Feature Extraction and Evaluation for Health Assessment and Failure Prognostics

K. Medjaher[1], F. Camci[2], and N. Zerhouni[1]

[1] *FEMTO-ST Institute, AS2M Department, UMR CNRS 6174-UFC/ENSMM/UTBM, 25000 Besançon, France*
kamal.medjaher@esn2m.fr

[2] *IVHM Centre School of Applied Sciences Cranfield University, UK*
f.camci@cranfield.ac.uk

ABSTRACT

Abstract - The estimation of Remaining Useful Life (RUL) of industrial equipments can be realized on their most critical components. Based on this assumption, the identified critical component must be monitored to track its health state during its operation. Then, the acquired data are processed to extract relevant features, which are used for RUL estimation.
This paper presents an evaluation method for the goodness of the features, extracted from raw monitoring signals, for health assessment and prognostics of critical industrial components. The evaluation method is applied to several simulated datasets as well as features obtained from a particular application on bearings.

1. INTRODUCTION

The availability, reliability and security of industrial equipments can be ensured by monitoring their most critical components to continuously assess their health condition and predict their future one leading to maintenance, life cycle and cost optimization. Examples of critical physical components can be bearings, gears, batteries, belts, etc. Bearings failure is considered as the one of the foremost causes of breakdown in rotating machinery (Li et al., 1999). Bearing faults account for the 40% of motor faults according to the research conducted by Electric Power Research Institute (EPRI) (Enzo & Ngan, 2010). Turbine engine bearing failures are the leading cause of class-A mechanical failures (loss of aircraft) (Richard, 2005). Even one aircraft saved with prognostics would pay its development cost (Marble & Morton, 2006). The identification of the most convenient time of maintenance after failure detection without reducing the safety requirements is crucial, which is possible with prognostics capability. Thus, bearing prognostics is very critical for effective operation and management.

Failure detection forces machinery to shut down that causes tremendous time, productivity and capital loss. In addition, it is not uncommon to replace a defected/used bearing with a new one that has shorter remaining useful life than the defected one. Each failure type (outer race, inner race, ball and cage defects) causes a distinct signature in the vibration frequency (Enzo & Ngan, 2010) and vibration analysis is considered as the most reliable method in bearing failure detection (Zhang, Sconyers, Patrick, & Vachtsevanos, 2010; Davaney & Eren, 2004; McFadden & Smith, 1984; Tandon & Choudhury, 1999). However, it is often difficult to extract the failure signature due to the noise in the data especially in early stages of the failure (Su, Wang, Zhu, Zhang, & Guo, 2010; Bozchalooi & Liang, 2008; He, Jiang, & Feng, 2009). These features are then used to do failure detection, diagnostic and prognostic.

Feature extraction is the common step in all types of prognostic approaches and one of the most critical steps in diagnostics and prognostics. The extracted features are first evaluated and then used by appropriate methods and algorithms to detect the faults and to predict the equipment's remaining useful life. In this framework, the goodness of the features affects the complexity of the diagnostic and prognostic methods. Features that represent healthy, close to failure machinery and their progression perfectly may lead to very simple diagnostic and prognostic methods. On the other hand, very complex diagnostic and prognostic methods using features that are ineffective in representation of failure and failure progression may lead to poor results. Thus, extraction of relevant features is a pre-requisite for effective diagnostics and prognostics.

This paper presents an evaluation method for the goodness of the features for prognostics. An effective feature evaluation method will achieve the selection of best features, which is critical for obtaining better prognostics results. The feature evaluation method is applied to bearings that were run until failure in a lab environment. The paper is organized as follows: section 2 presents a brief introduction to failure prognostic, section 3 deals with the quantification metric for the quality evaluation of features for prognostics, section 4 presents results and experiments and finally section 5 concludes the paper.

2. Failure prognostic paradigm

According to the International Standard Organization (ISO), failure prognostics corresponds to the estimation of the operating time before failure and the risk of future existence or appearance of one or several failure modes (AFNOR, 2005). In the scientific literature, the operating time before failure is called remaining useful life (RUL) for which a confidence value is often associated.
Several methods and tools for performing failure prognostics are proposed in the literature. This material can be grouped into three main approaches (Tobon-Mejia, Medjaher, & Zerhouni, 2012; Heng, Zhang, Tan, & Mathew, 2009; Jardine, Lin, & Banjevic, 2006; Vachtsevanos, Lewis, Roemer, Hess, & Wu, 2006), namely: model-based approach, data-driven approach and hybrid approach.

Model-based (also called physics of failure) methods deal

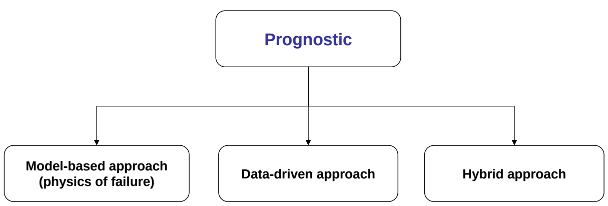

Figure 1. Main prognostic approaches.

with the exploitation of a mathematical model representing the behavior of the physical component including its degradation. The derived model is then used to predict the future evolution of the degradation. In this case, the prognostic 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 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 are concerned with 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. Generally, the raw data are first processed to extract features which are then used to build the diagnostic and prognostic models. The features can be temporal, frequency or both. In same 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 prognostics methods can use data provided by sensors or obtained through experience feedback (operation, maintenance, number of breakdowns, etc.).
The advantage of data-driven approach is its applicability, cost and implementation. Indeed, by 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.
Hybrid methods use both data-driven and model-based (or physics of failure) approaches. The application of each approach depends on the application and on the type of knowledge and data available.

3. Feature Extraction and evaluation

Fault detection, diagnostic and prognostic activities all use the notion of features, which are extracted from the raw monitoring signals provided by the sensors (temperature, vibration, force, etc.) installed on the system. Feature extraction is primordial in the process of health monitoring, health assessment and failure prognostic. Indeed, the relevant information which is related to the behavior of the component during its degradation is often hidden in the raw signals and needs to be extracted by means of appropriate methods. The figure 2 shows the steps involved in the process failure prognostic including feature extraction.

Diagnostics is a classification problem, whereas the prog-

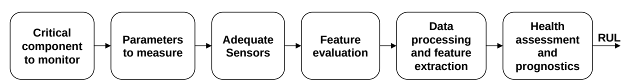

Figure 2. Steps for RUL estimation.

nostics is the process of forecasting the future health states. The goodness of the features for diagnostics is basically a measure of separability between data from healthy and faulty equipment. Good separability indicates that samples from different classes (i.e., healthy and faulty) are far apart from each other and samples from the same class are close to each other. The key point in prognostics is the continuity of the separation between time segments, whereas diagnostics focus on one separability measure between two static classes (i.e., failed and healthy). However, prognostics searches for separation between time segments for the whole the component where prognostics is aimed. Within class separability (parameters a and b in Fig 3 (Camci, Medjaher, Zerhouni, & Nectoux, 2012)) and between class separability (parameter c in Fig 3 (Camci et al., 2012)) are used to quantify the separability. Many class separation metrics have been reported in the literature (Calinski & Harabasz, 1974; Eker et al., 2011). These metrics focus on static classes; do not consider pro-

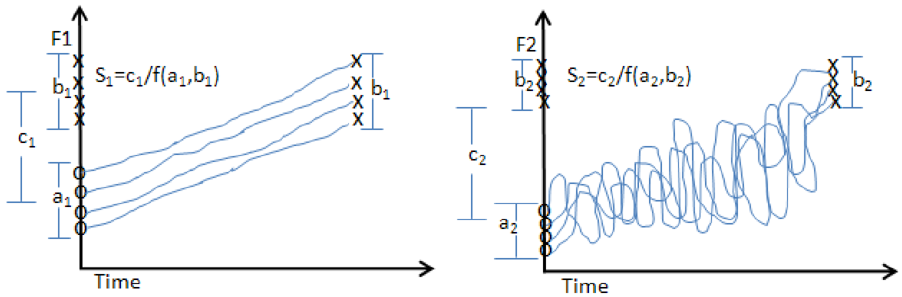

Figure 3. Feature quality for diagnostics and prognostics.

gression from one class to another. One feature may be good at separation of the classes, but not at representation of progression from one class to another. For example, separability measure (S2) of feature 2 (F2) is higher than in separability measure (S1) of feature 1 (F1) in Fig 3 (Camci et al., 2012). However, this does not mean that F2 is better in representing the failure progression. As seen from the figure, failure progression in F2 involves higher variation. Thus, a new quality measure should be employed for prognostics, which is a relatively new problem.

Monotonically non-increasing or non-decreasing: Mathematically, a function f is called monotonically increasing (monotonically non-decreasing), if for all x and y such that $x \leq y$ one has $f(x) \leq f(y)$ (f (y) $\leq f$ (x)).

It may be trivial to check the monotonicity for a single failure progression sample by analyzing the difference between consecutive points. When all the difference values are greater (less) than or equal to 0, then the function is defined as non-decreasing (non-increasing). However, monotonicity over all samples representing failure progression should be considered rather than individual analysis of samples. Example of several samples representing failure progression is displayed in Fig. 4 (Camci et al., 2012). As seen from the figure, the time is segmented for effective analysis of the failure progression. The effectiveness of a feature to represent the failure progression is calculated as the average separability of segments as represented in (1). The higher the total separability value (S) is, the better representation of the failure progression. Thus, the goal is to find the feature that has the highest S value. S basically is the average separation between time segments. High S value indicates that the difference between time segments are high. s_t value is the separability measure for consecutive time segments.

$$S = \frac{\sum_{t=1}^{T} s_t}{T} \tag{1}$$

where S is the average separability value, s_t is the separability at time t and T is the total number of time segments.

The distribution of the data points from different samples in each time segment should be used to measure the separability at a given time segment. The separability calculation is formulated in (2).

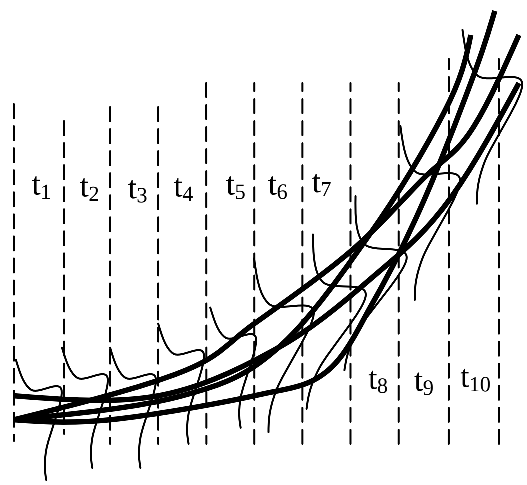

Figure 4. Failure progression for multiple samples.

$$s_t = \frac{a}{L} - \frac{\chi}{N_t} \tag{2}$$

with

$$\chi = \begin{cases} 0 & \text{if } \frac{a}{L} \neq 1 \\ \alpha & \text{if } \frac{a}{L} = 1 \end{cases} \tag{3}$$

where α is the number of samples overlapping with the distribution in consecutive time frame, N_t is the number of samples in time segment t and L represents the distance between 25th and 75th percentiles. The 25th and 75th percentiles were selected as a common sense to select the range to be able to capture the 50% of the data. The selection of the range may depend on signal to noise ratio and possible bias in the dataset.

The ratio of the length of the non-overlapped portion (called a) to L is a measure of the separability (a/L). The L and a parameters represent the distance between points in the given percentile. For example, if the overlapping occurs between 30th and 50th percentile, parameter a is the distance between samples in 30th and 50th percentile. When the separation is low, a/L ratio will be close to 0. When the separation is high, a/L becomes closer to 1. When there is no overlap between 25-75 percentiles of the distributions (a/L=1), there exist two different possibilities. In the first one, there is some overlap within data greater than 75th percentile or less than 25th percentile. The second one represents complete separation. When a/L becomes 1, then the ratio of number of data points causing overlap to the total number of data points in the distribution is subtracted in separability calculation.

4. Experiments and Results

4.1. Simulated Dataset

The presented evaluation method is applied to eight simulated datasets. These datasets have been developed to simulate various levels goodness for prognostics. The features with clear trend are considered to be good feature, whereas bad features do not include a trend with time. The datasets numbered from one to eight include increasing trend as shown in Figure 5.

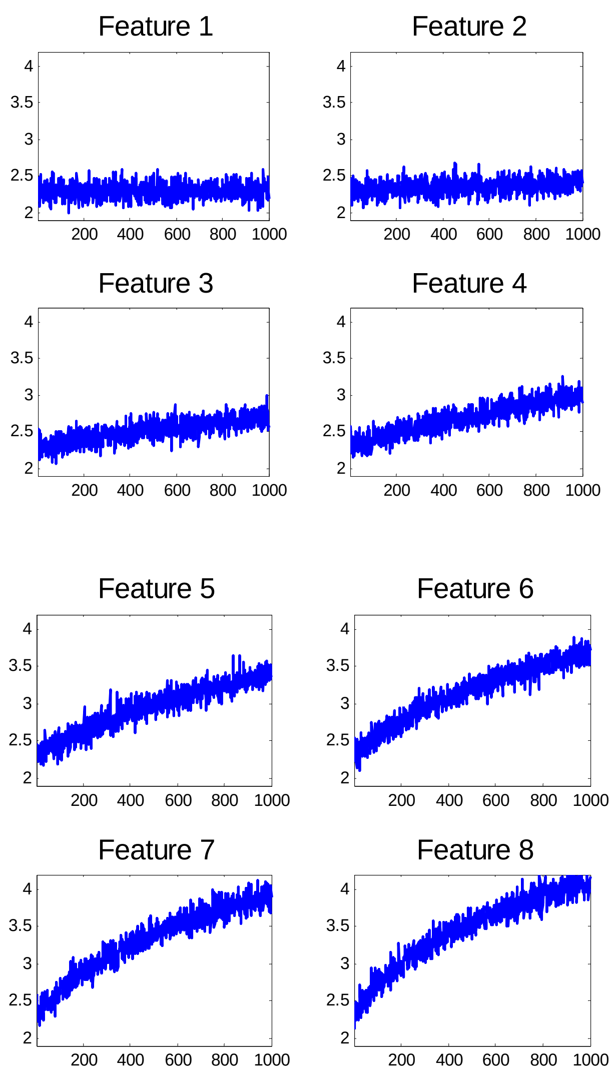

Figure 5. Simulated Features.

The trend in these datasets are formulated as logarithmically increasing mean with constant noise and shown in the formulation below. In these equations $\mu_{i,t}$ is the mean of feature i in time t and T is the final time point.

$$x(t) = \mu_{i,t} + \sigma \tag{4}$$

$$\mu_{i,1} = \log(10) \tag{5}$$

$$\mu_{i,T} = \log(i \times 10) \tag{6}$$

As seen from the Figure 5, the goodness of the features increases from feature 1 to feature 8. The trend in the later features can be seen better in later datasets. Figure 6 displays the goodness of features obtained with the presented evaluation metric. As seen from the figure, the goodness increases in the later features, which supports the increasing trend in Figure 5.

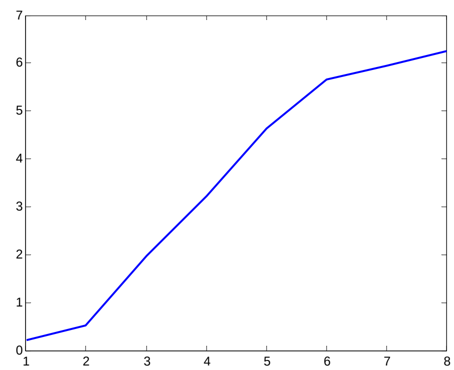

Figure 6. Goodness of features.

4.2. Bearing Example

The accelerated bearing life test bed is called PRONOSTIA, 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.

A pair of ball bearings is mounted on one end of the shaft to serve as the guide bearings and a NSK6307DU roller ball bearing is mounted on the other end to serve as the test bearings. The transmission of the movement between the motor and the shaft drive is coupled by a rub belt.

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 temperature probe (of type PT100) to record the temperature of the tested bearing. A speed sensor and a torque sensor are also available on the PRONOSTIA platform. The sampling frequency of the NI DAQCard-9174 data acquisition card is set to 25600 Hz and the vibration data provided by the two accelerometers are collected every 1 second.

The bearing operating conditions are determined by instantaneous 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.

Several features are extracted to be used for failure progression such as maximum, mean, standard deviation, skewness, kurtosis, root mean square error (RMS), crest factor and highest frequency.

Fig 7 displays two good (RMS and standard deviation); two bad features (Skewness and crest factor) for prognostics (in these plots, the x axis stands for time). As you can see from the figures, failure progression can be seen in the features with high separability measure.

Fig 8 displays the separability values of several features. In

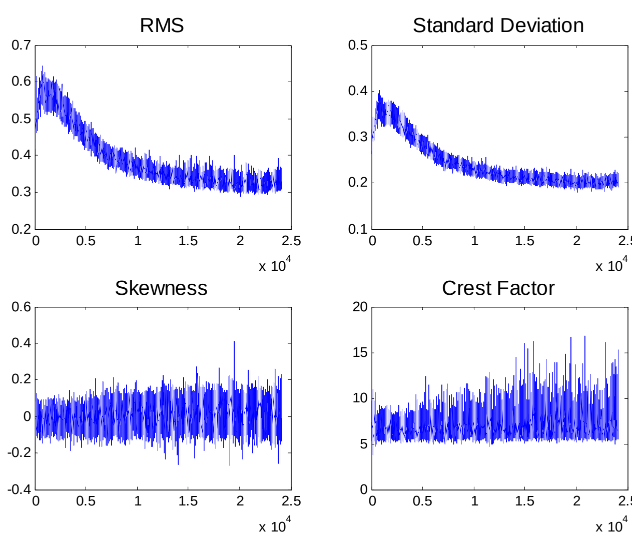

Figure 7. Examples of good/bad features for prognostics.

this figure three sensory signals were used each is represented by a line in the graph. The fluctuations show that the goodness may vary based on the sensory signal used. As seen from this figure, the goodness of skewness and crest factor (CF) are low, whereas the goodness of standard deviation and RMS are high. Thus, the evaluation method is able to differentiate the goodness of the features.

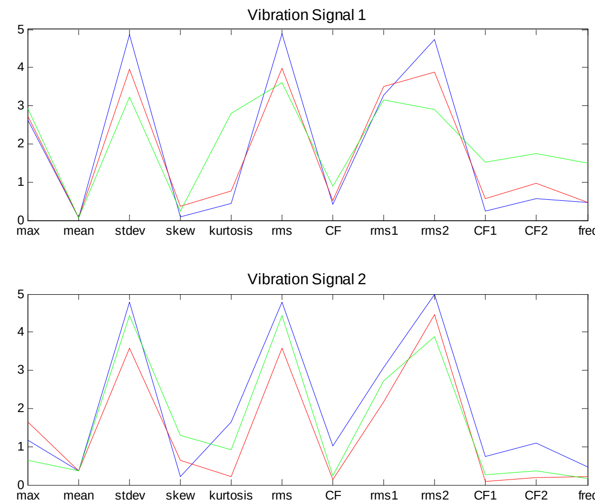

Figure 8. Separability values for the second type of degradation.

5. Conclusion

The quality of the features is critical for health assessment, diagnostics and prognostics. Feature extraction, selection and evaluation of the quality of features in diagnostics has been studied extensively. The nature of the prognostics problem is different from diagnostics. This paper presents quantification metric for evaluation of the quality of features for prognostics, which is relatively new problem compared to diagnostics. The presented metric is applied to features extracted from bearing vibration data collected in a lab environment. The features are plotted for visual evaluation to judge the quality of the evaluation metric. The results show that the metric is able to effectively quantify the quality of features for the purpose of prognostics.

References

AFNOR. (2005). *Condition monitoring and diagnostics of machines - Prognostics - Part 1: General guidelines.* NF ISO 13381-1.

Bozchalooi, I., & Liang, M. (2008). A joint resonance frequency estimation and in-band noise reduction method for enhancing the detectability of bearing fault signals. *Mechanical Systems and Signal Processing*, *22*, 915-933.

Calinski, R., & Harabasz, J. (1974). A Dendrite Method for Cluster Analysis. *Comm. in Statistics*, *3*, 1-27.

Camci, F., Medjaher, K., Zerhouni, N., & Nectoux, P. (2012). Feature Evaluation for Effective Bearing Prognostics. *Quality and Reliability Engineering International*. (In press)

Davaney, M., & Eren, L. (2004). Detecting Motor Bearing Faults. *IEEE Instrumentation and Measurement Magazine*, 30-50.

Eker, O., Camci, F., Guclu, A., Yilboga, H., Sevkli, M., & Baskan, S. (2011). A Simple State- based Prognostic Model for Railway Turnout Systems. *IEEE Transactions on Industrial Electronics*, *58*(5), 1718-1726.

Enzo, C. L., & Ngan, H. W. (2010). Detection of Motor Bearing Outer Raceway Defect by Wavelet Packet Transformed Motor Current Signature Analysis. *IEEE Transactions on Instruments and Measurement*, *59*(10), 2683-2690.

He, W., Jiang, Z., & Feng, K. (2009). Bearing fault detection based on optimal wavelet filter and sparse code shrinkage. *Measurement*, *42*, 1092-1102.

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

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

Li, Y., Billington, S., Zhang, C., Kurfess, T., Danyluk, S., & Liang, S. (1999). Adaptive Prognostics For Rolling Element Bearing Condition. *Mechanical Systems and Signal Processing*, *13*(1), 103-113.

Marble, S., & Morton, B. (2006). Predicting the Remaining Useful Life of Propulsion System Bearings. In *Proceedings of the 2006 IEEE Aerospace Conference.* Big Sky, MT, USA.

McFadden, P., & Smith, J. (1984). Vibration monitoring of rolling element bearings by the high frequency resonance technique - a review. *Tribology International*, *17*, 3-10.

Richard, A. W. (2005). A Need-focused Approach to Air Force Engine Health Management Research. In *Health Management Research IEEE Aerospace Conference.* Big Sky, Montana, USA.

Su, W., Wang, F., Zhu, H., Zhang, Z., & Guo, Z. (2010). Rolling element bearing faults diagnosis based on optimal Morlet Wavelet filter and autocorrelation enhancement. *Mechanical Systems and Signal Processing*, *24*, 1458-1472.

Tandon, N., & Choudhury, A. (1999). A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings. *Tribology International*, *32*, 469-480.

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

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

Zhang, B., Sconyers, C., Patrick, M. O. andR., & Vachtsevanos, G. (2010). Fault Progression Modeling: An Application to Bearing Diagnosis and Prognosis. In *Proceedings of American Control Conference.* MD USA.