

## Análise e detecção de faltas de alta impedância em sistemas de distribuição com a utilização de classificadores por aprendizado supervisionado

Gustavo da Silva Rocha Araújo, Thiago Americano do Brasil, Bernardo Almeida Vasconcellos de Souza, Jonathan Nogueira Gois e João Amin Moor Neto

\* Centro Federal de Educação Tecnológica Celso Suckow da Fonseca  
Rio de Janeiro, Brasil (autor correspondente: thiago.brasil@cefet-rj.br).

---

Abstract: The occurrence of high impedance faults (HIF) in primary distribution networks represents a danger to the safety of people, equipment, and animals. However, protection devices close to the distribution network are not able to be sensitized by this type of defect, in most cases. This work presents an integrated strategy for the classification and detection of HIF, based on the use of classifiers based on supervised learning algorithms, such as K-nearest neighbors, Support Vector Machines and Ensemble. An improved fault model was used to emulate random behaviors, and especially intermittency for high impedance faults. The residual current was monitored, and its characteristics were extracted using the Short-Time Fourier Transform (STFT). A temporal consistency logic was applied to the detection stage. The functioning of the presented algorithm was carried out through several simulations in a 20 kV distribution system.

Resumo: A ocorrência de faltas de alta impedância (FAI) em redes de distribuição primárias representa um perigo para a segurança das pessoas, equipamentos e animais. No entanto, dispositivos de proteção junto a rede de distribuição não são capazes de serem sensibilizados por esse tipo de defeito, na maioria das vezes. Este trabalho apresenta uma estratégia integrada de classificação e detecção de FAI, baseado no uso de classificadores baseados em algoritmos de aprendizado supervisionado, tais como *K-nearest neighbors*, *Support Vector Machines* e *Ensemble*. Um modelo de falta aprimorado foi usado para emular comportamentos de aleatoriedade, e especialmente intermitência de faltas de alta impedância. A corrente residual foi monitorada, e a extração de suas características foi realizada a partir da *Short-Time Fourier Transform (STFT)*. Uma lógica de consistência temporal foi aplicada ao estágio de detecção. O funcionamento do algoritmo apresentado foi realizado através de diversas simulações em um sistema de distribuição de 20 kV.

**Keywords:** High Impedance Fault; KNN; Detection; STFT; Machine Learning

**Palavras-chaves:** Falta de Alta Impedância; KNN; Detecção; STFT; Aprendizado Supervisionado.

---

### 1. INTRODUÇÃO

Faltas monofásicas são os defeitos mais comuns em redes primárias de distribuição. A partir do rompimento de um cabo energizado e de seu respectivo contato com uma determinada superfície, incorre-se na presença de arco elétrico, o que expõe vidas e patrimônio público e privado ao perigo de descarga elétrica. A falta de alta impedância (FAI) é geralmente proveniente do rompimento de um condutor energizado e seu posterior contato elétrico com uma superfície de alta resistividade, como o solo, o asfalto ou a grama. Por conta dessa característica, a corrente drenada durante a sua ocorrência geralmente apresenta baixa magnitude, acarretando que as configurações clássicas de proteção, como relés de sobrecorrente e fusíveis, presentes ao longo do trecho do alimentador, não sejam sensibilizados e consequentemente não promovam a extinção da falta (Michael Aucoin, 1996; Gautam and Brahma, 2013). Uma metodologia para se aprimorar a sensibilidade do sistema de proteção convencional é a utilização de faltas de alta impedância ensaiadas ao longo

do sistema de distribuição (Elkalashy *et al.*, 2007). No entanto, estes testes vêm sendo cada vez mais evitados, visto que um risco adicional acaba sendo imposto, desnecessariamente, à rede elétrica. Em suma, uma modelagem matemática que represente com fidelidade os mais diversos tipos de FAI mostra-se primordial no estudo da classificação e, posterior, detecção destas faltas, sem a necessidade do uso de faltas ensaiadas (Brasil, Gois and Neto, 2021). Partindo desta premissa, são encontrados na literatura diversas implementações matemáticas de FAI do tipo *black-box* (Khakpour *et al.*, 2015), as quais se mostram bastante efetivas, já que conseguem emular as capacidades de interrupção (ignição e extinção) do arco, bem como seu comportamento dinâmico. Tais modelos são baseados nos propostos por (Cassie, 1939), (Mayr, 1943) e (Schavemaker and Van Der Sluis, 2000). Entretanto, esta modelagem leva em consideração apenas o balanço térmico estabelecido na condição de falta, não prevendo possíveis períodos de instabilidades do arco elétrico.

Recentemente, um modelo aprimorado de arco foi proposto, sendo posteriormente utilizado em outros trabalhos (Brasil, Gois and Neto, 2021). Esta implementação inclui um fator randômico no comportamento randômico do arco elétrico, o qual possibilita reproduzir com maior fidelidade os períodos de instabilidade transitória das FAI, onde fatores tais como as condições do solo no qual é realizado o contato com o condutor rompido, sua temperatura e umidade não são adequadamente controlados (Wei *et al.*, 2020).

Em relação aos esquemas de detecção de faltas de alta impedância em redes primárias de distribuição, um conjunto de técnicas está disponível na literatura. (Soheili, Sadeh and Bakhshi, 2018) apresentou um algoritmo de detecção de FAI baseado no monitoramento de conteúdo harmônico da forma de onda da corrente, tomando como caso base uma rede de distribuição de 13 barras (Schneider *et al.*, 2018). (Costa *et al.*, 2015) discutiram a utilização da Transformada *Wavelet* para detecção de FAI. Um esquema de detecção não linear baseado em características de tensão-corrente foi proposto por (Wang, Geng and Dong, 2016). (Gautam and Brahma, 2013) propuseram uma técnica baseada em matemática morfológica para identificação de FAI. Técnicas inteligentes de reconhecimento de padrões que usam recursos eficazes para detecção de falhas também estão disponíveis (Chaitanya, Yadav and Pazoki, 2020). Um esquema de classificação baseado em rede neural artificial foi proposto por (Michalik *et al.*, 2006), utilizando a Transformada *Wavelet* para a extração das características espectrais das FAI. Recentemente, um modelo de decomposição empírica do arco elétrico foi proposto por (Lala and Karmakar, 2020), utilizando-se na etapa de classificação do algoritmo *k-nearest neighbour* (KNN), mostrando-se consistente na detecção de FAI.

Em geral, classificadores baseados em aprendizado supervisionado requerem menor esforço computacional quando comparados com demais classificadores, tais como o baseado em árvores de decisão (Samantaray, Tripathy and Dash, 2009), algoritmo genético (Sedighi, Haghifam and Malik, 2005), classificador baseado em sistema de tomada de decisão (Don Russell, 1990), classificadores *Bayesianos* (Sedighi *et al.*, 2005) e classificador baseado em regras (Tai and Chen, 2008).

Embora os métodos mencionados apresentem resultados de simulação bastante promissores, duas questões ainda permanecem obscuras. Primeiramente, praticamente nenhum estudo de detecção de FAI a partir do uso de classificadores com aprendizado supervisionado levou em consideração a detecção de faltas em condições de maior intermitência, randomicidade e alta variabilidade estatística. Deste modo, valores de acurácia de classificação irrealisticamente altos foram obtidos por diversos autores. Diante dos diversos classificadores de aprendizado supervisionado apresentados na literatura, o segundo questionamento trata de qual destas soluções é capaz de melhor se adequar a correta classificação de FAI que compreendem uma maior variabilidade que as utilizadas em outros trabalhos (Wei *et al.*, 2020; Brasil, Gois and Neto, 2021).

Tendo em vista estas lacunas presentes na literatura, o intuito desse trabalho é o de promover uma melhor solução de

classificação de FAI a partir do uso do modelo de arco aprimorado proposto por (Wei *et al.*, 2020) durante a extensa etapa de treinamento dos classificadores. Para aumentar a capacidade de generalização dos algoritmos na classificação de faltas de alta impedância, 3600 eventos transientes foram realizados ao longo de várias simulações em um sistema de distribuição de 20 kV, onde aproximadamente metade deles consistiu em FAI com os mais diversos comportamentos. De posse dos dados de treinamento, a acurácia de classificação a partir de diversos classificadores, dentre eles KNN (*Fine, Medium, Coarse, Cosine e Weighted*), *Support Vector Machines* (*Fine Gaussian e Medium Gaussian*) e *Ensemble* (*Boosted Trees e Subspace Discriminant*) foi estabelecida.

Um esquema de detecção de FAI usando *Short-Time Fourier Transform* (STFT) para extração de características e consistência temporal a partir de janela deslizante baseado em Brasil *et al.* (Brasil, Gois and Neto, 2021) foi apresentado, visando melhor distinguir as faltas de outros eventos transientes convencionais, aumentando a segurança e confiabilidade do sistema de detecção.

## 2. MODELAGEM E SIMULAÇÃO DA FALTA DE ALTA IMPEDÂNCIA

### 2.1 Modelo da FAI

A fim de aprimorar o modelo de classificação da FAI, um melhor entendimento do comportamento da corrente se faz necessário. Como já mencionado, a principal característica física da falta de alta impedância é a ocorrência de arco elétrico (Ghaderi, Ginn and Mohammadpour, 2017). Além de sua baixa magnitude de corrente e do arco elétrico associado, as FAI geralmente apresentam as seguintes características elétricas: assimetria no formato de onda da corrente (Sultan, Swift and Feddrchuk, 1994), *buildup e shoulder* (Ghaderi *et al.*, 2015), corrente não estacionária (Ghaderi *et al.*, 2015), não linearidade (Wang, Geng and Dong, 2016), baixa frequência das componentes harmônicas na corrente medida e no espectro de tensão (Emanuel *et al.*, 1990), componentes harmônicas de alta frequência medidas no espectro de corrente (Ghaderi, Ginn and Mohammadpour, 2017), intermitência do arco elétrico (Hamel, Gaudreau and Côté, 2004) e randomicidade (Wei *et al.*, 2020). A maioria dos modelos de arco apresentados na literatura normalmente considera apenas ignições de arco estáveis. Desconsiderando a maioria das características inerentes as FAI mencionadas acima (principalmente intermitência e aleatoriedade), é improvável obter um sistema de classificação robusto de faltas de alta impedância. Este problema pode ser corrigido inserindo um coeficiente aleatório (RC) para o modelo Mayr convencional:

$$\frac{1}{g_{arco}} \left( \frac{dg_{arco}}{dt} \right) = \frac{1}{\tau_0} \left( \frac{u_f i_f}{P_{perdas0} RC} - 1 \right), \quad (1)$$

onde  $g_{arco}$  se refere a condutância equivalente do arco ( $1/\Omega$ ),  $u_f$  e  $i_f$  representam a tensão (V) e a corrente (A) do arco respectivamente,  $\tau_0$  diz respeito a constante de tempo do arco estável (s) e  $P_{perdas0}$  retrata a potência dissipada durante o evento (W). De acordo com (Wei *et al.*, 2020), o coeficiente de aleatoriedade é dado a partir de:

$$\begin{cases} RC = RC_{ign}, & \text{se } RC_{sel} = 0 \\ RC = RC_{ext}, & \text{se } RC_{sel} = 1 \end{cases} \quad (2)$$

onde  $RC_{sel}$ ,  $RC_{ign}$  e  $RC_{ext}$  são variáveis aleatórias.  $RC_{sel}$  é determinado a partir de uma probabilidade que ocorre a cada  $xT$  (s) onde ( $x \in [0.25, 0.75]$ ), e  $T$  é estabelecido como 0.016s, visto que a frequência da rede considerada é 60 Hz.  $RC_{ign}$  e  $RC_{ext}$  representam a potência dissipada na ignição e extinção, respectivamente e são utilizadas para implementar variações limitadas no coeficiente de potência dissipada anterior, ou seja,  $P_{perdas} = P_{perdas0}RC$ . Essas atualizações nos coeficientes podem ser definidas como:

$$\begin{cases} RC_{ign} \in \left( k_1, \frac{k_2 P_{perdasB}}{P_{perdas0}} \right), \\ RC_{ext} \in \left( \frac{k_3 P_{perdasB}}{P_{perdas}}, \frac{k_4 P_{perdasB}}{P_{perdas0}} \right), \end{cases} \quad (3)$$

onde  $P_{perdasB}$  significa a potência dissipada limitante, a qual determina o coeficiente de potência dissipada a partir do qual o arco é extinto e cujo valor precisa ser determinado a partir de simulação. Em suma, a ocorrência do arco é dada quando  $P_{perdas} < P_{perdasB}$  e sua respectiva extinção quando  $P_{perdas} > P_{perdasB}$ . A aleatoriedade, bem como a intermitência, é definida a partir da escolha dos parâmetros  $k_1$ ,  $k_2$ ,  $k_3$  e  $k_4$ . O modelo de arco de Mayr modificado está representado na Fig. 1.  $R_{arco}$  denota a resistência do arco (inverso da condutância) e é responsável pelo comportamento dinâmico (com exceção do *buildup* e *shoulder*). Já  $R_C$  representa a resistência de contato do condutor com a superfície em que ele se projeta, sendo o responsável pelo efeito do *buildup* e *shoulder*. O intervalo a partir do qual ocorre a estabilidade do arco, bem como sua duração, presença/ausência de intermitência/aleatoriedade e a amplitude geral da falta são variáveis inerentes ao modelo de arco melhorado, o qual pode ser arbitrado ou disposto randomicamente para cada FAI emulada, de modo a simular o maior número de comportamentos possíveis na corrente de falta.

## 2.2 Tipos de Falta

As características da FAI podem variar de forma considerável. (Wei *et al.*, 2020) fez a divisão das FAI da maneira a seguir:

**Tipo-1** – Instantaneamente após a ocorrência da falta, há instabilidade no período de formação do arco, porém ele é capaz de se estabilizar de forma célere. Dependendo do tipo de superfície na qual se lançar, as distorções da falta e da amplitude da corrente se apresentem com características típicas. A Fig. 2 representa as características de três FAI do Tipo-1 recriadas em ambiente SimPowerSystems (MATLAB), com suas respectivas formas de onda utilizando o modelo de arco aprimorado.

**Tipo-2** - Esse tipo de falta possui um período de instabilidade do arco elétrico superior quando comparado ao Tipo-1. Além do mais, FAI que apresentem apenas oscilação aleatória são normalmente classificadas como Tipo-2A, enquanto faltas que induzam tanto o fenômeno de oscilação aleatória quanto de intermitência, são classificadas como Tipo-2B (Fig. 3).

**Tipo-3** – Para alguns tipos de falta, os valores mais expressivos de impedância da superfície de contato são determinantes. Essa particularidade pode fazer com que FAI do Tipo-3 se assemelhem a um leve desbalanço natural da rede, visto que a corrente drenada durante sua ocorrência apresenta baixa magnitude, o que ocasionalmente leva a uma maior dificuldade de detecção. Aliado a isso, esse tipo de falta pode apresentar uma distorção em alta frequência, caracterizado por ruído gaussiano branco (Fig. 3).

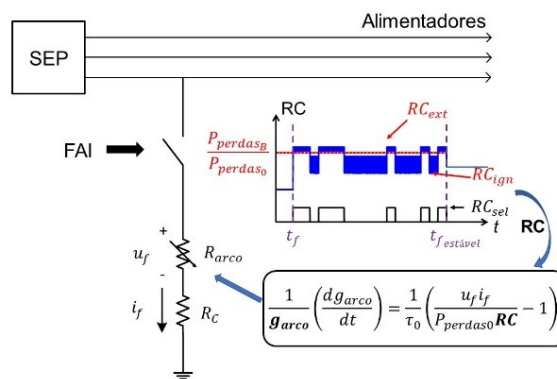


Fig. 1 Modelo aprimorado de FAI.

## 2.3 Método de detecção das FAI

O princípio de classificação da falta de alta impedância no presente trabalho se dá em três etapas: aquisição de dados, extração das componentes características e seleção, e posterior classificação a partir de técnicas de aprendizado supervisionado.

a) **Aquisição de dados:** A forma de onda da FAI também apresenta as características básicas observadas em um sistema de corrente alternada: componentes fundamental e harmônicas. Além disso, as características das FAI tendem a serem variantes no tempo. Logo, uma maneira de ser realizar a sua detecção é a partir da existência de conteúdo espectral relevante consistentemente durante um dado período de amostragem. Os transformadores de corrente (TC) são os equipamentos convencionais utilizados para a medição da corrente na saída dos alimentadores. De acordo com (Chen *et al.*, 2016), os TC podem ser utilizados, considerando que as componentes harmônicas relevantes estejam em espectros inferiores a 1.5 kHz, frequência a partir da qual a atenuação imposta por este tipo de equipamento passa a ser um fator limitante.

b) **Extração e seleção das características:** A fim de projetar apropriadamente um classificador de reconhecimento de padrões, deve-se previamente desenvolver uma ferramenta para a extração/seleção de características, que permita discriminar efetivamente a qual classe pertence um dado sinal (Sedighi *et al.*, 2005). Nesse trabalho, a *Short-Time Fourier Transform* (STFT) foi utilizada na etapa de extração de características. Essa transformada é responsável por subdividir o sinal em tamanhos iguais e, posteriormente, calcular a Transformada de Fourier tradicional em cada um desses segmentos individualmente. A STFT é dada por:

$$STFT(t, f) = \int_{-\infty}^{+\infty} [s(u) \text{win}(u - t)] e^{-j2\pi fu} du, \quad (4)$$

onde  $\text{win}(t)$  é a função de janelamento e  $s(t)$  é o sinal de entrada.

Devido à existência de vários processos estocásticos não correlacionados gerados pelas características aleatórias das FAI, torna-se impossível assumir algum tipo de tendência ou comportamento anterior sem uma análise mais aprofundada do conteúdo espectral destas faltas. Desse modo, a STFT apresenta-se como a melhor ferramenta para extração de informação relevante. Dessa maneira, a seleção de características é realizada considerando o conteúdo espectral de cada tipo de corrente de falta (Tipo-1, Tipo-2 e Tipo-3), iniciando em 0 Hz (componente CC) até a ordem de 25º harmônico (1.5 kHz).

c) *Etapa de classificação*: A rede de distribuição abarca uma ampla variedade de eventos inerentes à sua operação, tais como inserção/remoção de: banco de capacitores, motores trifásicos e cargas de diversos fatores de potência em pontos distintos do alimentador. Consequentemente, é necessário não só ter um sistema capaz de detectar FAI, mas também de diferenciar os eventos supracitados das faltas de alta impedância. Neste contexto, classificadores baseados em algoritmos de aprendizado supervisionado vêm sendo utilizados no suporte a detecção de FAI. (Sarlak and Shahrtash, 2008) propôs o uso de *Support Vector Machines* na classificação de faltas de alta impedância. (Lala and Karmakar, 2020) apresentou as vantagens de se utilizar o KNN para tal. Da mesma forma, os algoritmos *Ensemble*, baseados em árvores de decisão, foram sugeridos por alguns autores, tais como (Gomes, Ozansoy and Ulhaq, 2018) e (Le and Yao, 2019).

Além disto, não é consenso na literatura qual seria a melhor escolha de tamanho da janela da STFT, bem como quantas e quais características devem ser levadas em consideração na etapa de extração de sinal, como demonstrado por (Brasil, Gois and Neto, 2021). A fim de treinar os classificadores KNN, SVM e *Ensemble*, as características extraídas via STFT foram inseridas no *Classification Learner Toolbox* no MATLAB. Mais à frente, a Seção 4 tem o objetivo de avaliar a performance dos classificadores supracitados a partir da etapa de treinamento a ser apresentada no estudo de caso a seguir.

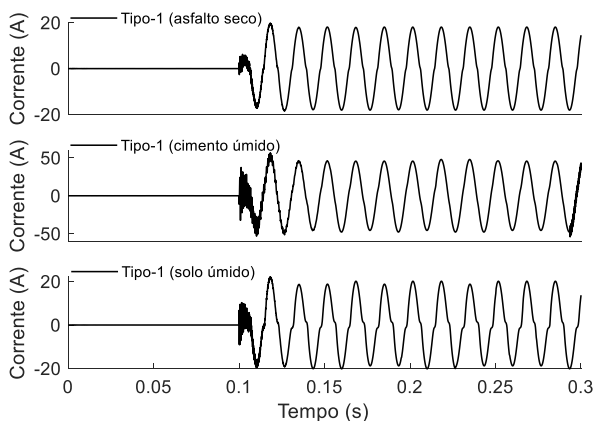


Fig. 2 Características simuladas das FAI do Tipo-1.

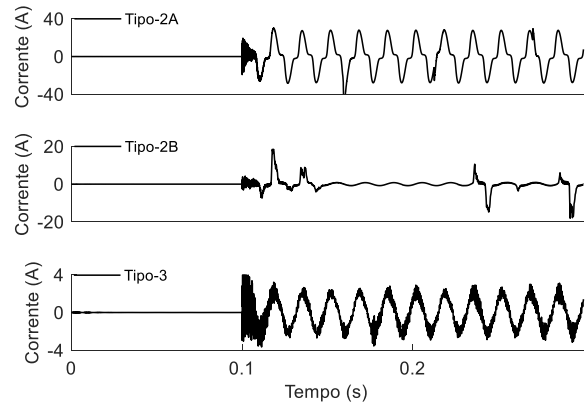


Fig. 3 Características simuladas das FAI do Tipo-2A, Tipo-2B e Tipo-3.

### 3. ESTUDO DE CASO

#### 3.1 Sistema de distribuição teste de 20 kV

O sistema da Fig. 4 foi baseado na topologia de rede proposta por Elkalashy et al. (Elkalashy *et al.*, 2007) e foi recriado em ambiente SimPowerSystems/MATLAB. A rede utilizada levou em consideração uma configuração típica de um sistema de distribuição, que é a utilização de um transformador abaixador de tensão, onde o lado de média tensão não apresenta ligação para o neutro. A componente residual da corrente é dada como a soma fasorial das correntes de um dado sistema. Para o modelo em questão, pode-se considerar a corrente residual como:

$$i_r = i_A + i_B + i_C, \quad (5)$$

e pode ser medida no início do alimentador. A partir da configuração proposta, pode-se supor que a medição da corrente residual realizada no lado de média tensão possui pouca ou nenhuma influência do carregamento presente no lado de baixa tensão, ou seja, para a operação normal do sistema, a corrente residual tende a ser nula. Para outros casos, tais como transitórios presentes em inserção de novas cargas, partida de motores, energização de transformadores e, principalmente, a ocorrência de FAI, a corrente residual medida assume valores diferentes de zero. Deste modo, faz-se necessário que o sistema de classificação seja capaz de diferenciar tais eventos.

#### 3.2 Treinamento

O treinamento foi realizado a partir de simulações, com todos os eventos ocorrendo no Alimentador A1 de 20 kV (média tensão). Nove eventualidades foram consideradas como acontecimentos transitórios que não configuram FAI no referido alimentador: inserção de cargas monofásicas, bifásicas e trifásicas, chaveamento de banco de capacitores monofásicos e trifásicos, energização de transformadores trifásicos sob carga leve (a fim de simular a corrente de magnetização deste), partida de motores monofásicos e trifásicos e variação do carregamento ao final do trecho do alimentador (no lado de baixa tensão). Durante as simulações

de FAI, considerou-se também a possibilidade de existência de certa magnitude de ruído gaussiano, cujo intervalo utilizado foi de  $P_{ruído} = [-15, -5]$  dB. O resumo de tais eventos, bem como seus respectivos intervalos, é apresentado na Tabela 1. A ocorrência de cada um destes eventos tem como consequência o desbalanço transitório das correntes no sistema modelado e será medida no ponto de medição indicado na Fig. 4. Os eventos foram limitados a ocorrerem um de cada vez, e apenas no Alimentador 1, em qualquer ponto do mesmo (de 0 a 45 km). O tempo definido para cada simulação foi de 5s, com uma média de ocorrência de 9 eventos (estipulados de forma randômica) por simulação e um total de 3600 eventos (400 simulações foram realizadas). Cada simulação realizou a extração de características da corrente residual a partir da STFT, dividindo-as também, de forma binária, entre eventos de falta ou não falta, associando cada uma das classificações realizadas a uma janela da STFT.

### 3.3 Detecção

Após obter o modelo de previsão com base no treinamento proposto na seção anterior, um sistema de detecção foi desenvolvido considerando uma suposição da probabilidade da ocorrência de falta. De maneira geral, a corrente residual é medida e seu espectro de características é extraído utilizando uma janela de tamanho N. Consequentemente, o estado de classificação para cada janela pode ser 0 (não-falta) e 1 (FAI). Com o objetivo de otimizar a etapa de detecção, uma estratégia de janela deslizante foi utilizada, onde cada janela analisada apresenta uma sobreposição de 15% sobre a janela anterior. Uma estimativa da ocorrência da falta pode ser calculada da seguinte maneira:

$$Estimativa = \sum_{k=1}^{k_{janela}} \frac{c(k)}{k_{janela}} \quad (6)$$

onde  $c(k)$  é o k-ésimo termo do estado de classificação da janela sobreposta  $k_{janela}$ , esta que é 30% maior que a janela original de tamanho N da STFT.

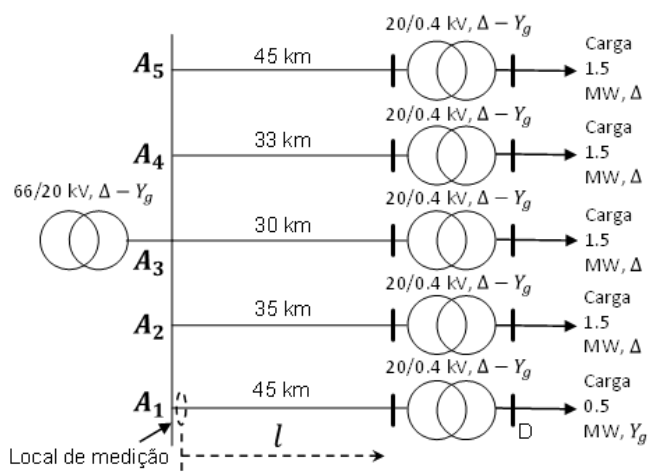


Fig. 4 Topologia da rede do sistema teste de 20 kV.

Tabela 1 Tipos de eventos simulados e seus respectivos intervalos.

Evento	Intervalo
Carga monofásica	$P_l = [10, 100]$ kW $Q_l = [-5, 5]$ kVAr
Capacitor monofásico	$Q_l = [-200, -20]$ kVAr
Motor monofásico	$T = [3.1, 19.5]$ N.m
Carga bifásica	$P_l = [10, 100]$ kW $Q_l = [-5, 5]$ kVAr
Carga trifásica	$P_l = [10, 100]$ kW $Q_l = [-5, 5]$ kVAr
Capacitor trifásico	$Q_l = [-200, -20]$ kVAr
Energização de transformadores trifásicos (carga leve)	$P_l = [10, 50]$ kW $Q_l = [-5, 5]$ kVAr
Motor trifásico	$T = [2.1, 280]$ N.m
Varição da carga na baixa tensão (D)	$P_l = [10, 50]$ kW
Falta de Alta Impedância	Tabela 2

Para evitar sinais espúrios devido aos eventos que possuam características espectrais semelhantes aos das FAI, uma consistência temporal de cada evento foi verificada neste trabalho. A variável TRIP\_STATE, que representa o alarme de detecção, apenas sofre variação em seu estado lógico caso a consistência temporal seja superior a 9N, como é representado na Fig. 5. As faltas do tipo Tipo-2B tendem a provocar maior flutuação na detecção devido à sua característica intrínseca de elevada intermitência e randomicidade. Como a duração do intervalo de intermitência tende a ser inferior a consistência temporal adotada neste trabalho, o sistema de detecção atua de maneira adequada. A Fig. 6 mostra o fluxograma do algoritmo proposto para detecção de HIF.

Tabela 2 Características das FAI simuladas.

FAI	$P_{perdas0}$	$P_{perdasB}$	$k_1$	$k_2$	$k_3$	$k_4$	$R_c(k\Omega)$
T1 (A)	7.5-12.5	$P_{perdas0}$	0.75	1.25	-	-	0.5-1.0
T1 (C)	1.0-3.0	$P_{perdas0}$	0.85	1.00	-	-	0.5-1.0
T1 (S)	2.0-4.0	$P_{perdas0}$	0.35	1.00	-	-	0.17-0.33
T2A	10	50	0.01	0.05	1.00	1.05	0.17-0.55
T2B	1	40	0.5	0.75	1	6	0.25-0.55
T3	1	4	0.85	1		1.25	2.5-10

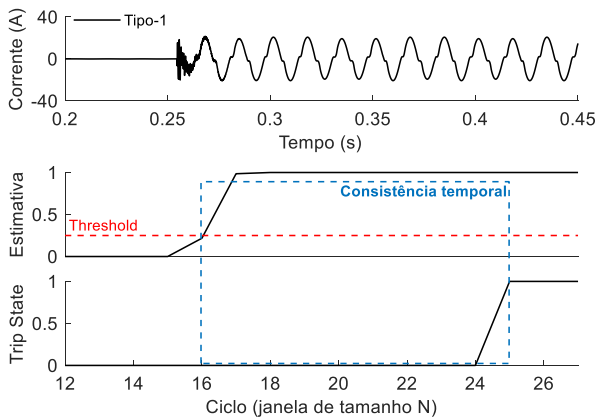


Fig. 5 Exemplo de detecção considerando consistência temporal de  $9N$  e uma FAI do Tipo-1.

#### 4. COMPARAÇÃO ENTRE CLASSIFICADORES

Visando analisar os resultados obtidos das simulações, a STFT foi utilizada na corrente residual considerando três tipos diferentes para o tamanho de janela de aquisição ( $N$ ): 833 pontos (16,67 ms ou um ciclo de 60 Hz), 2500 pontos (50 ms ou três ciclos de 60 Hz) e 5000 pontos (100 ms ou seis ciclos de 60 Hz). De modo correlato, o vetor de seleção das características também variou de acordo com o *range* entre 0 e 1.5 kHz.

Dessa maneira, o objetivo foi o de realizar a análise comparativa de desempenho dos classificadores tendo como parâmetros as variáveis mencionadas acima. Os classificadores estabelecidos foram: *KNN (Fine, Medium, Coarse, Cosine e Weighted)*, *Support Vector Machines (Fine Gaussian e Medium Gaussian)* e *Gradient Boosting (Boosted Trees e Subspace Discriminant)*. Tais algoritmos tem a capacidade de realizar de maneira adequada frente à classificação binária (falta ou não-falta) (Al-Shaher, Sabry and Saleh, 2003; Hastie, Tibshirani and Friedman, 2009; Lala and Karmakar, 2019). Em relação a teoria por trás de cada um destes tipos de classificadores, mais informações podem ser obtidas para o *Support Vector Machines* em (Al-Shaher, Sabry and Saleh, 2003; Livani and Evrenosoglu, 2013; Brasil, Gois and Neto, 2021) e para o *Ensemble* em (Hastie, Tibshirani and Friedman, 2009; Gomes, Ozansoy and Ulhaq, 2018). Para o *KNN*, a referência (Lala and Karmakar, 2020) pode ser proveitosa.

A obtenção do modelo leva em consideração a definição prévia dos parâmetros de cada. O particionamento via validação cruzada foi realizado a partir de um *kfold* de 10, o que permitiu avaliar a capacidade de generalização de cada um dos modelos de classificador obtidos, prevenindo contra a ocorrência de *overfitting*, a partir da estimação da acurácia de cada partição individualmente.

Como pode ser visualizado na Fig. 7, o modelo capaz de prever a ocorrência de FAI de forma mais efetiva foi o dotado de classificação do tipo *Weighted KNN* e uma etapa de extração de características via STFT com  $N = 833$  pontos. A

Tabela 4 demonstra que este classificador apresentou sua máxima efetividade quando foram utilizadas 24 características espectrais.

#### 5. RESULTADOS DE DETECÇÃO

##### 5.1 Detecção por validação cruzada

A fim de verificar a confiabilidade do sistema de detecção, um esquema de validação cruzada foi criado. A ocorrência de cada fenômeno transitório foi dada de maneira individual, isto é, cada novo evento se inicia a partir do término do evento anterior. Foram realizadas 20 simulações aleatórias com duração total de 5s cada, totalizando 153 diferentes eventos. Esta última etapa visa avaliar a robustez do algoritmo de detecção em relação à existência de falsos positivos (detecção errônea de FAI) e falsos negativos (recusa em detectar uma FAI). Cada evento ocorreu em um trecho aleatório ao longo do alimentador A1 que possui 45 km de extensão. Além disso, a escolha da fase de ocorrência do evento (fases a, b ou c) também foi efetuada de maneira randômica.

A partir dos resultados da Tabela 4, onde a diagonal principal da matriz diz respeito aos eventos detectados corretamente, foi possível verificar a detecção de todas as faltas do Tipo-1 e Tipo-3, onde apenas uma falta do Tipo-2B não foi detectada, representando assim um percentual de detecção de 91.67% (11 de 12). A maior dificuldade em caracterizar faltas do Tipo-2B pode estar associada às características extremamente oscilatórias deste tipo de falta, principalmente o fenômeno da intermitência. Para os demais tipos de falta, o índice de detecção foi de 100%, apresentando um percentual global de detecção de 98.55%.

Um aspecto importante da estratégia de detecção proposta neste trabalho foi a de manter uma taxa de detecção global superior a 98%, além de propiciar que nenhum evento transitório fosse confundido com uma FAI. Em outras palavras, a taxa de falsos positivos foi nula. Isto prova que a extração/seleção de características usadas na STFT foi corretamente conduzida e que a capacidade de generalização do algoritmo foi alcançada.

#### 6. CONCLUSÃO

Este trabalho mostrou que o processo de seleção dos recursos a serem extraídos e utilizados é crucial para a melhoria da capacidade de generalização dos algoritmos de aprendizado supervisionado. Sendo assim, espera-se que o esquema de detecção proposto possa ser eficaz no monitoramento de ocorrências de FAI em tempo real. Faltas Tipo-1, Tipo-2A e Tipo-3 foram efetivamente detectadas com a estratégia proposta nesse trabalho. Por outro lado, as faltas do Tipo-2B mostram-se problemáticas, especialmente caso apresentem intervalos de intermitência superiores a  $9N$  (9 ciclos de 60 Hz), o que pode levar à uma instabilidade no sinal de saída do TRIP\_STATE. Na prática, um aumento da consistência temporal pode ser facilmente realizado. Esta estratégia pode diminuir a presença de falsos positivos durante a ocorrência de transitórios inerentes ao sistema de distribuição.

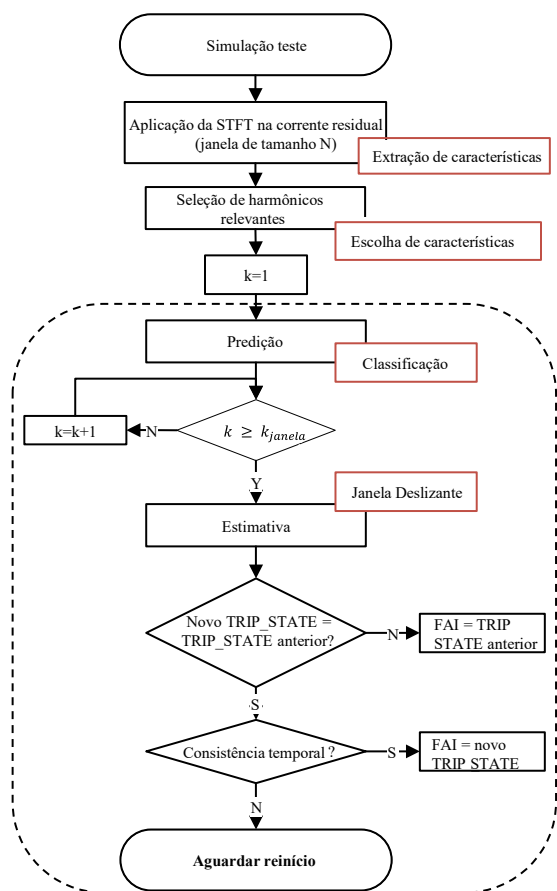


Fig. 6 Fluxograma do algoritmo de detecção de FAI.

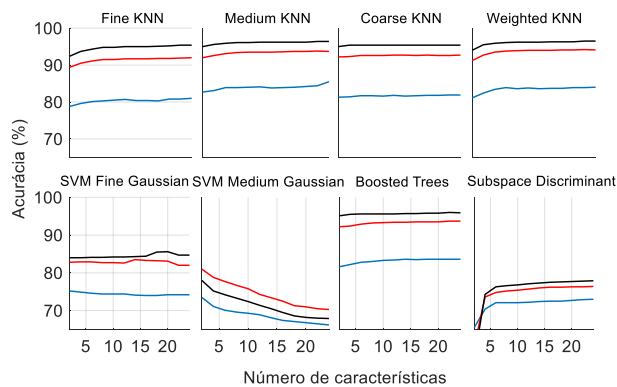


Fig. 7 Acurácia de diversos classificadores, considerando o número de características e a janela N de tamanho: 5000 pontos (azul), 2500 pontos (vermelho) e 833 pontos (preto).

Tabela 3 Acurácias de classificação do *Weighted KNN*.

Dim	Características selecionadas (harmônicos)	N=833	N=2500	N=5000
2	1,3	94.0	91.2	81.1
4	1,3,5,7	95.5	92.7	82.4
6	0,1,3,5,7,9	95.9	93.5	83.4
8	0,1,2,3,5,7,9,10	96.1	93.8	83.9
10	0,1,2,3,5,7,9,10,12,13	96.2	93.9	83.6
12	0,1,2,3,4,5,6,7,9,10,12,13	96.2	94.0	83.8
14	0,1,2,3,4,5,6,7,8,9,10,12,13,14	96.2	94.0	83.6
16	0,1,2,3,4,5,6,7,8,9,10,12,13,14,15,16	96.3	94.0	83.7
18	0,1,2,3,4,5,6,7,8,9,10,12,13,14,15,16,17,18	96.3	94.1	83.7

20	0,1,2,3,4,5,6,7,8,9,10,12,13,14,15,16,17,18,19,20	96.3	94.1	83.9
22	0,1,2,3,4,5,6,7,8,9,10,12,13,14,15,16,17,18,19,20,21,22	96.5	94.2	83.9
24	0,1,2,3,4,5,6,7,8,9,10,12,13,14,15,16,17,18,19,20,21,22,24,25	96.5	94.1	84.0

Tabela 4 Matriz de confusão, considerando eventos ocorrendo um a um, com N = 833.

Classe prevista	Classe verdadeira				
	Tipo-1	Tipo-2A	Tipo-2B	Tipo-3	Outros eventos
Tipo-1	33	0	0	0	0
Tipo-2A	0	13	0	0	0
Tipo-2B	0	0	11	0	0
Tipo-3	0	0	0	11	0
Outros eventos	0	0	1	0	83

## REFERÊNCIAS

- Al-Shaher, M. A., Sabry, M. M. and Saleh, A. S. (2003) Fault location in multi-ring distribution network using artificial neural network. Available at: [www.elsevier.com/locate/epsr](http://www.elsevier.com/locate/epsr).
- Brasil, T. A., Gois, J. N. and Neto, J. A. M. (2021) ‘Detection of High Impedance Faults in Primary Distribution Grid using Support Vector Machines Classification’, in. 14th IEEE/IAS International Conference on Industry Applications.
- Cassie, A. M. (1939) ‘Theorie Nouvelle des Arcs de Rupture et de la Rigidité des Circuits’, Cigre, Report, 102, pp. 588–608.
- Chaitanya, B. K., Yadav, A. and Pazoki, M. (2020) ‘An Intelligent Detection of High-Impedance Faults for Distribution Lines Integrated with Distributed Generators’, IEEE Systems Journal. Institute of Electrical and Electronics Engineers Inc., 14(1), pp. 870–879. doi: 10.1109/JSYST.2019.2911529.
- Chen, J. et al. (2016) ‘Detection of high impedance faults using current transformers for sensing and identification based on features extracted using wavelet transform’, IET Generation, Transmission and Distribution. Institution of Engineering and Technology, 10(12), pp. 2990–2998. doi: 10.1049/iet-gtd.2016.0021.
- Costa, F. B. et al. (2015) ‘Real-time detection of transients induced by high-impedance faults based on the boundary wavelet transform’, IEEE Transactions on Industry Applications. Institute of Electrical and Electronics Engineers Inc., 51(6), pp. 5312–5323. doi: 10.1109/TIA.2015.2434993.
- Don Russell, B. (1990) Computer Relaying and Expert Systems: New Tools for Detecting High Impedance Faults.
- Elkalashy, N. I. et al. (2007) ‘Modeling and experimental verification of high impedance arcing fault in medium voltage networks’, IEEE Transactions on Dielectrics and Electrical Insulation. IEEE, 14(2), pp. 375–383.
- Emanuel, A. E. et al. (1990) ‘High impedance fault arcing on sandy soil in 15kV distribution feeders: Contributions to the evaluation of the low frequency spectrum’, IEEE Transactions on Power Delivery, 5(2), pp. 676–686. doi: 10.1109/61.53070.
- Gautam, S. and Brahma, S. M. (2013) ‘Detection of high impedance fault in power distribution systems using

- mathematical morphology', *IEEE Transactions on Power Systems*, 28(2), pp. 1226–1234. doi: 10.1109/TPWRS.2012.2215630.
- Ghaderi, A. et al. (2015) 'High-impedance fault detection in the distribution network using the time-frequency-based algorithm', *IEEE Transactions on Power Delivery*. Institute of Electrical and Electronics Engineers Inc., 30(3), pp. 1260–1268. doi: 10.1109/TPWRD.2014.2361207.
- Ghaderi, A., Ginn, H. L. and Mohammadpour, H. A. (2017) 'High impedance fault detection: A review', *Electric Power Systems Research*. Elsevier Ltd, pp. 376–388. doi: 10.1016/j.epsr.2016.10.021.
- Gomes, D. P. S., Ozansoy, C. and Ulhaq, A. (2018) 'High-Sensitivity Vegetation High-Impedance Fault Detection Based on Signal's High-Frequency Contents', *IEEE Transactions on Power Delivery*. Institute of Electrical and Electronics Engineers Inc., 33(3), pp. 1398–1407. doi: 10.1109/TPWRD.2018.2791986.
- Hamel, A., Gaudreau, A. and Côté, M. (2004) 'Intermittent arcing fault on underground low-voltage cables', *IEEE Transactions on Power Delivery*, 19(4), pp. 1862–1868. doi: 10.1109/TPWRD.2003.822979.
- Hastie, T., Tibshirani, R. and Friedman, J. (2009) *Springer Series in Statistics The Elements of Statistical Learning Data Mining, Inference, and Prediction*.
- Khakpour, A. et al. (2015) 'Electrical arc model based on physical parameters and power calculation', *IEEE Transactions on Plasma Science*. Institute of Electrical and Electronics Engineers Inc., 43(8), pp. 2721–2729. doi: 10.1109/TPS.2015.2450359.
- Lala, H. and Karmakar, S. (2019) 'Classification of Arc Fault in Sphere-Gap and Rod-Gap Using Stockwell Transform and Machine Learning Based Approach', in 2019 International Conference on High Voltage Engineering and Technology (ICHVET), pp. 1–5. doi: 10.1109/ICHVET.2019.8724248.
- Lala, H. and Karmakar, S. (2020) 'Detection and Experimental Validation of High Impedance Arc Fault in Distribution System Using Empirical Mode Decomposition', *IEEE Systems Journal*. Institute of Electrical and Electronics Engineers Inc., 14(3), pp. 3494–3505. doi: 10.1109/JSYST.2020.2969966.
- Le, V. and Yao, X. (2019) 'Ensemble Machine Learning Based Adaptive Arc Fault Detection for DC Distribution Systems', in 2019 IEEE Applied Power Electronics Conference and Exposition (APEC), pp. 1984–1989. doi: 10.1109/APEC.2019.8721922.
- Livani, H. and Evrenosoglu, C. Y. (2013) 'A fault classification and localization method for three-terminal circuits using machine learning', *IEEE Transactions on Power Delivery*, 28(4), pp. 2282–2290. doi: 10.1109/TPWRD.2013.2272936.
- Mayr, O. (1943) 'Beiträge zur Theorie des statischen und des dynamischen Lichtbogens', *Archiv für Elektrotechnik*. Springer, 37(12), pp. 588–608.
- Michael Aucoin, B. (1996) High Impedance Fault Detection Implementation Issues, *IEEE Transactions on Power Delivery*.
- Michalik, M. et al. (2006) 'High-impedance fault detection in distribution networks with use of wavelet-based algorithm', *IEEE Transactions on Power Delivery*, 21(4), pp. 1793–1802. doi: 10.1109/TPWRD.2006.874581.
- Samantaray, S. R., Tripathy, L. N. and Dash, P. K. (2009) 'Combined EKF and SVM based high impedance fault detection in power distribution feeders', in 2009 International Conference on Power Systems, ICPS '09. doi: 10.1109/ICPWS.2009.5442697.
- Sarlak, M. and Shahrtash, S. M. (2008) 'High impedance fault detection in distribution networks using support vector machines based on wavelet transform', in 2008 IEEE Canada Electric Power Conference, pp. 1–6. doi: 10.1109/EPC.2008.4763380.
- Schavemaker, P. H. and Van Der Sluis, L. (2000) 'An improved Mayr-type arc model based on current-zero measurements', *IEEE Transactions on Power Delivery*, 15(2), pp. 580–584. doi: 10.1109/61.852988.
- Schneider, K. P. et al. (2018) 'Analytic Considerations and Design Basis for the IEEE Distribution Test Feeders', *IEEE Transactions on Power Systems*. Institute of Electrical and Electronics Engineers Inc., 33(3), pp. 3181–3188. doi: 10.1109/TPWRS.2017.2760011.
- Sedighi, A. R. et al. (2005) 'High impedance fault detection based on wavelet transform and statistical pattern recognition', *IEEE Transactions on Power Delivery*, 20(4), pp. 2414–2421. doi: 10.1109/TPWRD.2005.852367.
- Sedighi, A. R., Haghifam, M. R. and Malik, O. P. (2005) 'Soft computing applications in high impedance fault detection in distribution systems', *Electric Power Systems Research*, 76(1–3), pp. 136–144. doi: 10.1016/j.epsr.2005.05.004.
- Soheili, A., Sadeh, J. and Bakhshi, R. (2018) 'Modified FFT based high impedance fault detection technique considering distribution non-linear loads: Simulation and experimental data analysis', *International Journal of Electrical Power and Energy Systems*. Elsevier Ltd, 94, pp. 124–140. doi: 10.1016/j.ijepes.2017.06.035.
- Sultan, A. F., Swift, G. W. and Feddrchuk, D. J. (1994) DETECTING ARCING DOWNED-WIRES USING FAULT CURRENT FLICKER AND HALF-CYCLE ASYMMETRY, *IEEE Transactions on Power Delivery*.
- Tai, N. L. and Chen, J. J. (2008) 'Wavelet-based approach for high impedance fault detection of high voltage transmission line', *European Transactions on Electrical Power*, 18(1), pp. 79–92. doi: 10.1002/etep.169.
- Wang, B., Geng, J. and Dong, X. (2016) 'High-Impedance Fault Detection Based on Nonlinear Voltage-Current Characteristic Profile Identification', *IEEE Transactions on Smart Grid*, 9(4), pp. 3783–3791. doi: 10.1109/TSG.2016.2642988.
- Wei, M. et al. (2020) 'High Impedance Arc Fault Detection Based on the Harmonic Randomness and Waveform Distortion in the Distribution System', *IEEE Transactions on Power Delivery*. IEEE, 35(2), pp. 837–850. doi: 10.1109/TPWRD.2019.2929329.