BEARINGS PROGNOSTIC USING MIXTURE OF GAUSSIANS HIDDEN MARKOV MODEL AND SUPPORT VECTOR MACHINE

Fatima Ezzahra Sloukia[1], Rajae Bouarfa, Hicham Medromi[2] and Mohammed Wahbi[1]

[1]Electrical Engineering Department SIR2C2S/LASI-EHTP, Casablanca, Morocco

sloukia.fatimaezzahra@gmail.com
bouarfe.rajae@gmail.com
wahbi_mho@yahoo.fr

[2]Laboratoire d'Informatique, Systèmes et Énergies Renouvelables Team Architecture of Systems, ENSEM, Hassan II University, Casablanca, Morocco

hmedromi@yahoo.fr

ABSTRACT

Prognostic of future health state relies on the estimation of the Remaining Useful Life (RUL) of physical systems or components based on their current health state. RUL can be estimated by using three main approaches: model-based, experience-based and data-driven approaches. This paper deals with a data-driven prognostics method which is based on the transformation of the data provided by the sensors into models that are able to characterize the behavior of the degradation of bearings.

For this purpose, we used Support Vector Machine (SVM) as modeling tool. The experiments on the recently published data base taken from the platform PRONOSTIA clearly show the superiority of the proposed approach compared to well established method in literature like Mixture of Gaussian Hidden Markov Models (MoG-HMMs).

KEYWORDS

Prognostic, RUL, MoG-HMM, SVM.

1. INTRODUCTION

Manufacturers of modern complex industrial systems are facing a number of challenges including increased design complexity, high security and safety requirements, and the need to reduce life cycle costs. The emerging field of diagnostic, prognostic, and health management tackles these issues through the development of algorithms that use various sources of information to detect anomalies, diagnose faults, and predict the remaining useful life of maintainable components.

The main purpose of the prognostic is estimating the RUL of a failing component or subsystem so that maintenance can be executed to avoid catastrophic failures. ISO 13381-1: 2004 (ISO 13381-1, 2004) defines prognostic as the estimated time to failure and risk for one or more existing and future failure modes [1].
In general, prognostics can be divided into three main approaches: model-based, experience–based and data-driven prognostics [3-5]. The choice of a prognostic approach is mainly based on the type of knowledge available on the system.

The model based prognostic related to the use of analytical models, it provide more precise results however real systems are non-linear and degradation mechanisms are stochastic and difficult to obtain in the form of analytical models [6]. A method for health condition prediction of propulsion system bearings was developed by Marble and Morton [7] based on bearing spall propagation physical.

The experience-based methods use the feedback data assembled during a significant period of time, it's easier to implement when there is a significant among of data also it does not use on-line sensors data neither complex mathematical models, however precision decrease when there is a lack of experience data.

The data-driven methods attempt to transform the raw monitoring data into relevant information and behavior models of the system, it's easy to get reliable data than constructing physical or analytical behavior and the generated behavioral models from real monitoring data lead to more precise prognostic, however, these approaches are highly-dependent on the quantity and quality of monitoring.

In this paper, data driven approach is addressed in order to utilize the availability of condition monitoring data [25]. They can be divided into statistical methods (multivariate statistical methods, linear and quadratic regression …) and Artificial Intelligent (AI) methods which have been largely applied to machinery remaining life prediction [8]. The most used models for prognostics are artificial neural networks (ANNs) [9], support vector machine (SVM) [10], Fuzzy theory [11].

The proposed methods rely on two main phases: a learning phase and an exploitation phase (Figure 1). During the first phase, the raw data are used to extract reliable features, which are then used to learn behavioral models representing the dynamic of the degradation in the bearing. The modeling of the degradation is done by using Mixture of Gaussians Hidden Markov Model (MoG-HMM) and SVM. In the second phase, the learned models are exploited on line to assess the current health state of the bearing and estimate the RUL.
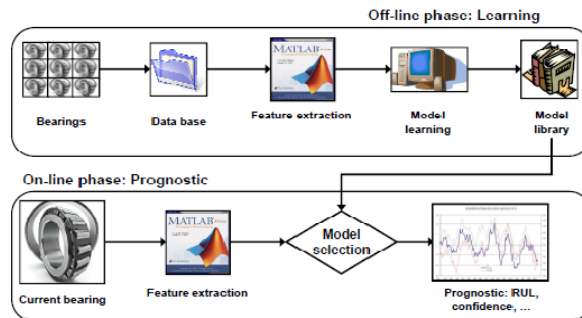

Figure 1. The two phases of the proposed method [12].

The choice of bearings can be explained by the fact that these components are considered as the most common mechanical elements in industry and are present in almost all industrial processes, especially in those using rotating elements and machines [13]. Moreover bearing failure is one of the foremost causes of breakdowns in rotating machinery and such failure can be catastrophic [14], resulting in costly downtime. Many previous studies [7,17,18,20] have developed theoretical foundation and tools to describe bearing failure modes.

Large variety of features can be extracted to describe the characteristics of vibration signal, which can be divided into time domain, frequency domain and time-frequency domain techniques [21]. The time domain analysis can be done either by visually examining portions of the time domain waveform (signal morphology) or by examining some statistical parameters related to the time domain vibration signal. Various time domain statistical parameters have been used in [22] as trend parameters to detect the presence of incipient bearing damage. The most commonly used ones are peak, root mean square (detect abnormally high energy dissipation [23]), crest factor, kurtosis, and skewness. In the frequency domain approach the dominant frequency components of the vibration signals and their amplitudes are used for trending purposes. As faults occur in one of the rolling element bearing components (inner race, outer race, or rolling elements) the vibration spectrum peaks located at the corresponding bearing defect frequency increase. The most used is the Fast Fourier Transform (FFT).

Time-frequency domain techniques use both time and frequency domain information allowing for the investigation of transient features, such as impacts. A number of time-frequency domain techniques have been proposed including the short-time Fourier transform (STFT), the Wigner-Ville distribution (WVD), and the wavelet transform (WT) [24].

In this work, we use the wavelet packet decomposition (WPD), a time–frequency domain technique and the root mean square (RMS), a time domain technique, in connection with the MoG-HMM or SVM modeling techniques for developing a method of bearing health and prognostics and elaborate a comparison of the performance of SVM and MoG-HMM. The proposed failure prognostic methods are tested on a condition monitoring data base [25] taken from the platform PRONOSTIA [26], related to bearings degradation tests, enabling the verification of condition monitoring, fault detection, fault diagnostic and prognostic approaches.
The remainder of this paper is organized as follows. Some of related works to SVM and Hidden Markov Model (HMM) are cited in section 2. A technical background is given in Section 3 in which the MoG-HMM and SVM models and the WPD are discussed. Section 4 presents experiments and results in which the proposed MoG-HMM and SVM methods are validated using experimental vibration monitoring data collected from bearings. Conclusions are given in Section 5.

2. RELATED WORK

SVM shows outstanding performance in the classification process compared with the other classifiers [15,16]. For accurate assessment of the residual life of bearings, Kim et al. [17] proposed a machine prognostics model based on health state estimation using SVM. Kankar et al. [18] have also shown the effectiveness of SVM for bearing faults classification.

Markov chain model based on the transition probability matrix is appropriate to the analysis of a random dynamic system [19]. HHM is a state of the art technique for time series modeling and classification and has also been successfully applied to many other fields such as tool wear condition monitoring [36] and bearing diagnosis [37]. Li et al. Chinam and Baruah [20] have HMM to assess the degradations of bearings and to estimate the RUL. In their method the authors considered the degradation as a stochastic process with several states representing different health states of the physical component.

3. TECHNICAL BACKGROUND

3.1. Wavelet Packet Decomposition (WPD)

In the *WPD* analysis of a signal, this latter is filtered with both low-pass (LP) and high-pass (HP) filters. The cut-off frequency of these filters is one fourth of the sampling frequency of the signal. The LP and the HP filtered signals are referred to the approximation (A) and the detail (D), respectively. Both the approximation and the detail are half the size of the original signal and represent the low frequency and the high-frequency content of the signal.

If the same decomposition procedure is applied to the first layer A and D signals, a new layer is obtained that consists of four signals each of which is one fourth of the size of the original signal. These new signals are named as AA (approximation of the approximation), DA (detail of the approximation), AD (approximation of the detail) and DD (detail of the detail). The energy of the signal in any layer is referred to as the node energy that we will later use as feature of the vibration signal.

A third level WPD of a signal is illustrated in Figure 3. In this representation, the third level signals AAA, DAA, ADA, DDA, AAD, DAD, ADD and DDD represent the frequency content of the original signal within the bands: $[\frac{kfs}{16}, \frac{(k+1)fs}{16}]$, k=0,1,...7, where fs is the sampling rate of the signal. A detailed review of WPD can be found in [27].
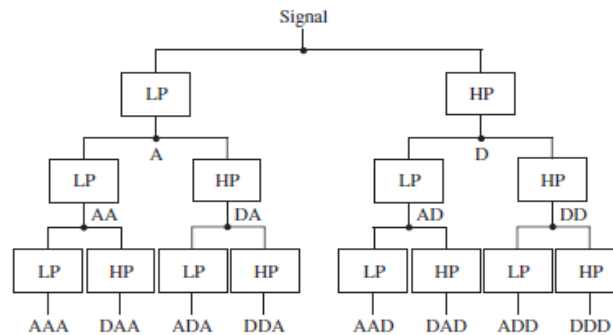

Figure 3. Third level wavelet packet decomposition of a signal.

3.2. Mixture of Gaussians Hidden Markov Model

The MoG-HMMs have proved to be a suitable tool as they model the physical component's degradation by using continuous observations provided by the monitoring sensors. They also permit the estimation of the stay durations in each health state leading to the prediction of the RUL value [28].

The HMM is defined by the parameters cited in [29]:

- The number of states, N. The state at time *t* is defined as s_t.
- The number of distinct observations for each state, M The observation symbols correspond to the physical output of the modeled system. The individual observation symbols are denoted as:

$$V= v_1,v_2,v_3,\ldots,v_M.$$

- The transition probability distribution, $A=\{a_{ij}\}$ where:
-

$$a_{ij}=P[s_{t+1}=j|s_t=i],1\leq i,j\leq N. \tag{1}$$

- Observation probability distribution of each state, $B=\{b_i(k)\}$, where:

$$b_i(k)=P[v_k|s_t=i],1\leq i\leq N,1\leq k\leq M. \tag{2}$$

- The initial state distribution, $\pi=\{\pi_i\}$, where:

$$\pi_i=P[s_t=i],1\leq i\leq N. \tag{3}$$

A compact notation $H=(A, B, \pi)$ is used for an HMM model.

In practice, HMMs are used to solve typical problems, which are explained in [29]:

- The evaluation problem: Given a sequence of observations and a HMM H, how to effectively evaluate P (V | H), it's solved by using the Forward-Backward algorithm.
- The decoding problem: Given an observation sequence and an HMM H, how to choose an optimal sequence of states, it's solved by using the Viterbi algorithm.
- The learning problem: find the parameters of a HMM H that maximize P (V | H), it's solved by using the Baum-Welch algorithm.

The problem with HMM approach is that it uses discrete observations, whereas in our case we use continuous features, so changes were made to the matrix B [29]:

$$b_i(O)=\sum_{m=1}^{M} C_{jm}\xi\left(O,\mu_{jm},U_{jm}\right), 1\ll j\ll N \tag{4}$$

With the observation vector O to model, C_{jm} is the coefficient of the mixture matrix M in state i, ξ is the Gaussian density, μ_{jm} is the mean vector and U_{jm} is the covariance matrix. During the learning phase we used an HMM 3 states left to right for each model MoG-HMM (Figure 4).
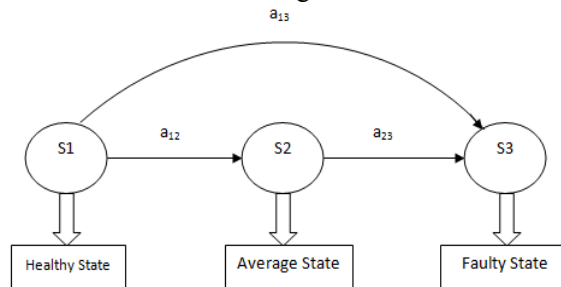

Figure 4. HMM left to right with three states

3.3. Machine Vector Support

SVM is based on the statistical learning theory introduced by Vapnik and his co-workers [30]. The basic idea is to find a hyperplane which separates the N dimensional data perfectly into its two classes. However, since example data is often not linearly separable, SVM's introduce the notion of a "kernel induced feature space" which casts the data into a higher dimensional space where the data is separable.

In a problem linearly separable we suppose that there is a given training sample set $G=\{(x_i,y_i)$, $i=1...n\}$, each sample $x_i \in R^d$ belongs to a class $y_i \in \{+1,-1\}$. The boundary can be expressed as follows:

$$w.x+b=0 \qquad\qquad (5)$$

where w is a weight vector and b is a bias. So the following decision function can be used to classify any data point in either class A or B:

$$f(x)=sign(w.x+b) \qquad\qquad (6)$$

The optimal hyperplane separating the data can be obtained as a solution to the following constrained optimization problem:

$$\text{Minimise} \qquad \frac{1}{2}||w||^2 \qquad\qquad (7)$$

$$\text{Subject to} \qquad y_i[(w.x+b)]-1\geq0, \; i=1,\ldots,n \qquad\qquad (8)$$

Introducing the Lagrange multipliers $\alpha_i\geq0$, the optimization problem can be rewritten as:

$$\text{Maximise} \quad L(w,b,\alpha)=\sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j (x_i.x_j) \qquad\qquad (9)$$

$$\text{Subject to} \quad \sum_{i=1}^{n} \alpha_i y_i =0 \qquad\qquad (10)$$

The decision function can be obtained as follows:

$$f(x)=sign \left(\sum_{i,j=1}^{n} \alpha_i y_i (x_i.x_j)+b \right) \qquad\qquad (11)$$

If the linear boundary in the input space is not enough to separate into two classes properly, it is possible to create a hyperplane that allows linear separation in the higher dimension. In SVM, it is achieved by using a transformation $\phi(x)$ that maps the data from input space to feature space. If a kernel function:

$$K(x,y)= \phi(x).\phi(y) \qquad\qquad (12)$$

is introduced to perform the transformation, the basic form of SVM can be obtained:

$$f(x)=sign\left(\sum_{i,j=1}^{n} \alpha_i y_i K(x_i,x_j)+b\right) \qquad\qquad (13)$$

Among the kernel functions in common use are linear functions, polynomials functions, gaussian basis functions and sigmoid functions. In this research, the polynomial and GBF functions were employed for classification of health states. SVMs were originally designed for binary classification and there are methods which are applicable to multiclass classification, such as "one-against-one", "one-against-all". The most suitable method is chosen according to the application constraints, the number of classes and the number of training samples [31].

4. EXPERIMENTS AND RESULTS USING BEARING FAILURE DATA

4.1. The (WPD or RMS) and MoG-HMM prognostic method

For both learning and exploitation phases, we extracted features (RMS, Crest Factor, Kurtosis, FFT, Cepstrum, WPD, ...etc) from raw data (vibration signal (Figure 5)).
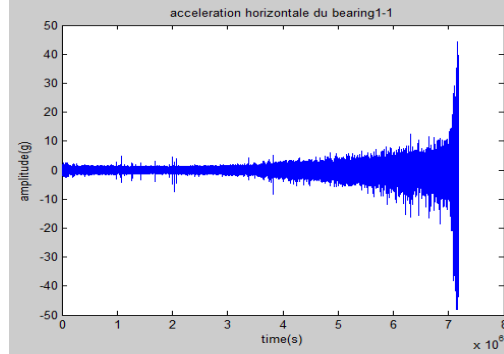

Figure 5. Example of vibration signal

After that we chosed two features: the RMS which is a temporal feature that detect abnormally high energy dissipation in the vibration signal, an example of RMS feature is shown in Figure 6. And the nodal energy of the WPD which has shown great results in many previous studies [28,32] which offers a richer range of possibilities for signal analysis and allows the best matched analysis to a signal. It provides level by level transformation of a signal from the time domain into the frequency domain, an example of the extracted nodal energy from the 3rd level of WPD is given in Figure 7, The choice of the 3rd level is motivated by Equation 14, where L is the decomposition level, F_s is the sampling frequency, and F_d is the defect frequency, which gives a frequency band of 3.2 KHz in each band.

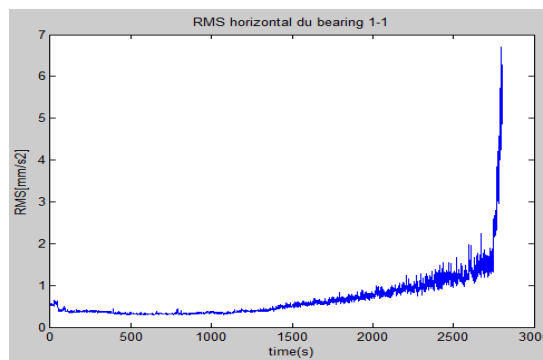$$L \leq \log_2 \frac{F_S}{3F_d} - 1 \tag{14}$$

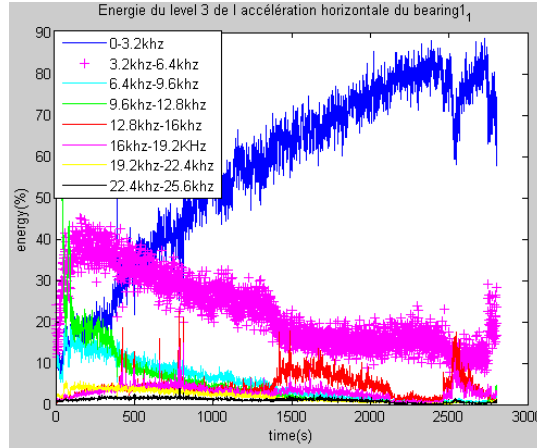Figure 6. Example of RMS extraction of a vibration signal.

Figure 7. Example of energy nodal of WPD extraction of a vibration signal.

In *the learning phase* the features extracted are used to estimate the (A,B,π) parameters using the Baum-Welch algorithm and temporal parameters (the stay duration in each state) using the Viterbi algorithm of the MoG-HMM model.

Three states are defined for each MoG-HMM. The number of mixtures is fixed at two. The parameters (A,B,π) are firstly arbitrary initialised then the feature is used to re-estimate the parameters which are adjusted by maximizing the probability P(V|H) of the model over the observations.

Six MoG-HMMs are obtained using Baum-Welch algorithm. The RMS feature has shown better results compared to WPD because WPD need a lot of models in the same condition to give precise results. An example of estimation of the parameters of bearing 1 condition 1 is illustrated hereafter:

$$\pi = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad A = \begin{bmatrix} 0.9985 & 0.0015 & 0 \\ 0.0015 & 0.9956 & 0.0029 \\ 0 & 0.0014 & 0.9986 \end{bmatrix} \quad M = \begin{bmatrix} 0.4176 & 0.5824 \\ 1 & 0 \\ 0.1271 & 0.8729 \end{bmatrix}$$

The Viterbi algorithm allows us to estimate the stay duration in each state and to identify the final state which corresponds in our case to the degradation state. An example of estimation of the decoded state sequence is illustrated in Figure 8.
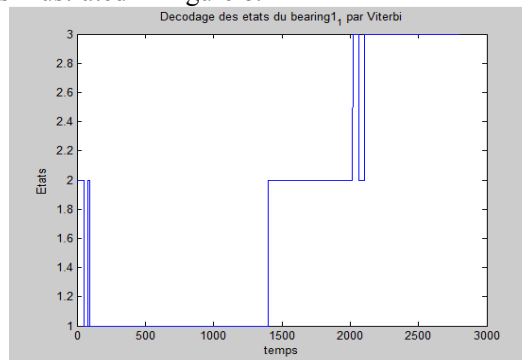

Figure 8. The decoded state sequence using Viterbi algorithm

The x-axis represents time and y-axis represents the state of the learning bearing. The duration of healthy state of this bearing is 1340s, the average state is 719s and the faulty state is 744s. Note that the time between two measurements is 10s.

The exploitation phase allows to characterize the health state of test data through the selection of the model that maximizes the probability P (V | H), the model must be able to estimate the remaining life of the bearing on-line.

To estimate the remaining useful life of the test bearing, we followed these steps:

- Define the failed state that corresponds to the state S3 (Faulty State),
- Estimate the stay duration in each state from the path estimated by Viterbi,
- Sum the lengths of stay durations,
- Estimate the current state, at each instant, of the test bearing using the Viterbi algorithm on the test data,
- Estimate the remaining time before reaching the final state based on stay durations previously estimated.

An example of calculation of the RUL and the corresponding error is shown in Figure 9. The x-axis represent current time of the test bearing and y-axis represent the failure time. The red line represents real RUL and the blue one represents the estimated RUL of the test bearing. We obtained: an estimated RUL of 2h 26min 10s and the relative error between the two RULs is 38.52%, with a real RUL of 3h 58min.
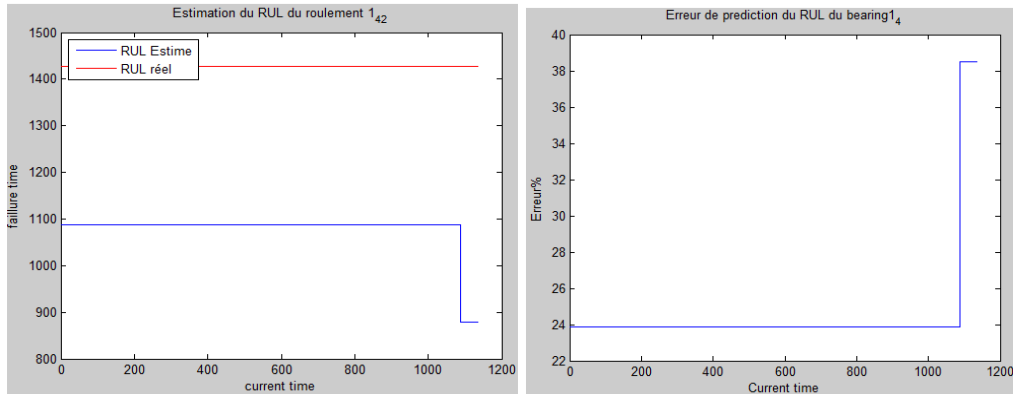

Figure 9. The estimated RUL and the corresponding relative error of the bearing 1_4 using MoG-HMM.

4.2. The (WPD or RMS) and SVM prognostic method

In *the learning phase* we formed learning models based on features (RMS and Nodal Energy of WPD) and using the library LibSVM [33] which contains many classification modules and supports multiclass classification and cross-validation (developed later), it also allows the use of different kernels: linear, polynomial ..., LibSVM uses an optimization called "Sequential Minimal Optimization" or SMO. More details of the algorithm are given in [34].

The resolution and performance of the SVM method involves the selection of several parameters: the type of kernel, kernel parameters (γ, ...) and the soft-margin constant, C, which controls the penalty associated to the examples and set the relative importance of maximizing the margin and minimizing the amount of slack [35]. Several methods exist to select these parameters. A classic

technique is to choose a grid-search for C and γ then applying a selection method (i.e. Cross validation) which optimizes the quality criterion associated to (C, γ) and can accurately predict unknown data.

Cross-validation is an approach that test the effectiveness of the learned model applied to training data on unseen data. Cross-validation is divided into several sub-methods. The most common method is the "k-Fold" which divides all the samples in K groups called folds (if K = n, this is equivalent to the *Leave One Out* strategy), of equal sizes (if possible). The prediction function is learned using (K − 1) folds, and the fold left out is used for test.

Cross-validation is simple to implement, it provides an estimation of the generalization error and helps to avoid the over-fitting [35].

In our case we use the one against all multiclass SVM classification (3 classes: healthy, average and faulty states). Because classifiers are sensitive to the way features are scaled, features must be normalized before the learning phase. The kernels chosen are polynomial and RBF kernel because they gave the most acceptable results.

After constructing learning models, we pass to *the exploitation phase* in which we classify test data based on the learned models, for that we used svmpredict() function in the LIBSVM library, this function gives the percentage of classification of test data regarding the learning model (Accuracy percentage), and then we made the choice of the most suitable model.

The obtained results for the bearings 1_3 to 1_7 condition 1 with the model based on bearing 1_2 applied on RMS and nodal energy of WPD features are given in Table 1. According to the results we find that SVM gives more accurate results using RMS than the nodal energy of WPD this is due to the fact that data provided from the WPD are correlated which degrade the performance of the SVM.

Table 1: Classification results of test data (RMS, WPD) towards the learning model based on bearing 1_2 condition 1.

		Accuracy in (%)				
	Bearing Feature + Parameters	1_3	1_4	1_5	1_6	1_7
Model based on Bearing 1_2	Horizontal RMS Best C=8, Best gamma=2 Kernel: polynomial	99.17%	99.30%	100%	99.91%	100%
	Horizontal Nodal energy of WPD Best C=2, Best gamma=0.0625 Kernel: polynomial	79.58%	95%	99.09%	97.48%	99.87%

After selecting the models we used them to determine the remaining useful life of each test bearing using the state sequence of test data and by following the steps listed in MoG-HMM.
An example of calculation of the RUL and the corresponding error is shown in Figure 10. For the same test bearing 1_4 (real RUL: 3h 58min) we find an estimated RUL of 3h 06 min and a relative error between the two RULs of 21.85%.
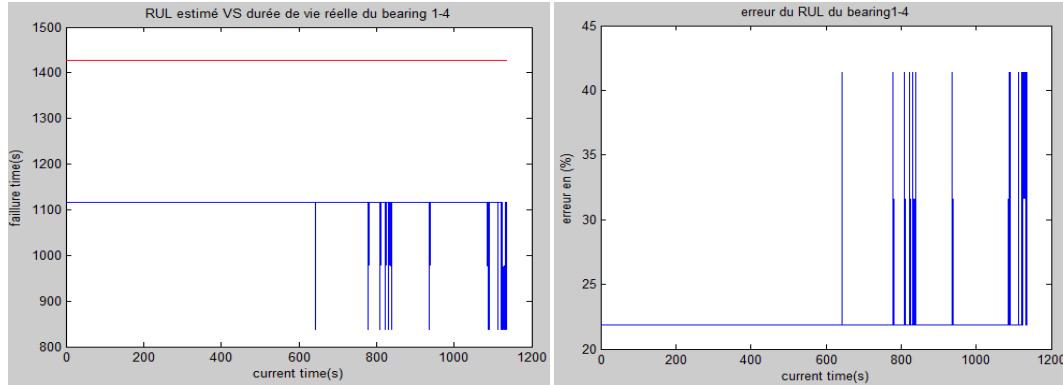
Fig.10: The estimated RUL and the corresponding relative error of the bearing1_4 using SVM method.

We note from the obtained results by using both methods (SVM and MoG-HMM) (Figures 9 and 10) that the estimated RUL using SVM is more closer to the real RUL than the one estimated using MoG-HMM and the relative error of this latter is higher than the one obtained using SVM. We conclude from the obtained results that the SVM presents better results compared to MoG-HMM because this latter requires more learning models to be able to give precise results and SVM is more easier to use.

5. CONCLUSIONS

This work was dedicated to the characterization of the health state of bearings degraded in an accelerated way by the platform PRONOSTIA, it presents a comparison of the performance of two well known methods: SVM and MoG-HMM, the choice of the MoG-HMM is motivated by its good capacity of modeling the temporal dependences which exist between the observations. It needs to make a hypothesis about the probability distribution of observations in the model's states: we presuppose a particular form of data distribution. The SVM is a good classifier with a good power of generalization, but the SVM does not allow representing the temporal evolution of the observations, while this temporal evolution is essential to discriminate the vibratory signals. Our perspective thus is to exploit the advantages of SVM and HMM to construct a single hybrid method: **SVM-HMM.**

REFERENCES

[1]   ISO 13381-1, Condition Monitoring and Diagnostics of Machines – Prognostics – Part 1: General Guidelines: International Standards Organization, 2004.

[3]   A. K. Jardine, D. Lin, and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance,"Mechanical Systems and Signal Processing, vol. 20, no. 7, pp. 1483– 1510, 2006.

[4]   A. Heng, S. Zhang, A. C. Tan, and J. Mathew, "Rotating machinery prognostics: State of the art, challenges and opportunities," Mechanical Systems and Signal Processing, vol. 23, no. 3, pp. 724 – 739, 2009.

[5]   Byington, C., Roemer, M., Kacprzynski, G. Galie, T. (2002). Prognostic enhancements to diagnostic systems for improved condition-based maintenance, In: Proc. IEEE Int. Conf. on Aerospace, vol. 6, pp. 2815-2824.

[6]   D.J. Inman, Charles R. Farrar, Vicente Lopes Jr., in: Damage Prognosis: For Aerospace, Civil and Mechanical Systems, John Wiley and Sons, 2005.

[7]   S. Marble, B.P. Morton, Predicting the remaining life of propulsion system bearings, in: Proceedings of the 2006 IEEE Aerospace Conference, Big Sky, MT, USA, 2006.

[8] A. K. Jardine, D. Lin, and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance,"Mechanical Systems and Signal Processing, vol. 20, no. 7, pp. 1483– 1510, 2006.

[9] N.Z. Gebraeel, M.A. Lawley, A neural network degradation model for computing and updating residual life distributions, IEEE Transactions on Automation Science and Engineering 5 (2008) 154– 163.

[10] C. Li, T. Li, Y.S. Bai, Condition residual life evaluation by support vector machine, in: Proceedings of the Eighth International Conference on Electronic Measurement and Instruments, Xi' an, China. (2007) 441–445.

[11] J. Fanegan, A. Ralescu, F. Ueda, A. Fuzzy, Model for estimating the remaining lifetime a diesel engine, in: Annual Meeting of the North American Fuzzy Information Processing Society, San Diego, California. (2007) 473–476.

[12] Tobon-Mejia D., Medjaher K., Zerhouni N., Tripot G. Estimation of the remaining useful life by using Wavelet Packet Decomposition and HMMs. In IEEE Aerospace Conference AIAA - 2011, Montana, USA. Pages 1 - 10, ISBN: 978-1-4244-7350-2, DOI: 10.1109/AERO.2011.5747561, 2011.

[13] P. O'Donnell, "Report of large motor reliability survey of industrial and commercial installations, part I, II & III," IEEE Transactions on Industry Applications, vol. 21, pp. 853–872, 1985.

[14] U. Pulkkinen, S. Uryas'ev, Optimal operational strategies for an inspected component. In: Petersen KE, Rasmussen B, Eds. Safety and Reliability '92, Proceedings of the European Safety and Reliability Conference '92. London: Elsevier, pp. 896–907, 1992.

[15] Niu, G., Han, T., Yang, B. S., & Tan, A. C. C. (2007a). Multi-agent decision fusion for motor fault diagnosis. Mechanical Systems and Signal Processing, 21, 1285–1299.

[16] Pal, M., & Mather, P. M. (2004). Assessment of the effectiveness of support vector machines for hyperspectral data. Future Generation Computer Systems, 20, 1215–1225.

[17] H.E. Kim, Andy C.C. Tan, J. Mathew, Eric Y. H. Kim and B.K. Choi, Machine prognostics based on health state estimation using SVM, Proceedings Third World Congress on Engineering Asset Management and Intelligent Maintenance Systems, Conference, vol. 199, pp. 834-845, 2008.

[18] P.K. Kankar, Satish C. Sharma and S.P. Harsha, Fault Diagnosis of Ball Bearings Using Continuous Wavelet Transform, Applied Soft Computing, vol. 11, pp. 2300-2312, 2011.

[19] T. Tien, A research on the prediction of machining accuracy by the deterministic grey dynamic model DGDM(1, 1, 1), Applied Mathematics and Computation 161 (2005) 923–945.

[20] R.B. Chinnam, P. Baruah, Autonomous diagnostics and prognostics through competitive learning driven HMM-based clustering, Proceedings of the International Joint Conference on Neural Networks, vol. 4, 2003, pp. 2466–2471.

[21] Feature extraction for bearing prognostics and health management (phm)–a survey (preprint); Weizhong Yan, Hai Qiu, and Naresh Iyer; Metals Branch; Metals, Ceramics, and NDE Division. May 2008.

[22] R.B.W. Heng, M.J.M. Nor, Statistical analysis of sound and vibration signals for monitoring rolling element bearing condition, Applied Acoustics 53 (1) (1998) 211-226.

[23] N.G. Gebraeel, A. Elwany, J. Pan, Residual life predictions in the absence of prior degradation knowledge, IEEE Trans. Reliab. 58 (1) (2009).

[24] G.Y. Yen, K.C. Lin, Wavelet packet feature extraction for vibration monitoring, Proceedings of IEEE International Conference on Control Applications, IEEE, New York, 1999, pp. 1573–1578.

[25] IEEE PHM 2012 Prognostic Challenge, http://www.femto-st.fr/ieee-PHM2012-data-challenge.

[26] P. Nectoux, R. Gouriveau, K. Medjaher, E. Ramasso, B. Morello, N. Zerhouni, C. Varnier, "PRONOSTIA: An Experimental Platform for Bearings Accelerated Degradation Tests" in IEEE International Conference on Prognostics and Health Management, PHM'12., Denver, Colorado : United States (2012)".

[27] C. Rowden, S. Hall, Speech Processing, McGraw-Hill, New York, 1992 (Chapter 5).

[28] Diego A. Tobon-Mejia, Kamal Medjaher, Noureddine Zerhouni and Gerard Tripot , A Data-Driven Failure Prognostics Method based on Mixture of Gaussians Hidden Markov Models.

[29] L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," in Proceedings of the IEEE, vol.77 (2), 1989, pp. 257–286.

[30] V.N. Vapnik, The Nature of Statistical Learning Theory, Springer, New York, 1995.

[31] Jonathan Milgram, Mohamed Cheriet, Robert Sabourin, "One Against One" or "One Against All": Which One is Better for Handwriting Recognition with SVMs?

[32] H. Ocak, K. A. Loparo, and F. M. Discenzo, "Online tracking of bearing wear using wavelet packet decomposition and probabilistic modeling: A method for bearing prognostics," Journal of sound and vibration, vol. 302, pp. 951–961, 2007.

[33] Chang, C-C. and Lin, C-J. (2001) LIBSVM: a library for support vector machines. Software available at http://www.csie.ntu.edu.tw/~cjlin/papers/libsvm.pdf

[34] John C. Platt, Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines.

[35] Asa Ben-Hur and Jason Weston, "A User's Guide to Support Vector Machines".

[36] H.M. Ertunc, K.A. Loparo, H. Ocak, Tool wear condition monitoring in drilling operations using hidden Markov models (HMMs), International Journal of Machine Tools and Manufacture 41 (9) (2001) 1363–1384.

[37] H. Ocak, K.A. Loparo, An HMM based fault detection and diagnosis scheme for rolling element bearings, Journal of Vibration and Acoustics, Transactions of the ASME 127 (4) (2005) 299–306