

Eliminação Recursiva de Características (SVM-RFE) para a Localização de Descargas Parciais em Transformadores

Arismar Morais Gonçalves Júnior*
Rogério Andrade Flauzino* Hélder de Paula**
Sofia Moreira de Andrade Lopes*

* *Departamento de Engenharia Elétrica e de Computação, Escola de Engenharia de São Carlos, Universidade de São Paulo, SP, (e-mails: arismarjr@usp.br; raflauzino@usp.br; sofia.moreira.lopes@usp.br)*

** *Faculdade de Engenharia Elétrica, Universidade Federal de Uberlândia, MG, (e-mail: drhelderdepaula@gmail.com)*

Abstract: Partial discharge (PD) investigations can provide early identification and localization of incipient failures in power transformers, a fact that can avoid considerable financial losses. For the development of these location techniques, feature extraction and selection of PD signals is a fundamental step, because it directly influences the location method's performance. In order to determine a better set of signal features for the location of the PDs, the SVM-RFE selection method is investigated in this work, based on parameters extracted from PD acoustic emission signals. The results show that the SVM-RFE method can select the most suitable features for the PDs localization, reducing the dimension of the input space, in addition to indicating that a smaller number of sensors can be employed in the measurements of the discharges.

Resumo: A investigação de descargas parciais (PDs) em transformadores pode proporcionar uma identificação antecipada e localização de falhas incipientes nestes equipamentos, fato que pode evitar consideráveis perdas financeiras. Para o desenvolvimento de tais técnicas de localização, no entanto, uma etapa fundamental se refere à extração e seleção de atributos dos sinais das PDs, parâmetros que podem influenciar diretamente no desempenho do método utilizado. Visando a determinação de um melhor conjunto de características para a localização das descargas parciais, o método de seleção SVM-RFE é investigado neste trabalho, com base em parâmetros extraídos de sinais de emissão acústica das PDs. Os resultados mostram que o método SVM-RFE é capaz de selecionar características mais adequadas para a localização das PDs, reduzindo o espaço de características inicial, além de indicar que um menor número de sensores pode ser empregado nas medições das descargas.

Keywords: Transformers; partial discharges; acoustic emission; features selection; SVM-RFE; localization.

Palavras-chaves: Transformadores; descargas parciais; emissão acústica; seleção de características; SVM-RFE; localização.

1. INTRODUÇÃO

A identificação e localização de descargas parciais (PDs - *Partial Discharges*) em transformadores de potência têm sido largamente utilizada como estratégia de manutenção preditiva destes equipamentos. A ocorrência das PDs pode ser um indicativo da degradação do isolamento dos enrolamentos, normalmente constituído por papel impregnado com óleo, de modo que, se antecipadamente detectadas, pode evitar falhas mais severas no transformador, assim como os elevados custos associados à parada inesperada do equipamento (Shuai et al., 2014; Ganeshan et al., 2017).

A atividade das PDs em transformadores pode ser captada por meio dos fenômenos advindos da ocorrência destas descargas, como os pulsos de corrente elétrica, as ondas eletromagnéticas irradiadas e as ondas acústicas meca-

nicamente excitadas pela pequenas “explosões” das PDs (Kunicki et al., 2018; Samimi and Ilkhechi, 2020). De modo geral, o método acústico tem sido mais amplamente empregado por constituir um método de medição não-invasivo, com possibilidade de aplicação sem a necessidade de retirada do transformador de operação e, principalmente, por propiciar a localização das fontes de PDs no interior no tanque. Este último fato é especialmente importante para a redução do tempo e do custo de reparo do equipamento.

Classicamente, a localização das PDs é realizada pela diferença entre os tempos de chegada dos sinais sonoros a diversos sensores piezoelétricos, em conjunto com estratégias de triangularização, técnica denominada TDOA (*Time Difference Of Arrival*) (Eleftherion, 1995; Ghosh et al., 2017). Apesar de propiciar uma correta localização das PDs, a TDOA apresenta algumas desvantagens, como

a dificuldade de estimação do real tempo de chegada dos sinais aos sensores, principalmente devido às diversas reflexões/refrações dos sinais acústicos no interior do tanque, e a necessidade de um método iterativo para a solução do conjunto de equações não-lineares. Por conseguinte, atualmente, técnicas de aprendizado de máquinas vem sendo aplicadas na localização das PDs (Homaei et al., 2014; Rahman et al., 2016; Gonçalves Júnior et al., 2019). Neste sentido, características extraídas das oscilografias dos sinais acústicos das PDs são investigadas por meio de métodos de classificação ou reconhecimento de padrões.

Uma vez que inúmeras características/atributos podem ser extraídos dos sinais das PDs, a determinação de um conjunto de características mais apropriado para a separação efetiva destes sinais em suas distintas classes (locais de ocorrência) é uma etapa fundamental. Com a seleção das características, o problema do sobreajustamento (*overfitting*) de modelos pode ser evitado, melhorando o desempenho dos mesmos na localização das PDs. Além do mais, os tipos de características mais relevantes para a localização das PDs podem ser determinados, aliviando a etapa de extração destes parâmetros (Webb and Copley, 2011).

Neste trabalho, um método recursivo de seleção de características (SVM-RFE) é empregado na determinação de atributos para a localização de descargas parciais em transformadores, a partir de sinais acústicos captados por diversos sensores piezoelétricos. Os resultados mostram que o método investigado é capaz de efetivamente selecionar as características mais relevantes para a localização das PDs, melhorando o desempenho de redes neurais artificiais na determinação dos locais de ocorrência destas descargas. Ademais, as características selecionadas sugerem que um menor número de sensores pode ser utilizado, reduzindo os custos relacionados à aquisição de sinais. Cabe ressaltar que nenhuma aplicação similar foi encontrada na literatura, no que se refere à seleção de características para a classificação/localização de PDs em transformadores, fato que justifica a originalidade deste trabalho.

2. EXTRAÇÃO DE CARACTERÍSTICAS

A extração de características é um processo necessário para a análise das formas de onda de sinais por meio de métodos de aprendizado de máquinas. De modo geral, a extração de características de sinais atua de forma a remover a informação redundante dos mesmos, reduzindo a dimensão dos dados e transformando-os em uma forma mais apropriada para aplicação dos métodos computacionais de classificação ou reconhecimento de padrões.

No contexto de localização das PDs, a extração de características proporciona a obtenção de parâmetros mais adequados para uma separação efetiva dos sinais das descargas em suas distintas classes, as quais estão intimamente relacionadas com diferenças nas formas de onda dos sinais acústicos, em virtude dos diferentes caminhos de propagação encontrados pelos mesmos desde o ponto de ocorrência das descargas até os sensores piezoelétricos.

De acordo com Wu et al. (2015) e Raymond et al. (2015), as seguintes características são normalmente utilizadas: amplitude dos sinais; tempos de chegada aos diferentes

sensores, tempos de subida e descida das formas de onda; parâmetros estatísticos, como o cálculo da média, variância, skewness e kurtosis; propriedades na frequência, como as componentes do espectro de frequências obtido pela transformada de Fourier; parâmetros de wavelets, como os coeficientes advindos desta transformada; características ortogonais, como os coeficientes resultantes da análise por componentes principais; aspectos fractais; características baseadas na teoria do caos; dentre outras.

3. SELEÇÃO DE CARACTERÍSTICAS

A seleção de características visa a identificação dos atributos mais adequados para os problemas de classificação ou reconhecimento de padrões, melhorando o desempenho de tais métodos de aprendizagem de máquinas.

Conforme exposto em Yan and Zhang (2015), três diferentes estratégias são utilizadas na literatura, com base em como o método de seleção interage com o classificador: i) filtros, em que as características são avaliadas de acordo com coeficientes de correlação ou critérios de informação, sendo independente de classificadores; ii) métodos envoltos (*wrappers*), onde um subconjunto de características adequado para a classificação dos dados é determinado a partir da validação cruzada de classificadores; e iii) métodos embutidos (*embedded*), em que a importância das características é investigada a partir dos coeficientes de classificadores. De modo geral, esta última metodologia, da qual o SVM-RFE faz parte, apresenta algumas vantagens como uma maior capacidade de evitar o *overfitting*, possibilidade de uso de todas as observações do conjunto de dados e velocidade de processamento.

3.1 SVM-RFE

O método recursivo de seleção de características por máquinas de vetores suporte (*SVM-RFE - Support Vector Machine Recursive Feature Elimination*) foi proposto por Guyon et al. (2002) para a seleção de características genéticas de cânceres. Tal método utiliza os coeficientes de classificadores SVM, isto é, o vetor peso descritivo do melhor hiperplano separador entre as classes do problema, no desenvolvimento de um critério de ranqueamento, removendo recursivamente as características que resultam nestes menores índices de ordenamento (Yan and Zhang, 2015; Adorada et al., 2018).

No SVM, o melhor hiperplano separador entre duas classes é definido como aquele que resulta na maior margem de separação entre estas classes. Em casos lineares, esta margem de separação pode ser obtida como duas vezes a distância entre o hiperplano separador e as amostras das classes mais próximas do mesmo, que constituem os denominados vetores suporte. A equação (1) apresenta o vetor peso (\mathbf{w}) que define o melhor hiperplano separador, obtido a partir da formulação Lagrangeana do problema de otimização que deve ser solucionado (Haykin, 1999):

$$\mathbf{w} = \sum_{i=1}^n \alpha_i y_i \mathbf{x}_i, \quad (1)$$

sendo α_i os multiplicadores de Lagrange e $\{\mathbf{x}_i, y_i\}$ as amostras do conjunto de treinamento do classificador, com

$\mathbf{x}_i \in \mathbf{R}^d$, $y_i \in \{-1, 1\}$ e $i = 1, 2, \dots, n$. O espaço d -dimensional está relacionado ao número de atributos do problema inicial, o qual será reduzido por meio da remoção das características menos relevantes.

Uma vez que a consideração (ou não-consideração) de determinadas características dos dados culmina em diferentes hiperplanos separadores, o vetor peso da SVM pode ser utilizado na identificação dos atributos que resultam nas melhores separações espaciais das classes. Assim, segundo (Guyon et al., 2002), o vetor peso pode ser empregado na constituição de um critério de ordenamento das características conforme equação (2), sendo k a k -ésima característica do conjunto de dados.

$$c(k) = w_k^2. \quad (2)$$

O uso recursivo do SVM e do critério descrito na equação (2) possibilita um ranqueamento dos atributos do conjunto de dados, dos melhores para os piores na separação das classes do problema, conforme descrito no pseudocódigo abaixo:

Entrada: Conjunto de dados; SVM

Saída: Lista de características ordenadas

início

```

   $s \leftarrow [1, 2, \dots, d]$ 
   $r \leftarrow \emptyset$ 
  enquanto  $s \neq \emptyset$  faça
    Treinar SVM com as características em  $s$ 
    Calcular  $\mathbf{w}$ 
    Calcular  $c$ 
    Ordenar  $s$  de acordo com  $c$ 
    Eliminar última característica de  $s$ 
    Atualizar  $r$  e  $s$ 
  fim

```

fim

A partir de um conjunto de dados com d características e de um algoritmo de treinamento de SVM, em cada iteração, a característica com o menor valor de c é eliminada, sendo a lista de características sobreviventes (s) e de características ordenadas (r) atualizadas.

4. BANCADA DE TESTES

O aparato experimental utilizado neste trabalho é constituído de um tanque metálico, no exterior do qual são instalados oito sensores piezoelétricos para a captação de sinais acústicos das PDs, conforme ilustrado na Figura 1. Tais sensores foram disponibilizados diagonalmente e de maneira intercalada nas faces do tanque, configuração que resulta em menores erros na localização das PDs (Hekmati and Hekmati, 2017). Desta forma, os sensores acústicos foram instalados nas seguintes coordenadas:

- $S1$: $x = 0 \text{ cm}$, $y = 123 \text{ cm}$ e $z = 15 \text{ cm}$;
- $S2$: $x = 0 \text{ cm}$, $y = 15 \text{ cm}$ e $z = 69 \text{ cm}$;
- $S3$: $x = 15 \text{ cm}$, $y = 0 \text{ cm}$ e $z = 15 \text{ cm}$;
- $S4$: $x = 123 \text{ cm}$, $y = 0 \text{ cm}$ e $z = 69 \text{ cm}$;
- $S5$: $x = 138 \text{ cm}$, $y = 15 \text{ cm}$ e $z = 15 \text{ cm}$;
- $S6$: $x = 138 \text{ cm}$, $y = 123 \text{ cm}$ e $z = 69 \text{ cm}$;
- $S7$: $x = 123 \text{ cm}$, $y = 138 \text{ cm}$ e $z = 15 \text{ cm}$;
- $S8$: $x = 15 \text{ cm}$, $y = 138 \text{ cm}$ e $z = 69 \text{ cm}$.

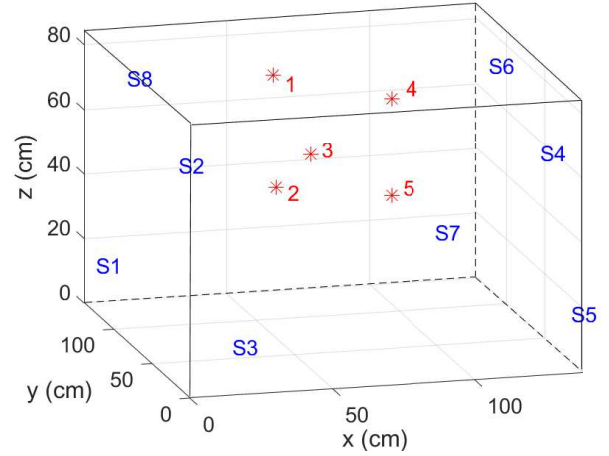


Figura 1. Posicionamento dos sensores piezoelétricos (S1 à S8) e locais de ocorrência das PDs (*).

Para a geração das descargas no interior do tanque, preenchido com óleo mineral isolante, foi utilizado um arranjo de eletrodos em configuração ponta-plano, o qual pode ser movimentado em três eixos por meio de um sistema de trilhos e roldanas. O *hardware* para a aquisição e condicionamento dos sinais foi desenvolvido por Navarro (2019), o qual consta de 16 canais analógico/digitais, 16 *bits* de resolução e taxa de amostragem de até 1,5 MS/s.

As PDs foram realizadas por meio da elevação da tensão no arranjo ponta-plano até o valor necessário para a disrupção e geração da descarga no *gap*, limiar este de aproximadamente 3kV. Cinco diferentes posições foram tomadas, nas seguintes coordenadas retangulares:

- Posição 1: $x = 52 \text{ cm}$, $y = 85 \text{ cm}$ e $z = 79 \text{ cm}$;
- Posição 2: $x = 53 \text{ cm}$, $y = 85 \text{ cm}$ e $z = 44 \text{ cm}$;
- Posição 3: $x = 53 \text{ cm}$, $y = 40 \text{ cm}$ e $z = 64 \text{ cm}$;
- Posição 4: $x = 83 \text{ cm}$, $y = 42 \text{ cm}$ e $z = 79 \text{ cm}$;
- Posição 5: $x = 82 \text{ cm}$, $y = 42 \text{ cm}$ e $z = 49 \text{ cm}$.

Em cada uma das diferentes posições, 50 oscilografias foram registradas por cada um dos sensores e armazenadas. Cada um desses registros retrata uma janela temporal de aproximadamente 0,035s sendo que, em geral, diversas descargas parciais se manifestam. A Figura 2 ilustra o sinal registrado pelo sensor S_5 para descargas na posição 3, assim como os envelopes superior (*envU*) e inferior (*envL*) dessa oscilografia.

4.1 Extração de características das oscilografias

Após o processamento inicial das oscilografias, para remoção do *offset* dos sinais e obtenção dos envelopes superior e inferior (por meio da transformada de Hilbert), características foram extraídas de forma a se obter um conjunto de parâmetros descritivos da localização das PDs. Neste sentido, os seguintes atributos foram estimados para os sinais de cada um dos sensores:

- Pico máximo (pk_{max}) e respectivo tempo (t_{max});
- Pico mínimo (pk_{min}) e respectivo tempo (t_{min});
- Número de picos positivos (npk_{pos}) e negativos (npk_{neg});
- Energia total (E);

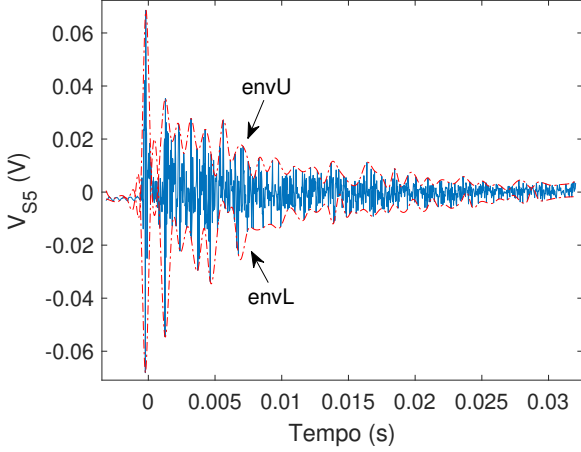


Figura 2. Oscilografia do sensor S_5 e envelopes para PDs na posição 3.

- Parâmetros estatísticos: média (μ), desvio padrão (σ), skewness (sk) e kurtosis (ku);
- Energia e parâmetros estatísticos advindos dos envelopes superior ($E_{envU}, \mu_{envU}, \sigma_{envU}, sk_{envU}$ e ku_{envU}) e inferior ($E_{envL}, \mu_{envL}, \sigma_{envL}, sk_{envL}$ e ku_{envL});
- Tempo de chegada dos sinais aos sensores ($t_{arrival}$);
- Diferenças dos tempos de máximo (Δt_{max}), mínimo (Δt_{min}) e de chegada ($\Delta t_{arrival}$), com relação ao sensor S_1 .

Desta forma, para cada teste de PDs realizado, tem-se um total de 197 características descritivas das oscilografias captadas pelos sensores. Tais atributos foram então disponibilizados nas colunas de uma matriz de características, de modo que cada linha desta matriz corresponde aos valores obtidos em cada teste realizado, organizados em sequência para cada posição de ocorrência das descargas. De modo a remover a influência da magnitude dos atributos no processo de seleção pelo SVM-RFE, cada coluna da matriz de características foi normalizada.

5. RESULTADOS

O método SVM-RFE de seleção de características foi desenvolvido no *software* MATLAB[®] utilizando a abordagem *one-against-all*, que transforma um problema de múltiplas classes em um problema de apenas duas classes. Nessa metodologia, o hiperplano separador é estimado considerando-se os dados de determinada classe (local de ocorrência das PDs) contra todas as outras possíveis. Assim, as características foram ordenadas pelo método SVM-RFE para cada ponto de ocorrência das descargas, com base na separabilidade para todos os outros locais. Para fins de redução do número de atributos, apenas as 10 primeiras características foram consideradas.

A tabela 1 apresenta as dez principais características selecionadas pelo SVM-RFE para a localização das PDs. Pode-se notar que os tipos de atributos mais importantes para a localização das descargas são: pk_{max} , que é selecionado 9 vezes (18% das 50 características da Tabela 1); σ_{envU} e σ_{envL} , selecionados 6 vezes cada (24%, no total); e pk_{min} , npk_{max} , σ e ku , selecionadas 4 vezes cada (32%, no total). Do mesmo modo, as características relacionadas ao

tempo (t_{max} , t_{min} e $t_{arrival}$), assim como as características estatísticas sk , sk_{envU} , ku_{envU} e ku_{envL} se mostram menos relevantes para a classificação das PDs com base nos seus locais de ocorrência, não sendo selecionadas entre os 10 primeiros atributos. Com relação aos principais sensores para a localização das PDs, pode-se perceber que as características relacionadas a S_5 e S_7 são mais recorrentes na Tabela 1, sendo selecionadas 12 vezes (48%, total). Outrossim, os atributos advindos dos sinais dos sensores S_2 e S_6 também se mostram importantes, estando presentes em 7 (14%) e 6 (12%) ocasiões nas características da Tabela 1, respectivamente.

Tabela 1. Dez primeiras características selecionadas pelo SVM-RFE.

	Pos. 1	Pos. 2	Pos. 3	Pos. 4	Pos. 5
1^a	$pk_{max}S_2$	$\mu_{envU}S_5$	$pk_{max}S_5$	$\sigma_{envL}S_7$	$E_{envU}S_2$
2^a	$\sigma_{envU}S_7$	$\mu_{envL}S_5$	$pk_{max}S_7$	$pk_{max}S_2$	$pk_{max}S_6$
3^a	$pk_{min}S_1$	$npk_{max}S_7$	$npk_{min}S_6$	$npk_{max}S_7$	$E_{envL}S_2$
4^a	$pk_{max}S_1$	σS_5	$\sigma_{envU}S_5$	$pk_{min}S_7$	$pk_{min}S_6$
5^a	$\mu_{envU}S_5$	$pk_{max}S_7$	kuS_3	kuS_1	$E_{envL}S_3$
6^a	$\sigma_{envU}S_1$	$\sigma_{envU}S_5$	$npk_{max}S_6$	$pk_{max}S_7$	$\mu_{envU}S_7$
7^a	σS_5	$\sigma_{envL}S_8$	$pk_{min}S_6$	$\sigma_{envL}S_2$	ES_3
8^a	$\sigma_{envL}S_2$	$sk_{envL}S_7$	σS_5	kuS_5	$\sigma_{envL}S_4$
9^a	$\sigma_{envL}S_7$	$npk_{max}S_8$	$npk_{min}S_5$	$\sigma_{envU}S_7$	$\sigma_{envU}S_6$
10^a	σS_2	kuS_4	μS_4	$\mu_{envL}S_5$	$pk_{max}S_3$

Para verificar o desempenho das características selecionadas pelo SVM-RFE na localização das PDs, redes neurais MLP (Multi-Layer Perceptron) foram desenvolvidas no MATLAB[®]. Para tanto, foi fixada uma topologia constituída por duas camadas escondidas, cada uma com 10 neurônios e função de ativação tangente sigmoial, treinada a partir do método *Levenberg-Marquardt backpropagation*. Os dados foram divididos em dois conjuntos, para treinamento (70%) e validação (30%) das redes. Neste último procedimento, a acurácia, isto é, o percentual de acertos na classificação/localização das descargas foi tomada como figura de mérito.

Quatro diferentes casos foram analisados, com relação ao número de características utilizadas:

- caso *i*: todos os 197 parâmetros do espaço de características original;
- caso *ii*: 37 diferentes características determinadas pelo SVM-RFE;
- caso *iii*: 24 características do SVM-RFE comuns aos 4 principais sensores (S_5, S_7, S_6 e S_2);
- caso *iv*: 20 características do SVM-RFE relativas aos cinco principais tipos (pk_{max} , σ_{envU} , σ_{envL} , npk_{max} e ku), considerando todos os sensores.

A Figura 3 apresenta as respostas na validação das redes MLP para cada um dos casos estudados, sendo que as observações 1 – 15, 16 – 30, 31 – 45, 46 – 60 e 61 – 75 dizem respeito aos locais 1, 2, 3, 4 e 5 de ocorrência das descargas, respectivamente. Pela figura, pode notar que o desempenho com a realização da seleção pelo SVM-RFE é superior ao desempenho sem nenhuma seleção prévia de características, ainda que apenas dados de 4 sensores ou apenas 5 diferentes tipos de atributos sejam utilizados. Tal fato confirma que a existência de características pouco relevantes no conjunto de dados realmente deteriora a capacidade de classificação/localização por métodos de aprendizagem de máquina.

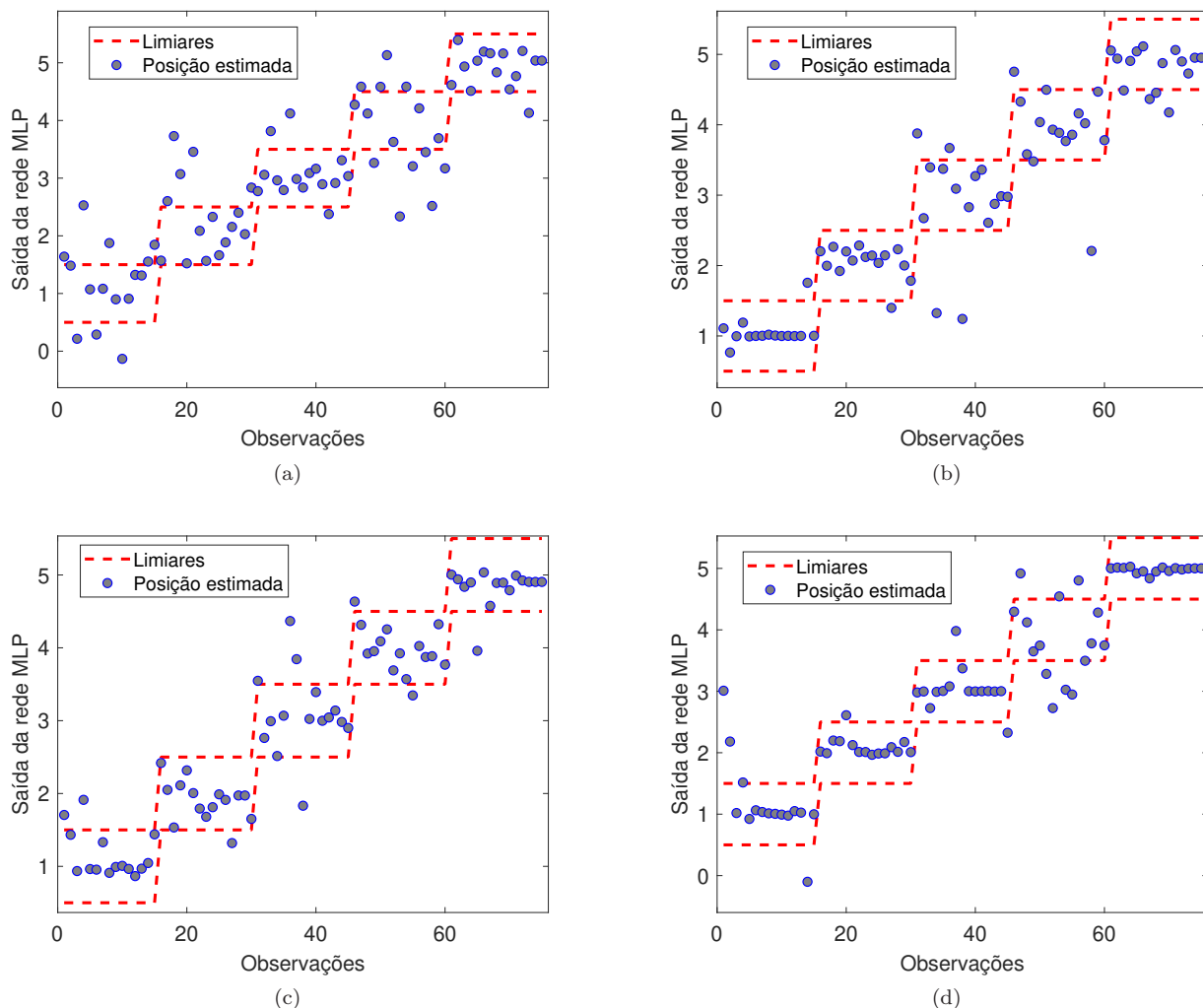


Figura 3. Desempenho das redes *MLP*: (a) caso *i*; (b) caso *ii*; (c) caso *iii*; (d) caso *iv*.

Para o cálculo da acurácia, foi admitido um limiar de $\pm 0,5$ na tomada das respostas corretas das redes, as quais estão relacionadas com os locais de ocorrência das descargas.

A Tabela 2 apresenta o percentual de acertos total e por localização das PDs para cada um dos casos investigados, onde observa-se que o SVM-RFE (caso *ii*) eleva substancialmente o número total de respostas corretas. Apesar de haver certa redução no percentual de acertos para as localidades 3 e 5, com relação ao caso sem seleção (caso *i*), 73% de acurácia ainda é um número significativo.

Tabela 2. Acurácia das redes *MLP* na localização das PDs.

Pos.	Caso <i>i</i>	Caso <i>ii</i>	Caso <i>iii</i>	Caso <i>iv</i>
1	46,67%	93,33%	86,67%	73,33%
2	66,67%	93,33%	93,33%	93,33%
3	80,00%	73,33%	73,33%	86,67%
4	33,33%	80,00%	86,67%	46,67%
5	93,33%	73,33%	93,33%	100,00%
Total	64,00%	82,67%	86,67%	80,00%

Pela Tabela 2 pode-se notar também que apenas 4 sensores (caso *iii*) podem ser empregados na localização das PDs por técnicas de aprendizado de máquinas, resultando em

um percentual de acertos ainda superior que os casos anteriores. Não obstante, a utilização de apenas os 5 principais tipos de características reduz o percentual de acertos total, principalmente pela redução na acurácia das observações relacionadas ao local 4 de ocorrência das PDs. Tal fato pode ser justificado pela possível desconsideração de alguma característica importante para a discriminação dessa classe/localidade.

6. CONCLUSÃO

A seleção de características para a localização de descargas parciais em transformadores é realizada neste trabalho por meio do SVM-RFE, uma método recursivo baseado em máquinas de vetores suporte.

Os resultados mostram que a técnica investigada pode reduzir efetivamente o conjunto de atributos obtidos dos sinais de emissão acústica, melhorando o desempenho de redes neurais *MLP* na localização das PDs. Ademais, as principais características selecionadas sugerem que um menor número de sensores piezoelétricos podem ser utilizados na captação das descargas.

Com a efetiva localização das PDs, falhas no isolamento de transformadores podem ser previamente identificadas, possibilitando o planejamento das atividades de manutenção e reparo do equipamento. Com isso, a árdua e demorada inspeção em busca das regiões de ocorrência das PDs pode ser evitada, influenciando diretamente na redução do tempo e no custo de reparo do transformador.

AGRADECIMENTOS

Os autores gostariam de agradecer ao CNPQ pelo suporte dado a este trabalho.

REFERÊNCIAS

- Adorada, A., Permatasari, R., Wirawan, P.W., Wibowo, A., and Sujiwo, A. (2018). Support Vector Machine - Recursive Feature Elimination (SVM-RFE) for feature selection of microRNA expression features of breast cancer. In *2nd International Conference on Informatics and Computational Sciences*.
- Eleftherion, P.M. (1995). Partial discharge XXI - Acoustic emission based PD source location in transformers. *IEEE Electrical Insulation Magazine*, vol. 11(6), pp. 22 - 26. doi:10.1109/57.475905.
- Ganeshan, S., Murugesan, J., Cavallini, A., Negri, F., Valecillos, B., and Piován, U. (2017). Identification of partial discharges in power transformers: An approach driven by practical experience. *IEEE Electrical Insulation Magazine*, 33(5), 23-31. doi:10.1109/MEI.2017.8014388.
- Ghosh, R., Chatterjee, B., and Dalai, S. (2017). A method for the localization of partial discharge sources using partial discharge pulse information from acoustic emissions. *IEEE Transactions on Dielectrics and Electrical Insulation*, 24(1), 237-245. doi:10.1109/TDEI.2016.006080.
- Gonçalves Júnior, A.M., de Paula, H., and Boaventura, W.d.C. (2019). Multiple Linear Regression Models for Partial Discharge Location Along Transformer Windings. In *Anais do 14º Simpósio Brasileiro de Automação Inteligente*. Ouro Preto, MG.
- Guyon, I., Weston, J., Barnhill, S., and Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Machine Learning*, 46, 389-422. doi:10.1023/A:1012487302797.
- Haykin, S. (1999). *Neural networks: a comprehensive foundation*. Upper Saddle River, New Jersey, 2nd edition.
- Hekmati, A. and Hekmati, R. (2017). Optimum acoustic sensor placement for partial discharge allocation in transformers. *IET Science, Measurement and Technology*, 11(5), 581-589. doi:10.1049/iet-smt.2016.0417.
- Homaei, M., Moosavian, S.M., and Illias, H.A. (2014). Partial Discharge Localization in Power Transformers Using Neuro-Fuzzy Technique. *IEEE Transactions on Power Delivery*, vol. 29(5), 2066 - 2076. doi:10.1109/TPWRD.2014.2339274.
- Kunicki, M., Cichon, A., and Borucki, S. (2018). Measurements on partial discharge in on-site operating power transformer: A case study. *IET Generation, Transmission and Distribution*, 12(10), 2487-2495. doi:10.1049/iet-gtd.2017.1551.
- Navarro, E.C. (2019). *Desenvolvimento de hardware de sistema de aquisição de sinais acústicos para detecção de descargas parciais*. Ph.D. thesis, Universidade de São Paulo.
- Rahman, M.S., Lewin, P.L., and Rapisarda, P. (2016). Autonomous localization of partial discharge sources within large transformer windings. *IEEE Transactions on Dielectrics and Electrical Insulation*, 23(2), 1088-1098. doi:10.1109/TDEI.2015.005070.
- Raymond, W.J.K., Illias, H.A., Bakar, A.H.A., and Mokhlis, H. (2015). Partial discharge classifications: Review of recent progress. *Measurement: Journal of the International Measurement Confederation*, 68, 164-181. doi:10.1016/j.measurement.2015.02.032.
- Samimi, M.H. and Ilkhechi, H.D. (2020). Survey of different sensors employed for the power transformer monitoring. *IET Science, Measurement and Technology*, 14(1), 1-8. doi:10.1049/iet-smt.2019.0103.
- Shuai, H., Qingmin, L., Chengrong, L., and Jiangyan, Y. (2014). Electrical and Mechanical Properties of the Oil-paper Insulation under Stress of the Hot Spot Temperature. *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 21(1), 179 - 185. doi:10.1109/TDEI.2013.003967.
- Webb, A.R. and Copsey, K.D. (2011). *Statistical pattern recognition*. John Wiley & Sons, West Sussex, United Kingdom, 3rd edition.
- Wu, M., Cao, H., Cao, J., Nguyen, H.L., Gomes, J.B., and Krishnaswamy, S.P. (2015). An overview of state-of-the-art partial discharge analysis techniques for condition monitoring. *IEEE Electrical Insulation Magazine*, 31(6), 22-35. doi:10.1109/MEI.2015.7303259.
- Yan, K. and Zhang, D. (2015). Feature selection and analysis on correlated gas sensor data with recursive feature elimination. *Sensors and Actuators, B: Chemical*, 212, 353-363. doi:10.1016/j.snb.2015.02.025. URL <http://dx.doi.org/10.1016/j.snb.2015.02.025>.