	99,44%	99,4%	99,4%	99,4%
KNN	91,65%	91,6%	91,6%	91,6%
<i>Naive Bayes</i>	22,82%	70,2%	22,8%	25,1%
<i>Random Forest</i>	99,64%	99,7%	99,6%	99,6%
<i>Bagging</i>	99,44%	99,5%	99,4%	99,4%
<i>AdaBoost</i>	93,57%	93,7%	93,6%	93,2%
<i>Random Subspaces</i>	98,84%	98,8%	98,8%	98,8%

Fonte: Elaborada pela autora.

5.4 Estudo dos Efeitos da Temperatura

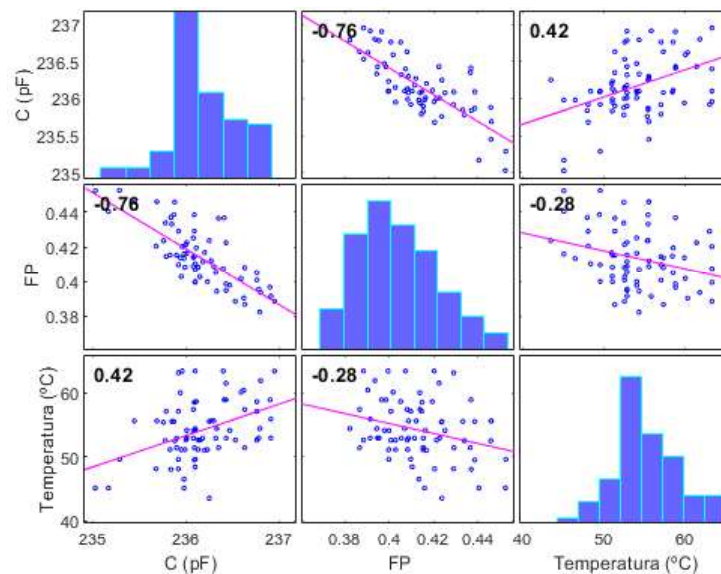
A partir dos resultados obtidos na etapa de identificação das anomalias utilizando a aprendizagem de máquina, é possível analisar alguns padrões nos dados. Para as buchas capacitivas, não foi possível encontrar um hiperplano que separasse os dados de acordo com a condição de operação. Portanto, é possível confirmar que, independente do estado da

bucha, os valores de capacitância e fator de potência podem ter variações que interferem na decisão a partir do limiar, dependendo da precisão do sistema e de quanto ruído está presente.

5.4.1 Correlação entre as Características

Com o objetivo de aprofundar o entendimento e a modelagem da bucha capacitiva, foram analisadas algumas características de correlação entre os parâmetros da bucha. A medida da correlação é utilizada para verificar se há alguma dependência específica entre os parâmetros monitorados nas subestações. Para os dados de Caxias foi possível analisar dados de cada um dos estados da bucha. Dessa forma, foi verificado se há dependência entre as características e se essa dependência se altera conforme o envelhecimento da bucha. Os padrões encontrados para cada um dos estados da bucha. Como cada estado apresentou características parecidas, foram escolhidas três correlações realizadas para ilustrar os resultados obtidos. Sendo assim, a Figura 27 representa um dos resultados das buchas novas, a Figura 28 representa as buchas em estado de alerta e a Figura 29 representa as buchas que necessitam ser trocadas.

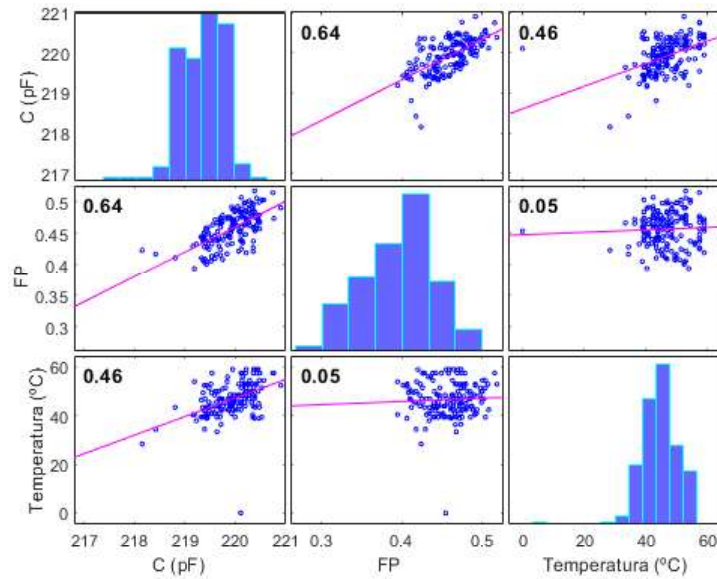
Figura 27 – Correlação dos dados para buchas em bom estado.



Fonte: Elaborada pela autora

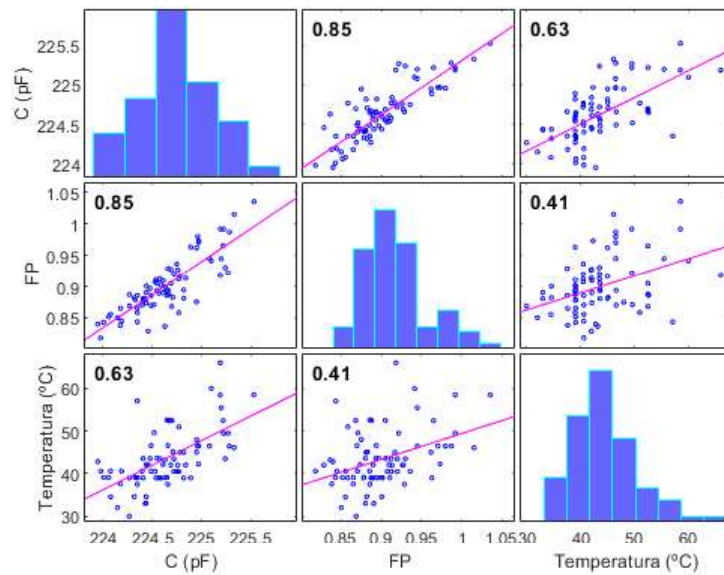
Para essas análises foi dado foco nas correlações entre capacitância, fator de potência e temperatura. Nesse contexto, foi verificado que a correlação de cada estado da bucha seguia um padrão, independente da fase ou unidade escolhida para analisar. Portanto, para representar os resultados obtidos foram escolhidas apenas algumas partes da base de Caxias para serem plotados os gráficos de correlação. Sendo assim, comparando as figuras geradas

Figura 28 – Correlação dos dados para buchas em estado de alerta.



Fonte: Elaborada pela autora

Figura 29 – Correlação dos dados para buchas que devem ser trocadas.



Fonte: Elaborada pela autora

da correlação de cada estado da bucha, é possível concluir que, para esta subestação, a capacitância será diretamente correlacionada com a temperatura, independente do tempo de operação. Além disso, essa correlação não se altera muito ao longo de toda a vida de uma bucha capacitiva.

A principal diferença encontrada foi que para buchas novas a correlação é negativa entre a capacitância e o fator de potência, em torno de 0,80. Já para buchas com algum nível de degradação, essa correlação fica positiva e também em torno de 0,80. Além disso, pelos resultados obtidos, o fator de potência parece ter uma correlação cada vez mais positiva ao longo do tempo em relação a temperatura do óleo.

A partir desses resultados foi possível compreender as limitações do trabalho para gerar a modelagem da bucha capacitiva com dependência da temperatura. Portanto, a primeira limitação visível é o comportamento da correlação entre as características ao longo do tempo. Entretanto, no modelo para simular a dependência da temperatura do óleo não contempla essa variação. Outra limitação é que esse comportamento foi verificado em apenas uma subestação, ou seja, o espaço amostral não é o suficiente para provar essa hipótese.

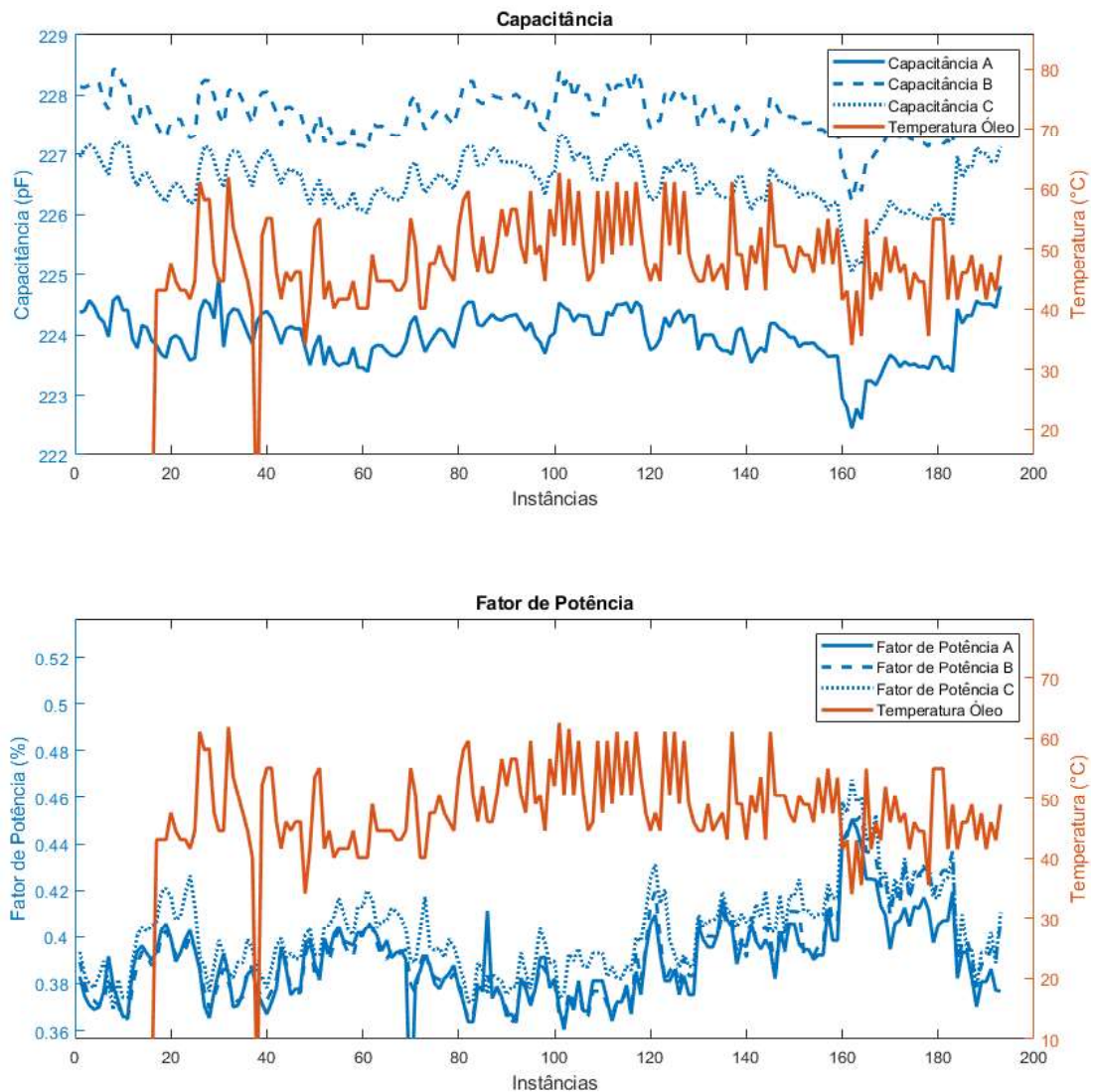
5.4.2 Modelagem do Efeito da Temperatura

Neste estudo, investigamos a modelagem do comportamento da bucha capacitiva, tendo como base o modelo ideal. O modelo proposto tem como objetivo avaliar o comportamento dos valores de capacitância e fator de potência com a adição dos efeitos da temperatura na geração dos dados. Para avaliar os resultados, foram usadas as temperaturas dos transformadores da subestação de Caxias para gerar essa dependência e comparar o comportamento dos dados simulados em função dos dados reais. O efeito da temperatura se baseia no estudo apresentado em (MURTY, 2017b), que relaciona a temperatura do óleo do transformador com os fatores de correção para o cálculo da $\tan(\delta)$. Esses valores de correção são apresentados na Tabela 7 e foram utilizados para gerar a variação nos dados de capacitância.

A Figura 30 apresenta os dados da Unidade 3 de Caxias comparado os padrões da capacitância e fator de potência com a variação da temperatura ao longo do tempo. Nota-se que a relação entre essas características é bem visível. Sendo que a capacitância está variando diretamente com a temperatura, enquanto o fator de potência varia de forma indireta, assim como avaliada na correlação mostrada na Figura 27.

A Figura 31 apresenta os dados gerados a partir da temperatura e a última medição *off-line* da Unidade 1 da subestação de Caxias. Analisando os gráficos gerados é possível verificar a influência da temperatura. Essa variação aparenta estar dentro dos padrões para o fator de potência e para capacitância, considerando o fator de correção de buchas novas disponibilizada pelas fabricantes. Para o fator de potência, nota-se que quanto maior é o fator de potência, maior é a influência da temperatura na sua variação. Entretanto, nessa simulação a variação de correlação entre as características não são abordadas nessa simulação para o estudo mais simplificado do efeito da temperatura.

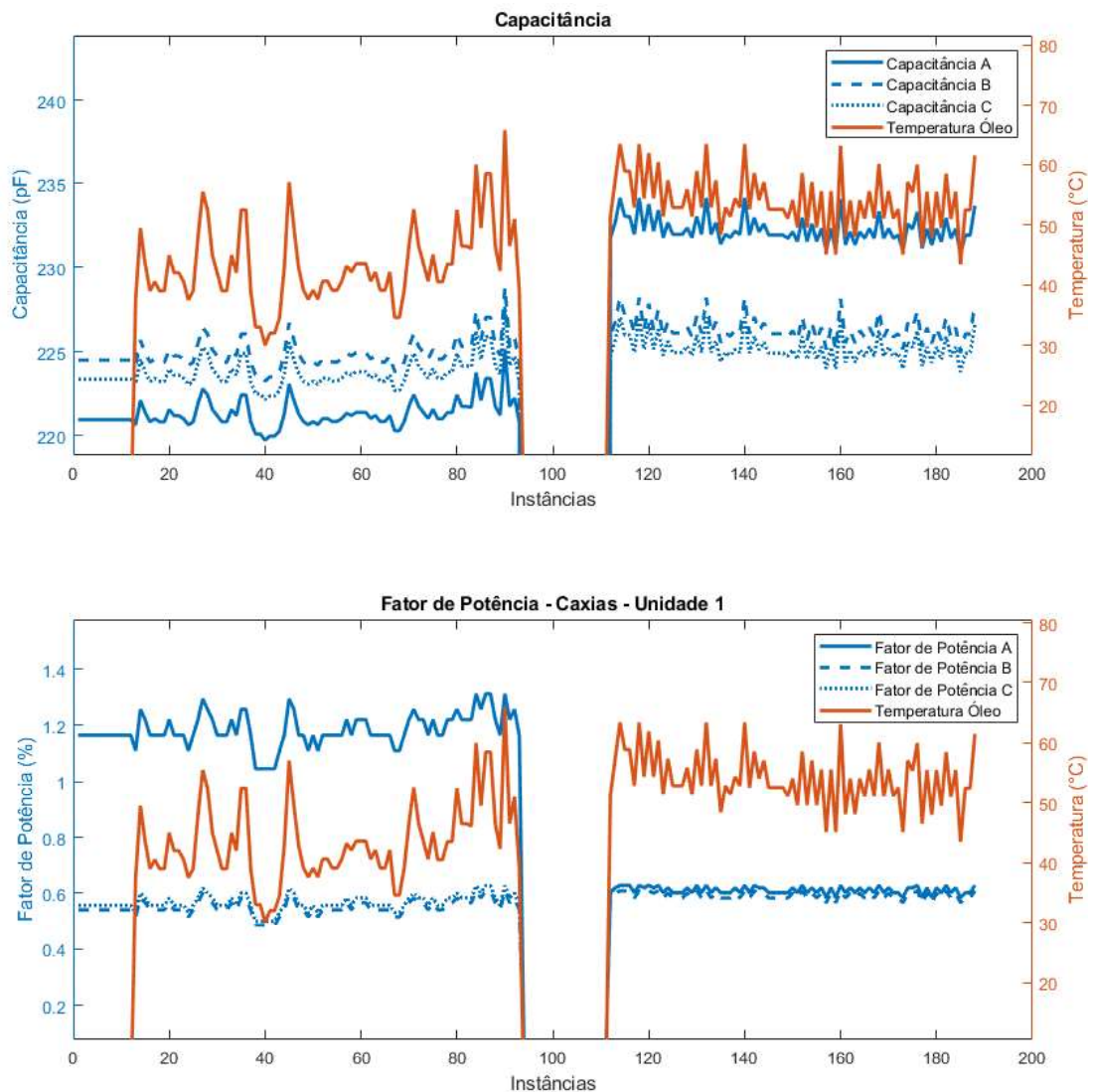
Figura 30 – Comparação da capacitância e fator de potência com a temperatura do óleo



Fonte: Elaborada pela autora

Nas Figuras 32 e 33 nota-se que a capacitância varia menos que o fator de potência. Nesse contexto, o modelo ainda não é a melhor representação para a variação de capacitância ou de fator de potência, visto que não é possível visualizar nenhuma correlação entre a temperatura e a capacitância e também não é imposto nenhuma regra de variação da correlação entre a temperatura e fator de potência, visto que esse padrão é verificado apenas na subestação de Caxias e necessitaria de mais amostras para validar esse padrão. Na Figura 32, a primeira metade representa o comportamento de uma bucha que deve ser trocada. Na segunda metade, é representado o comportamento de uma bucha que acabou de ser trocada, portanto é uma bucha nova. Já na Figura 33 é representado uma bucha em bom estado, mas que já está sendo utilizada por um tempo.

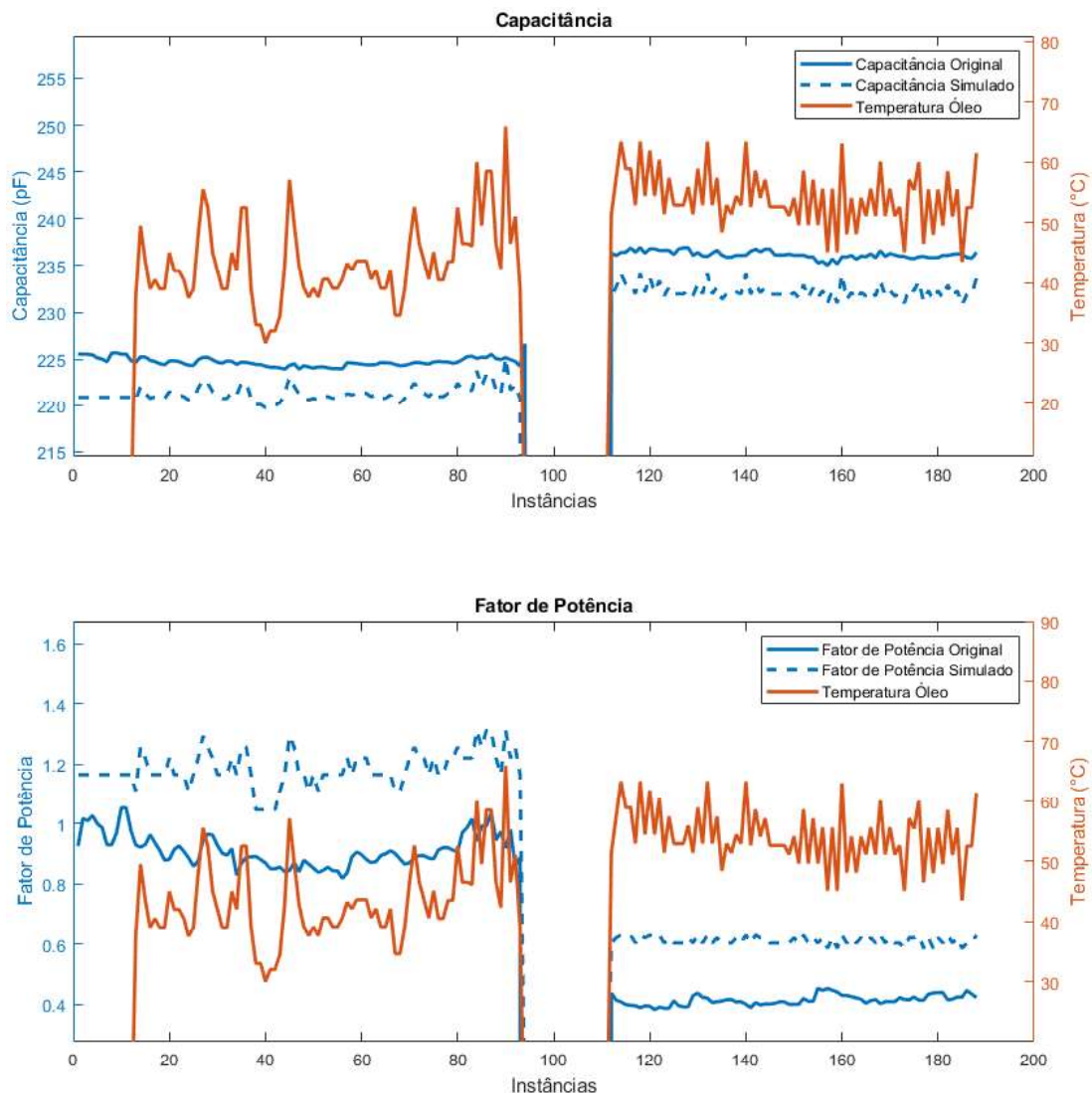
Figura 31 – Dados simulados.



Fonte: Elaborada pela autora

Comparando os dados reais com os dados simulados, nota-se que a variação dos fatores de potência e das capacitâncias estão coerentes. É necessário ressaltar que a correlação positiva ou negativa em relação a temperatura não fica tão nítida nos dados simulados quanto nos dados obtidos na unidade. Esse modelo consegue introduzir a variação da temperatura no modelo ideal, mas ainda há muitas limitações em relação a gerar um comportamento mais próximo do real. Dessa forma, ainda é preciso compreender como simular a variação da correlação entre características e validar se realmente esse padrão ocorre em outros sistemas. Entretanto, para este trabalho não foi possível encontrar uma forma para melhorar os valores de capacitância simulado.

Figura 32 – Comparação dos dados simulados com dados reais de buchas antes de trocar e depois de trocar.

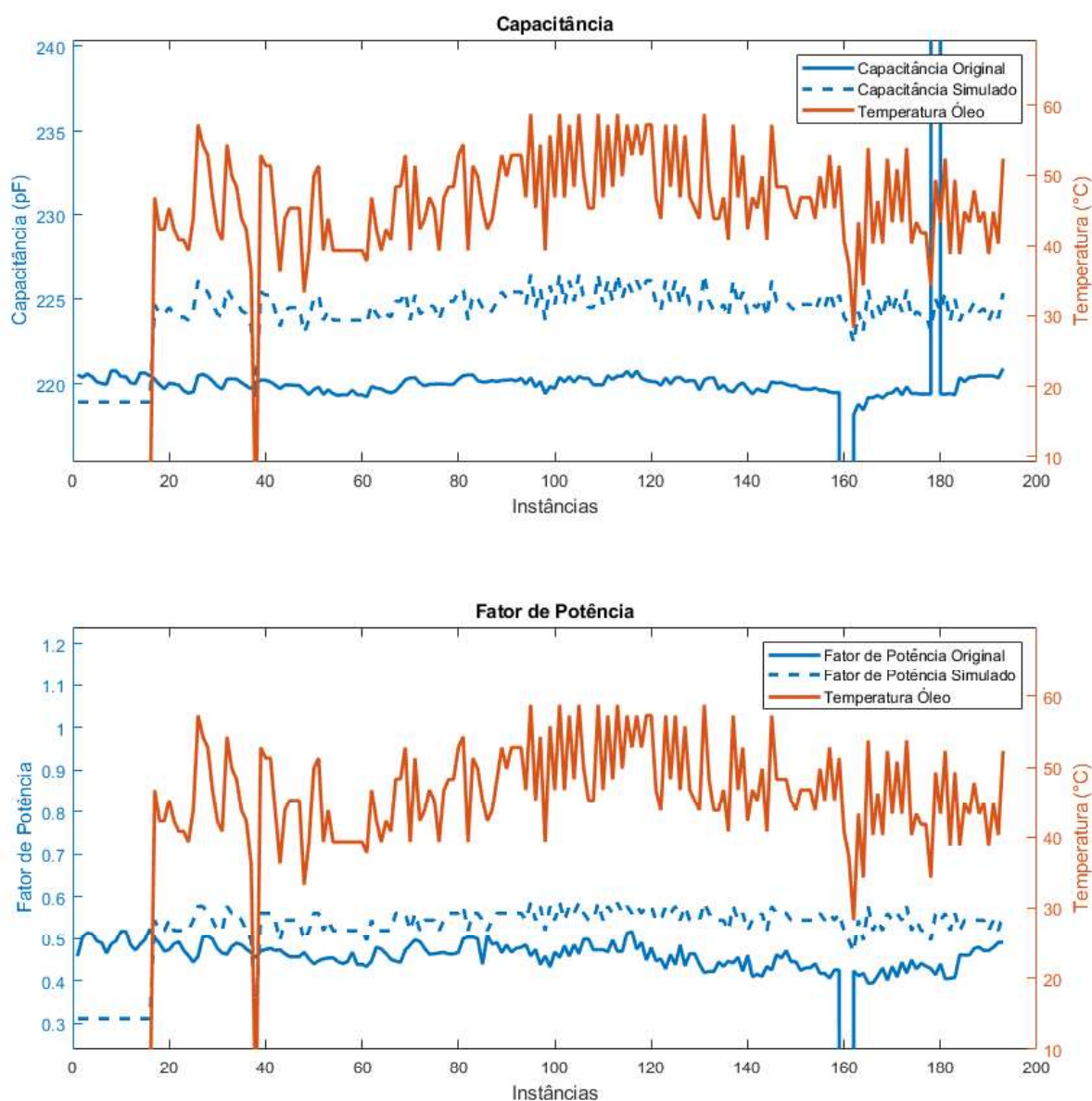


Fonte: Elaborada pela autora

5.5 Dados Adicionais

Atualmente o sistema de monitoramento *online* da subestação de Bateias está em operação e novos dados das buchas capacitivas estão sendo coletados. Assim que tivermos um período maior de coleta, pretendemos realizar as mesmas análises que desenvolvemos com os dados da subestação de Caxias. Nesta seção apresentamos, a título de exemplo, parte dos dados já coletados. Também é importante observar que os dados de Bateias obtidos até agora contemplam apenas medidas de buchas em bom estado. Assim que tivermos uma base mais completa poderemos aplicar as mesmas técnicas de identificação de anomalias usadas no cenário de Caxias.

Figura 33 – Comparação dos dados simulados com dados reais de uma bucha em estado normal de operação.

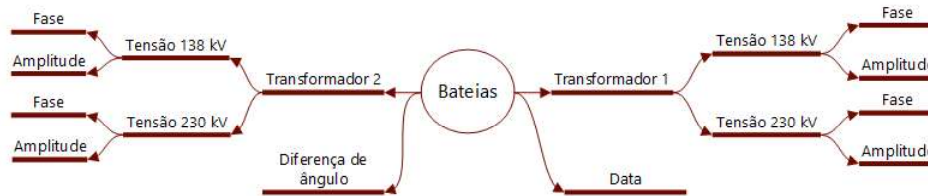


Fonte: Elaborada pela autora

A base de dados obtida na subestação de Bateias inclui os dois lados (baixa e alta tensão) de dois transformadores. Essa base de dados contém 17.700 instâncias que equivalem a aproximadamente 3 meses de medidas (julho, agosto e setembro). Cada instância é referente a uma fase de um lado do transformador, portanto em cada fase há 1.475 instâncias. As características obtidas nessa subestação são mostradas na Figura 34 e são: a *Data*, a *Diferença de ângulo*, a *Amplitude* das tensões e das correntes do lado 138 kV e do lado 230kV e a *Fase* das tensões e das correntes também do lado 138 kV e do lado 230kV. Portanto, para as análises, o primeiro processamento foi transformar os

valores medidos de tensão e corrente em *Fator de Potência* (Equação 2.3) e *Capacitância* (Equação 2.6).

Figura 34 – Características obtidas em Bateias.

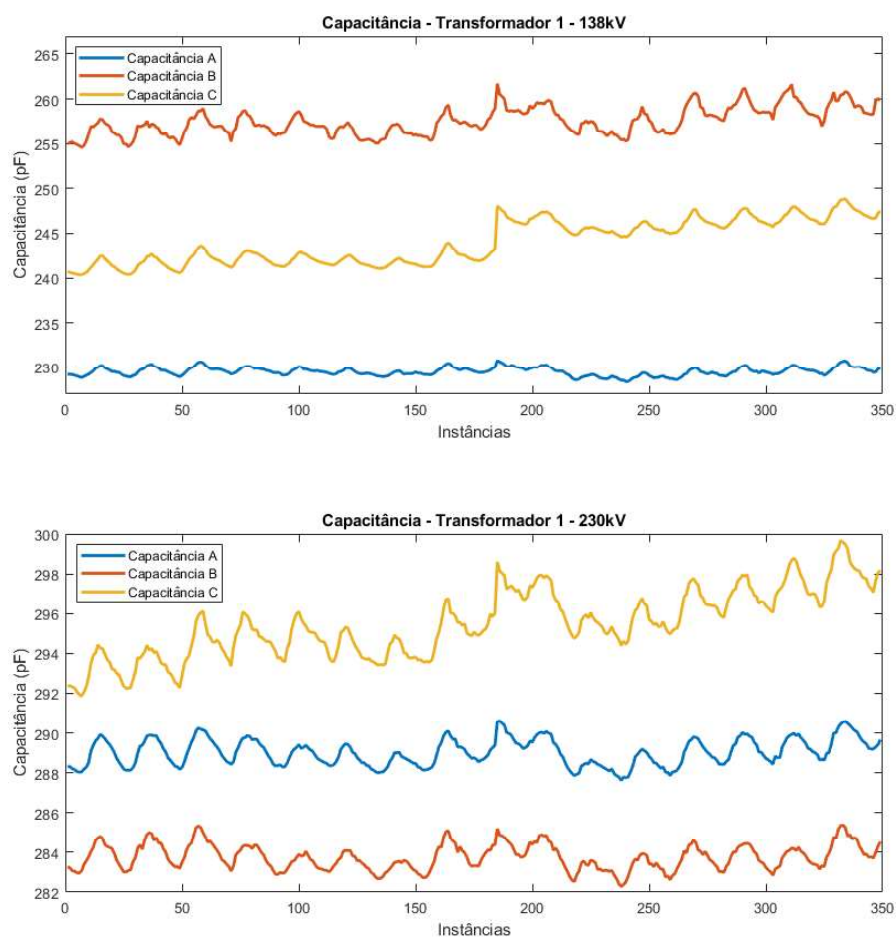


Fonte: Elaborada pela autora

As Figuras 35 e 36 apresentam os dados do transformador 1 e 2, para ambos os lados, 138kV e 230kV. Nas figuras é mostrado menos instâncias, pois foi realizado uma média móvel ao plotar os dados. Além disso, é possível verificar que as buchas instaladas em cada uma das três fases de cada lado do transformador estão com variações parecidas e há uma tendência dos valores aumentarem. Como já mencionado anteriormente, a análise que precisa ser realizado é em relação ao seu valor inicial, portanto o que deve ser considerado é a variação ao longo do tempo.

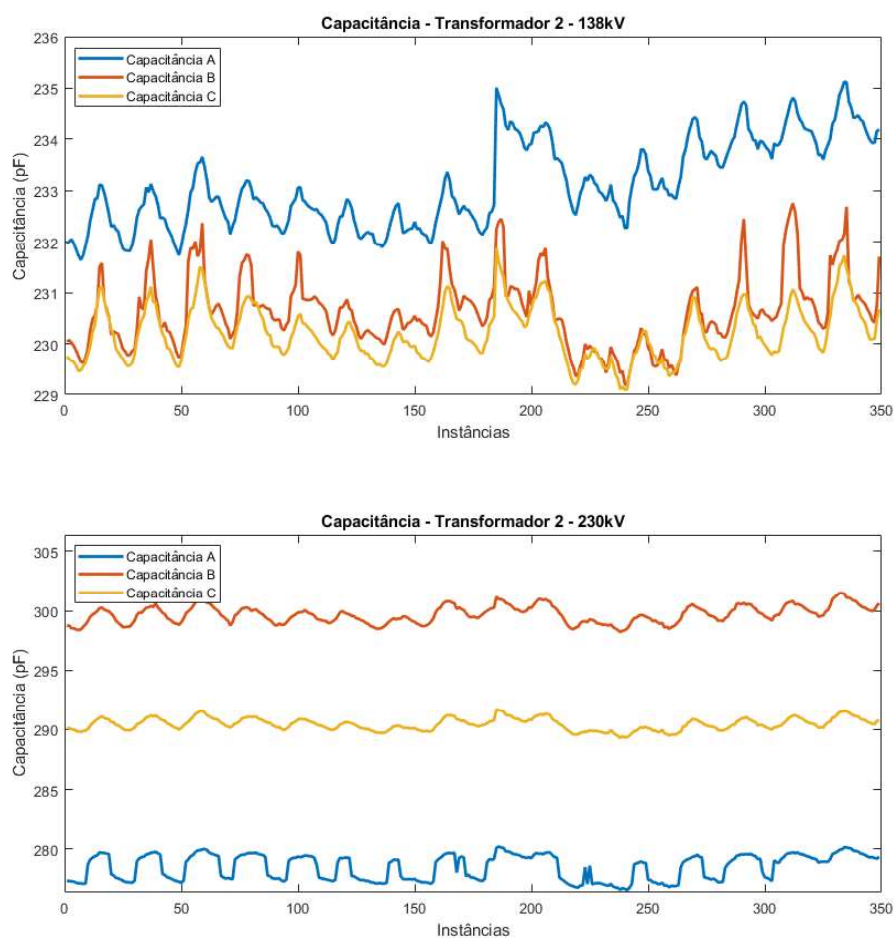
Independente do lado e do transformador, nota-se que a variação presente nas capacitâncias são parecidas. Neste contexto, é possível analisar se a bucha contém umidade em sua estrutura ou com alguma contaminação em seu sistema de isolamento. Entretanto como os gráficos não demonstram muita diferença entre as medições e são dados de um curto período, não é possível concluir se há algum problema com a bucha apenas com esses dados. Sendo assim, a concessionária precisaria tomar cuidados analisando as buchas ao longo de suas vidas para poder tomar decisões mais precisas.

Figura 35 – Dados do transformador 1 de Bateias para a tensão de 138kV.



Fonte: Elaborada pela autora

Figura 36 – Dados do transformador 1 de Bateias para a tensão de 230kV.



Fonte: Elaborada pela autora

6 Conclusão

Considerando o contexto de *smart grids*, nos últimos anos diversas abordagens para detecção de anomalias no sistema elétrico de potência foram propostas na literatura. Contudo, ainda existe pouca informação e documentação disponível sobre a modelagem e o comportamento das buchas capacitivas durante o seu ciclo de vida. Portanto, esta pesquisa foi realizada com o objetivo de se compreender melhor o comportamento da bucha e as variações das métricas envolvidas em função da sua condição de operação, além de fatores como a temperatura.

O primeiro desafio desta pesquisa foi a aquisição dos dados, pois o uso de monitoramento *on-line* das buchas é uma abordagem relativamente recente. Por esse motivo ainda não existem muitos dados de buchas em estágios mais degradados. Isso representa um desafio para a utilização de técnicas de aprendizagem de máquina, uma vez que é preciso considerar o desbalanceamento da base de dados. Para a avaliação dos algoritmos de aprendizagem de máquina, foi necessário preparar a base de dados. Inicialmente este procedimento de rotulação foi feito com o auxílio dos operadores da subestação. Contudo, é uma tarefa que exige profissionais da área e não é sempre que há alguém disponível para auxiliar nessa tarefa. Além disso, os dados de capacitância e fator de potência são analisados a partir de seu valor inicial. Muitas vezes não há todo o histórico da bucha, o que pode comprometer os resultados.

Os resultados obtidos demonstram que as métricas utilizadas são adequadas para analisar uma bucha. Contudo, as buchas capacitivas tem métricas muito sensíveis a ruídos e que necessitam ser analisadas de acordo com seu valor inicial. Considerando a primeira hipótese de pesquisa, esta hipótese foi validada com os resultados obtidos, que também auxiliaram na compreensão do comportamento e modelagem da bucha. A segunda hipótese de pesquisa não foi considerada falsa, visto que até o momento ainda não foi possível gerar um modelo de bucha capacitiva mais completo apenas adicionando os efeitos de temperatura. A geração de um modelo adicionando-se apenas a variação da temperatura e os resultados obtidos das correlações ainda precisa ser aprimorado.

A variação descrita nos manuais de fabricantes é referente a buchas novas e têm um comportamento perto do ideal. Entretanto, quanto mais degradada estiver a bucha, menos preciso é o modelo que representa o seu comportamento. No contexto da modelagem, apenas a adição do fator de correção da temperatura não foi o suficiente para gerar um modelo mais completo, visto que o modelo gera uma variação de capacitância muito grande para os padrões da bucha. Sendo assim, ainda é necessário encontrar uma maneira de simular a variação sem afetar tanto a capacitância. Portanto, para este trabalho foi possível

gerar a influência da temperatura, mas não foi possível amenizar os efeitos na capacitância, que será assunto de trabalhos futuros.

A medida que novos dados estejam disponíveis a partir do sistema de monitoramento *online*, com um histórico mais longo de operação das buchas capacitivas, será possível melhorar e atualizar o modelo de comportamento da bucha em seus diferentes estados de operação.

6.1 Trabalhos Futuros

A partir dos estudos desenvolvidos neste trabalho, identificamos alguns pontos importantes para trabalhos futuros:

- Avaliar os algoritmos de aprendizagem de máquina utilizando mais instâncias, além de dados de outras subestações para determinar melhor o comportamento das buchas capacitivas. Tendo em vista que cada bucha tem uma capacitância inicial diferente, quanto maior número de amostras melhor será a extração de características para determinar o seu estado e modelar variações do seu comportamento.
- Investigar outras métricas e o seu impacto no comportamento da bucha capacitiva e dos seus estágios de operação. Além dos valores de capacitância e fator de potência, outros parâmetros podem influenciar em menor ou maior grau no comportamento da bucha, como a temperatura ambiente, umidade do ar, temperatura do óleo do transformador e carga no transformador. Isso pode possibilitar uma modelagem mais completa da bucha capacitiva.
- Desenvolver novos métodos para avaliar a vida útil das buchas capacitivas, utilizando outros algoritmos de aprendizagem de máquina para gerar análise de anomalias transitórias. Essa avaliação pode auxiliar as concessionárias na melhoria do processo de manutenção.
- Desenvolver novo modelo para as buchas capacitivas que inclua a variação da temperatura de acordo com o estado das buchas, tendo em vista que em buchas mais degradadas são mais influenciadas por essa variação.

Referências

ABB. *Bushing diagnostics and conditioning*. 2021. Citado na página 51.

(ABNT), Associação Brasileira de Normas Técnicas. Buchas para tensões alternadas superiores a 1 kv. *ABNT NBR 5034-2014*, 10 2014. Citado na página 23.

AHMED, Z. Analysis of partial discharge in oip bushing models. In: . [s.n.], 2011. Disponível em: <<http://kth.diva-portal.org/smash/get/diva2:511553/FULLTEXT01.pdf>>. Citado 2 vezes nas páginas 19 e 20.

AJ, Christina; SALAM, M.A.; RAHMAN, Q.M.; WEN, Fushuan; ANG, S.P.; VOON, William. Causes of transformer failures and diagnostic methods – a review. *Renewable and Sustainable Energy Reviews*, v. 82, p. 1442–1456, 2018. ISSN 1364-0321. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1364032117307967>>. Citado 4 vezes nas páginas 34, 39, 41 e 42.

ANDRZEJEWSKI, M; GIL, W; MIKRONIKA, P Wronek. The high voltage bushing on-line monitoring in substation systems. In: IEEE. *2014 ICHVE International Conference on High Voltage Engineering and Application*. [S.l.], 2014. p. 1–4. Citado 2 vezes nas páginas 38 e 39.

ARAUJO, Daniel C. P.; G.ALVES, Marcos E.; TEIXEIRA, Romulo M.; MOREIRA, Cairo R. N. Prevenção de falha em transformador trifásico 525 kv pela monitoração on-line de tangente delta de buchas. *XXII SNPTEE SEMINÁRIO NACIONAL DE PRODUÇÃO E TRANSMISSÃO DE ENERGIA ELÉTRICA*, 10 2013. Disponível em: <http://tretech.com.br/wp-content/uploads/2015/01/XXIISNPTEE_Cemig_Mesquita_2013.pdf>. Citado na página 22.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *NBR 5034*: Buchas para tensões alternadas superiores a 1 kv. [S.l.], 2014. 28 p. Citado na página 19.

ATALLAH, M; FILHO, Orsino Borges; FUTINO, E; SANTOS, E; ZIRBES, R; LOPES, J; GONTIJO, G; BRANDÃO, J; ABREU, Fábio. Avaliação de buchas capacitivas para transformadores e reatores por meio de ensaios em alta tensão. In: . [S.l.: s.n.], 2000. p. 6. Citado na página 34.

BARANAUSKAS, JA; MONARD, MC. *Reviewing some machine learning concepts and methods*. *Relatório Técnico 102, ICMC-USP*. 2000. Citado na página 27.

BARANAUSKAS, José Augusto; MONARD, Maria Carolina. *Reviewing some machine learning concepts and methods*. [S.l.]: ICMC-USP, 2000. Citado 2 vezes nas páginas 27 e 28.

BHUTADA, Sagar; JOSHI, Shailesh; KARANDIKAR, Parashuram; HOLMUKHE, Rajesh. Insulation assessment of oil impregnated paper condenser bushings using dielectric frequency response technique. *Indian Journal of Science and Technology*, v. 10, p. 1–6, 06 2017. Citado na página 22.

BOTELHO, Marcos Paulo Giarola. Estratégias de monitoramento em tempo real de buchas condensivas. Universidade Federal do Rio de Janeiro, 2011. Citado na página 19.

BRANDÃO, Marcelo; LEAL, Aline. *Apagão no Amapá não foi provocado por raio, diz laudo*: Incêndio no transformador foi provocado por superaquecimento. 2020. Disponível em: <<https://agenciabrasil.ebc.com.br/geral/noticia/2020-11/apagao-no-amapa-nao-foi-provocado-por-raio-diz-laudo>>. Citado na página 15.

BREIMAN, Leo. Bagging predictors. *Machine learning*, Springer, v. 24, n. 2, p. 123–140, 1996. Citado na página 32.

CHEN, Lihua; ZHENG, Zhong; LIU, Shaoyu; GUO, Liang; DUN, Chun. Temperature prediction on power transformers and the guide on load dispatch. In: *2012 Annual Report Conference on Electrical Insulation and Dielectric Phenomena*. [S.l.: s.n.], 2012. p. 444–447. Citado 2 vezes nas páginas 49 e 50.

COHEN, Jacob. *Statistical power analysis for the behavioral sciences*. [S.l.]: Academic press, 2013. Citado na página 33.

COMISSION, International Electrotechnical. Insulated bushings for alternating voltages above 1 000 v. *IEC 60137 Edition 6.0*, 7 2008. Citado na página 22.

DUDA, Richard; HART, Peter; G.STORK, David. Pattern classification. In: _____. [S.l.: s.n.], 2001. xx. ISBN 0-471-05669-3. Citado na página 31.

ELECTRICAL, The Institute of; ENGINEERS, Electronic; C57.19.100-2012, ANSI/IEEE Std. Ieee guide for application of power apparatus bushings. 2 2013. Citado na página 23.

FREUND, Yoav; SCHAPIRE, Robert E et al. Experiments with a new boosting algorithm. In: CITESEER. *icml*. [S.l.], 1996. v. 96, p. 148–156. Citado na página 32.

FRONTIN, Sergio O.; NOMI, Yoshio; CABRAL, Sidnei; SILVA, Rivaiele. *Equipamentos de alta tensão - prospecção e hierarquização de inovações tecnológicas*. 1. ed. [S.l.: s.n.], 2013. 329-349 p. Citado 4 vezes nas páginas 13, 19, 20 e 21.

G1-AP. *Um mês do apagão no Amapá: o que ainda precisa para a segurança energética?* 2020. Disponível em: <<https://g1.globo.com/ap/amapa/noticia/2020/12/03/um-mes-do-apagao-no-amapa-o-que-ainda-precisa-para-a-seguranca-energetica.ghtml>>. Citado na página 15.

GÉRON, Aurélien. *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. [S.l.]: O'Reilly Media, 2019. Citado na página 27.

GUO, Jian; YANG, Xiaobo; REN, Ke; SI, Jiawen; YUE, Pengfei; KHAN, Roaidar. Modeling and simulation of asymmetric three-phase grid voltage phase-locking method based on improved extended kalman filter. In: IEEE. *2019 Chinese Control And Decision Conference (CCDC)*. [S.l.], 2019. p. 4897–4903. Citado 2 vezes nas páginas 40 e 41.

HO, Tin Kam. Random decision forests. In: *Proceedings of 3rd International Conference on Document Analysis and Recognition*. [S.l.: s.n.], 1995. v. 1, p. 278–282 vol.1. Citado na página 32.

HO, Tin Kam. Nearest neighbors in random subspaces. In: SPRINGER. *Joint IAPR International Workshops on Statistical Techniques in Pattern Recognition (SPR) and Structural and Syntactic Pattern Recognition (SSPR)*. [S.l.], 1998. p. 640–648. Citado na página 32.

HOLLANDER, Myles; WOLFE, Douglas A; CHICKEN, Eric. *Nonparametric statistical methods*. [S.l.]: John Wiley & Sons, 2013. v. 751. Citado na página 32.

IEEE Guide for Application of Power Apparatus Bushings. *IEEE Std C57.19.100-1995*, p. 1–32, 1995. Citado na página 45.

KHOUSSEI, S.; MATTAS, A. A brief introduction to smart grid safety and security. In: GRIFFOR, Edward (Ed.). *Handbook of System Safety and Security*. Boston: Syngress, 2017. p. 225–252. ISBN 978-0-12-803773-7. Disponível em: <<https://www.sciencedirect.com/science/article/pii/B9780128037737000115>>. Citado na página 12.

KO, Albert H.R.; SABOURIN, Robert; Britto, Jr., Alceu Souza. From dynamic classifier selection to dynamic ensemble selection. *Pattern Recognition*, v. 41, n. 5, p. 1718–1731, 2008. ISSN 0031-3203. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0031320307004499>>. Citado 4 vezes nas páginas 31, 32, 63 e 64.

MEHTA, Amit Kumar; SHARMA, RN; CHAUHAN, Sushil; AGNIHOTRI, SD. Study and diagnosis of the power transformer bushing insulation system. In: IEEE. *2011 IEEE Pulsed Power Conference*. [S.l.], 2011. p. 700–705. Citado na página 13.

MEIRA, Matias; RUSCHETTI, Cristian R; ÁLVAREZ, Raúl E; VERUCCHI, Carlos J. Power transformers monitoring based on electrical measurements: state of the art. *IET Generation, Transmission & Distribution*, IET, v. 12, n. 12, p. 2805–2815, 2018. Citado 2 vezes nas páginas 13 e 14.

MORALES, Emilio. A new approach for transformer bushing monitoring. 2012. Citado 6 vezes nas páginas 15, 34, 35, 36, 37 e 39.

MURTY, K.K. *Fundamentals of condenser bushings*. 2017. Citado na página 25.

MURTY, K. K. *Condenser bushings condition monitoring*. TRANSFORMERS MAGAZINE, 2017. 70–81 p. Disponível em: <<https://transformers-magazine.com/magazine/4935-condenser-bushings-condition-monitoring/>>. Citado 2 vezes nas páginas 49 e 68.

PAULINO, MEC; ALMEIDA, ATL. Avaliação de buchas de alta tensão com variação de frequência. In: *IV SBSE-Simpósio Brasileiro de Sistemas Eletricos*. [S.l.: s.n.], 2012. Citado na página 15.

PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V.; VANDERPLAS, J.; PASSOS, A.; COURNAPEAU, D.; BRUCHER, M.; PERROT, M.; DUCHESNAY, E. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, v. 12, p. 2825–2830, 2011. Citado na página 46.

- PITHWA, Varun J. Condition monitoring techniques and simulation of transformer differential protection scheme using matlab. *IJSRD - International Journal for Scientific Research & Development*, IJSRD Journal, v. 3, p. 830–833, 2015. Citado 2 vezes nas páginas 39 e 41.
- QUALITROL. *Transformer bushing monitoring*. 2017. [Online; accessed 3. Aug. 2021]. Disponível em: <<https://transformers-magazine.com/magazine/4940-transformer-bushing-monitoring>>. Citado na página 46.
- RAMBO, Marcos Vinicio Haas; KRAUSS, Carlos Cesar; JASINSKI, Roberto; CHUEIRI, Ivan Jorge. Monitoramento em tempo real de buchas capacitivas de transformadores. *XIX SNPTEE*, 2007. Citado 7 vezes nas páginas 13, 14, 24, 25, 26, 45 e 53.
- SAVREE. *Encyclopedia - saVRee*. 2020. Online; accessed 17. Apr. 2021. Disponível em: <<https://savree.com/en/encyclopedia>>. Citado 2 vezes nas páginas 20 e 21.
- SCHERING, H. *Bridges for Loss Measurements*. [S.l.]: Report on the Activities of the Physical-Technical Institute of Realm, Germany, 1919. Citado na página 23.
- SEPTYANI, Henny I.; ARIFIANTO, Indera; PURNOMOADI, Andreas P. High voltage transformer bushing problems. In: *Proceedings of the 2011 International Conference on Electrical Engineering and Informatics*. [S.l.: s.n.], 2011. p. 1–4. Citado na página 41.
- SIMON, P. *Too Big to Ignore: The Business Case for Big Data*. Wiley, 2013. (Wiley and SAS Business Series). ISBN 9781118642108. Disponível em: <<https://books.google.com.br/books?id=Dn-Gdoh66sgC>>. Citado na página 27.
- TAN; STEINBACH. Data mining classification : Basic concepts , decision trees , and model evaluation. In: . [S.l.: s.n.], 2004. Citado 2 vezes nas páginas 29 e 62.
- WAGENAAR, Loren B. Electrical bushings. In: _____. *The Electric Power Engineering Handbook: Electric power transformer engineering*. 3. ed. CRC Press, 2012. cap. 13. Disponível em: <<https://www.routledgehandbooks.com/doi/10.1201/b12110>>. Citado na página 19.
- WANG, Yaw-Juen. Simulation of random variation of three-phase voltage unbalance resulting from load fluctuation using correlated gaussian random variables. In: *PROC NATL SCI COUNC REPUB CHINA PART A PHYS SCI ENG*. [S.l.: s.n.], 2000. v. 24, n. 3, p. 216–225. Citado 2 vezes nas páginas 40 e 41.
- WANG, Y J. Modelling of random variation of three-phase voltage unbalance in electric distribution systems using the trivariate gaussian distribution. *IEE Proceedings-Generation, Transmission and Distribution*, IET, v. 148, n. 4, p. 279–284, 2001. Citado 3 vezes nas páginas 40, 41 e 48.
- WITTEN, Ian H; FRANK, Eibe; HALL, Mark A; PAL, CJ; DATA, MINING. Practical machine learning tools and techniques. In: *DATA MINING*. [S.l.: s.n.], 2005. v. 2, p. 4. Citado 2 vezes nas páginas 46 e 61.
- XIA, Jiuyun; QIAN, Zheng; YU, Hao; YAO, Junda. Research of on-line monitoring method for insulation condition of power transformer bushing. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. *Seventh International Symposium on Precision Mechanical Measurements*. [S.l.], 2016. v. 9903, p. 99031T. Citado 3 vezes nas páginas 37, 38 e 39.