



UNIVERSIDADE FEDERAL DO PARÁ
NÚCLEO DE DESENVOLVIMENTO AMAZÔNICO EM ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE
INFRAESTRUTURA E DESENVOLVIMENTO ENERGÉTICO

THIAGO BARROSO COSTA

**PROPOSTA DE METODOLOGIA PARA DIAGNÓSTICO DE FALHA EM
ROLAMENTOS DE BAIXA VELOCIDADE**

TUCURUÍ

2024

THIAGO BARROSO COSTA

**PROPOSTA DE METODOLOGIA PARA DIAGNÓSTICO DE FALHA EM
ROLAMENTOS DE BAIXA VELOCIDADE**

Dissertação apresentada ao Programa de Pós-graduação em Engenharia de Infraestrutura e Desenvolvimento Energético do Núcleo de Desenvolvimento Amazônico em Engenharia, da Universidade Federal do Pará, como requisito para a obtenção do título de Mestre em Infraestrutura e Desenvolvimento Energético. Área de concentração: Tecnologias Sustentáveis.

Prof. Dr. Alexandre Luiz Amarante Mesquita

Tucuruí
2024

Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD
Sistema de Bibliotecas da Universidade Federal do Pará
Gerada automaticamente pelo módulo Ficat, mediante os dados fornecidos pelo(a)
autor(a)

- C837p Costa, Thiago Barroso.
Proposta de metodologia para diagnóstico de falha em
rolamentos de baixa velocidade / Thiago Barroso Costa. —
2024.
86 f. : il. color.
- Orientador(a): Prof. Dr. Alexandre Luiz Amarante
Mesquita
Dissertação (Mestrado) - Universidade Federal do Pará,
Núcleo de Desenvolvimento Amazônico em Engenharia,
Programa de Pós-Graduação em Engenharia de
Infraestrutura e Desenvolvimento Energético, Tucuruí, 2024.
1. Vibração. 2. Rolamentos de baixa velocidade. 3.
Diagnóstico de falha. 4. Seleção de features. 5.
Aprendizagem de máquina. I. Título.

CDD 620.3

THIAGO BARROSO COSTA

PROPOSTA DE METODOLOGIA PARA DIAGNÓSTICO DE FALHA EM ROLAMENTOS DE BAIXA VELOCIDADE

Dissertação apresentada ao Programa de Pós-graduação em Engenharia de Infraestrutura e Desenvolvimento Energético do Núcleo de Desenvolvimento Amazônico em Engenharia, da Universidade Federal do Pará, como requisito para a obtenção do título de Mestre em Infraestrutura e Desenvolvimento Energético. Área de concentração: Tecnologias Sustentáveis.

Prof. Dr. Alexandre Luiz Amarante Mesquita

Aprovada em 19 de abril de 2024.

BANCA EXAMINADORA



Documento assinado digitalmente
ALEXANDRE LUIZ AMARANTE MESQUITA
Data: 01/11/2024 00:53:42-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Alexandre Luiz Amarante Mesquita/UFPA – Orientador



Documento assinado digitalmente
ANDRE LUIZ AMARANTE MESQUITA
Data: 01/11/2024 07:56:26-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. André Luiz Amarante Mesquita/UFPA



Documento assinado digitalmente
PAULO CEZAR MONTEIRO LAMIM FILHO
Data: 01/11/2024 22:12:41-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Paulo Cezar Monteiro Lamim Filho/UFSJ



Documento assinado digitalmente
LUIZ ALBERTO PINTO
Data: 04/11/2024 17:22:08-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Luiz Alberto Pinto/IFES

AGRADECIMENTOS

Grato a minha família e amigos por tudo. Grato aos colegas bolsistas do projeto UFPA-Dynamox pelo apoio, especialmente ao Jonatas Cruz pelo auxílio direto nos ensaios de laboratório e aquisição de dados.

Agradeço pela supervisão, orientação e conselhos dos professores Alexandre Mesquita, André Mesquita e Walter Sousa ao longo desses últimos dois anos de trabalho.

Agradecimentos à Dynamox® S.A. pelo financiamento ao projeto de pesquisa UFPA-Dynamox, especialmente ao Danilo Braga e à Jullyane Figueiredo por todo o suporte dado.

RESUMO

O monitoramento de rolamentos de baixa velocidade por análise de vibração torna-se desafiador devido ao baixo nível de energia do sinal que carrega os dados relativos à falha, tornando-o suscetível a interferências de outras fontes, prejudicando a interpretação da informação. Diante disso, uma alternativa é calcular características do sinal sensíveis a alterações de padrão relacionadas ao início e progressão de falha. Assim, o presente trabalho extraiu diferentes tipos de *features*, entre elas, duas não-lineares e onze extraídas do sinal no domínio do tempo. As quais foram ranqueadas e selecionadas por sua sensibilidade em diferenciar entre classes, o que se estimou por meio do valor estatístico do teste *t* de Welch. Entre elas, o Máximo Expoente de Lyapunov, que, no presente trabalho, passou por uma alteração em uma de suas etapas de cálculo, resultando em um aumento de sua sensibilidade em alguns casos. Ademais, a influência do tamanho da janela do sinal de vibração na separabilidade dos indicadores também foi analisada (conteúdo escasso na literatura de monitoramento de rolamentos de baixa velocidade). Após a seleção de *features*, os dados passaram por uma transformação linear mediante PCA (Análise de Componentes Principais), visando reduzir a dimensionalidade dos dados para três dimensões e diminuir os efeitos de redundâncias de preditores altamente correlacionados. Na sequência, os dados representados no espaço das componentes principais foram projetados na carta de controle estatístico T^2 de Hotelling. Esta permitiu a detecção e eliminação de potenciais *outliers*, os quais consistiram em pontos acima de uma linha limite estimada com base na distribuição estatística F. Finalmente, modelos de classificação binário e multiclasse de Máquina de Vetores de Suporte foram treinados com dados coletados em ensaio de laboratório com rolamentos sem e com defeito incipiente em três diferentes níveis. Os modelos apresentaram ótima performance, sobretudo o binário diante de dados de teste oriundos de mancais de tambores de transportadores de correia em plena operação industrial.

Palavras-chave: vibração; rolamento de baixa rotação; diagnóstico de falha; seleção de *features*; aprendizagem de máquina; máximo expoente de Lyapunov.

ABSTRACT

Monitoring low-speed bearings with vibration analysis is more challenging due to the low energy level of the vibration signal that carries the failure data, making it more susceptible to interference from other sources, impairing the interpretation of information. Thus, an alternative is to calculate signal predictors that may be sensitive to pattern changes relative to failure onset and progression. Hence, the present work extracted different types of features, among them two nonlinear features and eleven extracted from the signal in the time domain. Those features were ranked and selected based on their sensibility to class differentiation, which was estimated using the *t*-Welch statistic value. Among them is the Largest Lyapunov Exponent, which, in this work, had a modification in one of its calculation steps, improving its sensitivity in some cases. In addition, the influence of the vibration signal window size on the class separability of the indicators was evaluated (which is a scarce content in low-speed bearing monitoring literature). After feature selection, the data were subjected to a linear transformation through PCA (Principal Component Analysis), aiming to reduce the data dimensionality to three dimensions and to minimize the redundancy effects of highly correlated features. In sequence, the data represented in the space of principal components were projected on a Hotelling T^2 statistic control chart. The chart allowed the detection and rejection of potential outliers, which consisted of points above a limit line estimated based on F statistic distribution. Finally, binary and multiclass Support Vector Machine classification models were trained with experimental data acquired from normal conditions and three levels of incipient fault in bearing. The models performed well, mainly the binary model with test data obtained from belt conveyor pulley bearings in industrial operation.

Keywords: vibration; low-speed bearing; fault diagnosis; feature selection; machine learning; largest Lyapunov exponent.

LISTA DE ILUSTRAÇÕES

Figura 3.1 - Hiperplano e margem em função dos vetores de suporte.....	34
Figura 3.2 - Parâmetros geométricos relacionados à superfície de decisão	35
Figura 3.3 - Relação geométrica entre a margem e os vetores de suportes [Fonte: Adaptada de Bishop (2006)].....	36
Figura 3.4 – Relação geométrica entre o fator ξ e a margem	37
Figura 3.5 - Influência da variação do parâmetro C na direção do hiperplano	38
Figura 4.1 - Metodologia proposta: do processamento ao diagnóstico.	40
Figura 4.2 - Indicadores calculados (em cor verde são aqueles pouco explorados).	42
Figura 4.3 - Algoritmo para determinar a matriz inicial de distâncias.	44
Figura 4.4 - Ajuste da reta aos dados em função da faixa de expansão.....	46
Figura 4.5 – Comparação entre distribuições binárias com: a) t maior; b) t menor. ...	48
Figura 4.6 - Exemplo de gráfico biplot no plano das componentes principais.	49
Figura 4.7 - Carta T^2 de Hotelling com linha limite.	50
Figura 4.8 - Misturador MT120 utilizado como bancada de ensaios.	52
Figura 4.9 - Posicionamento do acelerômetro próximo ao rolamento testado.	53
Figura 4.10 - Etapas de aquisição de dados em ensaios com o misturador.	54
Figura 4.11 – Superfícies da pista externa do rolamento danificadas.	54
Figura 4.12 - Janelamento do sinal baseado no número de revoluções do eixo.....	56
Figura 4.13 - Bancada experimental sistema MEM.....	57
Figura 4.14 - Estrutura básica de um transportador de correia.	58
Figura 5.1 – Influência do janelamento do sinal com base no valor t de Welch.	60
Figura 5.2 – Janelamento dos sinais de vibração para 1 revolução.....	61
Figura 5.3 – Ranque dos 5 melhores preditores (parte 1).....	62
Figura 5.4 – Ranque dos 5 melhores preditores (parte 2).....	62
Figura 5.5 – Comparação entre features não-lineares.....	63
Figura 5.6 – Comparação entre os parâmetros de Hjorth.	63
Figura 5.7 – Conjuntos das classes pelos indicadores não-lineares.....	65
Figura 5.8 – Validação e teste do modelo sem carga.	66
Figura 5.9 – Validação e teste do modelo com 40% de carga.	66
Figura 5.10 – Validação e teste do modelo com 80% de carga.	67
Figura 5.11 - Validação e teste do modelo construído com L_{CONC}	67
Figura 5.12 – Acurácia de teste do Msvm com dados da bancada MEM.....	68

LISTA DE TABELAS

Tabela 2.1 – Outros métodos de monitoramento de rolamentos em baixa rotação. .	19
Tabela 4.1 – Continuação do algoritmo para o cálculo do MEL, por Caesarendra (2015).....	45
Tabela 4.2 - Algoritmo para o cálculo da DFH.....	47
Tabela 4.3 – Detalhamento do janelamento dos sinais.....	55
Tabela 4.4 - Rolamentos SKF® W33 acoplados aos mancais dos tambores.	57
Tabela 5.1 - Treinamento e teste de classificadores com sinais da direção horizontal.	69
Tabela 5.2 - Treinamento e teste de classificadores com sinais da direção vertical.	69

LISTA DE SIGLAS

- MEL – Máximo Expoente de Lyapunov
- E-MEL – Euclid-Máximo Expoente de Lyapunov
- P-MEL – Pearson-Máximo Expoente de Lyapunov
- DFH – Dimensão Fractal de Higuchi
- SVM – Máquina de Vetores de Suporte (do inglês: *Support Vector Machine*)
- MEM – Motor-Eixo-Mancal
- L0 – Sem carga
- L40 – Com 40% de carga
- L80 – Com 80% de carga
- DN – Conjunto de dados de rolamento normal (sem defeito)
- D25 – Conjunto de dados de rolamento com 25% de defeito
- D50 – Conjunto de dados de rolamento com 50% de defeito
- D100 – Conjunto de dados de rolamento com 100% de defeito
- RMS – Raíz Média Quadrada (do inglês: *Root Mean Square*)
- FF – Fator de Forma
- FI – Fator de Impulso
- FC – Fator de Crista
- FL – Fator de Liberação
- PCA – Análise de Componentes Principais (do inglês: *Principal Component Analysis*)
- CP1 – Primeira Componente Principal
- CP2 – Segunda Componente Principal
- AC – Corrente Alternada (do inglês: *Alternating Current*)
- BPFO – Frequência de Passagem do Elemento Rolante na Pista Externa do Rolamento (do inglês: *Ball Pass Frequency Outer Race*)
- SNR – Razão entre Sinal e Ruído (do inglês: *Signal-to-Noise Ratio*)
- ACC – Acurácia
- TVP – Taxa de Verdadeiros Positivos
- VPP – Valor Predito Positivo

LISTA DE SÍMBOLOS

γ_1 – Distorção

γ_2 – Curtose

A_0 – Atividade (1º parâmetro de Hjorth)

M_0 – Mobilidade (2º parâmetro de Hjorth)

C_0 – Complexidade (3º parâmetro de Hjorth)

σ – Desvio Padrão

t – Valor estatístico do teste de Welch

C – Penalidade (parâmetro do algoritmo SVM)

T_{lim}^2 – Valor da linha limite na carta T^2 de Hotelling

λ_{max} – Máximo Expoente de Lyapunov (cálculo do MEL)

M – Número de estados da matriz pseudo-espaco-de-estado (cálculo do MEL)

m – Dimensão implícita (cálculo do MEL)

N – Tamanho do sinal (cálculo do MEL)

J – Atraso na reconstrução da trajetória dinâmica do sinal (cálculo do MEL)

μ – Restrição temporal (cálculo do MEL)

k_{max} – Número máximo de iterações (cálculo do MEL e da DFH)

f_s – Taxa de amostragem

f_{max} – Limite máximo da faixa de frequência

T_s – Período de aquisição

M_{svm} – Modelo de classificação binário SVM treinado com dados do misturador

SUMÁRIO

1 INTRODUÇÃO	8
1.1 Considerações gerais.....	8
1.2 Objetivos.....	10
1.3 Estrutura do trabalho.....	10
2 REVISÃO BIBLIOGRÁFICA	11
2.1 Revisão geral.....	11
2.2 Contribuições deste trabalho	20
3 FUNDAMENTAÇÃO TEÓRICA	22
3.1 Atributos de sinais no domínio do tempo	22
3.2 Valor t do teste estatístico de Welch.....	24
3.3 Análise de Componentes Principais (PCA)	25
3.4 Carta T^2 de Hotelling.....	29
3.5 Máximo Expoente de Lyapunov (MEL).....	30
3.6 Dimensão Fractal de Higuchi (DFH)	31
3.7 Distância de Pearson.....	32
3.8 Máquina de Vetores de Suporte (SVM)	33
4 METODOLOGIA	40
4.1 Metodologia para o diagnóstico de falha.....	40
4.2 Extração de preditores dos sinais.....	41
4.3 Seleção de features	47
4.4 Máquina de Vetores de Suporte (SVM)	50
4.5 Coleta de dados do misturador	52
4.6 Sistema Motor-Eixo-Mancal (MEM)	56
4.7 Dados de tambores de transportadores de correia	57
5 RESULTADOS E DISCUSSÕES	59
5.1 Análise da influência do tamanho da janela do sinal	59
5.2 Ranqueamento e seleção de <i>features</i>	61
5.3 Treinamento do modelo de classificação SVM multiclasse	65
5.4 Treinamento de modelo binário de classificação	67
6 CONCLUSÃO	70
REFERÊNCIAS	72

1 INTRODUÇÃO

1.1 Considerações gerais

Segundo Bento (2012), a manutenção preditiva consiste no monitoramento dos sinais vitais de um ativo visando apontar uma falha potencial ainda em estágio inicial. Conforme Holanda (2016), sua principal vantagem é elevar o índice MTBF (Tempo Médio Entre Falhas), reduzindo custos e aumentando a confiabilidade ao evitar paradas desnecessárias. Nesse contexto, a Análise de Vibração é uma das técnicas mais conhecidas e usuais da manutenção preditiva em máquinas rotativas, pois o sinal de vibração carrega informação indicadora do nível de saúde da máquina (LIMA et al., 2006; LIMA, 2013; ANTONIOLLI, 1999).

A análise dos dados coletados do equipamento ainda é amplamente feita por profissionais qualificados e especialistas. Contudo a nova tendência da Indústria 4.0 e *IoT* (Internet das Coisas) demanda por avanços nas técnicas de extração automatizada de conhecimento em cima de *Big Data*, visando tornar as tomadas de decisão mais eficientes, evitando desperdícios de bens e serviços. Nesse cenário, destaca-se a Inteligência Artificial (IA), direcionando a computadores tarefas que antes eram exclusivas de analistas (LADEIRA et al., 2017; VASCONCELOS, 2017).

A técnica de Aprendizagem de Máquina (*Machine Learning*) está inserida no contexto de Inteligência Artificial. Tem sido usada para tarefas de reconhecimento de padrão, classificação e regressão sem haver programação explícita para tal função. Pois a máquina “aprende” ao ser submetida a diversos dados de treinamento, gerando um modelo com alta capacidade de generalização. Diante disso, a implementação de *machine learning* em processos reduz custos com recursos, além de aumentar a produtividade (FORTI, 2018; SILVA, 2021).

Nas últimas décadas, o monitoramento e diagnóstico de falha em mancais de rolamento de máquinas em rotações elevadas (acima de 1000 rpm, por exemplo) tem sido realizado com sucesso por métodos tradicionais de análise dos sinais temporais de vibração. Dentre os quais destacam-se a análise de parâmetros estatísticos no domínio do tempo e no domínio da frequência. Porém, segundo Caesarendra e Tjahjowidodo (2017), a maioria dos métodos tradicionais não são eficientes diante de baixas rotações, pois os sinais de vibração têm sua energia drasticamente reduzida e ruídos podem mascarar os sinais de interesse, além do sinal apresentar caráter não-linear, não-estacionário e caótico.

Segundo Jin et al. (2021), nas últimas décadas, equipamentos de grande porte tem sido o principal pilar da economia de diversas nações. A maioria desses maquinários dispõem de rolamentos grandes e de baixa rotação, sujeitos muitas vezes a cargas elevadas. Em máquinas rotativas em geral, as falhas em rolamentos representam cerca de 42% das ocorrências (SOUALHI et al., 2014). Segundo Liu e Zhang (2020), certas falhas podem ser catastróficas e causar prejuízos irreparáveis, por isso garantir a segurança de processos que envolvem máquinas rotativas tem atraído mais atenção. Além disso, em relação aos rolamentos de grande porte, a troca ou manutenção dos mesmos exige maiores custos e tempo. Nesse sentido, o controle e o monitoramento desses componentes são imprescindíveis para evitar falhas inesperadas e potencializar a redução de custos (HEMMER et al., 2018).

Além disso, cabe destacar que os rolamentos e demais componentes de uma máquina rotativa operam sob diversificada condição de contorno. Em outras palavras, um rolamento pode vir a falhar precocemente atuando em um meio com alto potencial de contaminação, choques inesperados, intemperismo, defeito de fabricação ou montagem (CAESARENDRA, 2015). Isso pode levar ao aumento de ocorrências de falhas inesperadas se o componente não estiver sendo monitorado. Outro ponto, conforme revela Caesarendra (2015), os rolamentos das máquinas que operam continuamente por longo tempo, podem vir a apresentar desgastes múltiplos e tornar mais desafiadora a análise eficaz dos sinais de vibração.

Então, como detectar e diagnosticar falhas em rolamentos em baixa rotação de maneira eficaz e eficiente?

Diante disso, levando em consideração que sob baixa rotação os sinais de vibração dos mancais de rolamento elevam seus níveis de irregularidades, desordem, não-linearidade e ruídos, torna-se mais dificultoso para um analista avaliá-lo pelos métodos tradicionais. Assim, verifica-se que algoritmos computacionais de *machine learning* se mostram aptos para resolver esse tipo de problema, seja para problemas de classificação ou de regressão. No entanto, é preciso definir o modelo mais apropriado e treiná-lo com *features* que representem adequadamente a(s) característica(s) desses sinais.

Neste sentido, este trabalho visa comparar o desempenho de preditores estatísticos, fatores e parâmetros de Hjorth extraídos do sinal no domínio do tempo, com os ditos não-lineares, a saber, Dimensão Fractal de Higuchi e Máximo Expoente

de Lyapunov. Ao final, foram selecionadas as melhores *features* dentre as não-lineares e as extraídas do sinal no domínio do tempo e, após combinadas, usadas como base para treinamento de um modelo de aprendizado supervisionado SVM. Além disso, a influência do tamanho da janela do sinal na performance dos preditores também foi analisada. E por fim, uma proposta de alteração no algoritmo para o cálculo do Máximo Expoente de Lyapunov também foi avaliada, apresentando resultados mais vantajosos na maioria dos casos considerados.

1.2 Objetivos

1.2.1 Objetivo Geral

Desenvolvimento de metodologia para diagnóstico de falha em rolamento de baixa rotação mediante algoritmo de aprendizado de máquina supervisionado.

1.2.2 Objetivos Específicos

- Estimar estatisticamente quais as melhores *features* em relação à capacidade de distinção entre classes de diferentes condições do rolamento em baixa rotação.
- Investigar a influência do tamanho da janela do sinal de vibração com base na separabilidade de classes por parte dos preditores.
- Verificar o impacto da substituição da distância euclidiana pela distância de Pearson no cálculo do preditor Máximo Expoente de Lyapunov.

1.3 Estrutura do trabalho

A seguir, é realizada uma revisão da bibliografia sobre monitoramento de rolamento em baixa rotação visando conhecer o estado da arte e direcionar esta pesquisa de modo a contribuir para a temática. Na sequência, a seção 3 sintetiza as bases teóricas das principais técnicas e métodos presentes na literatura que foram utilizados para alcançar os objetivos deste trabalho. A seção 4, por sua vez, apresenta em detalhes todos os passos da metodologia empregada neste trabalho. Em seguida, a seção 5 exibe os resultados organizados sucintamente em formas de figuras, gráficos ou tabelas; além das discussões e interpretações de tais resultados. Por fim, a conclusão do trabalho é abordada na seção 6, que traz as considerações finais acerca dos objetivos atingidos ao longo do trabalho, bem como propostas de trabalhos futuros.

2 REVISÃO BIBLIOGRÁFICA

2.1 Revisão geral

Segundo Bechhoefer et al. (2016), rolamentos de baixa rotação e de grande porte representam uma das mais frequentes ocorrências de falha em grandes máquinas rotativas. Devido a esta relevância, seus sinais de vibração têm sido amplamente estudados. Esse tipo de rolamento pode estar presente em turbinas eólicas, guindastes, recuperadoras, empilhadoras e escavadoras. (JIN et al., 2021; LIU e ZHANG, 2020; MERGL e KAŠPÁREK, 2022; ROTHSCHEDL, 2016; PRABHAKAR et al., 2022; CAESARENDRA et al., 2013b; SOKOLSKI e SMOLNICKI, 2021). Podem pesar acima de 11,3 toneladas com diâmetro externo variando de 200 mm a 8000 mm. Tipicamente operam com rotações de 0,5 a 100 rpm de modo constante, intermitente e/ou reversível. (MOODIE, 2009; CAESARENDRA, 2015).

Em geral, segundo Liu e Zhang (2020), tais rolamentos possuem os seguintes modos de falha: deformação plástica, lubrificação, arco elétrico (fuga de corrente), fraturas, rachaduras, desgaste e montagem inadequada. Além disso, tomando como exemplo rolamentos triaxiais de coroa interna, são comuns falhas nas pistas externa e interna por abrasão entre o dente e a superfície ou formação de pites, trinca por fadiga ou fratura do dente. (WANG et al., 2013). Nos elementos rolantes pode ocorrer desgaste de flanco, ou ainda trincas e pites (YE, 2005; ALBERT et al., 2017; ZONG et al., 2016; POTOČNIK et al., 2013). Os parafusos de fixação podem sofrer com trincas por fadiga ou fraturas (LIN et al., 2013; QIU et al., 2012). O anel de vedação pode falhar por distorção ou abrasão, enquanto que o reservatório de óleo pode ter vazamentos ou obstruções (WANG et al., 2017).

As técnicas mais comuns de monitoramento da condição desses rolamentos consistem em análise de vibração, emissão acústica, método de elementos finitos, temperatura, análise de óleo e ultrassom (HEMMER, 2020; POTOČNIK et al., 2013). Especificamente, conforme Liu e Wang (2017), os métodos de monitoramento e diagnóstico de falha baseados em sinais de vibrações adquiridos por sensores são diversos, incluindo técnicas no domínio do tempo, no domínio da frequência e no domínio do tempo-frequência.

Vale ressaltar que é mais desafiador analisar a condição de rolamentos de grande porte e de baixa rotação do que um rolamento típico em alta rotação, pois o sinal de vibração é de baixa energia. O sinal é não-linear, às vezes não-estacionário,

caótico e de difícil análise. Devido à baixa energia, a razão entre o ruído de fundo e o sinal requerido aumenta, e isso dificulta a extração de *features* do sinal. Por esta razão, o diagnóstico de falha desse tipo de rolamento torna-se um desafio maior (MOODIE, 2009; CAESARENDRA, 2015; SONG et al., 2018).

Em sua página na Internet, a empresa sueca SPM® afirma que:

Monitoramento de maquinário operando em baixa velocidade é mais complicado do que máquinas industriais em geral. Sinais oriundos de rolamentos e problemas de engrenagem são tipicamente de baixa energia, confuso e praticamente não-analisável. Distinguir esses sinais do ruído de fundo para extrair informações úteis do rolamento pode ser uma tarefa difícil. Dados de vibração coletados em baixa velocidade são muitas vezes erroneamente descartados como sendo 'apenas ruído'. Portanto, desgastes e danos do rolamento frequentemente deixam de ser detectados até que seja tarde demais. (SPM®, 2024).

A respeito dos métodos de monitoramento e diagnóstico de falha de grandes rolamentos sujeitos a cargas elevadas e operando em baixa rotação, Wang et al. (2017) apresentam uma revisão ampla, que inclui métodos de vibração, emissão acústica, termografia, análise de óleo, e propostas de desafios futuros. Especificamente para análise de vibração, são comparados métodos no domínio do tempo, da amplitude, da frequência e do tempo-frequência. Outrossim, Hemmer et al. (2018), abordam métodos já estabelecidos (como técnica do envelope do sinal e estatística RMS no domínio do tempo), e aqueles mais avançados: momentos estatísticos superiores; espectro de curtose e curtograma; Decomposição de *Wavelets*, *Empirical Mode Decomposition* (EMD), e análise de *cepstrum*.

Vale destacar alguns métodos de análise de vibração em rolamentos grandes de baixa rotação aplicados a maquinários industriais. Primeiramente, Caesarendra et al. (2015) utilizou o Máximo Expoente de Lyapunov como *feature* de monitoramento, Caesarendra et al. (2014) usaram Análise de *features* no Domínio Circular, e Caesarendra et al. (2013b) realizaram processamento do sinal por EMD/EEMD. Em turbinas eólicas, Zhao et al. (2016) realizaram Análise de Bi-Espectro.

Em relação às características do sinal de vibração, Caesarendra e Tjahjowidodo (2017) afirmam que é preciso ter parâmetros ou *features* de monitoramento confiáveis, permitindo a extração de informação precisa acerca da condição do rolamento. Diante disso, os mesmos publicaram uma revisão de métodos de extração de *feature* baseado em vibrações e suas aplicações, para estimar tendência de degradação de grandes rolamentos de baixa rotação. A pesquisa

consistiu em apresentar um estudo empírico de métodos de extração de *feature* para sinais de vibrações adquiridos de uma bancada de rolamento operando em baixa velocidade com falha naturalmente forçada. Assim, os métodos de extração de *features* foram elencados em seis categorias:

1. Domínio do tempo;
2. Domínio da frequência;
3. Representação tempo-frequência;
4. Medida de dissimilaridade no espaço de estado;
5. Medidas de complexidade;
6. outras features.

Entre as seis categorias, Caesarendra e Tjahjowidodo (2017), revelam que a terceira e a quarta são mais adequadas para sinais não-estacionários, não-lineares e caóticos. Os métodos da categoria 3 são: Transformada de Fourier em Tempo Curto; Transformada e Decomposição *Wavelet*; Transformada Hilbert-Huang baseada em EMD; e Distribuição Wigner-Ville. Grupo 4: Dimensão de Fractal; Dimensão de Correlação; Entropia Aproximada; e Máximo Expoente de Lyapunov.

O diagnóstico de falha pode ser visto como um problema de reconhecimento de padrão visando determinar a condição de uma máquina rotativa. Como uma poderosa ferramenta de reconhecimento de padrão, a Inteligência Artificial (AI) tem atraído bastante atenção de muitos pesquisadores e apresenta-se promissora em aplicações de reconhecimento de falha em máquinas rotativas (LIU et al., 2018).

As ferramentas mais comuns de *machine learning* para diagnóstico de falha de máquinas rotativas são KNN, SVM, Rede Neural Artificial e *Ensemble Learning* (WEI et al., 2019; ALTHUBAITI et al., 2022). Ademais, Althubaiti et al. (2022) apontam propostas baseadas em *Deep Learning*, são elas: CNN, *Auto-encoders*, DBN, RNN e GAN. Além disso, os mesmos abordam uma revisão dos principais modelos de prognósticos (previsões): modelos físicos, baseados em dados, ou híbridos; importantes para gerenciamento da segurança, agendamento e redução de custos na manutenção. Segundo Tobon-Mejia et al. (2010), prognóstico é definido como “Estimação do Tempo para Falha (ETTF) e o risco de existência ou surgimento posterior de um ou mais modos de falha.”

Para remover o ruído dos sinais de vibração a fim de identificar a característica de falha dos elementos de rolamento de baixa velocidade, Zvokelj et al. (2010)

propuseram a ICA multiescala baseada em EEMD (EEMD-MSICA), que permitiu a detecção de falhas em rolamentos e ofereceu um mecanismo de remoção de ruído de sinais multivariados que, combinado com a análise de envelope (EA), pôde ser usado como uma ferramenta de diagnóstico. Com o mesmo objetivo, Mishra et al. (2017) usaram um método híbrido baseado na eliminação de ruído mediante decomposição por *wavelet* e na análise de envelope para diagnosticar falhas em rolamento rígido de esferas operando a 60 rpm sujeito a uma carga radial de 50 N. Os dados experimentais de vibração foram adquiridos por vibrômetro com faixa de frequência de 0,05 Hz a 22 kHz e taxa de amostragem de 32.768 Hz durante 32 s. As falhas nas pistas externa e interna foram introduzidas como pequenas fendas retangulares cortadas através de eletroerosão. A velocidade de rotação foi de 60 rpm sujeita à carga radial de 50 N. O esquema de diagnóstico foi aplicado com sucesso para encontrar vários sintomas de falhas de rolamentos de esferas defeituosos usando dados experimentais.

Nikula et al. (2020) relataram que o diagnóstico em condições de baixa velocidade é prejudicado pelos longos períodos entre os impulsos relacionados a defeitos e o alto nível de ruído em relação à magnitude dos impulsos. Diante disso, os autores propuseram um critério baseado na autocorrelação para o diagnóstico de defeitos localizados em rolamentos rígido de esferas em baixa rotação. As etapas de cálculo do critério proposto incluíram filtragem, segmentação de sinal, seleção de *features* estatísticas e cálculo de autocorrelação. Os experimentos exploraram velocidades de rotação de 30, 60 e 90 rpm com ou sem carga. Os resultados revelaram que o critério baseado na autocorrelação forneceu indicações de defeitos associadas ao tipo correto de defeito. Já as abordagens automatizadas de diagnóstico de espectro do envelope com base na proporção de pico e nos indicadores de pico em relação à média e no curtograma rápido eram propensas a induzir a uma conclusão incorreta em condições de danos leves.

Wang et al. (2018) propuseram um método de diagnóstico de falha combinado baseado em MCKD-CEEMD-ApEn para rolamentos de grande porte operando a cargas elevadas. Os dados analisados foram obtidos de um rolamento de um guindaste portuário sob condições não-estacionárias e apresentando falhas combinadas. O rolamento iniciou sua operação em julho de 2013, e seus dados de vibração foram coletados de fevereiro de 2015 a dezembro de 2017,

aproximadamente a cada dois meses. Os sinais de vibração foram coletados a 1,5 rpm sob três condições: normal, falha inicial e falha grave. A frequência de amostragem foi de 512 Hz durante 40 s (um ciclo de rotação). O algoritmo multiclasse SVM foi treinado com um *dataset* de 25 vetores de *features* por classe. Os autores compararam sua abordagem baseada em *Approximate Entropy (ApEn)* com alternativa baseada em *Correlation Dimension (CD)*, *Sample Entropy (SE)* e *Largest Lyapunov Exponent (LLE)*. Os autores encontraram resultados superiores com o MCKD-CEEMD-ApEn usando um classificador SVM multiclasse.

Hemmer e Waag (2017) compararam o AE RMS e a aceleração RMS como indicadores de condição para o monitoramento da condição de um rolamento axial de rolos cônicos, com diâmetro de 650 mm, em um acionamento superior de uma perfuratriz *offshore*. Os dados de vibração foram adquiridos no sentido axial variando a velocidade de rotação de 50 rpm a 250 rpm em intervalos de 50 rpm, sujeitos a uma carga de 10 kN. Inicialmente foram coletados os dados relativos ao rolamento desgastado (que apresentava desgaste visível em forma de arco nas extremidades dos rolos). Em seguida, foram realizadas novas aquisições de dados para indentações artificiais, e esses danos artificiais foram combinados com lubrificação abaixo do nível mínimo indicado. Por fim, o rolamento danificado foi substituído por um novo, permitindo a coleta de dados de rolamentos saudáveis. O acelerômetro usado possuía sensibilidade de 100 mV/g, faixa de frequência de 2 Hz a 10 kHz e taxa de amostragem de 102,4 kHz durante 100 revoluções. Os autores concluíram que o AE RMS foi mais promissor e menos afetado pela velocidade de rotação, uma vantagem para máquinas sob condições operacionais variáveis. Assim, os sistemas de medição de emissão acústica podem complementar os sistemas baseados em vibração, de acordo com os autores.

Hemmer et al. (2018) buscaram diagnosticar defeitos em rolamentos por meio de três classificadores binários baseados em rede neural convolucional (CNN), máquina de vetor de suporte (SVM) e SVM baseado em *autoencoder* esparsa (SAE-SVM). No caso da CNN, os autores propuseram a transferência de aprendizado extraindo *features* de uma CNN pré-treinada em dados da imageNet. O equipamento de teste foi composto por um rolamento radial com falhas simuladas, cuja velocidade mínima testada foi de 100 rpm sujeita a 5 kN. O resultado da investigação mostrou que extrair *features* diretamente da CNN pré-treinada e depois usar o SVM para

classificação foi a melhor opção em relação à precisão, robustez, facilidade de implementação, e custo computacional.

Liu, Zhang e Carrasco (2020) propuseram a aplicação de limiarização de *wavelet* empírica como um método de redução de ruído de sinal de vibração para o diagnóstico de falhas em rolamentos de passo. O rolamento mencionado consiste em um rolamento de passo de baixa velocidade e de grande escala naturalmente danificado, tendo operado em um parque eólico por mais de 15 anos. A bancada de testes foi projetada com o referido rolamento de 1.129 mm de diâmetro externo operando em seis velocidades diferentes de 1 a 3,05 rpm. Os dados de vibração foram amostrados a 50 kHz com acelerômetro de sensibilidade de 1000 mV/g, e com resposta de frequência de 2 Hz a 10 kHz. Em relação à metodologia empregada, o limiar de *wavelet* empírico removeu os níveis mais severos de ruído, e os sinais de falha de menor energia extraídos foram analisados por meio da técnica de envelope, que permitiu a distinção das frequências de defeito.

Song et al. (2018) combinou as técnicas SF, M-PH e WPT para extrair *features* dos sinais de vibração com interferências de ruídos, enquanto o método de árvore de decisão (DT) foi usado para produzir regras para o diagnóstico automático. Quatro estados comuns de rolamentos foram investigados: normal, falha na pista externa, falha na pista interna e falha no elemento do rolo. O acelerômetro utilizado para aquisição dos dados de vibração tinha sensibilidade de 102,2 mV/g e faixa de frequência de 5 Hz a 60 KHz. Em todos os experimentos, a frequência de amostragem foi de 100 kHz e o eixo funcionou em velocidades fixas de 40, 50, 60, 70, 100, 150 e 200 rpm, respectivamente. Os autores também compararam o método proposto com SVM, HHT+SVM, ANN e ANN+HHT.

Wang, Tang e Huang (2014) apresentaram uma nova estratégia para o monitoramento da condição de rolamento de baixa velocidade com base em sinais de temperatura e torque. O equipamento de teste incluía um rolamento de giro usado em uma turbina eólica de 3 MW com um diâmetro de pista externa de 2.410 milímetros. A velocidade de rotação foi de 1,5 rpm, e foram simuladas diferentes condições de carga de milhares de kNm e centenas de kN. Em seguida, os autores concluíram que a ANN (Rede Neural Artificial) com otimização GA (Algoritmo Genético) pode ser usada para avaliar o estado de saúde do rolamento. Saufi et al. (2014) propuseram o método *Adaptive Resilient Stacked Sparse Autoencoder* (ArSSAE) usando *features* no

domínio do tempo extraídas de sinais de vibração e sinais de emissão acústica. *Features* estatísticas e curtograma foram usados como conjuntos de dados de entrada da rede neural. Os sinais de vibração foram registrados com uma taxa de amostragem de 25,6 kHz em uma bancada de testes de rolamentos em baixas velocidades de operação (48–480 rpm), simulando três níveis de danos. O método proposto superou os resultados do SVM, principalmente com o curtograma extraído do sinal de emissão acústica em comparação com outros conjuntos de dados em termos de erro de treinamento e previsão de teste.

Li et al. (2019) apresentaram um método aprimorado de aprendizado não-supervisionado para reconhecimento do estado de vida do rolamento de baixa rotação com base em um experimento sobre o teste de vida útil total do rolamento. O modelo denominado Função de Densidade de Pontos com Médias C Difusas (D-FCM) foi mais preciso do que a PCA para reconhecer o estado normal, o estado de degeneração e o estado de falha do rolamento.

Guo et al. (2022) propuseram a remoção de ruído de sinais de vibração por meio da Decomposição Modulada de Modo de Banda Larga (MBMD). Em seguida, uma rede neural BP foi aplicada para realizar o diagnóstico de falha de rolamento em baixa velocidade de um guindaste com um diâmetro de passo de 265 mm. Os dados de vibração foram coletados a uma taxa de amostragem de 1 kHz por 10 segundos, sob velocidade de rotação de 38 rpm, e submetidos a uma carga de 25 MPa. Foram avaliadas três condições do rolamento: normal, falha na pista interna e falha na pista externa. Para cada condição foram adquiridos 30 conjuntos de dados, dos quais 20 foram utilizados para treinamento e os 10 restantes serviram como amostras de teste. Por fim, os resultados mostraram que o modelo de rede neural BP baseado em MBMD teve maior precisão do que as abordagens baseadas em EEMD e VMD.

Klausen e Robbersmyr propuseram o método *Whitened Cross-Correlation Spectrum* (WCCS) para diagnosticar falhas em rolamentos. O método consiste em “clarear” o sinal de vibração usando a média síncrona e um modelo autorregressivo. O teste de vida útil do rolamento realizou 6 milhões de revoluções (14 dias de operação contínua) antes da parada completa. A faixa de frequência do acelerômetro usado para a coleta dos sinais de vibração ficou entre 2 Hz e 10 kHz, a sensibilidade foi de 100 mV/g, e o sinal de vibração foi amostrado em 51,2 kHz. A coleta dos dados foi realizada durante 50 rotações em velocidades do eixo de 500, 250, 100, 50 e 20

rpm. Os resultados mostraram que a falha precoce foi identificada em aproximadamente 8,33% da vida útil restante para todos os três tipos de falha.

Widodo et al. (2009) compararam a análise de emissão acústica e a análise de vibração para diagnóstico de falha usando técnicas de classificação envolvendo Máquina de Vetores de Relevância (RVM) e SVM. Os autores extraíram *features* estatísticas além de *features* baseadas em análise de componentes, incluindo ICA, PCA, KICA e KPCA. O experimento envolveu trincas simuladas e falhas por fragmentação em rolamentos de rolos cilíndricos sob velocidades de rotação de 20 e 80 rpm submetidos a uma carga de 5 kN. Os resultados sugeriram que RVM e emissão acústica superaram SVM e análise de vibração.

Shuuji et al. (2021) apresentaram uma rede neural convolucional com filtragem estatística aprimorada como um método de diagnóstico de falhas em rolamentos de baixa velocidade. Os autores empregaram três tipos de falhas em rolamentos feitas artificialmente por meio de usinagem por eletroerosão (falha na pista externa, falha entre pistas e falha nos rolos) adquiridas a partir de uma bancada de testes de falhas em rolamentos de baixa velocidade. O acelerômetro foi fixado na direção vertical do rolamento com frequência de amostragem de 100 kHz sob velocidades de rotação de 40, 50, 60, 70 e 100 rpm. Além disso, os autores investigaram o impacto dos níveis de ruído no sinal de vibração e compararam os resultados do SVM e diferentes métodos de aprendizagem profunda (CNN, SAE, RNN, DBN e BPNN).

Lu et al. (2016) propuseram um novo método de previsão baseado na Máquina de Vetores de Suporte de Mínimos Quadrados (LSSVM) para estimar a tendência de degradação do rolamento de baixa rotação. A PCA foi usada para fundir as múltiplas características para obter o indicador com base nas características do domínio do tempo e do domínio da frequência, refletindo a degradação de um certo tipo de rolamento de turbina eólica girando a 4 rpm. Um experimento de vida útil foi conduzido sob um teste de deterioração acelerada, impondo a carga limite durante 1122 min de vida até a fadiga. A pista do rolamento com diâmetro central de 730 mm apresentou muitas lascas, corrosão e estrias após o experimento. Havia também algumas reentrâncias e *pits* na superfície dos elementos rolantes.

Considerando ainda o prognóstico do rolamento de baixa velocidade e propostas de estimativa do tempo de sua vida útil restante, o método MFO-

CERLMDAN apresentou bons resultados na remoção de ruído de sinal, e o SPE baseado no MFO-CERLMDAN-KPCA indicou com sucesso a degradação do desempenho do rolamento de baixa frequência de rotação (PAN et al., 2022). O tamanho da falha em rolamentos de velocidade muito baixa foi estimado com precisão com base na medição da velocidade angular instantânea (IAS), além de considerar o comportamento diferente entre os momentos de entrada e saída do elemento rolante passando pelo defeito simulado (MOUSTAFA et al., 2016). Uma combinação de algoritmo de *Deep Autoencoder* (DAE) e algoritmo de filtro de partículas foi utilizada para extração de características e previsão de vida útil restante (RUL) (PAN et al., 2019). Um modelo de degradação não-linear de Wiener considerando a influência de covariáveis aleatórias foi estabelecido para a predição do RUL de rolamentos de ventilador usando dados de degradação obtidos de medição em campo (LIU et al., 2022).

Até este momento foram citados diversos trabalhos relacionados ao monitoramento e diagnóstico de falha de grandes rolamentos em baixa rotação. A Tabela 2.1 inclui outros ainda pertinentes ao tema, destacando os métodos que cada um empregou.

Tabela 2.1 – Outros métodos de monitoramento de rolamentos em baixa rotação.

Autoria	Descrição/Método(s)
Wang e Kootsookos (1998)	Envelope-Autocorrelação
Ompusunggu et al. (2013)	Filtro não-linear <i>Stochastic-Resonance</i> (SR)
Feng et al. (2014)	Predição de RUL
Kang et al. (2014)	<i>GA-Kernel Discriminative Feature</i> + OAA MCSVMs
Kang et al. (2015)	Análise de features baseadas em BBA
Prakash et al. (2017)	HMM para calcular RUL
Klausen et al. (2017)	Desenvolveram uma bancada experimental
Nguyen et al. (2018)	DWPT + EPS + HI
Caesarendra et al. (2018)	PANFIS para predição de features
Han et al. (2019)	TEO + CEEMD
Hemmer et al. (2020)	Indicador de saúde usando VAE
Sandoval et al. (2020)	<i>Spectral e Permutation Entropy</i>
Li et al. (2021)	IAC-MOMEDA para realçar de features fracas
Pan et al. (2023)	DBN aprimorada com FEPCD

2.2 Contribuições do trabalho

Após minuciosa revisão da bibliografia, definiu-se que este trabalho poderá contribuir para o campo de pesquisa em monitoramento em rolamento de baixa velocidade conforme os seguintes tópicos.

2.2.1 Determinação do melhor tamanho de janela do sinal para extração de features ponderando entre separabilidade e custo computacional

Análise da influência do tamanho da janela do sinal de vibração na sensibilidade dos preditores (*features*) em distinguir entre diferentes classes com base no número de revoluções do eixo. Para tanto, utilizando a ferramenta de teste estatístico *t* de Welch como indicador do grau de separabilidade.

2.2.2 Combinação de features não-lineares (LLE e HFD) e features tradicionais para diagnóstico de falha mediante aprendizado supervisionado

Uma vez que Sandoval (2021) encontrou melhores resultados associando indicadores de entropia (que também são *features* não-lineares) às *features* tradicionais no domínio do tempo (estatísticas, fatores e parâmetros de Hjorth). Este trabalho contribui em avaliar a associação de outros preditores não-lineares, a saber, o MEL e a DFH, àqueles tradicionais em tarefa de diagnóstico de falha.

2.2.3 Dados experimentais com rolamento em baixa rotação com falha incipiente simulando desgaste abrasivo na pista externa

A respeito da aquisição de dados, são comuns na literatura dados oriundos de rolamentos degradados ao longo de certo período de uso industrial (LIU E ZHANG, 2020; WANG et al., 2018; HEMMER, 2020); de ensaios experimentais provocando aceleração da perda do tempo de vida útil do rolamento (CAESARENDRA, 2015; KLAUSEN, 2017); e rolamentos com defeitos pontuais simulados (SANDOVAL, 2021; MISHRA et al., 2017; SAUFI et al., 2014).

Todavia, vale ressaltar que os estágios iniciais de diferentes modalidades de falha em rolamentos geralmente iniciam-se com desgaste 2 copos (com quebra do filme lubrificante), ou ainda de 3 corpos (presença de partículas abrasivas). Isso gera, *a priori*, arranhões nas superfícies dos elementos do rolamento em contato entre si. Neste sentido, Hemmer (2020) propôs em sua tese a obtenção de dados de bancada experimental com rolamento axial cônico de máquina perfuratriz com arranhões

arqueados em diferentes raios na pista de rolagem. O referido autor destaca que, defeitos como arranhões na superfície de elementos do rolamento torna o diagnóstico de falha com técnicas tradicionais de análise de vibração espectral menos úteis. Uma vez que não há um defeito pontual ensejando um comportamento periódico.

Portanto, trabalhos baseados em defeitos de desgaste com arranhões internos no rolamento são escassos – sendo visto somente em Hemmer (2020). Considerando que tais arranhões são defeitos de caráter incipiente que podem evoluir para trincas, fraturas e escamamento, este trabalho também elaborou uma metodologia de ensaios com rolamentos nessa condição. Neste caso, foram estabelecidos três níveis de defeitos por arranhamento na pista externa de um rolamento rígido de esferas, variando o nível de carga no equipamento.

2.2.4 Proposta de métrica alternativa para estimativa dos vizinhos mais próximos na trajetória dinâmica para cálculo do LLE de Rosenstein.

Caesarendra et al. (2015) reportaram que o Máximo Expoente de Lyapunov apresentou-se como o melhor preditor para tarefa de prognóstico de falha em rolamento de baixa velocidade. Visto que foi a *feature* extraída do sinal que melhor acompanhou a degradação da falha de um rolamento em ensaio experimental *run-to-failure*. Para calcular o Máximo Expoente de Lyapunov, os autores propuseram um algoritmo baseado no método de Rosenstein et al. (1993).

O qual, em suma, começa reconstruindo o pseudo-espaco-de-estado (trajetória dinâmica) do sinal de vibração; por conseguinte, encontra-se o vizinho mais próximo de cada ponto da trajetória por meio da distância euclidiana entre eles. E por fim, a separação exponencial entre a trajetória e sua vizinhança é mensurada e o sistema é considerado cada vez mais caótico quanto maior for o valor do expoente positivo. Para expoentes nulos ou negativos, o sistema é estável.

Além disso, vale ressaltar que as etapas de reconstrução da trajetória dinâmica e de encontrar os vizinhos mais próximos são as mais críticas. Este trabalho propôs avaliar o impacto da substituição da distância euclidiana pela distância de correlação de Pearson na etapa de busca pelos vizinhos mais próximos de cada ponto da trajetória. Dessa forma, comparando ambas (Euclid-MEL e Pearson-MEL) com base em sua sensibilidade em separar classes distintas.

3 FUNDAMENTAÇÃO TEÓRICA

3.1 Atributos de sinais no domínio do tempo

Uma série temporal consiste em uma sequência de valores numéricos discretos em função do tempo, da qual diversos atributos podem ser extraídos visando determinar índices que caracterizam o sinal. Diante disso, dado a série temporal discreta, \mathbf{x}_i ($i = 1, 2, 3, \dots, N$), em que N é a quantidade de elementos de \mathbf{x}_i , as expressões matemáticas seguintes representam características ou atributos que podem ser extraídos de uma série temporal, considerando o sinal no domínio do tempo (MATHWORKS®, 2024a; PÄIVINEN et al., 2005; CAESARENDRA, 2015):

3.1.1 Valor-de-Pico

$$x_{\text{pico}} = \max|\mathbf{x}_i| \quad (3.1)$$

O valor-de-pico corresponde ao valor absoluto do elemento de maior magnitude dentre os demais indivíduos da série (MATHWORKS®, 2024b).

3.1.2 Raíz Média Quadrada (RMS – Root Mean Square)

$$x_{\text{RMS}} = \sqrt{\frac{1}{N} \sum_{i=1}^N \mathbf{x}_i^2} \quad (3.2)$$

Consiste na raiz quadrada da média do quadrado de todos os elementos do sinal. Assim, esta média pode fornecer uma noção da energia global inerente ao sinal (CAESARENDRA, 2015).

3.1.3 Fator de Forma

$$x_{\text{FF}} = \frac{x_{\text{RMS}}}{\frac{1}{N} \sum_{i=1}^N |\mathbf{x}_i|} \quad (3.3)$$

O fator de forma consiste na razão entre o valor RMS e a média do valor absoluto de \mathbf{x}_i . Segundo Caesarendra (2015), refere-se ao valor que é afetado pela forma do sinal, mas é independente de suas dimensões.

3.1.4 Curtose

$$X_{Y_2} = \frac{\frac{1}{N} \sum_{i=1}^N (\mathbf{x}_i - \bar{\mathbf{x}})^4}{\left[\frac{1}{N} \sum_{i=1}^N (\mathbf{x}_i - \bar{\mathbf{x}})^2 \right]^2} \quad (3.4)$$

A curtose é uma medida do grau de achatamento da função densidade de probabilidade em relação ao seu centro (CAESARENDRA, 2015).

3.1.5 Distorção

$$X_{Y_1} = \frac{\frac{1}{N} \sum_{i=1}^N (\mathbf{x}_i - \bar{\mathbf{x}})^3}{\left[\frac{1}{N} \sum_{i=1}^N (\mathbf{x}_i - \bar{\mathbf{x}})^2 \right]^{3/2}} \quad (3.5)$$

A distorção mede a assimetria da função densidade de probabilidade do sinal em relação ao centro (CAESARENDRA, 2015).

3.1.6 Fator de Impulso

$$X_{FI} = \frac{X_{\text{pico}}}{\frac{1}{N} \sum_{i=1}^N |\mathbf{x}_i|} \quad (3.6)$$

O fator de impulso corresponde à razão entre o valor-de-pico e a média do valor absoluto de \mathbf{x}_i . Segundo MathWorks® (2024a), X_{FI} compara o valor-de-pico com o nível médio do sinal.

3.1.7 Fator de Crista

$$X_{FC} = \frac{X_{\text{pico}}}{X_{\text{RMS}}} \quad (3.7)$$

O fator de crista corresponde a razão entre o valor-de-pico e o valor RMS. Segundo Caesarendra (2015), trata-se de uma medida de impactos, apropriada para sinais com características de picos consecutivos.

3.1.8 Fator de Liberação

$$x_{FL} = \frac{x_{\text{pico}}}{\left(\frac{1}{N} \sum_{i=1}^N \sqrt{|x_i|}\right)^2} \quad (3.8)$$

O Fator de Liberação é a razão entre o valor-de-pico e o quadrado da média da raiz quadrada do valor absoluto de x_i (MathWorks®, 2024a).

3.1.9 Atividade

$$A_0 = x_{\sigma}^2 \quad (3.9)$$

A Atividade é conhecida como o primeiro parâmetro de Hjorth, e é igual à variância do sinal (PÄIVINEN et al., 2005).

3.1.10 Mobilidade

$$M_0 = \frac{x'_{\sigma}}{x_{\sigma}} \quad (3.10)$$

A Mobilidade, segundo parâmetro de Hjorth, relaciona o desvio padrão da primeira derivada do sinal e o desvio padrão do sinal original. A derivada do sinal é calculada mediante $x'(j)=x(j+1)-x(j)$, onde $j=1, \dots, N-1$ (PÄIVINEN et al., 2005).

3.1.11 Complexidade

$$C_0 = \frac{x''_{\sigma}/x'_{\sigma}}{x'_{\sigma}/x_{\sigma}} \quad (3.11)$$

A Complexidade é o terceiro parâmetro de Hjorth e corresponde à razão entre a Mobilidade da primeira derivada do sinal e a Mobilidade do sinal original (PÄIVINEN et al., 2005).

3.2 Valor t do teste estatístico de Welch

De acordo com Ruxton (2006), na pesquisa científica, frequentemente lida-se com problemas que demandam a utilização de testes estatísticos para mensurar a diferença relativa entre as tendências centrais (média ou mediana) de dois grupos, com base nas amostras de ambos. Além disso, Ruxton (2006) realizou uma busca de trabalhos que tivessem adotado tais testes, e o mesmo chegou à conclusão de que

poucos utilizam o teste de “variância desigual” ou teste t de Welch. O referido autor sugere ainda a hipótese de que isso se deve à falta de conhecimento por parte de muitos pesquisadores, fazendo com que muitos optem pelo teste t de Student ou teste U de Mann-Whitney, quando o teste t de Welch possa ser mais adequado.

Nesse sentido, Coombs et al. (1996) e Zimmerman e Zumbo (1993) apresentaram resultados que, qualitativamente, apontam que o teste de Welch e o de Student geram resultados equivalentes quando os dois grupos independentes avaliados possuem a mesma variância e o mesmo tamanho. Todavia, para conjuntos amostrais de variâncias e tamanhos diferentes, o teste t de Student realiza avaliações com menor confiabilidade em relação ao de Welch. Além disso, conforme afirma Ruxton (2006), o método para calcular o valor estatístico t é mais simples matematicamente para o caso do teste de Welch em relação ao de Student.

Fagerland e Sandvik (2009) apresentam a expressão para o cálculo do valor estatístico t de Welch, a qual está em função da média e da variância das amostras. Assim, dado dois grupos independentes $\{X_1, \dots, X_i\}$ e $\{Y_1, \dots, Y_j\}$, onde $i = 1, \dots, N$ e $j = 1, \dots, M$, sendo que N e M representam o tamanho (número de indivíduos) dos respectivos grupos. Ademais, assumindo que as variâncias de ambos os conjuntos não são equivalentes, o valor estatístico t é calculado pela seguinte expressão, em que, \bar{X} e \bar{Y} , e σ_X^2 e σ_Y^2 , são as médias e as variâncias dos grupos, respectivamente:

$$t = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_X^2}{N} + \frac{\sigma_Y^2}{M}}} \quad (3.12)$$

Por fim, West (2021), ao aplicar o método da Equação (3.12) para diferenciar dois grupos em problema de cunho bioquímico recomenda, ao final, o uso do teste de Welch em vez do teste de Student.

3.3 Análise de Componentes Principais (PCA)

A PCA talvez seja a técnica de análise estatística multivariada mais popular. Simplificadamente, trata-se de uma técnica estatística não-paramétrica que analisa um conjunto de dados definido por variáveis inter-correlacionadas. A técnica visa extrair dos dados as informações mais relevantes, representando-as em um novo conjunto de variáveis ortogonais entre si chamadas de *componentes principais* (ABDI

e WILLIAMS, 2010; CADAVID et al., 2009; SHLENS, 2014). Ainda, Abdi e Williams (2010) elencam os principais objetivos da técnica PCA:

1. Extrair a informação mais importante da distribuição de dados.
2. Reduzir a dimensionalidade dos dados mantendo apenas a informação relevante.
3. Simplificar a descrição do conjunto de dados original.
4. Analisar a estrutura das amostras e das variáveis.

Primeiramente, vale frisar que a definição a seguir foi baseada em Jolliffe e Cadima (2016). A qual se inicia com a suposição de uma matriz de dados \mathbf{X} , cujas “n” linhas representam as amostras, objetos ou indivíduos. Ao passo que as “p” colunas correspondem às variáveis ou dimensões da referida matriz. Diante disso, na PCA são analisadas as variáveis da matriz de dados, dessa forma o foco recai sobre o conjunto $\{\mathbf{x}_1, \dots, \mathbf{x}_p\} \in \mathbb{R}^n$. Assim, o objetivo consiste em determinar uma combinação linear dada por:

$$\sum_{j=1}^p a_j \cdot \mathbf{x}_j = \mathbf{X} \cdot \mathbf{a} \quad (3.13)$$

Onde \mathbf{a} corresponde ao vetor de constantes $\{a_1, \dots, a_p\}$. Em seguida, partindo do pressuposto de que a variância de qualquer combinação linear é dada por:

$$\text{var}(\mathbf{X} \cdot \mathbf{a}) = \mathbf{a}^T \cdot \mathbf{S} \cdot \mathbf{a} \quad (3.14)$$

Em que \mathbf{S} é a matriz de covariância dos dados, a qual pode ser obtida mediante $\check{\mathbf{X}}^T \cdot \check{\mathbf{X}}$, sendo $\check{\mathbf{X}}$ a matriz \mathbf{X} centrada na média, isto é, cada coluna \mathbf{x}_j é centrada na média tornando-se $\check{\mathbf{x}}_j$ mediante $\check{x}_{ij} = x_{ij} - \bar{x}_j$, onde $i = 1, \dots, n$, $j = 1, \dots, p$, e \bar{x}_j representa a média da variável \mathbf{x}_j . Ao final, a média de cada coluna da matriz $\check{\mathbf{X}}$ deve ser igual a zero.

Portanto, o objetivo é determinar a combinação linear que apresente máxima variância, por isso a Equação (3.14) deve ser maximizada sobre \mathbf{a} , ou seja, maximizar a forma quadrática $\mathbf{a}^T \cdot \mathbf{S} \cdot \mathbf{a}$, sujeita a restrição de que tais vetores possuam norma unitária, isto é, a condição $\mathbf{a}^T \cdot \mathbf{a} = 1$ deve ser satisfeita. Enfim, o problema é equivalente a maximizar a seguinte função de Lagrange:

$$\arg \max_{\|\mathbf{a}\|=1} \mathbf{a}^T \cdot \mathbf{S} \cdot \mathbf{a} - \lambda \cdot (\mathbf{a}^T \cdot \mathbf{a} - 1) \quad (3.15)$$

Onde λ é um multiplicador de Lagrange. Agora, derivando a Equação (3.15) e igualando ao vetor nulo, visando um ponto estacionário (platô) da função, encontra-se que:

$$\mathbf{S} \cdot \mathbf{a} - \lambda \cdot \mathbf{a} = \mathbf{0} \Leftrightarrow \mathbf{S} \cdot \mathbf{a} = \lambda \cdot \mathbf{a} \quad (3.16)$$

Diante disso, devido à matriz de covariância ser uma matriz quadrada $p \times p$, conclui-se que o vetor \mathbf{a} corresponde ao autovetor da matriz de covariância \mathbf{S} , enquanto que λ é o autovalor associado ao autovetor. Embora o interesse inicial seja encontrar o autovalor de maior magnitude λ_1 (e seu correspondente autovetor \mathbf{a}_1), sabe-se que a matriz \mathbf{S} possui “p” autovalores, ou seja, λ_k ($k = 1, \dots, p$). Enquanto que os autovetores correspondentes podem ser definidos como um conjunto de vetores ortonormais. Nesse sentido, esta ortogonalidade dos autovetores pode ser usada como restrição para o cálculo de outros vetores de outros coeficientes de novas combinações lineares. Em que o conjunto dos autovetores \mathbf{a}_k da matriz \mathbf{S} é considerado solução ótima para o problema de maximização da função de Lagrange da Equação (3.15) para cada k-nova combinação linear. Assim, a Equação (3.13) assume uma forma generalizada:

$$\mathbf{X} \cdot \mathbf{a}_k = \sum_{j=1}^p a_{jk} \cdot \mathbf{x}_j \quad (3.17)$$

Os elementos dos autovetores \mathbf{a}_k são chamados de *loadings*, que representam o peso de cada variável da matriz \mathbf{X} em relação a determinada componente principal, e os elementos da combinação $\mathbf{X} \cdot \mathbf{a}_k$ são chamados de *scores*, que correspondem à projeção de cada amostra de \mathbf{X} sobre certa componente.

Outrossim, tendo em vista a definição acima, Smith (2002) sintetiza o cálculo da PCA em 7 passos fundamentais:

1. Obter uma matriz de dados.
2. Centralizar a matriz de dados na média das variáveis.
3. Calcular a matriz de covariância da matriz centrada na média.

4. Calcular os autovalores e os autovetores da matriz de covariância.
5. Criar uma matriz formada pelos autovetores em ordem decrescente de significância, com base na magnitude do autovalor associado, respectivamente.
6. Reduzir a dimensionalidade dos dados originais selecionando apenas os autovetores mais relevantes e negligenciando os menos importantes.
7. Multiplicar a matriz original pela matriz de autovetores gerando uma representação dos dados em baixa dimensionalidade.

Como dito anteriormente, os *scores* são obtidos pelo produto escalar entre a matriz original de dados \mathbf{X} e a matriz de autovetores \mathbf{a} . Dessa forma, tem-se a seguinte expressão para os *scores* \mathbf{t} :

$$\mathbf{t} = \mathbf{X} \cdot \mathbf{a} \quad (3.18)$$

Nesse contexto, Bro e Smilde (2014) apresentam uma importante interpretação para a PCA, que parte da premissa de que os *scores* tendem a representar os dados originais em um subespaço com um menor número de dimensões, aliado a uma perda mínima de informação. Assim, quanto menor o erro entre a matriz de *scores* \mathbf{t} e a matriz original \mathbf{X} , mais próximo a matriz de *scores* está de representar bem a matriz original, apesar da redução de dimensionalidade. Diante disso, com base na Equação (3.18), e considerando que são retidos no modelo apenas “ r ” componentes mais relevantes e desconsideradas as restantes, a matriz original \mathbf{X} pode ser expressa em função das matrizes de *scores* \mathbf{t} , de *loadings* \mathbf{a} , e de resíduos \mathbf{E} (perda de informação).

$$\mathbf{X} = \mathbf{t} \cdot \mathbf{a}^T + \mathbf{E} \quad (3.19)$$

Portanto, sob a perspectiva da Equação (3.19), tem-se como objetivo minimizar a matriz de resíduos \mathbf{E} , de sorte que tem-se um problema de otimização semelhante ao caso da Equação (3.15). Enfim, dado que $\mathbf{E} = \mathbf{X} - \mathbf{t} \cdot \mathbf{a}^T$, os parâmetros ótimos \mathbf{t} e \mathbf{a} são:

$$\arg \min_{\mathbf{t}, \mathbf{a}} \|\mathbf{X} - \mathbf{t} \cdot \mathbf{a}^T\|^2 \quad (3.20)$$

3.4 Carta T^2 de Hotelling

Segundo Zaidi et al. (2023), monitoramento estatístico de processos é uma técnica utilizada para monitoramento contínuo de processos industriais visando controle de qualidade e identificação de problemas. O que se dá por meio de ferramentas como cartas de controle, histogramas, gráficos de Pareto, e *scatter plots* para monitoramento de processos e detecção de desvios de performances. As cartas de controle são úteis para a identificação de situações fora de controle, o que leva à investigação das causas especiais para variações inesperadas.

De acordo com Tomé et al. (2019), uma carta de controle pode ser usada para mapear a existência de valores anormais. Ademais, em vez de se utilizar uma carta para cada *feature*, adota-se a carta de controle multivariada T^2 de Hotelling, que analisa todas elas em um único plano. Conforme Asri et al. (2019), a carta T^2 de Hotelling é baseada na distância de Mahalanobis, ou seja, mensura a distância de cada indivíduo em relação ao centro (média) de uma distribuição normal multivariada. Portanto, segundo Pan e Zhou (2011), dada uma distribuição de dados, \mathbf{X} , a Equação (3.21) para o cálculo do valor T^2 pode se resumir a:

$$T^2 = (\mathbf{X}_i - \bar{\mathbf{X}})^T \cdot \mathbf{S}^{-1} \cdot (\mathbf{X}_i - \bar{\mathbf{X}}) \quad (3.21)$$

Em que \mathbf{S}^{-1} é o inverso da matriz de covariância. A qual consiste em $\mathbf{S} = \frac{1}{h} \sum_{j=1}^h (\mathbf{X}_j - \bar{\mathbf{X}}_j) \cdot (\mathbf{X}_j - \bar{\mathbf{X}}_j)^T$, ao passo que $\mathbf{X}_j = (X_{1j}, \dots, X_{ij})^T$, sendo $i = \{1, 2, \dots, n\}$ e $j = \{1, 2, \dots, h\}$, “n” o número de indivíduos do conjunto e “h” o número de variáveis.

De acordo com Mehmood (2016), uma linha limite pode ser definida com base na distribuição estatística F, mediante a qual é possível estabelecer um *threshold* acima do qual se pode detectar potenciais elementos anômalos. A Equação (3.22) estima o valor limite na carta T^2 de Hotelling, em que $F(\alpha, h, n - h)$ é o valor da distribuição F com confiança $(1 - \alpha)$, e “h” e “n - h” graus de liberdade.

$$T_{\text{lim}}^2 = \frac{(n - 1) \cdot h}{n - h} \cdot F(\alpha, h, n - h) \quad (3.22)$$

3.5 Máximo Expoente de Lyapunov (MEL)

Para a caracterização de sistemas dinâmicos não-lineares, é necessário o uso de métodos capazes de extrair significado físico de dados experimentais discretos de sistemas caóticos multidimensionais, tal como os expoentes de Lyapunov (EL's). Assim, a determinação dos mesmos torna-se importante para analisar possíveis sistemas caóticos, uma vez que tais expoentes podem apresentar qualitativamente a sensibilidade do sistema às condições iniciais. Por definição, EL's correspondem à taxa exponencial média de divergência ou convergência de órbitas (trajetórias) próximas no espaço-de-estado (atrator). Um atrator, por sua vez, caracteriza o comportamento dinâmico do sistema ao representar um conjunto de estados pelo qual o sistema dinâmico pode evoluir. Qualquer sistema que possua ao menos um EL positivo é considerado caótico, de forma que a magnitude do expoente diz respeito ao tempo em que o sistema passa a ser imprevisível (STOOP e PARISI, 1990; SANO e SAWADA, 1985; RUELLE, 1981; WOLF et al., 1985).

Matematicamente, segundo Vulpiani et al. (2009), os EL's são um conjunto de parâmetros associados à cada trajetória $\mathbf{x}(t)$, os quais estão relacionados à medição do grau de instabilidade da mesma. A evolução dessas trajetórias segue a ideia de um mapa multidimensional genérico como:

$$\mathbf{x}(t + 1) = f(\mathbf{x}(t)) \quad (3.23)$$

A estabilidade de uma trajetória pode ser analisada observando a evolução de suas trajetórias mais próximas $\mathbf{x}'(t)$, de modo que a distância inicial infinitesimal entre $\mathbf{x}(t)$ e $\mathbf{x}'(t)$ é dada por $\Delta(0)$, de maneira que:

$$\mathbf{x}'(0) = \mathbf{x}(0) + \Delta(0) \quad (3.24)$$

Em sistemas caóticos, com o passar do tempo, $\Delta(t)$ cresce exponencialmente:

$$\Delta(t) \sim \Delta(0) \cdot e^{\lambda t} \quad (3.25)$$

Onde λ é a taxa de expansão. A qual pode ser determinada isolando-se λ na Equação (3.25), gerando a expressão geral para a determinação do MEL, nos limites em que o tempo "t" tende ao infinito e a distância inicial entre as trajetórias é mínima, ou seja, em que $\Delta(0)$ tende a zero:

$$\lambda_{\max} = \lim_{t \rightarrow \infty} \lim_{\Delta(0) \rightarrow 0} \frac{1}{t} \cdot \ln \left(\frac{\Delta(t)}{\Delta(0)} \right) \quad (3.26)$$

Assim, quando $\lambda_{\max} \leq 0$, tem-se que o sistema não é considerado caótico, ao passo que, para $\lambda_{\max} > 0$, observa-se um comportamento caótico. Quanto a sua magnitude, para quando $\lambda_{\max} > 0$, segundo Wolf et al. (1985), diz respeito ao tempo em que o sistema passa a ser imprevisível, isto é, módulos elevados do MEL positivo traduzem-se em altos níveis de caoticidade.

3.6 Dimensão Fractal de Higuchi (DFH)

As ideias básicas de Dimensão Fractal (DF) foram propostas por Benoit Mandelbrot em 1967 visando desenvolver um método para estimar o comprimento de curvas geográficas. De acordo com Sharma et al. (2019), DF é utilizada para mensurar o grau de não-linearidade de respostas irregulares de sistemas dinâmicos. Ou ainda pode ser definida como um indicador quantitativo do grau de complexidade ao comparar variações sutis entre padrões. Assim, DF fornece uma medida quantitativa de caoticidade de respostas dinâmicas.

Higuchi (1988) propôs um método para o cálculo da DF de séries temporais de tamanho finito, N , coletadas em intervalos regulares:

$$\mathbf{X} = X(1), X(2), X(3), \dots, X(N) \quad (3.27)$$

O método consiste em, primeiramente, construir uma nova série temporal, \mathbf{X}_k^m , para k definido na faixa entre 1 e k_{\max} (um parâmetro a ser definido) e $m = 1, 2, \dots, k$:

$$\mathbf{X}_k^m = X(m), X(m+k), X(m+2 \cdot k), \dots, X\left(m + \left\lfloor \frac{N-m}{k} \right\rfloor \cdot k\right) \quad (3.28)$$

Em que $\lfloor \cdot \rfloor$ representa a notação de função de arredondamento para o valor inteiro mais próximo. Ambos “ m ” e “ k ” são inteiros positivos que indicam o tempo inicial e o intervalo de tempo, respectivamente. Vale destacar que para um intervalo de tempo igual a k , um conjunto de k novas curvas são criadas. Assim, o comprimento $L_m(k)$ de cada curva \mathbf{X}_k^m pode ser calculado como:

$$L_m(k) = \left\{ \left(\sum_{i=1}^{\lfloor \frac{N-m}{k} \rfloor} |X(m+i\cdot k) - X(m+(i-1)\cdot k)| \right) \frac{N-1}{\lfloor \frac{N-m}{k} \rfloor \cdot k} \right\} / k \quad (3.29)$$

O termo $N - 1 / \lfloor \frac{N-m}{k} \rfloor \cdot k$ representa o fator de normalização do valor do comprimento da curva \mathbf{X}_k^m . Finalmente, estima-se o comprimento médio da curva para cada valor de k , $\langle L(k) \rangle$, compreendendo ao valor médio de $L_m(k)$ (lembrando que $m = 1, 2, \dots, k$), e repetindo a iteração sobre k variando de 1 a k_{\max} . Portanto, se $\langle L(k) \rangle \propto k^{-D}$, a curva é um fractal de dimensão “D”.

3.7 Distância de Pearson

A correlação entre variáveis pode ser medida por meio de diversos índices, os três mais populares são os coeficientes de Pearson, de Spearman e de Kendall. O método de Pearson foi descoberto por Auguste Bravais em 1846, mas foi Karl Pearson que em 1896 descreveu o método com maior rigor, incluindo algumas contribuições, entre as quais propôs a normalidade das variáveis analisadas. Em suma, o coeficiente de Pearson é uma medida da força da correlação entre duas variáveis quantitativas. (HAUKE e KOSSOWSKI, 2011).

O cálculo da correlação entre duas variáveis, para Xiao et al. (2016), fundamenta-se no quanto que uma variável \mathbf{X} pode ser usada para prever o comportamento de outra \mathbf{Y} .

$$\hat{\mathbf{Y}} = f(\mathbf{X}) \quad (3.30)$$

Em que $\hat{\mathbf{Y}}$ denota os valores preditos para \mathbf{Y} com base em uma função f de \mathbf{X} . Assim, a precisão dessa previsão pode ser medida pelo erro entre o valor real e o valor predito:

$$e_i = Y_i - \hat{Y}_i \quad (3.31)$$

Em que $i = 1, \dots, N$, e N corresponde ao tamanho do conjunto. Assim, o erro quadrado médio pode ser descrito da seguinte forma:

$$s_{yx}^2 = \frac{\sum_{i=1}^N e_i^2}{N} = \frac{\sum_{i=1}^N (Y_i - \hat{Y}_i)^2}{N} \quad (3.32)$$

Agora, nesse sentido, estabelecendo como referências as médias \bar{X} e \bar{Y} de cada variável X e Y , sendo s_x^2 e s_y^2 suas respectivas variâncias, a definição padrão para o cálculo da correlação r é dada por:

$$r = \frac{\sum_{i=1}^N (X_i - \bar{X}) \cdot (Y_i - \bar{Y})}{N \cdot s_x \cdot s_y} \quad (3.33)$$

De acordo com Mukaka (2012), o resultado pode variar de -1 a +1, de modo que o sinal positivo indica que se uma variável cresce a outra também cresce, ou se uma decresce a outra também decresce. Quando o sinal é negativo, significa que quando uma variável cresce a outra decresce, e vice-versa. Além disso, quanto mais próximo de ± 1 , mais forte é a correlação; enquanto que próximo de 0 torna-se fraca.

Segundo Xiao et al. (2016), o coeficiente de correlação de Pearson mede a intensidade e a direção de relações estritamente lineares entre as variáveis. De acordo com Mukaka (2012), a distribuição deve ser normal, pois valores extremos podem prejudicar a precisão dos resultados. Enfim, Mukaka (2012) e Xiao et al. (2016) reportam que o coeficiente de Pearson r_p equivale à razão entre a covariância amostral das variáveis e seus respectivos desvios.

$$r_p = \frac{\sum_{i=1}^N (X_i - \bar{X}) \cdot (Y_i - \bar{Y})}{\sqrt{\left(\sum_{i=1}^N (X_i - \bar{X})^2\right) \cdot \left(\sum_{i=1}^N (Y_i - \bar{Y})^2\right)}} \quad (3.34)$$

Finalmente, segundo Xie et al. (2016), a distância de Pearson pode ser calculada pela seguinte expressão:

$$d_p = 1 - r_p \quad (3.35)$$

3.8 Máquina de Vetores de Suporte (SVM)

Uma Máquina de Vetores de Suporte (do inglês, SVM – *Support Vector Machine*) é um algoritmo de aprendizado supervisionado comumente usado para

solução de diversos problemas de classificação, regressão e identificação de *outliers*. Dessa forma, pode ser aplicado, por exemplo, em processamento de sinal, processamento de linguagem natural e reconhecimento de fala e imagem. O algoritmo SVM demonstra ótima performance em tarefas de classificação. Em que o objetivo é encontrar um hiperplano que possua a melhor direção (inclinação) possível, e que separe os pontos de dados de uma classe daqueles pertencentes a outra classe. O melhor hiperplano é aquele com a maior margem entre as duas classes (MATHWORKS®, 2024c; PEDREGOSA et al., 2011).

A Figura 3.1 ilustra um hiperplano (linha azul sólida) que separa a classe de sinal positivo da classe de sinal negativo. A margem corresponde à largura máxima entre as linhas azuis tracejadas paralelas ao hiperplano que não possuem pontos de dados em seu interior, e que são delimitadas pelos vetores de suporte (pontos circulados). Contudo, somente para problemas linearmente separáveis, o algoritmo pode encontrar um hiperplano como o da Figura 3.1, pois para a maioria dos problemas práticos, na verdade, o algoritmo busca maximizar uma margem considerada flexível (*soft margin*), que tolera um pequeno número de erros de classificação (MATHWORKS®, 2024c; PEDREGOSA et al., 2011).

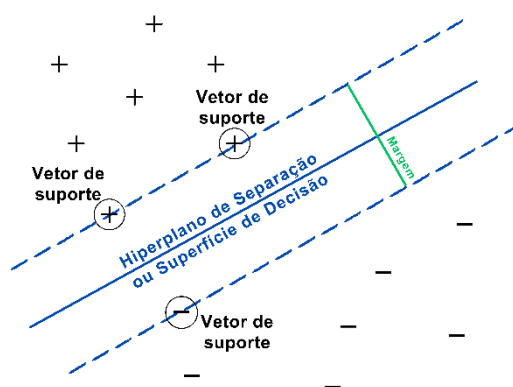


Figura 3.1 - Hiperplano e margem em função dos vetores de suporte
[Fonte: Adaptada de MathWorks® (2024c)].

Por sua vez, os vetores de suporte (que na Figura 3.1 correspondem aos pontos circulados que interceptam as linhas tracejadas) referem-se a um subconjunto das amostras de treinamento de cada classe que se encontra mais próximo do hiperplano de separação. Apenas os vetores de suporte escolhidos a partir dos dados de treinamento são necessários para construir a superfície de decisão. Uma vez treinado, o restante dos dados de treinamento torna-se irrelevante, resultando em uma

representação compacta do modelo. Diante de problemas não-lineares, a utilização de uma função kernel viabiliza uma solução linear. De modo que funções kernel mapeiam os dados para um espaço dimensional diferente, com a expectativa de que as classes sejam mais fáceis de separar após essa transformação (MATHEWORKS®, 2024c; PEDREGOSA et al., 2011).

Antes de adentrar nos conceitos matemáticos, vale ressaltar que as definições seguintes foram baseadas no conteúdo do capítulo 7 de Bishop (2006). Treinar uma máquina de vetores de suporte corresponde essencialmente a resolver um problema de otimização visando ajustar um hiperplano que maximize a margem entre as classes.

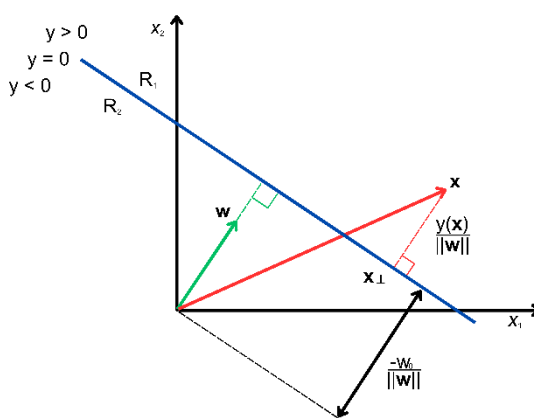


Figura 3.2 - Parâmetros geométricos relacionados à superfície de decisão
[Fonte: Adaptada de Bishop (2006)].

A Figura 3.2 mostra um hiperplano, $y = 0$, (linha azul) em um espaço bidimensional cuja direção é definida pelo vetor de pesos $\mathbf{w} \in \mathbb{R}^p$ (seta verde) e sua transladação no espaço é função do coeficiente $\omega_0 \in \mathbb{R}$. De sorte que a distância do hiperplano em relação à origem é dada por $\frac{-\omega_0}{\|\mathbf{w}\|}$, e a distância ortogonal de cada amostra, \mathbf{x} , para o hiperplano corresponde a $\frac{y(\mathbf{x})}{\|\mathbf{w}\|}$. Por sua vez, $y(\mathbf{x})$ é definido pela Equação (3.36), de modo que $\phi(\mathbf{x})$ representa uma transformação no espaço de *features*, e $b = \omega_0$ é o viés da função $y(\mathbf{x})$.

$$y(\mathbf{x}) = \mathbf{w}^T \cdot \phi(\mathbf{x}) + b \quad (3.36)$$

Os dados de treinamento são formados por N vetores de entrada $\{\mathbf{x}_1, \dots, \mathbf{x}_N\} \in \mathbb{R}^p$, cada qual está relacionado a um respectivo valor alvo $\{t_1, \dots, t_N\}$, tal

que $t_n \in \{-1, 1\}^N$. Ao passo que novos valores de \mathbf{x} (valores de testes, por exemplo) serão classificados de acordo com o sinal de saída da função da Equação (3.36), isto é, $y(\mathbf{x}_{\text{teste}}) = \text{sign}(\mathbf{w}^T \phi(\mathbf{x}_{\text{teste}}) + b)$. Assim, observa-se que, se tratando de dados linearmente separáveis no espaço de *features*, existe pelo menos uma combinação dos parâmetros \mathbf{w} e b , de maneira que para todo $t_n = +1$, implique em $y > 0$; e que, igualmente, para todo $t_n = -1$ ocorra que $y < 0$; de tal forma que a condição $t_n \cdot y(\mathbf{x}) > 0$ seja satisfeita para todas as amostras do conjunto de dados de treinamento.

Portanto, com base nessas definições, é necessário minimizar a expressão $t_n \cdot (\mathbf{w}^T \cdot \phi(\mathbf{x}_n) + b)$, visando encontrar a amostra que esteja mais próxima do hiperplano, a qual será considerada um “vetor de suporte”. De forma que a distância entre esta amostra e o hiperplano, dada por $\frac{t_n \cdot y(\mathbf{x}_n)}{\|\mathbf{w}\|}$, representa a margem que deve ser maximizada sobre os parâmetros \mathbf{w} e b . Este processo de otimização descrito acima consiste na Equação (3.37) abaixo:

$$\arg \max_{\mathbf{w}, b} \left\{ \frac{1}{\|\mathbf{w}\|} \cdot \min_n [t_n \cdot (\mathbf{w}^T \cdot \phi(\mathbf{x}_n) + b)] \right\} \quad (3.37)$$

Para simplificar a solução deste problema, considera-se um plano $y = 1$ da Figura 3.3 que intercepte os vetores de suporte (pontos vermelhos) delimitando, assim, a margem, e que seja paralelo ao hiperplano definido por $y = 0$.

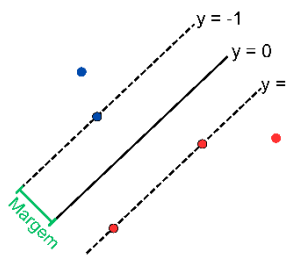


Figura 3.3 - Relação geométrica entre a margem e os vetores de suportes
[Fonte: Adaptada de Bishop (2006)].

Dessa forma, tem-se que $t_n \cdot (\mathbf{w}^T \cdot \phi(\mathbf{x}_n) + b) = 1$ para as amostras mais próximas do hiperplano (vetores de suporte). Todavia, para os casos em que os dados de treinamento não são linearmente separáveis no espaço de *features*, a determinação do hiperplano passa a considerar a tolerância de alguns erros de classificação. Para tanto, adota-se uma variável denominada de “variável de folga”, $\xi_n \geq 0$, $n = 1, \dots, N$. Com base na ilustração da Figura 3.4, é possível ver que, para

pontos que se encontram sobre ou abaixo da margem, $\xi_n = 0$. Enquanto que, para os demais pontos, vale a expressão $\xi_n = |t_n - y(\mathbf{x}_n)|$. Diante disso, quando o ponto se encontra sobre o hiperplano, $y(\mathbf{x}_n) = 0$, a folga será $\xi_n = 1$, e, portanto, quando $\xi_n > 1$, a amostra se encontra do lado oposto e tal indivíduo sofre erro de classificação.

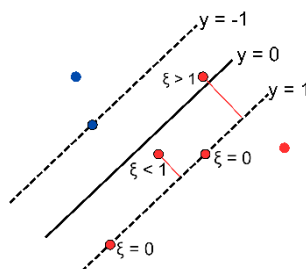


Figura 3.4 – Relação geométrica entre o fator ξ e a margem
[Fonte: Adaptada de Bishop (2006)].

Em outras palavras, enquanto $\xi_n = 0$, as amostras encontram-se do lado de fora da margem. Já para $0 < \xi_n \leq 1$ significa que os pontos estão dentro da margem. E quando $\xi_n > 1$, as amostras se encontram do lado oposto do hiperplano e são classificadas equivocadamente. Portanto, a restrição da Equação (3.38) passa a ser mais flexível admitindo alguns erros de classificação:

$$t_n \cdot (\mathbf{w}^T \cdot \phi(\mathbf{x}_n) + b) \geq 1 - \xi_n, \quad n = 1, \dots, N \quad (3.38)$$

Diante disso, a problema é simplificado para a maximização de $\|\mathbf{w}\|^{-1}$, que é equivalente a minimizar $\|\mathbf{w}\|^2$, de tal maneira que a Equação (3.37) pode ser resumida à expressão da Equação (3.39), sujeita à restrição da Equação (3.38). Além disso, ocorre a adição de um termo de penalização na equação em função da variável de folga ξ_n , gerando assim a expressão para a solução primal do problema:

$$\arg \min_{\mathbf{w}, b, \xi} \left(\frac{1}{2} \cdot \|\mathbf{w}\|^2 + C \cdot \sum_{n=1}^N \xi_n \right) \quad (3.39)$$

Em que a minimização acima está sujeita à restrição da Equação (3.38) e a $\xi_n \geq 0$. O parâmetro C da Equação (3.39) controla a força dessa penalidade e atua como a inversa de um parâmetro de regularização. Portanto, deve satisfazer a $C > 0$, e o mesmo controla o ajuste entre a margem e a variável de folga.

A Figura 3.5 ilustra a influência da magnitude do parâmetro C , mediante a qual é possível inferir que C diz respeito ao quanto se deseja evitar erros de classificação na etapa de treinamento do algoritmo. Conforme a Figura 3.5, para um valor elevado, $C = 1000$, a melhor solução resultou em uma margem rígida, que não tolera erros de classificação, mas tal resultado não parece ótimo, por apresentar uma margem ínfima, ter sofrido influência de *outlier*, e poder causar *overfitting*. Todavia, quando o parâmetro assume um valor menor, $C = 5$, observa-se um resultado aparentemente mais adequado, apesar de admitir erro de classificação para este fim, resultando em uma margem flexível, porém mais ampla.

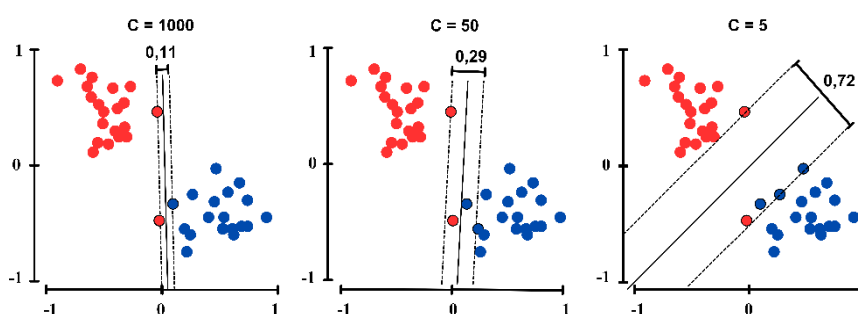


Figura 3.5 - Influência da variação do parâmetro C na direção do hiperplano [Fonte: Adaptada de Agarwal (2022)].

A solução dual para o problema primal da Equação (3.39) consiste na equação de Lagrange da Equação (3.40) abaixo, em que $a_n \geq 0$ e $\mu_n \geq 0$ são os multiplicadores de Lagrange introduzidos.

$$L(\mathbf{w}, b, \mathbf{a}) = \frac{1}{2} \cdot \|\mathbf{w}\|^2 + C \cdot \sum_{n=1}^N \xi_n - \sum_{n=1}^N a_n \cdot \{t_n \cdot y(\mathbf{x}_n) - 1 + \xi_n\} - \sum_{n=1}^N \mu_n \cdot \xi_n \quad (3.40)$$

Agora, visando encontrar o platô da equação acima, calcula-se o gradiente da mesma em relação às variáveis \mathbf{w} , b e $\{\xi_n\}$, e igualando a zero:

$$\frac{\partial L}{\partial \mathbf{w}} = 0 \Rightarrow \mathbf{w} = \sum_{n=1}^N a_n \cdot t_n \cdot \phi(\mathbf{x}_n) \quad (3.41)$$

$$\frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{n=1}^N a_n \cdot t_n = 0 \quad (3.42)$$

$$\frac{\partial L}{\partial \xi_n} = 0 \Rightarrow a_n = C - \mu_n \quad (3.43)$$

Finalmente, usando esses resultados para eliminar as variáveis \mathbf{w} , b e $\{\xi_n\}$ da Equação (3.40), obtém-se a representação dual para a maximização da margem:

$$\tilde{L}(\mathbf{a}) = \sum_{n=1}^N a_n - \frac{1}{2} \cdot \sum_{n=1}^N \sum_{m=1}^N a_n \cdot a_m \cdot t_n \cdot t_m \cdot k(\mathbf{x}_n, \mathbf{x}_m) \quad (3.44)$$

A qual é maximizada em relação a \mathbf{a} , sujeita às restrições abaixo:

$$0 \leq a_n \leq C \quad (3.45)$$

$$\sum_{n=1}^N a_n \cdot t_n = 0 \quad (3.46)$$

Em que a Equação (3.45) mostra que a_n torna-se limitado ao valor do parâmetro C , logo chamado de “caixa de restrição”. Nesse sentido, vale ressaltar que quando determinada amostra apresenta $a_n > 0$, significa que a mesma se trata de um vetor de suporte, ao passo que, $a_n = 0$, corresponde aos dados que se tornam irrelevantes para definição do hiperplano. Ainda, outra interpretação da Equação (3.45) relaciona o valor de C aos vetores de suporte, uma vez que quando, $a_n < C$, significa que os pontos estão sobre a margem, enquanto que $a_n = C$ implica que tais amostras ou se encontram dentro da margem ou do lado oposto do hiperplano.

Outro caso especial da Equação (3.44) é a presença da função kernel definida, neste contexto, por $k(\mathbf{x}, \mathbf{x}') = \phi(\mathbf{x})^T \cdot \phi(\mathbf{x}')$. Diante disso, torna-se viável representar uma equação com base na Equação (3.36) que forneça a resposta para novos dados, e que desta vez esteja em função do truque de kernel e dos parâmetros $\{a_n\}$. Assim, substituindo \mathbf{w} da Equação (3.41) na Equação (3.36):

$$y(\mathbf{x}) = \sum_{n=1}^N a_n \cdot t_n \cdot k(\mathbf{x}, \mathbf{x}_n) + b \quad (3.47)$$

4 METODOLOGIA

Esta seção apresenta a metodologia empregada neste trabalho, inicialmente abordando uma visão geral da metodologia no tópico 4.1. Na sequência, as etapas da metodologia proposta são explanadas com mais detalhes. Os itens 4.2 e 4.3 tratam da extração e seleção dos melhores preditores, respectivamente. Por sua vez, a seção 4.4 estabelece as principais considerações para definição dos parâmetros de treinamento do algoritmo de classificação SVM. Em relação à aquisição de dados, a seção 4.5 traz todos os detalhes a respeito da coleta de sinais de vibração de um misturador usado como bancada experimental. Ademais, o tópico 4.5.1 apresenta o método considerado para análise do janelamento do sinal. Ao final, os itens 4.6 e 4.7 descrevem os procedimentos experimentais para a aquisição de dados de um simples sistema rotativo motor-eixo-mancal e de mancais de tambores de transportadores de correia em plena atividade industrial.

4.1 Metodologia para o diagnóstico de falha

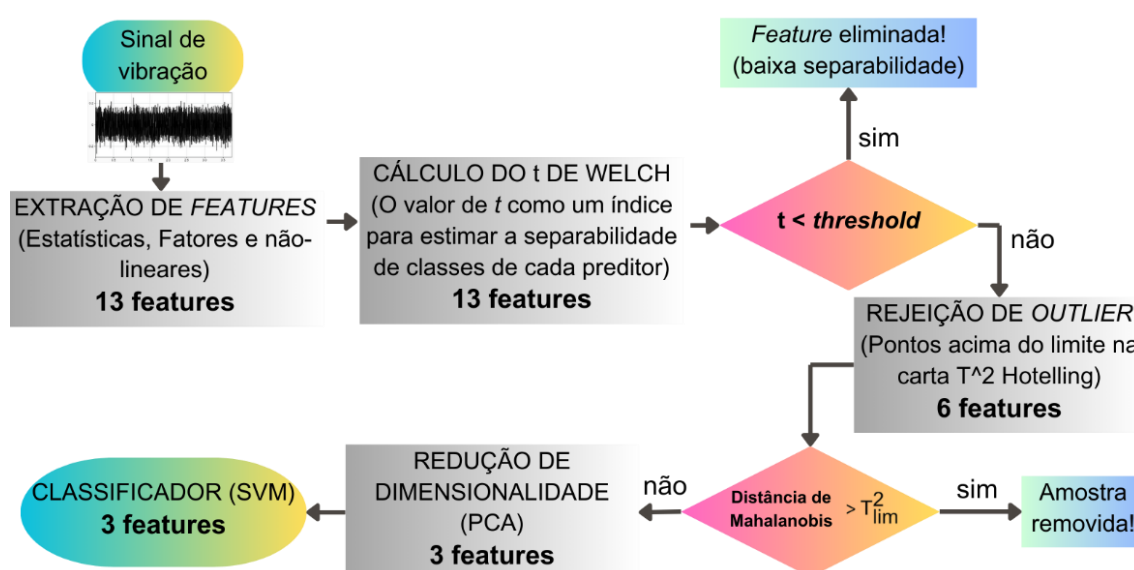


Figura 4.1 - Metodologia proposta: do processamento ao diagnóstico.

A Figura 4.1 apresenta o fluxograma da metodologia proposta neste trabalho para diagnóstico de falha de rolamentos em baixa rotação. De acordo com a Figura 4.1, após a aquisição dos dados, são calculados os indicadores (*features*) a partir da série temporal (sinal de vibração). Neste caso, foram extraídos 13 indicadores, os quais se enquadram em três categorias: estatísticos, fatores e não-lineares. Em seguida, é calculado o valor estatístico “t” do teste *t* de Welch para cada indicador.

Este índice “t” representa quantitativamente a capacidade que determinada característica extraída do sinal tem de distinguir classes, comparando a média e a variância das distribuições de cada classe. Assim, é possível selecionar as características que melhor separam as classes e desconsiderar as que não o fazem.

Após esta etapa de seleção, os dados, distribuídos agora em um espaço de seis dimensões, são representados em uma carta T^2 de Hotelling. Este gráfico representa a distância de Mahalanobis de cada observação em relação ao centro da distribuição dos dados. Portanto, é estimada uma linha limite com base na distribuição estatística F, em que os elementos acima do referido limite são considerados potenciais *outliers* e, por conseguinte, eliminados da distribuição. A eliminação de potenciais *outliers* deve-se ao fato do algoritmo de classificação SVM ser sensível a *outliers*. Na sequência, a matriz de *features* passa por uma redução de dimensionalidade mediante o uso da técnica algébrica de PCA. Nesta fase, optou-se por reduzir a dimensionalidade dos dados para 3 dimensões com o objetivo de simplificar a informação e reduzir os efeitos de redundância de indicadores altamente correlacionados, minimizando as chances de *overfitting* do algoritmo de aprendizagem de máquina na etapa de treinamento do modelo de classificação.

As seções seguintes apresentam mais detalhes das etapas descritas acima: extração e seleção de indicadores e uso do algoritmo de classificação SVM.

4.2 Extração de preditores dos sinais

Os atributos do sinal de vibração correspondem a um total de 13 (treze), dos quais 11 (onze) são atributos extraídos do sinal no domínio do tempo, e apenas dois preditores são baseados na estimativa do grau de não-linearidade do sinal. Com respaldo na pesquisa realizada por Caesarendra (2015), os atributos no domínio do tempo tendem a não apresentar ótimos resultados para a detecção de falhas incipientes em rolamentos de baixa rotação. Contudo, os parâmetros de Hjorth (Atividade, Mobilidade e Complexidade) não têm sido recorrentemente explorados para análise de falha em rolamentos de baixa rotação. A Figura 4.2 apresenta os preditores calculados e destaca aqueles pouco aplicados em problemas de monitoramento de rolamento em baixa rotação.

Estatísticas	Fatores	Não-lineares
<ul style="list-style-type: none"> •RMS •Valor de Pico •Distorção (γ_1) •Curtose (γ_2) •Parâmetros de Hjorth: <ul style="list-style-type: none"> ○Atividade (A_0) ○Mobilidade (M_0) ○Complexidade (C_0) 	<ul style="list-style-type: none"> •Fator de Crista (FC) •Fator de Impulso (FI) •Fator de Forma (FF) •Fator de Liberação (FL) 	<ul style="list-style-type: none"> •Máximo Expoente de Lyapunov (MEL) •Dimensão de Fractal de Higuchi (DFH)

Figura 4.2 - Indicadores calculados (em cor verde são aqueles pouco explorados).

A investigação dos parâmetros de Hjorth em baixa rotação pode ser vista apenas nos trabalhos de Caesarendra (2015) e Sandoval (2021). Todavia, Caesarendra (2015) estava mais focado na análise de *features* no domínio circular e em *features* não-lineares, ao passo que Sandoval (2021) estava mais empenhado em estudar indicadores de entropia. Diante disso, este trabalho teve o intuito de investigar a capacidade da combinação entre os parâmetros de Hjorth, o MEL e a DFH em diagnosticar sinais de vibração dos rolamentos em baixa rotação.

4.2.1 Atributos Estatísticos e Fatores

Os preditores estatísticos e os fatores calculados foram os 11 apresentados na seção 3.1, isto é, calculados com base nas Equações (3.1) à (3.11).

4.2.2 Máximo Expoente de Lyapunov

Como visto na seção 3.5, o cálculo do MEL requer a obtenção da trajetória dinâmica do sistema. Então, é necessário reconstruir o espaço-de-estado do sinal de vibração. O algoritmo utilizado tanto para a reconstrução do atrator do sinal, bem como para a estimativa do valor do MEL, foi o mesmo adotado por Caesarendra (2015), que por sua vez foi baseado na metodologia proposta por Rosenstein et al. (1993). Tal algoritmo faz uma estimativa do pseudo-espaço-de-estado mediante o “método do atraso”, que consiste em criar uma matriz de dimensão $M \times m$. Onde cada M -ésimo vetor corresponde à série temporal disposta em sequência com um atraso J , até o limite da dimensão implícita “ m ” da matriz. Atingindo tal limite, o procedimento segue para a próxima linha até alcançar o último elemento do sinal original. A Equação (4.1) relaciona esses parâmetros da matriz espaço-de-estado, de sorte que “ M ” representa

o número de linhas, “m” o número de colunas, “J” o atraso e “N” é o tamanho do sinal original.

$$M = N - (m - 1) \cdot J \quad (4.1)$$

Assim, dado o conjunto numérico em função do tempo, $\{x_1, \dots, x_N\}$, a matriz \mathbf{X} representativa da trajetória dinâmica da série temporal, tem cada linha sua correspondendo a um vetor: $\mathbf{X} = (\mathbf{X}_1 \dots \mathbf{X}_M)^T$. Sendo \mathbf{X}_i , $i = \{1, \dots, M\}$, um estado do sistema no tempo discreto “i”, sendo que cada \mathbf{X}_i , por sua vez, é dado por $\mathbf{X}_i = (x_i \ x_{i+J} \ \dots \ x_{i+(m-1) \cdot J})$. Caesarendra (2015) reporta que o atraso J pode ser determinado em função da frequência dominante do sinal de vibração. Além disso, Rosenstein et al. (1993) afirmam que, ainda assim, o método da frequência dominante pode não ser muito preciso, o que requer às vezes apelar para tentativas consecutivas até encontrar um valor mais adequado. E, depois tentativas, o valor escolhido foi $J = 1$.

Em relação à dimensão implícita “m”, Rosenstein et al. (1993) reportam que é comum adotar um valor que respeite o teorema de Takens, isto é, $m < 2 \cdot N$. Entretanto, Rosenstein et al. (1993) afirmam que o método proposto por eles tende a apresentar melhores resultados quando $m \ll 2 \cdot N$. Dessa forma, por meio do método do *False-Nearest-Neighbor* (FNN), estimou-se neste trabalho $m = 4$, que aliado ao valor de atraso $J = 1$, definido anteriormente, gerou uma matriz representativa da reconstrução do espaço-de-estado relativamente simples com 4 dimensões.

Após a reconstrução do atrator do sinal de vibração, o algoritmo procura pelo vizinho mais próximo de cada ponto da trajetória. De acordo com Rosenstein et al. (1993), o vizinho mais próximo, $\mathbf{X}_{\hat{j}}$, é definido buscando-se pelo ponto que minimiza a distância para um ponto referencial, \mathbf{X}_j . Antes, cabe destacar que a otimização da Equação (4.2) está sujeita a uma restrição temporal. Ou seja, segundo Mehdizadeh (2019), o ponto procurado deve ser o mais próximo, mas um limite mínimo deve ser respeitado, para evitar que seja selecionado um ponto na mesma trajetória. Essa restrição é definida por $|j - \hat{j}| > \mu$, sendo μ o período médio do sinal, que pode ser estimado mediante $\mu = \lceil f_s / \bar{f}_p \rceil$. Onde f_s é a taxa de amostragem, \bar{f}_p é a frequência média do espectro do sinal e $\lceil \cdot \rceil$ representa a função teto.

$$d_j(0) = \min_{\mathbf{X}_{\hat{j}}} \|\mathbf{X}_j - \mathbf{X}_{\hat{j}}\| \quad (4.2)$$

$$\text{Sujeito a: } |j - \hat{j}| > \mu = \left[\frac{f_s}{f_p} \right] \quad (4.3)$$

Observa-se por meio da Equação (4.2) que a distância entre os vizinhos mais próximos é mensurada com a norma euclidiana, como deveras proposto por Rosenstein et al. (1993). Entretanto, este trabalho investigou a influência da substituição da distância euclidiana por outra métrica baseada no cálculo do coeficiente de correlação de Pearson por meio da Equação (3.35), denominada aqui “distância Pearson”. Visando comparar o Euclid-MEL e o Pearson-MEL, foi adotado também o teste estatístico *t* de Welch, para comparar as médias das classes.

A Figura 4.3 mostra que o algoritmo realiza o cálculo das distâncias entre pontos vizinhos na trajetória X com o suporte de um *loop* (retângulo em linhas tracejadas). Fora do *loop*, observa-se a etapa inicial de reconstrução da trajetória, cuja matriz serve de dado de entrada para a sua etapa de replicação para X_{rep} . Ademais, é possível ver que a cada iteração i é realizada a diferença entre a matriz da trajetória (espaço-de-estado) pela matriz replicada.

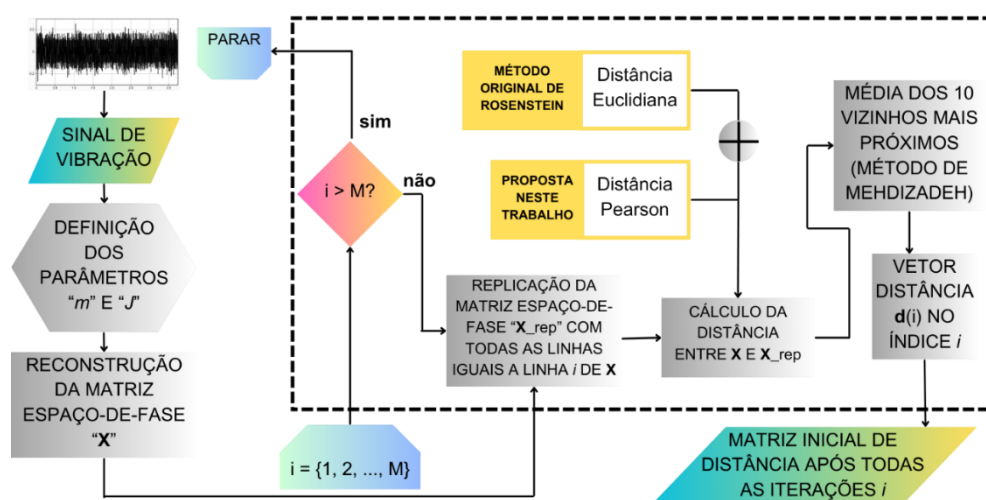


Figura 4.3 - Algoritmo para determinar a matriz inicial de distâncias.

Nesta etapa, a proposta de Mehdizadeh (2019) de determinar o vizinho mais próximo como a média de certo número de vizinhos mais próximos também foi considerada neste trabalho. Neste caso, Mehdizadeh (2019) constatou que a média calculada com mais de 15 vizinhos não se tem um ganho significativo. Neste trabalho, adotou-se a quantidade de 10 vizinhos devido à “saturação” ocorrer a partir do 10º vizinho. A diferença ou distância calculada em cada iteração é organizada em outra matriz de saída, que reúne os valores das distâncias iniciais calculadas.

Por fim, conforme o algoritmo da Tabela 4.1, a matriz inicial de distâncias \mathbf{d}_0 resultante do algoritmo da Figura 4.3 é atualizada por meio da condição imposta pela Equação (4.3), ou seja, a restrição temporal. Diante disso, uma nova matriz, \mathbf{d}_{0_nova} , é estabelecida, contendo apenas as distâncias que satisfazem à restrição temporal. Em seguida, o algoritmo calcula a distância entre a matriz \mathbf{X} original e uma nova matriz $\mathbf{X}(\Delta_b)$ cujas linhas correspondem à média dos dez vizinhos mais próximos de cada linha de \mathbf{X} . Por sua vez, Δ_b é um vetor que contém os índices dos vizinhos mais próximos de cada estado de \mathbf{X} .

Tabela 4.1 – Continuação do algoritmo para o cálculo do MEL, por Caesarendra (2015).

Continuação do algoritmo para o cálculo do MEL	
1:	\mathbf{X} e \mathbf{d}_0 são calculados conforme o algoritmo da Figura 4.3.
2:	μ é o período médio calculado usando a Equação (4.3).
3:	f_s é a taxa de amostragem do sinal de vibração.
4:	k_{max} é o limite máximo da faixa de expansão.
5:	Agora, antes de encontrar os vizinhos próximos para cada ponto da trajetória,
6:	torna-se necessário impedir que os pontos referenciem a eles mesmos. Para tanto,
7:	adota-se uma distância mínima de segurança baseada no período médio μ . Assim,
8:	cria-se uma nova matriz \mathbf{d}_{0_nova} , fazendo inicialmente $\mathbf{d}_{0_nova} = \mathbf{d}_0$. Em seguida:
9:	for $i=1$ to M do
10:	for $j=1$ to M do
11:	if $ j-i \leq \mu$ then
12:	$\mathbf{d}_{0_nova}(j) = \max \mathbf{d}_0 $
13:	end if
14:	end for j
15:	Encontrar o valor mínimo $\Delta_a(i)$ de $\mathbf{d}_{0_nova}(j)$ e seu índice posicional $\Delta_b(i)$:
17:	$[\Delta_a(i), \Delta_b(i)] = \min \mathbf{d}_{0_nova}(j) $
18:	end for i
19:	for $k=1$ to k_{max} do
20:	bound= $M-k$
21:	evolve=0
22:	point=0
23:	for $h=1$ to M do
24:	if $h \leq \text{bound}$ and $\Delta_b(h) \leq \text{bound}$ then
25:	$d = \sqrt{\sum [\mathbf{X}(h+(k-1), :) - \mathbf{X}(\Delta_b(h)+(k-1), :)]^2}$
27:	if $d \neq 0$ then
28:	evolve=evolve+ $\ln d $
29:	point=point+1
30:	end if
31:	end if
32:	end for h
33:	if point>0 then
34:	$\mathbf{d}_{nova}(k) = \text{evolve}/\text{point}$
35:	else
36:	$\mathbf{d}_{nova}(k) = 0$
37:	end if
38:	end for k
39:	Enfim, encontra-se o coeficiente angular α da curva que se ajusta a $\mathbf{d}_{nova}(k)$, logo:
40:	MEL= αf_s

O cálculo da distância entre as matrizes citadas acima tem o objetivo de encontrar a taxa média de separação dos vizinhos mais próximos ao longo da trajetória. Caesarendra (2015) aplica nesta fase um *loop* iterativo em função de uma Faixa de Expansão delimitada por k_{\min} e k_{\max} . Assim, o cálculo do MEL, conforme visto na Equação (3.26), com base no método proposto por Rosenstein et al. (1993) de extrair a média das distâncias dos pontos, feito em função de “k” iterações, pode ser dado pela seguinte expressão:

$$\lambda_{\max}(i) = \frac{f_s}{(M - k)} \sum_{j=1}^{M-k} \ln d_j(i) \quad (4.4)$$

Em que, na verdade, o MEL, λ_{\max} , consiste no coeficiente angular da reta que melhor se ajusta à curva da Equação (4.4) em uma dada Faixa de Expansão, como ilustrado na Figura 4.4. Neste trabalho, a Faixa de Expansão foi estabelecida entre $k_{\min} = 1$ e $k_{\max} = 4$, pois, nesta faixa, encontrou-se uma tendência mais monotônica. A Figura 4.4 trata-se de um exemplo elucidativo do procedimento necessário para a estimativa do valor do MEL, onde a Faixa de Expansão está delimitada entre 20,9 e 160,8 pelas linhas verdes tracejadas. O MEL é equivalente a inclinação da reta correspondente a linha vermelha tracejada, que melhor se ajusta aos dados representados pela curva azul.

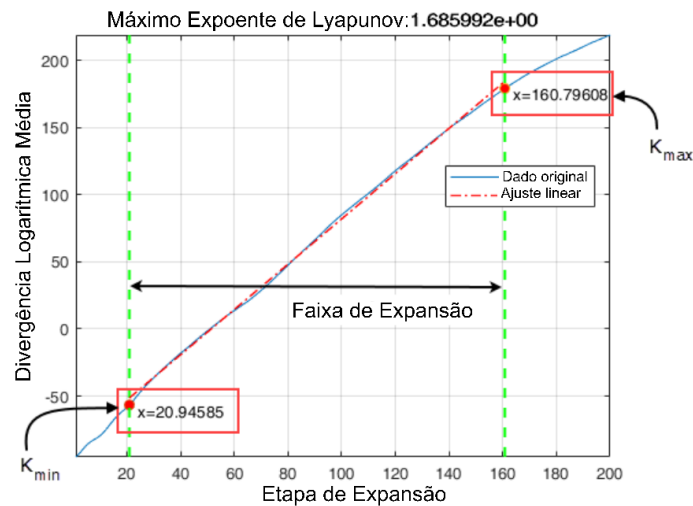


Figura 4.4 - Ajuste da reta aos dados em função da faixa de expansão
[Fonte: Adaptado de MathWorks® (2024d)].

4.2.3 Dimensão Fractal de Higuchi

Um sinal de vibração é considerado fractal uma vez que suas características se mantêm em maiores ou menores escalas de tempo (SMITS et al., 2016). Neste trabalho, a estimativa da dimensão fractal dos sinais segue a metodologia proposta por Tomoyuki Higuchi em 1988, a qual foi abordada na seção 3.6. A Tabela 4.2 corresponde ao algoritmo que pode ser utilizado para calcular o valor da Dimensão Fractal de Higuchi (DFH), o qual varia entre 1 e 2. Assim, quanto mais próximo de “2”, maior o nível de complexidade do sinal de vibração. Vale ressaltar que o valor de k_{max} é o único parâmetro a ser definido, e refere-se ao número de iterações. O método basicamente consiste em criar novas séries temporais a cada iteração “k” por reamostragem do sinal original (*downsampling*). Dessa forma, tal parâmetro influencia diretamente no resultado, e neste trabalho, após tentativas, foi definido como $k_{max} = 16$, pois verificou-se que este valor apresentou melhores resultados com um custo computacional relativamente baixo.

Tabela 4.2 - Algoritmo para o cálculo da DFH.

Algoritmo para o cálculo da DFH	
1:	N é o tamanho do sinal ou número de pontos.
2:	Os próximos três vetores nulos definidos serão preenchidos
3:	ao longo das iterações seguintes.
4:	$L = \text{zeros}(1, k_{max})$
5:	$x = \text{zeros}(1, k_{max})$
6:	$y = \text{zeros}(1, k_{max})$
7:	for k=1 to k_{max} do
8:	for m=1 to k do
9:	normFactor= $((N-1)) / ((N-m)k)$
10:	$X = \sum_{i=1}^{\lfloor \frac{N-m}{k} \rfloor} X(m+ik) - X(m+(i-1)k) $
11:	$L(m) = (X * \text{normFactor}) / k$
12:	end for m
13:	$y(k) = \ln(\sum L/k)$
14:	$x(k) = \ln(1/k)$
15:	end for k
16:	Finalmente, a dimensão fractal “DFH” é o coeficiente angular
17:	da curva que se ajusta à função y(k) no domínio x(k).
18:	

4.3 Seleção de features

Neste momento, tendo extraído todas as *features* requeridas dos sinais de vibração, segue-se com a etapa de seleção das melhores. E, inicialmente, foi utilizado o valor estatístico do teste *t* de Welch para medir o grau de separação entre as classes “normal” e “defeito”. E, ao final, aplicou-se a PCA aos atributos selecionados pelo teste

estatístico mencionado, de maneira a avaliar a viabilidade de reduzir a dimensionalidade dos dados.

4.3.1 Ranqueamento de indicadores mediante valor t de Welch

Como mencionado, o valor estatístico do teste t de Welch foi utilizado como um índice que mensura o quão separadas estão as classes. Conforme a Equação (3.12), observa-se que o valor t tende a ser elevado quando a diferença é grande entre as médias de ambas as distribuições, e suas variâncias são pequenas. Ao contrário, se as duas distribuições possuem uma pequena diferença entre as médias e/ou as variâncias são extensas, o valor de t tende a reduzir. Por esse motivo que o valor t do teste de Welch foi adotado para servir de indicador do grau de separabilidade das classes analisadas.

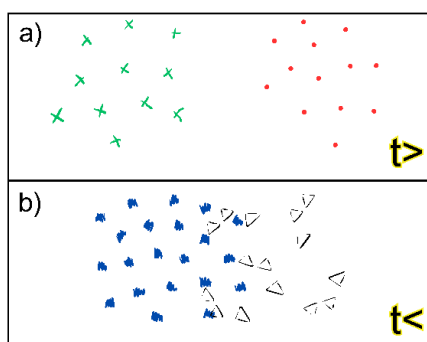


Figura 4.5 – Comparação entre distribuições binárias com: a) t maior; b) t menor.

A Figura 4.5 ilustra dois exemplos de distribuição com duas classes por meio dos quais observa-se que, em virtude da distribuição da Figura 4.5a ser linearmente separável, o valor estatístico t apresenta maior magnitude em relação à distribuição com classes parcialmente sobrepostas da Figura 4.5b.

Assim, foi calculado o valor de t para todos os atributos extraídos dos sinais, e, em seguida, foram selecionados aqueles que apresentaram a maior magnitude de t , isto é, a maior capacidade de distinção entre classes. Para tanto, adotou-se um critério de seleção, que consiste em eliminar os preditores que possuam valor de “ t ” abaixo de 50% da média dos três maiores valores. Contudo, este percentual pode ser alterado consoante especificidades de outros problemas.

A justificativa para o uso dessa técnica para a seleção de *features* deve-se ao fato de que aquelas com maiores índices são melhores para representar as características do sinal de vibração para o treinamento de um algoritmo de

aprendizado supervisionado. Uma vez que *features* de alta separabilidade podem elevar a capacidade de generalização do modelo.

4.3.2 Redução de dimensionalidade por PCA

Tendo escolhido os melhores indicadores pelo método do teste *t* de Welch, foi empregada ainda a técnica algébrica de PCA à distribuição composta pelos melhores atributos, visando a possibilidade de reduzir a dimensionalidade do espaço de *features*. O procedimento adotado para o cálculo da PCA segue exatamente o disposto na seção 3.3. Basicamente a PCA realiza uma transformação linear no espaço original dos preditores originando um “subespaço principal”. Ao passo que suas “componentes principais” compõem um novo sistema de coordenadas não-correlacionadas entre si. Além disso, o gráfico *biplot* da Figura 4.6 relaciona os *scores* e os *loadings* no plano das componentes principais, proporcionando uma ampla visualização do resultado da aplicação da PCA. Em ambos os eixos vertical e horizontal (CP1 e CP2) consta o percentual de representatividade da informação retida no modelo por parte de cada componente principal.

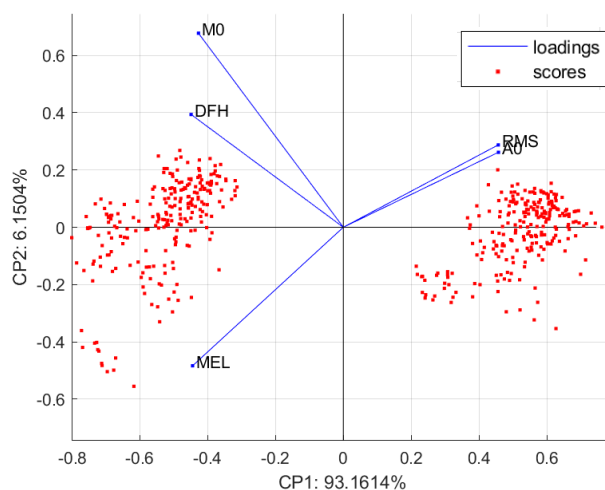


Figura 4.6 - Exemplo de gráfico biplot no plano das componentes principais.

Dessa forma, decidiu-se por reduzir a dimensionalidade dos dados para um espaço principal tridimensional. Em outras palavras, foram selecionados os três autovetores cujos respectivos autovalores associados eram os três maiores. Deve ser destacado que há uma perda de informação no processo. Todavia a soma das representatividades das três componentes principais geralmente retém altos percentuais de informação (isto quando o número inicial de atributos é relativamente

baixo, como neste trabalho). Além disso, o gráfico *biplot* da Figura 4.6 foi construído para verificar quais os preditores que são mais significativos para a explicação da informação, em virtude do peso (*loading*) de cada atributo projetado nas componentes. Diante disso, tem-se um novo espaço de *features*, cujas coordenadas são compostas agora pelas componentes principais, ou seja, houve uma espécie de transformação das *features* originais com a aplicação da PCA.

4.4 Máquina de Vetores de Suporte (SVM)

Os dados passados pela etapa de seleção de preditores estão prontos para servir de dados de treinamento para o algoritmo de aprendizado supervisionado SVM, uma vez que cada indivíduo da tabela possui seu respectivo rótulo, indicando a classe a que pertence. Contudo, os dados precisaram passar por mais um processamento: eliminação de *outliers*.

4.4.1 Eliminação de potenciais outliers

Para identificação de elementos como potenciais *outliers* foi utilizada a carta T^2 de Hotelling, que compreende a um gráfico que exhibe a distância de Mahalanobis de cada ponto em relação ao centro da distribuição. Neste caso, foi feita uma carta para cada classe analisada.

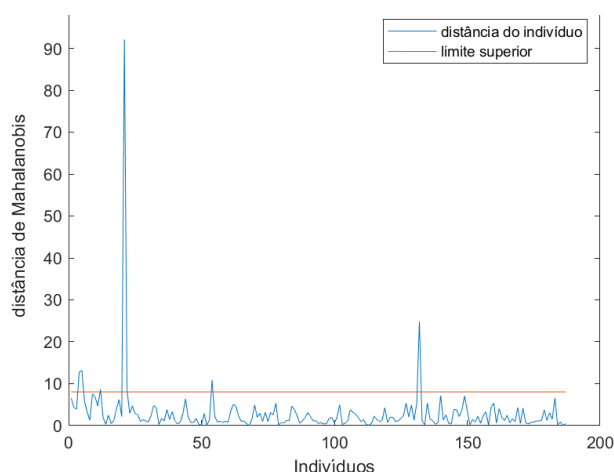


Figura 4.7 - Carta T^2 de Hotelling com linha limite.

Em problemas binários, por exemplo, haveriam duas cartas T^2 de Hotelling. Vale ressaltar que a medição da distância de cada indivíduo para o centro da distribuição é feita no espaço p -dimensional, sendo “ p ” o número de componentes ou

preditores. A Figura 4.7 ilustra, a título de exemplo, a aplicação da carta T^2 de Hotelling para identificação de *outliers*, na qual é possível identificar que um dos pontos se encontrava muito distante do centro do conjunto a que pertencia.

A linha limite é estimada estatisticamente pela Equação (3.22) com base na distribuição estatística F e um respectivo valor de significância α , dado em percentual. O qual influencia diretamente no grau de tolerância da linha limite. Ou seja, dependendo do valor percentual de α empregado, a linha limite pode subir ou descer sobre a carta. O valor de significância adotado na metodologia deste trabalho foi fixado em 5%, por se tratar de um valor comumente utilizado na literatura (o mesmo valor utilizado no exemplo da distribuição da Figura 4.7).

4.4.2 Treinamento do algoritmo SVM

Primeiramente, foi definido que a etapa de treinamento seria avaliada com Validação-Cruzada de 5 pacotes. Em seguida, a Equação (3.47) mostra a possibilidade de escolher uma função kernel, e, neste trabalho, a função kernel adotada foi simplesmente a linear. A justificativa se baseia no fato de que se procura treinar um modelo o mais generalista possível. Funções kernels de natureza não-linear podem gerar hiperplanos que se ajustem demais aos dados de treinamento, aumentando a possibilidade de *overfitting*.

Sobretudo considerando a natureza razoavelmente estocástica da distribuição dos dados extraídos de sinais de vibração. Tal fenômeno pode fazer com que conjuntos de treinamento e teste sejam bastante diferentes. Diante disso, um hiperplano de alto grau de não-linearidade que esteja extremamente ajustado aos dados de treinamento pode vir a falhar disparatadamente perante novos dados de teste. Assim, hiperplanos lineares podem estabelecer zonas de separações mais genéricas para a maioria dos casos de classificação de dados oriundos de sinais de vibração.

Outra questão extremamente importante é a parametrização do algoritmo SVM no tocante ao parâmetro de penalização "C". A definição do parâmetro C passa a ser relevante, pois ele controla o nível de tolerância a erros de classificação, e isso influencia diretamente na direção do hiperplano e na largura da margem. O valor utilizado foi $C = 1$, que se trata de um valor relativamente baixo. E isso significa que,

caso os dados não sejam linearmente separáveis, a superfície de decisão deverá ser definida de forma que alguns erros de classificação possam ser tolerados.

Novamente a preocupação em definir uma superfície de decisão mais generalista possível foi levada em consideração. Uma vez que elevados valores de “C” podem implicar em hiperplanos muito ajustados às bordas das distribuições, além de serem mais influenciados por *outliers*. Por outro lado, menores valores de “C” permitem a geração de superfícies de decisão com margens mais abrangentes, além de diminuir as possibilidades de *overfitting*.

A princípio, um modelo foi treinado em um cenário multiclasse *one-versus-all* com dados de 4 classes. Tais dados foram coletados de ensaio em um misturador de baixa rotação, cujo procedimento experimental está descrito na seção 4.5 a seguir. Outrossim, um modelo binário (presença ou ausência de defeito) foi treinado com base nos dados adquiridos no misturador. Para testar o modelo, empregou-se dados de uma bancada experimental relativamente simples composta por um sistema rotativo motor-eixo-mancal. Ademais, outros dados oriundos de mancais de tambores de correias transportadoras atuando em campo de mineração.

4.5 Coleta de dados do misturador

Os dados foram adquiridos do Misturador MT120 com um DynaLogger HF+ da Dynamox®, em que a Figura 4.8a ilustra uma visão ampla do equipamento, a Figura 4.8b mostra uma captura em operação, e a Figura 4.8c descreve suas especificações técnicas. O sensor foi fixado na luneta do misturador próximo ao rolamento rígido de esferas modelo 6308 2RS, conforme visto na Figura 4.9a (fotografia do rolamento dentro da luneta) e na Figura 4.9b (desenho da luneta).

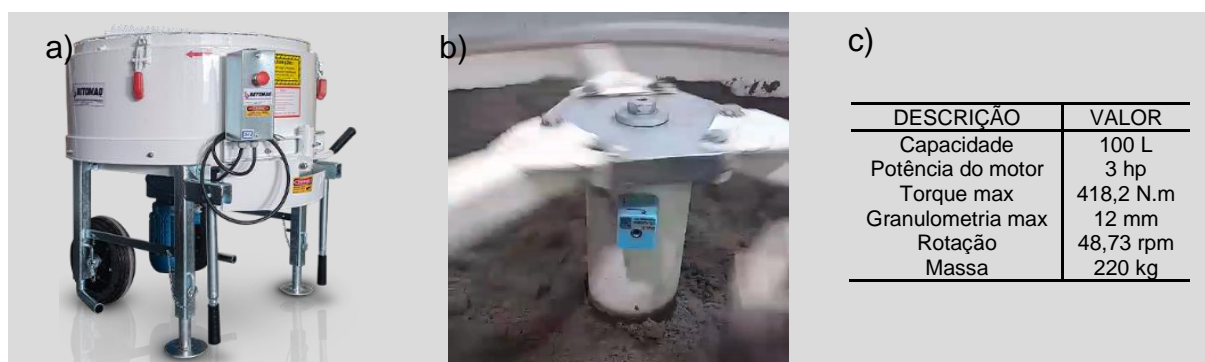


Figura 4.8 - Misturador MT120 utilizado como bancada de ensaios.

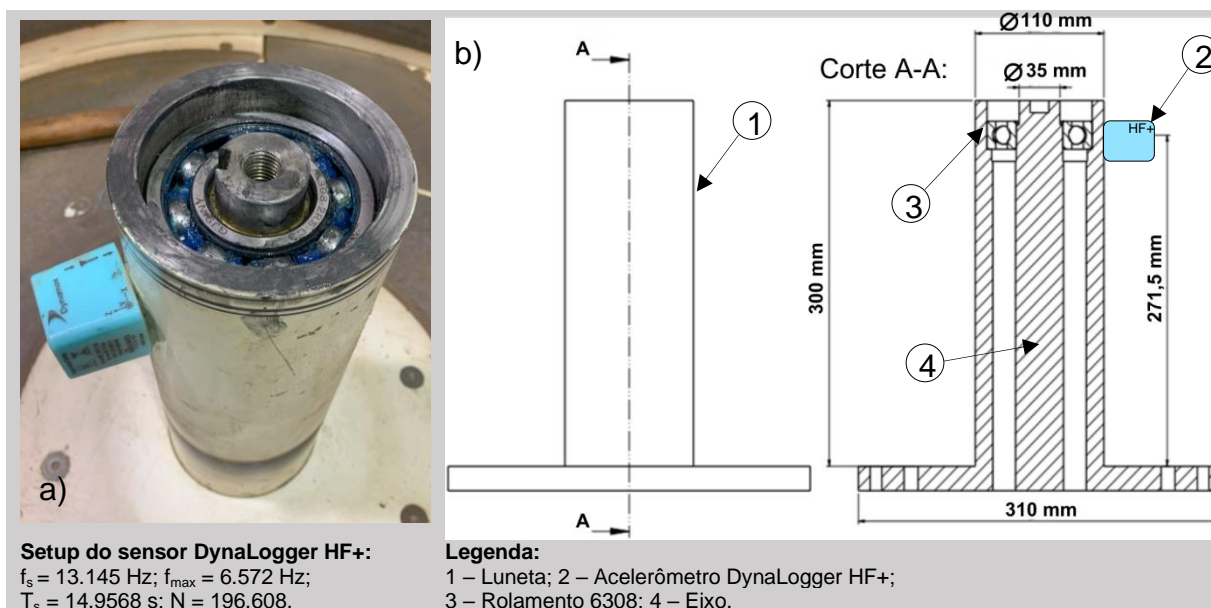


Figura 4.9 - Posicionamento do acelerômetro próximo ao rolamento testado.

Os ensaios consistiram principalmente em testar rolamentos **com** e **sem** defeito, variando o nível de carga do misturador em: sem carga (L0), e cargas de 40% (L40) e 80% (L80) da capacidade total. O defeito foi artificialmente provocado na superfície da pista externa do rolamento. E com o intuito de provocar danos incipientes, optou-se por realizar a simulação de riscos ou arranhões na superfície. Esse tipo de avaria geralmente está associado a uma das decorrências iniciais de falha do filme lubrificante entre os componentes internos do rolamento, sobretudo quando o mesmo opera sob altos níveis de carregamento. Dessa maneira, a lubrificação inadequada pode desencadear tais arranhaduras; as quais, *a posteriori*, podem evoluir para condições ou danos mais severos, como superaquecimento, trincas, fraturas, desgaste, descascamento, escamamento, corrosão, lascamento, esmagamento, entre outros.

Diante disso, visando inserir arranhaduras na pista externa do rolamento, optou-se por adotar breves sessões de lixamento da superfície com lixa especial para metais de granulação 60. Vale ressaltar que foi necessário retirar a lubrificação do rolamento e relubrificar-lo novamente após o lixamento. Além disso, executou-se três níveis de severidade da falha, isto é, primeiramente, a fração de 25% (D25) da superfície da pista externa foi danificada, em seguida 50% (D50), e, finalmente, 100% (D100), conforme apresenta a Figura 4.10. A execução do lixamento da superfície seguiu padrões semelhantes, visando manter certo homeomorfismo entre as diferentes zonas danificadas. Em função disso, a Figura 4.11 apresenta fotografias

das superfícies lesadas, nas quais é possível observar que a região de 25% - Figura 4.11a e Figura 4.11c – apresenta aspectos de riscos ou arranhaduras semelhantes àqueles vistos na região de 100% - Figura 4.11b e Figura 4.11d. Vale justificar que a presença de sombreamento nas fotografias da Figura 4.11 foi difícil de evitar devido à característica côncava da superfície da pista externa.

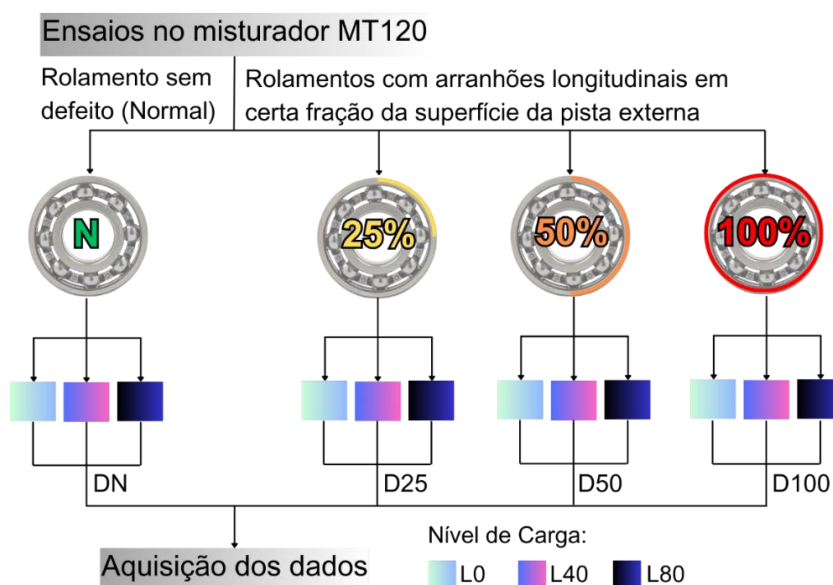


Figura 4.10 - Etapas de aquisição de dados em ensaios com o misturador.

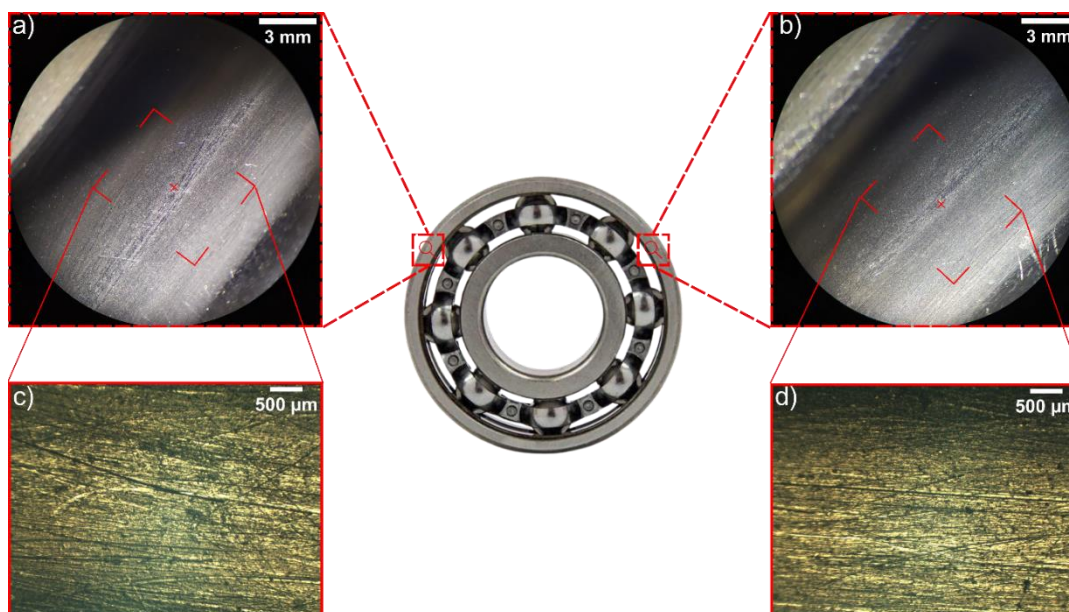


Figura 4.11 – Superfícies da pista externa do rolamento danificadas.

Ao final, foram realizadas 20 coletas de dados de vibração para cada uma das 12 condições de ensaio previstas na Figura 4.10.

4.5.1 Análise do impacto do tamanho da janela do sinal

Durante a coleta dos sinais de vibração, o *setup* do DynaLogger HF+ ocorreu conforme previsto na Figura 4.9, resultando na aquisição de uma série temporal de $T_s = 14,9568$ s e tamanho $N = 196.608$ pontos. O período de aquisição T_s corresponde ao tempo necessário para 12 revoluções completas do eixo do misturador, uma vez que o mesmo gira com velocidade em torno de 48,73 rpm (ver especificações do equipamento na Figura 4.8). Diante disso, realizou-se a análise de qual seria o melhor tamanho de janela do sinal para utilização na etapa de extração de *features* do mesmo. Dessa forma, iniciou-se analisando uma janela com período suficiente para cobrir 3 revoluções do eixo, resultando em um particionamento do sinal bruto de 14,9568 s em quatro fragmentos de 3,7391 s. Assim, em cada um dos 12 conjuntos de dados coletados, para os quais inicialmente havia 20 coletas, o número sobe para 80 depois da partição em janelas de três revoluções.

Após a etapa acima, cada um dos 80 fragmentos obtidos foi tratado como uma coleta individual a partir da qual as *features* deveriam ser extraídas. Além disso, outros tamanhos de janela foram analisados, a fim de efetivamente encontrar um janelamento mais adequado de se trabalhar. A Tabela 4.3 apresenta em detalhes os diferentes janelamentos baseados no número de revoluções do eixo, além do tempo requerido e do tempo utilizado (sendo que este sempre foi um valor ligeiramente acima do tempo requerido). Vale ressaltar que, primeiramente, alterou-se o tamanho da janela do sinal inicial bruto de 14,9568 s (em que se tinha 12 revoluções) para quatro fragmentos de 3,7391 s com 3 revoluções cada. Diante disso, os demais tamanhos de janela – da faixa entre 0,1 revolução e 2 revoluções – foram obtidos a partir da janela de 3 revoluções, conforme ilustra a Figura 4.12.

Tabela 4.3 – Detalhamento do janelamento dos sinais.

Nº de revoluções	Tempo requerido (s)	Tempo utilizado (s)	Nº de amostras
3	3,6938	3,7391	49.152
2	2,4625	2,4927	32.758
1	1,2313	1,2463	16.384
0,5	0,6156	0,6231	8.192
0,25	0,3078	0,3115	4.096
0,1	0,1231	0,1247	1.639

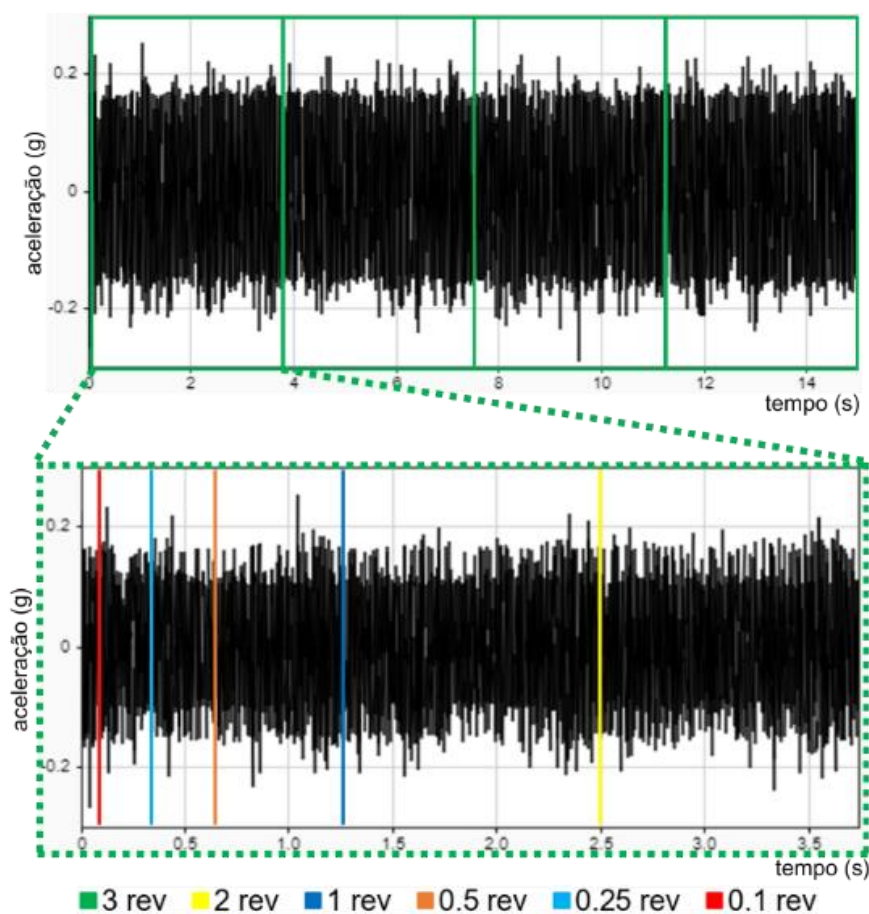


Figura 4.12 - Janelamento do sinal baseado no número de revoluções do eixo.

4.6 Sistema Motor-Eixo-Mancal (MEM)

A bancada apresentada na Figura 4.13 consiste em um sistema mecânico rotativo relativamente simples. O qual é composto basicamente por um motor elétrico trifásico AC de indução de 3 polos e potência de 3 hp (associado a um inversor de frequência WEG CFW08) acoplado a um eixo suportado por um mancal de rolamento do tipo 6202z.

Em relação aos ensaios, foram aqisitados dados de sinais de vibração de apenas duas condições do rolamento, a saber, com e sem defeito, a uma velocidade de rotação de 60 rpm. O defeito se trata de uma falha incipiente simulada em determinado ponto da pista externa do rolamento (BPFO). Foram aqisitados 125 observações para cada condição do rolamento, totalizando em 250. Finalmente, a Figura 4.13 também indica a posição em que se fixou o acelerômetro de sensibilidade de 101 mV/g, e sobretudo revela os parâmetros de aquisição configurados para a coleta dos sinais.

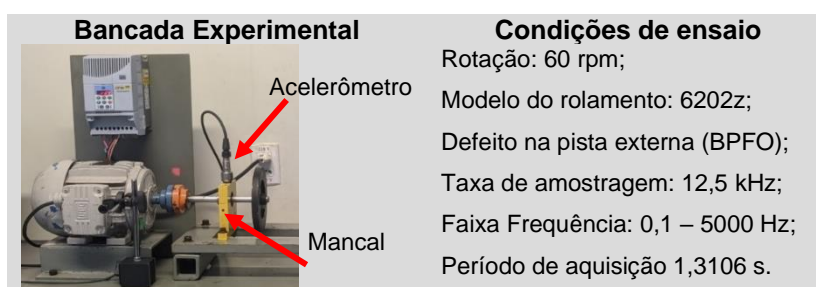


Figura 4.13 - Bancada experimental sistema MEM.

4.7 Dados de tambores de transportadores de correia

Foram coletados dados de vibração de mancais de tambores de transportadores de correias operantes em pátio de mineração, precisamente de tambores de acionamento, de retorno e de encosto. A Figura 4.14 ilustra uma estrutura básica de um equipamento industrial transportador de correia, na qual é possível identificar a posição de cada um dos tambores mencionados. No total, foram monitorados cinco mancais de tambores, dos quais: um de encosto, um de retorno e três de acionamento. A Tabela 4.4 apresenta informações específicas dos *spots* monitorados, como o tipo de tambor, o modelo do rolamento acoplado ao respectivo tambor, o diâmetro externo do rolamento, e a velocidade de rotação do eixo do tambor.

Tabela 4.4 - Rolamentos SKF® W33 acoplados aos mancais dos tambores.

Nº do Dataset	Tipo de tambor	Modelo do rolamento	Diâmetro Externo (mm)	Velocidade de Rotação (rpm)
1	Encosto	22234 CC	310	83
2	Retorno	22234 CCK	310	83
3	Acionamento	22244 CCK	400	59
4	Acionamento	23084 CA	620	39
5	Acionamento	23048 CCK	360	53

Todos os cinco rolamentos degradaram-se naturalmente, ou seja, ao longo do tempo de operação do equipamento. Conforme relatórios de prognósticos, as falhas apresentaram-se como defeito na pista externa do rolamento (BPFO). Por fim, as configurações dos sensores foram todas iguais, sendo a taxa de amostragem igual a 2.048 Hz, faixa de frequência até 1 kHz, e período de aquisição de 4 s. Ademais, foi

adotado um método de janelamento do sinal de modo que cada segmento contenham o período necessário para que o eixo execute uma rotação completa.

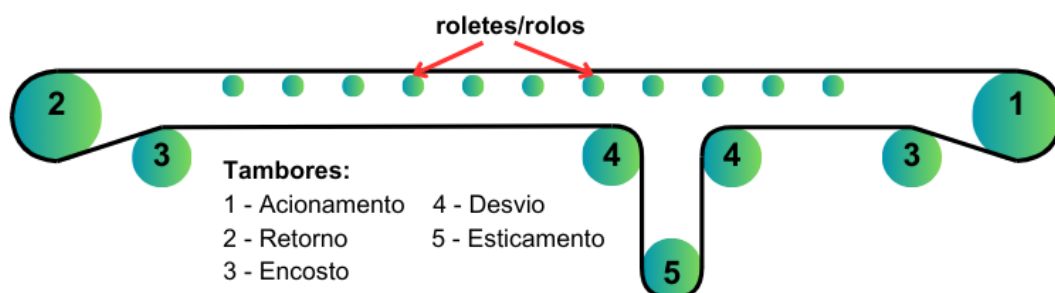


Figura 4.14 - Estrutura básica de um transportador de correia.

5 RESULTADOS E DISCUSSÕES

A presente seção apresenta os resultados obtidos, bem como as discussões e interpretações acerca dos mesmos. Inicialmente, a seção 5.1 traz o resultado da análise da influência do tamanho da janela do sinal de vibração na performance dos preditores em distinguir entre classes. Em seguida, o tópico 5.2 enfoca no ranqueamento das *features* extraídas dos sinais, visando aplicar, assim, o critério de seleção. Vale ressaltar que os resultados de comparação entre os indicadores Pearson-MEL e Euclid-MEL também são abordados. Ao final, o que se obteve com treinamento e teste de modelos de classificação SVM multiclasse e binário são vistos nos itens 5.3 e 5.4, respectivamente.

5.1 Análise da influência do tamanho da janela do sinal

A Figura 5.1 apresenta gráficos com todos os atributos extraídos dos sinais de vibração – inclusive o Pearson-MLE, proposta neste trabalho. Foram arranjados três cenários confrontando duas classes diferentes: DN *versus* D25; D25 *versus* D50; e D50 *versus* D100. Além disso, o nível de carregamento imposto ao misturador também foi um parâmetro analisado. O valor estatístico t de Welch foi utilizado de forma a mensurar a separabilidade de classes, e seus valores encontram-se no “eixo radar” de cada gráfico.

Ao observar os nove cenários da Figura 5.1, concluiu-se que quanto menor o tamanho da janela, menor a capacidade de distinção entre classe por todas as *features* analisadas. Além disso, a diminuição do valor de t torna-se ainda mais severa para janelas inferiores a 1 (uma) revolução completa. Por outro lado, nota-se que, de modo amplo, comparando janelas compostas por uma, duas ou três revoluções, a diferença entre elas não se apresentou de forma significativa.

Ainda sobre a Figura 5.1, vale destacar que já nesta análise é possível observar que as *features* não-lineares (sobretudo a DFH), parâmetros de Hjorth e RMS apresentaram maior significância, mesmo com a variação do janelamento dos sinais. Outro ponto que vale mencionar se refere a influência da carga utilizada no misturador durante os ensaios. Notou-se que o nível de carga altera expressivamente o desempenho dos preditores, embora não tenha sido detectado nenhum padrão claro da influência do nível de carga.

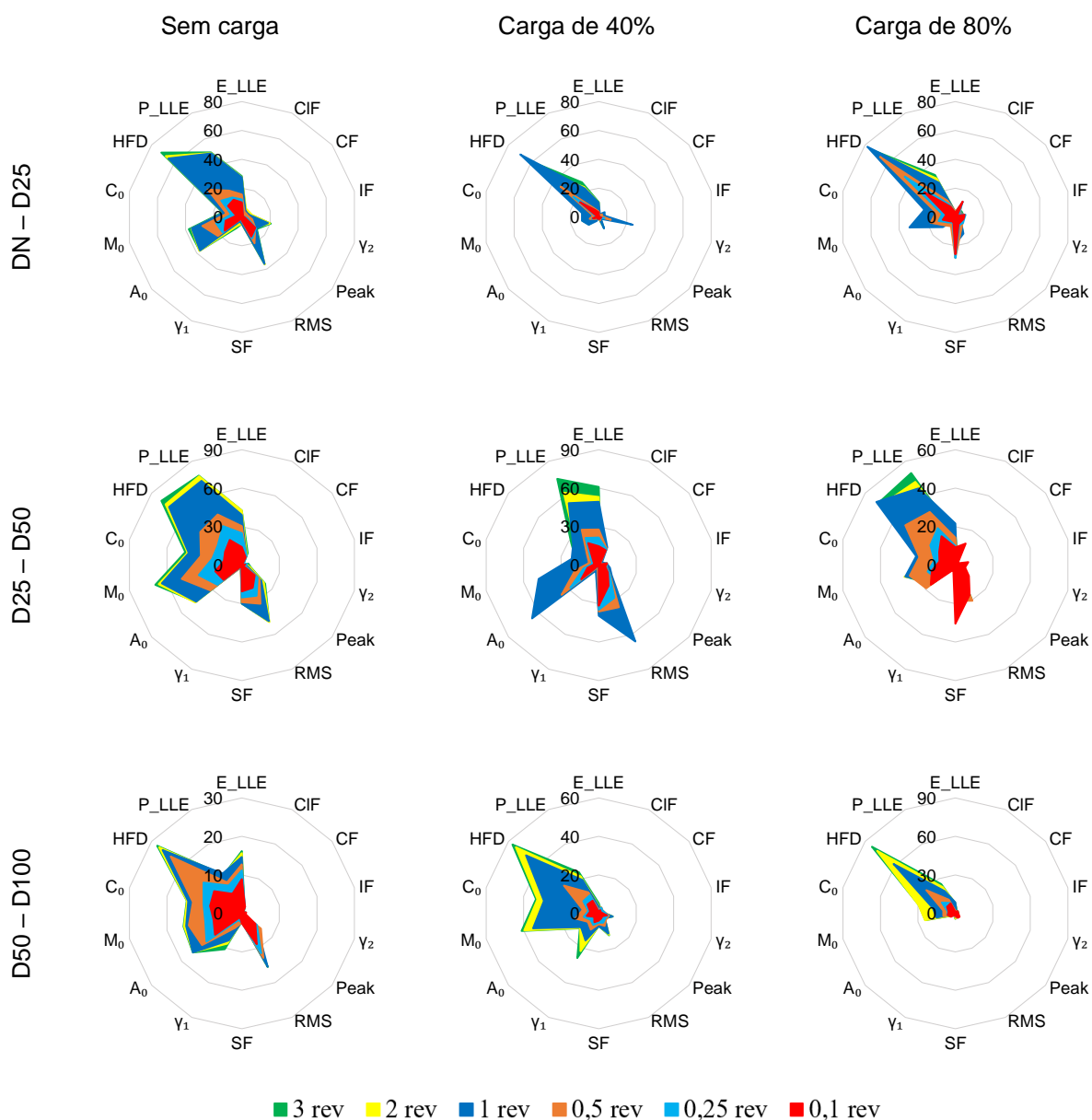


Figura 5.1 – Influência do janelamento do sinal com base no valor t de Welch.

Portanto, concluiu-se que, trabalhar com janelamento que abranja apenas uma revolução, seja o meio mais adequado. Uma vez que janelas maiores demandam maior custo computacional na etapa de extração de preditores do sinal, sem haver um ganho significativo na capacidade de distinção entre classes. Diante disso, decidiu-se por fragmentar os sinais de vibração (inicialmente com janela de 12 revoluções) em segmentos de 1 revolução, conforme mostra a Figura 5.2.

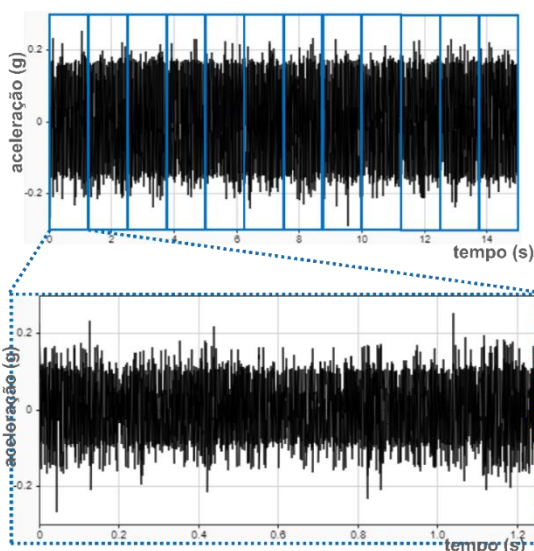


Figura 5.2 – Janelamento dos sinais de vibração para 1 revolução.

5.2 Ranqueamento e seleção de *features*

A Figura 5.3 e a Figura 5.4 apresentam o ranqueamento das 5 melhores *features* calculadas para 6 panoramas de classes binárias. A área de cada um dos três retângulos coloridos diz respeito a qual nível de carga reuniu indicadores de maior separabilidade. De outro modo, cada área colorida foi subdividida em retângulos menores para cada preditor, cujas áreas também são proporcionais ao valor t de Welch. Assim, tais gráficos são capazes de ilustrar, ao mesmo tempo, a influência do nível de carga na separabilidade das *features*, bem como ranqueá-las.

Em relação à influência do nível de carga, observa-se que nas combinações em que a classe de D25 esteve presente, os atributos obtiveram maior significância de t na ausência de carga no misturador. Neste caso, pode ter acontecido de, sem carga, os sinais terem sofrido menos com interferências externas provenientes do processo de mistura.

Em contrapartida, quando as combinações envolveram as classes D50 e/ou D100, observou-se que as *features* tiveram maior grau de separabilidade em condições com carga. Talvez por ter ocorrido aumento de torque no eixo, elevando tensões no rolamento, realçando, assim, as falhas mais severas. Em condição sem carga, as *features* não separaram bem as classes. Provavelmente por possuírem uma distribuição maior de arranhões ao longo da superfície, elevando o grau de irregularidade do sinal, fazendo com que os mesmos se confundam com um sinal característico de DN em que o índice SNR é normalmente elevado.

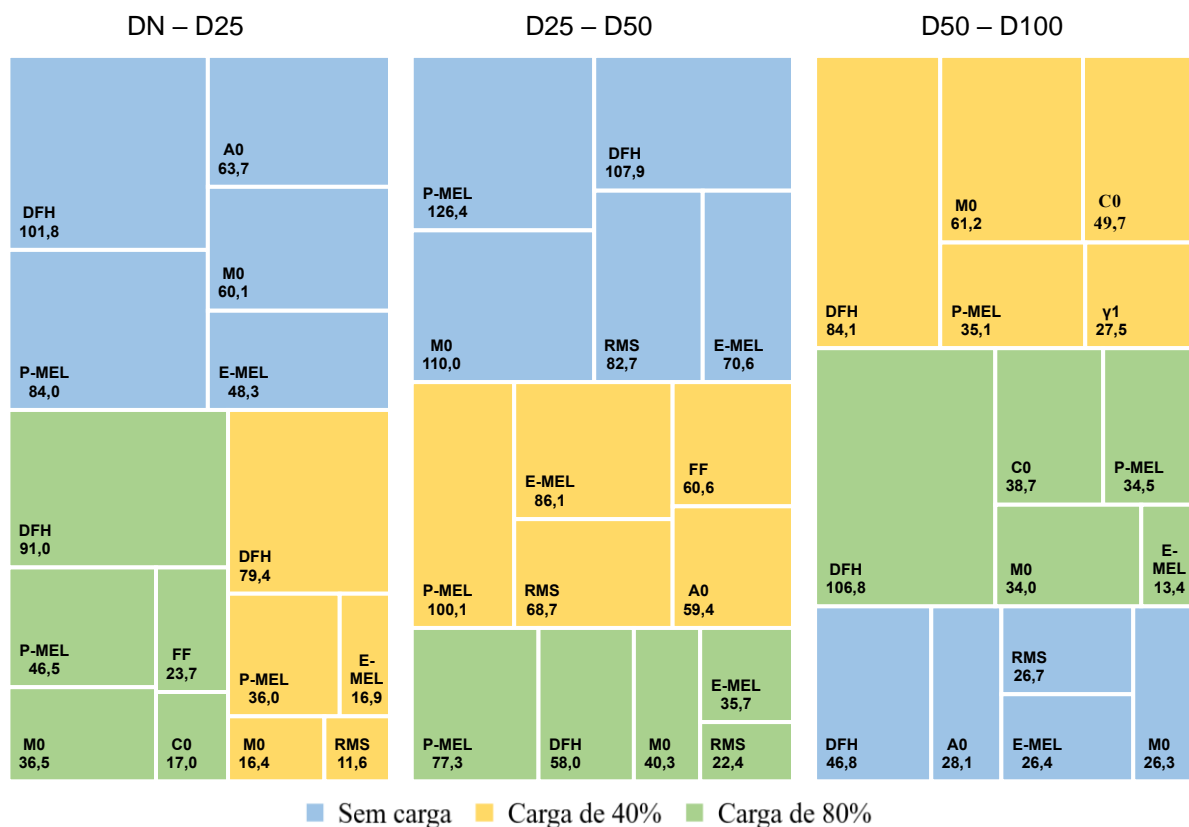


Figura 5.3 – Ranque dos 5 melhores preditores (parte 1).

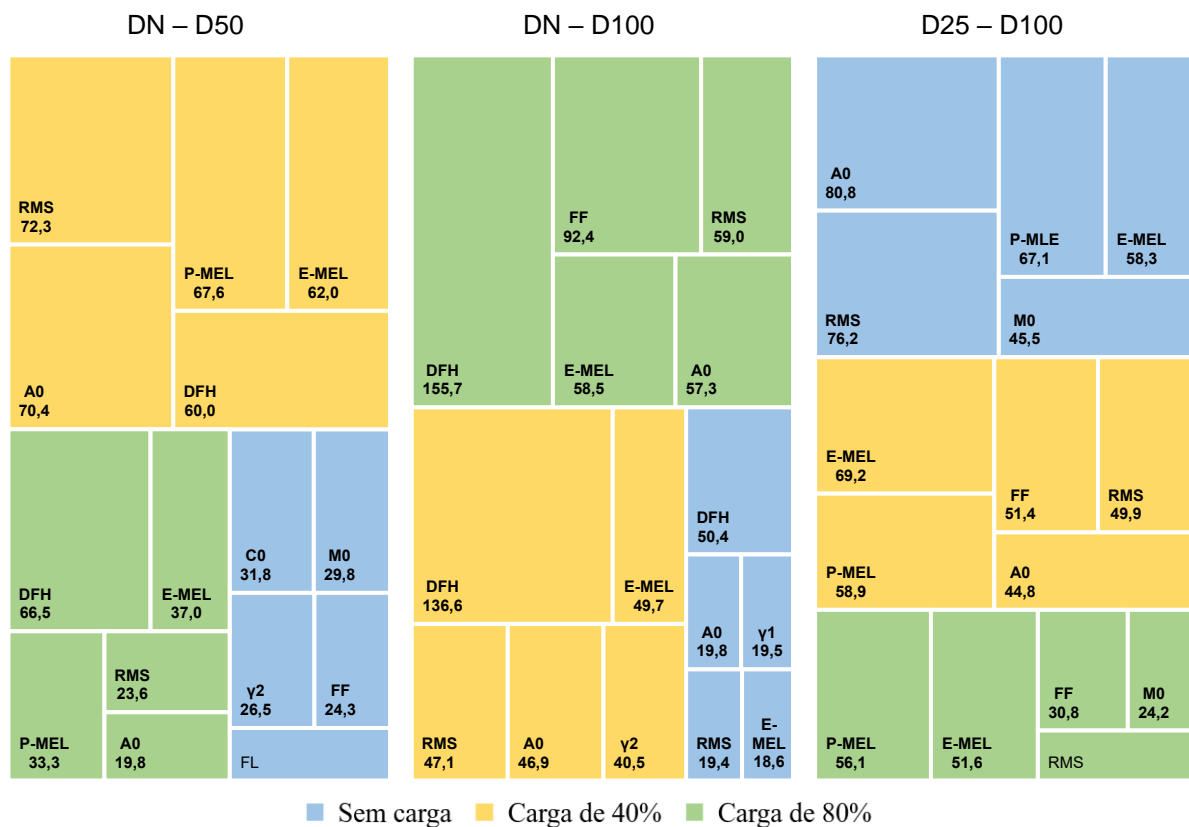


Figura 5.4 – Ranque dos 5 melhores preditores (parte 2).

Ainda, a respeito dos preditores selecionados, destaca-se a recorrência tanto das *features* não-lineares, do valor RMS, quanto dos parâmetros de Hjorth entre os 5 melhores indicadores. Evidentemente outros atributos como FF, FL, γ_1 e γ_2 também surgiram no ranqueamento. Todavia aquelas *features*, além de maior recorrência, na maioria dos casos vistos, ocuparam as primeiras posições. Portanto, a seção seguinte visa complementar a análise da separabilidade das características não-lineares e dos parâmetros de Hjorth.

5.2.1 Análise específica das *features* não-lineares e parâmetros de Hjorth

Devido às *features* não-lineares e os parâmetros de Hjorth terem se destacado em relação à capacidade de distinção de classes, elaborou-se gráficos como os da Figura 5.5 e da Figura 5.6, que permitem a comparação entre tais indicadores, com o objetivo de verificar a existência de algum padrão.

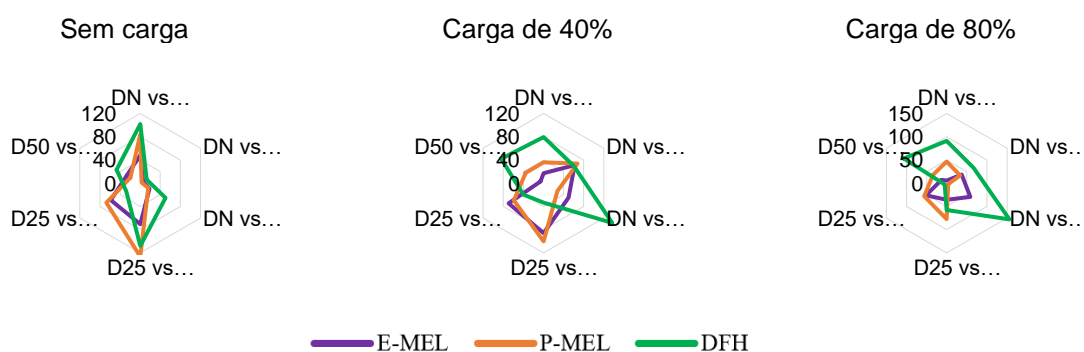


Figura 5.5 – Comparação entre *features* não-lineares.

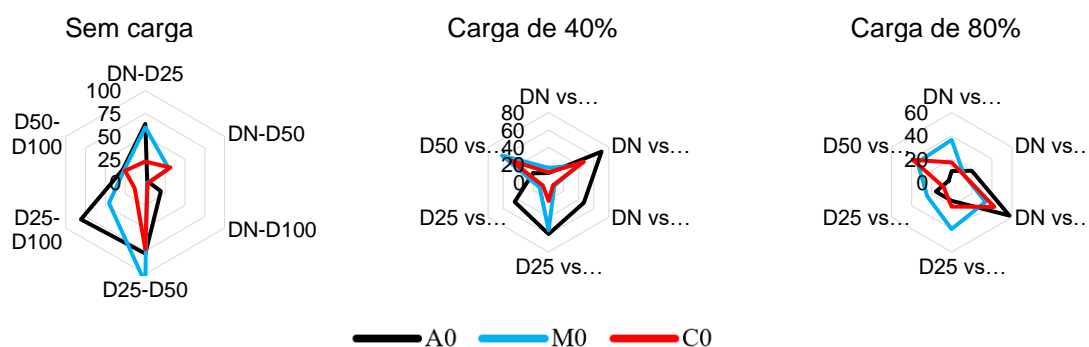


Figura 5.6 – Comparação entre os parâmetros de Hjorth.

Mediante as duas figuras acima, observa-se que, sem carga, tanto os preditores não-lineares quanto os parâmetros de Hjorth apresentaram maior grau de distinção entre os arranjos de classes DN-D25 e D25-D50, e em certo ponto também

para D25-D100. Isso mostra que na ausência de carga no misturador, a informação relativa à falha de 25% se sobressaiu bastante em relação às demais classes. Esse padrão indica que o baixo nível de tensão no rolamento devido à ausência de carga, faz com que as *features* analisadas reduzam sua capacidade de diferenciar entre as classes DN e as dos defeitos mais severos, além destes próprios também confundirem-se entre si. Esta conclusão fica mais evidente ao verificar que, com carga, foi possível verificar que houve certa diferenciação entre classes nas combinações que, sem carga, não havia ocorrido bem.

Ademais, é evidente a vantagem da combinação de preditores na tentativa de evidenciar *clusters*, facilitando posteriormente tarefas de classificação. Visto que, na maioria dos gráficos, ocorreu situações em que alguma *feature* foi incapaz de separar determinadas classes, enquanto outra foi mais sensível, ou vice-versa. Isso foi visto entre DFH com E-MEL ou P-MEL (Figura 5.5), e entre A_0 com M_0 ou C_0 (Figura 5.6). No mesmo sentido, Caesarendra (2015), verificou que A_0 foi melhor que M_0 e C_0 em seguir o aumento do nível de degradação em rolamento com rotação de 1 rpm, em ensaio *run-to-failure*.

Outrossim, por meio dos mesmos gráficos, é plausível inferir que o P-MEL (proposta deste trabalho) é, na maioria dos cenários vistos, mais sensível que o E-MEL. As distribuições das quatro classes dadas pelas três *features* não-lineares são ilustradas pela Figura 5.7, por meio da qual fica ainda mais evidente que o P-MEL, na maioria das vezes, foi capaz de separar melhor as classes em relação ao E-MEL. Entretanto, houve exceção ao distinguir entre as classes DN e D100 para carga de 40% e, sobretudo, de 80%, em que E-MEL sobrepujou P-MEL de modo mais significativo. Ainda assim, devido ao P-MEL ter sido superior nos outros exemplos, vale afirmar que isto levanta questões sobre a decisão de qual métrica utilizar para calcular o MEL.

Vale ressaltar que, a diferença nos resultados entre P-MEL e E-MEL se dá em virtude da discrepância de ambas no método empregado para encontrar os vizinhos mais próximos de cada estado da trajetória dinâmica. Visto que: enquanto E-MEL usa distância euclidiana, que pode ser vista como um coeficiente de dissimilaridade, cujo princípio se baseia no teorema de Pitágoras; a distância de Pearson, por sua vez, trata-se de um coeficiente de similaridade, o qual mensura o grau de correlação entre vetores que, neste caso, são os estados da trajetória.

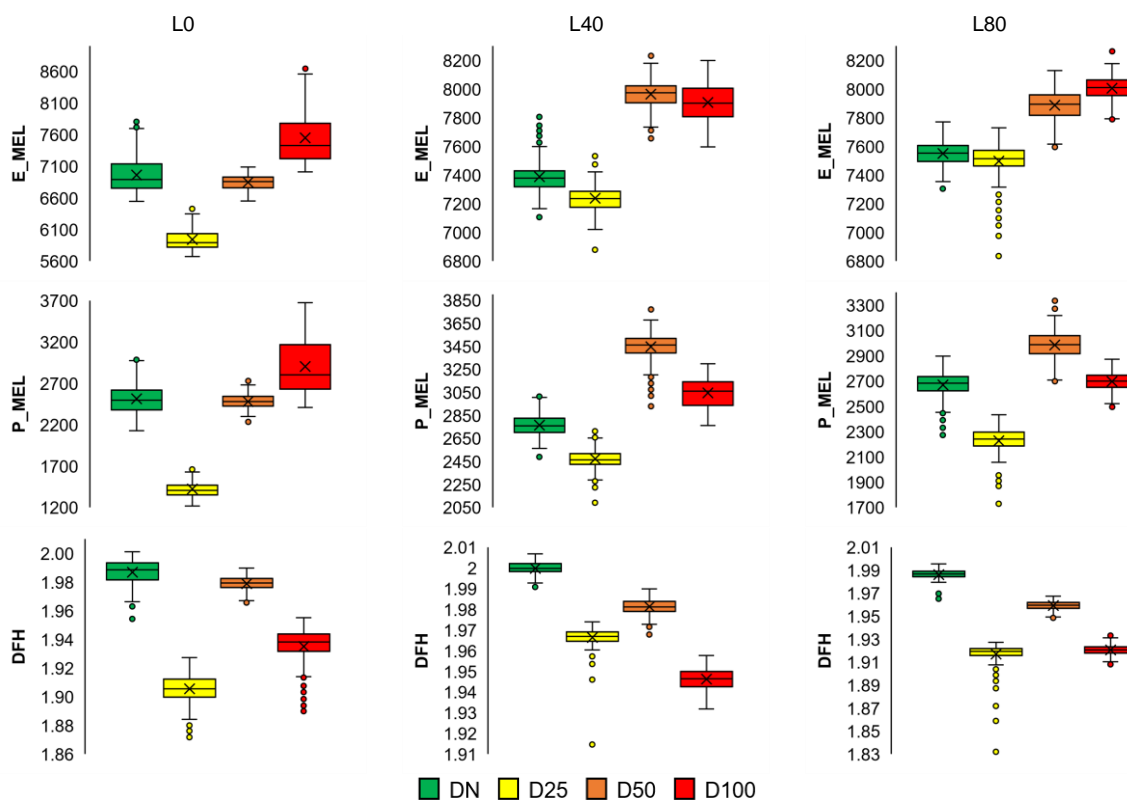


Figura 5.7 – Conjuntos das classes pelos indicadores não-lineares.

5.3 Treinamento do modelo de classificação SVM multiclasse

Baseado nos resultados acima, optou-se por treinar um algoritmo de classificação apenas com as *features*-não lineares, o valor RMS e os parâmetros de Hjorth. Diante disso, foi possível verificar a qualidade de tais preditores para efeitos de classificação de falha em rolamento de baixa rotação. Outros trabalhos na literatura também apresentaram ótimos resultados com *features* não-lineares e parâmetros de Hjorth. Sandoval et al. (2021) obtiveram melhores resultados com classificação de defeitos rolamentos em baixa rotação combinando indicadores de entropia e *features* no domínio do tempo (incluindo os parâmetros de Hjorth).

Caesarendra et al. (2013a) compararam as *features* não-lineares DFH, MEL, Entropia Aproximada (AppEn) e Dimensão de Correlação (DC) com RMS, Curtose, Distorção e Variância. Os autores também concluíram que as *features* não-lineares foram mais sensíveis, sobretudo o MEL, o qual evidenciou claramente a fase mais severa da falha. Wang et al. (2018) estudaram o uso das *features* não-lineares MEL, DC, AppEn e Sample Entropy (SE). Os autores reportaram obter altos percentuais de

acurácia dos modelos de classificação baseados em tais preditores, com destaque para AppEn.

Já em relação a este trabalho, a Figura 5.8, a Figura 5.9 e a Figura 5.10 apresentam os resultados de treinamento e de teste de modelos de classificação SVM multiclasse para os casos sem carga e com 40% e 80% de carga. Por meio das figuras mencionadas, conclui-se que a variação do nível de carga no misturador tem influência significativamente na assinatura vibracional adquirida. Pois observou-se que os hiperplanos dos três classificadores separam bem as classes do *dataset* de treinamento. Todavia, quando sujeitos a teste com dados provenientes de outro nível de carga, é notório que o classificador falha drasticamente. Dessa forma, conclui-se que construir um classificador com dados em determinado nível de carga pode acarretar em um modelo inapropriado para diagnosticar defeitos de sinais coletados em condição de carga diferente.

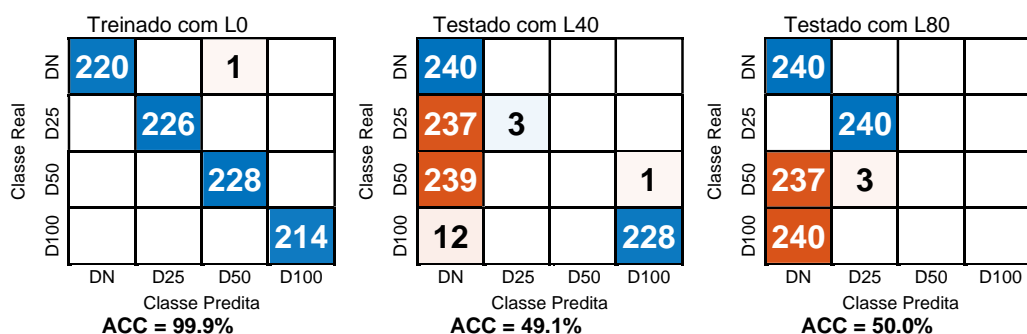


Figura 5.8 – Validação e teste do modelo sem carga.

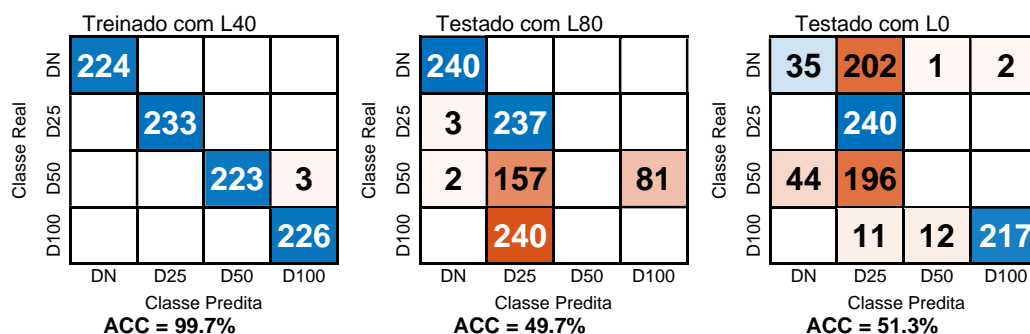


Figura 5.9 – Validação e teste do modelo com 40% de carga.

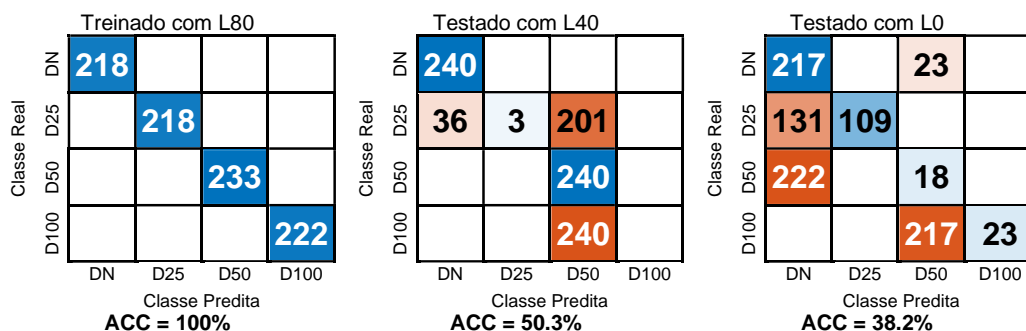


Figura 5.10 – Validação e teste do modelo com 80% de carga.

Diante disso, resolveu-se construir um classificador treinando-o com os dados dos três níveis de carga ao mesmo tempo, ou seja, concatenando os *datasets* L0, L40 e L80, formando uma única tabela de dados denominada por L_{CONC} . Em seguida, as matrizes-de-confusão de validação e de teste da Figura 5.11 mostram que, o classificador treinado com 70% de L_{CONC} não só apresentou elevada acurácia de validação, como também alta taxa de acurácia de teste com 30% de L_{CONC} . Assim, observou-se que treinar um classificador com dados em diferentes níveis de carga ao mesmo tempo forneceu um modelo de diagnóstico mais confiável do que os modelos discutidos previamente. E, sobretudo, os resultados reforçam a ideia de se construir classificadores usando dados em diferentes condições de operação, para que o modelo possa generalizar inobstante os cenários não-estacionários.

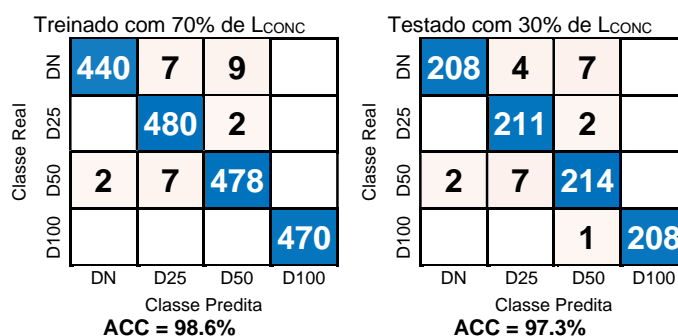


Figura 5.11 - Validação e teste do modelo construído com L_{CONC} .

5.4 Treinamento de modelo binário de classificação

Levando em consideração que a combinação feita com as features não-lineares, os parâmetros de Hjorth e o valor RMS apresentaram elevada separabilidade de classes; e que ao mesmo tempo permitiram a construção de um modelo de diagnóstico de diferentes níveis de defeito na pista externa do rolamento do

misturador; as mesmas *features* foram extraídas dos sinais de vibração de *datasets* binários, visando verificar se com elas seria possível construir modelos igualmente robustos em diferentes fontes de dados.

Primeiramente foi usado o conjunto de dados obtido a partir do sistema MEM, cujo detalhamento do procedimento experimental foi abordado no item 4.6 da seção Metodologia. A Figura 5.12 exibe a matriz-de-confusão resultante da validação de um classificador treinado com dados da bancada do sistema MEM. O resultado sugere que a combinação entre preditores não-lineares, os parâmetros de Hjorth e o valor RMS permitiu uma evidente separação da classe “normal” e “defeito”, fazendo com que o hiperplano de classificação pudesse classificar os dados com perfeição.

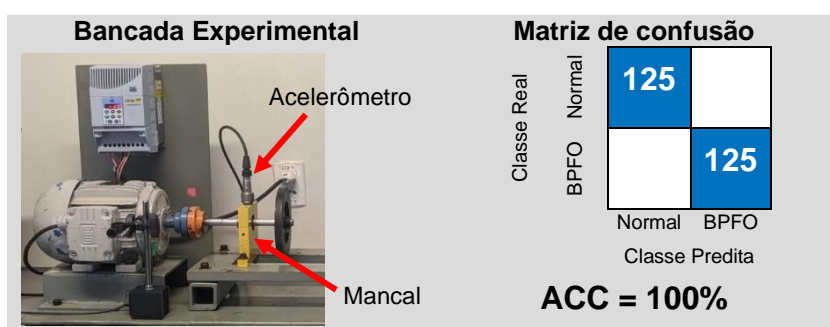


Figura 5.12 – Acurácia de teste do Msvm com dados da bancada MEM.

A Tabela 5.1 e a Tabela 5.2 apresentam os resultados de validação e teste de classificadores construídos a partir dos cinco conjuntos de dados adquiridos dos mancais de tambores de transportadores de correia em operação industrial, conforme detalhado no item 4.7 da seção Metodologia. A Tabela 5.1 especificamente exibe os resultados obtidos a partir dos sinais de vibração coletados na direção horizontal, enquanto a Tabela 5.2, por sua vez, na direção vertical. O principal objetivo de analisar ambas as direções consiste em constatar se porventura o diagnóstico de falha poderia ser realizado em qualquer uma dessas direções, ou se diferenças substanciais seriam identificadas.

Os resultados mostram que a combinação das *features* proposta neste estudo também permitiu construir modelos de classificação com taxas de acurácia relativamente altas, bem como elevados valores dos índices TVP (Taxa de Verdadeiro Positivo) e VPP (Valor Predito Positivo). Vale ressaltar que o TVP e o VPP foram calculados devido aos *datasets* de 2 a 5 estarem relativamente desbalanceados,

fazendo com que o uso somente da acurácia seja insuficiente e induza ao erro de se acreditar que um modelo altamente enviesado esteja otimizado.

Tabela 5.1 - Treinamento e teste de classificadores com sinais da direção horizontal.

Dataset	Treinamento			Teste				
	Matriz-de-Confusão		ACC (%)	Matriz-de-Confusão		ACC (%)	TVP (%)	VPP (%)
1	363	10	97,7	143	8	96,5	97,9	95,9
	7	355		4	191			
2	327	10	95,6	140	7	94,4	92	86,8
	9	83		4	46			
3	66	0	100	30	0	100	100	100
	0	177		0	81			
4	93	2	96,6	38	0	97,1	93,3	100
	3	51		2	28			
5	94	1	96,1	35	3	94,8	100	86,9
	4	30		0	20			

Tabela 5.2 - Treinamento e teste de classificadores com sinais da direção vertical.

Dataset	Treinamento			Teste				
	Matriz-de-Confusão		ACC (%)	Matriz-de-Confusão		ACC (%)	TVP (%)	VPP (%)
1	366	1	99,5	147	4	97,9	98,5	97,9
	3	356		3	192			
2	331	0	100	146	1	99,5	100	98
	0	94		0	50			
3	64	0	99,6	29	1	98,2	98,7	98,7
	1	173		1	79			
4	90	1	98,6	41	0	98,5	96,4	100
	1	54		1	27			
5	93	0	100	40	0	100	100	100
	0	34		0	17			

Os resultados vistos nesta seção exibem um alto desempenho dos modelos binários de classificação em diferenciar entre sinais de vibração de rolamento **sem** e **com** defeito na pista externa dos rolamentos monitorados. Sobretudo, é importante frisar que os atributos não-lineares e os parâmetros de Hjorth apresentaram-se em posições superiores no ranqueamento *features* de todos esses conjuntos de dados. E que apesar da redução da dimensionalidade, a composição dos preditores não-lineares com os parâmetros de Hjorth e valor RMS demonstrou alta performance em um modelo de classificação SVM.

6 CONCLUSÃO

O presente trabalho visou o desenvolvimento de metodologia para diagnóstico de falha em rolamento de baixa rotação mediante aprendizado supervisionado. Em que a etapa de seleção de preditores recebeu um tratamento estatístico, resultando em um ranqueamento das melhores *features*. Além disso, a influência do tamanho da janela do sinal de vibração na performance de tais indicadores foi analisada. No mesmo sentido, avaliou-se a viabilidade da alteração de uma etapa no algoritmo usado para o cálculo do MEL. Ao final, modelos de classificação SVM binário e multiclasse foram construídos com dados de bancadas experimentais e de transportadores de correia industriais.

Após o ranqueamento das *features*, as não-lineares combinadas com os parâmetros de Hjorth e o valor RMS apresentaram-se como ótimos preditores para diagnóstico de falha na pista externa de rolamento em baixa rotação. O que se confirmou com a aplicação do algoritmo SVM exibindo valores de acurácia elevados tanto com o modelo multiclasse quanto com os binários. Isto se deve à elevada separabilidade de classes por parte dos atributos selecionados. De sorte que, perante os dados em que os preditores escolhidos apresentaram baixa capacidade de distinção de classes, a performance do modelo treinado foi abaixo do esperado.

Além disso, é razoável considerar que, para diagnóstico de falha na pista externa de rolamentos, bastaria extrair somente as *features* não-lineares, o valor RMS e o parâmetro de Hjorth. Ou seja, dispensando os cálculos dos demais atributos do sinal no domínio do tempo e, por conseguinte, a etapa de seleção pelo ranqueamento de *features*, simplificando a metodologia. Outrossim, a redução de dimensionalidade com PCA simplificou o espaço de preditores para apenas 3 dimensões, além de ter contornado redundâncias de indicadores altamente correlacionados e minimizando chances de *overfitting*.

Ademais, a análise do tamanho ideal da janela do sinal a ser usado é um conteúdo escasso na literatura de rolamento em baixa rotação. Para tanto, concluiu-se que, em nove cenários diferentes, todas as características dos sinais exibiram aumento da capacidade de separação de classes quanto maior o tamanho da janela. Porém, para a maioria dos casos, a diferença não foi tão significativa ao comparar as janelas com 3, 2 e 1 revolução(ões). Por outro lado, o desempenho das *features* caiu

progressivamente com a redução da janela a valores abaixo de 1 revolução. Diante disso, concluiu-se que o janelamento mais adequado foi de 1 revolução.

Em relação a variação do nível de carga no misturador ficou evidente que tal variação influencia significativamente nos valores das *features* extraídas dos sinais de vibração, de modo que quando se treinou um modelo em determinado nível de carga, o classificador falhava em diagnosticar os sinais oriundos de outros níveis de carga. No tocante aos dados dos tambores de transportadores de correia, não se conhecia o nível de carga ou se houve variações substanciais. Contudo, verificou-se que com os dados do misturador foi necessário treinar um classificador com dados nos três níveis de carga ao mesmo tempo, para que fosse possível construir um modelo que generalizasse o diagnóstico de falha perante à tal condição não-estacionária. Portanto, concluiu-se que, se porventura houve grandes variações de carga durante a operação do transportador, o classificador foi treinado levando em consideração essas alterações, fornecendo altas taxas de acurácias apesar das possíveis mudanças nas condições de carga de operação.

Por fim, em relação à verificação dos efeitos da alteração no algoritmo tradicional para o cálculo do MEL, o P-MEL (proposta deste trabalho) mostrou-se superior ao E-MEL na maioria dos cenários analisados. No entanto, não é possível afirmar que em todas as situações o P-MEL apresentou performances superiores ao E-MEL. Pois, para o cálculo do MEL, a determinação dos seus parâmetros pode afetar severamente o resultado final, além da própria natureza caótica dos sinais de vibração. Todavia, é razoável concluir que as discrepâncias entre P-MEL e E-MEL levanta questões sobre qual seria a métrica mais adequada para o cálculo do MEL. Vale ressaltar que o cálculo do MEL e da DFH exigem um custo computacional maior em relação às *features* tradicionais, o que torna o tempo de treinamento maior.

Finalmente, conclui-se que o objetivo principal do trabalho foi alcançado com a propositura da metodologia completa para diagnóstico de falha em rolamento de baixa velocidade mediante algoritmo de aprendizado de máquina supervisionado. Para trabalhos futuros, fica a sugestão de se explorar outras *features* não-lineares dentre várias que existem na literatura, sobretudo no campo de análise de sinais de eletroencefalograma. Além disso, outras medidas de similaridade ou dissimilaridade poderiam ser testadas no algoritmo para cálculo do MEL, na etapa de encontrar os vizinhos mais próximos de cada ponto da trajetória dinâmica do sinal de vibração.

REFERÊNCIAS

- ABDI, Hervé; WILLIAMS, Lynne J. Principal component analysis. **Wiley interdisciplinary reviews: computational statistics**, v. 2, n. 4, p. 433-459, 2010.
- AGARWAL, Dishaa. Introduction to SVM (Support Vector Machine): Along with Python Code. *In: Analytics Vidhya*. [S.l.], 7 nov. 2022. Disponível em: <https://www.analyticsvidhya.com/blog/2021/04/insight-into-svm-support-vector-machine-along-with-code/>. Acesso em: 26 dez. 2022.
- ALTHUBAITI, Adnan; ELASHA, Faris; TEIXEIRA, Joao Amaral. Fault diagnosis and health management of bearings in rotating equipment based on vibration analysis—a review. **Journal of Vibroengineering**, v. 24, n. 1, p. 46-74, 2022.
- ANTONIOLLI, E. B. **Estudo comparativo de técnicas de medição e análise de vibrações para a manutenção preditiva em mancais de rolamentos**. 1999. Tese de Doutorado. Dissertação (Mestrado em Engenharia Mecânica) – Universidade Federal de Santa Catarina, Florianópolis, 1999. Disponível em: <https://scholar.googleusercontent.com/scholar>.
- ASRI, G. S. et al. The bread production process using application of the hotelling t2 control chart. *In: Journal of Physics: Conference Series*. IOP Publishing, 2019. p. 022047.
- BECHHOEFER, Eric; SCHLANBUSCH, Rune; WAAG, Tor Inge. Techniques for large, slow bearing fault detection. **International Journal of Prognostics and Health Management**, v. 7, n. 1, 2016.
- BENTO, F. D. S. **O uso da manutenção preditiva como subsídio à manutenção preventiva**. Orientador: Profª Drª Helen Silva Gonçalves. 2012. 56 f. TCC (Graduação) - Curso de Administração, CCSA, UFPB, João Pessoa, 2012.
- BISHOP, Christopher. **Pattern recognition and machine learning**. New York: Springer, 2006.
- BRO, Rasmus; SMILDE, Age K. Principal component analysis. **Analytical methods**, v. 6, n. 9, p. 2812-2831, 2014.
- CADAVID, Ana Christina; LAWRENCE, John K.; RUZMAIKIN, Alexander. Principal components and independent component analysis of solar and space data. **Solar Image Analysis and Visualization**, p. 37-51, 2009.
- CAESARENDRA, Wahyu et al. An application of nonlinear feature extraction – A case study for low-speed slewing bearing condition monitoring and prognosis. *In: 2013 IEEE/ASME International Conference on Advanced Intelligent Mechatronics*. IEEE, 2013a. p. 1713-1718.
- CAESARENDRA, Wahyu et al. Application of the largest Lyapunov exponent algorithm for feature extraction in low-speed slewing bearing condition monitoring. **Mechanical Systems and Signal Processing**, v. 50, p. 116-138, 2015.

CAESARENDRA, Wahyu et al. Circular domain features-based condition monitoring for low-speed slewing bearing. **Mechanical Systems and Signal Processing**, v. 45, n. 1, p. 114-138, 2014.

CAESARENDRA, Wahyu et al. Condition monitoring of naturally damaged slow speed slewing bearing based on ensemble empirical mode decomposition. **Journal of Mechanical Science and Technology**, v. 27, n. 8, p. 2253-2262, 2013b.

CAESARENDRA, Wahyu et al. Parsimonious network based on a fuzzy inference system (PANFIS) for time series feature prediction of low-speed slew bearing prognosis. **Applied Sciences**, v. 8, n. 12, p. 2656, 2018.

CAESARENDRA, Wahyu, **Vibration and acoustic emission-based condition monitoring and prognostic methods for very low speed slew bearing**, Doctor of Philosophy thesis, School of Mechanical, Materials and Mechatronic Engineering, University of Wollongong, 2015. <https://ro.uow.edu.au/theses/4479>.

CAESARENDRA, Wahyu; TJAHJOWIDODO, Tegoeh. A review of feature extraction methods in vibration-based condition monitoring and its application for degradation trend estimation of low-speed slew bearing. **Machines**, v. 5, n. 4, p. 21, 2017.

COOMBS, William T.; ALGINA, James; OLTMAN, Debra Olson. Univariate and multivariate omnibus hypothesis tests selected to control type I error rates when population variances are not necessarily equal. **Review of Educational Research**, v. 66, n. 2, p. 137-179, 1996.

FAGERLAND, Morten W.; SANDVIK, Leiv. Performance of five two-sample location tests for skewed distributions with unequal variances. **Contemporary clinical trials**, v. 30, n. 5, p. 490-496, 2009.

FENG, Yang et al. Reliability-based residual life prediction of large-size low-speed slewing bearings. **Mechanism and Machine Theory**, v. 81, p. 94-106, 2014.

FORTI, M. **Técnicas de machine learning aplicadas na recuperação de crédito do mercado brasileiro**. Orientador: Prof. Dr. João Luiz Chela. 2018. 74 f. Dissertação (Mestrado) - Curso de Economia, FGV, São Paulo, 2018.

GUO, Li-Hong et al. Fault Identification of Low-Speed Hub Bearing of Crane Based on MBMD and BP Neural Network. **Shock and Vibration**, v. 2022, n. 1, p. 5005263, 2022.

HAN, Tian et al. Fault feature extraction of low-speed roller bearing based on Teager energy operator and CEEMD. **Measurement**, v. 138, p. 400-408, 2019.

HAUKE, Jan; KOSSOWSKI, Tomasz. Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data. **Quaestiones geographicae**, v. 30, n. 2, p. 87, 2011.

HEMMER M.; WAAG T.I.; ROBBERSMYR K.G. A Review of Methods for Condition Monitoring of Large, Slow-rotating Bearings. Presented at: **30th Conference on**

Condition Monitoring and Diagnostic Engineering Management. University of Central Lancashire, Preston, UK, 2018.

HEMMER, Martin; WAAG, Tor I. A comparison of acoustic emission and vibration measurements for condition monitoring of an offshore drilling machine. In: **Annual Conference of the PHM Society.** 2017.

HEMMER, Martin et al. Fault classification of axial and radial roller bearings using transfer learning through a pretrained convolutional neural network. **Designs**, v. 2, n. 4, p. 56, 2018.

HEMMER, Martin et al. Health indicator for low-speed axial bearings using variational autoencoders. **IEEE Access**, v. 8, p. 35842-35852, 2020.

HEMMER, Martin. **Condition Monitoring Methods for Large, Low-speed Bearings.** Doctoral Dissertation for the Degree Philosophiae Doctor (PhD) at the Faculty of Engineering and Science, Specialisation in Mechatronics. University of Agder. 2020.

HJORTH, Bo. EEG analysis based on time domain properties. **Electroencephalography and clinical neurophysiology**, v. 29, n. 3, p. 306-310, 1970.

HOLANDA, S. M. S. **Aplicação da manutenção preditiva por análise de vibrações em equipamentos de trens urbanos com plano de manutenção proposto.** Orientador: Prof. Dr. Tiago Leite Rolim. 2016. 96 f. Dissertação (Mestrado) - Curso de Engenharia Mecânica, UFPE, Recife, 2016.

JIN, Xin et al. Failure prediction, monitoring and diagnosis methods for slewing bearings of large-scale wind turbine: A review. **Measurement**, v. 172, p. 108855, 2021.

JOLLIFFE, Ian T.; CADIMA, Jorge. Principal component analysis: a review and recent developments. **Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences**, v. 374, n. 2065, p. 20150202, 2016.

KANG, Myeongsu et al. Reliable fault diagnosis for low-speed bearings using individually trained support vector machines with kernel discriminative feature analysis. **IEEE Transactions on Power Electronics**, v. 30, n. 5, p. 2786-2797, 2014.

KANG, Myeongsu; KIM, Jaeyoung; KIM, Jong-Myon. Reliable fault diagnosis for incipient low-speed bearings using fault feature analysis based on a binary bat algorithm. **Information Sciences**, v. 294, p. 423-438, 2015.

KLAUSEN, A. et al. Accelerated bearing life-Time test rig development for low-speed data acquisition. **Modeling Identification and Control**, v. 38, n. 3, p. 143-156, 2017.

KLAUSEN, Andreas; ROBBERSMYR, Kjell G. Cross-correlation of whitened vibration signals for low-speed bearing diagnostics. **Mechanical Systems and Signal Processing**, v. 118, p. 226-244, 2019.

LADEIRA, R. W. S. *et al.* Automação de diagnóstico para manutenção preditiva baseada em análise de fluidos de equipamentos com machine learning. *In: SEMINÁRIO DE AUTOMAÇÃO E TI, 21^{o.}, 2017, São Paulo. Anais [...].* São Paulo: ABM Week, 2017. v. 21, p. 268-279, ISSN 2594-5335.

LI, Tengyu *et al.* Application of Adaptive MOMEDA with Iterative Autocorrelation to Enhance Weak Features of Hoist Bearings. **Entropy**, v. 23, n. 7, p. 789, 2021.

LI, Yuanyuan *et al.* Life state recognition of slewing bearing with heavy load and low speed based on point density function with fuzzy C means. **Transactions of the Institute of Measurement and Control**, v. 41, n. 5, p. 1256-1265, 2019.

LIMA, I. A. M. D. **Proposição de uma bancada didática para análise de vibração gerada por desalinhamento e desbalanceamento.** Orientador: Prof. Dr. Maria Alzira de Araújo Nunes. 2013. 103 f. TCC (Graduação) - Curso de Engenharia Automotiva, UnB, Brasília, 2013.

LIMA, W. D. C.; LIMA, C. R. C.; SALLES, A. A. Manutenção Preditiva, o Caminho para a Excelência: Uma Vantagem Competitiva. *In: SIMPÓSIO DE ENGENHARIA DE PRODUÇÃO, XIII., 2006, Bauru. Anais [...].* Franca: UNESP, 2006.

LIN, Wei Ming *et al.* Large pendulum bolt failure analysis and countermeasures. **Applied Mechanics and Materials**, v. 256, p. 600-603, 2013.

LIU, Chenxi; WANG, Fengtao. A review of current condition monitoring and fault diagnosis methods for low-speed and heavy-load slewing bearings. *In: 2017 9th International Conference on Modelling, Identification and Control (ICMIC).* IEEE, 2017. p. 104-109.

LIU, Mingjun *et al.* Remaining Useful Life Estimation of Fan Slewing Bearings in Nonlinear Wiener Process with Random Covariate Effect. **Shock and Vibration**, v. 2022, n. 1, p. 5441760, 2022.

LIU, Ruonan *et al.* Artificial intelligence for fault diagnosis of rotating machinery: A review. **Mechanical Systems and Signal Processing**, v. 108, p. 33-47, 2018.

LIU, Zepeng; ZHANG, Long. A review of failure modes, condition monitoring and fault diagnosis methods for large-scale wind turbine bearings. **Measurement**, v. 149, p. 107002, 2020.

LIU, Zepeng; ZHANG, Long; CARRASCO, Joaquin. Vibration analysis for large-scale wind turbine blade bearing fault detection with an empirical wavelet thresholding method. **Renewable Energy**, v. 146, p. 99-110, 2020.

LU, Chao *et al.* Degradation trend estimation of slewing bearing based on LSSVM model. **Mechanical Systems and Signal Processing**, v. 76, p. 353-366, 2016.

MATHWORKS, Inc. MATLAB[®] (Version R2023b). Help Center. Documentation, **[Support Vector Machine (SVM): Learn optimal hyperplanes as decision**

boundaries], Natick, Massachusetts, United States, 2024c. Disponível em: <https://www.mathworks.com/discovery/support-vector-machine.html>.

MATHWORKS, Inc. MATLAB® (Version R2023b). Help Center. Documentation, **[Signal Features]**, Natick, Massachusetts, United States, 2024a. Disponível em: <https://www.mathworks.com/help/predmaint/ug/signal-features.html>.

MATHWORKS, Inc. MATLAB® (Version R2023b). Help Center. Documentation, **[max]**, Natick, Massachusetts, United States, 2024b. Disponível em: <https://www.mathworks.com/help/matlab/ref/max.html>.

MATHWORKS, Inc. MATLAB® (Version R2023b). Help Center. Documentation, **[lyapunovExponent]**, Natick, Massachusetts, United States, 2024d. Disponível em: <https://www.mathworks.com/help/predmaint/ref/lyapunovexponent.html>.

MEHDIZADEH, Sina. A robust method to estimate the largest lyapunov exponent of noisy signals: a revision to the rosenstein's algorithm. **Journal of biomechanics**, v. 85, p. 84-91, 2019.

MEHMOOD, Tahir. Hotelling T2 based variable selection in partial least squares regression. **Chemometrics and Intelligent Laboratory Systems**, v. 154, p. 23-28, 2016.

MERGL, Václav; KAŠPÁREK, Jaroslav. Verifying the Lifting and Slewing Dynamics of a Harvester Crane with Possible Levelling When Operating on Sloping Grounds. **Forests**, v. 13, n. 2, p. 357, 2022.

MISHRA, Chandrabhanu; SAMANTARAY, Arun Kumar; CHAKRABORTY, Goutam. Rolling element bearing fault diagnosis under slow speed operation using wavelet denoising. **Measurement**, v. 103, p. 77-86, 2017.

MOODIE, Craig AS. **An investigation into the condition monitoring of large slow speed slew bearing**. Doctor of Philosophy thesis, School of Mechanical, Materials and Mechatronic Engineering, University of Wollongong. 2009. <http://ro.uow.edu.au/theses/3035>.

MOUSTAFA, Wael et al. Low speed bearings fault detection and size estimation using instantaneous angular speed. **Journal of Vibration and Control**, v. 22, n. 15, p. 3413-3425, 2016.

MUKAKA, MJMMJ. Statistics corner: a guide to appropriate use of correlation in medical research. **Malawi Med J**, v. 24, n. 3, p. 69-71, 2012.

NGUYEN, Hung Ngoc; KIM, Jaeyoung; KIM, Jong-Myon. Optimal sub-band analysis based on the envelope power Spectrum for effective fault detection in bearing under variable, low speeds. **Sensors**, v. 18, n. 5, p. 1389, 2018.

NIKULA, Riku-Pekka et al. Automation of low-speed bearing fault diagnosis based on autocorrelation of time domain features. **Mechanical Systems and Signal Processing**, v. 138, p. 106572, 2020.

OMPUSUNGGU, Agusmian Partogi; DEVOS, Steven; PETRE, Frederik. Stochastic-resonance based fault diagnosis for rolling element bearings subjected to low rotational speed. **International Journal of Prognostics and Health Management**, v. 4, n. 2, 2013.

PÄIVINEN, Niina et al. Epileptic seizure detection: A nonlinear viewpoint. **Computer methods and programs in biomedicine**, v. 79, n. 2, p. 151-159, 2005.

PAN, G. M.; ZHOU, Wang. Central limit theorem for hotelling's T 2 statistic under large dimension. **The Annals of Applied Probability**, p. 1860-1910, 2011.

PAN, Yubin et al. Fault recognition of large-size low-speed slewing bearing based on improved deep belief network. **Journal of Vibration and Control**, v. 29, n. 11-12, p. 2829-2841, 2023.

PAN, Yubin et al. Improved complete ensemble robust local mean decomposition with adaptive noise for slewing bearings performance degradation assessment. **IEEE Access**, v. 10, p. 78677-78690, 2022.

PAN, Yubin et al. Performance degradation assessment of slewing bearings based on deep auto-encoder and optimized particle filtering. **Journal of Vibration and Control**, v. 25, n. 17, p. 2380-2394, 2019.

PEDREGOSA et al. Scikit-learn: Machine Learning in Python. **Support Vector Machines**. JMLR 12, pp. 2825-2830, 2011. Disponível em: <https://scikit-learn.org/stable/modules/svm.html>.

POTOČNIK, Rok; FLAŠKER, Jože; GLODEŽ, Srečko. Numerical analysis of 3D subsurface crack propagation in large slewing bearing. **Proceedings of the ECF18**, 2013.

PRABHAKAR, S. N.; SAHOO, B.; CHANDRA, R. Condition Monitoring of Slew Bearing Stacker and Reclaimer using SPM HD Technology. **International Journal of Innovative Research in Science, Engineering and Technology**, v. 11, n. 4, 2022.

PRAKASH, G.; NARASIMHAN, S.; PANDEY, M. D. Condition based maintenance of low-speed rolling element bearings using hidden markov model. **International Journal of Prognostics and Health Management**, v. 8, n. 1, 2017.

QIU, Ming et al. A finite-element analysis of the connecting bolts of slewing bearings based on the orthogonal method. **Journal of mechanical science and technology**, v. 26, p. 883-887, 2012.

ROSENSTEIN, Michael T.; COLLINS, James J.; DE LUCA, Carlo J. A practical method for calculating largest Lyapunov exponents from small data sets. **Physica D: Nonlinear Phenomena**, v. 65, n. 1-2, p. 117-134, 1993.

ROTHSCHEDL, Christopher Josef. **Condition monitoring of large-scale slew bearings in bucket-wheel boom-type reclaimers**. 2016. Tese de Doutorado. University of Leoben.

RUELLE, David. Small random perturbations of dynamical systems and the definition of attractors. **Communications in Mathematical Physics**, v. 82, n. 1, p. 137-151, 1981.

RUXTON, Graeme D. The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney U test. **Behavioral Ecology**, v. 17, n. 4, p. 688-690, 2006.

SANDOVAL, Diego Aníbal. **Diagnosis of low-speed bearings via vibration-based entropy indicators and acoustic emissions**. Tesi doctoral, UPC, Departament d'Enginyeria Civil i Ambiental, 2021. Disponível em: <http://hdl.handle.net/2117/363068>.

SANDOVAL, Diego et al. Low-speed bearing fault diagnosis based on permutation and spectral entropy measures. **Applied Sciences**, v. 10, n. 13, p. 4666, 2020.

SANO, Masaki; SAWADA, Yasuji. Measurement of the Lyapunov spectrum from a chaotic time series. **Physical review letters**, v. 55, n. 10, p. 1082, 1985.

SAUFI, Syahril Ramadhan et al. Low-speed bearing fault diagnosis based on ArSSAE model using acoustic emission and vibration signals. **IEEE Access**, v. 7, p. 46885-46897, 2019.

SHLENS, Jonathon. A Tutorial on Principal Component Analysis. **arXiv e-prints**, p. arXiv: 1404.1100, 2014. Disponível em: <https://arxiv.org/pdf/1404.1100.pdf>.

SHUUJI, Miyazaki et al. Low-speed bearing fault diagnosis based on improved statistical filtering and convolutional neural network. **Measurement Science and Technology**, v. 32, n. 11, p. 115009, 2021.

SILVA, D. G. M. D. **Aspectos positivos do uso da Inteligência Artificial/Machine Learning na gestão e planejamento da manutenção aeronáutica**. Orientador: Luis Henrique Santos. 2020. TCC (Especialização) - Curso de Engenharia Aeronáutica, ITA, São José dos Campos, 2021.

SMITH, Lindsay I. **A tutorial on principal components analysis**. Department of Computer Science, University of Otago. New Zealand, 2002. Disponível em: <https://ourarchive.otago.ac.nz/bitstream/handle/10523/7534/OUCS-2002-12.pdf>.

SMITS, Fenne Margreeth et al. Electroencephalographic fractal dimension in healthy ageing and Alzheimer's disease. **PloS one**, v. 11, n. 2, p. e0149587, 2016.

SOKOLSKI, Piotr; SMOLNICKI, Tadeusz. A method for monitoring the technical condition of large-scale bearing nodes in the bodies of machines operating for extended periods of time. **Energies**, v. 14, n. 20, p. 6637, 2021.

SONG, Liuyang; WANG, Huaqing; CHEN, Peng. Vibration-based intelligent fault diagnosis for roller bearings in low-speed rotating machinery. **IEEE Transactions on Instrumentation and Measurement**, v. 67, n. 8, p. 1887-1899, 2018.

SOUALHI, Abdenour; MEDJAHHER, Kamal; ZERHOUNI, Nouredine. Bearing health monitoring based on Hilbert–Huang transform, support vector machine, and regression. **IEEE Transactions on instrumentation and measurement**, v. 64, n. 1, p. 52-62, 2014.

SPM - SHOCK PULSE METHOD. **SPM HD**. Estocolmo: SPM®, 2024. Disponível em: <https://www.spminstrument.com/measuring-techniques/shock-pulse-monitoring/spm-hd/>. Acesso em: 14 fev. 2024.

STOOP, Ruedi; PARISI, Jürgen. Calculation of Lyapunov exponents avoiding spurious elements. **Physica D: Nonlinear Phenomena**, v. 50, n. 1, p. 89-94, 1991.

TOBON-MEJIA, Diego A.; MEDJAHHER, Kamal; ZERHOUNI, Nouredine. The ISO 13381-1 standard's failure prognostics process through an example. In: **2010 Prognostics and System Health Management Conference**. IEEE, 2010. p. 1-12.

TOME, Emanuel Sousa; PIMENTEL, Mario; FIGUEIRAS, Joaquim. Damage detection under environmental and operational effects using cointegration analysis—application to experimental data from a cable-stayed bridge. **Mechanical Systems and Signal Processing**, v. 135, p. 106386, 2020.

VASCONCELOS, B. F. B. D. **Poder preditivo de métodos de Machine Learning com processos de seleção de variáveis: uma aplicação às projeções de produto de países**. Orientador: Victor Gomes. 2017. 90 f. Tese (Doutorado) - Curso de Economia, UNB, Brasília, 2017.

VULPIANI, Angelo; CECCONI, Fabio; CENCINI, Massimo. **Chaos: from simple models to complex systems**. World Scientific, 2009.

WANG, Fengtao et al. Combined failure diagnosis of slewing bearings based on MCKD-CEEMD-ApEn. **Shock and Vibration**, v. 2018, 2018.

WANG, Fengtao et al. Condition monitoring and fault diagnosis methods for low-speed and heavy-load slewing bearings: A literature review. **Journal of Vibroengineering**, v. 19, n. 5, p. 3429-3444, 2017.

WANG, G. X. et al. Shock response properties of crane slewing bearing under the condition of revolve braking. **Noise and Vibration Control**, v. 4, p. 30-34, 2013.

WANG, Hua; TANG, Mingmin; HUANG, Xiaodiao. Smart health evaluation of slewing bearing based on multiple-characteristic parameters. **Journal of mechanical science and technology**, v. 28, p. 2089-2097, 2014.

WANG, Y.-F.; KOOTSOOKOS, Peter J. Modeling of low shaft speed bearing faults for condition monitoring. **Mechanical systems and signal processing**, v. 12, n. 3, p. 415-426, 1998.

WEI, Yu et al. A review of early fault diagnosis approaches and their applications in rotating machinery. **Entropy**, v. 21, n. 4, p. 409, 2019.

WEST, Robert M. Best practice in statistics: Use the Welch t-test when testing the difference between two groups. **Annals of Clinical Biochemistry**, v. 58, n. 4, p. 267-269, 2021.

WIDODO, Achmad et al. Fault diagnosis of low-speed bearing based on relevance vector machine and support vector machine. **Expert systems with applications**, v. 36, n. 3, p. 7252-7261, 2009.

WOLF, Alan et al. Determining Lyapunov exponents from a time series. **Physica D: nonlinear phenomena**, v. 16, n. 3, p. 285-317, 1985.

XIAO, Chengwei et al. Using Spearman's correlation coefficients for exploratory data analysis on big dataset. **Concurrency and Computation: Practice and Experience**, v. 28, n. 14, p. 3866-3878, 2016.

XIE, Yaqin et al. An improved K-nearest-neighbor indoor localization method based on spearman distance. **IEEE signal processing letters**, v. 23, n. 3, p. 351-355, 2016.

YE, Y. X. **The research of large-diameter antifriction bearing's trouble mechanism and diagnosing technology of trouble**. 2005. Tese de Doutorado. MS dissertation, Department of Mechanical Engineering, Wuhan University of Technology, Wuhan, China.

ZAIDI, Fatima Sehar et al. Monitoring autocorrelated compositional data vectors using an enhanced residuals Hotelling T2 control chart. **Computers & Industrial Engineering**, v. 181, p. 109280, 2023.

ZHAO, Y. et al. Application and research in fault diagnosis of slewing bearing based on bispectrum analysis. **Machinery Design and Manufacture**, v. 4, p. 253-257, 2016.

ZIMMERMAN, Donald W.; ZUMBO, Bruno D. Rank transformations and the power of the student t test and Welch t'test for non-normal populations with unequal variances. **Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale**, v. 47, n. 3, p. 523, 1993.

ZONG, Haiyong; WANG, Hua; HONG, Rongjing. Dynamic research on single-row slewing bearing with local spalling of inner ring. **Advances in Mechanical Engineering**, v. 8, n. 2, p. 1687814016632488, 2016.

ŽVOKELJ, Matej; ZUPAN, Samo; PREBIL, Ivan. Multivariate and multiscale monitoring of large-size low-speed bearings using ensemble empirical mode decomposition method combined with principal component analysis. **Mechanical Systems and Signal Processing**, v. 24, n. 4, p. 1049-1067, 2010.