



**HAL**  
open science

# Investigation of uncertainty treatment capability of model-based and data-driven prognostic methods using simulated data

Piero Baraldi, Francesca Mangili, Enrico Zio

► **To cite this version:**

Piero Baraldi, Francesca Mangili, Enrico Zio. Investigation of uncertainty treatment capability of model-based and data-driven prognostic methods using simulated data. Reliability Engineering and System Safety, 2013, 112, pp.94-108. 10.1016/j.ress.2012.12.004 . hal-00934547

**HAL Id: hal-00934547**

**<https://centralesupelec.hal.science/hal-00934547>**

Submitted on 22 Jan 2014

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

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Investigation of uncertainty treatment capability of model-based and data-driven prognostic methods using simulated data

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

<sup>a</sup>*Dipartimento di Energia, Politecnico di 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, France*

## Abstract

We look at different prognostic approaches and the way of quantifying confidence in equipment Remaining Useful Life (RUL) prediction. More specifically, we consider: 1) a particle filtering scheme, based on a physics-based model of the degradation process; 2) a bootstrapped ensemble of empirical models trained on a set of degradation observations measured on equipments similar to the one of interest; 3) a bootstrapped ensemble of empirical models trained on a sequence of past degradation observations from the equipment of interest only.

The ability of these three approaches in providing measures of confidence for the RUL predictions is evaluated in the context of a simulated case study of interest in the nuclear power generation industry and concerning turbine blades affected by developing creeps.

The main contribution of the work is the critical investigation of the capabilities of different prognostic approaches to deal with various sources of uncertainty in the RUL prediction.

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

## 1 Introduction

In prognostics the current system condition is projected in time by a predictive model [1-2]. Since the prediction of the Remaining Useful Life (RUL) of degrading equipment is performed in the absence of future measurements concerning equipment degradation and operational conditions, the prognostic task is necessarily affected by large uncertainty. In this work, the sources of uncertainty affecting the RUL prediction are classified in three categories:

- A. Randomness in the future degradation of the equipment. This intrinsic uncertainty in the degradation process has several causes such as the unknown future load profile, and operation and environmental conditions.
- B. Modeling error, i.e., inaccuracy of the prognostic model used to perform the prediction. In model-based prognostic approaches, this source of uncertainty takes into account the assumptions and simplifications made on the form and structure of the model, and the uncertainty on the model parameters. In data-driven approaches, it relates to the incomplete coverage of the data set used to train the empirical model.

---

<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).

**Table I: nomenclature**

Symbol	Description	Symbol	Description
$t_i$	$i$ -th time instant	$h$	reference index of the ensemble models
$d_i$	equipment degradation at time $t_i$	$H$	number of empirical models
$\mathbf{z}_i$	physical observations related to $d_i$	$r_i^s$	prediction residual obtained in correspondence of the input $\mathbf{z}_i^s$
$RUL_i$	random variable (rv) representing system RUL at time $t_i$	$\chi$	
$\delta_i$	difference between $RUL_i$ and its expected value $E[RUL_i]$	$\hat{d}_i$	prediction of $d_i$ at time $t_i$
$rul_i$	realization of $RUL_i$	$\eta(t_i)$	empirical model for the prediction of $d_i$
$f(\mathbf{z}_i)$	prognostic model for the prediction of $rul_i$	$\mathbf{d}_{j,j'}$	vector of the degradation states at $t_j$ and $t_{j'}$
$\hat{r}ul_i$	prediction of $rul_i$	$\Delta t_{j,j'}$	rv representing the time interval to evolve from $d_j$ to $d_{j'}$
$\sigma_{\hat{r}ul_i}^2$	prediction error variance	$\tilde{\eta}(\mathbf{d}_{j,j'})$	
$\sigma_A^2$	prediction error variance due to randomness in the future degradation of the equipment	$\hat{\Delta}_{j,j'}$	estimate of $\Delta t_{j,j'}$
$\sigma_B^2$	prediction error variance due to modeling error	$\delta_{\Delta t_{j,j'}}$	difference between $\Delta t_{j,j'}$ and its mean value
$\sigma_C^2$	prediction error variance due to uncertainty in equipment degradation data	$r_{j,j'}$	prediction residual obtained in correspondence of the input $\mathbf{d}_{j,j'}$
$s$	index of the equipment	$\varepsilon_j$	creep strain at time $t_j$
$S$	number of equipments	$Q$	creep activation energy
$L_s$	failure time of the $s$ -th equipment	$\omega_j$	turbine rotational speed
$d_{th}$	failure thresholds	$K$	constant relating the load to $\omega_j$
$\Pr(x)$	probability distribution function of the rv $x$	$R$	ideal gas constant
$\Pr(x y)$	conditional probability distribution function of the rv $x$ given $y$	$T_j$	blade operating temperature at time $t_j$
$\mu_x$	mean value of the rv $x$	$\delta\varphi_j$	stress fluctuations
$\sigma_x$	standard deviation of the rv $x$	$v_j$	creep strain measurement noise
$g$	transition function of the degradation state	$\varepsilon_{th}$	creep strain failure threshold
$\gamma_j$	process noise vector	$\alpha_{0,1}^h / \beta_{0,1}^h$	coefficients of the $h$ -th empirical model of approach 2/3
$\hat{\sigma}_v^2$	estimate of $\sigma_v^2$	$\gamma_{0,1,2} / \lambda$	coefficients of the empirical model of the residual variance $\chi$ of approach 2/3
$p$	index of the Monte Carlo sampled particle	$\Delta\varepsilon_{j,j'}$	
$P$	number of particles	$\Delta\mathbf{z}_{j,j'}$	observed degradation increment from $t_j$ to $t_{j'}$
$w_i^p$	weight of the $p$ -th particle at time $t_i$	$\alpha$	significance level of the confidence interval
$\mathbf{D}$	dataset made by the observations $\mathbf{z}_i$	$C_i^{\text{inf/sup}}(\alpha)$	inferior/superior bound of the $(1-\alpha)$ -confidence interval for the RUL prediction
$\mathbf{D}_{i/o}$	dataset of the input/output pairs $(\mathbf{z}_i^s, rul_i^s)$	$c_{conf}^\alpha$	
$trn/val$	apex/subscript indicating training/ validation sets of data		
$N^{trn/val}$	number of training/validation patterns		

C. Uncertainty in current and past equipment degradation data, which are used by the prognostic model to elaborate the RUL prediction. These data are usually acquired by sensors with some measurement noise or derived from diagnostic systems assessing the equipment health state with some degree of uncertainty. Other possible sources of uncertainty, which are not considered in this work, are the imperfect knowledge of the value of degradation beyond which the equipment can no longer perform its functions (failure threshold),

or the time at which the degradation process starts (degradation initiation). Furthermore, in this work we assume that the equipment degradation is caused by a single degradation mechanism, not considering the uncertainty on the degradation caused by the onset of other, possibly competing, mechanisms [3].

The challenge of managing uncertainties associated with prognostics has been recently addressed in [1-2,4-5]. Uncertainty management in prognostics entails to identify, classify and analyze uncertainty sources with the aim of associating to the RUL predictions provided by a prognostic model an estimate of its uncertainty [4-7], i.e., a measure of the expected degree of mismatch between the real and predicted equipment failure time. This information, provided in the form of a probability distribution of the equipment RUL, can be used by the decision maker to confidently plan maintenance actions, according to the desired risk tolerance [2]. In this context, the objective of the present paper is to contribute to the way of investigating the capabilities of different prognostic approaches to deal with the uncertainty in the RUL prediction. To this aim, the analysis is performed with respect to three previously developed approaches [8].

A first approach, hereafter named “approach 1”, is based on a mathematical model of the degradation process for the RUL prediction [9], embedded in a filtering method capable of accounting for the stochasticity of the process (source of uncertainty A) and the noise affecting the measurements (source of uncertainty C). Most filtering approaches rely on Bayesian methods and provide the probability distribution of the RUL [10-11]. The exact Kalman filter has been largely used in case of linear state space models and independent, additive Gaussian measurements and modeling noises, whereas analytical or numerical approximations of the exact solution (such as the Extended Kalman filter, the Gaussian-sum filters or the approximate grid-based filters [12]) have been applied in cases where the dynamics of degradation is non-linear and/or the associated noises are non-Gaussian [13]. Numerical approximations based on the Monte Carlo sampling technique have gained popularity for their flexibility and ease of design [14-17]. Among them, Particle Filtering (PF) is often considered a state-of-the-art technology in the prognostic field and used as a term of comparison for newer approaches. The model-based particle filter approach here considered was firstly applied to state estimation for diagnostics [18-19] and then applied to prognostics [20-21]. According to the particle filtering scheme proposed in [20], the RUL distribution prediction is performed by considering the stochastic model of the degradation process and the on-line observations of the equipment degradation. In [22], this particle filtering-based prognostic approach is discussed with respect to the design of a predictive maintenance strategy, whose advantages are then compared with those of other maintenance strategies.

We consider also two data-driven approaches [23], based on statistical models that ‘learn’ trends from historical data. In particular, we consider bootstrap ensemble approaches [24-25], which are based on the aggregation of multiple model outcomes and have gained interest due to their ability of estimating the uncertainty in the predictions. These approaches allow estimating the model uncertainty (source of uncertainty B) by considering the variability in the predictions of the diverse models of the ensemble [24]. On the other hand, the estimate of the uncertainty due to the stochasticity of the degradation process (source A) and the input noise (source C), requires to investigate the relation between the input and the error of the

prognostic model based on its performance on a validation dataset. In what we will refer to as “approach 2”, a bootstrap ensemble model is built to estimate the equipment RUL based on sequences of observations of evolution to failure of a set of similar equipments operating under similar conditions; in what we will refer to as “approach 3”, a bootstrap ensemble model is built based on a sequence of degradation observations only of the equipment whose RUL we want to predict. Although approaches 2 and 3 are both based on the development of an ensemble of bootstrap models, they differ for the type of model used. Whereas in approach 2 we can directly model the RUL as a function of the observed parameters, in approach 3 we have to model the degradation evolution as a function of time, since direct RUL observations are not available. The three approaches are investigated with reference to the creep growth process in the turbine blades of a Gas Turbine Modular Helium nuclear Reactor (GT-MHR) [26-27]. The data used in this case study have been numerically simulated using a traditional model of the creep growth. Artificial data have been used in order to allow testing the three approaches on a large number of different blade degradation trajectories and thus evaluate their capability of correctly estimating the uncertainty on the provided RUL prediction. The remainder of the paper is organized as follows: Section 2 presents the terminology used in the paper and the problem setting; in Section 3, the decomposition of the prediction error variance into three terms corresponding to the randomness in the future degradation of the equipment, the modeling error, and the uncertainty in current and past equipment degradation measures is reported; in Section 4, the three considered prognostic approaches are described; in Section 5, the problem of blade creeping in high temperature turbines is illustrated and the capability of uncertainty management of the three prognostic approaches are discussed; finally, in Section 6 some conclusions are drawn and potential future work suggested.

## 2 Terminology and problem setting

In this work, we assume that the equipment is subject to a single degradation mechanism described as a random process; we do not consider the effects that other competing degradation mechanisms can have on the equipment degradation. Also, we assume that degradation cannot exceed a maximum acceptable level, hereafter referred to as “failure threshold”,  $d_{th}$ , which is fixed and identical among similar equipment. Notice that in prognostics the failure threshold does not necessarily indicate complete failure of the system, but, for safety margins, it is often set at a conservative value of the degradation limit beyond which the risk of complete failure exceeds tolerance limits or the performance of the system does not fulfill the requirements [28]. Since the failure threshold is usually derived from expert knowledge or from experimental measurements of the equipment degradation at failure, its estimate is typically affected by uncertainty which contributes to increase the RUL prediction error variance. Methods to deal with the uncertainty on the failure threshold have been proposed in [29-30], and, for this reason, this aspect is not considered in the present work.

We indicate by  $d_i$  the equipment degradation level at time  $t_i$  and we assume that its direct measure is not available, but some physical observations  $\mathbf{z}_i$  related to it are obtained. We indicate by  $\mathbf{z}_{1:i}=(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_i)$  the past and present observations taken at times  $t_1, t_2, \dots, t_i$  for the equipment whose RUL we want to predict. In

some cases, also the sequences of observations of evolution to failure of a set of  $S$  similar equipments operating under similar conditions are available and we will indicate by  $\mathbf{z}_{1:N_s}^s$ ,  $s=1, \dots, S$ , the observations taken at times  $t_1, \dots, t_{N_s}$  for the  $s$ -th equipment whose failure time is  $L_s$ , where  $L_s \geq t_{N_s}$ .

The objective of prognostics is the estimation of the equipment RUL, i.e., the time left from the current time  $t_i$  before the equipment degradation, currently of value  $d_i$ , crosses the failure threshold  $d_{th}$ . As degradation evolves randomly in time, the equipment RUL at time  $t_i$  is a random variable which will be referred to as  $RUL_i$ . Thus, the objective of applying prognostics to an equipment of current degradation level  $d_i$  is to estimate the probability density function (pdf)  $\Pr(RUL_i | d_i)$ . The uncertainty described by such distribution regards the future stochastic evolution of the equipment degradation and, thus, it is irreducible.

A realization  $rul_i$  of the random variable  $RUL_i$  can be written as:

$$rul_i = \mu_{RUL_i|d_i} + \delta_i \quad (1)$$

where  $\mu_{RUL_i|d_i}$  is the RUL expected value of the equipment with degradation  $d_i$  at time  $t_i$ , and  $\delta_i$  is a random variable with zero mean and variance  $\sigma_A^2$  which represents the uncertainty on the future evolution of degradation (source of uncertainty A).

Furthermore, in practice, the ‘exact’ model,  $g$ , of the equipment degradation process is not available (source of uncertainty B) and the degradation  $d_i$  at time  $t_i$  is not exactly known (source of uncertainty C). In this setting, the complete distribution of  $RUL_i$  cannot be derived and prognostics is limited to estimating:

- the expected value of  $RUL_i$
- the variance of the prediction error as a measure of the accuracy with which the estimated expected value predicts the actual RUL value.

With respect to the estimate of the expected value of  $RUL_i$ , it will be indicated by  $r\hat{u}l_i$  and considered as our RUL prediction. The prognostic model which generates at time  $t_i$  the estimate  $r\hat{u}l_i$  of  $RUL_i$  on the basis of the observations  $\mathbf{z}_i$  will be referred to as  $f$ , i.e.  $r\hat{u}l_i = f(\mathbf{z}_i)$ . Finally, we indicate by  $\sigma_{r\hat{u}l_i}^2$  the estimate of the prediction error variance, defined by  $\sigma_{r\hat{u}l_i}^2 = E[(r\hat{u}l_i - RUL_i)^2]$ .

### 3 Prediction error variance

According to [24], the prediction error variance  $\sigma_{r\hat{u}l_i}^2$  can be decomposed into two terms:  $\sigma_A^2$ , the variance related to the uncertainty on the future degradation of the equipment whose degradation at time  $t_i$  is  $d_i$  (source of uncertainty A), and  $\sigma_{B+C}^2$ , the variance related to the imprecision of the model  $f(\mathbf{z}_i)$  (source of uncertainty B) and the noise on the data  $\mathbf{z}_i$  (source of uncertainty C):

$$\begin{aligned} \sigma_{r\hat{u}l_i}^2 &= E[(RUL_i - r\hat{u}l_i)^2] = \sigma_A^2 + \sigma_{B+C}^2 = \\ &= E\left[\left(RUL_i - \mu_{RUL_i|d_i}\right)^2\right] + E\left[\left(\mu_{RUL_i|d_i} - f(\mathbf{z}_i)\right)^2\right] \end{aligned} \quad (2)$$

In some applications, it can be useful to distinguish the uncertainty due to the modeling error (source B) from that due to the noise on the input data (source C). To this aim, we introduce the quantity  $\mu_{RUL_i|z_i}$  which represents the  $RUL_i$  expected value of a degrading equipment for which at time  $t_i$  we have the observations  $\mathbf{z}_i$ , and we assume that  $f(\mathbf{z}_i)$  is an unbiased estimator of  $\mu_{RUL_i|z_i}$ . Thus we obtain:

$$\begin{aligned}\sigma_{B+C}^2 &= E\left[\left(f(\mathbf{z}_i) - \mu_{RUL_i|d_i}\right)^2\right] = E\left[\left(f(\mathbf{z}_i) - \mu_{RUL_i|z_i} + \mu_{RUL_i|z_i} - \mu_{RUL_i|d_i}\right)^2\right] \\ &= E\left[\left(f(\mathbf{z}_i) - \mu_{RUL_i|z_i}\right)^2\right] + E\left[\left(\mu_{RUL_i|z_i} - \mu_{RUL_i|d_i}\right)^2\right] \\ &= \sigma_B^2 + \sigma_C^2\end{aligned}\quad (3)$$

Combining eqs. (2) and (3), one obtains:

$$\begin{aligned}\sigma_{\hat{r}l_i}^2 &= E[(r\hat{l}_i - RUL_i)^2] = \\ &= E\left[\left(\mu_{RUL_i|d_i} - RUL_i\right)^2\right] + E\left[\left(f(\mathbf{z}_i) - \mu_{RUL_i|z_i}\right)^2\right] + E\left[\left(\mu_{RUL_i|z_i} - \mu_{RUL_i|d_i}\right)^2\right] \\ &= \sigma_A^2 + \sigma_B^2 + \sigma_C^2\end{aligned}\quad (4)$$

Notice that these results have been obtained by assuming that the different components of the prediction error are independent and thus the expected values  $E[(RUL_i - \mu_{RUL_i|d_i})(\mu_{RUL_i|d_i} - f(\mathbf{z}_i))]$  and  $E[(f(\mathbf{z}_i) - \mu_{RUL_i|z_i})(\mu_{RUL_i|z_i} - \mu_{RUL_i|d_i})]$  in eqs. (2) and (3), respectively, are zero.

## 4 Modeling approaches for RUL prediction

This Section illustrates briefly the three modeling approaches considered for RUL prediction.

### 4.1 Approach 1: Particle Filtering

In approach 1, a Monte Carlo-based filtering technique, called particle filtering [10,12], is used to predict the pdf  $\Pr(RUL_i | \mathbf{z}_{1:i})$  of the equipment RUL at time  $t_i$ . The prediction is based on the following information: a sequence of observations  $\mathbf{z}_{1:i}$  related to the equipment degradation at times  $t_1, t_2, \dots, t_i$ , the (observation) equation describing the relation between  $\mathbf{z}_i$  and the degradation level  $d_i$  at time  $t_i$ , the failure threshold  $d_{th}$ , and the (stochastic) model of the equipment degradation dynamics, e.g., described by a first-order Markov process:

$$d_j = g(d_{j-1}, \boldsymbol{\gamma}_{j-1}); d_0 \sim \Pr(d_0), j = 1, 2, \dots \quad (5)$$

where  $\Pr(d_0)$  is the initial distribution of the degradation at time  $t_0$ ,  $g$  is the possibly non-linear state transition function and  $\boldsymbol{\gamma}_j$  is the noise vector.

The estimation of the probability distribution  $\Pr(d_i | \mathbf{z}_{1:i})$  of the degradation  $d_i$  at time  $t_i$  given the set of observations  $\mathbf{z}_{1:i}$  is obtained by a recursive computational procedure divided into successive prediction and update stages [12]. In the prediction stage, supposing that the probability distribution function (pdf)  $\Pr(d_{i-1} | \mathbf{z}_{1:i-1})$  at time  $t_{i-1}$  is available, the transition probability distribution  $\Pr(d_i | d_{i-1})$  derived from the model in eq. (5) is used to obtain the prior pdf of the degradation state  $\Pr(d_i | \mathbf{z}_{1:i-1})$  at time step  $t_i$  via the Chapman–Kolmogorov equation [12]. In the update stage, the posterior distribution  $\Pr(d_i | \mathbf{z}_{1:i})$  is obtained

using the incoming measurement  $\mathbf{z}_i$  to update the prior distribution via the Bayes rule, based on the likelihood function  $\Pr(\mathbf{z}_i | d_i)$  defined by the observation equation [12]. The updated posterior probability distribution  $\Pr(RUL_i | \mathbf{z}_{1:i})$  can then be computed as the probability

$\Pr(RUL_i | \mathbf{z}_{1:i}) = \Pr[d(t_i + RUL_i) > d_{th} | \mathbf{z}_{1:i}]$  that the degradation level at time  $t_i + RUL_i$  exceeds the failure threshold  $d_{th}$  [20-21,31].

The recursive computation of the posterior  $\Pr(d_i | \mathbf{z}_{1:i})$  involves an integral which in practical cases does not have a closed-form solution. For this reason, approximated solutions have been proposed, like the Extended Kalman Filter, the Gaussian sum filter, and grid-based methods [32-33]. Also, Monte Carlo sampling techniques have become of increasing interest. Among these, particle filtering provides a solution by approximating the integrals in the Bayesian recursive procedure with weighted summations over a high number of samples called particles [10,12].

The application of the particle filtering procedure to the estimation of  $\Pr(RUL_i | \mathbf{z}_{1:i})$  is detailed in the pseudo-code given in Figure 1. The  $P$  particles  $p = 1, \dots, P$  are future degradation trajectories built by recursively sampling the particle degradation state  $d_j^p$  at time  $t_j$  from the transition probability distribution  $\Pr(d_j^p | d_{j-1}^p)$  derived from the degradation model, until the failure threshold  $d_{th}$  is exceeded and the length of life  $L^p$  of the particle is recorded. The value  $rul_i^p$  of the particle RUL at time step  $t_i$  can then be computed from  $rul_i^p = L^p - t_i$ . When an observation  $\mathbf{z}_i$  is collected, each particle is assigned a weight  $w_i^p$  proportional to the likelihood  $\Pr(\mathbf{z}_i | d_i^p)$  of observing  $\mathbf{z}_i$  given the degradation level  $d_i^p$  reached by the particle at the time  $t_i$  [20]. The distribution  $\Pr(RUL_i | \mathbf{z}_{1:i})$  is then approximated by an histogram of the  $P$  weighted values  $rul_i^p$  of the particle RULs at time  $t_i$ ; the weighted average and the weighted standard deviation of the values  $rul_i^p$ ,  $p = 1, \dots, P$  represent the prediction  $\hat{rul}_i$  of the expected value  $\mu_{RUL_i}$  of  $RUL_i$  and the estimate  $\hat{\sigma}_{\hat{rul}_i}^2$  of the prediction error variance  $\sigma_{rul_i}^2$ , respectively.

```

FOR p=1:P
  1. Sample  $d_0 \sim \Pr(d_0)$ 
  2.  $j=0$ ;  $t_j = t_0$ 
  WHILE  $d_j < d_{th}$ 
    3.  $j=j+1$ ;  $t_j = t_{j-1} + \Delta t$ 
    4. sample  $d_j^p \sim \Pr(d_j^p | d_{j-1}^p)$ 
  END WHILE
  5. Register the particle failure time  $t_f^p = t_j$ 
END FOR
FOR i=1:N
  6. Collect the observation  $z_i$ 
  FOR p=1:P
    7. Compute the RUL of the particle  $rul_i^p = t_f^p - t_i$ 
    8. Compute the weights:
      IF  $rul_i^p < 0$  set  $w_i^p = 0$ 
      ELSE  $w_i^p = w_{i-1}^p \cdot \Pr(z_i | d_i^p)$ 
    END FOR
  9. Normalize the weights  $wn_i^p = w_i^p / \sum_{k=1}^P w_i^k$ 
  10. Build the probability density function of the equipment RUL at
      time  $t_i$  as the histogram of the  $P$  weighted particle RULs  $rul_i^p$ .
  11. Compute  $\hat{rul}_i = \text{Mean}(rul_i^p)$ 
  12. Compute  $\hat{\sigma}_{rul_i}^2 = \text{Var}(rul_i^p)$ 
END FOR

```

**Figure 1: Particle filtering operative procedure for RUL estimation [8].**

The sampling importance resampling (SIR) algorithm is used to avoid the degeneracy problem of the particle filtering algorithm, which consists in having all but one of the importance weights close to zero after several weight updates [12]. This algorithm requires sampling, after one or more updates of the particle weights, a new set of particles from the old one with probability for a particle to be sampled proportional to its weight (see pseudo-code in Figure 2). New degradation trajectories have to be sampled starting from the degradation state  $d_i^p$  of each particle resampled at the observation time  $t_i$  and new values of the particles duration of life  $L^p$  are recorded. For more details the interested reader may refer to the specialized literature (e.g., [10,12]).

```

At time  $t_i$ 
1. Compute  $rul_i^p$  and  $wn_i^p$  as in Figure 1
FOR p=1:P
2. Sample a particle  $k$  with probability equal to its weight
    $wn_i^k$ 
3. set  $d_i^p = d_i^k$ ,  $j=i$  and  $t_j = t_i$ 
   WHILE  $d_j < d_{th}$ 
4.  $j=j+1$ ;  $t_j = t_{j-1} + \Delta t$ 
5. sample  $d_j^p \sim \Pr(d_j^p | d_{j-1}^p)$ 
   END WHILE
6. Register the new particle failure time  $t_j^p = t_j$ 
END FOR
7. Assign equal weights to each particle  $wn_i^{k:P} = 1/P$ 
8. Build the RUL probability density function and compute its
   moments  $\hat{rul}_i$  and  $\hat{\sigma}_{rul}^2$  as in Figure 1.

```

Figure 2: Procedure for performing resampling at time  $t_i$  [8].

Notice that by this approach the distribution  $\Pr(RUL_i | \mathbf{z}_{1:i})$  is estimated, which is different from the distribution  $\Pr(RUL_i | d_i)$  of the equipment RUL at time  $t_i$ , given that the equipment has degradation  $d_i$  at that time. However, in the Bayesian framework, this is the maximum information we can have on  $RUL_i$ . As for the uncertainty in the RUL prediction, in this approach the randomness of the degradation process (source of uncertainty A) is described by the model, whereas the observation equation accounts for the observation noise (source of uncertainty C). Thus, these two causes of uncertainty are accounted for in the RUL prediction through the procedure of particle sampling and weights updating, respectively. On the contrary, the contribution of model uncertainty to the RUL prediction uncertainty is not directly considered (source of uncertainty B), since it is assumed that the degradation dynamics model and the observation equation are exactly known. The effects of this uncertainty on the RUL prediction will be further discussed in Section 4.3.1. Notice, however, that if the uncertainty on the model parameters can be quantified and a probability distribution assigned to the value of the uncertain model parameter, the PF approach can be adjusted to handle also this source of uncertainty [34].

## 4.2 Approach 2: bootstrapped ensemble of empirical models trained on sequences of degradation observations and life time data

Approach 2 is based on the development of an empirical model  $f$  representing the relationship between the degradation observations  $\mathbf{z}_i$  available at time  $t_i$  and the corresponding equipment RUL. The empirical model is built considering the observations of a set of  $S$  trajectories  $\mathbf{z}_{1:N_s}^s$ ,  $s=1, \dots, S$ , of similar equipments which have each reached failure in a time  $L^s$  discretized in  $N_s$  steps. The empirical model receives in input the observations  $\mathbf{z}_i$  and produces as output the RUL prediction,  $\hat{rul}_i$ . In order to develop the model, a dataset of input/output pairs  $\mathbf{D}_{i/o} = \{(\mathbf{z}_i^s; rul_i^s); s=1, \dots, S; i=1:N_s\}$  is extracted from the set of observations  $\mathbf{z}_{1:N_s}^s$  by associating to the observations  $\mathbf{z}_i^s$  at time  $t_i$  along the  $s$ -th trajectory to failure the corresponding

realization of  $RUL_i$ , i.e.,  $rul_i^s = L^s - t_i$ . The dataset  $\mathbf{D}_{i/o}$  can then be used to train an empirical model based on one among the many data-driven modeling methods existing today (e.g., polynomial regression, non-parametric regression, neural networks, etc.). In their basic form, these methods provide in output a point prediction  $\hat{r}ul_i$  of the RUL without any information on the uncertainty of the estimate [35]. To overcome this limitation, the bootstrap method for estimating the accuracy in the prediction of a stochastic output whose mean value and variance are unknown functions of the input is used in this work. Under the hypothesis that the model  $f(\mathbf{z}_i) = \hat{r}ul_i$  is as an unbiased estimator of  $\mu_{RUL_i|\mathbf{z}_i}$ , i.e.,  $E[f(\mathbf{z}_i)] = \mu_{RUL_i|\mathbf{z}_i}$ , the model error variance  $\sigma_B^2$  can be rewritten as follows [24,36]:

$$\begin{aligned}\sigma_B^2(\mathbf{z}_i) &= E\left[\left(f(\mathbf{z}_i) - \mu_{RUL_i|\mathbf{z}_i}\right)^2\right] \\ &= E\left[\left(f(\mathbf{z}_i) - E[f(\mathbf{z}_i)]\right)^2\right] + E\left[\left(E[f(\mathbf{z}_i)] - \mu_{RUL_i|\mathbf{z}_i}\right)^2\right] \\ &= E\left[\left(f(\mathbf{z}_i) - E[f(\mathbf{z}_i)]\right)^2\right]\end{aligned}\quad (6)$$

An estimate of the model error variance,  $\sigma_B^2$ , is then obtained from an ensemble of models  $f^h(\mathbf{z}_i | \mathbf{D}_{i/o}^h)$ ,  $h=1, \dots, H$  trained using bootstrapped replicates  $\mathbf{D}_{i/o}^h$  of a training dataset  $\mathbf{D}_{i/o}^{tr}$ , drawn from  $\mathbf{D}_{i/o}$ . Given a generic input  $\mathbf{z}_i$ , the models of the ensemble generate  $H$  different predictions  $\hat{r}ul_i^h = f^h(\mathbf{z}_i | \mathbf{D}_{i/o}^h)$ ; their variance is assumed as the estimate  $\hat{\sigma}_B^2(\mathbf{z}_i)$  of the model error variance  $\sigma_B^2(\mathbf{z}_i)$  [24,37], whereas their average is taken as the best estimate  $\hat{r}ul_i$  of the equipment RUL.

With respect to the estimate of the remaining part of the RUL prediction variance, which is caused by the randomness of the degradation process and the observation noise (sources of uncertainty A and C), i.e.  $\sigma_{A+C}^2 = \sigma_A^2 + \sigma_C^2$ , an independent validation dataset  $\mathbf{D}_{i/o}^{val}$  is used. In particular, the ensemble of empirical models  $f^h(\mathbf{z}_i | \mathbf{D}_{i/o}^h)$  is applied to the observations in the validation dataset  $\mathbf{D}_{i/o}^{val}$ . The obtained RUL predictions  $\hat{r}ul_{i,N_s}^s$ ,  $s=S_{tr}+1, \dots, S$  are used to calculate, for each validation observation  $\mathbf{z}_i^{val,s}$ , the prediction residuals  $r_i^s$ :

$$r_i^s = (\hat{r}ul_i^s - rul_i^s)^2 - \hat{\sigma}_B^2(\mathbf{z}_i^{val,s}) \quad (7)$$

The set of input/output pairs obtained by associating to the observations  $\mathbf{z}_i^{val,s}$ ,  $s=1, \dots, S_{val}$ ,  $i=1, \dots, N_s$ , in  $\mathbf{D}_{i/o}^{val}$  the corresponding residuals  $r_i^s$  is used for training an empirical model  $\chi(\mathbf{z}_i) = \hat{\sigma}_{A+C}^2(\mathbf{z}_i)$  of the residual variance approximating the unknown relation between the input  $\mathbf{z}_i$  and the variance of the residuals [24,38].

When a new observation  $\mathbf{z}_i$  is collected, the following procedure is applied in order to obtain the estimate  $\hat{r}ul_i$  of the equipment RUL and of the corresponding variance  $\hat{\sigma}_{\hat{r}ul_i}^2$ :

- Compute the output  $\hat{r}ul_i^h$  of each models  $f^h(\mathbf{z}_i | \mathbf{D}_{i/o}^h)$  of the ensemble;
- Compute the point estimate of the RUL:

$$r\hat{u}_i = \frac{1}{H} \sum_{b=1}^H f^h(\mathbf{z}_i | \mathbf{D}_{i/o}^h) \quad (8)$$

- Compute the RUL prediction uncertainty as follows:

$$\hat{\sigma}_{r\hat{u}_i}^2 = \hat{\sigma}_B^2(\mathbf{z}_i) + \hat{\sigma}_{A+C}^2(\mathbf{z}_i) = \text{var}[f^h(\mathbf{z}_i | \mathbf{D}_{i/o}^h)] + \chi(\mathbf{z}_i) \quad (9)$$

Then,  $\hat{\sigma}_{r\hat{u}_i}^2$  accounts for all three sources of uncertainty listed in Section 1.

However, the degradation measurements depend on the entire past trajectory of degradation, which means that training and validation data taken from the same trajectory are not independent, causing an underestimation of the variance. For this reason, the validation dataset  $\mathbf{D}_{i/o}^{val}$  is made by input/output pairs  $(\mathbf{z}_i^s; rul_i^s)$ ,  $s=1, \dots, S_{val}$ ,  $i=1, \dots, N_s$ , taken from trajectories different from those used in training. Furthermore, to ensure enough diversity of the models in the ensemble, the bootstrapped training datasets  $\mathbf{D}_{i/o}^h$ ,  $h=1, \dots, H$ , are sampled from  $\mathbf{D}_{i/o}^{tm}$  as follows: first,  $S_{tm}$  training trajectories are randomly sampled with replacement from the  $S_{tm}$  different trajectories of  $\mathbf{D}_{i/o}^{tm}$ ; then,  $N^{tm}$  input/output pairs  $(\mathbf{z}_i^s; rul_i^s)$  are sampled with replacement from the total amount of input/output pairs in  $\mathbf{D}_{i/o}^{tm}$ .

### 4.3 Approach 3: Bootstrapped ensemble of empirical models trained on a sequence of past degradation observations from the equipment of interest only

Approach 3 is based on the development of an empirical model of the degradation process based on the time series of its past observations  $\mathbf{z}_{1:i}$  and used for identifying the time at which the degradation will exceed the failure threshold. For simplicity of illustration the observations  $\mathbf{z}_{1:i}$  are assumed to be direct measures of the degradation  $d_{1:i}$ , eventually affected by noise.

The approach differs from approach 1 in that the stochastic model describing the dynamics of the degradation process is not available and actually the point is to develop it empirically. Coherently, the estimate of the prediction error variance  $\sigma_{r\hat{u}_i}^2(\mathbf{z}_i)$  should account also for the error of approximation of the empirical model.

The approach differs from approach 2 in that there are no available pairs  $(\mathbf{z}_i; rul_i)$  for which  $rul_i$  is known for training and validating the prognostic model  $f(\mathbf{z}_i | \mathbf{D}_{i/o})$ .

Empirical modeling of the degradation process can be achieved by fitting the most suited degradation model, e.g., linear and non-linear regression models, general degradation path models, etc. [39] to the available data.

Let us call  $\hat{d}_j = \eta(t_j)$  a generic model of the equipment degradation, derived from the sequence of data  $\mathbf{z}_{1:i}$ .

The prediction  $r\hat{u}_i$  of the equipment RUL at time  $t_i$  can be simply obtained from the relation

$\eta(t_i + r\hat{u}_i) = d_{it}$ . Once again an estimate of the prediction error variance  $\sigma_{r\hat{u}_i}^2(\mathbf{z}_i)$  is needed, but cannot be

obtained by means of the method proposed for approach 2 since there are no available pairs  $(\mathbf{z}_i; rul_i)$  for

which  $rul_i$  is known, and thus eq. (7) cannot be used to calculate the value of the prediction residual  $r_i$  in correspondence of the observation  $\mathbf{z}_i$ .

Let us, instead, consider a model  $\hat{\Delta t}_{j,j'} = \tilde{\eta}(\mathbf{d}_{j,j'})$  receiving in input a vector of two degradation values  $\mathbf{d}_{j,j'} = [d_j, d_{j'}]$  and returning in output the estimate  $\hat{\Delta t}_{j,j'}$  of the time interval needed to reach degradation  $d_{j'}$  starting from  $d_j$ . Notice that model  $\tilde{\eta}(\mathbf{d}_{j,j'})$  can be derived, in general, from model  $\eta(t_j)$ :

$$\Delta t_{j,j'} = \tilde{\eta}(\mathbf{d}_{j,j'}) = \eta^{-1}(d_{j'}) - \eta^{-1}(d_j) \quad (10)$$

The prediction  $r\hat{u}_i$  is then obtained from this model by setting  $d_j = \mathbf{z}_i$  and  $d_{j'} = d_{th}$ , which means that the RUL prediction at time  $t_i$  corresponds to the estimate of the time interval  $\Delta t_{i,th}$  needed to increase the degradation from  $d_i$  to the failure threshold  $d_{th}$ .

The relation between the input  $\mathbf{d}_{j,j'} = [d_j, d_{j'}]$  and the output  $\Delta t_{j,j'}$  of the model in eq. (10) is:

$$\Delta t_{j,j'} = \mu_{\Delta t_{j,j'}} + \delta_{\Delta t_{j,j'}} \quad (11)$$

where  $\delta_{\Delta t_{j,j'}} = \Delta t_{j,j'} - \mu_{\Delta t_{j,j'}}$  is a zero mean random variable with variance  $\sigma_{\Delta t_{j,j'}}^2$  representing the uncertainty in the evolution of the degradation process from  $d_j$  to  $d_{j'}$  and  $\mu_{\Delta t_{j,j'}}$  is the mean value of the random variable  $\Delta t_{j,j'}$ . Both  $\sigma_{\Delta t_{j,j'}}^2$  and  $\mu_{\Delta t_{j,j'}}$  are, in general, functions of the input  $\mathbf{d}_{j,j'}$ .

The observations  $\mathbf{z}_{1:i}$  are used to build input/output pairs  $(\mathbf{d}_{j,j'} = [\mathbf{z}_j, \mathbf{z}_{j'}]; \Delta t_{j,j'} = t_{j'} - t_j)$ ,  $j = 1, \dots, i-1$ ;  $j' = j+1, \dots, i$  and the bootstrapping of approach 2 can be applied to estimate the variance of the prediction error of the model in eq. (10) by building training and validation datasets of input/output pairs. As underlined in Section 3.2, 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  $\mathbf{D}$  into two sequences of consecutive measurements,  $\mathbf{D}^{trn} = \{\mathbf{z}_{1:N_{trn}}\}$  and  $\mathbf{D}^{val} = \{\mathbf{z}_{N_{trn}+1:i}\}$  and to use  $\mathbf{D}^{trn}$  for building the model  $\hat{d} = \eta(t)$  and  $\mathbf{D}^{val}$  for building the dataset of input/output pairs  $\mathbf{D}_{i/o}^{val} = \{(\mathbf{d}_{j,j'}; \Delta t_{j,j'} = t_j - t_{j'})\}$ ,  $j = N_{trn} + 1, \dots, i-1$ ;  $j' = j+1, \dots, i$  to be used for estimating the prediction error.

An ensemble of models  $\eta^h(t | \mathbf{D}^h)$ ,  $h=1, \dots, H$ , is trained using bootstrap replicates  $\mathbf{D}^h$  of the training dataset  $\mathbf{D}^{trn}$ , and the ensemble of models  $\tilde{\eta}^h(\mathbf{d}_{j,j'} | \mathbf{D}^h)$  is derived from eq. (10). The average and variance of the ensemble model prediction are retained as the estimates  $\hat{\Delta t}_{j,j'}$  and  $\hat{\sigma}_B^2(\mathbf{d}_{j,j'})$  of, respectively, the time interval  $\Delta t_{j,j'}$  and the error variance of model  $\tilde{\eta}(\mathbf{d}_{j,j'} | \mathbf{D})$ . The ensemble of models is applied to the validation dataset  $\mathbf{D}_{i/o}^{val}$  in order to obtain a set of prediction residuals  $r_{j,j'}$ :

$$r_{j,j'} = (\hat{\Delta t}_{j,j'} - \Delta t_{j,j'})^2 - \hat{\sigma}_B^2(\mathbf{d}_{j,j'}) \quad (12)$$

Finally, an empirical model  $\chi(\mathbf{d}_{j,j'}) = \hat{\sigma}_{A+C}^2(\mathbf{d}_{j,j'})$ , estimating the part of the error variance  $\sigma_{\Delta t_{j,j'}}^2$  due to the stochasticity of the degradation process and the observation noise (sources of uncertainty A and C), is trained using the input/output pairs  $(\mathbf{d}_{j,j'}; r_{j,j'})$ . The sum of the RUL and noise variance equipments,  $\hat{\sigma}_{A+C}^2(\mathbf{z}_i)$ , is then obtained from this model by setting  $\mathbf{d}_{j,j'} = \mathbf{d}_{i,th} = [\mathbf{z}_i, d_{th}]$ .

Notice that the training data  $(\mathbf{d}_{j,j'}; r_{j,j'})$  used to build the models cover a range of values for the input  $\mathbf{d}_{j,j'}$  in general different from that of the input  $\mathbf{d}_{i,th} = [\mathbf{z}_i, d_{th}]$  to which the model is applied to obtain the estimate  $\hat{\sigma}_{A+C}^2(\mathbf{z}_i)$ . This can represent a limit to the quality of the estimate  $\hat{\sigma}_{A+C}^2(\mathbf{z}_i)$ , since in general the performance of empirical models are good when applied to input regions well described by training data, and decrease moving away from these regions.

When a new observation  $\mathbf{z}_i$  of the degradation  $d_i$  is collected at time  $t_i$ , the multiple RUL predictions

$r\hat{u}_i^h = \tilde{\eta}^h[\mathbf{d}_{i,th} | \mathbf{D}^h]$  and the RUL variance estimate  $\hat{\sigma}_{A+C}^2(\mathbf{z}_i) = \chi(\mathbf{d}_{i,th})$  are used to obtain the prediction  $r\hat{u}_i$  and the relative estimate  $\sigma_{r\hat{u}_i}^2$  of the prediction error variance:

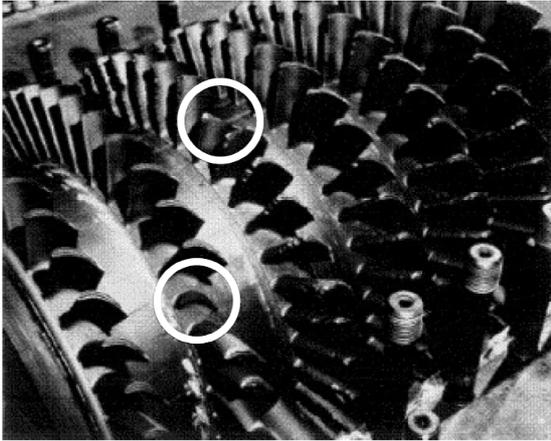
$$r\hat{u}_i = \frac{1}{H} \sum_{h=1}^H r\hat{u}_i^h \quad (13)$$

$$\sigma_{r\hat{u}_i}^2 = \hat{\sigma}_B^2(\mathbf{z}_i) + \hat{\sigma}_{A+C}^2(\mathbf{z}_i) = \text{var}[\tilde{\eta}^h(\mathbf{d}_{i,th} | \mathbf{D}^h)] + \chi(\mathbf{d}_{i,th}) \quad (14)$$

As for approach 2, all three sources of uncertainty listed in Section 1 are taken into account in the estimate  $\hat{\sigma}_{r\hat{u}_i}^2$ .

## 5 Numerical application

The three different approaches presented in Section 4 are verified with respect to the RUL prediction of a simulated turbine blade undergoing degradation. The application focuses on the turbine of a generation IV high temperature gas reactor, which is characterized by rather extreme turbine operational conditions such as working temperatures exceeding 900°C. The predominant damage mechanisms affecting turbines operating at such elevated temperatures include creep deformation, corrosion and fatigue [40]. The interaction of these and other mechanisms generates a degradation process that leads to crack initiation which rapidly leads to failure due to the quick accumulation of stress cycles caused by the high rotational speed. 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 equipment and in the reactor flow conditions [41]. Figure 3 shows an example of high-pressure turbine deblading occurred in a German power plant [42]. Also, fracture in rotary machines can result in turbine missiles, i.e., irregularly shaped projectiles travelling at high velocities which can impact on barriers in nuclear power plants causing severe damages to the facilities, and threatening public safety [43]. This and the high cost of turbine blade replacement are strong reasons for performing prognostics on creeping turbine blades.



**Figure 3: Deblading in a high pressure turbine [42].**

Indeed, health monitoring and prognostics may allow scheduling maintenance actions before the blade degradation evolves into cracks. In this context, the dimensionless quantity  $\varepsilon$ , defined as the percentage of elongation of the turbine blade in the longitudinal direction with respect to its original length, can be used as an indicator of the blade degradation state [44]. The blade is discarded when the accumulated elongation reaches a pre-determined value, namely the failure threshold, which assures that the risk of blade failure is below the desired safety limit. In [45] various sensing technologies for measuring blade deformations are outlined and their possible application to turbine blade and disk health monitoring are discussed. In particular, methods measuring blade tip clearance (BTP), which defines the gap between the tip of blades and the casing [45], appear to be promising for measuring blade plastic elongation. In this respect, the main technologies available today are based on capacitive measurements, eddy current or microwave sensing. Capacitive-based technologies [46-47] are already on the market [48-49] and provide clearance and vibration monitoring sensors to gas turbines used in the power generation and aerospace industries. In [50], a magnetic sensor relying on high-frequency eddy currents actively induced in passing blades, is tested on field trials with jet engines and it is demonstrated the capability of generating online clearance measurement for each blade. Similar results are obtained in [51] using a microwave sensing system, which is claimed to be capable of performing blade monitoring in the harsh environment of the first turbine stages.

In [45] it is observed that long-term trends in BTP can be measured and its future application to monitoring creep-related blade deformation is anticipated. In [44], the possibility of using BTP in blade failure risk analysis and diagnosis is analyzed and a blade prognostic approach based on BTP linear regression is proposed. In [50] a general framework for PHM of turbine blades is proposed, considering blade tip clearance and other damage indicators such as vibrations, blade angular position, etc.

Notice that, given the advancements in blade tip clearance sensing technology, a sufficient amount of data for training and validation of prognostic models can be expected to be available in the future. Furthermore, notice that the amount of real data necessary to validate the prognostic approaches developed can be largely reduced by resorting to the leave-one-out cross-validation procedure [52].

## 5.1 Degradation model

Modeling the degradation of a turbine blade is a hard task, especially if one needs to take into account all mechanisms involved and their interactions. For the purpose of this work, we limit ourselves to considering the accumulation of creep damage. Creep is an irreversible deformation process affecting materials exposed to a load below their elastic limit for a protracted length of time and at high temperatures. In the high pressure turbine first stage, blades creep is a major problem due to the high operational temperatures, and is often the life-limiting process [53]. Blade elongation,  $\varepsilon$ , is taken as a measure of the blade creep strain. In this work, the creep evolution is modeled using the Norton Law discretized with a step  $\Delta t = 5$  days, assuming that the dependence from the temperature follows the Arrhenius law [27,54]:

$$\varepsilon_{j+1} = \varepsilon_j + A \cdot \exp\left(-\frac{Q}{RT_j}\right) \cdot [K\omega_j^2 + \delta\varphi_j]^n \cdot \Delta t, \quad \varepsilon_0 = 0 \quad (15)$$

where  $\varepsilon_j$  is the creep strain at time  $t_j$ ,  $Q$  is the activation energy,  $A$  and  $n$  are material inherent characteristics varying from one blade to another,  $K$  is a constant relating the load to the rotational speed  $\omega_j$ ,  $R$  is the ideal gas constant,  $T_j$  is the blade operating temperature and  $\delta\varphi_j$  is a random variable modeling 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 equipment or turbulences of the gas flow. Oscillations of the rotational speed  $\omega_j$ , and of the blade operating temperature  $T_j$ , are represented by considering their deviations from the mean values  $\mu_\omega$  and  $\mu_T$  as noises. In practice,  $\delta\omega_j = \omega_j - \mu_\omega$  and  $\delta T_j = T_j - \mu_T$  are Gaussian random variables with mean values zero and assigned standard deviations. Thus, the noise vector  $\gamma_j$  in eq. (5) can be set equal to  $\gamma_j = [\delta\omega_j; \delta T_j; \delta\varphi_j]$ .

The values of the parameters  $T_j$ ,  $\omega_j$  and  $K$  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 [26]; the material inherent characteristics  $A$  and  $n$  are taken assuming that the blade is made of Ni-base cast Superalloy 713LC [26]. The distributions used for the parameters are reported in Table II.

**Table II: 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	$(\text{N/m}^2)^{-n}/\text{day}$	$\mu_A=7.2 \cdot 10^{-3}$ ; $\sigma_A=5\%$
	$n$	Normal	-	$\mu_n=6$ ; $\sigma_n=0.2\%$
Operating temperature	$T_j$	Normal	K	$\mu_T=1100$ ; $\sigma_T=1\%$
Rotational speed	$\omega_j$	Normal	rpm	$\mu_\omega=3000$ ; $\sigma_\omega=1\%$
Load parameter	$K$	Deterministic	Kg/m	$\rho=1068$
Stress fluctuations	$\delta\varphi$	Gamma	MPa	$\theta=2$ ; $k=10$

Eq. (15) represents a stochastic process whose unknowable future evolution (cause A, Section 1) produces an irreducible uncertainty in the RUL prediction. Parameters  $A$  and  $n$  instead represent an uncertainty in the model (source B, Section 1). In fact, to a specific blade correspond fixed parameters  $A$  and  $n$  but their exact

values are not known in practice; to include this source of uncertainty in the model, we assume to know with a certain precision the range of values of these parameters and associate to them a probability distribution (Table II).

For simplicity, it is assumed that it is possible to directly measure at inspection time  $t_j$  the value of the creep strain  $\varepsilon_j$ . Thus, the observation equation is:

$$\mathbf{z}_j = \varepsilon_j + v_j \quad (16)$$

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

$\Pr(\mathbf{z}_j | d_j^p)$  used in the particle filtering approach is Gaussian with mean  $d_j^p$  and variance  $\sigma_v^2$ . This noise represents a source of uncertainty (source C, Section 1) in the final RUL prediction.

The failure threshold for creep strain  $\varepsilon_{th}$  is set equal to the value of 1.5%.

Given the unavailability of real experimental data, a sequence of creep strain measurements  $\mathbf{z}_{1:i}$  on the blade of interest, hereafter called ‘test trajectory’, is simulated using eq. (15). 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 II) at each time instant  $t_j$ . Every 30 days a measurement  $\mathbf{z}_j$ , corresponding to the creep strain  $\varepsilon_j$ , is simulated by using eq. (16). 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 RUL of the turbine blade; it will be referred to as ‘true RUL’, and represented by the notation  $rul_i$ .

Also, a number  $S=13$  of historical creep growth trajectories of similar blades have been simulated using eq. (15). To induce variability in the behavior of the similar blades the values of the characteristic parameters  $A$  and  $n$  from one blade to another have been sampled from normal distributions (Table II) at the beginning of each new simulated degradation trajectory. Some examples of simulated creep growth trajectory are shown in Figure 4.

For each trajectory, a sequence of  $N_s$  direct creep strain measurement  $\mathbf{z}_{1:N_s}$ , one every 30 days, are simulated according to eq. (16).

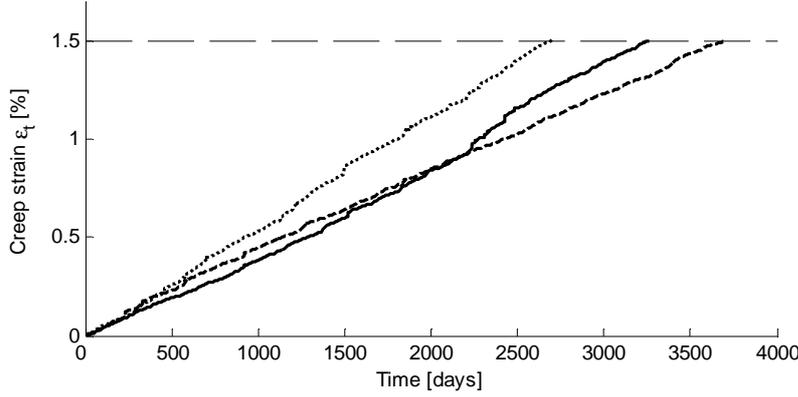


Figure 4: Examples of creep growth trajectories.

## 5.2 Results

During the life of the turbine blade, at every time  $t_i$ , the set of observations  $\mathbf{z}_{1:i}$  is assumed to be available; the objective of the analysis is to predict at time  $t_i$ ,  $i=1, \dots, 87$ , the RUL distribution for the test trajectory. Three situations have been artificially constructed for the turbine blade case study described in the previous Section 4.1, corresponding to the three prognostic approaches of Section 3.

In the PF approach 1, the model of eq. (15) is used to simulate a number  $P = 1000$  of particles starting from  $\varepsilon_0 = 0$ . Particle resampling is performed once every 5 measurements. Note that 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. (15).

In the bootstrapped ensemble approach 2, which uses multiple sequences of degradation observations,  $S_{tm}=10$  trajectories among the  $S = 13$  totally available are used for building an ensemble of  $H=25$  linear least square models  $\hat{r}ul^h(\mathbf{z}_i | D_{i/o}^h) = \alpha_0^h + \alpha_1^h \cdot \mathbf{z}_i$ , whereas the remaining  $S_{val}=3$  trajectories are used to validate the ensemble and build the training dataset for the least squares model  $\chi(\mathbf{z}_i) = \gamma_0 + \gamma_1 \cdot \mathbf{z}_i + \gamma_2 \cdot \mathbf{z}_i^2$  estimating the RUL variance  $\sigma_{A+C}^2$ .

In the bootstrapped ensemble approach 3, which uses the time series of degradation observations, 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  $\mathbf{D}^{tm}$  containing the first 75% of the available measurements and a validation dataset  $\mathbf{D}^{val}$  containing the remaining 25%. An ensemble of  $H=25$  linear least square models  $\eta^h(t_j | \mathbf{D}^h) = \beta_0^h + \beta_1^h \cdot t_j$  is built and the models  $\tilde{\eta}^h[\Delta \mathbf{z}_{j,j'} | \mathbf{D}^h] = \Delta \mathbf{z}_{j,j'} / \beta_1^h$  are derived from it. Notice that in a linear process, the time needed to increase the degradation level 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 values. The ensemble of models is tested on the validation dataset made of input/output pairs  $\mathbf{D}_{i/o}^{val} = \{\Delta \mathbf{z}_{j,j'} = \mathbf{z}_{j'} - \mathbf{z}_j; \Delta t_{j,j'} = t_{j'} - t_j\}$ ,  $j=N_{tm}+1, \dots, i-1, j'=j+1, \dots, i$  and the prediction residuals  $r_{j,j'}$  obtained are used to train the linear model  $\chi(\Delta \mathbf{z}_{j,j'}) = \lambda \Delta \mathbf{z}_{j,j'}$  for the

variance of  $\Delta t$ . After time  $t_{30}$ , each time  $t_i, i=31, \dots, 87$ , a new measurement becomes available, a new ensemble of models  $\tilde{\eta}^h$  and a new model  $\chi$  for the prediction error variance are built. The predictions  $r\hat{u}_i^h$  and the estimate  $\hat{\sigma}_{A+C}^2$  are obtained respectively from the models  $\tilde{\eta}^h$  of the ensemble and from model  $\chi$  in correspondence of the input  $\Delta \mathbf{z}_{i,th} = \varepsilon_{th} - \mathbf{z}_i$ . Since 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  $\Delta \mathbf{z}_{i,th} = \varepsilon_{th} - \mathbf{z}_i$  considered for obtaining the prognostic results, the empirical model  $\chi(\Delta \mathbf{z}_{j,j'})$  is used in an input region not described by the training data.

In all three approaches, in correspondence of each prediction  $r\hat{u}_i$  we estimate the prediction interval

$[C_i^{\text{inf}}(\alpha); C_i^{\text{sup}}(\alpha)]$ , i.e., the interval expected to contain the true RUL value  $rul_i$  with a probability of  $1-\alpha$ .

This interval is obtained as follows:

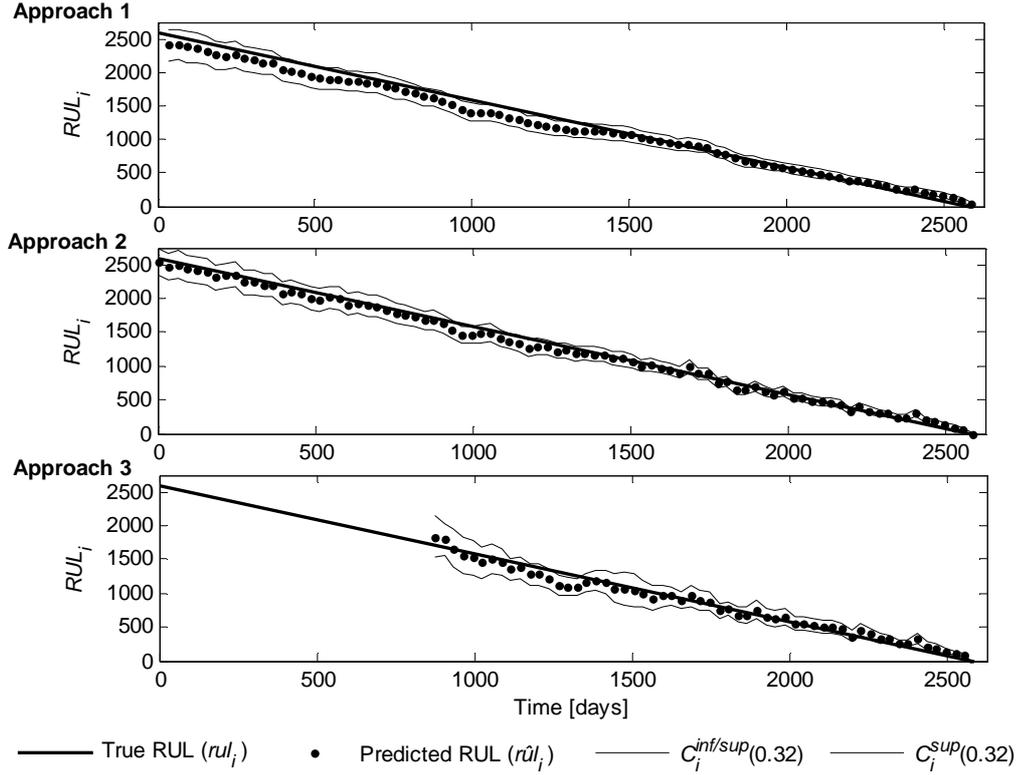
- In approach 1,  $C_i^{\text{inf}}(\alpha)$  and  $C_i^{\text{sup}}(\alpha)$  are the  $\alpha/2$  and  $1-\alpha/2$  percentiles, respectively, of the RUL distribution estimated by Particle Filtering.
- In approaches 2 and 3, assuming that the prediction error has a Gaussian distribution, the value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

$$C_i^{\text{inf}}(\alpha) = r\hat{u}_i - c_{conf}^{\alpha} \hat{\sigma}_{r\hat{u}_i} \quad \text{and} \quad C_i^{\text{sup}}(\alpha) = r\hat{u}_i + c_{conf}^{\alpha} \hat{\sigma}_{r\hat{u}_i} \quad (17)$$

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

Figure 5 shows the evolution of the true value  $rul_i$  of the blade RUL (continuous thick line), its estimated value  $r\hat{u}_i$  (dots) and the corresponding prediction interval for  $\alpha = 0.32$  (continuous thin line) obtained during the turbine blade life at times  $t_i, i=1, \dots, 87$ , by the three prognostic approaches.

The prediction intervals provided by approach 3 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 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 of the prediction, Kalman filtering can be applied [27].



**Figure 5:** true  $rul_i$  (continuous thick line) of a turbine blade, predicted value  $\hat{r}ul_i$  (dots) and prediction interval  $[C_i^{inf}(0.32); C_i^{sup}(0.32)]$  (continuous thin line) for the three prognostic approaches.

### 5.3 RUL distribution and prediction interval

The objective of this Section is to determine whether the estimates of the prediction intervals provided by the three approaches properly describe the uncertainty in the RUL predictions. In practice, we want to know whether the estimates  $\hat{\sigma}_{\hat{r}ul_i}^2$  are satisfactory approximations of the real  $\sigma_{\hat{r}ul_i}^2$ . According to eq. (4),  $\sigma_{\hat{r}ul_i}^2$  can be decomposed into the sum of three terms, due, respectively, to the process randomness,  $\sigma_A^2$ , the model error,  $\hat{\sigma}_B^2$ , and the noise on the observations,  $\sigma_C^2$ . It is also of interest to consider the term  $\sigma_{A+C}^2 = \sigma_A^2 + \sigma_C^2$ , since all three approaches proposed do not estimate these two terms separately. The computation of the true value of  $\sigma_{\hat{r}ul_i}^2$  would ideally require the availability of an infinite number  $P$  of equipment degradation trajectories which at time  $t_i$  are in the degradation state  $\varepsilon_i$ . Since in the case study here considered we can artificially generate degradation trajectories, an high number  $P=1000$  of degradation trajectories has been used to numerically approximate the variance  $\sigma_{\hat{r}ul_i}^2$ . For the  $p$ -th simulated trajectory, we have computed: 1) its true RUL,  $rul_i^p = L^p - t_i$ , with  $L^p$  being the equipment life duration along the  $p$ -th trajectory, 2) the equipment RUL prediction,  $\hat{r}ul_i^p$ , provided by the prognostic model in correspondence of the observations  $\mathbf{z}_i^p = \varepsilon_i + \mathbf{v}_i^p$  with  $\mathbf{v}_i^p$  a random Gaussian noise with variance  $\sigma_v^2$ . Then,  $\sigma_{\hat{r}ul_i}^2$  has been approximated by:

$$\sigma_{\hat{r}ul_i}^2 = E[(r\hat{u}l_i - RUL_i)^2] \cong \sum_{p=1}^P \frac{(r\hat{u}l_i^p - rul_i^p)^2}{P} \quad (18)$$

Similarly, the computation of the true value of  $\sigma_A^2 = E[(\mu_{RUL_i|\varepsilon_i^p} - RUL_i)^2]$  is approximated by:

$$\sigma_A^2 \cong \frac{1}{P} \sum_{p=1}^P [\mu_{RUL_i|\varepsilon_i^p} - rul_i^p]^2 \quad (19)$$

where the RUL expected value  $\mu_{RUL_i|\varepsilon_i}$  is approximated by  $\sum_{p=1}^P \frac{L^p - t_i}{P}$ .

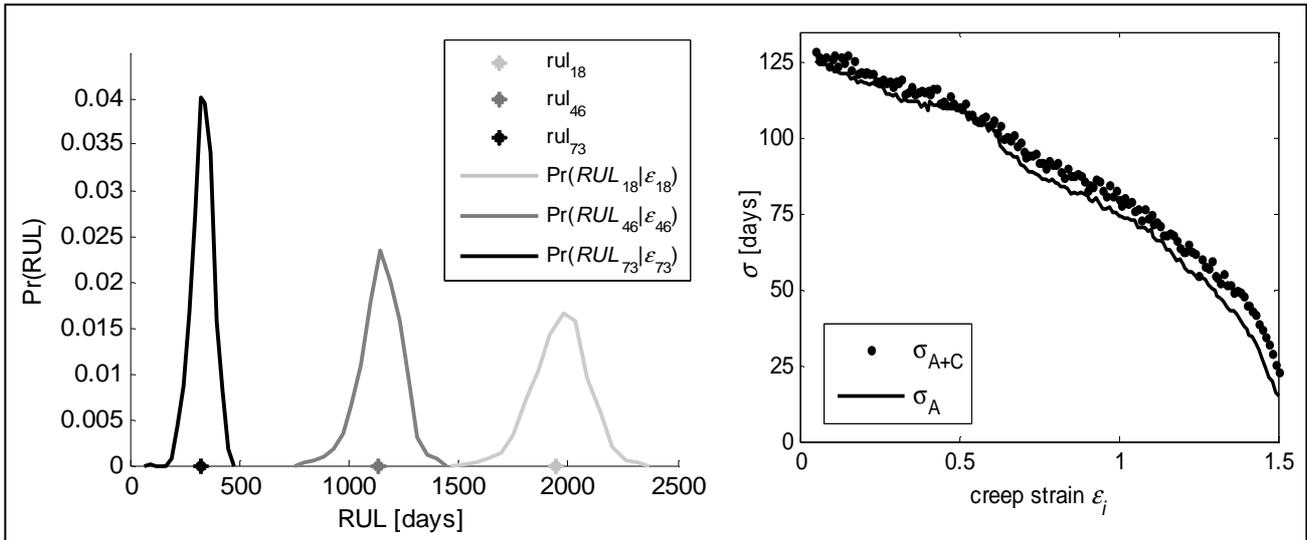
The real value of  $\sigma_{A+C}^2 = E[(\mu_{RUL_i|z_i} - RUL_i)^2]$  has been approximated by considering the  $P=1000$  equipment degradation trajectories which at time  $t_i$  are in the degradation state  $\varepsilon_i$  and for which the observations  $\mathbf{z}_i^p$  have been collected, and computing:

$$\sigma_{A+C}^2 \cong \frac{1}{P} \sum_{k=1}^P [\mu_{RUL_i|z_i^p} - rul_i^p]^2 \quad (20)$$

where  $\mu_{RUL_i|z_i^p}$  has been approximated by simulating  $P'=1000$  new degradation trajectories, each one starting from a different degradation state  $\varepsilon_i^{p'} = \mathbf{z}_i^p - \mathbf{v}_i^{p'}$ . This procedure allows to propagate the uncertainty on the true degradation state given the observation  $\mathbf{z}_i^p = \varepsilon_i + \mathbf{v}_i^p$  to the RUL mean value and to approximate  $\mu_{RUL_i|z_i^p}$  by:

$$\mu_{RUL_i|z_i^p} = \frac{1}{P'} \sum_{p'=1}^{P'} rul_i^{p'} \quad (21)$$

Notice that the terms  $\sigma_A^2$  and  $\sigma_{A+C}^2$  do not depend on the approach. In Figure 6 (left), the true RUL distribution  $\Pr(RUL_i | \varepsilon_i)$ , approximated by the distribution of the 1000 simulated  $rul_i^p$ , is shown at the times  $t_i = t_{18} = 515$  days,  $t_{46} = 1355$  days and  $t_{73} = 2135$  days, whereas in Figure 6 (right) the true values of the standard deviations  $\sigma_A$  (continuous line) and  $\sigma_{A+C}$  (dots) are reported as a function of the blade creep strain level  $\varepsilon_i$ .

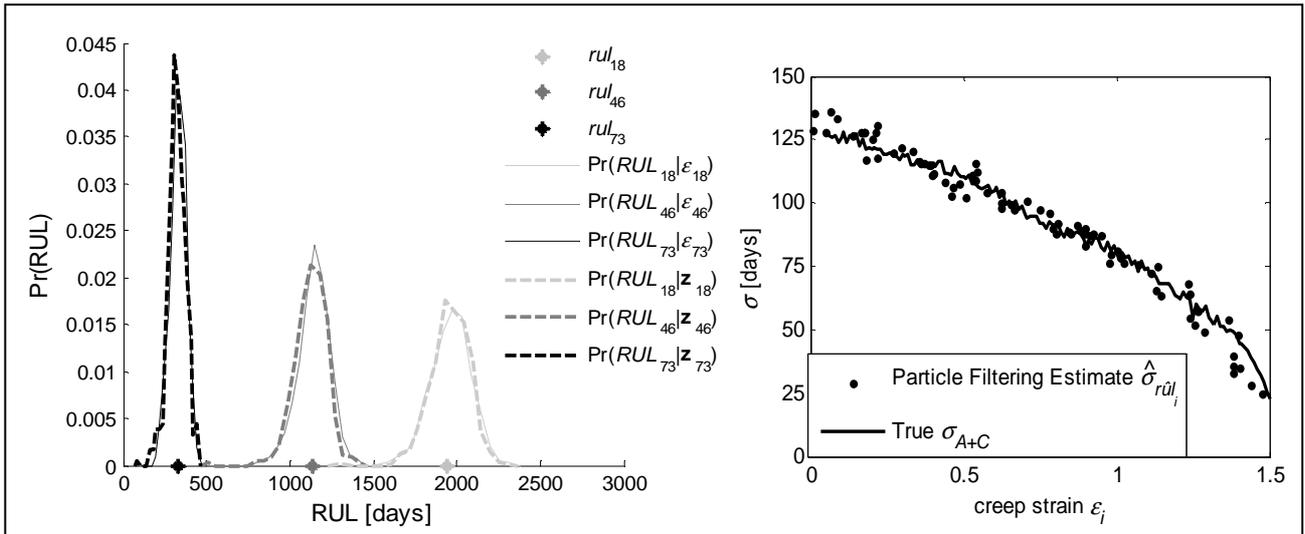


**Figure 6:** pdf  $\Pr(RUL_i | \varepsilon_i)$  of the RUL of a turbine blade at three different instances of the degradation trajectory (left) and evolution of the RUL standard deviations  $\sigma_A$  (continuous line) and  $\sigma_{A+C}$  (dots) as a function of the creep strain  $\varepsilon_i$  at time  $t_i$ .

Differently from the variance terms  $\sigma_A^2$  and  $\sigma_C^2$ , the model error variance term  $\sigma_B^2$  depends on the modeling approach used to estimate the RUL. Considerations on  $\sigma_B^2$  will be done in the following Sections 5.3.1, 5.3.2 and 5.3.3.

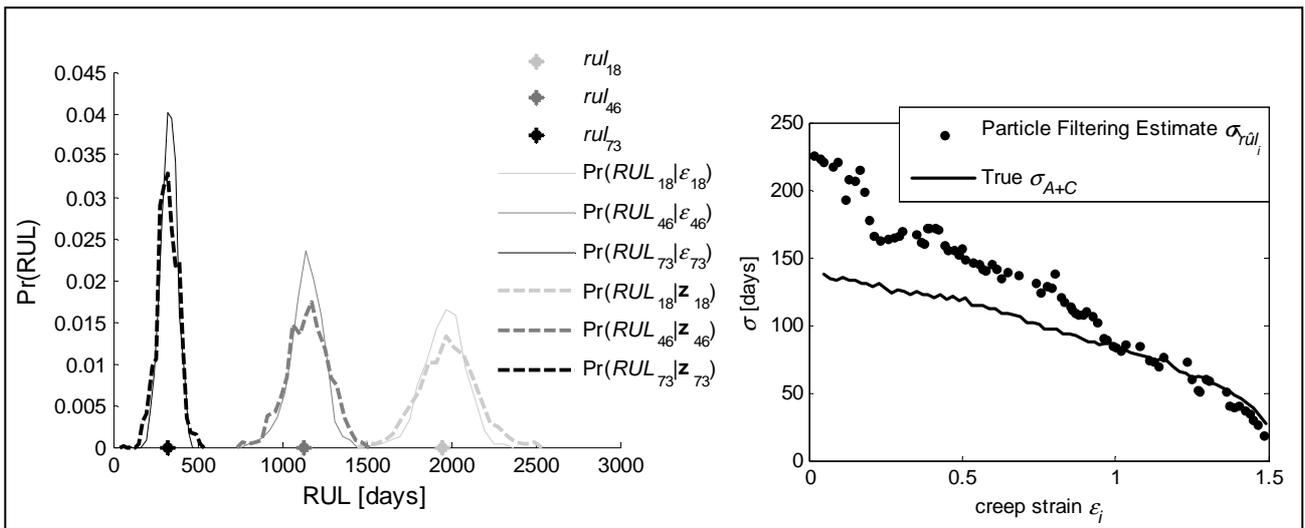
### 5.3.1 Prediction interval provided by the PF approach 1

The specific blade undergoing the creep degradation process is characterized by fixed values of the parameters  $A$  and  $n$  in eq. (15), which in general are not known. In this Section, in order to evaluate the PF performance in the estimate of the prediction uncertainty, the approach is firstly applied assuming to know the exact value of these parameters. In Figure 7, the distribution  $\Pr(RUL_i | \mathbf{z}_{1:i})$  predicted by the particle filtering method (left, dashed line) and the estimate of the prediction error standard deviation  $\hat{\sigma}_{r\hat{u}_i}$  (right, dots), are compared to the true RUL distribution  $\Pr(RUL_i | \varepsilon_i)$  (left, continuous line) and standard deviation  $\sigma_{A+C}$  (right, continuous line) of Figure 6. In can be noticed that the method supplies an accurate prediction of the RUL distribution, and correctly estimates the prediction uncertainty for all values of the creep strain  $\varepsilon_i$ .



**Figure 7:** comparison of the pdf  $\Pr(RUL_i | \varepsilon_i)$  (left, continuous line) and the standard deviation  $\sigma_{A+C}$  (right, continuous line) with, respectively, the pdf  $\Pr(RUL_i | \mathbf{z}_{1:i})$  (left, dashed line) and the estimated standard deviation  $\hat{\sigma}_{rul_i}$  (right, dots), obtained with the PF approach 1, assuming exact knowledge of the parameters  $A$  and  $n$ .

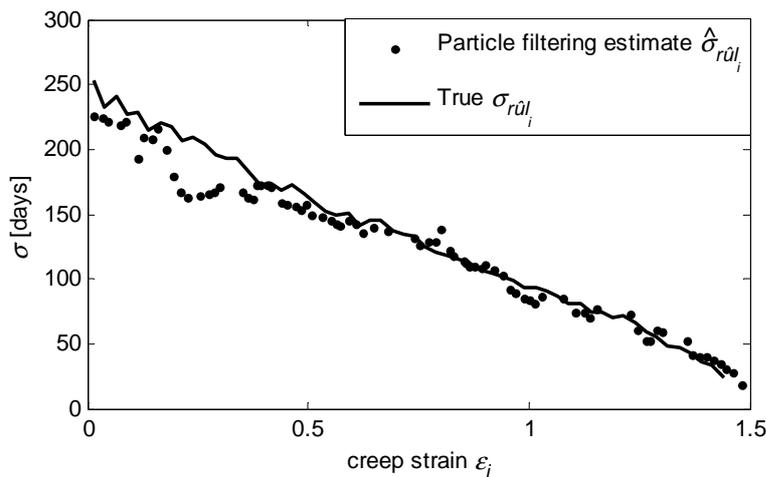
The more realistic case where the exact values of parameters  $A$  and  $n$  are not known has then been considered. In this case, uncertainty in the prognostic model (source of uncertainty B) is introduced. The particle filtering approach 1 handles it by generating particles characterized by different values of  $A$  and  $n$  randomly sampled from the distributions of Table II. In Figure 8, the true RUL distribution  $\Pr(RUL_i | \varepsilon_i)$  (left, continuous line) and standard deviation  $\sigma_{A+C}$  (right, continuous line) of Figure 6 are compared to the distribution  $\Pr(RUL_i | \mathbf{z}_{1:i})$  provided by the method (left, dashed line) and the estimate of the standard deviation  $\hat{\sigma}_{rul_i}$  (right, dots).



**Figure 8:** comparison of the pdf  $\Pr(RUL_i | \varepsilon_i)$  (left, continuous line) and the standard deviation  $\sigma_{A+C}$  (right, continuous line) with, respectively, the pdf  $\Pr(RUL_i | \mathbf{z}_{1:i})$  (left, dashed line) and the estimated standard deviation  $\hat{\sigma}_{rul_i}$  (right, dots), obtained with the PF approach 1, assuming to know only the distribution of the parameters  $A$  and  $n$ .

It can be noticed that the estimated prediction error standard deviation  $\hat{\sigma}_{r\hat{u}_i}$  is larger than the actual RUL standard deviation  $\sigma_{A+C}$ , especially for low values of  $\varepsilon_i$ , due to the model error variance  $\sigma_B^2$ . There are two main reasons for which the difference between  $\hat{\sigma}_{r\hat{u}_i}$  and  $\sigma_{A+C}$  decreases as the current creep strain gets closer to the failure threshold: *i*) the effect of the variability of the parameters  $A$  and  $n$  on the RUL distribution is lower if the gap between the degradation level  $\varepsilon_i$  and the failure threshold  $\varepsilon_{ih}$  is smaller; *ii*) the SIR particle filtering method selects among the large set of particles initially created with random values of  $A$  and  $n$  those having the values of these parameters closer to those of the specific blade undergoing the creep degradation process.

Figure 9 compares the estimated prediction error standard deviation  $\hat{\sigma}_{r\hat{u}_i}$  with the true prediction error standard deviation  $\sigma_{r\hat{u}_i}$  for the PF approach 1. The results confirm that, as expected, the PF approach 1 supplies accurate estimates of the prediction error variance  $\sigma_{r\hat{u}_i}^2$ , combining properly the contribution of the process stochasticity  $\sigma_A^2$ , the noise  $\sigma_C^2$  and the model error  $\sigma_B^2$ , described respectively by the degradation model (eq. (15)), the observation equation (eq. (6)) and the  $A$  and  $n$  parameters distributions (Table II).



**Figure 9:** comparison of the estimated prediction error standard deviation  $\hat{\sigma}_{r\hat{u}_i}$  (dots), with the true prediction error standard deviation  $\sigma_{r\hat{u}_i}$  (continuous line).

### 5.3.2 Prediction interval provided by approach 2

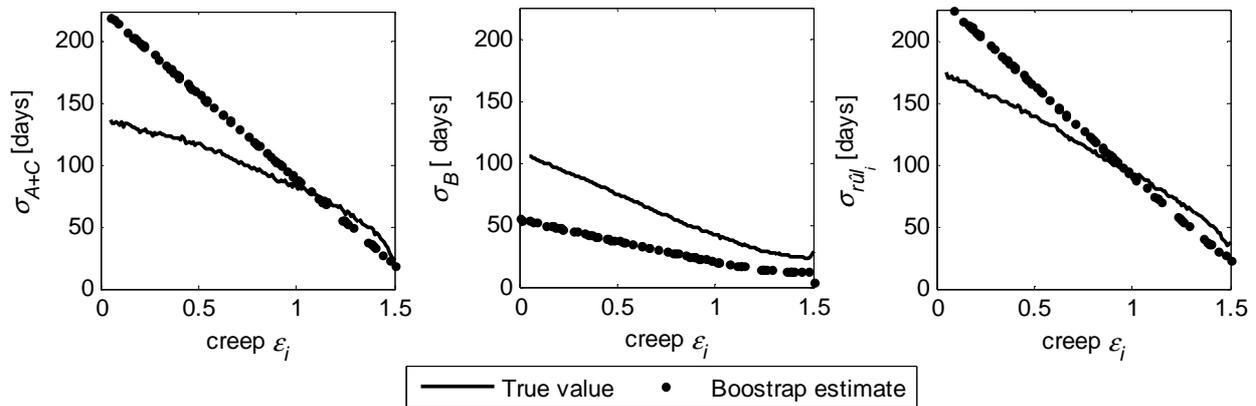
In order to estimate the real uncertainty affecting the RUL prediction of an ensemble of models, it is necessary to add the real model uncertainty  $\sigma_B^2$  to the noise and process randomness represented by  $\sigma_{A+C}^2$  (Figure 6, right). The real model uncertainty  $\sigma_B^2 = E\left[(\mu_{RUL_i|z_i} - r\hat{u}_i)^2\right]$  can be approximated by following the procedure reported in Appendix A. Basically,  $M$  different ensemble models are trained using different sets of creep growth trajectories; then, for  $P''$  test trajectories, the observations  $\mathbf{z}_i^{p''} = \varepsilon_i + v_i^{p''}$ ,  $p''=1, \dots, P''$  are simulated and the RUL predictions  $r\hat{u}_i^{p'',m}$  collected for each ensemble model  $m=1, \dots, M$ ; finally, the

corresponding errors  $\mu_{RUL_i|z_i^{p''}} - r\hat{u}_i^{p'',m}$  are computed, and the square values of these differences are averaged over the  $P''$  trajectories and the  $M$  models to supply the numerical approximation of  $\sigma_B^2$ :

$$\sigma_B^2 \cong \frac{1}{P''} \frac{1}{M} \sum_{p''=1}^{P''} \sum_{m=1}^M [\mu_{RUL_i|z_i^{p''}} - r\hat{u}_i^{p'',m}]^2 \quad (22)$$

The continuous line in Figure 10 shows the real values of  $\sigma_{A+C}$  (left),  $\sigma_B$  (middle) and  $\sigma_{r\hat{u}_i}$  (right) during the life of a turbine blade as a function of its creep strain  $\varepsilon_i$ . It can be observed that the term  $\sigma_{A+C}(\varepsilon_i)$  dominates the term  $\sigma_B(\varepsilon_i)$ , except in the very proximity of the failure threshold where  $\sigma_{A+C}$  goes to zero, whereas the model error remains larger due to the uncertainty related to the unknown value of the failure threshold.

The dotted lines in Figure 9 represent the estimates of these quantities provided by the bootstrap ensemble.

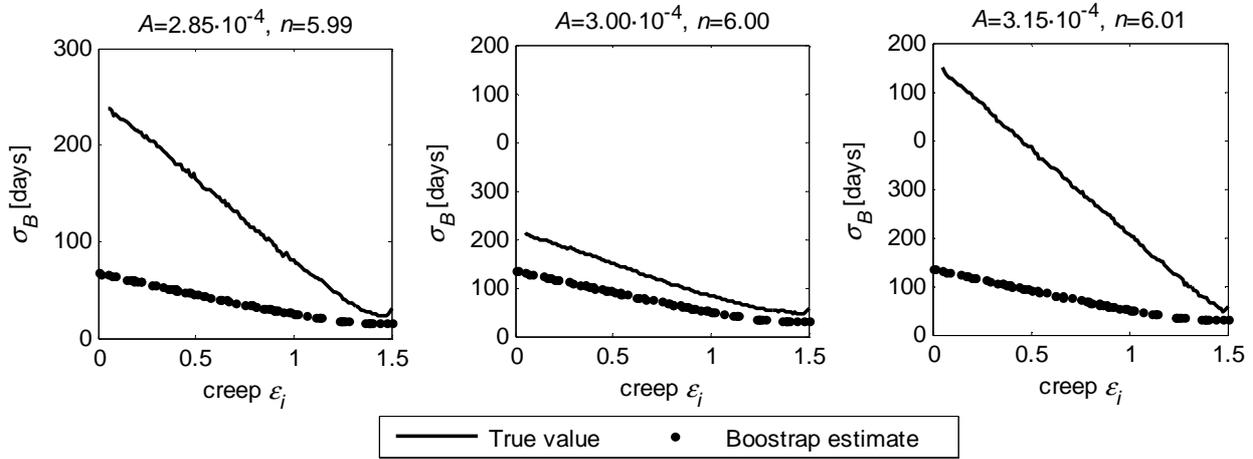


**Figure 10:** comparison of the bootstrap estimates (dashed line) and true values (continuous line) of  $\sigma_B(\varepsilon_i)$  (left),  $\sigma_{A+C}(\varepsilon_i)$  (middle) and  $\sigma_{r\hat{u}_i}(\varepsilon_i)$  (right) during the life of a turbine blade for different values of its creep strain  $\varepsilon_i$ .

The standard deviation  $\sigma_{A+C}$  estimated by the bootstrap ensemble is significantly larger than its real value. As in approach 1, this is due to the fact that the training trajectories have different values of parameters  $A$  and  $n$  and, thus, the empirical model learns the variance of a population of different blades instead of that of the specific blade with fixed values of  $A$  and  $n$ .

In Figure 10, the estimate of the model error variance  $\sigma_B^2$  appears to be not very accurate. Figure 11 shows that the inaccuracy can be even more remarkable if other test trajectories are considered, characterized by values of  $A$  and  $n$  far away from the mean value of their respective distributions ( $\mu_A = 3 \cdot 10^{-4}$  and  $\mu_n = 6$ ). Notice that the real model uncertainty depends on the test trajectory: the model trained on the historical trajectories tends to learn the ‘average’ behavior of the general creep growth trajectory; the consequence is that the model makes larger errors when the test trajectory is different from the ‘average’

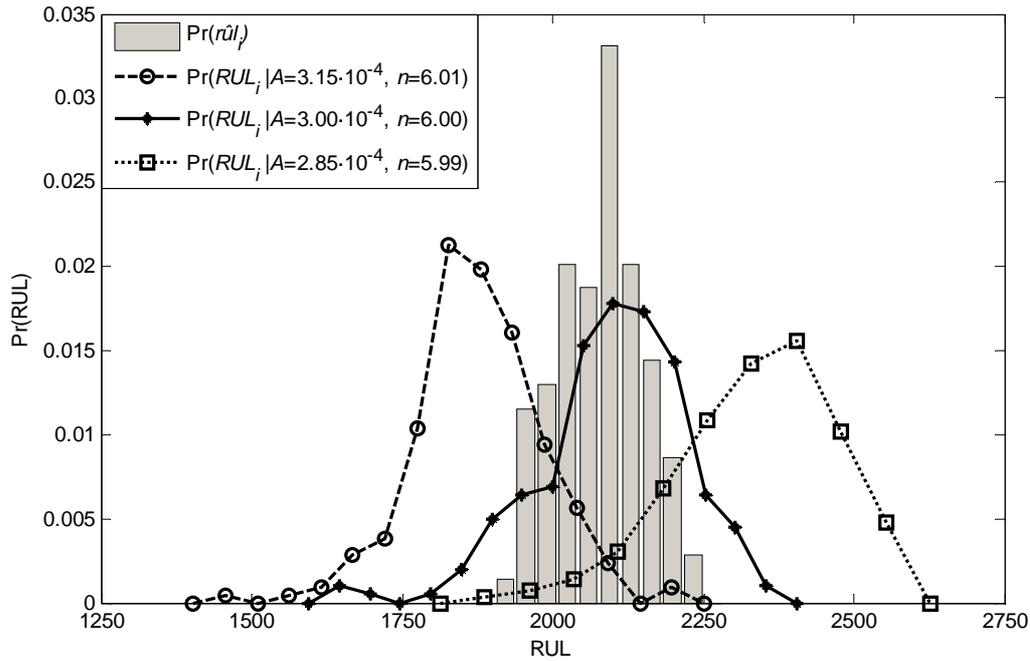
trajectory. On the contrary, the estimate,  $\hat{\sigma}_B^2$ , of the model uncertainty provided by the bootstrap ensemble depends only on the value of  $\varepsilon_i$ , being independent from the specific values of  $A$  and  $n$  of the test trajectory.



**Figure 11: comparison of the bootstrap estimates (dashed line) and true (continuous line) values of  $\sigma_B$  during the life of three turbine blades with different values of parameters  $A$  and  $n$ .**

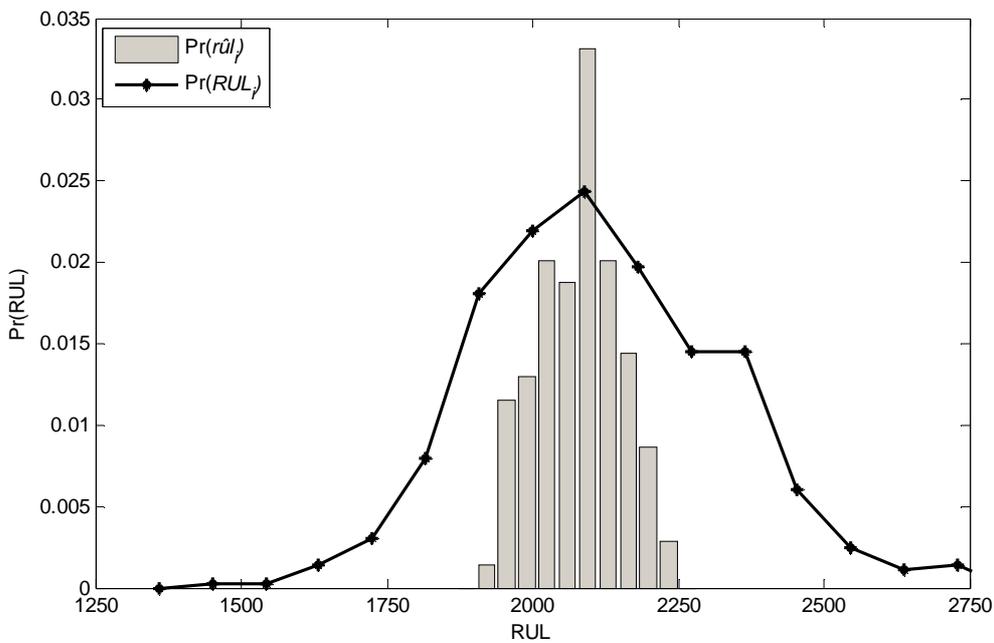
Differently from approach 1, the ensemble approach 2 is not able to learn the true values of  $A$  and  $n$  of the current test trajectory. Furthermore, this limitation of the model is not properly described by the prediction interval provided by the ensemble.

The reason for which the proposed bootstrap approach 2 is not able to correctly model the evolution of the error made by the model for a specific test trajectory is that the assumption that the predictive model  $f(\mathbf{z}_i)$  is an unbiased estimator of the RUL expected value  $\mu_{RUL_i}$ , is not fully verified. In fact, if we build several ensemble models trained with different randomly chosen datasets and perform the RUL prediction with each one of them, we notice that the average RUL prediction  $r\hat{u}_i$  over the different ensemble models is different from the RUL mean value computed over a set of creep growth trajectories with fixed values of  $A$  and  $n$ . This is shown in Figure 12 where the distribution  $\Pr(r\hat{u}_i | \mathbf{z}_i)$  of the prediction  $r\hat{u}_i$  obtained at  $\varepsilon_i = 0.4$  by several ensemble models is compared to the true RUL distribution  $\Pr(RUL_i | \varepsilon_i)$  for 3 different values of parameters  $A$  and  $n$ : 1)  $A = 2.85 \cdot 10^{-4}, n = 5.99$ ; 2)  $A = 3.00 \cdot 10^{-4}, n = 6.00$ ; 3)  $A = 3.15 \cdot 10^{-4}, n = 6.01$ .



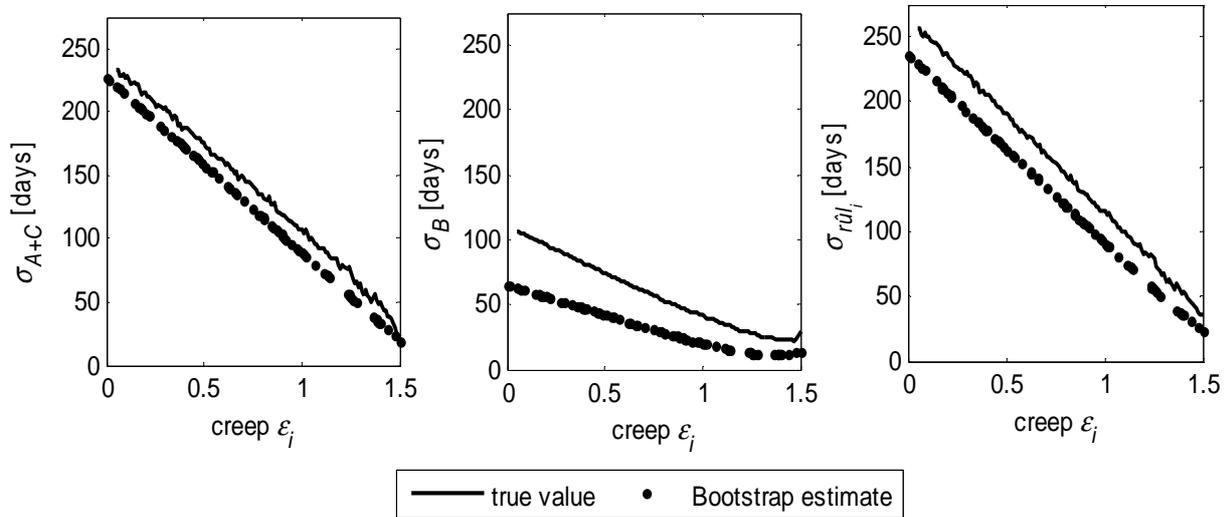
**Figure 12:** comparison of the distributions of the prediction  $\hat{r}ul_i$  with the distribution of the actual RUL of three turbine blades with different values of parameters  $A$  and  $n$  at  $\varepsilon_i = 0.4$ .

On the other side, the models trained on the historical trajectories are unbiased estimators of the RUL of the generic turbine blade with random values of  $A$  and  $n$  as it can be seen by comparing the distribution of the RUL prediction  $\hat{r}ul_i$  with the distribution of the actual RUL of a population of turbine blades with randomly sampled values of  $A$  and  $n$  (Figure 13).



**Figure 13:** comparison of the distributions of the prediction  $\hat{r}ul_i$  with the distribution of the RUL of a population of turbine blades with normally distributed values of  $A$  and  $n$  at  $\varepsilon_i = 0.4$ .

Figure 14 compares the bootstrap estimates (dots) of  $\sigma_{A+C}(\varepsilon_i)$  (left),  $\sigma_B(\varepsilon_i)$  (middle) and  $\sigma_{r\hat{l}_i}(\varepsilon_i)$  (right) with their true values obtained for a population of different turbine blades with parameters  $A$  and  $n$  normally distributed.

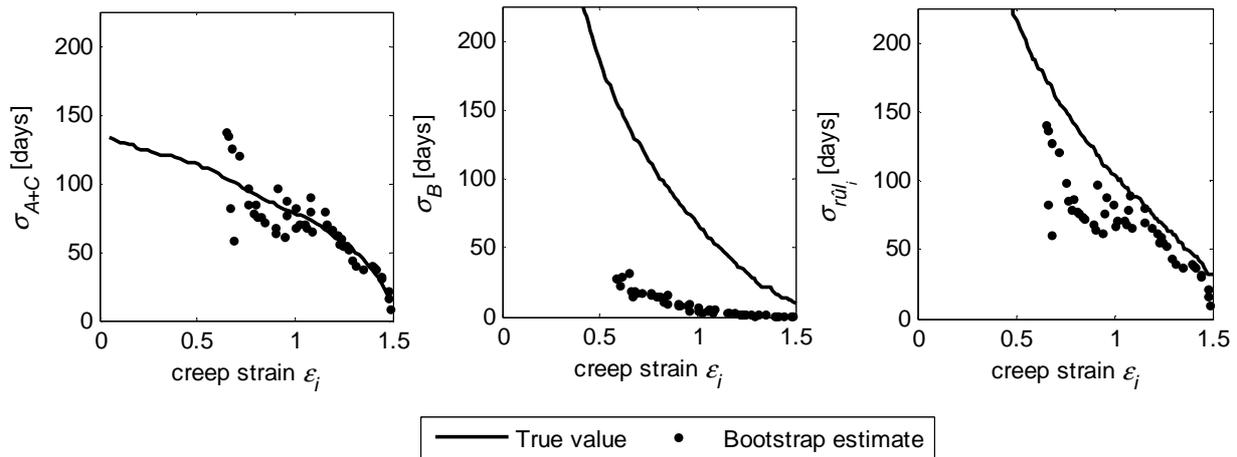


**Figure 14:** comparison of the bootstrap estimates (dashed line) of  $\sigma_{A+C}(\varepsilon_i)$  (left),  $\sigma_B(\varepsilon_i)$  (middle) and  $\sigma_{r\hat{l}_i}(\varepsilon_i)$  (right) with their true values (continuous line) for a population of turbine blades with normally distributed values of  $A$  and  $n$ .

These results confirm that the bootstrap approach 2 can actually provide a satisfactory estimates of  $\sigma_{A+C}(\varepsilon_i)$ ,  $\sigma_B(\varepsilon_i)$  and  $\sigma_{r\hat{l}_i}(\varepsilon_i)$  for a population of different turbine blades, and thus the proposed approach correctly quantifies the uncertainty of the prediction produced by the prognostic model for a generic creeping blade.

### 5.3.3 Prediction interval provided by approach 3

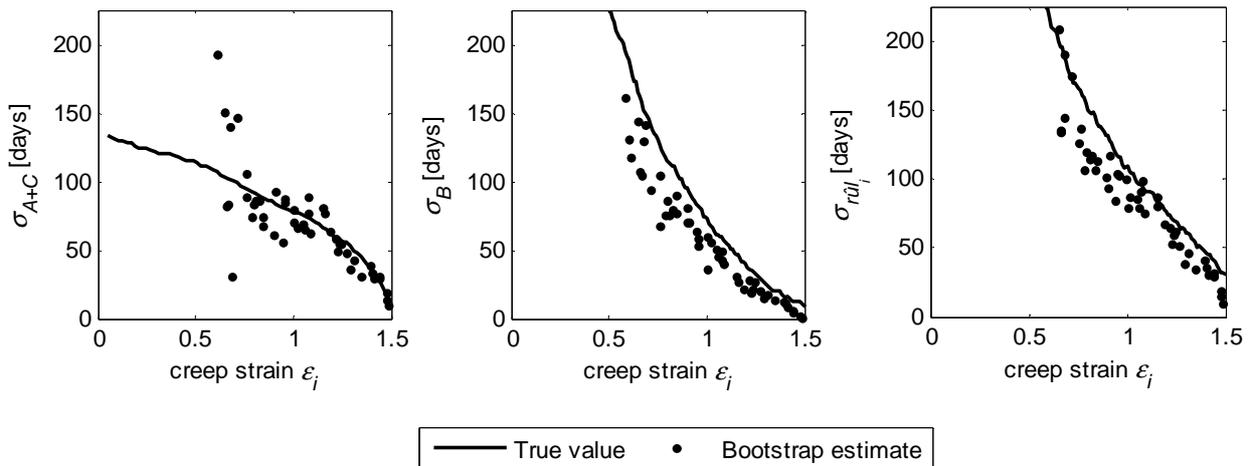
Figure 15 shows the analogous results of Figure 9 for approach 3. Notice that good estimates of  $\sigma_{A+C}(\varepsilon_i)$  are achieved only for large values of  $\varepsilon_i$ , i.e., when a large validation dataset is available and the model  $\chi(\Delta\varepsilon_{i,i'})$  is used in the same range of input values where it has been trained. We also notice that the value of  $\sigma_{A+C}(\varepsilon_i)$  is largely underestimated.



**Figure 15:** comparison of the bootstrap estimates (dots) and true (continuous line) values of  $\sigma_B(\varepsilon_i)$  (left),  $\sigma_{A+C}(\varepsilon_i)$  (middle) and  $\sigma_{r\hat{u}_i}(\varepsilon_i)$  (right) during the life of a turbine blade for different values of its creep strain  $\varepsilon_i$ .

To understand the reason for which the proposed bootstrap approach 3 tends to underestimate the model error variance in this case, it must be pointed out that, from a probabilistic point of view, the single available trajectory used for training the models is only one of an infinite number of possible trajectories, which may be drawn from the creep growth process we wish to model. Thus, bootstrap sampling of creep strain measurements from a single degradation trajectory does not account for the variability of all possible degradation trajectories.

In the case of a linear process, it is possible to overcome this limitation, by considering, instead of the sequence of creep strain measurements  $z_{1:i}$ , the set of independent creep strain increments for time unit  $\Delta z_j = (z_{j+1} - z_j) / (t_{j+1} - t_j)$ ,  $j = 1, \dots, i-1$ . In this way, the variability of the training data is increased, and a better representation of the intrinsic variability of the process is provided. An accurate model of the process can still be achieved by estimating the parameters  $\beta^h$ ,  $h=1, \dots, H$ , of the models of the ensemble  $\tilde{\eta}^h(\Delta \varepsilon_{j,j'}) = \Delta \varepsilon_{j,j'} / \beta^h$  as the average value of the creep strain increments of the bootstrap replicate  $\mathbf{D}_{incr}^h$  of the new training dataset  $\mathbf{D}_{incr}^{trn} = \{\Delta z_{1:i-1}\}$ . Figure 16 shows that a more accurate estimate of  $\sigma_B$  and  $\sigma_{r\hat{u}_i}$  is achieved using this new ensemble of models.



**Figure 16:** comparison of the bootstrap estimates (dots) and true values (continuous line) of  $\sigma_{A+C}$  (left),  $\sigma_B$  (middle) and  $\sigma_{r_{il}}$  (right) during the life of a turbine blade for different values of its creep strain  $\varepsilon_i$  obtained using as training data for the ensemble of models the creep strain increments between consecutive observations.

## 6 Conclusions

Three prognostic approaches have been investigated, particularly with respect to the treatment of the uncertainty in the predicted equipment RUL. Quantitative considerations have been made with regards to a simulated case study concerning the creep growth process in a high temperature turbine blade. The results show that the particle filtering approach provides a good approximation of the exact distribution of the equipment RUL in the case in which an accurate model reproducing the equipment degradation process is available. A limit of particle filtering is particle impoverishment, which relates to the failure of maintaining the diversity of particles and is caused by the resampling approach adopted to avoid particle degeneracy. Particle impoverishment implies the impossibility of the particles to correctly represent all possible evolutions of the degradation process, which include, for example, changes in time due to the variation of operating conditions. In this context, different resampling methods, such as the one proposed in [55], which samples particle considering not only the particle weight but also their spatial distribution (state values), can be considered.

When using model-based approaches, imprecision of the model in the reproduction of the degradation process due to simplifications, incorrect model structure or assumptions on the equipment specific geometries or material properties, etc. can be amplified over time, causing unreliable estimates of the RUL distribution. Using particle filtering, it is possible to include model parameters in the state vector and, thus, perform model adaptation in conjunction with state tracking. In any case, it is very difficult for a physics-based model to account for all aspects of a degradation process; for example, it is common to neglect some of the interactions between different degradation mechanisms or the possible existence of self-healing mechanisms which can reverse the degradation process and are likely to increase the uncertainty of the future degradation evolution. All these non-modeled phenomena can be accounted for by adding further noise to

the process model which will result in a larger confidence interval associated to the RUL estimate. Further research is needed to quantify the impact of modeling errors on the final prediction of model-based approaches.

In the bootstrap approaches 2 and 3 considered, it has been shown that a reliable prediction of the equipment RUL with a correct quantification of its uncertainty can be obtained. With respect to the ensemble of bootstrapped models trained with historical sequences of observations in approach 2, the main limitation is that it is not able to learn the peculiar characteristics of the equipment of interest but it tends to reproduce an ‘average’ behavior. To overtake this problem, a different modeling approach could be used, such as that based on the idea of fuzzy similarity [56], or a procedure for updating the ensemble with the information conveyed by new observations [27].

The application of the bootstrap ensemble in the time series scheme of approach 3, in which only direct measurements of the degradation experienced by the equipment of interest are available, has shown the importance of injecting diversity into the bootstrapped models by using independent training data, in order to correctly quantify the modeling error. The case study considered is characterized by a linear degradation process, so that independent training data can be obtained by considering the creep strain increments between consecutive measurements; on the contrary, this would not be feasible for non-linear degradation processes. Furthermore, in this case of very little information 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.

Contrarily to physics-based models, we expect that data-driven methods can automatically learn from data the effects on the equipment RUL of phenomena influencing the degradation process, such as self-healing and interactions between different degradation mechanisms. The capability of data-driven methods of providing correct estimates of the RUL and its uncertain distribution depends on the availability in the training set of examples of the phenomena that we want to represent.

In this work, the problem of detecting the initiation of a degradation process, which is usually achieved by using properly developed diagnostic systems, has not been addressed. Although none of the approaches presented in this work requires knowing the exact time at which degradation has initiated, late detection of an ongoing degradation process will reduce the number of degradation measures available for prognostics; this is expected to reduce the RUL prediction accuracy and increase its uncertainty especially in approaches 1 and 3 which, contrarily to approach 2, generate their RUL prediction on the basis of past degradation measurements. .

Since only artificial data have been used in the case study considered in this work, conclusions about the successful application of these approaches in the field cannot be directly drawn. The analyses performed have shown the potential of these methods in performing RUL prediction with adequate management of its uncertainty; in this sense, they hold promises for future research aimed at confirming this potential in the application to real data.

## Acknowledgments

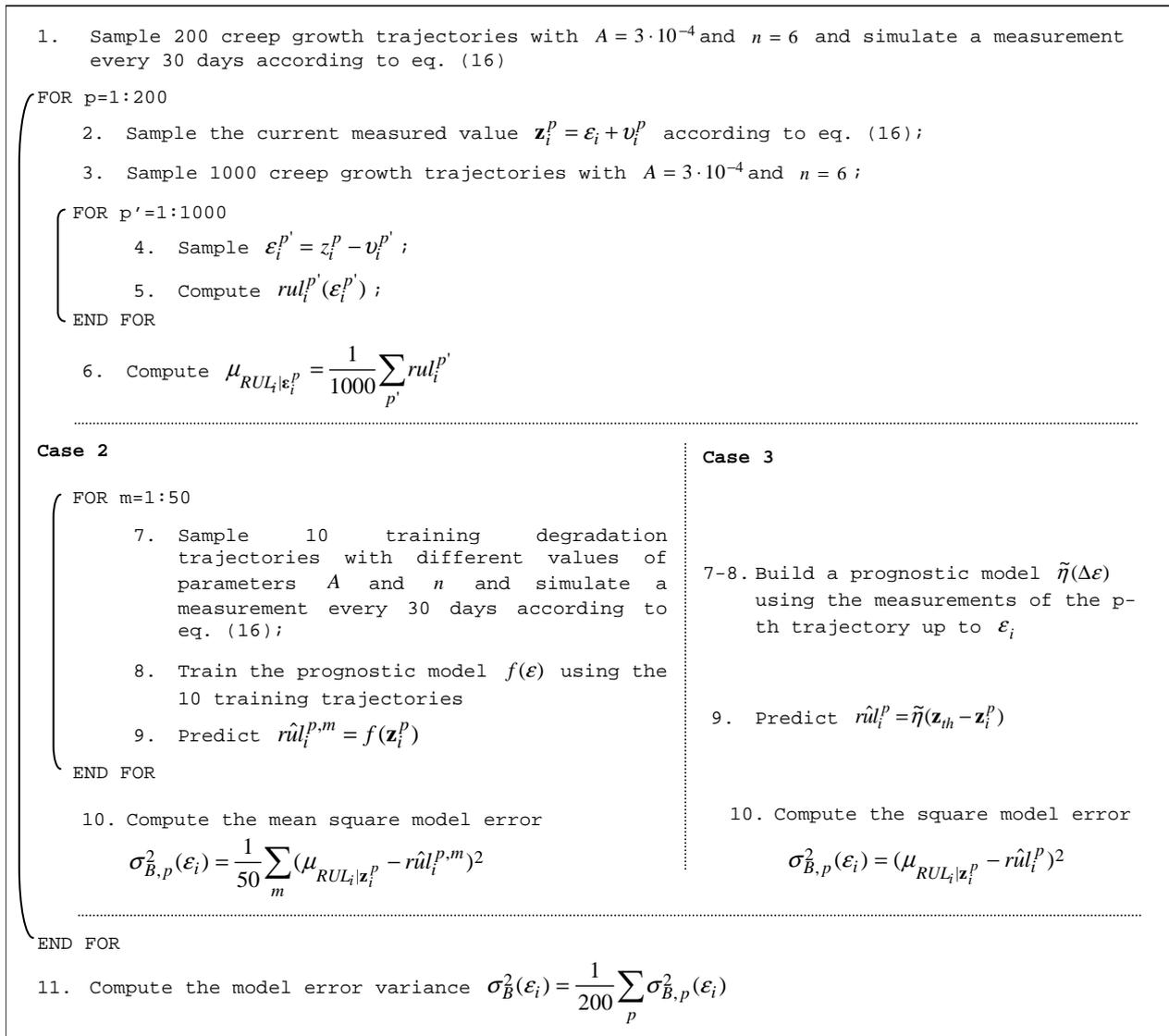
The work of Francesca Mangili has been supported by a PhD grant of the Institutt For Energiteknikk (IFE), OECD Halden Reactor Project.

The participation of Enrico Zio in this research is partially supported by the China NSFC under grant number 71231001.

Finally, the authors wish to thank the reviewers for their constructive comments which have allowed improving the paper.

## Appendix A: approximation of the model error and prediction error variances

In the empirical ensemble-based approaches 2 and 3, the true value of the model error variance  $\sigma_B^2$  has been approximated for different values of the creep strain  $\varepsilon_i$ , as the mean square value of the model error  $\mu_{RUL_i|\mathbf{z}_i} - r\hat{u}_i(\mathbf{z}_i)$  made by different ensembles on 200 creep growth trajectories sampled for a turbine blade with parameters  $A=3 \cdot 10^{-4}$  and  $n=6$ . The details of the procedure are sketched in the pseudo-code of Figure 1A for approaches 2 and 3.



**Figure 1A: procedure for approximating  $\sigma_B^2(\boldsymbol{\varepsilon}_i)$  in approaches 2 and 3.**

## References

- [1] Hess A, Calvello G, Frith P. Challenges, issues, and lessons learned chasing the ‘Big P’: real predictive prognostics Part 1. Proc IEEE Aerosp Conf, 2006 Mar 5-12; Big Sky, MT.
- [2] Tang L, Kacprzyński GJ, Goebel K, Vachtsevanos G. Methodologies for uncertainty management in prognostics. Proc IEEE Aerosp Conf, 2009 Mar 7-14; Big Sky, MT.
- [3] Li W, Pham H. An inspection-maintenance model for systems with multiple competing processes. IEEE Trans Rel, 2005;54(2):318–27.
- [4] Liu R, Ma L, Kang R, Wang N. The modeling method on failure prognostics uncertainties in maintenance policy decision process. Proc 9th Int Conf on Reliab, Maint and Saf (ICRMS), 2011 Jun 12-15; Guiyang, China.
- [5] Saxena A, Celaya J, Saha B, Saha S, Goebel K. Evaluating prognostics performance for algorithms incorporating uncertainty estimates. In Proc. IEEE Aerosp Conf, 2010.
- [6] Urbina A, Mahadevan S, Paez TL. Quantification of margins and uncertainties of complex systems in the presence of aleatoric and epistemic uncertainty, Reliab Eng Syst Saf 2011;96(9):1114-25.
- [7] Zio E. Prognostics and health management of industrial equipment. In: Kadry S, editor. Diagnostics and Prognostics of Engineering Systems: Methods and Techniques, IGI-Global, 2012 (to appear).

- [8] Baraldi P, Cadini F, Mangili F, Zio E. Model-based and data-driven prognostics under different available information. *Probab Eng Mech*. Submitted for publication, 2012.
- [9] Luo J, Pattipati K, Qiao L, Chigusa S. Model-based prognostic techniques applied to a suspension system. *IEEE Trans Syst, Man, and Cybern – Part C: Appl and Rev* 2008;38(5):1156-68.
- [10] Doucet A, de Freitas JFG, Gordon NJ. *Sequential Monte Carlo methods in practice*. New York: Springer-Verlag; 2001.
- [11] Myötyri E, Pulkkinen U, Simola K. Application of stochastic filtering for lifetime Prediction, *Reliab Eng Syst Saf* 2006;91(2):200-8.
- [12] Arulampalam MS, Maskell S, Gordon N, Clapp T. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Trans Sig Process* 2002;50(2):174-88.
- [13] Anderson BD, Moore JB. *Optimal filtering*. Englewood Cliffs (NJ): Prentice Hall; 1979.
- [14] Kitagawa G. Non-Gaussian state-space modeling of nonstationary time series. *J Am Stat Assoc* 1987;82:1032-63.
- [15] Crisan D, Doucet A, A survey of convergence results on particle filtering methods for practitioners. *IEEE Trans on Sig Process*, 2002;50(3):736-46.
- [16] Djuric PM, Kotecha JH, Zhang J, Huang Y, Ghirmai T, Bugallo MF, Miguez J. Particle filtering. *IEEE Sig Process Mag*, 2003;20(5):19-38.
- [17] Pitt MK, Shephard N. Filtering via simulation: Auxiliary particle filters. *J Am Stat Assoc* 1999;94(446):590-9.
- [18] Marseguerra M, Zio E. Monte Carlo simulation for model-based fault diagnosis in dynamic systems. *Reliab Eng Syst Saf* 2009;94(2):180-6.
- [19] Cadini F, Zio E, Avram D. Model-based Monte Carlo state estimation for condition-based component replacement. *Reliab Eng Syst Saf* 2009;94(3):752-8.
- [20] Cadini F, Zio E, Avram D. Monte Carlo-based filtering for fatigue crack growth estimation. *Probab Eng Mech* 2009;24:367-73.
- [21] Zio E, Peloni G. Particle filtering prognostic estimation of the remaining useful life of nonlinear components, *Reliab Eng Syst Saf* 2011;96(3):403-9.
- [22] Zio E, Compare M. Evaluating maintenance policies by quantitative modeling and analysis. *Reliab Eng Syst Saf* 2013;109:53-65.
- [23] Schwabacher MA. A survey of data-driven prognostic, *Proc Infotech@Aerospace*, 2005 September 26-29; Arlington, VA.
- [24] Heskes T. Practical confidence and prediction intervals in Advances. In: Mozer M, Jordan M, Heskes T, editors. *Neural Information Processing Systems 9*, Cambridge, MA: MIT Press; 1997, p. 466–72.
- [25] Raviv Y, Intrator N. Bootstrapping with noise: An effective regularization technique. *Connect Sci* 1996;8(3-4):355-72.
- [26] Couturier R, Escaravage C. High temperature alloys for the HTGR gas turbine: Required properties and development needs. *Proc IAEA Tech Comm Meet on Gas Turbine Power Convers Syst for Modular HTGRs*. 2000 Nov 14-16; Palo Alto, CA.
- [27] Baraldi P, Mangili F, Zio E. A Kalman filter - based ensemble approach with application to turbine creep prognostics. *IEEE Trans on Reliab* 2012, to appear.
- [28] Saxena A, Celaya J, Saha B, Saha S, Goebel K. Metrics for offline evaluation of prognostic performance. *Int. J. of PHM*, 2010;1:1-20.
- [29] Usynin A, Hines JW, Urmanov A. Uncertain failure thresholds in cumulative damage models. In: *Proc. of the Annual Reliability and Maintainability Symposium*. IEEE Conference Publications; 2008, p. 334–340.
- [30] Wang P, Coit DW. Reliability and degradation modeling with random or uncertain failure threshold. In: *Proc. of the Annual Reliability and Maintainability Symposium*. IEEE Conference Publications; 2008, p. 334–340.
- [31] Orchard M, Kacprzynski G, Goebel K, Saha B, Vachtsevanos G. Advances in uncertainty representation and management for particle filtering applied to prognostics. *Proc Int PHM Conf*, 2008 Oct 6-9; Denver, CO.
- [32] Kotecha JH, Djuric PM. Gaussian sum particle filtering. *IEEE Trans Sig process* 2003;51(10): 2592-601.

- [33] Saha B, Goebel K, Christophersen J. Comparison of prognostic algorithms for estimating remaining useful life of batteries. *Trans Inst of Meas & Control* 2009;31(3-4):293-308.
- [34] Lopes HF, Carvalho CM. Online Bayesian learning in dynamic models: An illustrative introduction to particle methods. In West M, Damien P, Dellaportas P, Polson NG, Stephens DA, editors. *Bayesian Dynamic Modelling Bayesian Inference and Markov Chain Monte Carlo: In Honour of Adrian Smith*, Clarendon: Oxford University Press; 2012 (to appear).
- [35] Wang X, Rabiei M, Hurtado J, Modarres M, Hoffman P. A probabilistic-based airframe integrity management model, *Reliab Eng Syst Saf* 2009;94(5):932-41.
- [36] Zio E. A Study of the bootstrap method for estimating the accuracy of artificial neural networks. Predicting nuclear transient processes. *IEEE Trans Nucl Sci* 2006;53(3):1460-78.
- [37] Efron B, Tibshirani RJ. *An Introduction to the Bootstrap*. New York: Chapman and Hall; 1993.
- [38] Carney JG, Cunningham P, Bhagwan U. Confidence and prediction intervals for neural network ensembles. *Proc. Int Joint Conf Neural Netw*, 1999 Jul 10-16; Washington, DC.
- [39] Gorjian N, Ma L, Mittinty M, Yarlagadda P, Sun Y. Review on degradation models in reliability analysis. *Proc 4th World Congr on Eng Asset Manag*. 2009 Sept 28-30; Athens, Greece.
- [40] Carter TJ, Common failures in gas turbine blades. *Eng Failure Analysis*, 2005;12:237-47.
- [41] Dunn TD, Lommers LJ, Tangira VE. Preliminary safety evaluation of the gas turbine-modular Helium reactor (GT-MHR). *Proc Int Top Meet Adv React Saf* 1994, Apr 17-21, Pittsburgh, PA.
- [42] Saez M, Tauveron N, Chataing T, Geffraye G, Briottet L, Alborghetti N. Analysis of the turbine deblading in an HTGR with the CATHARE code. *Nucl Eng and Des* 2006;236:574-86.
- [43] Made M, Mirmiran A, Walter TA. Local damage assessment of turbine missile impact on composite and multiple barriers. *Nucl Eng and Des*, 1997;178:145-56.
- [44] Goel N, Kumar A, Narasimhan V, Nayak A, Srivastava A. Health risk assessment and prognosis of gas turbine blades by simulation and statistical methods. *Can Conf on Electr and Comput Eng*. 2008 May 4-7; Niagara Falls, ON.
- [45] Flotow A, Mercadal M, Tappert P. Health monitoring and prognostics of blades and disks with blade tip sensors, *Proc IEEE Aerosp Conf*. 2000 Mar 18-25; Big Sky, MT.
- [46] Ye D, Duan F, Guo H, Li Y, Wang K. Turbine blade tip clearance measurement using a skewed dual-beam fiber optic sensor. *Opt Eng*, 2012;51(8).
- [47] Steiner A. Techniques for blade tip clearance measurements with capacitive probes. *Meas Sci Technol*, 2000;11:865-9.
- [48] <http://www.capacisense.com/>
- [49] <http://www.thermocoax.com>
- [50] Dowell M, Sylvester G, Krupp R, Zipfel G. Progress in turbomachinery prognostics and health management via eddy-current sensing, In *Proc. IEEE Aerosp Conf*, 2000, vol. 6, p. 133-43.
- [51] Kwapisz D, Hafner M, Rajamani R. application of microwave sensing to blade health monitoring. In *Proc 1st Eur Conf of PHM Soc*, 2012, vol. 3, p. 1-8.
- [52] Polikar R, Topalis A, Green D, Kounios J, Clark CM. Comparative multiresolution wavelet analysis of ERP spectral bands using an ensemble of classifiers approach for early diagnosis of Alzheimer's disease. *Comput in Biol and Med*, 2007;37:542-58.
- [53] Carter TJ, Common failures in gas turbine blades. *Eng Fail Anal* 2005;12(2):237-47.
- [54] Swindeman RW, Swindeman MJ. A comparison of creep models for nickel base alloy for advanced energy systems, *Int J of Press Vessels and Pip* 2008;85:72-9.
- [55] Li T, Sattar TP, Sun S. Deterministic resampling: Unbiased sampling to avoid sample impoverishment in particle filters. *Signal Process*, 2012;92:1637-45.
- [56] Zio E, Di Maio F. A data-driven fuzzy approach for predicting the remaining useful life in dynamic failure scenarios of a nuclear system. *Reliab Eng and Syst Saf* 2010;95(1):49-57.