A Possibilistic Approach for Uncertainty Representation and Propagation in Similarity-Based Prognostic Health Management Solutions

Loredana Cristaldi[1], Alessandro Ferrero[1], Simona Salicone[1], Giacomo Leone[2]

[1]Dipartimento di Elettronica, Informazione e Bioingegneria (DEIB), Politecnico di Milano, Milano, Italy
[2]Sorint.Tek, Grassobbio, Italy
Email: loredana.cristaldi@polimi.it, alessandro.ferrero@polimi.it, simona.salicone@polimi.it, gleone@latek.it

Abstract

In this paper, a data-driven prognostic model capable to deal with different sources of uncertainty is proposed. The main novelty factor is the application of a mathematical framework, namely a Random Fuzzy Variable (RFV) approach, for the representation and propagation of the different uncertainty sources affecting Prognostic Health Management (PHM) applications: measurement, future and model uncertainty. In this way, it is possible to deal not only with measurement noise and model parameters uncertainty due to the stochastic nature of the degradation process, but also with systematic effects, such as systematic errors in the measurement process, incomplete knowledge of the degradation process, subjective belief about model parameters. Furthermore, the low analytical complexity of the employed prognostic model allows to easily propagate the measurement and parameters uncertainty into the RUL forecast, with no need of extensive Monte Carlo loops, so that low requirements in terms of computation power are needed. The model has been applied to two real application cases, showing high accuracy output, resulting in a potentially effective tool for predictive maintenance in different industrial sectors.

Keywords

Data-Driven, Epistemic Uncertainty, Measurement Uncertainty, Future Uncertainty Prognostics and Health Management, Random Fuzzy Variable, Remaining Useful Life, Similarity

1. Introduction

In the last years, data-driven prognostic approaches have experienced a great

diffusion, mainly because of the increasing availability of condition monitoring data in substantial quantities, which is one of the pillars of the modern industrial paradigm Industry 4.0. Since such class of algorithms is based on the elaboration of data obtained by means of measurement processes, the development of approaches for the quantification, management and propagation of the measurement uncertainty is of fundamental importance. Nevertheless, since prognostics deals with predicting the future behaviour of a system and it is practically impossible to precisely predict future events, measurement uncertainty is not the only uncertainty source. Other sources such as the uncertainty about the future operational conditions the systems will face and model uncertainty, in fact, play a relevant role. In this regard, it is then necessary to account for the different sources that affect prognostics and develop a framework for uncertainty quantification and management.

Different approaches can be found in the literature. An example of model-based approaches, that rely on mathematical models to describe the degradation process and provide a RUL prediction [1], are filtering methods, which are capable of accounting for the stochasticity of the process and measurement uncertainty. The exact Kalman filter is widely used in case of linear state space models and independent, additive Gaussian measurements and modelling noises [2]. However, when the dynamic of degradation is nonlinear and/or the associated noises are non-Gaussian, other techniques must be adopted. In this regard, numerical approximations based on the Monte Carlo (MC) sampling technique are very popular, because of their flexibility and ease of design [3] [4]. Among them, Particle Filtering is extensively used for diagnostic and prognostic applications [5]. In the particle filtering scheme, the probability density function (PDF) of the model parameters are updated when new observations of the equipment degradation are acquired. The target posterior PDF is approximated by a large number of random samples termed particles, each of which has a likelihood weight assigned to it that represents the probability of that particle being sampled from the PDF.

However, the development of a physics-based model for the description of the degradation process of complex systems may be a very hard task, which usually requires a deep domain knowledge. Furthermore, even if a stochastic model is available, the application of MC sampling techniques may require high computational power, due to the large number of iterations needed for the statistical convergence. Convergence itself may be an issue for some of the cited techniques. For example, particle filtering is prone to the so-called particle degeneracy, a phenomenon for which, after a number of posterior PDF updates, only one particle has significant weight. This is quite common in high-dimensional problems, rendering traditional particle filter algorithms ineffective in such cases.

Focusing on the data-driven models domain, bootstrap ensemble approaches [6] [7], which are based on the aggregation of multiple model outcomes, have gained interest due to their ability of estimating the uncertainty in the predic-

tions. In most applications, these approaches are used to estimate only the model uncertainty, by considering the variability in the predictions of the different models of the ensemble. An effective solution to estimate also the contribution of the stochasticity of the degradation process and the measurements noise to the RUL forecast uncertainty is given in [5] [6]. In this work, a training set is used to define an ensemble of predictive models that receive in input a degradation observation z and produce a RUL forecast. Each model is trained by using a bootstrapped replicate of the training set. The different predictions provided by the models are exploited in order to estimate the model error variance σ_m^2. In order to estimate the remaining fraction σ_r^2 of the RUL prediction variance, caused by the randomness of the degradation process and the observation noise, an independent validation dataset is used. In particular, the ensemble of empirical models is applied to the observations z in the validation dataset. Denoting as ΔRUL_i^2 the squared difference between the RUL forecast and the actual RUL when the observed degradation is z_i, an estimate of σ_r^2, as function of z_i, is obtained as difference between ΔRUL_i^2 and σ_m^2. Computing such difference for each available degradation observation in the validation set, an empirical model $\chi(z) = \sigma_r^2$ can be defined.

Such solutions, however, present some problems. First, in case of complex empirical models characterized by large training time (f.i. deep neural networks), the training of an ensemble of them may be an issue. Nevertheless, a considerable amount of data is usually required to ensure heterogeneity in the bootstrapped replicas of the training set, and therefore in the resulting models.

Another factor to take into account when dealing with prognostics, is the epistemic uncertainty introduced by the incomplete knowledge and information on the parameters used to model the degradation and failure processes. Interesting methods for the representation and propagation of both aleatory (probabilistic) and epistemic uncertainty sources are found in [8]. In this work, the authors propose two methods: a pure probabilistic method and a hybrid method combining Monte Carlo and possibilistic methods. In the first solution, both epistemic and aleatory uncertainties are represented as probability distributions and a double-MC loop is performed: in the outer loop epistemic variables are sampled, whereas aleatory variables are sampled in the inner loop. In this way, a set of different cumulative distributions of the RUL is obtained (one for each iteration of the outer loop). Although all information about RUL uncertainty is preserved, its interpretation may not be straightforward in practical terms. Conversely, in the hybrid method the aleatory and epistemic uncertainties are represented respectively by probability and possibility distributions [9]. The method consists of a single MC loop: at each iteration, a realization of aleatory variables is performed and possibilistic interval analysis is carried out to process the epistemic uncertainty, so that a possibilistic random distribution of the RUL is obtained. Finally, at the end of the loop a set of possibilistic random distributions for the RUL is obtained and they can be combined into a single set of limiting cumulative distributions characterized by different degrees of confidence

[9].

Although operatively straightforward, such methods present some drawbacks. The first methods, relying on two MC loops requires a very large number of iterations, so that high computational power may be required in order to decrease the processing time. As for the second method, even if based only on a single MC loop, its main drawback is that it is based on a mix of two different mathematical domains, probability and possibility.

In order to overcome the limitations of the cited literature works, in this paper a novel data-driven prognostic model capable to deal with different sources of uncertainty is proposed. The main enhancement is the application of a unique mathematical framework, namely a Random Fuzzy Variable (RFV) approach, which allows the representation and propagation of the different aleatory and epistemic sources of uncertainty affecting Prognostic Health Management (PHM) applications: measurement uncertainty, present uncertainty, future uncertainty and model uncertainty.

The RFV approach, in fact, enables the representation and combination into a single mathematical object of the aleatory and epistemic contributions to uncertainty. Therefore, it results particularly suitable to deal not only with random measurement noise and model parameters uncertainty due to the stochastic nature of the degradation process, but also with systematic effects, such as systematic errors in the measurement process, incomplete knowledge of the degradation process, subjective belief about model parameters. Furthermore, the low analytical complexity of the employed prognostic model allows to easily propagate measurement and parameters uncertainty into the RUL forecast, with no need of extensive MC loops, so that low requirements in terms of computation power are needed. The model output is the RUL forecast, which, once again, is represented in terms of RFV, so that a confidence interval, at the desired confidence level, can be easily provided.

The rest of the paper is structured in the following way: in Section 2 the sources of uncertainty in PHM are introduced. In Section 3 the concept of similarity approach for prognostics is described, whereas the proposed prognostic model is illustrated in Section 4. Section 5 is dedicated to the introduction of the RFV approach for the representation of the measurement uncertainty through the possibility theory. The application of the RFV approach to the proposed prognostic model is described in Section 6. In Section 7, details about the tuning procedure for a crucial model parameter are given. Finally, in Section 8 the results obtained for two real case studies are presented, followed by conclusions in Section 9.

2. Sources of Uncertainty in Prognostic Health Management

Sources of error like modelling inconsistencies, system noise and degraded sensor fidelity can affect prognostic predictions. In PHM, considering [9] [10] [11], the following sources can be considered:

- Measurement uncertainty: the collected data are affected by a measurement uncertainty due to the employed sensors and instruments. Two kinds of uncertainty sources can be considered, typically referred to as systematic and random.

- Present uncertainty: Remaining Useful Life (RUL) prediction requires the current state estimation of the system. The system state may depend on multiple variables, which can be directly or indirectly monitored through sensors. If properly processed, such signals allow the extraction of features which are informative about the system health state and in most cases lead to a better interpretation of the data. The impossibility to perfectly estimate the state, as well as the propagation of the measurement uncertainty into the process of feature extraction contribute to the definition of the present uncertainty.

- Future uncertainty: it is due to the inability to predict exactly in advance the future operational conditions (like load conditions, environmental and usage conditions) of the system. It is often the most relevant source, as shown in [12].

- Model uncertainty: this source is strictly related to the application cases and to the applied approaches. Model uncertainty includes model parameters stochasticity, and process noise. Under-modelling is also an issue, due to missing failure modes in the analysis or, in case of application of data-driven approaches, the lack of data describing possible failure scenarios. Furthermore, epistemic uncertainty for the representation of expert's belief about model parameters is another source to account for.

3. Similarity-Based Prognostic Models

Many methods for uncertainty processing in PHM are present in the literature. In this regard, similarity-based prognostic algorithms represent an interesting class of Data Driven prognostic approaches, whose main contribution is that they easily account for future uncertainty.

The hypotheses for the application of a similarity-based approach are the followings:

1) Run-to-failure historical data from multiple units of a system/component are recorded (the term unit refers to an instance of a system/component).

2) The historical data covers a representative set of units of the system/component. Such set of products will be referred in the following as reference library (or simply library).

3) The history of each unit ends when it reaches a failure condition, or a preset threshold of undesirable conditions, after which no more runs will be possible or desirable (the history can start, however, from a variable degrading condition).

In order to estimate the RUL of a test item (for which the RUL has to be predicted), a similarity assessment between the test degradation pattern (monitored degradation pattern for the test item) and the reference trajectory patterns in the library is performed. An example can be found in [13], where the similarity is as-

sessed based on the computation of a distance value, as given in Equation (1):

$$d_i = \sqrt{\frac{1}{K} \sum_{k=1}^{K} \left(y_k - y_{ik} \right)^2} \qquad (1)$$

which corresponds to the Root Mean Square Error (RMSE) between the pattern of the test unit and the pattern of the *i*-th library specimen. In particular, y_k (y_{ik}) corresponds to the observed degradation for the test unit (library specimen *i*) at cycle (or generally time stamp) k and K is the total number of observed cycles for the test pattern. A small d_i means that the two profiles are characterized by similar degradation processes, whereas large distance values are related to products that are subject to different degradation mechanisms and therefore they should be excluded in the estimation of the RUL of the target product.

In the same paper the Authors exploit the degradation data of library items with higher similarity and process them through MC simulations in order to forecast the future degradation pattern of the test unit and obtain a Confidence Interval (CI) for the related RUL.

Another strategy is to evaluate the test unit RUL as function of the RUL of the reference units and their similarity degree with respect to the test unit, as shown in [14] [15]. In these works, a weight is assigned to each reference unit, according to Equation (2):

$$w_i = \exp\left(-\frac{sc_i}{\beta} \right) \qquad (2)$$

where the similarity coefficient sc_i, similarly to the distance d_i computed in Equation (1), is a function of the sum square error between the degradation patterns of test unit and library specimen *i*. As for parameter β, it is an arbitrary parameter that can be set by the analyst to introduce the desired degree of selectivity (*i.e.* if β is small, few specimens are influential). Finally, the test RUL is computed as the weighted sum of the RULs of the library units:

$$\text{RUL} = \frac{\sum_{i=1}^{N} w_i \cdot \text{RUL}_i}{\sum_{i=1}^{N} w_i} \qquad (3)$$

where *i* refers to the *i*-th reference unit and *N* is the number of units in the library.

A similar approach is suggested in [16]. The authors, in fact, propose a definition of a deterministic model M_i for each *i*-th training unit of the library, so that an estimated value of the degradation is provided at each given time. At this point, if a time-sequence Y of degradation values for a test unit is available, a distance measure between the model M_i and Y is defined as the sum of the squared errors, divided by the prediction variance of the model M_i. Then the RUL estimation for the test product is equal to a weighted sum of the RUL of the reference products, where the weights are assigned applying the k-nearest neighbour method, that is selecting the products with the k smallest distance values and apply a weight $1/k$

to their RULs.

The assumption is that the future operational conditions of the test product will be most likely similar to those of the units exhibiting higher similarity in correspondence of the observation window (*i.e.* the time interval for which the target unit degradation pattern has been monitored). Therefore, the higher the similarity between the test item and the *i*-th training reference pattern, the higher the weight w_i will be, so that the forecasted RUL will be closer to RUL_i, dealing in such a way with the uncertainty about the future operational, environmental and usage conditions.

However, some limitations are shared by the cited works and their analogous:

1) Their application may be precluded when there is no availability of run-to-failure degradation profile of a numerous, representative set of products, as it may occur with those systems characterized by long expected lifetime.

2) The literature investigation about similarity-based prognostic approaches has highlighted that measurement uncertainty is often neglected and not properly quantified and processed within the adopted prognostic model. If on one hand, this fact can be justified by considering that often measurement uncertainty represents a minor contributor of the overall uncertainty (especially at early life stages when the uncertainty about future operating conditions is unavoidably the main source), on the other hand it still represents a factor that may lead to a more accurate prognostic output if properly accounted.

As described in the next section, also the prognostic algorithm proposed in this paper estimates the RUL of a test item as weighted sum of the RUL of the library units, but it aims to overcome the cited limitations by applying the RFV approach. Such approach, in fact, is particularly suitable for the representation and propagation in a unique mathematical framework of both aleatory and epistemic uncertainties, and in this work, it has allowed to effectively deal with the measurement uncertainty associated to the degradation data and the epistemic uncertainty associated to the RUL of those library units whose time of failure is not known.

4. The Proposed Model

Analogously to [14] [15] [16], the prognostic model proposed in this paper evaluates the test RUL according to Equation (3). The main difference lies in how the weighting coefficients w_i are defined. Here, in fact, the distance value d_i (which gives information about patterns similarity) is mapped into a weight w_i through a mapping function $g(\cdot)$, defined as:

$$w_i = g(d_i) = \frac{1}{\sqrt{2\pi\sigma_g^2}}\exp\left(-\frac{(d_i - d_{\min})^2}{2\sigma_g^2}\right) \tag{4}$$

Function $g(\cdot)$ represents a Gaussian PDF, characterized by mean value equal to $d_{\min}$ (by definition, the minimum value among all d_i values, $i = 1, \cdots, N$) and standard deviation σ_g, such that higher weights are assigned to reference units at lower distance d_i. The motivation behind the definition of such function as a

Gaussian PDF is that, as shown in Section 6, it allows to easily propagate the measurement uncertainty into weights w_i.

A key factor in the application of the proposed algorithm is the value assigned to σ_g, which, similarly to parameter β in Equation (2), introduces the desired degree of selectivity. In Section 7, a strategy for a suitable choice of this parameter will be shown.

5. The RFV Approach

During the recent years, a more general approach to measurement uncertainty evaluation and propagation has been proposed by [17]. This new approach is framed within the mathematical theory of evidence, proposed by Shafer in the seventies, and represents a generalization of the probabilistic one recommended by the GUM. It allows one to represent and process any kind of incomplete information, both of random and systematic nature. In particular, the measurement results are expressed in terms of a particular class of type-2 fuzzy variables, the Random-Fuzzy Variables (RFV).

An example of RFV is shown in Figure 1. This figure shows that an RFV is composed by two functions, called possibility distribution functions (PD): $r_{int}(x)$ (cyan line in Figure 1) and $r_{ext}(x)$ (blue line in Figure 1). By considering these two PDs, it is possible to represent, in a single mathematical object, the effects of all possible contributions to uncertainty on the true value of the measurand. It was proved [17] that the external PD $r_{ext}(x)$ represents the effects of all contributions to uncertainty, whilst the internal PD $r_{int}(x)$ represents the effects of all non-random contributions to uncertainty, including the systematic ones.

It is also possible to prove that $r_{ext}(x)$ can be obtained by combining $r_{int}(x)$ with the random PD $r_{ran}(x)$ (magenta line in Figure 1) that represents the sole random contributions to uncertainty. It is hence possible, with a single RFV, to represent all contributions to uncertainty.

An interesting property of the RFVs is that their α-cut, for each level $\alpha \in [0, 1]$, provide all possible confidence intervals, at confidence levels $1 - \alpha$. Therefore, each α-cut provides an interval, within which the true value of the measured is expected to lie with a coverage probability $1 - \alpha$.
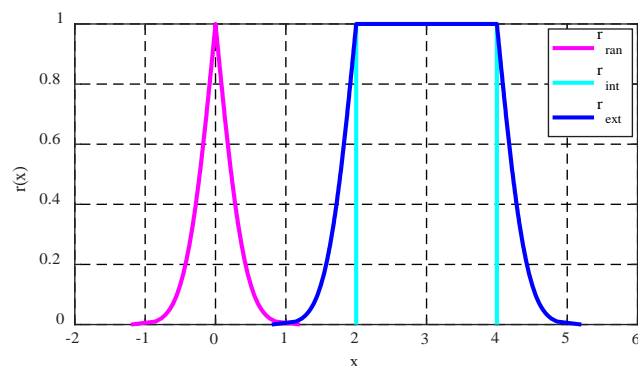
Figure 1. Example of Random-Fuzzy variable (RFV).

RFVs can be combined, according to a given measurement function, by means of appropriate operators, called *t*-norms, applied to the random PDs $r_{\mathrm{ran}}(x)$ and the internal PDs $r_{\mathrm{int}}(x)$. Therefore, it is possible to combine, in closed form [18] [19], different measurement results and obtain the final measurement result together with its associated uncertainty.

Two *t*-norms have been selected to process the PDs of the RFVs: the *min t*-norm and the *Frank t*-norm. The choice of which one should be used is done according to all the available metrological information related to both the nature of the uncertainty contributions and the way they affect the measurement procedure [20]. In particular, the *min t*-norm is used when the uncertainty contributions affect the measurement procedure in a systematic way and therefore, they do not compensate with each other. On the other hand, the *Frank t*-norm is used when the uncertainty contributions affect the measurement procedure in a random way and therefore, they do compensate with each other.

Next section shows how the RFV approach can be applied to the proposed model.

6. Application of the Random-Fuzzy Variable Approach to the Proposed Model

To apply the RFV approach to the proposed model, Equations (1), (3), (4) must be evaluated in terms of RFVs. This means that the measured values y_k, the distance values d_i, the weighing values w_i and the RUL values must be considered as RFVs, by also considering the different uncertainty contributions.

6.1. Computation of the RFV of the Measurement Values

To build the RFVs associated to each measured value y_k, random and systematic contributions to uncertainty are considered. It is supposed that the random contributions distribute according to a Gaussian PDF, having a standard deviation σ. It is also supposed that the systematic contributions distribute over an interval, but no information is available on the way they distribute, so that no PDF can be assigned in this case. The considered interval is centered on the measured value y_k and its width is supposed to be proportional to the measured value itself. If a relative error e is considered, the interval will have a semi-width $y_k \cdot e$.

Under the above assumptions, it is possible to build the RFV associated to each measured value. According to the available information:

- The random PD $r_{\mathrm{ran}}(x)$ represents the random contributions to uncertainty and therefore is built from the given Gaussian PDF, by applying a suitable transformation, called probability-possibility transformation [19] [21].

This transformation allows one to transform a PDF into an equivalent PD which preserves all the coverage intervals and corresponding coverage probabilities, thus maintaining the relevant metrological information associated with the initial PDF;

- The internal PD $r_{\mathrm{int}}(x)$ represents the systematic contributions to uncertainty and therefore is built according to the given interval. In particular, in Shafer's

theory of evidence, the considered situation when an interval of variation is given, but no PDF can be defined, is called total ignorance and is represented by a rectangular PD over the given interval.

It follows that the RFV shape of the RFV Y_k associated to each measured value is similar to the one shown in **Figure 1** in cyan and blue lines. Of course, the mean value and width of the RFV change with the measured value.

6.2. Computation of the RFV of the Distance between Degradation Curves

Equation (1), which provides the distance d_i, can be considered in two different ways.

The first consists in building all RFVs Y_k and Y_{ik} and apply the most appropriate *t*-norms in each operation considered in Equation (1). This procedure, which is mathematically correct, does not consider however all available metrological information and brings to an overestimation of the uncertainty associated to distance d_i (*i.e.* a very large RFV D_i). On the other hand, using all available metrological information, as shown by [22], it is possible to find, for each *i*-th curve, the RFV of the distance in a more immediate and accurate way.

In fact, let us consider that Equation (1) represents a mean square error. Let us start from the random contributions to uncertainty. By assumption, the standard deviation of the given PDFs (PDFs associated with the measured values) is the same for each measured value (see Sec. 6.1). It is known that the standard deviation σ_{d_i} of the mean square error is:

$$\sigma_{d_i} = \frac{\sigma}{\sqrt{K}} \tag{5}$$

Equation (5) allows us to directly build the random PD r_{ran} to be associated to the RFV of the distance d_i from the *i*-th curve, by simply applying the probability-possibility transformation to a Gaussian PDF having standard deviation equal to σ_{d_i}. This solution is indeed straightforward and avoids the combination of all different random PDs associated to the K measured values.

Similar considerations can be done when the systematic contributions to uncertainty are considered, associating in a straightforward way the final systematic contribution to uncertainty to the distance d_i, thus also avoiding the combination of all different internal PDs associated to the K measured values [22].

In fact, distance d_i presents a relative error e_{d_i} due to the systematic contributions to uncertainty:

$$e_{d_i} = \frac{e}{d_i \cdot K} \cdot \left| \sum_{k=1}^{K} y_k - \sum_{k=1}^{K} y_{ik} \right| \tag{6}$$

so that it is possible to directly build the internal PD r_{int} (to be associated to the RFV of the distance d_i from the *i*-th curve) by considering a rectangular PD around d_i, with semi-width equal to $e_{d_i} \cdot d_i$.

By combining the obtained internal and random PDs, RFV D_i associated to

the distance d_i is evaluated [17].

6.3. Computation of the RFV of the Weighting Coefficients

Once RFVs D_i are built for all curves $i = 1, \cdots, N$, also weights w_i (denoted W_i) can be evaluated in terms of RFVs. First of all, the mapping PDF $g(\cdot)$ is converted into an equivalent mapping PD $G(\cdot)$, as shown, as an example, by the red PD in Figure 2.

Then, RFV W_i is obtained by considering the intersection of D_i with the mapping PD. As an example, in Figure 2 it is shown how a single a α-cut of RFV W_i is obtained: a generic α-cut D_i^α of RFV D_i is considered (green line); this interval intersects the red PD and identifies the magenta interval, which represents the α-cut W_i^α, at the same level α, of RFV W_i.

By considering the same method for all α-cuts of RFV D_i, RFV W_i is built.

6.4. Computation of the RFV of the Test Unit RUL

Finally, from the weights W_i and according to Equation (2), it is possible to evaluate the RUL in terms of RFV. In particular, the parameter RUL$_i$ associated to the generic i-th reference unit will be either a scalar (if it is known) or a RFV (if not known). In the latter case, in fact, it is possible to introduce epistemic uncertainty about RUL$_i$; assuming that the test RUL forecast is performed at time stamp t' and that the user assumes that the maximum lifetime for i-th reference unit is equal to T' (based on personal or other experts' opinion), RUL$_i$ can be modeled as a rectangular RFV (which describes total ignorance) ranging in $[0, T' - t']$.

According to the available metrological information about the nature of the contribution and the measurement procedure, it is necessary to choose the more suitable t-norms to be applied. It is possible to state that the weights W_i are all uncorrelated with each other because they are related to the i-th curve and all N curves are independent from each other. Furthermore, as far as the systematic contributions to uncertainty affecting the weights W_i are concerned, there is no reason to suppose a probabilistic compensation between each other in Equation (2) and hence we can assume that they combine in a non-random way. Therefore, the *min* t-norm is chosen and a zero correlation factor applied when combining PDs r_{int}, while the Frank t-norm with the parametric value $\gamma = 0.1$ is chosen and a zero correlation factor applied when combining PDs r_{ran} [19].

Figure 3 represents an example of RFV RUL (in blue) obtained for a test unit and the corresponding CI (in red color), whose limits are denoted as RUL$_{\text{min}}$ and RUL$_{\text{max}}$, obtained selecting the α-cut at level $\alpha = 0$ of the RFV.

7. Tuning σ_g Parameter

As stated in Secetion 4, a key point in the application of the proposed algorithm is the choice of the standard deviation σ_g associated to the mapping function $g(\cdot)$ in Equation (3), as it controls the selectivity degree.
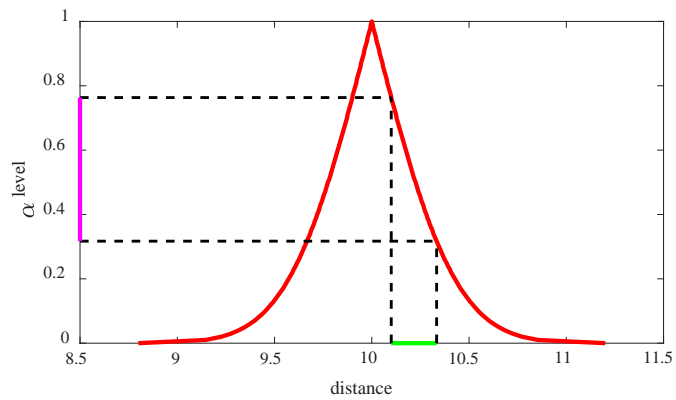
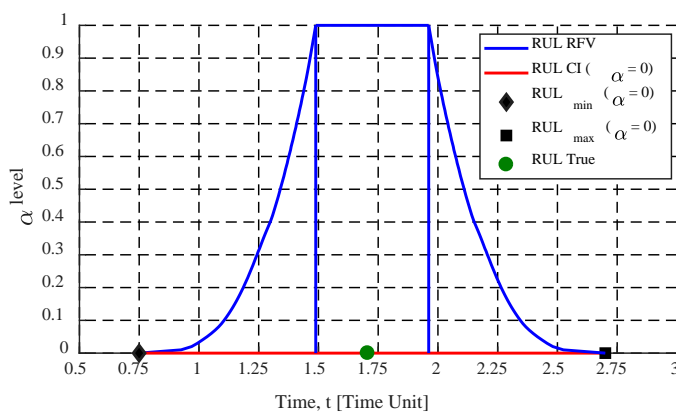**Figure 2.** Example of possibility distribution functions (PD).

Figure 3. RFV RUL (in blue) for a given unit and corresponding RUL CI (in red) obtained choosing $\alpha = 0$. Note that the actual RUL value (in green) is included in the CI provided by the algorithm.

In this work, a grid optimization approach is proposed for the determination of the value σ_g that maximizes the prognostic performances. Such approach consists in defining an *a-priori* set of admissible values for the parameter to be tuned, $\sigma^* = \left\{\sigma_1^*, \cdots, \sigma_J^*\right\}$, running the algorithm for each possible value and finally selecting the one which maximizes the performances.

Let as assume to be interested in forecasting the RUL of a test unit whose degradation pattern is known up to an observed level δ, and a reference library of N units is available. The following steps are performed:

1) First, the M units at lower distance d_i with respect to the test unit are identified.

2) The parameter σ_g is set equal to the generic j-th admissible value σ_j^* ($j = 1, 2, \cdots, J$).

3) A Leave-One-Out-Cross-Validation (LOOCV) is then run: the prognostic algorithm is run setting the m-th unit at lower distance ($m = 1, 2, \cdots, M$) as test sample (its degradation pattern is considered known up to the value δ) and the remaining $N - 1$ reference curves as training patterns. Two fundamental metrics are then computed. The first metric is a performance indicator PI, p_{jm}, which

informs about the correctness of the prediction:

$$p_{jm} = \begin{cases} 1 & \text{if } \text{RUL}_{act} \in [RUL_{min}, RUL_{max}] \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

where RUL_{act} corresponds to the actual RUL value. In other words, p_{jm} is equal to 1 when CI of the RUL prediction contains the actual RUL value, 0 otherwise.

The second metric, Δ_{jm}, is related to the width of the provided CI and is computed as:

$$\Delta_{jm} = RUL_{max} - RUL_{min} \tag{8}$$

4) Step 3 is repeated setting cyclically one of the M units as test unit. Finally, the average value of the performance indicators is computed for the j-th value of the parameter σ_g, according to:

$$p_j = \frac{\sum_{m=1}^{M} p_{jm}}{M} \tag{9}$$

$$\Delta_j = \frac{\sum_{m=1}^{M} \Delta_{jm}}{M} \tag{10}$$

5) The metrics p_j and Δ_j are computed for each of the admissible value of σ_g.

6) The optimal σ_g should lead to high value of PI, while keeping the CI width small (because the CI width reflects the uncertainty about the RUL output). In order to guarantee such conditions, first a threshold P^* is set and a subset of values σ_g is determined, by choosing those values for which the PI is higher than the given threshold, that is $p_j \geq P^*$. Let us denote with Σ this subset and with Δ_Σ the corresponding values of mean CI width Δ_j.

In case P^* is not achieved for any value of σ_g (*i.e.* Σ is an empty set), the value of σ_g providing the highest performance is automatically selected, without the need to perform the next point.

7) Once that a subset of optimal values of σ_g is determined, for what concerns the PI, the optimization of the CI width should be addressed. An idea could be selecting the value of σ_g providing the lowest value Δ_j. However, it must be considered that the metrics computed in steps 3-4 are obtained by means of a validation set of M units. In order the validation set to be representative of the system under analysis, M must be sufficiently large. Unfortunately, this is not always the case in practice. Therefore, selecting the value of σ_g as the one providing the lowest value Δ_j could be too restrictive, leading to RUL forecasts with small CI widths, but low PI. In this work, two selection strategies (SSs) will be considered and compared with each other. The two selection strategies are denoted as SS1 and SS2 and explained below:

- SS1: among the values in Σ, select the value σ_g providing the lowest CI width Δ_j.
- SS2: select the value of σ_g providing the confidence width Δ_j equal to the 5% quantile of Δ_Σ.

The aim is to determine which SS guarantees the best trade-off between high prognostic accuracy (*i.e.* the obtained CI for the RUL encloses the actual RUL value) and narrow confidence intervals, to provide valuable results from the predictive maintenance point of view.

8. Algorithm Validation

Two different application cases will be presented. The first one (AC1) is the same considered by [22] and consists in a database of 90 degradation patterns of Medium Voltage (MV) and High Voltage (HV) Circuit Breakers (CBs). Since the data used in the contribution are confidential information, the exact numerical values are not reported. The degradation pattern y for such units is shown in **Figure 4(a)**. The time is expressed as a generic Time Unit and the degradation level y as percentage. A degradation level $y = 0\%$ refers to a unit in a perfect healthy state, whereas $y = 100\%$ means that it has reached its End of Life (EoL), that is the time instant at which the unit is not anymore able to perform its intended function and a maintenance activity, refurbishment, replacement or the disposal of the unit is required.

The second application case (AC2) considers fatigue-crack-growth data, as presented by [23]. There are 21 sample paths, one for each test unit. It is assumed that the testing of the units is stopped at 0.12 million cycles and the units have failed if undergone a critical crack length of 0.04064 m (1.6 inches). All the units have at the beginning of the test an initial crack length of 0.02286 m (0.9 inches).

Figure 4(b) depicts the degradation pattern of the units. To be coherent with the representation of the first database, also in this case the degradation level y is
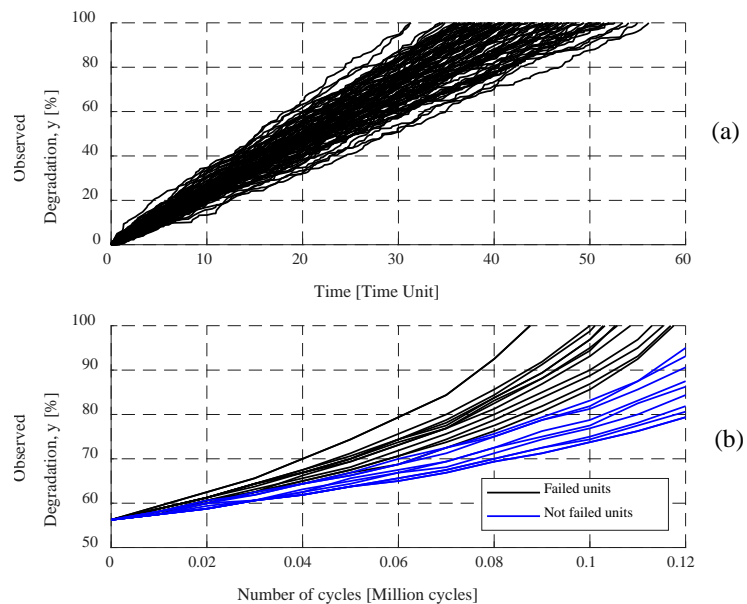
Figure 4. Degradation patterns for two different application cases. (a) A fleet of 90 Medium and High Voltage Circuit Breakers; (b) A fleet of 21 units undergone fatigue-crack-growth test.

expressed in percentage of the critical length previously reported. From **Figure 4(b)** it can be denoted that about half of the units have not failed by the end of the test. This and other features considerably differentiate the two application cases, as reported in **Table 1**.

The aim is to show the possibility to apply the proposed approach in a wide range of applications and its ability to overcome situations of scarce amount of degradation data.

No metrological information about the measurements involved in the two examples is available. According to authors' experience and personal assumptions, the standard deviation σ of the random contributions to measurement uncertainty and the relative error e are set equal to 0.1. Such quantities are adimensional, since the observed data are expressed as percentage of degradation.
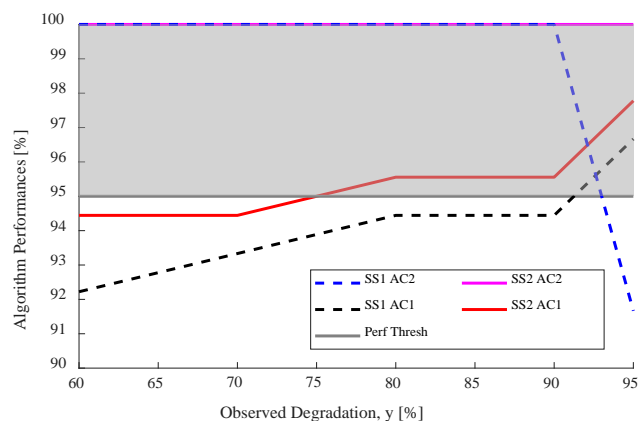
Figure 5 reports the prognostic performances (in terms of the Performance Indicator PI) of the proposed algorithm for the two application cases, at different levels of observed degradation. The PI is obtained as in Equation (7), averaging over the testing units. In this regard, for application case AC2, the performances have been computed only for the failed units, *i.e.* 12 units, whose RUL is known, and the validation of the provided results is viable.

The results refer to the performances obtained by considering, for each RUL, the CI corresponding to the α-cut at level $\alpha = 0$ of the RFV.

It is important to understand that the choice of the α-cut represents a trade-off between the width of the provided CI (amount of uncertainty about the RUL forecast) and the accuracy of the prognostic result (the provided CI includes the

Table 1. Difference in the two considered application cases.

Application case	Trend of degradation over time	Number of units in the reference library	Number of units in the reference library with known failure time
AC1	Linear	90	90
AC2	Exponential	21	12

Figure 5. Comparison of the prognostic performances obtained by the proposed algorithm for both application cases at different levels of observed degradation and through different selection strategies of the parameter σ_{g}.

actual RUL value). Higher levels of α-cut correspond to narrower CIs but also higher risk of incorrect forecasts. In this work, the level $\alpha = 0$ has been chosen, according to a pessimistic approach (worst case).

The dashed (solid) lines refer to the results obtained with the first (second) selection strategy SS1 (SS2). It is interesting to observe that the percentage of correct RUL forecasts for the first application case increases with the observed degradation, for both SS1 and SS2, but SS2 guarantees more accurate predictions. As an example, if a minimum threshold of 95% (grey solid line) is considered, SS1 allows to overcome it when the degradation is approximately 92%, while SS2 allows to overcome it in relevant advance than failure time, when approximately only 75% of degradation is observed.

The better performances achieved through SS2 are also confirmed in the second application case. In this case, as stated in Section 8, some of the units still have not reached the end of life. Therefore, for these units (whose failure time is unknown), the Authors have built the corresponding RFV of RUL according to their personal belief. A maximum lifetime equal to 0.15 million cycles is considered and the related RUL and associated epistemic uncertainty is modeled as a rectangular PD.

By applying SS2, the algorithm has provided correct predictions at each level of observed degradation for all the test units. Similar results have been obtained through SS1, except when RUL forecast has been performed at a level of observed degradation equal to 95%. In this case, the performances fall at 91.67%. However, one should observe that this decrease is due to an incorrect prediction for one single unit (1 unit of 12, indeed, corresponds to 8.33%). More in detail, for this particular case the incorrect prediction is due to a late prediction provided by the algorithm, such that the lower limit of the forecasted CI is larger than the actual RUL by only 13 cycles (*i.e.* the prediction error is very small).

Observing the results, it is important to highlight that the proposed algorithm have achieved excellent performances for both application cases. As already said, the second one is more complex than the first one, because of the exponential trend of the degradation over time and the smaller number of reference curves, for some of which the RUL is unknown. In this regard, one should not be misled by observing that the performances achieved for a more complex application case are higher. It is authors' opinion that testing the algorithm in AC2 for a larger set of units, the performances should normalize and exhibit a trend like the one exhibited by AC1.

At this point, once stated that SS2 guarantees more accurate predictions, it is useful to verify if the higher performances are counterbalanced with wider CIs. The benefit of better predictions would be vanished, indeed, if the provided CIs would be much wider. Narrow CIs for RUL, in fact, are of fundamental importance for an effective scheduling of the maintenance interventions, as they would reflect a lower uncertainty about the RUL forecast.

Figure 6 shows the average width of the CIs provided at different levels of observed degradation, for the two application cases and for the two considered
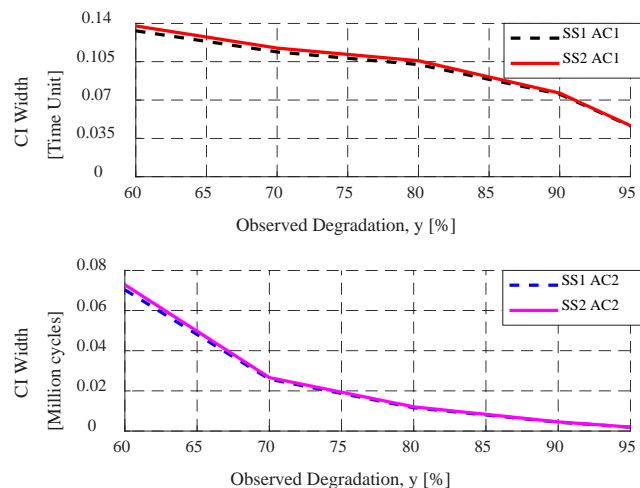
Figure 6. Comparison of the average CI width provided by the proposed algorithm for both application cases at different levels of observed degradation and through different selection strategies of the parameter σ_g.

selection strategies SS1 and SS2. As for the PI, the average width is obtained applying Equation (8) and averaging over the testing units. The results are satisfying, since SS2 seems to provide slightly wider CIs in both cases, but not in a significant way. As a matter of fact, the widths of the CIs are comparable. Nevertheless, independently on the chosen SS, the width of the CIs decreases as the observed degradation increases, so that the prognostic information becomes more valuable from the preventive maintenance point of view.

9. Conclusions

In this paper, a similarity-based data-driven prognostic algorithm for the estimation of the RUL of a unit is proposed. It is based on the exploitation of run-to-failure data of a representative set of units of the system/component under analysis, referred to as reference library. This allows one to implicitly introduce some knowledge about the future loading and operational conditions that the test unit will face in the rest of its life, mitigating the effect of the future uncertainty on the final prediction.

The core of the contribution is the application of a possibilistic framework, namely the RFV approach, for the representation and propagation of different crucial sources of uncertainty in PHM: the already cited future uncertainty; measurement uncertainty, whose role is particularly relevant in data-driven applications, since often the data are the results of measurement processes; epistemic uncertainty which arises by accounting for personal and experts' beliefs about model parameters.

Applying the mathematics of RFV, it is possible to evaluate the unit RUL in terms of RFV and extract the desired confidence interval. The results obtained for two real application cases have shown high prognostic performances of the proposed algorithm. In particular, a fundamental result is the high level of per-

formances achieved already at intermediate life stages (more than 95% of correct predictions when the degradation is equal to 75%), highlighting the ability of the algorithm to provide valuable results from the predictive maintenance point of view.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Luo, J., Pattipati, K., Qiao, L. and Chigusa, S. (2008) Model-Based Prognostic Techniques Applied to a Suspension System. *IEEE Transactions on Systems, Man, and Cybernetics Part A*: *Systems and Humans*, **38**, 1156-1168.
https://doi.org/10.1109/TSMCA.2008.2001055

[2] Gomes, J.P.P., Leao, B.P., Vianna, W.O.L., Galvao, R.K.H. and Yoneyama, T. (2012) Failure Prognostics of a Hydraulic Pump Using Kalman Filter. *Proceedings of the* 2012 *Annual Conference of the Prognostics and Health Management Society*, Minneapolis, 23-27 September 2012, 1-5.

[3] Pitt, M. and Shephard, N. (1999). Filtering via Simulation: Auxiliary Particle Filters. *Journal of the American Statistical Association*, **94**, 590-599.
https://doi.org/10.1080/01621459.1999.10474153

[4] Crisan, D. and Doucet, A. (2002) A Survey of Convergence Results on Particle Filtering Methods for Practitioners. *IEEE Transactions on Signal Processing*, **50**, 736-746. https://doi.org/10.1109/78.984773

[5] Orchard, M.E. and Vachtsevanos, G.J. (2009) A Particle-Filtering Approach for On-Line Fault Diagnosis and Failure Prognosis. *Transactions of the Institute of Measurement and Control*, **31**, 221-246. https://doi.org/10.1177/0142331208092026

[6] Heskes, T. (1997) Practical Confidence and Prediction Intervals. *Advances in Neural Information Processing Systems*, **9**, 466-472.

[7] Raviv, Y. and Intrator, N. (1996) Bootstrapping with Noise: An Effective Regularization Technique. *Connection Science*, **8**, 355-372.
https://doi.org/10.1080/095400996116811

[8] Baraldi, P., Mangili, F. and Zio, E. (2013) Investigation of Uncertainty Treatment Capability of Model-Based and Data-Driven Prognostic Methods Using Simulated Data. *Reliability Engineering and System Safety*, **112**, 94-108.
https://doi.org/10.1016/j.ress.2012.12.004

[9] Tang, L., Kacprzynski, G.J., Goebel, K. and Vachtsevanos, G. (2009) Methodologies for Uncertainty Management in Prognostics. *Proceedings of the* 2009 *IEEE Aerospace Conference*, Big Sky, 7-14 March 2009, 1-12.
https://doi.org/10.1109/AERO.2009.4839668

[10] Sankararaman, S. and Goebel, K. (2014) Uncertainty in Prognostics and Health Management: An Overview. *Proceedings of the* 2*nd European Conference of the Prognostics and Health Management Society*, Nantes, 8-10 July 2014, 1-11.

[11] Gu, J., Barker, D. and Pecht, M. (2007) Uncertainty Assessment of Prognostics of Electronics Subject to Random Vibration. *Proceedings of AAAI Fall Symposium on Artificial Intelligence for Prognostics*, Arlington, 9-11 November 2007, 50-57.

[12] Sankararaman, S. and Goebel, K. (2013) Why Is the Remaining Useful Life Prediction Uncertain. *Proceedings of the Annual Conference of the Prognostics and*

Health Management Society, New Orleans, 14-17 October 2013, 1-13.

[13] Leone, G., Cristaldi, L. and Turrin, S. (2017) A Data-Driven Prognostic Approach Based on Statistical Similarity: An Application to Industrial Circuit Breakers. *Measurement*, **108**, 163-170. https://doi.org/10.1016/j.measurement.2017.02.017

[14] Zio, E. and Di Maio, F. (2009) A Data-Driven Fuzzy Approach for Predicting the Remaining Useful Life in Dynamic Failure Scenarios of a Nuclear Power Plant. *Reliability Engineering and System Safety*, **95**, 49-57. https://doi.org/10.1016/j.ress.2009.08.001

[15] Guépié, B.K. and Lecoeuche, S. (2015) Similarity-Based Residual Useful Life Prediction for Partially Unknown Cycle Varying Degradation. *Proceedings of the* 2015 *IEEE Conference on Prognostics and Health Management* (*PHM*), Austin, 22-25 June 2015, 1-7. https://doi.org/10.1109/ICPHM.2015.7245054

[16] Wang, T., Yu, J., Siegel, D. and Lee, J. (2008) A Similarity-Based Prognostics Approach for Remaining Useful Life Estimation of Engineered Systems. *Proceedings of the* 2008 *International Conference on Prognostics and Health Management*, Denver, 6-9 October 2008, 1-6. https://doi.org/10.1109/PHM.2008.4711421

[17] Ferrero, A. and Salicone, S. (2009) The Construction of Random-Fuzzy Variables from the Available Relevant Metrological Information. *IEEE Transactions on Instrumentation and Measurement*, **58**, 365-374. https://doi.org/10.1109/TIM.2008.928873

[18] Ferrero, A., Prioli, M. and Salicone, S. (2014) The Construction of Joint Possibility Distributions of Random Contributions to Uncertainty. *IEEE Transactions on Instrumentation and Measurement*, **63**, 80-88. https://doi.org/10.1109/TIM.2013.2273596

[19] Ferrero, A., Prioli, M. and Salicone, S. (2013) Processing Dependent Systematic Contributions to Measurement Uncertainty. *IEEE Transactions on Instrumentation and Measurement*, **62**, 720-731. https://doi.org/10.1109/TIM.2013.2240097

[20] Salicone, S. and Prioli, M. (2018) Measurement Uncertainty within the Theory of Evidence. Springer Series in Measurement Science and Technology. Springer, New York. https://doi.org/10.1007/978-3-319-74139-0

[21] Klir, G.J. and Parviz, B. (1992) Probability-Possibility Transformations: A Comparison. *International Journal of General Systems*, **21**, 291-310. https://doi.org/10.1080/03081079208945083

[22] Cristaldi, L., Ferrero, A., Leone, G. and Salicone, S. (2018) A Possibilistic Approach for Measurement Uncertainty Propagation in Prognostics and Health Management. *Proceedings of the IEEE International Instrumentation & Measurement Technology Conference*, Houston, 14-17 May 2018, 1-6. https://doi.org/10.1109/I2MTC.2018.8409739

[23] Lu, C.J. and Meeker, W.Q. (1993) Using Degradation Measures to Estimate a Time-to-Failure Distribution. *Technometrics*, **35**, 161-174. https://doi.org/10.1080/00401706.1993.10485038