



Diagnósticos de falhas iniciais em rolamentos de geradores eólicos utilizando classificadores baseados nos parâmetros de Hjorth

Early fault diagnostics in wind generator bearings using classifiers based on Hjorth parameters

Arthur Carvalho Santos¹, Thiago Kenzo Mori², Wesley Angelino de Souza³,
Alessandro Goedel⁴, Marcelo Favoretto Castoldi⁵

RESUMO

Técnicas de aprendizado de máquina são uma abordagem amplamente utilizada para monitorar e diagnosticar falhas em máquinas elétricas. Essas técnicas extraem informações dos sinais coletados e classificam as condições de saúde dos componentes internos. Entre todos os componentes internos, o rolamento apresenta a maior taxa de falha. Classificadores frequentemente empregam dados de vibração adquiridos de máquinas elétricas, que podem indicar diferentes níveis de severidade da falha no rolamento. Diante dessa circunstância, este trabalho propõe uma metodologia para detectar falhas precoces em rolamentos de turbinas eólicas, aplicando classificadores baseados nos parâmetros de Hjorth. Os parâmetros de Hjorth são aplicados para analisar sinais de vibração coletados de experimentos a fim de distinguir estados de funcionamento normal e estados de mau funcionamento, permitindo assim a classificação de condições distintas. Os classificadores RF, SVM e k-NN apresentaram níveis de precisão acima de 95%. Notavelmente, a precisão dos classificadores é mantida mesmo após passar por um processo de redução de dimensionalidade. Dessa forma pode-se afirmar que os parâmetros de Hjorth fornecem uma alternativa viável para identificar falhas em geradores eólicos.

PALAVRAS-CHAVE: Geradores eólicos; Parâmetros de Hjorth; Sinais de Vibração

ABSTRACT

Machine learning techniques are a widely used approach to monitor and diagnose faults in electrical machines. These techniques extract information from collected signals and classify the health conditions of internal components. Among all internal components, bearings have the highest failure rate. Classifiers often employ vibration data acquired from electrical machines, which can indicate different levels of bearing fault severity. In this context, this work proposes a methodology to detect early faults in wind turbine bearings by applying classifiers based on Hjorth parameters. The Hjorth parameters are applied to analyze vibration signals collected from experiments to distinguish between normal operating states and malfunctioning states, thus allowing the classification of distinct conditions. The RF, SVM, and k-NN classifiers showed accuracy levels above 95%. Notably, classifier accuracy is maintained even after undergoing a dimensionality reduction process. Therefore, it can be stated that Hjorth parameters provide a viable alternative for identifying faults in wind generators.

KEYWORDS: Wind generator; Hjorth parameters; Vibration signal

INTRODUÇÃO

Os recentes eventos da pandemia e invasão da Ucrânia pela Rússia evidenciou os riscos de uma possível crise energética, aumentando o investimento em energias renováveis, como a energia eólica que apresenta uma previsão de aumento anual de 15% na geração de energia até 2027, representando 680 GW no final do período (HUTCHINSON M; ZHAO, 2023).

¹ Bolsista do PIVICT. Universidade Tecnológica Federal do Paraná, Cornélio Procópio, Brasil. E-mail: arthur.2019@alunos.edu.br. ID Lattes: 9950678454774434.

² Bolsista Fundação Araucária. Universidade Tecnológica Federal do Paraná, Cornélio Procópio, Brasil. E-mail: thiagomori@alunos.utfpr.edu.br. ID Lattes: 4819625262739741.

³ Docente no Departamento Acadêmico de Elétrica (DAELE). Universidade Tecnológica Federal do Paraná, Cornélio Procópio, Paraná, Brasil. E-mail: wesleyangelino@utfpr.edu.br. ID Lattes: 8594457321079718.

⁴ Docente no Departamento Acadêmico de Elétrica (DAELE). Universidade Tecnológica Federal do Paraná, Cornélio Procópio, Paraná, Brasil. E-mail: agoedel@utfpr.edu.br. ID Lattes: 1920650157123774.

⁵ Docente no Departamento Acadêmico de Elétrica (DAELE). Universidade Tecnológica Federal do Paraná, Cornélio Procópio, Paraná, Brasil. E-mail: marcastoldi@utfpr.edu.br. ID Lattes: 6178029384175205.



XIII Seminário de Extensão e Inovação XXVIII Seminário de Iniciação Científica e Tecnológica da UTFPR

Ciência e Tecnologia na era da Inteligência Artificial: Desdobramentos no Ensino Pesquisa e Extensão
20 a 23 de novembro de 2023 - Campus Ponta Grossa, PR



SEI-SICITE
2023

As turbinas eólicas instaladas em campos *offshore* são capazes de gerar 1,7 mais energia que em campos *onshore*, porém os custos de manutenção podem representar até 50% dos custos de geração, sendo o subsistema elétrico e de geração responsável pela maior parte das falhas, representando 35% do total das falhas, onde um dos componentes desse subsistema é uma máquina elétrica (ATTALLAH; IBRAHIM; ZAKZOUK, 2023).

Entre os componentes que constituem as máquinas de indução, o rolamento apresenta o maior índice de problemas, sendo responsável por 41% das falhas (VAIMANN; BELAHGEN; KALLASTE, 2015), e dado esse alto índice, métodos de monitoramento e diagnósticos são desenvolvidos com base em dados de vibração, emissão acústica, temperatura, tensão, entre outros, para minimizar os custos de manutenção (DUAN et al., 2018).

Dado este cenário, este trabalho apresenta um método para monitorar a condição de saúde dos rolamentos, através de atributos no domínio do tempo extraídos do sinal de vibração da evolução da degradação do rolamento, aplicando os parâmetros de Hjorth como uma ferramenta capaz de rotular as amostras perante a condição de saúde do rolamento, para a aplicação em classificadores de aprendizado supervisionado.

A metodologia proposta permite identificar o momento da ocorrência da falha desde o seu estágio inicial com auxílio dos parâmetros de Hjorth, permitindo a rotulação de amostras que inicialmente compõem um conjunto de dados não rotulados, para a sua aplicação em classificadores de aprendizado supervisionados utilizando atributos extraídos no domínio do tempo, sendo capaz de identificar as falhas em estágio inicial, reduzindo os custos de operação e manutenção.

BANCO DE DADOS

Os sinais de vibração foram coletados pelo Center Intelligent Maintenance Systems da Universidade de Cincinnati. O experimento foi realizado utilizando 4 rolamentos com lubrificação forçada acoplados em um eixo conectado em um motor por meio de correias, cuja a velocidade permaneceu constante em 2000 rpm. Um plugue magnético foi instalado no tanque de óleo de tal forma que o acúmulo de detritos provenientes da degradação encerra o experimento (QIU et al., 2006).

O processo foi realizado 3 vezes, sendo que o experimento 1 utilizou acelerômetros nas direções x e y, enquanto que os experimentos 2 e 3 utilizaram apenas um acelerômetro por rolamento cuja o seu posicionamento não é especificado. Os sinais possuem duração de 1 segundo e coletados a cada 10 minutos. A Tabela 1 apresenta a quantidade de amostras coletadas em cada experimento e os rolamentos que apresentaram falha. Os dados provenientes do “Experimento 3” não foram considerados nesse trabalho dado que os sinais apresentaram inconstâncias em relação a ocorrência de falha ou sua localização (GOUSSEAU et al., 2016).

Tabela 1 – Organização dos Dados

	Quantidade de amostras	Rolamento com falha	Local da falha
Experimento 1	2156	Rolamento 3 Rolamento 4	Pista Interna Elemento Rolante
Experimento 2	984	Rolamento 1	Pista Externa
Experimento 3	4448	Rolamento 3	Pista Externa

Fonte: Elaborado pelos autores (2023).



PARÂMETROS DE HJORTH

Os parâmetros de Hjorth tem como objetivo realizar a análise de sinais de eletroencefalograma diretamente no domínio do tempo, obtendo informações das propriedades básicas do sinal sem a necessidade de realizar transformadas para o domínio da frequência (HJORTH, 1970). Os parâmetros de Hjorth são apresentados a seguir.

Atividade é definida como o momento espectral de ordem zero (m_0), representando a potência do sinal do domínio do tempo, sendo expressa pela variância da amplitude do sinal, dada Eq. (1).

$$\text{Atividade} = m_0 = \sigma^2(y) \quad (1)$$

Mobilidade representa o momento espectral de 2ª ordem (m_2), expressa pela Eq. (2), como a raiz quadrada da razão entre variância da derivada de 1ª ordem do sinal pela variância da amplitude do sinal, sendo uma medida do desvio padrão da inclinação em relação ao desvio padrão da amplitude, também sendo interpretada como a frequência média.

$$\text{Mobilidade} = m_2 = \sqrt{\frac{\sigma^2(\dot{y})}{\sigma^2(y)}} \quad (2)$$

Complexidade representa o momento espectral de 4ª ordem (m_4), sendo uma medida para a similaridade da forma de onda com uma onda senoidal, expressando uma mudança na frequência do sinal analisado por meio da Eq. (3), dada pela a raiz quadrada da razão entre variância da derivada de 2ª ordem do sinal pela variância derivada de 1ª ordem da amplitude do sinal.

$$\text{Complexidade} = m_4 = \sqrt{\frac{\sigma^2(\ddot{y})}{\sigma^2(\dot{y})}} \quad (3)$$

METODOLOGIA

Os parâmetros de Hjorth foram aplicados sobre os sinais de vibração dos rolamentos que apresentaram falha, para determinar o momento em que ocorre a falha em cada rolamento e assim rotular os sinais entre amostras saudáveis e amostras com falha. Após os processos de rotulação das amostras utilizando os parâmetros de Hjorth, foram extraídos 10 atributos do domínio do tempo de cada sinal vibração com duração de 1 segundo. Os 10 atributos junto do rótulo do sinal formam uma instância do bando de dados apresentado para os classificadores.

Os classificadores tiveram seus hiperparâmetros ajustados por meio da busca em grade, com o objetivo de obter a melhor estrutura para realizar a classificação das amostras. Um processo de redução de dimensionalidade foi aplicado para obter o subconjunto dos atributos com desempenho similar ao apresentado pelo classificador utilizando todos os atributos.

EXTRAÇÃO DE ATRIBUTOS E CLASSIFICADORES

Os atributos extraídos dos sinais de vibração foram: Desvio Padrão (DP), RMS, Assimetria, Curtose, Valor de pico (V_p), Comprimento de Onda (WL), Fator de Crista (FC), Fator K (FK), Fator de Impulso (FI) e Fator de Forma (FF), sendo apresentados pelo Quadro 1, onde x indica o sinal de vibração completo, \bar{x} expressa o valor médio a amplitude do sinal, L o comprimento do sinal e x_i um ponto do sinal.



XIII Seminário de Extensão e Inovação XXVIII Seminário de Iniciação Científica e Tecnológica da UTFPR

Ciência e Tecnologia na era da Inteligência Artificial: Desdobramentos no Ensino Pesquisa e Extensão
20 a 23 de novembro de 2023 - Campus Ponta Grossa, PR



SEI-SICITE
2023

Os classificadores implementados são de aprendizado supervisionado, portanto, além de apresentar características de cada amostra, se faz necessário apresentar as classes que cada amostra pertence para realizar o processo de treinamentos dos modelos. Os classificadores implementados são:

Quadro 1 – Atributos extraídos no domínio do tempo.

Atributo	Descrição	Equação
DP	Medida da dispersão dos dados em relação ao valor médio	$\sqrt{\frac{1}{L} \sum_{i=1}^L x_i - \bar{x} ^2}$
RMS	Quantifica a potência média do sinal	$\sqrt{\frac{\sum_{i=1}^N x_i^2}{N}}$
Assimetria	Verifica o quanto os dados se distanciam de uma distribuição normal	$\frac{\frac{1}{L} \sum_{i=1}^L x_i - \bar{x} ^3}{\left(\sqrt{\frac{1}{L} \sum_{i=1}^L x_i - \bar{x} ^2}\right)^3}$
Curtose	Medida da concentração dos dados em relação a média de uma distribuição normal	$\frac{\frac{1}{L} \sum_{i=1}^L x_i - \bar{x} ^4}{\left(\frac{1}{L} \sum_{i=1}^L x_i - \bar{x} ^2\right)^2}$
V_p	Maior valor em modulo do sinal	$ max(x) $
WL	Fornecer informações sobre a frequência média do sinal	$\sum_{i=1}^P x_{i+1} - x_i $
FC	Possui sensibilidade para detectar falhas iniciais	$\frac{x_{pico}}{RMS}$
FK	Combina a sensibilidade do valor de pico e do valor RMS	$x_{pico} \cdot RMS$
FI	Compara o valor máximo com o valor média do sinal	$\frac{x_{pico}}{\bar{x}}$
FF	Métrica dependente do formato de onda apresentada pelo sinal	$\frac{RMS}{\bar{x}}$

Fonte: Elaborado pelos autores (2023).

1. **Floresta Aleatória**, do inglês *Random Forest* (RF), é um modelo que realiza a criação de diversas árvores de decisão independentes entre si, gerando conjunto de regras até que realiza a classificação das amostras.
2. **Máquinas de Vetores de Suporte**, do inglês *Support Vector Machine* (SVM), é um algoritmo que busca realizar a separação entre as amostras buscando o melhor hiperplano separador dentre todos os possíveis.
3. **k-ésimo Vizinho mais Próximo**, do inglês *k-Nearest Neighbors* (*k*-NN), realiza a classificação com base na distância entre amostras classificadas e a amostra a ser classificada, realizando a classificação com base na similaridade com as amostras já classificadas.

RESULTADOS E DISCUSSÕES

SEPARAÇÃO DO SINAL

Aplicando os parâmetros de Hjorth na evolução temporal dos sinais de cada rolamento, é possível visualizar um momento pré-falha, onde é observada um aumento gradativo da Atividade e Mobilidade, e uma redução na Complexidade. O momento da falha é detectado com uma variação abrupta dos parâmetros, sendo esse instante de tempo utilizado para estabelecer a fronteira de separação. A Tabela 2, apresenta o dia da falha e a quantidade de amostras saudáveis e com falha que cada rolamento gerou.



XIII Seminário de Extensão e Inovação XXVIII Seminário de Iniciação Científica e Tecnológica da UTFPR

Ciência e Tecnologia na era da Inteligência Artificial: Desdobramentos no Ensino Pesquisa e Extensão
20 a 23 de novembro de 2023 - Campus Ponta Grossa, PR



SEI-SICITE
2023

Os dados dos 3 rolamentos analisados foram concatenados para dar maior poder de generalização para os classificadores, dessa forma totalizando 4154 amostras saudáveis e 1142 amostras com falha.

Tabela 2 – Separação das amostras

	Dia da Falha	N° de amostras saudáveis	N° de amostras com falha
Rolamento 3 - Experimento 1	33	1.910	246
Rolamento 4 - Experimento 1	25	1.540	616
Rolamento 1 - Experimento 2	4,8	704	280

Fonte: Elaborado pelos autores (2023).

CLASSIFICAÇÃO

Os dados foram separados de forma aleatória, reservando 80% dos dados para treinamento e 20% para validação dos classificadores. Os classificadores implementados obtiveram acurácia acima de 95% utilizando os 10 atributos. O redução de dimensionalidade por meio da busca exaustiva de atributos, um método do tipo *wrapper*, foi aplicado para encontrar o subconjunto formado por 2 atributos que alcançasse o melhor desempenho dos classificadores.

Os classificadores RF e *k*-NN selecionaram os atributos RMS e WL, enquanto o SVM utilizou DP e WL, os subconjuntos selecionados pelos classificadores tiveram os dados separados novamente utilizando a proporção de 80% dos dados para treinamento e 20% para teste. A Tabela 3 apresenta as métricas de desempenho obtidos, a variação do desempenho comparada com a classificação utilizando os 10 atributos, assim como o tempo empenhado para treinamento e validação em cada classificador.

Os classificadores apresentaram reduções de 1% nas métricas de desempenho dos classificadores após o processo de redução de dimensionalidade, porém a acurácia se manteve acima de 95% para todos os modelos. Quando comparado o desempenho entre os resultados obtidos utilizando os dados do eixo x e y, tem-se que o desempenho dos classificadores utilizando os dados provenientes do eixo y foi superior em relação a classificação utilizando os dados provenientes do eixo x.

O tempo empenhado para realizar o treinamento dos classificadores RF e *k*-NN sofreram uma redução significativa em relação ao treinamento utilizando todos os atributos. A variação apresentada para o tempo de classificação das amostras também foi analisando, sendo o *k*-NN o que apresentou a maior redução no tempo de classificação, porém as diferenças se deram na ordem de milissegundos para todos os classificadores.

Tabela 3 – Separação das amostras

Métrica	Dados do eixo x			Dados do eixo y		
	Classificador			Classificador		
	RF	SVM	<i>k</i> -NN	RF	SVM	<i>k</i> -NN
Acurácia	0,98 (↓ 1%)	0,96 (↓ 1%)	0,98 (↓ 1%)	0,99 (-0%)	0,96 (↓ 1%)	0,98 (↓ 1%)
Precisão	0,98 (-0%)	0,96 (↓ 1%)	0,97 (↓ 1%)	0,98 (↓ 1%)	0,96 (↓ 1%)	0,98 (-0%)
Recall	0,97 (↓ 1%)	0,92 (↓ 2%)	0,97 (↑ 1%)	0,98 (↓ 1%)	0,93 (↓ 1%)	0,97 (↓ 1%)
F1-Score	0,97 (↓ 1%)	0,94 (↓ 1%)	0,97 (-0%)	0,98 (↓ 1%)	0,94 (↓ 2%)	0,98 (-0%)
$T_{\text{treino}}(s)$	89,52 (↓ 27%)	24,33 (↓ 5%)	0,51 (↓ 30%)	86,14 (↓ 18%)	24,39 (↑ 4,8%)	0,54 (↓ 27,4%)
$T_{\text{validação}}(s)$	0,018 (↑ 20%)	0,065 (↓ 10%)	0,004 (↓ 64%)	0,025 (↑ 25%)	0,06 (↑ 20%)	0,004 (↓ 55,5%)

Fonte: Elaborado pelos autores (2023).



CONCLUSÃO

Os parâmetros de Hjorth se apresentam como uma ferramenta eficaz para determinar o instante de tempo de uma série temporal para realizar a rotulação dos sinais, possibilitando a sua utilização para o treinamento de modelos de aprendizado supervisionado para realizar o monitoramento da condição de saúde dos rolamentos. Os modelos implementados nesse trabalho compartilharam o WL como atributo selecionado pelo processo de redução de dimensionalidade, isso pode ter se dado devido o WL ser uma medida da frequência média do sinal. O desempenho obtido utilizando apenas 2 atributos leva a uma redução de 80% na quantidade de dados a serem processados e armazenados, sendo um cenário interessante caso se deseje embarcar os classificadores.

Os melhores resultados foram apresentados utilizando os dados da direção y, sendo assim a provável direção que os sensores estavam posicionados durante os experimentos 2 e 3. Dada a pequena diferença apresentada entre as métricas de desempenho, o monitoramento pode ser realizado utilizando apenas 1 sensor por rolamento.

Agradecimentos

Os autores agradecem ao Centro Integrado de Pesquisa em Controle e Automação (CIPECA) a Universidade Tecnológica Federal do Paraná – Campus Cornélio Procópio pelo suporte financeiro ou estrutural para o desenvolvimento deste projeto.

Conflito de interesse

Os autores declaram que não há conflito de interesse neste trabalho.

REFERÊNCIAS

ATTALLAH, Omneya; IBRAHIM, Rania A.; ZAKZOUK, Nahla E. CAD system for inter-turn fault diagnosis of offshore wind turbines via multi-CNNs & feature selection. en. **Renewable Energy**, v. 203, p. 870–880, fev. 2023.

DUAN, Zhihe et al. Development and trend of condition monitoring and fault diagnosis of multi-sensors information fusion for rolling bearings: a review. en. **The International Journal of Advanced Manufacturing Technology**, v. 96, n. 1-4, p. 803–819, abr. 2018.

GOUSSEAU, William et al. Analysis of the Rolling Element Bearing data set of the Center for Intelligent Maintenance Systems of the University of Cincinnati. In: CM2016. Charenton, France: [s.n.], out. 2016.

HJORTH, Bo. EEG analysis based on time domain properties. en. **Electroencephalography and Clinical Neurophysiology**, v. 29, n. 3, p. 306–310, set. 1970.

HUTCHINSON M; ZHAO, F. **GWEC Global Wind Report, Hutchinson, M. & Zhao, F.** [S.l.]: GWEC: Brussels, Belgium, 2023.

QIU, Hai et al. Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics. en. **Journal of Sound and Vibration**, v. 289, n. 4-5, p. 1066–1090, fev. 2006.

VAIMANN, Toomas; BELAHCEN, Anouar; KALLASTE, Ants. Necessity for implementation of inverse problem theory in electric machine fault diagnosis. In: IEEE International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives (SDEMPED). Guarda, Portugal: IEEE, set. 2015. P. 380–385.