Sensor-Based Degradation Prediction and Prognostics for Remaining Useful Life Estimation:
Validation on Experimental Data of Electric Motors

Federico Barbieri[1], J. Wesley Hines[2], Michael Sharp[2], Mauro Venturini[1]

[1]*Dipartimento di Ingegneria, Università degli Studi di Ferrara, Ferrara, Italy*
federico02.barbieri@student.unife.it
mauro.venturini@unife.it

[2]*Department of Nuclear Engineering, The University of Tennessee Knoxville, Knoxville, USA*
jhines2@utk.edu
msharp6@utk.edu

ABSTRACT

Prognostics is an emerging science of predicting the health condition of a system and/or its components, based upon current and previous system status, with the ultimate goal of accurate prediction of the Remaining Useful Life (RUL). Based on this assumption, components/systems can be monitored to track the health state during operation. Acquired data are generally processed to extract relevant features related to the degradation condition of the component/system. Often, it is beneficial to combine several of these degradation parameters through an optimization process to develop a single parameter, called prognostic parameter, which can be trended to estimate the RUL. The approach adopted in this paper consists of a prognostic procedure which involves prognostic parameter generation and General Path Model (GPM) prediction. The Genetic Algorithm (GA) and Ordinary Least Squares (OLS) optimization methods will be used to develop suitable prognostic parameters from the selected features. Both time and frequency domain analysis will be investigated. Steady-state data obtained from electric motor accelerated degradation testing is used for method validation. Ten three-phase 5HP induction were run through temperature and humidity accelerated degradation cycles on a weekly basis. Of those, five presented similar degradation pathways due to bearing failure modes. The results show that the OLS method, on average, generated the best prognostic parameter performance using a combination of time domain features. However, the best single model performance was obtained using the GA methodology. In this case, the estimated RUL nearly coincided with the true RUL with an absolute percent error averaging under 5% near the end of life.

1. INTRODUCTION

The growing need to increase the competitiveness of industrial systems requires a reduction of maintenance costs, without compromising safe plant operation. Therefore, forecasting the future behavior of a system or component allows more optimal maintenance planning and cost savings, because unexpected repairs and downtime can be avoided. More generally, the prediction of future failures can provide key information to the decision-making process. Since business interruption costs usually prove to be significantly higher than the cost of performing the repairs required to return to service (Orme & Venturini, 2011), the maximization of machine availability is essential to promote profitable operation. On-line condition-based maintenance (CBM) is a new maintenance philosophy involving real-time analysis of equipment sensor data to infer maintenance condition or health. Maintenance activities are then performed on the basis of necessity, as identified by a condition-based maintenance system. In comparison to traditional maintenance philosophies, CBM offers the potential for minimizing instances of equipment failures, reducing scheduled maintenance activities, maximizing serviceable life of life-limited components, and increasing equipment availability. However, critical to the success of implementing a condition-based maintenance philosophy, are the necessary technical capabilities to infer equipment condition from real-time process measurements, so that informed maintenance decisions can be made. The development of the technical capabilities to implement a condition-based maintenance philosophy, including the development of predictive prognostic models, are of major interest across almost all industrial environments in which the availability, reliability and performance of machinery is

critical. However, developing such capabilities is a significant technical challenge. Some of the benefits provided by such an "on-condition" based maintenance are 1) improved plant safety; 2) less time spent on inspection; 3) better ability to plan maintenance; 4) improved fault detection; and 5) increased asset availability.

Condition-based maintenance is certainly facilitated by a complete health monitoring system. The term "health monitoring system" pertains to methods that allow the practitioner to evaluate a system's actual health/damage conditions, predict the onset of failure, and mitigate the risks associated with an abnormal system behavior. In published literature, this is traditionally considered to consist of several modules, including Monitoring, Detection, Diagnostics and Prognostics. For several decades, researchers have been investigating and developing different techniques for failure detection, isolation, and identification across a wide range of application domains in science, medicine and engineering. A comprehensive review of the techniques and methods used in fault diagnostics is beyond the scope of this work; however, some review publications are provided by Venkatasubramanian et al. (2003) and Jardine et al. (2006).

While the detection and diagnostics portions have been well established for several decades, the prognostics-related techniques have only recently attracted attention in research studies. Prognostics involves predicting the amount of time or cycles that a system or component will continue to meet its design specifications. The ultimate goal of most prognostic systems is accurate prediction of the Remaining Useful Life (RUL) of individual systems or components, on the basis of their use and performance. In general, making accurate prognostic predictions allows benefits such as advanced scheduling of maintenance activities, proactive allocation of replacement parts and enhanced fleet deployment decisions based on the estimated progression of component life consumption (Li & Nilkitsaranont, 2009; Watson et al., 2011). For these reasons, prognostic health management is now widely recognized as the direction of the future (Saxena et al., 2009a).

According to (Nam et al., 2012), methods for calculating the remaining useful life depend on the type and amount of data available for the particular component under inspection. Accordingly, Type I, II and III prognostics can be classified. At the basic level, Type I considers the time to failure of components and is determined by a best fit of the historical or estimated failure data. Among the parametric models, the Weibull distribution is commonly used to calculate the best fit, because it is flexible to different probability distribution shapes. Type II uses stress-based data and therefore provides the average life of an average component under given usage conditions. The last type, Type III, is a degradation-based analysis, which improves Type II by monitoring and analyzing the specific component's response under its specific usage. A component can be expected to fail when the degradation pathway meets a failure threshold. In Type III prognostics, the degradation parameter, which is a measure that characterize the progression to failure, can be either sensed measurements, or inferred variables. For Type III prognostics, data related to how the system is responding to the operating environment and how it is degrading is called degradation parameters and they are often fused into a prognostic parameter: a parameter that can be used as a measure of damage. This prognostic parameter is then used to construct a prognostic model.

Several papers have been recently published in this research area, with the aim of providing a comprehensive overview of the available prognostic techniques and the related application areas. Bryg et al. (2008) apply logistic regression to aircraft engine takeoff data and control system fault information to provide failure probability over time. Heng et al. (2009a) identified the merits and weaknesses of the different prognostic techniques and also discussed the identified challenges, such as the requirement of utilizing incomplete trending data, considering the effects of variable operating conditions and accounting for failure interactions. Heng et al. (2009b) developed a prognostic model which consists of a neural network, used as probability density function estimator and applied to pump vibration data. Li and Nilkitsaranont (2009) developed both a linear and a quadratic regression technique to predict the remaining useful life of gas turbine engines. Si et al. (2011) reviewed the most recent modeling developments for estimating the remaining useful life by means of statistical data driven approaches. Watson et al. (2011) conducted a model-based analysis of electro-mechanical actuators: the actuator life model and the probabilistic prognostic approach are used to determine the remaining useful life of the system.

The first stage in any health monitoring system typically involves appropriate preprocessing of equipment sensor data. This stage is often referred to as feature extraction. Feature extraction is the process of extracting useful information from raw signal data. The feature extraction stage within a health monitoring system is designed to generate a vector of data features, which can be used to infer the current fault status of a monitored system. However, as equipment degrades, measured parameters of the system tend to change; therefore, sensed measurements, or appropriate transformations thereof, may be used to characterize the system degradation. Traditionally, Type III prognostic methods use some measure of degradation to make RUL estimates. It is beneficial to combine several measures of degradation, into a single parameter, identified as prognostic parameter, to provide a more robust prognostic model. Selection of an appropriate prognostic parameter is key for making useful individual-based RUL estimates.

In this paper, a prognostic procedure for Remaining Useful Life (RUL) estimation, composed of two separate steps, is outlined. First, the extracted features are fused to produce a prognostic parameter which is designed to be correlated with RUL. This parameter is then modeled through a General Path Model (GPM) and extrapolated to a failure threshold to estimate the RUL. The Genetic Algorithm (GA) and Ordinary Least Squares (OLS) estimation methods are subsequently used to develop suitable prognostic parameters from the selected features. Finally, advanced prognostic metrics are used to accurately evaluate model performance and validate the methodology against experimental data taken on a group of three-phase motors with similar degradation pathways, mainly due to bearing failures. This paper is organized as follows: Section 2 outlines some basic concepts about condition-based maintenance and health monitoring systems; Section 3 describes the prognostic procedure used in this paper for remaining useful life estimation; Section 4 reports the metrics used for evaluating prognostic methodology reliability; Section 5 presents the experimental data used for validating the prognostic methodology; Section 6 presents the results and discusses the capability of the prognostic methodology; and Section 7 summarizes the most significant conclusions.

2. HEALTH MONITORING FOR CONDITION BASED MAINTENANCE

Traditionally, maintenance activities have taken one of two approaches: preventive and corrective. Between these two extremes lies condition based maintenance, wherein maintenance actions are performed as needed based on the condition of the equipment. It is important to accurately trend the effect of a fault on system performance through a correct and appropriate typical health monitoring system. Figure 1 shows a diagram of a typical health monitoring system.
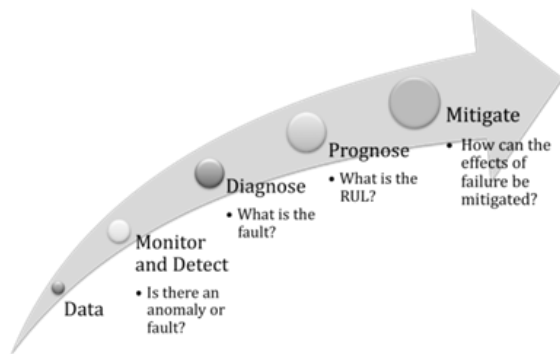

Figure 1. Health Monitoring System (Hines, 2009).

At first, data collected from a system of interest is monitored for deviations from normal behavior. Monitoring can be accomplished by several methods, such as first principle models, empirical models, statistical analysis and this module can be considered an error correction routine

(Hines et al., 2006a; Loboda & Feldshteiyn, 2010; Palmè et al., 2009). An error correction routine means that the model gives its best estimate of the true value of the system variables under unfaulted conditions and these estimates are compared to the data collected from the system to generate a time series of residuals. Residuals characterize system deviations from normal behavior and can be used to determine if the system is operating in an abnormal state. Finally, a prognostic model is employed to estimate the Remaining Useful Life (RUL) of the system.

As done in this research, it is possible to utilize accelerated degradation testing data, collected during increased stress conditions, for prognostics model development. However, care must be given to ensure that the failures seen during accelerated testing are analogous to real world failures. Accelerated testing conditions can result in fault modes which only occur under the accelerated conditions. Several methods are available for extracting useful data which describes actual operation from accelerated testing data (Carey & Koenig, 1991; Elsayed & Chen, 1998; Park & Padgett, 2006; Tang & Chang, 1995).

2.1. System Reliability and RUL Estimation

Accurate, real time, prognostic models certainly represent the holy grail of reliability engineering. Prognostics is the last stage, since it utilizes all available information and contributes to the system reliability information. However, prognostics is a fairly immature field compared to the more established areas of condition monitoring, fault detection, and diagnostics. Publications concerning prognostics have focused on the need for prognostics and the challenges in prognostic model development (Greitzer et al., 1999; Hess et al., 2005), the many and varied applications of prognostics (Ferrell, 2000; Kalgren et al., 2007; Keller et al., 2006; Orchard & Vachtsevanos, 2007; Puggina & Venturini, 2012; Roemer et al., 2005), etc.. A key aspect is the identification of the nature of RUL. In general, at any time before failure, the RUL is given by the time between the current time and the failure time as shown in Figure 2 for a component that fails at 100 cycles.
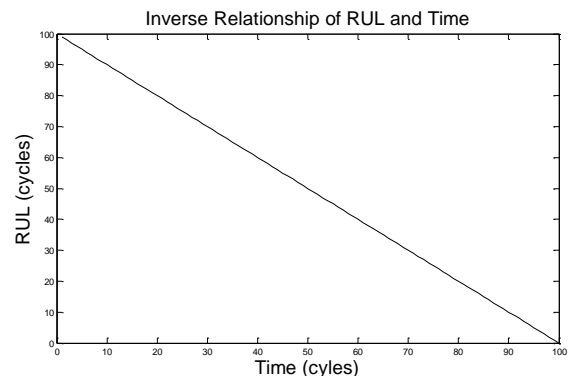

Figure 2. RUL vs Time (cycles) (Coble, 2010).

A significant portion of the published literature in prognostics and RUL estimation research focuses on solutions to specific problems, such as electronic prognostics (Mishra & Pecht, 2002; Vichare & Pecht, 2006), vibration analysis (Carden & Fanning, 2004), helicopter gearbox monitoring (Kacprzynski et al.,2004; Vachtsevanos et al., 1997; Wang & Vachtsevanos, 2001), JSF applications (Ferrell, 1999; Roemer et al., 2005), turbines prognostics (Cavarzere & Venturini, 2012; Li & Nilkitsaranont, 2009; Lipowsky et al., 2010; Venturini & Puggina, 2012; Venturini & Therkorn, 2013), etc..

The goal in developing generic prognostic algorithms is to develop methods which may be rapidly configured for a new system to allow for effective and efficient deployment of CBM technology on complex systems; because, while a specific approach may result in very good point solutions for specific problems, generic prognostic algorithms which may be more broadly applicable are clearly of higher interest.

2.2. Type III: Effects Based Prognostics

Type III prognostic algorithms attempt to characterize the lifetime of the specific system operating in its specific environment. These methods are able to utilize a myriad of information related to unit degradation and failure. A prognostic path is a trajectory along which the prognostics parameter is evolving in time towards the critical level corresponding to a failure event. The prognostic model can then be used to predict the RUL of that unit. Type III prognostic model development must also take into consideration the method by which degradation is accumulated considering a cumulative damage model (Ramakrishnan & Pecht, 2003), which assumes that all damage incurred remains until some external source actively repairs the system. From this assumption, it follows that the prognostics parameter cannot spontaneously move towards a less degraded state; i.e. systems do not self-heal, and any indication of such is due strictly to measurement error. In other words, all damage incurred by a unit is cumulative and builds toward a threshold beyond which the unit will no longer meet its design specifications to some prescribed confidence. Beyond this assumption of no self-healing, another common assumption is that of a common failure threshold. One of the main difficulties associated with Type III prognostics is obtaining the most representative prognostics parameter for the system. Often, obtaining this parameter is neither simple nor direct because rarely a prognostics parameter can be explicitly measured. More often the degradation incurred by a system must be inferred from one or more monitored system parameters. There are several mathematical approaches to model cumulative damage such as Markov Chain-based Models, Shock Models, General Path Models. The most common method is the General Path model (GPM), which is described in detail in Section 3 and has been utilized in this research for RUL estimation. Basically, the GPM attempts to track a measure of degradation called prognostic parameter and extrapolates it to failure (Byington et al., 2004; Hines et al., 2006b; Luo et al. 2003).

3. PROGNOSTIC PROCEDURE FOR REMAINING USEFUL LIFE ESTIMATION

The RUL estimation is the output generated by a prognostic algorithm. This estimation is not often directly inferred from raw data, but several processing steps have to be applied to extract useful information from the gathered data. Sensors provide measurements which are processed and features related to degradation are extracted. These extracted features are fused to produce a single prognostic parameter which is designed to be correlated with RUL. The rationale behind the fusion of the features is to combine several information sources into one parameter that has improved robustness and an underlying trend related to overall condition. This parameter is then modeled, usually through a General Path Model, and extrapolated to a failure threshold to estimate the RUL.

In this paper, 5 HP electrical motors are degraded through several cycles of an accelerated degradation process of heating and quenching. After each cycle, operational data is collected. From this data, time and frequency domain features are extracted which are fused to create a prognostic parameter, which is modeled and trended to failure. This section provides a general overview of the methods used to develop the prognostic model which is used to calculate RUL estimates.

3.1. Signal Processing for Feature Extraction

The first stage in any Prognostic Health Management system typically involves appropriate preprocessing of equipment sensor data. This is termed the Data Manipulation phase in the OSA-CBM architecture (Sreenuch, 2013). The feature extraction stage within a Prognostic Health Management system is designed to generate a group of data features, which can be used to infer the current health status of a monitored system. Gathered signals from machine components generally contain large samples of data which require a large amount of memory and computation time to be analyzed. Instead, this data can be reduced into a lower but informative representation by extracting meaningful features from raw signals, since not all of them provide useful information for RUL estimation. Sometimes, if the trend is not clear and defined over the lifetime, the difficulty of analysis can increase and degrade the accuracy. Therefore reducing the dimension of data features by selecting the "best features" is necessary to remove the irrelevant and erroneous features.

3.1.1. Feature Extraction from Time Domain Analysis

One simple method of identifying several useful features is calculating and tracking the evolution of some simple statistical moments applied to data, and examining any trend throughout the entire lifetime of the component (Gu et al., 2013; Sharp, 2012). The most common statistical moments used in practice are Mean, Standard Deviation, Root Mean Square (RMS), Skewness and Kurtosis. The time-domain approach alone is often not capable of identifying sufficient features to clarify the health status of the system, especially for rotating equipment. For this reason, frequency-domain techniques are used to overcome the shortcomings of time-domain analysis because they can easily identify and isolate frequency components and trends.

3.1.2. Feature Extraction from Frequency Domain Analysis

Frequency domain analysis is typically used to extract features for rotating equipment that exhibit a marked difference between baseline and faulty data. Frequency domain analysis is based on transforming the time series signals into the frequency domain. The main advantage of frequency domain analysis over time domain analysis is its ability to identify and isolate the amplitude of certain frequency components of interest. Features regarding frequency information can generally indicate machinery faults better than time domain features, especially in the case of vibration signals, because characteristic frequency components such as resonance frequency components or defect frequency components can be relatively easily detected and matched to faults (Dyer & Stewart, 1978; Li et al., 2012; McInerny & Dai, 2003; Poyhonen et al., 2004; Tandon & Choudhury, 1999; Yang et al., 2003; Zhang et al., 2005). A conventional frequency-domain technique is spectrum analysis by using Fast Fourier Transform (FFT).

3.2. Prognostic Parameter Generation

Several methods can be adopted to develop suitable prognostic parameters from the obtained and selected features and they will be examined in the next sections. One of the most important key to correctly develop and apply a prognostic model to a system is selecting and identifying an appropriate prognostic parameter. This is sometimes based on engineering judgment, expert analysis, and visual inspection; but an optimization routine is also advisable to identify the most representative parameter for modeling system degradation. The useful characteristics of a good prognostic parameter are substantially three (Coble & Hines, 2011; Coble & Hines, 2012): monotonicity, prognosability and trendability.

3.2.1. Genetic Algorithm Approach for Prognostic Parameter Generation

One of the most common stochastic optimization methods and also one of the methodologies used in this research is the Genetic Algorithms (GA). The goal is to combine several features into an optimal prognostic parameter that can be easily modeled with a chosen function and which failure occurs at near the same threshold. The GA does this by optimizing a combination of three performance indices: monotonicity, prognosability, and trendability. The GA has a unique ability to evaluate many combinations of features in a randomly generated iterative process that mimics natural selection. A theoretical review is described in detail in (Coble & Hines, 2012). The PEP "Process and Equipment Prognostics" toolbox developed at the University of Tennessee uses genetic algorithms that are contained in the MATLAB Global Optimization Toolbox® to generate a near-ideal prognostic parameter from the previously selected features.

3.2.2. Prognostic Parameter Generation Using Ordinary Least Squares (OLS) Estimation

Another method to estimate the prognostic parameter is Ordinary Least Squares (OLS) estimation (Welz et al., 2014). It is conceptually simple and computationally straightforward, and for this reason it has been adopted in several engineering fields, and was also used in this research to generate a prognostic parameter from the features selected in the previous stage.

As in a bivariate linear regression model, in a simulation model obtained by using the multivariate least squared residual approach, the sum of squared residuals is minimized in order to identify the set of estimators or weights. Considering this method to obtain a prognostic parameter we can define X to be a matrix of features. Features are collected into a single matrix by concatenating each test case. This creates an n x s matrix, X, where n is the number of total data points in all test cases, and s is the number of features in the model. This X matrix is regressed against the n x 1 vector y where each y_i (i=1,2,…,n) corresponds to the percent of the total unit life at that observation. This means that the features of each test case are fitted to a linear curve from 0 to 1. The linear weights are then obtained as in Eq. (1):

$$w=(X^T X)^{-1} X^T y \tag{1}$$

where w is an s x 1 vector.

This solution involves the inversion of what is commonly called the Hessian matrix ($X^T X$). If the predictors (X) are highly correlated, this matrix is ill-conditioned and slight changes in the predictor data cause significant changes in the solution weights. A measure of condition is the ratio of the largest eigenvalue to the smallest eigenvalue. The

numerical inversion of ill-conditioned matrices causes unstable solutions, i.e. the problem is ill-posed. The difficulty of constructing prediction models with correlated data is not related to linear regression models but also occurs, with even greater instabilities, in non-linear techniques such as neural networks. Regularization methods such as ridge regression and ICOMP have been proved to provide repeatable, low noise OLS solutions for monitoring applications (Gribok et al., 2002).

3.3. Prognostic Model: General Path Model (GPM)

This section focuses on the most common Type III method for RUL estimation, i.e. the extrapolation of a general path model (GPM). The GPM fits past prognostic parameter trends with a functional form and uses that form to extrapolate from new data to a failure threshold. Other methods, such as particle filter methods, have proven to be capable prognostic models; however, the GPM was chosen in this research. The GPM employed in this work is a formulation of the model proposed by Lu and Meeker (1993), where a complete review of this methodology can be found. Since data contains information useful to GPM forecasting, it is not necessary that all historical units are run to failure. Indeed GPM uses degradation patterns instead of failure times. One of the most important assumptions of the GPM is that there are some defined critical levels of degradation, beyond which the component is considered as failed and therefore no longer meets its design specification. In order to quantify this level, some components should be run to failure, or otherwise engineering judgment might be used if the nature of the degradation is clearly known.

A natural extension of GPM reliability methodology can be used in order to estimate Remaining Useful Life (RUL) of an individual component or system. Upadhyaya, et al. (1994) proposed a type of degradation extrapolation, where the component's time of failure estimation is made by extrapolating the degradation path model to the failure threshold. In (Upadhyaya, et al., 1994), the authors used both neural networks and nonlinear regression models to predict the RUL of a small induction motor. A source of both difficulty and uncertainty is the definition of this failure threshold because systems rarely have a hard failure threshold which holds for each unit. Often there is an associated failure distribution which must be taken into account when estimating RUL and its uncertainty (Usynin et al., 2008).

As discussed in (Carlin & Louis, 2000), a common method for integrating prior population-based historical data with current individual data is Bayesian updating. As described in Figure 3, historical data is used to estimate the model parameter. As new data are collected, they are used to update the model fit resulting in a new posterior distribution of the model parameters. This posterior is then used as the new prior distribution for further updates. When new data

are collected, they are used to update the parameter distribution again.

In this paper, Bayesian methods are used to include prior information for linear regression problems. For a complete discussion of Bayesian statistics including other Bayesian update methods, the reader is referred to (Carlin & Louis, 2000; Gelman et al., 2004; Lindely & Smith, 1972).

4. METRICS FOR EVALUATING PROGNOSTIC PREDICTIONS

In general, performance metrics address the issue of how well the RUL prediction estimates improve over time as more measurement data become available. Prognostic predictions inherently incorporate temporal aspects of the system. Many methods for quantifying prognostic performance have been developed (Saxena 2010). One of the variants that can be used for capturing the temporal aspect of errors in prognostic predictions is to group them into bins based on the current lifetime of the system, ideally created from a series of historical cases and each with continuous predictions starting early in unit life, known total unit lifetime and with total number of bins that can vary based on the available data (approximately 100 bins is expected to provide good information for most cases (Sharp, 2013)).

4.1. Standard Prognostic Model Metrics

Standard and simple prognostic metrics are largely known and used in practice for their simplicity, although most of these metrics are often not accurately considered regarding their usefulness for evaluation prognostic predictions.

It is important to note that the methods described are applied off-line and require knowledge of the true time to failure.
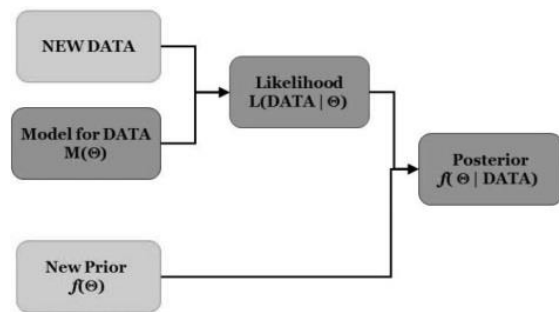

Figure 3. Bayesian Updating Procedure (Hines, 2009).

One of the most intuitive and easy to understand performance metric presented here is the Mean Absolute Error (MAE). The MAE is the average absolute difference between the model prediction and the true Remaining Useful Life at all times t and for all historic query cases i, and is defined by Eq.(2) (Sharp, 2013):

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{T_i}\sum_{t=1}^{T_i}|\Delta(i,t)| \qquad (2)$$

where:

- $\Delta(i,t)$: the error between the predicted and the true RUL at time index t for the unit under test i;

- N: the number of historic query cases.

4.2. Improvements of Prognostic Model Metrics

One of the major shortcoming of standard metrics, also presented and introduced by (Saxena, 2009b), is that they are designed to evaluate the prognostic estimations of a single query case. Considering only a single query case, the metrics report only aspects of that case, while averaging individual query based metrics over a large set of query cases could be used to evaluate the suitability of a model that produced them.

Variants on some well-known performance metrics seek to remedy this oversight to overcome and fill in the gaps left by standard metrics (Sharp, 2013). Advanced performance metrics will be defined to sufficiently characterize the output predictions of a prognostic model: Weighted Error Bias (WEB), Weighted Prediction Spread (WPS), Confidence Interval Coverage (CIC), and the Confidence Convergence Horizon (CCH). A detailed review of these metrics can be found in (Sharp, 2013). Each one captures a key aspect and desirable quality prognostic predictions that can be quickly, easily, and intuitively compared amongst separately developed models to rank and rate there output performance.

The Weighted Error Bias (WEB) represents a measure indicating the effective bias in all predictions as a percentage of total unit lifetime T_{TUL} and is defined by Eq. (3) (Sharp, 2013):

$$WEB = \frac{100}{N}\sum_{i=1}^{N}\sum_{t=1}^{T} w_i(t)\frac{\Delta(i,t)}{(T_{TUL})_i} \qquad (3)$$

where:

- $\Delta(i,t)$: the error between the predicted and the true RUL at time index *t* for the unit under test i;

- N: number of historic query cases;

- *w*: a weight factor errors based on their time in the lifecycle of the historic unit. This weighting factor can be altered to fit the specific needs of any system, but is generally set as a Gaussian function centered on the end of life with bandwidth of 50% of the unit total lifetime. This serves to accentuate the errors at end of life, generally the most critical portion of lifetime to have accurate predictions.

The WEB is a measure of the bias of model predictions (positive or negative). A small bias percentage throughout prediction time is desired. This metric also allows a simple method for model improvement, by subtracting the indicated percent bias from all the model predictions.

The predictions of Remaining Useful Life (RUL) made by a model are clearly more important near the end of the system's life than at the beginning of life. The prediction spread for each binned point of system life is calculated as the difference between the upper and lower bounds of the corresponding 95% confidence intervals from the binned error values. Using the same weighting function as the Weighted Error Bias (WEB), the Weighted Prediction Spread (WPS) can be defined by Eq.(4) (Sharp, 2013):

$$WPS = 100 * \frac{\sum_{bi=1}^{n_B} W_{bi} * CI_{bi}}{\sum_{bi=1}^{n_B} W_{bi}} \qquad (4)$$

where:

- n_B: Number of bins;

- CI_{bi}: Confidence Interval for bi^{th} bin;

- W_{bi}: Weighting function based on the center value Bin_{bi} for each reference bin. Each bin importance weighting, W_{bi}, is defined by the Gaussian kernel with a kernel bandwidth of 50% (Eq. (5)) (Sharp, 2013):

$$W_{bi} = \exp\left(-\left(\frac{Bin_{bi}-100\%}{50\%}\right)^2\right) \qquad (5)$$

This weighting factor is equivalent to that used for WEB, but keyed to bin location times instead of per observation time locations. The WPS Metric provides the Weighted Prediction Spread in percentage.

A more explicit and useful metric evaluating this coverage is the Confidence Interval Coverage (CIC). This metric incorporates information relating to both the error bias and the error variance at given points in life. It is defined by Eq. (6) as the total percentage of binned error sets whose 95% confidence interval contains the true RUL (Sharp, 2013):

$$CIC = 100 * \frac{\sum_{bi=1}^{n_B} TPRUL_{bi}}{n_B} \qquad (6)$$

where:

- n_B: Number of bins;

- $TPRUL_{bi}$: true percent RUL values that are contained within their corresponding error bin set (i.e. B_{bi}).

This additional metric verifies the total accuracy of the prediction set. An optimal coverage of 100% shows that the true value of any prediction is contained within the prediction spread or approximate confidence interval of the prognostic model's predictions. The explicit end of life accuracy and precision of a prediction set is estimated by another important metric, the Confidence Convergence Horizon (CCH). A 10% Confidence Convergence Horizon

(CCH), or simply the Convergence Horizon (CH), identifies the percentage of system Remaining Useful Life (RUL) beyond which all prediction confidence intervals are both less than 10% of the total system life and contain the true RUL.

Each of the metrics detailed here provide information on different aspects of the output error for a particular model. By combining the information provided in each of the metrics, a complete picture of the form and magnitude of the model errors can be made. For example, if a model exhibits a low WEB but a proportionally high MAE, this would indicate that there are high early lifetime errors which dissipate within the critical zone as specified by the WEB weighting factor. Note that these particular metrics are not in the same units, so direct comparisons of them must be done carefully. Specifically, and for more rapid and convenient comparisons of overall model errors, the percentage based metrics (WEB, WPS, CIC, CH) can be directly incorporated into a single aggregate scoring metric to rank the overall performance of a particular prognostic model (Sharp, 2013).

4.3. Absolute Percent Error

For a given model with a specific combination of features (M_1 through M_11 in Table 4), the quality of the RUL estimations changes for each motor because it depends essentially on the shape of the generated prognostic parameter. Another important influencing factor is the similarity of the estimated prognostic parameter to the prognostic parameters of the motors used to build the model.

To evaluate the RUL estimation throughout the lifetime, for each motor i at the time step t, the Absolute Percent Error (APE) can be used, as defined by Eq. (7):

$$\text{APE} = \frac{\Delta(i,t)}{T_{tF}} \cdot 100 \qquad (7)$$

where:

- $\Delta(i,t)$: the error between the predicted and the true RUL at time t for the unit under test i;

- T_{tF}: Time to Failure.

It should be noted that, while the metrics defined in Section 4.2 allow the evaluation of RUL estimation throughout the entire lifetime for all motors for each model (see Table 4), the APE refers to a specific motor at a given time point. In other words, the metrics in Section 4.2 provide the overall performance of a model, while the APE tracks the predicted RUL of a specific motor as time passes, compared to the true RUL.

5. EXPERIMENTAL DATA FOR METHODOLOGY VALIDATION

Ten 5HP U.S. Electrical Motors/Emerson general-purpose industrial motors were chosen as low cost analogs to the high power induction motors found throughout industry. The motors were run through a degradation cycle on a weekly basis. A cyclic thermal aging process, designed to induce accelerated insulation breakdown and corrosion within the motors, was applied to each of these three-phase, 3600 rpm motors. First, the motors were heated for three days in an oven. After the heating cycle, the motors were placed in a moisture testing bed with high humidity for further degradation. Then the motors were allowed to cool for a few hours before being placed in the second heating cycle for three additional days. After the second heating cycle, the motors were placed on a test bed and run for one hour. The accelerated aging plan has been adapted from a previous work performed by (Upadhyaya et al., 1997) and as suggested by IEEE Standard 117 (1974). According to IEEE Standard 117 several testing procedures may be performed in order to determine accelerated degradation testing of motors. For the type of insulation in the tested motors, the recommended testing time is 32 days at 170 °C. In this testing, the motors have been divided into two groups, one heated to 160 °C, and one with the temperature at 140 °C. Motors from #1 to #7 were heated at 160 °C, while motors from #8 to #10 were heated at 140 °C. The lower than IEEE Standard 117 temperature for the "hot group" provides a slower and more realistic evolution of any degradation mechanisms and related features of the motors. This also provides more data points, making a more accurate tracking and estimation of the degradation curve of the testing motors. The IEEE Standard 117 also recommends that the motors undergo moisture testing as well as thermal degradation to better simulate normal operating conditions. In order to achieve the moisture testing, the motors were placed in a condensation chamber consisting of temperature-regulated coolant in a sealed container for a total of 48 hours at 100 % humidity. The moisture should be uniform across the testing motor and no voltage should be applied at this time. After the condensation testing, the motors should be allowed to dry overnight. Each accelerated aging cycle has been designed to take just over one week. After a thermal aging cycle, each motor was mounted on a test bed, connected through an elastomeric coupling to a generator, and instrumented with a data collection system to collect various key signals. The following thirteen variables were monitored during this test using the data collection setup:

- Motor Current (three phases)

- Motor Voltage (three phases)

- X & Y Direction Accelerometer

- Industrial Microphone

- Tachometer

- Temperature

- Output Current (Generator)

- Output Voltage (Generator)

Steady-state data, taken for 2 seconds at 10,240 Hz every 15 minutes for one hour, gives a total of 4 steady state data files a day. Each data file, extracted for each test, contains 17,000 values. The data used in this research was obtained from five of the motors, which are #2, #3, #5, #6 and #7 because these motors all present similar degradation due to bearing failures. Each motor was run to failure after a different number of accelerated degradation cycles and tests. The number of tests for each motor before the failure is reported in Table 1.

6. VALIDATION OF THE PROGNOSTIC METHODOLOGY AGAINST EXPERIMENTAL DATA OF ELECTRIC MOTORS

A schematic summary of the prognostic procedure is shown in Figure 4. All the steps required by the methodology are outlined in the following.

6.1. Feature Extraction from Degradation Data

The steady state motor data was analyzed in both time and frequency domain in order to extract useful features for prognostic parameter generation.

Motor	Tests to Failure
#2	108
#3	103
#5	110
#6	109
#7	107

Table 1. Tests to failure for each analyzed motor.

The feature trend over the motor's lifetime can be either increasing or decreasing. Usable features, with clear trends over time in each motor, have been extracted from time domain by using statistical moments applied to the considered data and examining all trends for all motors. For the sake of brevity, they are not reported graphically in the paper, but are listed in Table 2. The Table 2 lists all the time series features extracted from the degradation data of the motors. Each of them presents an overall monotonic trend over time (either increasing or decreasing) and therefore can be utilized for prognostic parameter generation.

Frequency domain analysis is based on transforming the time series signals into the frequency domain. Features have

been obtained by evaluating a certain peak value in the frequency spectrum and by calculating the RMS values in the frequency domain over a specific frequency band of the frequency spectrum. Table 3 lists the frequency features extracted from the degradation data.
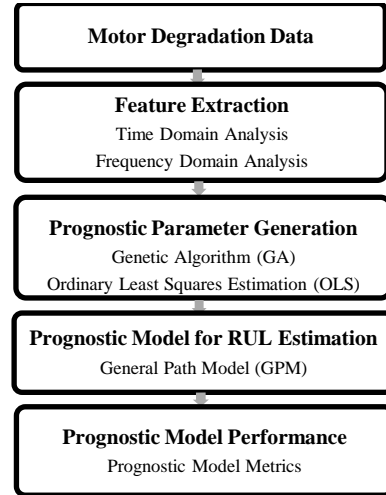

Figure 4. Steps of the prognostic methodology.

Feature Number	Description
[1;3]	Features from RMS of the Current Signals
[4;6]	Features from RMS of the Voltage Signals
[7;9]	Features from Kurtosis of the Current Signals
[10;12]	Features from Kurtosis of the Voltage Signals
[13;14]	Features from Kurtosis of the Current and Voltage Output Signals
[15;16]	Features from RMS of the Vibration Signals (X-Y)

Table 2. Usable features extracted from time domain analysis.

Feature Number	Description
[17;19]	Features from Peak Tracking of the Bearing Fault Frequencies (X-Direction Vibration Signal)
[20;22]	Features from Peak Tracking of the Bearing Fault Frequencies (Y-Direction Vibration Signal)
[23;24]	Features from RMS Values of X & Y Vibration Signals [Bearing Fault Band Frequency].

Table 3. Usable features extracted from frequency domain analysis.

For example, the RMS values are calculated in the band containing bearing fault frequency information. The frequencies investigated in the frequency spectrum, which are indicative of bearing failures, are the inner and outer race and the general ball pass frequency:

- Ball pass frequency of the inner race: 325 Hz;

- Ball pass frequency of the outer race: 215 Hz;

- General ball pass frequency: 283 Hz.

The bearings used in this experiment were SKF 6205-2Z, and the previous relevant quantities were calculated by using the equations provided by the manufacturer (SKF Catalog, 2011).

6.2. Prognostic Parameter Generation from Extracted Features

For validation purposes, the "leave one out cross validation" (LOOCV) method was used. This means that the linear combination weighting values generated using Genetic Algorithm (GA) and Ordinary Least Squares (OLS) estimation methods, applied to extracted features, were calculated for four of the motors while leaving the data source for one of the motors out. The resulting weight values, as described in 3.2.1. and 3.2.2. are then multiplied by the removed motor features data to create prognostic predictions from the generated model. Models and prognostic parameters are developed by means of features extracted both from time and frequency domain analysis and by using different combinations. Table 4 lists the eleven models developed in this paper and the corresponding combination of features used for model tuning. Models M_1 through M_7 only include features extracted from time domain analysis. In particular, models M_1 through M_5 consider different combinations of groups of features, while models M_6 and M7 consider particular combinations of specific features which, on the basis of their trend over time (not shown in this paper), were likely to be relevant for prognostic parameter estimation. Otherwise, models M_8 and M_9 only include features extracted from frequency domain analysis. Finally, model M_11 accounts for all the available features, while model M_10 does not include the two features from RMS of the Vibration Signals (X-Y). In this manner, all the most significant combinations of features are evaluated.

Generally, as expected, the prognostic parameters obtained by means of the GA approach shows better values of the three metrics (monotonicity, prognosability and trendability) than the ones obtained by means of the OLS Estimation, since the GA function uses a fitness function that sums the three parameter characteristics. However, this kind of optimization can provide prognostic parameters less smooth and less linear in trend than the OLS estimation.

Model	Features
M_1	[1;16]
M_2	[1;14]
M_3	[1;6]
M_4	[1;9]
M_5	[7;14]
M_6	[1 4 5 6 9 12]
M_7	[1 4 5 6 9 12 13 14]
M_8	[17;22]
M_9	[17;24]
M_10	[1;14 17;24]
M_11	[1;24]

Table 4. Models developed by using different combinations of features.

In fact, high values of the three prognostic parameter characteristics do not necessarily imply a linear and smooth shape.

Using a linear General Path Model, instead of other prognostic methodologies for RUL estimation, has proved that prognostics parameters with these characteristics may provide better results (Coble & Hines, 2012). The prognostic parameters, estimated for each motor by means of a model built with the same combination of features (Model M_11), which present an almost linear and smooth shape, are reported in the following. Figure 5 reports the prognostic parameters estimated by means of OLS estimation and all the features extracted from both time and frequency domain analysis. This figure can be compared directly to Figure 6, which instead shows the Prognostic Parameters estimated by means of GA approach.
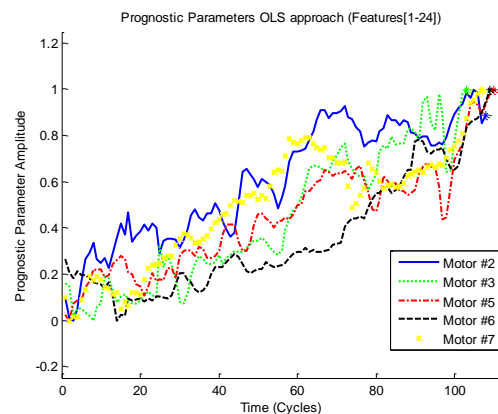

Figure 5. Estimated Prognostic Parameters (OLS estimation; M_11).

It can be noted that the prognostic parameters generated by means of the GA approach present more spikes and fluctuations over the lifetime than those derived from OLS estimation.

6.3. RUL Estimation and Model Performance Evaluation

A linear General Path Model (GPM) has been chosen in this paper as the resulting prognostic parameters followed this basic trend. If the trend was quadratic or exponential, those models would have been used. In this research Bayesian methods are used to include prior information for GPM extrapolation and RUL estimation. The inclusion of prior information improves the predictive performance when very little data has been collected (Welz, 2014).

6.3.1. OLS estimation of the prognostic parameter

Figure 7 shows the RUL estimation for a specific motor, in this case Motor #3, by using different combinations of features. The prognostic parameter is obtained by means of OLS estimation. Figure 8 shows the resulting APE.

The model built by the means of the time domain features (M_1; Features [1-16]) presents a generally good estimation, especially near the middle of the lifetime where the APE decreases considerably. The model built with the frequency domain features (M_9; Features [17-24]) provides the worst estimation near the middle of the lifetime, where the APE increases, although the performance improves near the end of testing and the APE becomes very low and comparable to the other motors. This fact is essentially due to the shape of the prognostic parameter and the presence of spikes which alter the accuracy of the RUL estimation. The model built using the means of features from both time and frequency domain (M_11; Features [1-24]) provides the best estimation near the end of the lifetime, where the APE reaches the minimum value and the estimated RUL remains good over all the entire lifetime. Figure 9 shows the estimated RUL for Motor #2, by using the same three combinations of features.
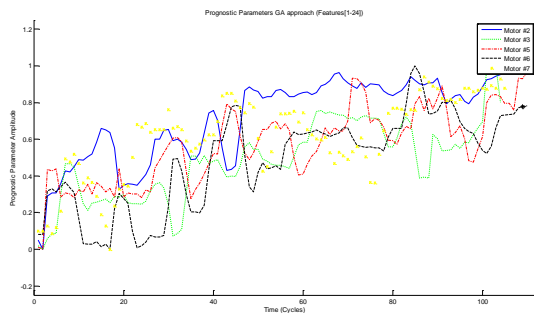

Figure 6. Estimated Prognostic Parameters
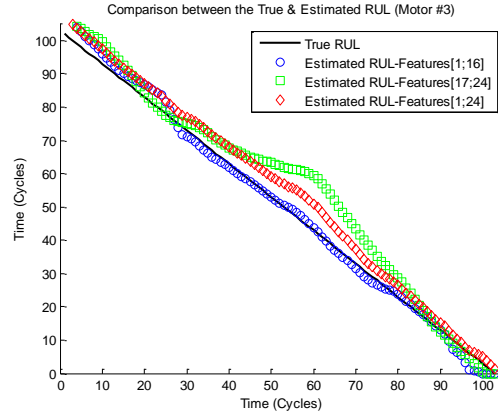
(GA estimation; M_11).

Figure 7. RUL for Motor #3 using OLS prognostic parameters and different combinations of features.
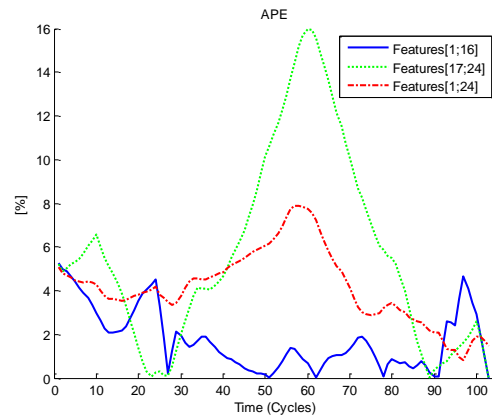

Figure 8. Resulting APE for Motor #3 using OLS prognostic parameters and different combinations of features.

In this case the best estimation over the lifetime is also provided by the model built using the means of all the features extracted both from time and frequency domain analysis (M_11; Features [1-24]). The estimated RUL nearly coincides with the true RUL throughout the lifetime, but the quality of the estimate decreases near the end of life. In this case, the model built with features derived from frequency domain (M_9; Features [17-24]) provides the worst estimation at the beginning of the testing but considerably improves after the middle of life. The model built by means of features derived from time domain (M_1; Features [1-16]) does not present an accurate estimation, except at the beginning of the lifetime. Figure 10 shows the evolution of the APE for each model considered in Figure 9. It is clear that the model M_11 presents the lowest values of APE. The APE remains under 4% until the end of life where the RUL estimations worsen.

Figure 11 shows the estimated RUL for the other motors, by using the prognostic parameters obtained by means of the

OLS estimation and by using all the features from time and frequency domain analysis (M_11; Features [1-24]).
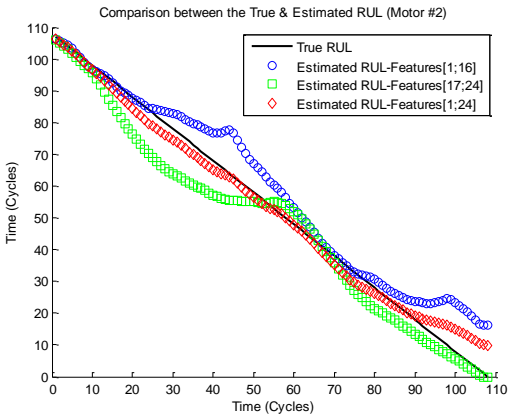

Figure 9. RUL for Motor #2 using OLS prognostic parameters and different combinations of features.
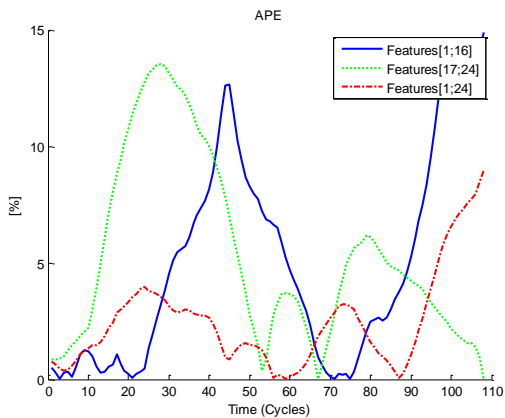

Figure 10. APE for Motor #2 using OLS prognostic parameters and different combinations of features.

It can be noted that the estimation provided by this model nearly coincides with the true RUL. As can be seen in Figure 12, the APE values remain low over the entire lifetime for all the motors (mostly below 5%). There is an increase in value after the middle of life, but the quality of the estimation returns appreciable near the end of life. For motors #5 and #6, the APE is even lower than 3% in the middle and near the end of life. This result proves the reliability of the estimation for the model M_11 with the prognostic parameter obtained by means of OLS estimation.

6.3.2 GA estimation of the prognostic parameter

Figure 13 shows the RUL estimation for motor #3, using a different combination of features, found by using the GA approach. In this case, it is clear that the RUL estimation becomes appreciable near the end of life for each

combination of features. This is to be expected since the prognostic model is linear and the GA is not constrained to find a combination of features that is linear. The GA simply attempts to optimize the three prognostic metrics. The use of a linear model may not be ideal for this model, but seems to work well. For a direct comparison between OLS and GA estimation, the Figure 13 can be compared to Figure 7.
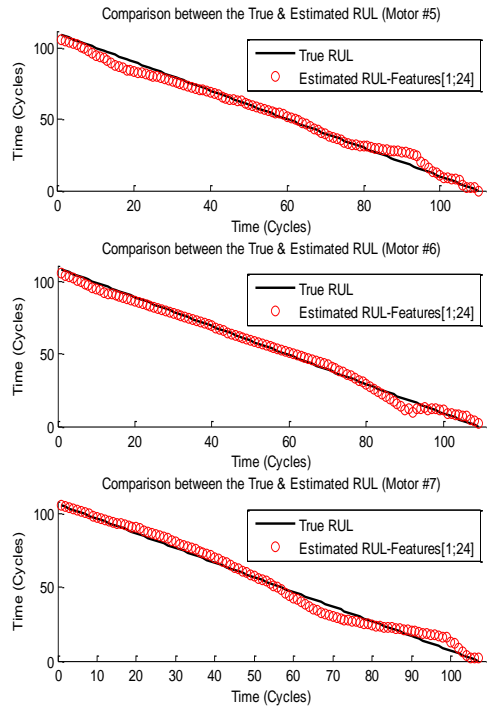

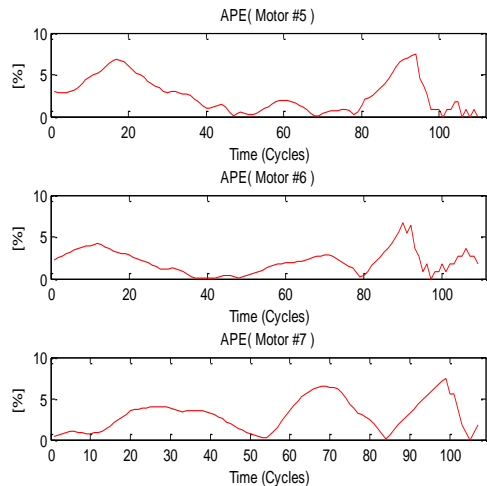Figure 11. RUL Estimation for motors #5, #6, #7 (OLS estimation; M_11).

Figure 12. APE for Motor #5, #6, #7 (OLS estimation- M_11).

It can be easily observed that the estimated RUL from the GA derived model results in a more oscillatory behavior during the lifetime. Figure 14 shows the resulting APE. The decrease of the APE value near the end of life is expected because the GA optimizes the prognosibility which constrains the endpoint to a small range.
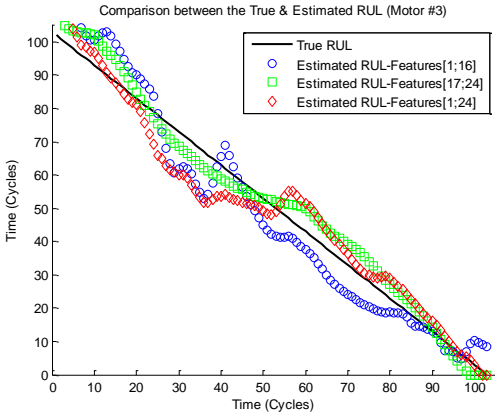

Figure 13. RUL Estimation for motor #3 using GA prognostic parameter and different combinations of features.
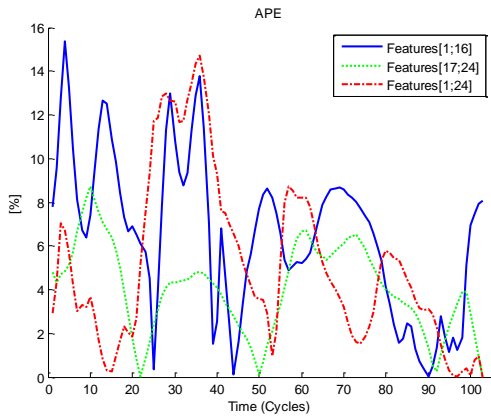

Figure 14. Resulting APE for motor #3 using GA prognostic parameter and different combinations of features.

The models built by means of features derived from the frequency domain (M_9; Features [17-24]) and features from time and frequency domain (M_11; Features [1-24]) generally present a good estimation. In particular, the RUL estimation for the model built by means of features derived from the frequency domain (M_9; Features [17-24]) remains remarkable during the lifetime and this model provides low APE values. Comparing to the other considered models, the model built using the features derived from the time domain (M_1; Features [1-16]) provides the worst estimation, characterized by the highest APE values, as is evident in Figure 14. Figure 14 can be compared to the case of the

prognostic parameters obtained for Motor #3 by means of OLS estimation reported in Figure 8.

Figure 15 shows the RUL estimation for each motor by using model M_3, which proved effective while tuning the GA approach (see Table 5).
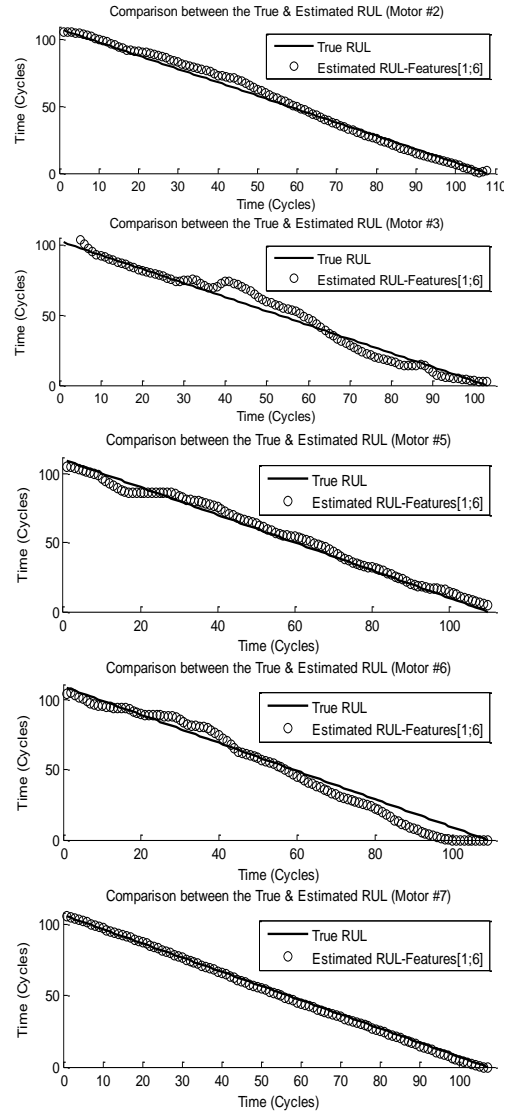

Figure 15. RUL estimation by using GA (M_3).

It can be noted that the estimated RUL provided by the model, nearly coincides with the true RUL throughout the lifetime for all the examined motors. The Figure 16 shows the corresponding trend of APE. The values in Figure 16 remain noticeably low over the lifetime for all the examined motors. For each motor, except for Motor #6, the APE values are below 5% near the end of the lifetime. It has to be noted that for Motor #7 the APE values are below 5% for the entire lifetime.

To compare the overall performance of the prognostic models, the true and predicted RUL and the actual Time to

Failure (T_{tF}) can be used to calculate the previously described prognostics metrics. All of the metrics can be fused to contribute to the aggregate score of the model, and the results are dependent on the user inputs. This means that well chosen and defined features will lead the user to create a good prognostic parameter and thus an improved RUL predictions. Table 5 and Table 6 show the results of the models using prognostic parameters generated using the means of the OLS estimation and GA methods. The highlighted entries represent the models that performed the best.

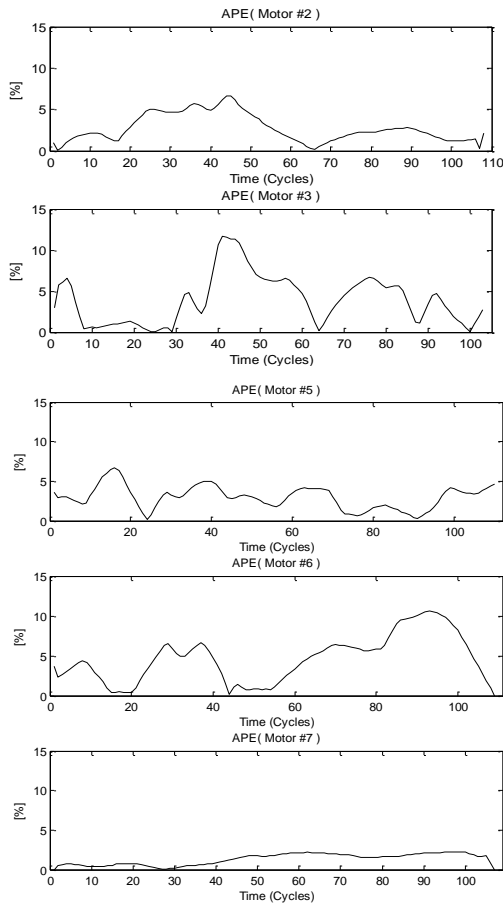In the OLS results shown in Table 5, the model M_11

Figure 16. APE by using GA (M_3).

developed by means of the features extracted from time and frequency domain provides the best aggregate score and the least amount of error in the RUL predictions. The other models also perform well in that the difference in the MAE values and aggregate scores are low.

This means that the combination of these features generates a prognostic parameter using OLS which provides approximately the same information and therefore the same results in RUL estimation.

Ordinary Least Squares Estimation Approach						
Model	Aggregate Score	MAE	WEB [%]	WPS [%]	Confidence interval coverage [%]	Convergence Horizon [%]RUL
M_1	68.76	4.99+/-4.10	-1.02	23.95	100	0
M_2	68.26	6.02+/-4.14	-1.78	25.17	100	0
M_3	68.13	5.51+/-2.16	-2.16	20.56	95.2	0
M_4	68.59	5.51+/-2.16	-0.90	24.75	100	0
M_5	66.79	11.37+/-2.76	-4.75	28.09	100	0
M_6	66.95	9.45+/-3.49	-3.39	28.79	100	0
M_7	65.12	10.44+/-5.99	-3.19	36.34	100	0
M_8	64.65	10.34+/-7.36	-4.43	36.98	100	0
M_9	67.35	8.29+/-2.96	2.19	28.40	100	0
M_10	70.58	3.71+/-0.78	1.63	16.06	100	0
M_11	71.06	3.49+/-1.78	-0.32	15.46	100	0

Table 5. Prognostic Model Results using OLS to generate prognostic parameters.

Genetic Algorithm Approach						
Model	Aggregate score	MAE	WEB [%]	WPS [%]	Confidence interval coverage [%]	Convergence Horizon [%]RUL
M_1	61.79	21.65+/-9.13	-9.43	33.89	81	9.5
M_2	62.11	18.43+/-11.87	0.23	51.32	100	0
M_3	71.51	3.35+/-1.34	-1.19	12.78	90.5	9.5
M_4	56.39	22.78+/-12.27	6.09	68.36	100	0
M_5	55.86	30.32+/-22.41	-2.44	64.59	90.5	0
M_6	61.45	16.45+/-9.14	-4.26	35.36	85.7	0
M_7	57.22	23.97+/-6.61	-2.49	54.35	85.7	0
M_8	64.01	11.03+/-6.28	1.42	42.56	100	0
M_9	65.95	9.96+/-5.81	1.35	34.87	100	0
M_10	65.43	10.04+/-4.09	-4.62	24.12	90.5	0
M_11	68.03	9.02+/-3.46	-1.95	25.91	100	0

Table 6. Prognostic Model Results using GA to generate prognostic parameters.

Using only features from frequency domain (M_8, M_9) provides results in aggregate score similar to the cases of time domain features but a clear difference in WPS values. This can be explained by considering the differences in trend, among the motors, in the features extracted from frequency domain. These differences negatively influence the prognostic parameter generation and subsequently model performance. By using the OLS method, on average, the best results in RUL estimation are achieved by increasing the number of features utilized for parameter generation. In this specific case, the best results are obtained

14

by using prognostic parameters generated from a combination of time and frequency domain features. Furthermore, the prognostic parameter generated from M_11 showed the best shape and trend over time. It has to be noted that none of the models were able to obtain the 10% Convergence Horizon (CH).

Figure 17 shows the results for the OLS case when M_11 is used for RUL estimation. The average RUL prediction (blue line in Figure 17) closely follows the true RUL (red line). There is a slight deviation near the end of life, or at 100% life consumption, but this difference is not so large that RUL predictions cannot be made.

In the GA results reported in Table 6, the model M_3 that uses only the first six features from the time domain, which are derived from the RMS values of the motor current and voltage signals, has the best overall aggregate score. In this case, unlike the best result obtained by using OLS estimation approach to generate the prognostic parameter, the model is able to obtain the 10% convergence horizon (CH) near the end of the lifetime, meaning that the RUL estimation becomes more precise when the time of failure approaches. A similar convergence horizon result is obtained by model M_1 tuned by using all the time domain features. However, in this case, the aggregate score is decisively worse than the aggregate score obtained with the same combination of features by means of OLS for parameter generation. Comparing the model results obtained by means of a combination of time and frequency domain features, the GA results are clearly worse in aggregate score, WPS and WEB values. This result can be explained by considering that the prognostic parameter generated from all the features by OLS estimation showed in general a better trend over time than the prognostic parameter generated from the same features by GA approach. As previously discussed, this provides a better and more constant RUL estimation when using a linear model. Figure 18 shows the results for the GA case when M_11 is used for RUL estimation.

Clearly the OLS method for generating the prognostic parameters is on average preferable because it provides more similar results by changing the combination of features (so, it is more robust, because it is less sensitive to feature selection). The GA method also produces very high MAE results, except for the best case identified by model M_3. The poorer results in the GA method can be attributed to the random nature of the optimization process during prognostic parameter generation. Furthermore, this can be explained by the subsequent oscillation in RUL estimation over the lifetime, which increases in quality only near the end of testing producing sometime a good result in Convergence Horizon (CH) value.

Figure 19 depicts the results for the GA approach for prognostic parameter generation when M_3 is used for RUL estimation.
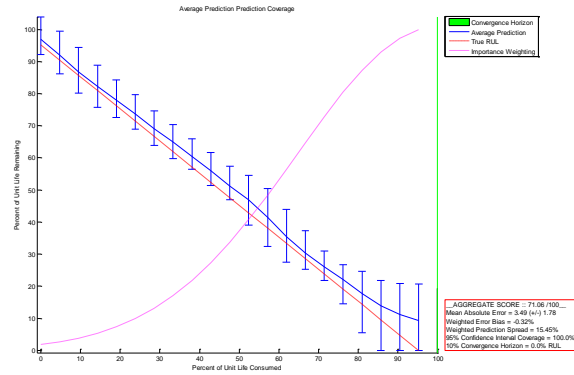

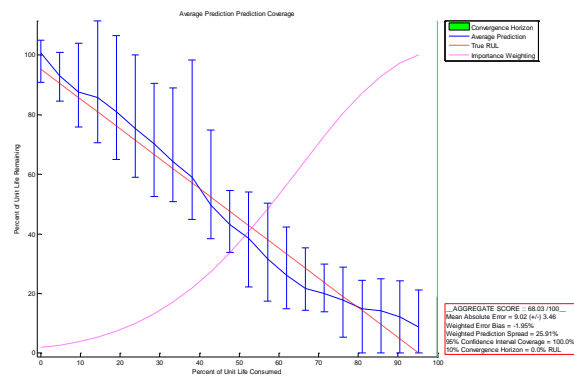Figure 17. Model M_11 results for OLS case.
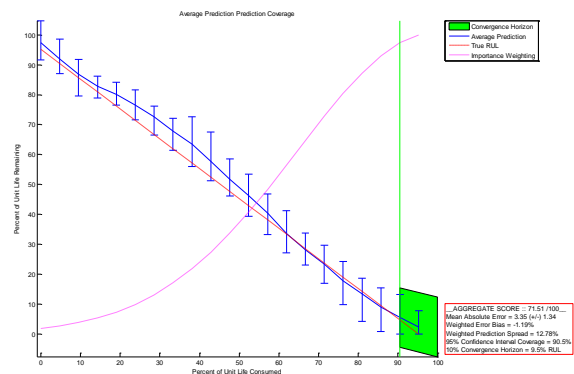

Figure 18. Model M_11 results for GA case.

Figure 19. Model M_3 results for GA case.

This represents the best result obtained by means of this approach and provides the best aggregate score value among all the developed models by means of the extracted features. This Figure can be compared to Figure 17, that depicts the best model results obtained by means of OLS to generate the prognostic parameters. The RUL estimation for each motor is shown for this model in Figure 15. It is clear that a good Convergence Horizon (CH) result at the end of

15

lifetime is due to the goodness of the RUL estimation near the end of life for each motor.

6.4. Influence of Using a Higher Number of Data Points for Feature Extraction

Since the features used for prognostic parameter generation were extracted using only one characteristic value from each of the steady state data files (each file is composed of 17,000 values), an investigation about the influence of using more data values from each motor test has been also considered. This analysis has been carried out in order to determine whether considering more data values during the feature extraction phase would result in better prediction performance. For the sake of brevity, the analysis has been focused only on the time domain features and on the OLS prognostic parameter construction method, which has previously shown better performance. The same features were selected from the motor degradation data but four and ten data points were extracted for each test. It should be noted that increasing the number of data points during the extraction step does not change the feature trend. This important characteristic has been found in each extracted feature from time domain analysis. To provide an idea of this useful information, Figure 20 shows a comparison between extracted features derived from the RMS of Voltage Signals for Motor #2, by using one, four and ten values for each test.
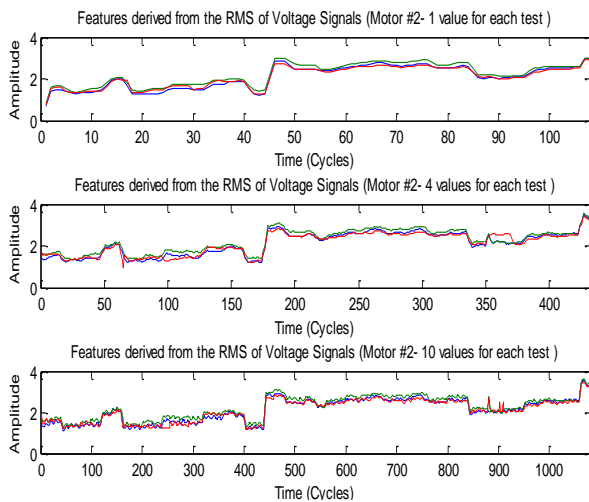

Figure 20. Features extracted using more data points.

In general, increasing the number of data points decreases model performance. This is probably due to an addition of noise in the extracted features, as shown in Figure 20, which negatively influences prognostic parameter generation and provides a worse RUL estimation, especially at the beginning of lifetime. Therefore, in this particular data set,

the use of more data points during feature extraction does not offer better predictive performance.

7. CONCLUSIONS

In this paper, motor degradation data was used for feature extraction, prognostic parameter generation, and prognostic model development. The developed approach consists of a prognostic procedure for Remaining Useful Life (RUL) estimation and involves two separate steps. First, the extracted features are fused to produce a prognostic parameter which is designed to be correlated with RUL. Then, this parameter is modeled, through a linear General Path Model (GPM), and extrapolated to a failure threshold to estimate the RUL.

The degradation data used in this research is experimental steady-state data typical of industrial electric motor degradation. Five three-phase motors were run through a degradation cycle on a weekly basis, to cause bearing failure. These degradation cycles lasted approximately seven months until failure.

Both time and frequency domain analysis were investigated. In the time domain, 16 features were extracted from the motor current, voltage, and vibration; as well as features from the generator output. The features included windowed time series moments such as the mean, standard deviation, root mean square, skewness and kurtosis. In the frequency domain, 8 additional features were extracted from the vibration signals by means of peak tracking techniques. In the frequency spectrum, peaks that are indicative of bearing failure, such as the inner and outer race and the general ball pass frequency, were investigated.

Two methods for prognostic parameter generation were evaluated: the Genetic Algorithm (GA) approach, which is random in nature, and Ordinary Least Squares (OLS) estimation. Different combinations of features were used for prognostic parameter generation and OLS usually provided prognostic parameters that were smoother and more linear.

Once the prognostic parameters were generated, they were used in the General Path Model for RUL estimation. Models were built with parameters generated by the OLS and GA methods. The results showed that the OLS method was, on average, preferable with respect to the GA method. In fact, generating prognostic parameters by means of OLS estimation proved less sensitive to feature combination. The somewhat poorer results obtained using the GA method may be attributed to the random nature of the optimization process during prognostic parameter generation. However, the single model with the best performance was obtained using the GA methodology with a specific combination of time domain features (M_3; Features [1-6]). In this case, the estimated RUL provided by the model nearly coincided with the true RUL throughout all the lifetime for all the examined motors. The error values, between the true and estimated

RUL, remained noticeably low, under 10% over the lifetime for all the examined motors. Furthermore, the absolute percent error values were, on average, under 5% near the end of life. The top GA model resulted in the best aggregate score value of 71.51.

The effect of using more data points per motor test for feature extraction was also investigated. The same features were selected from the motor degradation data with both four and ten data points extracted for each test. The feature trend did not statistically change by increasing the number of data points in feature extraction, and the fewer number of data points provided smoother features over time. Based on the obtained results, it is possible to conclude that, for the considered data set, the use of more data points during feature extraction does not offer better predictive performance and may add more noise or variance to the features, which increases modeling error. This is an unexpected result because using more data points in an OLS model usually produces a smoothing result. A possible reason for this behavior is data overfitting, which can be reduced by regularizing the solution through ridge regression or PCA technique. Further investigation may be the topic of future research.

REFERENCES

Bryg, D. J., Mink, G. & Jaw, L. C. (2008). *Combining Lead Functions and Logistic Regression for Predicting Failures on an Aircraft Engine*. ASME Paper No. GT2008-50118.

Byington, C.S., Watson, M., Edwards, D. & Stoelting, P. (2004). *A Model-Based Approach to Prognostics and Health Management for Flight Control Actuators*. Proceedings of the IEEE Aerospace Conference, 3551-3562.

Carden, E.P. & Fanning, P. (2004). *Vibration Based Condition Monitoring: A Review*. Structural Health Monitoring 3 (4): 355 - 377.

Carey, M.B. & Koenig, R.H. (1991). *Reliability Assessment Based on Accelerated Degradation: A Case Study*. IEEE Transactions on Reliability 40 (5) 1991: 499 - 506.

Carlin, B.P. & Louis, T.A. (2000), *Bayes and Empirical Bayes Methods for Data Analysis*. 2nd ed. Boca Raton: Chapman and Hall/CRC.

Cavarzere, A. & Venturini, M. (2012). *Application of Forecasting Methodologies to Predict Gas Turbine Behavior Over Time*. ASME J. Eng. Gas Turbines Power, 134(1), p. 012401.

Coble, J.B. (2010). *Merging Data Sources to Predict Remaining Useful Life-An Automated Method to Identify Prognostic Parameters*. A Doctoral Dissertation, The University of Tennessee, Knoxville TN.

Coble, J. & Hines J.W. (2011). *Applying the General Path Model to Estimation of Remaining Useful Life*.

International Journal of Prognostics and Health Monitoring (IJPHM), Vol 2 (1) 007, pages: 13.

Coble, J. & Hines, J.W. (2012). *Identifying Suitable Degradation Parameters for Individual-Based Prognostics in Diagnostics and Prognostics of Engineering Systems: Methods and Techniques*, Editor, Seifedine Kadry, pp 135-150.

Dyer, D. & Stewart, R.M. (1978). *Detection of rolling element bearing damage by statistical vibration analysis"*, Journal of Mechanical design.

Elsayed, E.A. & Chen, A.C-K. (1998). *Recent Research and Current Issues in Accelerated Testing*. Proceedings of the IEEE Conference on Systems, Man, and Cybernetics, 1998: 4704 - 4709.

Ferrell, B.L. (1999). *JSF Prognostics and Health Management*. Proceedings of the IEEE Aerospace Conference, 471.

Ferrell, B.L. (2000). *Air Vehicle Prognostics and Health Management*. Proceedings of the IEEE Aerospace Conference, 145 - 146.

Gelman, A., Carlin, J., Stern, H. & Rubin, D. (2004). *Bayesian Data Analysis*. 2nd ed. Boca Raton: Chapman and Hall/CRC.

Greitzer, F.L., Stahlman, E.J., Ferryman, T.A., Wilson, B.W., Kangas, L.J. & Sisk, D.R. (1999). *Development of a Framework for Predicting Life of Mechanical Systems: Life Extension Analysisand Prognostics (LEAP)*. International Society of Logistics (SOLE) Symposium, Las Vegas, Aug 30 - Sept 2, 1999.

Gribok, A., J. Hines, A. Urmanov, & R. Uhrig. (2002). *Regularization of Ill-Posed Surveillance and Diagnostic Measurements*. Power Plant Surveillance and Diagnostics, eds. Da Ruan and P. Fantoni, Springer.

Gu, H., Zhao, J. & Zhang, X. (2013). *Hybrid methodology of degradation feature extraction for bearing prognostics*. Maintenance and Reliability; 15 (2): 195-201.

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

Heng, A., Tan, A. C. C., Mathew, J., Montgomery, N.,Banjevic, D. & Jardine, A. K. S. (2009b). *Intelligent condition based prediction of machinery reliability*. Mechanical Systems and Signal Processing, 23, pp. 1600-1614.

Hess, A., Calvello, G. & Frith, P. (2005). *Challenges, Issues, and Lessons Learned Chasing the 'Big P': Real Predictive Prognostics Part 1*. Proceedings of the IEEE Aerospace Conference, pp. 3610 - 3619.

Hines, J.W., Seibert, R. & Arndt, S.A. (2006a). *Technical Review of On-Line Monitoring Techniques for Performance Assessment*. (NUREG/CR-6895) Vol. 1, State-of-the-Art , Published January.

Hines, J.W., Usynin, A. & Urmanov, A. (2006b). *Prognosis of Remaining Useful Life for Complex Engineering Systems*. American Nuclear Society International Topical Meeting on Nuclear Plant Instrumentation, Control, and Human-Machine Interface Technologies (NPIC&HMIT), Albuquerue, NM.

Hines, J. W. (2009). *Empirical Methods for Process and Equipment Prognostics*. Reliability and Maintainability Symposium RAMS.Jardine, A. K. S., Lin, D. & Banjevic, D. (2006). *A review on machinery diagnostics and prognostics implementing condition-based maintenance*. Mechanical Systems and Signal Processing, 20, pp. 1483-1510.

Kacprzynski, G.J., Sarlashkar, A., Roemer, M.J., Hess, A. & Hardman, W. (2004). *Predicting Remaining Life by Fusing the Physics of Failure Modeling with Diagnostics*. Journal of the Mineral, Metals, and Materials Society 56 (3): 29 - 35.

Kalgren, P.W., Baybutt, M., Ginart, A., Minnella, C., Roemer, M.J. & Dabney, T. (2007). *Application of Prognostic Health Management in Digital Electronic Systems.* Proceedings of the 2007 IEEE Aerospace Conference, March: 1-9.

Keller, K., Swearingen, K., Sheahan, J., Baily, M., Dunsdon, J., Cleeve, B., Przytula, K.W. & Jordan, B. (2006). *Aircraft Electrical Power Systems Prognostics and Health Management*. IEEE Aerospace Conference, Big Sky, MT.

IEEE Standard 117. (1974). *IEEE Standard Test Procedure for Evaluation of Systems of Insulating Materials for Random-Wound AC Electric Machinery Degradation and Testing Plan*.

Li, R., Sopon, P. & He D. (2012). *Fault features extraction for bearing prognostics*. Springer Online Text.

Li, Y.G. & Nilkitsaranont, P. (2009). *Gas turbine performance prognostic for condition-based maintenance*. Applied Energy, 86, pp. 2152 – 2161.

Lindely, D.V. & Smith, A.F. (1972). *Bayes Estimates for Linear Models*. Journal of the Royal Statistical Society (B) 34 (1): 1-41.

Lipowsky, H., Staudacher, S., Bauer, M. & Schmidt, K. J. (2010). *Application of Bayesian Forecasting to Change Detection and Prognosis of Gas Turbine Performance*. ASME J. Eng. Gas Turbines Power, 132 (3), p. 031602.

Loboda, I. & Feldshteiyn, Y. (2010). *Polynomials and Neural Networks for Gas Turbine Monitoring*. Proceedings of ASME Turbo Expo 2010. GT2010-23749.

Lu, C.J. & Meeker, W.Q. (1993). U*sing Degradation Measures to Estimate a Time-to-Failure Distribution*. Technometrics Vol 35 No 2, pp 161-173.

Luo, J., Namburu, M., Pattipati, K., Qiao, L., Kawamoto M., & Chigusa, S. (2003). *Model-based Prognostic Techniques*. IEEE Systems Readiness Technology Conference (AUTOTESTCON), Sep 22-25: Anaheim, CA.

McInerny, S.,A. & Dai, Y. (2003). *Basic Vibration Signal Processing for Bearing Fault Detection*. IEEE Transactions on Education, Vol. 46, No. 1.

Mishra, S. & Pecht, M. (2002*). In-situ Sensors for Product Reliability Monitoring*. Proceedings of SPIE, vol 4755 pp 10-19.

Nam, A., Sharp, M., Hines, J. W. & Upadhyaya, B. R. (2012). *Bayesian Methods for Successive Transitioning Between Prognostic Types: Lifecycle Prognostics*. 8th International Topical Meeting on Nuclear Plant Instrumentation, Control and Human Machine Interface Technologies, San Diego, CA, pp. 1-8, Paper #262.

Orchard, M. & Vachtsevanos, G., (2007). *A particle filtering approach for on-line failure prognosis in a planetary carrier plate*. International Journal of Fuzzy Logic and Intelligent Systems, vol. 7, no. 4, pp. 221-227.

Orme, G. J. & Venturini, M. (2011). *Property Risk Assessment for Power Plants: Methodology, Validation and Application*. Energy, 36, pp. 3189-3203.

Palmè, T., Fast, M., Assadi, M. & Pike, A. (2009). *Different Condition Monitoring Models for Gas Turbines by means of Artificial Neural Networks*. Proceedings of ASME Turbo Expo 2009. GT2009-59364.

Park, C. & Padgett, W.J. (2006). *Stochastic Degradation Models with Several Accelerating Variables*. IEEE Transactions on Reliability 55 (2) 2006: 379 - 390.

Poyhonen, S., Jover, P. & Hyotyniemi, H. (2004). *Signal processing of vibrations for condition monitoring of an induction motor*. ISCCSP: 2004 First International Symposium on Control, Communications and Signal Processing, New York, pp. 499 - 502.

Puggina, N. & Venturini, M. (2012). *Development of a Statistical Methodology for Gas Turbine Prognostics*. ASME J. Eng. Gas Turbines Power, 134(2), 022401 (9 pages).

Ramakrishnan, A. & Pecht, M. (2003). *A Life Consumption Monitoring Methodology for Electronic Systems*. IEEE Transactions on Computer Packaging Technologies, Vol. 26, No. 3. pp.625-634.

Roemer, M.J., Dzakowic, J., Orsagh, R.F., Byington, C.S. & Vachtsevanos, G. (2005). *Validation and Verification of Prognostic and Health Management Technologies*. Proceedings of the IEEE Aerospace Conference, 3941 - 3947.

Saxena, A., Celaya, J., Saha, B., Saha, S. & Goebel, K. (2009a). *On Applying the Prognostic Performance Metrics*. Annual Conference of the Prognostics and Health Management Society 09, San Diego, CA, pp. 1-16, Paper #039.

Saxena, A., Celaya, J., Saha, B., Saha, S. & Goebel, K. (2009b). *Evaluating Algorithm Performance Metrics Tailored for Prognostics*. Proceedings of IEEE Aerospace Conference, Big Sky, MT.

Saxena, A, Celaya, J, Saha, B., Saha, S., & Goebel, K. (2010). *Metrics for Offline, Evaluation of Prognostic*

Performance. International Journal of Prognostics and Health Management. Vol 1 (1) 001.

Sharp, M. (2012). *Prognostic Approaches Using Transient Monitoring Methods*. A Doctoral Dissertation, The University of Tennessee, Knoxville TN.

Sharp, M. (2013). *Simple Metrics for Evaluating and Conveying Prognostic Model Performance To Users With Varied Backgrounds*. International Journal of Prognostics and Health Management.

Si, X. S., Wang, W., Hua, C. H. & Zhou, D. H. (2011). *Remaining useful life estimation - A review on the statistical data driven approaches*. European Journal of Operational Research, 213, pp. 1-14.

SKF Group. Deep Groove Ball Bearing Catalog (Single Row). (2011). Catalog Number 6205-2Z.

Sreenuch, T., A. Tsourdos & I.K. Jennions (2013). *Distributed embedded condition monitoring systems based on OSA-CBM standard*. Computer Standards & Interfaces, 35 (2), pp 238–246.

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(8): 469-480.

Tang, L.C. & Chang, D.S. (1995). *Reliability Prediction Using Nondestructive Accelerated Degradation Data: Case Study on Power Supplies*. IEEE Transactions on Reliability 44 (4) 1995: 562-566.

Upadhyaya, B.R., Naghedolfeizi, M. & Raychaudhuri, B. (1994). *Residual Life Estimation of Plant Components*. P/PM Technology, p 22-29.

Upadhyaya, B.R., Erbay, A.S. & McClanahan J. P. (1997). *Accelerated Aging Studies of Induction Motors and Fault Diagnostics*. Maintenance and Reliability Center, The University of Tennessee, Knoxville.

Usynin, A., Hines, J.,W. & Urmanov, A. (2008), *Uncertain Failure Thresholds in Cumulative Damage Models*. Nuclear Engineering Department, University of Tennessee Knoxville.

Vachtsevanos, G., Kim, W., Al-Hasan, S., Rufus, F., Simon, M., Schrage, D. & Prasad, J.V.R. (1997). *Mission Planning and Flight Control: Meeting the Challenge with Intelligent Techniques*. Journal of Advanced Computational Intelligence, Vol. 1, No. 1, pp. 62-70.

Venkatasubramanian, V., Rengaswamy, R., Yin, K., & Kavuri, S. N. (2003). *A review of process fault detection and diagnosis: Part i: Quantitative model-based methods*. Computers & Chemical Engineering, 27(3):293-311.

Venkatasubramanian, V., Rengaswamy, R. & Kavuri, S. N. (2003). *A review of process fault detection and diagnosis: Part ii: Qualitative models and search strategies*. Computers & Chemical Engineering, 27(3):313-326.

Venkatasubramanian, V., Rengaswamy, R., Yin, K. & Kavuri, S. N. (2003). *A review of process fault detection and diagnosis: Part iii: Process history based methods*. Computers & Chemical Engineering, 27(3):327-346.

Venturini, M. & Puggina, N. (2012). *Prediction Reliability of a Statistical Methodology for Gas Turbine Prognostics*. ASME J. Eng. Gas Turbines Power, 134(10), 101601 (9 pages).

Venturini, M. & Therkorn, D. (2013). *Application of a Statistical Methodology for Gas Turbine Degradation Prognostics to Alstom Field Data*. J. Eng. Gas Turbines Power 135(9), 091603 (10 pages), doi:10.1115/1.4024952.

Vichare, N. & Pecht, M. (2006). *Prognostics and Health Management of Electronics*. IEEE Transactions on Components and Packaging Technologies, 29 (1), pp 222 - 229.

Wang, P. & Vachtsevanos, G. (2001). *Fault Prognostics Using Dynamic Wavelet Neural Networks*. Journal of Artificial Intelligence for Engineering Design, Analysis and Manufacturing, vol.15, pp. 349-365.

Watson, M.J., Smith, M. J., Kloda, J., Byington, C. S. & Semega, K.. (2011). *Prognostics and Health Management of Aircraft Engine EMA Systems*. ASME Paper No. GT2011-46537.

Welz, Z., Nam, A, Sharp, M., Hines, J.W & Upadhyaya, B. R. (2014). *Prognostics for Light Water Reactor Sustainability: Empirical Methods for Heat Exchanger Prognostic Lifetime Predictions*. 2nd European Conference of the Prognostics and Health Management Society (PHME'14), Nantes, France, July 8-10.

Yang, H., Mathew, J. & Ma, L. (2003). *Vibration Feauture Extraction Techniques for Fault Diagnosis of Rotating Machinery-A Literature Survey*. Asia - Pacific Vibration Conference, 12-14 November, Gold Coast, Australia.

Zhang, X., Xu, R., Kwan, C., Liang, S.Y, Xie, Q. & Haynes, L. (2005). *An Integrated Approach to Bearing Fault Diagnostics and Prognostics*. American Control Conference June 8-10, 2005. Portland, OR, USA.

BIOGRAPHIES

Federico Barbieri received his Bachelor Degree in Mechanical Engineering in 2011 and his Master Degree in Mechanical Engineering in 2013 from the Università degli Studi di Ferrara (Italy). As a Visiting Research Scholar, he developed his MSc thesis at the Nuclear Engineering Department of the University of Tennessee (USA), where he carried out applied research in the area of prognostics algorithms, system signal processing and optimization methods.

J. Wesley Hines is Professor and Head of Nuclear Engineering at the University of Tennessee (UT). He received the B.S. degree in Electrical Engineering from Ohio University and was a nuclear qualified submarine officer in the Navy. He has received an MBA, an M.S. in Nuclear Engineering, and a Ph.D. in Nuclear Engineering

from The Ohio State University. Dr. Hines teaches and conducts research in process diagnostics, prognostics and condition-based maintenance. His most recent research projects explore the use of advanced technologies for sensor calibration monitoring, plant fault detection, and plant performance optimization for nuclear, fossil fuel, and chemical process facilities.

Michael E. Sharp is an Assistant Research Professor at the University of Tennessee Nuclear Engineering Department. He received his PhD in nuclear engineering from the University of Tennessee in 2012. Research interests include empirical modeling, system signal processing, optimization methods, and automated pattern matching techniques. Current research includes optimization of generic prognostic extrapolation techniques and information merging.

Mauro Venturini is an Associate Professor of "Energy Systems and Power Generation" at the Dipartimento di Ingegneria dell'Università degli Studi di Ferrara (Italy), where he has also served as an Assistant Professor since 2000. He graduated in Nuclear Engineering with honor at the Università di Bologna (Italy) in 1997. At present, his research activity is mainly focused on (i) gas turbine diagnostics and prognostics, (ii) energy production from renewable sources and (iii) micro-CHP systems for household applications.