

# Model-Based and Data-Driven Prognostics under Different Available Information

Piero Baraldi<sup>a1</sup>, Francesco Cadini<sup>a</sup>, Francesca Mangili<sup>a</sup>, Enrico Zio<sup>b,a</sup>

<sup>a</sup>*Politecnico di Milano, Milano, Italy*

<sup>b</sup>*Chair on Systems Science and the Energetic challenge, European Foundation for New Energy-Electricite' de France, Ecole Centrale Paris and Supélec, Paris, France*

## Abstract

In practical industrial applications, different prognostic approaches can be used depending on the information available for the model development. In this paper, we consider three different cases: 1) a physics-based model of the degradation process is available; 2) a set of degradation observations measured on components similar to the one of interest is available; 3) degradation observations are available only for the component of interest.

The objective of the present work is to develop prognostic approaches properly tailored for these three cases and to evaluate them in terms of the assumptions they require, the accuracy of the Remaining Useful Life (RUL) predictions they provide and their ability of providing measures of confidence in the predictions. The first case is effectively handled within a particle filtering (PF) scheme, whereas the second and third cases are addressed by bootstrapped ensembles of empirical models.

The main methodological contributions of this work are i) the proposal of a strategy for selecting the prognostic approach which best suits the information setting, even in presence of mixed information sources; ii) the development of a bootstrap method able to assess the confidence in the RUL prediction in the third case characterized by the unavailability of any degradation observations until failure.

A case study is analyzed, concerning the prediction of the RUL of turbine blades affected by a developing creep.

**Keywords:** Prognostics, particle filtering, bootstrapped ensemble, turbine blade, creep.

## 1 Introduction

Prognostics aims at supplying reliable predictions about the Remaining Useful Life (RUL) of a component or system undergoing degradation. This is expected to improve planning of maintenance actions, increase safety and lower costs [1] [2].

---

<sup>1</sup> P. Baraldi is with the Politecnico di Milano, Dipartimento di Energia, via Ponzio 34/3, 20133 Milano, Italy (phone:+39 02 23996355; fax: +39 02 23996309; e-mail: piero.baraldi@polimi.it).

Different forms of information and data may be available for the assessment of the evolution to failure of a degrading system, e.g., time-to-failure data of similar systems, direct or indirect measures of the degradation states reached during its evolution or during the evolution of a set of similar systems under similar operating conditions, information on exogenous operational and environmental parameters, deterministic, empirical or semi-empirical models of the degradation process, etc. Depending on the situation, different prognostic methods may be applied [3] [4].

In this work, we consider three practical situations with decreasing information available for the prognostic task, and propose accurate and robust prognostic methods for each of them.

In general, prognostic methods can be classified in model-based and data-driven methods [5]. Model-based methods use an explicit mathematical model of the degradation process to predict the future evolution of the degradation state and, thus, the RUL of the system [6]. Examples of degradation models are the non-linear stochastic model of fatigue crack dynamics [7] [8] or the creep growth model based on the Norton law [9]. In practice, even when the model of the degradation process is known, the RUL estimate may be difficult to obtain, since the degradation state of the system may not be directly observable and/or the measurements may be affected by noise and disturbances. In these cases, model-based estimation methods aim at inferring the dynamic degradation state and provide a reliable quantification of the estimation uncertainty on the basis of the sequence of available noisy measurements. Many approaches rely on Bayesian methods [10] [11]: the exact Kalman filter has been largely used in case of linear state space models and independent, additive Gaussian noises, whereas analytical or numerical approximations of the Kalman filter (such as the Extended Kalman filter, the Gaussian-sum filters or the grid-based filters) have been applied in most realistic cases where the dynamics of degradation is non-linear and/or the associated noises are non-Gaussian [12].

Numerical approximations based on the Monte Carlo sampling technique have gained popularity for their flexibility and ease of design [13].

In the first case considered in this work, hereafter referred to as case 1, we have available a stochastic model of the degradation process and we know the value of the failure threshold, i.e., the maximum degradation beyond which the system loses its function. Also, a sequence of observations of the system degradation state are available and an observation equation describes the relation between the observations and the system degradation state. On this basis, a Monte Carlo-based filtering technique, called particle filtering (PF), is set-up to predict the distribution of the system RUL and online-update it when new observations are collected. The proposed approach improves the one previously proposed in [14] and [15] by taking into account the uncertainty on the parameters of the model of the degradation process and addressing the particle degeneration problem by means of the resampling algorithm [16].

On the other side, data-driven methods are used when an explicit model of the degradation process is not available, but sufficient historical data have been collected. These methods are based on statistical models that ‘learn’ trends from the data [17]. In this respect, artificial neural networks are often used [5] [18] [19]; other examples are Autoregressive Moving Average techniques [20], Relevance Vector Machines [18] [20] [21], fuzzy similarity-based methods [22]. Recently, ensemble approaches, based on the aggregation of

multiple model outcomes, have been introduced due to the superior robustness and accuracy with respect to single models [23] and the possibility of estimating the uncertainty of the predictions [24].

In this work, data-driven methods have been developed to tackle two different situations of information available (hereafter referred to as cases 2 and 3). In case 2, a number of observations of degradation evolution and the failure times of a set of similar systems operating under similar conditions are available; in case 3 only observations of the degradation of the system for which we want to predict the RUL and the value of the failure threshold are available. In both cases, the proposed prognostic approaches are based on the regression of the system degradation state by using an ensemble of bootstrapped models [24] which allows providing the uncertainty of the estimated RUL caused by the uncertainty in the data, the variability of the system behavior and the empirical model error. From the methodological point of view, the main contribution of the present work consists in the approach developed to deal with case 3 which, differently from case 2, is characterized by the unavailability of degradation data until the component failure and, thus, of the input (degradation value)/ output (RUL) pairs used in case 2 for estimating the uncertainty in the RUL prediction.

The three cases are studied with reference to the creep growth process in the blades of a helium gas turbine of a Gas Turbine Modular Helium Reactor (GT-MHR) [25] [26].

The problem of selecting the most appropriate prognostic approach in the case in which a mix of the information considered in cases 1,2 and 3 is available, has finally been addressed by comparing the performance of the three proposed approaches and by investigating their sensitivity to the accuracy of the model of the degradation process and to the amount and accuracy of the empirical data available.

The remainder of the paper is organized as follows: in Section 2 the objectives of the prognostic activity are presented; in Section 3, the sources of information for prognostics are discussed; in Section 4, the three cases considered are described; in Section 5, the prognostic methods developed to tackle the three cases are presented; in Section 6, the problem of blade creeping in high temperature turbines is illustrated and the prognostic results obtained in the three different cases considered are discussed; in Section 7 the problem of selecting the correct approach for specific situations of information available is discussed; finally, in Section 8 some conclusions are drawn and potential for future work suggested.

## 2 Information and data for prognostics

Let us discretize, for ease of exposition, the continuum time variable  $t$  into a sequence of time instants  $t_i$ ,  $i=1,2,\dots$  assumed to be equally spaced.

The aim of prognostics is to estimate the Remaining Useful Life  $RUL_i$  of a degrading system, i.e., the time left from the current time  $t_i$  before the system degradation crosses the failure threshold. Since degradation evolution is intrinsically random, the system  $RUL_i$  is a random variable and, thus, the objective of applying a prognostic method to a system whose current degradation state is  $d_i$  is to estimate the probability distribution  $p_{RUL_i}(rul_i | d_i)$  of  $RUL_i$  at time  $t_i$ .

Table I summarizes the main sources of information upon which prognostics can be based [4]:

- A physical model of the degradation mechanism (source A, Table I), e.g., described by a first-order Markov process:

$$d_j = g(d_{j-1}, \gamma_{j-1}); d_0 \sim p_{D_0}(d_0) \quad (1)$$

where  $d_j$  is the degradation state at time  $t_j$ ,  $p_{D_0}(d_0)$  is the initial distribution of the degradation at time  $t_0$ ,  $g$  is a possibly non-linear function describing the value of a one-time-step degradation increment and  $\gamma_j$ ,  $j=1,2,\dots$  is a sequence of mutually independent vectors of state noises. The model  $g$  can contain parameters referring to system inherent characteristics (material, physical, chemical, geometrical, etc.), which may vary from one individual system to another of the same type: this variability is described by probability distribution functions. The model can also describe the dependence of the degradation process from external parameters (environmental, operational, etc.), which may vary during the system life. Although these parameters are not directly related to the system degradation state, they may influence its evolution. Some of these parameters may be directly observable and measured by sensors, others may not; for some, there may be a priori knowledge of their behavior in time or statistical knowledge of their distribution.

- A set of observations  $\mathbf{z}_{1:i}$ , collected at different time instants  $t_{1:i}$ , during the life of the system whose RUL we want to predict (source of information B, Table I) or of a population of identical or similar systems (source of information C, Table I). Among the observable process parameters in  $\mathbf{z}$  there can be a direct measure of the degradation state of the system (e.g., depth of a crack fracture, elongation of a creeping component, etc.) or they can be only indirectly related to it (e.g., the time of travel or the intensity of ultrasonic waves for non-destructive inspections).
- The value of the failure threshold  $d_{th}$  (source D, Table I).
- The observation equation (source E, Table I), i.e., the physical model describing the relation between the observation  $\mathbf{z}_j$  containing the values of the observable process parameters measured by sensors at some time instant  $t_j$  and the actual degradation state  $d_j$  of the system:

$$\mathbf{z}_j = h(d_j, \mathbf{n}_j) \quad (2)$$

where  $h$  is a known function, in general non-linear, and  $\mathbf{n}_j$  is a vector of measurement noises.

- The life durations  $\{L_s\}_{s=1}^S$  of a number  $S$  of similar systems which have failed (source F, Table I); notice that, the actual value of the RUL of the  $s$ -th failed system can be computed at any time  $t_i < L_s$  as

$$rul_i^s = L_s - t_i \quad (3)$$

### 3 Three prognostic cases with different sources of information

Three cases are considered in this work, in which a set of measurements  $\mathbf{z}_{1:i}$  collected during the life of the system whose RUL we want to predict (source of information B) is available in combination with other different sources of information (Table II).

In case 1, the physical model of the evolution of the degradation state is known, as well as the distribution and evolution in time of all its characteristic and external parameters (source A). Other sources of information available are the value of the failure threshold  $d_{th}$  (source D) and the observation equation (source E) linking the observations with the degradation state. This situation is typical for well known degradation mechanisms, such as the crack or creep growth processes, which have been widely studied in laboratory.

In case 2, a set of observations  $\{\mathbf{z}_{1:N_s}^s\}_{s=1}^S$ , of  $S$  similar systems (source C) and the duration of their lives (source F) are available. This situation is typical for short-life systems, for which many trajectories to failure can be observed.

Finally, in case 3, the information available is the observation equation (source E) and the value of the failure threshold  $d_{th}$  (source D). This situation can occur in case of very reliable systems, e.g., those used in the nuclear industry, which have a very long life duration and are usually renewed before failure happens.

### 4 Modeling approaches

This Section illustrates the three modeling approaches undertaken to cope with the three prognostic cases outlined in Section 3 (Table II).

#### 4.1 Case 1: Particle Filtering

In case 1, at time  $t_i$ , the current degradation state  $d_i$  is not directly known, but the stochastic system dynamic model of eq. (1), the observation equation of eq. (2), the sequence of  $i$  observations  $\mathbf{z}_{1:i}$  related to the system degradation state and the value of the failure threshold  $d_{th}$  are available. Thus, instead of estimating  $p_{RUL_i}(rul_i | d_i)$  we are forced to restrict our objective to estimating the probability density function (pdf)  $p_{RUL_i}(rul_i | \mathbf{z}_{1:i}, D(t_i) < d_{th})$ , conditioned on the observations  $\mathbf{z}_{1:i}$  and on the fact that at time  $t_i$  the equipment has not yet failed, i.e.,  $D(t_i) \leq d_{th}$ ,

In this setting, defining  $D(t_j)$  the random variable which describes the degradation state at time  $t_j$ , it is desired to infer the unknown pdf  $p_{D(t_j)}(d_j | \mathbf{z}_{1:i}, D(t_i) < d_{th})$  of the degradation  $d_j$  at the future times  $t_j > t_i$  on the basis of all the previously estimated distribution of the state values  $p_{D(t_{0:j-1})}(d_{0:j-1} | \mathbf{z}_{1:i})$  and of all the observations  $\mathbf{z}_{1:i}$ . The RUL cumulative probability distribution  $F_{RUL_i}(rul_i | \mathbf{z}_{1:i}, D(t_i) \leq d_{th})$  is then computed from  $p_{D(t_j)}(d_j | \mathbf{z}_{1:i}, D(t_i) < d_{th})$  as the probability that the failure threshold  $d_{th}$  is exceeded

before time  $t_i + rul_i$  :

$$\begin{aligned}
F_{RUL_i}(rul_i | \mathbf{z}_{1:i}, D(t_i) < d_{th}) &= \text{Prob}(RUL_i < rul_i | \mathbf{z}_{1:i}, D(t_i) \leq d_{th}) \\
&= \text{Prob}(L - t_i < rul_i | \mathbf{z}_{1:i}, D(t_i) \leq d_{th}) \\
&= \text{Prob}(L < t_i + rul_i | \mathbf{z}_{1:i}, D(t_i) \leq d_{th}) \\
&= \int_{d_{th}}^{+\infty} p_{D(t_j=t_i+rul_i)}(d_j | \mathbf{z}_{1:i}, D(t_i) \leq d_{th}) \mathrm{d}d_j
\end{aligned} \tag{4}$$

In the prognostic problem, we resort to PF for estimating  $p_{D(t_j)}(d_j | \mathbf{z}_{1:i}, D(t_i) \leq d_{th})$  and solving the integral in eq.(4). In particular, the Sampling Importance Resampling (SIR) version of PF is here adopted, whose analytical details are provided in Appendix A.

The SIR PF method is based on sampling a large number  $K$  of trajectories  $\{d_{0:i}^k\}_{k=1}^K$  (called particles), by recursively sampling the state  $d_j^k$  from the transition pdf  $p_{D(t_{j+1})}(d_{j+1}^k | d_j^k)$  which can be derived from the physical model in eq. (1). Then, the posterior pdf  $p_{D(t_i)}(d_i | \mathbf{z}_{1:i})$  can be approximated as [10]:

$$p_{D(t_i)}(d_i | \mathbf{z}_{0:i}) \approx \sum_{k=1}^K w_i^k \delta(d_i - d_i^k) \tag{5}$$

where  $w_i^k$  is the importance weight associated to the sampled state sequence  $d_{0:i}^k$ ,  $k = 1, 2, \dots, K$ . The weight  $w_i^k$  can, then, be recursively computed as:

$$w_i^k = \frac{p_{Z(t_i)}(\mathbf{z}_i | d_i^k) w_{i-1}^k}{\sum_{k=1}^K p_{Z(t_i)}(\mathbf{z}_i | d_i^k) w_{i-1}^k} \tag{6}$$

where  $p_{Z(t_i)}(\mathbf{z}_i | d_i^k)$  is the *likelihood* of the observation  $\mathbf{z}_i$ .

To predict the pdf of the degradation states at future times  $t_j, j=i+1, i+2, \dots$ , the prediction stage is iterated for each particle, by recursively appending the sampled trajectory  $d_{0:i}^k$  with a new degradation states  $d_{i+1}^k, d_{i+2}^k, \dots, d_j^k$ , while keeping the weights fixed to their values  $w_i^k$  calculated at the time  $t_i$  of the last observation. Indeed, the pdf  $p_{D(t_j)}(d_j | \mathbf{z}_{1:i})$  can be approximated as:

$$p_{D(t_j)}(d_j | \mathbf{z}_{0:i}) \approx \sum_{k=1}^K w_i^k \delta(d_j - d_j^k) \tag{7}$$

Finally, the pdf  $p_{D(t_j)}(d_j | \mathbf{z}_{0:i}, d_i^k \leq d_{th})$  conditioned on the fact that  $D(t_i) \leq d_{th}$  can still be approximated resorting to eq. (7) but taking into account only those particles whose degradation at time  $t_i$  is below the threshold, i.e.,  $d_i^k \leq d_{th}$ . Operatively, this entails setting to zero the weights of these particles and normalizing the remaining ones, thus getting a new set of weights  $\tilde{w}_i^k$ .

Notice that the approximated pdf thus obtained is a discrete probability mass function where only the degradation values  $d_{t_i+rul_i}^k$  assumed by the particles at time  $t_i + rul_i$  have a finite probability equal to their weights  $\tilde{w}_i^k$ ; then, the integral in eq. (4) corresponds to the summation of the weights of the particles whose degradation at time  $t_i + rul_i$  exceeds the threshold  $d_{th}$ :

$$\begin{aligned} F_{RUL_i}(rul_i | \mathbf{z}_{1:i}, d_i < d_{th}) &= \int_{d_{th}}^{+\infty} \sum_{k=1}^K \tilde{w}_i^k \delta(d_j - d_{t_i+rul_i}^k) dd_j \\ &= \sum_{k=1}^K \tilde{w}_i^k H(d_{t_i+rul_i}^k - d_{th}) \end{aligned} \quad (8)$$

where  $H(d_{t_i+rul_i}^k - d_{th})$  is the Heaviside step function.

The application of the particle filtering procedure to the estimation of  $p_{RUL_i}(rul_i | \mathbf{z}_{1:i})$  is detailed in the pseudo-code of Figure 1.

Unfortunately, the procedure illustrated suffers from the so called degeneracy phenomenon: after few samplings, the weight variance increases and most of the  $K$  weights in eq. (8) become negligible so that the corresponding trajectories do not contribute to the estimate of the pdf of interest [10][16]. As a result, the approximation of the target distribution  $p_{RUL_i}(rul_i | \mathbf{z}_{1:i}, D(t_i) < d_{th})$  becomes very poor and significant computational resources are spent trying to update particles with minimum relevance.

A possible solution to this problem is offered by the bootstrap resampling algorithm, which is detailed in the pseudo-code of Figure 2 [11]. When degeneracy occurs, e.g. after few iterations of the weight updating procedure,  $K$  samples are drawn with replacement from the swarm of  $K$  particles; the  $k$ -th particle is sampled with a probability proportional to its weight value  $\tilde{w}_i^k$  and the sequence of degradation state  $d_{1:i}^k$  until time  $t_i$  is retained for the resampled particle  $k'$  and recursively augmented with new degradation states  $d_{j+1}^{k'}$ . The  $K$  resampled particles are then assigned the same weight  $1/K$  is assigned to all of them. Then, the filtering procedure continues with the original trajectories  $d_{1:j}^k$  and the associated weights  $\tilde{w}_i^k$  replaced by new trajectories  $d_{1:j}^{k'}$  with weights  $\tilde{w}_i^{k'} = 1/K$ .

## 4.2 Case 2: Data-driven prognostics based on an ensemble of bootstrapped models trained on degradation and life duration data

In case 2, the information available at time  $t_i$  is a set of degradation observations  $\{\mathbf{z}_{1:N_s}^s\}_{s=1}^S$ , taken during the trajectory to failure of  $S$  similar systems, the duration of their life  $L^s$ , and the observations  $\mathbf{z}_{1i}$  related to the degradation state of the system of interest.

In this context of information available, we are not able to estimate the probability distribution,

$p_{RUL_i}(rul_i | \mathbf{z}_{1i})$ , of  $RUL_i$  for a system that at time  $t_i$  is in the degradation state  $d_i$ . In practice, our objective is limited to obtain:

1. an estimate  $r\hat{u}_i$  of the expected value  $\mu_{RUL_i}$  of  $RUL_i$ ;
2. an estimate  $\hat{\sigma}_{r\hat{u}_i}^2$  of the variance of the prediction error  $\sigma_{r\hat{u}_i}^2 = E[(r\hat{u}_i - rul_i)^2]$ ; this quantity can be interpreted as a measure of the accuracy with which the predicted value  $r\hat{u}_i$  is expected to describe the actual  $rul_i$ .

The idea is to develop an empirical model:

$$f(\mathbf{z}_i) = r\hat{u}_i \quad (9)$$

of the relationship between the degradation observation available at time  $t_i$ ,  $\mathbf{z}_i$ , and  $RUL_i$ . This empirical model receives in input the current observation  $\mathbf{z}_i$  and produces as output the RUL prediction,  $r\hat{u}_i$  and an estimate  $\hat{\sigma}_{r\hat{u}_i}^2$  of the variance of the predicted error.

In order to develop the model, a dataset:

$$\mathbf{D}_{i/o} = \{ \{ \mathbf{z}_i^s; rul_i^s \}_{i=1}^{N_s} \}_{s=1}^S \quad (10)$$

is extracted from the set of observations  $\{\mathbf{z}_{1:N_s}^s\}_{s=1}^S$ , by associating to the observation  $\mathbf{z}_i$  taken at time  $t_i$  during the  $s$ -th trajectory to failure, the corresponding RUL:

$$rul_i^s = L^s - t_i \quad (11)$$

The dataset  $\mathbf{D}_{i/o}$  can be used to train an empirical model built using one among the many data-driven modeling methods existing today (e.g., polynomial regression, non-parametric regression, neural networks, etc.).



In general, the regression problem can be framed as follows: given a set of data pairs  $\{\mathbf{z}_i, y_i\}_{i=1}^N$ , generated from:

$$y_i = h(\mathbf{z}_i) + v_i \quad (12)$$

where  $y_i$  is the target value,  $h(\mathbf{z}_i)$  the true input/output relation and  $v_i$  a process noise with zero mean and standard deviation  $\sigma_{v_i}^2(\mathbf{z}_i)$ . Our aim is to train a model  $f(\mathbf{z}_i)$  using the data  $\{\mathbf{z}_i, y_i\}_{i=1}^N$ , which approximate  $h(\mathbf{z}_i)$ . According to [24] the developed empirical model can be interpreted as an estimate of the mean distribution of the target values given an input vector  $\mathbf{z}_i$ . In this context, Heskes [24] proposed a method, for providing a measure of confidence in the prediction  $f(\mathbf{z}_i)$  (see appendix B for a detailed description of the method). In our application, assuming  $RUL_i$  a random variable with mean  $\mu_{RUL_i}(\mathbf{z}_i)$  and variance  $\sigma_{RUL_i}^2(\mathbf{z}_i)$ , we can write the relationship between pairs  $\{\mathbf{z}_i^s; rul_i^s\}$  as:

$$rul_i = \mu_{RUL_i}(\mathbf{z}_i) + v_i \quad (13)$$

where  $\mu_{RUL_i}(\mathbf{z}_i)$  is the expected value of the RUL value given the observation  $\mathbf{z}_i$  and  $v_i$  is a random variable with zero mean and standard deviation  $\sigma_{v_i}^2(\mathbf{z}_i) = \sigma_{RUL_i}^2(\mathbf{z}_i)$ . Comparing eq. (13) with eq. (12) we have  $y_i = rul_i$  and  $h(\mathbf{z}_i) = \mu_{RUL_i}(\mathbf{z}_i)$ . Accordingly,  $f(\mathbf{z}_i)$  is interpreted as an estimator of  $\mu_{RUL_i}(\mathbf{z}_i)$  and, as described in [24], the uncertainty in the prediction  $f(\mathbf{z}_i)$  is quantified by the prediction error variance  $\sigma_{\hat{rul}_i}^2(\mathbf{z}_i)$  which can be decomposed into two terms:

$$\begin{aligned} \sigma_{\hat{rul}_i}^2(\mathbf{z}_i) &= E[(\hat{rul}_i - rul_i)^2] = \\ &= E\{[f(\mathbf{z}_i) - \mu_{RUL_i}(\mathbf{z}_i)]^2\} + E\{[\mu_{RUL_i}(\mathbf{z}_i) - rul_i]^2\} \\ &= \sigma_m^2(\mathbf{z}_i) + \sigma_{RUL_i}^2(\mathbf{z}_i) \end{aligned} \quad (14)$$

where the term  $\sigma_m^2(\mathbf{z}_i)$  is the model error variance describing the regression error made by the model  $f(\mathbf{z}_i)$  in estimating the true RUL mean value  $\mu_{RUL_i}(\mathbf{z}_i)$ , and the term  $\sigma_{RUL_i}^2(\mathbf{z}_i)$  is the RUL variance caused by the uncertainty on the future degradation of the system and describing the accuracy of  $\mu_{RUL_i}(\mathbf{z}_i)$  in predicting the target  $rul_i$ .

Notice that the application of the method described in Appendix B to the estimate of the model error variance  $\sigma_m^2(\mathbf{z}_i)$ , and the RUL variance  $\sigma_{RUL_i}^2(\mathbf{z}_i)$  requires the partition of the input/output dataset  $\mathbf{D}_{i/o}$

into a training and a validation datasets,  $\mathbf{D}_{i/o}^{tm}$  and  $\mathbf{D}_{i/o}^{val}$ : the training dataset is used to train the regression model  $f(\mathbf{z}_i)$ , whereas the validation dataset is used to test  $f(\mathbf{z}_i)$  and collect examples of its prediction error. Since the two datasets  $\mathbf{D}_{i/o}^{tm}$  and  $\mathbf{D}_{i/o}^{val}$  have to be independent in order to avoid underestimating the variance  $\sigma_{RUL_i}^2(\mathbf{z}_i)$ , we have considered a validation dataset  $\mathbf{D}_{i/o}^{val} = \{\{\mathbf{z}_i^s; rul_i^s\}_{i=1}^{N_s}\}_{s=1}^{S_{val}}$  made by input/output pairs taken from trajectories different from those used to build the training dataset  $\mathbf{D}_{i/o}^{tm}$ .

In practice, the overall approach to estimate  $r\hat{ul}_i$  and  $\hat{\sigma}_{r\hat{ul}_i}^2$  requires to:

- train an ensemble of models  $\{f^b(\mathbf{z}_i | \mathbf{D}_{i/o}^b)\}_{b=1}^B$  using bootstrapped replicates  $\mathbf{D}_{i/o}^b$  of the training dataset  $\mathbf{D}_{i/o}^{tm}$ ;
- test the bootstrapped ensemble on the validation dataset  $\mathbf{D}_{i/o}^{val}$  to compute the prediction residuals  $r^2(\mathbf{z}_i^s)$  as in eq. (B4) of Appendix B (with  $y_i = rul_i^s$ );
- use the set of residuals input/output pairs  $\{\{\mathbf{z}_i^s; r^2(\mathbf{z}_i^s)\}_{i=1}^{N_s}\}_{s=1}^{S_{val}}$  to train the model  $\chi(\mathbf{z}_i) = \hat{\sigma}_{RUL_i}^2(\mathbf{z}_i)$  describing the dependence of  $\hat{\sigma}_{RUL_i}^2(\mathbf{z}_i)$  from  $\mathbf{z}_i$  [24][27];
- when a new observation  $\mathbf{z}_i$  about the degradation state of a functioning system is collected compute the output  $r\hat{ul}_i^b$  of each models  $f^b(\mathbf{z}_i | \mathbf{D}_{i/o}^b)$  of the ensemble;
- compute the prediction  $r\hat{ul}_i$ , i.e., the estimate of the RUL expected value  $\mu_{RUL_i}$ :

$$r\hat{ul}_i = \frac{1}{B} \sum_{b=1}^B f^b(\mathbf{z}_i | \mathbf{D}_{i/o}^b) \quad (15)$$

- compute the estimate  $\hat{\sigma}_m^2(\mathbf{z}_i)$  of the model error variance:

$$\hat{\sigma}_m^2(\mathbf{z}_i) = \frac{1}{B} \sum_{b=1}^B [f^b(\mathbf{z}_i | \mathbf{D}_{i/o}^b) - r\hat{ul}_i]^2 \quad (16)$$

- apply the model  $\chi(\mathbf{z}_i)$  to the input  $\mathbf{z}_i$  to obtain the RUL variance estimate  $\hat{\sigma}_{RUL_i}^2(\mathbf{z}_i)$ ;
- sum up the two variance components to obtain the prediction error variance  $\sigma_{r\hat{ul}_i}^2$ :

$$\sigma_{r\hat{ul}_i}^2(\mathbf{z}_i) = \hat{\sigma}_m^2(\mathbf{z}_i) + \hat{\sigma}_{RUL_i}^2(\mathbf{z}_i) \quad (19)$$

### 4.3 Case 3: Data-driven prognostics based on an ensemble of bootstrapped models trained on degradation data only

This case is characterized by the availability of the observations  $\mathbf{z}_{1:i}$  related to the degradation state of the system of interest at  $j$  different measurement time instants up to the current time  $t_i$ , the relative observation equation and the value of the failure threshold  $d_{th}$ . Given the observation equation, an estimate of the

degradation state  $d_j$  can always be derived from the observation  $\mathbf{z}_j$ . For simplicity of illustration, we consider here only a situation where the observation  $\mathbf{z}_j$  is a direct measure of the degradation state  $d_j$ , eventually affected by a zero-mean noise, and thus no further estimate of  $d_j$  is needed.

The modeling approach proposed in the previous Section, based on the availability of input/output pairs formed by the observations  $\mathbf{z}_i$  and the corresponding RUL value  $rul_i$ , cannot be directly applied to this case. For this reason, an approach which uses the time series  $\mathbf{z}_{1:i}$  of the past observations to build a model of the time evolution of the degradation process is proposed. Notice that the approach differs from that used in case 1 since the physical stochastic model describing the true dynamics of the degradation process (eq. (1)) is unknown and should be replaced by an empirical deterministic model derived from the few available data. Coherently, the estimate of the prediction error variance  $\sigma_{rul_i}^2(\mathbf{z}_i)$  should account also for the error made when approximating the true degradation process with the empirical model.

A generic model of the evolution of the degradation state of the system, achieved by fitting the most suited degradation model, e.g., linear and non-linear regression models, general degradation path models, etc. [28] to the sequence of data  $\mathbf{z}_{1:i}$ , can be written as:

$$\hat{d}_j = \eta(t_j) \quad (17)$$

where  $\hat{d}_j$  is the degradation value at time  $t_j$  predicted by the model.

The prediction  $r\hat{u}_i$  of the system RUL at time  $t_i$  is then obtained from the relation

$$\eta(t_i + r\hat{u}_i) = d_{th} \quad (18)$$

An estimate of the prediction error variance  $\sigma_{rul_i}^2(\mathbf{z}_i)$  cannot be obtained by means of the method proposed in case 2, since there are no available pairs  $\{\mathbf{z}_j, rul_j\}_{j=1}^{N'}$  for which  $rul_j$  is known, and thus the prediction residuals  $\{r^2(\mathbf{z}_j)\}_{j=1}^{N'}$  cannot be computed in correspondence of any of the observation  $\mathbf{z}_{1:i}$ . To overcome this problem, we consider a model

$$\Delta\hat{t}_{j,j'} = \tilde{\eta}(\mathbf{d}_{j,j'}) \quad (19)$$

which receives in input a vector of two degradation states  $\mathbf{d}_{j,j'} = [d_j \ d_{j'}]$  and returns in output the estimate  $\Delta\hat{t}_{j,j'}$  of the time interval needed to reach the degradation state  $d_j$  starting from  $d_{j'}$ . In general, model  $\tilde{\eta}(\mathbf{d}_{j,j'})$  can be derived from model  $\eta(t_j)$  according to:

$$\Delta \hat{t}_{j,j'} = \tilde{\eta}(\mathbf{d}_{j,j'}) = \eta^{-1}(d_{j'}) - \eta^{-1}(d_j) \quad (20)$$

The prediction  $\hat{r}ul_i$  is obtained from this model by setting  $d_j = \mathbf{z}_i$  and  $d_{j'} = d_{th}$ ; in this view, the RUL prediction at time  $t_i$  corresponds to the estimate of the time interval  $\Delta t_{i,th}$  needed to increase the degradation state from  $d_i$  to the failure threshold  $d_{th}$ . Model  $\tilde{\eta}(\mathbf{d}_{j,j'})$  is assumed to be an unbiased estimator of the mean value  $\mu_{\Delta t_{j,j'}}(\mathbf{d}_{j,j'})$  of the random variable  $\Delta t_{j,j'}$ ; the variance  $\sigma_m^2(\mathbf{d}_{j,j'})$  of the difference between the estimate  $\tilde{\eta}(\mathbf{d}_{j,j'})$  and  $\mu_{\Delta t_{j,j'}}(\mathbf{d}_{j,j'})$  represents the uncertainty associated to the model  $\tilde{\eta}$ ; the variance  $\sigma_{\Delta t_{j,j'}}^2$  of the difference between  $\mu_{\Delta t_{j,j'}}(\mathbf{d}_{j,j'})$  and the actual target value  $\Delta t_{j,j'}$  represents the uncertainty in the evolution of the degradation process from  $d_j$  to  $d_{j'}$ . Being  $\sigma_{\Delta t}^2$  and  $\mu_{\Delta t}$  functions of the input  $\mathbf{d}_{j,j'}$ , eq. (12) becomes:

$$\Delta t_{j,j'} = \mu_{\Delta t_{j,j'}}(\mathbf{d}_{j,j'}) + v_{j,j'} \quad (21)$$

where  $v_{j,j'}$  represent a process noise with zero mean and standard deviation  $\sigma_{\Delta t_{j,j'}}^2(\mathbf{d}_{j,j'})$ .

The bootstrap method used for case 2 and described in Appendix B, can now be applied considering, instead of the quantities  $y_i$ ,  $h(\mathbf{z}_i)$  and  $f(\mathbf{z}_i)$ , the quantities  $\Delta t_{j,j'}$ ,  $\mu_{\Delta t_{j,j'}}(\mathbf{d}_{j,j'})$  and  $\tilde{\eta}(\mathbf{d}_{j,j'})$ , respectively.

As underlined in Section 4.2, to avoid underestimating the prediction error, the validation datasets should not contain measurements belonging to degradation trajectories used for training. Since only a single trajectory is now available, the solution proposed is to partition the dataset  $\mathbf{z}_{1:i}$  into two sequences of consecutive measurements,  $\mathbf{D}^{trn} = \{\mathbf{z}_{1:N_{trn}}\}$  and  $\mathbf{D}^{val} = \{\mathbf{z}_{N_{trn}+1:i}\}$ . An ensemble of  $B$  models  $\{\tilde{\eta}^b(\mathbf{d}_{j,j'})\}_{b=1}^B$  is then generated by training each model on a bootstrapped replicate  $\mathbf{D}^b$  of  $\mathbf{D}^{trn}$  and validated on validation dataset  $\mathbf{D}_{i/o}^{val}$  derived from  $\mathbf{D}^{val}$ :

$$\mathbf{D}_{i/o}^{val} = \{ \{ \mathbf{d}_{j,j'} = [\mathbf{z}_j \quad \mathbf{z}_{j'}], \Delta t_{j,j'} = [t_j \quad t_{j'}] \}_{j=N_{trn}+1}^{i-1} \}_{j'=j+1}^i \quad (22)$$

The prediction residuals  $r^2(\mathbf{d}_{j,j'})$  are then computed as in eq. (B4) of Appendix B, where  $y_i$  is replaced by  $\Delta t_{j,j'}$  and  $f^b(\mathbf{z}_i)$  is replaced by  $\tilde{\eta}^b(\mathbf{d}_{j,j'})$  and used to build the empirical model  $\hat{\sigma}_{\Delta t_{j,j'}}^2 = \chi(\mathbf{d}_{j,j'})$  estimating the variance  $\sigma_{\Delta t_{j,j'}}^2(\mathbf{d}_{j,j'})$  of  $\Delta t_{j,j'}$ . The RUL variance  $\hat{\sigma}_{RUL_i}^2(\mathbf{z}_i)$  is then obtained from model  $\chi$  fed with the input  $\mathbf{d}_{i,th} = [\mathbf{z}_i \quad d_{th}]$ .

Thus, when a new observation  $\mathbf{z}_i$  is collected at time  $t_i$ , the outcomes of the ensemble models are use to generate the RUL prediction:

$$r\hat{u}_i = \frac{1}{B} \sum_{b=1}^B \tilde{\eta}^b(\mathbf{d}_{i,th}) \quad (23)$$

and the prediction error variance estimate

$$\sigma_{r\hat{u}_i}^2 = \hat{\sigma}_m^2(\mathbf{z}_i) + \sigma_{RUL_i}^2(\mathbf{z}_i) = \frac{1}{B} \sum_{b=1}^B [\tilde{\eta}^b(\mathbf{d}_{i,th}) - r\hat{u}_i]^2 + \chi(\mathbf{d}_{i,th}) \quad (24)$$

Notice that the training data used to build model  $\chi$  cover a range of values for the input  $\mathbf{d}_{j,j'}$ , in general different from that of the input  $\mathbf{d}_{i,th}$  to which the model is applied to obtain the estimate  $\hat{\sigma}_{RUL_i}^2$ . This represents a limit to the quality of the estimate  $\hat{\sigma}_{RUL_i}^2$ , since the performances of empirical models tend to degrade when they are applied to input patterns belonging to regions far away from those containing the patterns used to train the model.

## 5 Numerical application

In this Section, the three different cases presented in Section 3 are considered with reference to the prognostics of a turbine blade in which creep damage is developing [26]. Creep is an irreversible deformation process affecting materials exposed to a load below the elastic limit for a protracted length of time and at high temperature. Notice that a turbine undergoing this degradation process can experience the loss of its blades, one of the most feared failure modes of turbomachinery since it is accompanied by abrupt changes in the power conversion system and in the reactor flow conditions [29]. Figure 3 shows an example of high-pressure turbine deblading occurred in a German power plant [29].

As shown in Figure 4, the uniaxial creep deformation consists in an augmentation of the original length and a reduction of the diameter. In this work, the adimensional quantity  $\varepsilon$ , defined as the percentage of elongation of the turbine blade in the longitudinal direction with respect to its original length, is considered as measure of the creep strain.

### 5.1 Information available for prognostics

The main sources of degradation-related information for the creep growth process, listed in Table III, are further detailed in this Section.

*Information A: creep growth model*

Creeping in turbine blades is a stochastic degradation process which can be modeled through the Norton Law, assuming that the dependence from the temperature follows the Arrhenius law [9]:

$$\frac{d\varepsilon}{dt} = A \cdot \exp\left(-\frac{Q}{RT}\right) \cdot \varphi^n \quad (25)$$

where  $d\varepsilon/dt$  is the creep strain rate,  $Q$  is the activation energy,  $A$  and  $n$  are material characteristics varying from one blade to another,  $R$  is the ideal gas constant,  $T$  is the blade operating temperature and  $\varphi$  is the applied stress. For simplicity, the blade temperature is supposed equal to the gas temperature and the stress  $\varphi$  is derived from the rotational speed  $\omega$  of the turbine:

$$\varphi = \rho \frac{r_{tip}^2 - r_{hub}^2}{2} \omega^2 \quad (26)$$

where  $\rho$  is the blade density and  $r_{hub}$  and  $r_{tip}$  are the hub and tip radiuses, respectively. The rotational speed  $\omega$  and the gas temperature  $T$  are external parameters depending on the power setting of the gas turbine. For  $\Delta t$  sufficiently small compared to the time horizon of the analysis (here  $\Delta t = 5$  days, with respect to the time horizon of several thousands), the state space model in eq. (26) can be discretized to give:

$$\varepsilon_{j+1} = \varepsilon_j + A \cdot \exp\left(-\frac{Q}{RT_j}\right) \cdot \varphi_j^n \cdot \Delta t_{j,j+1}, \quad \varepsilon_0 = 0 \quad (27)$$

The characteristic parameters  $A$  and  $n$  vary from one blade to another, whereas the external parameters, i.e., the rotational speed  $\omega$  and the gas temperature  $T$  vary continuously in time; all these parameters are assumed to have normal distributions. Finally, the fluctuations in the stress applied to a specific blade, which are due to fabrication defects, aging and corrosion of the blade, vibrations of the system or turbulences of the gas flow, are modeled through a random variable  $\delta\varphi$  added to the stress  $\varphi$  in eq. (27).

The values and distributions of the parameters  $T$ ,  $\omega$ ,  $r_{hub}$  and  $r_{tip}$  have been set with reference to the helium gas turbine of a Gas Turbine Modular Helium Reactor (GT-MHR) developed by an international consortium, with a targeted 286MWe generation per module[25]; the material inherent characteristics  $A$ ,  $n$  and  $\rho$  are taken assuming that the blade is made of Ni-base cast Superalloy 713LC [25]. The distributions used for the parameters are reported in Table III.

#### *Information B: creep strain measurements*

This source of information consists in a sequence of observations  $\mathbf{z}_{1:i}$  of creep strain performed on the blade on which we want to apply the three prognostic approaches, hereafter called ‘test trajectory’. Given the unavailability of real experimental data, in this work the creep growth trajectory is simulated using eq. (27).

The variation in time of the rotational speed  $\omega$ , the gas temperature  $T$  and the stress fluctuations  $\delta\phi$  are simulated by sampling their values  $\omega_j$ ,  $T_j$  and  $\delta\phi_j$  from the relative distributions (Table III) at each time instant  $t_j$ . Every 30 days a creep strain measurement  $\mathbf{z}_j$ , corresponding to the creep strain  $\varepsilon_j$ , is simulated by using eq. (28). A total number of 87 creep strain measurements have been simulated for a turbine blade with parameters  $A=3 \cdot 10^{-4}$  and  $n=6$ .

In order to verify the performance of the prognostic approaches, the simulation of the test trajectory has been conducted until the time  $L$  at which the creep strain reaches the failure threshold. The difference between  $L$  and the time  $t_i$  at which the prognosis is performed is the actual remaining useful life of the turbine blade and will be referred to as “true RUL”, and represented by the notation  $rul_i$  (Column 1, Table V).

*Information C: historical creep strain measurements*

This source of information consists in a number  $S=13$  of historical sequences of creep growth observations from similar blades. In analogy to what is done for information B, the degradation trajectories have been simulated using eq. (27). The variations of the characteristic parameters  $A$  and  $n$  from one blade to another have been simulated by sampling their values from normal distributions at the beginning of each new simulated degradation path. Some examples of simulated creep growth paths are shown in Figure 5.

For each trajectory, a number  $N_s$  of direct creep strain measurement  $\mathbf{z}_{1:N_s}$ , one every 30 days, are simulated according to eq. (28) (Information E).

*Information D: failure threshold*

A turbine blade is considered within its useful life if the creep elongation strain in the longitudinal direction of the turbine blade is less than 1% or 2% of its initial length. Thus, the failure threshold for creep strain  $\varepsilon_{th}$  is set equal to the value of 1.5%.

*Information E: measurement equation*

For simplicity, we assume to be able to directly measure the value of the creep strain once every 30 days. Consequently, the observation equation is:

$$\mathbf{z}_j = \varepsilon_j + v_j \quad (28)$$

where  $v_j$  is a white Gaussian measurement noise with standard deviation  $\sigma_v = 0.02$ .

*Information F: life duration data*

The time step at which the creep strain  $\varepsilon_j$  exceeds the failure threshold  $\varepsilon_{th}$  is collected for each of the  $S=13$  simulated degradation trajectories and represents the life duration  $L_s$  of the simulated turbine blade.

## 5.2 The three prognostic problems and corresponding modeling approaches

According to the three cases presented in Section 2, three prognostic problems have been tackled with respect to the turbine blade case study described in the previous Section 5.1. In all cases, the objective of the analysis is to predict at time  $t_i$ ,  $i=1, \dots, 87$ , the RUL distribution for the test trajectory. At every time  $t_i$  during the life of the turbine blade, the set of observations  $\mathbf{z}_{1:i}$  is assumed to be available (source of information B) and the predictions of the RUL is updated according to the new available information, i.e., the last observation  $\mathbf{z}_i$ .

In case 1, the sources of information A, B, D and E in Table IV are available. The particle filter has been applied and a number  $U = 1000$  of particles are simulated starting from  $\varepsilon_0 = 0$ . Particle resampling is performed once every 5 measurements. The particle filter has been preferred to the Kalman filter since the distribution of the process noise is not Gaussian as a consequence of the combination of speed, temperature and stress fluctuations in the creep growth process described by eq. (27)

In case 2, sources of information B, C and F in Table IV are available.  $S_{trn} = 10$  trajectories among the  $S = 13$  totally available are used for building an ensemble of  $B = 25$  linear least square models

$$r\hat{u}^b(\mathbf{z}_i | D_{i/o}^b) = \alpha_0^b + \alpha_1^b \cdot \mathbf{z}_i, \quad (29)$$

whereas the remaining  $S_{val} = 3$  trajectories are used to validate the ensemble and build the training dataset for the model

$$\chi(\mathbf{z}_i) = \hat{\sigma}_{RUL_i}^2 = \gamma_0 + \gamma_1 \cdot \mathbf{z}_i + \gamma_2 \cdot \mathbf{z}_i^2 \quad (30)$$

estimating the RUL variance  $\sigma_{RUL_i}^2(\mathbf{z}_i)$ .

In case 3, prognostic results are achieved based on the sources of information D and E of Table IV together with the information on the test trajectory (source of information B). In this case, the prognostic model has been developed only after time  $t_{30}$  in order to have available a dataset  $\mathbf{D} = \{\mathbf{z}_{1:i}\}$  of at least  $i = 30$  direct creep strain measurements. This dataset has been partitioned into a training dataset  $\mathbf{D}^{trn}$  containing the first 75% of the available measurements and a validation dataset  $\mathbf{D}^{val}$  containing the remaining 25%. An ensemble of  $B = 25$  linear least square models

$$\eta^b(t_j | \mathbf{D}^b) = \beta_0^b + \beta_1^b \cdot t_j \quad (31)$$

is built and the models

$$\tilde{\eta}^b[\Delta\mathbf{z}_{j,j'} | \mathbf{D}^b] = \Delta\mathbf{z}_{j,j'} / \beta_1^b \quad (32)$$



are derived from it. Notice that, in a linear process, the time needed to increase the degradation state from  $\varepsilon_j$  to  $\varepsilon_{j'}$  is proportional to the degradation increment  $\Delta\varepsilon_{j,j'} = \varepsilon_{j'} - \varepsilon_j$  and does not depend on the initial and final degradation states. The ensemble of models is tested on the validation dataset made of input/output pairs

$$\mathbf{D}_{i/o}^{val} = \left\{ \left\{ \Delta\mathbf{z}_{j,j'} = \mathbf{z}_{j'} - \mathbf{z}_j; \Delta t_{j,j'} = t_{j'} - t_j \right\}_{j=N_{im}+1}^{i-1} \right\}_{j=j+1}^i, \quad (33)$$

and the prediction residuals  $r_{j,j'}$  obtained are used to train the linear model

$$\chi(\Delta\mathbf{z}_{j,j'}) = \lambda\Delta\mathbf{z}_{j,j'} \quad (34)$$

for the variance of  $\Delta t$ . The predictions  $\hat{r}ul_i^b$  and the estimate  $\hat{\sigma}_{RUL_i}^2$  are obtained respectively from the model ensemble  $\tilde{\eta}^b$  and from model  $\chi$  in correspondence of the input  $\Delta\mathbf{z}_{i,th} = \varepsilon_{th} - \mathbf{z}_i$ . This way, the data used for training model  $\chi$  concern creep strain increments which for the first two thirds of the trajectory are smaller than the increment  $\Delta\mathbf{z}_{i,th}$  considered for obtaining the prognostic results, so that the empirical model  $\chi(\Delta\mathbf{z}_{j,j'})$  is used in an input region not described by the training data.

Finally, the prediction  $\hat{r}ul_i$  and the relative prediction error variance are obtained from eqs. (23) and (24). Each time  $t_i, i=31, \dots, 87$ , a new measurement becomes available, a new ensemble of models is built and a new RUL prediction is obtained.

### 5.3 Results

Table V reports the RUL predictions obtained by applying the three prognostic approaches of cases 1, 2 and 3 to a degrading blade. The first row refers to the RUL prediction performed at time  $t_{50} = 1475$  days on the basis of the measurements  $\mathbf{z}_{1:50}$  of the test trajectory, the second to the prediction performed at time  $t_{80} = 2375$  days on the basis of the measurements  $\mathbf{z}_{1:80}$ . Column 1 reports the true RUL value,  $rul_i$ , observed for the turbine blade under test, whereas columns 2 and 3 report the expected value  $\mu_{RUL_i}$  and the variance  $\sigma_{rul_i}^2$  of the distribution  $\Pr(RUL_i | \varepsilon_i)$ . This latter distribution represents the irreducible uncertainty of the RUL prediction which is caused by the stochastic future evolution of the creep strain.  $\Pr(RUL_i | \varepsilon_i)$  has been obtained by simulating  $P=1000$  degradation trajectories all characterized by the values  $A$  and  $n$  of the blade under test and by a creep strain  $\varepsilon_i$  at time  $t_i$ . Notice that the predictions  $\hat{r}ul_i$  of

the three approaches provide satisfactory estimates of  $\mu_{RUL_i}$ , whereas in all the cases the prediction error variances  $\hat{\sigma}_{r\hat{l}_i}^2$  tend to overestimate  $\sigma_{r\hat{l}_i}^2$ . This is due to the fact that according to eq. (14),  $\hat{\sigma}_{r\hat{l}_i}^2$  takes into consideration both the uncertainty due to the future stochastic evolution of the test trajectory represented by  $\sigma_{RUL_i}^2$  and the uncertainty due to the prognostic model regression error  $\sigma_m^2$ . It is interesting to observe that an analyst which has to decide the maintenance policy to be applied to the turbine blade would like to have the least uncertain prediction of the RUL. Thus, in the case in which the analyst were in the position to choose one of the three prognostic approaches, he/she would prefer the one which guarantees the lowest uncertainty, i.e., the one whose prediction error variance is smaller.

In correspondence of each prediction  $r\hat{l}_i$ , it is also possible to estimate the prediction interval  $PI(\alpha)$

$$C_i^{\text{inf}}(\alpha) < r\hat{l}_i < C_i^{\text{sup}}(\alpha), \quad (35)$$

i.e., the interval expected to contain the true RUL value  $rul_i$  with a probability of  $1 - \alpha$ . According to the three approaches, this interval can be obtained as follows:

- In case 1,  $C_i^{\text{inf}}(\alpha)$  and  $C_i^{\text{sup}}(\alpha)$  are the  $\alpha/2$  and  $1 - \alpha/2$  percentiles, respectively, of the RUL distribution estimated with the particle filtering method.
- In cases 2 and 3, assuming that the prediction error has a Gaussian distribution, the interval can be computed according to the theory of the bootstrap method [24] as:

$$r\hat{l}_i - c_{\text{conf}}^{\alpha} \hat{\sigma}_{r\hat{l}_i} < r\hat{l}_i < r\hat{l}_i + c_{\text{conf}}^{\alpha} \hat{\sigma}_{r\hat{l}_i} \quad (36)$$

where  $c_{\text{conf}}^{\alpha}$  is the  $1 - \alpha/2$  percentile of a Student's  $t$ -distribution with number of degrees of freedom equal to the number  $B$  of bootstrapped models.

Figure 6 shows the evolution of the true value of the blade RUL (continuous thick line), its estimated value  $r\hat{l}_i$  (dots) and the corresponding prediction interval for  $\alpha = 0.32$  (continuous thin line) obtained during the turbine blade life at times  $t_i, i=1, \dots, 80$ . In case three, since a minimum number of historical data must be available to build the predictive model, the prediction is performed only after  $t_{30}$ . Notice that in this case the prediction intervals are characterized by large oscillations and low accuracy, especially at the beginning of the trajectory, i.e., when few training data are available. Furthermore, the RUL prediction itself is very noisy. This effect can be reduced by properly filtering the predictions. To this purpose, since the time evolution of the RUL is a linear process ( $rul(t) = rul(t-1) - 1$ ), and assuming a Gaussian noise affecting the prediction, Kalman filtering can be applied [26].

In order to perform a robust analysis of the performances of the three approaches, the model in eq. (27) has been used to generate 250 different creep growth trajectories. For each trajectory, the prognostic indicators

$\hat{r}ul_i(\mathbf{z}_i)$  and the confidence interval  $C_i^{\text{inf}}$  (0.32) have been computed at  $N_{tst}$  different time steps, once every 150 days, based on the past measurements collected once every 30 days.

For each degradation test trajectory, two performance indicators are computed:

1. the mean relative absolute error  $rMAE$ :

$$rMAE = \frac{1}{N_{tst}} \sum_{i=1}^{N_{tst}} \left| \frac{rul_i - \hat{r}ul_i}{rul_i} \right| \quad (37)$$

which evaluates the accuracy of the estimate  $\hat{r}ul_i$  with respect to the true  $rul_i$  of the system. Notice that since  $\hat{r}ul_i$  estimates the expected value of  $\mu_{RUL_i}$  and not the true value of the  $rul_i$ , this value is not expected to be zero even for the best possible prognostic model.

2. the coverage:

$$Cov = \frac{1}{N_{tst}} \sum_{i=1}^{N_{tst}} c_i, \quad c_i = \begin{cases} 1 & C_i^{\text{inf}}(0.32) < rul_i < C_i^{\text{sup}}(0.32) \\ 0 & \text{otherwise} \end{cases} \quad (38)$$

This indicator is used to verify whether the estimation of the prediction interval  $PI(0.32)$  actually contains with probability  $1-0.32=0.68$  the true RUL of the system. Coverage values around 0.68 indicate satisfactory estimation of the prediction interval.

The average values  $\overline{rMAE}$  and  $\overline{Cov}$  of the performance indicators obtained in the three cases over the 250 test trajectories are reported in Table VI.

The best results are obtained in case 1, which is the one with the maximum amount of information available. In this case the prediction is accurate (low  $rMAE$ ) and the uncertainty of the prediction well estimated (coverage close to 0.68). The accuracy of the prediction in case 3 slightly outperforms the one in case 2, although a smaller amount of information is available to build the model. This can be explained by considering that in case 2 the prediction is based on knowledge about the creeping behaviors of a population of similar, but not identical blades, i.e., characterized by different values of parameters  $A$  and  $n$ ; on the contrary, in case 3 the empirical model is trained using degradation data concerning only the turbine blade of interest and thus all training data refer to the same values of parameters  $A$  and  $n$ . To confirm this hypothesis, in Table VII two cases 1b and 2b analogous to cases 1 and 2 are considered: in case 1b it is assumed that the exact values of parameters  $A$  and  $n$  are known for each blade, whereas in case 2b the degradation trajectories used to build the training dataset are simulated using the same value of  $A$  and  $n$  considered in the test trajectory. We observe that the accuracy of the prediction is increased and in both cases 1b and 2b the results are better than in case 3 given the larger amount of information available.

## 6 Different information settings

In the previous Sections, we have considered three well defined situations of information available and we have developed three, properly tailored, prognostic approaches. However, in real applications, it is common to face hybrid situations characterized by the availability, at the same time, of multiple sources of information. Furthermore, since some sources of information can be partially inaccurate or affected by large uncertainty, the identification of the correct prognostic approach to be applied can become a non trivial problem. In order to provide some indications to the decision maker, we consider a case in which all the sources of information listed in Table II are available (so that all the three proposed approaches can be applied) and we perform an analysis of the sensibility of the performance of the three prognostic approaches to the quality and quantity of the information available. To this purpose, we have considered the following indicators of the quality and quantity of the information:

- A. the amplitude of the noise affecting the creep strain measurements;
- B. the number of past measurements of the current trajectory available for making the RUL prediction;
- C. the accuracy of the physical model of the degradation process;
- D. the number of historical degradation trajectories available.

The performance of the three approaches is evaluated considering 250 test trajectories. Figure 7 (upper-left) shows the variation of the relative mean square error ( $rMAE$ ) when the amplitude of the noise affecting the creep strain measurements is varied from 0 to 0.03. Notice that, as expected, the performances of the three approaches decrease as the amplitude of the noise increases, and that the third approach is the most sensible to this parameter.

With respect to the second indicator, Figure 7 (upper-right) shows the mean absolute error of the approaches when they are applied at different time instants during the evolution of the degradation trajectories. Since every 30 days a new measurement is collected, the number of measurements available for making the RUL prediction increases as time passes. In this case, the performance is evaluated using the mean absolute error ( $MAE$ ) instead of the  $rMAE$  which tends to be very noisy at the blade end-of-life, when the denominator  $rul_i$  of eq. (37) gets close to zero. The performance of approach 3 is the most affected by the time at which the prognosis is made: the performance is very poor at the beginning of the degradation trajectory, when very few measurements are available for the construction of the empirical degradation model, but it increases significantly as time passes and, finally, when the creep strain becomes close to the failure threshold, approach 3 outperforms the other approaches.

The sensibility of approach 1 to the accuracy of the physical model (indicator 3.) has been estimated by using a biased value in the PF process model of eq. (27) of parameter  $A$ : its mean value has been taken in the range  $[2.7 \cdot 10^{-4}; 3.6 \cdot 10^{-4}]$  instead of equal to its true value  $\mu_A = 3 \cdot 10^{-4}$  (Table III). Figure 7 (bottom, left) shows that the particle filtering approach outperforms the others only if the available degradation model is very accurate: an error of 10% in the estimate of  $A$  is sufficient to decrease approach 1 performance below those of the other two methods. Similar results have been obtained varying the values of other parameters of the process model.

Finally, with respect to indicator 4, we have verified the performance of approach 2 varying the number of historical trajectories available for training the empirical model from 2 to 30. Figure 7 (bottom, right) shows that the  $rMSE$  made by approach 2 decreases when the number of historical trajectories increases from 2 to 12 and then tends to stabilize around a value very close to the performance of approach 1.

Considering the results obtained performing this sensibility analysis, we can provide the following guidelines for the choice of the prognostic approach: if one is very confident about the accuracy of the available physical degradation model, approach 1 should be preferred; on the contrary, if one doubts about the model accuracy, approach 2 is, in general, the most accurate, especially if the number of historical trajectories available is large. However, if the measurement noise is small, the system is close to failure and many degradation measurements have been taken during the current degradation trajectory, approach 3 can provide better accuracy.

When multiple approaches with comparable degree of accuracy are available, an alternative strategy to the choice of the best performing approach consists in the combination of the different approaches outcomes. This requires the development of a weighting strategy for the aggregation of the predictions made by the different approaches, based on their performances in the different situations of information available. For this, aggregation techniques proposed in literature for ensemble models [23][26] will be considered in a future work.

## 7 Conclusions

Different forms of information and data may be available for the prognosis of the RUL of a system undergoing degradation. In this work, we have considered three practical situations with decreasing amount of information available: in the first case the model of the degradation process is available, in the second case the model is not available but can be empirically derived from a number of observations collected during the degradation trajectories to failure of similar systems, in the third case only direct measurements of the degradation state reached during the life of the system of interest are available.

In this work, we have discussed the choice of the prognostic method in different information settings, considering the accuracy and the ability of providing measures of confidence in the RUL prediction of different prognostic approaches. In the first considered case, where a physical model of the degradation process is available, a particle filtering approach has been properly tailored to the prognostic problem, whereas a bootstrapped ensemble-based technique has been proposed and further developed to estimate the uncertainty of the RUL prediction in those situations where a priori knowledge of the mechanisms and models of the degradation process are missing (cases 2 and 3). For this, the prognostic problem has been reformulated, so that it was possible to define training and validation datasets of input/output pairs necessary for the construction of the prognostic model and the assessment of its accuracy; furthermore, solutions to ensure the independency between these two datasets have been developed. The merit of the proposed approach is that it allows producing a confidence interval for the RUL prediction, even when a measure of the prediction accuracy is not automatically provided by the regression method adopted.

The approaches proposed have been tested on a case study concerning the creep growth process in a high temperature turbine blade. The results show that both the particle filter and the bootstrapped ensemble methods provide a reliable prediction of the system RUL with a quantification of its uncertainty, the particle filter being the best performing method.

With respect to the ensemble of bootstrapped models trained with historical measurements of the degradation process in similar systems, the main limitation of the method is that it is not able to learn the peculiar characteristics of the system of interest but it tends to reproduce an ‘average’ degradation trajectory. To overcome this problem, a different modeling approach could be used, such as that based on the idea of fuzzy similarity [22], or a procedure for updating the ensemble with the information conveyed by new observations could be implemented [26].

For the application of the ensemble to the last case, in which only direct measurements of the degradation state reached during the life of the system of interest are available, the bootstrap method requires building an empirical model for the RUL variance estimate which is then used outside the region covered by the training data. Although good extrapolations have been obtained in the linear creep growth case study, the feasibility of the approach on more complex models should be verified.

Finally, we have investigated how the performance of the three proposed approaches varies depending on the quantity and quality of the available information. This sensitivity analysis has driven the development of a decision making policy for the identification of the prognostic approach which best suits the information setting, even in presence of mixed information sources.

## Acknowledgments

The work of Francesca Mangili has been supported by a PhD grant of the Institut For Energiteknikk (IFE), OECD Halden Reactor Project. The authors wish to thank the reviewers for their constructive comments which have allowed improving the paper.

## Appendix A: Particle Filtering

Given the first-order Markov process in eq. (1) and a set of observations  $\mathbf{z}_{1:i}$  related to the equipment degradation state  $d_j$  by eq. (2), we aim at predicting the filtered posterior pdf  $p_{D(t_i)}(d_i | \mathbf{z}_{1:i})$  at time  $t_i$ .

Within a Bayesian framework, the filtered posterior pdf  $p_{D(t_i)}(d_i | \mathbf{z}_{1:i})$  is recursively computed in two stages: prediction and update [16][31]. Given the pdf  $p_{D(t_{i-1})}(d_{i-1} | \mathbf{z}_{1:i-1})$  at time  $t_{i-1}$ , the prediction stage involves using the transition probability  $p_{D(t_i)}(d_i | d_{i-1})$  defined by the system equation (1) and the known distribution of the noise vector  $\gamma_{i-1}$  to obtain the prior probability distribution of the system state  $d_i$  at time  $t_i$  via the Chapman-Kolmogorov equation:

$$\begin{aligned}
p_{D(t_i)}(d_i | \mathbf{z}_{1:i-1}) &= \int p_{D(t_i)}(d_i | d_{i-1}, \mathbf{z}_{1:i-1}) p_{D(t_{i-1})}(d_{i-1} | \mathbf{z}_{1:i-1}) \mathrm{d}d_{i-1} \\
&= \int p_{D(t_i)}(d_i | d_{i-1}) p_{D(t_{i-1})}(d_{i-1} | \mathbf{z}_{1:i-1}) \mathrm{d}d_{i-1}
\end{aligned} \tag{A1}$$

where the Markovian assumption underpinning the system model (1) has been used.

At time  $t_i$ , a new observation  $\mathbf{z}_i$  is collected and used to update the prior distribution via Bayes rule, so as to obtain the required posterior pdf of the current state  $d_i$  [33]:

$$p_{D(t_i)}(d_i | \mathbf{z}_{1:i}) = \frac{p_{D(t_i)}(d_i | \mathbf{z}_{1:i-1}) p_{Z(t_i)}(\mathbf{z}_i | d_i)}{p_{Z(t_i)}(\mathbf{z}_i | \mathbf{z}_{1:i-1})} \tag{A2}$$

where the normalizing constant is

$$p_{Z(t_i)}(\mathbf{z}_i | \mathbf{z}_{1:i-1}) = \int p_{D(t_i)}(d_i | \mathbf{z}_{1:i-1}) p_{Z(t_i)}(\mathbf{z}_i | d_i) \mathrm{d}d_i \tag{A3}$$

The recurrence relations (A1) and (A2) form the basis for the exact Bayesian calculation of the pdf

$p_{D(t_i)}(d_i | \mathbf{z}_{1:i})$  at time  $t_i$ .

Unfortunately, except for a few cases, including linear Gaussian state space models (Kalman filter) and hidden finite-state space Markov chains (Wohnam filter), it is not possible to evaluate analytically these distributions, since they require the evaluation of complex high-dimensional integrals.

An alternative and effective approach is that of resorting to Monte Carlo sampling methods for integration.

This solution is based on sampling a large number  $K$  of trajectories  $\{d_{0:i}^k\}_{k=1}^K$  (called particles), from a suitably introduced importance function  $q(d_{0:i} | \mathbf{z}_{1:i})$ . In the following we briefly describe how these simulated trajectories can be utilized for filtering out the unobserved trajectory of the real degradation process. For more details, one can refer to the specialized literature, e.g., [11] and [16].

The posterior probability  $p_{D(t_i)}(d_i | \mathbf{z}_{1:i})$  we wish to calculate is the marginal of the probability

$p_{D(t_{0:i})}(d_{0:i} | \mathbf{z}_{1:i})$ , i.e., the multiple integral of this latter with respect to  $d_0, d_1, \dots, d_{i-1}$  in  $[-\infty, \infty]$ , which may be formally extended to include also the variable  $d_i$  by means of a  $\delta$ -function, viz.,

$$p_{D(t_i)}(d_i | \mathbf{z}_{0:i}) = \int p_{D(t_{0:i-1})}(d_{0:i-1}, u | \mathbf{z}_{1:i}) \delta(d_i - u) \mathrm{d}d_{0:i-1} \mathrm{d}u \tag{A4}$$

Then, by using the large number of trajectories  $\{d_{0:i}^k\}_{k=1}^K$  sampled from the importance function  $q(d_{0:i} | \mathbf{z}_{1:i})$ , the integral can be approximated as [10]:

$$\begin{aligned}
p_{D(t_i)}(d_i | \mathbf{z}_{0:i}) &= \int \left[ \frac{p_{D(t_{0:i-1})}(d_{0:i-1}, u | \mathbf{z}_{1:i})}{q(d_{0:i-1}, u | \mathbf{z}_{1:i})} \delta(k_i - u) \right] q(d_{0:i-1}, u | \mathbf{z}_{1:i}) \mathbf{d}d_{0:i-1} \mathbf{d}u \\
&\approx \frac{1}{K} \sum_{k=1}^K w_i^{*k} \delta(d_i - d_i^k)
\end{aligned} \tag{A5}$$

where  $w_i^k$  is the importance weight associated to the state sequence  $d_{0:i}^k$ ,  $k=1,2,\dots,K$ , sampled from  $q(d_{0:i} | \mathbf{z}_{1:i})$  and is given by

$$w_i^k = \frac{p_{D(t_i)}(d_{0:i}^k | \mathbf{z}_{1:i})}{q(d_{0:i}^k | \mathbf{z}_{1:i})} = \frac{p_{Z(t_i)}(\mathbf{z}_{1:i} | d_{0:i}^k) p_{D(t_i)}(d_{0:i}^k)}{p_{Z(t_i)}(\mathbf{z}_{1:i}) q(d_{0:i}^k | \mathbf{z}_{1:i})} \tag{A6}$$

Typically  $p_{Z(t_i)}(\mathbf{z}_{1:i})$  cannot be expressed in closed form. However, in [10] it is shown that the approximation in (A5) is equivalent to:

$$p_{D(t_i)}(d_i | \mathbf{z}_{0:i}) \approx \sum_{k=1}^K \frac{w_i^{*k}}{\sum_{k=1}^K w_i^{*k}} \delta(d_i - d_i^k) = \sum_{k=1}^K w_i^k \delta(d_i - d_i^k) \tag{A7}$$

where  $\tilde{w}_i^k$  and  $w_i^k$  are, respectively, the *unnormalised* and *normalized* importance weights:

$$\tilde{w}_i^k = \frac{p_{Z(t_i)}(\mathbf{z}_{1:i} | d_{0:i}^k) p_{D(t_{0:i})}(d_{0:i}^k)}{q(d_{0:i}^k | \mathbf{z}_{1:i})}, \quad w_i^k = \frac{w_i^{*k}}{\sum_{k=1}^K w_i^{*k}} \tag{A8}$$

It is often convenient to choose the importance density to be the transition probability:

$$q(d_i^k | \mathbf{z}_{1:i}) = p_{D(t_i)}(d_i^k | d_{i-1}^k) \tag{A9}$$

so that the importance function factorizes as follows

$$q(d_{0:i}^k | \mathbf{z}_{1:i}) = p_{D(t_i)}(d_i^k | d_{i-1}^k) p_{D(t_{0:i-1})}(d_{0:i-1}^k | \mathbf{z}_{1:i-1}) \tag{A10}$$



and one can obtain samples by augmenting each of the existing ones  $d_{0:i-1}^k$  with the new state  $d_i^k$  sampled from  $p_{D(t_i)}(d_i^k | d_{i-1}^k)$ .

Using the Bayes rule, the hypothesis of Markovianity of the process and the fact that the observation  $\mathbf{z}_i$  depends on the state  $d_i$  only, i.e.,  $p_{Z(t_i)}(\mathbf{z}_i | d_{0:i}^k) = p_{Z(t_i)}(\mathbf{z}_i | d_i^k)$ , the wights  $\tilde{w}_i^k$  defined in eq. (A8) can be rewritten as (details of the calculations can be found in [16]):

$$\begin{aligned} \tilde{w}_i^k &= \frac{p_{Z(t_i)}(\mathbf{z}_i | d_i^k) p_{D(t_i)}(d_i^k | d_{i-1}^k) p_{Z(t_{0:i-1})}(\mathbf{z}_{1:i-1} | d_{0:i-1}^k) p_{D(t_{0:i-1})}(d_{0:i-1}^k)}{p_{D(t_i)}(d_i^k | d_{i-1}^k) p_{D(t_{0:i-1})}(d_{0:i-1}^k | \mathbf{z}_{1:i-1})} \\ &= p_{Z(t_i)}(\mathbf{z}_i | d_i^k) w_{i-1}^k \end{aligned} \quad (\text{A11})$$

where  $p_{Z(t_i)}(\mathbf{z}_i | d_i^k)$  is the *likelihood* of the observation  $\mathbf{z}_i$ , which can be derived from the observation equation (2).

The resulting normalized weights  $w_i^k$  are then:

$$w_i^k = \frac{p_{Z(t_i)}(\mathbf{z}_i | d_i^k) w_{i-1}^k}{\sum_{k=1}^K p_{Z(t_i)}(\mathbf{z}_i | d_i^k) w_{i-1}^k} \quad (\text{A12})$$

## Appendix B: Bootstrapped ensemble-based estimate of the prediction uncertainty

Assume we are given a set of data pairs  $\{\mathbf{z}_i, y_i\}_{i=1}^N$ , generated according to

$$y_i = h(\mathbf{z}_i) + v_i \quad (\text{B1})$$

where  $y_i$  is the target value,  $h(\mathbf{z}_i)$  true input/output relation and  $v_i$  a process noise with zero mean and standard deviation  $\sigma_{v_i}^2(\mathbf{z}_i)$ . When we train a model  $f(\mathbf{z}_i)$  on such data, our aim is to approximate  $h(\mathbf{z}_i)$ ; such a model can be interpreted as an estimate of the mean distribution of the target values given an input vector  $\mathbf{z}_i$ . In many practical application, all the more in prognostics, it is highly desirable to have a measure of confidence in the prediction  $f(\mathbf{z}_i)$ . As described in [24] the uncertainty in the prediction  $f(\mathbf{z}_i)$  is quantified by the prediction error variance  $\sigma_f^2(\mathbf{z}_i)$  which can be decomposed in two terms:

$$\begin{aligned}
\sigma_f^2(\mathbf{z}_i) &= E[(f(\mathbf{z}_i) - y_i)^2] = \\
&= E\{[f(\mathbf{z}_i) - h(\mathbf{z}_i)]^2\} + E\{[h(\mathbf{z}_i) - y_i]^2\} \\
&= \sigma_m^2(\mathbf{z}_i) + \sigma_{v_i}^2(\mathbf{z}_i)
\end{aligned} \tag{B2}$$

where the term  $\sigma_m^2(\mathbf{z}_i)$  is the variance of the distribution of  $f(\mathbf{z}_i) - h(\mathbf{z}_i)$  and is concerned with the accuracy of the model  $f(\mathbf{z}_i)$  in estimating the true function  $h(\mathbf{z}_i)$ , whereas the term  $\sigma_{v_i}^2(\mathbf{z}_i)$  is the variance of the distribution of  $v_i = y_i - h(\mathbf{z}_i)$ , and is concerned with the accuracy of  $h(\mathbf{z}_i)$  in predicting the target  $y_i$  itself.

To generate an estimate of  $\sigma_m^2(\mathbf{z}_i)$  one can resort to a bootstrapped ensemble of models  $\{f^b(\mathbf{z}_i)\}_{b=1}^B$  built on bootstrapped replicates of the original set of data pairs  $\{\mathbf{z}_i, y_i\}_{n=1}^N$ . These replicates are obtained by randomly sampling with replacement  $N$  data pairs from the original dataset. As derived in [32], the bootstrapped outputs  $\{f^b(\mathbf{z}_i)\}_{b=1}^B$  provide us with the empirical estimate of the distribution of  $f(\mathbf{z}_i) - h(\mathbf{z}_i)$ . This estimate is given by the distribution of  $f(\mathbf{z}_i) - \overline{f^b(\mathbf{z}_i)}$ , where  $\overline{f^b(\mathbf{z}_i)}$  is the average value of the predictions  $\{f^b(\mathbf{z}_i)\}_{b=1}^B$  and is retained as the estimate of the true function  $h(\mathbf{z}_i)$  to which we have no access. Thus, under the hypothesis that  $\overline{f^b(\mathbf{z}_i)}$  is an unbiased estimator of  $h(\mathbf{z}_i)$  and that the distribution  $f(\mathbf{z}_i) - h(\mathbf{z}_i)$  is Gaussian (see [32] for more details), the term  $\sigma_m^2(\mathbf{z}_i)$  can be estimated as

$$\hat{\sigma}_m^2(\mathbf{z}_i) = \frac{1}{B} \sum_{b=1}^B [f^b(\mathbf{z}_i) - \overline{f^b(\mathbf{z}_i)}]^2 \tag{B3}$$

To estimate total variance  $\sigma_f^2(\mathbf{z}_i)$ , we need to build a model that provide an estimate for the noise term  $\sigma_{v_i}^2(\mathbf{z}_i)$  in correspondence of an input  $\mathbf{z}_i$ . Such a model is found by fitting the residuals

$$\{r^2(\mathbf{z}_i) = [y_i - \overline{f^b(\mathbf{z}_i)}]^2 - \sigma_m^2\}_{j=1}^{N'} \tag{B4}$$

which should be calculated on a validation dataset different than the training set to avoid overfitting and underestimating of the variance  $\sigma_{v_i}^2(\mathbf{z}_i)$ . In other words, we train a model  $\chi(\mathbf{z}_i)$  on the set of input output pairs  $\{\mathbf{z}_i, r^2(\mathbf{z}_i)\}_{n=1}^{N'}$  by maximizing the loglikelihood function

$$LL = \sum_{i=1}^{N'} \log \left( \frac{1}{\sqrt{2\pi} \chi(\mathbf{z}_i)} \exp \left( -\frac{r^2(\mathbf{z}_i)}{2\chi^2(\mathbf{z}_i)} \right) \right) \tag{B5}$$

written under the hypothesis that the residuals  $r^2(\mathbf{z}_i)$  have a Gaussian distribution with zero mean and variance which can be shown to be equal to  $\sigma_{v_i}^2(\mathbf{z}_i)$ :

$$\begin{aligned}
 E[r^2(\mathbf{z}_i)] &= E\{[y_i - f(\mathbf{z}_i)]^2 - \sigma_m^2(\mathbf{z}_i)\} \\
 &= E\{[y_i - f(\mathbf{z}_i)]^2 - \sigma_m^2(\mathbf{z}_i)\} \\
 &= \sigma_f^2(\mathbf{z}_i) - \sigma_m^2(\mathbf{z}_i) \\
 &= \sigma_{v_i}^2(\mathbf{z}_i)
 \end{aligned} \tag{B6}$$

where eq. (B2) has been used in the last equivalence of eq. (B6).

Finally, the prediction error variance in correspondence of the input  $\mathbf{z}_i$  can be approximated by:

$$\sigma_f^2(\mathbf{z}_i) = \hat{\sigma}_m^2(\mathbf{z}_i) + \hat{\sigma}_{v_i}^2(\mathbf{z}_i) = \left\{ \frac{1}{B} \sum_{b=1}^B [f^b(\mathbf{z}_i) - \overline{f^b(\mathbf{z}_i)}]^2 \right\} + \chi^2(\mathbf{z}_i) \tag{B7}$$

## References

- [1] G. Vachtsevanos, F. L. Lewis, M. Roemer, A. Hess and B. Wu, *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*, 1st edition, John Wiley & Sons, Hoboken, 2006.
- [2] D. B. Jarrell, D. R Sisk and L. J. Bond, *Prognostics and Condition-Based Maintenance: A New Approach to Precursive Metrics*, *Nucl. Technol.*, 145, 2004, 275-286.
- [3] J. W. Hines and A. Usynin, *Current Computational Trends in Equipment Prognostics*, *Int. J. Comput. Intell. Syst.*, 1(1), 2008, 94-102.
- [4] E. Zio, *Prognostics and Health Management of Industrial Equipment*, submitted for publication.
- [5] T. Brotherton, G. Jahns, J. Jacobs and D. Wroblewski, *Prognosis of faults in gas turbine engines*, *IEEE Aerosp. Conf. Proc.*, 6, 2000, 163-171, 18-25 Mar, Big Sky, MT, USA.
- [6] J. Luo, K. Pattipati, L. Qiao, and S. Chigusa, *Model-based Prognostic Techniques Applied to a Suspension System*, *IEEE Trans. on Syst., Man, and Cybern.*, 38(5), 2008, 1156-1168.
- [7] U. Pulkkinen, *A stochastic model for wear prediction through condition monitoring*, in: K. Holmberg, A. Folkesson (Eds.), *Operational reliability and systematic maintenance*, Elsevier, London (New York), 1991, pp. 223-43.
- [8] A. Ray and S. Tangirala, *A nonlinear stochastic model of fatigue crack dynamics*, *Prob. Engng Mech.*, 12(1), 1997, 33-40.
- [9] R.W. Swindeman, and M.J. Swindeman, *A comparison of creep models for nickel base alloy for advanced energy systems*, *Int. J. of Press. Vessel. and Pip.*, 85, 2008, 72-79.
- [10] A. Doucet, *On sequential simulation-based methods for Bayesian filtering*, Technical report. Dept. of Engng, University of Cambridge, CUED-F-ENGTR310, 1998.
- [11] A. Doucet, J.F.G. de Freitas and N.J. Gordon. *Sequential Monte Carlo methods in practice*. Springer-Verlag, New York, 2001.
- [12] B.D. Anderson, J.B. Moore, *Optimal filtering*, Prentice Hall, Englewood Cliffs (NJ), 1979.
- [13] G. Kitagawa, *Non-Gaussian state-space modeling of nonstationary time series*, *J. of the Am. Stat. Assoc.*, 82, 1987, 1032-63.
- [14] F. Cadini, E. Zio and D. Avram, *Monte Carlo-based filtering for fatigue crack growth estimation*, *Prob. Engng Mech.*, 24, 2009, 367-373.
- [15] T. Khan, *Particle filter based prognosis study for predicting remaining useful life of steam generator tubing*, *IEEE Conference on Prognostics and Health Management (PHM)*, 20-23 June 2011, Montreal, QC.
- [16] M.S. Arulampalam, S Maskell, N Gordon, and T Clapp. *A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking*. *IEEE Trans. Signal Process*, 50(2), 2002, 174-88.
- [17] M.A. Schwabacher, *A Survey of Data-Driven Prognostic*, *Infotech@Aerospace*, 26 - 29 September 2005, Arlington, Virginia.

- [18] K. Goebel, B. Saha, and A. Saxena, A Comparison of Three Data-Driven Algorithms for Prognostics, Proc. of the 62nd Meet. of the Soc. for Mach. Fail. Prev. Technol., 2008, pp.119-131.
- [19] L. Peel, Data driven prognostics using a Kalman filter ensemble of neural network models, Proc. Int. Conf. on prong. and health manag., 2008.
- [20] B. Saha, K. Goebel, and J. Christophersen, Comparison of Prognostic Algorithms for Estimating Remaining Useful Life of Batteries, Trans. of the Inst. of Meas. & Control, 31(3-4), 2009, 293-308.
- [21] F. Di Maio, J. Hu, P. Tse, K. Tsui, E. Zio, and M. Pecht, Ensemble-approaches for clustering health status of oil sand pumps, Expert Syst. with Appl., doi: 10.1016/j.eswa.2011.10.008.
- [22] E. Zio, and F. Di Maio, A Fuzzy Similarity-Based Method for Failure Detection and Recovery Time Estimation. Int. J. of Perform. Engng, 6(5), 2010, 407-424.
- [23] R. Polikar, Ensemble Based Systems in Decision Making, IEEE Circuits and Syst. Mag., 6(3), 2006, 21-45.
- [24] T. Heskes, Practical confidence and prediction intervals, in: M.Mozer, M. Jordan, and T. Heskes (Eds.), Advances. Neural Information Processing Systems 9, MIT Press, Cambridge, 1997, pp. 466–472.
- [25] R. Couturier, and C. Escaravage, High temperature alloys for the HTGR Gas Turbine: Required properties and development needs, IAEA-TECDOC-1238, 2000.
- [26] P. Baraldi, F. Mangili, and E. Zio, A Kalman Filter - based Ensemble Approach for Turbine Creep Prognostics. Submitted for publication, 2011.
- [27] E. Zio, A Study of the Bootstrap Method for Estimating the Accuracy of Artificial Neural Networks in Predicting Nuclear Transient Processes. IEEE Trans. Nucl. Sci., 53(3), 2006, 1460-1478.
- [28] N. Gorjian, L. Ma, M. Mittinty, P. Yarlalagadda, and Y. Sun, Review on Degradation Models in Reliability Analysis, Proc. of the 4th World Congr. on Engng Asset Manag., Athens, 28-30 Sept. 2009.
- [29] M. Saez, N. Tauveron, T. Chataing, G. Geffraye, L. Briottet, and N. Alborghetti, Analysis of the turbine deblading in an HTGR with the CATHARE code, Nucl. Engng and Design, 236, 2006, 574–586.
- [30] M. Orchard, A Particle Filtering-based Framework for On-line Fault Diagnosis and Failure Prognosis, Ph.D. Thesis, Dept. of Electrical and Computer Engng, Georgia Institute of Technology, 2007.
- [31] P. Li and V. Kadiramanathan, “Particle Filtering based Likelihood Ratio Approach to Fault Diagnosis in Nonlinear Stochastic Systems”, in IEEE Trans. On Systems, Man, And Cybernetics, Part C: Applications and Reviews, vol. 31, no 3, pp. 337-343, 2001.
- [32] Carney, J.G. Cunningham, P., Bhagwan, U., Confidence and prediction intervals for neural network ensembles, International Joint Conference on Neural Networks (IJCNN), Jul 1999, Vol. 2, pp. 1215-1218

## Figures

```

FOR k=1:K
  1. Sample  $d_0 \sim p(d_0)$ 
  2.  $j=0$ ;  $t_j = t_0$ 
  WHILE  $d_j \leq d_{th}$ 
    3.  $j=j+1$ ;  $t_j = t_{j-1} + \Delta t$ 
    4. sample  $d_j^k \sim p(d_j^k | d_{j-1}^k)$ 
  END WHILE
END FOR

FOR i=1:N
  5. Collect the observation  $z_i$ 
  FOR k=1:K
    6. Compute the weights:
      IF  $d_i^k < 0$  set  $w_i^k = 0$ 
      ELSE  $w_i^k = w_{i-1}^k \cdot \Pr(z_i | d_i^k)$ 
    END FOR
  7. Normalize the weights  $wn_i^k = w_i^k / \sum_{k=1}^K w_i^k$ 
  8. Build the cumulative density function of the
     system RUL at time  $t_i$  as:
     
$$F_{RUL_i}(rul_i) = \sum_k w_i^k \text{ with } k | d_{t_i+rul_i}^k > d_{th}$$

  END FOR

```

Figure 1: Particle filtering operative procedure for estimation of the RUL cumulative distribution.

At time  $t_i$

```

1. Compute  $wn_i^k$  as in Figure 1
FOR k=1:K
  2. Sample a particle  $u$  with probability equal
     to its weight  $wn_i^u$ 
  3. set  $d_i^k = d_i^u$ ,  $j=i$  and  $t_j = t_i$ 
  WHILE  $d_j < d_{th}$ 
    4.  $j=j+1$ ;  $t_j = t_{j-1} + \Delta t$ 
    5. sample  $d_j^k \sim \Pr(d_j^k | d_{j-1}^k)$ 
  END WHILE
END FOR
6. Assign equal weights to each particle  $wn_i^{k:K} = 1/K$ 
7. Build the cumulative density function of the
   system RUL at time  $t_i$  as:
   
$$F_{RUL_i}(rul_i) = \sum_k w_i^k \text{ with } k | d_{t_i+rul_i}^k > d_{th}$$


```

Figure 2: Procedure for performing resampling at time  $t_i$ .

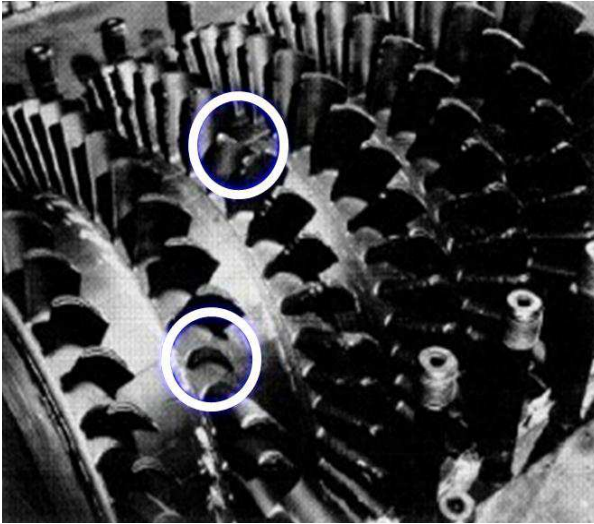


Figure 3: Deblading in a high pressure turbine [30].

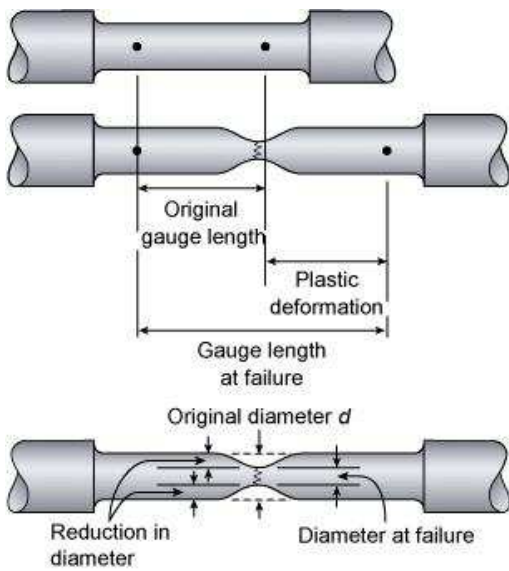


Figure 4: Schematics of a specimen before and after a creep test (<http://www.twi.co.uk/content/jk69.html>).

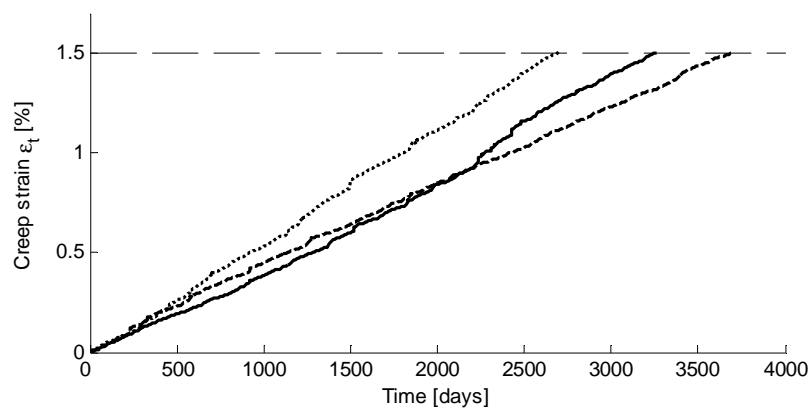


Figure 5: Examples of creep growth paths.

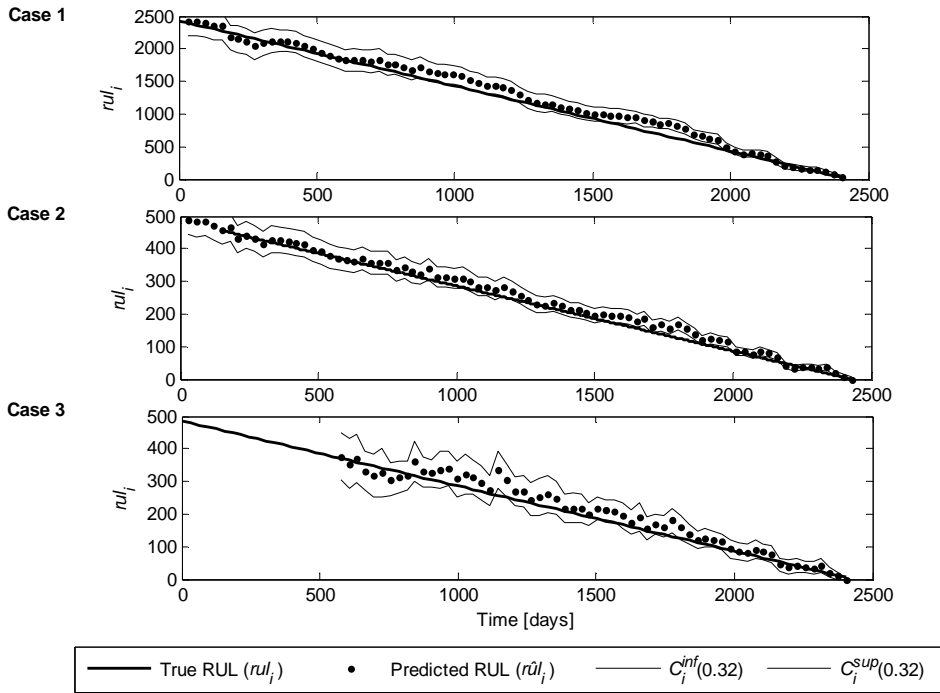


Figure 6: true RUL (continuous thick line) of a turbine blade with its predicted value  $\hat{rul}_i$  (dots) and prediction interval (continuous thin line) for the three prognostic cases.

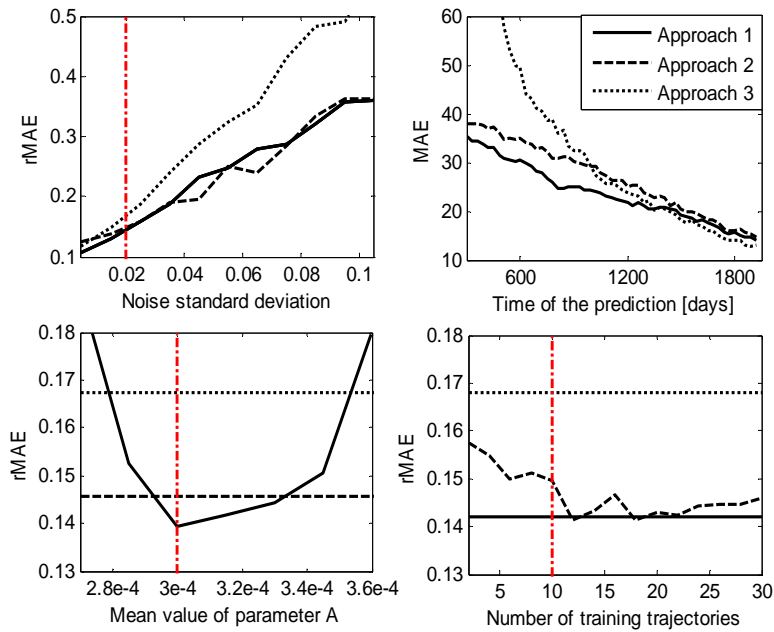


Figure 7: performance of the proposed approaches in different settings of information available. The vertical (red) line indicates the value assigned to the parameter in the numerical application of Section 5.

## Tables

Table I: main sources of information for prognostics

Source	Description	Mathematical representation
A	Dynamic model of the degradation process	eq.(1)
B	Sequence of observations related to the degradation of the system collected at $t_j = 1, 2, \dots, i$	$\mathbf{z}_{1:i}$
C	Historical sequences of observations related to the degradation of a set of $S$ failed systems collected at $N_s$ time instants $t_j$ ; $s=1, \dots, S$ ; $j=1, \dots, N_s$	$\mathbf{z}_{1:N_s}^s, s=1, \dots, S$
D	Value of the failure threshold	$d_{th}$
E	Measurement equation	eq.(2)
F	Durations of lives of the set of $S$ failed systems $s$ .	$L^s, s=1, \dots, S$

Table II: information available in each of the three prognostic cases considered

Source of information	Case 1	Case 2	Case 3
A	X		
B	X	X	X
C		X	
D	X		X
E	X		X
F		X	

Table III: type of distribution, mean value and standard deviation used for the creep growth model parameters

Variable	Symbol	Distribution	Units	Parameters of the distribution
Activation energy	$Q$	Deterministic	kJ/mol	$Q=290$
Norton Law parameters	$A$	Normal	$(\text{N/m}^2)^{-n}/\text{h}$	$\mu_A=3 \cdot 10^{-4}$ ; $\sigma_A=5\%$
	$N$	Normal	-	$\mu_n=6$ ; $\sigma_n=0.2\%$
Operating temperature	$T_i$	Normal	K	$\mu_T=1100$ ; $\sigma_T=1\%$
Rotational speed	$\omega_i$	Normal	rpm	$\mu_\omega=3000$ ; $\sigma_\omega=1\%$
Density	$P$	Deterministic	$\text{Kg/m}^3$	$\rho=8000$
Hub radius	$r_{hub}$	Deterministic	m	$r_{hub}=0.7$
Tip radius	$r_{tip}$	Deterministic	m	$r_{tip}=0.87$
Stress fluctuations	$\delta\varphi$	Gamma	MPa	$\theta=2$ ; $k=10$

Table IV: main sources of information for prognostics of a creeping turbine blade

Source	Description	Mathematical representation
A	The creep growth model and the distributions of the model parameters	eq. (27) and Table III
B	Measurements of the creep strain of the currently creeping blade taken at $i$ different time instants $t_j$	$\mathbf{z}_{1:i} = \boldsymbol{\varepsilon}_{1:i} + \mathbf{v}_{1:i}$
C	Historical measurements of the creep strain of a set of $S$ blades failed for creeping, taken at $N_s$ different time instants $t_j$	$\mathbf{z}_{1:N_s}^s, s=1, \dots, S$
D	The value of the failure threshold	$d_{th} = \boldsymbol{\varepsilon}_{th}$
E	The measurement equation and the noise distribution	eq. (28) and $\text{Pr}(\boldsymbol{\sigma}_v)$
F	The length of life $L^s$ of the set of $S$ failed blades.	$L^s, s=1, \dots, S$



Table V: estimates of  $\mu_{RUL_i}$  and  $\sigma_{rUL_i}^2$  at time  $t_{50} = 1475$  (first row) and  $t_{80} = 2375$  (second row) in the three prognostic cases considered.

$t_i$	$rUL_i$	$\mu_{RUL_i}$	$\sigma_{RUL_i}^2$	Case 1		Case 2		Case 3	
				$r\hat{UL}_i$	$\hat{\sigma}_{r\hat{UL}_i}^2$	$r\hat{UL}_i$	$\hat{\sigma}_{r\hat{UL}_i}^2$	$r\hat{UL}_i$	$\hat{\sigma}_{r\hat{UL}_i}^2$
1475	1110	1092	90	1085	107	1079	109	1075	238
2375	210	264	42	247	45	167	63	248	57

Table VI: prognostic performance in the three prognostic cases considered

	$\overline{rMAE}$	$\overline{Cov}$
<b>Case 1</b>	0.150±0.009	0.663±0.018
<b>Case 2</b>	0.172±0.009	0.613±0.019
<b>Case 3</b>	0.170±0.009	0.682±0.014

Table VII: prognostic performance in case 1 when parameters  $A$  and  $n$  are assumed known and in case 2 when parameters  $A$  and  $n$  are kept constant for all historical training trajectories

	$\overline{rMAE}$	$\overline{Cov}$
<b>Case 1b</b>	0.135±0.009	0.669±0.019
<b>Case 2b</b>	0.145±0.007	0.623±0.016