COBEM 2021 26th International Congress of Mechanical Engineering
ABCM Associação Brasileira de Engenharia e Ciências Mecânicas
November 22-26, 2021, Virtual Congress, Brazil

COB-2021-0193
SPECTRAL KURTOSIS-BASED UNSUPERVISED LEARNING METHOD FOR ESTIMATING REMAINING USEFUL LIFE OF ROTATING MACHINERY

Leonardo Franco de Godói
Eurípedes Guilherme de Oliveira Nóbrega
University of Campinas
leonardofrancogodoi@gmail.com, egon@fem.unicamp.br

Abstract. *Predicting the remaining useful life of rotating machines has been established as one of the biggest challenges in the area of machine health monitoring. In the last few years, a wide variety of techniques have been developed aiming at the prognosis of rotating machines, given their importance in different types of industrial processes and the subsequent criticality of failures related to specific mechanical components such as bearings and gearboxes. In this scenario, supervised and unsupervised machine learning methods have assumed great relevance due to their performance and flexibility compared to traditional ones. Recent works have studied the applicability of statistical tools for extracting features from raw signals collected from the vibration of such machines. Kurtosis, as the fourth central statistical moment, may indicate non-Gaussian tail behavior in a signal. The spectral kurtosis can find the respective components in the frequency domain that are responsible for this behavior. Therefore, it is possible to generate a spectral kurtosis-based health indicator model for the machines to be used to estimate a prognosis. The objective of this work is the proposal of a novel method for the estimation of the remaining useful life of rotating machinery through submission to an unsupervised neural network of the spectral kurtosis analysis of vibration signals, as a time-varying sequence of features, in order to identify any abnormal behavior and predict operation failures. The analysis is based on kurtogram images, which represent the spectral kurtosis calculated for multiple frequency domain window sizes along the machine operation. The images are used as inputs to a deep autoencoder neural network. The remaining useful life is estimated based on the distance between a set of samples at a given time and a normal behavior model, defined by the error resulting from the autoencoder reconstruction process. Vibration data from a public dataset containing signals extracted from faulty rolling bearings is analyzed to assess the method performance. The achieved results permit to affirm that the proposed method may be successfully adopted to estimate the remaining useful life of rotating machines.*

Keywords: *Spectral kurtosis, artificial intelligence, neural networks, rotating machines, prognostics.*

1. INTRODUCTION

Prognostic systems focused on rotating machines have become essential in different industry sectors, since they increase the process reliability and assist in decision making to optimize the use of critical equipments in the short and long term (Okoh *et al.*, 2014).

One of the most important information that can be obtained from a rotating machine prognostic system is its remaining useful life (RUL), which, if accurately predicted, can avoid catastrophic failures and significantly minimize the maintenance costs. This analysis is usually done using health indicators, whose role in RUL prediction is to clearly present the progression of potential faults in a specific component by detecting and describing the trend of its degradation process (Qin *et al.*, 2017).

Several techniques can be applied for the prognosis of faults in mechanical components by extracting robust, meaningful health indicators. Such techniques may be based on the analysis of signals in the time-domain, frequency-domain or both (Atamuradov *et al.*, 2020). The main advantage of the latter approach is the ability to handle non-stationary signals, whose statistical properties vary over time, which is very common in the operation of mechanical systems (Vicuña and Acuña, 2014). The work of Yoo and Baek (2018) is an example of the use of the continuous wavelet transform as a time-frequency domain processing technique for the RUL prediction of bearings. Other methods, such as those proposed by Liu *et al.* (2016) and Soualhi *et al.* (2014), make use of the short-time Fourier transform (STFT) and the Hilbert-Huang Transform (HHT), which are also very popular time-frequency techniques applied for the condition monitoring of rotating machines. Lately, the spectral kurtosis (SK), which consists in characterizing the local fault information of non-stationary components by calculating the kurtosis in the frequency domain, has received prominence due to its versatility, and started to be applied as an alternative to the previously mentioned techniques (Udmale *et al.*, 2018).

Many of the recently developed methods combine processing spectral analysis with deep learning tools, including neural networks trained in an unsupervised fashion. In the study of rotating machines, the analysis in the frequency domain helps in the task of extracting attributes from raw vibration signals, which are then refined and interpreted using neural networks.

This work proposes a method for estimating the RUL of rolling element bearings from features extracted by the calculation of the spectral kurtosis of vibration signals, and the subsequent processing by an autoencoder (AE) neural network, aimed at learning patterns of the input data through unsupervised training. The health indicator is here defined as the error given by the root mean square difference between the current signal and the signal reconstructed by the network model trained with normal vibration data. Exponential regression is applied on the error to describe a degradation model and to predict the future health condition of the bearing. A threshold indicates the point in time at which it is considered that the machine must have its operation interrupted. Therefore, the RUL is the time between the instant for which the threshold is reached and the current analysis time. The method is evaluated using a rolling bearing vibration public dataset, whose data where collected from an experimental test rig containing a degradation generation system, which simulates component wear in a few hours of operation. The obtained results are compared to the expected RUL values reported by the dataset providers for multiple runs carried out with different rolling bearings on the same test rig under similar operating conditions.

2. THEORETICAL BACKGROUND

The main concepts that were applied in this work, including feature extraction with SK technique and the generation of a health indicator through the use of an AE network, are presented in the following.

2.1 Spectral kurtosis of vibration signals

The spectral kurtosis, originally introduced by Dwyer (1983) as a statistical technique to locate non-Gaussian frequency components in a signal, has proven to be efficient in detecting transients in noisy signals. The SK has also been defined by Antoni (2004) as the normalised fourth-order spectral cumulant of a conditionally non-stationary (CNS) process Y:

$$K_Y(f) = \frac{S_{4Y}(f)}{S_{2Y}^2(f)} - 2, \quad f \neq 0 \tag{1}$$

where S_{4Y} and S_{2Y} are the fourth and second-order spectral moments of Y calculated along the frequency f.

Although this definition was developed for stationary cases, it can be maintained for the analysis of CNS signals, such as those on which estimators have been proposed. The most popular estimators are based on the concepts of spectrogram, which is a matrix composed of STFT coefficients arranged as columns. Considering a hop of size M and an analysis window $w(n)$, the spectrogram is described as (Sutawanir, 2017):

$$S(j,l) = \frac{1}{MW_nN} \left| \sum_{n=1}^{N} y(n+lM)w(n)e^{-j\frac{2\pi nj}{N}} \right|^2 \tag{2}$$

where l is the time frame index and j is the frequency bin. Thus, S has the format $J \times L$, considering that J is the number of frequency bins of the one-sided STFT and L is the number of time frames.

According to Antoni (2006), an STFT-based estimator of the SK has the role of connecting theoretical results with real practice, but it is affected by two sources of bias, one due to finite sample effects, and other due to leakage effects arising from the analysis window. An unbiased estimator was proposed by Vrabie *et al.* (2003) for using the SK to measure the distance to gaussianity of different spectral components, in which the SK at all frequency bins $0 \leq j \leq J-1$ is calculated as follows:

$$\hat{K}_S(j) = \frac{L}{L-1}\left[\frac{(L+1)\sum_{l=1}^{L}|S(j,l)|^4}{(\sum_{l=1}^{L}|S(j,l)|^2)^2} - 2\right], \quad 0 \leq j \leq J-1 \tag{3}$$

In this case, it is established that the L frames are non-overlapping, independent stochastic realizations of J-sample signal.

2.2 Kurtograms

The kurtogram is an analysis tool based on the SK calculation commonly used to determine the optimal filter, which has proven to be a powerful alternative to filter vibration signals in several applications involving bearing monitoring (Huang *et al.*, 2018). A so-called 1/3-binary tree expands the concept of a binary tree by increasing the number of filters in order to enrich the structure of the representation. Figure 1 presents a kurtogram paving of the plane. The kurtogram aims at computing the kurtosis at each frequency line in order to detect the hidden non-stationarities in a signal and their respective frequency bands (Tafinine and Mokrani, 2012). Since the SK gives the optimal combination of a frequency (f) and a frequency resolution (Δf), the kurtogram may be represented in (f, Δf) plane with:

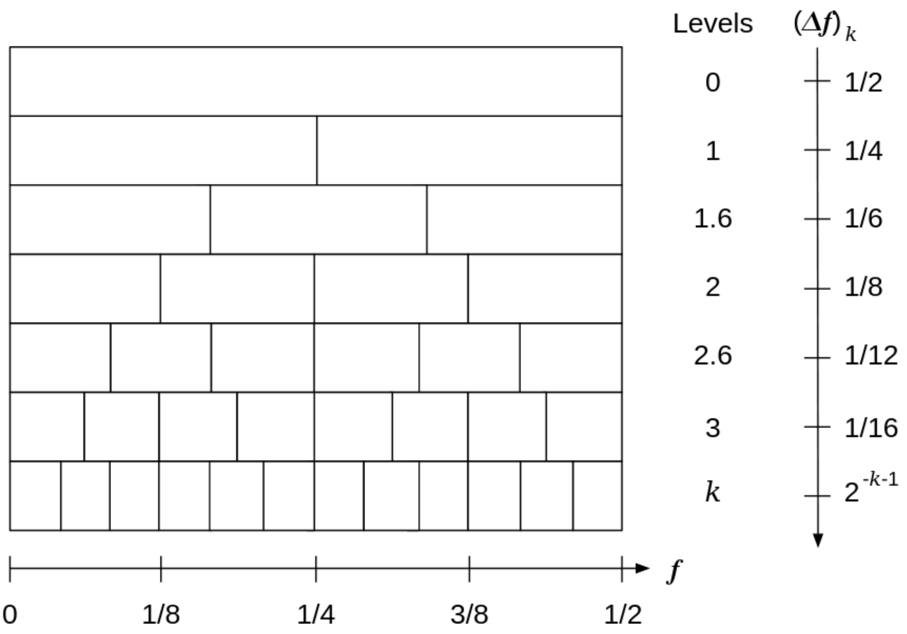

Figure 1. Frequency/frequency resolution plane of the 1/3-binary tree kurtogram.

$$f_i = (i + 2^{-1})2^{-k-1} \tag{4}$$

and

$$(\Delta f)_k = 2^{-k-1} \tag{5}$$

The kurtogram has been seen as an efficient approach for both online and offline automatic fault detection without manual localization of the associated frequency band, overcoming the innefectiveness of the power spectral density to characterize transients in a signal (Udmale *et al.*, 2018).

2.3 Autoencoder neural networks

An autoencoder is a special type of neural network that is trained in an unsupervised fashion to learn features from a given dataset for achieving dimensionality reduction (Shrestha and Mahmood, 2019). The unsupervised learning paradigm consists in training the network to produce as output the data which was provided as input. In comparison to the principal component analysis (PCA), which is a linear dimensionality reduction method, a deep AE, composed of multiple hidden layers, is an approach able to extract non-linear features from data (Ren *et al.*, 2018).

The basic AE structure consists of two stages: an encoder for mapping the input data into a latent space, and a decoder for reconstructing the input data from the latent representation (Liu *et al.*, 2018). In a shallow, fully connected AE, the encoded features, represented by the vector h, are obtained from an input vector x by the following equation:

$$h = \phi(W^{(e)}x + b^{(e)}) \tag{6}$$

where ϕ is a non-linear activation function, $W^{(e)}$ is the weight matrix of the connections between the input and the hidden layer, and $b^{(e)}$ is the bias vector of the hidden layer.

The decoder can then be described from the encoded data as follows:

$$\hat{x} = \phi(W^{(d)}h + b^{(d)}) \tag{7}$$

where $W^{(d)}$ is the weight matrix of connections between the hidden layer and the last layer, $b^{(d)}$ is the bias vector of the last layer, and $\hat{x}$ is the reconstructed input vector.

In the recent years, deep AEs have been widely applied for the analysis of rotating machines because they are able to perform automatic feature extraction, reducing dependence on traditional signal processing techniques and minimizing the influence of noise on the vibration signal (Shao *et al.*, 2017).

3. THE PROPOSED METHOD

The developed methodology aims to take advantage from the kurtogram capacity of characterizing vibration signals and the power of the autoencoder neural network to extract complex features from time series.

3.1 Obtaining the health indicator

The degradation of a component health must be described by a specific indicator capable of representing transients and trends resulting from its operation over time. Such indicator is produced by performing a sequence of steps that takes into account both normal and degrading operations. Figure 2 presents the pipeline adopted to implement the proposed method. This process consists of two phases, starting with the providing of raw vibration signals and ending with the obtaining of the degradation curve of the analyzed component. In the first phase, the first 25% measured samples of a run-to-failure vibration series are used for building a model. For this purpose, the respective kurtograms to be used as inputs are generated and, at the end of this phase, the model is defined by an AE trained to reconstruct such kurtograms and a vector with the averages of the information contained in the central layer.

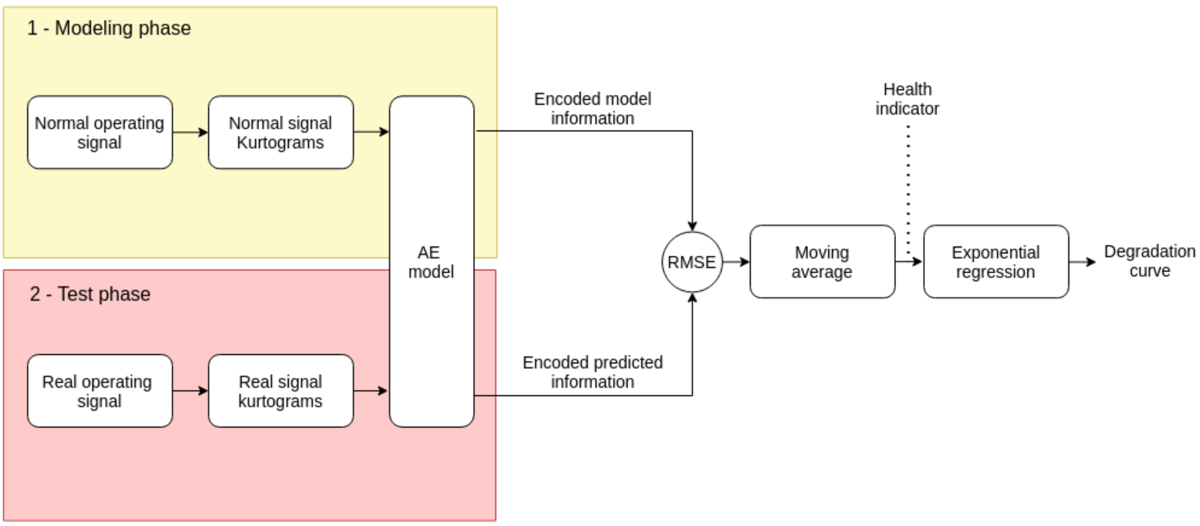

Figure 2. Two-phase pipeline of the proposed method.

It can be seen in the diagram in Figure 2 that the indicator at a certain instant is given by the root mean square error (RMSE) resulting from the comparison between the encoded vector of the normal operation model of the component and the encoded vector of the segment at that instant. The sequence of indicators is subsequently submitted to the moving average (MA) calculation, so that variations are smoothed and transients that do not contribute to the characterization of the trend for the RUL are eliminated. The degradation curve is then generated by applying an exponential regression on the sequence of obtained indicators.

3.2 Kurtogram processing with an autoencoder

One of the crucial parts of the pipeline presented in the previous subsection is the AE model. The neural network must be trained in order to capture intrinsic, deep patterns of the input data, so that the difficulty in reconstructing unknown signals, such as those that indicate the development of some failures, results in higher error values, characterizing a behavior that can be easily verified through the corresponding health indicator.

The scheme shown in Figure 3 illustrates the processing of kurtograms by using an AE with three hidden layers. The encoder contains the layers represented by the vectors h_1 and h_2, and the decoder contains the layers h_2 and h_3. It is important to note that the sizes of these layers must respect the relation $n > m > p$, where n denotes the number of neurons in the input, m denotes the number of neurons in the first and in the third hidden layer, and p denotes the number of neurons in the central hidden layer of the AE. Therefore, the encoded information from the neural network, which is used to characterize both the model and the test signals, is located at the output of the layer $h2$.

Because it is a fully connected neural network with one-dimensional input, each kurtogram is flattened before being effectively processed by the AE. Thus, the reconstructed kurtogram can be obtained by resizing the vector observed at the output of the network.

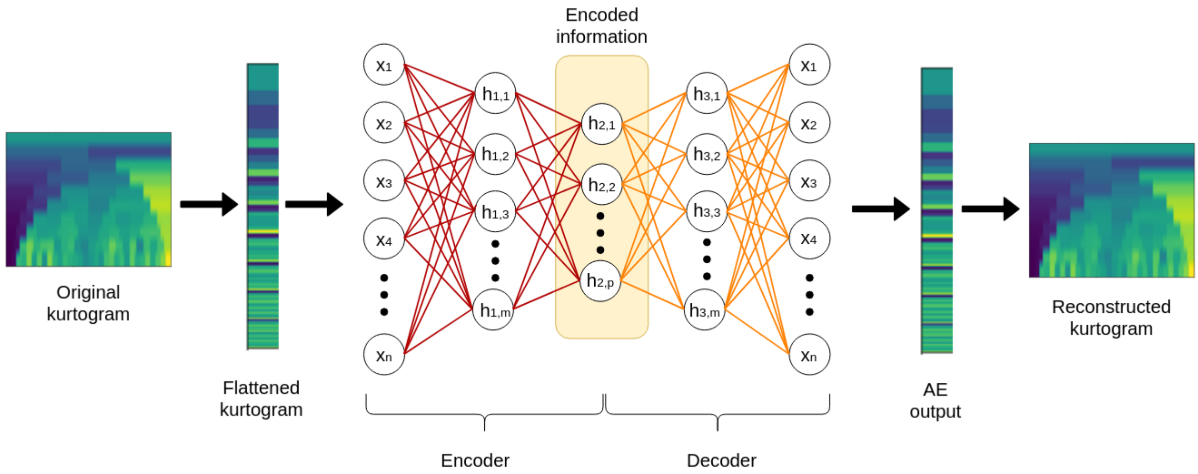

Figure 3. Autoencoder scheme applied to kurtogram analysis.

4. CASE STUDY

Data from the PRONOSTIA experimental test rig, publicly available by the FEMTO-ST Institute (Nectoux *et al.*, 2012), were used for the assessment of the proposed method. Figure 4 presents an overview of the PRONOSTIA test rig. This platform was created to consistently simulate the degradation process of ball bearings. For this purpose, it consists of a rotating part, which includes an asynchronous motor with a gearbox and its two shafts, a degradation generation part, which applies a radial force on the tested bearing, and a measuring part responsible for the acquisition of vibration and temperature signals.

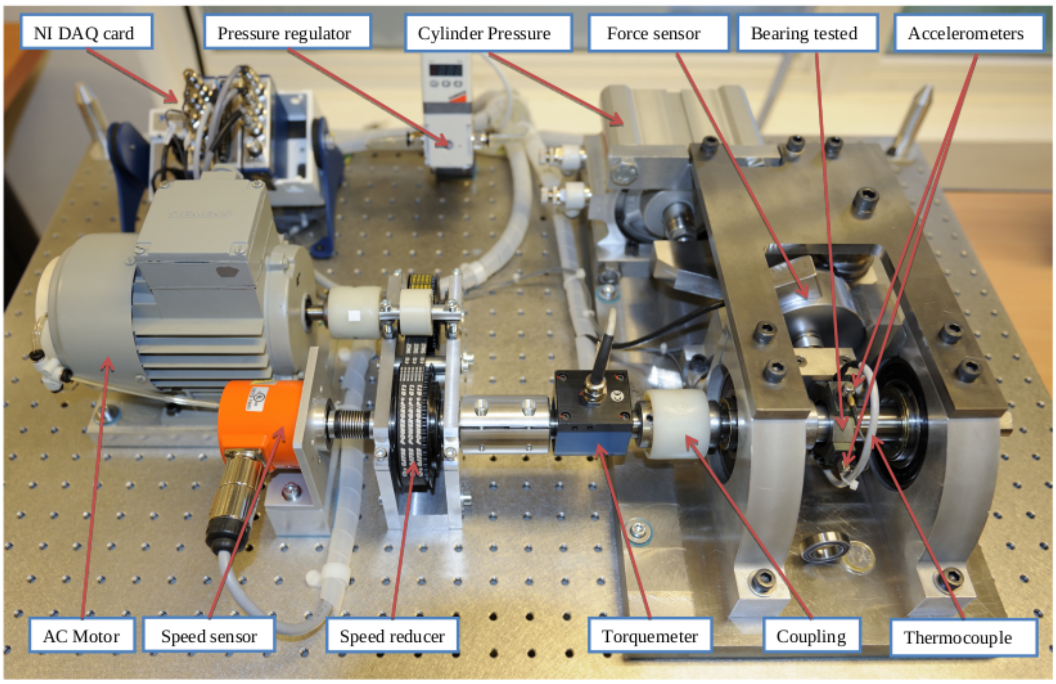

Figure 4. The PRONOSTIA experimental platform.

4.1 Data description

The dataset created with the PRONOSTIA platform is composed of run-to-failure experiments, in which the tests were interrupted when the vibration signals exceeded 20 g to prevent damage from spreading to the platform. Data were obtained with three different load and speed conditions. The experiments described in this work were based on data collected under operating conditions corresponding to the speed of 1800 rpm and load of 4000 N. Temperature and vibration signals (horizontal and vertical) were collected from eleven different bearings. The vibration signals were sampled at a frequency of 25.6 kHz, with each sample measured every 10 s with duration of 1/10 s, resulting in 2560 points sampled per measurement.

Four different datasets were used in this analysis. The first dataset was used as a training set, and the other three datasets were used as test sets to assess the proposed RUL estimation method. The trained set is here referred to as (A), and the test sets are identified as (B), (C) and (D). The network trained model, calculated offline, is adopted to process the kurtograms generated with the signals from the test datasets, simulating an online real time measurement for the sequence of analyzed signals.

4.2 Application of the method for the RUL estimation

A procedure was carried out for the selection of the best setting for the AE architecture presented in Figure 3 considering the selected operating condition. The evaluated hyperparameters were the number of neurons in the hidden layers and the activation function. Options between 200 and 600 neurons were compared for the 1st and 3rd hidden layers, and options between 100 and 300 neurons were compared for the central layer. The activation functions considered in this procedure were the rectified linear unit (relu) and the hyperbolic tangent (tanh), two of the most used in neural network design.

Both in training and test, all performed studies took into account only the vibration signals measured in the horizontal direction. Table 1 presents the results obtained from the training of the AE under multiple combinations of hyperparameters. Training loss refers to the neural network loss calculated with respect to the training data portion. Validation loss refers to the loss calculated with respect to the portion of validation data, which is used during the training phase to simulate unknown data, such as those that will be observed in the test phase.

Table 1. Training loss and validation loss of the AE for different numbers of neurons and activation functions, obtained from the training with the learning set.

#	Neurons	Activation	Training loss	Validation loss	Loss ratio
1	600, 300, 600	tanh	0.000440	0.000700	1.5909
2	500, 250, 500	tanh	0.000312	0.000375	1.2019
3	400, 200, 400	tanh	0.003089	0.003513	1.1373
4	300, 150, 300	tanh	0.000674	0.000832	1.2344
5	200, 100, 200	tanh	0.000576	0.000883	1.5330
6	600, 200, 600	tanh	0.000363	0.000624	1.7190
7	450, 150, 450	tanh	0.000628	0.001013	1.6131
8	300, 100, 300	tanh	0.001305	0.001760	1.3487
9	600, 150, 600	tanh	0.000690	0.001027	1.4884
10	500, 125, 500	tanh	0.000483	0.000706	1.4617
11	400, 100, 400	tanh	0.000501	0.000852	1.7006
12	600, 300, 600	relu	0.000544	0.001297	2.3842
13	500, 250, 500	relu	0.000676	0.001729	2.5577
14	400, 200, 400	relu	0.000813	0.001643	2.0209
15	300, 150, 300	relu	0.001068	0.002000	1.8727
16	200, 100, 200	relu	0.001830	0.002769	1.5131
17	600, 200, 600	relu	0.001026	0.002566	2.5010
18	450, 150, 450	relu	0.000956	0.002075	2.1705
19	300, 100, 300	relu	0.001583	0.002705	1.7088
20	600, 150, 600	relu	0.000872	0.002254	2.5849
21	500, 125, 500	relu	0.001219	0.003030	2.4856
22	400, 100, 400	relu	0.001817	0.003637	2.0017

The criterion adopted for selecting the best settings was based on the ratio between validation loss and training loss, since small training loss and high validation loss values indicate that the network has become too specialized for the signals used for the construction of the normal model, therefore more sensitive to variations in the test data pattern.

Based on this criterion, the setting #20 was selected for the model. According to this setting, the chosen activation is relu, the first and last hidden layers contain 600 neurons each, and the central layer contains 150 neurons, resulting in a compression ratio of 4. It is important to note that since only the encoder is used for estimating the RUL, only the first two hidden layers make up the prediction model. Thus, the second hidden layer, containing the smallest number of neurons, produces the output to be used for the RMSE calculation.

After selecting the hyperparameters, the model was tested with its respective learning set. This time, the complete dataset was used, including the first 25% samples used for training the network, so that the other 75% samples correspond to data unknown to the trained model. Figure 5 presents the result obtained from the analysis of the complete learning set. The raw vibration signal was compared to the respective resulting health indicator and its moving average. In general, it is possible to notice that health indicator is able to characterize trends and changes in the vibration behavior more clearly, allowing early detection of relevant degradation levels, including the one for which the threshold is set.

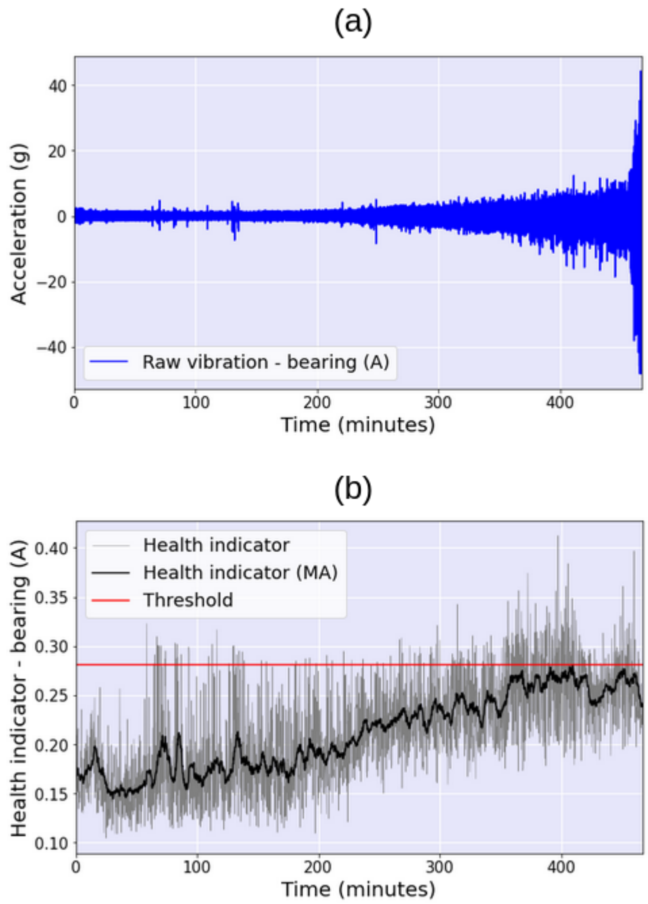

Figure 5. Analysis of the complete training set: (a) raw vibration signal and (b) obtained health indicator with its respective moving average.

The maximum value of indicator moving average series was adopted as a threshold, represented by the red line in Figure 5b, in order to be later applied for the evaluation of other datasets collected under similar operating conditions.

Using the trained model and its respective failure threshold, the next step consisted of estimating the RUL based on vibration signals collected in other bearings subjected to the same load and speed parameters. The estimation of the component RUL is done using two regression approaches. Long-term regression takes into account samples of the entire dataset, while short-term regression considers multiple segments that composes the dataset. Unlike long-term regression, where only one pair of coefficients is calculated, in short-term regression one pair of coefficients is calculated for each segment, and the last one is applied for the future projection.

Figures 6, 7 and 8 present the curves resulting from the analysis of the signals corresponding to the test sets, collected under the same operating conditions as the learning set. The RUL is here given by the time between the dashed gray vertical line, which denotes the limit between the past and the future projections, and the threshold line. Therefore, two RUL values are obtained in each analysis. The dashed orange line indicates the time at which the failure actually occurred. With this, it is possible to determine the time error between the prediction and the actual RUL.

In all cases, short-term regression was able to provide a more accurate RUL estimate. This aspect comes from the fact that the future projection of the short-term regression takes into account only the latest measurements, closer to the time of failure, and is, therefore, less influenced by older measurements, less related to the failure behavior.

Table 2 presents the comparisons between the expected RUL values and those obtained through the application of the proposed method for each operation condition.

In the analyzed bearings, the RUL estimation error remained below 30%, which is considered satisfactory given the level of difficulty of this type of prognostic task (Xinghui *et al.*, 2014; Sutrisno *et al.*, 2012). Furthermore, the more pessimistic estimated RUL with respect to the actual one is important in practical applications, due to safety aspects that

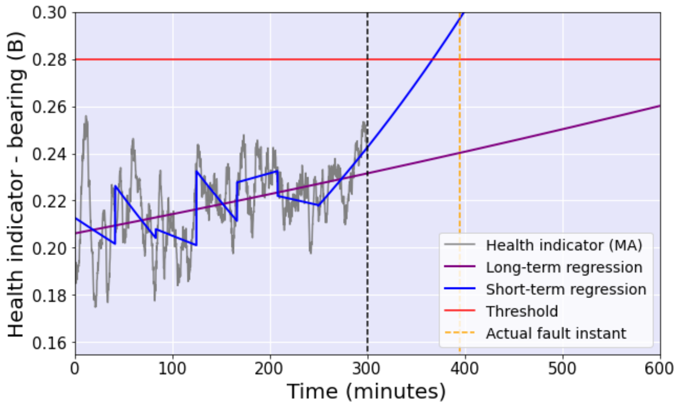

Figure 6. Health indicator and regressions obtained from the bearing test set (B).

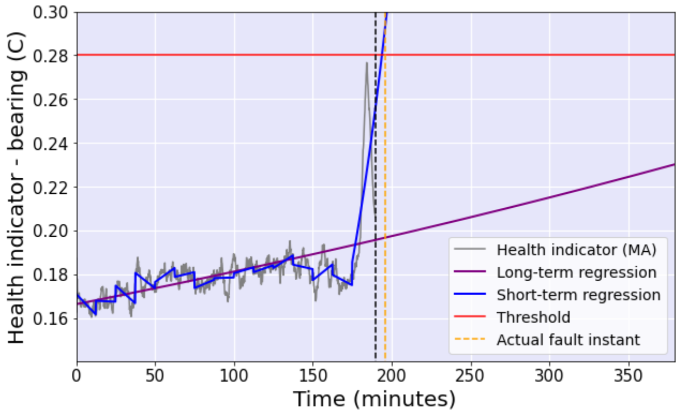

Figure 7. Health indicator and regressions obtained from the bearing test set (C).

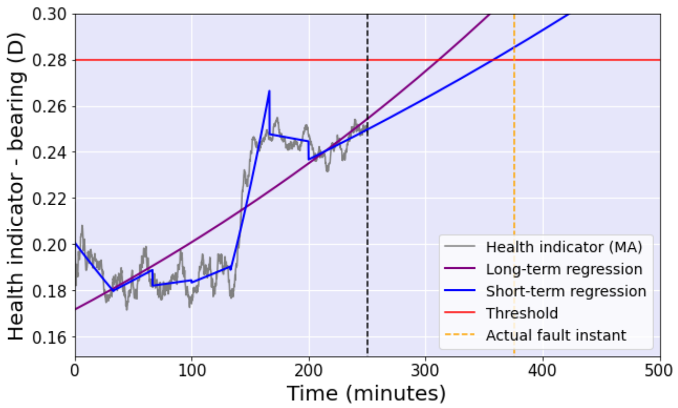

Figure 8. Health indicator and regressions obtained from the bearing test set (D).

Table 2. Comparison between actual and estimated RUL time for each test bearing.

Bearing	**Actual RUL (min)**	**Predicted RUL (min)**	**Error (min)**	**Relative error (%)**
B	95	68	27	28.42
C	6	5	1	16.67
D	126	109	17	13.49

make the early interruption of the system a crucial action to avoid reaching the failure threshold.

5. CONCLUSIONS

Throughout the previous sessions, a method based on the concept of spectral kurtosis and kurtograms for estimating the RUL of rotating machine components, more specifically rolling element bearings, was presented. The health indicator is generated from predictions made by the encoder of an AE neural network trained to reconstruct kurtograms of a signal corresponding to a normal vibration model. The number of neurons in the hidden layers and their respective activation function, which are important hyperparameters to guarantee a satisfactory performance, were selected from the test of a set of multiple combinations, using a score that measures the level of specialization of the network to the training data.

For validation purposes, the trained model was then applied to different bearings subjected to the same operating conditions. The resulting RUL values can be considered satisfactory if compared to other works carried out on the same datasets, and were close to the expected values, as it was made available by the datasets providers. Additionally, the most pessimistic estimations observed in all studied cases may guarantee greater security in practical decision making.

Many aspects are worth exploring in future works involving the techniques described here. The performance of different unsupervised neural network architectures can be verified, such as those composed of convolutional layers, suitable for image analysis, as well as the effects of other model hyperparameters, such as the optimizer, the number of epochs and the batch size. With regard to the spectral kurtosis technique, other estimators and available options for the production of kurtogram images are promising research approaches.

6. ACKNOWLEDGEMENTS

The authors thank the finantial support of São Paulo Research Foundation (FAPESP), grant #2020/10301-1, and the AS2M department of FEMTO-ST Institute for making the bearing datasets publicly available.

7. REFERENCES

Antoni, J., 2004. “The spectral kurtosis of nonstationary signals: Formalisation, some properties, and application”. In *12th European Signal Processing Conference*. pp. 1167–1170.

Antoni, J., 2006. “The spectral kurtosis: A useful tool for characterising non-stationary signals”. *Mechanical Systems and Signal Processing*, Vol. 20, No. 2, pp. 282–307.

Atamuradov, V., Medjaher, K., Camci, F., Zerhouni, N., Dersin, P. and Lamoureux, B., 2020. “Machine health indicator construction framework for failure diagnostics and prognostics”. *Journal of Signal Processing Systems*, Vol. 92, No. 6, pp. 591–609.

Dwyer, R.F., 1983. “Detection of non-gaussian signals by frequency domain kurtosis estimation”. *International Conference on Acoustic, Speech, and Signal Processing (ICASSP)*, pp. 607–610.

Huang, W., Khorasgani, H., Gupta, C., Farahat, A. and Zheng, S., 2018. “Remaining useful life estimation for systems with abrupt failures”. *Annual Conference of the PHM Society*, Vol. 10, No. 1.

Liu, G., Bao, H. and Han, B., 2018. “A stacked autoencoder-based deep neural network for achieving gearbox fault diagnosis”. *Mathematical Problems in Engineering*, Vol. 2018, pp. 1–10.

Liu, H., Li, L. and Ma, J., 2016. “Rolling bearing fault diagnosis based on STFT-deep learning and sound signals”. *Shock and Vibration*, Vol. 2016, pp. 1–12.

Nectoux, P., Gouriveau, R., Medjaher, K., Ramasso, E., Chebel-Morello, B., Zerhouni, N. and Varnier, C., 2012. “PRONOSTIA: an experimental platform for bearings accelerated degradation tests”. In *IEEE International Conference on Prognostics and Health Management, PHM’12*. IEEE Catalog Number : CPF12PHM-CDR, Denver, United States, pp. 1–8. URL `https://hal.archives-ouvertes.fr/hal-00719503`.

Okoh, C., Roy, R., Mehnen, J. and Redding, L., 2014. “Overview of remaining useful life prediction techniques in through-life engineering services”. *Procedia CIRP*, Vol. 16.

Qin, A., Zhang, Q., Hu, Q., Sun, G., He, J. and Lin, S., 2017. “Remaining useful life prediction for rotating machinery based on optimal degradation indicator”. *Shock and Vibration*, Vol. 2017, pp. 1–12.

Ren, L., Sun, Y., Cui, J. and Zhang, L., 2018. “Bearing remaining useful life prediction based on deep autoencoder and deep neural networks”. *Journal of Manufacturing Systems*, Vol. 48, pp. 71–77.

Shao, H., Jiang, H., Zhao, H. and Wang, F., 2017. “A novel deep autoencoder feature learning method for rotating machinery fault diagnosis”. *Mechanical Systems and Signal Processing*, Vol. 95, pp. 187–204.

Shrestha, A. and Mahmood, A., 2019. “Review of deep learning algorithms and architectures”. *IEEE Access*, Vol. 7, pp. 53040–53065.

Soualhi, A., Medjaher, K. and Zerhouni, N., 2014. “Bearing health monitoring based on hilbert-huang transform, support vector machines and regression”. *IEEE Transactions on Instrumentation and Measurement*, pp. 1–11.

Sutawanir, 2017. “Estimation of spectral kurtosis”. In *AIP Conference Proceedings*.

Sutrisno, E., Oh, H., Vasan, A.S.S. and Pecht, M., 2012. “Estimation of remaining useful life of ball bearings using data driven methodologies”. In *2012 IEEE Conference on Prognostics and Health Management*. IEEE. doi:

10.1109/icphm.2012.6299548. URL https://doi.org/10.1109/icphm.2012.6299548.

Tafinine, F. and Mokrani, K., 2012. "Real time automatic detection of bearing fault in induction machine using kurtogram analysis". *The Journal of the Acoustical Society of America*, Vol. 132, No. 5, pp. EL405–EL410.

Udmale, S.S., Patil, S.S., Phalle, V.M. and Singh, S.K., 2018. "A bearing vibration data analysis based on spectral kurtosis and ConvNet". *Soft Computing*, Vol. 23, No. 19, pp. 9341–9359.

Vicuña, C.M. and Acuña, D.Q., 2014. "Cyclostationary processing of vibration and acoustic emissions for machine failure diagnosis". In *Lecture Notes in Mechanical Engineering*, pp. 141–156.

Vrabie, V., Granjon, P. and Seviere, C., 2003. "Spectral kurtosis: From definition to application". In *6th IEEE International Workshop on Nonlinear Signal and Image Processing (NSIP 2003)*.

Xinghui, Z., Jianmin, Z., Jianshe, K., Haiping, L. and Hongzhi, T., 2014. "Bearing prognostics with non-trendable behavior based on shock pulse method and frequency analysis". *Journal of Vibroengineering*, Vol. 16, No. 8, pp. 3963–3976. URL https://www.jvejournals.com/article/15488.

Yoo, Y. and Baek, J.G., 2018. "A novel image feature for the remaining useful lifetime prediction of bearings based on continuous wavelet transform and convolutional neural network". *Applied Sciences*, Vol. 8, No. 7, p. 1102.

8. RESPONSIBILITY NOTICE

The author(s) is (are) solely responsible for the printed material included in this paper.