Algoritmo semiautomático para classificação de faltas em geradores síncronos

Gustavo Felipe. Martin Nascimento
GRUCAD/EEL/CTC
Universidade Federal de Santa Catarina
Florianópolis, Santa Catarina, Brasil
gustavo.nascimento@posgrad.ufsc.br

Luis O. S. Grillo
GRUCAD/EEL/CTC
Universidade Federal de Santa Catarina
Florianópolis, Santa Catarina, Brasil
luis.otavio.grillo@posgrad.ufsc.br

Eduardo S. M. de Souza
GRUCAD/EEL/CTC
Universidade Federal de Santa Catarina
Florianópolis, Santa Catarina, Brasil
eduardo.seiji@grad.ufsc.br

Paulo Cesar Hilleshein
GRUCAD/EEL/CTC
Universidade Federal de Santa Catarina
Florianópolis, Santa Catarina, Brasil
paulo.hilleshein@grad.ufsc.br

Nelson Jhoe Batistela
GRUCAD/EEL/CTC
Universidade Federal de Santa Catarina
Florianópolis, Santa Catarina, Brasil
jhoe.batistela@ufsc.br

Luciano Mendes de Freitas
Eng. de Man. de Ativos de Sist. (EMAS)
ENGIE Brasil Energia
Florianópolis, Santa Catarina, Brasil
luciano.freitas@engie.com

***Resumo*— Geradores síncronos estão entre os principais ativos do sistema elétrico de potência, e a sua indisponibilidade de operação pode afetar a confiabilidade de todo o sistema. Por isso, os investimentos realizados no aprimoramento de técnicas de manutenção preditiva, principalmente não intrusivas, são cada vez maiores. É nesse contexto que se insere o monitoramento de alterações da assinatura do campo magnético externo à máquina. O estado da arte dessa abordagem permite, atualmente, a detecção de faltas. Contudo, a classificação das faltas ainda depende de uma análise de especialistas e, mesmo assim, é uma tarefa árdua. Assim, este trabalho visa à implementação de métodos de machine learning para a classificação de diferentes tipos de faltas em geradores síncronos, utilizando dados históricos de campo magnético externo com o intuito de auxiliar na determinação do tipo de defeito presente na máquina. Com isso, em um primeiro momento, um método de classificação supervisionado, o *Random-Forest*, foi empregado em dados sintéticos, provenientes de modelos de geradores síncronos, para avaliar o sucesso dessa abordagem. Observada a possibilidade de se empregar técnica, o algoritmo é então utilizado em dados coletados de um gerador síncrono instalado no Laboratório de Máquinas e Acionamentos Elétricos (LABMAQ) da Universidade Federal de Santa Catarina (UFSC). É importante ressaltar que métodos desse tipo não substituem o trabalho de um especialista, mas podem ser importantes aliados na classificação das faltas.**

Palavras-chave*— *Geradores síncronos, faltas incipientes, machine-learning, classificação de faltas

I. INTRODUÇÃO

Geradores síncronos (GS) são máquinas elétricas comumente empregadas em usinas hidro e termoelétricas, sendo responsáveis por grande parte da geração de energia no Brasil e no mundo. Desse modo, paradas não programadas, seja pela ocorrência de defeitos nesses equipamentos ou pela execução de manutenções, podem reduzir a confiabilidade do sistema elétrico e acarretar importantes prejuízos, não só econômicos, mas também sociais. Por outro lado, a realização de manutenções preventivas, além de custosa, pode ser inconveniente em momentos de escassez energética, e sua eficácia é limitada pelos critérios estatísticos utilizados.

Nesse contexto, torna-se imperativo o desenvolvimento de metodologias preditivas para detecção de faltas em seus estágios iniciais e de maneira não-invasiva. O estado da arte da detecção dessas faltas se utiliza de medições de sinais de vibração, de corrente e de campo magnético. Dessa maneira, é possível fazer com que as paradas de manutenção sejam mais espaçadas e realizadas em períodos de maior conveniência sem aumentar o risco de ocorrência de eventos catastróficos.

A detecção de faltas a partir do monitoramento de vibração mecânica de geradores síncronos é uma técnica madura comercialmente, já fundamentada em normas técnicas [1][2], e baseia o diagnóstico realizado por uma série de equipamentos comerciais [3][4]. Além disso, outras grandezas, como temperatura e fluxo magnético no entreferro, são utilizadas para complementar o diagnóstico [3][5]. Contudo, a dificuldade envolvida na instalação de sensores representa uma das principais desvantagens destas técnicas, uma vez que, geralmente, envolve processos invasivos, como a desmontagem de partes da máquina e a realização de soldagem ou de perfuração de estruturas do gerador para fixação de sensores.

Neste contexto, foi desenvolvido no âmbito de um programa de P&D da ANEEL um sistema [6] não invasivo para detecção de faltas incipientes a partir do monitoramento do campo magnético disperso no entorno do gerador. Dessa forma, o acesso ao interior da máquina não é mais necessário, pois facilita a instalação do equipamento, tornando-a mais rápida e com menor custo.

No estágio atual, essa abordagem permite a indicação da ocorrência de possíveis faltas ao acompanhar a evolução da magnitude de certas componentes harmônicas da FFT (*Fast Fourier Transform*) do sinal de fluxo magnético [7]. Assim, a partir do conhecimento da assinatura magnética da máquina sadia, pode-se inferir o aparecimento de algum defeito ao se detectar variações significativas na amplitude de algumas harmônicas de interesse. A metodologia de detecção de faltas a partir do monitoramento do histórico de amplitudes das componentes harmônicas do sinal do campo magnético no entorno de uma máquina síncrona pode ser consultada em mais detalhes nas teses de doutorado [8] e [9] e nas dissertações de mestrado [10], [11], [12] e [13].

Apesar de a detecção de faltas por meio do monitoramento do histórico de amplitudes do campo magnético externo da máquina já apresentar resultados promissores, a classificação (indicação do tipo e da localização da falta) ainda é um

Este trabalho foi financiado parcialmente pelas empresas Engie Brasil Energia e Itá Energética S.A. e pelo CNPq.

desafio. Neste contexto, este trabalho visa à implementação de métodos de *machine learning* para a classificação de diferentes tipos de faltas em geradores síncronos, utilizando dados históricos de campo magnético externo. Em um primeiro momento, um método de classificação supervisionado, o *Random-Forest* [14], foi empregado em dados sintéticos, provenientes de modelos de geradores síncronos, para avaliar o sucesso dessa abordagem. Uma vez observada essa possibilidade, o algoritmo foi então utilizado em dados coletados de um gerador síncrono instalado no Laboratório de Máquinas e Acionamentos Elétricos (LABMAQ) da Universidade Federal de Santa Catarina (UFSC).

Nesse artigo, a seção II exibe uma breve introdução ao algoritmo *Random-Forest*. A seção III, então, apresenta a bancada de ensaios utilizada para adquirir os dados experimentais e discute o processo de geração dos dados sintéticos utilizados para o treinamento do modelo de classificação. Por fim, a seção IV apresenta os resultados obtidos.

II. *Random-Forest* Como Algoritmo De Classificação

O *Random-Forest* [14] é um método de *machine learning* supervisionado baseado em árvores de decisão, que pode ser utilizado tanto para problemas de regressão quanto para problemas de classificação. É importante ressaltar que métodos supervisionados são aqueles que se utilizam de dados de treinamento bem conhecidos e rotulados para desenvolver os modelos de regressão/classificação. Na fase de treinamento, estes modelos são determinados a partir de um conjunto de parâmetros conhecidos (*features*) buscando atingir a variável alvo.

No caso específico do *Random-Forest*, ele cria múltiplas árvores de decisão (uma ferramenta de apoio à decisão que representa um conjunto de escolhas na forma gráfica de uma árvore) independentes durante a fase de treinamento, formando uma floresta. O resultado de cada árvore de decisão criada é usado no resultado final. Em problemas de regressão, o resultado final é a média dos resultados de cada árvore, enquanto em problemas de classificação, utiliza-se a moda (resultado que mais se repete) [15]. Essas árvores são formadas escolhendo-se as variáveis (*features*) que compõem os ramos de maneira aleatória. Portanto, o *Random-Forest* corrige o hábito de as árvores de decisão se ajustarem excessivamente (*overfitting*) ao seu conjunto de treinamento [16]. A Fig. 1 ilustra um esquema de uma árvore de decisão, enquanto a Fig. 2 mostra a representação do algoritmo *Random-Forest*.

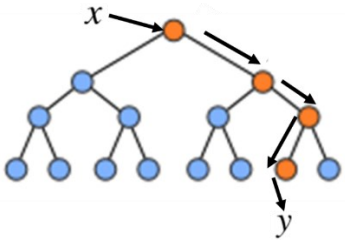


Fig. 1. Desenho esquemático de uma árvore de decisão. [17]

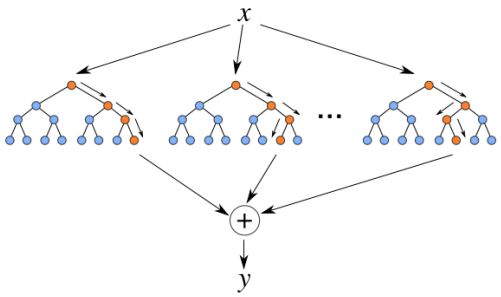


Fig. 2. Desenho esquemático de várias árvores de decisão, formando uma "floresta". [17]

Um exemplo da utilização desse método em problemas de regressão é a previsão do perfil de consumo de energia elétrica de uma edificação utilizando como *features* a temperatura externa, o calendário (dias da semana, mês do ano) e o horário [18] [19].

Neste trabalho, contudo, o *Random-Forest* foi aplicado como método de classificação com o intuito de indicar, com base nas harmônicas sensibilizadas, a localização da falta em geradores síncronos. Ou seja, buscou-se desenvolver um modelo que indicasse se a falta está localizada no rotor ou no estator, utilizando como *features* as harmônicas sensibilizadas pela falta.

III. Determinação do Modelo de Classificação

Como mencionado no item anterior, algoritmos de classificação baseados em métodos supervisionados de *machine learning* necessitam de dados bem rotulados para o seu treinamento e a consequente determinação do modelo de previsão. Nesse contexto, essa seção tem como objetivo apresentar a bancada de testes de imposição de faltas existente no Laboratório de Máquinas e Acionamentos Elétricos da UFSC, de onde os dados experimentais são provenientes, bem como o processo para a geração de dados sintéticos.

A. A bancada de ensaios

Durante o desenvolvimento da tese de Rigoni [8], foi projetada uma bancada, instalada no LABMAQ da UFSC, que possibilitasse a imposição de diferentes tipos de falta em máquinas síncronas.

Essa bancada é constituída, basicamente, de um motor de corrente contínua (CC), utilizado como acionamento mecânico para dois geradores síncronos, um com oito e outro com dois polos, ambos de 10 kVA. Além das diferentes velocidades síncronas, é importante ressaltar que o gerador de maior velocidade (dois polos) é de polos lisos, enquanto o outro é de polos salientes. Neste trabalho, o grupo gerador de oito polos foi objeto de estudo. A Fig. 3 apresenta uma fotografia da bancada, enquanto a Tabela I mostra as especificações das máquinas CC e da síncrona de oitos polos.



Fig. 3. Bancada para imposição de faltas em máquinas síncronas.

Tabela I ESPECIFICAÇÕES DOS MOTORES QUE COMPÕEM A BANCADA

	Motor CC	*MS 8 polos*
Potência	10 kW	10 kVA
Velocidade [rpm]	700 a 4000	750 a 900
Corrente de armadura [A]	56	15,2
Tensão de armadura [V]	220	380
Tensão de campo [V]	220	220
Corrente de campo [A]	5,5/1,6	3,5

Devido aos seus aspectos construtivos, a bancada possibilita a imposição das seguintes faltas:

- Curto-circuito no enrolamento de campo
- Curto-circuito no enrolamento de armadura
- Circuito aberto da gaiola de amortecimento
- Curto-circuito entre lâminas do núcleo
- Excentricidade estática
- Desbalanceamento mecânico.

Neste trabalho, as faltas de curto-circuito nos enrolamentos, tanto do rotor como do estator, são objetos de estudo e podem ser impostas a partir de *taps* existentes para tal fim. No caso do rotor, são três *taps* que permitem o curto-circuito de 20%, 50% ou 100% das espiras de um polo. Já para o estator, os *taps* existentes permitem impor o curto-circuito de 16,7%, 50% ou 100% das espiras de um polo de uma fase.

É importante lembrar que métodos supervisionados de *machine learning* se baseiam em dados bem conhecidos para o seu treinamento e, consequentemente, para a determinação dos modelos de previsão. Portanto, é intuitivo pensar que quanto mais dados houver, e mais diversos eles forem, melhor será o desempenho de um modelo de previsão. Muitas vezes, são usadas milhares de amostras para realizar o treinamento de um modelo que seja confiável. Entretanto, é importante ressaltar que a metodologia de detecção de faltas abordada neste trabalho leva em conta o acompanhamento do histórico de amplitudes de componentes harmônicas do campo magnético.

Por conseguinte, a montagem de um *dataset* de treino demanda diversos desses históricos de amplitudes. Vale lembrar que para a montagem de apenas um histórico de amplitudes, são necessárias diversas aquisições do campo magnético externo. Para cada uma dessas aquisições, necessita-se realizar a FFT, rastrear as componentes harmônicas de interesse e extrair as suas amplitudes. Todo esse processo torna a montagem de um *dataset* de treino com dados experimentais, que demandaria pelo menos dezenas destes históricos, quase inexequível.

B. Geração dos dados sintéticos

Uma das maneiras de se contornar esse problema é com a geração de dados sintéticos, provenientes de softwares de simulação de máquinas elétricas. O trabalho de Aurich [20] propõe um software que simula o funcionamento e a aquisição de dados de sensores de campo magnético externo a uma máquina síncrona em operação. Nesse mesmo software, é possível inserir alguns tipos de falta, tanto no rotor quanto no estator da máquina, com diferentes graus de severidade. Assim, as faltas de curto-circuito de espiras nos enrolamentos do rotor e do estator foram escolhidas para serem inseridas, separadamente, nas simulações, uma vez que são as faltas que podem ser impostas tanto no software quanto na bancada.

Desse modo, foram gerados mais de 300 arquivos, cada um contendo o histórico de 135 componentes harmônicas do sinal simulado de campo magnético externo a uma máquina com características construtivas idênticas à máquina de oito polos existente na bancada descrita no item anterior. Cada um desses históricos contou com 300 amostras, e a falta foi imposta a partir do instante 150, de forma a se conhecer o comportamento da máquina sadia. A Fig. 4 apresenta o exemplo de um dos históricos gerados de uma falta de curto-circuito de espiras no estator, mostrando as primeiras 15 (de 135 selecionadas) harmônicas, enquanto a Tabela II resume as características das faltas impostas nos dados sintéticos.

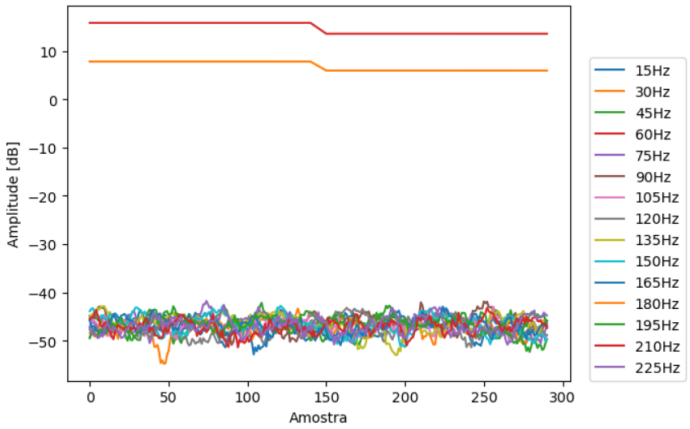


Fig. 4. Exemplo que um histórico de amplitudes de componentes harmônicas com a imposição de uma falta de curto-circuito de espiras do estator.

Tabela II ESPECIFICAÇÕES DOS MOTORES QUE COMPÕEM A BANCADA

Tipo de falta	**Localização**	**Severidade [espiras de um polo]**	**Número de históricos gerados**
Curto-circuito no enrolamento	Rotor	10%	56
		50%	51
		100%	51
	Estator	16,7% de uma fase	50
		50% de uma fase	50
		100% de uma fase	50

Um algoritmo de detecção de anomalias [21], que detecta uma possível falta, é então aplicado a cada um dos históricos gerados, verificando quais são as harmônicas alteradas pela falta imposta. Caso a harmônica seja sensibilizada, ela recebe um *flag* positivo. Como os históricos foram gerados com diferentes ruídos, mesmo faltas idênticas podem sensibilizar componentes diferentes. A Fig. 5 apresenta um exemplo dos dados utilizados para o treinamento do modelo de classificação em forma de uma tabela de dados. Nessa tabela, as colunas representam as componentes harmônicas (que são

	15Hz	30Hz	45Hz	60Hz	75Hz	90Hz	105Hz	120Hz	135Hz	150Hz	...	1905Hz	1920Hz	1935Hz	1950Hz	1965Hz	1980Hz	1995Hz	2010Hz	2025Hz	y
0	0	1	0	1	1	0	0	0	0	1	...	1	0	0	0	0	0	0	1	0	1
1	1	1	0	1	0	1	0	0	1	0	...	0	1	0	0	0	0	0	1	0	1
2	0	1	0	1	1	0	0	0	1	0	...	0	0	1	1	1	1	1	0	0	1
3	1	1	0	1	0	1	0	1	1	0	...	0	1	1	1	0	1	0	0	0	1
4	1	1	0	1	1	1	1	0	1	0	...	1	1	1	1	1	1	1	1	0	1
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
303	1	0	0	1	0	0	0	0	0	0	...	0	0	1	1	1	1	1	0	0	2
304	0	0	0	1	0	0	1	0	1	1	...	1	0	0	0	1	1	0	0	1	2
305	1	0	1	1	0	0	1	0	1	1	...	0	0	0	0	1	1	1	1	0	2
306	0	1	0	1	0	0	1	0	0	1	...	1	1	0	1	1	1	0	1	0	2
307	0	0	0	1	0	0	1	0	0	0	...	0	0	1	0	0	1	1	1	1	2

308 rows × 136 columns

Fig. 5. Exemplo dos dados simulados utilizados para treinamento e teste de um algoritmo de classificação.

utilizadas como *features* no modelo) do sinal de campo magnéticos no entorno da máquina avaliada, com exceção da última coluna, denominada "y" que indica o tipo de falta, se no rotor ou no estator (variável alvo). Em cada linha, está representada um histórico de amplitudes, sendo que o valor 1, indica que aquela frequência foi sensibilizada pela falta.

IV. Classificação de Faltas

Basicamente, um algoritmo de classificação supervisionado busca prever qual é a classe de uma amostra com base nas suas características. No caso específico deste trabalho, os tipos de falta foram divididos em duas classes, considerando a sua localização no rotor ou no estator da máquina.

Desse modo, buscou-se a implementação de um algoritmo que pudesse classificar as faltas com base na sensibilização, ou não, de determinadas frequências em relação à falta imposta. Devido à sua facilidade de implementação e de compreensão, escolheu-se o *Random-Forest*, já apresentado no item 2. Como linguagem de programação, utilizou-se o *Python*, por ser uma linguagem *open-source*, com livre acesso à diversas bibliotecas, no caso do *Random-Forest* a *sklearn*, com funções já previamente implementadas. Dessa forma, montou-se uma floresta, definindo-se como hiperparâmetros, 500 árvores, cada uma com profundidade máxima igual ao número de variáveis.

Após a seleção dos hiperparâmetros, parte-se para a etapa de treinamento do algoritmo. É praxe que, dos dados disponíveis, se selecione 80% deles de forma aleatória para treinamento, sendo que os 20% restante ficam para teste do modelo.

A. Dados Sintéticos

Os históricos de amplitudes que compõem os dados de teste são então aplicados a todas as árvores geradas, obtendo como resultado final a classe que mais se repete, considerando os resultados individuais de cada árvore.

É relevante lembrar que o modelo é determinado a partir frequências das componentes harmônicas sensibilizadas pela falta. Dessa maneira, é de se esperar que cada componente tenha um peso distinto no modelo proposto. A Fig. 6 mostra as dez componentes harmônicas mais determinantes no modelo de classificação utilizado.

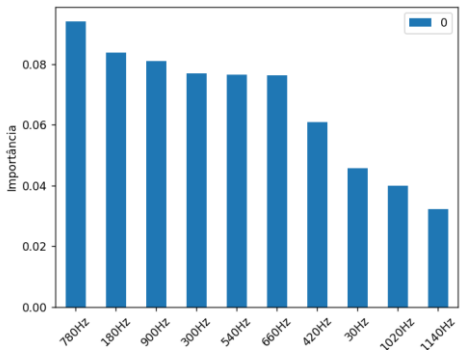


Fig. 6. Dez componentes harmônicas mais determinantes no modelo de classificação.

Os dados apresentados na figura anterior mostram que, dentre as frequências mais determinantes, as duas que mais contribuem para o modelo são as de 780 Hz e 180 Hz. Por outro lado, as componentes com menor peso no modelo são as de 2010 Hz e 60 Hz. Isso leva a crer que as componentes de 780 Hz e 180 Hz são sensibilizadas de maneira diferente por cada tipo de falta, enquanto nas de 2010 Hz e 60 Hz não haveria uma clara distinção entre as faltas de rotor e estator. Tal comportamento pode ser visualizado na Fig. 7, que apresenta o *flag* indicativo de falta em função da amostra para essas quatro frequências mencionadas.

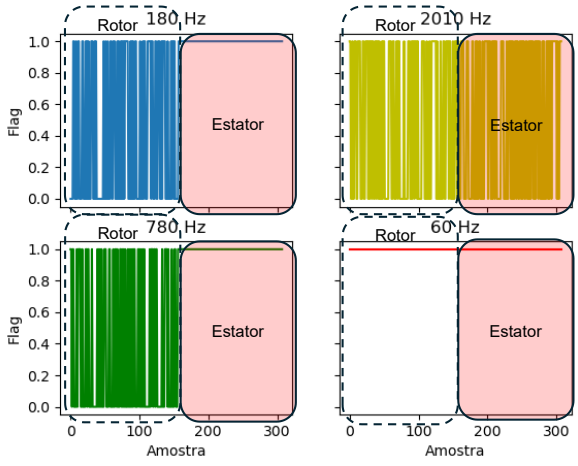


Fig. 7. Amostras sensibilizadas por faltas no rotor e no estator para as frequências de 180 Hz, 780 Hz, 2010 Hz e 60 Hz.

Nessas curvas, as últimas 150 amostras correspondem a faltas no estator, enquanto as demais a faltas no rotor. Assim, nota-se que, para faltas no estator, as frequências de 780 Hz e 180 Hz estão sempre sensibilizadas, o que não ocorre para faltas no rotor. Já para as frequências de 2010 Hz e 60 Hz, não é possível notar diferenças evidentes na sensibilidade destas frequências às diferentes faltas impostas.

Uma das melhores maneiras de visualizar o desempenho de um algoritmo de classificação é a partir da montagem de uma matriz de confusão. Ela permite, de maneira gráfica, comparar os resultados esperados (*True label*) com os resultados obtidos pelo algoritmo de classificação (*Predicted label*). A Fig. 8 mostra a matriz de confusão montada a partir dos resultados obtidos.

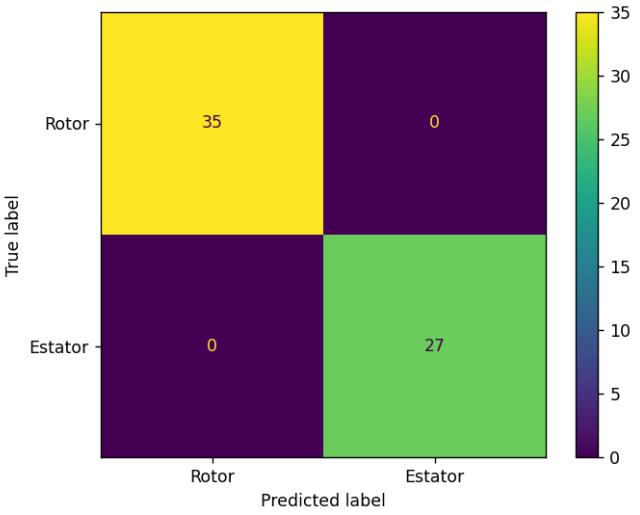


Fig. 8. Matriz de confusão dos dados sintéticos.

Os dados expostos na Fig. 8 indicam que o algoritmo foi capaz de diferenciar faltas de curto-circuito de espiras do rotor de curto-circuito de espiras do estator com base nas frequências das componentes harmônicas sensibilizadas pela falta. Contudo, é importante lembrar que, embora promissores, esses resultados foram obtidos a partir de dados simulados e que demandam ainda uma validação com dados experimentais.

B. Dados de Ensaios

Com o intuito de validar o modelo de classificação determinado, buscou-se empregá-lo a dados experimentais, obtidos de ensaios previamente realizados na bancada de testes existente no Laboratório de Máquinas e Acionamentos Elétricosda UFSC, utilizando a máquina de 8 polos sincronizada à rede de 60 Hz. Nesses ensaios, foram impostas faltas de curto-circuito de espiras do rotor e de curto-circuito de espiras do estator, separadamente. Com o intuito de preservar a máquina, apenas faltas correspondentes ao curto-circuito de 20% e 50% das espiras de um polo do rotor foram impostas, enquanto para o estator, impôs-se apenas as faltas que correspondem ao curto-circuito de 16,7% e 50% das espiras de um polo de uma das fases. Além disso, variou-se o ponto de operação, conforme a Tabela III

Tabela III PONTOS DE OPERAÇÃO AVALIADOS.

Ponto de Operação	Potência Aparente	Fator de Potência
P1	4,5 kVA	1
P2		0,8 indutivo
P3		0,5 indutivo
P4		0,3 indutivo
P5		0,8 capacitivo
P6		0,5 capacitivo
P7		0,3 capacitivo
P8	3 kVA	1
P9		0,8 indutivo
P10		0,5 indutivo
P11		0,3 indutivo
P12		0,8 capacitivo
P13		0,5 capacitivo
P14		0,3 capacitivo
P15	7 kVA	1
P16		0,8 indutivo
P17		0,5 indutivo
P18		0,3 indutivo
P19		0,8 capacitivo
P20		0,5 capacitivo
P21		0,3 capacitivo

Considerando os 21 pontos de operação descritos na tabela anterior, havia, então, 84 históricos de amplitudes a serem avaliados. Os resultados obtidos com a aplicação do modelo de classificação, determinado a partir dos dados sintéticos, aos dados provenientes de ensaios indicaram uma acurácia de 64,3%. A matriz de confusão dos resultados obtidos está ilustrada na Fig. 9.

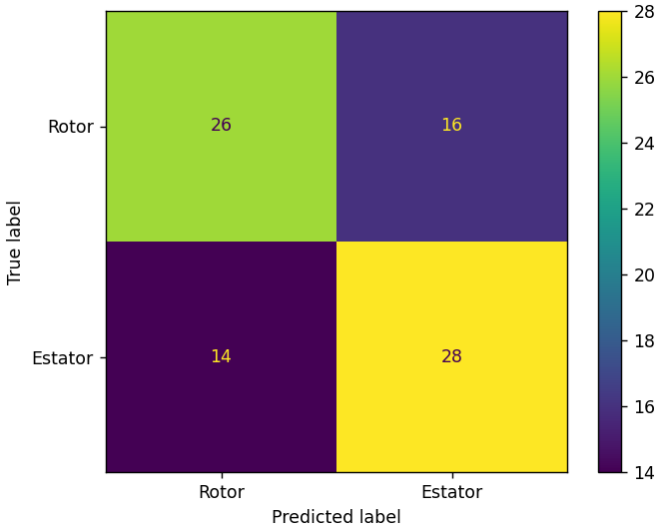


Fig. 9. Matriz de confusão dos dados experimentais.

Apesar do valor relativamente baixo de acurácia, abaixo de 70%, a análise da utilidade de um modelo de previsão passa pela determinação da probabilidade de classificar uma amostra corretamente a partir de um previsor de base. Nesse caso, uma vez que o conjunto de dados experimentais era composto por duas classes (falta no rotor e falta no estator) e pelo mesmo número de amostras correspondentes a cada classe (42), um previsor de base poderia ser um que classificasse todas as amostras como falta no rotor. Esse previsor de base apresentaria, então, uma acurácia de 50%. Portanto, a utilização do modelo determinado a partir dos

dados sintéticos foi capaz de aumentar a acurácia em quase 15%, comprovando a sua utilidade.

Além disso, buscou-se, nesse primeiro momento apenas verificar a possibilidade de utilização de um algoritmo desse tipo. Espera-se que, com o refinamento do algoritmo de detecção de faltas e com a melhor escolha das *features* (neste trabalho foram utilizadas as primeiras 135 harmônicas do sinal de fluxo magnético no entorno da máquina), o desempenho do algoritmo de classificação seja melhorado.

V. Conclusões

Este trabalho buscou determinar um modelo de classificação capaz de diferenciar faltas incipientes em geradores síncronos. Para isso, utilizou-se dados de uma detecção prévia de uma falta a partir da análise de dados históricos de amplitudes das componentes harmônicas do sinal de campo magnético externo dos geradores. Para a determinação do modelo, foi utilizado um método supervisionado de *machine learning*, o *Random-Forest*.

Contudo, métodos supervisionados necessitam de *datasets* rotulados para o seu treinamento e a consequente determinação do modelo. Esses *datasets*, usualmente, têm um grande volume de dados, muitas vezes com centenas, ou até milhares, de amostras. No caso deste trabalho, foi utilizada como amostra a sensibilização, ou não, de componentes harmônicas do campo magnético a uma possível falta. A sensibilização de uma componente harmônica por uma falta é definida com base na análise de históricos das amplitudes das componentes harmônicas do campo magnético. A geração de um desses históricos demanda diversas aquisições de campo magnético, tantas quanto o número de pontos desejados no histórico. Após as aquisições, é realizada, então, a FFT dos sinais adquiridos, para, em seguida, rastrear as componentes harmônicas de interesse e obter as amplitudes dessas componentes, para aí então formar o histórico de amplitudes. Todo esse processo torna a formação de históricos a partir de dados experimentais, com o intuito de utilizá-los como dados de treinamento, impraticável. Dessa maneira, optou-se por utilizar dados provenientes de simulações. Assim, foi possível utilizar mais de 300 históricos de amplitudes para formar um *dataset* de treinamento e teste do modelo de classificação.

Utilizando-se, então, do *dataset* formado, foi desenvolvido um modelo de classificação de faltas com base no algoritmo *Random-Forest*. Os resultados obtidos mostraram que foi possível classificar as faltas de curto-circuito entre espiras no enrolamento do rotor e no do estator com base nas frequências das componentes harmônicas sensibilizadas pela falta. O mesmo modelo foi então aplicado a dados obtido de ensaios prévios, alcançando uma acurácia de cerca de 65%, aumentando em 15% a acurácia de um modelo de classificação de base e comprovando, assim, a sua utilidade.

Contudo, é importante lembrar que esse tipo de algoritmo deve ser visto como uma ferramenta de auxílio, que visa a reduzir a carga de trabalho, mas não substitui a análise de um especialista.

Agradecimentos

Este trabalho foi motivado e parcialmente custeado pelos projetos regulados pela ANEEL e desenvolvidos no âmbito do Programa de P&D da ENGIE Brasil Energia e da Itá Energética S.A., PD-0403-0048/2019 e PD-00403-0057/2023, ambos intitulados "Equipamento não Invasivo para Detecção de Falhas em Geradores Síncronos através do Campo Magnético Externo", e realizado com apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Além disso, os autores gostariam de agradecer Msc. Luciano Bortoloto Antunes por ceder os dados dos ensaios para o teste do modelo.

Referências

[1] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO). Condition monitoring and diagnostics of machines — Vibration condition monitoring. ISO 13373, 2002.

[2] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO). Condition monitoring and diagnostics of machines - prognostics - Part 1: General guidelines. ISO 13381-1, 2004.

[3] AQTECH. Soluções para hidrelétricas | AQTech. Disponível em: https://www.aqtech.com/hidro. Acesso em: 23 abr. 2024.

[4] SKF. Sistemas de monitoramento da condição | SKF. Disponível em: https://www.skf.com/br/products/condition-monitoring-systems. Acesso em 23/04/2024.

[5] IRIS POWER LP. RFAII-R - Rotor Flux Monitoring (2 and 4 Pole Round Rotors). Disponível emhttps://irispower.com/products/rfaii-r/. Acesso em: 23 abr. 2024.

[6] L. O. S. Grillo, C. A. C. Wengerkievicz, H. F. dos Santos, N. J. Batistela, P. Kuo-Peng, N. Sadowski, T. K. Matsuo, C. L. de Souza, L. M. de Freitas, R. J. Nascimento. "Equipment for Monitoring Synchronous Generators Condition through External Magnetic Field Waveforms." **Journal Of Microwaves, Optoelectronics And Electromagnetic Applications**, [S.L.], v. 21, n. 4, p. 527-548, dez. 2022. FapUNIFESP (SciELO). Disponível em: http://dx.doi.org/10.1590/2179-10742022v21i4268102.

[7] H. F. dos Santos, M. Rigoni, R. de A. Elias, L. A. Feler, C. A. C. Wengerkievicz, N. J. Batistela, P. Kuo-Peng, J. P. A. Bastos, L. M. de Freitas. Non-invasive monitoring system of synchronous generator using external field. **Journal Of Microwaves, Optoelectronics And Electromagnetic Applications**, [S.L.], v. 16, n. 1, p. 70-89, mar. 2017. FapUNIFESP (SciELO). http://dx.doi.org/10.1590/2179-10742017v16i1872.

[8] M. Rigoni. **Desenvolvimento de um sistema de detecção e avaliação de faltas em geradores síncronos por meio do campo magnético externo.** 2014. Tese (Doutorado em Engenharia Elétrica) – Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Santa Catarina, Florianópolis, 2014.

[9] H. F. dos Santos. "**Detecção e avaliação de faltas em geradores síncronos por meio de históricos de harmônicos do campo magnético externo.**" 2021. Tese (Doutorado em Engenharia Elétrica) – Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Santa Catarina, Florianópolis, 2021.

[10] H. F. dos Santos. **Desenvolvimento de um equipamento para monitoração de geradores síncronos através do campo magnético externo.** Dissertação (Mestrado em Engenharia Elétrica) - Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Santa Catarina, 2016.

[11] C. S. P. Gameiro. **Desenvolvimento de metodologia para detecção automática de faltas em máquinas elétricas.** Dissertação (Mestrado em Engenharia Elétrica) - Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Santa Catarina, 2019.

[12] L. A. Feler. **Estudo e desenvolvimento de circuitos para condicionamento e medição de sinais de campo magnético.** Dissertação (Mestrado em Engenharia Elétrica) - Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Santa Catarina, 2019.

[13] M. Hoegen. **Estratégias de detecção de alterações no comportamento do campo magnético externo de geradores síncronos.** Dissertação (Mestrado em Engenharia Elétrica) - Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Santa Catarina, 2021.

[14] T. K. Ho. "Random decision forests". In **Proceedings of the 3rd International Conference on Document Analysis and Recognition**, Montreal, QC, Canada, 14–16 August 1995; Volume 1, pp. 278–282, doi:10.1109/ICDAR.1995.598994.

[15] L. Breiman, J. Friedman, C. J. Stone, R. A. Olshen. "**Classification and Regression Trees**", Monterey, CA, USA, Wadsworth and Brooks, 1984.

[16] HARP, **Harp Random Forests**, [Online], Disponível em: https://dsc-spidal.github.io/harp/docs/examples/rf/. Acesso em 23/04/2024

[17] B. Tierney. "Random Forest Machine Learning in R, Python and SQL - Part 1. **Toad World Blog.** Disponível em: https://blog.toadworld.com/2018/08/31/random-forest-machine-learning-in-r-python-and-sql-part-1. Acesso em: 23 abr. 2024.

[18] G. F. Martin Nascimento. **Optimisation des ressources et consommation des « smart buildings » en vue de l'efficacité énergétique**. 2023. 205 p. Tese elaborada em regime de cotutela entre a Universidade Federal de Santa Catarina e a Université Grenoble Alpes, Grenoble, 2023. Disponível em: https://bu.ufsc.br/teses/PEEL2100-T.pdf.

[19] G. F. Martin Nascimento, F. Wurtz, P. Kuo-Peng, B. Delinchant, N. J. Batistela. Outlier Detection in Buildings' Power Consumption Data Using Forecast Error. **Energies**, [S.L.], v. 14, n. 24, p. 8325, 10 dez. 2021. MDPI AG. http://dx.doi.org/10.3390/en14248325.

[20] D. G. Aurich. **Modelagem analítica para simulação de interações entre grandezas eletromagnéticas em geradores síncronos**. 2022. 88 f. TCC (Graduação) - Curso de Engenharia Elétrica, Centro Tecnológico, Universidade Federal de Santa Catarina, Florianópolis, 2022.

[21] L. O. S. Grillo, C. A. C. Wengerkievicz, N. J. Batistela, P. Kuo-Peng; N. Sadowski, L. M. de Freitas. A Method for Statistical Processing of Magnetic Field Sensor Signals for Non-Invasive Condition Monitoring of Synchronous Generators. **Sensors**, [S.L.], v. 22, n. 22, p. 8631, 9 nov. 2022. MDPI AG. http://dx.doi.org/10.3390/s22228631