



Diagnóstico e Prognóstico de falha de rolamentos via Support Vector Machines e técnicas de pré-processamento de dados

Monalisa Cristina Moura dos Santos^a

monalisamoura24@gmail.com

Isis Didier Lins^a

isis.lins@ceerma.org^a

Caio Bezerra Souto Maior^a

caiomaior@hotmail.com

Márcio das Chagas Moura^a

marcio@ceerma.org

Enrique López Droguett^b

elopezdroguett@ing.uchile.cl

^aCentro de Estudos e Ensaios em Risco e Modelagem Ambiental – Departamento de Engenharia de Produção – Universidade Federal de Pernambuco
Av. da Arquitetura, s/n - Cidade Universitária, Recife - PE, 50740-550

^bDepartamento de Engenharia Mecânica – Universidade do Chile
Av. Beauchef 851, Santiago, Chile

RESUMO

Este artigo realiza o diagnóstico e prognóstico para dados de rolamentos que vibram até a falha advindos do IEEE PHM 2012 *Prognostic Challenge*. O diagnóstico é desenvolvido através da aplicação da FFT – Transformada Rápida de Fourier – observando-se a presença de frequências características de defeito nos dados de vibração. Na primeira amostragem, observa-se a ocorrência da BSF (frequência de rotação dos rolamentos), na segunda, múltiplos da BPFI (frequência na pista interna) e, na última, múltiplos da BPFO (frequência de passagem na pista externa). Para o prognóstico, houve o pré-processamento dos dados através das *Wavelets* e a previsão do RUL (*Remaining Useful Life*) utilizando o SVR (*Support Vector Regression*) + PSO (*Particle Swarm Optimization*). No prognóstico, observa-se um erro de previsão de RUL menor com os dados pré-processados quando comparado ao erro encontrado com os dados originais. Isso indica que o pré-processamento via *Wavelets* melhora a capacidade preditiva do SVR+PSO.

PALAVRAS CHAVE: Manutenção baseada na condição. Vibração. Support vector regression. Transformada Rápida de Fourier. Wavelets.

Tópicos: IND – PO na Indústria

ABSTRACT

This paper performs the diagnostics and prognostics from a run-to-failure data by the IEEE PHM 2012 Prognostic Challenge. The diagnostics is developed through the application of FFT – Fast Fourier Transform – observing the presence of characteristic frequencies of defect in vibration data. For the first sampling, we observe the occurrence of the BSF (frequency of rotation of the bearing), for the second sample, the BPFI (frequency of the inner race) multiples, and, for the last one, BPFO (frequency of the outer race) multiples were present. In the prognostics, the data were pre-processed through *Wavelets* and the prediction of RUL (*Remaining Useful Life*) was performed by SVR (*Support Vector Regression*) + PSO (*Particle Swarm Optimization*). In the prognosis, we observe a smaller RUL prediction error with pre-processed data than the error obtained with original data. This indicates that the data pre-processing via *Wavelets* enhances the SVR+PSO predictive capacity.

KEYWORDS: Condition-based maintenance. Vibration. Support vector regression. Fast Fourier transform. Wavelets.

Paper topics: IND- Operational Research in Industry



1. Introdução

A manutenção baseada na condição (*condition-based maintenance* - CBM) [Jardine et al. 2006] utiliza dados de funcionamento da máquina a fim de determinar em que estado ela se encontra e em que condição o equipamento falhou e, com isso, determinar como será realizado o reparo do equipamento. Para uma completa realização da CBM, é necessário haver as etapas de prognóstico e diagnóstico. A descrição do diagnóstico é dividida em: detecção, isolamento e identificação de uma falha [Dong e He 2007]; o prognóstico consiste em prever a condição do equipamento em situações futuras a partir da análise de dados de monitoramento. O presente artigo visa estimar a *remaining useful life* (RUL), ou seja, o quanto a máquina ainda permanecerá funcionando até ocorrer a próxima falha. Assim, essas etapas permitem que haja um estudo efetivo do equipamento permitindo tomadas de decisões mais adequadas.

Segundo Mirshawa e Olmedo (1993), os custos gerados pela manutenção são mínimos – custos com a mão de obra, ferramentas e instrumentos – se comparados aos custos de indisponibilidade do equipamento. Tais custos geram a perda de produção, a falta de qualidade dos produtos e de possíveis penalidades na imagem da empresa [Mirshawa e Olmedo 1993], sendo de suma importância práticas de manutenção e estudos de confiabilidade a fim de que a produção não sofra com a indisponibilidade dos equipamentos. Logo, o diagnóstico e o prognóstico dos equipamentos se tornam essenciais para que a produção flua e supra a demanda.

Por utilizar dados coletados diretamente da máquina em funcionamento, é comum que eles cheguem sem qualquer tipo de processamento prévio e com ruídos que escondem as características verdadeiras dos sinais. Algumas técnicas de pré-processamento, a exemplo das *Wavelets* [Morlet et al. 1982] – utilizadas no presente artigo, são utilizadas para processar os dados e retirar os ruídos existentes.

A partir da eliminação dos ruídos, a série temporal processada é utilizada em algoritmos de previsão a fim de melhorar seus resultados previstos. O artigo dá enfoque à previsão do RUL, uma métrica de confiabilidade bastante utilizada na CBM, por meio de-SVM (*Support Vector Machines*) [Moura et al. 2011] [Lins et al. 2015].

Para o diagnóstico, a transformada rápida de Fourier (FFT) [Cooley e Tukey 1965] foi o método utilizado para analisar a condição do rolamento durante o funcionamento até a quebra. Esse tipo de análise possibilita identificar as frequências das partes dos rolamentos que indicam a presença de algum modo de falha, tornando possível a realização de manutenções, evitando a parada total do sistema.

Os dados de vibração foram obtidos do IEEE PHM 2012 *Prognostic Challenge*, um desafio de confiabilidade que tinha por objetivo estimar o RUL – intervalo tempo restante até que a falha ocorra - de rolamentos levados a condições extremas em testes experimentais. A estimação do RUL e o diagnóstico da condição dos rolamentos são problemas críticos em indústrias.

Para o diagnóstico, há o cálculo das Frequências Características de Defeito (FCDs) que são utilizadas para a comparação com a saída da análise espectral realizada pelo FFT e, segundo [Tandon e Choudhury 1999], são selecionados os picos dos dados processados que estão relacionados aos múltiplos das FCDs. Essa análise torna possível localizar onde está o possível defeito no rolamento.

Já para o prognóstico, o pré-processamento através das *Wavelets* foi adotado, pois de acordo com Percival e Walden [2006], o método retira o ruído dos dados a partir de filtros que melhoram a qualidade do sinal. Para a previsão, o SVM – uma técnica de aprendizagem supervisionada [Schölkopf e Smola 2002] que visa, neste artigo, estimar a equação de regressão que representa o comportamento do equipamento até a falha – é uma metodologia preferível, pois não há a ocorrência de ótimos locais como pode ser observado nas Redes Neurais Artificiais [Boyd e Vandenberghe 2004][Schölkopf e Smola 2002]. A fim de melhorar o desempenho da metodologia empregada, utiliza-se o PSO (*Particle Swarm-Optimization*) – uma metaheurística de otimização probabilística inspirada no comportamento social de organismos vivos (por exemplo, aves e peixes) [Bratton e Kennedy 2007] –, para escolher os coeficientes do SVM [Lins et al 2011].



O artigo visa combinar o diagnóstico e o prognóstico de falha a partir da análise espectral (saídas do FFT), do pré-processamento dos dados pelas *Wavelets* e a previsão do RUL utilizando o SVM para um rolamento que funcione até a quebra. Essa metodologia visa alinhar o prognóstico juntamente com o diagnóstico a fim de realizar uma análise completa do comportamento do rolamento que sofre condições extremas e consequentemente prevenir a parada total do sistema sendo necessárias somente pequenas manutenções para trocar componentes específicos.

2. Abordagem Teórica

2.1. Estudo de Vibração em Rolamentos

Existem diversos métodos de monitoramento da condição operacional de um rolamento que ajudam a determinar quando ele está prestes a falhar ou não, evitando paradas dispendiosas do equipamento. Dados fornecidos pelos medidores fornecem a condição atual do equipamento [Tandon e Choudhury 1999]. Para realizar tais medições, o monitoramento da vibração do equipamento é a técnica mais difundida na previsão de falhas em mancais de rolamento e baseia-se na identificação de frequências características de cada componente do rolamento. Impactos periódicos causam aumento da vibração dos componentes e o consequente aumento da amplitude. Os principais componentes de um mancal de rolamento são: pista externa, gaiola, elementos rolantes e pista interna como se pode observar na **Figura 1** [Prabhakar et al. 2002].

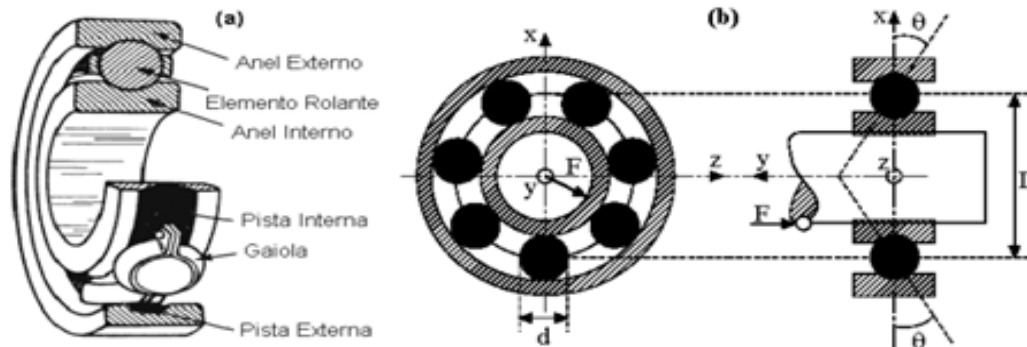


Figura 1:(a) Componentes de um rolamento; (b) dimensões do rolamento. **Fonte:** [Mesquita et al. 2002]

As frequências características de defeito (FCDs) são:

BPMI – Frequência de passagem dos elementos rolantes por um ponto da pista interna (*Ball Pass Frequency Inner*):

$$\omega_{id} = \frac{z\omega_s}{2} \left(1 + \frac{d}{D} \cos(\alpha)\right) \quad (1)$$

BPMO – Frequência de passagem dos elementos rolantes por um ponto da pista externa (*Ball Pass Frequency Outer*):

$$\omega_{od} = \frac{z\omega_s}{2d} \left(1 - \frac{d}{D} \cos(\alpha)\right) \quad (2)$$

BSF – Frequência de rotação dos elementos rolantes (*Ball Spin Frequency*):

$$\omega_b = \frac{D\omega_s}{2d} \left(1 - \frac{d^2}{D^2} \cos^2(\alpha)\right) \quad (3)$$

FTF – Frequência de rotação da gaiola (*Fundamental Train Frequency*):

$$\omega_f = \frac{\omega_s}{2} \left(1 - \frac{d}{D} \cos(\alpha)\right) \quad (4)$$



Nas Equações (1)-(4), ω_s é a velocidade de rotação do eixo, em rad/s; d é o diâmetro do elemento rolante; D é o diâmetro da gaiola; Z é o número de elementos rolantes; α é o ângulo de contato. Essas equações permitem o cálculo da frequência característica de vibração de cada parte do rolamento quando está em operação. O experimento realizado pelo IEEE PHM 2012 *Data Challenge* constituiu-se de um rolamento submetido à operação contínua e os dados de sua condição durante o funcionamento foram monitorados até a falha. Para o diagnóstico dos dados foi adotada a FFT, que faz uma análise no domínio da frequência permitindo a comparação das vibrações processadas pela FFT e as frequências características do rolamento.

2.2. Transformada Rápida de Fourier

A transformada de Fourier leva os dados do domínio do tempo para o domínio da frequência, o que possibilita a visualização do espectro de vibração e a identificação das frequências características dos componentes. A partir da análise espectral, pode-se identificar as frequências, compará-las com as FCDs e localizar qual parte dos componentes do rolamento estão sofrendo com algum tipo de defeito [Prabhakar et al. 2002].

A transformada rápida de Fourier é um algoritmo da transformada discreta de Fourier que reduz o número de cálculos necessários para um conjunto de N dados de $2N^2$ para $2N \log_2 N$. Considere a sequência $x[n]$ que é periódica em tempo T , ou seja:

$$x[n] = x[n + rT], \forall r \in \mathbb{Z} \quad (5)$$

Esta sequência pode ser representada por uma série de Fourier, dada por:

$$x[n] = \frac{1}{N} \sum_{m=0}^{\infty} x[m] e^{\frac{i2\pi mn}{N}} \quad (6)$$

A série de Fourier de um sinal periódico contínuo no tempo, em geral, precisa de infinitas exponenciais complexas. Porém, a série de Fourier para qualquer sinal discreto no tempo com período N requer apenas N exponenciais complexas, pois elas são periódicas. Assim, a representação da transformada torna-se:

$$x[n] = \frac{1}{N} \sum_{m=0}^{N-1} x[m] e^{\frac{i2\pi mn}{N}} \quad (7)$$

2.3. Wavelets

As *Wavelets* possuem aplicabilidade em dados não estacionários e que necessitam de um tratamento para retirada de ruídos e podem ser classificadas em dois tipos: a contínua e a discreta. Utiliza-se a transformada as *Wavelets* para realizar a análise dos dados de vibração.

O sinal analisado através das *Wavelets* pode ser dividido em duas partes: a de aproximação ($c_{j,k}$) – componentes de alta escala e baixa frequência, que mostram informações gerais sobre o sinal – e a de detalhe ($d_{j,k}$) – componentes de baixa escala e alta frequência, que mostram partes do sinal que não eram vistas antes devido aos ruídos. Tais coeficientes são dados a partir do algoritmo piramidal e são calculados a partir do algoritmo piramidal, mostrado da **Figura 2**, [Morettin 2014]. Esses coeficientes são usados para compor a função da transformada juntamente com as componentes chamadas de *Wavelet* mãe (Ψ) (**Equação 10**) e *Wavelet* pai (ϕ) (**Equação 9**) [Percival e Walden 2006].

$$f(t) = \sum_k c_{j_0,k} \phi_{j_0,k}(t) + \sum_{j \geq j_0} \sum_k d_{j,k} \psi_{j,k}(t) \quad (8)$$



$$\phi(t) = \sqrt{2} \sum_k l_k \phi(2t - k). \quad (9)$$

$$\psi(t) = \sqrt{2} \sum_k h_k \phi(2t - k). \quad (10)$$

j : Nível de resolução
 k : Fator de translação
 n : Máximo nível de resolução
 t : Tempo

Para a criação das *Wavelets*, faz-se necessária a escolha da base de filtros que será usada para analisar os dados. Dentre as bases existentes nas *Wavelets* discretas, foi escolhida a base de *Daubechies*, pois possui aplicações de sucesso no processamento de dados discretos [Daubechies 1993]. A base de *Daubechies*, calculada de forma iterativa, é constituída por uma família de *Wavelets* ortogonais representada por filtros de tamanho $N = 2p$, sendo p o número de momentos nulos – momentos nulos são elementos de valor zero colocados antes de realizar o algoritmo a fim de que o número de coeficientes da saída seja o desejado – e os coeficientes de saída localizados entre $[-p + 1, p]$ (suporte compacto) [Morettin 2014].

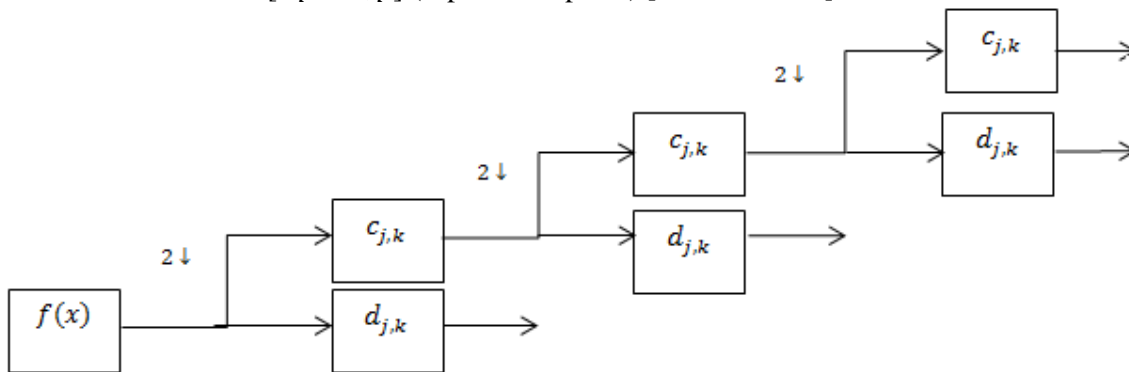


Figura 2: Algoritmo piramidal. **Fonte:** Elaborada pelo autor.

2.4. Support Vector Machines

SVM consiste em uma técnica de aprendizado que se baseia na Teoria Estatística de Aprendizagem (*Statistical Learning Theory* – SLT). A técnica será utilizada no estudo de confiabilidade, mais especificamente para a previsão de RUL. No SVM, o problema possui informação dos dados a partir do conjunto $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_T, y_T)\}$, em que T é a quantidade total de dados em D .

Os problemas de SVM podem ser divididos em dois tipos: os problemas de classificação e os problemas de regressão. O SVM empregado será o de regressão (SVR – *Support Vector Regression*) e é dividido em duas etapas: a de treinamento e a de teste. Na fase de treinamento, o algoritmo aprende o tipo de comportamento dos dados e escolhe quais os melhores coeficientes para a equação de regressão. No teste, é medido o grau de efetividade do treino. Nesta fase é possível realizar previsões dos comportamentos dos dados em termos de reconhecimento de padrão e de eventos futuros. Após a fase de aprendizagem, os dados de entrada do SVM são usados num problema de programação matemática quadrática e convexa, expresso a seguir [Schölkopf e Smola 2003]:

$$\min_{w,b,\xi,\xi^*} \frac{1}{2} \|w^2\| + C \sum_l (\xi_l + \xi_l^*), \quad l = 1, 2, \dots, T, \quad (11)$$

$$\text{s.a.} \quad y_l - w^T \phi(x_l) - b \leq \varepsilon + \xi_l, \quad \forall l, \quad (12)$$

$$-y_l + w^T \phi(x_l) + b \leq \varepsilon + \xi_l^*, \quad \forall l, \quad (13)$$

$$\xi_l, \xi_l^* > 0, \quad \forall l, \quad (14)$$

em que:

ε : Largura da função ε -insensitive.

C : Peso entre as duas partes da função objetivo.

ξ_l ou ξ_l^* : Variáveis de folga do erro relativo ao ponto l .



Após a aplicação das condições de KKT (Karush-Kuhn-Tucker), o problema primal transforma-se em um problema dual de maximização e, quando resolvido, encontra-se equação:

$$f(x) = \sum_i (\alpha_i - \alpha_i^*) K(x_i, x_j) + b \quad (15)$$

Se tanto α_i ou α_i^* (multiplicadores de Lagrange) são estritamente positivos, então o i -ésimo ponto de treinamento é um vetor de suporte e pode ser usado na função de regressão da Eq. (19). A fim de lidar com a não linearidade do SVM, aplica-se a função kernel para realizar a seleção de uma função de mapeamento adequada e o cálculo explícito dos produtos dos pontos com vetores de entradas mapeados, o que baixa o custo computacional. A função *kernel* $K(x_i, x_j)$ adotada é a função “*Gaussian Radial Basis*” (RBF), dada por $K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$ em que γ é parâmetro do modelo.

Para a escolha dos parâmetros mais adequados de entrada da equação do SVR, foi escolhido o PSO, que tinha por objetivo encontrar valores adequados para o γ , que representa a largura do *kernel*, o C , que mede a capacidade de previsão do modelo e o ε , que é a largura da função perda “ $\varepsilon - insensitive$ ”. Essa metodologia é utilizada com a finalidade de melhorar a previsão realizada pelo SVM e pode ser utilizada na resolução de problemas de confiabilidade.

3. Metodologia

Em relação ao estudo de diagnóstico (**Figura 4**), analisam-se os espectros de vibração que são comparados com as frequências fundamentais das partes do mancal de rolamento [Tandon e Choudhury 1999]. Para o prognóstico há o pré-processamento da série temporal (no caso específico deste artigo a série corresponde a dados de vibração) através das *Wavelets* e, logo em seguida, esses dados serão aplicados ao algoritmo de previsão, o SVR, quando há o cálculo da equação de regressão do RUL em função dos dados de vibração. Após tais etapas, serão obtidas as estimativas pontuais a partir da regressão linear da previsão do RUL através do SVR+PSO *versus* o índice – que tem relação com o tempo de operação do rolamento – e o cálculo do RUL estimado (**Figura 3**). Para a modelagem, foram utilizados os *softwares* MATLAB (*Wavelet Toolbox*) e R (pacote *kernelab*).

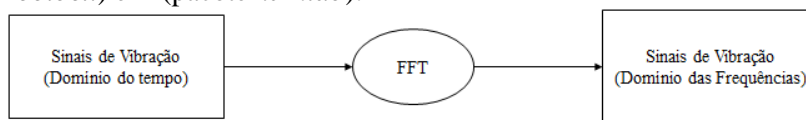


Figura 3: Metodologia do diagnóstico. **Fonte:** Elaborada pelo autor.

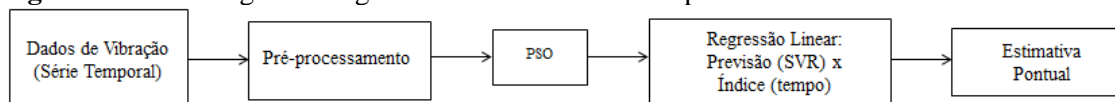


Figura 4: Metodologia do prognóstico. **Fonte:** Elaborada pelo autor.

4. Aplicação Prática: IEEE PHM 2012 Data Challenge

4.1. Diagnóstico

Para realizar o diagnóstico – como observado na **Tabela 1**, o conjunto de dados original foi analisado através do comportamento de determinado minuto durante a parte do começo, do meio e do fim dos dados, sem nenhum tipo de processamento.

Para realizar a avaliação dos dados foram escolhidos os pontos de picos observados em intervalos de frequências que possuíam visualização mais relevante com relação ao gráfico, pois indica que ao ocorrer tal pico, a vibração teve mais impacto em relação a outras frequências. Essas frequências observadas durante as amostragens indicam em que parte do rolamento, provavelmente, está ocorrendo o defeito. Por exemplo, nas amostras em que ocorrem frequências relacionadas ao BSF, deve-se levar em consideração, durante a manutenção, os elementos rolantes. Esse procedimento evita a troca de componentes do rolamento de forma desnecessária e consequentemente gastos com a substituição do rolamento como um todo.



Tabela 1: Amostragem para o diagnóstico.

	Tempo (em segundos)	
	<i>Bearing1</i>	<i>Bearing3</i>
Começo	0 ao 60 s	0 ao 60 s
Meio	14000 ao 14050 s	12870 ao 12920 s
Fim	27980 ao 28030 s	23700 ao 23750 s

Fonte: Elaborada pelo autor.

A partir da análise espectral (saída da FFT), as frequências de rotação do rolamento são observadas. Os dados analisados foram advindos do conjunto original de treino e teste do *Bearing1* e *Bearing3*, e foram divididos em três partes: foi analisado o primeiro minuto do experimento, depois foram processados os conjuntos que correspondiam a um minuto que se localizava no meio do conjunto de dados e, por último, os arquivos que compõem o último minuto da rotação do rolamento.

Essa amostragem foi escolhida, pois analisa o sinal original observando-se a mudança da vibração ao longo do tempo. A saída do FFT foi relacionada às FCDs a partir do uso na **Tabela 2**. Comparando-se o primeiro minuto com o último, nota-se um aumento na amplitude de rotação do rolamento, indicando a proximidade da falha do rolamento. Em contra partida, percebe-se algumas frequências durante o funcionamento que indicam a degradação de máquina rotativa em determinadas partes do equipamento, mais especificamente na pista interna, pista externa, na gaiola e no elemento rolante. Tais frequências são comparadas com as FCDs e indicam defeitos devido ao mau funcionamento de uma das partes do rolamento citadas na **Figura 1**.

De uma forma geral, os defeitos nos rolamentos são classificados em locais e distribuídos. Os defeitos distribuídos são rugosidades superficiais, ondulações, desalinhamento entre as pistas rolantes e elementos rolantes que foram montados de maneira incorreta, esses tipos de defeitos são causados por erros de fabricação, instalação inapropriada ou desgastes dos componentes do rolamento. Já os defeitos localizados, são trincas, corrosões, lascas ou estilhaços nas superfícies de rotação [Tandon e Choudhury 1999].

Na segunda amostragem, associada-ao *Bearing1* - como nota-se na **Tabela 3** e a **Figura 5** – a qual mostra alguns pontos da saída do FFT -, constatou-se a ocorrência de um múltiplo da BPFI (678,17 Hz), porém não há registros de nenhuma frequência que seja igual à BPFI. Nesse mesmo conjunto, houve a ocorrência de dois pontos que se aproximavam do valor da BSF (389,51 Hz e 354,20 Hz), já no último conjunto de pontos notam-se altas frequências de rotação indicando a quebra do rolamento e um provável defeito na pista externa devido ao aumento tanto da aceleração quanto da frequência. Para os dados do *Bearing3*, a partir da análise da **Tabela 4**, na primeira amostragem, observou-se a ocorrência de um ponto próximo à BSF (139,01 Hz). Já para o segundo conjunto, houve a ocorrência de dois pontos múltiplos do BPFI um próximo ao outro e, para a última amostragem, nota-se uma frequência de valor próximo ao BPFO e outras que indicam a possível presença de um defeito na pista externa do rolamento.

Tabela 2: Frequências características de defeito do rolamento estudado.

Frequência Característica do Defeito (em Hz) ($\alpha = 0^\circ$)	
BPFI	218,4
BPFO	49014,3
BSF	122,9
FTF	13,2

Fonte: Elaborado pelo autor.

Tabela 3: Frequências de pico para cada conjunto do *Bearing1*.



Frequências Começo (Hz)	
Ponto	Frequência
1	358,99 (x3BSF)
2	287,91 (x3BSF)
3	598,66
4	518,57
5	206,77
6	149,10

Frequências Meio (Hz)	
Ponto	Frequência
1	678,17 (x3BPFI)
2	440,4 (x2BPFI)
3	389,50 (x3BSF)
4	354,20 (x3BSF)
5	282,47
6	105,79

Frequências Fim (Hz)	
Ponto	Frequência
1	2775,00
2	7891,59
3	6423,76
4	4506,93
5	3506,79
6	2313,89

Fonte: Elaborado pelo autor.

Tabela 4: Frequências de pico para cada conjunto do *Bearing3*.

Frequências Começo (Hz)	
Ponto	Frequência
1	192,11
2	311,32
3	166,06
4	270,23
5	212,22
6	139,01 (x1BSF)

Frequências Meio (Hz)	
Ponto	Frequência
1	581,82
2	227,78 (x1BPFI)
3	407,35 (x2BPFI)
4	167,31
5	299,33
6	130,04

Frequências Fim (Hz)	
Ponto	Frequência
1	3038,38
2	4818,00 (x1BPFO)
3	10757,93
4	6531,71
5	4255,64
6	2749,64

Fonte: Elaborado pelo autor.
Frequências do Meio.

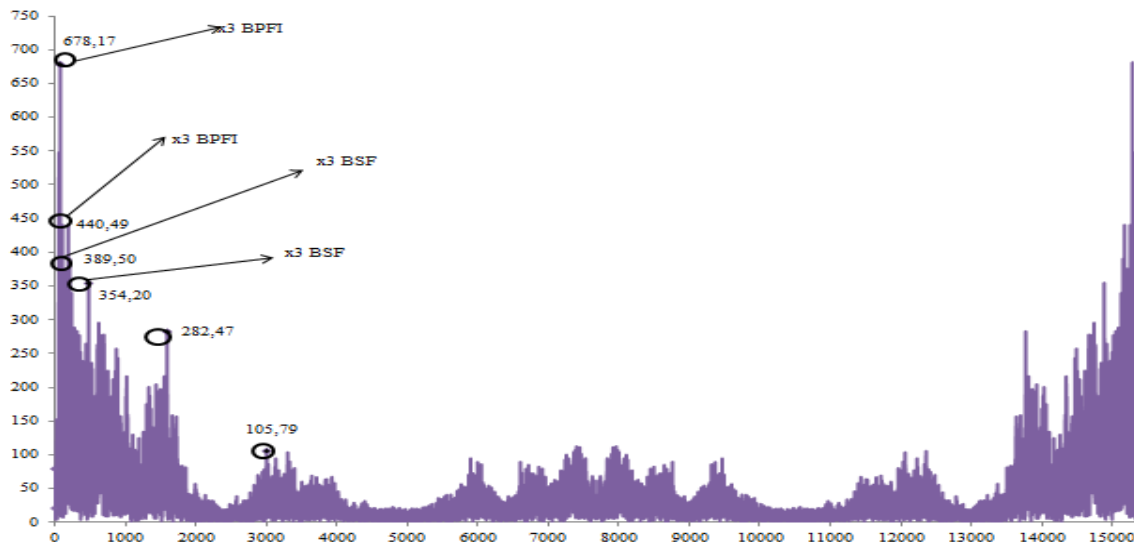


Figura 5: Segunda amostra do *Bearing1* processado pela FFT. Fonte: Elaborado pelo autor.



4.2. Prognóstico

Os dados utilizados para teste na pesquisa foram provenientes de um desafio focado na estimação do RUL de rolamentos realizado pelo IEEE PHM 2012 *Data Challenge* (IEEE Reliability Society e FEMTO-ST Institute) e as condições do experimento foram: uma velocidade de 1800 rpm juntamente com uma carga de 4000N. Os dados de vibração do rolamento foram analisados no MATLAB, sendo o conjunto de treinamento chamado “*Bearing1*” – que teve medidas do sinal de vibração até a falha completa – e o conjunto de teste de “*Bearing3*”. Foi utilizado o filtro de *Daubechies* com quatro níveis de decomposição para os dois sinais.

Primeiramente, treina-se com os dados do *Bearing1* para calcular a equação de regressão que irá ser utilizada na fase de teste. Os dados do *Bearing1* são formados por pares (x_t, y_t) em que x_t é um valor de aceleração e y_t , o RUL correspondente no tempo t . Para a escolha dos coeficientes do SVR, utilizou-se o PSO a fim de treinar os dados com e sem processamento. Para realizar a previsão do RUL através dos dados de teste, usa-se o conjunto de dados truncados (obtidos até um momento prévio ao da falha do rolamento) fornecidos no desafio. Há um total de 1802 pares (x_t, y_t) no conjunto de dados truncados.

As previsões de RUL obtidas do SVR+PSO são relativas a cada uma das acelerações e seus respectivos parâmetros de entrada são representados pela **Tabela 5**. No entanto, é necessário obter uma estimativa pontual para o RUL. Para isso, realiza-se uma regressão linear das previsões do RUL vs. índice (que guarda relação direta com o tempo). Como resultado da regressão linear, duas estimativas pontuais de RUL são encontradas: (i) o valor que cruza o eixo y e (ii) valor que cruza o eixo x. Ambas são usadas para o cálculo do erro percentual absoluto dado pela Equação 16:

APE

$$= \frac{(|\text{Real RUL} - y \text{ cross value}| + |\text{Real RUL} - (x \text{ cross value})|)}{2 \cdot \text{Real RUL}} \quad (16)$$

RealRUL: valor total do RUL

Tabela 5: Valores dos parâmetros usados no SVR.

Valores dos Parâmetros do SVR		
Dados truncados sem processamento (SVR + PSO)		
γ	C	ε
84.2407	15.0837	0.0070
Dados truncados com processamento (SVR+PSO)		
γ	C	ε
81.3929	15.1844	0.04874

Fonte: Elaborada pelo autor.

γ :	parâmetro do <i>kernel</i>
C :	mede a capacidade de previsão do modelo
ε :	largura da função perda " ε – insensitive"

Antes de alimentar o modelo de regressão linear RUL vs. índice, foi aplicada uma amostragem dos dados, que ocorreu de tal forma, segundo a **Tabela 6**:

Tabela 6: Amostragem realizada no conjunto de 1802 pontos.

Tempo de Amostragem (em s)	Intervalo de Amostragem (em s)
500	12750 a 23750
200	5950 a 12751
100	5730 a 5951



Fonte: Elaborada pelo autor.

Após essa amostragem, os dados foram sofrerem *upsampling*, ou seja, voltaram ao tamanho inicial da amostra, sendo os locais que não possuíam dados da amostra anterior completados com espaços em branco. Posteriormente, foram retirados os 1120 primeiros pontos dos resultados do RUL previsto através do SVR+PSO, pois, no início da rotação, a vibração possui pouco impacto no rolamento em si, o que não é tão significativo para a representação da falha e, em seguida, foram atribuídos pesos aos dados, como descrito na **Tabela 7**. Essa amostragem e atribuição de pesos têm por finalidade adequar melhor os dados ao modelo de regressão linear, que é realizada entre o RUL previsto vs. índice, representados na **Figura 6** e na **Figura 7** para o caso sem processamento e com processamento, respectivamente.

Tabela 7: Pesos atribuídos aos dados amostrados.

Pesos	Intervalo de amostragem (índice em s)
3	1121 a 1790
4	1791 a 1802

Fonte: Elaborada pelo autor.

Ao realizar a regressão linear RUL vs. índice, obteve-se duas estimativas para o RUL – uma para abcissa e outra para a ordenada – e foram calculadas as médias entre elas gerando um RUL de 3121,06s para os dados originais truncados e um RUL de 2840,73s para os dados processados truncados, como observado na **Tabela 8**. Para um RUL de 5730s fornecido pela organização do desafio, obteve-se a partir da (**Equação 16**), utilizando os RULs previstos tanto na abcissa quanto na ordenada, um erro percentual absoluto de 67,13% para os dados originais truncados e um erro de 45,53% para os dados processados truncados, havendo uma melhora da previsão a partir do pré-processamento dos dados por meio das *Wavelets*.

Se comparado ao vencedor do desafio, que obteve um erro de 37% Bearing3 [Sutrisno et al 2012], a previsão realizada em tal artigo pode ser considerada condizente com o modelo proposto. O erro tende a ter valores altos, pois os dados de vibração possuem uma variação muito grande, sendo a intenção principal obter um modelo que represente o comportamento do rolamento. Por exemplo, em Sutrisno et al. [2012], observa-se um erro de 80% para o resultado da previsão do quarto conjunto dos dados de teste que possuíam as mesmas condições de operação do *Bearing3*.

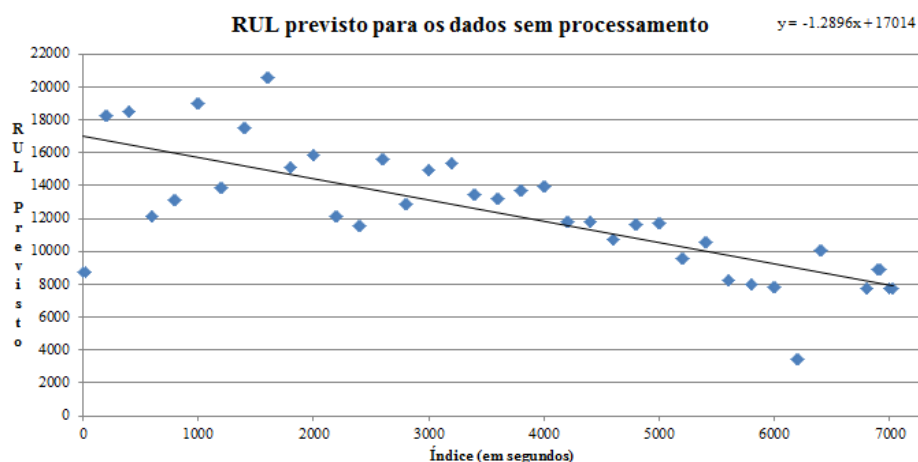


Figura 6: RUL previsto a partir dos dados *Bearing3* truncados e sem processamento e a reta tendência. **Fonte:** Elaborado pelo autor.

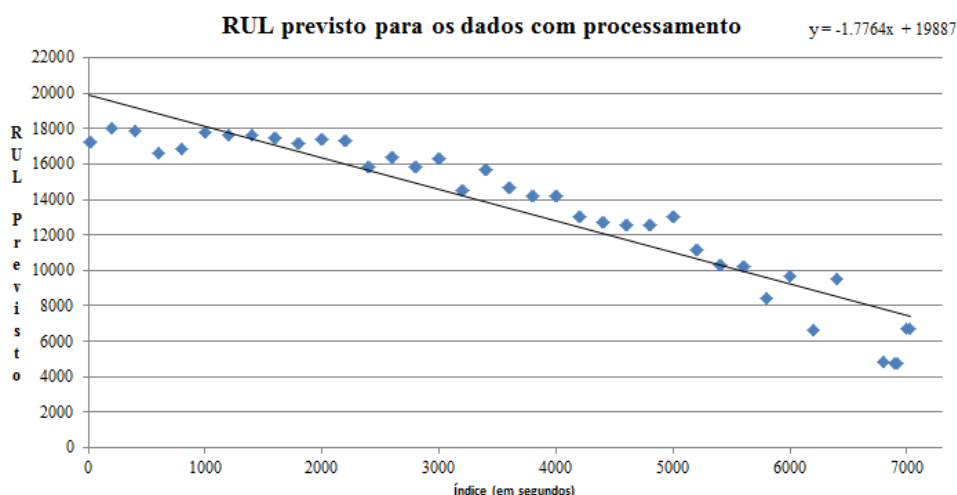


Figura 7: RUL previsto a partir dos dados *Bearing3* truncados com processamento e a reta tendência. **Fonte:** Elaborado pelo autor.

Tabela 8: Resultados estimação do RUL.

Previsão dados originais truncados	Previsão dados processados truncados
$RUL_{estimado} = 3121,06s$	$RUL_{estimado} = 2840,23s$
Erro Percentual Absoluto = 67,13%	Erro Percentual Absoluto = 45,53%
$y = -1,260x + 17014$	$y = -1,776x + 19887$

Fonte: Elaborado pelo autor.

5. Conclusão

Neste trabalho, realizou-se o diagnóstico e o prognóstico de falha a partir de sinais de vibração de rolamentos. Para o diagnóstico, realizou-se a análise espectral a partir da FFT onde foram observadas as frequências da rotação do rolamento. Notou-se uma semelhança entre quais tipos de FCDs ocorreram nas amostragens selecionadas. Para os dados do começo houve a ocorrência de múltiplo da BSF frequência associada ao elemento-rolante, já para o segundo conjunto da amostragem, houve a ocorrência de múltiplos do BPFI – frequência associada à pista interna do rolamento- e para a amostragem do fim foram registradas frequências associadas à pista externa (BPFO). A identificação de tais frequências sinaliza a possibilidade de falhas ligadas a esses pontos do rolamento, o que ajuda na realização da manutenção baseada na condição atual do equipamento.

Já para o prognóstico dos dados do IEEE PHM 2012 *Prognostic Challenge*, primeiramente foi realizado , o pré-processamento dos dados através das *Wavelets* a fim de eliminar ruídos. O RUL foi previsto a partir do SVR + PSO [Lins et al 2012], um algoritmo de regressão advindo do SVM, que faz uso de aprendizado supervisionado para calcular a equação de regressão e realizar previsões. A FFT foi o método escolhido para o diagnóstico do rolamento e tem a capacidade de levar os dados do domínio do tempo para o domínio da frequência, o que ajuda a realizar o estudo da FCDs e relatar as possíveis causas da falha. Os dados da pesquisa são acelerações de rolamentos de um experimento até a falha e são divididos em conjunto de treinamento e teste a fim de realizar a previsão.

Para a previsão realizada através do SVR+PSO foi observado um erro percentual absoluto de 67,13% para os dados originais truncados e um erro de 45,53% para os dados processados truncados, observando-se um erro de 37% para o vencedor do desafio [Sutrisno et al 2012], havendo uma melhora no ajuste dos dados a partir do pré-processamento. Se comparado aos resultados das outras previsões obtidas por Sutrisno et al. [2012], por exemplo, um erro de 80% para a previsão de outro rolamento submetido as mesmas condições do *Bearing3*, contata-se que o pré-processamento dos dados trouxe uma melhoria para a estimação do RUL no rolamento analisado, concluindo que o pré-processamento dos dados é de significativa importância para a melhora dos resultados da previsão.



Referências

- Ali J. B., Chebel-Morello B., Saidi L., Malinowski S. (2015). Accurate bearing remaining useful life prediction based on Weibull distribution and artificial neural network. *Mechanical Systems and Signal Processing*, 56-57: 150–172.
- Boyd S., Vandenberghe L. (2004). *Convex Optimization*. Cambridge University Press, Cambridge.
- Bratton, D., & Kennedy, J. (2007). Defining a Standard for Particle Swarm Optimization. *2007 IEEE Swarm Intelligence Symposium*, 120-127. <https://doi.org/10.1109/SIS.2007.368035>.
- Cooley, J.W. & Tukey J.W.(1965). An Algorithm for the Machine Computation of Complex Fourier Series. *Mathematics of Computaton*, 19: 297-301.
- Daubechies, I. (1993). Ten Lectures on Wavelets. *SIAM Review*. <https://doi.org/10.1137/1035160>.
- IEEE PHM 2012 Prognostic Challenge. Scoring of results and application procedure. Web page. <http://www.femto-st.fr/f/d/IEEE-Challenge-Appli.pdf>. Acessado: 2016-08-15.
- Jardine, A.K.S., Lin, D. & Banjevic, D., (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, 20: 1483–1510.
- Karatzoglou A., Smola A., Kurt H. (2016). Kernel-Based Machine Learning Lab. CRAN-R.
- Lins, I.D., Moura, M.C., Zio E., Droguett, E.L. (2011). A particle swarm-optimized support vector machine for reliability prediction. *Quality and Reliability Engineering International*. 28: 141-158.
- Maior, C.B.S. (2016) Remaining useful life prediction analysis via emd and wavelets preprocessing a particle swarm-optimized support vector machine model. Dissertação de Mestrado- Universidade Federal de Pernambuco.
- Mesquita A.L.A., Santiago D.F.A., Bezerra R.A., Miranda U.A., Dias M., Pederiva R. (2002), Detecção de falhas em rolamentos usando transformada tempo-frequência – comparação com análise envelope. *Mecânica Computacional*. 21: 1938-1954.
- Mirshawka, V. & Olmedo, N.C., (1993). Manutenção – combate aos custos na não-eficácia – a vez do Brasil. São Paulo: Editora McGraw-Hill Ltda.
- Misiti M., Misiti Y., Oppenheim G., Poggi J.M. (2015). Wavelet Toolbox User’s Guide. *The MathWorks Inc*.
- Morettin, P. A. (2014). ONDAS E ONDALETAS. Da Análise de Fourier à Análise de ondaletas de Séries Temporais. Editora da Universidade de São Paulo, São Paulo.
- Morlet. J., Arens, G., Fougéau, E., Giard, D. (1982). Wave propagation and sampling theory – Part II: Sampling theory and complex waves. *Geophysicis*. 47(2): 222-236. <https://doi.org/10.1190/1.1441329>.
- Moura, M. C.; Zio, E.; Lins, I. D.; Droguett, E. L.(2011) Failure and reliability prediction by support vector machines regression of time series data. *Reliability Engineering & System Safety*, 96: 1527-1534.
- Percival, D.B. & Walden, A.T., (2006). *Wavelet methods for time series analysis*. Cambridge University Press, New York.
- Prabhakar S., Mohanty A.R., Sekhar A.S. (2002). Application of discrete wavelet transform for detection of ball bearing race faults. *Tribology International*. 35:793–800.
- Schölkopf, B. & Smola, A.J. (2002). *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MA: The MIT Press, Cambridge.
- Shumway R. H. e Stoffer D. S.(2011). *Time series analysis and its applications with R examples*. Springer, 3rd Editon, California.
- Sutrisno, E., Oh, H., Vasan, A. S. S., e Pecht, M. (2012). Estimation of remaning useful ball bearings using data driven methodologies. *2012 IEEE Conference on Prognostics and Health Management*, 2, 1-7. <https://doi.org/10.1109/ICPHM.2012.6299548>.
- Tandon N. & Choudhury A. (1999). A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings. *Tribology Internacional*. 32: 469- 480.
- Torrence C. & Compo G. (1998). A Practical Guide to Wavelet Analysis. *Program in Atmospheric and Oceanic Sciences*, 79: 61- 78.