Review of prognostic problem in condition-based maintenance.
Otilia Elena Dragomir, Rafael Gouriveau, Florin Dragomir, Eugénia Minca, Noureddine Zerhouni

To cite this version:

HAL Id: hal-00418761
https://hal.archives-ouvertes.fr/hal-00418761
Submitted on 21 Sep 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Abstract—Prognostic is nowadays recognized as a key feature in maintenance strategies as it should allow avoiding inopportune maintenance spending. Real prognostic systems are however scarce in industry. That can be explained from different aspects, on of them being the difficulty of choosing an efficient technology: many approaches to support the prognostic process exist, whose applicability is highly dependent on industrial constraints. Thus, the general purpose of the paper is to explore the way of performing failure prognostics so that manager can act consequently. Different aspects of prognostic are discussed. The prognostic process is (re)defined and an overview of prognostic metrics is given. Following that, the “prognostic approaches” are described. The whole aims at giving an overview of the prognostic area, both from the academic and industrial points of views.

I. INTRODUCTION

Maintenance activity combines different methods, tools and techniques to reduce maintenance costs while increasing reliability, availability and security of equipments. Thus, one usually speaks about fault detection, failures diagnostic, response development (choice of management actions - preventive and/or corrective) and scheduling of these actions. Briefly these steps correspond to the need, firstly, of "perceiving" phenomena, next, of "understanding" them, and finally, of "acting" consequently. However, rather than understanding a phenomenon which has just appeared like a failure (a posteriori comprehension), it seems convenient to "anticipate" it's manifestation in order to consequently and, as soon as possible, resort to protective actions. This is what could be defined as the "prognostic process" and which the object of this paper is.

Prognostic reveals to be a very promising maintenance activity as it should permit to improve safety, plan successful missions, schedule maintenance, reduce maintenance cost and down time [1]. Also, industrials show a growing interest in this thematic which becomes a major research framework; see recent papers dedicated to "CBM", condition-based maintenance [2], [3] and [4]. However, considering the benefits that such technology may bring to the security, economics and resource management fields, the research community still doesn't propose a formal framework to instrument the prognostic process and real prognostic systems are scarce in industry. That can be explained from different aspects. Firstly, "prognostic" still is not a stabilized concept: there is no consensual way of understanding it which makes harder the definition of tools to support it in real applications. Secondly, many approaches for prediction exist whose applicability is highly dependent of the available knowledge on the monitored system. Thirdly, the vagueness of prognostic process definition impedes to point out the inherent challenges for scientists. Thus, the purpose of this paper is to analyze and discuss the prognostic process from different points of view: the concept, the measures and the tools. The whole aims at giving a frame to perform (and develop) real prognostic systems.

The paper is organized as follows. First of all, the concept of "prognostic" is briefly defined and positioned within the maintenance strategies. Some developments have been led to improve the proactive capacities of maintainers. So, the next part is dedicated to the analysis of the tools used in prognostic and prediction. At this stage, the advantages and the drawbacks of the identified research approaches for prognostic's purpose are pointed out.

II. PROGNOSTIC CONCEPT (RE) DEFINITION

A. Concept of prognostic

The European Standard on maintenance terminology does not define "prognostic" [5]. It doesn't appear either on the IFAC keywords list. This reveals that prognostic is a quite new area of interest.

Prognostic is traditionally related to fracture mechanics and fatigue. It started to be brought up by the modal analysis community as a field of interest [6]. In this "meaning", prognostic is called the prediction of a system’s lifetime and corresponds to the last level of the classification of damage detection methods introduced by [7]. Prognostic can also be defined as a probability measure: a way to quantify the chance that a machine operates without a fault or failure up to some future time. This "probabilistic prognostic value" is all the more an interesting indication as the fault or failure can have catastrophic consequences (e.g. nuclear power plant) and maintenance manager need to know if inspection intervals are appropriate. However, a small number of papers address this acceptation for prognostic [6], [8].

Finally, although there are some divergences in literature, prognostic can be defined as recently proposed by the
International Organization for Standardization: "prognostic is the estimation of time to failure and risk for one or more existing and future failure modes" [9]. In this acceptance, prognostic is also called the "prediction of a system's lifetime" as it is a process whose objective is to predict the remaining useful life (RUL) before a failure occurs given the current machine condition and past operation profile [2].

All definitions proposed here before assimilate prognostic to a "prediction process": a future situation must be caught. In addition, this obviously supposes: a first to predict the evolution of a situation at a given time, and a second one to assess this predicted situation with regards to an evaluation referential (Fig. 1.). A central problem can be pointed out: the accuracy of prognostic system is related to its ability to approximate and predict the degradation of equipment; the prediction phase is a critical one. A look at prognostic metrics enables to point it out.

![Fig. 1. Prognostic as a prediction and assessment process [10].](image)

**B. Prognostic metrics**

There is no general agreement as to an appropriate and acceptable set of metrics that can be employed in prognostic applications, and researchers and maintenance practitioners are still working on this [12]. Various measures emerge however from literature: as for any industrial task, prognostic can be evaluated at least in two ways: 1) the main objective of prognostic is to provide the efficient information that enables the underlying decision process, i.e., the choice of maintenance actions. Also, a first set of metrics are those that quantify the risks incurred by the monitored system. This kind of metrics can be called the **prognostic measures**, 2) assuming that prognostic is in essence an uncertain process, it is useful to be able to judge from its "quality" in order to imagine more suitable actions. In this way, **prognostic system performance measures** can be constructed.

![Fig. 2. Some prognostic metrics.](image)

**Prognostic measures**

The main prognostic measure pursued is the predicted time to failure (TTF), also called the remaining useful life (RUL). In addition, a confidence measure can be built to indicate the degree of certainty of the future predicted failure time. By extension, and considering that practitioners can be interested in assessing the system with regard to any performance limit, RUL and confidence can be generalized: in Fig. 2a, TTxx refers to the remaining time to overpass the performance limit Perf/xx, and Conf/xxT is the confidence with which can be taken the asset TTxx > T.

**Prognostic system performance measures**

The **timeliness** of the predicted time to failure (TTF) is the relative position of the probability density function (pdf) of the prediction model along the time axis with respect to the occurrence of the failure event. This measure evolves as more data are available and reveals the expected time to perform preventive actions [12] (Fig. 2b). According to [13], one has to define two different boundaries for the maximum acceptable late and early predictions. **Accuracy** measures the closeness of the predicted value to the actual one. It has an exponential form and is as higher as the error between the predicted value of TTF and the real one is smaller. **Precision** reveals how close predictions are clustered together and is a measure of the narrowness of the interval in which the remaining life falls. Precision follows from the variance of the predicted results for many experiments. Complementary of accuracy and precision is illustrated in Fig. 2c.

III. PROGNOSTIC APPROACHES

**A. Overview**

Various approaches to prognostics have been developed that range in fidelity from simple historical failure rate
models to high-fidelity physics-based models [14]. The required information (depending on the type of prognostics approach) include: engineering model and data, failure history, past operating conditions, current conditions, identified fault patterns, transitional failure trajectories, maintenance history, system degradation and failure modes.

Let’s have a first view of prognostics approaches that have successfully been applied on different types of problems.

- **Experience-Based Prognostics.** Use statistical reliability to predict probability of failure at any time.

- **Evolutionary/Statistical Trending Prognostics.** Multi-variable analysis of system response and error patterns compared to known fault patterns.

- **Artificial Intelligence Based Prognostics.** Mechanical failure prediction using reasoners trained with failure data.

- **State Estimator Prognostics.** System degradation or diagnostic feature tracking using Kalman filters and other predictor-corrector schemes.

- **Model-Based or Physics of Failure Based Prognostics.** Fully developed functional and physics-of-failure models to predict degradation rates given loads and conditions.

Similar to diagnosis, prognostic methods can be classified as being associated with one of the following two approaches: model-based and data-driven. Each one of these approaches has its own advantages and disadvantages, and, consequently, they are often used in combination in many applications. Next paragraphs present a synthesis of it.

### B. Model based approaches

The model-based methods assume that an accurate mathematical model can be constructed from first principles. As an example, physics-based fatigue models have been extensively employed to represent the initiation and propagation of structural anomalies. These methods often use residuals as features, where the residuals are the outcomes of consistency checks between the sensed measurements of a real system and the outputs of a mathematical model. The premise is that the residuals are large in the presence of malfunctions, and small in the presence of normal disturbances, noise and modeling errors. Statistical techniques are used to define thresholds to detect the presence of faults. Several techniques are proposed in the literature to generate residuals: parity space, parameters estimation, observers, bond graph, etc.

**Model-based literature.** Model-based approaches to prognostic require specific failure mechanism knowledge and theory relevant to the monitored machine. The existing papers propose different model based solution for the industrial problems.

Bartelmus and Zimroz [15] estimated through a demodulation process, the vibration signal for a planetary gearbox in good and bad conditions.

Kacprzynski et al. [16] proposed fusing the physics of failure modeling with relevant diagnostic information for helicopter gear prognostic.

Chelidze and Cusumano [17] proposed a general method for tracking the evolution of a hidden damage process given a situation where a slowly evolving damage process is related to a fast, directly observable dynamic system.

Luo et al. [18] introduced an integrated prognostic process based on data from model-based simulations under nominal and degraded conditions.

Oppenheimer and Loparo [19] applied a physical model for predicting the machine condition in combination with a fault strengths-to-life model, based on a crack growth law, to estimate the RUL.

Adams [20] proposed to model damage accumulation in a structural dynamic system as first/second order nonlinear differential equations.

Chelidze [21] modeled degradation as a "slow-time" process, which is coupled with a "fast-lime", observable subsystem. The model was used to track battery degradation (voltage) of a vibrating beam system.

Li et al. [22] and [23] introduced two defect propagation models via failure mechanism modeling for RUL estimation of bearings.

Ray and Tangirala [24] used a nonlinear stochastic model of fatigue crack dynamics for real-time computation of the time-dependent damage rate and accumulation in mechanical structures.

A different way of applying model-based approaches to prognostic is to derive the explicit relationship between the condition variables and the lifetimes (current lifetime and failure lifetime) via failure mechanism modeling. Two examples of research along this line are [25] for machines considered as energy processors subject to vibration monitoring and [26] for bearings with vibration monitoring. In [27] and [28] the problem of forecasting machine failure is illustrated for a high power fan bearing and a railroad diesel engine.

Engel et al. [29] discussed some practical issues regarding accuracy, precision and confidence of the RUL estimates.

Lesieutre et al. [30] developed a hierarchical modeling approach for system simulation to assess the RUL.

**Advantage and drawback.** The main advantage of model-based approaches is their ability to incorporate physical understanding of the monitored system. In addition, in many situations, the changes in feature vector are closely related to model parameters [21], and a functional mapping between the drifting parameters and the selected prognostic features can be established [31]. Moreover, if the understanding of the system degradation improves, the model can be adapted to increase its accuracy and to address subtle performance problems. Consequently, they can significantly outperform data-driven approaches (next section). But, this closed relation with a mathematical model may also be a strong
weakness: it can be difficult, even impossible to catch the system's behavior. Further, some authors think that the monitoring and prognostic tools must evolve as the system does...

C. Data driven approaches

Data-driven approaches use real data (like on-line gathered with sensors or operator measures) to approximate and track features revealing the degradation of components and to forecast the global behavior of a system. Indeed, in many applications, measured input/output data is the major source for a deeper understanding of the system degradation. Data-driven approaches can be divided into two categories: artificial intelligence (AI) techniques (neural networks, fuzzy systems, decision trees, etc.), and statistical techniques (multivariate statistical methods, linear and quadratic discriminators, partial least squares, etc.). Case-based Reasoning (CBR), intelligent decision-based models and min-max graphs have been considered as potential candidates for prognostic algorithms too.

- artificial intelligent techniques
  - neural networks (multi-layers perceptron, probabilistic neural networks, learning vector quantization, self-organizing maps, etc.),
  - fuzzy rule-based systems and neuro-fuzzy systems,
  - decision trees,
  - graphical models (Bayesian networks, hidden Markov models).

- statistical techniques:
  - multivariate statistical methods (static and dynamic principle components (PCA),
  - linear and quadratic discriminant,
  - partial least squares (PLS),
  - canonical variates analysis (CVA),
  - signal analysis (niters, auto-regressive models, FFT, etc.).

Artificial intelligence techniques.

Within the field of maintenance problems, Artificial Neural Networks (ANNs) and neuro-fuzzy systems (NFs) have successfully been used to support the detection, diagnostic and prediction processes, and research works emphasize on the interest of using it [32], [33], [34], [35], [36]: ANNs and NFs are a general and flexible modeling tool, especially for prediction problems. Let's point out the principle arguments of this assumption (non exhaustive list).

1- Adaptable tools. ANNs are data-driven self-adaptive methods in that they learn from examples and capture subtle functional relationships among the data, even if the underlying relationships are unknown. Thus, they are well suited for problems whose solutions require knowledge that is difficult to specify but for which there are enough data or observations. As examples, Recurrent Radial Basis Neural Networks (RRBF) have been used for time series prediction, detection and prognostic of nonlinear systems states (gas ovens) and for dynamic detection of breakdowns [37]. The dynamic wavelet neural networks were applied by Wang and Vachtsevanos [34] to predict the fault propagation process and estimate the RUL as the time left before the fault reaches a given value.

2- Robust tools. After the learning phase, ANNs can often correctly infer the unseen part of a population even if the sample data contain noisy information. Wang et al. [38] proved the robustness of the ANN in his researches. He compared the results of applying recurrent neural networks and neural-fuzzy inference systems to predict the fault damage propagation trend. Neuro-fuzzy networks have been use in robust prognostic systems for real time industrial applications like mechanical gears cracking by [39].

3- General tools. ANNs are capable of performing nonlinear modeling which is a really interesting characteristic as many real world systems are nonlinear too. Recurrent neural network were applied by Yam et al. [35] for predicting the machine condition trend. The dynamic wavelet neural network (DWN) was implemented to transform sensor data to the time evolution of a fault pattern and predict the remaining useful time of a bearing [40]: the DWN model was first trained by using vibration signals of defective bearings with varying depth and width of cracks, and then was used to predict the crack evolution until the final failure. Self-organizing neural networks were used by Zhang and Ganesan [36] for multivariable trending of the fault development, to estimate the residual life of a bearing system.

4- Open tools. In recent works, extensions of ANNs like neuro-fuzzy systems (NFs) have been developed in order to overpass the performance of classical neural networks, in particular for prediction problems. See [39] and [3] for an example. Chinnam and Baruah [41] presented a neuro-fuzzy approach to estimate the RUL for the situation where neither failure data nor a specific failure definition model is available, but domain experts with strong experiential knowledge are on hand.

Statistical techniques.

Statistical techniques require, due to the algorithms involved, quantitative data measurements. This information related to the sources is combined and the result is a stochastic estimation of the future state. Following paragraphs give a non exhaustive list of these techniques.

Yan et al. employed the logistic regression model to calculate the probability of failure for given condition variables [42]. A predetermined level of failure probability
was used in combination with an ARMA (autoregressive moving average) time series model to trend the condition and to estimate the RUL.

Lin and Makis [43] introduced a partially observable continuous-discrete stochastic process model to describe the hidden evolution process of the machine state associated with the observation process and to estimate the RUL.

HMM (Hidden Markov Model) and PIM (Proportional Intensity Model) are two statistical models in survival analysis that enable having trending models for the fault propagation process to estimate the future states. HMM describes the failure mechanism which depends both on age and condition variables. In [44] and [45], HMM and PIM are considered as powerful tools for RUL estimation. Vlok et al. [46] applied PIM with covariate extrapolation to estimate bearing residual life.

Wang [47] used the residual delay time concept and stochastic filtering theory to derive the residual life distribution and used an AR process to model a vibration signal for prognostic [48].

Phelps et al. [49] proposed to track sensor-level test-failure probability vectors instead of the physical system or sensor parameters for prognostics. A Kalman filter with an associated IMM (interacting multiple model) was used to perform the tracking.

In [50], a prognostic process for transmission gears is proposed by modeling the vibration signal as a Gaussian mixture. By adaptively identifying and tracking the changes in the parameters of Gaussian mixture, it is possible to predict gear faults.

Swanson [51] proposed to use a Kalman filter to track the dynamics of the mode frequency of vibration signals in tensioned steel band (with seeded crack growth).

Wang et al. [52] proposed a stochastic process, called a “gamma process”, with hazard rate as the residual life prediction criterion. The condition information considered was expert judgment based on vibration analysis.

Goode et al. [53] used the statistical process control (SPC) to separate the whole machine life into two intervals, the I-P (Installation-Potential failure) interval in which the machine is running correctly and the P-F (Potential failure-Functional failure) in which the machine is running with a problem. Based on two Weibull distributions assumed for the I-P and P-F time intervals respectively, failure prediction was derived in the two intervals and the RUL was estimated.

In [54], Garga et al. proposed a signal analysis approach for prognostics of an industrial gearbox. The main features used included the root mean square (RMS) value, Kurtosis and Wavelet magnitude of vibration data.

Zhang and Ganesan proposed a parameter estimation approach for a nonlinear model with temperature measurements of gas turbines [36]. The on-line detection procedure presented can track small variations in parameters for early warning.

Advantage and drawback. The strength of data-driven techniques is their ability to transform high-dimensional noisy data into lower dimensional information for diagnostic/prognostic decisions. AI techniques have been increasingly applied to machine prognostic and have shown improved performances over conventional approaches.

In practice however, it isn't easy to apply AI techniques due to the lack of efficient procedures to obtain training data and specific knowledge. So far, most of the applications in the literature just use experimental data for model training. Thus, data-driven approaches are highly-dependent on the quantity and quality of system operational data.

IV. CONCLUSION

Durable development involves the integration of economical strategies beside social and environmental ones for the optimization of processes. This major provocation of tripe performance outlined an interesting development area in industrial world: concepts like statically corrective or preventive maintenance were completed by predictive and proactive maintenance. In a word "prognostic" is nowadays recognized as a key feature in maintenance strategies. Obviously, considering the benefits that a "prognostic system" may bring to the security, economics and resource management fields, the industrial interest in prognostic is also perceptible in the scientific community for which it represents an exciting research area, with applicable perspectives. Thereby, although prognostic is still a novel axes of development (a few decades); many works have been led to define and develop accurate tools for that purpose. The variety of potential prognostic tools as well as the diversity of published works is of good omen for industrials that may be interested in using such technologies. However, knowing that techniques are suited to the prediction problem is not sufficient to make a choice: one must have a closer look on implementation requirements and constraints.

REFERENCES
