

# Chapter 8

## Review on Health Indices Extraction and Trend Modeling for Remaining Useful Life Estimation



Mohand Arab Djeziri, Samir Benmoussa, and Enrico Zio

### 8.1 Introduction

Because of the growing demands of equipment availability, performance, and maintenance, the scientific community has been developing methods for forecasting failures, and for the estimation of the Remaining Useful Life (RUL) for scheduling Condition-Based Maintenance (CBM) and Predictive Maintenance (PM). The National Aeronautics and Space Administration (NASA) was among the first to work on prognosis, because in the aerospace field, prognosis of failure can avoid catastrophes. Performance evaluation is a key element in fault prognosis and several methods have been proposed, based on different evaluation criteria. The work presented in [63, 100, 101] goes in the direction of a standardization of these criteria and proposes performance metrics applicable to different methods of fault prognosis. These metrics allow, on the one hand, establishing design requirements by quantifying acceptable performance limits and on the other hand, comparing different methods. In [101], a structured synthesis of the used metrics for the evaluation of the performance of fault prognosis methods that adapt to different application domains is presented, including Prognosis Horizon (PH),

---

M. A. Djeziri (✉)  
LIS, UMR 7020, Marseille, France  
e-mail: [mohand.djeziri@lis-lab.fr](mailto:mohand.djeziri@lis-lab.fr)

S. Benmoussa  
LASA, Badji-Mokhtar Annaba University, Annaba, Algeria  
e-mail: [samir.benmoussa@univ-annaba.dz](mailto:samir.benmoussa@univ-annaba.dz)

E. Zio  
Energy Department, Politecnico di Milano, Milan, Italy

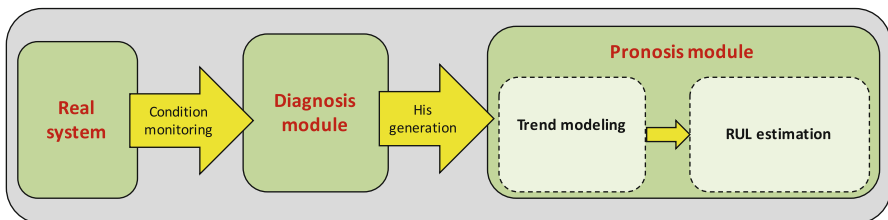
Foundation Électricité de France (EDF), Centrale Supélec, Université Paris Saclay,  
Chatenay-Malabry, France  
e-mail: [enrico.zio@polimi.it](mailto:enrico.zio@polimi.it); [enrico.zio@ecp.fr](mailto:enrico.zio@ecp.fr)

Alpha-Lambda Performance, Relative Accuracy (RA), and Convergence Rate. Fault prognosis methods are compared with respect to these metrics.

From the methodological viewpoint, several classifications of fault prognosis methods are proposed, like the pyramidal classification proposed in [19], which provides a classification into three approaches: expert approaches, physical model-based approaches, and data-driven approaches. The originality of this classification is related to the fact that these approaches are positioned in a pyramidal organization chart according to the scope of application, cost, and complexity of each approach. The evolution of hardware and software resources for data acquisition, storage, and processing has favored the widening of the application scope and the accuracy of the data-driven methods of fault prognosis. In addition, hybrid approaches have emerged to benefit from the combination of these approaches [35, 36].

This paper focuses on fault prognosis with a horizontal approach, which offers the advantage of relating fault diagnosis and prognosis. Unlike other existing reviews [45, 57, 103, 108], which only focus on RUL prediction, this paper offers a review of approaches that deal with the problem including fault diagnosis and allowing RUL estimation also when the only available data relate to normal operation. Since the health indices (HIs) generated by the methods initially used for fault diagnosis are not all usable for failure prognosis, evaluation methods of the properties that a HI must satisfy to be usable for RUL estimation, namely the Monotonicity, Trendability, and Prognosability [10, 24], are presented and then used to evaluate the usability for failure prognosis of HIs generated by fault diagnosis methods. The methods are presented in their basic version to facilitate the understanding of the ideas and the formal analysis, followed by indications and references to their extensions for particular practical cases.

As illustrated in Fig. 8.1, the structure of the horizontal approaches has two main parts: HI generation based on condition monitoring and HI trend modeling for RUL estimation. Correspondingly, this paper is organized as follows: the studied framework and definitions are presented in Sect. 8.2. Section 8.3 is devoted to formal description of methods for the definition of HIs and an analysis of their use for the estimation of the RUL. Then, the techniques for the estimation of the RUL by trend modeling are presented in Sect. 8.4, with an analysis of their complexity and performance. The purpose is to provide an overview of existing techniques and guidance for choosing approaches according to the field of application and available



**Fig. 8.1** Structure of the horizontal approach of fault prognosis

knowledge (physical knowledge, expert knowledge, data-driven). In the horizontal approach, the methods used for generating the HIs can be completely different from the method used to model trends for RUL estimation. For this reason, in this work, we have opted for a separate classification of the two parts: a classification of the methods used for the generation of HIs and a classification of the HIs trend modeling techniques for RUL estimation. In both parts, metrics are proposed for performance evaluation.

## 8.2 Study Framework

By definition, a fault is an unauthorized and unexpected deviation from the normal condition, whereas a degradation refers to the deterioration of performance in an irreversible manner. Degradation becomes failure when performance falls below a critical threshold defined in the functional specification of the equipment: the system is no longer able to perform the required function. According to the international standard (ISO 13381-1:2004), fault prognosis is defined as the estimation of the Remaining Useful Life (RUL) or the End of Life (EoL), and the estimation of the risk of subsequent development or existence of one or more faulty modes. However, in the literature, the definition of the fault prognosis concept is adapted to the context, the objectives, and the field of application, among these interpretations:

- Wang et al. [114]: In the industrial and manufacturing areas, prognosis is interpreted to answer the question: what is the RUL of a machine or a component once an impending failure condition is detected and identified.
- Mathur et al. [78]: Prognosis is an assessment of the future health.
- Lebold et al. [64]: Prognostics is the ability to perform a reliable and sufficiently accurate prediction of the RUL of equipment in service. The primary function of prognostics is the projection into the future of the current health state of equipment, taking into account the estimate of future usage profiles.
- Byington et al. [19]: Prognostics is the ability to predict the future condition of a machinery based on the current diagnostic state of the machinery and its available operating and failure history data.
- Jardine et al. [57]: Prognostics deals with fault prediction before it occurs. Fault prediction is a task to determine whether a fault is impending and estimate how soon and how likely a fault will occur.
- Muller et al. [87]: Prognostics is the ability to predict the future state of an item from its present, its past, its degradation laws, and the maintenance actions to be investigated.

In recent publications, the notion of prognosis is increasingly associated with the estimation of the RUL:

- Tobon et al. [109]: Fault prognostics can be defined as the prediction of when a failure might take place.

- Gucik-Derigny [50]: The prognosis consists in predicting the evolution of the future state of health of a system and estimating the remaining lifetime of a system before one or more failures appear on the system.
- Singleton et al. [104]: Effective diagnostic and prognostic tools are essential for timely fault detection and Remaining Useful Life prediction.
- Sun et al. [107]: Prognostics usually focuses on the prediction of the failure time or the Remaining Useful Life of a system or component in service by analysis of data collected from sensors.
- Lee et al. [66]: Prognostics can be interpreted as the process of health assessment and prediction, which includes detecting incipient failures and predicting RUL.
- Lim et al. [69]: Prognostics is the analysis of the symptoms to predict future conditions and Remaining Useful Life.

It can be noticed that the references cited above define prognosis as the prediction of the RUL based on an analysis of the monitoring condition data and the current state of the system.

## 8.2.1 Formal Definitions of the RUL

The RUL is sometimes also called Remaining Service Life, Residual Life, or Remnant Life [57], and refers to the time left before observing a failure given the current machine age and condition, and the past and future operation profile. In [103], the RUL of an asset or system is defined as the time-span from the current time to the end of the useful life. In [105], the RUL at any time  $t$  is defined as the remaining lifetime of a unit given that it is running at time  $t$  and given all the available information related to the unit at time  $t$ . Two main mathematical definitions of the RUL can be found in the literature, depending on the method used for estimating this quantity and depending on the available information: a definition of the RUL as a function of the condition monitoring ( $CM$ ) and a definition of the RUL as a function of the reliability function ( $RF$ ).

### 8.2.1.1 Definition of the RUL as a Function of CM

The definitions of the RUL given above are in agreement with the formal definition given in Jardine et al. [57], where the RUL is defined as a function of the CM of the system ( $Z(t)$ ), which gathers all the prior knowledge on the past operating state of the system as well as the co-variables that describe its current operating state, and is expressed as follows:

$$RUL(t|Z_t) = T - t | T > t, Z(t) \quad (8.1)$$

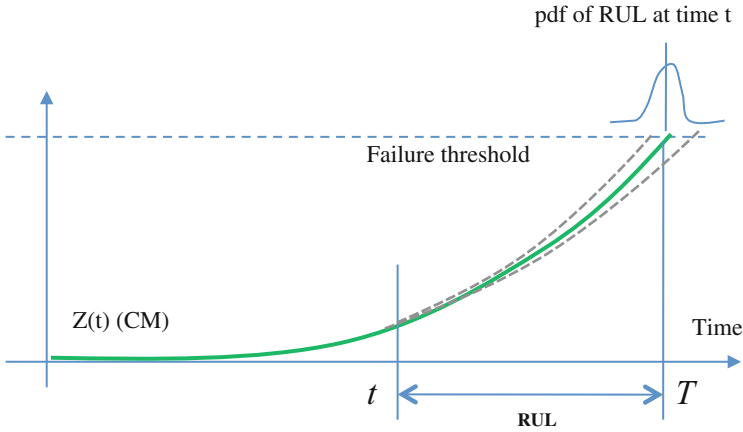


Fig. 8.2 Illustration of RUL

with:

- T: random variable of time to failure
- t: current age
- Z(t): past condition profile up to the current time.

This expression is illustrated in Fig. 8.2. According to the knowledge in  $Z(t)$ , the RUL can be calculated as a deterministic, statistic (as an expectation), or probabilistic variable (as a probability density function). The dashed gray envelope defines the margin of uncertainty about future operating conditions and the system environment.

**8.2.1.2 Definition of the RUL as a Function of RF**

In [8, 9], it is stated that information from condition monitoring can be included in reliability analysis by considering the hazard rate function as a probabilistic function. There are several methods for calculating the conditional and unconditional reliability functions (RFs) and for computing the Remaining Useful Life (RUL) as a function of the current conditions. In classical reliability, the RF is calculated mainly for two cases: as an unconditional RF, assuming that the item has not yet been put into operation ( $P(T > t)$ ), and as a conditional RF, assuming that the item has not yet failed up to sometime  $x$  ( $P(T > t | T > x)$ ).

Assuming that the system is operating at time  $t$ , the RUL is expressed in [112] as a time  $v$  for which the probability that the state of the system  $Z$  at time  $t + v$ , noted  $P[Z(t + v) \geq L | Z(t)]$  approximates the probability of failure  $q$  assumed to be known. The RUL expression is given as follows:

$$RUL(t, q) = \sup\{v : P[Z(t + v) \geq L | Z(t)] \leq q\} \tag{8.2}$$

where  $L$  is the failure threshold and  $P[Z(t + v) \geq L|Z(t)]$  is defined as the reliability of the system.

### 8.3 Health Indices Definition Methods

Existing methods of fault diagnosis have been reviewed in recent years, such as [45, 46, 108]. The most recent is proposed by D. Gao et al. [45], in which the authors propose a first classification of fault diagnosis methods into two groups: hardware redundancy-based fault diagnosis and analytical redundancy-based fault diagnosis. The fault diagnosis techniques based on analytical redundancy are structured in five classes according to the mathematical tools and prior knowledge used: model-based fault diagnosis, signal-based fault diagnosis, knowledge-based fault diagnosis, hybrid fault diagnosis, and active fault diagnosis. In this work, the methods for defining health indices are gathered in two classes: physical model-based, data analysis and signal processing. The focus is on methods that can be used for the generation of health indices sensitive to progressive degradation and whose trend can be modeled for RUL estimation. The presented methods of HI generation are summarized in Table 8.1.

#### 8.3.1 Physical Model-Based Methods

These methods are based on a physical representation of the process. They require a good understanding of the behavior of the system but does not require the availability of data on the operation of the system in degraded modes. Physical models are usually described by partial differential equations [3, 67] or state representation equations [72, 98]. Once the physical model is available, the behavior of the current process is compared with that of the model in normal operation to detect the start of degradation. After degradation has been detected, trend models are used to predict the evolution of degradation over time until reaching a failure threshold, usually predefined. The RUL corresponds to the time taken by the degradation to reach this failure threshold.

##### 8.3.1.1 Analytical Redundancy Relations

The analytical redundancy relations (ARRs) are obtained from an over-constrained system by eliminating the unknown variables [18], assuming that all functions are differentiable with respect to their arguments. For a nonlinear deterministic system (Eq. (8.3)) where  $x \in \mathfrak{N}^n$  is the state vector,  $u \in \mathfrak{N}^{m_u}$  is the control vector,  $d \in \mathfrak{N}^{m_d}$  is the disturbance vector,  $\theta \in \mathfrak{N}^{m_f}$  is a fault vector, and  $y \in \mathfrak{N}^p$  is the measurement vector:

**Table 8.1** Analysis table of HI generation methods

Approach	Physical knowledge on the system	Physical knowledge on the degradation process	Data describing the system features	Data describing the degradation	Complexity of implementation	Needed instrumentation	Consider changes in operating modes
Model-based	ARRs	Necessary	Data in normal operation are necessary for parameter identification	Modeling and parameter estimation steps are complex	Modeling and parameter estimation steps are complex	Applicable on over-constrained systems	Yes
	Parity space	Necessary	Data in normal operation are necessary for parameter identification	Modeling and parameter estimation steps are complex	Modeling and parameter estimation steps are complex	Applicable on over-constrained systems	Yes
	Observers theory	Necessary	Data in normal operation are necessary for parameter identification	Modeling and parameter estimation steps are complex	Modeling and parameter estimation steps are complex	Applicable on observable systems	Yes
Algebraic methods	Parameter estimation	Necessary	Data in normal operation are necessary for parameter identification	Modeling and parameter estimation steps are complex	Modeling and parameter estimation steps are complex	Applicable on observable systems	Yes
		Necessary	Data in normal operation are necessary for parameter identification	Modeling and parameter estimation steps are complex	Modeling and parameter estimation steps are complex	Applicable on observable systems	Yes

(continued)

**Table 8.1** (continued)

	Approach	Physical knowledge on the system	Physical knowledge on the degradation process	Data describing the system features	Data describing the degradation	Complexity of implementation	Needed instrumentation	Consider changes in operating modes
Data-based	Time	Not necessary	Not necessary	Data describing all the system features is necessary	Necessary to check sensitivity to degradation	Easy	No instrumentation constraints	No need a learning of all operating modes
	PCA SI EMD							
Time-Frequency	Time-Frequency EMD	Not necessary	Not necessary	Data describing all the system features is necessary	Necessary to define decomposition level	Easy	No instrumentation constraints	No need a learning of all operating modes
	FT DWT-HT							
Data-driven	ARR	Not necessary	Not necessary	Data describing all the system features is necessary	Necessary for clustering	Learning step is complex	No instrumentation constraints	Yes by learning of all operating modes
	AR-models NN SVR							



$$\begin{aligned}\dot{x} &= h(x, u, d, \theta) \\ y &= g(x, u, d, \theta)\end{aligned}\quad (8.3)$$

the derivative of order  $q$  of the output  $y$  gives rise to the following set of  $(q + 1)p$  constraints:

$$y^{(q)} = \gamma^q \left( x, \bar{u}^{(q)}, \bar{d}^{(q)}, \bar{\theta}^{(q)} \right) \quad (8.4)$$

where  $\bar{u}^{(q)} \in \mathfrak{R}^{(q+1)m_u}$ ,  $\bar{d}^{(q)} \in \mathfrak{R}^{(q+1)m_d}$ , and  $\bar{\theta}^{(q)} \in \mathfrak{R}^{(q+1)m_\theta}$ .

Under the condition that  $(q + 1)p > n + (q + 1)m_d$  and the Jacobian  $\left[ \frac{\partial \gamma^{(q)}}{\partial x} \quad \frac{\partial \gamma^{(q)}}{\partial d} \right]$  is of rank  $n + (q + 1)m_d$  [22], both the state  $x$  and the unknown input  $d$  can be eliminated, leading to the set of ARRs.

$$HI = r \left( \bar{y}^{(q)}, \bar{u}^{(q)}, \bar{\theta}^{(q)} \right) = 0 \quad (8.5)$$

In normal operation, Eq.(8.5) is true, whereas it is not in presence of a fault. Equation (8.5) shows that the set of residuals  $r$  is a function of the set of parameters  $\theta$  identified on the system in normal operation, and corresponding to well-identified hardware components or physical phenomena. The appearance of a progressive degradation in the system manifests a progressive deviation of the parameter affected by the degradation from its nominal value identified during normal operation. Thus, the residuals that are a function of this parameter progressively deviate from zero, enabling the detection of the start of degradation. The use of the failure signature matrix makes it possible to check the isolability of the degradation: even if the isolation of the degraded component or physical phenomenon is not always possible, the subsystem that degrades in a complex system is often possible. The identification of the component or subsystem at the origin of the degradation is a relevant knowledge for practical purposes, exploitable in the modeling of the degradation trend for RUL estimation. In addition, ARRs can be generated automatically using a bipartite graph [18] or a bond graph model [60], and are easy to implement once the parameters of the state model have been identified.

### 8.3.1.2 Parity Space

This method is applicable to linear state models and consists in eliminating the internal variables of the system by projection onto an input-output representation space, called parity space [48]. It is generally applied in a discrete time space, taking measurements over a time interval called observation window. Information redundancy is, thus, created without resorting to successive derivations of the measurements. Consider the following example of a system described by the following state model (linear or linearized around an operating point):

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ x(t) = Cx(t) \end{cases} \quad (8.6)$$

First, the observability matrix  $O_{obs}$  of the system is computed by Eq. (8.7), using the individual observability matrix of each sensor [73]:

$$\begin{aligned} O_{obs} &= (C_1 \ C_2 \ \cdots \ C_p) \\ \text{with } C_i &= (c_i \ c_i A \ \cdots \ c_i A^{n-1}) \end{aligned} \quad (8.7)$$

The observability matrix is, then, used to calculate the left null observability matrix noted  $W$ , which is not unique. In practice, it is not possible to calculate a matrix  $W$  that is perfectly orthogonal to the matrix  $O_{obs}$ , adding thus a further uncertainty to the structured and unstructured uncertainties of the model and giving rise to non-zero residuals.

After computing the matrix  $W$ , the observability matrix is reformulated in terms of inputs, outputs, and their derivatives. The derived observability given for the  $i^{th}$  output is

$$O_i = \begin{bmatrix} y(t) \\ \frac{d}{dt}y(t) \\ \vdots \\ \frac{d^n}{dt^n}y(t) \end{bmatrix} - \begin{bmatrix} 0 & 0 & \cdots & 0 \\ c_i B & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ c_i A^{n-1} B & c_i A B & \cdots & c_i B \end{bmatrix} \begin{bmatrix} u(t) \\ \frac{d}{dt}u(t) \\ \vdots \\ \frac{d^{n-1}}{dt^{n-1}}u(t) \end{bmatrix} \quad (8.8)$$

The residuals are calculated using Eq. (8.9), multiplying the global derived observability noted  $O_D$ , computed for all outputs with the matrix  $W$ :

$$\begin{aligned} HI = R &= W O_D \\ \text{with } O_D &= [o_1 \ o_2 \ \cdots \ o_p]^T \end{aligned} \quad (8.9)$$

After analyzing the equations, especially the observability matrix  $O_{obs}$  and the global derived observability matrix  $O_D$ , it is noted that the  $C_i$  and the  $O_i$  are functions of the state matrix  $A$ , whose parameters represent the physical elements of the system (physical components or physical phenomena). Thus, the occurrence of a system degradation will cause a variation of the parameters of the matrix  $A$  and, consequently, a deviation of the residuals from their values in normal operation. The residuals are, therefore, sensitive to the degradation of the system. However, the causal relationship between the residuals  $R$  and the variations of the parameters of the matrix  $A$  is not explicit: it is drawn in the process of projection in the parity space. For this reason, the parity space is used only for the detection of sensor faults, with an extension to the actuators, under the strong assumption that there are no system faults.

### 8.3.1.3 Observer Methods

Observers theory is widely used in the literature for the estimation of observable but unmeasured states [43, 82]. It has been used in fault diagnosis for the generation of health indices through the development of the unknown inputs observers. The stability and convergence analysis, the gains calculation, and assumptions on matrix rank and model inversion have been the subject of several research works [82, 84, 86] and will not be detailed in this work. Only the relevance of the use of the health indices generated using an observer in the context of fault prognosis is analyzed.

Two kinds of models are most used for the synthesis of observers in the framework of fault diagnosis and prognosis. The first model, given in Eq.(8.10) below, allows the simultaneous description of the state of the system and the degradation. These models are interconnected and are of multiple time scales, in order to highlight the difference in the evolution between the fast dynamics of the system behavior and the slow evolution of the degradation:

$$\begin{aligned}\dot{x} &= f(x, \lambda(\theta), u) \\ \dot{\theta} &= \epsilon g(x, \theta) \\ y &= Cx + Du + v\end{aligned}\tag{8.10}$$

where  $x \in \mathbb{R}^n$  is the set of state variables associated with the fast dynamics of the system;  $\theta \in \mathbb{R}^m$  is the set of slow-dynamic variables related to the degradation of the system;  $u \in \mathbb{R}^l$  is the input vector. The parameter vector  $\lambda \in \mathbb{R}^q$  is a function of  $\theta$ . The rate constant  $0 < \epsilon \ll 1$  defines the time scale separation between fast dynamics and slow drift.  $y \in \mathbb{R}^p$  is the output vector and  $v$  is the measurements noise.

The general pattern of the observer-based fault prognosis begins with the joint estimation of the state and the unknown input, with precision, and in a finite time. Then, the estimation of the RUL is carried out by a time projection of the evolution of the slow and fast dynamics until the total failure. The finite-time convergence of an observer is a less common notion in the literature than asymptotic convergence; yet it is necessary in the context of fault prognosis. New conditions of stability and convergence in finite time have been developed in Lyapunov theory, and presented in [16, 17, 85] and [86]. Methods for the synthesis of observers with finite-time convergence have been proposed in [38, 56] and [82] for linear systems, and in [79–81] and [84] for nonlinear systems. In the case of observers with unknown input and finite-time convergence, synthesis work has been developed in [97] for the linear case, and in [43] for sliding mode observers.

Although the model of Eq.(8.10) is closest to the reality of the degradation phenomenon and its interaction with the state of the system, it is strongly nonlinear and, moreover, the interaction between the state and the degradation is difficult to formalize (to model). Thus, the most used model for the generation of health indices is the second type of model given in Eq.(8.11) below [71], where the nonlinear system considered consists of a linear part exploited for the synthesis

of the observer's gain, and a nonlinear part satisfying some more or less restrictive assumptions:

$$\begin{cases} \dot{x} = l(x, f, u) \\ y = h(x, u) + W(u) f \end{cases} \quad (8.11)$$

$x \in R^n$  is the state vector.  $u \in U$  is the known input vector.  $y \in R^p$  is the output vector.  $f \in R^m$  is the vector of unknown inputs whose number is equal to or less than the number of measurements ( $mp$ ).  $W$  is the transfer matrix of the degradation to the output and it is a function of the conditions of use (of the input  $u$ ). Once the state  $x$  is estimated, it is possible to derive an estimate of the unknown input as follows:

$$HI = \hat{f} = W_1^+(u) (y_1 - h_1(\hat{x}, u)) \quad (8.12)$$

where  $W_1^+$  is the pseudo-inverse of  $W$ .

Thus, finite-time convergence is a necessary condition for the use of the unknown input observer for fault prognosis. Equation (8.10) also shows that the prior knowledge of the effect of degradation on the system is necessary, as it makes it possible to calculate the matrix  $W$ , which must be invertible. As the name implies, degradation is considered to be an unknown input, implying that any change in the dynamics of the system is considered as a degradation. Since the degradation is introduced in an additive way into the model of the system, it can be assumed that:

- The health index aggregates all the faults that may occur in the system.
- The prior knowledge necessary for the calculation of the matrix  $W$  can be used to identify the nature of the degradation.

### 8.3.1.4 Algebraic Methods

In the algebraic framework, the HIs are expressed as an algebraic equation of the system's variables ( $u$  and  $y$ ) and their derivatives. In fact, in Fliess's theory [40], and differently from Kalman's theory, a nonlinear system is defined by a differential field extension  $L/K$  finitely generated, where  $L$  is the system field which contains the system variables and the algebraic equations that links the variables;  $K$  is the ground field that contains the coefficients of the system.

The input  $u$  of a system  $L/K$  is a set  $u = \{u_1, \dots, u_m\}$  of  $L$  such that the extension  $L/K \langle u \rangle$  is differentially algebraic, which means that any element  $\omega \in L$  satisfies an algebraic differential equation over  $K \langle u \rangle$  of the form  $P(\omega, u_1, \dots, u_m, \dots, \dot{u}_1, \dots, \dot{u}_m)$ . The output of a system  $L/K$  is a set  $y = \{y_1, \dots, y_p\}$  of  $L$ . A transcendence basis  $x = \{x_1, \dots, x_n\}$  of the differential field extension  $L/K \langle u \rangle$  is the state of the system  $L/K$ . Any component of the

derivative  $\dot{x} = \{\dot{x}_1, \dots, \dot{x}_n\}$  and  $y$  are  $K < u >$  algebraic on  $x$ , which leads to the following generalized state variables representation:

$$\begin{cases} F_1(\dot{x}_1, x, \dots, u, \dot{u}, \dots, u^{(\alpha_1)}) = 0 \\ \vdots \\ F_n(\dot{x}_n, x, \dots, u, \dot{u}, \dots, u^{(\alpha_n)}) = 0 \\ H_1(y_1, x, \dots, u, \dot{u}, \dots, u^{(\beta_1)}) = 0 \\ \vdots \\ H_p(y_p, x, \dots, u, \dot{u}, \dots, u^{(\beta_p)}) = 0 \end{cases} \quad (8.13)$$

where  $F_i, H_j$  are polynomials over  $k$  and  $\alpha_i, \beta_j \in N$ ;  $x$  is called a generalized state.

In the presence of faults ( $f$ ), the nonlinear system is denoted as an algebraic differential field extension  $k(u : y, f)/k(u, f)$  [27]. If the fault  $f$  is a differential algebraic equation with coefficients over the field  $K(u, y)$ , then it is said to be diagnosable. In other words, the fault variable is written in polynomial form, as function of the input variables, the output variables and their respective derivatives as follows:

$$f = h(u, \dot{u}, \ddot{u}, \dots, u^{(m)}, y, \dot{y}, \ddot{y}, \dots, y^{(n)}) \quad (8.14)$$

The residuals ( $\hat{r}$ ) given by Eq. (8.15) below are used to evaluate the obtained fault indicator (Eq. (8.14)):

$$\hat{r} = s^{-n} \frac{d^n}{ds^n} \hat{F} \quad (8.15)$$

where the sign  $\hat{}$  means that the variable is written in the Laplace domain.

The following steps are used to obtain the residuals [11, 41]:

- Put the fault indicator in Laplace domain,
- Differentiate the result  $n$  times with respect to  $s$  in order to eliminate the initial conditions, which may be unknown,
- Multiply by  $s^{-n}$  and return back to time domain.

These residuals are performed by using the integrals of the measured signals. In the case of noisy signals, these integrals produce a filtering effect. The derivative with respect to  $s$  of order  $n$  ( $\frac{d^n}{ds^n}$ ) in Laplace domain results in a multiplication by  $(-1)^n t^n$  in the time domain, and the multiplication by  $s^{-n}$  in Laplace domain corresponds to an integration of order  $n$  in the time domain.

The fault diagnosis based on the algebraic approach is mainly applied to linear systems, and some nonlinear systems for actuator and sensor fault diagnosis. In [15], the algebraic approach in association with a bond graph tool was extended to component fault diagnosis, under the assumption that the system inputs and outputs are fault free. The residuals of Eq. (8.15) reflect only the cumulative sum of the

fault indicator from the degradation start until failure time, which means that the residuals do not reflect the degradation dynamics but only its effect. It should also be pointed out that this method does not need prior knowledge on the degradation nature and it has not been used yet for fault prognosis. Finally, this method can only handle additive faults and any change in the system dynamics can be considered as degradation.

### 8.3.1.5 Parameter Estimation Method

The fault diagnosis based on the parameter estimation consists to parametric identification of the system model using the system inputs ( $u$ ) and outputs ( $y$ ), and monitoring the estimated system parameters. For a nonlinear system described by the following state-space model:

$$\begin{cases} \dot{x} = h(x, u, \theta) + d \\ y = g(x) \end{cases} \quad (8.16)$$

where  $x$  is the state vector,  $h$  is the state function,  $g$  is the output function, and  $d$  represents the system disturbances that are assumed to be a bounded signal, and under the assumption that the system model parameters vary depending on the occurrence of a fault on the physical system. In the normal operation,  $\theta$  takes the nominal values of the physical parameters; however, in faulty operation, the value of  $\theta$  varies as a function of fault severity on the physical system. The model of Eq. (8.16) is, then, used for an on-line nonlinear parameter estimation problem, for which unknown fault parameters are estimated using system inputs and outputs, and appropriate approaches such as neural networks [115], fuzzy models [13], and Takagi–Sugeno (TS) models [88], and for linear systems, least-squares (LS) approaches are used. The estimation error (Eq. (8.17)) between the reference model parameters estimated in normal operation and the parameters estimated under faulty conditions is taken as HIs for diagnosis purposes [21].

$$HI = \hat{\theta} - \theta_n \quad (8.17)$$

The fault diagnosis via parameter estimation can handle only additive faults on parameters with slow rate dynamic. The main limit of this method is the difficulty of concluding on fault isolability conditions, since the parameters being estimated are model parameters and they do not represent the system physical parameters. This problem has been partially solved by studying the influence of each physical parameter on the model parameters [55]. As the gradual variations in degradation cause progressive changes in system parameters from their nominal values, which leads to a gradual deviation of HIs from zero, the HI trend can be analyzed to construct prediction models for RUL estimation.

### 8.3.1.6 Practical Constraints

**Data Availability** At the beginning of operation, data describing the system degradation process and expert knowledge are not available. In this case, the HIs generated from the physical knowledge are the most suitable. In the majority of the practical cases, the faulty operation is defined by thresholds that the parameters of the system must not exceed. These thresholds are used in the literature to estimate the failure thresholds for HIs whose parameters have a clear physical meaning [14, 31]. The estimation of the RUL is then performed using trend modeling methods that do not require prior knowledge of the dynamics of degradation presented in the second part of this chapter.

**System Instrumentation** The observability is a necessary condition for the implementation of the methods presented above, for the implementation of the AR method the system must be, in addition, over-constrained. This property is easily verifiable on a state model or a bond graph model. The identification of an optimal placement of sensors to obtain an observable system (or over-constrained for the application of the AR method) is also possible. But in practice, it is not always possible to place all the necessary sensors for the observability of a system, for reasons of cost, lack of space on the system, the non-availability of the sensor, and the consequences of placing a given sensor on the system. Among the practical cases of systems on which the authors of this paper have encountered difficulties of instrumentation for the application of the methods presented above: electric motors [14], where the torque sensor is rarely available, which makes it impossible to generate HIs whose electrical part and mechanical part are decoupled. Indeed, in the example of the HIs generated for the mechanical transmission system, the torque  $\Gamma(t)$  is not measured, but, rather, it is estimated using the model of the interaction between the electrical part and the mechanical part of the brushless motor, given as follows:

$$\begin{aligned} e(t) &= k_e \dot{\theta}(t) \\ \Gamma(t) &= k_t i(t) \end{aligned} \tag{8.18}$$

where  $k_e$  is the electromotive force (EMF) and  $k_t$  is the motor torque constant. The use of the current variable to calculate the torque creates a matching of the HIs generated from the electrical part of the motor with the HI generated from the mechanical part. Consequently, it is not possible to locate the degraded subsystem. On thermal engines, in particular marine diesel engines [61], there are many sensors, but the system is not observable. The available sensors are mostly effort sensors (temperature, pressure), while flow sensors (volume flow, mass flow, heat flow, entropy flow) are not available. This is due to the unavailability of some sensors (such as the entropy flow sensors) and the consequence that the sensor placement may have on the system (for example, a mass-flow sensor must be inserted in the pipe, which may promote the appearance of fluid leaks). The same instrumentation constraints are encountered on the electricity production and management systems, and industrial equipment [75, 122].

### 8.3.2 *Data-Driven and Signal Processing Methods*

Among the data-driven methods of *HIs* generation for fault diagnosis, several are used also for the generation of *HIs* for failure prognosis [124]. Methods of multivariate analysis, such as principal component analysis (PCA) and its variants (IsoMap, PCA-Kernel, ...), are widely used [2, 68, 90, 96] due to the fact that they allow, in addition to generating the *HI*, extracting a degradation profile from raw data, assuming that this information is initially contained in the raw data. The time and frequency attributes of the measurement signals are also widely used, especially when the instrumentation of the system is poor, and is limited to just one or two sensors. These techniques allow a separation of the features contained in the signal highlighting dynamics which are not perceptible on the raw signal. The features presenting progressive trends that are not related to the normal operation of the system are often related to the process of degradation of the system, and can be exploited for the prognosis of failure [49, 51, 94]. Signal processing methods the most used for generating *HIs* are: statistical indices [74], empirical modes decomposition (EMD) [54], low pass filters [92], fast Fourier transform (FFT) [77], and wavelet decomposition [76]. In the area of failure prognosis, Ref. [74] uses statistical indices to extract the characteristics susceptible to failure and robust to noise from vibration data pump oil sands; Ref. [44] also uses these statistical indices on the raw data measured on bearings; Ref. [119] applies the EMD to bearings vibration signals to identify and diagnose faults. For the diagnosis of bearing faults from acceleration signals, Ref. [4] applies filtering with several levels of bandwidth to improve the signal-to-noise ratio. For the application of the wavelet transformation, Ref. [70] applies it to the voltage data of the rolling elements of a gearbox to characterize symptoms of early fatigue and cracking.

The principle of analytical redundancy can also be applied by creating redundancy through data-driven models, such as neural networks, support vector machine (SVM), and auto-regressive models. This technique is applied in [31] for the prognosis of failures of the embedded electronic systems, where a NARX neural network is used for the estimation of the consumed power and an ARMAX model is used for the estimation of the temperature. These two estimates are then compared to the measured values to generate health indices for fault diagnosis and failure prognosis. A multilayer perceptron (MLP) neural network is used in [47] for health condition monitoring of a wind turbine gearbox, and a recurrent neural network (RNN) is proposed in [7] for or early fault detection of gearbox bearings. In [102] an adaptive network-based fuzzy inference system (ANFIS) is implemented for wind turbine condition monitoring using normal behavior models.

#### 8.3.2.1 **Practical Constraints**

**Data Availability** The data-driven methods presented above are all based on the assumption that data containing the degradation process is available, so they are



complementary to the physical model-based methods that only require data from the normal operation of the system, used for parameter identification, and a physical knowledge of the system. These two types of approaches are complementary, covering thus a wide field of application.

**Properties of the Generated HIs** Unlike HIs generated using physical model-based methods, the properties analysis (Monotonicity, Trendability, and Prognosability) of HIs generated by data-driven methods has been the subject of several research works [10, 24, 25, 52]. Monotonicity is related to the irreversibility assumption of degradation phenomena. Trendability is related to the degradation profile, i.e. it is related to the fact that the HI value is representative of the degradation value at any moment of the evolution of the degradation. Prognosability is related to the amplitude of the HI corresponding to the total failure; this property is respected when the threshold of HI corresponding to the total failure is constant. The metrics presented below are the ones proposed in [24] and [10], as their score is easily interpretable (between 0 and 1), where 1 indicates the most satisfactory and 0 the less satisfactory level of the specific HI property:

$$\text{Monotonicity} = \frac{1}{N} \left| \sum_{i=1}^N M_i \right| \quad (8.19)$$

where  $M_i$  is the monotonicity of a single run-to-failure trajectory given by:

$$M_i = \frac{n_i^+}{n_i - 1} - \frac{n_i^-}{n_i - 1}, i = 1, \dots, N \quad (8.20)$$

$n_i$  is the total number of observations in the  $i$ th run-to-failure trajectory and  $n_i^+$  ( $n_i^-$ ) the number of observations characterized by a positive (negative) first derivative.

$$\text{Trendability} = \min(|\text{corrcoef}_{ij}|), i, j = 1, \dots, N \quad (8.21)$$

$\text{corrcoef}_{ij}$  is the linear correlation coefficient between the  $i$ th and the  $j$ th run-to-failure trajectories. The computation of the correlation coefficient between two vectors requires that they are formed by the same number of patterns.

$$\text{Prognosability} = \exp \left( \frac{-\text{std}(HI_{fail})}{\text{mean}|HI_{start} - HI_{fail}|} \right) \quad (8.22)$$

where  $HI_{start}$  and  $HI_{fail}$  are the HI values at the beginning and end of the run-to-failure trajectories, respectively;  $\text{std}(HI_{fail})$  is standard deviation of the HI values at the end of the trajectories.  $\text{mean}|HI_{start}|$  and  $\text{mean}|HI_{fail}|$  are the average variation of the HI values between the beginning and the end of the trajectories, respectively.

Recent researches are directed towards the development of methods allowing the extraction of a set of features optimizing the scores of the three properties, as in [10], where the HIs identification is formulated as the problem of selecting the best combination of features to be used, and a multi-objective optimization that considers as objectives the metrics of Monotonicity, Trendability, and Prognosability. The proposed method is based on a binary differential evolution (BDE) algorithm for the multi-objective optimization.

The HI generation methods presented in this paper have been applied by the authors of this review paper on real cases. The details of the application of each method can be found in [14] and [35] for AR method, [33] for parity space method, [15] for algebraic methods, [36, 90] for PCA method, [89] for EMD method, [34, 89] for WD method, and [31, 32] for HI generated using machine learning methods. Table 8.1 summarizes the constraints of use, the advantages, and limitations of the methods presented above.

## 8.4 HI Trend Modeling for RUL Estimation

As illustrated in Fig. 8.3, the modeling approaches of HI trends for the estimation of RUL can be decomposed into three main families: physical approaches, data-driven approaches, and expert methods. Another classification is proposed in [2], where the RUL estimation methods are classified into: reliability based, similarity based, model based, and data-driven based approaches. The most used physical model form is the differential one, whose order and parameters are identified according to the physical knowledge and data available on the degradation process. The updating of the parameters makes it possible to compensate the modeling uncertainties and the adaptation of the model to changes in the degradation rate. Data-driven approaches are the most used and can be decomposed into five families: statistical models, stochastic models, deterministic models, probabilistic models, and machine learning model. The third family of trend modeling comprises those approaches that formalize the knowledge of industry experts through the tools of fuzzy logic and Bayes probabilities. Only data-driven methods able to include expert knowledge into the prediction models are presented in this paper.

### 8.4.1 Data-Driven Models

Four kinds of models are presented in this section:

- Stochastic models, especially continuous and discrete Markov processes.
- Probabilistic models, based on Bayes probability theory.

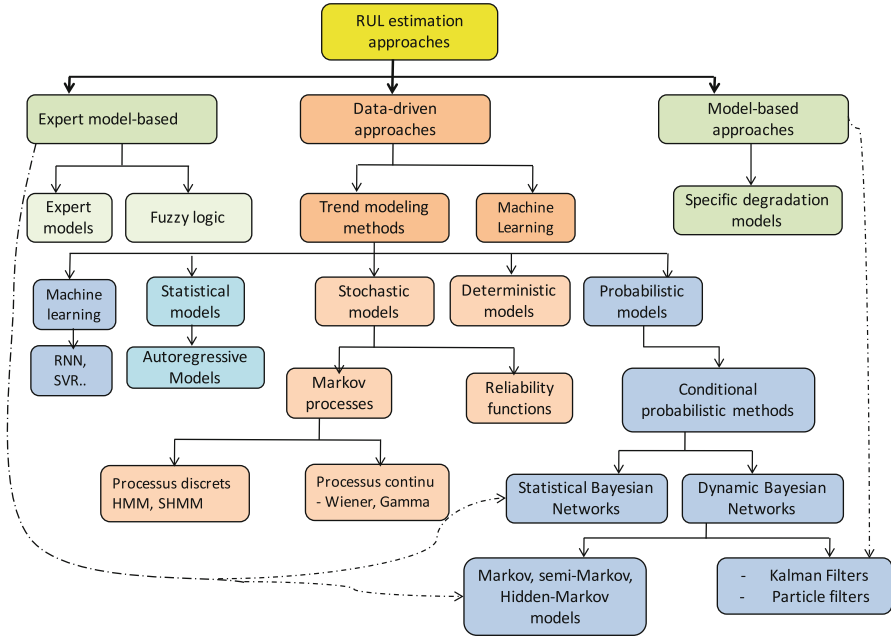


Fig. 8.3 Classification of RUL estimation approaches

- Statistical models, with a focus on the auto-regressive (AR) models and the auto-regressive moving average (ARMA) model, which is representative of these methods.
- Deterministic models, which are geometric models allowing to estimate the RUL as a deterministic variable.

**8.4.1.1 Stochastic Models**

Markov processes are the most used stochastic models for fault prognosis. These models describe processes without memory, where the probability of the future state  $X_n$  depends only on the current state  $X_{n-1}$  as shown in the following equation:

$$P(X_n = x_n | X_{n-1} = x_{n-1}, \dots, X_0 = x_0) = P(X_n = x_n | X_{n-1} = x_{n-1}) \quad (8.23)$$

$x_1, \dots, x_n$  are linked to the different states of the system. These processes can be divided into two categories: continuous Markov processes and discrete Markov processes.

**Continuous Markov Processes**

The most common continuous Markov processes used in the literature for prognosis are Wiener and Gamma processes. The hypothesis of independent increments leads

these processes to Markov properties because:  $X(t + \Delta t) - X(t)$  is independent of  $X(t)$  and  $X(t + \Delta t) = X(t) + (X(t + \Delta t) - X(t))$ ; the process  $\{X(t), t \geq 0\}$  is, therefore, a Markov process [91].

- Wiener processes are continuous Markov processes  $[X_t, t > 0]$ , with a drift parameter  $\mu$  and a variance parameter  $\sigma^2, \sigma > 0$ . They are well adapted to the modeling of degradation processes which vary over time with a Gaussian noise. These processes are described as follows [103]:

$$X_t = x_0 + \mu t + \sigma B(t) \tag{8.24}$$

where  $B(t)$  is the Brownian motion. The RUL  $H_{t_i}$  at a time  $t_i$  is defined as the time taken by the variable  $X_t$ , with  $t > t_i$  to reach a predefined failure threshold  $w$  such that:

$$H_{t_i} = \inf\{\Delta t_i : X_{t_i+\Delta t_i} \geq w | X_{t_i} < w\} \tag{8.25}$$

In the literature, the RUL is often given with a confidence interval, obtained by the calculation of a probability density function given by the following expression [26]:

$$f_{H_{t_i}}(h_{t_i}) = \frac{w - X_{t_i}}{\sqrt{2\pi t_i^3 \sigma^2}} \exp\left(-\frac{(w - X_{t_i} - \mu t_i)^2}{2t_i \sigma^2}\right) \tag{8.26}$$

Many work apply this process and its variants [95, 105, 110, 111, 113], and particularly [116] which proposes a Wiener process with an updated drift parameter  $\mu_t$ .

- Gamma process is the most appropriate for modeling a monotonic and gradual deterioration [91]. Reference [1] proposed to use it as a deterioration model randomly introduced over time [6, 23, 42, 62]. Mathematically, a random quantity  $X$  follows a Gamma distribution with a shape parameter  $\nu > 0$  and a scale parameter  $u > 0$  if its probability density function (PDF) is given as follows:

$$Ga(x|\nu, u) = \frac{u^\nu}{\Gamma(\nu)} x^{\nu-1} \exp(-ux) I_{(0,\infty)}(x) \tag{8.27}$$

where  $I_{(0,\infty)}(x) = 1$  for  $x \in (0, \infty)$  and  $I_{(0,\infty)}(x) = 0$  for  $x \notin (0, \infty)$ ,  $\Gamma(\nu) = \int_{z=0}^{\infty} z^{\nu-1} e^{-z} dz$  is the Gamma function for  $\nu > 0$ .

Given a non-decreasing function  $\nu(t)$ , Gamma process with the form function  $\Upsilon(t) > 0$  and the scale parameter  $u > 0$  is a continuous stochastic process with the following characteristics:

1.  $X(0) = 0$  with a probability of 1
2.  $X(\tau) - X(t) \sim Ga(\nu(\tau) - \nu(t), u)$  for all  $\tau > t \geq 0$
3.  $X(t)$  has independent increments.

Let  $X(t)$  be the deterioration at time  $t$ ,  $t \geq 0$ : the PDF of  $X(t)$  is as follows:

$$f_{X(t)}(x) = Ga(x|v(t), u) \quad (8.28)$$

HI's expectation and variance are as follows:

$$E(X(t)) = \frac{v(t)}{u}, \quad VAR(X(t)) = \frac{v(t)}{u^2} \quad (8.29)$$

A system is said to be faulty when its degradation reaches a predefined threshold  $S$ . From Eq. (8.28), the distribution of the failure time at time  $t$  is written as follows:

$$\begin{aligned} F(t) &= Pr\{T_S \leq t\} = Pr\{X(t) \geq S\} \\ &= \int_{x=S}^{\infty} f_{X(t)}(x)dx = \frac{\Gamma(v(t), Su)}{\Gamma(v(t))} \end{aligned} \quad (8.30)$$

where  $\Gamma(a, x) = \int_{z=x}^{\infty} z^{a-1} e^{-z} dz$  is the incomplete gamma function with  $x \geq 0$  et  $a > 0$ . The PDF of the failure time at time  $t$  is, thus:

$$f(t) = \frac{\partial}{\partial t} \left[ \frac{\Gamma(v(t), Su)}{\Gamma(v(t))} \right] \quad (8.31)$$

The mean failure time and the average RUL are given in the following equations:

$$\mathcal{T}_t = \int_{t=0}^{+\infty} t f(t) dt \quad (8.32)$$

$$RUL_t = \mathcal{T}_t - t \quad (8.33)$$

The two Markov processes presented above are widely used to model degradation, covering the majority of degradation profiles: linear and nonlinear, noisy and monotonous. However, these processes require the calculation of a HI  $X(t)$ , which estimates the current level of degradation of the system and which can be calculated using one of the methods presented in Sect. 8.3. The main limitation is related to the central property of Markov models, called memoryless assumption, which is a relatively strong assumption and thus may lead to strong approximation for real applications. To overcome this issue, a reliability model can be developed to consider the changes in the operating modes of the systems [89]. This model is based on two assumptions: (1) the future value of the HI is a function of the current state of the system, given by the present value of the HI, time, operating modes assumed known, and external noises supposed to follow a Gaussian law; (2) the HI is non-negative and monotonous.

Given these two hypotheses, the dynamics of the HI can be described as follows:

$$\Delta X_t = \frac{\beta t^{\beta-1}}{\eta^\beta} \exp(\gamma Z_t + \varepsilon) \quad (8.34)$$

where  $\beta > 0$  is the shape parameter of the model,  $\eta > 0$  is its scale parameter,  $\gamma = [\gamma_1, \dots, \gamma_m] \in \mathbb{R}^m$  is a vector of  $m$  elements, describing the influence of changes in operating modes  $Z_t = [Z_{1,t}, \dots, Z_{m,t}]$  on the degradation. The uncertainties of the model are represented by the variable  $\varepsilon$  assumed to follow a normal distribution  $N(0, Q)$ . The first term  $\beta t^{\beta-1}/\eta^\beta$  depends on time and means that  $\Delta X_t$  depends on the system aging.

The HI evolution  $X_t$  is defined as the accumulation of all segments  $\Delta X_t$ :

$$X_t = \sum_{\tau=0}^t \Delta X_\tau \quad (8.35)$$

Based on the linearity of mathematical expectation, the value of the mathematical expectation of  $X(t)$  is calculated as follows:

$$\begin{aligned} E[X_t] &= E \left[ \sum_{\tau=0}^t \Delta X_\tau \right] \\ &= \sum_{\tau=0}^t E[\Delta X_\tau] \\ &= \sum_{\tau=0}^t E \left[ \frac{\beta \tau^{\beta-1}}{\eta^\beta} \exp(\varepsilon) \right] \\ &= \sum_{\tau=0}^t \frac{\beta \tau^{\beta-1}}{\eta^\beta} E[\exp(\varepsilon)] \end{aligned} \quad (8.36)$$

$\varepsilon \sim N(0, Q)$  being a normal distribution variable,  $\exp(\varepsilon)$  is a log-normal distribution variable with mean value  $\exp(Q/2)$ :

$$\begin{aligned} E[X_t] &= \sum_{\tau=0}^t \frac{\beta \tau^{\beta-1}}{\eta^\beta} \exp(Q/2) \\ &= \exp(Q/2) \sum_{\tau=0}^t \frac{\beta \tau^{\beta-1}}{\eta^\beta} \end{aligned} \quad (8.37)$$

Assuming that  $k$  is the RUL and  $L > X_t$ , the predefined failure threshold, the RUL can be estimated as follows:

$$\begin{aligned}
 P(k|X_t < L) &= P(X_{t+k} < L|X_t < L) \\
 &= P\left(X_t + \sum_{i=t+1}^{t+k} \Delta X_i < L\right) \\
 &= P\left(\sum_{i=t+1}^{t+k} \Delta X_i < L - X_t\right) \\
 &= F_{\sum_{i=t+1}^{t+k} \Delta X_i}(L - X_t)
 \end{aligned} \tag{8.38}$$

$F_{\sum_{i=t+1}^{t+k} \Delta X_i}(L - X_t)$  is the distribution function ( $fr$ ) of the sum  $S_k = \sum_{i=t+1}^{t+k} \Delta X_i$  to the value  $L - X_t$ .

### Discrete Markov Processes

These methods are based on the principle of Markov chains for modeling processes that evolve through a finite number of states [37, 59]. By definition, it is assumed that the probability associated with each state, the probability associated with the transition from one state to another, and the probability of future failure can be estimated. The main property of Markov models is the assumption that the future state depends only on the current state, called conditionally independent or memoryless assumption. The most commonly used models for fault prognosis are the Hidden Markov models (*HMMs*), characterized by two parameters: (1) number of states of the system, (2) number of observations by state, and three probability distributions: (1) probability distribution of transitions between states, (2) probability distribution of observations, and (3) an initial probability distribution of states [12, 37, 83]. The HMM presents an appropriate mathematical model to describe the failure mechanisms of systems, which evolve in several degraded health states over the time prior to failure, as it can estimate the unobservable health states using observable sensor signals. The word “hidden” is related to the fact that the states are hidden from direct observations, so they only manifest themselves via a probabilistic behavior. HMM can exactly capture the characteristics of each state of the failure process, which is the basis of HMM prognosis [37]. These methods allow, thus, to model several operating conditions of the system and failure scenarios. However, their implementation requires a large amount of data and knowledge for learning, and the calculation intensity, which is proportional to the number of states, can become important for systems with several operating states. The three basic issues in HMMs implementation are: (1) Evaluation/Classification that represents what is the probability to get the model given an observation sequence, (2) Decoding/Recognition that represents what sequence of hidden states is the most optimal or is most probably the one that generates the given sequence of observations, and (3) Learning/Training that represents how to adjust the model parameters.

In the fault prognosis area, a widely used algorithm is the backward–forward algorithm, where the RUL at the time  $n$  can be defined as:

$$X_n = \inf \{x_n : Y_{n+x_n} = N / Y_n \neq N\} \tag{8.39}$$

where  $Y_n$  is the  $n$ th observation. The calculation of RUL using Markov chains usually involves the use of the phase-type distribution. As a result, the distribution and the expectation of the RUL are given as:

$$\Pr(X_n = k) = \alpha_n \tilde{P}^{k-1} (I - \tilde{P})e \quad (8.40)$$

$$E(X_n) = \alpha_n (I - \tilde{P})^{-1} e \quad (8.41)$$

where

$$P = \begin{pmatrix} \tilde{P} & P_0 \\ 0 & 1 \end{pmatrix} \text{ with } P_0 = (1 - \tilde{P})e \quad (8.42)$$

HMM is suitable for nonlinear systems. It can estimate the data distribution of normal operation with nonlinear and multimodal characteristics, assuming that predictable fault patterns are not available. It is applicable to non-stationary systems. It has been widely applied in real applications. The main reason is that the plant operation condition can be divided into several meaningful states, such as “Good,” “OK,” “Minor Defects Only,” “Maintenance Required,” “Unserviceable,” so that the state definition is closer to what is used in industry. It can be used for fault and degradation diagnosis on non-stationary signals and dynamical systems. It is appropriate for multi-failure modes [37, 59, 66].

The main limitation is related to the property of Markov models, i.e. the memoryless assumption. The health state visit time is assumed to follow an exponential distribution, which could be inappropriate for some cases. The transition probability among the system states in Markov models is often determined by empirical knowledge or by a large number of samples, which is not always available. A large amount of data is needed for accurate modeling [37, 59, 66].

The hidden semi-Markov model (HSMM) is an improved HMM, which overcomes the inherent limitation of assuming exponential distributions. Unlike the HMM, the HSMM does not follow the unrealistic Markov chain assumption and therefore provides more powerful modeling and analysis capability for real problems. In addition, the HSMM allows modeling the time duration of the hidden states and therefore is well suited for fault prognosis. A practical example is given in [28], where an approach for RUL estimation from heterogeneous fleet data under variable operating conditions is proposed in three steps:

- Identification of the degradation states of an homogeneous discrete-time finite-state semi-Markov model using unsupervised ensemble clustering approach.
- The maximum likelihood estimation (MLE) method and the Fisher information matrix (FIM) are used for parameter identification of the discrete Weibull distributions describing the transitions among the states and their uncertainties.
- The direct Monte Carlo (MC) simulation based on the degradation model is used to estimate the RUL of fleet equipment.



The proposed approach is applied to two case studies: heterogeneous fleets of aluminum electrolytic capacitors and turbofan engines. Another solution proposed in [53] to overcome the lack of knowledge on the condition monitoring is the online updating of the parameters of the degradation model formulated as a first-order Markov process. The originality of this work consists of the combination of Particle Filtering (PF) technique with a Kernel Smoothing (KS) one, for simultaneously estimating the degradation state and the unknown parameters in the degradation model, while significantly overcoming the problem of particle impoverishment.

#### 8.4.1.2 Conditional Probabilistic Models

These models are based on Bayes theorem, which describes relationship between the conditional and marginal probabilities of two stochastic events  $A$  and  $B$  as follows:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (8.43)$$

These methods describe the current state as a conditional probability function and, then, apply Bayes theorem to update the probability assessment of future behavior. The most used modeling tool is the Bayesian network, which is a probabilistic graphical model representing random variables in the form of an acyclic oriented graph. In the field of aeronautics, [39] uses the network with variables such as aircraft weight, landing speed, and brake operation to predict brake failure. In other research work, Bayesian networks are associated with the Kalman filter [22, 69] or particle filter [20, 93, 106] for failure prognosis.

#### 8.4.1.3 Statistical Models

The ARMA, the ARIMA (*Auto-Regressive Integrated Moving Average*), and the ARMAX (*Auto-Regressive Moving Average exogenous inputs*) models, initially used for time series prediction, have been used to estimate the RUL by considering the future value of the degradation as a linear function of system inputs, past observations, and random noise. To show how these methods are used for prognostics, let us take the example of the ARMA model. A time series  $\{x_t | t = 1, 2, \dots\}$  is generated by an ARMA model  $(p, q)$  as follows:

$$x_t = \sum_{i=1}^p \phi_i x_{t-i} + \sum_{j=0}^q \theta_j \epsilon_{t-j}, \quad (\theta_0 = 1) \quad (8.44)$$

where  $x_t$  is a series at the instant  $t$ ,  $p$  and  $q$  are non-zero integers,  $p$  is the order of the auto-regressive part,  $q$  is the order of the moving average part,  $\{\epsilon_t\}$  indicates

the noise series,  $\{\phi_i, i = 1, \dots, p\}$  et  $\{\theta_j, j = 1, \dots, q\}$  are the parameters to be estimated.

To use this model for fault prognosis, the variable  $x_t$  is considered as the HI which represents the system condition state and the failure threshold  $D$  is supposed known. The RUL at instant  $t$  of the system is calculated by the following equation:

$$rul_t = \min\{\Delta t : x_{t+\Delta t} \geq D | x_t < D\} \quad (8.45)$$

Yan et al. [118] have used ARMA model for fault prognosis. An ARMA model is incorporated in a software for data fusion and prognosis [65]. An extension of an ARMA model by using bootstrap forecasting is proposed in [117].

The use of these models is simple for prognosis. However, they assume that the future state of the system is a linear function of the system inputs, past observations, and noise, which is not often the case in reality. Moreover, their results are sensitive to the initial conditions, thus leading to an accumulation of systematic errors in the prediction.

#### 8.4.1.4 Deterministic Models

This approach is supervised by the calculation of the Euclidean distance ( $d$ ) between the actual status of the system, given by the actual HIs values and the faulty HIs identified offline. The degradation speed ( $v$ ), which indicates how the degradation moves from the normal operation to the faulty one, is used to compute the RUL as follows:

$$rul(t) = \left| \frac{d(t)}{v(t)} \right| \quad (8.46)$$

To compute the distance  $d(t)$  between the  $n$  HIs in real time operation and the barycenter of the identified faulty operating cluster ( $C_f(c_1, c_2, \dots, c_n)$ ), the following Euclid metric is considered:

$$d(t) = \sqrt{(r_1(t) - c_1)^2 + (r_2(t) - c_2)^2 + \dots + (r_n(t) - c_n)^2} \quad (8.47)$$

where  $r_1(t), r_2(t), \dots, r_n(t)$  is a set of HIs defining the monitoring space. This set can be generated using one of the HI generation methods presented in the previous section.

The numerical differentiation of the distance variable  $d$  is taken to compute the degradation speed  $v$ :

$$v(t) = \frac{d(t + \Delta T) - d(t)}{\Delta T} \quad (8.48)$$

where  $\Delta T$  is the sampling time for speed degradation computation. whose value is chosen so that the noise is not amplified.

The main interest of this method is the fact that no knowledge is required on the tendency or the pattern of degradation. It is accurate when the faulty operation is clearly identified and successfully applied to a wind turbine system [36] for RUL estimation.

Other deterministic models are used in the literature, especially when the degradation profile is known. The models are identified by using the fitting methods applied on the available profile of the degradation. Linear, exponential, and polynomial models are the most used [120].

#### 8.4.1.5 Learning Techniques

Learning techniques are widely used in the literature for trend modeling in the field of failure prognosis. These regression models, such as neural networks and support vector regression (SVR), are scalable and able to accommodate nonlinear dynamics, but require a large amount of data for learning. For unsupervised learning cases, an example of using SVR for RUL prediction is proposed in [32] for failure prognosis of embedded systems. The prediction is realized using a SVR at a step of the evolution of the health index. The SVR expression is given as follows:

$$\hat{HI}(k+1) = \sum_{[i=(m-1)\tau+1]}^{N-1} \alpha_i^* K(HI(i), HI(k)) + b^* \quad (8.49)$$

where  $\alpha_i^*$  are Lagrange multipliers and  $\tau$  the delay. In this work, the standard SVR toolbox is used without making any special changes to the prediction of temporal overlays. The free parameters,  $C$ ,  $\epsilon$ , the size of the kernel (Gaussian), and the dipping dimension  $m$  are selected from a comprehensive search in the parameter space to optimize the performance of the prediction on the validation set. The available  $N$  observations are therefore shared between two sets of training and validation of respective sizes  $Ne$  and  $Nv$ . Values for which the prediction error at a step on the validation set is minimal are retained for the final prediction. Once the parameters are fixed, the prediction is made using all the  $N$  observations available. The predictions at several steps, i.e. for the values ( $k \geq N + 1$ ), are realized by the ratio of the prediction at one step, using the estimated vectors  $\hat{HI}(k)$  at the previous iterations and not the observations themselves.

In addition to machine learning techniques, deep learning techniques like long short-term memory (LSTM), which can remember information for long periods of time, are used for trend modeling and RUL prediction. An application case is proposed in [121], where a long short-term memory recurrent neural network is used for RUL prediction of lithium-ion batteries.

### 8.4.2 Physical Models

RUL estimation based on a physical model consists of considering that degradation follows a parametric trend, which can take one of the following ordinary differential equations (ODE):

$$\begin{aligned}
 \dot{F} &= \beta_1 F \\
 \dot{F} &= \beta_2 F^2 \\
 \dot{F} &= \beta_3 \dot{F} + F + \beta_4 \\
 \ddot{F} &= \beta_5 \dot{F}^2 + \beta_6 F + \beta_7
 \end{aligned} \tag{8.50}$$

where  $F$  is the fault component value describing the degradation and  $\beta_i$  ( $i = 1 \dots 7$ ) represent the degradation model coefficients which are identified on-line by using the least square method [30] or particle filter [29, 58]. For example, the RUL associated with the degradation model of the form  $\dot{F} = \beta_1 F$  is given by:

$$\text{RUL} = \frac{\ln\left(\frac{1-th}{1-HI_i}\right)}{\beta} - NT_s \tag{8.51}$$

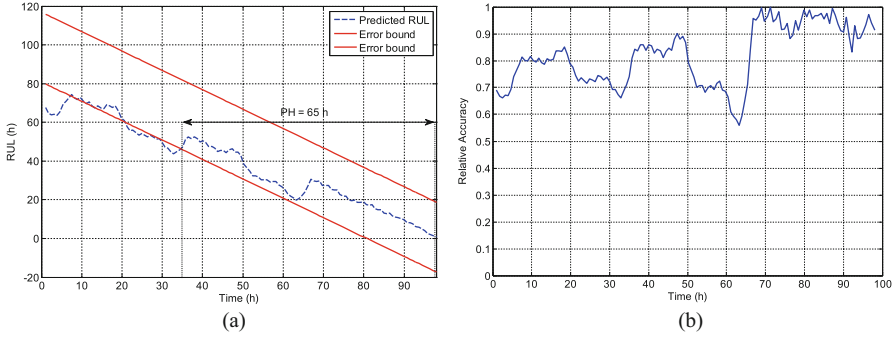
where  $N$  represents the sample data,  $T_s$  is the sampling time, and  $th$  is the failure threshold.

Other trend modeling approaches for RUL estimation can be found in the literature, such as [99] where the RUL estimation is treated as an uncertainty propagation problem, [103] where the review is focused on statistical data-driven approaches, relying only on available past observed data and statistical models, [5] which provides practical options for prognostics so that beginners can select appropriate methods for their fields of application.

### 8.4.3 Practical Constraints

The purpose of the RUL estimate is to give to the maintenance experts' two pieces of information: the first is that the system will undergo the occurrence of a total failure, the second is to give a sufficient time horizon for the maintenance experts in order to plan a maintenance strategy. A metric proposed in [101], called Prognosis Horizon ( $PH$ ), is used to evaluate this time horizon in a confidence interval which can be defined by the user. Prognosis Horizon ( $PH$ ) ranges within  $[0, \infty[$  and is calculated as follows:

$$PH(i) = EoP - CT \tag{8.52}$$



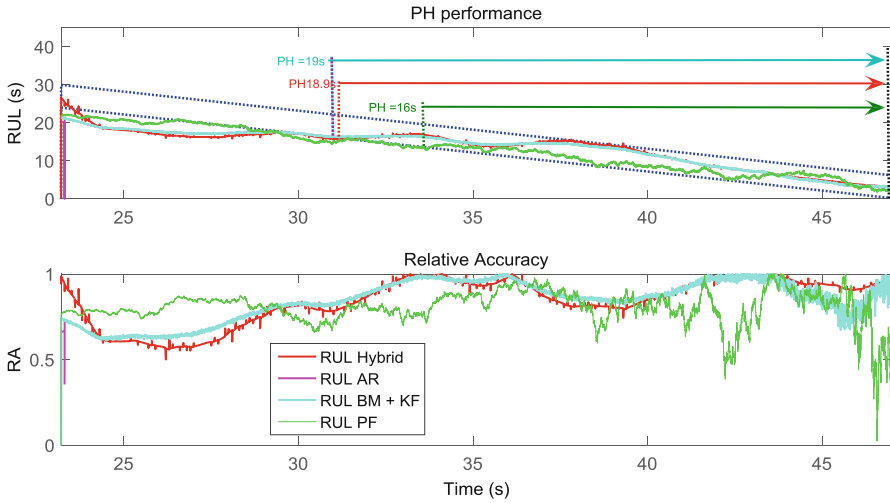
**Fig. 8.4** Illustration of the PH RA metrics calculated for a wind turbine system in presence of an unbalance fault caused by a progressive deformation of the blade

It represents the difference between the Current Time index (*CT*) and the End of Prediction time index (*EoP*), obtained when the prediction crosses the failure threshold. A practical example of the *PH* calculation is illustrated in Fig. 8.4a which represents the *PH* calculated for a wind turbine system in presence of an unbalance fault caused by a progressive deformation of the blade. In this practical example, the HIs are generated using the PCA method and the trend modeling for RUL estimation is performed using a deterministic model, based on Euclidean distance [36]. Figure 8.4a shows that the obtained PH, by considering a confidence interval of 18%, is equal to 65 h. The maintenance expert can, then, make a decision, whether or not this HP is sufficient to plan a maintenance strategy in good conditions. If it judges that this PH is insufficient it is possible to increase it but by increasing the confidence interval.

To give the user an easily interpretable measurement tool of the confidence that can be given to the PH metric, another metric is proposed in [101] where the accuracy is quantified according to the real RUL. This metric is called relative accuracy (*RA*) and expressed as follows:

$$RA(t) = 1 - \frac{|RUL^*(t) - RUL(t)|}{RUL^*(t)} \tag{8.53}$$

*RUL\** is the real RUL. The range score of the RA metric is between [0, 1], and the best score is close to 1. A practical example of the results of this metric applied to the estimated RUL before the total degradation of the wind turbine system is given in Fig. 8.4b. It shows that the *RA* is greater than 0.7 in average over the PH, but has a great variability. All these measures will enable maintenance experts to assess the risks and make the right decisions for the maintenance of the systems.



**Fig. 8.5** Practical illustration of the use of PH and RA metrics for the performance comparison of the RUL estimation methods

**Table 8.2** Performance comparison of the considered RUL estimation methods by universal metrics

	PH	RAmin	RAmax	RAmean
RUL Hybrid	18.9 s	0.5	1	0.8509
RUL AR	19 s	0.6207	1	0.8504
RUL BM + KF	19 s	0.2702	1	0.8487
RUL PF	16 s	0	1	0.7155

In addition, these two metrics can be used to compare the performances of different methods in a given context. A practical example is given in Fig. 8.5, where the PH and RA are calculated for four prognostic methods applied to the RUL prediction on the mechanical transmission system presented in [14]. In this paper, the analytical redundancy method is used for HIs generation and four trend modeling methods are applied for RUL estimation: an auto-regressive (AR) model, whose parameters are estimated using the least square methods, an updated Wiener process, whose drift parameter is updated using a Kalman filter, a first-order differential model whose parameter is updated using a particle filter (PF), and a deterministic model based on the calculation of the Euclidean distance [14]. The performance results of the considered RUL estimation methods are given in Table 8.2, which shows that the AR model and the Wiener model have the largest PH, thus giving the user more time to react, whereas the Wiener model is less stable since it presents a greater variability in its RA.

It would be also better to evaluate the RUL prediction accuracy using—accuracy and cumulative relative accuracy (CRA) proposed in [100, 101].

## 8.5 Discussion and Future Challenges

### 8.5.1 Discussion

After analyzing the methods described above, one can see that discrete Markov models are the most complex to implement because they require expert knowledge and rich databases on the previous operation of the system and its failures; the uncertainty brought by expert knowledge is often taken into account using fuzzy logic. The memoryless assumption, which is the main property of discrete Markov processes, and that they also share with continuous processes, is a major limitation in the use of these processes for the estimation of the RUL. To overcome it, the hidden semi-Markov models (HSMMs), that do not follow the unrealistic Markov chain assumption, to provide more powerful modeling and analysis capabilities for real fault prognosis problems.

Continuous Markov processes, especially the Wiener and Gamma processes, are widely used in the literature as they are easy to implement and are well adapted to modeling the progressive dynamics of degradation phenomena. The updating of the parameters by increasingly powerful techniques such as maximum likelihood, the Kalman filter, and the particle filter makes it possible to adapt the estimation to the possible changes in the rate of degradation and provide in part a solution to the limit related to the memoryless assumption. Research works have gone even further in modeling, drawing on the Cox model, by proposing a reliability function that takes into account the covariates representing changes in the operating modes of a system: the limit of this model is related to the fact that the evolution of the covariates must be known beforehand, which is difficult to obtain on systems such as energy and transport systems where covariates are environment variables that are not controlled.

The representation of degradation processes by adaptive differential models makes it possible to take into account the physical knowledge available on these phenomena for the choice of the order of the models. Continuous parameter updating adapts to the change in degradation rate, and structured and unstructured uncertainties are taken into account by the generation of normal operating thresholds and total failure thresholds. The main limitation of this type of model is related to problems of amplification of the noises generated by the successive derivations of the outputs, as well as the lack of physical knowledge about the degradation processes, which generally leads to an arbitrary choice of the order of the model. Geometric models are efficient and accurate, but require complex classification work to identify clusters, using the physical model for generating the useful databases for learning.

The choice of the HI modeling approach depends on the context of use, the complexity of the system, and the information available on its previous operating modes, especially for the definition of the structure of the model as well as for the identification of its parameters. Table 8.3 summarizes a set of criteria, not

**Table 8.3** Trend modeling approaches for RUL estimation

	Choice of model structure	Estimation of model parameters	Data in degraded operation	Expert knowledge	Physical knowledge	Accuracy of estimation	Adaptation to setpoints change
<i>Discrete Markov processes</i>							
HMM	Expert knowledge	By training	Necessary	Necessary	Not necessary	Depends on prior knowledge	No
HSMM	Expert knowledge	By training	Necessary	Necessary	Not necessary		No
<i>Continuous Markov processes</i>							
Wiener process	Predefined	Maximum likelihood Kalman filter particle filter	Required to estimate parameters	Not necessary	Not necessary	More accurate when the rate of degradation is not very variable	Not very suitable
Updated Wiener process	Predefined	Maximum likelihood Kalman filter particle filter	Required to estimate parameters	Not necessary	Not necessary	Better adapted to frequent changes in degradation rate	Yes by updating the settings
Gamma process	Predefined	Maximum likelihood Kalman filter particle filter	Required to estimate parameters	Not necessary	Not necessary	More precise on processes of monotonous degradation and not very noisy	Yes by updating the current state



<i>Reliability functions</i>							
<i>Modified Cox models</i>	Predefined	Maximum likelihood Kalman filter particle filter	Required to estimate parameters	Required to estimate parameters	Knowledge of covariates is necessary	More precise on monotonous degradation and not very noisy	Yes, is the most suitable model
<i>Physical models</i>							
<i>Differential models</i>	Chosen or deduced from physical knowledge	Least squares maximum likelihood Kalman filter particle filter	Not necessary if physical knowledge is available	Not necessary if physical knowledge is available	Necessary to choose the model order and parameters initialization	Accurate when physical knowledge is sufficient	Yes the model can consider the system setpoints
<i>Geometric models</i>							
<i>Distance model</i>	Depends on the size of the monitoring space	Not necessary	Not necessary	Not necessary	Not necessary	More accurate when classification is accurate	Yes By updating the speed and the direction of the degradation

exhaustive, that can be used as a basis for choosing the HI modeling methodology for the estimation of RUL.

### 8.5.2 Future Challenges

The general formulation of the Remaining Useful Life (RUL) of a system can be expressed in a general form, as a function of the time  $t$ , the current condition monitoring  $CM(t)$  and the current health state  $HS(t)$  (Eq. (8.54) below). However, in practice, the state of degradation is neither available nor measurable in the majority of cases, health indices must be deduced from the physical knowledge, expert knowledge, and available measurements [90]:

$$RUL = g(t, HS(t), CM(t)) \quad (8.54)$$

- The use of the techniques initially developed for fault detection and isolation (*FDI*) to estimate the health state (*HS*) of the system is a good idea, but in the context of failure prognosis, the early detection becomes a major issue, as it is necessary to estimate the RUL well in advance to allow maintenance operators to plan their maintenance interventions. FDI techniques treat uncertainties in a probabilistic or deterministic manner to generate thresholds that provide a better compromise between false alarms and non-detections. In the context of failure prognosis, it is necessary to take into account also the Prognosis Horizon [100, 101]. In addition, the problems related to the occurrence of multiple faults, their interaction, their effect on the HS of the system remain.
- Condition monitoring ( $CM(t)$ ), necessary for RUL estimation, is not always known especially in systems operating in a randomly variable environment, such as offshore wind turbines and transportation systems. The solution proposed in the literature consists in taking account of the known or controlled CMs by using, for example, the modified Cox model [89], and in compensating the lack of knowledge about the unknown CMs by an online update of the model parameters. This solution is effective in some application cases, but in the case where these CMs vary strongly and continuously, the RUL estimate will change considerably and continuously, which will prevent the use of the estimated RUL for planning the maintenance. The solution may be to associate the RUL estimate with risk analysis methods taking into account several operating and degradation scenarios [123, 125].

## 8.6 Conclusion

This review of horizontal approaches for RUL estimation has highlighted the diversity of methods proposed in the literature as well as their formal description and context of use. The analysis shows that the two major stages of the procedure

can be synthesized independently, but that the context of use, the complexity of the systems as well as the history of available data and expertise are common elements that govern the relevance of the choice of the HI generation methods and the method of modeling its tendency for the estimation of the RUL. The diversity of methods for generating health indices can cover a wide range of applications. The PCA makes it possible to simultaneously reduce the size of the data and generate health indices from large databases with linear, bilinear, or nonlinear dependencies. When the instrumentation of the systems is not rich, the statistical, frequency, and time-frequency attributes can be extracted from the signals and, then, analyzed to make them HIs for the estimation of the RUL. When physical knowledge is relevant enough to take modeling assumptions, build and validate physical models, these latter are, then, associated with health index generation methods such as analytical redundancy, observers, and parameter estimation. In this paper, we have also presented the trend modeling methods that are the simplest to implement and that are adapted to the physical properties of degradation processes, such as the progressive aspect and the influence of the environment and operating modes of the systems. The choice of the model depends on the context of use, the physical knowledge available on degradation processes, the expert feedback, and available data. The RUL can be presented as a stochastic, probabilistic, or deterministic variable.

## References

1. M. Abdel-Hameed, A gamma wear process. *IEEE Trans. Reliab.* **24**(2), 152–153 (1975)
2. K. Abid, M.S. Mouchaweh, L. Cornez, Fault prognostics for the predictive maintenance of wind turbines: state of the art, in *In Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 113–125 (2018)
3. D. Adams, M. Nataraju, A nonlinear dynamical systems framework for structural diagnosis and prognosis. *Int. J. Eng. Sci.* **40**(17), 1919–1941 (2002)
4. J. Altmann, J. Mathew, Multiple band-pass autoregressive demodulation for rolling-element bearing fault diagnosis. *Mech. Syst. Signal Process.* **15**(5), 963–977 (2001)
5. D. An, N.H. Kim, J.H. Choi, Practical options for selecting data-driven or physics-based prognostics algorithms with reviews. *Reliab. Eng. Syst. Saf.* **133**, 223–236 (2015)
6. J. Bakker, J.V. Noortwijk, Inspection validation model for life-cycle analysis, in *Proceedings of the 2nd International Conference on Bridge Maintenance, Safety and Management (IABMAS)*, pp. 18–22 (2004)
7. P. Bangalore, L.B. Tjernberg, An artificial neural network approach for early fault detection of gearbox bearings. *IEEE Trans. Smart Grid* **6**(2), 980–987 (2015)
8. D. Banjevic, Remaining useful life in theory and practice. *Metrika* **69**, 337–349 (2009)
9. D. Banjevic, A.K.S. Jardine, Calculation of reliability function and remaining useful life for a Markov failure time process. *IMA J. Manag. Math.* **17**, 115–130 (2006)
10. P. Baraldi, G. Bonfanti, E. Zio, Differential evolution-based multi-objective optimization for the definition of a health indicator for fault diagnostics and prognostics. *Mech. Syst. Signal Process.* **102**, 382–400 (2018)
11. J. Barbot, M. Fliess, T. Floquet, An algebraic framework for the design of nonlinear observers with unknown inputs, in *46th IEEE Conference on Decision and Control*, pp. 384–389 (2007)

12. E. Bechhoefer, A. Bernhard, D. He, P. Banerjee, Use of hidden semi-Markov models in the prognostics of shaft failure, in *Annual forum proceedings – American Helicopter Society*, pp. 1330–1335 (2006)
13. B. Bellali, A. Hazzab, I.K. Bousserhane, D. Lefebvre, Parameter estimation for fault diagnosis in nonlinear systems by ANFIS. *Procedia Eng.* **29**, 2016–2021 (2012)
14. S. Benmoussa, M.A. Djeziri, Remaining useful life estimation without needing for prior knowledge of the degradation features. *IEEE IET Sci. Meas. Technol.* **11**(8), 1071–1078 (2017)
15. S. Benmoussa, B.O. Bouamama, R. Merzouki, Bond graph approach for plant fault detection and isolation: application to intelligent autonomous vehicle. *IEEE Trans. Autom. Sci. Eng.* **11**(2), 585–593 (2014)
16. S. Bhat, D. Bernstein, Geometric homogeneity with applications to finite-time stability. *Math. Control Signals Syst.* **17**, 101–127 (2000)
17. S. Bhat, D. Bernstein, Finite-time stability of continuous autonomous systems. *SIAM J. Control. Optim.* **38**(3), 751–766 (2005)
18. M. Blanke, M. Kinnaert, J. Lunze, M. Staroswiecki, *Diagnosis and Fault-Tolerant Control* (Springer, Berlin, 2006)
19. C. Byington, M. Roemer, T. Galie, Prognostic enhancements to diagnostic systems for improved condition-based maintenance, in *In Proceedings of IEEE Aerospace Conference* (2002)
20. F. Cadini, E. Zio, D. Avram, Model-based Monte Carlo state estimation for condition-based component replacement. *Reliab. Eng. Syst. Saf.* **94**(3), 752–758 (2009)
21. J. Chen, R. Patton, *Robust Model-Based Fault Diagnosis for Dynamic Systems* (Springer, Berlin, 1999)
22. A. Christer, W. Wang, J. Sharp, A state space condition monitoring model for furnace erosion prediction and replacement. *Eur. J. Oper. Res.* **101**(1), 1–14 (1997)
23. E. Cinlar, E. Osman, Z. Bazant, Stochastic process for extrapolating concrete creep. *J. Eng. Mech. Div.* **103**(6), 1069–1088 (1977)
24. J. Coble, Merging data sources to predict remaining useful life – an automated method to identify prognostic parameters. Ph.D. Diss., University of Tennessee (2010)
25. J. Coble, J. Hines, Identifying optimal prognostic parameters from data: a genetic algorithm approach, in *Annual Conference of the Prognostics and Health Management Society* (2009)
26. D. Cox, H. Miller, *The Theory of Stochastic Processes*, vol. 134 (CRC Press, Boca Raton, 1977)
27. J. Cruz-Victoria, R. Martinez-Guerra, J. Rincon-Pasaye, On nonlinear systems diagnosis using differential and algebraic methods. *J. Frankl. Inst.* **345**, 102–117 (2008)
28. S.A. Dahidi, F.D. Maio, P. Baraldi, E. Zio, Remaining useful life estimation in heterogeneous fleets working under variable operating conditions. *Reliab. Eng. Syst. Saf.* **3156**, 109–124 (2016)
29. M.J. Daigle, K. Goebel, A model-based prognostics approach applied to pneumatic valves. *Int. J. Prognosis Health Manage.* **2**, 84–99 (2011)
30. W. Danwei, Y. Ming, L. Chang, A. Shai, *Model-based Health Monitoring of Hybrid Systems* (Springer, Berlin, 2013)
31. O. Djedidi, M.A. Djeziri, N. MSirdi, Data-driven approach for feature drift detection in embedded electronic devices, in *IFAC Proceeding of the IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes (SAFEPROCESS)*, pp. 964–969 (2018)
32. O. Djedidi, M.A. Djeziri, S. Benmoussa, Failure prognosis of embedded systems based on temperature drift assessment, in *Proceedings of the International Conference on Integrated Modeling and Analysis in Applied Control and Automation*, pp. 11–16 (2019)
33. M.A. Djeziri, A. Aitouch, B.O. Bouamama, Sensor fault detection of energetic system using modified parity space approach, in *Proceeding of the IEEE Control and Decision Conference*, pp. 2578–2583 (2007)

34. M.A. Djeziri, S. Benmoussa, M. Ouladsine, B.O. Bouamama, Wavelet decomposition applied to fluid leak detection and isolation in presence of disturbances, in *IEEE Proceeding of the 18th Mediterranean Conference on Control and Automation (MED)*, pp. 104–109 (2012)
35. M. Djeziri, S. Benmoussa, L. Nguyen, N. MSirdi, Fault prognosis based on physical and stochastic models, in *Proceeding of the 2016 European Control Conference* (2016)
36. M.A. Djeziri, S. Benmoussa, R. Sanshez, Hybrid method for remaining useful life prediction in wind turbine systems. *Renew. Energy* (2017). <https://doi.org/10.1016/j.renene.2017.05.020>
37. M. Dong, D. He, A segmental hidden semi-Markov model (HSMM)-based diagnostics and prognostics framework and methodology. *Mech. Syst. Signal Process.* **21**, 2248–2266 (2007)
38. R. Engel, G. Kreisselmeier, A continuous-time observer which converges in finite time. *IEEE Trans. Autom. Control* **47**(7), 1202–1204 (2002)
39. S. Ferreira, A. Arnaiz, B. Sierra, I. Irigoien, Application of Bayesian networks in prognostics for a new integrated vehicle health management concept. *Expert Syst. Appl.* **39**(7), 6402–6418 (2012)
40. M. Fliess, Some basic structural properties of generalized linear systems. *Syst. Control Lett.* **15**(5), 391–396 (1990)
41. M. Fliess, C. Join, H. Sira-Ramirez, Robust residual generation for linear fault diagnosis: an algebraic setting with examples. *Int. J. Control.* **77**, 1223–1242 (2004)
42. D. Frangopol, M. Kallen, J.V. Noortwijk, Probabilistic models for life cycle performance of deteriorating structures: review and future directions. *Prog. Struct. Eng. Mater.* **6**(4), 197–212 (2004)
43. L. Fridman, Y. Shtessel, C. Edwards, X.G. Yan, Higher order sliding-mode observer for state estimation and input reconstruction in nonlinear systems. *Int. J. Robust Nonlinear Control* **18**, 339–412 (2007)
44. D. Galar, U. Kumar, J. Lee, W. Zhao, Remaining useful life estimation using time trajectory tracking and support vector machines. *J. Phys. Conf. Ser.* **364**(1), 012063 (2012)
45. Z. Gao, C. Cecati, S.X. Ding, A survey of fault diagnosis and fault-tolerant techniques; part I: fault diagnosis with model-based and signal-based approaches. *IEEE Trans. Ind. Electron.* **62**(6), 3757–3767 (2015)
46. Z. Gao, C. Cecati, S.X. Ding, A survey of fault diagnosis and fault-tolerant techniques; part II: fault diagnosis with knowledge-based and hybrid/active approaches. *IEEE Trans. Ind. Electron.* **62**(6), 3768–3774 (2015)
47. M.C. Garcia, M.A. Sanz-Bobi, J.D. Pico, SIMAP: intelligent system for predictive maintenance: application to the health condition monitoring of a wind turbine gearbox. *Comput. Ind.* **57**(6), 552–568 (2006)
48. J. Gertler, Fault detection and isolation using parity relations. *Control. Eng. Pract.* **5**, 653–661 (1997)
49. J.C. Gomez-Mancilla, L. Palacios-Pineda, V. Nosov, Software package evaluation for Lyapunov exponent and others features of signals evaluating the condition monitoring performance on nonlinear dynamic system. *J. Energy Power Eng.* **9**(5), 443–551 (2015)
50. D. Gucik-Derigny, Contribution au pronostic des systemes à base de modles : theorie et application. Ph.D. Thesis, Universite Paul Cezanne Aix-Marseille (2011)
51. D. He, R. Li, J. Zhu, Plastic bearing fault diagnosis based on a two-step data mining approach. *IEEE Trans. Ind. Electron.* **60**(8), 3429–3440 (2013)
52. W. He, Q. Miao, M. Azarian, M. Pecht, Health monitoring of cooling fan bearings based on wavelet filter. *Mech. Syst. Signal Process* **64–65**, 149–161 (2015)
53. Y. Hu, P. Baraldi, F.D. Maio, E. Zio, A particle filtering and kernel smoothing-based approach for new design component prognostics. *Reliab. Eng. Syst. Saf.* **134**, 19–31 (2015)
54. N. Huang, Z. Shen, S. Long, M. Wu, H. Shih, Q. Zheng, H. Liu, The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis. *Proc. R. Soc. London A Math. Phys. Eng. Sci.* **454**(1971), 903–995 (1998)
55. R. Isermann, *Fault Diagnosis Systems* (Springer, Berlin, 2006)
56. M. James, Finite time observers and observability, in *Proceedings of IEEE Conference on Decision and Control* (1990)

57. A. Jardine, D. Lin, D. Banjevic, A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mech. Syst. Signal Process.* **20**(7), 1483–1510 (2006)
58. M. Jha, G. Dauphin-Tanguy, B. Ould-Bouamama, Particle filter based hybrid prognostics for health monitoring of uncertain systems in bond graph framework. *Mech. Syst. Signal Process.* **75**, 301–329 (2016)
59. M. Kan, A. Tan, J. Mathew, A review on prognostic techniques for non-stationary and non-linear rotating systems. *Mech. Syst. Signal Process.* **62**, 1–20 (2013)
60. D.C. Karnopp, D. Margolis, R. Rosenberg, *Systems Dynamics: A Unified Approach*. (John Wiley, Hoboken, 1990)
61. Y. Khellil, G. Graton, M.A. Djeziri, M. Ouladsine, Fault detection and isolation in marine diesel engines, a generic methodology, in *IFAC Proceeding of the 8th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes (SAFEPROCESS)*, pp. 964–969 (2012)
62. J. Lawless, M. Crowder, Covariates and random effects in a gamma process model with application to degradation and failure. *Lifetime Data Anal.* **10**(3), 213–227 (2004)
63. B.P. Leao, T. Yoneyama, G. Rocha, K.T. Fitzgibbon, Prognostics performance metrics and their relation to requirements, design, verification and cost-benefit, in *2008 International Conference on Prognostics and Health Management*, pp. 1–8 (2008)
64. M. Lebold, M. Thurston, Open standards for condition-based maintenance and prognostic systems, in *Maintenance and Reliability Conference*, pp. 6–9 (2001)
65. J. Lee, J. Ni, D. Djurdjanovic, H. Qiu, H. Liao, Intelligent prognostics tools and e-maintenance. *Comput. Ind.* **57**(6), 476–489 (2006)
66. J. Lee, F. Wu, W. Zhao, M. Ghaffari, L. Liao, D. Siegel, Prognostics and health management design for rotary machinery systems – reviews, methodology and applications. *Mech. Syst. Signal Process.* **42**(1), 314–334 (2014)
67. C. Li, H. Lee, Gear fatigue crack prognosis using embedded model, gear dynamic model and fracture mechanics. *Mech. Syst. Signal Process.* **19**(04), 836–846 (2005)
68. G. Li, S.J. Qin, Y.J.D. Zhou, Reconstruction based fault prognosis for continuous processes. *Control. Eng. Pract.* **18**(10), 1211–1219 (2010)
69. C. Lim, D. Mba, Switching Kalman filter for failure prognostic. *Mech. Syst. Signal Process.* **52**, 426–435 (2015)
70. J. Lin, M. Zuo, Gearbox fault diagnosis using adaptive wavelet filter. *Mech. Syst. Signal Process.* **17**(6), 1259–1269 (2003)
71. F. Liu, Synthèse d'observateurs à entrées inconnues pour les systèmes non linéaires. Ph.D. Thesis, Université de Basse-Normandie (2007)
72. J. Luo, M. Namburu, K. Pattipati, L. Qiao, M. Kawamoto, S. Chigusa, Model-based prognostic techniques, in *Proceedings of IEEE Systems Readiness Technology Conference* (2003), pp. 330–340
73. M. Luschen, Derivation and application of nonlinear analytical redundancy techniques with applications to robotics. Ph.D. Thesis, Houston, TX (2001)
74. F.D. Maio, J. Hu, P. Tse, M. Pecht, K. Tsui, E. Zio, Ensemble-approaches for clustering health status of oil sand pumps. *Expert Syst. Appl.* **39**(5), 4847–4859 (2012)
75. F.D. Maio, F. Antonello, E. Zio, Condition-based probabilistic safety assessment of a spontaneous steam generator tube rupture accident scenario. *Nucl. Eng. Des.* **326**, 41–54 (2018)
76. S. Mallat, A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **11**(7), 674–693 (1989)
77. C. Marcelo, J.P. Fossatti, J.I. Terra, Fault diagnosis of induction motors based on FFT, in *Fourier Transform-Signal Processing*. InTech (2012)
78. A. Mathur, K. Cavanaugh, K. Pattipati, P. Willett, T. Galie, Reasoning and modeling systems in diagnosis and prognosis, in *Aerospace/Defense Sensing, Simulation, and Controls, International Society for Optics and Photonics* (2001), pp. 194–203

79. T. Menard, E. Moulay, W. Perruquetti, A global high-gain finite-time observer. *IEEE Trans. Autom. Control* **55**(6), 1500–1506 (2010)
80. P. Menold, Finite time and asymptotic time state estimation for linear and nonlinear systems. Ph.D. Thesis, Institute for Systems and Automatic Control, University of Stuttgart, Allemagne (2004)
81. P. Menold, R. Findeisen, F. Allgower, Finite time convergent observers for nonlinear systems, in *Proceedings of the IEEE Conference on Decision and Control* (2003)
82. P. Menold, R. Findeisen, F. Allgöwer, Finite time convergent observers for linear time-varying systems, in *Proceedings of the Mediterranean Conference on Control and Automation* (2003)
83. Q. Miao, V. Makis, Condition monitoring and classification of rotating machinery using wavelets and hidden Markov models. *Mech. Syst. Signal Process.* **21**(2), 840–855 (2007)
84. M. Moisan, O. Bernard, Semi-global finite-time observers for nonlinear systems. *Automatica* **44**(12), 3152–3156 (2008)
85. E. Moulay, W. Perruquetti, Finite-time stability and stabilization of a class of continuous systems. *J. Math. Anal. Appl.* **323**(2), 1430–1443 (2003)
86. E. Moulay, W. Perruquetti, Finite-time stability of nonlinear systems, in *42nd IEEE International Conference on Decision and Control* (2003), pp. 3641–3646
87. A. Muller, M. Suhner, B. Iung, Formalisation of a new prognosis model for supporting proactive maintenance implementation on industrial system. *Reliab. Eng. Syst. Saf.* **93**(2), 234–253 (2008)
88. A. Nagy-Kiss, G. Schutz, J. Ragot, Parameter estimation for uncertain systems based on fault diagnosis using Takagi–Sugeno model. *ISA Trans.* **56**, 65–74 (2015)
89. L. Nguyen, M.A. Djeziri, B. Ananou, M. Ouladsine, J. Pinaton, Degradation modelling with operating mode changes, in *IEEE International Conference on Prognostics and Health Management*, Austin, TX (2015)
90. L. Nguyen, M.A. Djeziri, B. Ananou, M. Ouladsine, J. Pinaton, Health indicator extraction for fault prognosis in discrete manufacturing processes. *IEEE Trans. Semicond. Manuf.* **28**(3), 306–317 (2015)
91. J.V. Noortwijk, A survey of the application of gamma processes in maintenance. *Reliab. Eng. Syst. Saf.* **94**(1), 2–21 (2009)
92. A. Oppenheim, R. Schaffer, J. Buck, *Discrete-Time Signal Processing*, vol. 2 (Prentice-Hall, Englewood Cliffs, 1989)
93. M. Orchard, B. Wu, G. Vachtsevanos, A particle filtering framework for failure prognosis, in *World Tribology Congress III*, pp. 883–884 (American Society of Mechanical Engineers, New York, 2005)
94. Y. Pan, J. Chen, L. Guo, Robust bearing performance degradation assessment method based on improved wavelet packet and support vector data description. *Mech. Syst. Signal Process* **23**(3), 669–681 (2009)
95. C. Park, W. Padgett, Accelerated degradation models for failure based on geometric Brownian motion and gamma processes. *Lifetime Data Anal.* **11**(4), 511–527 (2005)
96. M.A. Patil, P. Tagade, K.S. Hariharan, S.M. Kolake, T. Song, T. Yeo, S. Doo, A novel multistage support vector machine based approach for li ion battery remaining useful life estimation. *Appl. Energy* **159**(1), 285–297 (2015)
97. T. Raff, F. Lachner, F. Allgower, A finite time unknown input observer for linear systems, in *Mediterranean Conference on Control and Automation* (2006)
98. M. Roemer, G. Kacprzyński, M. Kawamoto, S. Chigusa, Advanced diagnostics and prognostics for gas turbine engine risk assessment, in *IEEE, Aerospace Conference Proceedings*, pp. 345–353 (2000)
99. S. Sankaraman, Significance, interpretation, and quantification of uncertainty in prognostics and remaining useful life prediction. *Mech. Syst. Signal Process.* **52**, 228–247 (2015)
100. A. Saxena, J. Celaya, E. Balaban, B. Saha, S. Saha, K. Goebel, Metrics for evaluating performance of prognostic techniques, in *International Conference on Prognostics and Health Management (PHM08)*, pp. 1–17 (2008)

101. A. Saxena, J. Celaya, B. Saha, S. Saha, K. Goebel, On applying the prognostic performance metrics, in *Annual Conference of the Prognostics and Health Management Society* (2009)
102. M. Schlechtingen, I.F. Santos, S. Achiche, Wind turbine condition monitoring based on SCADA data using normal behavior models. Part 1: system description. *Appl. Soft Comput.* **13**(1), 259–270 (2013)
103. X. Si, W. Wang, C. Hu, D. Zhou, Remaining useful life estimation: a review on the statistical data driven approaches. *Eur. J. Oper. Res.* **213**(1), 1–14 (2011)
104. R. Singleton, E. Strangas, S. Aviyente, Extended Kalman filtering for remaining useful life estimation of bearings. *IEEE Trans. Ind. Electron.* **62**(3), 1781–1790 (2015)
105. K.L. Son, M. Fouladirad, A. Barros, E. Levrat, B. Jung, Remaining useful life estimation based on stochastic deterioration models: a comparative study. *Reliab. Eng. Syst. Saf.* **112**, 165–175 (2013)
106. J. Sun, H. Zuo, W. Wang, M. Pecht, Prognostics uncertainty reduction by fusing on-line monitoring data based on a state-space-based degradation model. *Mech. Syst. Signal Process.* **42**(2), 396–407 (2014)
107. J. Sun, H. Zuo, W. Wang, M. Pecht, Prognostics uncertainty reduction by fusing on-line monitoring data based on a state-space-based degradation model. *Mech. Syst. Signal Process.* **45**(2), 396–407 (2015)
108. K. Tidiri, N. Chatti, S. Verron, T. Tiplica, Bridging data-driven and model-based approaches for process fault diagnosis and health monitoring: a review of researches and future challenges. *Annu. Rev. Control.* **42**, 63–81 (2016)
109. D. Tobon-Mejia, K. Medjaher, N. Zerhouni, The ISO 13381-1 standard's failure prognostics process through an example, in *IEEE Prognostics and Health Management Conference (PHM'10)* (2010), pp. 1–12
110. S. Tseng, C. Peng, Stochastic diffusion modeling of degradation data. *J. Data Sci.* **5**(3), 315–333 (2007)
111. S. Tseng, J. Tang, I. Ku, Determination of burn-in parameters and residual life for highly reliable products. *Nav. Res. Logist.* **50**, 1–14 (2003)
112. D. Van, C. Berenguer, Condition based maintenance model for a production deteriorating system, in *Conference on Control and Fault-Tolerant Systems* (2010)
113. X. Wang, Wiener processes with random effects for degradation data. *J. Multivar. Anal.* **101**(2), 340–351 (2010)
114. P. Wang, G. Vachtsevanos, Fault prognostics using dynamic wavelet neural networks. Technical Report. AAAI Technical Report (1999)
115. A.P. Wang, H. Wang, Fault diagnosis for nonlinear systems via neural networks and parameter estimation, in *2005 International Conference on Control and Automation*, vol. 1 (2005), pp. 559–563
116. W. Wang, M. Carr, W. Xu, K. Kobayashi, A model for residual life prediction based on Brownian motion with an adaptive drift. *Microelectron. Reliab.* **51**(2), 285–293 (2011)
117. W. Wu, J. Hu, J. Zhang, Prognostics of machine health condition using an improved ARIMA-based prediction method, in *2nd IEEE Conference on Industrial Electronics and Applications* (2007), pp. 1062–1067
118. J. Yan, M. Koc, J. Lee, A prognostic algorithm for machine performance assessment and its application. *Prod. Plan. Control* **15**(8), 796–801 (2004)
119. D. Yu, J. Cheng, Y. Yang, Application of EMD method and Hilbert spectrum to the fault diagnosis of roller bearings. *Mech. Syst. Signal Process.* **19**(2), 259–270 (2005)
120. Y. Zhang, R.H.H. Xiong, M.G. Pecht, Lithium-ion battery remaining useful life prediction with box-cox transformation and Monte Carlo simulation. *IEEE Trans. Ind. Electron.* **66**(2), 1585–1597 (2018)
121. Y. Zhang, R.H.H. Xiong, M.G. Pecht, Long short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries. *IEEE Trans. Veh. Technol.* **67**(7), 5695–5705 (2018)
122. E. Zio, Prognostics and health management of industrial equipment, in *Diagnostics and Prognostics of Engineering Systems: Methods and Techniques* (IGI Global, Pennsylvania, 2012), pp. 333–356. <https://doi.org/10.4018/978-1-4666-2095-7>



123. E. Zio, Challenges in the vulnerability and risk analysis of critical infrastructures. *Reliab. Eng. Syst. Saf.* **152**, 137–150 (2016)
124. E. Zio, F.D. Maio, M. Stasi, A data-driven approach for predicting failure scenarios in nuclear systems. *Ann. Nucl. Energy* **37**, 482–491 (2010)
125. Z. Zng, E. Zio, A classification-based framework for trustworthiness assessment of quantitative risk analysis. *Saf. Sci.* **99**, 1215–226 (2017)