Data-Driven Prognostics Based on Health Indicator Construction: Application to PRONOSTIA's Data

K. Medjaher, N. Zerhouni, J. Baklouti

Abstract— **Failure prognostics can help improving the availability and reliability of industrial systems while reducing their maintenance cost. The main purpose of failure prognostics is the anticipation of the time of a failure by estimating the Remaining Useful Life (RUL). In this case, the fault is not undergone and the estimated RUL can be used to take appropriate decisions depending on the future exploitation of the industrial system. This paper presents a data-driven prognostic method based on the utilization of signal processing techniques and regression models. The method is applied on accelerated degradations of bearings performed under the experimental platform called PRONOSTIA. The purpose of the proposed method is to generate a health indicator, which will be used to calculate the RUL. Two acceleration sensors are used on PRONOSTIA platform to monitor the degradation evolution of the tested bearings. The vibration signals related to the degraded bearings are then compared to a nominal vibration signal of a non-degraded bearing (nominal bearing). The comparison between the signals is done by calculating a correlation coefficient (which is considered as the health indicator). The values of the calculated correlation coefficient are then fitted to a regression model which is used to estimate the RUL.**

I. INTRODUCTION

Prognostics and Health Management (PHM) [1]–[3] is a process involving data acquisition and processing, fault detection and diagnostic, fault prognostics and decision support. Its main purpose is to detect, diagnose and anticipate the faults on the system and take appropriate decisions to maintain it in time. PHM can thus help improving the availability, the reliability and the security of systems. It helps also reducing the maintenance costs.

Contrary to fault detection and isolation (or fault diagnostic) which is well developed and spread within the research and the industrial communities [4], [5], fault prognostics is relatively a recent activity which is gaining more and more recognition [3], [4], [6], [7]. Fault prognostics aims at estimating the Remaining Useful Life (RUL) of a system. The RUL estimation can be done by using three main approaches: model-based prognostics (also called physics of failure prognostics), data-driven prognostics and hybrid prognostics. Model-based prognostics uses mathematical models of the system including degradation models to estimate the current health status of the system and predict its future one leading to RUL estimation. Data-driven prognostics is based on the utilization of data provided by monitoring sensors to built a model in terms of states or trends, which are then used to predict the RUL. Finally, the hybrid prognostics merges

K. Medjaher and N. Zerhouni are with FEMTO-ST Institute, AS2M department, UMR CNRS 6174 - UFC / ENSMM / UTBM, 25000 Besançon, France. E-mail: kamal.medjaher@ens2m.fr

both previous approaches. Model-based prognostics has the advantage of giving more precise RUL predictions than data-driven prognostics. However, in terms of implementation, data-driven approach is more easier than model-based approach. This is because for complex systems, with presence of several components and nonlinearities, it is not trivial to derive mathematical models of their behavior. In these cases, data-driven prognostics can be the solution which makes the trade-off between precision, complexity and cost of implementation.

This paper presents a data-driven method for the estimation of RUL of bearings. These latter components are present in rotating machines and their failures can cause unavailability of the machine and loss in productivity. They can thus be considered as critical components for which the fault progression needs to be monitored and assessed during time. The main idea of the proposed method is to build a health indicator which can be used to track the fault progression of the bearing and predict its RUL. To build the health indicator, two bearings from a same category are considered: the first bearing is normal (without faults) and is taken as a reference and the second bearing presents a degradation. The signal of the degraded bearing is then continuously compared to the signal of the nominal bearing and the difference between them is a sort of residual evolving during time and which is taken as a health indicator of the bearing. Finally, the projection of the health indicator can be exploited to predict the health status of the bearing and calculate its RUL. The proposed method is applied on experimental acceleration signals acquired from the platform PRONOSTIA [8] and the results show the effectiveness of the method.

The paper is organized as follows. After the introduction, section 2 deals with a brief recall of fault prognostics paradigm. Section 3 presents the method proposed and section 4 gives details on the application and the obtained results. Finally, section 5 concludes the paper.

II. PROGNOSTICS AND HEALTH MANAGEMENT

Prognostics and Health Management (PHM) is a key process for Condition-Based Maintenance (CBM) or Predictive Maintenance (PM). The modules of a CBM are shown in Fig. 1. A brief description of the activity of each module is given hereafter.

- Sensors: the number and type of sensors are chosen according to the system and its physical phenomena and degradations to monitor.
- Data acquisition: the signals related to the measured parameters are gathered and stored by using appropriate
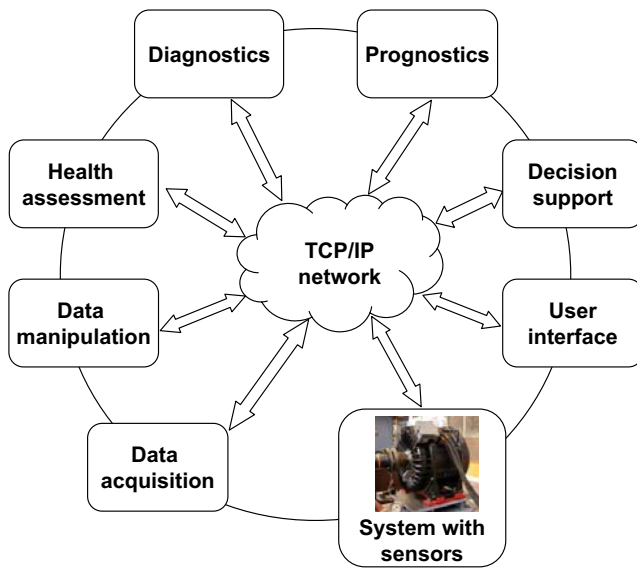
Fig. 1. Modules of a CBM.

Fig. 2. Illustration of remaining useful life.

Fig. 3. Fault prognostics approaches.

acquisition equipments and sampling frequencies depending on the dynamic of the monitored parameters (the sampling frequency of vibration is not the same than that one of temperature). Furthermore, some pre-processings can be applied on the data at this stage (re-sampling for example).

- Data manipulation: the acquired data are processed to extract, reduce and select relevant features and indicators that can be used to estimate and predict the fault progression in the system.
- Health assessment: this module deals with the classification and the detection of the states of the system. It can be assimilated to fault detection.
- Diagnostics: this module concerns the isolation and the identification of faults' causes.
- Prognostics: the aim of the module is to predict the remaining useful life of the system.
- Decision support: provides a set of decisions and actions which can help conducting the system to reach the defined objectives or maintaining the system by preparing the necessary resources.

The following of the paper concerns the fault prognostics module. Fault prognostics is defined by the international standard organization as the estimation of the Time To Failure (ETTF) and the risk of existence or later appearance of one or more failure modes [9]. In the reported research and application works, the terminology ETTF is commonly called Remaining Useful Life (RUL). An illustration of RUL progression is shown in Fig. 2. Fault prognostics can be done by using three main approaches: model-based, data-driven and hybrid prognostics (Fig. 3).

Model-based (also called physics of failure) methods deal with the exploitation of a mathematical model representing the behavior of the physical component including its degradation. The deri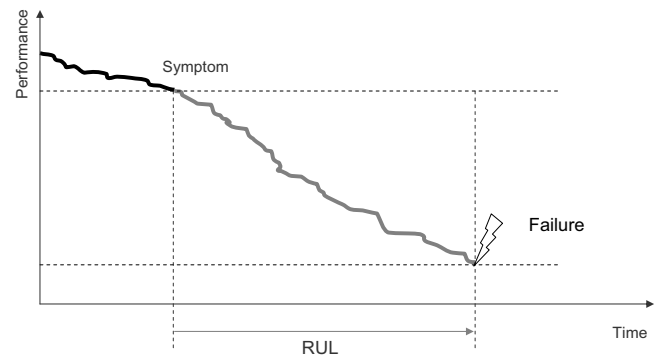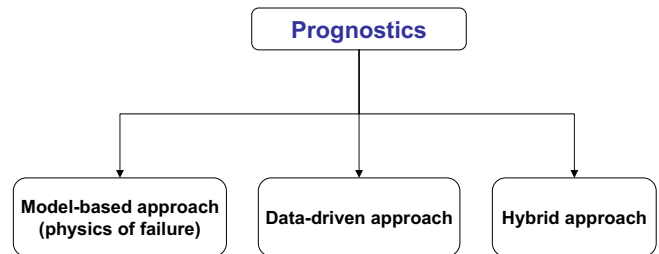ved model is then used to predict the future evolution of the degradation and to estimate the RUL [10], [11]. In this case, the prognostics consists in evolving the degradation model until a determined future instant from the actual deterioration state and by considering the future use conditions of the corresponding component.

The main advantage of this approach is its precision, since the predictions are achieved based on a mathematical model of the degradation. However, the derived degradation model is specific to a particular kind of component or material, and thus, can not be generalized to all the system's components. In addition to that, getting a mathematical model of degradation is not an easy task and needs well instrumented test-benches which can be expensive.

Data driven methods concern the transformation of the monitoring and/or the exploitation data into relevant models, which can be used to assess the health state of the industrial system and predict its future one leading to the estimation of its RUL [1], [12], [13]. Generally, the raw data are first processed to extract features which are then used to build the prognostic models. The features can be temporal, frequency or both. In some applications, individual features are not sufficient and one needs to combine them in order to build what can be called health indicators. Note that data-driven prognostic methods can use data provided by sensors or obtained through experience feedback (operation, maintenance, number of breakdowns, etc.).

The advantage of data-driven prognostic approach is its applicability, cost and implementation. Indeed, by using these methods, it is possible to predict the future evolution of degradation without any need of prior mathematical model of the degradation. However, the results obtained

by this approach are less precise than those obtained by using model-based methods. Compared to model-based methods, the data-driven methods offer a trade-off in terms of complexity, cost, precision, and applicability. They are suitable for systems where it is easy to obtain monitoring data and transform them into behavior models of the degradation phenomena. The following of the paper presents a data-driven method to predict the RUL of bearings.

III. HEALTH INDICATOR FOR PROGNOSTICS

The proposed method is a component-oriented prognostics. For this purpose, we suppose that the RUL of the whole system corresponds to the RUL of its critical components (more precisely to the types of the degradations of the critical components). A general scheme of a component-oriented PHM is shown in Fig. 4.

To apply the method, the user has first to identify the critical components of the system and the main parameters related to the fault progression to measure. For example, for a lift the critical components can be the electrical motor and the opening and closing mechanism of the door. Then, the user selects accordingly the number and the types of sensors to install on the identified components in order to gather appropriate signals representative of the degradations. The sensors can be temperature, vibration and acoustic emission for bearings, current and voltage for electrical motors, displacement, velocity and effort for mechanical parts, etc. The signals are finally processed to build health indicators which can be used to track the progression of the faults on the system. Different processing techniques are available and can be used. Some examples of signal processing techniques are: temporal and frequency feature extraction and Wavelet Packet Decomposition (WPD). The process of health indicator construction is shown in Fig. 5.

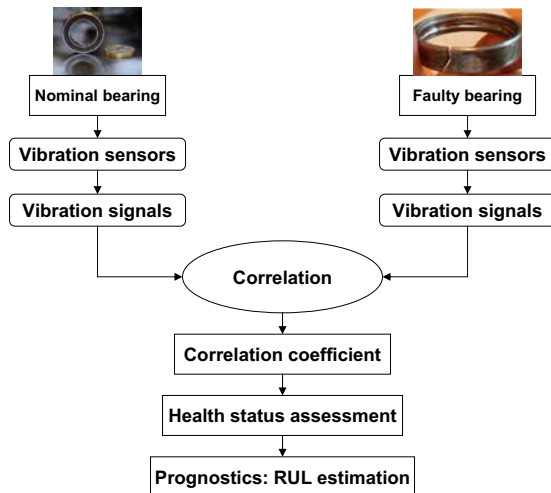The health indicator corresponds to the difference (or dis-

Fig. 5.   Process of construction of the health indicator.

tance) between the vibration signals of the degraded bearing and the nominal bearing. It is obtained by calculating the

correlation coefficient between the two vibration signals (Eq. 1). In equation (1), x_i and y_i are two vectors of size N, and $\bar{x}_i$ and $\bar{y}_i$ their corresponding means.

$$HI = \frac{\sum\limits_{i=1}^{N} (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum\limits_{i=1}^{N} (x_i - \bar{x})^2} \cdot \sqrt{\sum\limits_{i=1}^{N} (y_i - \bar{y})^2}} \qquad (1)$$

To build the health indicator, several accelerated degradation experiments should be conducted on a set of bearings of the same type. The vibration signals obtained from each experiment are then compared to the vibration signals corresponding to nominal bearings (without faults or degradations). The correlation coefficient between the two categories of bearings (nominal and degraded) should decrease in time as the degradation increases. This indicator is then used to assess the progression of the degradation and to estimate the RUL. In practice, due to high sampling frequency of the vibration signals, the health indicator needs some pre-processings before using it. The pre-processings are: re-sampling, filtering (or smoothing) and non-linear regression fitting (Fig. 6)
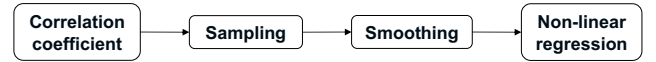

Fig. 6.   Pre-processing of the health indicator.

IV. APPLICATION AND RESULTS

A. Description of the platform PRONOSTIA

The accelerated bearing life test bed is called PRONOSTIA (7), which it is an experimentation platform dedicated to test and to validate bearing health assessment, diagnostic and prognostic. In the present experimental setup a natural degradation process of bearings is performed. During the experiments any failure types (inner race, outer race, ball, or cage) or their combinations could occur. This is allowed in the system to better represent a real industrial situation.

The experimental platform PRONOSTIA is composed of two main parts: a first part related to the speed variation and a second part dedicated to load profiles generation. The speed variation part is composed of a synchronous motor, a shaft, a set of bearings and a speed controller. The synchronous motor develops a power equal to 1.2 kW and its operational speed varies between 0 and 6000 rpm. The second part is composed of a hydraulic jack connected to a lever arm allowing to create different loads on the bearing mounted on the platform for degradation.

Two high frequency accelerometers (DYTRAN 3035B) are mounted horizontally and vertically on the housing of the test roller bearing to pick up the horizontal and the vertical accelerations. In addition, the monitoring system includes one PT100 to measure the temperature of the tested bearing. A speed sensor and a torque sensor are also available on the PRONOSTIA platform. A data acquisition card is also
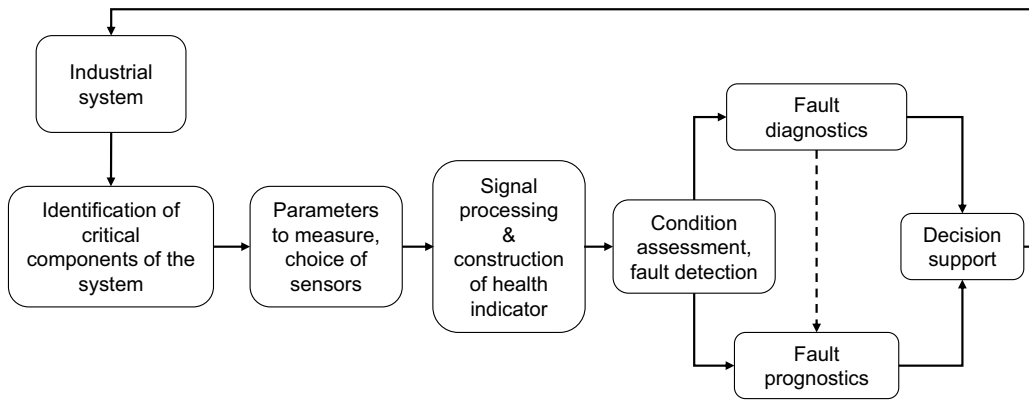
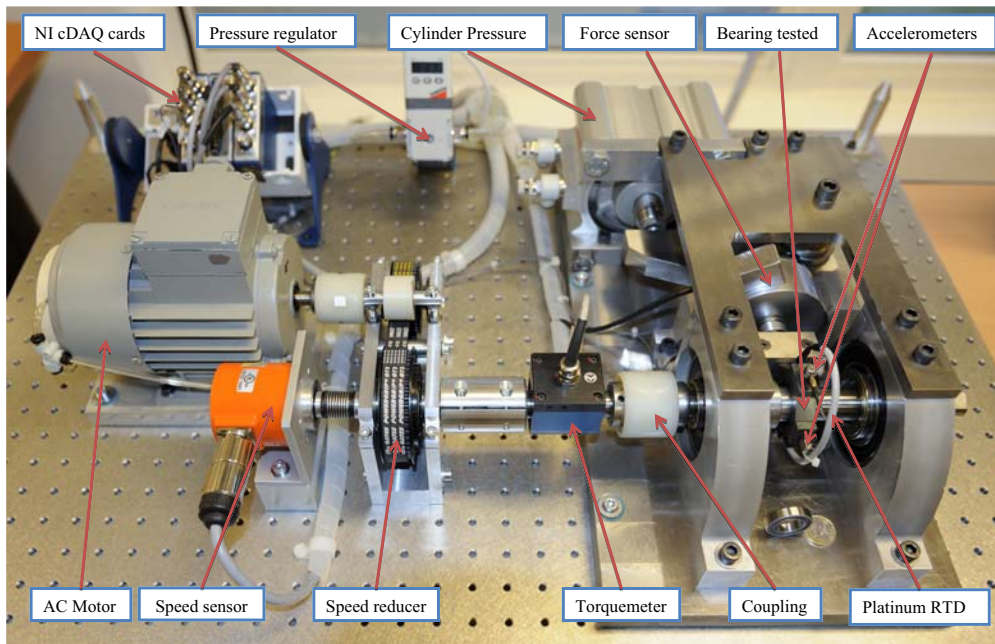Fig. 4.   Component-oriented prognostic process.

Fig. 7.   The experimental platform PRONOSTIA.

used, which allows the integration of three modules: the first includes the two acceleration sensors, the second is for the temperature probe and the third for the torque transducer.

The monitoring data are transmitted to a dedicated computer where they are stored. The data acquisition software is programmed by using a LabView interface. The readings can be directly taken from the digital readout on the analyzer, graphical representation of the data can be displayed on the screen and finally, the data can be analyzed online or offline during later processing. The sampling frequency of the data acquisition card is set to 25600 Hz and the vibration data provided by the two accelerometers are collected every 1 second. Each record is stored in a matrix format where the following parameters are defined: the time, the horizontal acceleration, the vertical acceleration, the temperature, the speed and the torque.

The bearing operating conditions are determined by instan-taneous measures of the radial force applied on the bearing, the rotation speed of the shaft handling the bearing and of the torque inflicted to the bearing. Thus, three sensors are used: a load cell and its transducer amplifier, an incremental encoder and its analogue signal converter, and a torque transducer with its converter. With this experimental platform, several profiles can be created by varying the load and the speed. This is very important as it allows simulating constant as well as variable operating conditions for bearing's degradation. For the measurement of the degradation's characteristics during the test, the bearing starts from its nominal operating mode until its faulty operating mode or state. The bearing's behavior is captured during its whole degradation process by using the different sensors.

B. Simulation results

The prognostic method described in section III is ap-plied on two bearings with same physical characteristics:
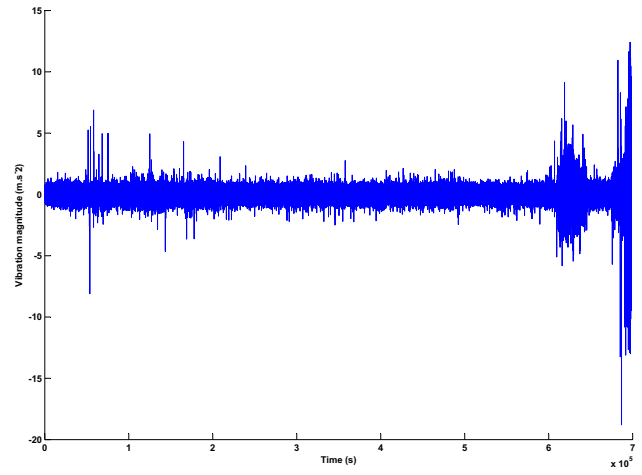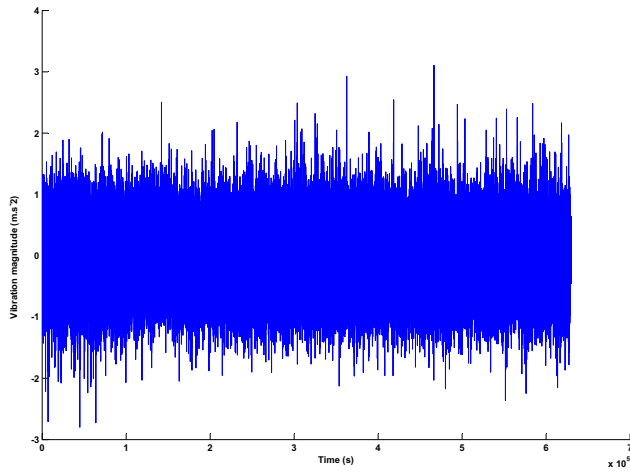
Fig. 8. Raw vibration signals: normal bearing (left) and degraded bearing (right).

a nominal bearing (without degradation) and a degraded bearing. The duration of the experiments is identical for both bearings, it is equal to 4 hours and 33 minutes. The figure 8 shows the raw vibration signals obtained from the tested bearings. From Fig. 8 one can see that it is difficult to make conclusions about the degradation. Thus, more processing is needed on these raw signals.

The health indicator progression calculated by using the correlation equation (Eq. 1) is shown in Fig. 9. To use the

Fig. 10. Re-sampling of the health indicator.

distance between the current time and the time for which the regression model given in Eq. 2 reaches the failure threshold (Eq. 3). The threshold is defined as the acceptable limit of the vibration magnitude of each degraded bearing, which corresponds in this application to the end of each experiment. The exponential fitting of the smoothed health indicator is shown in Fig. 12. The exponential model obtained after fitting is given by the following equation:

$$HI\left(t\right) = 0.2394.e^{-2.245 \times 10^{-5}.t} \tag{2}$$

$$RUL\left(t\right) = t_{final} - HI^{-1}\left(t\right) \tag{3}$$

where t_{final} is the time when the fault occurs and HI^{-1} the inverse of $HI(t)$ used to get the current time (t). The estimated RUL for the tested bearing is shown in Fig. 13. From this latter figure it can be seen that the predicted RUL is less than the real RUL. This is in fact a pessimistic RUL which is the one accepted in practice by the industrials. Indeed, in the case of fault prognostics, it is better to predict the fault before its real occurrence than predicting it after
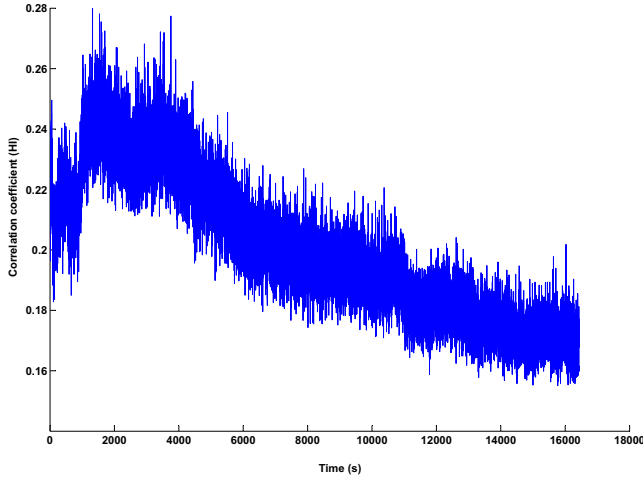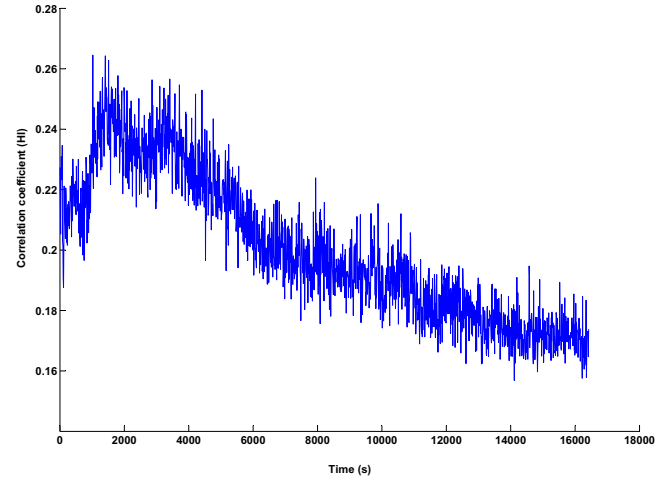

Fig. 9. Progression of the health indicator.

calculated health indicator for RUL estimation, it is necessary to fit it to a mathematical model. This fitting is done in this paper by using a non-linear regression (polynomial, exponential, etc.). Before fitting, the health indicator is first re-sampled and then smoothed to avoid over-fittings. The re-sampling process consists of reducing the number of recorded points (one point over ten is kept). The smoothing process is done by implementing a simple moving average over the re-sampled health indicators. The results of re-sampling and smoothing are shown in Fig. 10 and Fig. 11.

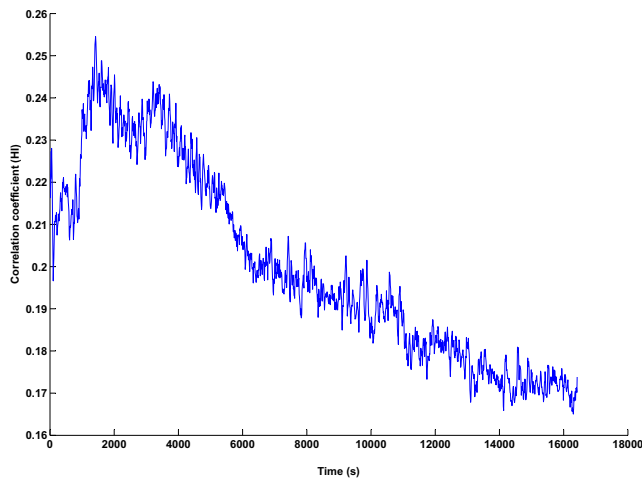The remaining useful life of the degraded bearing is the
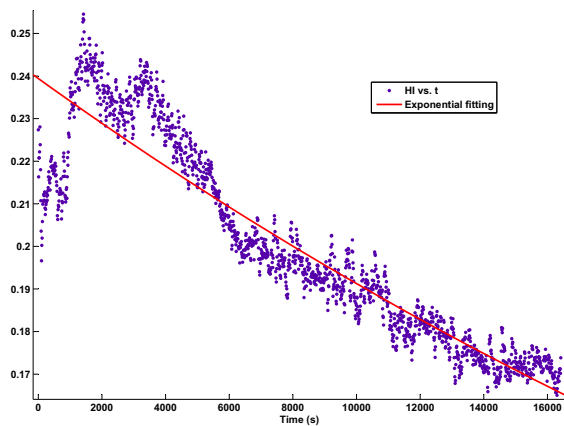
Fig. 11. Smoothing of the health indicator.

Fig. 12. Exponential fitting of the smoothed health indicator.

its occurrence. A pessimistic RUL prediction allows the operator to better prepare his/her maintenance interventions whereas an optimistic RUL prediction may lead to negative consequences. Note that in real applications the online RUL estimation is done by extrapolating the health indicator built from the data acquired until the time (t).

V. CONCLUSION

A prognostic method for remaining useful life estimation is proposed in this paper. The method is within the data-driven framework and is based on the calculation of a health indicator which permits to continuously assess the health status of the physical component. The health indicator corresponds to the difference between the signals of a nominal component (without faults) and the signals of a same type of component with a degradation. It is calculated in this contribution by using a correlation between the nominal and the degraded signals. The method is applied on vibrations signals acquired from the experimental platform PRONOSTIA. The obtained results show the effectiveness of the method for RUL estimation.

Further works may concern the calculation of the health indicator by using other techniques (such as the distance between
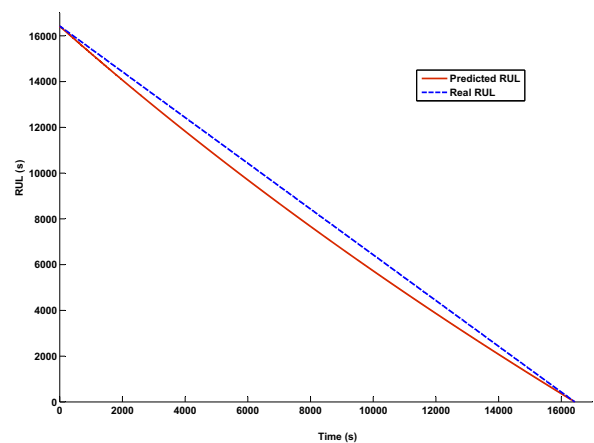

Fig. 13. The estimated RUL for the tested bearing.

the nominal and the degraded clusters) and the application of the method to other types of physical components (gears, belts, electronic components, etc.).

REFERENCES

[1] K. Medjaher, D. A. Tobon-Mejia, and N. Zerhouni, "Remaining useful life estimation of critical components with application to bearings," IEEE *Transactions on Reliability*, vol. 61, no. 2, pp. 292 – 302, 2012.

[2] D. A. Tobon-Mejia, K. Medjaher, and N. Zerhouni, "CNC machine tool's wear diagnostic and prognostic by using dynamic bayesian networks," *Mechanical Systems and Signal Processing*, vol. 28, pp. 167 – 182, 2012.

[3] 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.

[4] 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.

[5] V. Venkatasubramanian, "Prognostic and diagnostic monitoring of complex systems for product lifecycle management: Challenges and opportunities," *Computers & Chemical Engineering*, vol. 29, no. 6, pp. 1253 – 1263, 2005.

[6] G. Vachtsevanos, F. L. Lewis, M. Roemer, A. Hess, and B. Wu, *Intelligent fault diagnosis and prognosis for engineering systems.* Wiley, 2006.

[7] D. Tobon-Mejia, K. Medjaher, N. Zerhouni, and G. Tripot, "A data-driven failure prognostics method based on mixture of gaussians hidden markov models," *IEEE Transactions on Reliability*, vol. 61, no. 2, pp. 491 – 503, 2012.

[8] P. Nectoux, R. Gouriveau, K. Medjaher, E. Ramasso, B. Morello, N. Zerhouni, and C. Varnier, "Pronostia: An experimental platform for bearings accelerated life test," in IEEE *International Conference on Prognostics and Health Management*, Denver, CO, USA, 2012.

[9] AFNOR, "Condition monitoring and diagnostics of machines - prognostics - part 1: General guidelines. NF ISO 13381-1," 2005.

[10] J. Luo, K. R. Pattipati, L. Qiao, and S. Chigusa, "Model-based prognostic techniques applied to a suspension system," *Transactions on Systems, Man, and Cybernetics*, vol. 38, pp. 1156–1168, 2003.

[11] D. Chelidze and J. Cusumano, "A dynamical systems approach to failure prognosis," *Journal of Vibration and Acoustics*, vol. 126, pp. 2 – 8, 2004.

[12] A. Heng, A. C. Tan, J. Mathew, N. Montgomery, D. Banjevic, and A. K. Jardine, "Intelligent condition-based prediction of machinery reliability," *Mechanical Systems and Signal Processing*, vol. 23, no. 5, pp. 1600 – 1614, 2009.

[13] M. Dong and D. He, "A segmental hidden semi-markov model (hsmm)-based diagnostics and prognostics framework and methodology," *Mechanical Systems and Signal Processing*, vol. 21, pp. 2248–2266, 2007.