Unsupervised Trend Extraction for Prognostics and Condition Assessment

Ahmed Mosallam [**] **Kamal Medjaher** [***]
Nourredine Zerhouni [***]

FEMTO-ST Institute
AS2M department
25000 Besançon, France.
[**] *e-mail: Ahmed.Mosallam@femto-st.fr*
[***] *e-mail: First Name.Last Name@ens2m.fr*

Abstract: Maintenance is becoming more expensive nowadays due to the increased complexity in design and function of industrial systems. Continuous health monitoring is thus of high importance to increase the availability of industrial systems and consequently reduce the costs. This paper presents an algorithm for unsupervised trends extraction from multidimensional sensory data so as to use such trend in machinery health monitoring and maintenance needs. The proposed method does not assume any prior knowledge about the nature and type of the input signals. It is based on extracting successive multi-dimensional features from machinery sensory signals. Then, unsupervised feature selection on the features domain is applied without making any assumptions concerning the source of the signals and the number of the extracted features. Finally, empirical mode decomposition algorithm (EMD) is applied on the projected features with the purpose of following the evolution of data in a compact representation over time. The algorithm is demonstrated on accelerated degradation dataset of bearings acquired from PRONOSTIA experimental platform and a second dataset acquired form NASA repository where it is shown to be able to extract interesting signal trends.

Keywords: Feature Extraction, Health Indicator, Trend Construction, Health State Detection, Prognostics.

1. INTRODUCTION

The degradation of critical components is one of the main reasons of machines breakdown. Moreover, faulty components can also affect other components in the system which might lead to sever consequences. Effective condition monitoring and health assessment of machinery deterioration is therefore crucial for reducing the downtime and the operation costs while increasing the availability.

Machine Prognostics and Health Management (PHM) is the process of detecting abnormal conditions, diagnosis of the fault and their cause and prognostics of future fault progression, as reviewed by Tobon-Mejia et al. (2012), Heng et al. (2009) and Schwabacher (2005). Maintenance strategies can be classified into three categories 1) Breakdown maintenance, 2) Preventive maintenance and 3) Condition-based maintenance (CBM) Jardine et al. (2006). In breakdown maintenance no actions are taken to maintain the equipment until it breaks down. Preventive maintenance only includes setting a periodic intervals for machine inspection. CBM observes the condition measurements of the system and maintenance actions are performed once indicator(s) show(s) that equipment performance is deteriorating. Compared to other maintenance strategies, CBM attempts to avoid unnecessary tasks by taking actions only when there is evidence of abnormal behaviour in the machine. Thus, CBM can produce signif-

icant savings through more efficient scheduling and therefore has been widely explored in research and industry.

PHM methods can be classified into three categories: anomaly detection, diagnostics and prognostics. Machinery condition monitoring research has been shifting lately from fault detection and diagnostics to prognostics Jardine et al. (2006). Prognostics can be defined as the estimation of time to failure and the risk for one or more failure modes ISO (2004). It usually requires the past history of the machinery condition sensory data to extract health progression index. The extracted health progression indexes are then used to predict the future machinery health status. The process of extracting representative health progression data from sensory time series data is known as trend extraction which can be performed using single or multiple parameters. Multiple parameter analysis is critical in that relations between parameters; not only the parameters themselves, that can be utilised to represent the deterioration. It can be very challenging to process and to extract informative representation as raw signals usually exhibit tremendous amount of oscillations and partly or completely hidden by noise.

In this work, time and frequency domain features have been extracted from two sensory data series. The features have been grouped according to the quality of the relations between each of which. Then, each group of features has been projected into a compact two dimensional represen-

tation using principal component analysis (PCA) Jolliffe (2002). In order to get monotonic signals, the final trends were extracted from each projected group using the empirical mode decomposition algorithm (EMD) Huang et al. (1998).

This paper is structured as follows: a brief literature review is presented in Section 2. In section 3, the proposed method is explained. The experimental set-up and results are presented Section 4. Finally, conclusion of this paper and suggestions for future work is depicted in Section 5.

2. BACKGROUND

In this section a brief literature review on PHM is provided:

2.1 Anomaly detection

anomaly detection is the simplest part of PHM process; which can be defined as the process of identifying when a fault/outlier has occurred. This section is dedicated to review some of the research work which conducted for this task. The algorithm proposed by Filev and Tseng (2006) can be used for online anomaly detection. It is not based on any learning algorithm, instead, the algorithm decides different operation modes once it starts running on N observations. Any change on the number of the clusters or the behaviour of the new data in an already defined cluster will be considered as suspicious. The proposed method is divided into two main parts, Initialisation and Monitoring. For initialisation step, 27 features were extracted, however, the authors did not specify exactly all the features. Then, data standardisation has been performed by applying unit variance and mean centring. PCA has been used for data dimensionality reduction. Finally estimating the unknown numbers of operating modes (OM) clusters, by using greedy expectation maximisation clustering algorithm. For monitoring part, first feature extraction has been applied. Then, calculating of the online data and updating the modelled mean and variance of the old dataset have been performed. For the online readings, data standardisation and dimensionality reduction have been performed similar to the initialisation part. Then, updating OM clusters by using Mahalanobis distance measure. Finally, tracking OM by using evolving Takagi-Sugeno model in order to predict the dynamics of the data within each cluster. Tavares et al. (2011) performed a comparison study between the performance of PCA and partial least square (PLS) for fault detection. The proposed method is divided into three parts, feature extraction, modelling and deviation detection. For feature extraction, 16 signals have been measured from moving gate-type incinerator. Building the model on healthy datasets using PCA and PLS. Finally, T square and Q statistics have been proposed to detect faults. The results show that both PCA and PLS gave good results in detecting faults.

2.2 Diagnostics

Another important aspect in PHM systems is diagnostics and evaluation of wear advancements. Cempel (2002) proposed a multidimensional diagnostics approach for mechanical systems. Vibration signals were acquired from diesel engine, heavy fan and rubbing blades in a turbo-set to validate the approach. The authors applied centring and normalisation for the signals. Then, using Singular Value Decomposition (SVD), the multidimensional dataset were reduced to a lower dimension. Finally, health evolution indexing and fault diagnostics has been proposed by choosing the most informative SVD indexes. Self-organised maps (SOM) have been proposed by Moshou et al. (2010) for fault diagnostics and assessment of fault severity. Kurtosis and line integral of acceleration signal have been extracted from bearings vibration signals for different faults. Then, a SOM has been trained using the extracted features. Empirical Mode Decomposition (EMD) was also proposed in many works as it returns smooth monotonic signals. Miao et al. (2010) proposed an algorithm for machinery health monitoring and early fault detection. The method decomposes the input signal into Intrinsic Mode Function (IMF) using EMD, then Combined Mode Function (CMF) is applied to mix neighbouring IMFs to obtain the best signal. Feature extraction is then performed using Fourier Transform (FT) on the acquired signal. Finally, the method uses HMM to build a model of the normal condition of the gearbox and Average Probability Index (API) is constructed as an index for machinery health status. W. and M. (1998) performed a comparison study on the performance of 4 statistical indexes such as crest factor, kurtosis, skew and beta distribution function. The study shows that there is no significant advantage in using beta function compared to using kurtosis and crest factor for detecting and identifying different machinery defects. The paper also shows that the statistical parameters are affected by the shaft speed. The proposed method is only based on extracting four features mentioned earlier from vibration and sound signals acquired from test rig.

2.3 Prognostics

In the work proposed by Zhang et al. (2006) 16 time and frequency features have been extracted from raw vibration signals acquired from double suction pump. The authors then used PCA to merge features and to project the multidimensional features vector into a compact indicator. Finally, fault threshold has been calculated using Best Efficiency Point. Li et al. (2009) applied Auto Regression (AR) filter on vibration signals from bearing test rig. The model parameters have been estimated using Levinson-Durbin recursion (LDR). Then, the energy ratio between the random parts and the original signal was used as fault indicator. The algorithm is simple to implement and suitable for highly accelerated signals. Zhang et al. (2005) presents an integrated framework for fault detection, diagnosis and prognosis using Hidden Markov Model (HMM). The proposed method starts the data preprocessing by using frame blocking, frequency spectral analysis and noise filtering. Then, using PCA the dimensionality of the dataset has been reduced. Next the health status estimator has been built using HMM and health index interpolation by using Paris's formula. The method was defied on vibration data acquired from test bed.

From the review it can be seen that the processing of sensory signals depends on the nature of the signals and the purpose of the research. Most of the work assume specific type of input signals and the number of the sensors

used in the experiments. The goal of this work is to extract signals that represent critical components' health evolution over the time. In this work the relationships between the different signals acquired from the sensors have been used without making any assumption on the number and the type of the acquired signals.

3. THE METHOD

The idea is to extract indexes from successive multi-dimensional features acquired from machinery sensory signals, without making any assumptions concerning the source of the signals and the number of the extracted features, which reflect the machinery deterioration over time. The algorithm consists of four main steps as shown in Figure 1
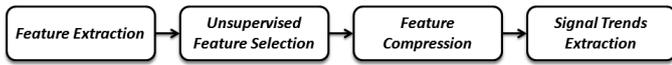
Fig. 1. Steps of the proposed method

3.1 Feature extraction

Gathered signals from machine components contain generally immense number of data which require a large amount of memory and computation power to be analysed as shown in Figure 2. Instead input data can be reduced into a lower but informative representation by extracting meaningful features from raw signals as shown in Figure 3. These features can be derived from time domain, frequency domain or joint time-frequency domain. In this work features from both time and frequency domains have been extracted from raw signals and are listed in table 1.
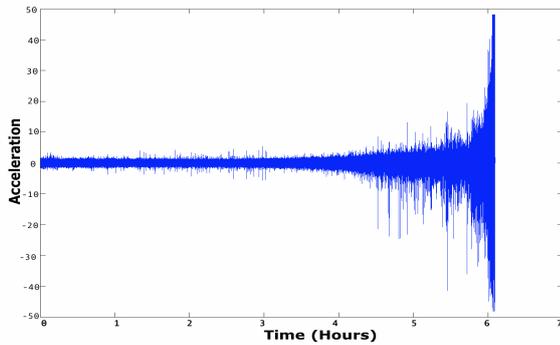

Fig. 2. Example of raw signal extracted from bearing testbed

3.2 Unsupervised feature selection

Not all extracted features acquired in the previous step are interesting, that is, contain information about the degradation of the system. We are interested in signals that have non random relationships and consequently contain information about system degradation. To select such signals, an unsupervised feature selection algorithm proposed by Mosallam et al. (2011) is applied. The algorithm first calculates pairwise symmetrical uncertainty for all the input signals, defined by:
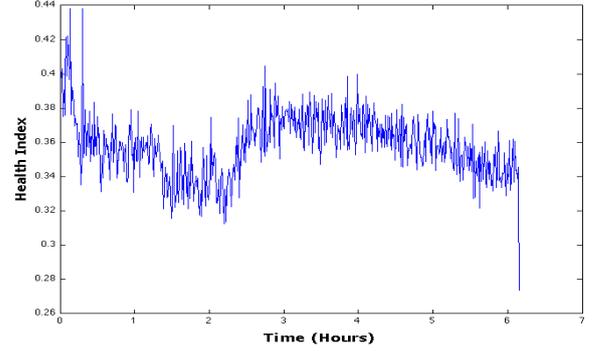

Fig. 3. Example of RMS signal extracted from sensory bearing testbed

$$SU(X, Y) = 2 \frac{I(X, Y)}{H(X) + H(Y)} \qquad (1)$$

where I(X,Y) is the mutual information between two random variables X and Y; H(X) and H(Y) represent the information entropy of a random variable. Then, the algorithm measures the distance between all the pairs using hierarchical clustering. The algorithm finally ranks the resulting clusters according to the quality of the included signals in representing interesting relationships, that is, contain information about machinery degradation. The selected features are then compressed into lower dimension using PCA as explained in the proceeding step.

3.3 Feature compression

In order to follow the trajectories of selected features over time, the number of features has to be reduced to a compact form. The goal in this step is to compress the n features selected in the previous step onto one-dimensional space. One way to compress the variables, i.e. reducing the number of their dimensions, without much loss of information is by using PCA. It projects data from features to principal component domain while keeping the highest variance by any projection of the data on the first principal component and the second greatest variance on the second principle and so on. In this work we use standard PCA method where eigenvalue λ_i and eigenvectors v_i have been calculated for covariance matrix C of the selected features

$$C\lambda_i = \lambda_i v_i \qquad (2)$$

Then, the first component is used to represent the health status evolution with respect to time as shown in Figure 4. The compressed features are then further processed to get a smooth trend of the health status.

3.4 Trends extraction

So far, the signals produced from the PCA compression step exhibit high level of oscillations and noise. Such oscillations make the signal hard for visualisation and difficult to model. In this section, the internal structure of the data is extracted in a way which best explains the degradation in a simple monotonic signal and the rest of oscillations are discarded using empirical mode decomposition algorithm "EMD". EMD, which was originally proposed by Huang
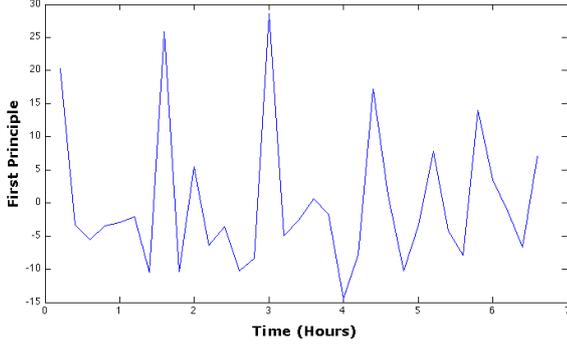
Fig. 4. Example of group of selected feature compressed using PCA

Table 1.

Summary of features extracted from raw signals

Feature	Formula
Peak-to-Peak	$mean(upper_{pks}) + mean(lower_{pks})$
Maximum peak value	$max(findpeaks(signal))$
Root mean square	$\sqrt{\frac{1}{n}(x_1^2 + x_2^2 ... + x_n^2)}$
Kurtosis	$\frac{E(x-\mu)^4}{\sigma^4}$
Skewness	$\frac{\sum_{i=1}^{N}(x_i - \bar{x})^3}{(N-1)\sigma^3}$
Mutual information	$I(X,Y) = H(X) + H(Y) - H(X,Y)$
Entropy	$H(X) = -\sum_{x \in X} p(x) \log p(x)$
Arithmetic mean of PSD	$20 log_{10} \frac{\frac{1}{n}\sum abs(fft(x))}{10^{-5}}$
Line integral	$i = \sum_{i=0}^{n} abs(x_{i+1} - x_i)$
Autoregressive model	$x_t = c + \sum_{i=1}^{p} \phi_i x_{t-i} + \epsilon_t$
Energy	$e = \sum_{i=0}^{n} x_i^2$

et al. (1998), is a method to decompose a signal into a successive Intrinsic Mode Functions (IMF), such as depicted in Figure 5, and is composed of the following steps:

- Find all the local maxima and minima of the input signal and compute the corresponding upper and lower envelopes using cubic spline respectively
- Subtract the mean value of the upper and lower envelopes from the original signal.
- Repeat until the signal remains nearly unchanged and obtain IMF_i
- Remove IMF_i from the signal and repeat if it is neither a constant nor a trend.

4. EXPERIMENTS

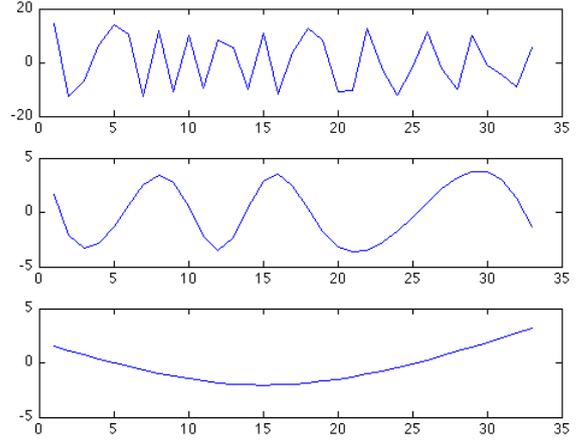The algorithm has been verified using two different dataset, PRONOSTIA and NASA datasets; the details

Fig. 5. Example of decomposition process using EMD

of the datasets and the experiments are explained in the following subsection:

4.1 PRONOSTIA dataset

PRONOSTIA is an experimentation platform dedicated to test and validate bearing fault detection, diagnostic and prognostic approaches, and it is located in AS2M department, FEMTO-ST institute. The main objective of the platform is to provide accelerated experimental data sets that characterise the degradation of ball bearings along their operational life. It is composed of four main parts: a rotating part, a degradation generation part (with a radial force applied on the tested bearing) a measurement part and test bearings as shown in Figure 6. Bearing's dataset consists of data acquired from three sensors: 2 vibration and one temperature sensors. The vibration sensors consists of two miniature accelerometers positioned at 90 degree to each other; the first is placed on the vertical axis and the second is placed on the horizontal axis. The two accelerometers are placed radially on the external race of the bearing. In this work three data sets, acquired from this platform, have been used to validate the algorithm, where only the acceleration data has been used. For the
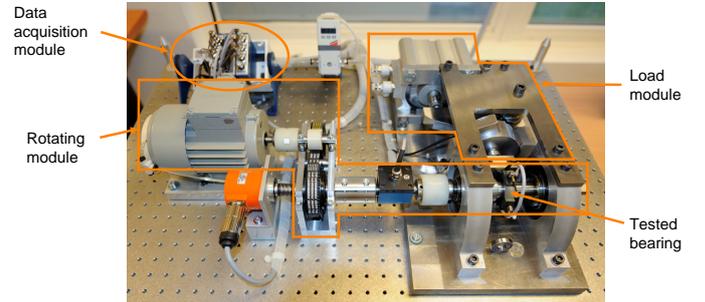

Fig. 6. PRONOSTIA experimentation platform

first dataset, the algorithm generates 9 signals, each of that represents evolution of the machinery degradation over the time with different trajectories. Figure 7 shows two selected monotonic health indices for the first experiment. As can be seen from the figure the trends show a smooth increasing signals over the time which can be used to

deduce health status of the machinery. Figure 7a is the result of fusing two AR parameters from the first sensor and one parameter from the second sensor. Figure 7b shows another indicator which was a result of fusing also three features namely, skewness for the first sensor along with the second parameter of AR model and line integral feature extracted from the second sensor signal.
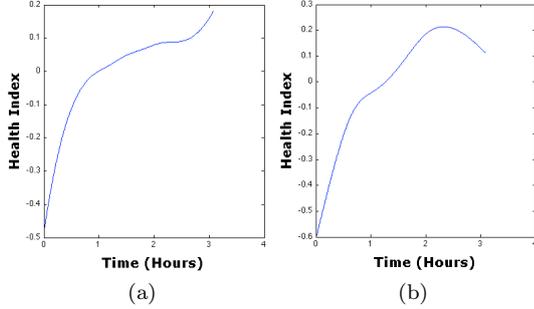

Fig. 7. Results of PRONOSTIA first data set

For the second dataset, the algorithm generates 10 signals, each of that indicates the health degradation over the time. Figure 8 shows two selected monotonic health indices for the first experiment. Figure 8a; which is almost linear function, is the result of fusing two features, maximum peak value and AR parameter, both have been extracted from the first sensor. Figure 8b shows a smooth monotonic function. It is a result of fusing also two features, RMS and line integral acquired from the second sensor.
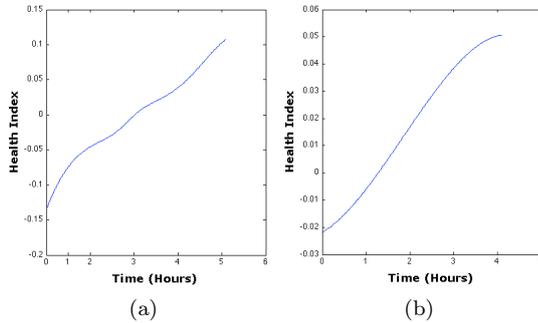

Fig. 8. Result of PRONOSTIA second data set

For the third dataset, the algorithm generates 9 signals that indicate the health degradation over the time. Figure 9 shows two selected monotonic health indices for the aforementioned experiment. Figure 9a; which is almost linear function, is the result of fusing three features, peak-to-peak, maximum peak value and AR parameter where all the features have been extracted from the first sensor. Figure 9b shows a smooth decreasing monotonic function. It is a result of fusing three features, arithmetic mean of PSD extracted from both sensors and kurtosis which has been extracted from second sensor.

4.2 NASA dataset

Three data sets acquired from four bearings were installed on one shaft. The rotation speed was kept constantly at 2000 RPM. A radial load is placed onto the shaft and bearing by a spring mechanism. In this work only the first
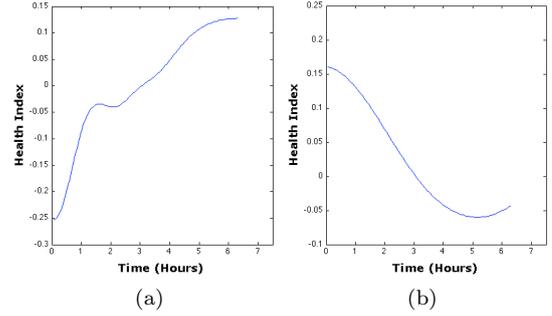

Fig. 9. Result of PRONOSTIA third data set

two sensors readings have been used to validate the algorithm. The data set is fully described on NASAs DASH-link web (c3.ndc.nasa.gov/dashlink/). For the first dataset the algorithm generates 10 trends from the raw data. Figure 10 shows two selected monotonic health indices for the aforementioned experiment. Figure 10a shows the result of fusing four features; Peak-to-Peak, Maximum peak value and Energy acquired from the second sensor and Entropy from the first sensor. Figure 10b shows a smooth decreasing monotonic function. It is a result of fusing two features, Mutual information between both sensors and entropy which has been extracted from second sensor.
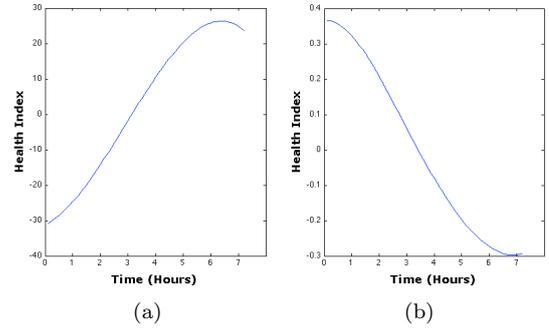

Fig. 10. NASA's prognostics first data set results

For the second dataset the algorithm generates 6 trends from the raw data. Figure 11 shows two selected monotonic health indices for the aforementioned experiment. Figure 11a shows the result of fusing two features; Maximum peak value acquired from the second sensor and Energy acquired from the first sensor. Figure 11b shows a smooth decreasing monotonic function. It is a result of fusing two features, Entropy of the first sensor and Arithmetic mean of PSD acquired from the second sensor.

For the final dataset the algorithm generates 10 trends from the raw data and non of them was monotonic. Figure 12 shows two selected non-monotonic health indices for the same experiment. Figure 12a is the result of fusing two features, Peak-to-Peak; acquired from first sensor, and Entropy of both sensors. Figure 12b shows another indicator which was a result of fusing two features namely, Root mean square and AR coefficient acquired from second sensor.
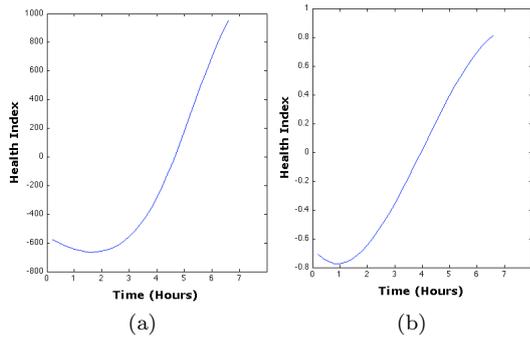
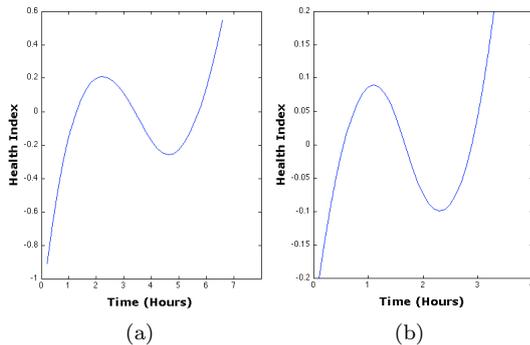Fig. 11. NASA's prognostics second data set results

Fig. 12. NASA's prognostics third data set results

5. CONCLUSION

In this work a trend extraction algorithm from sensory data has been proposed. The generated trends have to represent the progression of machinery health status and shall be used for later fault detection, diagnostics and prognostics. The proposed method is based on extracting successive multi-dimensional features from machinery sensory signals acquired from machines critical components. Then, unsupervised feature selection on the features domain is applied without making any assumptions concerning the source of the signals and the number of the extracted features. Finally EMD is applied on the projected features with the purpose of following the evolution of data in a compact representation over time. The algorithm is demonstrated on accelerated degradation dataset of bearings acquired from PRONOSTIA experimental platform and NASA dataset where it is shown to be able to extract interesting signal trends. The results show that the algorithm is generic and can extract the progression of the machinery health status in a compact form. The generated trends could be used for machinery prognostics and health assessment.

REFERENCES

Cempel, C. (2002). Multidimensional condition monitoring of mechanical systems in operation. *Mechanical Systems and Signal Processing*, 17(6), 1291–1303.

Filev, D.P. and Tseng, F. (2006). Novelty detection based machine health prognostics. *Evolving Fuzzy Systems, 2006 International Symposium on*, 193–199.

Heng, A., Zhang, S., Tan, A.C.C., and Mathew, J. (2009). Rotating machinery prognostics: State of the art, challenges and opportunities. *Mechanical Systems and Signal Processing*, 23(3), 724–739.

Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.C., Tung, C.C., and Liu, H.H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1971), 903–995.

ISO (2004). Condition monitoring and diagnostics of machines - prognostics - part 1: General guidelines. int. standard iso13381-1. *Techincal report*.

Jardine, A.K.S., Lin, D., and Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, 20(7), 1483–1510.

Jolliffe, I.T. (2002). Principal Component Analysis and Factor Analysis. In *Principal Component Analysis*, Springer Series in Statistics, chapter 7, 150–166. Springer New York, New York.

Li, R., Sopon, P., and He, D. (2009). Fault features extraction for bearing prognostics. *Journal of Intelligent Manufacturing*, 1–9. doi:10.1007/s10845-009-0353-z.

Miao, Q., Wang, D., and Pecht, M. (2010). A probabilistic description scheme for rotating machinery health evaluation. *Journal of Mechanical Science and Technology*, 24(12), 2421–2430.

Mosallam, A., Byttner, S., Svensson, M., and Rognvaldsson, T. (2011). Nonlinear relation mining for maintenance prediction. *Aerospace Conference, 2011 IEEE*, 1–9.

Moshou, D., Kateris, D., Sawalhi, N., Loutridis, S., Gravalos, I., Diamantaras, K., Duch, W., and Iliadis, L. (2010). *Fault Severity Estimation in Rotating Mechanical Systems Using Feature Based Fusion and Self-Organizing Maps*, volume 6353, 410–413. Springer Berlin / Heidelberg.

Schwabacher, M. (2005). A survey of data-driven prognostics. *AIAA InfoTech Aerospace*.

Tavares, G., Zsigraiová, Z., Semiao, V., and Carvalho, M.d.G. (2011). Monitoring, fault detection and operation prediction of msw incinerators using multivariate statistical methods. *Waste Management*, 31(7), 1635–1644.

Tobon-Mejia, D.A., Medjaher, K., and Zerhouni, N. (2012). Cnc machine tool's wear diagnostic and prognostic by using dynamic bayesian networks. *Mechanical Systems and Signal Processing*, 28, 167–182.

W., H.R.B. and M., N.M.J. (1998). Statistical analysis of sound and vibration signals for monitoring rolling element bearing condition. *Applied Acoustics*, 53(1), 211–226.

Zhang, S., Hodkiewicz, M., Ma, L., Mathew, J., Kennedy, J., Tan, A., and Anderson, D. (2006). *Machinery Condition Prognosis Using Multivariate Analysis*, 847–854. Springer London.

Zhang, X., Xu, R., Kwan, C., Liang, S., Xie, Q., and Haynes, L. (2005). An integrated approach to bearing fault diagnostics and prognostics. In *Proceedings of the 2005, American Control Conference, 2005.*, 2750–2755. IEEE, Portland, OR, USA.