Maintenance policy performance assessment in presence of imprecision based on Dempster-Shafer Theory of Evidence

Piero Baraldi, Michele Compare, Enrico Zio

▶ **To cite this version:**

Piero Baraldi, Michele Compare, Enrico Zio. Maintenance policy performance assessment in presence of imprecision based on Dempster-Shafer Theory of Evidence. Information Sciences, 2013, 245, pp.112 - 131. 10.1016/j.ins.2012.11.003 . hal-00757462

HAL Id: hal-00757462

https://centralesupelec.hal.science/hal-00757462

Submitted on 27 Nov 2012

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

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

Maintenance policy performance assessment in presence of imprecision based on Dempster-Shafer Theory of Evidence

P. Baraldi[1], M. Compare[1], E. Zio[1,2,*]

[1] Politecnico di Milano, Italy

[2] Ecole Centrale Paris - Supelec, France

*Corresponding Author: enrico.zio@ecp.fr, enrico.zio@supelec.fr, enrico.zio@polimi.it

Abstract: The aim of this work is to assess the performance of a maintenance policy when a stochastic model of the life of the component of interest is known, but relies on parameters that are imprecisely known, and only through information elicited from experts. The case in which the information used to feed the model comes from a single expert has been investigated by the authors in a previous work. This paper deals with the different situation in which a number of experts are involved in the elicitation of the uncertain parameters; in particular, each expert provides an interval he/she believes containing the unknown value of the parameter which he/she is knowledgeable about. The different type of available information calls for the development of a different method to represent and propagate the associated uncertainty. Resorting to Probability theory to address this issue is questionable. Then, a technique based on the Dempster-Shafer Theory of Evidence (DSTE) is embraced in this work, which allows facing a practical case study concerning the check valve of a turbo-pump lubricating system in a Nuclear Power Plant. The output of such method consists of couples of Lower and Upper cumulative distributions describing the uncertainty in the maintenance performance indicators of interest (i.e., unavailability and costs), which accounts for both the aleatory and epistemic contributions.

Keywords: Maintenance, Uncertainty, Dempster-Shafer Theory of Evidence (DSTE).

1 Introduction

The performance of a given maintenance policy can be a priori evaluated by modeling the behavior of the maintained component. The models developed to this aim rely on a number of parameters which may be weakly known in real applications, due to lack of real/field data collected during operation or properly designed tests. In these cases, the main source of information to estimate these parameters becomes the experts' judgment, which is in general poorly refined. A number of methods and techniques have been proposed in the literature to address the maintenance performance assessment issue in the presence of the imprecision, from different angles:

- Probability distributions have been used to represent the uncertainty in the parameters of the stochastic models of the degradation mechanisms (e.g.,[33]). However, resorting to probability distributions to represent uncertainty due to a lack of knowledge may result in a set of assumptions and biases, with loss of generality ([16], [18], [44]).

- Stochastic Flowgraphs [25] and Hidden Markov Models [35] (both based on the Maximum Likelihood Estimation method) have been proposed to estimate the unknown parameters of the stochastic model of the maintained component, when some field data are missing (e.g., [26], [43]).

- Fuzzy Logic ([47]) 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.,[1]-[3], [27]).

- Theoretical (e.g., [22]) and computational (e.g., [28]) methods have been developed to incorporate the imprecise parameters (e.g., represented by interval probabilities [11], [37], [39] or by fuzzy stets [20]) into markov models.

On the other side, there are other techniques such as DSTE and Possibility Theory (see [5]-[8], [13]-[16], [18], [21], [23], [38], [40], [45], [46], for detailed surveys and comparisons) which are emerging to be more appropriate in describing epistemic uncertainty, though they have never been adopted in Markov models (as pointed out in [37]) nor in other models of the degradation mechanisms ([4]).

To plug this gap, in a previous work ([4]) the authors have applied one such technique, based on the concept of Fuzzy Random Variables (FRVs), to the maintenance policy performance assessment issue. The situation investigated in [4] can be summarized as follows:

- the stochastic model that describes the life of the component of interest, in terms of degradation process, failure behavior and maintenance interventions, is known without any uncertainty.

- The model of the component's behavior depends on a number of ill-known parameters.

- Information about the ill-known parameters is elicited from a single expert, who provides for every uncertain parameter a set of intervals, which contain its true value with different degrees of confidence.

With reference to the latter point, the situation considered in [4] requires that a single expert is knowledgeable, at least qualitatively, on all uncertain parameters and, what is more, he/she is able to provide intervals with associated confidence levels: this may be difficult in some practical cases. In the present work, this restrictive condition is relaxed: different teams of experts, with diverse skills and competences, are involved in the elicitation of the information about the parameters of the model. In particular, each expert is asked to provide an interval that he/she supposes containing the true value of the uncertain parameter. This calls for a proper technique to represent and aggregate the experts' knowledge. Namely, the intervals are treated as random sets, and then combined to build an Evidence Space, according to the DSTE.

The uncertainty representation technique influences the method for uncertainty propagation. To this aim, an hybrid Monte Carlo-DSTE approach is adopted in this work which maps the different combinations of uncertain parameters into some selected summary measures (mean, quantiles, etc.) of the quantities of interest (e.g., unavailability and cost) [23]. Then, a description of the uncertainty on these values is provided in terms of Belief and Plausibility measures. Notice that in this work the issue of establishing an optimal maintenance policy is not addressed. In fact, this requires the availability of logic, mathematical and computational models to perform the additional step of selecting the optimal policy from the point of view of the identified performance indicators, while fulfilling constraints such as those regarding safety and regulatory requirements. In practice, this multi-objective optimization problem has to be faced in a situation in which some constraints and/or the objective functions are affected by uncertainty. To effectively tackle this problem, a number of approaches have been already propounded in the literature considering different framework for uncertainty representation: probability distributions in [12], [17], [24], fuzzy sets in [29] and [42], and plausibility and belief functions in [30]. This problem is not considered in this work.

The remainder of the paper is organized as follows: the main features of the DSTE framework are briefly recalled in Section 2; Section 3 describes the method to represent and propagate the uncertainties focusing on

the representation of the information elicited from the experts. The method is illustrated in Section 4 with reference to the practical case study investigated in [4], concerning a check valve of a turbo-pump lubricating system in a Nuclear Power Plant With the final goal of investigating the potential of this technique for maintenance performance assessment. Section 4 also briefly recalls the case study investigated in [4] with the results obtained in the case in which the epistemic uncertainty on the parameters are neglected. An overall comparison of the method here investigated with that proposed in [4] is provided in Section 5. Some conclusions are given in the last Section.

2 Basics of DSTE

DSTE (also called Theory of Belief Functions) provides a formal structure to process information which is at the same time of random and imprecise nature [8]. In this Section, we provide some basics of this theory to help the understanding of the method proposed in the paper. For further theoretical details, the interested reader may refer to [7], [19], [23], [38], [45], [46].

Assume X is a variable, then DSTE involves the specification of a triplet (S, I, m) ([23]), where S (called 'sample space') is the domain of X; I is a collection of non-empty subsets of S (referred to as 'focal elements'), and m (Basic Probability Assignment, BPA) is a mapping function from the power set of S into the unit interval such that

$$\begin{cases} m(E) > 0 & \text{if } E \in I \\ m(E) = 0 & \text{if } E \notin I \quad \text{and} \quad E \subset S \end{cases} \tag{1.}$$

and

$$\sum_{E \in I} m(E) = 1 \tag{2.}$$

The value of the BPA for any set A represents the portion of all relevant and available evidence that supports the claim that the true value of X lies in A, but to no particular subset of A. That is, the value $m(A)$ pertains to the set A only, and makes no additional claims about any part of A. This entails that if any additional evidence on a subset B of A is available, then it must be represented by another BPA, i.e., $m(B)$ [19].

BPA is analogous to the probability mass function (pmf), which assigns probability masses to a discrete number of points on the real axis. Unlike a pmf, the focal elements of a Dempster-Shafer structure may be intervals overlapping one another, and this is the fundamental difference that distinguishes Dempster-Shafer theory from traditional probability theory [19].

Given a set A included in S, there are two measures, called Belief and Plausibility, that are obtained from m as [45]:

$$Bel(A) = \sum_{E \subseteq A} m(E) \tag{3.}$$

$$Pl(A) = \sum_{E \cap A \neq \varnothing} m(E) \tag{4.}$$

The belief of A is quantified as the sum of the probability masses assigned to all sets enclosed by it; hence, it is a lower bound representing the amount of belief that directly supports the fact that the true value of X lies in A. The plausibility of A is, instead, the sum of the probability masses assigned to all sets whose intersection with the proposition is not empty; hence, it is the mass associated to the possibility that the true value of X is included in A [19], [45].

In Dempster's view [13], BPA encodes the probability family $\mathbf{P}(\nu) = \{P, \forall A \text{ measurable } Bel\ A \le P(A)\} = \{P, \forall A \text{ measurable } P(A) \le Pl(A)\}$. Then,

$$Bel(A) = \inf_{P \in \mathbf{P}(\upsilon)} P(A) \text{ and } Pl(A) = \sup_{P \in \mathbf{P}(\upsilon)} P(A) \quad 5.$$

This entails also that the pair $[Bel(A), Pl(A)]$ can be interpreted as lower and upper probabilities. That is, $\forall P \in \mathbf{P}(\upsilon), Bel(A) \le P(A) \le Pl(A)$ [20]. On this basis, we can define the upper $\overline{F}(x)$ and lower $\underline{F}(x)$ cumulative distribution functions such that $\forall x \in S, \underline{F}(x) \le F(x) \le \overline{F}(x)$, with $\underline{F}(x) = Bel(]-\infty, x])$ and $\overline{F}(x) = Pl(]-\infty, x])$ (i.e., the generic set A in Equations 3.- 5. assumes here the form of $]-\infty, x]$). For the sake of brevity, in the present work these Plausibility and Belief functions are called distributions and are indicated with abuse of notation by $Pl(x)$ and $Bel(x)$, respectively.

3 Uncertainty setting

Let us consider a model $\underline{Z}=g(\underline{Y})$, where $\underline{Z}=(Z^1,...,Z^o)$ is the vector containing the O output variables of interest, and $g(\cdot)$ is a function that models how $\underline{Z}$ depends on the k uncertain variables Y^j, $j=1,2,\cdots,k$, of vector $\underline{Y}$; the uncertainty on these variables is characterized by known probability distributions $F_{Y^j}(y^j;\boldsymbol{\theta}^j)$, $j=1,2,\cdots,k$, where $\boldsymbol{\theta}^j = \{\theta^{j,1},...,\theta^{j,M^j}\}$, are the vectors containing the M^j hyper-parameters of the corresponding probability distributions. Also these parameters are uncertain and the information to characterize them is drawn from experts. This framework of analysis where the aleatory and epistemic components of the uncertainty are separated into two hierarchical levels is often referred to as 'level 2' approach or setting [31]. This makes the present study different from other works of the literature, in which DSTE is embraced to treat epistemic uncertainty directly entering system reliability (e.g., [41]), i.e., with no stochastic variables.

In the problem of assessing the performance of a maintenance policy, g is the model of the life of the component of interest, which encodes random variables (Y^j, $j=1,\cdots,k$) such as the failure time, the repair time, the time of transition from a degradation state to another, etc. The output vector $\underline{Z}$ encompasses the variables $Z^1,...,Z^o$ that are usually considered when assessing the performance of the given maintenance policy, e.g., portion of the mission time in which the component is unavailable, cost, etc. In this work, we are interested in some measures $\underline{\Xi} = \Xi^1,...,\Xi^Q$ such as mean, quantiles, etc., representative of the random variables $\underline{Z}$.

Information elicited from the experts is used to estimate the hyper-parameters $\boldsymbol{\theta}^j$, j=1,…,k of the model g, such as the failure rate of an exponential distribution describing the component failure time. The uncertainty affecting the hyper-parameters is represented and propagated by means of the method discussed in the following.

For the sake of clarity, the uncertainty treatment is described by ways of a simple case study concerning a non-repairable, binary component (i.e., at any time its state can be either working or failed, and once failed the component cannot be repaired), whose Time To Failure (TTF) is Weibull-distributed. In this example, there is *k=1* uncertain variable , i.e., $\underline{Y}=(Y^1)=(TTF)$, described by the Cumulative Distribution Function (CDF) $F_{TTF}(ttf;a,b)=1-e^{-\left(\frac{ttf}{a}\right)^b}$, with $M^1=2$ uncertain parameters $\boldsymbol{\theta}^1=\{a,b\}$, which are the shape and scale parameters of the Weibull distribution, respectively. Since this uncertain setting comprises only one random variable, there is no need to specify that we are referring to it. This allows to simplify the notation, by removing the first superscript from the symbols of both variables and parameters that pertain to a specific random variable, e.g., $\boldsymbol{\theta}$ is used instead of $\boldsymbol{\theta}^j$, and θ^p instead of $\theta^{j,p}$. The output is the portion *D* of the mission time *T*=10^{5}h in which the component is in the down state, whereas $\underline{\Xi}$ contains the value of the mean and the 95th percentile of *D* (i.e., *O*=1 and *Q*=2). Notice that the first quantity is the component average unavailability over the mission time.

The function *g* that links *TTF* to *D* is given by:

$$D=g(TTF)=\begin{cases}\dfrac{T-TTF}{T} & if \quad TTF\le T\\ 0 & otherwise\end{cases} \qquad 6.$$

Then, *D* is also a random variable, whose range of variability is the interval [0,1], and whose distribution, for given values of *a* and *b*, is:

$$F_D(d)=P(D\le d)=P\left(\frac{T-TTF}{T}\le d\right)=P(TTF\ge T(1-d))=e^{-\left(\frac{T(1-d)}{a}\right)^b} \qquad 7.$$

where *d* represents the generic value taken by the variable *D*. Figure 1 shows the shape of this function for *a*=1855h and *b*=7.5, jointly with the values of its mean (circle on the abscissas axis) and 95th percentile (asterisks on the abscissas axis): the component is unavailable at least for 97% of the mission time.

3.1 Information elicited from experts

Let us consider a generic uncertain parameter θ; a number e of experts are asked to provide the intervals $I_i=[\underline{\iota}_i,\overline{\iota}_i]$, $i=1,...,e$ that are believed to contain the true value of θ. The assignments are made on the basis of the experts' experience and independently from one another.

With reference to the simple case study of the non-repairable and binary component with Weibull-distributed *TTF*, there are two uncertain parameters $\theta^1=a$ and $\theta^2=b$. Let us assume that three experts provide estimations of *a* and three experts estimations of *b*. In this respect, notice that the method proposed in this work does not require that the same number of experts are involved in the elicitation of the different parameters. In fact, it allows addressing the more general situation in which the number of experts for each parameter is different; this entails also that the experts are different. Thus, for the sake of generality, we assume that the experts knowledgeable about the first parameter are different from those skilled in estimating the second parameter. Nonetheless, the approach illustrated in this work is capable of addressing the situation where each expert estimates both parameters; this requires simple modifications in some parts of the algorithm we are going to show.

Each of the experts is asked to provide the interval he/she believes containing the true value of the uncertain parameter. Let us suppose that the gathered information is that summarized in Table 1: the interval provided by the *i-th* expert for the *p*-th parameter of the random variable is identified by its lower and upper bounds $\underline{\iota_i^p}$ and $\overline{\iota_i^p}$, respectively.

For the sake of clarity, we remind that in the uncertain settings with *k* random variables $Y^1,...,Y^k$, the notation has to indicate the random variable which the parameters are associated to. Thus, 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.

3.2 Uncertainty representation

According to the procedure proposed in [23], the evidence space (S, I, m) defining the generic uncertain parameter θ is defined by assuming that the sets I_i, $i=1,...,e$ constitute its focal elements. More precisely, S is defined as the domain of the parameter, whereas the set of focal elements results $I=\{I_i, i=1,..., e\}$. Since the BPA assigned to a focal element I_i represents the portion of all available evidence that supports the claim that the true value of the parameter lies in the interval I_i (Section 2), assuming in this case to have available *e* independent and equally credible sources of information, i.e. the *e* experts, the BPA associated to the focal element I_i is assumed to be the fraction of the sources that specified that focal element:

$$m = Kr(I_i)/e \qquad 8.$$

where $Kr(I_i)$ is the number of experts that specified the set I_i. Thus, if all the experts provide different intervals I_i, $i=1,\ldots,e$ the BPA assignment will be $1/e$. In order to further investigate the uncertainty representation provided by this method, let us consider a simplified case where there are two experts providing overlapping intervals $I_1=[a_1,b_1]$ and $I_2=[a_2,b_2]$ with $a_1<b_1<a_2<b_2$. Applying Eqs. (3) and (4), one gets that the probability that the parameter true value lies in $I_1 \cap I_2=[a_2,b_1]$ is between $Bel(I_1 \cap I_2)= 0$ and $Pl(I_1 \cap I_2)=1$. Since for expert 1 the parameter true value can be in the interval $[a_1,a_2]$, (i.e., before the intersection), and for expert 2 the parameter true value can be in the interval $[b_1,b_2]$, (i.e., after the intersection), we do not have any direct evidence that the true value of the parameter belongs to the intersection and thus $Bel(I_1 \cap I_2)= 0$. On the other hand, since both experts do not exclude that the parameter true value lies in $I_1 \cap I_2$, the plausibility of such interval is 1.

With reference to the simple example illustrated above (binary, non-repairable component with Weibull-distributed *TTF*), we have assumed that three experts ($e^1=e^2=3$) provide the intervals they suppose containing the scale and shape parameters of the Weibull distribution (Figure 2 (a) and (c)). These intervals form the set of focal elements $I^1=\{I_1^1, I_2^1, I_3^1\}$ with associated BPA $m^1(I_1^1)=m^1(I_2^1)=m^1(I_3^1)=1/3$ for the scale parameter, and $I^2=\{I_1^2, I_2^2, I_3^2\}$ with associated BPA $m^2(I_1^2)=m^2(I_2^2)=m^2(I_3^2)=1/3$, for the shape parameter. Such assignments determine the corresponding Plausibility (most left) and Belief distributions (most right) $Pl(a)$ and $Bel(a)$, $Pl(b)$ and $Bel(b)$, reported in Figure 2 (b) and (d), respectively.

Again, when the uncertainty setting comprises *k* random variables, the notation needs to be changed. The evidence spaces are indicated as $(S^{j,p}, I^{j,p}, m^{j,p})$, and are 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:

$$m^{j,p} = Kr(I_i^{j,p}) / e^{j,p} \qquad 9.$$

In the case in which the experts are asked to estimate all the M^j parameters of the *j*-th random variable, $j = 1,..., k$, then we consider the focal sets I^j (defined as the M^j-dimensional intervals $I_1^j \times I_2^j \times ... \times I_{M^j}^j$) provided by the experts, which are associated to the probability masses given by:

$$m^j = Kr(I^j) / e^j \qquad 10.$$

3.3 Uncertainty propagation

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 a *O*-dimensional space; this entails that the first step to propagate the uncertainty is to build an evidence space on such *Nu*-dimensional space. According to [23], 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 $I = \{E = I^{1,1} \times .... \times I^{k,M^k} \mid I^{i,j} \subseteq S^{i,j}; \forall i, j\}$. 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:

$$m_I(\mathrm{E}) = \begin{cases} \prod_{\substack{j=1,...k, \\ p=1,...M^j}} m^{j,p}(I_{i^{j,p}}^{j,p}) & \text{if } E = I_{i^{1,1}}^{1,1} \times .... \times I_{i^{k,M^k}}^{k,M^k} \in I \\ 0 & \text{otherwise} \end{cases} \qquad 11.$$

being $i^{j,p} \in \{1,..., e^{j,p}\}$. Notice that the independence assumption adopted in this case, usually referred to as 'random set independence' [10], comes from the assumption that the experts providing estimations of a parameter, are different from those of any other parameter. This assumption has been used in [23] with respect to the mean and standard deviation of a lognormal distribution. Generally speaking, Equation 11. cannot be applied to cases in which different forms of independence hold between the parameters; other formula have been proposed to address these different situations [10], [46], [40]. In particular, when the experts are asked to provide all the M^j parameters related to the *j*-th random variable, thus introducing correlations among them, then the set of focal elements becomes $I = \{E \mid E \in I^1 \times .... \times I^k\}$, and the associated probability masses are given by:

$$m_I(\mathrm{E}) = \begin{cases} \prod_{j=1,\ldots k} m^j(I^j) & \text{if } E = I^1 \times \ldots \times I^k \in I \\ 0 & \text{otherwise} \end{cases} \qquad 12.$$

The universe $S_{\underline{\vartheta}}$ of the evidence space is given by the Cartesian product of the six intervals provided by the experts; *I* is made up of 3*3=9 sets:

$$I = \{E^1 = I_1^1 \times I_1^2, E^2 = I_1^1 \times I_2^2, E^3 = I_1^1 \times I_3^2, E^4 = I_2^1 \times I_1^2, E^5 = I_2^1 \times I_2^2, E^6 = I_2^1 \times I_3^2,$$
$$E^7 = I_3^1 \times I_1^2, E^8 = I_3^1 \times I_2^2, E^9 = I_3^1 \times I_3^2\}.$$

The BPA, m_I, assigns to each of these sets the quantity $\frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}$, which is the product of the probability masses m^1 and m^2 assigned to the intervals.

The methodology to propagate the uncertainties from $\underline{\theta}$ to $\underline{\Xi}$ consists of the following steps [23]:

1. Define a probability distribution *d* on S_I to be used for generating a sample $\underline{\vartheta} = [\vartheta^{1,1}, \ldots, \vartheta^{k,M^k}]$ of $\underline{\theta}$. One way is to define the distributions $d^{j,p}(\vartheta^{j,p})$ for sampling in each $S^{j,p}$, $j = 1, \ldots, k$ and $p = 1, \ldots, M^j$; assuming independence between the parameters, the distribution $d(\underline{\vartheta})$ for sampling $\underline{\vartheta}$ is then defined as $d(\underline{\vartheta}) = \prod_{\substack{j=1,\ldots k \\ p=1,\ldots M^j}} d^{j,p}(\vartheta^{j,p})$. 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:

   $$d^{j,p}(\vartheta^{j,p}) = \sum_{i=1}^{e^{j,p}} m^{j,p}(I_i^{j,p}) U_i^{j,p}(\vartheta^{j,p}) \qquad 13.$$

   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 $\theta^{j,p}$.
   With reference to the example of the binary, non-repairable component with Weibull-distributed *TTF*, a probability mass of 1/9 is assigned to each of the nine elements of the set *I*; in the simplified notation with no reference to the random variable, the uniform distributions $U_i^p(\vartheta^p)$, for *p*=1, 2 and $i = 1, .., e^p$ are given by:

$$U_i^p(\vartheta^p) = \begin{cases} \dfrac{1}{\overline{\iota_i^p} - \underline{\iota_i^p}} & \vartheta^p \in I_i^p \\ 0 & \text{otherwise} \end{cases}$$

where $\underline{\iota_i^p}$ and $\overline{\iota_i^p}$ are the lower and upper bounds, respectively, of the interval I_i^p, reported in Table 1. Then, for example, the probability density function (pdf) $d^1(\vartheta^1)$ of the first parameter is given by:

$$d^1(\vartheta^1) = \sum_{i=1}^{e^1=2} m^1(I_i^1) U_i^1(\vartheta^1) =$$

$$= \begin{cases} \frac{1}{3}\cdot\frac{1}{(\overline{\iota_2^1}-\underline{\iota_2^1})} = 0.0066 & \text{if} \quad \vartheta^1 \in I_2^1 \setminus (I_1^1 \cup I_3^1) = [1820,1830] \\ \frac{1}{3}\cdot\frac{1}{(\overline{\iota_2^1}-\underline{\iota_2^1})} + \frac{1}{3}\cdot\frac{1}{(\overline{\iota_3^1}-\underline{\iota_3^1})} = 0.0122 & \text{if} \quad \vartheta^1 \in (I_2^1 \cap I_3^1) \setminus I_1^1 = [1830,1840] \\ \frac{1}{3}\cdot\frac{1}{(\overline{\iota_1^1}-\underline{\iota_1^1})} + \frac{1}{3}\cdot\frac{1}{(\overline{\iota_2^1}-\underline{\iota_2^1})} + \frac{1}{3}\cdot\frac{1}{(\overline{\iota_3^1}-\underline{\iota_3^1})} = 0.2056 & \text{if} \quad \vartheta^1 \in I_1^1 \cap I_2^1 \cap I_3^1 = [1840,1870] \\ \frac{1}{3}\cdot\frac{1}{(\overline{\iota_3^1}-\underline{\iota_3^1})} + \frac{1}{3}\cdot\frac{1}{(\overline{\iota_1^1}-\underline{\iota_1^1})} = 0.0139 & \text{if} \quad \vartheta^1 \in (I_1^1 \cap I_3^1) \setminus I_2^{1,} = [1870,1880] \\ \frac{1}{3}\cdot\frac{1}{(\overline{\iota_3^1}-\underline{\iota_3^1})} = 0.0055 & \text{if} \quad \vartheta^1 \in I_3^1 \setminus (I_1^1 \cup I_2^1) = [1880,1890] \\ 0 & \text{otherwise} \end{cases}$$

where the symbol '\' indicates the set of all elements belonging to the set on the left but not to that on the right. Figure 3 shows the pdfs corresponding to both shape and scale parameters of the Weibull distribution. For visualization, these are reported in different scales.

Notice that in practical applications, the actual definition of the distribution is unimportant in the sense that the procedure will converge towards correct estimations of the uncertainty on the variables of interest as the sample size increases without bound, although the rate of convergence may be significantly affected by the choice of the sampling distribution [23].

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 is done by sampling, for each $j = 1,...,k$, two uniform random numbers r_1 and r_2 from [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 inverse transform method [32], a value $\vartheta^{j,p}$ in consistency with the definition of the density function $U_i^{j,p}$. Notice that in the case in which the experts provide all the parameters related to a random variable, this step consists in sampling first the multi-dimensional interval I^j (with a probability m^j) and then a point from it.
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, $j = 1,\cdots,k$ in order to obtain the uncertainty on the output variables Z^o, $o=1,...O$. This requires simulating the model behavior a large number N_T of times. Since 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 $\underline{\Xi}=(\Xi^1,...,\Xi^Q)$.

   In regard to the reference example, the sample $\underline{\theta}$=[1855,7.5] defines the Weibull distribution plotted in Figure 1. The values of its mean and 95th percentile (also plotted in the Figure) are considered representative of the entire distribution.

4. Repeat steps 2-3 a large number N_S of times. Notice that N_S must be large enough to assure that at least one point is sampled in each of the focal elements E of the multi-dimensional space $S_{\underline{\theta}}$. In particular, in the reference example N_S=10000. This allows giving full account to the imprecision in the parameters.
5. Estimate the Plausibility and Belief distributions $Pl_{\Xi^q}(\xi^q)$ and $Bel_{\Xi^q}(\xi^q)$ of the Ξ^q components of $\underline{\Xi}$, $q=1,...,Q$. This is done by identifying for every ξ^q the set $INV(\Xi)^q = \{E \subset I : E \cap \Xi^{q-1}(]-\infty, \xi^q]) \neq \varnothing\}$, that is, 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 distributions are given by:

$$Pl_{\Xi^q}(\xi^q) = \sum_{E \in INV(\Xi^q)} m_I(E) \qquad 14.$$

and

$$Bel_{\Xi^q}(\xi^q) = 1 - Pl^c(\xi^q) \qquad 15.$$

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

Finally, notice that both distributions $Pl_{\Xi^q}(\xi^q)$ and $Bel_{\Xi^q}(\xi^q)$ reach one in correspondence of high values of ξ^q (i.e., Equations 14. and 15. must sum to 1 as ξ^q approaches the upper limit of its UoD).

With respect to the binary, non-repairable component with Weibull-distributed *TTF*, the Plausibility and Belief distributions of both mean and 95th percentile of *D* are reported in Figure 4. The first step of the Plausibility function of the 95th percentile of *D* falls in correspondence of d=0.9853, and is of 1/9. The two samples out of N_T=10000 of uncertain parameters that lead to Weibull distributions whose 95th percentile is smaller than d are reported in Table 2. These points belong to the focal set E^7. The fact that there are only two samples from this set is due to the small probability mass associated to the interval [1880,1890] (i.e., the value of the pdf is small and the interval is short, Figure 3, left). The next 1/9 step of the Plausibility distribution is in correspondence of d=0.9854; the set of points $\underline{\theta}$ leading to distributions with 95th percentile smaller than d comprises 35 points belonging to both E^7 and E^1. Finally, at d= 0.991, $INV(\Xi^2)$ encodes points belonging to all focal sets, and the Plausibility distribution reaches 1 (Equation 11). In this respect, notice that if the number N_T of sampled combinations of parameters is low, then Equations 14. and 15. do not sum to 1.

A final consideration deals with the representation of the uncertainty provided by the DSTE-based method. Namely, the Plausibility and Belief distributions encode both epistemic and aleatory uncertainty; the former is due to the uncertainty in the parameters of the model, the second to the stochastic behaviour of the component. However, it should be borne in mind that these functions are affected by the estimation error of the MC method, which can be reduced by increasing the number of simulations ([32], [34]). To wit, Figure 5 shows 10 different couples of the Plausibility and Belief distributions of the mean of *D*, in correspondence of N_T=1e3 (left) and N_T=1e5 (right). In the former case, the curves appear more sensitive to the particular simulation than in the latter case. This stems from the reduction of the MC method estimation error when the

number of simulations increases, which leads to more stable values of the mean of D and ultimately to shapes of the distributions more precisely defined. Moreover, the larger variability of the mean values of D due to MC estimation error, which may lead to both overestimations and underestimations of the mean, reflects into more spread out cumulative distributions (Figure 5, left).

4 Case Study

In this Section, the case study considered in [4] is briefly reported; it is derived from [48] and concerns the degradation and maintenance of a check valve of a turbo-pump lubricating system in a Nuclear Power Plant. This component is affected by one principal degradation mechanism, i.e., fatigue, and only one failure mode, i.e., rupture. The degradation process is modelled as a discrete-state, continuous-time stochastic process which evolves among the following three degradation levels (Figure 6):

1. 'Good': a component in this state is new or almost new (no crack is detectable by maintenance operators).
2. 'Medium': if the component is in this degradation level, then it is convenient to replace it.
3. 'Bad': a component in this degradation state is very likely to experience a failure in few working hours.

The choice of describing the degradation process by means of a small number of levels, or degradation 'macro-states', is driven by industrial practice: experts usually adopt a discrete and qualitative classification of the degradation states based on qualitative interpretations of symptoms.

The probability density functions (pdfs) of the transition times are Weibull distributions, with scale parameters η_{ij} and shape parameters β_{ij} for the transitions from state i towards state j ($i, j \in \{1, 2, 3\}$, and $i<j$). The Weibull distribution is commonly applied in fracture mechanics, especially under the weakest-link assumption[36].

A further state, 'Failed', can be reached from every degradation state upon the occurrence of a shock event. The exponential distribution with constant failure rate λ_j describes the failure behaviour of the component while it is in state j, for every j=1, 2, 3. The choice of assigning a constant failure rate to every degradation state is driven by industrial practice: experts are familiar with this setting and comfortable with providing information about the failure rates values.

A Condition Based Maintenance (CBM) policy is applied to the component, which is composed by the following tasks:

- Inspections: these actions, aimed at detecting the degradation state of the component, are considered to last 5h for a cost of 50€. They are the only scheduled actions.
- CBM actions: Preventive Maintenance (PM) actions which are dependent on the result of an inspection action. More precisely, if the component is found to be in state 'Good', no action is performed, whereas if the degradation state is 'Medium' or 'Bad', then the component is replaced and, consequently, the degradation state is taken back to 'Good'. Both these replacement actions are supposed to take 25h and cost 500€, each.
- Corrective Maintenance (CM) actions. The corrective action, performed after a component failure, is assumed to be the replacement of the component. Due to the fact that this event is unscheduled, this action brings an additional duration of 75h and an additional cost of 3500€, with respect to the replacement after an inspection, leading to a total duration of 100h and to a total cost of 4000€. In particular, the additional time may be caused by the supplementary time needed for performing the

procedure of replacement after failure or to the time elapsed between the occurrence of the failure and the start of the replacement actions.

The Inspection Interval (*II*), which is the time span between two successive planned inspections, is the only decision variable considered in this case study; optimization is then directed to the search for the value of the *II* that minimizes the costs and maximizes the availability of the component.

4.1 Maintenance policy performance assessment in case of no epistemic uncertainty on the parameters

The results of the considered case study, obtained in [4] in the unrealistic situation in which the parameters of the model are not affected by epistemic uncertainty, are briefly reminded. Table 3 reports the values of these parameters, which have been taken from [48].

Figure 8 shows the values of the unavailability of the component over time, with the related 68.3% confidence intervals, obtained by applying the MC method with $5*10^4$ trials; the length *Dt* of the bins partitioning the time horizon is 500h, and *II* = 2000h. Notice that the 68.3% confidence intervals are so narrow that they seems to reduce to points. This is due to the combination of small variability of the mean unavailability in this study and the large number of MC simulations performed. The ordinate of Figure 8 reports the average unavailability corresponding to the bins [*Dt*i*, *Dt*(i+1)*[, for *i=0,1,…,T/Dt-1*, which are associated to the points *Dt*(i+1)* in abscissa. Namely, when the MC method is applied, a statistics of the portion of bin in which the component is unavailable is collected in every bin. This statistics describes how the portion of downtime is influenced by the aleatory variability associated to the stochastic model of the component behavior. For example, Figure 7 shows the distribution of the portion *d* of the bin [2000h,2500h[in which the component is unavailable, when the uncertain parameters take the nominal values of Table 3. The collected values of *d* in every bin are then averaged to get the estimation of the mean unavailability in the bin (the values reported in Figure 8). However, the MC method provides just an estimation of the true distribution and thus of its mean; these are affected by an error that can be reduced by increasing the number of MC simulations. According to the central limit theorem [34], the estimation error on the mean value of the unavailability distribution is described by a normal distribution, which tends to a crisp value in its mean as the number of MC simulation increases. When the standard deviation of this normal distribution is added and subtracted to the estimated mean value, then the 68.3% confidence value is determined [32].

The CDF in Figure 7 has two main steps, which can be interpreted by considering a population of identical components:

1. At *d*=0.01, due to the components still working at the end of the previous bin, which are inspected at *t*=2000h and found in degradation state Good. These components, which constitute almost 17% of the entire population, overtake maintenance actions that last 5h (= 1% of *Dt*).

2. At *d*=0.05, due to the components that are found in degradation state Medium or Bad, and thus require maintenance actions that take 25h; these make the component unavailable for 5% of *Dt*.

Obviously, there are other contributions to *D*, which are related to:

- replacement actions of the components failed in the previous bin, and re-set into operation in the current bin; these components cause the smoothly increasing behavior of the CDF between *d*=0.05 and *d*=0.2;

- unavailability due to maintenance actions on components that have failed in the first bin ([0h, 500h[) and are thus inspected between 2000h and 2500h;

- unavailability due to the failure of the components that have already experienced one or more failures in the previous bins.

Notice that the downtime in the bin is always smaller than 20% of its length; this is due to the fact that none of the components of the considered population has experienced more than one failure in the same bin, and the duration of a replacement action is the 20% of the bin.

An oscillation between four main levels of mean bin unavailability can be recognized in Figure 8(a); this behavior can be explained by looking at Figure 8(b), which breaks up the total unavailability into its different constituents: unavailability due to the inspection of the component while it is in degradation state Good, unavailability due to the preventive replacements if it is found in states Medium or Bad at inspection, and unavailability due to the corrective maintenance actions that are performed upon failure.

By considering a population of components of the same type, the comparison of Figure 8(a) and Figure 8(b) shows that the first increase of unavailability, at t=2000h, is mainly due to the corrective maintenance actions that replace the components failed within the time interval [1500h,2000h[. In this respect, notice that at t=1861h, about 63% of the components have already entered in degradation state Medium (by definition, the scale parameter η_{12} of the Weibull distribution is coincident with the $63.21^{th} \approx 63^{rd}$ percentile), and a small number of components have even experienced a further transition towards the state Bad. The values of the failure rates associated to these latter states ($10^{-4}h^{-1}$and $10^{-2}h^{-1}$, respectively), which are larger than that associated to the Good state ($10^{-6}h^{-1}$), explain the increase in the number of components that fail in the interval [1500h,2000h[.

The unavailability in Figure 8(a) reaches the maximum at t=2500h, which refers to the bin [2000h,2500h[; the sources of unavailability in this bin have been discussed above.

In the successive bins, there is an increase of the number of components whose inspection and failure times are shifted with respect to the 'crowd' (i.e., the large number of components experiencing the same behavior), and this leads to more and more smoothed peaks due to replacement of components in Medium degradation state and larger and larger unavailability in the bins that follow these peaks due to the replacement of components both failed and in Medium degradation state. The unavailability due to replacement of components in Bad state and inspections in state Good remains small.

Figure 9 shows the mean unavailability of the component in the mission time, with the related 68.3% confidence interval, for different values of the *II*. Initially, there is a decreasing behavior that reaches a minimum in correspondence of *II*=1000h/1500h; after this point, the unavailability starts rapidly increasing. This is the result of two conflicting trends: on one side, the more frequent are the inspections the larger is the probability to find the component in degradation states Medium and Bad: this prevents the component to fail and thus saves the corresponding large time to replacement. On the other side, frequent replacements are ineffective, since the component life is not completely exploited in this case. The minimum at *II*=1500h represents the optimal balance between these two tendencies.

Figure 10 shows the maintenance costs associated to different choices of the *II*, which have a shape similar to that of the mean unavailability.

4.2 Representation and propagation of epistemic uncertainties

The aim of this Section is to apply the method illustrated in Section 3 to the case study described above, when the parameters of the distributions that model the transitions of the component among the four states of Figure 6 are ill-known, and their evaluation comes (with imprecision) from teams of experts. To sum up, the uncertainty situation is the following:

- there are $k=5$ uncertain variables, which define the 5 transition times reported in Table 4;
- the distributions associated to the variables are known, and depend on the set of the uncertain parameters θ^j, $j=1,...,5$ reported in Table 4. In turn, there are $Nu=7$ uncertain parameters, which are the shape and scale parameters of the two Weibull distributions and the failure rates pertaining to the three degradation levels (see Table 5).

However, the uncertainty on the third failure rate is not considered. In fact, a sensitivity analysis carried out in [4] has showed that the output of the model does not appreciably change when the value of the third failure rate ranges in a wide interval, whereas accounting for a further uncertain parameter strongly increases the computational effort. Namely, in the considered case study there are 3 intervals for each of the 7 uncertain parameters; this entails that the number of focal sets of the evidence space (S_X, I, m_I) is B=3^7=2187.

Thus, if the random sampling method is applied at point 3) of the procedure in Section 3.3, then a number of samples N_S larger than B (e.g., N_S = 15000 ≈ 6, 7 times B) must be drawn to be reasonably sure to get at least one point in every focal set. On the contrary, neglecting the uncertainty on the third failure rate strongly reduces the number of samples (B=729) and the computational time, with a small impact on the estimation of the value of $\underline{Z}$.

With regards to the choice of the number of samples N_S, notice that the larger N_S the larger the number of output mean unavailability values $\underline{\Xi}$, and thus the higher the precision in the identified pair of distributions [*Bel*, *Pl*]. Therefore, setting N_s requires to find an optimal trade off between the precision of the distributions and the need of reducing the computational time. Such optimization will be the subject of future work.

Notice also that the simulation of a single MC history (step 3 of the procedure in Section 3.3) requires that the model g encodes a number of random variables $k>>5$, since the history corresponding to a given sample of these 5 times in general do not cover the entire time horizon T. For example (Figure 11(a)), if the first transition is from state 1 to state 'Failed' and occurs at t=2000h, then the interval time between t=2000h+100h (the time instant at the end of the replacement action that starts after the failure) and T remains not investigated. This problem can be overcome by thinking of g as a function that depends on a number K of 5-ple, and not just on 5 variables; the number K that allows to cover the entire mission time is also a random variable, since it depends on the sampled times, which produce histories of different lengths. However, this is not a problem in practice: the number K can be chosen such that it is reasonably sure that the sampled times simulate histories of duration larger than the time horizon T. Then, the analysis focuses only on the interval [0, T] (Figure 11(b)).

On the other side, once the combination of uncertain parameters relevant to the first 5-ple has been sampled, it remains the same for the entire duration of the history (i.e., all the K 5-ple of samples are drawn from the same probability distributions). This is equivalent to assuming that the components considered in the different simulations have the same stochastic behavior, which is exactly described by the 5 distributions of the transitions among the four states of Figure 6, although their parameters are not exactly known (see also[36]).

The output vector $\underline{Z}$ is made up of portions of downtime in each of the N_{bin} bins partitioning the time axis and in the mission time, and the cost associated to the given maintenance policy; thus $O=2+ N_{bin}$. The summary measures $\underline{\Xi}$ of $\underline{Z}$ we are interested in, are the mean values; thus, also $Q=2+ N_{bin}$. Notice that the mean value of the portion of downtime over a given period represents the average unavailability over the period, i.e., a quantity with a direct interpretation for maintenance decision makers.

4.2.1 Information elicited from experts

In this work, it is supposed that for every uncertain parameter, the same number $e^{j,p}=3$ of experts are involved in the elicitation phase. Each of the experts provides the extreme values of the interval he/she believes containing the true value of the uncertain parameter he/she is asked to estimate. These are reported in Table 5, for every $j=1,...,5$ and $p=1,...,M^j$. For example, with reference to the fourth row of Table 5, the cells of the most right two columns tell us that the third expert involved in the elicitation of the scale parameter of the Weibull distribution that describes the transition from degradation state Good to Medium believes that almost the 63% of the components experience such transition at a time instant in the interval [1720h,2000h]. The other two experts involved in the elicitation of η_{12} are less vague and provide smaller intervals [1815h,1908h] and [1843h,1880h], respectively. Notice that, for the sake of simplicity, for every uncertain parameter, in this case study the intervals provided by the experts are assumed nested (i.e., $I_i^{j,p} \subseteq I_{i+1}^{j,p}$ for every $j=1,...,5$, $p=1,...,M^j$ and $i=1,2$). Finally, the uncertainty on the failure rate corresponding to the degradation state 'Bad' has been not accounted for, given that it has been assumed exactly known.

Finally, notice that for every uncertain parameter, the value considered in Section 4.1 is the middle point of the corresponding intervals provided by the expert.

4.2.2 Description of the epistemic uncertainties

The information elicited from the experts has been used to build, for every uncertain parameter $\theta^{j,p}$, $j=1,...,5$ and $p=1,...,M^j$, an evidence space $(S^{j,p}, I^{j,p}, m^{j,p})$. The sample space $S^{j,p}$ is the union of the three intervals provided by the experts, which in this case is coincident with the largest interval provided by the third expert, whereas the set of focal elements $I^{j,p}$ is made up of these three intervals. Finally, the BPA assigns to every interval the same mass value $m^{j,p}(I_i^{j,p}) = \frac{1}{e^{j,p}} = \frac{1}{3}$.

4.2.3 Uncertainty propagation

The uncertainty propagation procedure described in Section 3.3 has been applied to the considered case study. Figure 12 shows the obtained Plausibility and Belief distributions of the unavailability over the different bins in which the mission time has been divided. In the first bins (i.e., from t=500h to t=1500h), the Plausibility and Belief distributions are very close to each other and reach 1 in correspondence of a value of the mean unavailability very close (or even equal) to 0; this tells us that in those bins, the mean unavailability remains very small for any combination of the values of the uncertain parameters ranging in the intervals provided by the experts.

The situation is different at t=2000h, where both Plausibility and Belief distributions are shifted towards higher values of the unavailability. This is due to the increase of the number of components that experience a failure in the bin [1500h,2000h[, due to components' transitions towards degradation states Medium and Bad, as explained above (see Figure 8). Notice also that the 'distance' between the Plausibility and Belief

distributions is quite large, if compared to those of the first bins. This is due to the fact that the behaviour of the components is heavily influenced by the particular combination of the uncertain parameters. For example, considering that the scale parameter represents the 63rd percentile, a combination of the values $\eta_{12} \cong 720$h and $\eta_{23} \cong 690$h leads to simulated histories in which it is very likely that the components experience a failure before t=2000h, with the consequent unavailability; on the contrary, the combination $\eta_{12} \cong 2000$h and $\eta_{23} \cong 800$h results in histories in which more rarely there is a failure in the bin [1500h,2000h[.

In the next bin, t=2500h, the distributions are even more shifted toward the right part of the unavailability axis, which is in agreement with the behaviour of the unavailability in the case with no uncertainty on the model parameters (see Figure 8). In the successive bins (Figure 12) the Plausibility and Belief distributions follow the 'cycle' of the first bins; for example, the curves relevant to the bin [1500h,2000h[are similar to the corresponding ones in the bin [3500h,4000h[; the differences between the Plausibility distributions and the Belief distributions pertaining to 'similar' bins are due to the increase in the number of components that experience a life different from that of the 'crowd', as explained in Section 4.1.

In an effort to render more concise the information presented by the distributions in Figure 12, Figure 13 shows, for every bin, the intervals bounded by the values of the median of the Plausibility and Belief distributions of the average unavailability in the bins. That is, the extremes of the intervals constitute the lower and upper bounds, respectively, of the 50th percentile of the average unavailability in the bins. For comparison, the estimations of the average unavailability over the bins found in Section 4.1 are also provided in Figure 13.

Figure 14 and Figure 15 report the Plausibility and Belief distributions of the mean unavailability and cost over the mission time, respectively, corresponding to three different values of the II, i.e. II=1000h, II=1500h and II=2000h. In particular, the small amount of uncertainty on the values of both unavailability and costs, when the component is inspected every 1000h, derives from the fact that the 'crowd' remains very compact in this case. From these Figures, it clearly appears that when the maintenance optimization problem is faced in presence of uncertainty, the identification of the best maintenance policy is not a trivial problem. For example, establishing whether the performance corresponding to II=1500h is better than that associated to II=2000h is an open issue which needs to be addressed.

5 Comparison with the FRVs-based method

In this Section, the method proposed in this work to represent and propagate uncertainties is firstly compared to that considered in [4], based on the concept of FRVs. First of all, we consider how demanding these methods are with respect to the elicitation of the information from experts. With regards to the hybrid MC-DSTE method discussed in this work, each expert is asked to provide just an interval, which he/she believes the true value of the ill-known parameter belongs to. In particular, different teams of experts can be involved in the elicitation of the information about different parameters, which allows to exploit the diverse skills that may be needed for estimating them. On the contrary, in the FRVs-based method applied in [4] it is required that a single expert is knowledgeable, at least qualitatively, on all uncertain parameters and, what is more, is familiar with the statistical meaning of confidence levels when providing the weighted families from which the possibility distributions are built: this may be very difficult in practice. However, the FRVs-based method can be also applied when the expert gives just one interval per parameter; in this case, the information is equivalent to that given for the hybrid MC-DSTE method in case of only one expert per parameter, and both methods can be applied based on the same information. Furthermore, if the possibility distributions are transformed into evidence spaces using one of the techniques proposed in the literature to

this aim (e.g.,[16]), then the FRV-based method can be used also with the information provided for the MC-DSTE method.

To conclude the considerations about the elicitation of the information from experts, notice that in this case study, independently on the applied method, the experts are supposed to be able to provide information on the parameters of both the Weibull and exponential distributions, and this may be very difficult in practice. In fact, as pointed out in [4], while it is plausible that an expert is able to estimate the time until which almost the 65% of the components have experienced a transition (i.e., the scale parameter of the Weibull distribution), it seems very unlikely that he/she knows the shape parameters of these distributions (which are the slopes of the Weibull probability plots). Also the estimation of the mean time to failure of the components in a given degradation state (i.e., the inverse of the failure rate) may not be easy; in fact, failure from the first degradation state is usually a rare event, whose frequency is difficult to estimate even in a qualitative way, whereas the lack of precise knowledge of the time instants in which the components transit towards the other degradation states affects the evaluation of the mean times to failure associated to these states; that is, if the time instant since one has to start to count is unknown, then the resulting measure of the time to failure is biased, especially if the component is rarely inspected. Moreover, both the assumptions that the transitions between the degradation states are Weibull-distributed and that in a given degradation level the failure times are exponentially distributed, may not hold. But this is not the matter of this work.

From the point of view of the output, the two methods provide different kinds of information. Namely, the hybrid MC-DSTE method first collects in every bin partitioning the time axis a statistics of the unavailability of the component, which is associated to a certain combination of the uncertain parameters and thus to a certain probability mass, and then summarizes this statistics by its mean value (or another measure such as median, 95^{th} percentile, etc.); this is assumed to suffice to describe the whole informational content of the collected data. On the contrary, the method of FRVs considers the entire range of values taken by the unavailability in the bins, and on this basis identifies the lower and upper probability distributions. Given this difference, a comparison between the results does not make any sense.

As pointed out in [4], one of the main shortcomings of the FRV-based method is that the results provided (i.e., the entire range of values of downtime over a period) are difficult to be understood. On the contrary, the MC-DSTE based method provides an estimation of the mean unavailability (i.e., a quantity which the maintenance decision makers are familiar with) over the period of interest. This seems to be an advantage of the MC-DSTE based method with respect to the FRV-based one.

A drawback of both methods lies in the very large memory demand and computational times required, which ensue from the complexity of the algorithm. In fact, this requires that a number N_T of MC trials are simulated to capture the aleatory uncertainty of the system for each of the N_S samples from the Nu-dimensional space of the uncertain parameters (i.e., N_S*N_T simulations). Moreover, the mapping between the output and the Nu-dimensional space (step 5 of the procedure in Section 3.3) is burdensome.

Table 6 reports the computational times of the method in case of N_T=2000 and N_S=10000 combinations of values of the uncertain parameters. These are similar to those relevant to the FRV-based method ([4]). Notice that the choice of the number of combinations, which heavily influences the computational time, should be driven by the number Nu of uncertain parameters considered in the model and by the number of nested intervals used to describe their uncertainty. However, being Matlab an interpretative language, a tool developed in other environments may be more performing; on the other side, the application of the latin hypercube sampling technique at point 3) of the procedure in Section 3.3 or a technique that forces the sampling in the regions in which a more refined investigation is required, could be considered to increase the efficiency of the computation. These issues will be tackled in future works.

6 Conclusions

Uncertainty affects the parameters of the models used to assess the performance of a given maintenance policy. An incorrect treatment of such uncertainty may lead to serious bias of the outcome of the analysis, possibly non-conservatively. The method investigated in this work offers an effective tool to give due account to the uncertainties on the parameters of the maintenance model of the component of interest. Compared with another method already investigated by the authors, it seems less demanding from the point of view of the information to be elicited from experts, and provides results that are more understandable for maintenance practitioners.

The methodology has been applied to a case study concerning the degradation model of a check valve of a turbo-pump lubricating system in a Nuclear Power Plant. The study has shown that neglecting uncertainty may drive the maintenance decision maker towards incorrect conclusions. In this case, if the unavailability computation were performed without taking into account the uncertainty on the input parameters, the decision maker would set the inspection intervals between maintenance actions to the value of II=1000h, whereas a proper consideration of the uncertainties suggests that, on the basis of the available knowledge, this choice for the maintenance inspection interval is not better than other intervals such as II=1500h.

To conclude, some interesting issues remain open:

- The information provided by both methods in general does not allow to make a decision in a simple way. Thus, how to exploit these results from the decision maker point of view remains an open issue, which needs to be addressed in future works.
- The computational time can be very large. For its reduction, two research directions have been identified: development of methods to select and disregard the parameters whose uncertainties weakly affect the output of the model, and techniques to choose the optimal sample numbers N_s, which conciliates the demand for precision in the results and the need for small computational times.

7 References

[1]. An, M., Chen, Y., Baker, C.J. A fuzzy reasoning and fuzzy-analytical hierarchy process based approach to the process of railway risk information: A railway risk management system. Information Sciences Volume 181 (18), (2011) 3946–3966.

[2]. Baraldi, P., Balestrero, A., Compare, M., Benetrix, L., Despujols, A., Zio, E., A Modeling Framework for Maintenance Optimization of Electrical Components Based on Fuzzy Logic and Effective Age, Accepted on Quality and Reliability Engineering International.

[3]. Baraldi, P., Compare, M., Rossetti, G., Despujols, A., Zio, E. 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.

[4]. Baraldi, P., Compare, M., Zio E. Representation and propagation of the uncertainty in expert information on the parameters of degradation models for maintenance policy assessment. Submitted for publication to Applied Soft Computing (2011).

[5]. Baraldi, P., Zio, E. A comparison between probabilistic and Dempster-Shafer Theory approaches to Model Uncertainty Analysis in the Performance Assessment of Radioactive Waste Repositories. Risk Analysis 30 (7) (2010) 1139-1156.

[6]. Baudrit, C., Dubois, D. Comparing methods for joint objective and subjective uncertainty propagation with an example in a risk assessment, In Fabio Gagliardi Cozman, Robert Nau, Teddy Seidenfeld (Eds.), proc. of the 4th Int. Symp. on Imprecise Probabilities and Their Applications (ISIPTA 2005), Carnegie Mellon University, Pittsburgh, PA, USA, 2005, pp. 31-40.

[7]. Baudrit, C., Dubois, D., Guyonnet D. Joint Propagation and Exploitation of Probabilistic and Possibilistic Information in Risk Assessment. IEEE transactions on fuzzy systems 14 (5) (2006) 593-608.

[8]. Baudrit, C., Dubois, D., Perrot, N. Representing parametric models tainted with imprecision. Fuzzy Sets and Systems 159 (2008) 1913-1928.

[9]. Borgonovo, E., Apostolakis, G.E., Tarantola, S., Saltelli, A. Comparison of global sensitivity analysis techniques and importance measures in PSA. Reliability Engineering and System Safety 79 (2003) 175-185.

[10]. Couso, I., Moral S., Walley, P. Examples of Independence for Imprecise Probabilities. In Gert De Cooman, Fabio Gagliardi Cozman, Serafín Moral, Peter Walley (Eds.), proc of the 1st Int. Symp. on Imprecise Probabilities and Their Applications (ISIPTA 1999), Universiteit Gent, Ghent, Belgium, 1999, pp. 121-130

[11]. De Cooman, G., Hermans, F., Quaeghebeur, E. Imprecise Markov Chains And Their Limit Behaviour, Probability in the Engineering and Informational Sciences 23 (2009) 597-635.

[12]. Deb, K., Gupta, S., Daum, D., Branke, J., Mall, A.K., Padmanabhan, D. Reliability-Based Optimization Using Evolutionary Algorithms. IEEE Transactions on Evolutionary Computation13 (5) (2009) 1054-1074.

[13]. Dempster, A.P. Upper and lower probabilities induced by a multi-valued mapping. Annals of Mathematical Statistics 38 (1967) 325-339.

[14]. Dubois, D. Possibility Theory and Statistical Reasoning. Computational Statistics and Data analysis 51 (2006) 47-69.

[15]. Dubois, D., Prade, H. Possibility theory, probability theory and multiple valued-logics: A clarification. Annals o Mathematics in Artificial Intelligence 32 (2001) 35-66.

[16]. Dubois, D., Prade, H., Smets, P. Representing partial ignorance. IEEE Transaction System Man, Cybernetic 26 (3) (1996) 361–377.

[17]. Eskandari, H., Geiger, C.D., Bird, R. Handling uncertainty in evolutionary multiobjective optimization: SPGA, in Proceedings of IEEE Congress on Evolutionary Computation, CEC 2007, 25-28 Sept. 2007, Singapore, pp. 4130 – 4137. (2007)

[18]. Ferson, S, Ginzburgb L.R. Different methods are needed to propagate ignorance and variability. Reliability Engineering & System Safety 54 (2-3) (1996) 133-144.

[19]. Ferson, S., Kreinovich, V., Ginzburg, L.R., Myers, D.S., Sentz, K. Constructing Probability Boxes and Dempster-Shafer Structures. SAND REPORT SAND2002-4015, (2003).

[20]. Ge, H., Asgarpoor, S., Reliability Evaluation of Equipment and Substations with Fuzzy Markov Processes, IEEE Transaction on Power System, 25 (3) (2010) 1319-1328.

[21]. Haenni, R., Lehmann, N. Implementing belief function computations. Journal of intelligent systems 18 (1) (2003) 31-49.

[22]. Hartfiel, D.J. Markov set-chains, Springer, Berlin, 1998.

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

[24]. Hughes, E.J. Evolutionary multi-objective ranking with uncertainty and noise. Lecture Notes in Computer Science LNCS 1993, E. Zitzler et al. (Eds), Springer-Verlag Berlin Heidelberg. (2001)

[25]. Huzurbazar, A.V., Flowgraph Models for Multistate Time-to-Event Data, John Wiley & Sons, 2005.

[26]. Huzurbazar, A.V., Williams B.J., Incorporating Covariates in Flowgraph Models: Applications to Recurrent Event Data, Technometrics, 52(2) (2010) 198-208.

[27]. Jones, B., Jenkinson, I., Wang, J. 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 224 (2010) 35-48.

[28]. Kozine, I.O., Utkin, L.V. Interval-valued finite Markov chains, Reliable Computing 8 (2) (2002) 97–113.

[29]. Li, J., Kwan, R.S.K. A fuzzy genetic algorithm for driver scheduling. European Journal of Operational Research, 147 (2) (2003) 334-344.

[30]. Limbourg, P. Multi-Objective Optimization of Problems with Epistemic Uncertainty. Lecture Notes in Computer Science LNCS, 3410 (2005) 413-427.

[31]. Limbourg, P., de Racquigny, E. Uncertainty analysis using evidence theory- confronting level-1 and level-2 approaches with data availability and computational constraints. Reliability Engineering and System Safety 95 (2010) 550-564.

[32]. Marseguerra, M., Zio, E. Basics of Monte Carlo Method with application to System Reliability. LiLoLe-Verlag, 2002.

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

[34]. Papoulis, A., Pillai, U. Probability, Random Variables, and Stochastic Processes. 4th Edition. Mc Graw-Hill, 2002.

[35]. Rabiner, L.R., Juang, B.H. An Introduction to Hidden Markov Models, IEEE ASSP Magazine 3 (1) (1986) 4-16.

[36]. Remy, E., Idée, E., Briand, P., François, R. Bibliographical review and numerical comparison of statistical estimation methods for the three-parameters Weibull distribution. In Ale, B., Papazoglu, I. & Zio, E. (eds.), proc. of the European Safety and Reliability Conference (ESREL 2010), Rhodes, Greece, 2010, pp.219-228.

[37]. Rocco S.C.M. 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 226 (1) (2011) 234-245.

[38]. Shafer, G. A mathematical theory of evidence. Princeton University Press. Princeton, NJ, 1976.

[39]. Škulj, D., Discrete time Markov chains with interval probabilities. International Journal of Approximate Reasoning 50 (2009) 1314–1329.

[40]. Su, Z.G., Wang, P.G., Yu, X.J. Lv, Z.Z. Maximal confidence intervals of the interval-valued belief structure and applications. Information Sciences Volume 181 (9), (2011) 1700–1721.

[41]. Tonon, F., Bae, H., Grandhi, R.V., Pettit, C.L. Using random set theory to calculate reliability bounds for a wing structure. Structure and Infrastructure Engineering 2 (3-4) (2006) 191-200.

[42]. Trebi-Ollennu, A., White, B.A. Multiobjective Fuzzy Genetic Algorithm Optimization Approach to Nonlinear Control System Design. IEE Proceedings in Control Theory and Applications, 144 (2) (1997) 137-142.

[43]. Vrignat, P., Avila, M., Duculty, F., Aupetit, S., Slimane, M., Kratz, F. Maintenance policy: degradation laws versus Hidden Markov Model availability indicator, Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 226(2) (2011) 137-155.

[44]. Wu, J.S., Apostolakis, G.E., Okrent, D. Uncertainty in System Analysis: probabilistic versus non-probabilistic theories. Reliability Engineering and System Safety 30 (1990) 163-181.

[45]. Yager, R.R. Modeling uncertainty using partial information. Information Sciences 121 (1999) 271-294.

[46]. Yager, R.R. On the fusion of imprecise uncertainty measures using belief structures. Information Sciences 181 (2011) 3199-3209.

[47]. Zadeh, L.A. Fuzzy Sets. Information and Control 8 (1965) 338-353.

[48]. Zille, V., Despujols, A., Baraldi, P., Rossetti, G., Zio, E. A framework for the Monte Carlo simulation of degradation and failure processes in the assessment of maintenance programs performance.

In Bris, R., Guedes Soares, C & Martorell, S. (eds), proc. of the European Safety and Reliability Conference (ESREL 2009), Praha, Czech Republic, 2009, pp. 653-958.

Figure Captions

Figures

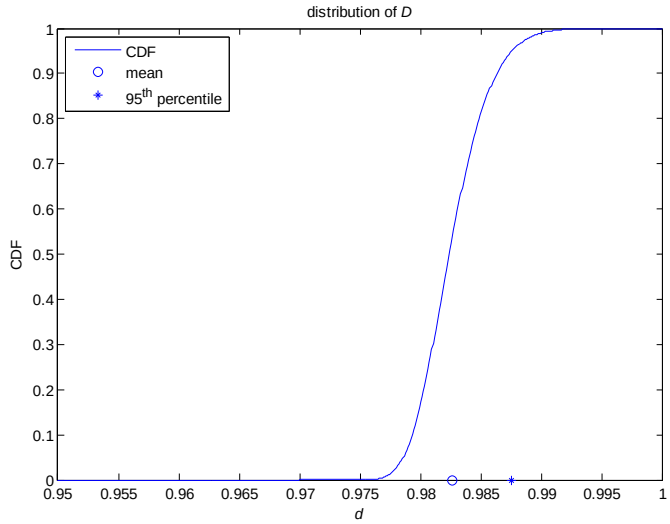

Figure 1

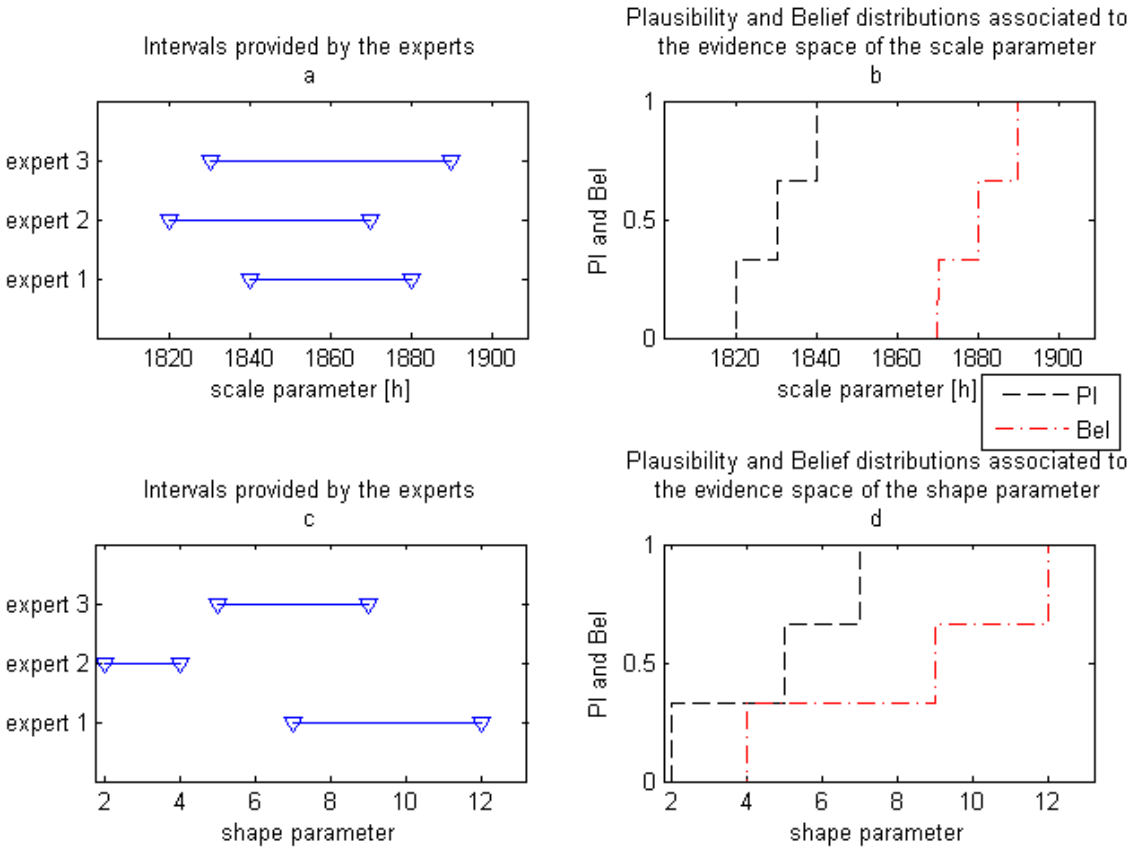

Figure 2

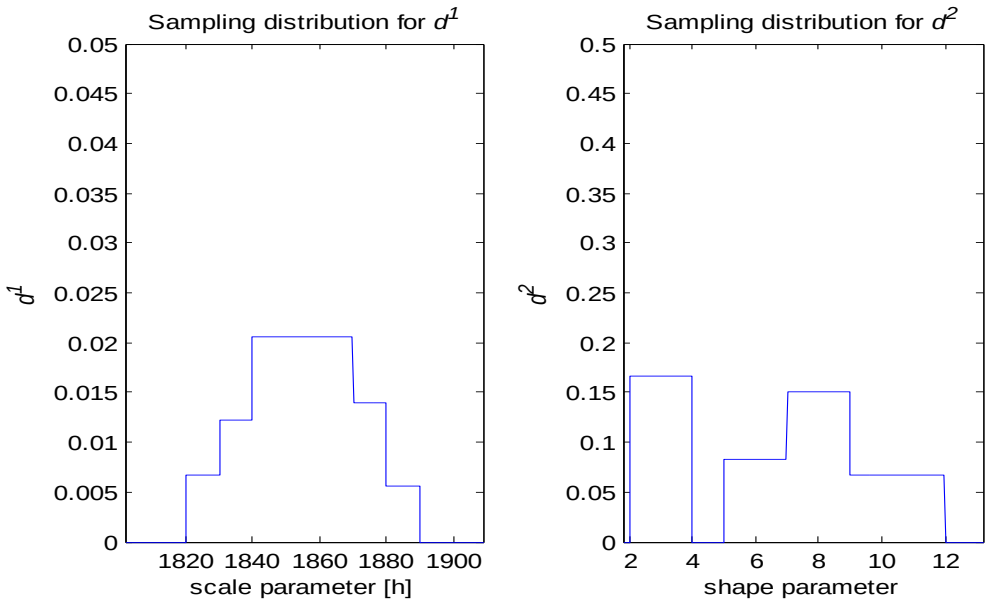

Figure 3 (a) and (b)

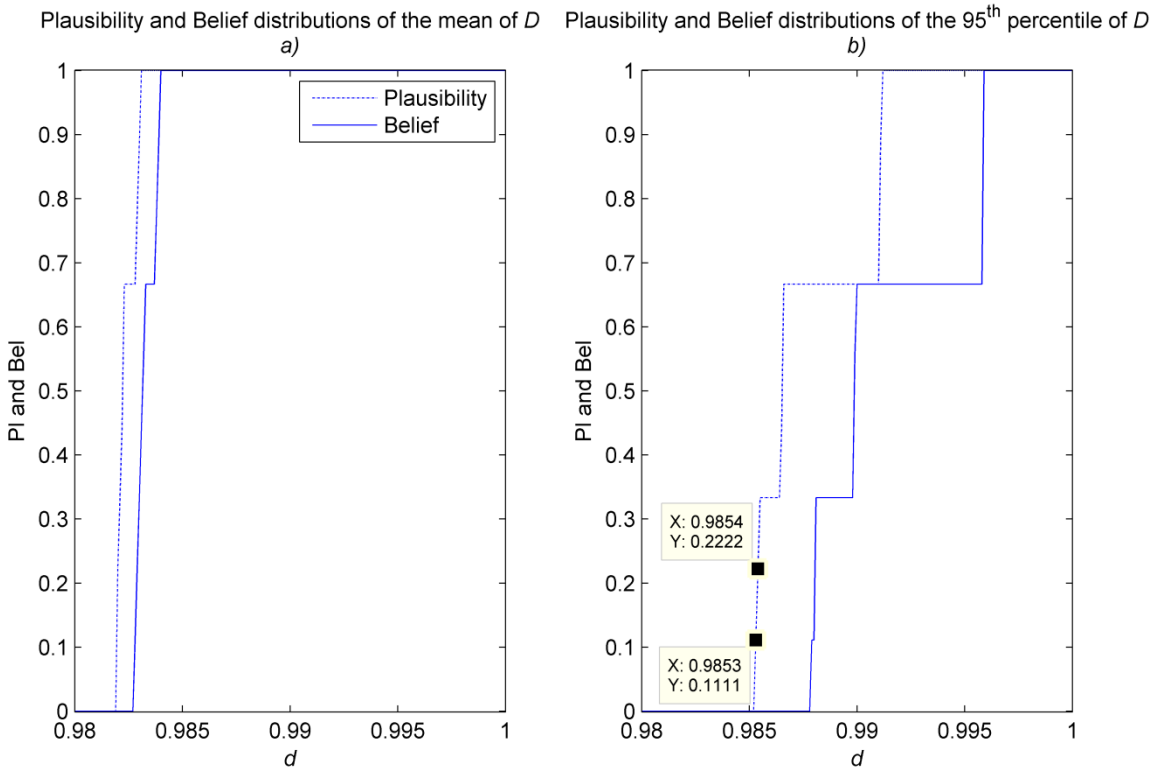

Figure 4 (a) and (b)

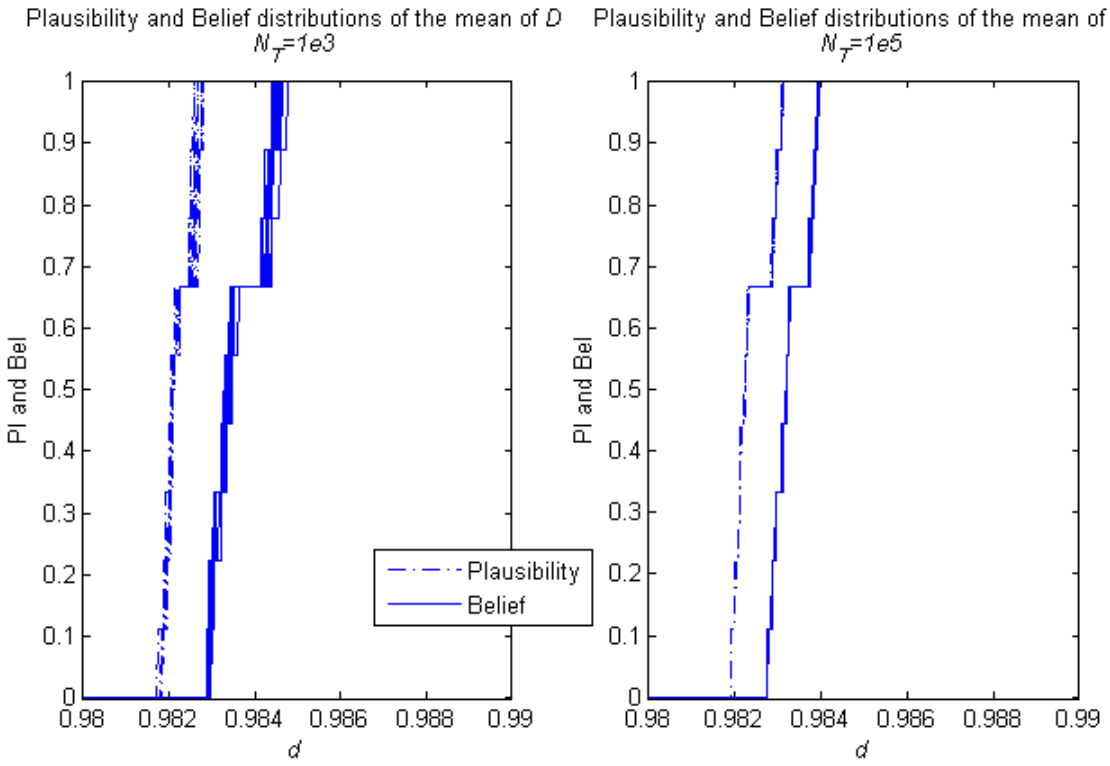

Figure 5

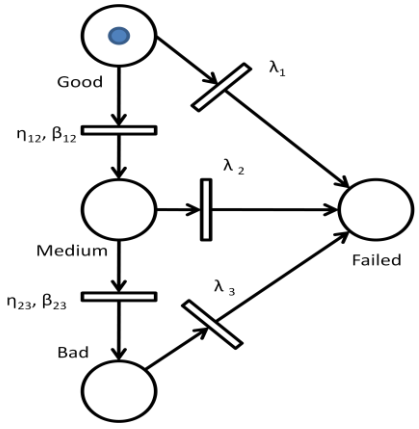

Figure 6

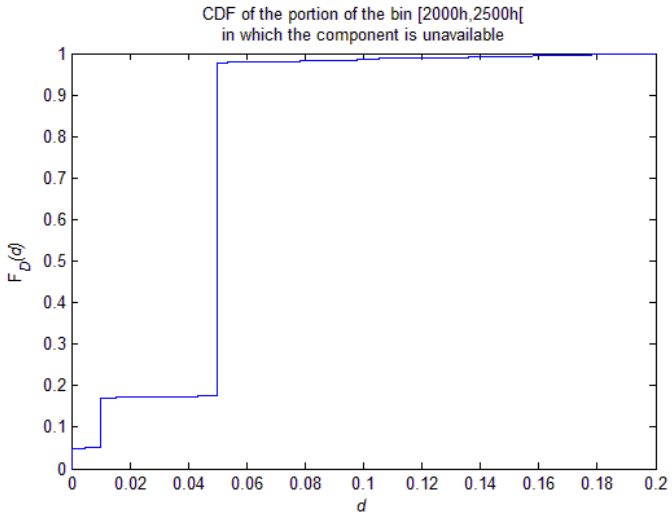

Figure 7

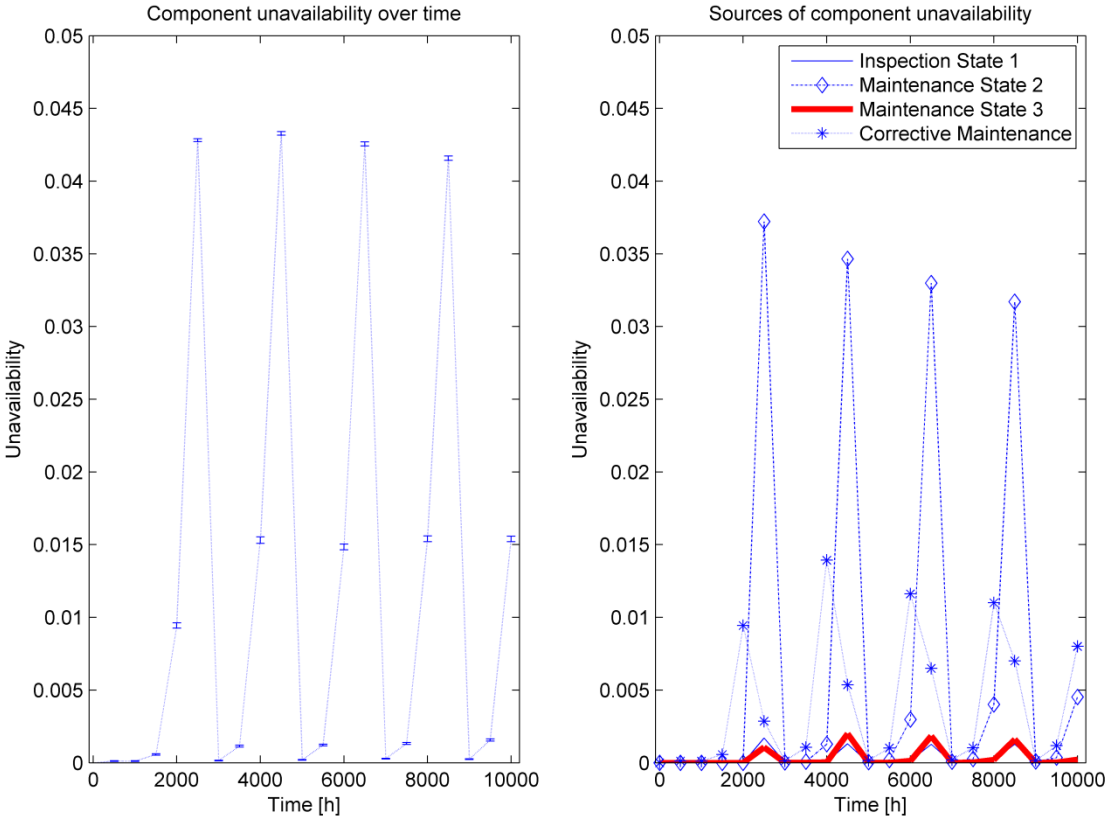

Figure 8 (a) and (b)

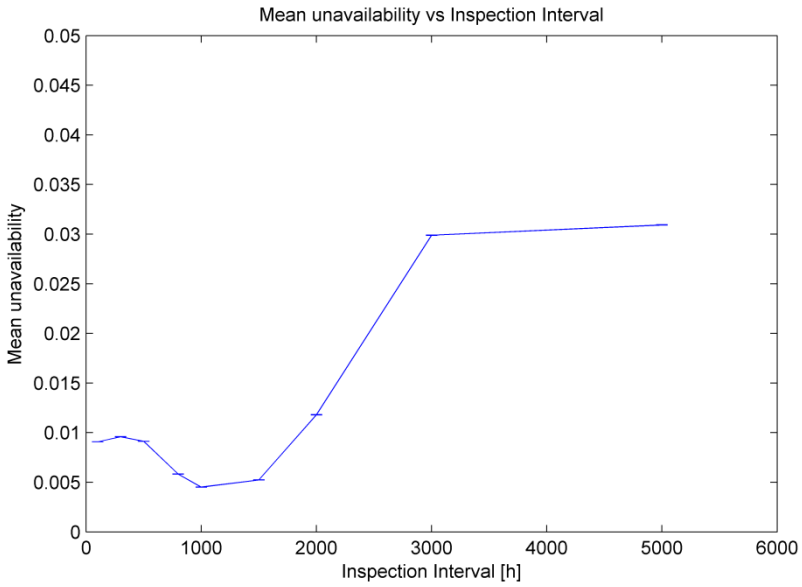

Figure 9

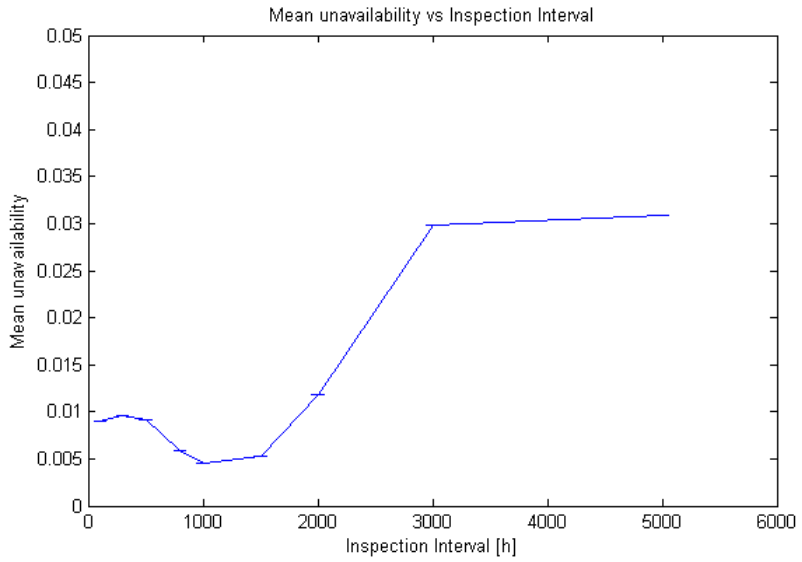

Figure 10

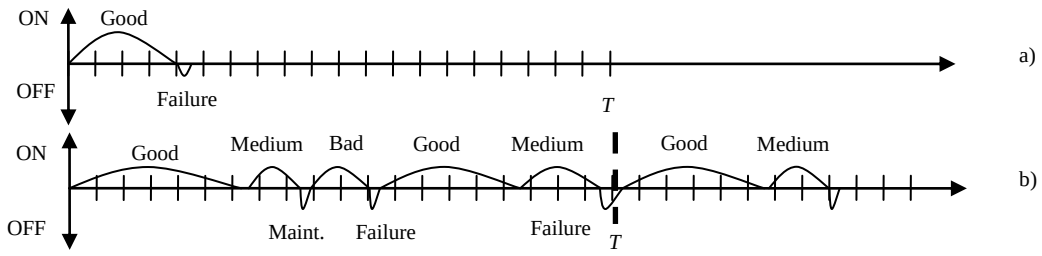

Figure 11

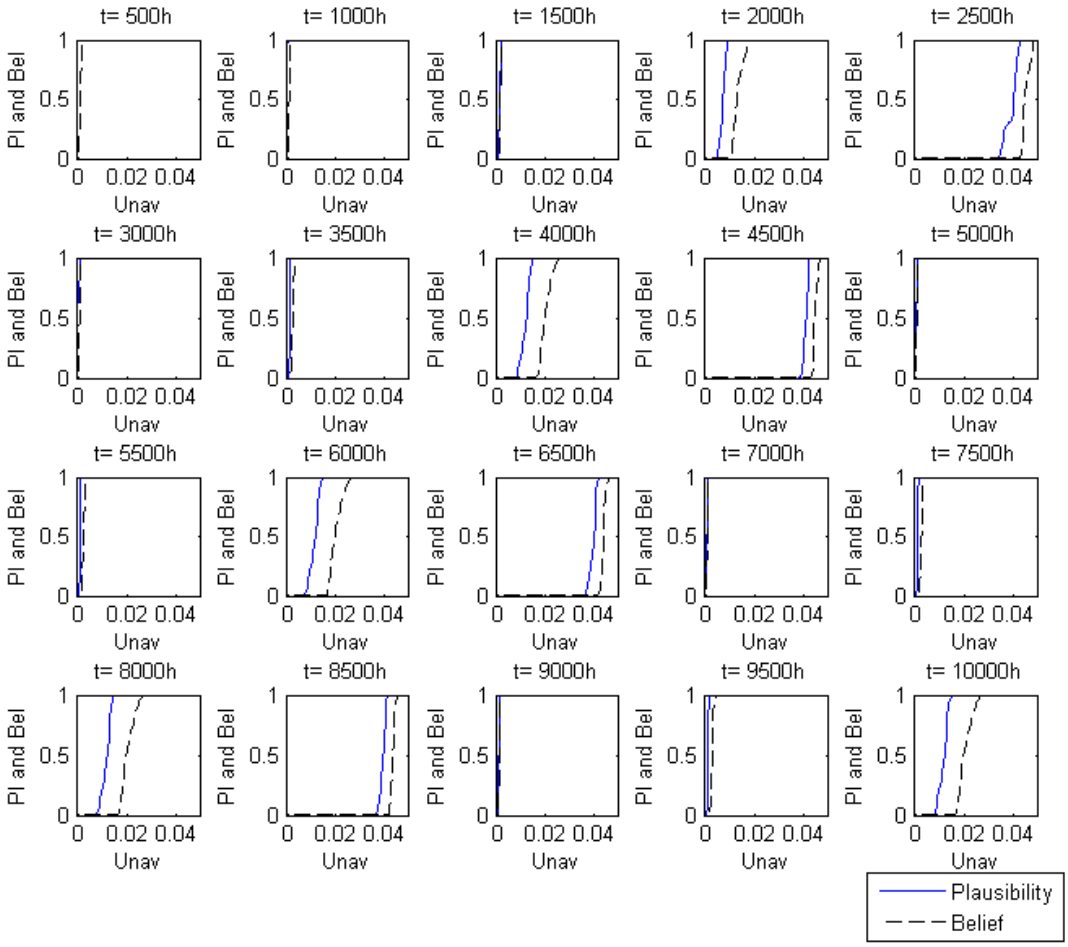

Figure 12

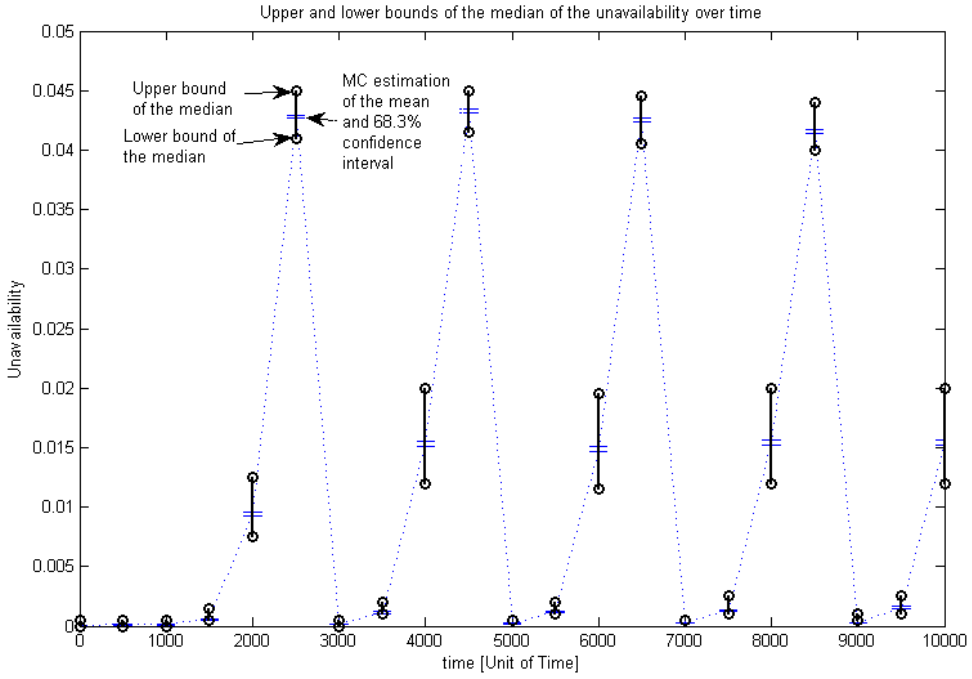

Figure 13

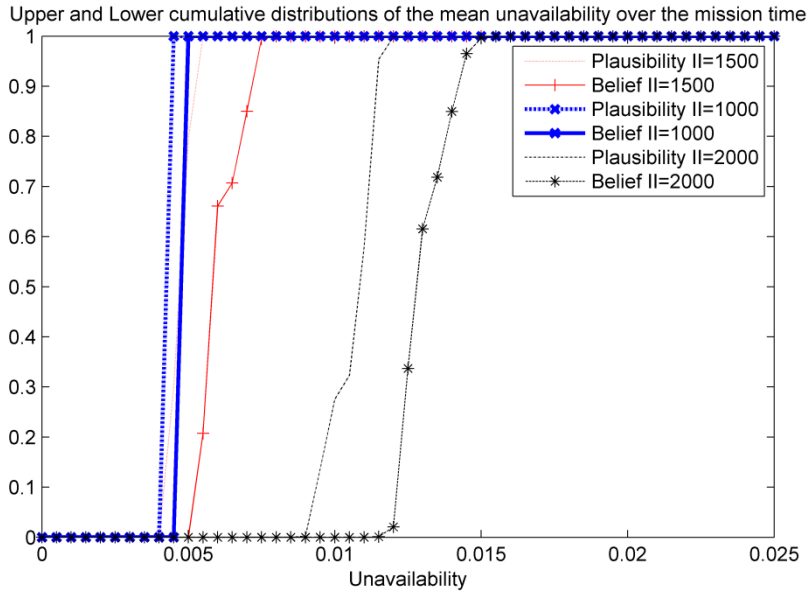

Figure 14

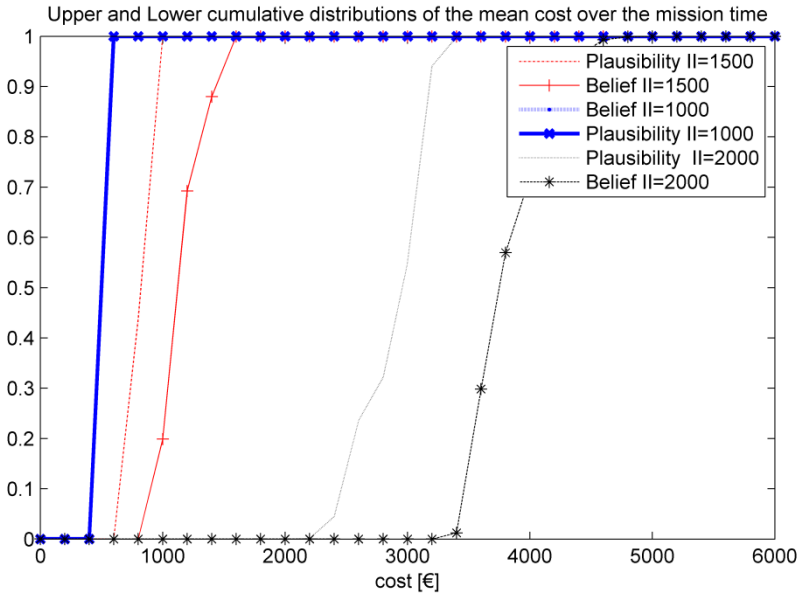

Figure 15

Tables

Parameters	Expert Knowledge					
	Expert 1		Expert 2		Expert 3	
	min	max	Min	max	min	max
a	$\underline{\iota_1^1}$ =1840	$\overline{\iota_1^1}$ =1880	$\underline{\iota_2^1}$ =1820	$\overline{\iota_2^1}$ =1870	$\underline{\iota_3^1}$ =1830	$\overline{\iota_3^1}$ =1890
b	$\underline{\iota_1^2}$ =7	$\overline{\iota_1^2}$ =12	$\underline{\iota_2^2}$ =2	$\overline{\iota_2^2}$ =4	$\underline{\iota_3^2}$ =5	$\overline{\iota_3^2}$ =9

Table 1: uncertainty ranges for the parameters provided by independent sources

sample	A	b
1	1889.6	11.97
2	1887.3	11.89

Table 2: samples drawn from S, which belong to the focal set E^1

Parameters	Nominal Values
η_{12}	1861h
β_{12}	8
η_{23}	743h
β_{23}	8
λ_1	$10^{-6}h^{-1}$
λ_2	$10^{-4}h^{-1}$
λ_3	$10^{-2}h^{-1}$

Table 3: Parameters of the probability distributions

Random Variables	Uncertain Parameters	Description
Y^1	$\boldsymbol{\theta}^1 = \{\theta^{1,1},\theta^{1,2}\}$	Transition time from degradation level 'Good' to 'Medium'
Y^2	$\boldsymbol{\theta}^2 = \{\theta^{2,1},\theta^{2,2}\}$	Transition time from degradation level 'Medium' to 'Bad'
Y^3	$\boldsymbol{\theta}^3 = \{\theta^{3,1}\}$	Transition time from degradation level 'Good' to 'Failed'
Y^4	$\boldsymbol{\theta}^4 = \{\theta^{4,1}\}$	Transition time from degradation level 'Medium' to 'Failed'
Y^5	$\boldsymbol{\theta}^5 = \{\theta^{5,1}\}$	Transition time from degradation level 'Bad' to 'Failed'

Table 4: tailoring of the general model to the considered case study

Parameters		Expert Knowledge					
		Expert 1		Expert 2		Expert 3	
		min	max	Min	max	min	max
$\theta^{1,1}$	η_{12}	1843	1880	1815	1908	1720	2001
$\theta^{1,2}$	β_{12}	7.92	8.08	7.8	8.2	7.4	8.6
$\theta^{2,1}$	η_{23}	735	750	725	762	687	800
$\theta^{2,2}$	β_{23}	7.92	8.08	7.8	8.2	7.4	8.6
$\theta^{3,1}$	λ_1	9.9e-7	1.01e-6	9.75e-7	1.03e-6	9.25e-7	1.075e-6
$\theta^{4,1}$	λ_2	0.99e-4	1.01e-4	9.75e-5	1.03e-4	9.25e-5	1.075e-4
$\theta^{5,1}$	λ_3	1e-2	1e-2	1e-2	1e-2	1e-2	1e-2

Table 5: uncertainty ranges for the parameters provided by three independent sources

Parameters	Values
Number of MC trials	2000
Number of combinations of uncertain parameters	8000
CPU time (Intel Core 2 duo, 3.17 GHz, 2GB RAM)	≈30h

Table 6: DSTE-based method parameters