



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

JUSLEY DA SILVA SOUZA

**ANÁLISE DE ATRIBUTOS DE CLASSIFICAÇÃO PARA O DIAGNÓSTICO DE FALHAS EM
ROLAMENTOS BASEADO EM SVM**

JUSLEY DA SILVA SOUZA

**ANÁLISE DE ATRIBUTOS DE CLASSIFICAÇÃO PARA O DIAGNÓSTICO DE
FALHAS EM ROLAMENTOS BASEADO EM SVM**

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: Desenvolvimento Energético.

Orientador: Alexandre Luiz Amarante Mesquita
Coorientador: Rafael Suzuki Bayma

Tucuruí – PA

2019

JUSLEY DA SILVA SOUZA

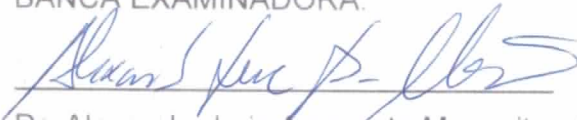
**ANÁLISE DE ATRIBUTOS DE CLASSIFICAÇÃO PARA O DIAGNÓSTICO DE
FALHAS EM ROLAMENTOS BASEADO EM SVM**

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: Desenvolvimento Energético.

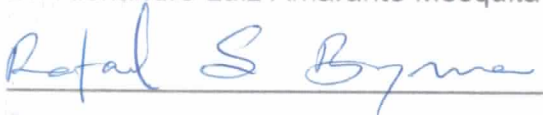
Orientador: Alexandre Luiz Amarante Mesquita
Coorientador: Rafael Suzuki Bayma

Aprovada em 06 de Agosto de 2019.


BANCA EXAMINADORA:



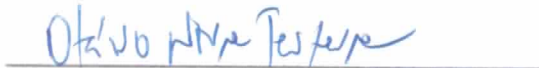
Dr. Alexandre Luiz Amarante Mesquita - Universidade Federal do Pará - Orientador



Dr. Rafael Suzuki Bayma - Universidade Federal do Pará - Coorientador



Dr. Leopoldo Pacheco Bastos - Universidade Federal do Pará – Examinador Interno



Dr. Otávio Noura Teixeira - Universidade Federal do Pará/PPCA – Examinador Externo



Dr. Paulo Cezar Monteiro Lamim Filho - Universidade Federal de São João del-Rei/PPGEL – Examinador Externo

“A Ele seja a glória para sempre!”

Romanos 11.36

AGRADECIMENTOS

Primeiramente ao Deus da minha vida, por ter me sustentado e me guiado até aqui, por ter me protegido em todos os dias da minha vida, por ter iluminado meus pensamentos e me dado graça para conquistar essa vitória. Deus é bom o tempo todo. Sem Ele, nada disso seria possível. À Ele toda honra e glória.

Minha gratidão a todos os meus familiares, e em especial meus pais, Gerson Barbosa e Orlanda Telles, pelo amor, educação, suporte espiritual e financeiro e incentivo em todos os momentos. Minha mamãe obrigada por cada palavra de incentivo, força, fé e amor ditas a mim, isso me encorajou a continuar em momentos que pensei em desistir. Você foi essencial nessa conquista. És minha maior inspiração. Essa conquista é nossa! Te amo infinitamente.

Ao meu irmão Jefferson, como você é importante pra mim, me ensina tanto com suas conversas. Obrigada por sempre acreditar em mim e me fazer tão bem. Seu amor me deu força para chegar até aqui. Amo você.

Agradeço ao meu namorado e companheiro Rogério Mattos, que sempre esteve ao meu lado. Você acreditou em mim, desde o dia da inscrição do mestrado. Sou grata a você pelas palavras de apoio, pela paciência em me ensinar, por cada sorriso que você tirou de mim, em situações que eu achava impossível, pela compreensão e pelo amigo que és. Eu amo ter você em minha vida. Aos meus sogros, Joana e Ribamar, serei eternamente grata a vocês, obrigada pelo apoio e por cuidarem de mim. Deus continue abençoando vocês.

Aos professores e amigos da Pós-graduação muito obrigada pelos ensinamentos e experiências a nós apresentados. Ao meu orientador, Alexandre Mesquita, obrigada pela orientação, pela ajuda em esclarecer minhas dúvidas, pelo apoio e por acreditar no meu trabalho. Ao meu coorientador, Rafael Bayma, obrigada pela disponibilidade em me ajudar, pela paciência em cada reunião e por cada ensinamento, serão de grande valia para minha vida profissional e pessoal. Ao professor Walter dos Santos, sou grata pelo seu apoio e ajuda no Laboratório de Engenharia Mecânica de Tucuruí. Muito obrigada a todos.

Meus agradecimentos à Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior (CAPES), pelo suporte financeiro a mim concedido para que a pesquisa do mestrado fosse realizada.

RESUMO

Nas indústrias, a preocupação em disponibilidade total de máquinas e equipamentos mecânicos no setor produtivo é assunto de pesquisas e testes para obtenção de técnicas mais eficientes aplicadas ao monitoramento e diagnóstico de falhas. Rolamentos são elementos de máquina de grande aplicação no setor industrial, e apresentam alto índice de falhas, que geram paradas da máquina para realizar manutenção. Por esse motivo este trabalho apresenta técnica de Inteligência Artificial aplicada aos sinais de vibrações de uma máquina rotativa para diagnóstico de falhas em seus rolamentos. Os sinais de vibração fazem parte de um banco de dados aberto, oferecido pela *Case Western Reserve University*. Neste trabalho é aplicado de duas formas o algoritmo de classificação Máquina de Vetores de Suporte (*Support Vector Machine, SVM*) para o diagnóstico de defeitos em rolamento. No primeiro caso são usados preditores estatísticos (Valor RMS, Fator de Crista, Fator K, Curtose e *Skewness*) como atributos para o classificador SVM. No segundo caso, o processamento do sinal é feito aplicando o método EEMD (*Ensemble Empirical Mode Decomposition*), que gera vários sinais denominados de Funções de Modos Intrínsecos (IMFs). Para cada IMF faz-se a sua modelagem usando Modelagem Autoregressiva (AR), e os coeficientes dos modelos AR de cada IMF são usados como atributos para o classificador SVM. As análises são realizadas para grupos de treinamentos e validação, com janelas escolhidas aleatoriamente e com janelas escolhidas na sequência temporal, considerando-se dois problemas de classificação dentro dos mesmos dados: um considera-se a mesma severidade e muda apenas o tipo de defeito e a outra em que tanto a severidade quanto o tipo de defeito variam. Como resultados, as metodologias apresentaram excelentes resultados de confiabilidade para diagnóstico de falhas em rolamentos.

Palavras-chave: Diagnóstico de falhas. Mancais de rolamento. EEMD. Modelagem AR. SVM.

ABSTRACT

In industries, the concern in total availability of machines and the mechanical equipment in the productive area it's subject of research and tests to obtain more efficient techniques to be applied for monitoring and faults' diagnosing. Bearings are machine elements of great application in the industrial area and they present high fault index that generate machine's stops to carry out maintenance. For this reason, this paper presents Artificial Intelligence technique applied to the vibration signals of a rotary machine for fault diagnosis in its bearings. The vibration signals are part of an open database offered by Case Western Reserve University. In this paper the Support Vector Machine (SVM) classification algorithm is applied in two ways for the rolling bearings faults' diagnosis. In the first case statistical predictors (Root Mean Square Value, Crest Factor, K Factor, Kurtosis and Skewness) are used as features for the SVM classifier. In the second case, the signal processing is performed by applying the Ensemble Empirical Mode Decomposition (EEMD), which generates several signals called Intrinsic Mode Functions (IMFs). For each IMF, it's modeled using Autoregressive Modeling (AR), and the AR modeling coefficients of each IMF are used as features for the SVM classifier. The analyzes are performed for training and validation groups, with randomly chosen window and with temporal sequence chosen window, considering two classification problems within the same data, the first one considers the same severity and only changes the fault type and the other vary both severity and fault type. As result, both methodologies presented excellent reliability results for bearing faults' diagnosis.

Keywords: Faults' diagnosis. Rolling bearings. EEMD. AR modeling. SVM.

LISTA DE ILUSTRAÇÕES

Figura 2.1 – Esquema do processo de diagnóstico de defeitos em rolamentos.	18
Figura 3.1 – Elementos rolamento rígido de esferas.	24
Figura 3.2 – Formas de carregamentos em rolamentos.	25
Figura 3.3 – Surgimento da falha na superfície de um rolamento.	25
Figura 3.4 – Etapas método EMD.	30
Figura 3.5 – Aplicação da (a) EMD e (b) EEMD a um sinal.	32
Figura 3.6 – Classificação linearmente separável.	37
Figura 3.7 – Cálculo da distância d entre os hiperplanos H_1 e H_2 .	38
Figura 3.8 – (a) Conjunto de dados não linear; (b) Fronteira não linear nos dados de entrada e (c) Fronteira linear no espaço de atributos.	44
Figura 4.1 – Bancada de teste – Banco de dados <i>Case Western Reserve University</i> .	47
Figura 4.2 – Conjunto de dados de treinamento e validação a) janelas aleatórias b) janelas sequenciais.	49
Figura 4.3 – Sinal no domínio do tempo do rolamento: a) sem defeito b) com defeito 0,007" na pista interna c) com defeito 0,007" na pista externa d) com defeito 0,007" na esfera.	50
Figura 4.4 – Atributos estatísticos – condições do rolamento a)RMS b)Fator de Crista c)Fator K d)Curtose e) <i>Skewness</i> .	53
Figura 4.5 – Acurácia – 4 classes – defeito 0,007" – dados aleatórios.	55
Figura 4.6 – Gráfico de espalhamento defeito 0,007" – Atributos estatísticos, dados aleatórios – janela 200 amostras.	57
Figura 4.7 – IMFs rolamento com defeito pista interna 0,007", dados aleatórios – janela 200 amostras a) IMF1 b) IMF2.	58
Figura 4.8 – Gráfico de espalhamento defeito 0,007" – Atributos EEMD+AR, dados aleatórios – janela 200 amostras.	59
Figura 4.9 – Acurácia – 4 classes – defeito 0,007" – dados sequenciais.	60
Figura 4.10 – Gráfico de espalhamento defeito 0,007" – Atributos	61

estatísticos, dados sequenciais – janela 200 amostras.	
Figura 4.11 – Gráfico de espalhamento defeito 0,007" – Atributos EEMD+AR, dados sequenciais – janela 200 amostras.	63
Figura 4.12 – Acurácia – 4 classes – defeito 0,021" – dados aleatórios.	64
Figura 4.13 – Gráfico de espalhamento defeito 0,021" – Atributos estatísticos, dados sequenciais – janela 400 amostras.	65
Figura 4.14 – Gráfico de espalhamento defeito 0,021" – Atributos EEMD+AR, dados sequenciais – janela 400 amostras.	66
Figura 4.15 – Acurácia – 4 classes – defeito 0,021" – dados sequenciais.	67
Figura 4.16 – Gráfico de espalhamento defeito 0,021" – Atributos estatísticos, dados sequenciais – janela 400 amostras.	68
Figura 4.17 – Gráfico de espalhamento defeito 0,021" – Atributos EEMD+AR, dados sequenciais – janela 400 amostras.	69
Figura 4.18 – Acurácia combinação atributos estatísticos – dados aleatórios.	70
Figura 4.19 – Acurácia, dados aleatórios – 7 classes.	71
Figura 4.20 – Gráfico de espalhamento 7 classes – Atributos estatísticos, dados aleatórios – janela de 800 amostras.	73
Figura 4.21 – Gráfico de espalhamento 7 classes– Atributos estatísticos, dados aleatórios – janela de 1000 amostras.	73
Figura 4.22 – Gráfico de espalhamento 7 classes – Atributos EEMD+AR, dados aleatórios – janela de 1000 amostras.	75
Figura 4.23 – Acurácia, dados sequenciais – 7 classes.	75
Figura 4.24 – Gráfico de espalhamento 7 classes – Atributos estatísticos, dados sequenciais – janela de 1000 amostras.	77
Figura 4.25 – Gráfico de espalhamento 7 classes – Atributos EEMD+AR, dados sequenciais – janela de 1000 amostras.	78
Figura 4.26 – Acurácia combinação atributos estatísticos – dados aleatórios.	79

LISTA DE TABELAS

Tabela 4.1 – Matriz confusão defeito 0,007" – Atributos estatísticos, dados aleatórios – janela 200 amostras.	56
Tabela 4.2 – Matriz confusão defeito 0,007" – Atributos EEMD+AR, dados aleatórios – janela 200 amostras.	59
Tabela 4.3 – Matriz confusão defeito 0,007" – Atributos estatísticos, dados sequenciais – janela 200 amostras.	61
Tabela 4.4 – Matriz confusão defeito 0,007" – Atributos EEMD+AR, dados sequenciais – janela 200 amostras.	62
Tabela 4.5 – Matriz confusão defeito 0,021" – Atributos estatísticos, dados aleatórios – janela 400 amostras.	65
Tabela 4.6 – Matriz confusão defeito 0,021" – Atributos EEMD+AR, dados aleatórios – janela 400 amostras.	66
Tabela 4.7 – Matriz confusão defeito 0,021" – Atributos estatísticos, dados sequenciais – janela 400 amostras.	68
Tabela 4.8 – Matriz confusão defeito 0,021" – Atributos EEMD+AR, dados sequenciais – janela 400 amostras.	69
Tabela 4.9 – Matriz confusão 7 classes – Atributos estatísticos, dados aleatórios – janela 800 amostras.	72
Tabela 4.10 – Matriz confusão 7 classes – Atributos EEMD+AR, dados aleatórios – janela 800 amostras.	74
Tabela 4.11 – Matriz confusão 7 classes – Atributos estatísticos, dados sequenciais – janela 800 amostras.	76
Tabela 4.12 – Matriz confusão – Atributos EEMD+AR, dados sequenciais – janela 800 amostras.	77

LISTA DE QUADROS

Quadro 4.1 – Especificações rolamento SKF 6205-2RS.	48
Quadro 4.2 – Classes – condição do rolamento.	51

LISTA DE ABREVIATURAS

SD – Sem defeito

PI7 – Com defeito pista interna 0,007"

PE7 – Com defeito pista externa 0,007"

ESF7 – Com defeito esfera 0,007"

PI21 – Com defeito pista interna 0,021"

PE21 – Com defeito pista externa 0,021"

ESF21 – Com defeito pista interna 0,021"

LISTA DE SIGLAS

AR – *Autoregressive*

ARMA – *Autoregressive Moving Average*

CWRU – *Case Western Reserve University*

DFT – *Discrete Fourier Transform*

EDM – *Electro Discharge Machining*

EMD – *Empirical Mode Decomposition*

EEMD – *Ensemble Empirical Mode Decomposition*

FFT – *Fast Fourier transform*

IMFs – *Intrinsic Mode Functions*

RMS – *Root Mean Square*

SVM – *Support Vector Machine*

SVs – *Support Vectors*

SUMÁRIO

1.	INTRODUÇÃO	14
1.1	Considerações gerais	14
1.2	Objetivos	16
1.2.1	Objetivo geral	16
1.2.2	Objetivos específicos	16
1.3	Estruturação do trabalho	17
2.	REVISÃO BIBLIOGRÁFICA	18
3.	FUNDAMENTAÇÃO TEÓRICA	24
3.1	Falhas em rolamento	24
3.1.1	Técnicas de detecção de falhas em rolamento	26
3.2	<i>Ensemble Empirical Mode Decomposition - EEMD</i>	28
3.3	<i>Autoregressive Modeling - AR</i>	33
3.4	<i>Support Vector Machine - SVM</i>	35
3.4.1	SVM lineares com margens rígidas	37
3.4.2	SVM lineares com margens suaves	42
3.4.3	SVM não linear	44
4.	RESULTADOS E DISCUSSÕES	47
4.1	Descrição da bancada	47
4.2	Processamento	48
4.3	Aplicação SVM com uso de atributos estatísticos e atributos EEMD+AR	54
4.3.1	Atributos estatísticos e atributos EEMD+AR – 4 classes	55
4.3.1.1	Defeito 0,007" – dados aleatórios e sequenciais	55
4.3.1.2	Defeito 0,021" – dados aleatórios e sequenciais	63
4.3.1.3	Combinação de atributos estatísticos – dados aleatórios – 4 classes	70
4.3.2	Atributos estatísticos e atributos EEMD+AR – 7 classes	70
4.3.2.1	Defeito 0,007" e 0,021" – dados aleatórios e sequenciais	71
4.3.2.2	Combinação de atributos estatísticos – dados aleatórios – 7 classes	78
5.	CONCLUSÃO	80
6.	REFERÊNCIAS	82

1. INTRODUÇÃO

1.1. Considerações gerais

Diversos elementos de máquina compõem um equipamento, e dentre esses, alguns são de suma importância para seu bom desempenho. Um exemplo são os rolamentos, cerca de 90% dos maquinários industriais possuem rolamentos, compõem um importante componente de ligação entre componentes com movimentos de rotação (SANTOS, 2017). Sua aplicação e dimensões variam conforme necessidade da operação do sistema mecânico.

Os mancais de rolamento são projetados para operar por diversas horas de trabalho. Portanto, a manutenção mecânica desses elementos de máquinas, em qualquer setor da indústria, é uma questão de grande relevância para o funcionamento e vida útil da máquina. Neste cenário, os diagnósticos preditivos de defeitos se tornam imprescindíveis para obtenção do melhor custo/benefício para o setor produtivo.

Na busca por redução de custos, eficiência nas intervenções de manutenção mecânica e aumento da disponibilidade dos equipamentos, técnicas de qualidade são empregadas para minimizar e/ou eliminar tais problemas e gerar benefícios para a indústria (SANTOS *et al.*, 2017). Em um sistema mecânico há a possibilidade de presença de danos ou falhas, sendo esses causados por trincas, desbalanceamento ou desalinhamento, entre outros. Com isso, as propriedades dinâmicas e estruturais são comprometidas. Dessa forma esse defeito deve ser identificado e corrigido, para que o sistema opere em bom estado e cumpra sua função.

Ao longo dos anos, várias técnicas de manutenção e monitoramento vêm sendo empregadas para realizar diagnósticos, e com o advento da tecnologia, a técnica de Inteligência Artificial têm possibilitado diagnósticos mais eficientes e sem a intervenção do especialista.

Em um sistema mecânico a falha do componente rolamento está associada a fatores como: desgaste, corrosão, trincas, utilização inadequada de procedimentos para montagem e manutenção, lubrificação incorreta, erosão elétrica, efeitos do ambiente ao qual está inserido e vibração (MARÇAL, 2000). Os defeitos desenvolvidos no rolamento podem estar na pista externa, interna, nos elementos rolantes ou gaiola (BARSZCZ e SAWALHI, 2012). Esses defeitos acarretam danos tanto ao componente quanto ao sistema, que pode ocasionar

problemas físicos: quebra e inutilização de seus componentes; e financeiro: indisponibilidade do sistema na área produtiva.

Três tipos de estratégia de manutenção são aplicados pelas indústrias, a corretiva, preventiva e preditiva. A manutenção corretiva é aquela em que se utiliza o equipamento até sua quebra, e somente após apresentar defeito é realizada a sua parada e manutenção e/ou troca. Já a manutenção preventiva necessita de uma base estatística e requer um planejamento baseado no tempo esperado de durabilidade do componente, um ponto a ser considerado é que pode haver a troca de um componente, sem tal necessidade, pois sua vida útil pode não ter sido atingida. A estratégia mais eficiente para a detecção de falhas, que tem por objetivo quando e como fazer intervenções na máquina é a preditiva, pois analisa a condição de operação e funcionamento do sistema mecânico através de instrumentação de medição e análise de sinais. Para realizar esta medição, acompanhamento da evolução das falhas e avaliação segura das condições do sistema mecânico há vários métodos: análise de vibração, emissão acústica, análise de partícula, termografia e análise de óleo (TREVISAN, 2011).

Diante de diversas técnicas de previsões de falhas disponíveis no mercado para a identificação de falhas em rolamento, a análise de vibrações, que através do processamento de sinais identifica características próprias de cada defeito, tem sido tema de diversos estudos na área de engenharia (SANTOS, 2017), e conforme Sidar (2015) é a técnica mais popular e utilizada, devido à confiabilidade e sensibilidade à severidade da falha.

Sendo assim o diagnóstico e monitoramento são possíveis no controle de falhas e programação adequada para os setores de produção e manutenção, para evitar gastos excessivos com substituição precipitada do rolamento e indisponibilidade do equipamento (MARÇAL, 2000).

As técnicas de processamento de sinais podem ser aplicadas no domínio do tempo, no domínio da frequência e no domínio tempo-frequência. Cada um restrito as limitações de aplicação e diagnóstico. À exemplo da aplicação da FFT e Técnica do Envelope, que são utilizadas apenas para sinais estacionários, em que a velocidade de rotação da máquina é constante durante o período de aquisição, e isso impõem limite para diversas aplicações (NAPOLITANO, 2014). Sendo assim a utilização da Inteligência Artificial em classificar e prever defeitos auxilia no diagnóstico e eliminação de limitações. E conforme Santos (2017), na última década, a pesquisa de análise de sinal foi dominada por modelos que englobam não estacionaridade como uma característica importante.

Neste trabalho propõe-se aplicar a técnica de Inteligência Artificial, mais especificamente, Aprendizagem de Máquina (*Machine Learning*) por meio do classificador

Máquina de Vetores de Suporte (*Support Vector Machine*, SVM) para diagnóstico de falhas em rolamentos do banco de dados da *Case Western Reserve University*, cujos sinais são publicamente disponíveis e amplamente utilizados (SMITH e RANDALL, 2015).

Na primeira análise os atributos para o classificador são obtidos através das técnicas estatísticas (Valor RMS, Fator de crista, Fator K, Curtose e *Skewness*) dos sinais de vibração do rolamento com defeito e sem defeito. E na segunda análise, que pode ser aplicado tanto para sinais estacionários quanto não estacionários, o processamento do sinal é realizado por meio do método EEMD (*Ensemble Empirical Mode Decomposition*), para extração das IMFs (*Intrinsic Mode Functions*), seguido da aplicação da modelagem Autorregressiva (AR) em cada IMF, para que assim os coeficientes do modelo AR de cada IMF sejam os atributos para alimentar o SVM Classificador.

1.2. Objetivos

1.2.1. Objetivo geral

Analisar e comparar atributos de classificação voltados para diagnóstico de falhas em rolamentos por meio da técnica de Aprendizado de Máquina (*Machine Learning*) com o classificador SVM.

1.2.2. Objetivos específicos

- Obtenção dos atributos estatísticos e aplicação no classificador SVM;
- Aplicação do método EEMD (*Ensemble Empirical Mode Decomposition*) ao sinal de vibração obtido;
- Utilização do modelo AR em sinais temporais das IMFs;
- Inserir os atributos provenientes da modelagem AR no classificador SVM para verificar se o rolamento está sem defeito ou com defeito.
- Investigar o desempenho do Classificador SVM através: do tamanho da janela de cada atributo, número de classes inseridas e diferentes combinações de atributos estatísticos.

1.3. Estruturação do trabalho

Este trabalho apresenta outros cinco capítulos, conforme estruturação a seguir.

No capítulo 2 é feita a revisão bibliográfica, em que é possível descrever a evolução no tempo das técnicas de detecção de falhas, e o monitoramento da condição da máquina, bem como trabalhos recentes que tratam da detecção de falhas em rolamentos por meio do método EEMD, modelagem AR e Classificador SVM.

No capítulo 3 apresenta-se a fundamentação teórica dos temas tratados nesta dissertação, iniciando pelo elemento de máquina em estudo, rolamento, e suas formas de carregamento. Detalham-se as técnicas de detecção de falhas em rolamento: Valor Nível Global RMS, Fator de Crista, Fator K, Curtose e *Skewness*. Em seguida apresentam-se as teorias do método EEMD, AR e SVM.

O estudo de caso é descrito no capítulo 4, em que se descreve a bancada de teste utilizada e o processamento de sinais detalhado para cada análise. Também neste capítulo apresentam-se os resultados e discussões de cada atributo inserido no classificador SVM. Desta forma, possibilitou-se verificar o desempenho de cada metodologia investigada.

A Conclusão está no quinto capítulo, em que apresenta a contribuição obtida a partir dos resultados e a importância do diagnóstico de falhas em rolamento no setor produtivo. Sugestões para trabalhos futuros são propostas no âmbito de realizar novas investigações com diferentes parâmetros.

2. REVISÃO BIBLIOGRÁFICA

Esta dissertação apresenta métodos de detecção e diagnóstico de falhas em rolamento através da Inteligência Artificial; o primeiro utiliza técnicas estatísticas como atributos para o classificador. O segundo utiliza como atributos os coeficientes da Modelagem Autoregressiva (*Autoregressive Modeling - AR*) das funções IMFs (*Intrinsic Mode Functions*) provenientes da aplicação do método EEMD (*Ensemble Empirical Mode Decomposition*). Em outras palavras, inicialmente, obtêm-se os sinais de vibração do rolamento em bom estado e danificados na pista interna, pista externa e esferas. O processo de decompor os sinais de vibração em funções de modo intrínseco (IMFs) e extrair atributos é introduzido para desenvolver um vetor de atributos para alimentar o modelo matemático autoregressivo. Por fim, o classificador SVM é usado para classificar e visualizar esse vetor de atributos, em que separam-se as características danificadas dos componentes em bom estado do rolamento. A Modelagem AR melhora o desempenho da classificação selecionando os atributos, para maximizar o efeito de visualização da separação e agrupamento de vetores de parâmetros no classificador. A Figura 2.1 apresenta o processo do segundo método de diagnóstico deste trabalho.

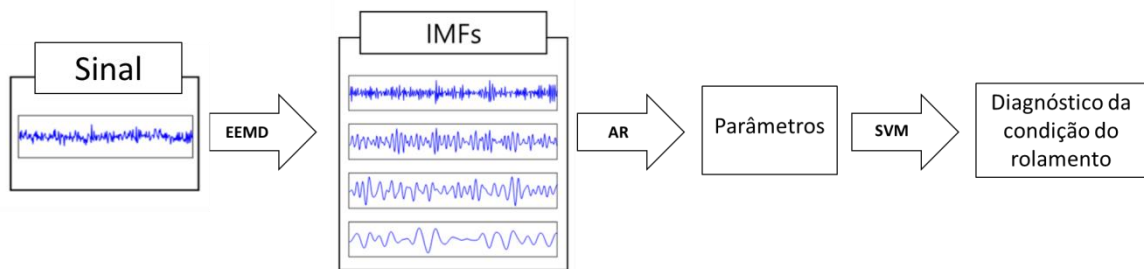


Figura 2.1 – Esquema do processo de diagnóstico de defeitos em rolamentos (Autoria própria).

Diante das estratégias de manutenção disponíveis, as indústrias utilizam aquela que melhor atende à sua necessidade e que possa evitar possíveis falhas no sistema. De acordo com a aplicação de cada estratégia, o tempo de medição varia, pode ser periódico ou contínuo. Segundo Otani e Machado (2008) a manutenção preditiva atende alguns requisitos, dentre eles, permite a coleta de dados com o equipamento em funcionamento, ou com o mínimo de interferência possível no processo de produção e permite coletar dados que possibilitem a análise de tendência.

Para realizar o monitoramento da condição da máquina alguns parâmetros devem ser considerados, como: vibração, detritos no óleo, emissões acústicas. Alterações nesses parâmetros ajudam na detecção do desenvolvimento de falhas, diagnóstico de causas de problemas e antecipação de falhas (TANDON e PAREY, 2006). Ainda de acordo com Tandon e Parey (2006), a principal função do monitoramento de condições é fornecer um conhecimento da condição da máquina e sua taxa de mudança, para que medidas preventivas possam ser tomadas em um momento apropriado; e que a aplicação do monitoramento de condições nas indústrias resulta em economia nos custos de manutenção e maior disponibilidade e segurança.

Devido a grande porcentagem de uso do rolamento em sistemas mecânicos e para uma boa performance do sistema, este deve operar desempenhando sua função, porém se o rolamento apresentar falhas, o funcionamento da máquina estará comprometido. Por essa razão a necessidade de aplicação de técnicas de processamento de sinais e monitoramento do sistema mecânico, para diagnóstico de falhas, a fim de aumentar a vida útil dos equipamentos, desenvolver manutenção planejada, minimizar custos e elevar a qualidade do conjunto.

A técnica mais aplicada para o monitoramento das condições do equipamento é a análise de vibração. Nela pode-se realizar análise no domínio do tempo, no domínio da frequência, e no domínio tempo-frequência. Alfredson e Mathew (1985) apresentaram técnicas de monitoramento da condição de rolamentos, no domínio do tempo e frequência. Neste trabalho utiliza-se a análise no domínio do tempo, apesar de que muitos pesquisadores também classificarem o EEMD como uma técnica no domínio tempo-frequência.

Após a obtenção dos sinais de vibração do rolamento em bom estado e com defeitos em sua pista interna, pista externa e esferas, a primeira abordagem é feita através do cálculo das técnicas estatísticas (Valor RMS, Fator de Crista, Fator K, Curtose e *Skewness*) e outra aplica-se EEMD, para decompor cada sinal em IMFs, e posteriormente em cada IMF é aplicado a modelagem AR para então gerar os atributos para inserir no SVM classificador. Para fins de treinamento, cada conjunto de dados pertencente à condição do rolamento foi dividido em dois grupos com o mesmo número de exemplos, um para treinamento e outro para validação. O treinamento é realizado de duas maneiras: com janelas escolhidas aleatoriamente e com janelas escolhidas na sequência temporal.

A EMD é baseada nas seguintes hipóteses: (1) o sinal tem pelo menos dois extremos – um máximo e um mínimo; (2) a escala de tempo característica é definida pelo lapso de tempo entre os extremos; e (3) se os dados forem totalmente desprovidos de extremos, mas apenas contem pontos de inflexão, então eles podem ser diferenciados uma ou mais vezes

para revelar os extremos. Os resultados finais podem ser obtidos por integrações dos componentes. A essência do método é identificar os modos oscilatórios intrínsecos por suas escalas de tempo características nos dados empiricamente e depois decompor os dados de acordo. (HUANG *et al.*, 1998).

Wu e Huang (2005) afirmam que a Decomposição do Conjunto de Modo Empírico (EEMD) de fato pode separar sinais de diferentes escalas sem o indevido efeito de mixagem (combinação) de modos.

É possível encontrar diversos trabalhos que aplicam a EEMD, à exemplo tem-se Mahgoun *et al.* (2012) que afirma que o sinal de vibração é geralmente ruidoso e para detectar a falha em um estágio inicial de seu desenvolvimento, geralmente o sinal residual é usado. Mahgoun *et al.* (2012) aplicam o EEMD com o objetivo de detectar falhas localizadas de danos em um estágio inicial.

Caesarendra *et al.* (2013) utilizaram o método EMD e EEMD em dois casos reais de rolamento de giro de baixa velocidade (1 rpm) com progressão de falha natural e os resultados mostram que o EMD e EEMD foram melhores que a aplicação da FFT na identificação de frequências de falhas.

Lei *et al.* (2013) utilizaram o método do Conjunto de Decomposição em Modo Empírico (*Ensemble Empirical Mode Decomposition – EEMD adaptativo*) para o diagnóstico de falhas de máquinas rotativas. Neste método, o número de filtragem foi selecionado de forma adaptativa, e a amplitude do ruído adicionado muda com os componentes de frequência do sinal durante o processo de decomposição. A simulação, a etapa experimental e os resultados da aplicação demonstraram que a EEMD adaptativa fornece os melhores resultados em comparação com a EEMD original no diagnóstico de máquinas rotativas.

Liu *et al.* (2015) aplicaram o método EEMD e Rede Bayesiana para diagnóstico de falhas em uma bomba de engrenagens, o método proposto utilizou todas as informações úteis além dos sinais do sensor e desenvolveu uma metodologia para detecção de falhas. Já Yi *et al.* (2018) propuseram três índices de estado estacionário baseados no método EEMD para monitoramento de condições e diagnóstico de falhas de rolamentos de eixos ferroviários.

Através da particularidade de cada problema de engenharia vivenciado diariamente, obtém-se dados de diversas formas e às vezes inúmeras variáveis para solucioná-lo, e é a partir de medições reais das propriedades físicas que é possível construir equações (modelos matemáticos). E a partir da Identificação de Sistemas é possível simular, otimizar ou prever esses comportamentos reais, pois a identificação compreende técnicas e algoritmos capazes de modelar e analisar cada modelo matemático a partir dos dados experimentais observados e

coletados. Segundo Ljung (2010), identificação de sistema é um tópico extenso, com diferentes técnicas que dependem das características dos modelos a serem estimados: linear, não linear, híbrido, não paramétrico e etc. Conforme Aguirre (2014) na última década tem-se verificado uma tendência geral do uso de técnicas de identificação e análise de sistemas desejável e até mesmo necessário em praticamente todas as áreas do conhecimento humano, e refere-se à identificação de sistemas como sendo uma área do conhecimento que estuda maneiras de modelar e analisar sistemas a partir de observações, ou seja, de dados.

As características dos fenômenos dinâmicos podem ser descritas usando modelos de séries temporais. Conforme Parmezan e Batista (2016), a análise de séries temporais consiste na aplicação de técnicas, provenientes de áreas relacionadas à Estatística e a Inteligência Artificial, que buscam descrever os componentes que se mostram como características de uma série em particular.

Desde o início dos anos 80, várias abordagens baseadas em modelos foram introduzidas no monitoramento de condições de máquinas, principalmente para o diagnóstico de falhas de fabricação e processamento do equipamento (CHEN, 1992). Vários métodos paramétricos estão disponíveis para modelagem de sistemas mecânicos. Estes incluem modelagem autorregressiva – AR e modelagem média autorregressiva – ARMA (CHEN, 1992 e ROFE, 1997).

Os modelos autorregressivos (AR) estabeleceram seu valor em uma ampla gama de aplicações na análise espectral digital. A capacidade dos modelos AR em termos de manipulação de segmentos curtos de dados, em que oferece melhor resolução de frequência e espectros de potência suave é vantajosa em comparação com a Transformada Discreta de Fourier (DFT) e a Transformada Rápida de Fourier (FFT) (PALANIAPPAN, 2006).

O método autorregressivo AR pode ser considerado como um filtro linear que pode ser usado para análise de sinais que são corrompidos pelo ruído branco (ATYABI *et al.*, 2017).

A modelagem Autoregressiva é um método de análise de sequência temporal cujos parâmetros compreendem informação significativa da condição do sistema. Um modelo AR preciso pode refletir as características de um sistema dinâmico. Adicionalmente, é indicado que os parâmetros de autoregressão do modelo AR são muito sensíveis à variação da condição. Entretanto quando o modelo AR é aplicado a sinais não estacionários, é difícil estimar os parâmetros de autoregressão. O modelo ARMA (autoregressivo e média móvel) pode ser aplicado a sinais não estacionários, porém é necessário maior tempo computacional (HONG *et al.*, 1989; YA e SHUZI, 1989).

Para utilização do modelo AR é necessário o pré-processamento do sinal de vibração, sendo assim para analisar os defeitos por vibração e diagnosticar o estado do rolamento é aplicado o método EEMD (*Ensemble Empirical Mode Decomposition*) ao sinal obtido do sistema mecânico (CHENG *et al.*, 2008). EEMD é um método de análise de dados assistido por ruído. Consiste em "peneirar" um conjunto de sinal de ruído branco e pode separar escalas naturalmente sem qualquer critério de seleção subjetivo a priori como no teste de intermitência para o algoritmo EMD original (WU e HUANG, 2005). Na técnica EMD qualquer sinal pode ser decomposto em um número finito de componentes (sinais) chamados funções de modos intrínsecos – IMFs – (HUANG *et al.*, 1998). Após a obtenção das IMFs pela decomposição EEMD, que agora são estacionárias, o pós-processamento do método EEMD combinado à modelagem AR pode fornecer uma ferramenta útil para analisar os sinais de vibração com falhas. E posteriormente os parâmetros do modelo AR obtidos são utilizados como dados de entrada para um classificador inteligente SVM (*Support Vector Machine*), este é um método de classificação, e servirá para classificar se o componente em estudo está sem defeito ou com defeito.

Recentemente houve avanços expressivos na área de Estatística e Inteligência Computacional que permitiram o desenvolvimento de programas capazes de detectar e classificar padrões. Um problema de Reconhecimento de Padrões consiste de ler atributos de um objeto com o intuito de classificá-lo como membro de uma classe pré-definida (classificação supervisionada) ou usá-lo para definir grupos de objetos com características semelhantes (classificação não supervisionada) (BIANCHI, 2006). Um problema de diagnóstico de rolamentos pode ser entendido sob este mesmo paradigma, em que os atributos de uma forma de onda de vibração (objeto) devem ser utilizados para determinar sob qual tipo (classe) de defeito o rolamento que a produziu se encontra.

SVM (*Support Vector Machine*), desenvolvida por Vapnik (1995) é uma técnica de aprendizagem de máquina derivada de duas fundamentações sólidas: Teoria da Aprendizagem Estatística e Otimização Matemática. SVM tem sido recentemente aplicada com sucesso a uma variedade de problemas que vão desde o reconhecimento de caracteres ao reconhecimento de objetos baseado na aparência. Alguns dos motivos para esse sucesso estão relacionados ao fato dessa técnica exibir bom desempenho de generalização em muitas bases de dados reais, e é bem fundamentada teoricamente (BARBOSA, 2015).

A técnica aplicada neste trabalho é do tipo classificação supervisionada, em que se supõe conhecida a condição de defeito sob a qual cada forma de onda coletada se encontra. Destes sinais é extraído um conjunto de atributos que são posteriormente utilizados para

treinar um algoritmo classificador. O classificador testado foi do tipo SVM (*Support Vector Machine*) e segundo Steinwart (2008), classificação é uma das principais áreas de aplicação para SVM.

Conforme Singla *et al.* (2011) classificadores SVM são normalmente treinados para problemas binários, onde se deseja distinguir entre apenas duas classes (“com defeito” e “sem defeito”, por exemplo). Problemas multiclasss normalmente combinam múltiplos classificadores deste tipo (comumente chamada de estratégia “todos contra um”) para se chegar a uma conclusão definitiva. Esta é a abordagem adotada neste trabalho.

3. FUNDAMENTAÇÃO TEÓRICA

3.1. Falhas em rolamento

Um dos componentes do sistema mecânico que pode acarretar custos à indústria é o rolamento, pois pode gerar vibrações à máquina, por apresentar tanto falhas de montagem, de fabricação, desalinhamento, passagem de corrente elétrica ou devido à vida útil. A seguir está descrito os componentes do rolamento e suas formas de carregamento.

O rolamento é composto pelos elementos: pista interna, externa, elemento girante (esferas ou rolos) e gaiola, conforme Figura 3.1.

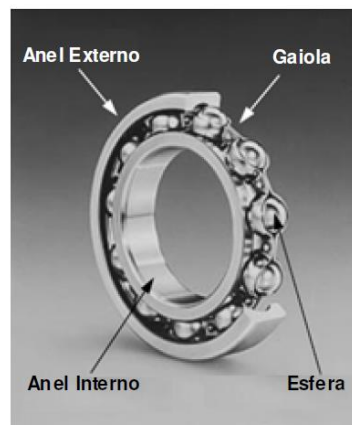


Figura 3.1 – Elementos rolamento rígido de esferas (Carvalho, 2010).

De acordo com a carga ao qual o rolamento está sujeito, este irá gerar uma zona de carregamento com distribuições distintas, sendo possível identificar na Figura 3.2 como a zona de carregamento se comporta à uma carga radial. Mesmo quando o rolamento não apresenta defeito, a vibração desenvolve-se quando os elementos girantes entram e saem desta zona. Também é possível relacionar a zona de carregamento com a geometria do rolamento, o tipo de material utilizado na confecção de seus elementos, o tipo de montagem (com ou sem pré-carga), com a espessura dos anéis do rolamento, com a folga existente no rolamento e com as características do lubrificante utilizado (BEZERRA, 2004).

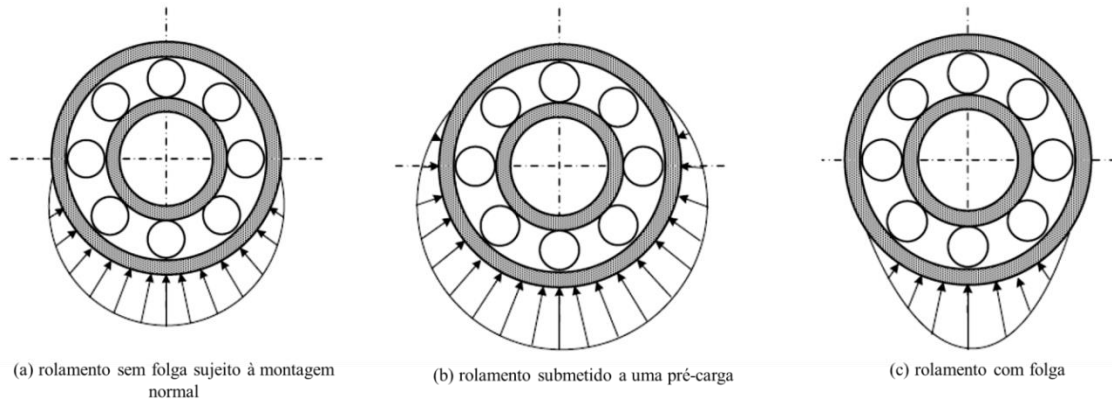


Figura 3.2 – Formas de carregamentos em rolamentos (Bezerra, 2004).

Conforme o tempo passa é natural que o rolamento apresente desgastes, independente se a montagem foi incorreta, lubrificação inadequada, desalinhamento, defeitos de fabricação ou contaminação. E conseqüentemente ao desenvolvimento de folgas, ilustrada na Figura 3.2 (c), a zona de carga irá diminuir.

Neste trabalho para aplicação da técnica de diagnóstico de falhas no rolamento, considera-se que a zona de carregamento não sofre movimentação e que a pista externa é estacionária. Com isso, o deslocamento dos elementos girantes pela pista irá provocar tensões cíclicas de cisalhamento na camada abaixo da pista, ao passarem pela zona de carregamento.

As tensões cíclicas de cisalhamento desenvolvidas ocasionam microfissuras, que surgem em pontos de pouca resistência, ou onde o material é anisotrópico ou em pontos onde ocorrem inclusões de materiais não metálicos. É possível que gradativamente, as microfissuras evoluam para a superfície da pista onde surgirá micro trincas (HARRIS, 1991; JUVINALL e MARSHEK, 1991). A Figura 3.3 ilustra esse processo.

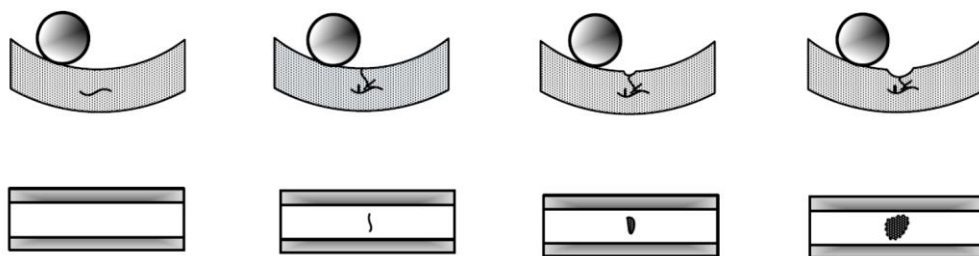


Figura 3.3 – Surgimento da falha na superfície de um rolamento (Bezerra, 2004).

Essa evolução das falhas pode desenvolver-se em qualquer elemento do rolamento; na Figura 3.3, pode-se observar o surgimento e evolução da falha na pista externa do rolamento. Isto devido à passagem contínua dos elementos rolantes na trinca, que atingiu a

superfície da pista. E como consequência tem-se a formação de covas (*pits*) e/ou descascamento (*spalls*) que evoluirá, gradualmente, até que o rolamento sofra uma falha que impossibilite o seu uso (JUVINALL e MARSHEK, 1991).

3.1.1. Técnicas de detecção de falhas em rolamento

Para realizar diagnóstico de falhas há disponível no mercado uma diversidade de técnicas para tais detecções, conforme Tandon e Choudhury (1999), o método mais utilizado é a que envolve medidas de vibração. Sendo assim, deve ser objeto de monitoramento.

Na análise de vibração para detecção das falhas em rolamento, há diversas técnicas que podem ser empregadas para diagnóstico do defeito. Essas técnicas podem ser: no domínio do tempo – método qualitativo e que é possível apenas a identificação do defeito e não a sua localização; no domínio da frequência – capaz de identificar a localização do problema (técnica do envelope, filtros passa banda) e no domínio tempo-frequência – técnica empregada quando a rotação é variável.

As técnicas no domínio do tempo englobam diversos métodos que identificam o surgimento e desenvolvimento do defeito, porém não localizam a região que apresenta a falha (KOKIL *et al.*, 2014). O método utilizado para investigação primária de integridade da máquina é o Valor do Nível Global RMS. Esses níveis são medidos a partir de um sinal $x(t)$ e comparados com valores tabelados para indicar a severidade do sinal de vibração. O valor RMS fornece apenas a indicação de que houve um aumento na energia do sinal, devido o desenvolvimento do defeito (KOKIL, 2014; SIKORA, 2016). Para um sinal discreto $x(n)$ com N amostras, o valor RMS pode ser obtido conforme Equação 3.1 (BEZERRA, 2004).

$$X_{rms} = \sqrt{\frac{1}{N} \sum_{i=1}^N [x(i)]^2} \quad (3.1)$$

Bezerra (2004) define o Fator de Crista – Equação 3.2 – como a razão do nível do pico da vibração pelo nível RMS.

$$F_{cr} = \frac{X_{pico}}{X_{rms}} \quad (3.2)$$

Este método parte do princípio que quando o rolamento sofre um dano, o nível de pico da aceleração aumenta mais rapidamente do que o nível RMS. Conforme Mesquita *et al.* (2002) na vibração normal do rolamento, sem defeitos, o fator de crista varia entre 2 e 6. Quando há evolução das falhas, o fator de crista atinge valores altos, acima de 6.

Conforme Bezerra (2004) à medida que o defeito se espalha pela superfície do rolamento o nível de ruído se eleva, aumentando o valor do nível RMS mais rápido que o valor de pico. Isto produz uma redução no Fator de Crista, de modo que este pode não ser um bom indicador de falhas em um estágio de falha severa. Por essa razão, o Fator K, definido como o produto entre o valor RMS e o valor de pico do sinal – Equação 3.3 – pode ser utilizado como tentativa de contornar a limitação da informação proveniente do Fator de Crista (BEZERRA, 2004).

$$F_k = X_{rms} X_{pico} \quad (3.3)$$

A densidade de probabilidade também é utilizada para detecção de falhas em sinais de vibração de rolamentos. Nos rolamentos sem falha, a densidade de probabilidade do sinal tem uma distribuição Gaussiana de aceleração e já o rolamento com falhas apresenta distribuição não-Gaussiana, devido ao aumento no número dos altos níveis de aceleração (MESQUITA *et al.*, 2002). Para indicar a forma da densidade de probabilidade vários momentos estatísticos podem ser usados. Na prática, os momentos de ordem 3 ou 4 são, por vezes, de interesse, mas geralmente há pouca razão estatística para examinar momentos mais altos do que estes (CASELLA, 2002). O terceiro momento estatístico é normalizado pelo cubo do desvio padrão e é conhecido como coeficiente de assimetria (*skewness coefficient*) – Equação 3.4, e conforme Finotti, *et al.* (2019) representa a simetria da distribuição de dados, e os valores negativos de assimetria indicam os dados que estão assimétricos para à esquerda e valores positivos para a assimetria à direita (AMIT *et al.*, 2014). E o Fator de Curtose é o quarto momento normalizado em relação ao desvio padrão elevado à quarta potência, que indica o achatamento da distribuição considerando uma curva Gaussiana (FINOTTI *et al.*, 2019). É calculado conforme Equação 3.5.

$$S = \frac{\frac{1}{N} \sum_{i=1}^N [x(i) - \bar{x}]^3}{\sigma^3} \quad (3.4)$$

$$K = \frac{\frac{1}{N} \sum_{i=1}^N [x(i) - \bar{x}]^4}{\sigma^4} \quad (3.5)$$

sendo $x(i)$ o sinal temporal de vibração, N o número de amostras do sinal, σ é o desvio padrão e \bar{x} é o valor médio do sinal temporal.

Conforme Kokil *et al.* (2014) para mancais em bom estado o valor do Fator de Curtose teórico é de 3, e quando há falhas este valor aumenta. E à medida que a falha se espalha pela superfície da pista ou esfera o valor do Fator de Curtose diminui podendo atingir valores iguais ou menores que três mascarando uma possível falha no rolamento (TANDON e CHOUDHURY, 1999).

O valor eficaz ou RMS é o valor relacionado ao nível de energia ou potência de um dado sinal. O parâmetro Curtose é um parâmetro adimensional que representa o momento estatístico de quarta ordem do sinal, e indica o grau de “achatamento” da função de distribuição de probabilidade do sinal (BEZERRA, 2004).

É importante ressaltar que o cálculo de tais atributos estatísticos normalmente é feito a partir de amostras, o que pode introduzir erros nas estimativas se o registro não for suficientemente longo, e conseqüentemente levar a diagnósticos incorretos. Por essa razão, uma das propostas deste trabalho é investigar como o diagnóstico é afetado pelo tamanho da janela sobre a qual os atributos são calculados.

Observa-se que há uma diversidade de medidas estatísticas que podem ser usadas para caracterizar formas de onda de vibração e os padrões de anormalidade estão relativamente bem estabelecidos em literatura. No entanto, é natural esperar que o uso simultâneo de todas estas informações torne-se confuso, sobretudo ao considerar fatores agravantes como o nível de incerteza das medições (provocado pelo ruído de medição ou o tamanho limitado da série temporal) e rotação da máquina. Por este motivo torna-se importante treinar sistemas computacionais para automatizar a tarefa do reconhecimento do defeito. Nos tópicos seguintes o método de reconhecimento de padrões é abordado.

3.2. Ensemble Empirical Mode Decomposition - EEMD

EEMD (*Ensemble Empirical Mode Decomposition*) é um método de análise de dados assistido por ruído. Consiste em "peneirar" um conjunto de sinal de ruído branco e pode

separar escalas naturalmente sem qualquer critério de seleção subjetivo a priori, como no teste de intermitência para o algoritmo EMD original (WU e HUANG, 2005).

Conforme Huang *et al.* (1998), o método EMD é desenvolvido a partir da suposição de que qualquer sinal consiste em diferentes modos intrínsecos simples de oscilação. Cada modo linear ou não linear terá o mesmo número de extremo e cruzamento zero. Existe apenas um extremo entre sucessivos cruzamentos zero. Cada modo deve ser independente dos outros e deve satisfazer a seguinte definição:

(1) Em todo o conjunto de dados, o número de extremos e o número de cruzamentos nulos devem ser iguais ou diferentes no máximo por um.

(2) Em qualquer ponto, o valor médio do envelope definido pelo máximo local e o envelope definido pelo mínimo local é zero.

Uma IMF representa um modo oscilatório simples comparado com a função harmônica simples. E com esta definição qualquer sinal $x(t)$ pode ser decomposto da seguinte maneira (CHENG *et al.*, 2008):

1. Identificar todos os extremos locais e, em seguida, conectar todos os máximos locais por uma linha cúbica *spline* como o envelope superior;

2. Repetir o procedimento para o mínimo local para produzir o envelope inferior. Os envelopes superior e inferior devem cobrir todos os dados entre eles;

3. A média do valor de envelope superior e inferior é designada como m_1 e a diferença entre o sinal $x(t)$ e m_1 é o primeiro componente, h_1 , conforme Equação 3.6

$$x(t) - m_1 = h_1 \quad (3.6)$$

Idealmente, se h_1 é uma IMF, então h_1 é a primeira IMF componente de $x(t)$.

4. Se h_1 não for uma IMF, h_1 é tratada como o sinal original e repetem-se as etapas 1, 2, 3, e obtém-se a Equação 3.7.

$$h_1 - m_{11} = h_{11} \quad (3.7)$$

A Figura 3.4 apresenta as etapas 1, 2, 3 e 4 explanadas acima para o método EMD.

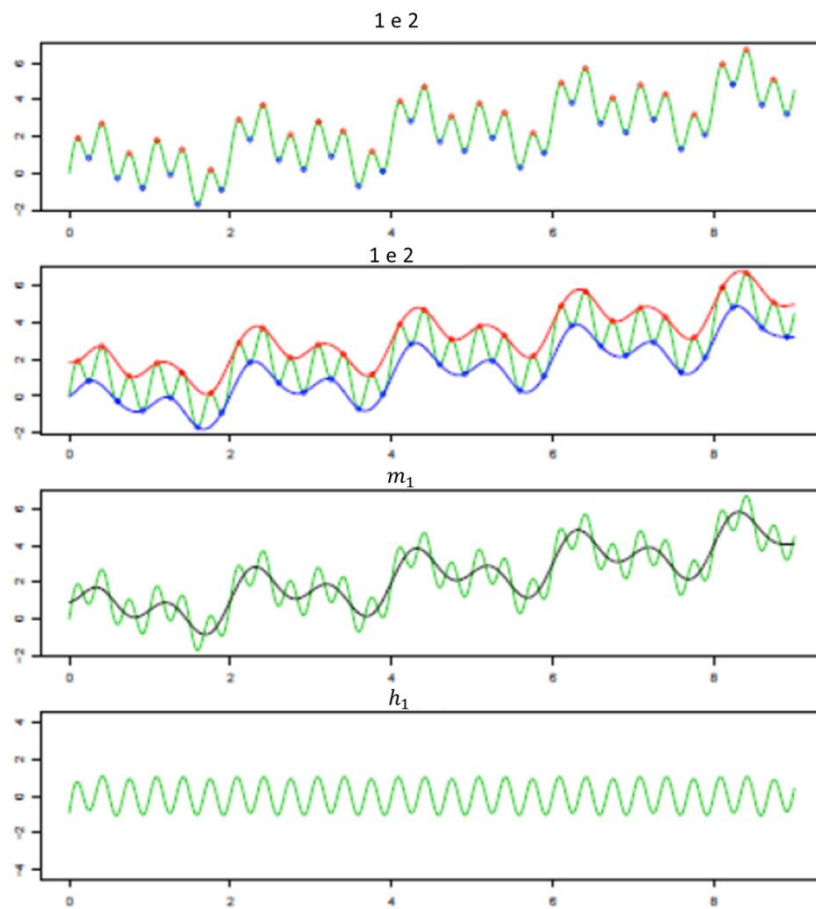


Figura 3.4 – Etapas método EMD (Kim, 2006).

Após filtragem repetida, ou seja, até k vezes, h_{1k} se torna uma IMF:

$$h_{1(k-1)} - m_{1k} = h_{1k} \quad (3.8)$$

Então é designado como $c_1 = h_{1k}$, a primeira componente IMF a partir dos dados originais.

5. Separa-se c_1 de $x(t)$, é possível obter $r_1 = x(t) - c_1$. Agora r_1 é considerado o dado original e repetem-se as etapas acima, sendo assim pode-se obter a segunda componente IMF c_2 de $x(t)$. Repete-se o processo descrito acima por n vezes, então n -IMFs do sinal $x(t)$ podem ser obtidas. Então:

$$\begin{aligned} r_1 - c_2 &= r_2 \\ &\vdots \\ r_{n-1} - c_n &= r_n \end{aligned} \quad (3.9)$$

O processo de decomposição pode ser interrompido quando o r_n se tornar uma função monótona da qual nenhuma IMF pode ser extraída. E obtém-se a Equação 3.10.

$$x(t) = \sum_{j=1}^n c_j + r_n \quad (3.10)$$

Sendo assim, pode-se conseguir uma decomposição do sinal em n -modos empíricos e um resíduo r_n , que é a tendência média de $x(t)$. Cada uma das IMFs c_1, c_2, \dots, c_n inclui diferentes bandas de frequência que vão de alto a baixo e são estacionárias (CHENG *et al.*, 2008).

Porém Gaci (2016) afirma que ao se utilizar o algoritmo EMD, este pode apresentar o efeito de mixagem de modos, efeito este que ocorre quando as oscilações com escalas de tempo distintas são preservadas em uma IMF, ou as oscilações com a mesma escala de tempo são filtradas (peneiradas) em diferentes IMFs. Para que fosse solucionado este problema Wu e Huang (2005) sugeriram um algoritmo EMD auxiliado por ruído, denominado EEMD, que permite uma melhor aptidão de separação em escala do que o método EMD padrão. O método EEMD consiste em adicionar diferentes séries de ruído branco ao sinal em várias tentativas. Como o ruído adicional é diferente em cada tentativa, as IMFs resultantes não exibem qualquer correlação com as IMF correspondentes de uma tentativa para outra. Se o número de tentativas for adequado, o ruído adicionado pode ser eliminado pela média dos conjuntos das IMFs obtidas relacionadas às diferentes tentativas.

Portanto o EEMD é proposto para resolver a questão da mixagem de modo, em que se insere o ruído gaussiano no sinal original antes de prosseguir para a EMD, e itera-se o processo da média das IMFs para eliminar a mixagem de modos (WU e HUANG, 2005).

Na Figura 3.5 observa-se o método EMD e EEMD aplicado a um sinal. Na parte (a) temos que na primeira IMF do resultado da EMD, ondas senoidais de diferentes frequências foram misturadas. Devido à presença do modo de mixagem em baixas IMFs, as IMFs altas também não conseguiram decompor dois sinais sinusoidais diferentes. Já na parte (b) a EEMD separou com sucesso a porção de alta frequência nas IMFs baixas, de modo que os sinais de alta frequência e de baixa frequência foram claramente isolados na IMF5 e IMF7, respectivamente.

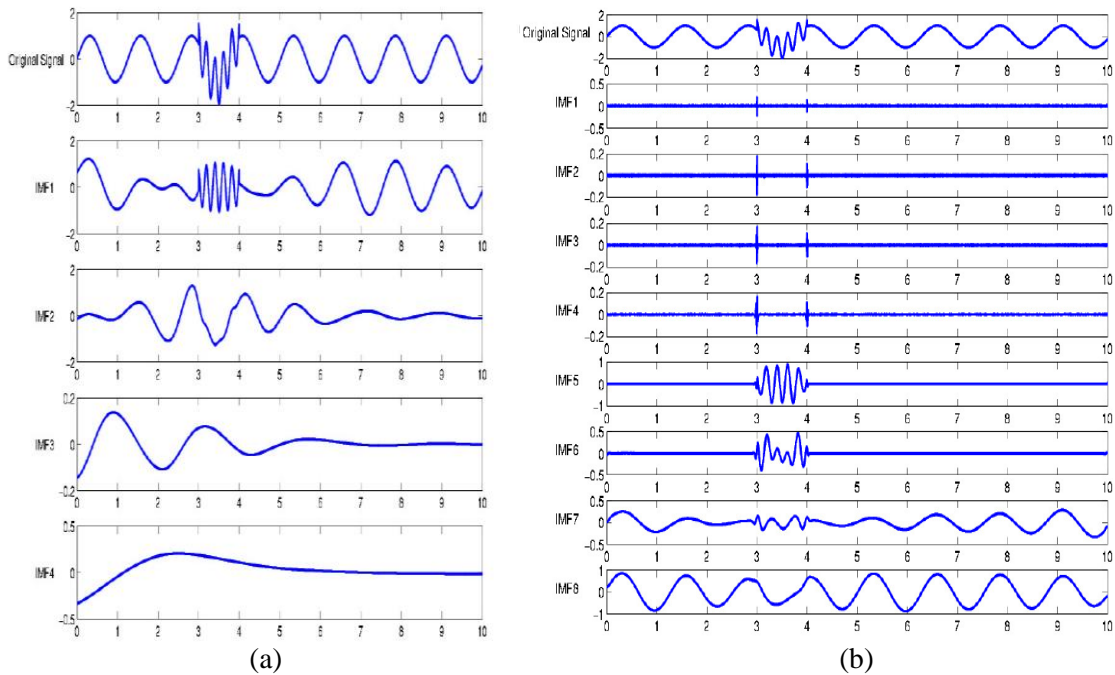


Figura 3.5 – Aplicação da (a) EMD e (b) EEMD a um sinal (Lee *et al.*, 2017).

Sendo assim, neste trabalho aplica-se o EEMD ao sinal de vibração do rolamento. Conforme Wu e Huang (2005) ao aplicar o método EEMD as etapas seguintes são necessárias:

Etapa 1: Na n ésima tentativa, uma nova série temporal é gerada pela adição de uma série temporal de ruído branco $u_n(t)$ a um determinado sinal $x(t)$ $Y_n(t) = X(t) + u_n(t)$, para $n = 1, 2, \dots, N$, sendo N o número de conjunto.

Etapa 2: Baseado no EMD original, o sinal contaminado por ruído $Y_n(t)$ é decomposto em um conjunto de IMFs e um resíduo, conforme Equação 3.11.

$$Y_n(t) = \sum_{m=1}^{M-1} IMF_m^{(n)}(t) + r_M^{(n)}(t) \quad (3.11)$$

Onde $M-1$ é o número total de IMFs resultando em cada decomposição de $Y_n(t)$, e $IMF_m^{(n)}$ é a m^{th} IMF e $r_M^{(n)}$ é o resíduo obtido na n^{th} tentativa.

Etapa 3: As etapas 1 e 2 são reiterados para N tentativas. Em cada tentativa, uma série diferente de ruído branco $u_n(t)$ é adicionada ao sinal original.

Etapa 4. A IMF final da EEMD (IMF_m^{ave}) é obtida pela média do total m de IMF relacionado à N tentativas, obtendo assim a Equação 3.12.

$$IMF_m^{ave}(t) = \frac{1}{N} \sum_{n=1}^N IMF_m^{(n)}(t) \quad (3.12)$$

Os resultados obtidos pela EEMD dependem da escolha do número de conjunto (N) e da amplitude do ruído adicionado (A). Wu e Huang (2005) mostram que a seguinte relação deve ser satisfeita:

$$\varepsilon = \frac{A}{\sqrt{N}} \quad (3.13)$$

Sendo ε o desvio padrão final do erro, calculado como a diferença entre o sinal original e a soma das IMFs resultantes da EEMD.

Conforme Wang *et al.* (2014), o número de conjunto (N) deve se aproximar do infinito para suavizar o ruído branco assistido. E que na prática, a faixa de ruído escolhida é em torno da faixa de 0,1 – 0,4, e o número de conjunto (N) na ordem de 100 geralmente produz resultados satisfatórios e torna o ruído residual inferior a uma fração de 1% de erro.

3.3. Autoregressive modeling – AR

Neste trabalho obtém-se um modelo matemático, em sinais temporais das IMFs, capaz de representar o comportamento do rolamento em falha, através de dados de vibração do sistema mecânico, e em seguida realiza-se a modelagem autorregressiva nas IMFs encontradas. O termo autoregressão indica que é uma regressão da variável contra ela mesma.

Freitas (2007) afirma que, na prática, sabe-se que a maioria das séries temporais resulta de experiências que não poderão ser repetidas e geralmente ficam melhor representadas por processos estocásticos. Um processo estocástico ergódico e estacionário pode ser representado por um modelo projetado a partir de uma amostra apropriadamente coletada. Caso contrário, uma amostra somente não é suficiente para generalizar a respeito do processo. Assim, a análise preliminar das bases de dados possibilita assumir hipóteses necessárias para a especificação e a implementação de modelos mais eficazes. Os modelos lineares e estacionários de uma série no tempo têm duas características particularmente desejáveis: podem ser mais facilmente compreendidos e existem métodos bem direcionados para implementá-los. A contrapartida para estas conveniências é que podem ser inteiramente

impróprios para sistemas com moderadas complicações. Entretanto, estes métodos ainda são bastante utilizados na predição de séries temporais.

Neste trabalho é abordado o modelo paramétrico para séries temporais estacionária AR, modelo em que o número de parâmetros é finito e a análise é realizada no domínio do tempo.

Para Jachan (2006) uma ampla variedade de estimadores de parâmetros e ordem foram desenvolvidos e aplicados com sucesso em muitos problemas de processamento de sinal. A técnica AR pode ser aplicada à predição, modelagem de correlação, extração de características e a muitos outros problemas de processamento de sinais.

Neste trabalho, as características do AR foram extraídas das IMFs, após aplicação da EEMD aos sinais de vibração. Na análise estatística de séries temporais, dada uma série temporal de dados $y(t)$, a modelagem Autorregressiva (AR) é uma ferramenta para entender e, talvez, prever valores futuros nesta série. O modelo AR é um método de análise de sequência temporal cujos parâmetros compreendem informações importantes da condição do sistema. Um modelo AR preciso pode refletir as características de um sistema dinâmico. Os parâmetros do modelo AR são indicados como muito sensíveis às variações de condição (CHENG, 2008).

Para Atyabi *et al.* (2017) identificar a ordem correta do modelo AR é um amplo desafio. Baixas ordens do modelo representam o sinal de forma insatisfatória, enquanto ordens mais altas aumentam o ruído. Os métodos convencionais para estimar a ordem do modelo incluem o Critério de Informação de Akaike (*Akaike Information Criterion* - AIC), o Critério de Informação Bayesiano (*Bayesian Information Criterion* - BIC) e o Erro de Previsão Final (*Final Prediction Error* - FPE).

O processo do modelo autoregressivo consiste em se explicar o valor presente da série $y(t)$ através de uma função de valores no passado L , $y[t - 1], y[t - 2] + \dots + y[t - L]$. Portanto o modelo autoregressivo pode ser escrito conforme equação 3.14.

$$y(t) = \theta_1 y[t - 1] + \theta_2 y[t - 2] + \dots + \theta_L y[t - L] \quad (3.14)$$

Em que y representa o sinal de saída, $t \in \mathbb{Z}$ é o tempo discreto, L representa o atraso do modelo e $\theta_1, \theta_2, \dots, \theta_L$ são os coeficientes a se ajustar aos dados. Para resolver o modelo utiliza-se sua forma matricial, Equação 3.15.

$$\mathbf{Y} = \mathbf{\Psi}\boldsymbol{\theta} + \boldsymbol{\xi} \quad (3.15)$$

Em que \mathbf{Y} o vetor de observações, que contem os dados de saída do sistema:

$$y[L + 1] = \begin{bmatrix} y[L] & y[L - 1] \\ y[L + 1] & y[L] \\ y[L + 2] & y[L + 1] \\ \vdots & \vdots \end{bmatrix} \quad (3.16)$$

$\mathbf{\Psi}$ é a matriz de regressão

$$\mathbf{\Psi} = (y[t - 1], y[t - 2], \dots, y[t - L]) \quad (3.17)$$

E $\boldsymbol{\theta}$ é o vetor de parâmetros que se quer determinar

$$\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_L)^T \quad (3.18)$$

Para estimar os parâmetros $\boldsymbol{\theta}$ (“theta”) utiliza-se o método de mínimos quadrados, em que é formulado como um produto vetorial de matriz, este método consiste em minimizar o quadrado das diferenças entre os valores observados e seus respectivos valores esperados. A ideia principal do método baseia-se em minimizar o erro quadrático total da amostra.

3.4. Support Vector Machine - SVM

Aprendizado de Máquina (*Machine Learning*) é uma subárea da Inteligência Artificial, onde estuda algoritmos que aprendem com exemplos e experiências, e desenvolve algoritmos que a partir do aprendizado de dados faz a predição. Esta área está associada à estatística, cálculo, álgebra linear e probabilidade. A técnica usada para escrever automaticamente um classificador é chamada de Aprendizado Supervisionado, onde coleta-se dados de treinamento, treina o classificador e faz predições. Apresenta-se neste trabalho o SVM classificador, método supervisionado e categórico de Aprendizado de Máquina (*Machine Learning*) desenvolvido por Vapnik (1995) em seu livro sobre a Teoria da Aprendizagem Estatística com ênfase em SVM, para classificação binária. A Máquina de Vetor de Suporte é uma ferramenta poderosa para classificação binária, capaz de gerar rápidas funções do classificador após um período de treinamento (SINGLA *et al.*, 2011). E conforme

Selina *et al.* (2014) cinco dos dez algoritmos populares são projetados para classificação e regressão e a formulação mais intuitiva para o diagnóstico de falhas é a classificação.

De acordo com Ding *et al.* (2013), SVM usa o princípio de minimização de risco estrutural em vez do princípio de minimização de risco empírico, com isto busca o melhor compromisso entre a complexidade do modelo e a habilidade com informações de amostra limitada.

As principais ideias são que, ao assumir classes linearmente separáveis criam-se hiperplanos de separação com margem máxima. E para casos não linearmente separáveis, os pontos de dados são projetados, geralmente em um espaço de maior dimensão onde os pontos de dados efetivamente se tornam linearmente separáveis (essa projeção é realizada através de técnicas *Kernel*) (MEYER, 2017). A utilização deste classificador vem crescendo nos últimos anos devido aos resultados obtidos na maioria das vezes ser superiores aos comparados à outras técnicas de aprendizado. Algumas das aplicações são: identificação de rostos, de objetos em figuras, detectar fraudes nas transações bancárias, detectar se um e-mail é spam ou não.

Neste trabalho é abordado o aprendizado supervisionado, algoritmo para classificar um conjunto de dados binários linearmente separáveis.

Ao explanar os tópicos seguintes, apresentam-se as SVMs como lineares e não lineares. Nas SVM lineares há a eficácia na classificação de dados linearmente separáveis, na sua primeira formulação, têm-se as margens rígidas (que definem fronteiras lineares a partir de dados linearmente separáveis) e na segunda formulação há margens suaves. Neste caso as margens rígidas são estendidas para lidar com um conjunto de treinamento mais geral e que tolera a presença de alguns ruídos e *outliers*. SVM não lineares ocorrem em casos em que não é possível dividir satisfatoriamente os dados de treinamento por um hiperplano.

Na Figura 3.6 é apresentado o princípio do SVM no plano bidimensional. Há a classificação de uma série de pontos para duas diferentes classes, classe A (pontos pretos) e classe B (pontos azuis). O objetivo do SVM é desenvolver um hiperplano de separação ideal que classifica todos os vetores de treinamento em duas classes, maximizando a margem entre os pontos mais próximos das classes (SCHÖLKOPF e SMOLA, 2002). Esses pontos situados nos limites da margem são chamados de Vetores de Suporte, e o meio da margem é o hiperplano de separação ideal, conforme Figura 3.6.

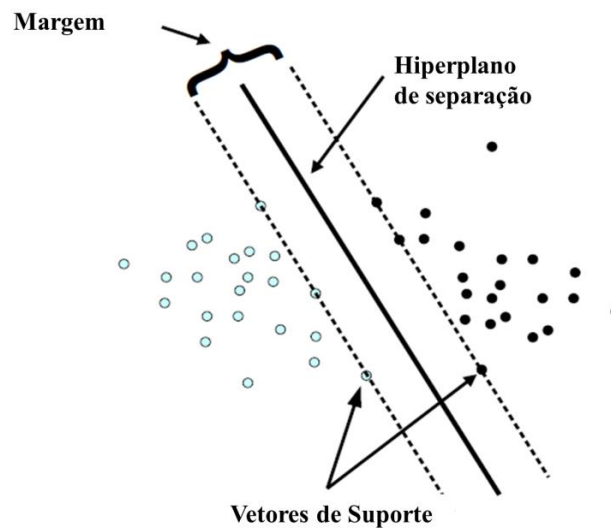


Figura 3.6 – Classificação linearmente separável (Meyer, 2017).

3.4.1. SVM lineares com margens rígidas

A seguir é apresentada a formulação do SVM com margens rígidas para problemas linearmente separáveis de dados pertencentes a duas classes. Considerando um conjunto de treinamento, T , com n dados $x_i \in X$ ($X =$ espaço dos dados) e seus respectivos rótulos $y_i \in Y$ ($Y = -1, +1$). T é linearmente separável se for possível separar os dados das classes $+1$ e -1 por um hiperplano (SCHÖLKOPF e SMOLA, 2002).

O hiperplano é definido pela equação 3.19:

$$f(x) = \omega \cdot x + b = 0 \quad (3.19)$$

Onde ω é o vetor peso, $\omega \cdot x$ é o produto escalar entre os dois vetores e b é a bias (CHENG *et al.*, 2008).

A equação 3.19 divide o espaço de dados em duas regiões: $\omega \cdot x + b > 0$ e $\omega \cdot x + b < 0$. Conforme Smola *et al.* (2000) para produzir a classificação final, uma função sinal $g(x) = \text{sign}(f(x))$ pode então ser empregada, conforme Equação 3.20

$$g(x) = \text{sign}(f(x)) = \begin{cases} +1 & \text{se } \omega \cdot x + b > 0 \\ -1 & \text{se } \omega \cdot x + b < 0 \end{cases} \quad (3.20)$$

A partir de $f(x)$, é possível obter um número infinito de hiperplanos equivalentes pela multiplicação de ω e b por uma mesma constante (PASSERINI, 2004). Define-se o

hiperplano canônico em relação ao conjunto T como aquele em que ω e b são escalados de forma que os exemplos mais próximos ao hiperplano $\omega \cdot x + b = 0$ satisfaçam a Equação 3.21 (MÜLLER *et al.*, 2001).

$$|\omega \cdot x_i + b| = 1 \quad (3.21)$$

Sendo assim implica nas inequações 3.22, apresentadas na expressão da Equação 3.23.

$$\begin{cases} \omega \cdot x_i + b \geq +1 \text{ se } y_i = +1 \\ \omega \cdot x_i + b \leq -1 \text{ se } y_i = -1 \end{cases} \quad (3.22)$$

$$y_i(\omega \cdot x_i + b) - 1 \geq 0, \quad \forall (x_i, y_i) \in T \quad (3.23)$$

Conforme Figura 3.7, sendo x_1 um ponto no hiperplano $H_1: \omega \cdot x + b = +1$ e x_2 um ponto no hiperplano $H_2: \omega \cdot x + b = -1$, ao projetar-se $x_1 - x_2$ na direção de ω , perpendicular ao hiperplano separador $\omega \cdot x + b = 0$, é possível obter a distância entre os hiperplanos H_1 e H_2 (CAMPBELL, 2000). Projeção apresentada na Equação 3.24.

$$x_1 - x_2 \left(\frac{\omega}{\|\omega\|} \cdot \frac{(x_1 - x_2)}{\|x_1 - x_2\|} \right) \quad (3.24)$$

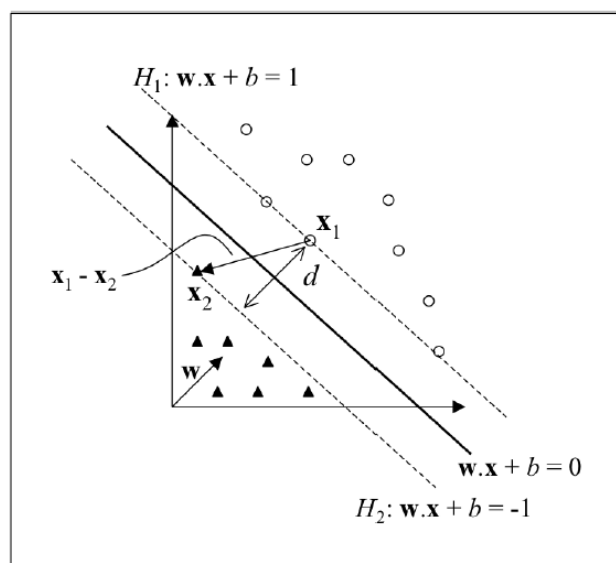


Figura 3.7 – Cálculo da distância d entre os hiperplanos H_1 e H_2 (Lorena e Carvalho, 2007).

Portanto tem-se que $\omega \cdot x_1 + b = +1$ e $\omega \cdot x_2 + b = -1$. Conforme Hearst *et al.* (1998) a diferença entre essas equações fornece $\omega \cdot x_1 - x_2 = 2$ e ao substituir esse resultado na Equação 3.24, obtém-se a Equação 3.25.

$$\left(\frac{2(x_1 - x_2)}{\|\omega\| \|x_1 - x_2\|} \right) \quad (3.25)$$

Para obter o comprimento do vetor projetado, toma-se a norma da Equação 3.25, obtendo assim $\frac{2}{\|\omega\|}$. Na Figura 3.7 pode-se visualizar distância d entre os hiperplanos H_1 e H_2 , paralelos ao hiperplano separador. Como ω e b foram escalados de forma a não haver exemplos entre H_1 e H_2 , $\frac{1}{\|\omega\|}$ é a distância mínima entre o hiperplano separador e os dados de treinamento. E de acordo com Campbell (2000), essa distância é definida como a margem geométrica do classificador linear.

Posterior a essas considerações, verifica-se que a maximização da margem de separação dos dados em relação à $\omega \cdot x + b = 0$ pode ser obtida pela minimização de $\|\omega\|$ (BURGES, 1998). E conforme Schölkopf e Smola (2002), recorre-se ao seguinte problema de otimização.

$$\text{Minimizar } \frac{1}{2} \|\omega\|^2 \quad (3.26)$$

sendo restrito à:

$$y_1(\omega \cdot x_1) + b - 1 \geq 0, \forall_1 = 1, \dots, n \quad (3.27)$$

SVM com margens rígidas é assim definida devido às restrições que são impostas de maneira a assegurar que não haja dados de treinamento entre as margens de separação das classes. Conforme Schölkopf e Smola (2002), o problema de otimização obtido é quadrático. Como a função objetivo sendo minimizada é convexa e os pontos que satisfazem as restrições formam um conjunto convexo, esse problema possui um único mínimo global (PASSERINI, 2004). Problemas desse tipo podem ser solucionados com a introdução de uma função Lagrangiana, que engloba as restrições à função objetivo, associadas a parâmetros denominados multiplicadores de Lagrange α_1 (Equação 3.28) (SCHÖLKOPF e SMOLA, 2002).

$$L(\omega, b, \alpha) = \frac{1}{2} \|\omega\|^2 - \sum_{i=1}^n \alpha_i (y_i (\omega \cdot x_i) + b - 1) \quad (3.28)$$

Conforme Schölkopf e Smola (2002), a função Lagrangiana deve ser minimizada, o que provoca uma maximização das variáveis α_i e em minimização de ω e b . Portanto temos que:

$$\frac{\partial L}{\partial b} = 0 \text{ e } \frac{\partial L}{\partial \omega} = 0 \quad (3.29)$$

Então,

$$\sum_{i=1}^n \alpha_i y_i = 0 \quad (3.30)$$

$$\omega = \sum_{i=1}^n \alpha_i y_i x_i \quad (3.31)$$

Ao substituir as equações 3.30 e 3.31 na equação 3.28, obtém-se o problema de otimização:

$$\underset{\alpha}{\text{Maximizar}} \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \quad (3.32)$$

Sendo as restrições $\alpha_i \geq 0, i = 1, \dots, n$ e $\sum_{i=1}^n \alpha_i y_i = 0$. E de acordo com Lorena e Carvalho (2007), esta formulação é denominada forma dual, enquanto o problema original é referenciado como forma primal. E que o problema dual é formulado apenas ao utilizar os dados de treinamento e os seus rótulos.

Seja α^* a solução do problema dual e ω^* e b^* as soluções da forma primal. Após obter o valor de α^* , ω^* pode ser determinado pela Equação 3.31. O parâmetro b^* é definido por α^* e por condições de Kühn-Tucker, provenientes da teoria de otimização com restrições e que devem ser satisfeitas no ponto ótimo. Para o problema dual formulado, tem-se (PONTIL e VERRI, 1998).

$$\alpha^* (y_i (\omega^* \cdot x_i + b^*)) - 1 = 0, \quad \forall_i = 1, \dots, n \quad (3.33)$$

Observa-se nessa equação que α^* pode ser diferente de 0 somente para os dados que se encontram sobre os hiperplanos H_1 e H_2 . Estes são os exemplos que se situam mais próximos ao hiperplano separador, exatamente sobre as margens. Para os outros casos, a condição apresentada na Equação 3.33 é obedecida apenas com $\alpha_i^* = 0$. Esses pontos não participam então do cálculo de ω^* (Equação 3.31). Os dados que possuem $\alpha_i^* > 0$ são denominados vetores de suporte (*Support Vectors-SVs*) e podem ser considerados os dados mais informativos do conjunto de treinamento, pois somente eles participam na determinação da equação do hiperplano separador.

O valor de b^* é calculado a partir dos SVs e das condições representadas na Equação 3.33. Computa-se a média apresentada na Equação 3.34 sobre todos x_j tal que $\alpha_j^* > 0$, ou seja, todos os SVs. Nessa equação n_{sv} denota o número de SVs, e SV representa o conjunto dos SVs.

$$b^* = \frac{1}{n_{sv}} \sum_{x_j \in SV} \frac{1}{y_j} - \omega^* \cdot x_j \quad (3.34)$$

Ao substituir ω^* pela expressão na equação 3.31, tem-se:

$$b^* = \frac{1}{n_{sv}} \sum_{x_j \in SV} \left(\frac{1}{y_j} - \sum_{x_i \in SV} \alpha_i^* y_i x_i \cdot x_j \right) \quad (3.35)$$

Então tem-se o classificador $g(\mathbf{x})$ apresentado na Equação 3.36, em que sgn representa a função sinal, ω^* é calculado pela Equação 3.31 e b^* pela Equação 3.35.

$$g(\mathbf{x}) = \text{sgn}(f(x)) = \text{sgn} \left(\sum_{x_i \in SV} y_i \alpha_i^* \mathbf{x}_i \cdot \mathbf{x} + b^* \right) \quad (3.36)$$

A função linear apresentada na equação 3.36 representa o hiperplano considerado com melhor capacidade de generalização, conforme Teoria do Aprendizado Estatístico- TAE, pois este separa os dados com maior margem (LORENA e CARVALHO, 2007).

3.4.2. SVM lineares com margens suaves

Na prática dificilmente encontra-se aplicações para dados linearmente separáveis, devido à diversos fatores, entre eles a presença de ruídos e *outliers* nos dados, ou à própria natureza do problema, que pode ser não linear. Para tanto, utiliza-se margens rígidas estendidas para lidar com um conjunto de treinamento mais geral, que permite assim modificações na restrição da Equação 3.27 (SCHÖLKOPF e SMOLA, 2002). Há então a introdução de variáveis de folga ξ_i , para todo $i = 1, \dots, n$, que suavizam as restrições impostas ao problema de otimização primal, e torna-se:

$$y_i(\omega \cdot x_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0, \quad \forall_i = 1, \dots, n \quad (3.37)$$

A aplicação desse procedimento suaviza as margens do classificador linear, permite que alguns dados permaneçam entre os hiperplanos H_1 e H_2 e também a ocorrência de alguns erros de classificação.

Um erro no conjunto de treinamento é indicado por um valor de ξ_i maior que 1. E a soma dos ξ_i representa um limite no número de erros de treinamento. Para levar em consideração esse termo, minimiza-se assim o erro sobre os dados de treinamento, a Equação 3.26 é reformulada como (BURGES, 1998):

$$\underset{\omega, b, \xi}{\text{Minimizar}} \frac{1}{2} \|\mathbf{w}\|^2 + C \left(\sum_{i=1}^n \xi_i \right) \quad (3.38)$$

Conforme Passerini (2004) a constante C é um termo de regularização que impõe um peso à minimização dos erros no conjunto de treinamento em relação à minimização da complexidade do modelo. A presença do termo $\sum_{i=1}^n \xi_i$ no problema de otimização também pode ser vista como uma minimização de erros marginais, pois um valor de $\xi_i \in (0,1)$ indica

um dado entre as margens. Tem-se então uma formulação de acordo com os princípios da Teoria do Aprendizado Estatístico TAE.

O problema de otimização gerado é quadrático, conforme restrições lineares apresentadas na Equação 3.32. E para solucionar introduz uma função Lagrangiana e torna suas derivadas parciais nulas. Então:

$$\text{Maximizar}_{\alpha} \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j X_i \cdot X_j \quad (3.39)$$

$$\text{Sendo as restrições} \begin{cases} 0 \leq \alpha_i \leq C, \forall i=1, \dots, n \\ \sum_{i=1}^n \alpha_i y_i = 0 \end{cases}$$

Pode-se observar que essa formulação é igual à apresentada para as SVMs de margens rígidas, a não ser pela restrição nos α_i , que agora são limitados pelo valor de C .

Seja α^* a solução do problema dual, enquanto ω^* , b^* e ξ^* denotam as soluções da forma primal. O vetor ω^* continua sendo determinado pela Equação 3.31 e as variáveis de folga ξ_i^* podem ser calculadas pela Equação 3.45 (CRISTIANINI e SHAW-TAYLOR, 2000).

$$\xi_i^* = \max \left\{ 0, 1 - y_i \sum_{j=1}^n y_j \alpha_j^* x_j \cdot x_i + b^* \right\} \quad (3.40)$$

Conforme Pontil e Verri (1998) a variável b^* provém α^* e de condições de Kuhn-Tucker:

$$\alpha_i^* (y_i (\omega^* \cdot x_i + b^*) - 1 + \xi_i^*) = 0 \quad (3.41)$$

e

$$(C - \alpha_i^*) \xi_i^* = 0 \quad (3.42)$$

3.4.3. SVM não linear

A limitação do SVM linear é que ele é aplicável diretamente somente para problemas binários (duas classes). Para utilizá-lo em problemas não lineares, deve ser feita uma redução para diversos problemas binários.

Hoshyar *et al.* (2018) afirmam que em 1992 Guyon, Boser e Vapnik apresentaram uma abordagem para gerar classificadores não lineares que utiliza funções *Kernel* para casos de separação não linear do espaço de entrada original. Primeiramente o SVM transforma os dados originais em um espaço de atributos (*feature space*) de dimensão maior. Esta transformação pode ser obtida através de diferentes mapeamentos não lineares e a função *Kernel* $K(x, x_i)$ pode ser escolhida para se adequar ao problema. Após essa transformação, o hiperplano ideal pode ser facilmente encontrado. O hiperplano alcançado é o caso ideal em relação a uma margem máxima (HOSHYAR *et al.*, 2018).

A Figura 3.8 apresenta um exemplo de SVM não linear, em que é possível visualizar que no conjunto de dados de entrada a fronteira curva seria mais adequada na separação das classes, porém mais complexa. Por isso a utilização de mapeamento para transformar os dados de entrada em um espaço de atributos (*feature space*) de dimensão maior e assim obter o hiperplano ideal.

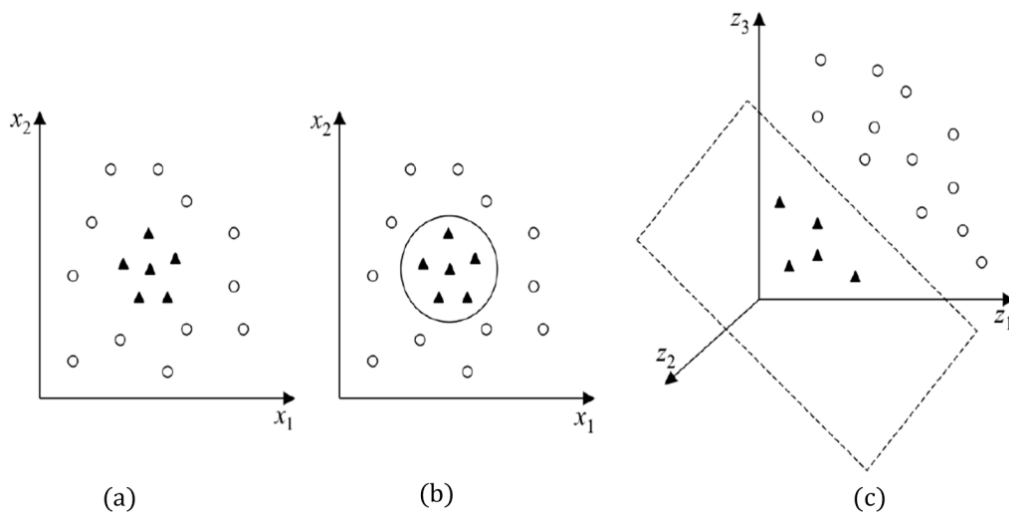


Figura 3.8 – (a) Conjunto de dados não linear; (b) Fronteira não linear nos dados de entrada e (c) Fronteira linear no espaço de atributos. (Lorena e Carvalho, 2007).

A equação 3.28 é a função para decisões lineares, mas não pode ser aplicada a casos não lineares, portanto SVM não linear através de um espaço de dimensão maior resolve este problema ao utilizar o pré-processamento dos dados conforme Equação 3.43 (HASSEIM *et al.*, 2013).

$$x \rightarrow \varphi(x) \quad (3.43)$$

E então

$$\omega \cdot \varphi(x) + b = 0 \quad (3.44)$$

A partir da equação acima o hiperplano ideal não apenas separa corretamente os dois pontos de dados de classe, mas também faz com que a margem (distância do ponto mais próximo ao hiperplano) seja máxima. E ao aplicar a Transformada de Lagrange, a função de classificação ótima é derivada (SINGLA *et al.*, 2011).

$$f(x) = \text{sgn} \left(\sum_{i=1}^l a_i^* y_i K(x, x_i) + b \right) \quad (3.45)$$

Onde:

x_i é o autovetor da amostra de treinamento;

x é o autovetor da amostra de reconhecimento;

a_i^* é o operador Lagrange e

$K(x, x_i) = \varphi(x_i) \cdot \varphi(x)$ é a função *Kernel*.

As funções *Kernel* fornecem um método conveniente para obter as características de alta dimensão mapeadas a partir dos dados sem computar a transformação não linear (SELINA *et al.*, 2014). Elas representam um produto escalar de pontos de dados de entrada mapeados em um espaço de características dimensional superior pela transformação φ . As equações 3.46, 3.47, 3.48 e 3.49 apresentam as funções comuns do *kernel*: linear, sigmoidal, polinomial e de função base radial (RBF) (HASSEIM *et al.*, 2013).

$$\textit{Kernel} \text{ Função base radial (RBF): } K(x, x_i) = \exp(-\sigma \cdot \|x - x_i\|^2) \quad (3.46)$$

$$\textit{Kernel} \text{ polinomial: } K(x, x_i) = (\delta(x \cdot x_i) + k)^d \quad (3.47)$$

$$\textit{Kernel} \text{ linear: } K(x, x_i) = (x \cdot x_i) \quad (3.48)$$

$$\textit{Kernel} \text{ sigmoidal } K(x, x_i) = \tanh(\delta(x \cdot x_i) + k) \quad (3.49)$$

Para resolver problemas de classificação com três ou mais classes, a classificação SVM utiliza as seguintes estratégias: SVMs de classificação multiclases, em que uma função de decisão SVM tenta classificar todas as classes, as mais utilizadas são: estratégia *One-versus-All*, em que constrói classificadores binários que distinguem entre uma das classes e as demais. E estratégia *One-versus-One*, que constrói um classificador para cada par de classes.

4. RESULTADOS E DISCUSSÕES

Para desenvolver este trabalho utilizou-se um banco de dados de rolamentos da *Case Western Reserve University (CWRU)*, que possibilitou a aquisição de sinais de vibração de rolamentos de esferas rígidas de um motor em uma bancada. Os sinais adquiridos foram obtidos em condições diferentes de funcionamento: rolamento em bom estado, com falha na pista interna, falha na pista externa e falha na esfera. Para cada defeito, uma única forma de onda encontra-se disponível. Os sinais são estacionários, porém foram coletados sob diferentes rotações. Para esta análise a pista externa é considerada como estacionária.

4.1. Descrição da bancada

Conforme Smith e Randall (2015) e dados fornecidos pelo banco de dados, a bancada de testes é composta por: um motor de 2 HP *Reliance Electric*, acelerômetro e transdutor de torque. O acelerômetro foi instalado com base magnética na carcaça do motor na posição vertical (12 horas), Figura 4.1. Para a aquisição dos dados de aceleração utilizou-se o *DAT recorder* de 16 canais e os sinais de vibração foram coletados a uma taxa de 12.000 amostras/s. O transdutor de torque foi utilizado para coletar dados de velocidade e potência, mas não é utilizado nessa análise.

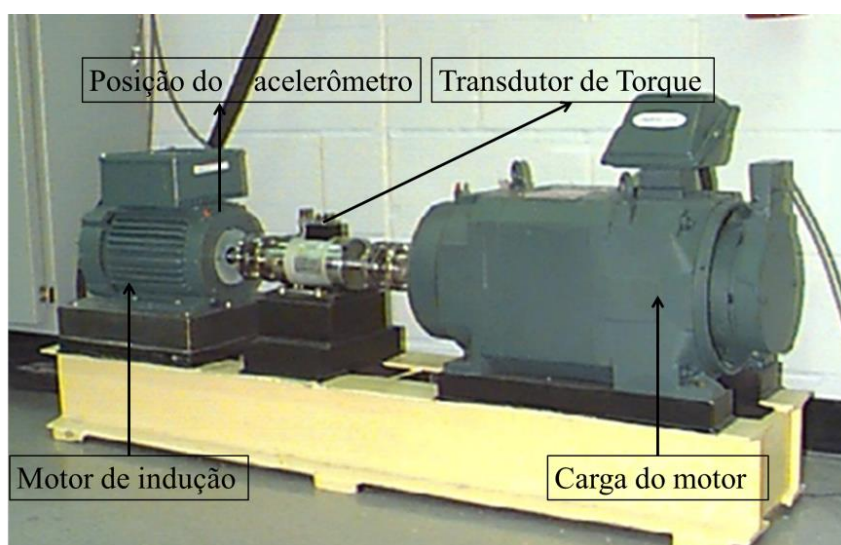


Figura 4.1 – Bancada de teste – Banco de dados *Case Western Reserve University (CWRU)*, (2008).

Os rolamentos de esferas rígidas especificados para este experimento foi o 6205-2RS JEM SKF, cuja especificação é resumida no Quadro 4.1. Os rolamentos suportam o eixo do motor, e considerou-se o rolamento posicionado no lado de acionamento (*drive end bearing*).

Quadro 4.1 – Especificações rolamento SKF 6205-2RS (CWRU, 2008).

Diâmetro interno (mm)	Diâmetro Externo (mm)	Espessura (mm)	Diâmetro das Esferas (mm)	Diâmetro Primitivo (mm)	Número de esferas	Ângulo de contato (°)
25,001	51,9988	15,0012	7,94004	39,0398	9	0

Para inserção das falhas nos rolamentos do motor utilizou-se a usinagem por eletroerosão (*electro-discharge machining – EDM*). As falhas geradas foram introduzidas separadamente na pista interna, pista externa e esferas, e apresentam valores que variam de 0,007" a 0,028" de diâmetro (SMITH e RANDALL, 2015). Conforme literatura, as falhas evoluem gradativamente, portanto avaliaram-se os dados referentes a um diâmetro menor (0,007") e outro diâmetro três vezes maior (0,021"); ambos com profundidade de falha de 0,011". O banco de dados disponibiliza dados de vibração registrados para cargas de 0 a 3 HP, e velocidades de rotação de 1730 a 1797 rpm. Para análise deste trabalho utilizou-se a rotação de 1730 rpm (28,83 Hz) e carga do motor de 3HP.

4.2. Processamento

Os sinais provenientes do banco de dados são relativamente longos (da ordem de dezenas de milhares de amostras cada um) e, por esta razão, os atributos foram calculados ao longo de “janelas”. Janelas provenientes da mesma forma de onda foram consideradas exemplos do mesmo defeito. Foram testadas janelas de diferentes tamanhos, desde 100 até 2000 amostras. Para fins de treinamento e validação, cada conjunto de dados pertencentes a uma classe, com o mesmo número de amostras e janelas escolhidas aleatoriamente – Figura 4.2 – é utilizado para treinar os classificadores enquanto que outro é utilizado para validá-los. E no conjunto de treinamento e validação sequencial, as janelas são escolhidas na sequência temporal, conforme ilustra Figura 4.2. Notar que aumentar o tamanho da janela implica em uma diminuição do número de exemplos, o que tende a produzir um treinamento mais difícil para os classificadores.

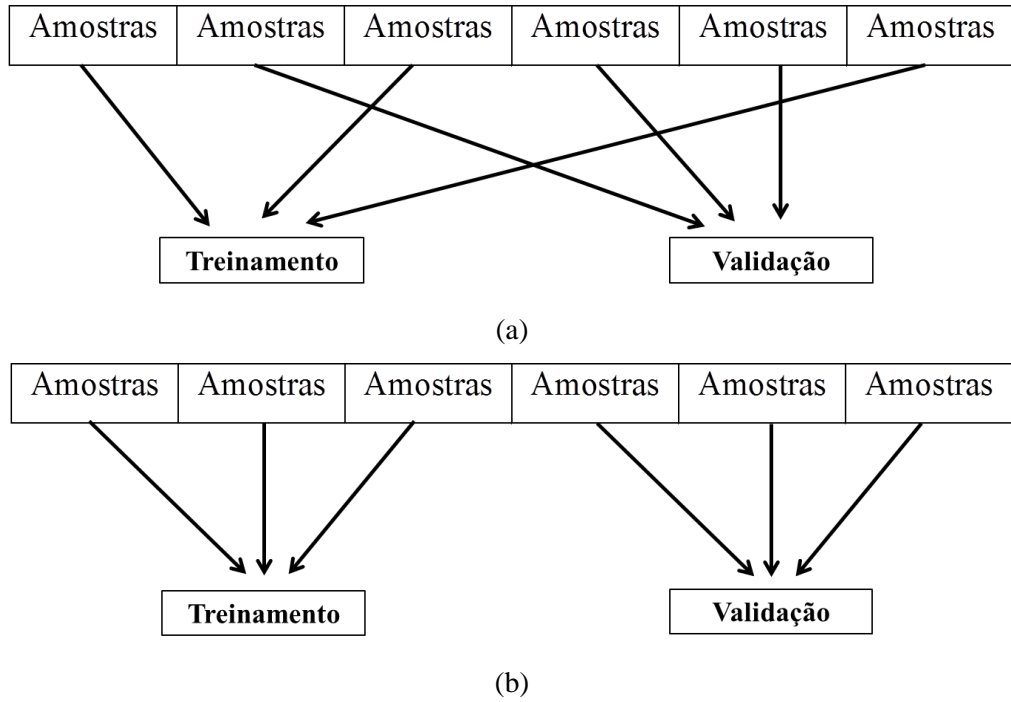
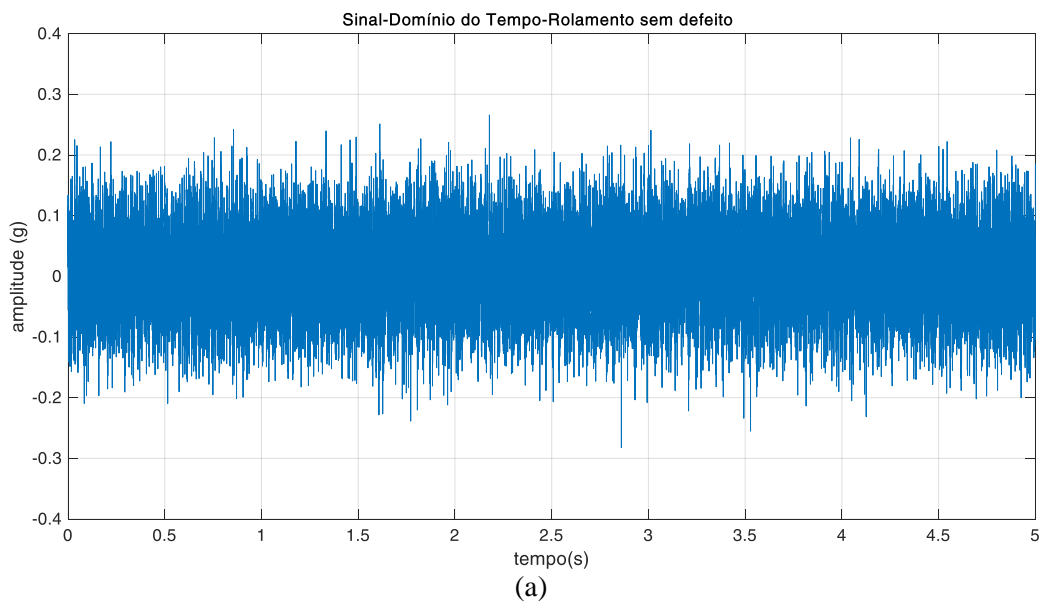


Figura 4.2 – Conjunto de dados de treinamento e validação a) janelas aleatórias b) janelas sequenciais (Autoria própria).

A Figura 4.3 apresenta cada condição do rolamento no domínio do tempo para o diâmetro de defeito de 0,007" – sinal sem defeito, com defeito pista interna, com defeito pista externa e com defeito na esfera.



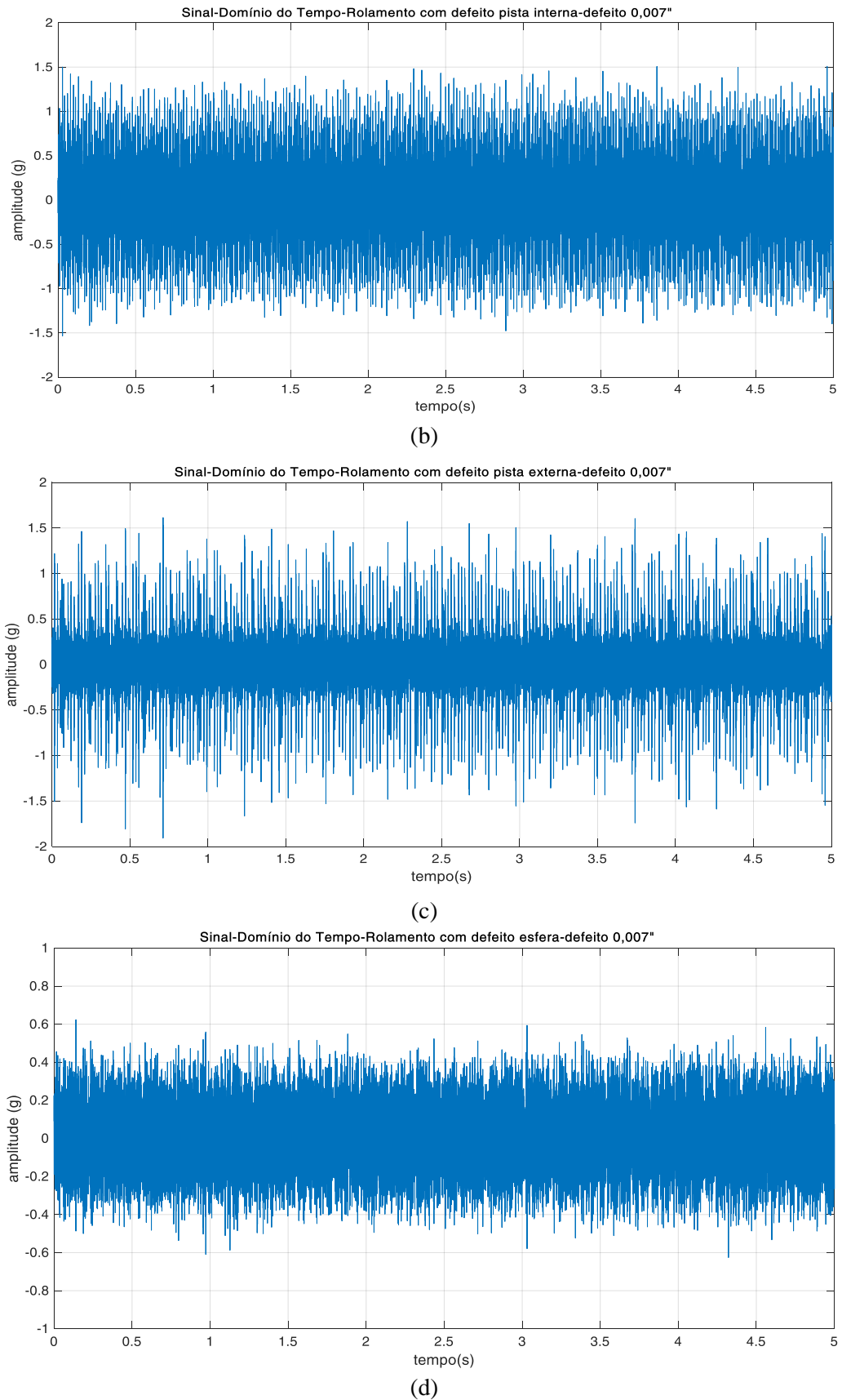


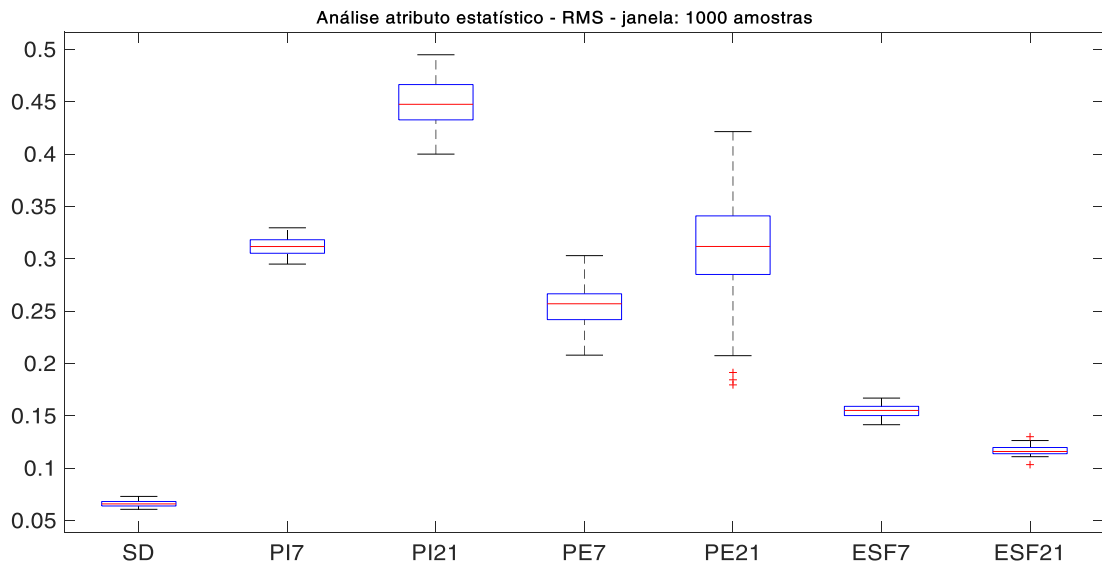
Figura 4.3 – Sinal no domínio do tempo do rolamento: a) sem defeito b) com defeito 0,007" na pista interna c) com defeito 0,007" na pista externa d) com defeito 0,007" na esfera (Autoria própria).

De acordo com a classificação supervisionada as classes são pré-definidas, portanto as classes do problema de diagnóstico de falhas em rolamentos foram estabelecidas conforme Quadro 4.2.

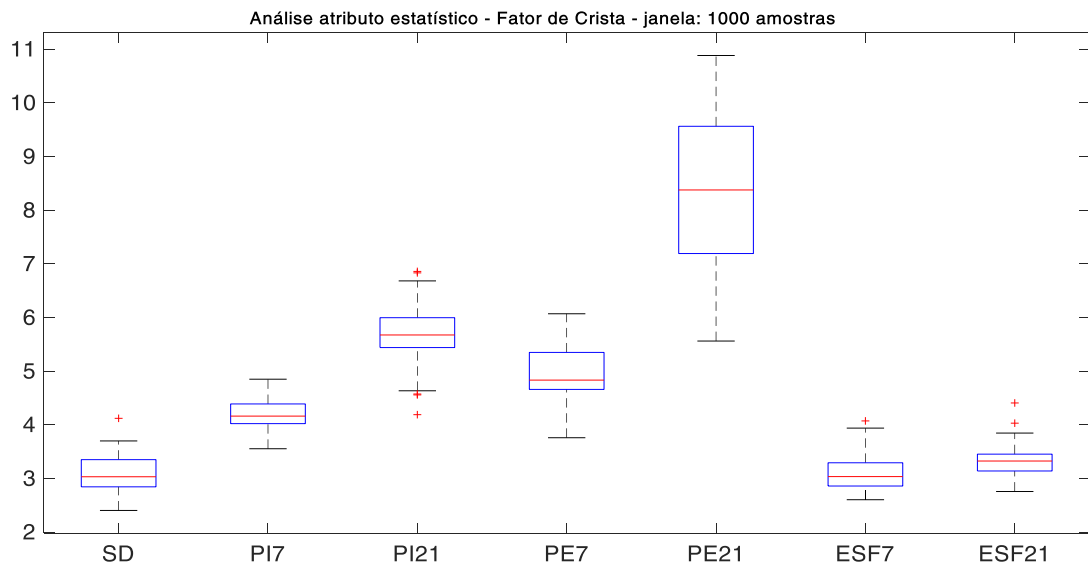
Quadro 4.2 – Classes – condição do rolamento (Autoria própria).

Classe	Condição
1	Sem defeito (SD)
2	Com defeito 0,007" na pista interna (PI7)
3	Com defeito 0,007" na pista externa (PE7)
4	Com defeito 0,007" na esfera (ESF7)
5	Com defeito 0,021" na pista interna (PI21)
6	Com defeito 0,021" na pista externa (PE21)
7	Com defeito 0,021" na esfera (ESF21)

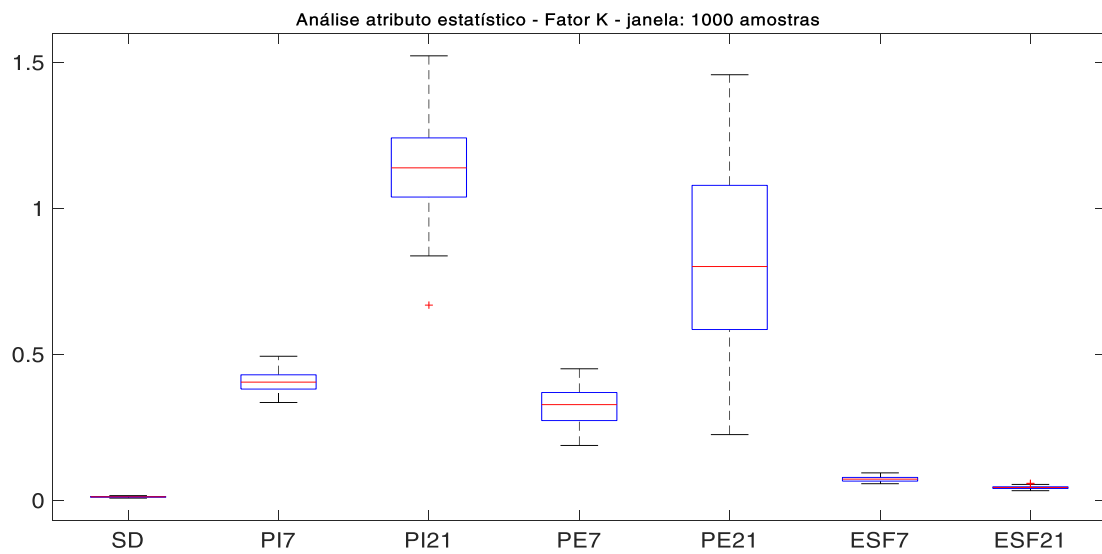
E para cada janela estabelecida, foram calculados os atributos estatísticos – Valor do Nível Global RMS, Fator de Crista, Fator K, Fator Curtose e *Skewness* – conforme equações 3.1, 3.2, 3.3, 3.4 e 3.5. A Figura 4.4 ilustra diagramas *boxplot* de cada um ao utilizar janelas aleatórias de 1000 amostras para exemplificar, separados de acordo com as sete classes estabelecidas. Nestes gráficos é possível perceber a dificuldade em diferenciar alguns dos defeitos utilizando estes atributos. Por exemplo, na Fig. 4.4(a), as medianas das classes PI7 e PE21 são muito próximas, indicando que seria muito difícil diferenciar estes defeitos apenas pelo nível RMS. Já na Fig. 4.4(d) o espalhamento nas medidas de curtose sugere que a probabilidade de se confundir defeitos ao utilizar este atributo deve ser alta. Apesar disso, a variabilidade com cada classe, mostra que, se usados simultaneamente, estes atributos podem ser preditores adequados dos defeitos propostos, desde que utilizados a partir de um algoritmo classificador bem treinado.



(a)



(b)



(c)

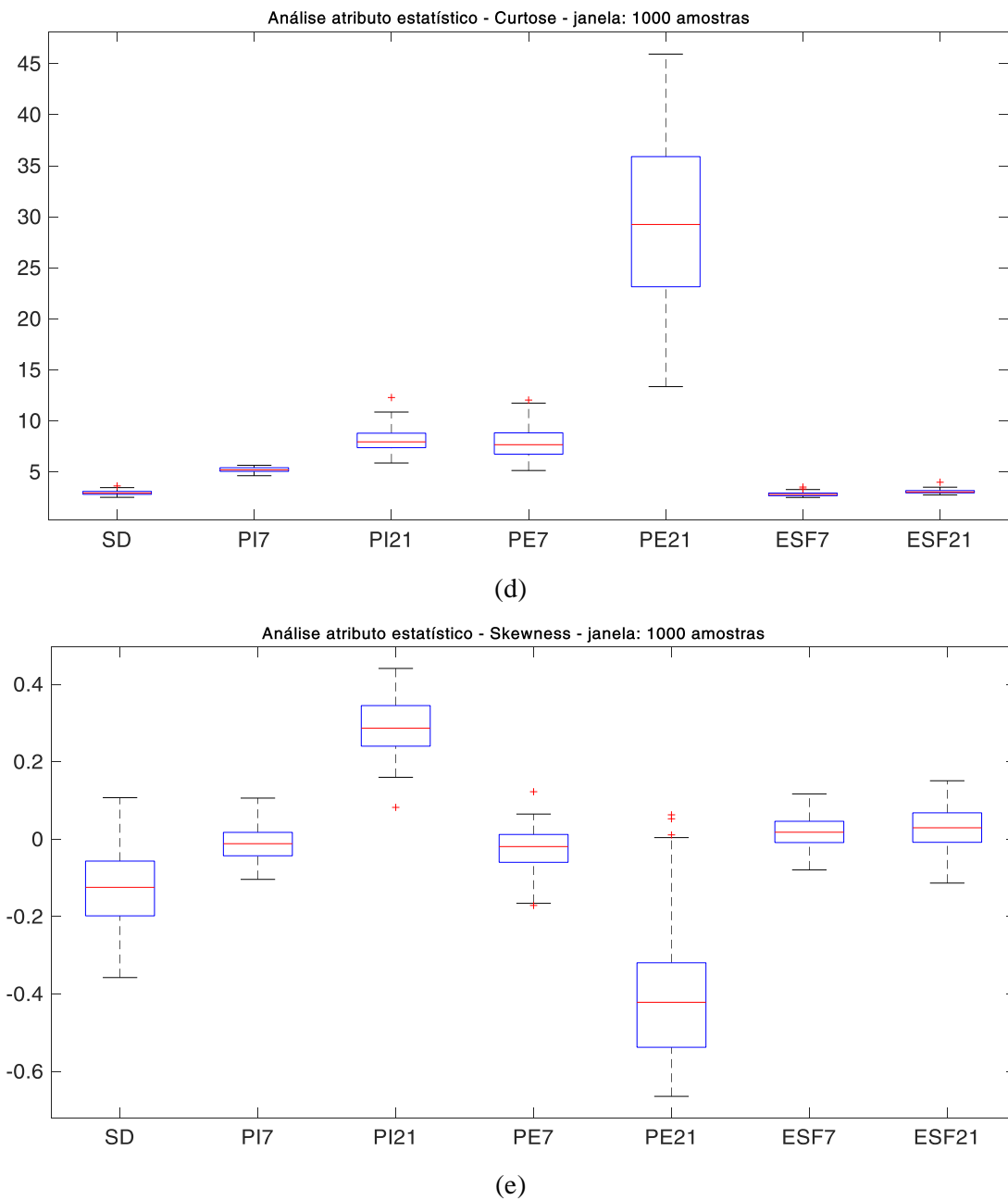


Figura 4.4. Atributos estatísticos – condições do rolamento a)RMS b)Fator de Crista c)Fator K d)Curtose e) *Skewness* (Autoria própria).

Através das medianas observadas na Figura 4.4 é possível perceber que o valor RMS varia consideravelmente quando o rolamento sai da condição normal, mas não permite distinguir entre alguns defeitos. Observando-se o Fator de Crista, verifica-se que a orientação geral de que o defeito pode ser constatado a partir do limiar de 6 nem sempre é válida, considerando que a condição não é satisfeita para alguns casos de defeito na pista interna, externa e esfera de diâmetro 0,007" e na pista interna e esfera de diâmetro 0,021". Analisando apenas o Fator K e sua dispersão de dados, observa-se que os casos de rolamento sem defeito

e com defeito de diâmetros 0,007" e 0,021" na esfera também são ambíguos. Para o valor de Curtose a orientação é que o defeito se verifica para valores maiores que 3, mas ao analisar a dispersão dos valores, não ficam evidentes os defeitos de 0,007" e 0,021" na esfera. Esta análise demonstra que, isoladamente, os atributos estatísticos não permitem distinguir de forma inequívoca entre os defeitos propostos. Contudo, ao considerar a variação expressiva dos atributos diante de cada uma das situações, é natural cogitar que uma análise combinada dos atributos possa resolver as ambiguidades existentes entre as classes. Isto, portanto, justifica uma abordagem automática baseada em aprendizagem de máquina.

4.3. Aplicação SVM com uso de atributos estatísticos e atributos EEMD+AR

Para realizar a classificação automática, foram treinados diversos classificadores SVM, que diferiram em: 1) número de classes inseridas; 2) tamanho da janela sobre a qual os atributos são calculados; 3) combinações de atributos e 4) métodos de processamento aplicados aos sinais (estatístico ou EEMD+AR). Conforme Quadro 4.2, cada classe foi separada em 2 grupos, um para realizar o treinamento e outro para validação, cada grupo contém 30 arquivos com respectivas janelas de amostras. Analisou-se treinamento com janelas escolhidas aleatoriamente e com janelas escolhidas na sequência temporal. Portanto, cada linha representa uma classe da condição do rolamento e cada coluna corresponde aos atributos calculados. Os atributos estatísticos foram organizados em um vetor, segundo a ordem: RMS, Fator de crista, Fator K, Curtose e *Skewness*. E para os atributos obtidos a partir dos coeficientes da modelagem AR, foi utilizado o método EEMD (número de conjuntos $N=2$ e ruído $\varepsilon=0,3$) no processamento para obtenção das IMFs, seguida da aplicação da modelagem Autoregressiva com $L=2$ atrasos; e assim o vetor foi obtido com 4 coeficientes: 2 coeficientes da IMF1 e 2 coeficientes da IMF2.

Após cada classe apresentar suas matrizes de treinamento e validação, realizou-se a junção das matrizes e acrescentou-se uma nova coluna, esta foi destinada às classes que cada grupo pertencia (1 à 7), Quadro 4.2, obtém-se uma nova matriz de 210 linhas e 6 colunas. Para a matriz de validação apenas realizou-se a junção das sete condições do rolamento, sem acrescentar a coluna de classes, pois após o treinamento, o classificador é responsável, através do aprendizado supervisionado, em classificar cada amostra.

Para investigar a influência do número de classes inseridas, do tamanho da janela sobre a qual os atributos são calculados, combinações de atributos e métodos de

processamento aplicados aos sinais, avaliou-se o desempenho dos classificadores em termos de acurácia. Para primeira avaliação considerou-se que a severidade do problema é a mesma e muda-se apenas o tipo de defeito; já na segunda, tanto a severidade quanto o tipo de defeito variam. O resultado do Classificador SVM é mostrado no tópico a seguir.

4.3.1. Atributos estatísticos e atributos EEMD+AR – 4 classes

4.3.1.1. Defeito 0,007" – dados aleatórios e sequenciais

Primeiramente para investigar a influência do tamanho da janela e o número de classes inseridas para treinamento com janelas escolhidas aleatoriamente, utilizou-se apenas o diâmetro do defeito 0,007", sendo assim tem-se apenas 4 classes (sem defeito-SD, com defeito pista interna-PI7, com defeito pista externa-PE7 e com defeito esfera-ESF7) e realizou-se o cálculo para janelas de 100 até 1000 amostras. O resultado é apresentado na Figura 4.5, em que foi possível plotar o desempenho global em distinguir entre as quatro classes (acurácia) em função do tamanho da janela.

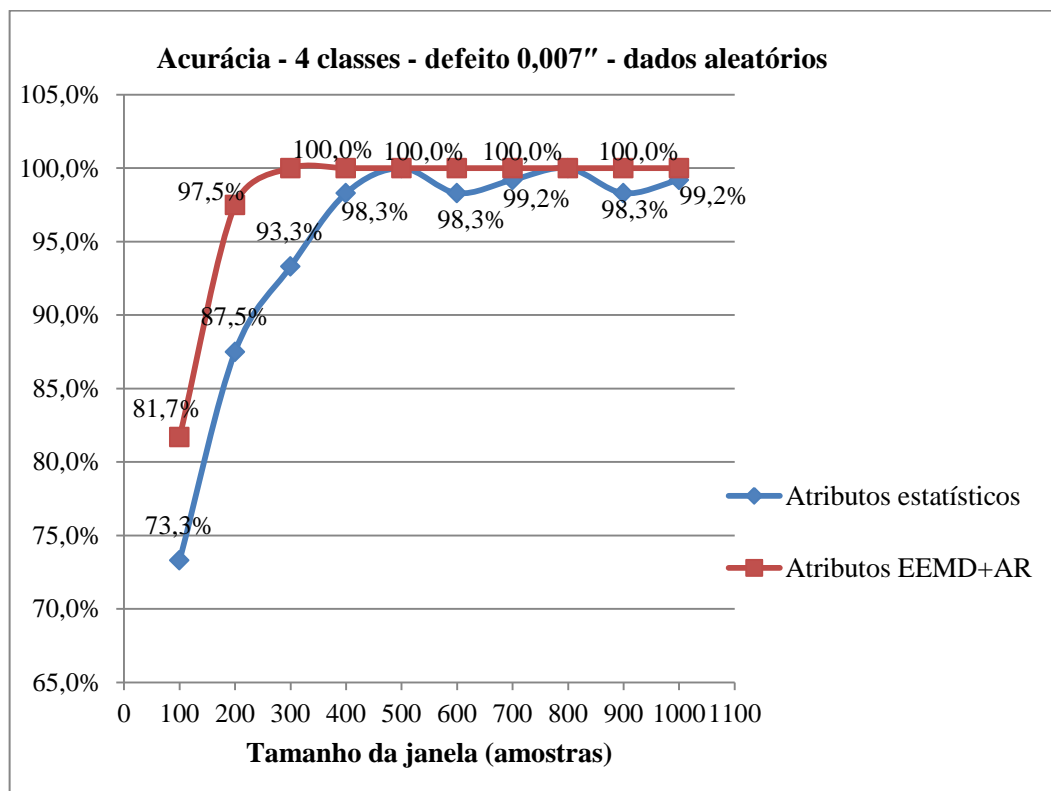


Figura 4.5 – Acurácia – 4 classes – defeito 0,007" – dados aleatórios (Autoria própria)

Ao inserir atributos EEMD+AR, estes se sobressaem aos atributos estatísticos em todos os tamanhos das janelas avaliados, e apresenta máximo desempenho do classificador a partir da janela de 300 amostras e mantêm-se até a janela de 1000 amostras. Resultado coerente, pois foi utilizado apenas um tamanho de diâmetro e, portanto o classificador tem mais facilidade em distinguir as classes. Outro ponto a ser considerado é que, mesmo em janelas com poucas amostras, o classificador apresenta boa categorização.

A Tabela 4.1 ilustra a matriz confusão para o classificador treinado com os cinco atributos estatísticos, calculado para janela de 200 amostras. E observa-se que o classificador realizou 15 erros entre as classes PI7 e PE7 (defeito pista interna 0,007" e defeito pista externa 0,007"). Três amostras pertencentes à classe PI7 foram categorizadas como classe PE7 (defeito pista externa 0,007"), portanto houve dificuldade em distinguir a localização do defeito que estava na pista interna, mas foi detectado na pista externa. É possível que isto seja devido ao fato de defeitos pequenos produzirem formas de onda cujas características podem ser facilmente confundidas. Já na classe PE7 (defeito pista externa), 7 amostras foram reconhecidas como classe PI7 (defeito pista interna) e 5 amostras foram categorizadas como classe ESF7 (defeito esfera). Ou seja, o classificador teve dificuldades em distinguir a localização do defeito.

Tabela 4.1 – Matriz confusão defeito 0,007" – Atributos estatísticos, dados aleatórios – janela 200 amostras (Autoria própria).

		Classe verdadeira			
		SD	PI7	PE7	ESF7
Previsão	SD	30 25%	0	0	0
	PI7	0	27 22,5%	7	0
	PE7	0	3	18 15%	0
	ESF7	0	0	5	30 25%

Observa-se que houve erros tanto em distinguir defeitos de pistas como esfera. Isto pode ser devido à distribuição dos dados, como pode ser visto na Figura 4.6, em que é possível analisar o gráfico de espalhamento. Nele, cada exemplo do conjunto de validação

é ilustrado em termos de apenas duas coordenadas. Foram utilizados os atributos RMS e Fator de Crista como coordenadas principais. Conforme ilustrado (observar que as linhas desenhadas não representam o classificador, e servem apenas de ilustração), o classificador encontrou dificuldade para distinguir as amostras das classes PI7(2), PE7(3) e ESF7(4).

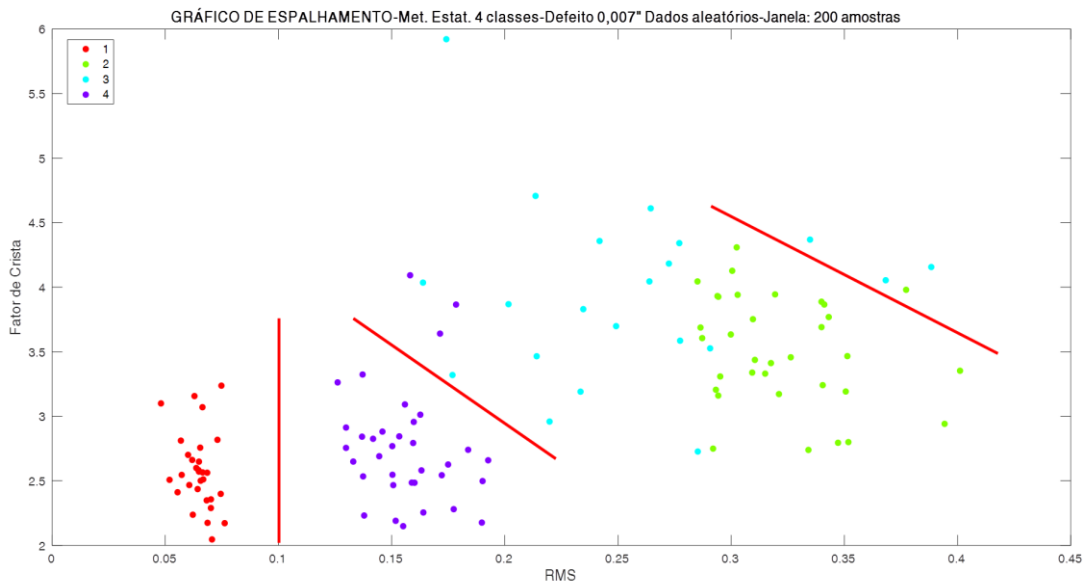


Figura 4.6 – Gráfico de espalhamento defeito 0,007" – Atributos estatísticos, dados aleatórios – janela 200 amostras (Autoria própria).

Para obtenção dos atributos EEMD+AR, primeiramente é necessário aplicar o método EEMD ao sinal de vibração do banco de dados, para obtenção das duas IMFs; seguida então da aplicação da modelagem AR em cada IMF, para obtenção de dois coeficientes cada uma. Portanto o vetor desta análise é composto por quatro atributos, organizados segundo a ordem: "theta1", "theta2", "theta3" e "theta4". Para visualização das IMFs, a Figura 4.7 ilustra duas IMFs do sinal com defeito na pista interna 0,007" para dados aleatórios na janela de 200 amostras.

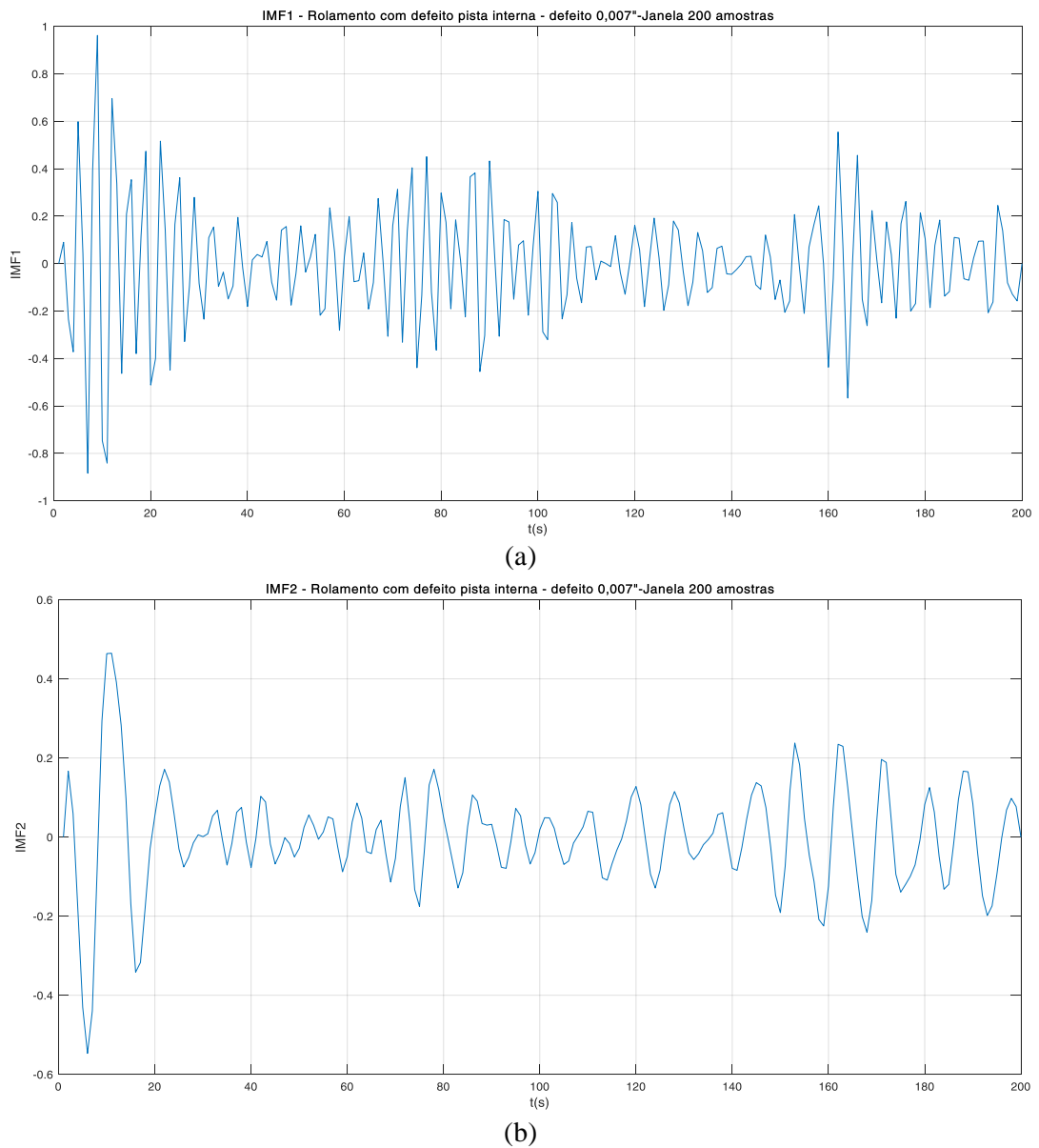


Figura 4.7 – IMFs rolamento com defeito pista interna 0,007", dados aleatórios – janela 200 amostras a) IMF1 b) IMF2 (Autoria própria).

A matriz confusão para o classificador treinado com os atributos EEMD+AR, também para janela de 200 amostras, é apresentada a seguir na Tabela 4.2. E verificou-se melhor desempenho de acurácia comparado aos atributos estatísticos, pois houve apenas 3 erros entre os dados de validação.

Tabela 4.2 – Matriz confusão defeito 0,007” – Atributos EEMD+AR, dados aleatórios – janela 200 amostras (Autoria própria).

		Classe verdadeira			
		SD	PI7	PE7	ESF7
Previsão	SD	30 25%	0	0	0
	PI7	0	29 24,2%	1	0
	PE7	0	1	29 24,2%	1
	ESF7	0	0	0	29 24,2%

Uma amostra pertencente a classe PI7 foi reconhecida como classe PE7 (defeito pista externa); uma amostra da classe PE7 foi categorizada como classe PI7 (defeito pista interna); e uma amostra da classe ESF7 (defeito esfera) foi categorizada como classe PE7 (defeito pista externa). Observa-se que os 3 erros de defeitos de pistas e esfera, é devido a distribuição dos dados, e uma forma de ilustrar é o gráfico de espalhamento em função de apenas duas coordenadas arbitrárias (coeficientes “theta 1” e “theta 2”), Figura 4.8. A separação entre as classes é mais nítida, sendo possível isolá-las linearmente, conforme ilustrado (observar que as linhas desenhadas servem apenas de ilustração).

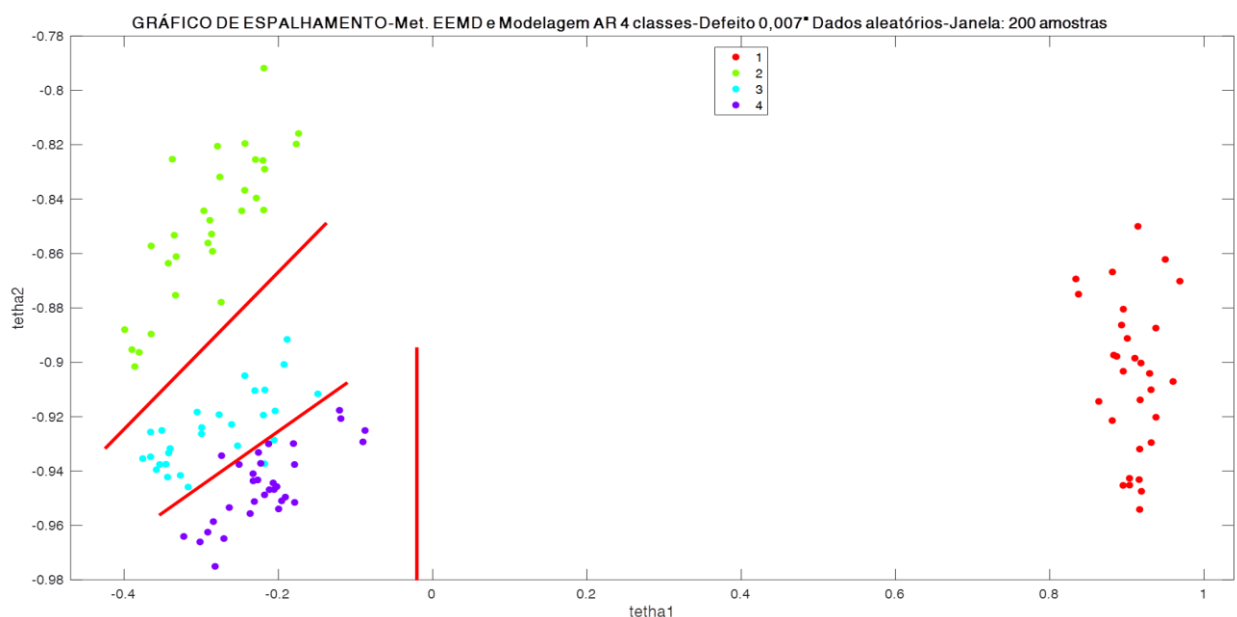


Figura 4.8. Gráfico de espalhamento defeito 0,007” – Atributos EEMD+AR, dados aleatórios – janela 200 amostras (Autoria própria).

Ao utilizar o mesmo tamanho de defeito (0,007"), porém aplicando o treinamento com janelas escolhidas na sequência temporal – Figura 4.9 – observa-se também que, os atributos EEMD+AR se sobressaem aos atributos estatísticos em todas as janelas, e que os atributos estatísticos apresentam máximo desempenho a partir da janela de 700 amostras até 1000 amostras. Interessante que nesta análise, os atributos estatísticos apresentaram desempenho máximo apenas em janelas maiores.

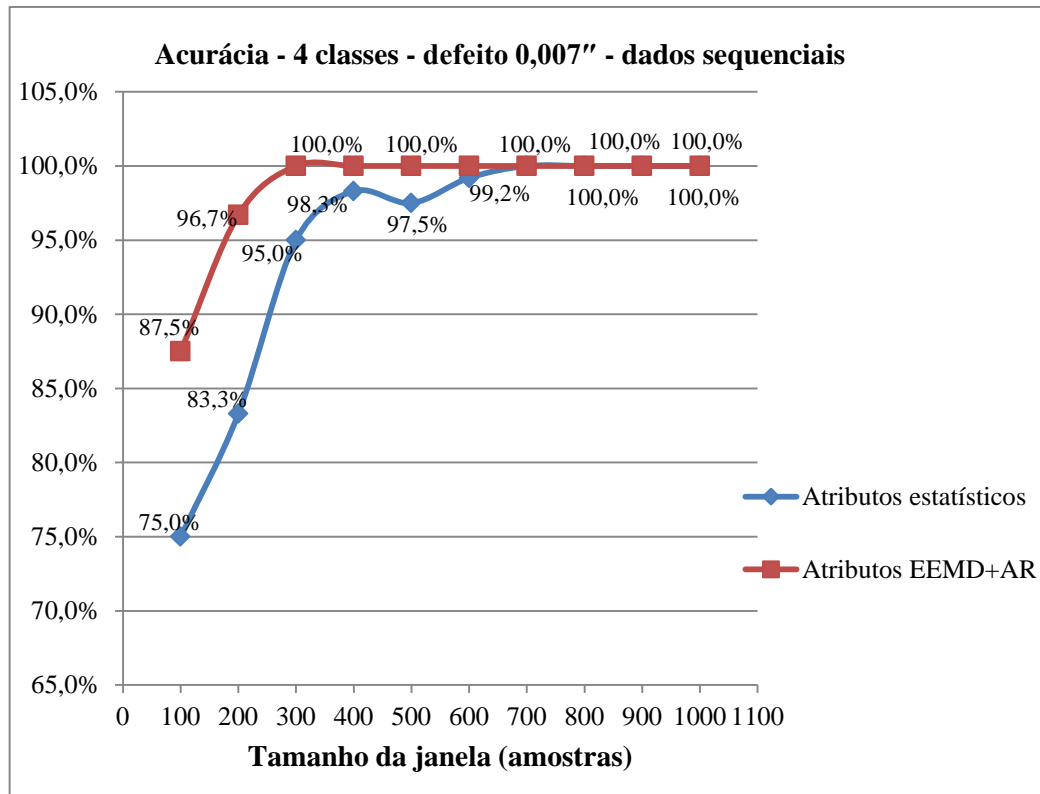


Figura 4.9 – Acurácia – 4 classes – defeito 0,007" – dados sequenciais (Autoria própria).

A Tabela 4.3 apresenta a matriz confusão dos cinco atributos estatísticos, calculado para janela de 200 amostras. Nesta análise houve 20 erros entre as classes com defeitos na pista e na esfera. Analisando a matriz verificou-se que 6 amostras pertencentes à classe PI7 foi categorizada como classe PE7 (defeito pista externa 0,007"). E na classe PE7, 8 amostras foram reconhecidas como classe PI7 (defeito pista interna 0,007") e 5 amostras foram categorizadas como classe ESF7 (defeito esfera), ou seja, o classificador teve dificuldades em distinguir a localização do defeito. E na classe ESF7, uma amostra foi categorizada como PE7 (defeito pista externa).

Tabela 4.3 – Matriz confusão defeito 0,007" – Atributos estatísticos, dados sequenciais – janela 200 amostras (Autoria própria).

		Classe verdadeira			
		SD	PI7	PE7	ESF7
Previsão	SD	30 25%	0	0	0
	PI7	0	24 20%	8	0
	PE7	0	6	17 14,2%	1
	ESF7	0	0	5	29 24,2%

É visível a dificuldade em distinguir a localização do defeito de pista interna que foi categorizado como defeito pista externa, e na esfera que foi reconhecida como defeito pista externa, fato que pode ser explicado e visualizado pelas linhas ilustrativas do classificador na Figura 4.10, na distribuição dos dados no gráfico de espalhamento, que está em função dos atributos RMS e Fator de Crista. Fica visível também que há a necessidade de mais atributos para distinção das classes.

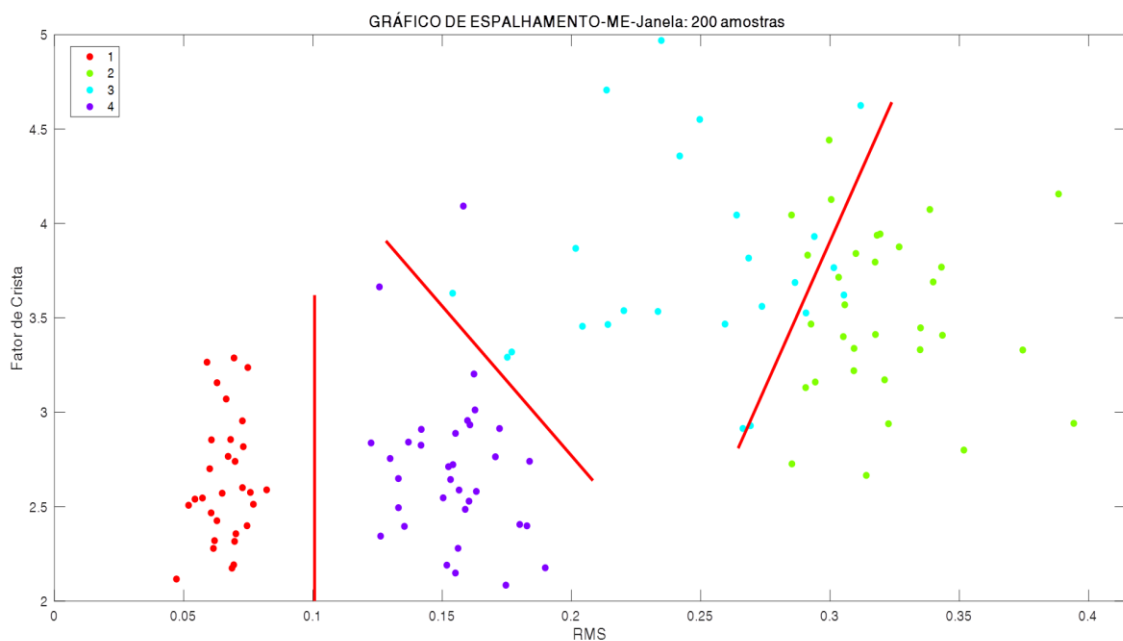


Figura 4.10 – Gráfico de espalhamento defeito 0,007" – Atributos estatísticos, dados sequenciais – janela 200 amostras (Autoria própria).

Ao utilizar atributos EEMD+AR, e dados sequenciais para janela de 200 amostras, é possível obter a matriz confusão, conforme Tabela 4.4. Observou-se que para essa janela, o valor de acurácia diminuiu comparado ao mesmo atributo, porém usando dados aleatórios, pois houve 4 erros entre os dados de teste (01 amostra errada a mais).

Tabela 4.4 – Matriz confusão defeito 0,007" – Atributos EEMD+AR, dados sequenciais – janela 200 amostras (Autoria própria).

		Classe verdadeira			
		SD	PI7	PE7	ESF7
Previsão	SD	30 25%	0	0	0
	PI7	0	30 25%	2	0
	PE7	0	0	27 22,5%	1
	ESF7	0	0	1	29 24,2%

Duas amostras pertencentes a classe PE7 foi reconhecida com classe PI7 (defeito pista interna) e uma amostra da classe PE7 foi reconhecida como classe ESF7 (defeito esfera); e uma amostra da classe ESF7 foi categorizada como classe PE7 (defeito pista externa). Verificou-se erros em distinguir defeitos de pista externa e esfera, como pode ser observado na distribuição dos dados, utilizando os coeficientes “theta 1” e “theta 2” como coordenadas principais do gráfico de espalhamento da Figura 4.11. É possível realizar a separação entre as classes, através das linhas ilustrativas, porém apresenta dificuldades entre as classes PE7(3) e ESF7(4), defeitos de pista externa e esfera.

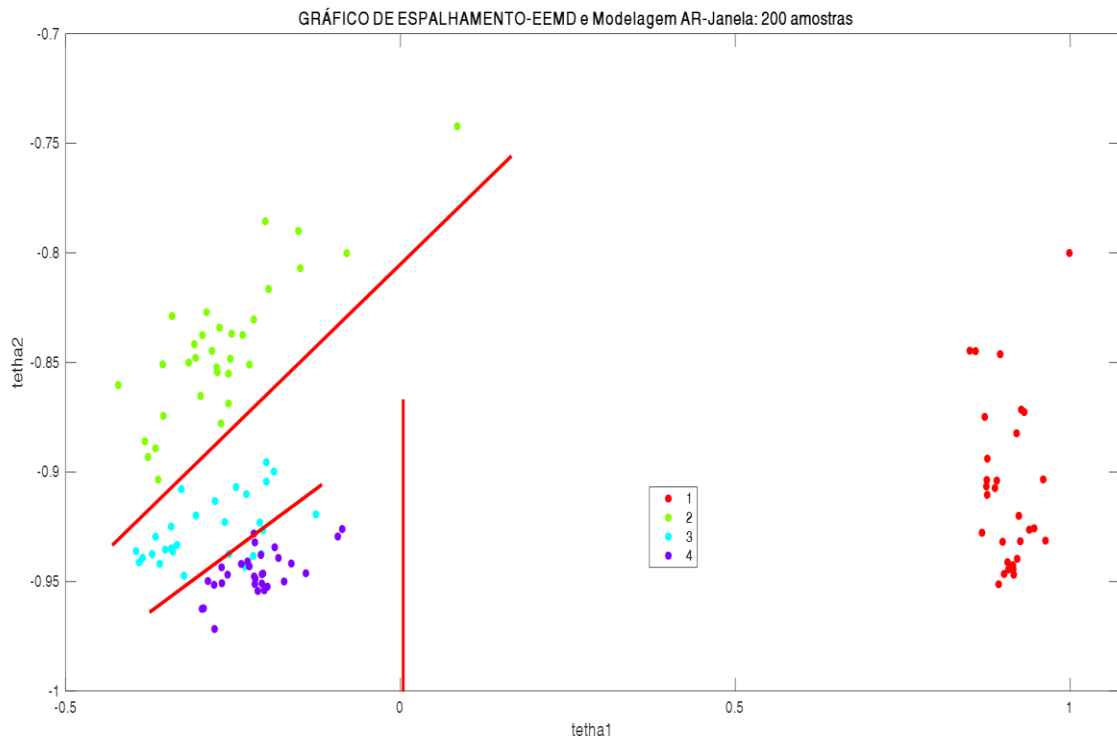


Figura 4.11 – Gráfico de espalhamento defeito 0,007" – Atributos EEMD+AR, dados sequenciais – janela 200 amostras (Autoria própria).

4.3.1.2. Defeito 0,021" – dados aleatórios e sequenciais

Para avaliar a evolução das falhas, as classes dessa análise diferem apenas no tamanho do diâmetro do defeito do rolamento, que agora é 0,0021". A quantidade de classes, número de janelas, atributos estatísticos e atributos EEMD+AR permanecem inalterados. Sendo assim, utilizando o treinamento com janelas escolhidas aleatoriamente, o resultado de desempenho global do classificador em função do tamanho da janela é mostrado na Figura 4.12

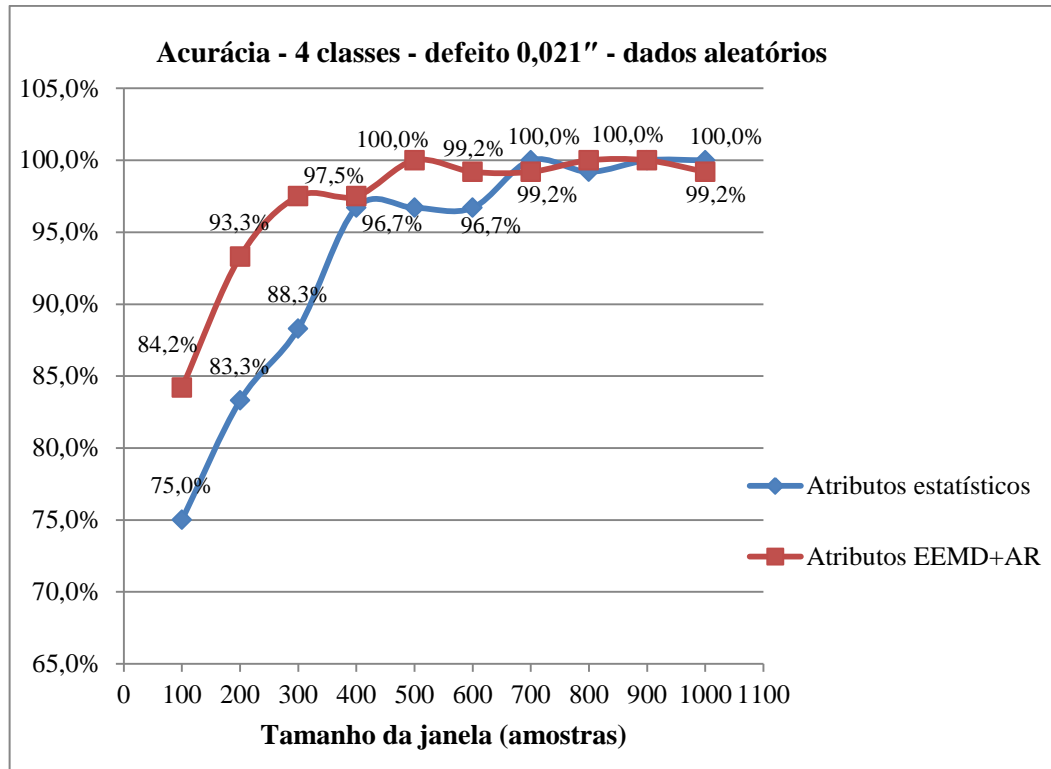


Figura 4.12 – Acurácia – 4 classes – defeito 0,021" – dados aleatórios (Autoria própria).

Observa-se que ao utilizar atributos estatísticos, para este diâmetro maior, o desempenho aumentou significativamente nas janelas menores; e em relação aos atributos EEMD+AR, apresentou bom resultado de desempenho, porém o valor de acurácia comparado ao diâmetro menor (0,007") oscilou em décimos – janelas de 500 a 1000 amostras. Resultado positivo, pois utilizou-se apenas um diâmetro de defeito, e com isso fica mais fácil o classificador realizar a categorização.

E é visível que ao aumentar a janela, o desempenho do classificador melhora, isto devido à presença de mais informações relevantes a categorização das classes.

Para visualizar o desempenho do classificador na forma de matriz confusão, a curva da Figura 4.12 converge para janela de 400 amostras, a Tabela 4.5 ilustra os cinco atributos estatísticos calculados.

Tabela 4.5 – Matriz confusão defeito 0,021" – Atributos estatísticos, dados aleatórios – janela 400 amostras (Autoria própria).

		Classe verdadeira			
		SD	PI21	PE21	ESF21
Previsão	SD	30 25%	0	0	0
	PI21	0	29 24,2%	0	0
	PE21	0	0	27 22,5%	0
	ESF21	0	1	3	30 25%

Nos dados de validação, este apresentou 4 erros de defeitos de pistas: uma amostra pertencente a classe PI21 foi reconhecida como classe ESF21 (defeito esfera 0,021") e três amostras da classe PE21 foram categorizadas como classe ESF21 (defeito esfera 0,021"). Ambos defeitos de pistas foram categorizados errados, como defeito de esfera. No gráfico de espalhamento – Figura 4.13 – é possível visualizar a dificuldade em reconhecer os defeitos de pistas interna e externa, observando apenas os dois atributos (RMS e Fator de Crista).

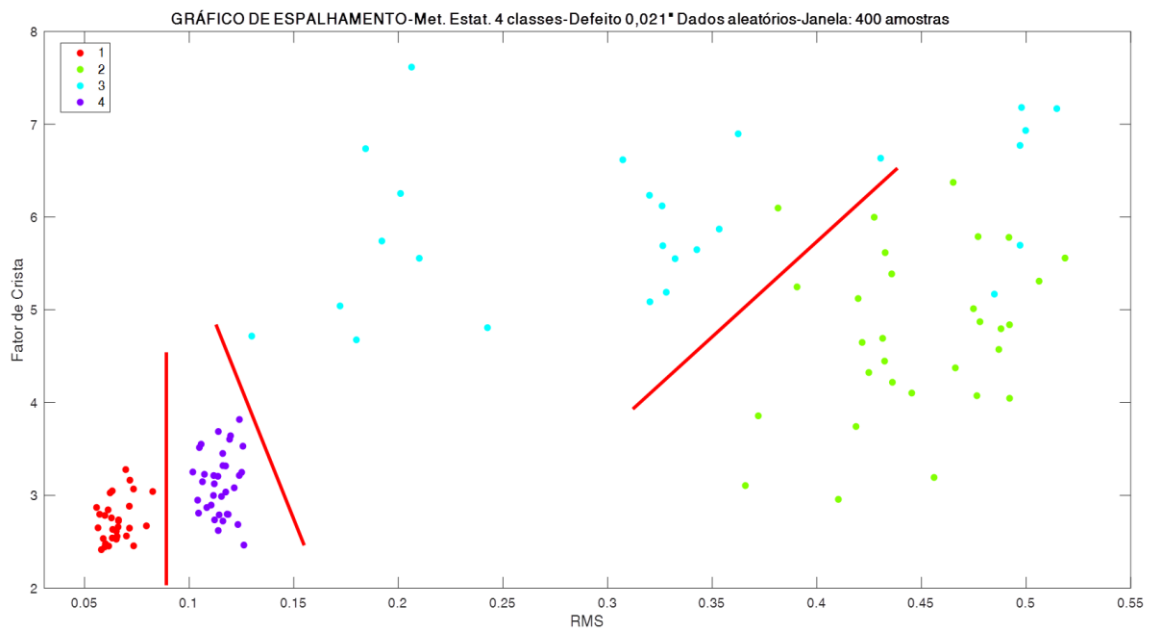


Figura 4.13 – Gráfico de espalhamento defeito 0,021" – Atributos estatísticos, dados sequenciais – janela 400 amostras (Autoria própria).

A Tabela 4.6 mostra o resultado da matriz confusão para o classificador treinado com os atributos EEMD+AR, também para janela de 400 amostras. E apresentou apenas 3 erros entre os dados de teste.

Tabela 4.6 – Matriz confusão defeito 0,021" – Atributos EEMD+AR, dados aleatórios – janela 400 amostras (Autoria própria).

		Classe verdadeira			
		SD	PI21	PE21	ESF21
Previsão	SD	30 25%	0	0	0
	PI21	0	30 25%	0	0
	PE21	0	0	29 24,2%	2
	ESF21	0	0	1	28 23,3%

Uma amostra pertencente à classe PE21 foi reconhecida com classe ESF21 (defeito esfera 0,021") e duas amostras da classe ESF21 foram categorizadas como classe PE21 (defeito pista externa 0,021"). A dificuldade apresentada foi em relação a distinção de defeito de pista externa e esfera, o que sugere ser devido a distribuição dos dados de validação das condições do rolamento presentes no gráfico de espalhamento, Figura 4.14.

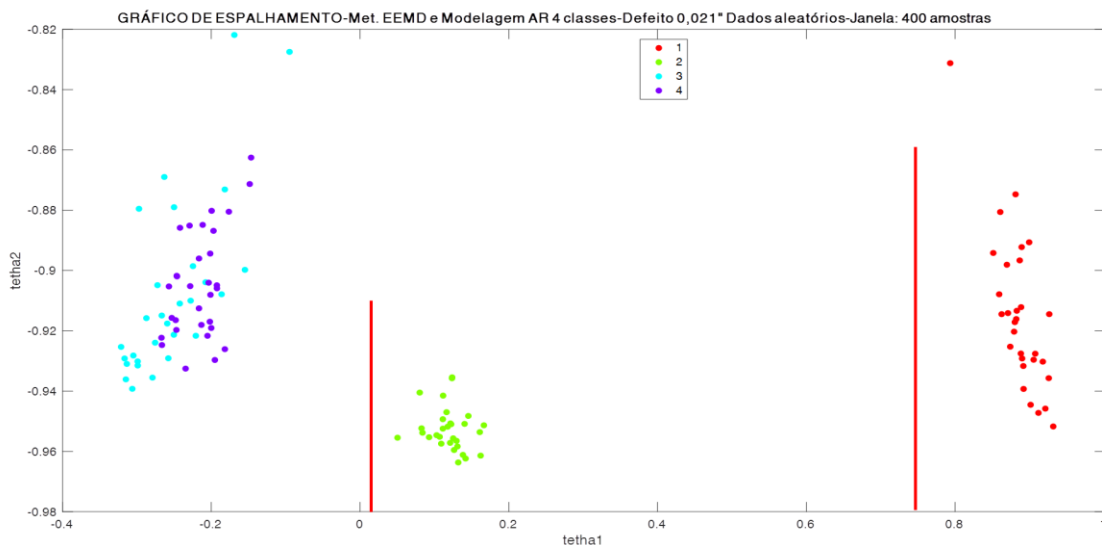


Figura 4.14 – Gráfico de espalhamento defeito 0,021" – Atributos EEMD+AR, dados sequenciais – janela 400 amostras (Autoria própria).

Na Figura 4.15, para o mesmo tamanho de defeito (0,021"), porém aplicando treinamento com janelas escolhidas sequencialmente, verifica-se que, os atributos EEMD+AR se sobressaem aos atributos estatísticos em algumas janelas.

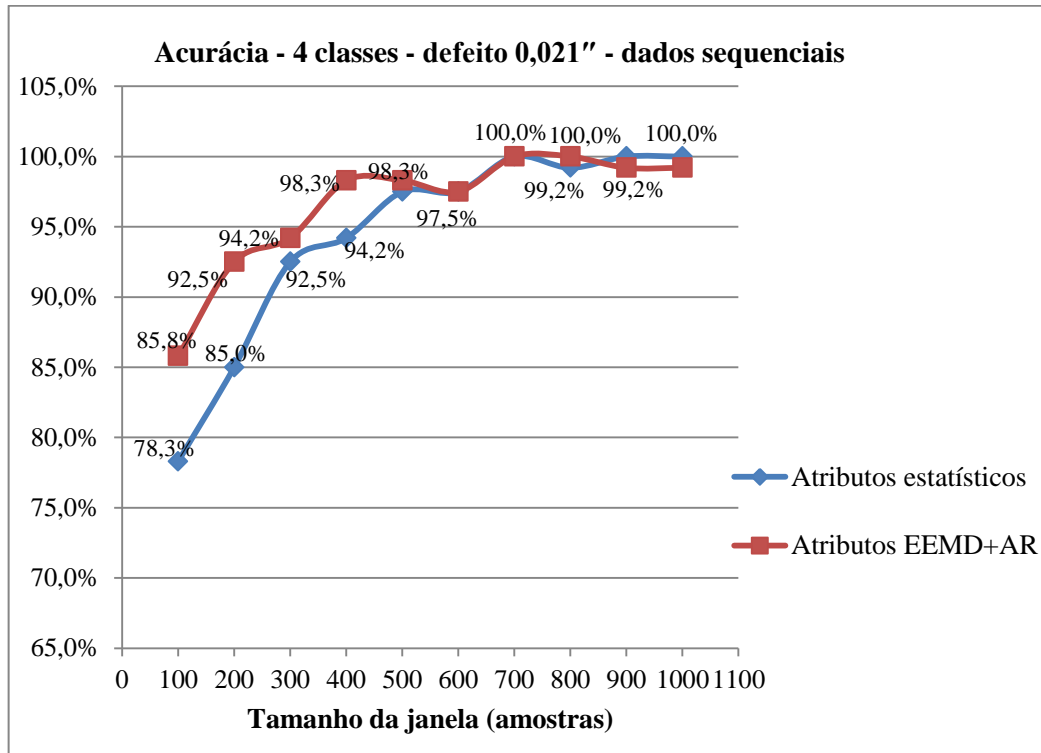


Figura 4.15 – Acurácia – 4 classes – defeito 0,021" – dados sequenciais (Autoria própria).

A matriz confusão dos cinco atributos estatísticos, calculado para janela de 400 amostras, é apresentada na Tabela 4.7. Observa-se que houve 7 erros entre os dados de validação. Em que duas amostras pertencentes à classe PI21 foram categorizadas: uma como classe PE21 (defeito pista externa 0,021") e a outra como classe ESF21 (defeito esfera 0,021"). Na classe PE21 (defeito pista externa), 5 amostras foram reconhecidas como classe ESF21 (defeito esfera 0,021"). O classificador teve dificuldades em distinguir a localização do defeito entre pistas e esferas.

Tabela 4.7 – Matriz confusão defeito 0,021" – Atributos estatísticos, dados sequenciais – janela 400 amostras (Autoria própria).

		Classe verdadeira			
		SD	PI21	PE21	ESF21
Previsão	SD	30 25%	0	0	0
	PI21	0	28 23,3%	0	0
	PE21	0	1	25 20,8%	0
	ESF21	0	1	5	30 25%

Para ilustrar a distribuição dos dados de validação desta análise, a Figura 4.16 apresenta o gráfico de espalhamento utilizando os atributos RMS e Fator de Crista.

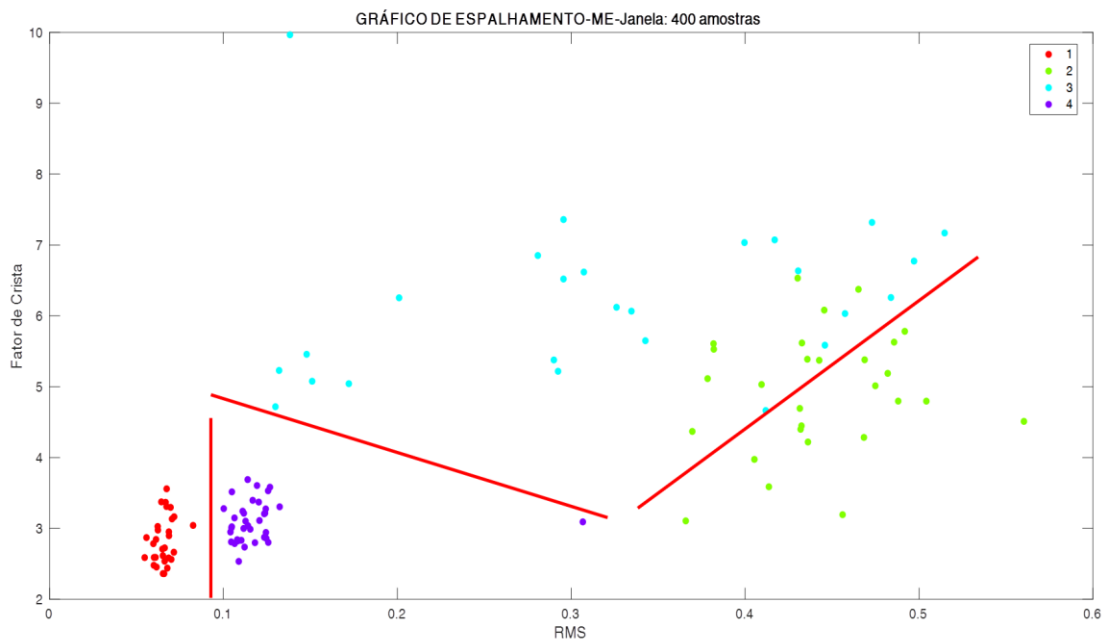


Figura 4.16 – Gráfico de espalhamento defeito 0,021" – Atributos estatísticos, dados sequenciais – janela 400 amostras (Autoria própria).

Para a mesma janela de 400 amostras, porém utilizando atributos EEMD+AR, a Tabela 4.8 mostra a matriz confusão. O valor de acurácia aumentou, comparado ao mesmo atributo, porém com dados aleatórios, pois apresentou apenas 2 erros entre os dados de teste (01 amostra errada a menos).

Tabela 4.8 – Matriz confusão defeito 0,021" – Atributos EEMD+AR, dados sequenciais – janela 400 amostras (Autoria própria).

		Classe verdadeira			
		SD	PI21	PE21	ESF21
Previsão	SD	30 25%	0	0	0
	PI21	0	30 25%	0	0
	PE21	0	0	28 23,3%	0
	ESF21	0	0	2	30 25%

Duas amostras pertencentes a classe PE21 foram categorizadas como classe ESF21 (defeito esfera 0,021"), o classificador apenas apresentou dificuldade em distinguir o defeito de pista externa em defeito de esfera. A Figura 4.17 apresenta o gráfico de espalhamento para o conjunto de validação dos dados sequenciais, e é possível observar a dificuldade de separação das classes PE21(3) e ESF21(4), devido à concentração das amostras.

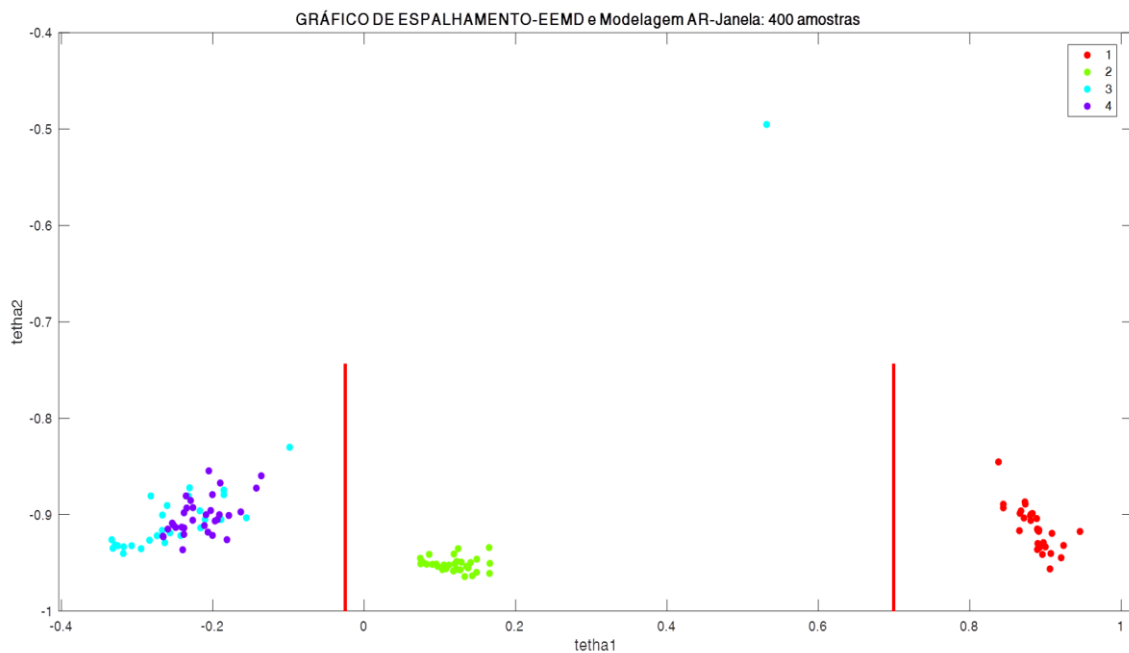


Figura 4.17 – Gráfico de espalhamento defeito 0,021" – Atributos EEMD+AR, dados sequenciais – janela 400 amostras (Autoria própria).

4.3.1.3. Combinação de atributos estatísticos – dados aleatórios – 4 classes

Após análise do desempenho global dos cinco atributos estatísticos contendo janelas escolhidas aleatoriamente para quatro classes, cogitou-se um classificador treinado apenas com dois atributos, para avaliar a capacidade de apresentar o mesmo desempenho para janelas de 400 amostras. Foram treinados classificadores cujo atributo primário fosse o valor RMS, combinado a apenas um dos outros. O resultado é apresentado na Figura 4.18.

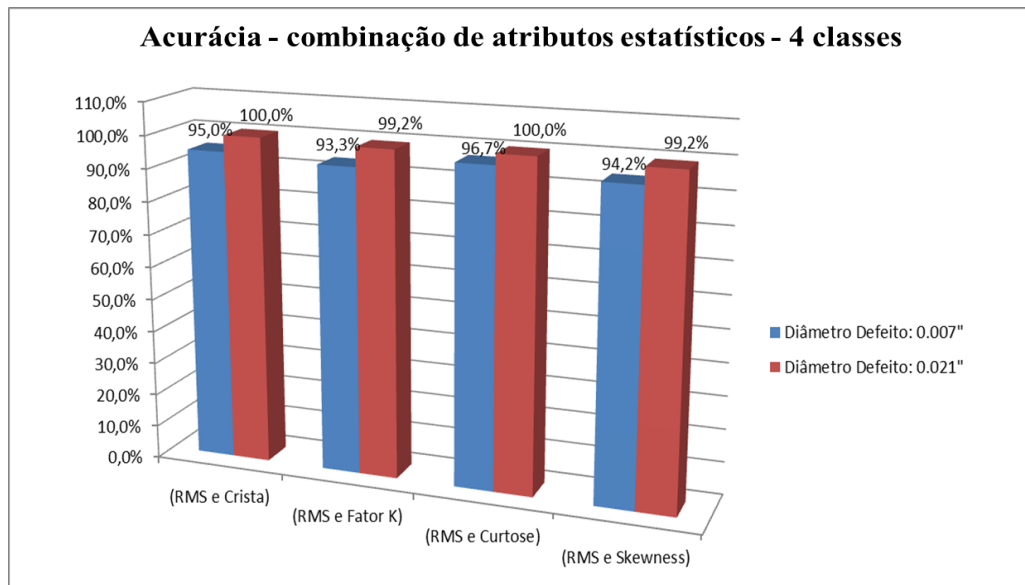


Figura 4.18 – Acurácia – combinação atributos estatísticos – dados aleatórios (Autoria própria).

A conjectura parece válida ao observar que a combinação dos atributos RMS e Crista e RMS e Curtose apresentaram elevada acurácia. Uma investigação mais profunda é recomendada, entretanto, para verificar se estes resultados podem estar sofrendo de tendenciosidade nos dados ou dependência de parâmetros, como a velocidade de rotação.

4.3.2. Atributos estatísticos e atributos EEMD+AR – 7 classes

Após a investigação da utilização de quatro classes em função do tamanho da janela já apresentada, realizou-se análise de todas as classes da condição do rolamento juntas, para avaliar o desempenho do classificador. Sendo assim, os dois diâmetros de defeitos (0,007" e 0,021") foram inseridos no Classificador SVM e janelas desde 100 até 2000 amostras foram testadas. A seguir são apresentados os resultados do classificador treinados com janelas escolhidas aleatoriamente e com janelas escolhidas na sequência temporal.

4.3.2.1. Defeito 0,007" e 0,021" – dados aleatórios e sequenciais

O Classificador foi treinado com sete classes, estas foram estabelecidas conforme Quadro 5.1 e as janelas escolhidas de forma aleatória, o resultado obtido é apresentado na Figura 4.19.

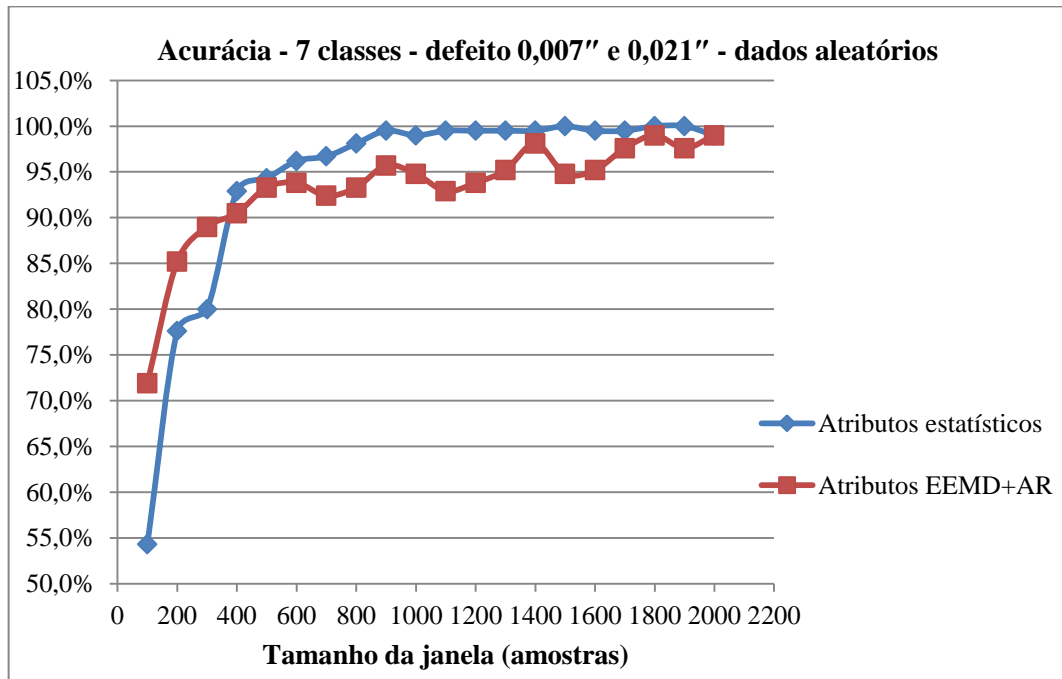


Figura 4.19 – Acurácia, dados aleatórios – 7 classes (Autoria própria).

A Figura 4.19 mostra que usando atributos estatísticos, quanto maior a janela, melhor é o desempenho do classificador, o que é esperado, já que uma janela mais longa tende a conter mais informações relevantes para caracterizar a distribuição do sinal (ergodicidade). A curva dos atributos estatísticos – Figura 4.19 – converge para janelas de 800 amostras, para as quais o desempenho não melhora muito significativamente. Esta seria, portanto, a janela que melhor concilia comprimento e acurácia.

A Tabela 4.9 ilustra a matriz confusão para o classificador treinado com os cinco atributos estatísticos calculados para janela de 800 amostras.

Tabela 4.9 – Matriz confusão 7 classes – Atributos estatísticos, dados aleatórios – janela 800 amostras (Autoria própria).

		Classe verdadeira						
		SD	PI7	PE7	ESF7	PI21	PE21	ESF21
Previsão	SD	30 14,3%	0	0	0	0	0	0
	PI7	0	30 14,3%	1	0	0	0	0
	PE7	0	0	29 13,8%	0	0	2	0
	ESF7	0	0	0	30 14,3%	0	0	0
	PI21	0	0	0	0	30 14,3%	0	0
	PE21	0	0	0	0	0	28 13,3%	0
	ESF21	0	0	0	0	0	0	30 14,3%

Observa-se que houve apenas 3 erros entre os dados de teste. Uma amostra pertencente a classe PE7 foi categorizada como classe PI7 (defeitos de pista, mesmo tamanho) e duas da classe PE21 foram reconhecidas como classe PE7 (ambos pista externa). É interessante observar que no primeiro caso, houve uma dificuldade em distinguir a localização do defeito, que estava na pista externa, mas foi detectado como sendo na pista interna, embora a amplitude (0.007") tenha sido correta. É possível que isto seja devido ao fato de defeitos pequenos produzirem formas de onda de baixa energia, cujas características podem ser facilmente confundidas. No entanto, a baixa ocorrência deste erro sugere que isto tende a ocorrer raramente. No segundo caso, o defeito foi localizado corretamente, mas o diâmetro do defeito foi subestimado.

Nos gráficos de espalhamento, Figura 4.20 e Figura 4.21, é visível o comportamento da separação entre classes, devido ao tamanho da janela. A Figura 4.20 ilustra o gráfico da janela de 800 amostras.

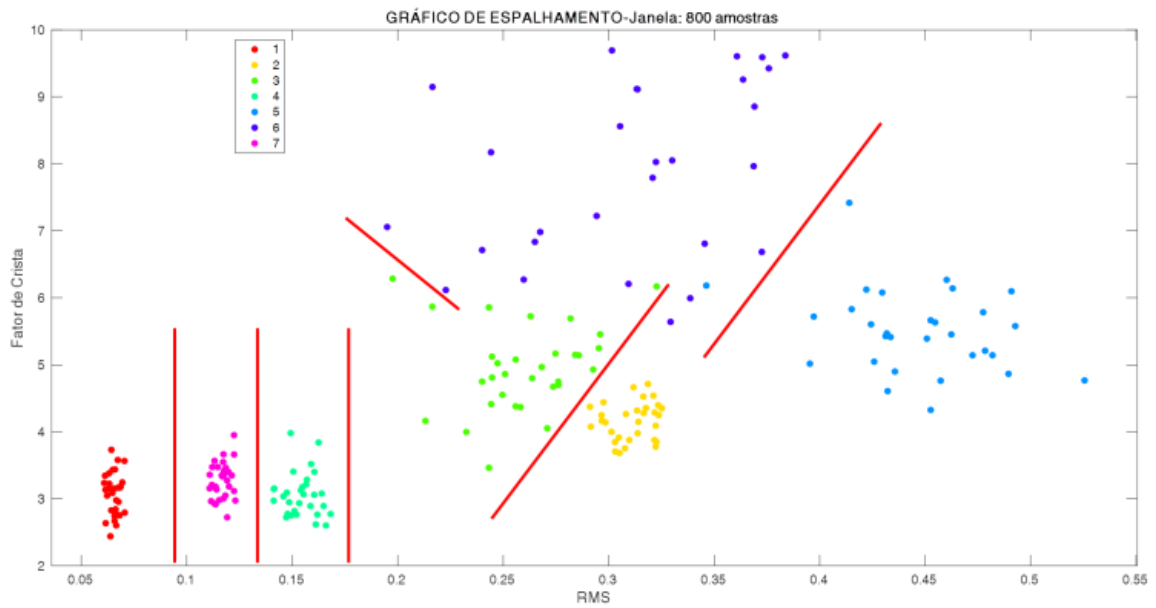


Figura 4.20 – Gráfico de espalhamento 7 classes – Atributos estatísticos, dados aleatórios – janela de 800 amostras (Autoria própria).

E ao aumentar o tamanho da janela para 1000 amostras, a separação entre as classes é bastante nítida, Figura 4.21, sendo possível isolá-las linearmente, conforme ilustrado; as linhas desenhadas servem apenas para ilustração.

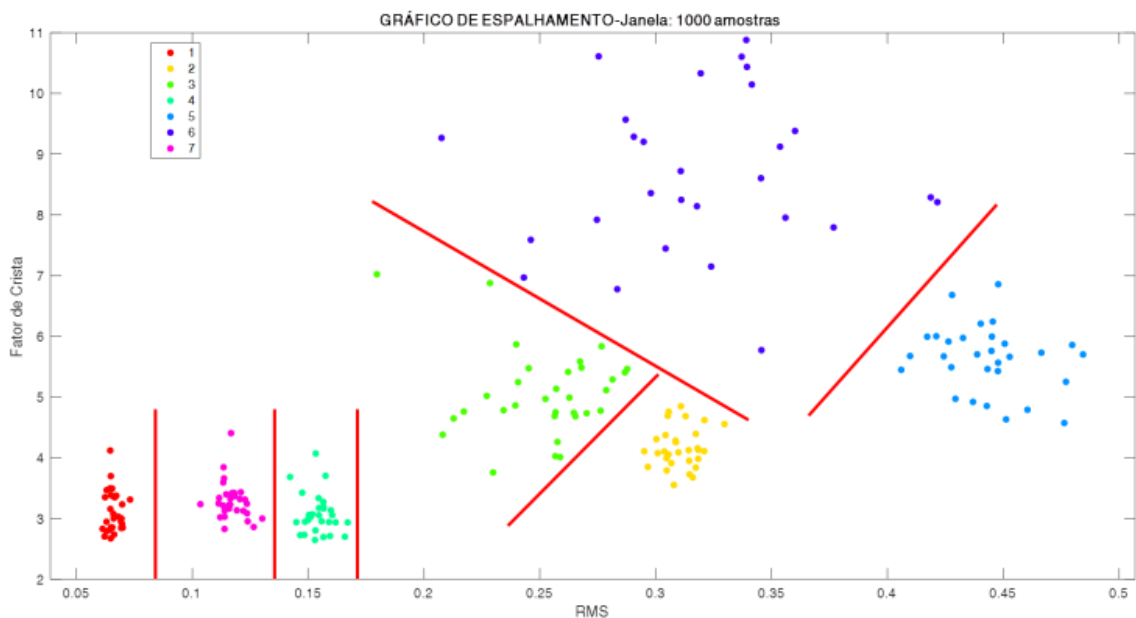


Figura 4.21 – Gráfico de espalhamento 7 classes– Atributos estatísticos, dados aleatórios – janela de 1000 amostras (Autoria própria).

Observa-se que no geral, o treinamento não tem muitas dificuldades de gerar classificadores com alta acurácia, o que sugere que os dados provenientes do banco de

dados são bem representativos das condições do rolamento. Isto ficou claro ao analisar os dois gráficos de espalhamento, Figura 4.20 e Figura 4.21, e corrobora com a tese sobre a qualidade dos dados do experimento, assim como o sucesso do classificador SVM com núcleos lineares.

Ao considerar os atributos EEMD+AR, observa-se que estes se sobressaem aos atributos estatísticos em tamanho de janelas menores; porém a partir da janela de 400 amostras os atributos estatísticos começam a superá-los. O que podemos cogitar é que, os atributos EEMD+AR, são muito parecidos, porém em classes diferentes; fato devido à inserção de mais classes ao classificador, contemplando diferentes diâmetros do defeito. Essa conjectura é percebida ao analisar a matriz confusão apresentada na Tabela 4.10, pois o classificador realiza a categorização correta da localização do defeito em pista externa, porém há dificuldade em distinguir entre os diâmetros de 0,007" e 0,021".

Tabela 4.10 – Matriz confusão 7 classes – Atributos EEMD+AR, dados aleatórios – janela 800 amostras (Autoria própria).

		Classe verdadeira						
		SD	PI7	PE7	ESF7	PI21	PE21	ESF21
Previsão	SD	30 14,3%	0	0	0	0	0	0
	PI7	0	30 14,3%	0	0	0	0	0
	PE7	0	0	29 13,8%	0	0	12	0
	ESF7	0	0	0	30 14,3%	0	0	0
	PI21	0	0	0	0	30 14,3%	0	0
	PE21	0	0	1	0	0	17 8,1%	0
	ESF21	0	0	0	0	0	1	30 14,3%

As amostras para a janela de 1000 amostras estão representadas no gráfico de espalhamento, Figura 4.22. É visível que considerando apenas dois atributos (“theta1” e “theta 2”), o classificador não consegue realizar a categorização entre as classes de defeito de pista externa e esfera.

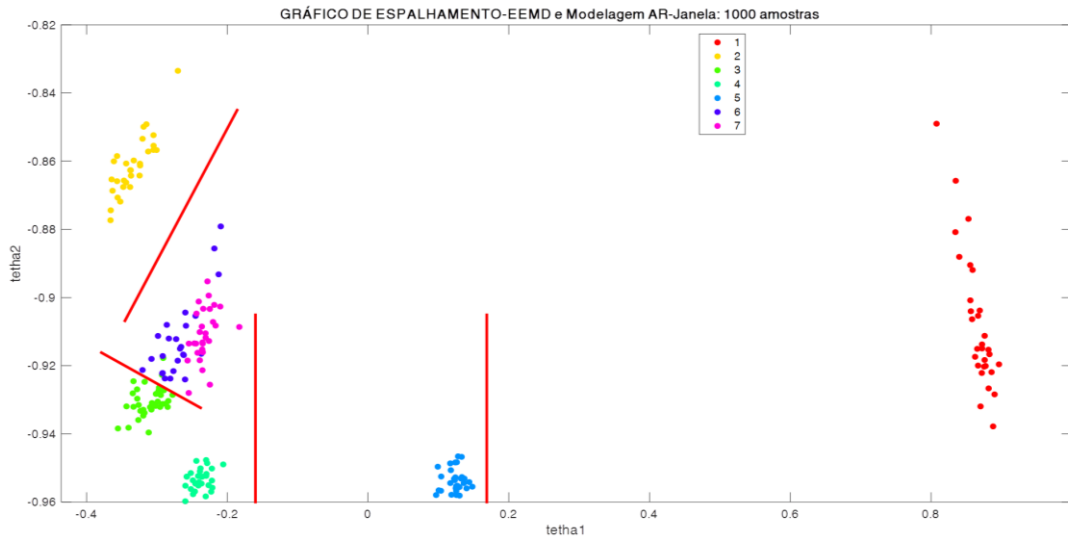


Figura 4.22 – Gráfico de espalhamento 7 classes – Atributos EEMD+AR, dados aleatórios – janela de 1000 amostras (Autoria própria).

Ao realizar a investigação do desempenho do Classificador com janelas escolhidas na sequência temporal, para as sete classes do rolamento, o resultado obtido é similar aos dados obtidos aleatoriamente, pois os atributos EEMD+AR também se sobressaem aos atributos estatísticos, mas nesta análise apenas nas duas primeiras janelas – Figura 4.23 – e a partir da terceira janela, os atributos estatísticos começam a superar os atributos EEMD+AR.

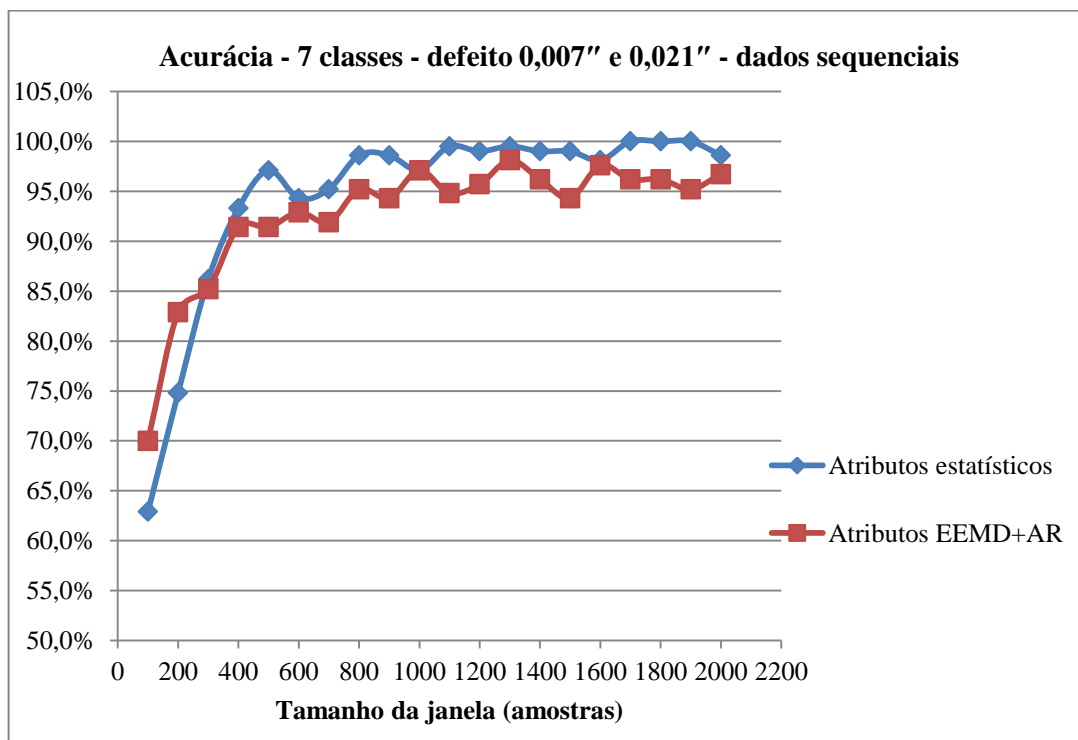


Figura 4.23 – Acurácia, dados sequenciais – 7 classes (Autoria própria).

Os atributos estatísticos comparados ao processo de dados aleatórios apresentam pequena variação no valor de acurácia, como pode ser observado na Tabela 4.11, que ilustra a matriz confusão do classificador treinado com os cinco atributos estatísticos calculados de forma sequencial, para a janela de 800 amostras. E observa-se que houve apenas 3 erros de pista interna e externa diâmetro de 0,0021". O classificador reconheceu corretamente a localização, porém apresentou dificuldade em distinguir o diâmetro do defeito de pista de 0,021" do diâmetro de 0,007".

Tabela 4.11 – Matriz confusão 7 classes – Atributos estatísticos, dados sequenciais – janela 800 amostras (Autoria própria).

		Classe verdadeira						
		SD	PI7	PE7	ESF7	PI21	PE21	ESF21
Previsão	SD	30 14,3%	0	0	0	0	0	0
	PI7	0	30 14,3%	0	0	2	0	0
	PE7	0	0	30 14,3%	0	0	1	0
	ESF7	0	0	0	30 14,3%	0	0	0
	PI21	0	0	0	0	28 13,3%	0	0
	PE21	0	0	0	0	0	29 13,8%	0
	ESF21	0	0	0	0	0	0	30 14,3%

E ao analisar a matriz confusão da Tabela 4.12, é possível ver que o resultado dos atributos EEMD+AR, apresentaram 10 erros de pistas externas, estes foram localizados corretamente, porém há dificuldade em distinguir entre os diâmetros de 0,007" e 0,021".

Tabela 4.12 – Matriz confusão 7 classes – Atributos EEMD+AR, dados sequenciais – janela 800 amostras (Autoria própria).

		Classe verdadeira						
		SD	PI7	PE7	ESF7	PI21	PE21	ESF21
Previsão	SD	30 14,3%	0	0	0	0	0	0
	PI7	0	30 14,3%	0	0	0	0	0
	PE7	0	0	26 12,4%	0	0	6	0
	ESF7	0	0	0	30 14,3%	0	0	0
	PI21	0	0	0	0	30 14,3%	0	0
	PE21	0	0	4	0	0	24 11,4%	0
	ESF21	0	0	0	0	0	0	30 14,3%

Os gráficos de espalhamentos dos atributos estatísticos e dos atributos EEMD+AR estão ilustrados na Figura 4.24 e 4.25.

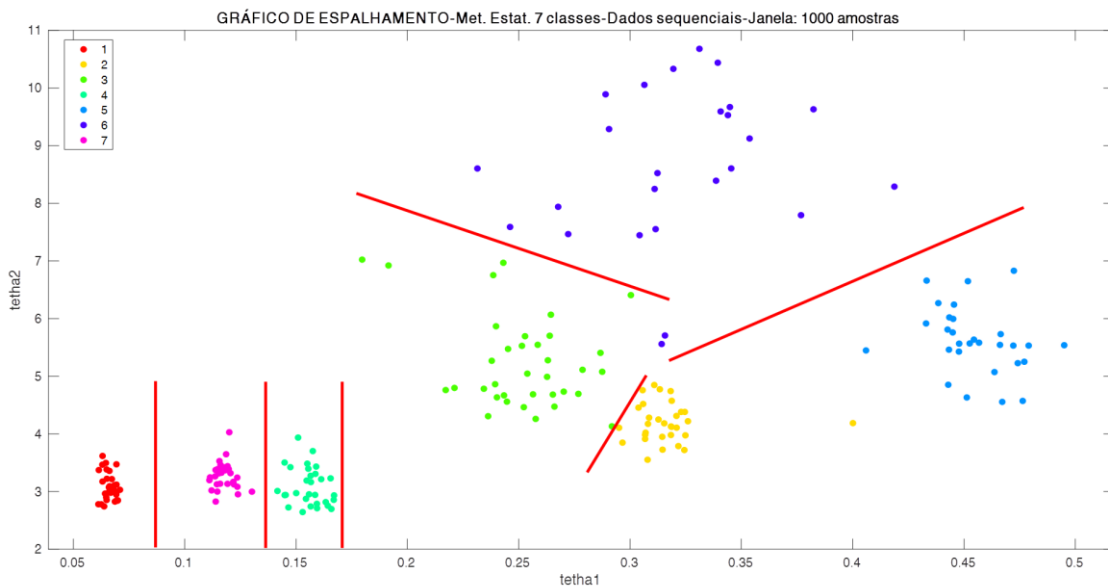


Figura 4.24 – Gráfico de espalhamento 7 classes– Atributos estatísticos, dados sequenciais – janela de 1000 amostras (Autoria própria).

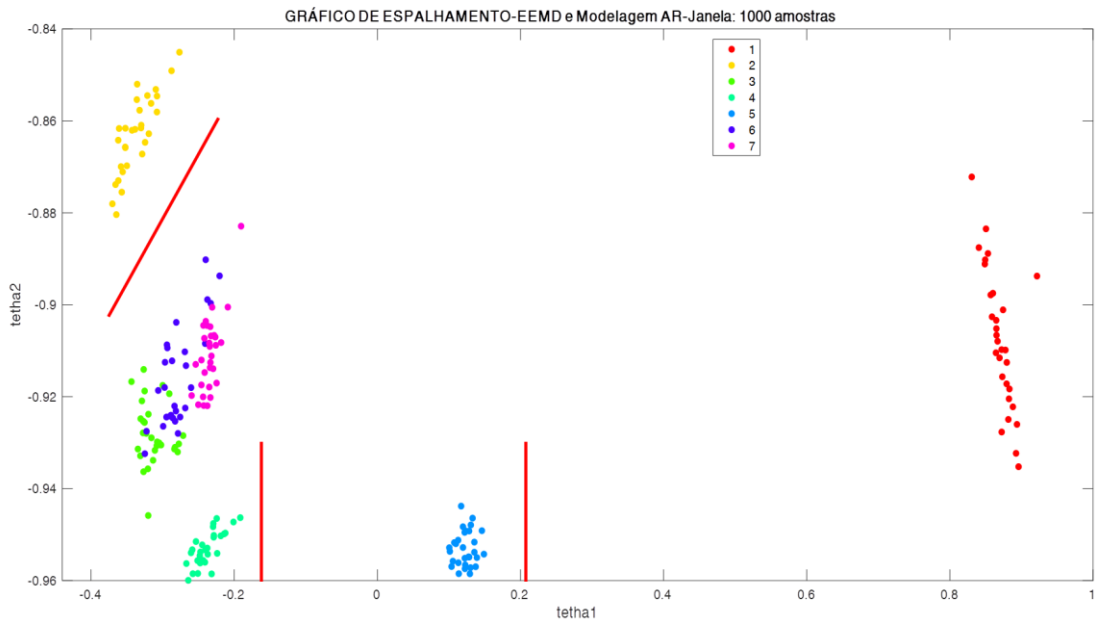


Figura 4.25 – Gráfico de espalhamento 7 classes – Atributos EEMD+AR, dados sequenciais – janela de 1000 amostras (Autoria própria).

4.3.2.2. Combinação de atributos estatísticos – dados aleatórios – 7 classes

Após apresentação do desempenho do classificador, utilizando atributos estatísticos contendo dados aleatórios para sete classes, e ao observar a Figura 4.21, verifica-se que o nível RMS é suficiente para distinguir entre quase todas as classes, exceto a classe PE21, para a qual o Fator de Crista torna-se absolutamente necessário para distingui-la das classes PI7(2) e PE7(3). Então, é natural que se cogite se um classificador treinado apenas com dois atributos seria capaz de ter o mesmo desempenho.

Com isso, a análise da Figura 4.21 levantou a conjectura de que um classificador treinado apenas com dois atributos seria capaz de ter o mesmo desempenho, o que levou a esta investigação, em que foram treinados classificadores cujo atributo primário fosse o valor RMS, combinado a apenas um dos outros atributos estatísticos. A Figura 4.26 apresenta o resultado do classificador SVM.

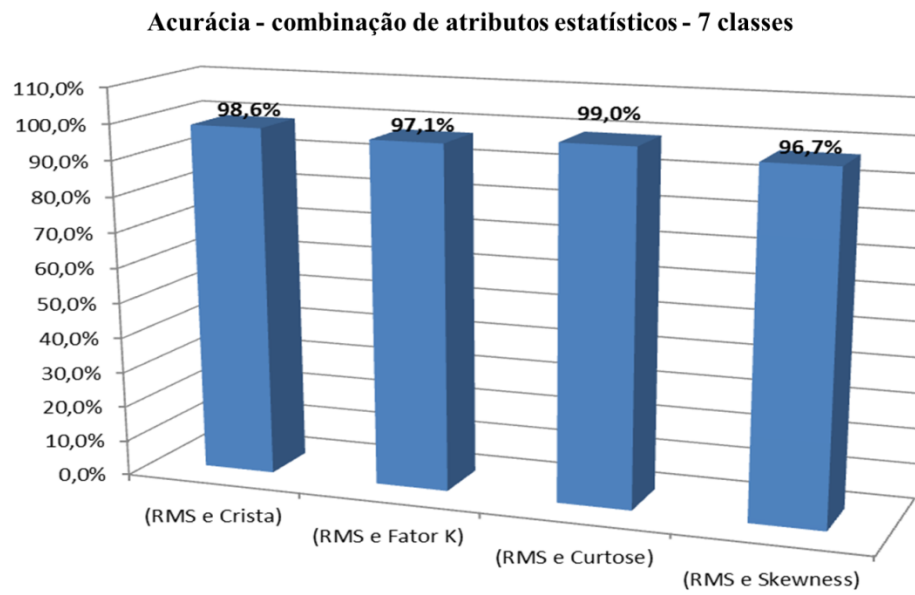


Figura 4.26 – Acurácia combinação atributos estatísticos – dados aleatórios (Autoria própria).

Conforme os resultados de acurácia obtidos para as sete condições do rolamento, a hipótese parece válida ao observar que a combinação dos atributos RMS e Curtose apresentaram elevada acurácia. Porém é recomendada uma investigação mais profunda para verificar se estes resultados podem estar sofrendo de tendenciosidade nos dados ou dependência de parâmetros.

5. CONCLUSÃO

Devido à importância do diagnóstico de falhas em rolamento no processo de manutenção de um sistema mecânico no setor industrial, este trabalho colabora na apresentação de investigações de inserção de atributos estatísticos e atributos EEMD+AR em diferentes tamanhos de janelas temporais, diferentes diâmetros de defeitos e combinações de atributos estatísticos, para que o classificador SVM realize a predição do diagnóstico do defeito. E então verificou-se que:

- Atributos estatísticos são formas eficazes para distinguir entre diferentes condições do rolamento, desde que analisados juntamente com outros atributos estatísticos, e não apenas sozinho. As elevadas acurácias obtidas sugerem que há forte indício de que os sinais de vibração possuem padrões estatísticos plenamente caracterizados por momentos de até ordem 4;
- Classificadores SVM podem ser treinados com alta eficiência (99%) ao usar atributos estatísticos estudados como preditores, desde que sejam calculados em uma janela de tempo suficiente longa;
- A análise da quantidade de classes inseridas no classificador influenciou no desempenho global, pois os atributos EEMD+AR em classes de apenas um diâmetro, produziram acurácia melhor que a análise das sete classes juntas;
- Ao inserir atributos EEMD+AR, com os parâmetros selecionados, o classificador categorizou a localização dos defeitos, porém teve dificuldades em distinguir os diâmetros de 0,007" e 0,021";
- A análise de combinações de atributos estatísticos para sete classes mostrou que o valor RMS e Curtose do sinal oferecem um conjunto minimalista de atributos capazes de distinguir entre sete classes diferentes de estados do rolamento, nas condições oferecidas pelo experimento;
- Conjectura também que a modelagem AR pode não ter conseguido distinguir entre os tamanhos do defeito, devido à presença de semelhanças nas IMFs;
- A análise dos dados de treinamento e validação, com janelas escolhidas aleatoriamente e com janelas escolhidas na sequência temporal permitiu analisar o comportamento da separação entre as classes e a tendenciosidade dos dados, e verificou-se que produzem melhorias pouco significativas no desempenho global do classificador.

Os resultados demonstram a viabilidade de se realizar diagnóstico de rolamentos através de classificadores SVM, desde que realizado várias análises e hipóteses. Apesar de resultados positivos, é importante expandir esta investigação para analisar a influência de circunstâncias, tais como a velocidade de rotação, parâmetros do modelo EEMD, diferentes atrasos na modelagem AR, e presença de outros tipos de defeitos sobre a acurácia, para uma visão mais definitiva sobre o emprego deste tipo de técnica em sistemas reais de diagnóstico.

Diante disto, para continuidade das análises de diagnóstico de falhas em rolamentos, propõem-se as seguintes sugestões para pesquisas futuras:

- Coletar sinais de vibração da bancada de testes existente no laboratório de Engenharia Mecânica do CAMPUS Tucuruí-PA e usar os sinais para diagnóstico do Classificador SVM;
- Realizar diferentes combinações de atributos estatísticos;
- Testar novos atributos estatísticos;
- Aplicar ao classificador SVM sinais não estacionários;
- Aplicar outras técnicas de Classificação;
- Inserir combinações de defeitos simultaneamente (pista interna, pista externa, esfera e gaiola) para investigar o desempenho do Classificador SVM.

6. REFERÊNCIAS

- [1] Aguirre, L. A. (2014). Introdução à identificação de sistemas: técnicas lineares e não lineares: teoria e aplicação (4. ed. rev.). Belo Horizonte: Editora UFMG.
- [2] Alfredson, R. J. e Mathew, J. (1985). Time domain methods for monitoring the condition of rolling element bearings. Institution of Engineers (Australia), Mechanical Engineering Transactions (ISSN 0727-7369), vol. ME10, 102-107.
- [3] Amit, R. B., Gajanan, K. A. e Sachin, P. U. (2014). Comprehensive bearing condition monitoring algorithm for incipient fault detection using acoustic emission. *Jurnal Tribologi* 2, pp.1-30.
- [4] Atyabi, A. Shic, F. e Naples, A. (2017). Mixture of autoregressive modeling orders and its implication on single trial EEG classification. *Expert Syst Appl.* 65, 164–180. doi: 10.1016/j.eswa.2016.08.044.
- [5] Barbosa, T. S. (2015). Detecção de falhas em vigas em balanço através da análise de sinais de vibração e estatísticas de ordem superior. (Dissertação). UFLA, Lavras.
- [6] Barszcz, T. e Sawalhi, N. (2012). Fault Detection Enhancement in Rolling Element Bearings Using the Minimum Entropy Deconvolution. *Archives of Acoustics.* 37(2). doi: 10.2478/v10168-012-0019-2. 131–141 pp.
- [7] Bezerra, R. A. (2004). Detecção de Falhas em Rolamentos por Análise de Vibração. (Tese de doutorado). Pós Graduação da Faculdade de Engenharia Mecânica, Campinas, São Paulo.
- [8] Bianchi, M. F. de. (2006). Extração de características de imagens de faces humanas através de Wavelets, PCA e IMPCA. Dissertação. Escola de Engenharia de São Carlos da Universidade de São Paulo.
- [9] Burges, C. J. C. (1998). A tutorial on support vector machines for pattern recognition. *Knowledge Discovery and Data Mining*, 2(2), 1–43.

- [10] Caesarendra W., Kosasih, P. B., Tieu, A. K., Moodie, C. A. S. e Choi, B.-K. (2013). Condition monitoring of naturally damaged slow speed slewing bearing based on ensemble empirical mode decomposition. *Journal of Mechanical Science and Technology* 27(8), 2253-2262. doi: 10.1007/s12206-013-0608-7.
- [11] Campbell, C. (2000). An introduction to kernel methods. *Radial Basis Function Networks: Design and Applications*, 155–192, Berlin: Springer Verlag.
- [12] Carvalho, R. V. (2010). Análise dinâmica de rolamentos de esfera. (Dissertação de Mestrado). Universidade Estadual de Campinas, Faculdade de Engenharia Mecânica. Campinas, SP.
- [13] Case Western Reserve University Bearing Data Center <<https://cseggroups.case.edu/bearingdatacenter/home>>. Acesso em 17.12.18.
- [14] Casella, G. e Berger, R. L. (2002). *Statistical Inference*, 2nd ed. Duxbury, Thomson Learning, chapter 2. ISBN: 0-534-24312-6.
- [15] Cheng, J., Yu, D. e Yang, Y. (2008). A Fault Diagnosis Approach for Gears Based on IMF ARModel and SVM. *EURASIP Journal on Advances in Signal Processing*. 2008, 7. doi: 10.1155/2008/647135.
- [16] Chen, Y. (1992). Machine Condition Monitoring by Inverse Filtering and Statistical Analysis. *Mech. Syst. Signal Process.*, 6(2), 177–189.
- [17] Cristianini, N. e Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and other kernel-based learning methods*. Cambridge University Press.
- [18] Ding, G., Lide Wang, Peng Yang, Ping Shen, Shuping Dang. (2013). Diagnosis Model Based on Least Squares Support Vector Machine Optimized by Multi-swarm Cooperative Chaos Particle Swarm Optimization and Its Application. *Journal of Computers*, 8(4), 975-982.

- [19] Finotti, R. P. et al. (2019). An SHM approach using machine learning and statistical indicators extracted from raw dynamic measurements. *Latin American Journal of Solids and Structures*, 16(2), e165. doi: 10.1590/1679-78254942.
- [20] Freitas, A. A. C. (2007). Previsão de Séries Temporais via Seleção de Variáveis, Reconstrução Dinâmica, ARMA-GARCH e Redes Neurais Artificiais. (Tese de Doutorado). Universidade Estadual de Campinas - Faculdade de Engenharia Elétrica e de Computação. Campinas, SP.
- [21] Gaci, S. (2016). A new ensemble empirical mode decomposition (EEMD) denoising method for seismic signals. *Energy Procedia*, 97 (2016), 84–91. doi: 10.1016/j.egypro.2016.10.026.
- [22] Harris, T. A. (1991). *Rolling Bearing Analysis*. (3 ed). New York, USA: John Wiley & Sons.
- [23] Haseim, A. A., Sudirman, R. e Khalid P. I. (2013). Handwriting Classification Based on Support Vector Machine with Cross Validation. *Scientific Research. Engineering*, 5, 84-87. doi: 10.4236/eng.2013.55B017.
- [24] Hearst, M. A., Schölkopf, B., Dumais, S., Osuna, E., e Platt, J. (1998). Trends and controversies - support vector machines. *IEEE Intelligent Systems*, 13(4), 18–28.
- [25] Hong, D., Ya, W. e Shuzi, Y. (1989). *Fault diagnosis by time series analysis*. Applied Time Series Analysis, World Scientific. River Edge, NJ, USA.
- [26] Hoshyar, A. N., Kharkovsky, S. e Samali, B. (2018). Statistical Features and Traditional SA-SVM Classification Algorithm for Crack Detection. *Journal of Signal and Information Processing*, 9, 111-121. doi: 10.4236/jsip.2018.92007.
- [27] Huang, N. E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.-C., Tung, C.C., e Liu, H.H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Lond. A*, 454, 903–995.

- [28] Jachan, M. (2006). Time-Frequency-Autoregressive-Moving-Average Modeling of Nonstationary Processes. Thesis. Vienna University of Technology, Vienna, Austria.
- [29] Juvinall, R. C. e Marshek, K. M. (1991). Fundamentals of Machine Component Design. (2 ed). New York, USA: John Wiley & Sons.
- [30] Kim, D. (2006). Aula: Introduction to EMD (Empirical Mode Decomposition) with application to a scientific data. Online: <http://dasan.sejong.ac.kr/~dhkim/main/research/talks/EMDintroSeminar.pdf>. Acesso em 20.04.2018.
- [31] Kokil, S., Gajjal, S. Y., Shah, M. M. e Kokil, S. D. (2014). Detection of Fault in Rolling Element Bearing using Condition Monitoring by Experimental Approach. International Journal of Engineering Research & Technology (IJERT). 3 (8).
- [32] Lee, D. H., Ahn, J. H. e Koh, B. H. (2017). Fault Detection of Bearing Systems through EEMD and Optimization Algorithm. Sensors, 17, 2477. doi: 10.3390/s17112477.
- [33] Lei, Y., Li, N., Lin, J. e Wang, S. (2013). Fault Diagnosis of Rotating Machinery Based on an Adaptive Ensemble Empirical Mode Decomposition. Sensors, 13, 16950-16964. doi:10.3390/s131216950.
- [34] Liu Z., Liu Y., Shan H., Cai B. e Huang Q. (2015). A fault diagnosis methodology for gear pump based on EEMD and Bayesian Network. PLOS ONE 10(5). doi:10.1371/journal.pone.0125703.
- [35] Ljung, L. (2010). Perspectives on System Identification. (Technical report SE-581 83). Linköping, Sweden: Division of Automatic Control, Automatic Control, Linköpings University.
- [36] Lorena, A. C. e Carvalho, A. C. P. L. F. de (2007). Uma Introdução às Support Vector Machines. RITA. Volume XIV. Número 2. doi: 10.22456/2175-2745.5690.

- [37] Mahgoun, H. Bekka R. E. e Felkaoui, A. (2012). Gearbox fault diagnosis using ensemble empirical mode decomposition (EEMD) and residual signal. *Mechanics & Industry*, 13(1), 33-44. doi: 10.1051/meca/2011150.
- [38] Marçal, R. F. M. (2000). Um método para detectar falhas incipientes em máquinas rotativas baseado em análise de vibrações e lógica Fuzzy. Tese. Universidade Federal do Rio Grande do Sul. Porto Alegre.
- [39] Mesquita, A. L. A., Santiago, D. F. A, Bezerra, R. A., Miranda, U. A., Dias Jr, M. e Pederiva, R. (2002). Detecção de falhas em rolamentos usando Transformadas Tempo-Frequência – Comparação Com Análise De Envelope. *Mecânica Computacional*, 21, 1938-1954.
- [40] Meyer, D. (2017). Support Vector Machines. The Interface to libsvm. Technikum Wien, Austria.
- [41] Müller, K. R., S. Mika, G. Rätsch, K. Tsuda, and B. Schölkopf. (2001). An introduction to kernel-based learning algorithms. *IEEE Transactions on Neural Networks*, 12(2), 181–201.
- [42] Napolitano, A. Cyclostationary Signal Processing and its Generalizations, 2014.
- [43] Otani, M. e Machado, W. V. (2008). A proposta de desenvolvimento de gestão da manutenção industrial na busca da excelência ou classe mundial. *Revista Gestão Industrial*, 4 (2), 01-16.
- [44] Palaniappan R. (2006). Towards Optimal Model Order Selection for Autoregressive Spectral Analysis of Mental Tasks Using Genetic Algorithm. *IJCSNS International Journal of Computer Science and Network Security*, 6(1A),153–162.
- [45] Parmezan, A. R. S. e Batista, G. E. A. P. A. (2016). Descrição de Modelos Estatísticos e de Aprendizado de Máquina para Predição de Séries Temporais. (Relatórios Técnicos N° 412 - ISSN - 0103-2569). Instituto de Ciências Matemáticas e de Computação São Carlos.

- [46] Passerini, A. (2004). Kernel Methods, multiclass classification and applications to computational molecular biology. (PhD thesis), Università Degli Studi di Firenze.
- [47] Pontil M. e Verri, A. (1998). Support vector machines for 3-D object recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(6), 637–646.
- [48] Rofe, S. (1997). Signal Processing Methods for Gearbox Fault Detection. Defence Science and Technology Organisation (Technical Report, DSTO-TR- 476). Australia.
- [49] Santos, M. C. M, Lins, I. D., Maior, C. B. S., Moura, M. das C. e Droguett, E. L. (2017). Diagnóstico e Prognóstico de falha de rolamentos via Support Vector Machines e técnicas de pré-processamento de dados. XLIX Simpósio Brasileiro de Pesquisa Operacional.
- [50] Santos, R. de S. (2017). Detecção de falhas em rolamentos de máquinas rotativas utilizando técnicas de processamentos de sinais. Tese. Universidade Estadual Paulista, Faculdade de Engenharia de Guaratinguetá, São Paulo.
- [51] Schölkopf, B. and Smola, A. J. (2002). Learning with Kernels. The MIT Press, Cambridge, Massachusetts. London, England.
- [52] Selina S. Y. Ng, Peter W. T. and Kwok L. T. (2014). A One-Versus-All Class Binarization Strategy for Bearing Diagnostics of Concurrent Defects. *Sensors*, 14, 1295-1321. doi: 10.3390/s140101295.
- [53] Sidar, R., Sen, P. K., Sahu, G. (2015). Review of Vibration Based Fault Diagnosis in Rolling Element Bearing and Vibration Analysis Techniques. *International Journal of Scientific Research Engineering & Technology (IJSRET)*, 4 (10), 998-1003. ISSN 2278 – 0882.
- [54] Singla, R., Chambayil, B., Khosla, A. e Santosh, J. (2011). Comparison of SVM and ANN for classification of eye events in EEG. *J. Biomedical Science and Engineering*, 4, 62-69. doi: 10.4236/jbise.2011.41008.

- [55] Sikora, E. A. (2016) Detection of bearing damage by statistic vibration analysis. IOP Conf. Series: Materials Science and Engineering 124 012167 doi:10.1088/1757-899X/124/1/012167.
- [56] Smith, W. A. e Randall, R. B. (2015). Rolling element bearing diagnostics using the Case Western Reserve University data: a benchmark study. Mechanical Systems and Signal Processing. doi: 10.1016/j.ymssp.2015.04.021.
- [57] Smola, A. J., P. Barlett, P. L., Schölkopf, B. e Schuurmans, D. (2000). Introduction to large margin classifiers. *Advances in Large Margin Classifiers*, 1–28. MIT Press.
- [58] Steinwart, I. e Christmann, A. (2008). *Support Vector Machine* (1st ed.). New York, USA: Springer-Verlag.
- [59] Tandon, N. e Choudhury, A. (1999.) A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings. *Tribology International*, 32(8), 469-480. doi: 10.1016/S0301-679X(99)00077-8.
- [60] Tandon, N. e Parey, A. (2006). Condition Monitoring of Rotary Machines. *Condition Monitoring and Control for Intelligent Manufacturing*. 109-136. Springer, London. doi: 10.1007/1-84628-269-1.
- [61] Trevisan, R. (2011). Utilização de medições ultrassônicas de vibração e da emissão acústica para implantação de lubrificação base-condição visando a prevenção de falhas em mancais de rolamento. (Dissertação de Mestrado). Universidade Federal do Rio Grande do Sul. Pós Graduação em Engenharia de Minas, Metalúrgica e de Materiais. Porto Alegre, RS.
- [62] Vapnik, V. (1995). *The nature of statistic learning theory*. New York: Springer.
- [63] Wang, Y. H., Yeh, C. H., Young, H. W. V., Hu K. e Lo, M. T. (2014). On the computational complexity of the empirical mode decomposition algorithm. *Physica A*, 400, 159–167.

[64] Wu, Z. e Huang, N. E. (2005). Ensemble Empirical Mode Decomposition: A Noise Assisted Data Analysis Method. *Advances in Adaptive Data Analysis* 01(01), 1-41. doi: 10.1142/S1793536909000047.

[65] Ya, W. e Shuzi, Y. (1989). Application of several time series models in prediction. *Applied Time Series Analysis*, World Scientific. River Edge, NJ, USA.

[66] Yi, C., Wang, D., Fan, W., Tsui, K.-L. e Lin, J. (2018). EEMD-Based Steady-State Indexes and Their Applications to Condition Monitoring and Fault Diagnosis of Railway Axle Bearings. *Sensors*, 18, 704. doi:10.3390/s18030704.