Genetic algorithms in the framework of Dempster-Shafer Theory of Evidence for maintenance optimization problems

Michele Compare, Enrico Zio

Genetic algorithms in the framework of Dempster-Shafer Theory of Evidence for maintenance optimization problems

Michele Compare[a], Enrico Zio*[a,b], senior member, IEEE

[a]Energy Department, Politecnico di Milano, Milan, Italy
[b]Chair on Systems Science and the Energetic Challenge, European Foundation for New Energy-Electricité de France, Ecole Centrale Paris and Supelec, France
enrico.zio@polimi.it; enrico.zio@ecp.fr

Abstract – The aim of this paper is to address the maintenance optimization problem when the maintenance models encode stochastic processes, which rely on parameters that are imprecisely known, and when these parameters are only determined through information elicited from experts. A genetic algorithms (GA)-based technique is proposed to deal with such uncertainty setting; this approach requires addressing three main issues: i) the representation of the uncertainty in the parameters and its propagation onto the fitness values; ii) the development of a ranking method to sort the obtained uncertain fitness values, in case of single-objective optimization; and iii) the definition of Pareto dominance, for multi-objective optimization problems. A known hybrid Monte Carlo - Dempster-Shafer Theory of Evidence method is used to address the first issue, whereas two novel approaches are developed for the second and third issues. For verification, a practical case study is considered concerning the optimization of maintenance for the nozzle system of a turbine in the Oil & Gas industry.

Index Terms – Evidence theory, genetic algorithms, pareto dominance, maintenance optimization.

Acronyms and Abbreviations

DSTE	Dempster-Shafer Theory of Evidence
PT	Possibility Theory
CBM	Condition Based Maintenance
GA	Genetic Algorithm
BPA	Basic Probability Assignment
MC	Monte Carlo
RAMS	Reliability, Availability, Maintainability, and Safety
CDF	Cumulative distribution Function

Notation

$F_{Y^j}(y^j;\boldsymbol{\theta}^j)$	Distribution of the uncertain variable Y^j
$W^t=\{W_1,...,W_H\}$	Offspring population at step t
$X^t=\{X_1,...,X_H\}$	GA population at step t
$\underline{Y}=Y^j,...,Y^j,...,Y^k$	Vector of uncertain variables Y^j
$\underline{Z}=(Z^1,...,Z^o)$	Output variables of interest
$\boldsymbol{\theta}^j=\{\theta^{j,1},...,\theta^{j,p},....,\theta^{j,M^j}\}$	Vector of the M^j uncertain parameters of distribution F_{Y^j}
$\underline{\Xi}=\Xi^1,...,\Xi^q,...,\Xi^Q$	Output vector of measures of $\underline{Z}$
$\underline{\Xi}_h=\left(\Xi_h^1,...,\Xi_h^q,....,\Xi_h^Q\right)$	Output vector $\underline{\Xi}$ computed in correspondence of the h-th member of the current population

1 Introduction

In the last few decades, the economic relevance of maintenance has grown in all sectors of industry. Nowadays, establishing an optimal maintenance policy is a key factor for safety, production, and asset management; and is fundamental to guarantee competitiveness. Given the dimension, complexity, and economic relevance of the problem, maintenance optimization must be supported by modeling [1]. The behavior of the failure-degradation processes affecting the equipment, their impact on the system functionalities, the effects of the (possibly imperfect) maintenance actions on such processes, the maintenance decision rules, etc. are all examples of the facets that a reliable, precise, and robust maintenance model is expected to encode. See [1], [2], for surveys on and issues in maintenance modeling.

Generally speaking, the more complex the maintenance model, the larger the number of parameters it relies on. These parameters may be poorly known in real applications, due to a lack of real data collected during operation or properly designed tests, especially when dealing with new products and new technology. In these cases, the main source of information to estimate the model parameters becomes the experts' judgment.

From these considerations, it emerges that maintenance models encode uncertainty. In the practice of reliability and maintenance engineering, uncertainty is usually divided into two constituent parts: i) the uncertainty due to the inherent variability of the phenomena of interest, which is referred to as aleatory uncertainty; and ii) the uncertainty due to lack of precise knowledge of quantities or processes of the system or the environment, which is usually named epistemic uncertainty [3]. The correct processing of both uncertainty types in the maintenance models is crucial [4], as witnessed by the large amount of literature produced on this topic. These works tackle the maintenance performance assessment issue in the presence of uncertainty from different perspectives.

- Probability distributions have been used to represent the uncertainty in the parameters of the stochastic models of the degradation mechanisms (e.g., [5], [6]). However, the capability of the probabilistic approach to represent the epistemic uncertainty associated with the expert judgments has been questioned [7]-[8].

- Fuzzy Logic ([9]) has been applied to address the cases in which the lack of knowledge concerns both the degradation model of a component and its parameters (e.g., [10]-[13]).

- Theoretical and computational methods have been developed to incorporate the imprecise parameters (e.g., represented by interval probabilities [14]-[17], fuzzy sets [18], possibility distributions [19]-[20], and probability assignments [21]-[22]) into Markov or semi-Markov models.

In spite of the interest in the correct representation and treatment of the aleatory and epistemic uncertainties in the maintenance models, it seems fair to say that the problem of how to use these models to optimize maintenance has not received the same attention. In practice, the maintenance optimization problem based on stochastic models (aleatory uncertainty), which rely on epistemically uncertain parameters, can be framed as a multi-objective optimization problem with uncertain objective functions (e.g., unavailability, cost, etc.). This problem is undoubtedly difficult [23], and a few approaches have been proposed in the literature to effectively tackle it. These approaches consider different frameworks for uncertainty representation: probability distributions [23]-[24], fuzzy sets [25]-[26], and plausibility and belief functions [27].

Also, the authors of this work have proposed methodologies in the framework of Possibility Theory (PT) [19]-[20], and Dempster-Shafer Theory of Evidence (DSTE) [21]-[22] to represent and propagate the uncertainty in the maintenance models. But these works have left open the issue of how to optimize a CBM policy based on the output of such models, which are pairs of plausibility and belief functions.

In this context, the objective of the present work is to propose a solution to that open issue, which is based on an enhancement of the GA technique. The proposed approach provides the maintenance decision maker with a set of

optimal maintenance settings, from which he or she can select the preferred one. In details, a hybrid MC-DSTE method is used to propagate the epistemic and aleatory uncertainties onto the fitness values, which turn out to be affected by noise and epistemic uncertainty. Then, a technique is proposed to rank the obtained noisy and uncertain fitness values in case of single objective problems. Finally, a generalization of the Pareto dominance concept is given in the multi-objective setting, which allows defining the Pareto optimal set. This set is at the basis of additional decision criteria to guide the decision maker in selecting the preferred solution. An application of the methodology to a practical case study is proposed, which concerns the optimization of the maintenance of the nozzle system of a turbine for the oil & gas industry.

The remainder of the paper is organized as follows. Section 2 describes the uncertainty setting in which this work is positioned. The ranking criterion and the Pareto dominance definition are introduced in Section 3, together with a brief literature review on the GA with uncertain fitness. The proposed GA approach is applied to the case study presented in Section 4. Concluding remarks are given in Section 5.

2 Uncertainty modeling and propagation

Consider a model $\underline{Z}=g(\underline{Y})$, where $\underline{Z}=(Z^1,...,Z^o)$ is the vector of the O output variables of the model, and $g(\cdot)$ is the function that links the output $\underline{Z}$ and the k uncertain variables Y^j, $j=1, 2,\cdots, k$, of the input vector $\underline{Y}$. Aleatory uncertainty on these variables is described by probability distributions $F_{Y^j}(y^j;\boldsymbol{\theta}^j)$, where $\boldsymbol{\theta}^j=\{\theta^{j,1},...,\theta^{j,M^j}\}$ are vectors of M^j epistemically uncertain hyper-parameters, j=1, 2,…, k. The values of these parameters are elicited from experts, in the form of intervals of plausible values. This situation is typical in industry. For example, it has been investigated in flood risk analysis [28], in reliability engineering [29], in rock engineering [30], and in maintenance engineering [21], [22], to cite a few.

Specifically to our maintenance optimization problem, g is the model of the life evolution of the component of interest, described by random variables (Y^j, $j=1,\cdots,k$) such as the time of transition from a degradation state to another, the failure time, the repair duration, etc. For example, consider the discrete-state stochastic degradation model of Fig. 1, which would be a part of the overall maintenance model $g(\cdot)$. There are three stochastic transitions, whose firing times are distributed according to $F_{Y^j}(y^j;\boldsymbol{\theta}^j)$, j=1, 2, 3. Each of these distributions (e.g., Weibull) depends on the set of parameters $\boldsymbol{\theta}^j$, $j=1,...,k$ (e.g., scale and shape parameters), whose values are provided by experts in the form of intervals to account for the uncertainty in their values. This degradation model has been used in different contexts; for example, to optimize maintenance strategies [21], [22], in support to reliability analysis [31], [32], [33], and in prognostics applications [34], as it provides an estimation of the remaining useful life of the component.

In addition to the considerations above, it is our experience that in some cases (e.g., the degradation behaviour of some components of the turbines used in the oil & gas industry [35]), the experts of the maintenance engineering department have a relatively precise knowledge about both the stochastic process to be used to model the evolution of the degradation mechanisms, and the values of their parameters, too. In details, the choice of the process is usually justified by physical considerations; for example, the corrosion mechanism affecting the nozzles of gas turbines is often modeled as a discrete-state, continuous time process. The observed aging dynamics of these components leads to considering the transition rates from a degradation state to a more degraded state as increasing in time. The Weibull distribution provides a flexible tool to model such behavior. The knowledge of the model parameters, instead, comes from the outcomes of the statistical analysis previously performed on similar systems, and on qualitative considerations about the impact that some influencing factors (e.g., work load, location, etc.) would have had on the degradation mechanism behavior. However, the experts are able to estimate the boundaries of the intervals they suppose contain the true values of the parameters, only. This consideration justifies the use of the DSTE to describe such epistemic uncertainty.

The variables $Z^1,...,Z^o$ of the output vector $\underline{Z}$ are those relevant for assessing the performance of maintenance policies, e.g., the component downtime, the unavailability, the cost associated to the maintenance policy, etc. Given the stochastic character of the processes involved in the component life, we characterize the variability of the output variables by some measures $\underline{\Xi} = \Xi^1,...,\Xi^Q$ such as mean, percentiles, etc., Then, the maintenance policies are evaluated with respect to variables $\Xi^1,...,\Xi^Q$ such as the mean unavailability, the mean cost over a defined time horizon, etc. These variables are the set objectives of the optimization problem.

In this work, the hybrid MC-DSTE method proposed in [36], and further developed in [21] and [22], is adopted to propagate the uncertainty. Details of the method are given in the Appendix. For further practical and theoretical aspects, the interested readers can refer to [22], and [36]. Such a method generates multiple realizations of the uncertain parameters of the model, and computes summary measures $\underline{\Xi}$ representative of the quantities $\underline{Z}$. The uncertainty in the values of $\underline{\Xi}$ is described in terms of Belief and Plausibility measures (see the Appendix). Hence, in the end, we get a set of pairs $\{[\, Bel_{\Xi^1}(\xi^1)\,, Pl_{\Xi^1}(\xi^1)\,],\ldots,[\, Bel_{\Xi^Q}(\xi^Q)\,, Pl_{\Xi^Q}(\xi^Q)\,]\}$ (Fig. 1).

The choice of using the DSTE framework to represent and propagate the uncertainty in the imprecise information retrieved from the experts about the stochastic model parameters is not new in the Reliability, Availability, Maintainability, and Safety (RAMS) field. In fact, it has been pointed out (e.g., [37]-[39]) that, in industry, reliability studies assume that the probability values are precisely known. Indeed, this condition is rarely fulfilled, especially when the component reliabilities are inferred from databases, or when components fail rarely. In fact, in the former case, the problem of imprecision becomes critical, as incoherency and incompleteness of data very often affect the data collection (e.g., [40], [41]), whereas in the latter case failure data are necessarily poor (e.g., nuclear industry, aeronautic industry, etc.). As mentioned before, the imprecision associated with this situation cannot be suitably handled by probability theory, as it underlies an epistemic uncertainty. Thus, the theory of evidence proposes an interesting, suitable formalism to handle this type of uncertainty, especially for reliability engineers as it is rather close to the theory of probability in some ways [37]. This consideration also explains the increasing use of such theory in reliability engineering.

On the other hand, the DSTE has been challenged on particular aspects by some experts in the area of safety and reliability analyses (e.g., [42]-[43]), as it seems to suffer from major drawbacks such as the computational efforts and the difficulties that may be encountered when eliciting the probability masses from a number of experts [44]-[45].

For the sake of clarity, we stress the fact that the hybrid MC-DSTE method propagates both epistemic and aleatory uncertainties, and the output $\underline{\Xi}=(\Xi^1,...,\Xi^Q)$ encodes the epistemic uncertainty due to the parameters of the model *g* and the noise due to the finite sample of realization of the MC method. The aleatory uncertainty is summarized in the values of $\underline{\Xi}$.

Finally, notice that in Dempster's view [46], for any measurable set *A*, the pair $[\, Bel_{\Xi^q}(A), Pl_{\Xi^q}(A)\,]$ can be interpreted as lower and upper probabilities encoded by Basic Probability Assignment (BPA). Then, we can indicate by $\underline{F}_{\Xi^q}(\xi^q) = Bel_{\Xi^q}(]-\infty,\xi^q])$, and $\overline{F}_{\Xi^q}(\xi^q) = Pl_{\Xi^q}(]-\infty,\xi^q])$ as the lower, and upper cumulative distribution functions (CDF), respectively, such that $\forall \xi^q \in S \;\; \underline{F}_{\Xi^q}(\xi^q) \le F_{\Xi^q}(\xi^q) = P(\Xi^q \le \xi^q) \le \overline{F}_{\Xi^q}(\xi^q)$.

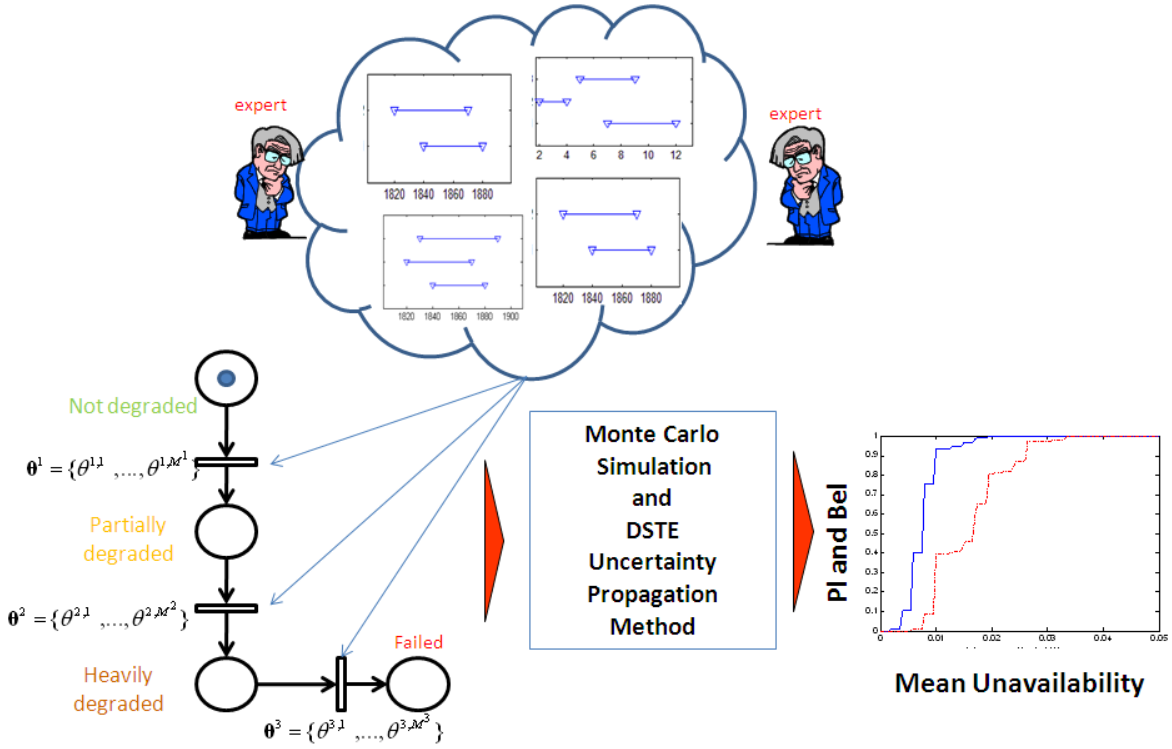

Fig. 1. Uncertainty propagation scheme.

3 Genetic Algorithms with noisy and uncertain objectives

Genetic Algorithms (GA) are now a well-established optimization tool, successfully used to address a broad variety of issues in several areas of engineering and life science [47]. In particular, GA have been used in RAMS, the discipline focus of this work, e.g., to optimize reliability allocation [48], [49], maintenance scheduling [50], risk based maintenance strategies [47], etc.

As the application of GA pervasively enters industrial contexts, and with more computational capability available, the complex realism of the optimization problems being solved by GA increases: objective functions are more complex, non-linear, and affected by uncertainty [24]. With respect to this latter aspect, on the one hand, processing of uncertainty is nowadays fundamental in industrial applications, for its impact on the decision making processes [4]. On the other hand, the issue of extending the applicability of GA to optimization problems in which the objective functions are affected by uncertainty is undoubtedly difficult [23].

A number of works have tackled optimization problems whose objective functions are expected values of random variables, i.e., affected by aleatory uncertainty. For instance, GA have been successfully used in reliability engineering to optimize plant design (e.g., [48], [49]) with respect to objective functions like system reliability or availability; such objective functions depend on the stochastic failure and repair behavior of the system components, and are evaluated as point values, typically expectations. Then, the application of GA to solve the optimization problem is straightforward.

On the contrary, the GA need to be modified when accounting for the uncertainty on the fitness values. For example, when the objective functions are computed by Monte Carlo (MC) simulation, then noise is introduced: two successive evaluations of the same solution return two different values of fitness [24]. Studies on the robustness of GA to noise have shown that satisfactory solutions of the optimization problem can be obtained also in the case of noise in the fitness values, although the GA evolution dynamics tend to slow down [24]. Research efforts have been mainly devoted to the development of methods for reducing noise and computing time [51], [52].

For example, a powerful integration of GA and MC simulation for optimizing the reliability design of complex plants has been proposed in [52]. To reduce computing time, the framework proposed in [52] exploits the fact that during the GA search the solutions (particularly the good ones) appear repeatedly in the population: if the values of fitness are stored, then statistical significance is obtained with a relatively low computational expenditure.

More recently, some approaches have been developed to solve both single-objective and multi-objective optimization problems in stochastic environments, where fitness values are affected not only by noise but also by epistemic uncertainty (e.g., [23], [27], [53]). The main issues to be addressed when extending GA to such problem settings are: i) the ranking of the solutions based on their fitness values, in the case of single objective optimization, and ii) the definition of Pareto dominance in the multi-objective case [27]. A few methods have been proposed in the literature (e.g., [23], [24], [53], [54]), based on the pair-wise comparison of the probability distributions that describe the uncertainty (i.e., noise and epistemic uncertainty, together) in the fitness values of pairs of solutions. The main limitation of these algorithms lies in the fact that epistemic uncertainty is treated exclusively within the framework of probability theory.

Extensions of GA to optimization problems in which epistemic uncertainty is not described within the probability theory framework have been proposed in the literature. For example, in [26], both objectives and constraints are described by fuzzy sets, and every member of the population belongs to each of these sets with a degree of membership. An aggregated fitness is calculated on the basis of these memberships, which is then used to rank the members of the evolving population.

A different approach has been developed in [27], where epistemic uncertainty is described in the framework of the DSTE [36], [46], [55]-[57]. There, the ranking procedure in the single objective optimization problem is based on the pair-wise comparison of the intervals that represent the uncertainty in the fitness values associated to the pair of solutions, to check whether the intervals overlap each other. The Pareto dominance definition is an extension to the multi-dimensional space of such ranking criterion.

Notice that both these latter works ([26] and [27]) give account to epistemic uncertainty, only; noise is not considered. Moreover, notice that, in the context of the present work, epistemic uncertainty refers to poor knowledge about the model parameters, which impacts the ranking of the solutions, even though two successive evaluations of the same chromosome give the same result.

3.1 Single Objective Genetic Algorithm in the DSTE framework

In this section, a general procedure for the GA is given as follows (further details can be found in [58], [59]).

Step 1: Initialization. Set t=1. Randomly generate H solutions to form the first population $X^{t=1} = \{X_1, ..., X_H\}$. In the specific case of the maintenance optimization problem, a solution is generally a vector of decision variables such as the time interval between two successive inspections, the type of maintenance action to be performed, etc.

Step 2: Fitness evaluation. Evaluate the fitness of the solutions in X^t. In the uncertainty setting considered in this work, performing this step requires running the uncertainty propagation procedure explained in the Appendix for every solution $X_h \in X^t$ to obtain the pair $\left[Bel_{\Xi_h^1}(\xi^1), Pl_{\Xi_h^1}(\xi^1) \right]$, for h=1,…, H.

Step 3. Breeding. Generate an offspring population $W^t = \{W_1, ..., W_H\}$ as follows.

I. **Selection**. Choose two solutions X_s and X_l from X^t. Usually, this choice is based on the fitness values (e.g., standard and hybrid selection, fit-fit or fit-weak selection and mating). Notice that the selection algorithm heavily influences the performance of the GA, which is usually evaluated in terms of effectiveness (i.e., the capability of finding the optimal value), and efficiency (i.e., the speed of convergence towards the optimal solution) [60]. Generally speaking, these two attributes may be either strictly connected to each other, or even conflicting, depending on the particular optimization problem at hand. Namely, preserving the genetic diversity (e.g., random or fit-weak selection algorithms) on one side favors the effectiveness of the algorithm, as it prevents the algorithm from attaining a non-

global local minima; on the other side, it may entail poorer efficiency. On the contrary, selection policies less disruptive of the genetic codes, such as the fit-fit or the hybrid roulette, improve the algorithm efficiency by favoring the fittest individuals, but this condition may lead to non-global local minima.

II. **Crossover.** Using a crossover operator, generate offspring, and add them to W^t.

III. **Mutation.** Mutate each solution $\{W_1,...,W_H\}$ with a predefined mutation rate.

IV. **Fitness assignment.** Evaluate the fitness value for each solution $\{W_1,...,W_H\}$.

V. **Replacement.** Based on the fitness values, select *H* solutions from W^t, and copy them to X^{t+1} (e.g., fittest individuals or weakest individuals policies). Again, the replacement policy influences the performance of the GA in terms of effectiveness and efficiency.

Step 4. If the stopping criterion is satisfied, terminate the search, and return to the current population; else, set $t=t+1$, and go back to Step 3.

From this procedure, it clearly emerges that performing steps 3.I and 3.V in the uncertainty setting considered in this work requires addressing the issue of developing a ranking criterion to establish whether a solution is better or worse than the others at reaching the uncertain objective.

3.2 Ranking of uncertain values

The uncertainty propagation procedure recalled in Section 2 yields the fitness associated to solution X_h, which is represented by the pair of measures $\underline{F}_{\Xi_h^q}(\bar{\xi}^q)$, $\overline{F}_{\Xi_h^q}(\bar{\xi}^q)$, for $q=Q=1$, and $h=1,...,H$. Then, to evolve the search towards improved solutions, one needs to develop a method for comparing the individuals of the population on the basis of such pairs. To this aim, a novel method is proposed in this work, which is based on that developed by the authors in [68] within the probability theory framework. This method is here briefly recalled.

Consider two possible solutions X_s and X_l, and assume that the uncertainty in the corresponding values of Ξ^q are expressed by the random variables and Ξ_l^q, respectively, for $q=Q=1$. Then, the distribution of the random variable $\Delta_{sl}^q = \Delta_{sl}^q(\Xi_s^q, \Xi_l^q) = \Xi_s^q - \Xi_s^q$ can be computed. The complement to 1 of its value in 0 (i.e., $1-F_{\Delta_{sl}^q}(0)$) gives the probability that Ξ_s^q is larger than Ξ_l^q. The relation order between solutions X_s and X_l is obtained by comparing $1-F_{\Delta_{sl}^q}(0)$ to a threshold range $[T_l, 1-T_l]$, symmetric around 0.5, and considering the following criteria.

- If $F_{\Delta_{sl}^q}(0) \le T_l$, then X_s is larger than X_l (i.e., X_s is worse than X_l in a minimization problem).
- If $1-T_l \le F_{\Delta_{sl}^q}(0)$, then X_l is larger than X_s (i.e., X_s is better than X_l in a minimization problem).
- If $T_l \le F_{\Delta_{sl}^q}(0) \le 1-T_l$, then X_s and X_l are equivalent.

In simple words, the relation order is defined between two solutions X_s and X_l when the decision maker judges as large enough the probability that solution X_s is better or worse than X_l (e.g., >0.7).

In this work, this procedure is adapted to the case in which uncertain quantities are represented by pairs of Belief and Plausibility measures. To do this adaptation, the pair-wise based ranking method proposed in this work imports, within the uncertainty setting described in Section 2, the definition of the difference between uncertain numbers of interval arithmetic [61].

Definition 1: Consider two interval variables $s=[a, b]$, and $l=[c, d]$; then the difference $z=s-l = [a-d, b-c]$.

That is, the leftmost value of *z* is given by the difference of the smallest value of the minuend *s* and the largest value of the subtrahend *l*, whereas the rightmost value of *z* is given by the difference between the largest value of *s* and the smallest value of *l*.

In this work, this concept is applied to the case where the bounds of the intervals in Definition 1 are CDFs instead of real points. In details, consider two uncertain variables Ξ_s^q, and Ξ_l^q described by the pairs $\left[Bel_{\Xi_s^q}\left(]-\infty,\xi^q]\right), Pl_{\Xi_s^q}\left(]-\infty,\xi^q]\right)\right]$, and $\left[Bel_{\Xi_l^q}\left(]-\infty,\xi^q]\right), Pl_{\Xi_l^q}\left(]-\infty,\xi^q]\right)\right]$, respectively, which are interpreted as the pairs $\left[\underline{F}_{\Xi_s^q}\left(\xi^q\right),\overline{F}_{\Xi_s^q}\left(\xi^q\right)\right]$, and $\left[\underline{F}_{\Xi_l^q}\left(\xi^q\right),\overline{F}_{\Xi_l^q}\left(\xi^q\right)\right]$ of the lower, and upper CDFs encoded by the corresponding BPA [46]. In this respect, notice that the lower, and the upper CDFs are the rightmost, and the leftmost bounds, respectively, although they are represented as the leftmost, and rightmost, respectively, according to [46]. This representation is due to the fact that, in Dempster's view of the evidence theory, lower and upper bounds are defined with respect to the ordinate axis.

Bearing this difference in mind, if we treat these bracketing CDFs similarly to the points *a*, *b*, *c*, and *d* of the intervals in Definition 1, then we get that the upper bound of the difference Ξ_s^q - Ξ_l^q is $\overline{\Delta}_{sl}^q\left(\xi^q\right)=\overline{F}_{\Xi_s^q}\left(\xi^q\right)-\underline{F}_{\Xi_l^q}\left(\xi^q\right)$, whereas the lower bound is $\underline{\Delta}_{sl}^q\left(\xi^q\right)=\underline{F}_{\Xi_s^q}\left(\xi^q\right)-\overline{F}_{\Xi_l^q}\left(\xi^q\right)$.

To extend the applicability of Definition 1 to the setting considered in this work, we need to prove that $\overline{\Delta}_{sl}^q$, and $\underline{\Delta}_{sl}^q$ are actually the upper (i.e., leftmost), and lower (i.e., rightmost) bounds, respectively, of the difference Ξ_s^q - Ξ_l^q. That is, we have to prove that it is not possible that the difference between any distribution among those bounded by $\left[\underline{F}_{\Xi_s^q}\left(\xi^q\right),\overline{F}_{\Xi_s^q}\left(\xi^q\right)\right]$ and any distribution bounded by $\left[\underline{F}_{\Xi_l^q}\left(\xi^q\right),\overline{F}_{\Xi_l^q}\left(\xi^q\right)\right]$ gives rise to a distribution not bounded by $\left[\underline{\Delta}_{sl}\left(\xi^q\right),\overline{\Delta}_{sl}\left(\xi^q\right)\right]$.

We here give an intuitive proof. Recall that the CDF $F_{\Delta_{sl}^q}(\delta^q)$ of the difference Δ_{sl}^q of two *s*-independent random variables Ξ_s^q and Ξ_l^q is given by the convolution of the probability density function (PDF) $f_{\Xi_l^q}(\xi^q)$ and the CDF $F_{\Xi_s^q}(-\delta^q)$; this convolution is defined as the integral that expresses the amount of overlap of $F_{\Xi_s^q}(\xi^q)$ over $f_{\Xi_l^q}(\xi^q)$, as δ^q pulls backward (i.e., from right to left) the first function over the second [67]. Intuitively, the larger the distance between these two curves, the larger the value of δ^q at which the convolution integral starts increasing. This consideration tells us that, if we subtract from $\overline{F}_{\Xi_s^q}\left(\xi^q\right)$ any distribution in $\left[\underline{F}_{\Xi_l^q}\left(\xi^q\right),\overline{F}_{\Xi_l^q}\left(\xi^q\right)\right]$ which is shifted on the left with respect to $\underline{F}_{\Xi_l^q}\left(\xi^q\right)$, then we get a distribution shifted on the right with respect to $\overline{\Delta}_{sl}^q$ (i.e., the distribution we want to prove being the furthest left). With the necessary modifications, if we subtract from $\underline{F}_{\Xi_l^q}\left(\xi^q\right)$ any distribution in $\left[\underline{F}_{\Xi_l^q}\left(\xi^q\right),\overline{F}_{\Xi_l^q}\left(\xi^q\right)\right]$ which is positioned on the right with respect to $\overline{F}_{\Xi_l^q}\left(\xi^q\right)$, then we get a distribution positioned on the left with respect to $\underline{\Delta}_{sl}^q$ (i.e., the distribution we want to prove being the furthest right).

For example, assume that two solutions X_s, and X_l are associated to values of Ξ^q uniformly distributed in [3,5], and [-1,1], respectively (Fig. 2 (a), and (b)); then, the convolution of $f_{\Xi_l^q}(\xi^q)$ and $F_{\Xi_s^q}(-\delta^q)$ is zero up to δ^q =3-1=2 (Fig. 2 (c)). When the distance between the distributions increases, then the shift needed to have a function

overlap increases. For example, if Ξ_s^q, and Ξ_l^q are uniformly distributed in [4,6], and [-2,0], respectively (Fig. 2 (d), and (e)), then the convolution is zero up to δ^q =4-0=4 (Fig. 2(e)).

For the sake of clarity, Fig. 3 shows the amount of overlap (filled area) of the distributions $F_{\Xi_s^q}(\xi^q)$, over $f_{\Xi_l^q}(\xi^q)$, in Fig. 2 (a), and (b), respectively, when $F_{\Xi_s^q}(\xi^q)$ is shifted backward by δ^q =3.

Moreover, notice that any change, even slight or partial, in the shape of any $F_{\Xi_s^q}(\xi^q)$, inevitably implies a change in the shape of the convolution integral in the same direction (bold lines in Fig. 2 (d) and (f)). This consideration entails that, even if we consider a pair of distributions $F_{\Xi_s^q}(\xi^q)$ and $F_{\Xi_l^q}(\xi^q)$, very close to the corresponding bounding CDFs, their difference is always contained in $\left[\underline{\Delta}_{sl}\left(\xi^q\right), \overline{\Delta}_{sl}\left(\xi^q\right)\right]$.

Thus, the upper bound $\overline{\Delta}_{sl}^q$ of the difference between variables Ξ_s^q and Ξ_l^q is given by $Pl_{\Xi_s^q} - Bel_{\Xi_l^q}$, which is the pair among all possible distributions bracketed by $[\,Bel_{\Xi_s^q}(]-\infty,\xi^q]), Pl_{\Xi_s^q}(]-\infty,\xi^q])\,]$ and $\left[Bel_{\Xi_l^q}(]-\infty,\xi^q]), Pl_{\Xi_l^q}(]-\infty,\xi^q])\right]$ with the smallest distance between each other, whereas the lower bound $\underline{\Delta}_{sl}^q$ is given by $Bel_{\Xi_s^q} - Pl_{\Xi_l^q}$, which is the pair of distributions with the largest distance.

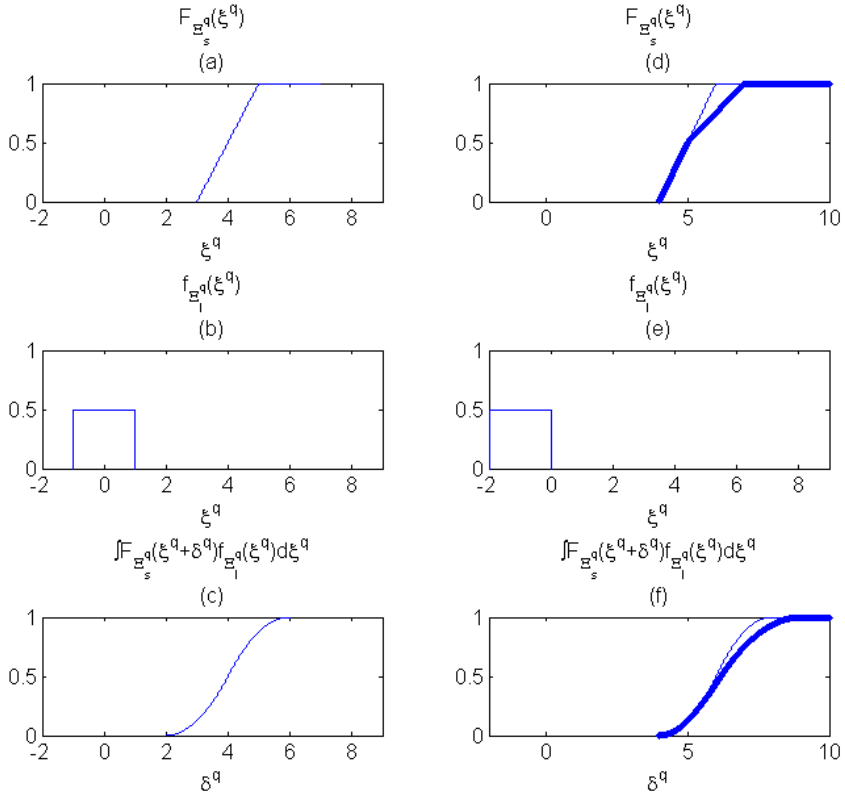

Fig. 2. Subtraction of probability distributions.

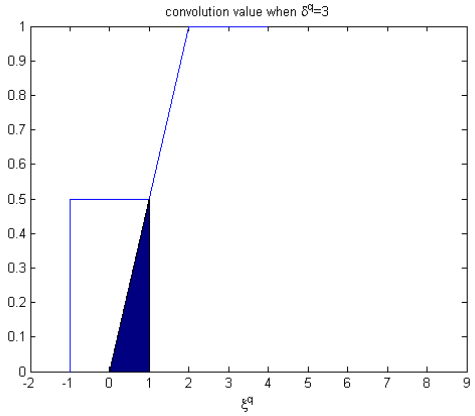

Fig. 3. Convolution in correspondence of δ=-3.

The proposed algorithm is based on the following steps.

1. Compute two distributions:
   $\overline{\Delta}_{sl}^{q}(\xi^q) = Pl_{\Xi_s^q}(]-\infty,\xi^q]) - Bel_{\Xi_l^q}(]-\infty,\xi^q])$, and
   $\underline{\Delta}_{sl}^{q}(\xi^q) = Bel_{\Xi_s^q}(]-\infty,\xi^q]) - Pl_{\Xi_l^q}(]-\infty,\xi^q])$.
2. The relation order between variables Ξ_s^q and Ξ_l^q is established on the basis of the following criteria.
   i. If $\overline{\Delta}_{sl}^{q}(0) \le T_l$, then variable Ξ_s^q is larger than Ξ_l^q.
   ii. If $\underline{\Delta}_{sl}^{q}(0) \ge 1 - T_l$, then variable Ξ_l^q is larger than Ξ_s^q.
   iii. Solutions X_s and X_l are equally important in the other cases.

T_l is the lower threshold defined above (e.g., T_l =0.3).

To sum up, if $\overline{\Delta}_{sl}^{q}(0) \le T_l$, then the probability that Ξ_s^q is larger than Ξ_l^q lies in the interval [1-T_l ,1], whereas the probability of the opposite case is a value between 0 and T_l. In this situation, in which we are confident on the relevance of one solution with respect to the other, it is reasonable to decide that $\Xi_s^q > \Xi_l^q$.

Notice that, from an intuitive point of view, a large overlap of the pairs of curves corresponding to Ξ_s^q and Ξ_l^q determines their equivalence, whereas the situation in which either $Pl_{\Xi_s^q}$ and $Bel_{\Xi_l^q}$, or $Pl_{\Xi_l^q}$ and $Bel_{\Xi_s^q}$, are far away from each other implies that one solution is better than the other. Thus, the number of solutions that will be classified as equivalent is generally larger than what we would get if the uncertainty in the fitness were described by a single distribution. However, this situation does not necessarily lead to a difficulty with distinguishing between similar items.

In fact, in practical RAMS applications, there are groups of components whose fitness values (e.g., reliabilitiy, importance measures) are similar to each other, but noticeably different from those of the other groups. For example, the Authors in [64] showed practical case studies in which the system components form clusters with similar values of Birnbaum Importance Measure.

Finally notice also that, as pointed out in [64], there may be cases in which the pair-wise comparisons of three generic random variables leads to $\Xi_s^q > \Xi_l^q$, and $\Xi_l^q > \Xi_j^q$, but $\Xi_j^q > \Xi_s^q$. This is a contradictory ranking, as the transitive property does not hold. However, it has been proven in [64] that by setting T_l smaller than 1/3, such

contradictory ranking is avoided, and at worst it can happen that $\Xi_s^q > \Xi_l^q$, $\Xi_l^q > \Xi_j^q$, and $\Xi_j^q = \Xi_s^q$. In this case, the three uncertain variables are considered equivalent.

The problem of the 'contradictory' ranking, which arises when sorting algorithms are based on pair-wise comparisons of probability distributions, did not emerge in the works of the literature that propose extensions of GA to treat noisy fitness values (e.g., [24], [23], [53]). This situation is due to the fact that assigning different ranking positions to solutions with equal fitness values does not significantly affect the effectiveness of the GA search for the optimal solution; rather, the GA efficiency (i.e., speed of convergence) may be weakened. For example, assume that the fit-fit approach is considered in the reproduction phase, and that there are *n* solutions with equal fitness values. When we sort them in the corresponding ranking positions *i*, *i*+1,…, *i*+*n*-1, each solution occupies a rank, which depends on the sorting algorithm, or even on the particular run of the algorithm (e.g., the Quicksort algorithm may randomly choose the pivot element [65]). Now, the fit-fit algorithm selects and mates members of these *n* solutions. This is a locally hybrid reproduction approach, which is between the fit-fit and random selection approaches, in the sense that, for those *n* positions, and at most the two neighborhoods in positions *i*-1 and *i*+*n*, there is a random facet behavior entering the selection of the parents. This may be even beneficial for GA, as it combines the speed of the fit-fit technique with the capability of preserving genetic diversity, typical of the random selection method (see [58] for references). However, the systematic study to assess the impact that such local-hybridization of the selection algorithm has on efficiency and effectiveness is outside the scope of this work.

With regards to the replacement phase, similar considerations can be done for both the fittest and weakest individuals; considering equal solutions as if they were different steers the replacement approach towards a hybrid method.

Particular care should be paid to the choice of the final solution. In fact, the fitness values of the solutions are affected by uncertainty; then, when these fitness values are sorted on the basis of the proposed method,the contradictory ranking problem may arise. Thus, equivalent solutions are associated to different ranking positions. To address this issue, a sorting algorithm has been proposed in [64], which gives back the same ranking positions to all the equivalent solutions. In details, the final output of the GA comprises the last population $\underline{X}^t = \{X_1,...,X_H\}$, and a vector $\underline{R} = \{r_1,...,r_H\}$, whose elements are the ranking positions of the solutions in $\underline{X}^t$. The same rank value may appear more than once in vector $\underline{R} = \{r_1,...,r_H\}$, depending on the number of equivalent solutions. In particular, we assume that the best solutions are those corresponding to $r = 1$.

3.3 Multi-objective optimization

Generally speaking, addressing multi-objective optimization problems requires some modifications to the procedure described in Section 3. In particular, the changes concern Steps 3.I and 3.V, where solutions are compared and ranked on the basis of their performance in achieving the multiple objectives. To this aim, the concept of Pareto dominance is introduced in traditional multi-objective GA (i.e., for which objectives are not affected by uncertainty or noise): assuming that all objectives $\Xi^1,...,\Xi^q,...,\Xi^Q$ are to be minimized, a feasible solution X_s is said to dominate another feasible solution X_l ($X_s \succ X_l$), iff $\Xi_s^q \le \Xi_l^q$ for all q=1,…,Q, and $\Xi_s^q < \Xi_l^q$ for at least one q=1,…,Q [52], [59].

A solution is said to be Pareto optimal if it is not dominated by any other solution in the solution space. A Pareto optimal solution cannot be improved with respect to any objective without worsening at least one other objective.

The set of all feasible non-dominated solutions in $\underline{X}$ is referred to as the Pareto optimal set, and for a given Pareto optimal set the corresponding objective function values in the objective space form the Pareto front.

The Pareto dominance concept is used at Step 3.I of the procedure in Section 3 in one of the following three typical dominance-based ranking methods:

- dominance rank, based on assessing the number of individuals an individual is dominated by;
- dominance depth, based on assessing which dominance front an individual belongs to; or
- dominance count, based on assessing how many individuals an individual dominates.

Step 3.V is modified by introducing an archive of vectors, each one constituted by a non dominated solution and corresponding fitness values. This archive represents the current Pareto optimal set, which is dynamically updated at the end of each generation. The fitness values of non-dominated individuals in the current population are compared with those already stored in the archive, and the following archival rules are implemented.

- If the new individual dominates existing members of the archive, then those dominated members are removed, and a new one is added.
- If the new individual is dominated by any member of the archive, it is not stored.
- If the new individual neither dominates nor is dominated by any member of the archive, then check the following.
    - If the archive is not full, then the new individual is stored.
    - If the archive is full, then the new individual replaces the most similar one in the archive. An appropriate concept of distance is that of the Euclidean distance based on the values of the fitness values of the chromosomes normalized to the respective mean values in the archive.

3.3.1 Pareto dominance definition

In this section, we extend the definitions given in Section 3.3 to the case in which the fitness values associated to every solution h=1,…, H are uncertain and noisy, and thus represented by a set of pairs $\left[Bel_{\Xi_h^1}(\xi^1), Pl_{\Xi_h^1}(\xi^1) \right]$, …, $\left[Bel_{\Xi_h^Q}(\xi^Q), Pl_{\Xi_h^Q}(\xi^Q) \right]$.

A first consideration concerns the fact that the simple extension of the ranking criterion discussed in Section 3 to multiple objectives is not suitable, because this would lead to an even more insidious problem of contradictory ranking. In details, consider the following definition of Pareto dominance.

Definition 2: $X_s \succ X_l$ iff $\overline{\Delta}_{sl}^q > T_l$ for all q=1,…,Q, and $\underline{\Delta}_{sl}^q(0) \geq 1 - T_l$ for at least one q=1,…,Q.

In simple words, $X_s \succ X_l$ when Ξ_l^q is at most equal to Ξ_s^q for all q=1,…,Q, and there exists at least a q for which $\Xi_l^q > \Xi_s^q$.

This definition is faulty. To prove this definition is faulty, for the sake of simplicity, we can refer to the case of Q=2. On one side, one may have $\underline{\Delta}_{sl}^1(0) \geq 1 - T_l$ and $T_l \leq \overline{\Delta}_{sl}^2 \leq 1 - T_l$, which leads us to conclude that $X_s \succ X_l$. On the other side, one may have also that $\underline{\Delta}_{sl}^1(0) \geq 1 - T_l$, $\underline{\Delta}_{lj}^1(0) \geq 1 - T_l$, and $\underline{\Delta}_{sj}^1(0) > T_l$ (i.e., contradictory ranking in the first objective). Thus Ξ_l^q is equivalent to Ξ_s^q, with respect to objective 1, and thus X_s does not dominate X_l .

To overcome this issue, the sorting algorithm shown in [64] is first applied to every objective q= 1,…,Q. This algorithm exploits the ranking criterion described above, and assigns the same ranking position to all the solutions that are equivalent with respect to objective q (for example, if H=4, and the second and third solutions are equivalent, then the final ranking is 1, 2, 2, 4). Then, the following definition of Pareto dominance is introduced to identify the Pareto front.

Definition 3: $X_s \succ X_l$ iff $r_s^q \le r_l^q$ for all q=1,…,Q, and $r_s^q < r_l^q$ for at least one q=1,…,Q.

r_s^q, and r_l^q are the ranking positions of the solutions X_s, and X_l with reference to objective q, respectively.

This definition of Pareto dominance is applied both to the dominance rank criterion at Step 3.I of the procedure in Section 3, and at Step 3.5 to update the archive of non-dominated solutions. In this respect, the size of this archive is here set large enough to avoid its filling. In fact, managing this situation calls for the development of the concept of distance between fitness values, which will be faced in future works.

4 Case study

The case study considered in this work concerns the nozzle system of a turbine installed in an Oil & Gas plant, which is made up of N=22 similar nozzles. These nozzles are affected by a number of degradation mechanisms such as oxidation, erosion, cracking, etc. This work focuses on erosion, whose stochastic behavior is modeled by a continuous-time discrete-state transport process, which monotonously evolves within four states, d_i, i=0,…, 3 (Fig. 4). The stochastic transition time T_i from d_{i-1} to d_i, is exponentially distributed, with mean $\overline{T}_i$, for any i= 1, 2, 3. That is,

$$F_{T_i}(t_i) = 1 - e^{-\frac{t_i}{\overline{T}_i}} \quad i=1, 2, 3.$$

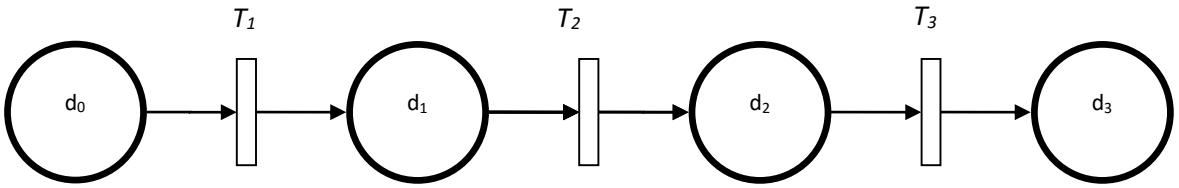

Fig. 4. four-states degradation process.

The values of the mean transition times $\overline{T}_i$ are not precisely known, and only understood via expert judgment. We assume that there are two experts who provide estimations of $\overline{T}_i$ i=1, 2, 3 on the basis of their experience. According to the approach discussed in [22], each expert gives the triplet of intervals which he or she believes contain the three unknown values $\overline{T}_i$, i=1, 2, 3.

In relation to this process, recall that the rate of the exponential distribution is by definition the inverse of the mean transition time, before which almost 65% (i.e., 63.2%) of the components of a homogeneous population have experienced a transition. Then, the expert is expected to have knowledge about the mean transition time.

Table reports the estimations provided by the experts. For example, expert 1 believes that the mean time to have a transition from d_0 to d_1 is between 10 and 11 years. Expert 2 is more optimistic and precise, believing that this transition does not occur before 10.8 years, and cannot occur after more than 11.2 years.

Table I

Values provided by experts

$\overline{T}_1$	[10,11]	[10.8,11.2]
$\overline{T}_2$	[3,3.5]	[2.9,3.1]

$\overline{T}_3$	[1.5,2]	[1.9,2.1]

A Condition-Based Maintenance (CBM) approach is applied to the nozzle system under study. Turbine efficiency is continuously monitored by processing the information provided by sensors which trace physical variables such as pressure, temperature, etc. When the efficiency value drops below a given threshold T_E, then the nozzle system is replaced. Replacement makes the system unavailable for U_R=4 days. The cost of the consequent business interruption is given by the product of the duration of the unavailability period times the annual income *I*, which is defined as the income corresponding to one turbine working continuously at full capacity for one year. In this paper, *I*= 500K€ (i.e., thousand euro). Thus, the total cost C_R associated to a replacement action upon the achievement of T_E is the sum of the business interruption cost due to system unavailability, and the cost C_S=25K€ of replacing the nozzle system.

Notice that the choice of replacing the nozzle system in the case of threshold T_E achievement is justified by the maintenance operators' lack of knowledge about the degradation state of the nozzles, and thus about the time required to fix it. Thus, on the basis of their experience, they assume that it is convenient to replace the nozzle system when efficiency loss reaches T_E.

The nozzle system is also periodically inspected, with period *II*. Every inspection is performed by one maintenance operator, who takes T_{insp}=1.5 days for carrying out the machine disassembling and re-assembling operations necessary to check the health state of the nozzles. Obviously, larger values of *II* steer the policy towards a full exploitation of the components, and avoid ineffective machine stops. On the contrary, smaller values of *II* yield the machine working in better health conditions with larger efficiency values. For this reason, *II* is an important decision variable for optimizing the maintenance policy.

The duration $t(D^\gamma)$ of the preventive maintenance action performed on component γ depends on the degradation state D^γ in which it is found. More precisely, fixing nozzles in degradation state d_1 requires one operator working for 0.5 day; 1 day is needed for a maintenance operator to repair nozzles in degradation state d_2. Finally, if a nozzle is heavily degraded (in degradation state d_3), then a maintenance operator takes 4 days to repair it.

From these considerations, it appears that the preventive maintenance time T_M required for repairing all the *N* nozzles is given by $T_M = T_{insp} + \sum_{\gamma=1}^{N} t(D^\gamma)$.

Table summarizes the values of the durations of the maintenance actions corresponding to the different degradation states. For the sake of brevity, we indicate by $t(D^\gamma = d_i)$ the time required to fix the γ-th nozzle when it is found in degradation state d_i.

Table II

Values of the duration of the maintenance actions, and of the loss of efficiency

	$t(D^\gamma)$ [days]	$l_E(D^\gamma)$
$D^\gamma = d_1$	0.5	0.5%
$D^\gamma = d_2$	1	0.5%
$D^\gamma = d_3$	4	1.2%

Obviously, the system is unavailable during inspections and repairs. This unavailability causes a business interruption whose cost is given by the part of the annual income *I* that the maintenance actions prevent from being gained. Thus, reducing the amount of time spent in repairing the nozzle system has a beneficial effect on the maintenance costs. In this respect, a larger number of maintenance operators N_{mo} can be involved in repairing actions. The effect on the time reduction is given by

$$T_M = T_{insp} + \frac{\sum_{\gamma=1}^{N} t(D^\gamma)}{N_{mo}}.$$

On the other side, reducing maintenance time has its own cost, as maintenance operators must be paid for their work. We assume that their daily cost is C_O=500€/person. Then, N_{mo} is another decision variable that enters the optimization of the CBM policy.

Generally speaking, nozzle degradation entails a loss in turbine efficiency, whose magnitude depends on the degradation state. In this work, we assume that, when the generic nozzle γ enters degradation state d_1, it causes a loss $l_E(D^\gamma = d_1) = 0.5\%$ in turbine efficiency. An additional drop of 0.5% is associated to each component in degradation state d_2 (i.e., $l_E(D^\gamma = d_2) = 0.5\%$), whereas each nozzle in state d_3 brings about a further, large loss of 1.2% (i.e., $l_E(D^\gamma = d_3) = 1.2\%$). These data are summarized in Table . Thus, the loss L_E in turbine efficiency in a cycle (i.e., the time between two maintenance actions) is given by

$$L_E = \sum_{\gamma=1}^{N} \sum_{i=1}^{3} l_E(D^\gamma = d_i) \cdot (T_{stop} - T_i^\gamma)$$

where T_{stop} is the end time of the cycle (i.e., either the end of the interval *II*, or the time at which the turbine efficiency reaches the threshold T_E, whichever comes first), whereas T_i^γ is the stochastic transition time T_i of component γ. Notice that the simplified scheme considered in this work entails that the worst condition (i.e., the N=22 nozzles are all in degradation state d_3) determines a total loss in turbine efficiency of at most 22*(1.2+0.5 +0.5)% = 48.4%.

Turbine inefficiency entails a cost, which is due to the production loss with respect to the full capacity production conditions. This loss is given by the part of the annual income that inefficiency prevents from being gained. That is, $I_C = L_E \cdot I$ is the inefficiency cost in a cycle.

Finally, for clarity, all the parameters and variables of the case study, with relevant explanations, values, and formulas, are summarized in Table .

Table III

Case study parameters and variables

Name	Description	Value
C_O	Maintenance operator daily cost	500€
C_S	Cost for replacing the nozzle system	25K€
C_S	Total Cost for replacing the nozzle system	$C_S + U_R \cdot I$
d_i	i-th degradation state	i=1,…, 4
D^γ	Variable indicating the degradation state of component γ	γ=1,…, N
I_C	Inefficiency cost in a cycle	$L_E \cdot I$
I	Annual Income	500K€
II	Inspection Interval	Decision variable (Table IV)
$l_E(D^\gamma = d_i)$	Loss in turbine efficiency when component γ enters degradation state d_i	See Table
L_E	Loss in turbine efficiency in a cycle	$L_E = \sum_{\gamma=1}^{N} \sum_{i=1}^{3} l_E(D^\gamma = d_i) \cdot (T_{stop} - T_i^\gamma)$
N	Number of similar components in the system	22
N_{mo}	Number of maintenance operators involved in repairing actions	Decision variable (Table IV)

$t(D^{\gamma} = d_i)$	Time to repair nozzle γ when it is found in state d_i	See Table
T_E	Efficiency Threshold	Decision variable (Table IV)
T_{insp}	Duration of inspections	1.5d
T_M	Duration of preventive maintenance actions	$T_M = T_{insp} + \frac{\sum_{\gamma=1}^{N} t(D^{\gamma})}{N_{mo}}$
T_{stop}	Time instants at which the cycle ends	Either the inspection time at the end of the interval II or the time in which the turbine efficiency reaches the threshold TE, whichever comes first
T_i^{γ}	Stochastic transition time T_i from d_i to d_{i+1} experienced by component γ	
$\overline{T_i}$	Mean transition time from state d_{i-1} to state d_i	Uncertain parameter (Table I)
U_R	System Unavailability owing to the replacement task	4d

4.1 Results: single objective optimization

In this section, the extended GA approach proposed in this work is applied to minimize the mean system unavailability. In simple words, we are searching for the combination of the three variables *II*, T_E, and N_{mo} that minimizes the mean system unavailability over a time horizon of 20 years.

Table reports the characteristics of the search space in which the optimal solution is searched. No further constraints are considered. Table explicitly describes the GA rules and control parameters used in this case study, as well as the setting of the parameters of the MC-DSTE method used to represent and propagate the uncertainty.

Table IV

GA settings

	Search space	Number of bits
II [y]	[0.5,8.5]	5
T_E	[0.05, 0.50]	4
N_{mo}	{1,…,8}	3

Table V

Parameters and rules

		Search space
GA parameters	Population size *H*	30
	Number of generations (termination criterion)	50
	Selection	Fit-Fit
	Replacement	Children-Parents
	Mutation probability	0.001
	Crossover probability	1
Hybrid MC-DSTE Merthod (see Appendix)	Number of MC simulation N_T	200
	Number of combinations of parameters N_S	40

Number of focal sets E in the evidence space $(S_{\underline{\theta}}, I, m_I)$, B	8

The results of the proposed GA method are summarized in Table , which reports the values of the decision variables corresponding to the optimal solutions of the first rank order. These results are all characterized by the large number N_{mo} of operators performing the maintenance actions. This situation means that setting such decision variables to small values entails a significant worsening of the maintenance policy performance in terms of mean system unavailability. From Table , it also emerges that the minima are positioned around II=3y and 4y, and their multiples (6 and 8, respectively).

Another interesting piece of information can be discovered. Consider solutions X_6 and X_7; these solutions have the same values of II and N_{mo}, but different values of T_E (41.56 and 35.94, respectively). As significant changes in the maintenance policy performance are not expected to be encountered when T_E varies in the interval [35.94, 41.56], it seems reasonable to say that the unavailability values correspond to the same setting of II and N_{mo}, and values of T_E in that the intervals are similar to each other. Fig. 5 shows the pairs $[Bel_{\Xi^1}(\xi^1), Pl_{\Xi^1}(\xi^1)]$ corresponding to values of T_E=0.38 and T_E=0.40, compared with those of solutions X_6 and X_7. The overlap of these curves confirms that the solutions with values of T_E that lie in the interval [35.94, 41.56] are not worse than those found by GA. From this result, we can conclude that the output of the proposed GA does not include all the possible solutions of the same rank; certainly, a longer running of the algorithm would be expected to add new, more performing solutions in the final population, in substitution of those with smaller fitness values. This effect is typical of the GA global search, especially when the search space is very large; in fact, GA is a heuristic that efficiently identifies the areas of the search space in which the optimal solutions are located, thus avoiding the burden of evaluating the fitness values corresponding to every point of the search space.

Table VI

Set of solutions of rank 1

	II [y]	T_E [%]	N_{mo}
X_1	4	10.62	8
X_2	5.5	47.19	8
X_3	3	44.37	8
X_4	6	24.69	8
X_5	3	47.19	7
X_6	8	41.56	8
X_7	8	35.94	8

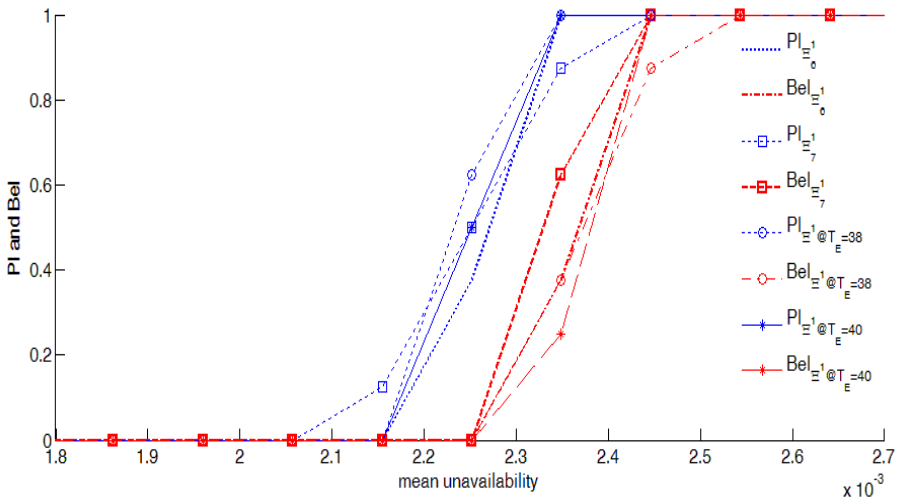

Fig. 5. Comparison of the fitness values of solutions X_6 and X_7, with those corresponding to other possible solutions of the search space.

As pointed out in Section 3.1, sorting the solutions on the basis of pair-wise comparisons may lead to the contradictory ranking issue, which calls for an additional investigation when making the final choice. As anticipated above (i.e., Section 3.2), this sorting is done by running the algorithm shown in [64] on the final population, which explicitly assigns the rank order to the elements of the population, based on a cross-check of all the pairwise relationships. From this output, we can select the solutions tagged with rank order r=1, which are the 7 solutions reported in Fig. (a-g). Notice that the execution of this algorithm is not required during the Single-Objective optimization based on the fit-fit reproduction phase. In fact, we have the vector of solutions of the current population sorted according to their fitness values, and we mate them into pairs; we don't need to know what their rank orders are.

Notice also that Fig. (h) shows the fitness values corresponding to the first two solutions among those of the same rank order r=1. Their direct comparison would have led to conclude that the second solution is better than the first one. On the contrary, this couple of solutions are considered equivalent by the proposed GA technique. This equivalence is due to the fact that there are other solutions with fitness values in the middle of the two pairs of curves in Fig. (h). In fact, Fig. (a-g) show the 7 solutions that the pair $[Bel_{\Xi_1^1}(\xi^1), Pl_{\Xi_1^1}(\xi^1)]$ associated to solution X_1 has a large overlap with $[Bel_{\Xi_3^1}(\xi^1), Pl_{\Xi_3^1}(\xi^1)]$, of solution X_3. According to the ranking criterion given in Section 3.2, in this situation, X_1 and X_3 have to be considered equivalent. Yet, the pair of lower and upper probability bounds corresponding to solution X_3 overlap those of solution X_4 (and then X_3 is equivalent to X_4), which is itself equivalent to X_2.

On this basis, further criteria not accounted for in the maintenance model, as they are difficult to be quantified (e.g., perceived reliability, difficulties in organizing the maintenance teams, etc.), can be considered to identify the best solution among those belonging to the best cluster.

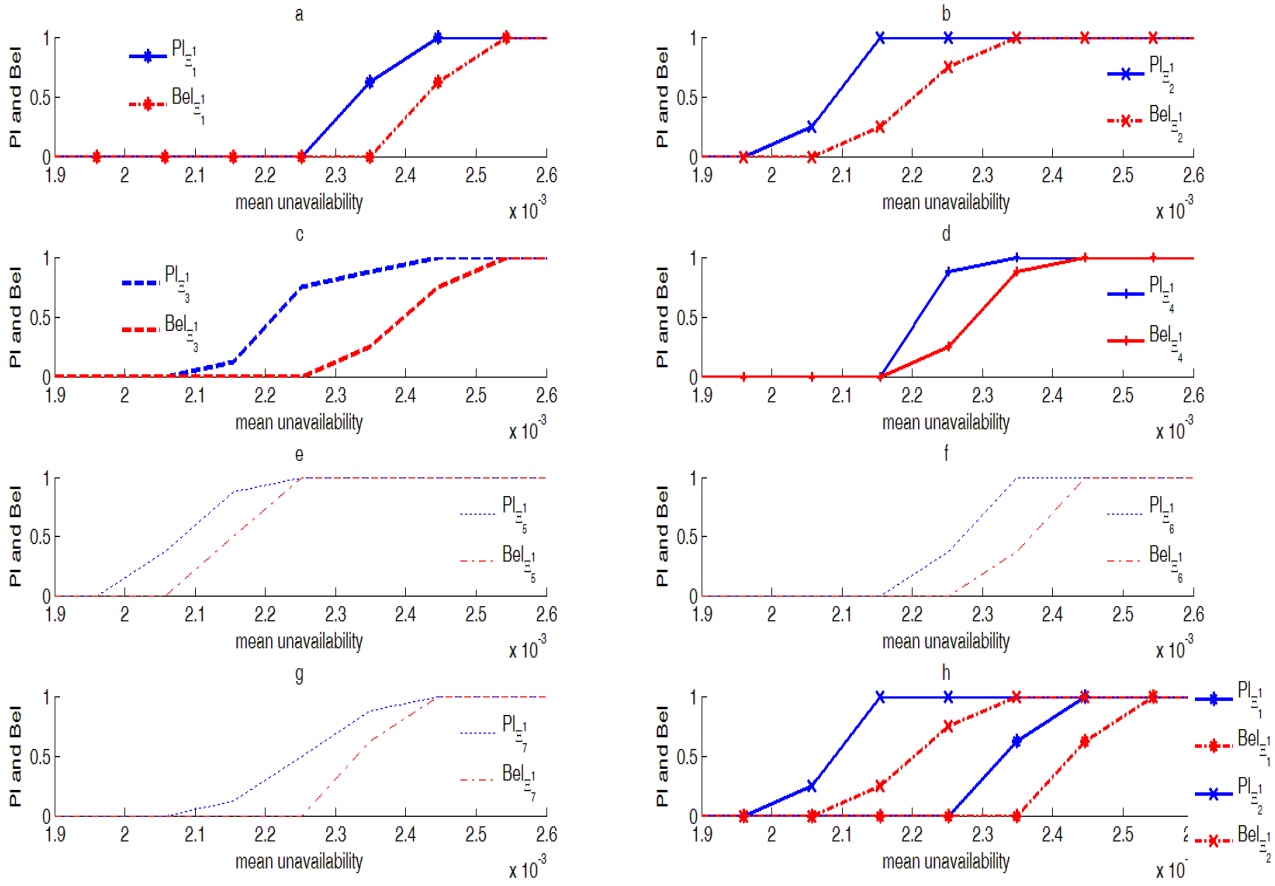

Fig. VI. Fitness values of the solutions of rank 1.

The main limitation of the algorithm lies in the large computational times required. As pointed out in [22], these times mainly depend on the number N_T of MC trials, and the number N_s of combinations of parameters, which itself depends on the number B of focal sets; see the Appendix. To reduce the algorithm running time, the solution opposite to that considered in [52] has been implemented in this case study: the uncertainty propagation procedure is performed only once, at the first time the solution enters the current population, with statistical significance; the corresponding fitness values are evaluated and recorded, so that they can be re-used when the same solution appears again. This computational choice can be memory demanding. For example, in the present case study, 860 solutions out of the 8*16*32=4096 possible solutions in the search space have been handled by the algorithm. The size of the array storing the corresponding fitness is (860, 2, 150), where 2 refers to the two measures *Pl* and *Bel*, and 150 is the number of bins partitioning the mean unavailability axes. Obviously, the larger the number of fitness values, and the larger the search space, the larger the memory required. Then, the method here proposed to reduce the computational times needs to be traded off against that developed in [52], depending on the specific application. This issue will be tackled in future works.

In the setting described above, the CPU time required by the algorithm implemented in Matlab is 10 hours on a Pentium Dual Core (1.73GHz). This performance can be heavily improved by coding the algorithm in a non-interpretative language such as C++, Fortran, etc.

4.2 Results: Multi-objective optimization

The results relevant to the multi-objective case study are summarized in Table , which reports the Pareto optimal set. Some differences emerge with respect to the solutions of the single-objective optimization case (Table). Namely, the solutions of the Pareto front are shifted towards smaller values of the inspection interval *II;* for

example, the solutions with *II*=6 & *II*=8, which are contained in Table , do not belong to the Pareto front. This condition is due to the fact that less frequent inspections on one side allow saving repair time, but on the other side they entail a loss in the turbine efficiency, with a consequent loss of money. This latter aspect is not taken into account in the single-objective optimization.

Moreover, with respect to Table , the solutions corresponding to *II*=3 & *II*=3.25 are associated with smaller values of T_E in the Pareto front. This difference stems from the fact that, if one chooses to stop the machine more frequently (smaller values of *II*), then setting larger values of the threshold T_E (decrement in the turbine efficiency) becomes anti-economic. In turn, the advantage of the multi-objective optimization is that two drivers (i.e., unavailability and cost), instead of just one, enter the decision making process.

Notice that the Pareto front cannot be plotted, as two pairs of distributions are associated to every solution of the Pareto optimal set.

Notice also that the computational times relevant to the multi-objective setting are similar to those of the single objective setting. In fact, the sorting algorithm, whose application is here doubled, does not take significant amounts of time.

Table VII

Pareto optimal set

II [y]	T_E [%]	N_{mo}
5	41.57	7
5	44.37	7
4,25	30.31	8
4	41.56	7
3	30.31	7
3,25	30.31	8

To highlight the contribution of the proposed method, we have also run the standard GA available in Matlab, in the case in which the model parameters are not affected by uncertainty and take the values $\overline{T}_1 = 11$, $\overline{T}_2 = 3$, $\overline{T}_3 = 2$. The algorithm converged to the Pareto set, which turned out to be made up of one point only: *II*= 2.87, T_E = 29.2%, N_{mo} = 7.07. Fig. 6 shows the corresponding Pareto front, indicated by the star. This result allows making two considerations, as follows.

- The traditional GA technique suffers from the risk of providing incomplete Pareto fronts; even if additional investigations were done on how to set the algorithm parameters, one would not get the certainty of having found all the points of the Pareto front [69]. Fig. 6 shows, indicated by circles, the points of the front corresponding to the optimal set in Table , as they would have been calculated in the setting with no uncertainties in the degradation model parameters. Now, two out of these 6 points, indicated by bold circles, do not dominate the star point, and thus belong to the Pareto front, but they are not found by the GA.
- Accounting for the uncertainty in the model parameters gives the possibility of considering additional solutions to the maintenance optimization problem, which would have been otherwise disregarded. In general, the larger the cardinality of the Pareto front, the larger the number of decision alternatives, and the more informed the final decision. Certainly, a method to select a solution from the Pareto front is needed to complete the study. For example, preference-based techniques ([70]) can be adopted to manipulate the Pareto front, which allow introducing additional conditions (i.e., lexical constraints introduced by the decision maker), or considering just the extreme solutions of the front (i.e., those corresponding to the smallest values of the single objectives), or considering compromise solutions. This issue will be tackled in future works.

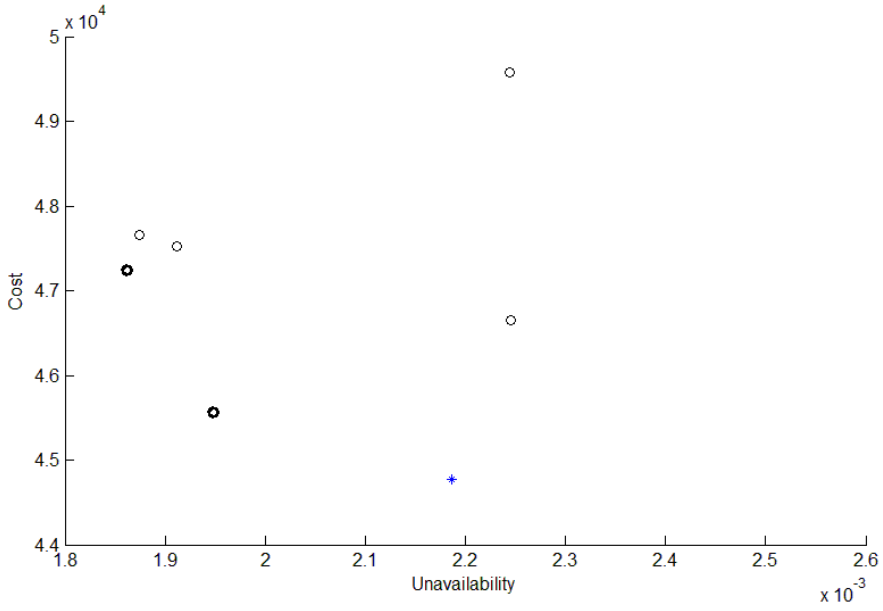

Fig. 6. Pareto fronts.

5 Conclusions

In this work, a novel approach has been proposed to extend GA for applications in which the objective functions depend on stochastic processes described by distributions whose parameters are imprecisely known, and only through information elicited from experts. Technically, this situation introduces noise and uncertainty in the optimization problem.

A novel ranking method has been developed to address single-objective problems under this uncertainty setting, and a novel definition of Pareto dominance has been provided for the multi-objective optimization. Particular attention has been given to the problem of the contradictory ranking, especially in the multi-objective case, which has been disregarded by other works in the literature. This issue requires an additional analysis by the decision maker for selecting the best solution, possibly on the basis of an additional distinctive criterion.

The proposed algorithm has been applied to a practical case study concerning the optimization of the maintenance policy of the nozzle system of an oil & gas turbine. This example shows the potential of the technique, and its limitations, the main one being the computational times required. In this respect, although a computational solution has been implemented to reduce the algorithm running time, additional effort will be dedicated in future works to devise further methods for reduction.

Future works will focus also on the development of a sharing technique in the considered uncertainty setting, to prevent the population evolution against the loss of genetic diversity [58].

6 Appendix

Every expert is asked to provide the interval that he or she believes contains the true value of the uncertain parameter. For the sake of clarity, we specify that the number of experts involved in the quantification of the *p*–th

parameter $\theta^{j,p}$ of the *j*-th random variable Y^j is indicated by $e^{j,p}$; and the interval $I_i^{j,p}$ provided by the *i*-th expert is identified by its lower, and upper bounds $\underline{\iota_i^{j,p}}$, and $\overline{\iota_i^{j,p}}$, respectively.

According to the procedure proposed in [36], the evidence space defining the generic uncertain parameter $\theta^{j,p}$ is indicated as ($S^{j,p}$, $I^{j,p}$, $m^{j,p}$), and is defined by assuming that the sets $I_i^{j,p}$, $i=1,...,e^{j,p}$ constitute the focal elements. Thus, $S^{j,p}$ is the domain of the *p*-th parameter, whereas the set of focal elements $I^{j,p}=\{I_i^{j,p}, i=1,...,e^{j,p}\}$. Finally, the BPA associated to a particular focal element $I_i^{j,p}$ is given by

$$m^{j,p}=Kr(I_i^{j,p})/e^{j,p}. \quad (1)$$

The model *g* depends on a number *Nu* of parameters affected by epistemic uncertainties, where $Nu=\sum_{j=1}^{k} M^j$; for convenience, these parameters are organized in the vector $\underline{\theta}=\theta^{1,1},...,\theta^{1,M^1},\theta^{2,1},...,\theta^{2,M^2},...,\theta^{k,1},...,\theta^{k,M^k}$. Hence, *g* maps points of a *Nu*-dimensional space into an *O*-dimensional space; this mapping entails that the first step to propagate the uncertainty is to build an evidence space on such *Nu*-dimensional space.

According to [36], the evidence space ($S_{\underline{\theta}}$, *I*, m_I) characterizing the uncertainty in this multi-dimensional space of $\underline{\theta}$ is constructed on the basis of the mono-dimensional evidence spaces of the single parameters of $\underline{\theta}$. Specifically, S_θ is the set containing the points $\underline{\vartheta}=[\vartheta^{1,1},...,\vartheta^{k,M^k}]$ that belong to the Cartesian product of the sample spaces of the *Nu* uncertain parameters; that is, $\{\underline{\vartheta} \mid \underline{\vartheta}=[\vartheta^{1,1},...,\vartheta^{k,M^k}] \in S^{1,1}\times....\times S^{k,M^k}\}$. The set of focal elements is $I=\{E \mid E\in I^{1,1}\times....\times I^{k,M^k} \in S^{1,1}\times....\times S^{k,M^k}\}$, of cardinality B. Under the assumption that the parameters of $\underline{\theta}$ are stochastically independent, m_I is defined in analogy to the case of probability spaces, where the probability of the combination of events pertaining to different spaces is given by the product of the probabilities of the single events; that is,

$$m_I(\mathrm{E})=\begin{cases}\prod_{\substack{j=1,...k,\\ p=1,...M^j}} m^{j,p}(I_{i^{j,p}}^{j,p}) & \text{if } E=I_{i_{1,1}}^{1,1}\times....\times I_{i_{k,M^k}}^{k,M^k}\in I\\ 0 & \text{otherwise}\end{cases} \quad (2)$$

being $i^{j,p}\in\{1,...,e^{j,p}\}$.

The methodology to propagate the uncertainties from $\underline{\theta}$ to $\underline{\Xi}$ consists of the following steps [36].

1. Define a probability distribution *d* on S_I to be used for generating a sample $\underline{\vartheta}=[\vartheta^{1,1},...,\vartheta^{k,M^k}]$ of $\underline{\theta}$. One way is to define the distributions $d^{j,p}(\vartheta^{j,p})$ for sampling in each $S^{j,p}$, $j=1,...,k$, and $p=1,...,M^j$; assuming statistical independence between the parameters, the distribution $d(\underline{\vartheta})$ for sampling $\underline{\vartheta}$ is then defined as $d(\underline{\vartheta})=\prod_{\substack{j=1,...k\\ p=1,...,M^j}} d^{j,p}(\vartheta^{j,p})$. The construction of the distributions $d^{j,p}(\vartheta^{j,p})$ is based on the assumption that the sets $I_i^{j,p}$ contained in $I^{j,p}$ can be treated as discrete outcomes with probabilities $P(I_i^{j,p})=m^{j,p}(I_i^{j,p})$. Conditional on its occurrence, a uniform distribution $U_i^{j,p}(\vartheta_{j,p})$ over $I_i^{j,p}$ is considered. Then, the density function associated with ($S^{j,p}$, $I^{j,p}$, $m^{j,p}$) is given by

$$d^{j,p}(\vartheta^{j,p})=\sum_{i=1}^{e^{j,p}} m^{j,p}(I_i^{j,p})U_i^{j,p}(\vartheta^{j,p}) \quad (3)$$

where $\vartheta^{j,p} \in S^{j,p}$, with the convention that $U_i^{j,p}(\vartheta^{j,p}) = 0$ if $\vartheta^{j,p} \notin I_i^{j,p}$. In turn, the distribution $d^{j,p}$ is, for every value $\vartheta^{j,p}$, the weighted mean of the values of the uniform distributions $U_i^{j,p}(\vartheta^{j,p})$, where the weights $m^{j,p}$ are the number of experts that agree on including the value $\vartheta^{j,p}$ among the possible values of the ill-known parameter $\theta^{j,p}$.

2. Generate a random or Latin hypercube sample from the *Nu*-dimensional space $\underline{\theta}$, coherently with the distribution defined in the previous step 1. This generation is done by sampling, for each $j = 1,...,k$, two uniform random numbers r_1 and r_2 from the half-closed interval [0,1); the first number r_1 is used to select a set $I_i^{j,p}$ with probability $m^{j,p}(I_i^{j,p})$, whereas r_2 is used to select, via the inverse transform method [66], a value $\vartheta^{j,p}$ in consistency with the definition of the density function $U_i^{j,p}$.
3. Each sample of $\underline{\theta}$ gives rise to a probability space characterizing the aleatory uncertainty in the output $\underline{Z}$. In practice, once the values of the parameters in $\underline{\theta}$ have been fixed, one can perform a standard MC propagation of the uncertainty affecting the stochastic variables Y_j, $j = 1,\cdots,k$ to obtain the uncertainty on the output variables Z^o, $o=1,...O$. This action requires simulating the model behavior a large number N_T of times. Because probability spaces are too complex to be considered graphically or numerically as single, distinct entities, various summary measures (e.g., mean, percentiles, etc.) that can be derived from the definition of a probability space are often used to lump the information of the space. Such measures are computed in this step, and form the output vector $\underline{\Xi}=(\Xi^1,...,\Xi^Q)$.
4. Repeat Steps 2 and 3 a large number of times N_S. Notice that the number B of focal sets in *I* influences the choices of N_S [22].
5. Estimate the Plausibility, and Belief measures, respectively $Pl_{\Xi^q}(]-\infty,\xi^q])$, and $Bel_{\Xi^q}(]-\infty,\xi^q])$, of the Ξ^q components of $\underline{\Xi}$, $q=1,...,Q$. This estimation is done by identifying for every ξ^q the set $INV(\Xi)^q = \{E \subset I : E \cap \Xi^{q-1}(]-\infty,\xi^q]) \neq \varnothing\}$, i.e., we first search the points $\underline{\vartheta} = [\vartheta^{1,1},...,\vartheta^{k,M^k}]$ of the multi-dimensional space $S_{\underline{\theta}}$ which determine probability spaces whose measure Ξ^q (e.g., mean) falls in the interval $]-\infty,\xi^q]$. Then, we identify the subsets *E* of *I* which these points belong to. On this basis, the Plausibility, and Belief measures are computed as

$$Pl_{\Xi^q}(]-\infty,\xi^q]) = \sum_{E \in INV(\Xi^q)} m_I(E), \tag{4}$$

and

$$Bel_{\Xi^q}(]-\infty,\xi^q]) = 1 - Pl^c(]-\infty,\xi^q]) \tag{5}$$

being $Pl^c_{\Xi^q}(\bar{\xi}^q) = \sum_{E \in \overline{INV}(\Xi^q)} m_I(E)$ the Plausibility of the interval $]\xi^q,\infty[$ (i.e., the complement of $]-\infty,\xi^q]$), and the set $\overline{\mathrm{INV}}(\Xi^q) = \{E \subset I : E \cap \Xi^{q-1}(]\xi^q,\infty[) \neq \varnothing\}$.

7 Acknowledgements

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

8 References

[1]. E. Zio and M. Compare, "Evaluating maintenance policies by quantitative modeling and analysis," *Reliability Engineering and System Safety*, Vol. 109, pp. 53-65, 2013.

[2]. E. Zio and M. Compare, "A snapshot on maintenance modeling and applications," *Marine Technology and Engineering*, Vol. 2, pp. 1413-1425, 2013.

[3]. G.E. Apostolakis, "The concept of probability in safety assessments of technological systems," *Science*, Vol. 250, pp. 1359-1364, 1990.

[4]. C. Guedes Soares, J. Caldeira Duarte, Y. Garbatov, E. Zio and J.D. Sorensen, "Framework for maintenance planning," *Safety and Reliability of Industrial Products, Systems and Structures*, C. Guedes Soares (Ed.), CRC Press/Balkema, Leiden, The Netherlands, 2010.

[5]. R.P. Nicolai, J.B.G. Frenk and R. Dekker, "Modeling and optimizing imperfect maintenance of coatings on steel structures," *Structural Safety*, Vol. 31, pp. 234-244, 2009.

[6]. P. Baraldi, M. Compare, A. Despujols, W. Lair and E. Zio, "A practical analysis of the degradation of a nuclear component with field data," *Proceedings of European Safety and Reliability Conference 2013-ESREL 2013*.

[7]. D. Dubois, "Possibility Theory and Statistical Reasoning," *Computational Statistics and Data analysis*, Vol. 51, pp. 47-69, 2006.

[8]. C. Baudrit, D. Dubois and D. Guyonnet, "Joint Propagation and Exploitation of Probabilistic and Possibilistic Information in Risk Assessment," *IEEE Transactions on Fuzzy Systems*, Vol. 14, No. 5, pp. 593-608, 2006.

[9]. L.A. Zadeh, "Fuzzy Sets," *Information and Control*, Vol. 8, pp. 338-353, 1965.

[10]. M. An, Y. Chen and C.J. Baker, "A fuzzy reasoning and fuzzy-analytical hierarchy process based approach to the process of railway risk information: A railway risk management system," *Information Sciences*, Vol. 181, No. 18, pp. 3946–3966, 2011.

[11]. P. Baraldi, A. Balestrero, M. Compare, L. Benetrix, A. Despujols and E. Zio, "A Modeling Framework for Maintenance Optimization of Electrical Components Based on Fuzzy Logic and Effective Age," *Quality and Reliability Engineering International*, Vol. 29, No. 3, pp. 385-405, 2013.

[12]. P. Baraldi, M. Compare, G. Rossetti, A. Despujols and E. Zio, "A modelling framework to assess maintenance policy performance in electrical production plants," *Maintenance Modelling and Applications, ESREDA-ESRA Project Group Report*. Andrews, Berenguer and Jackson Eds., pp 263-282, 2011.

[13]. B. Jones, I. Jenkinson and J. Wang, "The use of fuzzy set modelling for maintenance planning in a manufacturing industry," *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering*, Vol. 224, pp. 35-48, 2010.

[14]. G. De Cooman, F. Hermans and E. Quaeghebeur, "Imprecise Markov Chains And Their Limit Behaviour," *Probability in the Engineering and Informational Sciences*, Vol. 23, pp. 597-635, 2009.

[15]. I.O. Kozine and L.V. Utkin, "Interval-valued finite Markov chains," *Reliable Computing*, Vol. 8, No. 2, pp. 97–113, 2002.

[16]. S.C.M. Rocco, "Effects of the transition rate uncertainty on the steady state probabilities of Markov models using interval arithmetic," *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, Vol. 226, No. 1, pp. 234-245, 2011.

[17]. D. Škulj, "Discrete time Markov chains with interval probabilities," *International Journal of Approximate Reasoning*, Vol. 50, pp. 1314–1329, 2009.

[18]. H. Ge and S. Asgarpoor, "Reliability Evaluation of Equipment and Substations with Fuzzy Markov Processes," *IEEE Transaction on Power System*, Vol. 25, No. 3, pp.1319-1328, 2010.

[19]. Baraldi, P., Compare, M., Zio, E. "Uncertainty analysis in degradation modeling for maintenance policy assessment," Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, Vol. 227, No. 3, pp. 267-278, 2013.

[20]. P. Baraldi, M. Compare and E. Zio, "Uncertainty treatment in expert information systems for maintenance policy assessment," Applied Soft Computing, Vol. 22, pp. 297-310, 2014.

[21]. P. Baraldi, M. Compare and E. Zio, "Dempster-Shafer theory of evidence to handle maintenance models tainted with imprecision," *Proceedings of the 11th International Probabilistic Safety Assessment and*

Management Conference and the Annual European Safety and Reliability Conference 2012, PSAM11 ESREL 2012, 25-29 June 2012, Helsinki, Finland, pp. 61-70, 2012.

[22]. P. Baraldi, M. Compare and E. Zio, "Maintenance policy performance assessment in presence of imprecision based on Dempster–Shafer Theory of Evidence," *Information Sciences*, Vol. 245, pp. 112-131, 2013.

[23]. H. Eskandari, C.D. Geiger and R. Bird, "Handling uncertainty in evolutionary multiobjective optimization: SPGA," *Proceedings of the IEEE Congress on Evolutionary Computation, 2007 - CEC 2007*, 25-28 September 2007, Singapore, pp. 4130,4137, 2007.

[24]. J.E. Hughes, "Evolutionary Multi-objective Ranking with Uncertainty and Noise," *Proceedings of Evolutionary Multi-Criterion Optimization, First International Conference, EMO 2001*, Zurich, Switzerland, March 7-9, pp. 329-343, 2001.

[25]. J. Li and R.S.K. Kwan, "A fuzzy genetic algorithm for driver scheduling," *European Journal of Operational Research*, Vol. 147, No. 2, pp. 334-344, 2003.

[26]. A. Trebi-Ollennu and B.A. White, "Multiobjective fuzzy genetic algorithm optimisation approach to nonlinear control system design," *IEE Proceedings - Control Theory and Applications*, Vol. 144, pp. 137-142, 1997.

[27]. P. Limbourg, "Multi-objective Optimization of Problems with Epistemic Uncertainty," *Proceedings of Evolutionary Multi-Criterion Optimization, Third International Conference, EMO 2005*, Guanajuato, Mexico, March 9-11, pp. 413-427, 2005.

[28]. N. Pedroni, E. Zio, E. Ferrario, A. Pasanisi and M. Couplet, "Hierarchical propagation of probabilistic and non-probabilistic uncertainty in the parameters of a risk model," *Computers & Structures*, Vol. 126, pp. 199-213, 2013.

[29]. F. Tonon, "Using random set theory to propagate epistemic uncertainty through a mechanical system," *Reliability Engineering and System Safety*, Vol. 85, pp. 169–181, 2004.

[30]. F. Tonon, A. Bernardini and A. Mammino, "Determination of parameters range in rock engineering by means of Random Set Theory," *Reliability Engineering and System Safety*, Vol. 70, pp. 241-261, 2000.

[31]. P. Baraldi, M. Compare and E.Zio, "A practical analysis of the degradation of a nuclear component with field data," *Reliability and Risk Analysis: Beyond the Horizon*, Eds. R.D.J.M. Steenbergen, P.H.A.J.M. van Gelder, S. Miraglia, A.C.W.M. Vrouwenvelder, CRC Press, 2013.

[32]. A. Lisnianski, D. Elmakias, D. Laredo and H.B. Haim, "A multi-state Markov model for a short-term reliability analysis of a power generating unit," *Reliability Engineering and System Safety*, Vol. 98, No. 1, pp. 1-6, 2012.

[33]. R. Moghaddass and M. J Zuo, "A parameter estimation method for condition monitored equipment under multi-state deterioration," *Reliability Engineering and System Safety*, Vol. 106, pp. 94-103, 2012.

[34]. R. Moghaddass, and M.J. Zuo, "Multistate Degradation and Condition Monitoring for Devices with Multiple Independent Failure Modes," *Applied Reliability Engineering and Risk Analysis: Probabilistic Models and Statistical Inference*, Eds. I. B. Frenkel, A. Karagrigoriou, A. Lisnianski and A. Kleyner, John Wiley & Sons, Ltd, Chichester, UK, 2013.

[35]. F. Di Maio, M. Compare, S. Mattafirri and E. Zio, "A double-loop Monte Carlo approach for Part Life Data Base reconstruction and scheduled maintenance improvement", *Safety and Reliability: Methodology and Applications*, Nowakowski et al. (Eds), pp. 1877- 1884, 2015.

[36]. J.C. Helton, J.D. Johnson and W.L. Oberkampf, "An exploration of alternative approaches to the representation of uncertainty in model predictions," *Reliability Engineering and System Safety*, Vol. 85 pp. 39-71, 2004.

[37]. C. Simon and P. Weber, "Evidential Networks for Reliability Analysis and Performance Evaluation of Systems With Imprecise Knowledge," *IEEE Transactions on Reliability*, Vol. 58, No. 1, 2009, pp. 69-87.

[38]. L. Utkin and F. Coolen, "Imprecise reliability: An introductory overview," *Studies in Computational Intelligence*, Vol. 40, 2007, pp. 261-306.

[39]. M. Beer, S. Ferson and V. Kreinovich, "Imprecise probabilities in engineering analyses," *Mechanical Systems and Signal Processing*, Vol. 37, No. 1-2, 2013, pp. 4-29.

[40]. Statistical Policy Office, Office of Information and Regulatory Affairs, Office of Management and Budget: 'Statistical Policy-Working Paper 31: Measuring and Reporting Sources of Error in Surveys', 2001.
[41]. P. Baraldi, M. Compare, E. Zio, M. de Nigris, and G. Rizzi, "Identification of contradictory patterns in experimental datasets for the development of models for electrical cables diagnostics," *International Journal of Performability Engineering*, Vol. 7, No. 1, pp. 43-60, 2011.
[42]. I.O. Kozine and Y.V. Filimonov, "Imprecise reliabilities: experiences and advances," *Reliability Engineering and System Safety*, Vol. 67, pp. 75–83, 2000.
[43]. J.S. Wu, G.E. Apostolakis, and D. Okrent, "Uncertainty in System Analysis: probabilistic versus non-probabilistic theories," *Reliability Engineering and System Safety*, Vol. 30, No. (1-3), pp. 163-181, 1990.
[44]. F.Aguirre, M. Sallak, and W. Schon, "Construction of Belief Functions From Statistical Data About Reliability Under Epistemic Uncertainty," *IEEE Transactions on Reliability*, Vol. 62, No. 3, 2013, pp. 555-568.
[45]. M. Sallak, W. Schon, and F. Aguirre, "Extended Component Importance Measures Considering Aleatory and Epistemic Uncertainties," *IEEE Transactions on Reliability*, Vol. 62, 2013, pp. 49-65.
[46]. A.P. Dempster, "Upper and lower probabilities induced by a multi-valued mapping," *Annals of Mathematical Statistics*, Vol. 38, pp. 325-339, 1967.
[47]. M. Marseguerra, E, Zio, L. Podofillini, "Optimal Reliability/Availability of Uncertain Systems via Multi-Objective Genetic Algorithms," *IEEE Trans. On Reliability*, Vol. 53, pp. 424-434, 2004.
[48]. D.W. Coit, A.E. Smith, "Reliability Optimization of Series-Parallel Systems Using a Genetic Algorithm," *IEEE Transactions On Reliability*, Vol. 45, pp. 254-266, 1996.
[49]. M. Gen, YS. Yun, "Soft computing approach for reliability optimization: State-of-the-art survey," *Reliability Engineering and System Safety*, Vol. 91, pp. 1008–1026, 2006.
[50]. M. Marseguerra, E. Zio, L. Podofillini, "Condition-based maintenance optimization by means of genetic algorithms and Monte Carlo simulation," *Reliability Engineering and System Safety*, Vol. 77, pp. 151-165, 2002.
[51]. K. Chellapilla, D.B. Fogel, "Anaconda Defeats Hoyle 6-0: A Case Study Competing an Evolved Checkers Program against Commercially Available Software," *Proceedings of Congress on Evolution Computation - CEC2000*, San Diego, CA, 16-19 July 2000, Vol. 1, pp. 857-863, 2000.
[52]. M. Cantoni, M., Marseguerra, and E. Zio, "Genetic Algorithms and Monte Carlo Simulation for Optimal Plant Design," *Reliability Engineering and System Safety*, Vol. 68, pp. 29-38, 2000.
[53]. J. Teich, "Pareto-Front Exploration with uncertain objectives," *Proceedings of Evolutionary Multi-Criterion Optimization, First International Conference, EMO 2001*, Zurich, Switzerland, March 7-9, pp. 314-328, 2001.
[54]. G. Rudolph, "A Partial Order Approach to Noisy Fitness Functions," *Proceedings of the 2001 Congress on Evolutionary Computation*, 27-30 May 2001, Seoul, Korea, pp. 318 – 325, 2001.
[55]. S. Ferson, and L.R. Ginzburg, "Different methods are needed to propagate ignorance and variability," *Reliability Engineering & System Safety*, Vol. 54, No. (2-3), pp. 133-144, 1996.
[56]. S. Ferson, V. Kreinovich, L.R. Ginzburg, D.S. Myers and K. Sentz, "Constructing Probability Boxes and Dempster-Shafer Structures," SAND REPORT SAND2002-4015, 2003.
[57]. G. Shafer, "A mathematical theory of evidence," Princeton University Press, Princeton, NJ, 1976.
[58]. M. Marseguerra, E. Zio and S. Martorell, "Basics of Genetic Algorithms Optimization for RAMS Applications," *Reliability Engineering and System Safety*, Vol. 91, pp. 977-991, 2006.
[59]. A. Konaka, D.W. Coit and A.E. Smith, "Multi-objective optimization using genetic algorithms: A tutorial," *Reliability Engineering and System Safety*, Vol. 91, pp. 992-1007, 2006.
[60]. K. Sugihara, "Measures for performance evaluation of genetic algorithms," in 3rd Joint Conference on Information Science, JCIS '97 ,1997, pp. 172–175, extended Abstract.
[61]. T. Hickey, Q. Ju and M.H. van Emden, "Interval Arithmetic: from Principles to Implementation," *Journal of the ACM*, Vol. 48, No. 5, pp. 1038-1068.
[62]. R. D. Luce, "Semi-orders and a Theory of Utility Discrimination," *Econometrica*, Vol. 24, No. 2, pp. 178-191, 1956.

[63]. S. Greco, B. Matarazzo, R. Slowinski, "Rough Sets theory for multicriteria decision analysis," *European Journal of Operational Research*, Vol. 129, pp. 1-47, 2001.
[64]. P. Baraldi, M. Compare and E. Zio, "Component Ranking by Birnbaum Importance in Presence of Epistemic Uncertainty in Failure Event Probabilities," *IEEE Transactions On Reliability*, Vol. 62, No. 1, pp. 37-48, 2013.
[65]. D. E. Knuth, "The Art of Computer Programming, Volume 3: Sorting and Searching", Addison-Wesley, 1998.
[66]. E. Zio, "The Monte Carlo Simulation Method for System Reliability and Risk Analysis," Springer series in Reliability Engineering, Springer-Verlag, London, 2013.
[67]. A. Papoulis, U. Pillai, "Probability, Random Variables, and Stochastic Processes," 4th Edition. Mc Graw-Hill, 2002.
[68]. P. Baraldi, E. Zio, and M. Compare, "A method for ranking components importance in presence of epistemic uncertainties", *Journal of Loss Prevention in the Process Industries,* Vol. 22, No. 5, pp. 582-592, 2009.
[69]. K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A Fast and Elitist Multiobjective Genetic Algorithm NSGA-II," *IEEE transactions on evolutionary computation*, Vol. 6, No. 2, pp. 182-197, 2002.
[70]. U. Junker, "Preference-Based Problem Solving for Constraint Programming", *Recent Advances in Constraints*, 12th Annual ERCIM International Workshop on Constraint Solving and Constraint Logic Programming, CSCLP 2007, Rocquencourt, France, June 7-8, 2007, Revised Selected Papers, Fages, Rossi and Soliman Eds., pp. 109-126, 2008.
[71]. M. Marseguerra, E. Zio, L. Podofillini and D.W. Coit, "Optimal Design of Reliable Network Systems in Presence of Uncertainty," *IEEE Transactions On Reliability,* Vol. 54, pp. 243-253, 2005.

9 Biographies

Michele Compare (BS in mechanical engng., University of Naples Federico II, 2003, PhD in nuclear engng., Politecnico di Milano, 2011) is a research consultant and CEO at Aramis. He was a post-doc fellow at the Politecnico di Milano. He worked as RAMS engineer, and risk manager. His main research efforts are devoted to the development of methods and techniques in support of the maintenance modeling and decision making in complex systems.

Enrico Zio (BS in nuclear engng., Politecnico di Milano, 1991; MSc in mechanical engng., UCLA, 1995; PhD, in nuclear eng'g., Politecnico di Milano, 1995; PhD, in nuclear engng., MIT, 1998) is Director of the Chair in Complex Systems, and the Energetic Challenge of Ecole Centrale Paris and Supelec, full professor, Rector's delegate for the Alumni Association and past-Director of the Graduate School at Politecnico di Milano, and adjunct professor at University of Stavanger. He is the Chairman of the European Safety and Reliability Association ESRA, member of the Korean Nuclear society and China Prognostics and Health Management society, and past-Chairman of the Italian Chapter of the IEEE Reliability Society. He is serving as Associate Editor of IEEE Transactions on Reliability, and as editorial board member in various international scientific journals. He has functioned as Scientific Chairman of three International Conferences, and as Associate General Chairman of two others. His research topics include analysis of the reliability, safety, and security of complex systems under stationary and dynamic conditions, particularly by Monte Carlo simulation methods; and development of soft computing techniques for safety, reliability and maintenance applications, system monitoring, fault diagnosis, and prognosis. He is an author or co-author of five international books, and more than 170 papers in international journals.