Prediction of Industrial Equipment Remaining Useful Life by Fuzzy Similarity and Belief Function Theory

Piero Baraldi

Dipartimento di Energia, Politecnico di Milano, Via La Masa 34, 20156 Milano, Italy, piero.baraldi@polimi.it

Francesco Di Maio

Dipartimento di Energia, Politecnico di Milano, Via La Masa 34, 20156 Milano, Italy, francesco.dimaio@polimi.it

Sameer Al-Dahidi

Dipartimento di Energia, Politecnico di Milano, Via La Masa 34, 20156 Milano, Italy, sameer.aldahidi@polimi.it

Enrico Zio[a,b]

[a]Dipartimento di Energia, Politecnico di Milano, Via La Masa 34, 20156 Milano, Italy, enrico.zio@polimi.it

[b]Chair System Science and the Energy Challenge, Fondation Electricité de France (EDF), CentraleSupélec, Université Paris Saclay, Grande Voie des Vignes, 92290 Chatenay-Malabry, France, enrico.zio@ecp.fr

Francesca Mangili

Istituto Dalle Molle di studi sull'Intelligenza Artificiale (IDSIA), SUPSI, USI, Galleria 2, Via Cantonale 2c CH-6928, Manno, Switzerland, francesca@idsia.ch

Abstract

We develop a novel prognostic method for estimating the Remaining Useful Life (*RUL*) of industrial equipment and its uncertainty. The novelty of the work is the combined use of a fuzzy similarity method for the *RUL* prediction and of Belief Function Theory for uncertainty treatment. This latter allows estimating the uncertainty affecting the *RUL* predictions even in cases characterized by few available data, in which traditional uncertainty estimation methods tend to fail. From the practical point of view, the maintenance planner can define the maximum acceptable failure probability for the equipment of interest and is informed by the proposed prognostic method of the time at which this probability is exceeded, allowing the adoption of a predictive maintenance approach which takes into account *RUL* uncertainty. The method is applied to simulated data of creep growth in ferritic steel and to real data of filter clogging taken from a Boiling Water Reactor (BWR) condenser. The obtained results show the effectiveness of the proposed method for uncertainty treatment and its superiority to the Kernel Density Estimation (KDE) and the Mean-Variance Estimation (MVE) methods in terms of reliability and precision of the *RUL* prediction intervals.

Keywords: Prognostics; Remaining Useful Life; Uncertainty; Fuzzy similarity; Belief function; Boiling Water Reactor condenser.

1 Introduction

Various data-driven methods have been proposed for predicting the Remaining Useful Life (RUL) of degrading equipment (Hines & Usynin, 2008; Vachtsevanos, 2006; Zio, 2012), i.e., the time left before the equipment will stop fulfilling its functions. Data-driven methods are of interest when an explicit model of the degradation process is not known; they are built based on observations of the degradation process of one or more similar equipment, and usually perform the regression of the future equipment degradation path until pre-defined criteria of failure are met (Niu et al., 2010; Baraldi et al., 2012a-b, Baraldi et al., 2013a-c; Di Maio 2012; Zio & Di Maio 2012). Among data-driven methods one can distinguish between (i) degradation-based approaches, modeling the future equipment degradation evolution and (ii) RUL prediction approaches, directly predicting the RUL (Wang et al. 2008).

Degradation-based approaches use statistical models that *learn* the equipment degradation evolution from time series of the observed degradation states (Gorjian et al., 2009; Wang et al., 2011; Zhao et al., 2013); the predicted degradation state is, then, compared with the failure criteria, e.g., the threshold of the degradation parameter beyond which the equipment fails performing its function (failure threshold). Examples of modeling techniques used in degradation-based approaches are Auto-Regressive models (Benkedjouh et al., 2013; Gorjian et al., 2009), multivariate adaptive regression splines (Lee et al., 2006), Relevance Vector Machines (Nystad, 2009; Di Maio et al., 2012) and Gaussian Processes (Rasmussen, 2006; Baraldi et al., 2013a).

RUL prediction approaches, instead, typically resort to artificial intelligence techniques that directly map the relation between the observable parameters and the equipment RUL, without the need of predicting the equipment degradation state evolution towards a failure threshold (Peel, 2008; Schwabacher et al., 2007). Techniques used in direct RUL prediction approaches are most often similarity-based (also known as instance-based) learning algorithms (Zio et al., 2010a; Zhang et al., 2015). As these methods avoid performing explicit generalization, they have proved to be effective also when few training data with no clear patterns of regularity are available for training. Others regression methods, such as ANNs, could be used to perform direct RUL prediction, however, due to the large number of parameters to be tuned in these models, they typically require large training samples to provide accurate models that do not overfit the data.

Degradation–based prognostics provides more informative and transparent outcomes than direct RUL prediction prognostics, since it supplies a prediction not only of the current equipment RUL, but also of the entire degradation trajectory that the equipment will follow. However, degradation-based prognostics, differently than direct RUL prediction prognostics, requires identifying a degradation indicator and fixing a failure threshold, which may be difficult in practice, especially in cases where only few and/or irregular degradation trajectories are available, and may introduce further uncertainty and sources of errors. In fact, the information available for modeling the equipment degradation may be scarce and incomplete, e.g., few examples of similar equipment degradation trajectories may be available, the degradation state of the equipment may not be directly measured, and the failure criteria may not be known with precision. Therefore, the RUL estimate should take into account the intrinsic uncertainty due to the variability of the degradation

process (caused, for example, by the micro-structural differences between pieces of the same equipment, or by unforeseen future loads, operational settings and external conditions) (Baraldi et al. 2012), which implies that we cannot be sure that two identical pieces of equipment, having experienced the same degradation path up to the present time, will keep following exactly the same path also in the future.

Thus, given the scarcity of information typically available and the different sources of uncertainty to which the RUL estimate is subject to (i.e., due to different environmental conditions, measurement noise, process noise, etc. (Al-Dahidi et al., 2014)), data-driven models can commit errors in the RUL estimate (Yan et al., 2004), and uncertainty management becomes a fundamental task in prognostics. Indeed, it is necessary to provide maintenance planners with an assessment of the expected mismatch between the real and predicted equipment failure times, in order to allow them confidently planning maintenance actions, according to the maximum acceptable failure probability (Tang et al., 2009).

However, in spite of the recognized potential of the data-driven approaches, they still face difficulties in providing a measure of confidence on the RUL predictions, i.e., the uncertainty affecting the predictions. For example, fuzzy similarity-based model (Zio & Di Maio, 2010a) and regression methods such as ANNs (Vachtsevanos & Wang, 2001) typically do not provide an explicit and direct quantification of the RUL prediction uncertainty, whereas other methods such as Relevance Vector Machine (Nystad, 2009; Di Maio et al., 2012) or Gaussian Process Regression (Rasmussen, 2006: Baraldi et al., 2013a) have been shown capable of quantifying RUL prediction uncertainty in cases in which a training set made by a large number of examples of the phenomena that we want to represent is available (Baraldi et al., 2013c), but they may experience difficulties in cases of scarce available data.

In this context, the objective of the present work is to develop a novel method for properly representing the uncertainty in the RUL prediction. In practice, the maintenance planner defines the maximum acceptable failure probability and is informed by the prognostic method of the time at which this probability will be exceeded. To this purpose, we consider the direct RUL similarity-based prognostic model proposed in (Zio & Di Maio, 2010a), which uses a set of degradation trajectories collected in a reference library and performs a data-driven similarity analysis for predicting the RUL of a newly developing degradation trajectory (hereafter called test trajectory). The matching process is based on the evaluation of the distance between the reference and test trajectories (Angstenberger, 2001). This method has been selected because of its favorable characteristics in terms of capability of dealing with few and/or irregular degradation trajectories in comparison with other time-series approaches for direct RUL prediction. This prognostic model is here extended in order to provide a measure of confidence in the RUL prediction. To address this issue, we adopt a solution based on belief function theory (BFT) (also called Dempster-Shafer or evidence theory) (Dempster, 1967; Shafer, 1976; Su et al. 2011). The BFT allows combining different pieces of (uncertain) evidence, based on the assignment of basic belief masses to subsets of the space of all possible events, which are, in this case, the possible values that the equipment RUL can take. In practice, the proposed method considers each reference trajectory as a piece of evidence regarding the value of the RUL of the test trajectory. These pieces of evidence are discounted

based on their similarity to the test trajectory and pooled using Dempster's rule of combination (Altinçay, 2007, Petit-Renaud & Denoeux, 2004). The result is a basic belief assignment (BBA) that quantifies one's belief about the value of the RUL for the test trajectory given the reference trajectories. From the BBA, the total belief (i.e., the amount of evidence) supporting the hypothesis that the RUL will fall in any specific interval can be computed. In this context, we propose to define a prediction interval as an interval to which a sufficiently large total belief has been assigned.

The method is applied to two case studies considering simulated data generated by a non-linear model of creep growth in ferritic steel and real industrial data concerning the clogging of filters used to clean the sea water pumped in a Boiling Water Reactor (BWR). The performance of the proposed method is verified with respect to three performance indicators (i.e., Mean Square Error (MSE) for estimating the accuracy of the RUL predictions, Coverage (Cov) for the reliability of the prediction intervals and Mean Amplitude (MA) for their precision (Baraldi et al., 2015). For comparison, the Kernel Density Estimation (KDE) (Botev et al., 2010) and the Mean-Variance Estimation (MVE) (Nix & Weigend, 1994) methods which have already been successfully applied for estimating RUL predictions uncertainty in different prognostic applications on industrial components such as turbofan engines (Wang, 2010) and turbine blades (Baraldi et al., 2012a), are applied to the same case studies and their results are compared to those obtained by the proposed method.

The remaining part of the paper is organized as follows: in Section 2, the methodology for the direct RUL similarity-based prediction of equipment RUL is described and a method for integrating it with belief function theory is proposed to provide a measure of confidence in the similarity-based RUL prediction; in Section 3, two numerical applications concerning the growth of creep damage in ferritic steel and the clogging of sea water filters are presented, and the results obtained by the proposed method are discussed and compared with those obtained by two alternative methods. Finally, some conclusions are drawn in Section 4.

2 Methodology

We assume to have R reference trajectories, which contain measurements collected during the degradation of R pieces of equipment similar to the one currently monitored (test equipment). Let $\boldsymbol{z}_{1:n^r}^r = [\boldsymbol{z}_1^r, \dots, \boldsymbol{z}_i^r, \dots, \boldsymbol{z}_{n^r}^r]$, $r = 1, \dots R$, be a reference trajectory, where $\boldsymbol{z}_i^r = [z_1^r(\tau_i), \dots, z_p^r(\tau_i), \dots, z_P^r(\tau_i)]$ and $z_P^r(\tau_i)$ is the value of parameter z_p^r measured at time τ_i for trajectory r, and let τ_F^r be its failure time.

Let $\boldsymbol{z}_{1:I}^{test} = [\boldsymbol{z}_1, \dots, \boldsymbol{z}_i, \dots, \boldsymbol{z}_I]$ be the test trajectory, containing I observations for the equipment of interest from τ_i to the present time τ_I.

2.1 Similarity-based RUL prediction

The idea underpinning the RUL estimation method is to evaluate the similarity between the test trajectory and the R reference trajectories, and to use the $RULs$ corresponding to the latter to estimate the RUL corresponding to the test trajectory (Guha and Chakraborty, 2010; Liu et al. 2012; Petit-Renaud & Denoeux, 2004; Wang et al. 2008; Zio & Di Maio, 2010a).

Trajectory similarity is evaluated considering the pointwise difference between n-long sequences of observations. Let $\boldsymbol{z}_{I-n+1:I}^{test}$ be the sequence of the n latest observations available for the test trajectory and $\boldsymbol{z}_{j-n+1:j}^{r}$, a sequence of the same length extracted from the reference trajectory r, we take as measure of the distance between $\boldsymbol{z}_{I-n+1:I}^{test}$ and $\boldsymbol{z}_{j-n+1:j}^{r}$ the quantity

$$\delta_j^r = \sqrt{\sum_{i=1}^{n} d^2(\boldsymbol{z}_{I-n+i}^{test}, \boldsymbol{z}_{j-n+i}^{r})} \tag{1}$$

where $d^2(\mathbf{x}, \mathbf{y})$ is the square Euclidean distance between vectors $\mathbf{x}$ and $\mathbf{y}$. Then, the similarity s_j^r between $\boldsymbol{z}_{I-n+1:I}^{test}$ and $\boldsymbol{z}_{j-n+1:j}^{r}$ is defined as a function of the distance measure δ_j^r (Zio & Di Maio, 2010b):

$$s_j^r = \exp(-\frac{\left(\delta_j^r\right)^2}{\lambda}) \tag{2}$$

The value of the arbitrary parameter λ is set by the analyst based on an optimization procedure, which will be explained in Section 3 directly on the case studies: the smaller is the value of λ, the stronger the definition of similarity. A strong definition of similarity implies that the two segments under comparison have to be very close in order to receive a similarity value s_j^r significantly larger than zero. In practice, the parameter λ is often set to the value that minimizes the error of the similarity-based prediction computed on a validation dataset.

Based on this definition of similarity, for each reference trajectory. we can identify the n-long sequences of observations with highest similarity with the test sequence $\boldsymbol{z}_{I-n+1:I}^{test}$. Let $\tau_{j^*}^r, r = 1, \dots, R$, be the last time instant of such most similar sequences. Then, for each reference trajectory, we retain its RUL at time $\tau_{j^*}^r$

$$RUL^r = \tau_F^r - \tau_{j^*}^r \tag{3}$$

as a prediction of the RUL of the test trajectory. Finally, the similarity-based prediction $\widehat{RUL}$ of the test equipment RUL at time τ_I is given by the weighted sum of the values RUL^r:

$$\widehat{RUL} = \frac{\sum_{r=1}^{R} s_{j^*}^r RUL^r}{\sum_{r=1}^{R} s_{j^*}^r} \tag{4}$$

The idea behind the weighting of the predictions RUL^r associated to the individual trajectories is that: i) all failure trajectories in the reference library can, in principle, bring useful information for determining the RUL of the trajectory currently developing; ii) those segments of the reference trajectories which are most similar to the latest part of the test trajectory should be the most informative for its RUL computation.

2.2 Prediction interval based on belief function theory

Uncertainty affects the RUL estimate and, thus, maintenance plans cannot be based only on the RUL prediction provided by eq. (4): a reliable indicator of its uncertainty must be also considered. In this Section, we assume that the maintenance planner is able to specify a maximum acceptable equipment failure probability, α, and we propose a method, based on the Belief Function (or Dempster-Shafer) Theory (BFT) (Dempster, 1976; Shafer, 1976), to identify the latest time at which, according to the available information, we can guarantee that the probability of the equipment to be failed is lower than α. Since in this work we consider situations

characterized by degradation processes affected by large variability and we use an empirical model developed using few degradation trajectories, we expect RUL predictions to be characterized by large uncertainty. In this work, we adopt an uncertainty representation method based on BFT because its capability of representing limited knowledge on an uncertain quantity (Yager, 1987; Helton, 2004). If we consider, for example, an extreme case in which the only information available on the equipment RUL is that it will lie in the interval $[0, \tau_F^{max}]$, the classical probabilistic representation of the uncertainty will be, according to the principle of indifference, an uniform distribution with range $[0, \tau_F^{max}]$. However, as it has been shown in (Yager, 2011), this assignment causes the paradox that it assigns a precise probability value to an event such as "RUL in the interval $[0, \tau_F^{max}/2]$", whereas, according to the available knowledge, the probability of this event can have any value between 0 and 1. For these reasons, in the presence of large uncertainty on the RUL prediction, we suggest to use an approach based on the BFT.

For the ease of clarity and for completeness of the paper, the notions of BFT necessary for understanding the proposed method will be recalled in the following. For further details about the mathematical developments and the possible interpretations of the theory, the interested reader is referred to Dempster (1976), Shafer (1976) and Smets (1998).

The BFT represents the belief of an agent about the value of an uncertain variable Y assuming values y in the frame of discernment Ω_Y. Based on the available information and knowledge, the agent provides a basic belief assignment (BBA) made of a set of masses $m_Y(\mathrm{Y}_k)$ assigned to subsets $\mathrm{Y}_k, k = 1,2,\ldots$ of Ω_Y, based on the available information. The mass $m_Y(\mathrm{Y}_k)$ represents the belief that the value of Y belongs to the subset Y_k. Any subset Y_k with associated a finite mass $m_Y(\mathrm{Y}_k) > 0$ is called focal element; the BBA verifies the condition that the sum of all its masses is 1.

Let us assume that two agents, with two different sources of information and knowledge, provide two BBAs m_Y^1 and m_Y^2. According to the Dempster's rule of combination, the two BBAs can be aggregated into the BBA $m_Y^{1\oplus 2}$:

$$m_Y^{1\oplus 2}(\mathrm{Y}_k) = \frac{1}{K}\sum_{\mathrm{Y}_{k'}\cap\mathrm{Y}_{k''}=\mathrm{Y}_k} m_Y^1(\mathrm{Y}_{k'})m_Y^2(\mathrm{Y}_{k''}),\ \forall \mathrm{Y}_k \in \Omega_Y, Y_k \neq \emptyset \tag{5}$$

$$m_Y^{1\oplus 2}(\emptyset) = 0$$

where

$$K = 1 - \sum_{\mathrm{Y}_{k'}\cap\mathrm{Y}_{k''}=\emptyset} m_Y^1(\mathrm{Y}_{k'})m_Y^2(\mathrm{Y}_{k''}) \tag{6}$$

is a normalization factor introduced to convert a possibly subnormal BBA (i.e., a BBA assigning a finite mass to the empty set $\emptyset$) into a normal one.

It may occur that one doubts the reliability of a source of information inducing the BBA m_Y. In this case, the discounting operation can be used to reduce by some factor $\chi \in [0,1]$ the belief assigned by m_Y to the evidence conveyed by that information (Petit-Renaud & Denoeux, 2004):

$$\widetilde{m}_Y(\mathrm{Y}_k) = (1-\chi)m_Y(\mathrm{Y}_k), \forall \mathrm{Y}_k \in \Omega_Y, \mathrm{Y}_k \neq \Omega_Y \tag{7}$$

$$\tilde{m}_Y(\Omega_Y) = \chi + (1-\chi)m_Y(\Omega_Y)$$

Notice that the mass assigned to the frame of discernment Ω_Y represents the *ignorance* about the value of Y because it indicates the absence of evidence that the value of Y belongs to any subset Y_k of Ω_Y.

The BFT has been applied to treat uncertain information in classical nonparametric regression by associating to each training pattern of input/output pairs (x_i, v_i) the BBA $m_Y(\mathrm{Y}_i = \{y_i\}) = 1$ having as single focal element the pattern output v_i (Petit-Renaud & Denoeux, 2004). In a similarity-based approach each training pattern is treated as an expert whose opinion is assumed to be the more relevant the more similar the pattern is to the test input x (i.e., the larger the similarity, the more useful the information for RUL estimation). Such belief is well modeled by a discounting operation that reduces the belief $m_Y(\mathrm{Y}_i = \{y_i\}) = 1$ of a training pattern (x_i, y_i) proportionally to its dissimilarity to the test pattern. In particular, in the application to prognostics, we assign to each input/output pair $(\boldsymbol{z}^r_{1:n^r}, RUL^r)$, $r = 1{:}\,R$, of a reference trajectory and corresponding RUL prediction RUL^r, the BBA $m^r_{RUL}(\{RUL^r\}) = 1$ and the discounting factor χ defined by $\chi = 1 - \gamma.s^r_{j^*}$ where $\gamma \in [0,1]$ represents the degree of trust given to the entire set of reference trajectories and the similarity $s^r_{j^*}$ is given by eq. (2). Thus, from eq. (7), the discounted BBAs $\tilde{m}^r_{RUL}(\{RUL^r\})$, $r = 1{:}\,R$ are obtained:

$$\begin{aligned} \tilde{m}^r_{RUL}(\{RUL^r\}) &= \gamma.s^r_{j^*} \\ \tilde{m}^r_{RUL}(\Omega_{RUL}) &= 1 - \gamma.s^r_{j^*} \end{aligned} \tag{8}$$

The frame of discernment Ω_{RUL} is the domain of RUL defined by the interval $[0, \tau_F^{max} - \tau_I]$, where τ_F^{max} is the maximum possible life duration of the equipment provided by an expert. The quantity $\tau_F^{max} - \tau_I = RUL^{max}$ is the maximum value that can be assumed by the variable RUL at the present time τ_I, whereas 0 is, obviously, the minimum possible value of the equipment RUL.

It is important to notice that $\gamma < 1$ implies that a part of belief is assigned to the *ignorance* represented by Ω_{RUL}, even in the unrealistic case of a reference trajectory exactly identical to the test one. Then, parameter γ represents the analyst prior opinion about the maximum information that can be derived from a reference trajectory about the test trajectory. In fact, the belief assigned to the event $\widehat{RUL} = RUL^r$ when the two trajectories $\boldsymbol{z}^r_{1:n^r}$ and $\boldsymbol{z}^{test}_{1:I}$ are identical, that is when $s^r_{j^*} = 1$, is equal to γ.

Finally, by combining the discounted BBAs $\tilde{m}^r_{RUL}$, $r = 1{:}\,R$ by the Dempster's rule of combination, we obtain the combined BBA m_{RUL}:

$$\begin{aligned} m_{RUL}(\{RUL^r\}) &= \frac{\gamma.s^r_{j^*}}{K}\prod_{r' \neq r}\left(1 - \gamma.s^{r'}_{j^*}\right), r = 1{:}\,R \\ m_{RUL}(\Omega_{RUL}) &= \frac{1}{K}\prod_{r=1}^{R}\left(1 - \gamma.s^r_{j^*}\right) \end{aligned} \tag{9}$$

where

$$K = \prod_{r=1}^{R}\left(1 - \gamma.s^r_{j^*}\right) + \gamma.\sum_{r=1}^{R} s^r_{j^*}\prod_{r' \neq r}\left(1 - \gamma.s^{r'}_{j^*}\right) \tag{10}$$

Given the BBA in eq. (9), we can finally calculate the belief associated to any interval $[RUL^{inf}, RUL^{sup}]$ as the sum of the belief masses associated to all subsets included in $[RUL^{inf}, RUL^{sup}]$: this represents the amount of belief that directly supports the hypothesis $RUL^{test} \in [RUL^{inf}, RUL^{sup}]$ where RUL^{test} is the true RUL of the test equipment, and it has been interpreted as a lower bound for the probability that $RUL^{test} \in [RUL^{inf}, RUL^{sup}]$, or, analogously, as an upper bound for the probability that $RUL^{test} \notin [RUL^{inf}, RUL^{sup}]$.

In conclusion, a left-bounded interval $\Delta^+(\alpha) = [RUL^{inf}(\alpha), +\infty]$, such that a belief $1-\alpha$ is assigned to it, provides the following information about the probability distribution of the true equipment RUL: $P\left(RUL^{test} > RUL^{inf}(\alpha)\right) > 1-\alpha$ or, equivalently, $P\left(RUL^{test} < RUL^{inf}(\alpha)\right) < \alpha$. The advantage of this latter interpretation of $RUL^{inf}(\alpha)$ is that it can be used to plan the maintenance action: performing maintenance before $RUL^{inf}(\alpha)$ guarantees a probability of failure lower than α.

The predictive interval $\Delta^+(\alpha)$ depends in large measure on the value assigned to parameter γ by the analyst, based on her/his opinion about the relevance of the information derived from historical trajectories when making predictions about a new one. As it may be difficult for the analyst to express a reliable opinion about γ, we suggest to set its value considering the coverage of the resulting prediction intervals $\Delta^+(\alpha)$., i.e., the probability that given a trajectory with RUL equal to RUL^{true} and the corresponding credible interval $\Delta^+(\alpha)$, the condition $RUL^{true} \in \Delta^+(\alpha)$ is verified. Indeed, a desirable property for $\Delta^+(\alpha)$ is that its coverage, which can be estimated using training data, is greater than $1-\alpha$. This procedure will be discussed in more detail in the next Section.

3 Numerical application

In this Section, we verify the proposed method for the uncertainty quantification of a similarity-based prognostic approach and compare its effectiveness with that of two alternative methods, i.e., the Kernel Density Estimation (KDE) and Mean-Variance Estimation (MVE) methods, on simulated and real data. In Section 3.1, the similarity-based method is applied to simulated data concerning the evolution of creep damage in ferritic steel. The influence on the prognostic performance of parameters λ of Eq. (2) and γ of Eq. (8) is also investigated, and a procedure for setting their values is proposed. On the basis of these results, in Section 3.2, the method is applied to real data taken from a case study about the clogging of filters in a BWR condenser.

3.1 Simulated data: creep growth in ferritic steel

Ferritic steels are widely used for welded steam pipes in the construction of power plant components that operate under high temperature and stress conditions. In such conditions, the creep deformation and rupture are important factors in determining the equipment lifetimes.

3.1.1 Creep growth models

We have simulated the evolution of the creep strain ε in ferritic steel exposed to a load σ, by using the uni-axial form of the non-linear creep constitutive equations proposed within the framework of Continuum Damage Mechanics by Mustata & Hayhurst (2005):

$$\begin{cases} \dot{\varepsilon} = A\sinh\left[\dfrac{B\sigma(1-H)}{(1-\phi)(1-\omega)}\right] \\ \dot{H} = \dfrac{h\dot{\varepsilon}}{\sigma}\left(1-\dfrac{H}{H^*}\right) \\ \dot{\phi} = \dfrac{K_c}{3}(1-\phi)^4 \\ \dot{\omega} = C\dot{\varepsilon} \end{cases} \tag{11}$$

where ε is the creep strain, i.e., the percentage of elongation of the turbine blade in the longitudinal direction with respect to its original length, ϕ and ω are two damage state variables describing, respectively, the coarsening of the carbide precipitates, and the inter-granular creep constrained cavitation damage, H is the hardening state variable used to represent the strain hardening effect attributed to primary creep, and A, B, H^*, h, K_c and C are material inherent characteristics. Each characteristic $\varphi_m = A, B, H^*, h, K_c, C$ varies with the temperature according to the Arrhenius law, i.e., $\varphi_m = \varphi_{m0}\exp(-Q_m/T)$, $m = 1{:}\,6$, where T is the operating temperature and φ_{m0} and Q_m are parameters whose values have to be determined experimentally.

To generate different trajectories, the intrinsic variability of the degradation process is simulated by sampling the values of the load σ and temperature T to which the steel is exposed at each time step from a normal distribution centered on their mean values, whereas the variability of the degradation process of similar pieces of equipment is simulated by sampling the values of parameters φ_{m0} and Q_m, $m = 1{:}\,6$, at the beginning of each new simulated degradation trajectory. Finally, in order to generate the sequence of observations $\boldsymbol{z}_{1:n^r}^r = \{\varepsilon(\tau_i) + v_i\}_{i=1:n^r}^r$, with n^r being the time of the last observation before failure of the r-th degradation trajectory, a white Gaussian noise v_i is added to the simulated creep strain $\varepsilon(\tau_i)$ at the observation time τ_i. We assume failure to happen when the limiting creep strain value of 2% is reached. Figure 1 shows an example of simulated creep growth trajectory (upper) and the corresponding sequence of observations $\boldsymbol{z}_{1:n^r}^r$ (bottom).

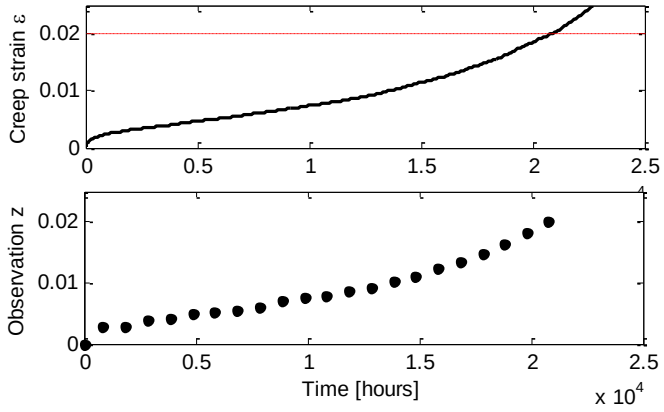

Figure 1: example of simulated creep growth trajectory (upper) with the corresponding sequence of observations (bottom).

3.1.2 Results

All the degradation trajectories used in this Section have been simulated by iteratively applying the simulation model of eq. (11). Using the simulated trajectories, we have developed $N^{trn} = 50$ different prognostic models,

each one built using a different training set $\{\boldsymbol{z}_{1:n^r}^r, r = 1{:}\,R\}$, made by $R = 7$ different training trajectories. Each model is, then, verified with respect to $N^{tst} = 50$ different test trajectories, $\boldsymbol{z}_{1:l}^q, q = 1{:}\,N^{tst}$.

Let us define $\widehat{RUL}_l$ and $RUL_l^{inf}(\alpha)$, $l = 1{:}\,N^{trn}$, as the predictions of the RUL and of its left bound with belief $1-\alpha$, obtained by the model developed using the trajectories in the l-th training set, for the q–th test trajectory, $q = 1{:}\,N^{tst}$. The three following performance indicators, obtained by simple average of the N^{tst} performances of each l-th model on all the test trajectories, will be considered for quantifying the BTF similarity-based method capabilities for RUL estimation:

- The Mean Square Error (MSE), i.e., the mean value of the square error $\left(\widehat{RUL}_l - RUL^{true}\right)^2$ made in predicting the true RUL, RUL^{true} of the q-th test equipment, $q = 1{:}\,N^{tst}$. The MSE measures the accuracy of the prediction $\widehat{RUL}^l$ and is desired to be as small as possible.
- The Coverage (Cov_α) of the prediction interval $\Delta_l^+(\alpha) = [RUL_l^{inf}(\alpha), +\infty]$, i.e., the percentage of times the condition $RUL^{true} > RUL_l^{inf}(\alpha)$ is verified, where $1-\alpha$ is the belief associated by the RUL BBA to the interval $\Delta_l^+(\alpha)$. This indicator measures the reliability of the interval and we expect to obtain values of Cov_α larger than $1-\alpha$, since the belief $1-\alpha$ associated to the interval is a lower bound for the probability that the test equipment true RUL is in the interval, i.e., is greater than $RUL_l^{inf}(\alpha)$.
- The mean amplitude (MA_α) of the interval $[RUL_l^{inf}(\alpha), \widehat{RUL}_l]$, which gives a measure of the precision of the RUL prediction. In order to have a high precision, we wish to keep the value of MA_α as small as possible.

In Figure 2, the variation of the square root of the MSE indicator with parameter λ is shown for the three life values of $\beta = 25\%, 50\%,$ and 75% of the equipment life fraction $\beta = \tau_I/\tau_F$. Notice that the RUL predictions, $\widehat{RUL}$, are obtained using the similarity-based weighted average in eq. (4), whereas the prediction intervals are estimated using the target belief $1-\alpha = 0.8$. As expected, the prediction error decreases as the life fraction β increases, i.e., as we get closer to failure. Results in Figure 2 show that the maximum accuracy of $\widehat{RUL}$ is obtained for values of the parameter λ around 5×10^{-5}.

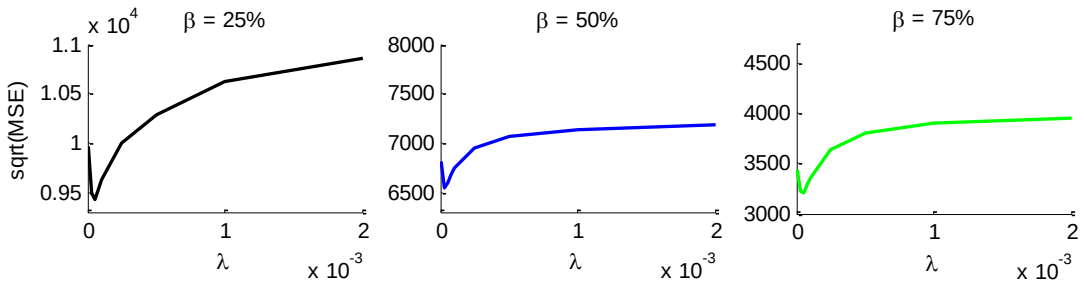

Figure 2: square root of the *MSE* of the prediction $\widehat{RUL}$ as a function of parameter λ.

The precision of the prediction, which is evaluated by the indicator MA_α, is also an important aspect to be considered in the optimization procedure. However, the choice of parameters λ and γ should be subordinate to the verification that the coverage Cov_α is actually larger than $1-\alpha$. Lower values of the coverage would indicate that a too large belief mass has been assigned to the predictions RUL^r provided by the reference trajectories most similar to the test trajectory, so that the belief $1-\alpha$ assigned to the prediction interval is not justified by the experimental evidence.

Figure 3 shows the coverage $Cov_{0.2}$ of the left bounded prediction interval $\Delta^+(0.2)$ (upper), the square root of the MSE made by the prediction $\widehat{RUL}$ (middle) and the mean amplitude $MA_{0.2}$ of the interval $[RUL^{inf}(0.2), \widehat{RUL}]$ (bottom) in correspondence of three different values of parameter λ as a function of the parameter γ.

For the value of $\lambda = 5\text{x}10^{-5}$ that maximizes the accuracy of the prediction $\widehat{RUL}$, the coverage is always larger than the minimum accepted value of $1-\alpha = 0.8$. However, for such a small value of λ the precision, represented by the indicator $MA_{0.2}$, is much lower than for $\lambda = 5\text{x}10^{-4}$ and $\lambda = 5\text{x}10^{-3}$. This is due to the fact that if λ is small, the similarity of a reference trajectory tends to be small, except in the rare case of a trajectory very similar to the test trajectory. As a consequence, for very small values of λ, it is often hard to support with sufficient evidence the hypothesis that the RUL value belongs to any subset of the RUL domain Ω_{RUL}.

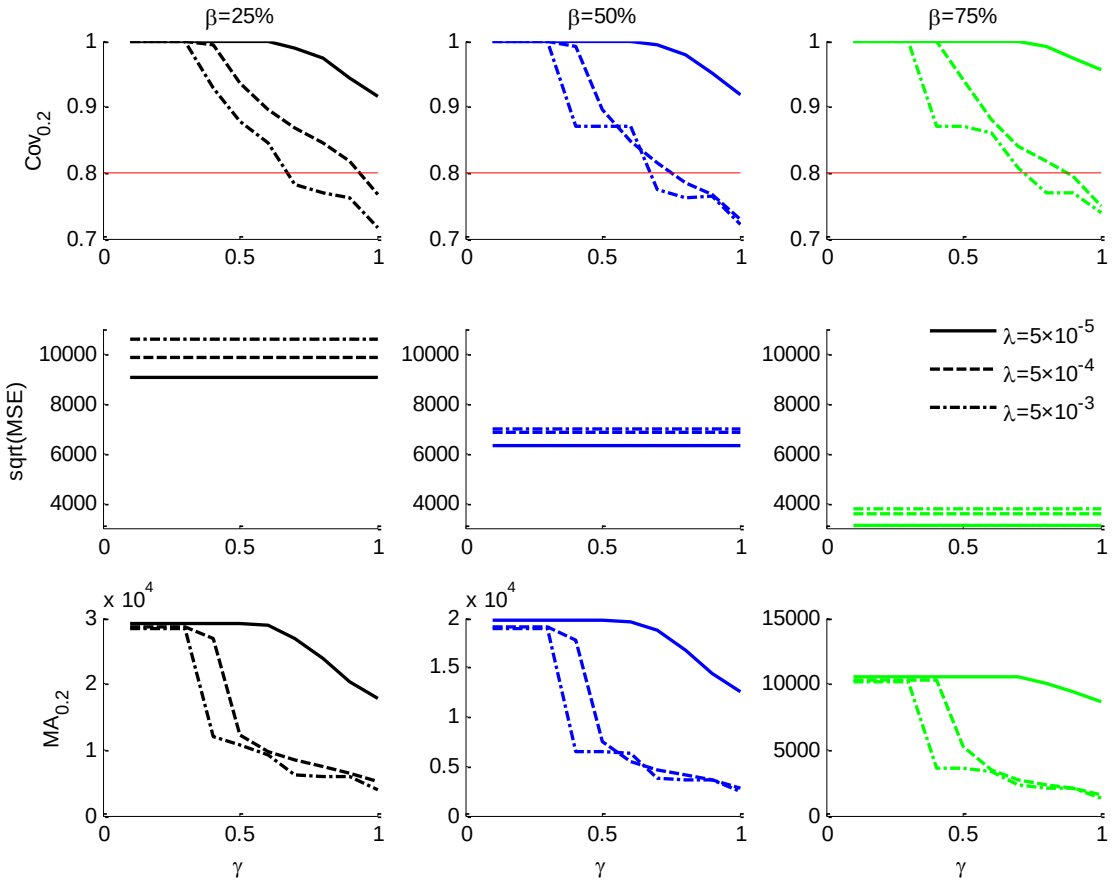

Figure 3: value of the three performance indicators as a function of γ at three fractions β of the trajectory life durations and for three values of λ.

These observations have motivated the adoption of the following procedure for setting the parameters γ and λ:

1. We identify some possible values of λ (e.g., in this case study, $\lambda_1 = 0.5.10^{-5}, \lambda_2 = 0.5.10^{-4}, \lambda_2 = 0.5.10^{-3}$).
2. For each value of λ in 1., we derive a condition for parameter γ by imposing a coverage, $Cov_{0.8}$ greater than 0.8 (e.g., $\gamma \leq 1$ if $\lambda = 5x10^{-5}$, $\gamma \leq 0.7$ if $\lambda = 5x10^{-4}$ and $\gamma \leq 0.6$ if $\lambda = 5x10^{-3}$).
   Since the precision tends to monotonically increase (amplitude of MA tends to decrease) as γ increases, we choose, for each value of λ, the maximum γ value which satisfies the condition in 2. ($\gamma = 1.0$ if $\lambda = 5x10^{-5}$, $\gamma = 0.7$ if $\lambda = 5x10^{-4}$, and $\gamma = 0.6$ if $\lambda = 5x10^{-3}$).
3. Among the identified pairs of values of λ and γ in 2., we choose the pair with the most satisfactory trade-off between prediction accuracy and precision.

With respect to the last step of the procedure, Table 1 reports the performance of the three identified pairs of values of λ and γ in terms of prediction accuracy (square root of the MSE indicator), coverage ($Cov_{0.2}$), and precision ($MA_{0.2}$). Based on these results, we set the parameters to the values $\lambda = 5x10^{-4}$ and $\gamma = \gamma_{max} = 0.7$, since performances are better for this value of λ than for $\lambda = 5x10^{-3}$, both in terms of accuracy and

precision, whereas, and with respect to $\lambda = 5x10^{-5}$, a large improvement of the precision is obtained at the expenses of a small reduction in the accuracy.

Table 1: performance indicators for three different values of λ in correspondence of γ_{max}.

λ		5×10^{-5}	5×10^{-4}	5×10^{-3}
γ_{max}		1.0	0.7	0.6
$Cov_{0.2}$	$\beta = 25\%$	0.915	0.868	0.846
	$\beta = 50\%$	0.918	0.814	0.870
	$\beta = 75\%$	0.955	0.840	0.861
$\sqrt{MSE}$ (10^3)	$\beta = 25\%$	9.065	9.851	10.552
	$\beta = 50\%$	6.272	6.807	6.976
	$\beta = 75\%$	3.089	3.589	3.793
$MA_{0.2}$ (10^3)	$\beta = 25\%$	17.871	6.217	9.202
	$\beta = 50\%$	12.571	4.568	6.233
	$\beta = 75\%$	8.646	2.685	3.304

Figure 4 shows the predictions obtained at all measurement time instants τ_I for 4 new test trajectories different from those used for parameter settings. More results and a more detailed discussions can be found in Appendix A, where it is shown that the large oscillations of the confidence bound that are observed in Figure 4 can be attenuated by increasing the value of λ or reducing the value of γ, at the price of a lower accuracy and precision. Also, their amplitude becomes smaller when the density of reference trajectories available is larger (for instance because a larger number of degradation trajectories have been observed or because their variability is smaller).

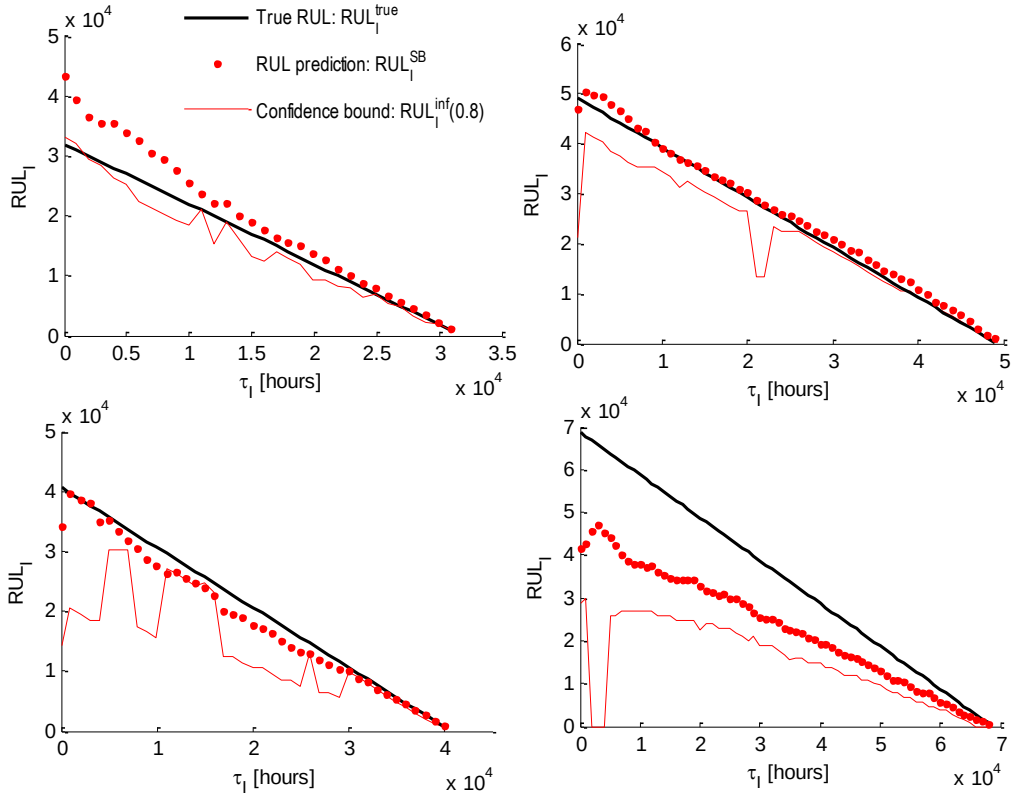

Figure 4: predictions obtained for 4 different test trajectories using $\lambda = 5\times10^{-4}$ and $\gamma = 0.7$.

3.1.3 Comparison with other uncertainty estimation methods

In this Section, we apply the Kernel Density Estimation (KDE) (Botev et al. 2010) and the Mean-Variance Estimation (MVE) (Nix & Weigend, 1994) methods for the quantification of the RUL prediction uncertainty in the numerical case study. The obtained results are compared with those obtained by the proposed method considering the MSE, Cov and MA performance indicators.

KDE is a non-parametric method used for estimating the Probability Density Function (PDF) of a random variable (Botev et al. 2010). The basic idea is to assign a kernel function to each observation in a data set, and then, to sum all kernels to obtain the PDF (Botev et al. 2010). In this work, the KDE is employed for estimating the PDF of the RUL prediction provided by the SB model at each time instant τ_i. The reader interested in more details about the KDE method can refer to (Botev et al. 2010).

MVE has been originally proposed in Nix & Weigend, (1994) for constructing prediction intervals of an uncertain variable using a feedforward ANN properly developed for this purpose. In this work, the MVE is employed for constructing the $1-\alpha = 0.8$ prediction intervals of the RUL predictions provided by an ensemble of H bootstrapped ANNs models (Carney et al., 1999; Polikar, 2006). The reader interested in more details about the MVE method can refer to (Nix & Weigend, 1994). In this application, an ensemble of $H = 5$ ANNs models has been built considering a training set formed by $N^{trn} = 50$ training trajectories. Each ANN is characterized by an architecture with three layers (input, hidden and output) and 10 hidden neurons. Different ANN configurations characterized by M inputs taken from a time window of M consecutive measurements have been considered. The optimum value of $M = 1$ has been identified by trials and errors considering the

MSE, Cov and MA performance indicators on a validation set. The ensemble output (i.e., the predicted RUL) is obtained by averaging the outputs of the $H = 5$ ANNs. A further ANN with 3 layers and 60 hidden neurons has been built to estimate the RUL prediction uncertainty.

Table 2 reports the average values of the performance indicators over the $N^{tst} = 50$ test trajectories obtained by the proposed, the KDE and the MVE methods.

Table 2: Average value of the performance indicators over $N^{tst} = 50$ test trajectories obtained by applying the proposed and the two alternative methods.

	Proposed method (SB-BFT)	Alternative method (SB-KDE)	Alternative method (ANN-MVE)
$\sqrt{MSE}$ (10^3)	3.597	3.597	3.359
$Cov_{0.2}$	0.936	0.962	0.857
$MA_{0.2}$ (10^3)	2.594	6.527	3.882

The results show that the proposed method provides more precise RUL predictions (i.e., lower $MA_{0.2}$ which corresponds to narrower prediction intervals) and more reliable prediction intervals (i.e., larger $Cov_{0.2}$) satisfying the desired coverage level of 0.8, whereas the KDE and the MVE, even though they assure the desired coverage level of 0.8, they provide less precise RUL predictions (i.e., larger $MA_{0.2}$ values which correspond to larger prediction intervals). One can also recognize that the proposed and the KDE methods, based on the use of the SB model for the RUL point estimator, provide slightly less accurate RUL predictions. This is due to the fact that the ensemble approach used in this case in combination with the MVE method to provide the RUL point estimator is more robust and accurate than the individual SB model used by the proposed and the SB-KDE methods.

Figure 5 shows the estimates of the RUL and the associated lower confidence bounds provided by the three methods for one test trajectory. One can easily recognize that the proposed method provides narrower confidence bound (lower $MA_{0.2}$) than those of the SB-KDE and the ANN-MVE methods.

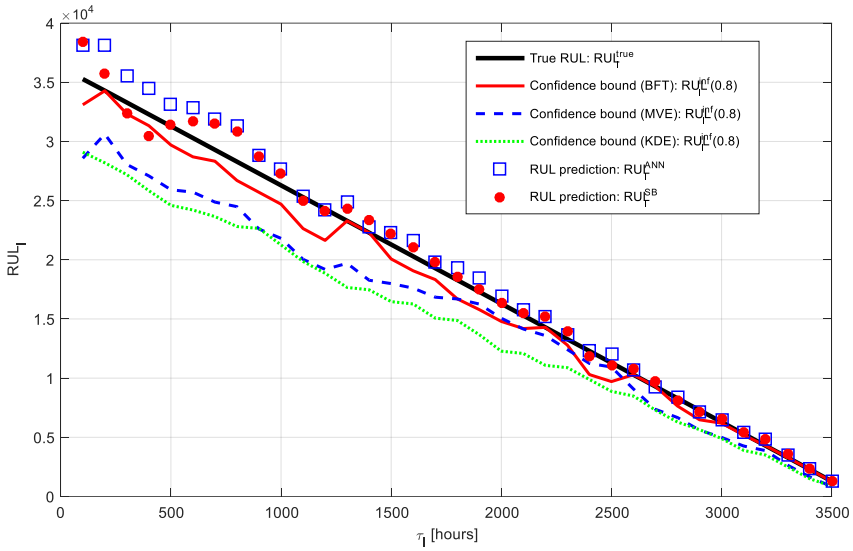

Figure 5: Comparison of the obtained *RUL* predictions and the corresponding confidence bounds for one test trajectory by the proposed method (dots and continuous line, respectively), SB-KDE (dots and dot line, respectively) and ANN-MVE (squares and dash line, respectively).”

This analysis shows the capability of the proposed method of properly quantifying the uncertainty affecting the RUL predictions with narrower confidence bounds (lower $MA_{0.2}$) and larger coverage values (larger $Cov_{0.2}$) compared to the two alternative methods. Notice, however, that the proper setting of the parameters of the proposed method for balancing accuracy and precision of the predictions might be time-consuming, as discussed in Section 3.1.2. The capability of the proposed method for uncertainty treatment in case of few and/or irregular degradation trajectories is verified by its application on the real data regarding the clogging of BWR condenser filters of Section 3.2.

3.2 Real data: clogging of BWR condenser filters

In this Section, we consider the heat exchanger filters used to clean the sea water entering the condenser of the BWR reactor of a Swedish nuclear power plant. During operations, filters undergo clogging and, once clogged, can cumulate particles, seaweeds, and mussels from the cooling water in the heat exchanger. For this reason, prompt and effective cleaning of the filters is desirable. Predictive maintenance can help achieving this result, keeping maintenance costs reasonably low.

From data collected on field, we have available sequences of observations $\boldsymbol{z}_{1:n^q}^q$, $q = 1{:}\,8$ taken during the clogging process of $Q = 8$ historical filters. Each observation $\boldsymbol{z}_i^q = [\Delta P_i^q, \dot{M}_i^q, T_i^q]$ contains the measurements of the pressure drop ΔP_i^q, the flow across the filter $\dot{M}_i^q$ and the sea water temperature T_i^q collected at time τ_i during the clogging process of the q-th filter.

For clarification purposes, Figure 6 shows the sequences of observations $\boldsymbol{z}_i^q$ collected during the clogging process of filters, $q = 1, 2$ and 4 from the beginning of their life ($\tau_i = 0$) to the failure ($\tau_i = \tau_F^q$). It is worth noticing that:

- the typical behavior of filter clogging characterized by an increase of the pressure drop ΔP_i^q (Figure 6 (top)) and a decrease of the flow rate across the filter $\dot{M}_i^q$ (Figure 6 (middle)) is clearly observable,
- the larger the sea water temperature, the faster the clogging process,
- the large variability of the filter lifetimes due to the variability of the sea water temperature.

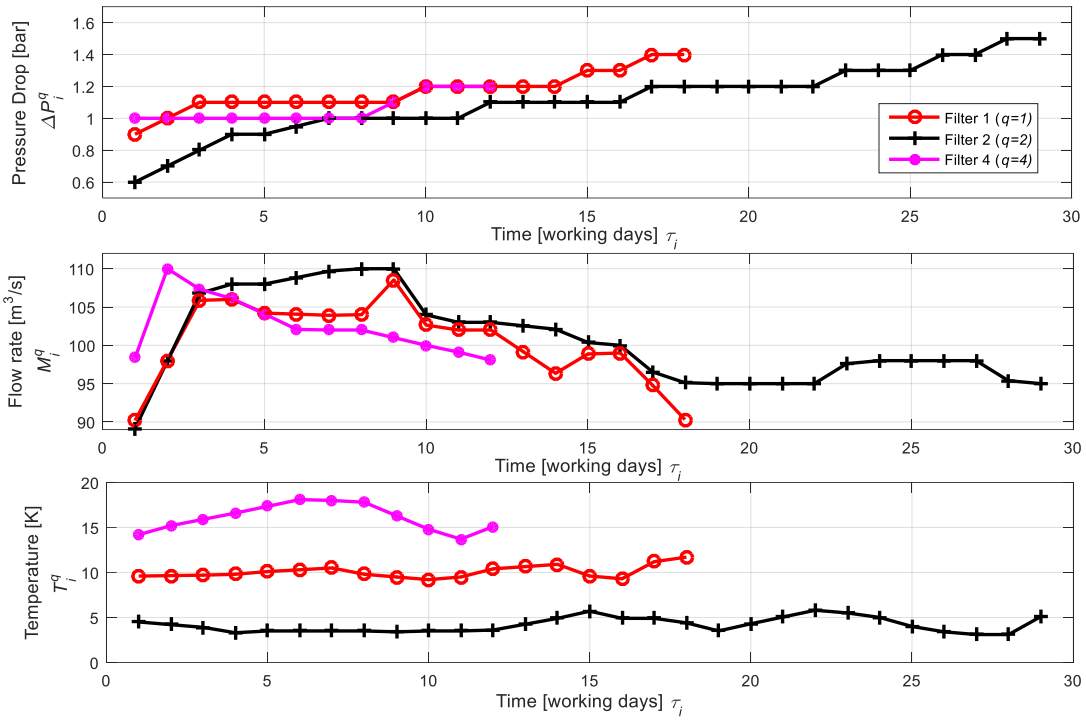

Figure 6: The sequences of observations of the pressure drop (ΔP_i^q) (top), the flow rate ($\dot{M}_i^q$) (middle) and the temperature (T_i^q) (bottom) collected during the life of three filters.

To further investigate the large variability in the clogging process of the Q=8 filters, we consider the degradation indicator I_i^q which quantifies the amount of clogging of filter q at time τ_i and is defined by (Nystad, 2009):

$$I_i^q = \frac{\Delta P_i^q}{\left(\dot{M}_i^q\right)^2} \tag{12}$$

Figure 7 shows the evolution of the degradation indicator I_i^q during the lives of the Q=8 filters. It can be observed that the clogging process is, indeed, affected by large uncertainties, which, according to the analysis of Figure 6 and the opinions of plant experts, is caused by the variability of the sea water conditions such as temperature and other factors influencing the life cycle of mussels, algae and other sea organisms; in this context, the challenge is to provide sufficiently narrow confidence intervals for the value of the predicted filters RUL.

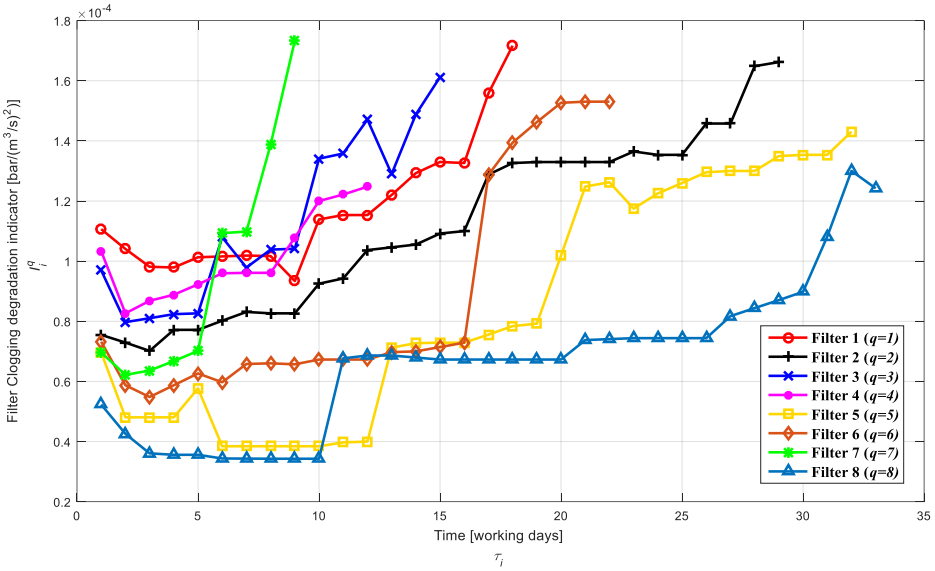

Figure 7: Evolution of the filter clogging degradation indicator during the lives of $Q = 8$ filters.

The prognostic method proposed is applied to each trajectory q at the three life fractions β as in Section 3.1.2, using the remaining $R = 7$ trajectories as reference trajectories in a leave-one-out scheme. Figure 8 shows how the three performance indicators MSE, Cov_α and MA_α of Section 3.1.2 vary with parameter γ for three values of parameter λ (0.1, 0.05 and 0.01). These results confirm those obtained for the simulated creep growth data of Section 3.2: the MSE has a minimum around $\lambda = 0.05$ and the value of the $MA_{0.2}$ indicator decreases with both λ and γ. Notice also that, for the values of λ considered in Figure 8, almost all possible values of γ are acceptable since the coverage $Cov_{0.2}$ is always larger than 0.8, except for $\lambda = 0.1$ and $\gamma = 1$. The precision obtained for $\lambda = 0.05$ when $\gamma = 1$ is very close to that obtained for $\lambda = 0.1$, whereas the error is lower. Then, this optimization leads us to set $\lambda = 0.05$ and $\gamma = 0.95$ for generating the prognostic predictions with uncertainty in correspondence of each observation available. We do not set $\gamma = 1$ to account for the fact that the information provided by a piece of equipment about another one is uncertain even when their degradation paths up to the present time are identical, that is when $s_{j^*}^r = 1$. Appendix B discusses the counterintuitive results obtained by setting $\gamma = 1$.

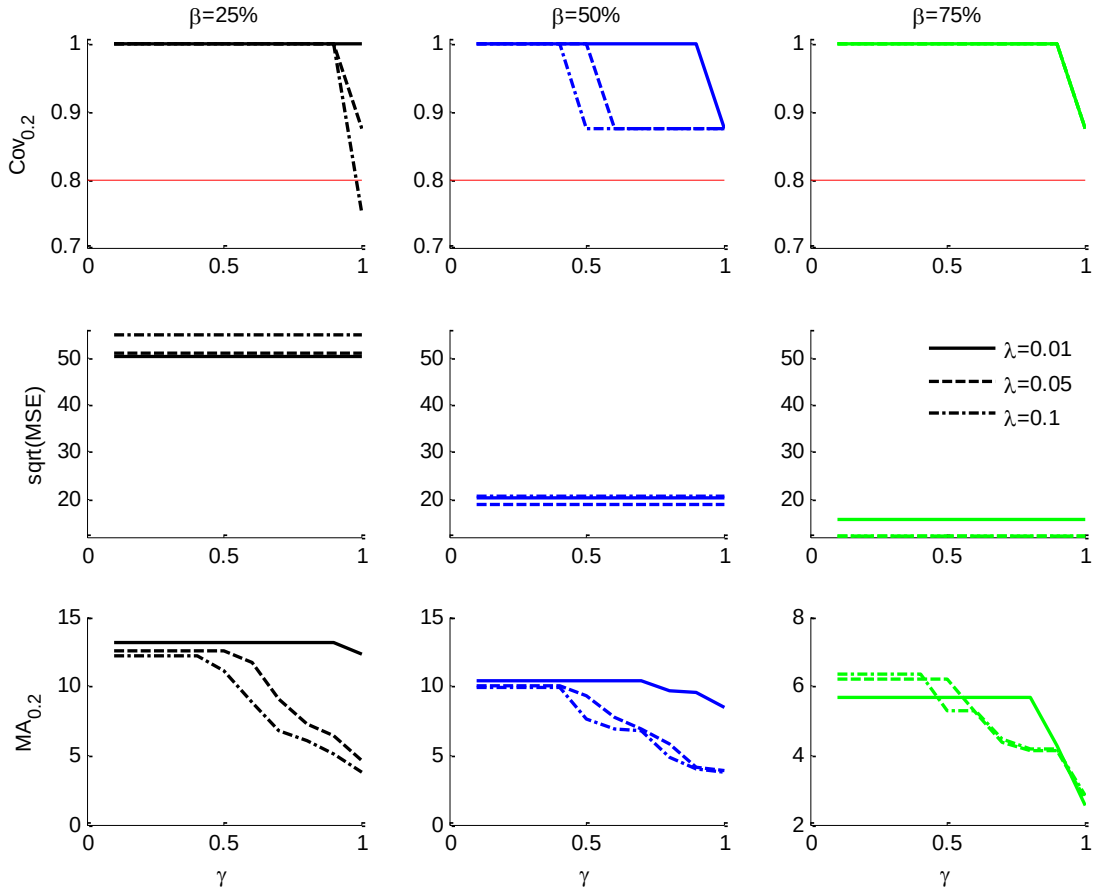

Figure 8: value of the three performance indicators as a function of γ at three fractions β of the trajectory life durations and for three values of λ.

Results obtained with $\lambda = 0.05$ and $\gamma = 0.95$ are shown in Figure 9 for all $Q = 8$ test trajectories available. In trajectory 4, the confidence bound is for most of the time equal to zero. This means that its similarity with all reference trajectories is rather low and, thus, the prediction is very uncertain. Also in many other cases, the prediction accuracy is rather low and the prediction interval large. However, due to the small number of training trajectories available and the large uncertainties affecting the clogging process, we can be satisfied with this result.

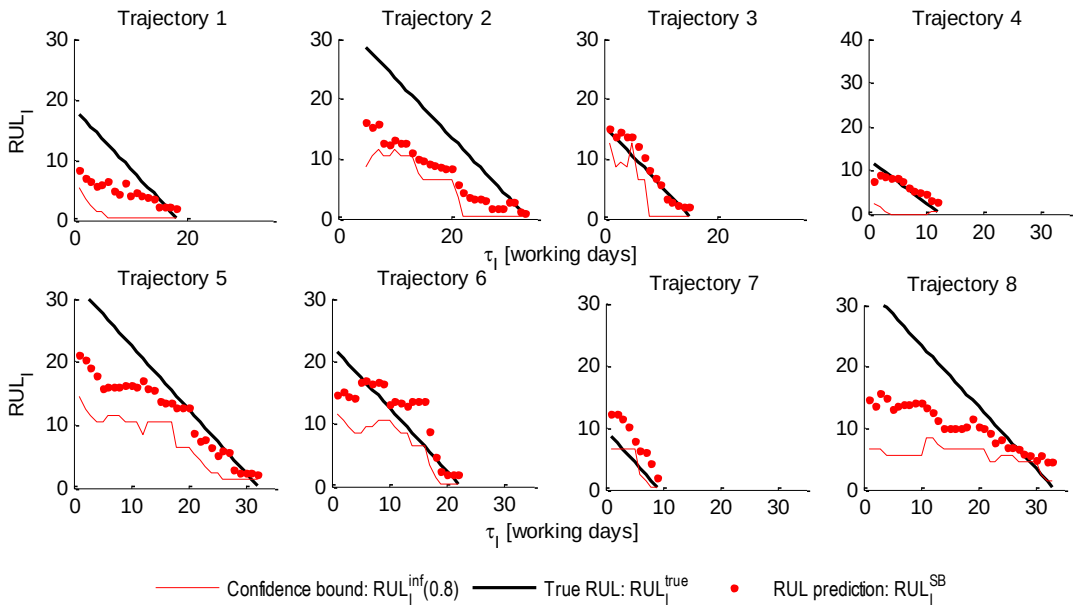

Figure 9: predictions obtained for the $Q = 8$ filter clogging trajectories available, using parameters $\lambda = 0.05$ and $\gamma = 0.95$.

The obtained results are compared with those obtained by the SB-KDE and the ANN-MVE methods. The ANN-MVE method is applied to each q-th trajectory, using the remaining $R = 7$ trajectories as training and validation trajectories (with a time window length of $M = 5$ consecutive measurements optimized by trials and errors) in a leave-one-out scheme.

Table 3 reports the average values of the performance indicators over the $Q = 8$ test trajectories.

Table 3: Average value of the performance indicators for the $Q = 8$ test trajectories obtained by the proposed method and the two alternative methods.

	Proposed method (SB-BFT)	Alternative method (SB-KDE)	Alternative method (ANN-MVE)
$\sqrt{MSE}$	3.743	3.743	5.710
$Cov_{0.2}$	0.986	0.815	0.734
$MA_{0.2}$	3.742	4.167	3.944

The results show that the proposed and the SB-KDE methods allow obtaining the desired coverage level of 0.8, although the proposed method provides narrower prediction intervals (smaller $MA_{0.2}$) than those provided by the SB-KDE method. With respect to the RUL accuracy, the SB is more accurate than the ensemble of ANN used by the ANN-MVE method. This confirms the ability of the proposed method to deal with few and irregular degradation trajectories and provide more accurate RUL predictions compared to an ensemble method, which typically requires more training data.

This analysis proves the effectiveness of the proposed method, when few training data are available, in *i)* accurately predicting the RUL of the filters and *ii)* properly quantifying the uncertainty affecting the RUL predictions with narrower confidence bounds (lower $MA_{0.2}$) and larger coverage values (larger $Cov_{0.2}$). As

already mentioned, the limitation of the proposed method lies in the difficulty of properly setting its parameters while balancing the accuracy and the precision of the predictions.

4 Conclusions

In this work, we have considered the problem of directly predicting the RUL of a degrading equipment and providing a measure of confidence on the prediction, based on a set of reference degradation trajectories experienced by similar equipment failed in the past. To this aim, a similarity-based approach is integrated within the framework of belief function theory.

Two key elements in the application of the method are the parameter λ, which defines how strong is the desired interpretation of similarity, and the parameter γ, which defines the degree of trust given to the reference trajectories. Using artificial data simulated by a non-linear model for creep growth in ferritic steel, we have analyzed how the values of these two parameters influence the performance of the method and given some indications on how to set their values.

Finally, we have applied the method to the problem of predicting the RUL of clogging filters used in nuclear power plants, obtaining prediction intervals for the values of the RUL with satisfactory accuracy, considering the large uncertainties affecting the clogging process.

Furthermore, the Kernel Density Estimation (KDE) and the Mean-Variance Estimation (MVE) methods have been applied to the same case studies to quantify the uncertainty affecting the RUL predictions. The comparison of the obtained results confirms the superiority of the proposed method with respect to the two alternative methods in terms of reliability (i.e., Cov) and precision (i.e., MA). More specifically, the proposed method has been proved to be effective also when few training data are available thanks to the capability of the BFT of properly representing and treating the uncertainty when scarce information is available.

We expect that the use of a method able to associate to a RUL point estimation also a reliable and narrower prediction interval can help the building the maintenance decision maker confidence towards prognostics and allow adopting predictive maintenance approaches in real industrial applications. To this purpose, it would be important to quantify the benefits that can be obtained in terms of metrics such as Return on Investment (ROI) or Total Lifecycle Costs. Future work will devoted to this aim.

A limitation of the proposed method is the presence of possibly large oscillations in the confidence bounds, which may be confusing for the maintenance planner. It has been shown that such oscillations can be reduced by conveniently setting the parameter values; this, however might be time-consuming and could also reduce the accuracy and precision of the prediction. Notice also, that the amplitude of the oscillation decreases as the density of the reference trajectories increases.

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 to this research is partially supported by the China NSFC under grant number 71231001.

The authors would like to thank all the reviewers for their valuable comments to improve the quality of this paper.

References

Al-Dahidi, S., Baraldi, P., Di Maio, F., Zio, E. (2014). A novel fault detection system taking into account uncertainties in the reconstructed signals. Annals of Nuclear Energy, 73, 131–144.

Altinçay, H, (2007). Ensembling evidential k-nearest neighbor classifiers through multi-modal perturbation. Applied Soft Computing Journal, 7 (3), 1072-1083.

Angstenberger, L. (2001). Dynamic Fuzzy Pattern Recognition, International Series in Intelligent Technologies, 17, Kluwer Academic Publishers.

Baraldi, P., Mangili, F., Zio, E. (2012)a. Ensemble of Bootstrapped Models for the prediction of the Remaining Useful Life of a Creeping Turbine Blade. IEEE Int. PHM Conf., Denver, Colorado, 18-21 June 2012.

Baraldi, P., Mangili, F., Zio, E. (2012)b. A Kalman Filter-based Ensemble Approach for Turbine Creep Prognostics. IEEE Transactions on Reliability, 61 (4), 966 – 977.

Baraldi, P., Cadini, F., Mangili, F., Zio, E. (2013)a. Model-Based and Data-Driven Prognostics under Different Available Information. Probabilistic Engineering Mechanics, 32, 66-79.

Baraldi, P., Cadini, F., Mangili, F., Zio, E. (2013)b. Prognostics under different available information. 2013 Prognostics and System Health Management Conference (PHM), Milan, Italy, 8-11 September 2013.

Baraldi, P., Mangili, F., Zio, E. (2013)c. Investigation of uncertainty treatment capability of model-based and data-driven prognostic methods using simulated data. Reliability Engineering and System Safety, 112, pp. 94-108.

Baraldi, P., Mangili, F., Zio, E. (2015). A belief function theory based approach to combining different representation of uncertainty in prognostics. Information Sciences, 303, 134-149.

Benkedjouh, T., Medjaher, K., Zerhouni, N., Rechak, S. (2013). Health assessment and life prediction of cutting tools based on support vector regression. Journal of Intelligent Manufacturing, article published online.

Botev, Z. I., Grotowski, J. F., & Kroese, D. P. (2010). Kernel density estimation via diffusion. The Annals of Statistics, 38(5), 2916-2957.

Carney, J., Cunningham, P., Bhagwan, U. (1999). Confidence and prediction intervals for neural network ensembles. International Joint Conference on Neural Networks, IJCNN '99, 10-16 July 1999, Washington, DC, 1215-1218.

Dempster, A.P. (1976). Upper and lower probabilities induced by a multivariate mapping. Annals of Mathematical Statistics, 38, 325-339.

Di Maio F., Tsui K.L., Zio E. (2012). Combining Relevance Vector Machines and Exponential Regression for Bearing RUL estimation. Mechanical Systems and Signal Processing, Mechanical Systems and Signal Processing, 31, 405–427.

Gorjian, N., Ma, L., Mittinty, M., Yarlagadda, P., Sun, Y. (2009). Review on Degradation Models in Reliability Analysis, Proceedings of the 4th World Congress on Engineering Asset Management, Athens, 28-30 Sept 2009.

Guha, D., Chakraborty, D. (2010). A new approach to fuzzy distance measure and similarity measure between two generalized fuzzy numbers. Applied Soft Computing Journal, 10(1), 90-99.

Helton, J. C. (2004). Alternative representations of epistemic uncertainty. Special Issue of Reliability Engineering and System Safety, 85, 1–10

Hines, J.W., Usynin, A. (2008). Current Computational Trends in Equipment Prognostics. Int. J. Comput. Intell. Syst., 1(1), 94–102.

H., Zhao, F., Jiao, L. (2012). Fuzzy spectral clustering with robust spatial information for image segmentation. Applied Soft Computing Journal, 12 (11), pp. 3636-3647.

Lee, J., Ni, J., Djurdjanovic, D., Qiu, H., Liao, H. (2006). Intelligent prognostics tools and e-maintenance. Computers in Industry, 57(6), 476–489.

Mustata, R. Hayhurst, D.R. (2005). Creep constitutive equations for a 0.5Cr 0.5 Mo 0.25V ferritic steel in the temperature range 565 8C–675 8C. International Journal of Pressure Vessels and Piping, 82, 363–372.

Niu, G., Yang, B.S. (2010). Intelligent condition monitoring and prognostics system based on data-fusion strategy. Expert Systems with Applications, . 37(12), 8831-8840.

Nix, D. A., Weigend, A. S. (1994). Estimating the mean and variance of the target probability distribution. in Proc. IEEE Int. Conf. Neural Netw., 1. Orlando, FL, Jun.–Jul. 1994, 55–60.

Nystad, B.H. (2009). Condition-Based Maintenance (CBM) – filter clogging at OKG 1, a case study, HWR-961, OECD Halden Reactor Project.

Peel, L. (2008). Data Driven Prognostics using a Kalman Filter Ensemble of Neural Network Models. International Conference on Prognostics and Health Management, 6-9 Oct 2008, Denver, CO.

Petit-Renaud, S., Denoeux, T. (2004). Nonparametric regression analysis of uncertain and imprecise data using belief functions. International Journal of Approximate Reasoning, 35, 1-28.

Polikar, R. (2006). Ensemble based systems in decision making. *Circuits Syst. Mag. IEEE*, 6(3), 21–45.

Rasmussen, C., Williams, C. (2006). Gaussian processes for machine learning. MIT Press, Cambridge, MA.

Santosh, T.V., Srivastava, A., Sanyasi Rao, V.V.S., Gosh, A. K., Kushwaha, H.S. (2009). Diagnostic System for Identification of Accident Scenarios in Nuclear Power Plants using Artificial Neural Networks. Reliability Engineering and System Safety, 94, 759-762.

Shafer, G. (1976). A mathematical theory of evidence. Princeton University Press. Princeton, NJ.

Smets, P. (1998). The transferable belief model for quantified belief representation. in: D.M. Gabbay, P. Smets (Eds.), Handbook of Defeasible Reasoning and Uncertainty Management Systems, Kluwer Academic Publishers, Dordrecht, vol. 1, 267–301.

Smets, P. (1994). What is Dempster–Shafers model?, in: Advances in the Dempster–Shafer Theory of Evidence, Wiley, pp. 5–34.

Su, Z.-G., Wang, P.-H., Shen, J., Yu, X.-J., Lv, Z.-Z., Lu, L. (2011). Multi-model strategy based evidential soft sensor model for predicting evaluation of variables with uncertainty. Applied Soft Computing Journal, 11 (2), 2595-2610.

Schwabacher, M., Goebel, K. (2007). A Survey of Artificial Intelligence for Prognostics. Association for the Advancement of Artificial Intelligence Fall Symposium, 9-11 Nov 2007, Arlington (VA).

Tang, L., Kacprzynski, G.J., Goebel, K., Vachtsevanos, G. (2009). Methodologies for Uncertainty Management in Prognostics. Proc IEEE Aerosp Conf, 2009 Mar 7-14 2009; Big Sky, MT.

Vachtsevanos, G., Lewis, F.L., Roemer, M., Hess, A., Wu, B. (2006). Intelligent Fault Diagnosis and Prognosis for Engineering Systems. 1st edition, John Wiley & Sons, Hoboken.

Wang, T., Yu, J., Siegel, D., Lee, J. (2008). A Similarity-Based Prognostics Approach for Remaining Useful Life Estimation of Engineered Systems. International Conference on Prognostics and Health Management, 6-9 Oct 2008, Denver (CO).

Wang, T. (2010). Trajectory similarity based prediction for remaining useful life estimation (Doctoral dissertation, University of Cincinnati).

Wang, W., Carr, M., Xu, W., Kobbacy K. (2011). A model for residual life prediction based on Brownian motion with an adaptive drift. Microelectronics Reliability, 51, pp. 285–293

Yager R. (1987). On the Dempster–Shafer framework and new combination rules. Information Sciences, 41,93–137.

Yager R. (2011). On the fusion of imprecise uncertainty measures using belief structures. Information Sciences, 181 (15), 3199–3209.

Yan, J., Koç, M., Lee, J. (2004). A Prognostic Algorithm for Machine performance Assessment and its Application. Production Planning and Control, 15(8), 796-801.

Zhao W., Tao T., ZhuoShu D., Zio, E., (2013). A dynamic particle filter-support vector regression method for reliability prediction. Reliability Engineering & System Safety, 119, 109-116.

Zhang, Q., Tse, P.W.-T., Wan, X., Xu, G., (2015). Remaining useful life estimation for mechanical systems based on similarity of phase space trajectory. Expert Systems with Applications, 42, 2353–2360.

Zio, E., Di Maio, F., (2010)a. A data-driven fuzzy approach for predicting the remaining useful life in dynamic failure scenarios of a nuclear system. Reliability Engineering & System Safety, 95(1), 49‑57.

Zio E., Di Maio F., Stasi M., (2010)b. A Data-driven Approach for Predicting Failure Scenarios in Nuclear Systems. Annals of Nuclear Energy, 37, 482–491.

Zio, E. (2012). Prognostics and Health Management of Industrial Equipment. In: Kadry S, editor. Diagnostics and Prognostics of Engineering Systems: Methods and Techniques, IGI-Global.

Zio E., Di Maio F. (2012). Fatigue Crack Growth Estimation by Relevance Vector Machines. Expert Systems with Applications, 39, (12), 10681–10692.

Appendix A

In this Appendix, we extend the discussion in Section 3.1.2 about the performance of the similarity-based prognostic approach on the simulated data. First, we provide an illustrative example that explains the low precision of predictions obtained using small values of parameter λ. Then, we discuss choices of the parameters values different than the one proposed in in Section 3.1.2, that is $\lambda = 5\text{x}10^{-4}$ and $\gamma = 0.7$.

Figure A1 shows the RUL prediction with the relative prediction interval for a specific trajectory (left) in correspondence of two different values of parameter λ: $\lambda = 5\text{x}10^{-5}$ (upper) and $\lambda = 5\text{x}10^{-4}$ (bottom). Notice that for $\lambda = 5\text{x}10^{-5}$, the lower bound of the prediction interval is equal to 0 for large part of the trajectory (Figure A1, upper, left); this does not mean that the evidence of very early failure is high (as demonstrated by the fact that the predicted RUL is far from 0), but only that the evidence drawn from the reference trajectories is not sufficient to assert with the desired belief $1 - \alpha = 0.8$ that the RUL value is actually larger than 0. In other words, the prediction $RUL^{inf} = 0$ is a statement of *ignorance* about the value of RUL. Contrarily, in the case of $\lambda = 5\text{x}10^{-4}$ (Figure A1, bottom, left) the lower bound of the prediction interval is always higher than 0. Figure A1, right shows the values of the similarity $s_{j^*}^r$ assigned to each reference trajectory $r = 1{:}\,7$ and the BBA m_{RUL} assigned to the corresponding prediction RUL^r and to the RUL domain Ω_{RUL} at time $\tau_{23} = 21811$ hours, which is characterized by a confidence bound equal to 0 using $\lambda = 5\text{x}10^{-5}$. Notice that the similarities $s_{j^*}^r$ obtained using $\lambda = 5\text{x}10^{-5}$ are significantly lower than those obtained using $\lambda = 5\text{x}10^{-4}$, and, consequently, the mass $m_{RUL}(\Omega_{RUL})$ assigned to the RUL domain using $\lambda = 5\text{x}10^{-5}$ is larger than 0.2, so that the total belief assigned to the trajectories predictions RUL^r does not reach the required value of 0.8.

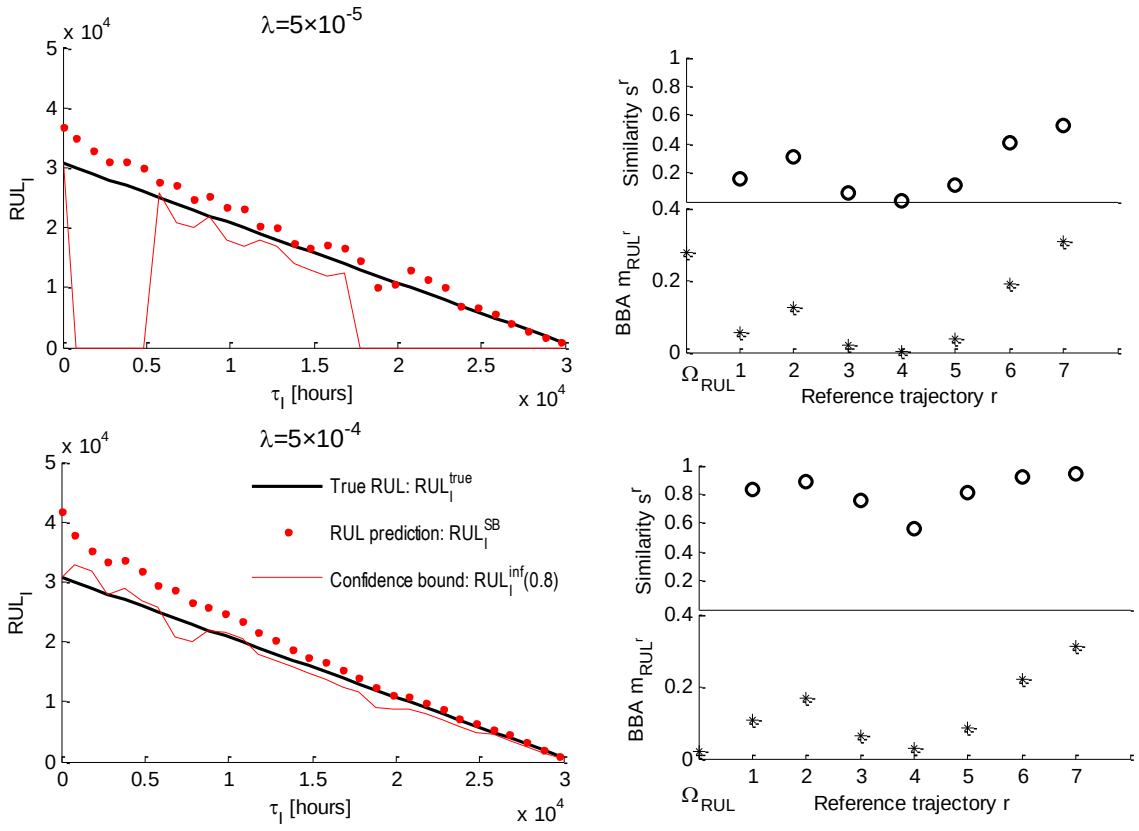

Figure A1: comparison of the RUL prediction with confidence bound (left) and the similarity values and BBAs assigned to the different trajectories at τ_{23}=21811 hours (right) for two values of λ: $\lambda = 5\text{x}10^{-5}$ (upper) and $\lambda = 5\text{x}10^{-4}$ (bottom).

In Figure 4 in Section 3.1.2, the predictions obtained with $\lambda = 5\text{x}10^{-4}$ and $\gamma = 0.7$, are shown and two phenomena can be observed: first, some situations of *ignorance* about the value of RUL where $RUL^{inf} = 0$, are still encountered. This is due to the fact that the information provided by the reference trajectories is not relevant for a specific test trajectory, e.g., because they are too dissimilar. Another noticeable phenomenon in Figure 4 is the presence of large jumps of the confidence bound RUL^{inf}. These jumps occur when the reference trajectory corresponding to the minimum RUL prediction RUL^r included in the prediction interval in order to attain the desired belief $1 - \alpha = 0.8$ changes.

Although justified by the method, the oscillations of the confidence bound may be confusing for the maintenance planner. A reduction in the oscillations can be obtained by increasing the value of λ or reducing the value of γ, at the price of a lower accuracy and precision.

Figure A2 shows the RUL predictions obtained for the same four trajectories of Figure 4 using the parameters values $\lambda = 5\text{x}10^{-3}$ and $\gamma = 0.5$. Table A1 compares the performance of the prediction computed on $N^{tst} = 50$ test trajectories different from those used for optimizing the parameters, in this case and in the case of Figure 4 where $\lambda = 5\text{x}10^{-4}$ and $\gamma = 0.7$. In the Table, the mean value of the RUL, $\overline{RUL}$, for different values of the life fraction β is also shown, and the performance indicators $\sqrt{MSE}$ and $MA_{0.2}$ are expressed also as a percentage of $\overline{RUL}$.

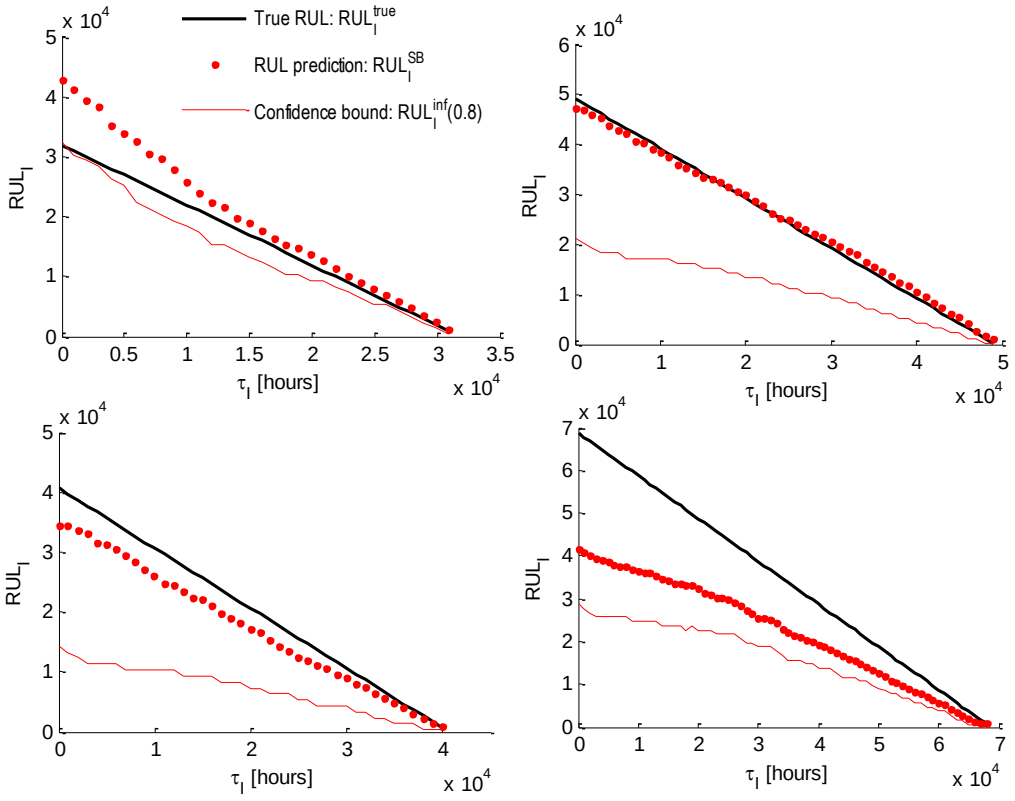

Figure A2: predictions obtained for 4 different test trajectories of Figure 4 using $\lambda = 5\text{x}10^{-3}$ and $\gamma = 0.5$.

Table A1: *RUL* prediction performance.

		$\beta = 25\%$		$\beta = 50\%$		$\beta = 75\%$	
$\overline{RUL}$ (10^4)		2.711		1.827		0.975	
$Cov_{0.2}$	$\lambda = 5x10^{-4}$; $\gamma = 0.7$	0.782		0.814		0.849	
	$\lambda = 5x10^{-3}$; $\gamma = 0.5$	0.838		0.850		0.853	
$\sqrt{MSE}$ (10^3)	$\lambda = 5x10^{-4}$; $\gamma = 0.7$	9.152	33.8%	5.965	32.6%	3.191	32.7%
	$\lambda = 5x10^{-3}$; $\gamma = 0.5$	9.822	36.2%	6.160	33.7%	3.411	35.0%
$MA_{0.2}$ (10^3)	$\lambda = 5x10^{-4}$; $\gamma = 0.7$	8.445	31.2%	5.594	30.6%	3.228	33.1%
	$\lambda = 5x10^{-3}$; $\gamma = 0.5$	11.300	41.7%	7.167	39.2%	3.960	40.6%

The results of Figure A2 and Table A1 confirm that the oscillation of the confidence bound can be damped down by increasing the value of λ or reducing the value of γ, but this choice increases the prediction error and the amplitude $MA_{0.2}$. Clearly, to an increased $MA_{0.2}$ corresponds also a higher value of the coverage indicator $\text{Cov}_{0.2}$.

When a situation with a larger density of reference trajectories is considered, the oscillations of the lower bound become of smaller amplitude, although more frequent. This happens, for example, when a larger number R of reference trajectories is available or when the variability within the degradation trajectories becomes smaller. To show this, we have reduced the variance of the parameters φ_{m0} and Q_m, $m = 1{:}\,6$, and of the load σ and temperature T used in the model of eq. (11) to simulate $N^{tst} = 50$ test trajectories and $N^{trn} = 50$ training sets of $R = 50$ reference trajectories. The optimization procedure applied for the case with $R = 7$ has been used to set the parameters to the values $\lambda = 5\text{x}10^{-5}$ and $\gamma = 0.95$. Four examples of the predictions obtained are shown in Figure A3, whereas the values of the performance indicators are presented in Table A2. As expected, with a higher density of training trajectories available, the prediction is both more accurate and precise.

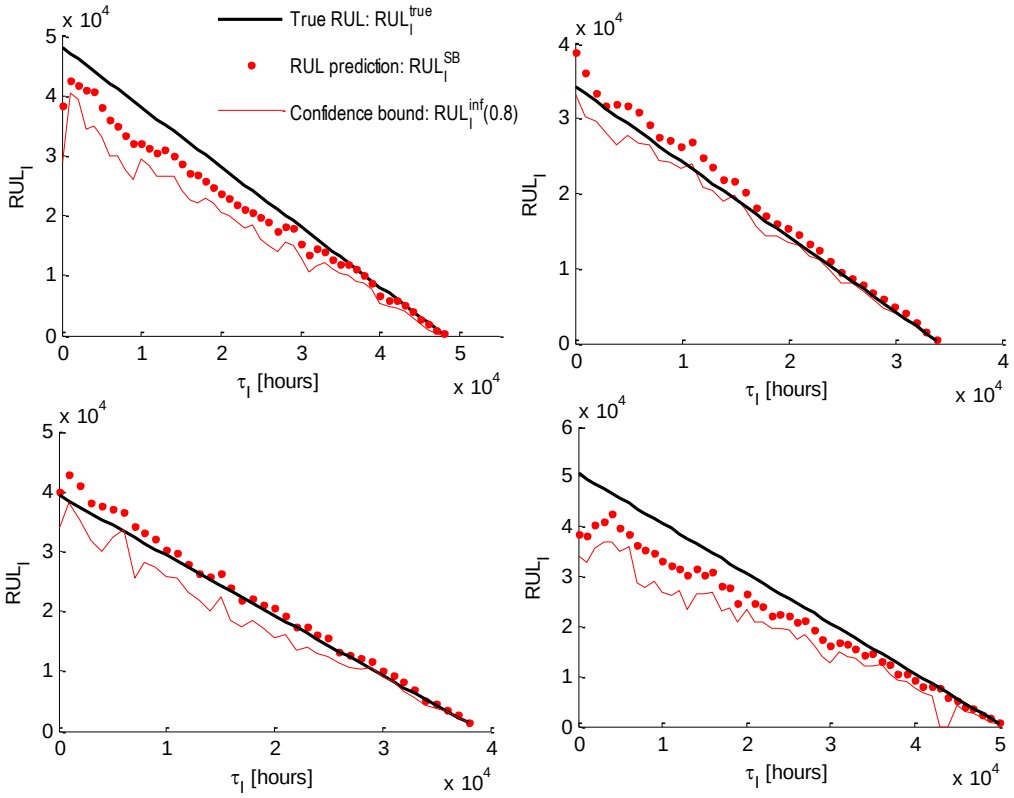

Figure A3: predictions obtained for 4 different test trajectories using R=50 reference trajectories and parameters $\lambda = 5\text{x}10^{-5}$ and $\gamma = 0.95$.

Table A2: RUL prediction performance with $\lambda = 5\text{x}10^{-5}$ and $\gamma = 0.95$

	$\beta = 25\%$		$\beta = 50\%$		$\beta = 75\%$	
$\overline{RUL}$ (10^4)	3.108		2.090		1.102	
$Cov_{0.2}$	0.814		0.832		0.808	
$\sqrt{MSE}$ (10^3)	5.313	17.1%	3.444	16.5%	1.659	15.1%
$MA_{0.2}$ (10^3)	4.961	16.0%	3.187	15.2%	1.788	16.2%

Appendix B

Figure A4 shows the predictions obtained for the $Q = 8$ trajectories, when parameters $\lambda = 0.05$ and $\gamma = 1$ are used. In trajectory 6, we notice that the confidence bound is higher than the RUL prediction. This is an example of the counterintuitive results that can be obtained by setting $\gamma = 1$ if two trajectories are very similar. Figure A5 shows the similarities s_{j*}^{r} and the BBAs m_{RUL} assigned to the reference trajectories for the test trajectory 6 at time $\tau_{15} = 15$ working days (upper). We notice that trajectory 8 receives the belief assignment $m_{RUL}(\{RUL^8\}) = 0.937$. Figure A5 also shows the evolution of the observable parameters ΔP_i^q, $\dot{M}_i^q$, and T_i^q (bottom), for the test trajectory 6 and the reference trajectory 8 receiving the maximum belief assignment. We notice that all three parameters ΔP_i^q, $\dot{M}_i^q$, and T_i^q of the two trajectories are very similar around time $\tau_{15} = 15$ working days, but evolve very differently after that time. To correct this problem, it is sufficient to reduce the value of parameter γ as can be seen from Section 3.2, Figure 9.

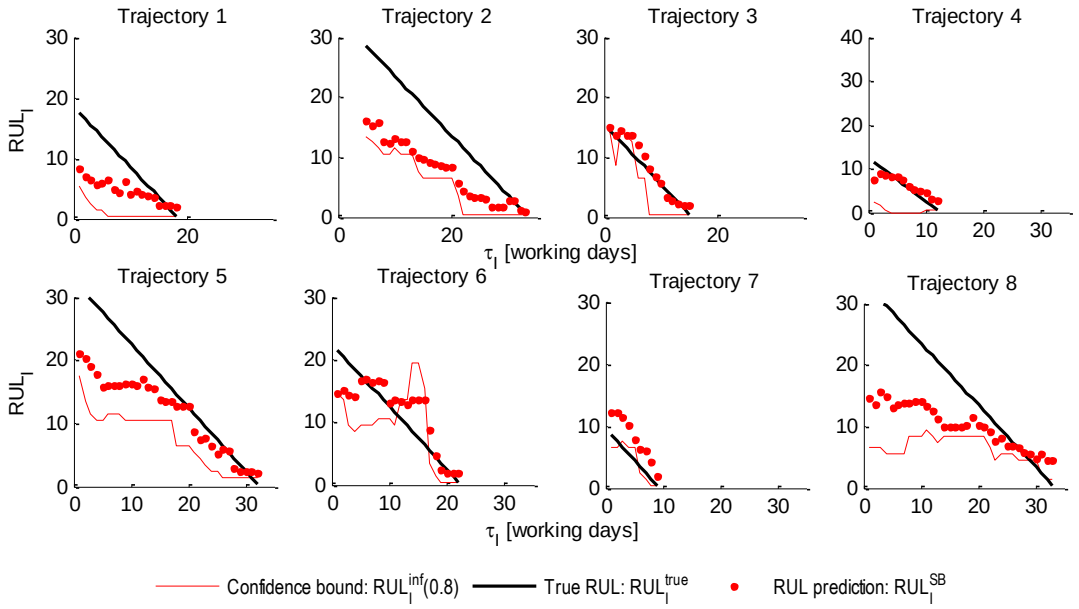

Figure A4: predictions obtained for the $Q = 8$ filter clogging trajectories available using parameters $\lambda = 0.05$ and $\gamma = 1$.

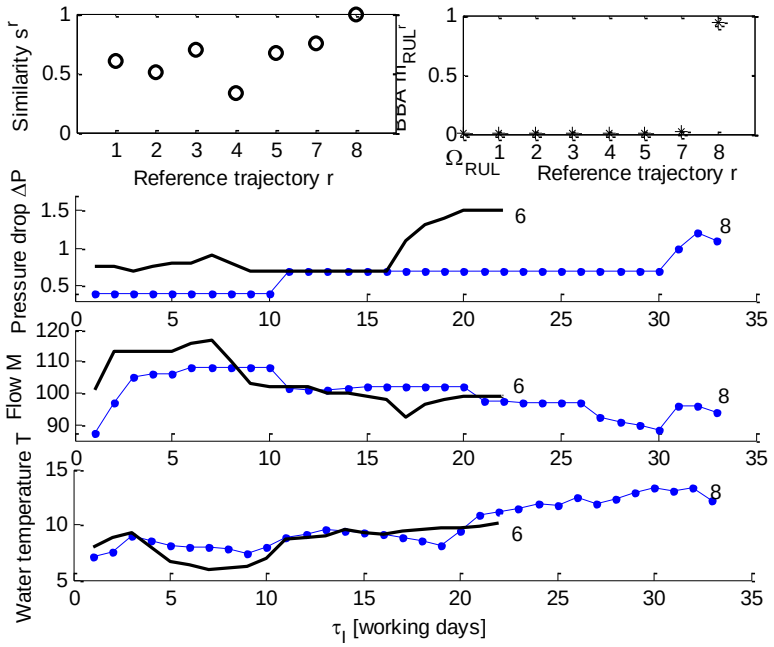

Figure A5: evolution of the three observable parameters ΔP_i^q, $\dot{M}_i^q$, and T_i^q (bottom) for trajectories $q = 6$ and $q = 8$, with similarities $s_{j^*}^r$ and BBAs m_{RUL} at time $\tau_{15} = 15$ working days (upper).