Representing parametric probabilistic models tainted with imprecision

C. Baudrit[a], D. Dubois[b], N. Perrot[a]

[a] *UMR782 Génie et Microbiologie des Procédés Alimentaires. AgroParisTech, INRA, F-78850 Thiverval-Grignon.*
[b] *Institut de Recherche en Informatique de Toulouse, Université Paul Sabatier 31062 Toulouse, Cedex 4, France*

Abstract

Numerical possibility theory, belief function have been suggested as useful tools to represent imprecise, vague or incomplete information. They are particularly appropriate in uncertainty analysis where information is typically tainted with imprecision or incompleteness. Based on their experience or their knowledge about a random phenomenon, experts can sometimes provide a class of distributions without being able to precisely specify the parameters of a probability model. Frequentists use two-dimensional Monte-Carlo simulation to account for imprecision associated with the parameters of probability models. They hence hope to discover how variability and imprecision interact. This paper presents the limitations and disadvantages of this approach and propose a fuzzy random variable approach to treat this kind of knowledge.

Key words: Imprecise Probabilities, Possibility, Belief functions, Probability-Boxes, Monte-Carlo 2D, fuzzy random variable.

1 Introduction

The processing of uncertainties has become crucial in industrial applications and consequently in decision making processes. Uncertainties are often captured within a purely probabilistic framework. It means that uncertainty pertaining to the parameters of mathematical models representing physical or biological processes can be described by a single probability distribution. However, this method requires substantial knowledge to determine the probability law associated with each parameter.

Email addresses: `cbaudrit@grignon.inra.fr` (C. Baudrit[a]), `dubois@irit.fr` (D. Dubois[b]), `nathalie.perrot@grignon.inra.fr` (N. Perrot[a]).

Due to time and financial constraints, information regarding model parameters is often vague, imprecise or incomplete. It is more and more acknowledged that uncertainty regarding model parameters has essentially two origins [23]. It may arise from randomness (often referred to as "stochastic uncertainty") due to natural variability of observations resulting from heterogeneity or the fluctuations of a quantity in time. Or it may be caused by imprecision (often referred to as "epistemic uncertainty") due to a lack of information. This lack of knowledge may stem from a partial lack of data, either because collecting this data is too difficult or costly, or because only experts can provide some imprecise information. For example, it can be quite common for an expert to estimate numerical values of parameters in the form of confidence intervals according to his/her experience and intuition. The uncertainty pervading model parameters is thus not of a single nature: randomness and incomplete knowledge may coexist, especially due to the presence of several, heterogeneous sources of knowledge, as for instance statistical data jointly with expert opinions. One of the main approaches capable of coping with incompleteness as a feature distinct from randomness is the imprecise probabilities calculus developed at length by Peter Walley [40]. In this theory, sets of probability distributions capture the notion of partial lack of probabilistic information. In practice, while information regarding variability is best conveyed using probability distributions, information regarding imprecision is more faithfully conveyed using families of probability distributions. At the practical level, such families are most easily encoded either by probability boxes [22] or by possibility distributions (also called fuzzy intervals) [16] or yet by belief functions introduced by Dempster [12] (and elaborated further by Shafer [36] and Smets [38] in a different context).

Faced with information such as “I am sure the value of x lies in an interval $[a,b]$ but the value $\{c\}$ seems to be the most likely”, it is common to use triangular possibility distribution with core $\{c\}$ and support $[a,b]$. The choice of a linear interpolation between the core $\{c\}$ and the support $[a,b]$ looks debatable. However, it has been shown in [3,17] that the family of probabilities encoded by a triangular possibility distribution contains all probability distributions of mode $\{c\}$ and support $[a,b]$. It has also been shown in [3] that the probability family encoded by triangular possibility distribution can contain bimodal probability distributions. We can question the physical interpretation of such bimodal distributions if we know that the probability model associated with x is unimodal. Nevertheless, non-linear shapes could be used instead, for example if the expert had knowledge suggesting that while values located outside the core are possible, they are nevertheless very unlikely (in which case convex functions on each side would be used). While it can be argued that the choice of the shapes of possibility distributions is subjective, this subjectivity has far less consequences on the results than the fact of arbitrarily selecting single probability distributions in the presence of such partial information. Actually, uncertainty might be reduced by imagining that expert can sometimes provide the class of parametric distributions (e.g normal, lognormal...) without being able to specify the parameters of probability model in an exact way (e.g. mean, standard deviation, median ...) [32]. Frequentists use two-dimensional Monte-Carlo [9,26] simulation to account for uncertainty associated with the parameters of a probability model. This approach assumes that single precise probability distributions are used to represent uncertainty related to the parameters of the probability model. Because of its mathematical simplicity, 2MC simulation is routinely used and recommended as a convenient and natural approach. A two dimensional Monte-Carlo simulation is a nesting of two ordinary Monte-Carlo simulations [25]. By nesting one Monte-Carlo simulation within another, experts hope to discover how variability and imprecision interact and produce uncertain outputs. Given the imprecise nature of information regarding parameters, it sounds more faithfully to use intervals or confidence intervals for representing parameter knowledge, than full-fledged probability distributions (especially considering the assumed non-stochastic nature of parameter distributions). In order to represent a class of probability distributions tainted with imprecision, it seems natural to combine Monte-Carlo technique [25] with the extension principle of fuzzy set theory [15]. This process generates a fuzzy random variable [34]. This kind of approach has already received some attention in the literature for computing output of functions with probabilistic and fuzzy arguments [2,24,32], or handling fuzzy parametric models [29,32].

Section 2 is dedicated to the basic concepts of probability-boxes, possibility theory, belief functions and fuzzy random variables in connection with imprecise probabilities. The main disadvantages and limitations of the two-dimensional Monte-Carlo simulation are discussed in Section 4. Next, we present how the framework of fuzzy random variables allows to represent the class of probability distributions with imprecise parameters faithfully. Our approach is different from those proposed by Kentel et al. [29] and Moller et al. [32] in the sense where we process fuzzy random variables in the belief function framework. Lastly, in Section 5 we compare the fuzzy random variable approach with the two-dimensional Monte-Carlo simulation on an academic example.

2 Representing imprecise probabilities

Let $(\Omega,\mathcal{A})$ be a measurable space where $\mathcal{A}$ is an algebra of measurable subsets of Ω. Let $\mathcal{P}$ be a set of probability measures on $(\Omega,\mathcal{A})$. Such a family may be natural to consider if a parametric probabilistic model is used but the parameters such as the mean value or the variance are ill-known (for instance they lie in an interval). It induces upper and lower probability functions respectively defined by:

$$\overline{P}(A) = \sup_{P\in\mathcal{P}} P(A) \quad \text{and} \quad \underline{P}(A) = \inf_{P\in\mathcal{P}} P(A) \quad \forall A \subseteq \Omega.$$

The upper probability of A is equal to one minus the lower probability of the complement of A. So, the lower probability is a measure of how much family $\mathcal{P}$ supports event A and upper probability of A reflects the lack of information against A. In a subjectivist tradition, the lower probability for an event A can be interpreted, in accordance with the so-called betting method, as the maximum price that one would be willing to pay for the gamble that wins 1 unit of utility if A occurs and nothing otherwise. The probability family

$$\mathcal{P}(\underline{P} < \overline{P}) = \{P, \forall A \subseteq \Omega, \underline{P}(A) \leq P(A) \leq \overline{P}(A)\}$$

induced from upper and lower probability induced from $\mathcal{P}$, is generally a proper superset of $\mathcal{P}$. It is clear that representing and reasoning with a family of probabilities may be very complex. In the following we consider four frameworks for representing special sets of probability functions, which are more convenient for a practical handling.

2.1 Probability boxes

Let $X : \Omega \to \mathbb{R}$ be a random variable and $F_X : \mathbb{R} \to [0, 1]$ be its associated cumulative distribution function defined by $F_X(x) = P(X \in (-\infty, x]), \forall x \in \mathbb{R}$. Suppose $\overline{F}_X$ and $\underline{F}_X$ are nondecreasing functions from the real line $\mathbb{R}$ into $[0, 1]$ such that $\underline{F}_X(x) \leq F_X(x) \leq \overline{F}_X(x), \forall x \in \mathbb{R}$. The interval $[\underline{F}_X, \overline{F}_X]$ is called a "probability box" or "p-box" [22]. It encodes the class of probability measures whose cumulative distribution functions F_X are restricted by the bounding pair of cumulative distribution functions $\underline{F}_X$ and $\overline{F}_X$.
A p-box can be induced from the probability family $\mathcal{P}$ by:

$$\underline{F}_X(x) = \underline{P}((-\infty, x]) \quad \text{and} \quad \overline{F}_X(x) = \overline{P}((-\infty, x]), \quad \forall x \in \mathbb{R}.$$

Let $\mathcal{P}(\underline{F}_X \leq \overline{F}_X)$ be the probability family containing $\mathcal{P}$ and defined by

$$\mathcal{P}(\underline{F}_X \leq \overline{F}_X) = \{P, \forall x \in \mathbb{R}, \underline{F}_X(x) \leq F(x) \leq \overline{F}_X(x)\}.$$

Generally $\mathcal{P}(\underline{F}_X \leq \overline{F}_X)$ strictly contains $\mathcal{P}(\underline{P} < \overline{P})$, hence also the set $\mathcal{P}$ it is built from. The probability box $[\underline{F}_X, \overline{F}_X]$ provides a bracketing of some ill-known cumulative distribution function and the gap between $\underline{F}_X$ and $\overline{F}_X$ reflects the incomplete nature of the knowledge, thus picturing the extent of what is ignored.

2.2 Numerical possibility theory

Possibility theory [16] is relevant to represent consonant imprecise knowledge. A possibility distribution on a state space S can model imprecise information regarding a fixed unknown parameter and it can also serve as an approximate representation of incomplete observation of a random variable. The basic notion is the possibility distribution, denoted π, an upper semi-continuous mapping from the real line to the unit interval. A possibility distribution describes the more or less plausible values of some uncertain variable X. Possibility theory provides two functions (the possibility Π and the necessity N) allowing to evaluate the confidence that we can have in the assertion: the value of a real variable X does lie within a certain interval. The normalized measure of possibility Π (respectively necessity N) is defined from the possibility distribution $\pi : S \to [0,1]$ such that $\sup_{x\in S} \pi(x) = 1$ as follows:

$$\Pi(A) = \sup_{x\in A} \pi(x), \quad N(A) = 1 - \Pi(\overline{A}) = \inf_{x\notin A} (1 - \pi(x)).$$

The following basic properties hold:

$$\Pi(A \cup B) = \max(\Pi(A), \Pi(B)), \quad N(A \cap B) = \min(N(A), N(B)) \quad \forall A, B \subseteq \mathbb{R}.$$

A possibility distribution π is the membership function μ_F of a normalized fuzzy set F. Faced with information expressing that an unknown quantity is restricted by a fuzzy set F, the identity $\pi(s) = \mu_F(s)$ means that if t is the membership degree of s in F, t is interpreted as the possibility degree that the value of this unknown quantity is s. In the following the α-cut of a fuzzy subset F of a set S is the subset $F_\alpha = \{s, \mu_F(s) \geq \alpha\}$. A numerical possibility distribution may also be viewed as a nested family of subsets, which are the α-cuts π. The degree of certainty that F_α contains X is $N(F_\alpha)$ ($= 1 - \alpha$ if $S = \mathbb{R}$ and π is continuous). Conversely, suppose a nested family of subsets $(A_i)_{i=1,\dots,n}$ (such that $A_1 \subset \ldots \subset A_n$) with degrees of certainty λ_i that A_i contains X is available. Provided that λ_i is interpreted as a lower bound on $N(A_i)$ and π is chosen as the least specific possibility distribution satisfying these inequalities [18], this is equivalent to knowing the possibility distribution

$$\pi(x) = \min_{i=1\dots n}\{1 - \lambda_i, x \notin A_i\}$$

with convention $\pi(x) = 1$ in case $x \in A_i$ for all i.We can interpret any pair of dual functions necessity/possibility $[N, \Pi]$ as upper and lower probabilities induced from specific probability families.

- Let π be a possibility distribution inducing a pair of functions $[N, \Pi]$. We define the probability family $\mathcal{P}(\pi) = \{P, \forall A$ measurable, $N(A) \leq P(A)\} = \{P, \forall A$ measurable, $P(A) \leq \Pi(A)\}$. In this case, $\sup_{P\in\mathcal{P}(\pi)} P(A) = \Pi(A)$ and $\inf_{P\in\mathcal{P}(\pi)} P(A) = N(A)$ (see [11,18]) hold. In other words, the family $\mathcal{P}(\pi)$ is entirely determined by the probability intervals it generates.
- Suppose pairs (interval A_i, necessity weight λ_i) supplied by an expert are interpreted as stating that the probability $P(A_i)$ is at least equal to λ_i where A_i is a measurable set. We define the probability family as follows: $\mathcal{P}(\pi) = \{P, \forall A_i, \lambda_i \leq P(A_i)\}$. We thus know that $\overline{P} = \Pi$ and $\underline{P} = N$ (see [18], and in the infinite case [11]).

We can define a particular p-box $[\underline{F}_X, \overline{F}_X]$ from the possibility distribution π such that $\underline{F}_X(x) = N((-\infty, x])$ and $\overline{F}_X(x) = \Pi((-\infty, x])$, $\forall x \in \mathbb{R}$. But this p-box contains many more probability functions than $\mathcal{P}(\pi)$ (see [3] for more details about the compared expressivity of a p-box and a possibility distribution).

2.3 Belief functions induced from random sets

A random set on a finite set S is defined by a mass assignment ν which is a probability distribution on the power set of S. We assume that ν assigns a positive mass only to a finite family of subsets of S called the set $\mathcal{F}$ of focal subsets. Generally $\nu(\emptyset) = 0$ and $\sum_{E \in \mathcal{F}} \nu(E) = 1$. In the context of this paper, we consider *disjunctive* random sets, whose focal elements E contain mutually exclusive elements. A focal element represents imprecise information about some quantity X such that all that is known about X is that it lies in E. The weight $\nu(E)$ is then the probability that the state of information is of the form $X \in E$ (and not more precise). A random set induces set functions called plausibility and belief measures respectively denoted by Pl and Bel, and defined by Shafer [36] as follows:

$$Bel(A) = \sum_{E, E \subseteq A} \nu(E) \quad \text{and} \quad Pl(A) = \sum_{E, E \cap A \neq \emptyset} \nu(E) = 1 - Bel(A^c).$$

$Bel(A)$ gathers the imprecise evidence that asserts A ; $Pl(A)$ gathers the imprecise evidence that does not contradict A.

This approach initiated by Shafer [36] and further elaborated by Smets [38] in a different context allows imprecision and variability to be treated separately within a single framework. Indeed, it provides mathematical tools to process information that is at the same time of random and imprecise nature. These set-functions can be interpreted as families of probability measures, even if this view does not match the original motivation of Shafer [36] and Smets [38] for belief functions. A mass distribution ν may encode the probability family $\mathcal{P}(\nu) = \{P \in \mathcal{P} / \forall A \subseteq \Omega, Bel(A) \leq P(A)\} = \{P \in \mathcal{P} / \forall A \subseteq \Omega, P(A) \leq Pl(A)\}$. This family generates lower and upper probability functions that coincide with the belief and plausibility functions, i.e.

$$Pl(A) = \sup_{P \in \mathcal{P}(\nu)} P(A), \quad Bel(A) = \inf_{P \in \mathcal{P}(\nu)} P(A)$$

Originally, such imprecise probabilities were introduced by Dempster [12], who considered a probability space and a set-valued mapping Γ from a probability space $(\Omega, \mathcal{A}, P)$ to S yielding a random set. For simplicity assume $\forall \omega \in \Omega, \Gamma(\omega) \neq \emptyset$. Let $X : \Omega \to S$ be a measurable selection from Γ such that $\forall \omega \in \Omega$, $X(\omega) \in \Gamma(\omega)$ and P_X be its associated probability measure such that $P_X(A) = P(X^{-1}(A))$. Define upper and lower probabilities as follows:

$$\overline{P}(A) = \sup_{X \in s(\Gamma)} P_X(A) \quad \underline{P}(A) = \inf_{X \in s(\Gamma)} P_X(A)$$

where $s(\Gamma)$ is the set of measurable selections of Γ. For all measurable subsets $A \subseteq \Omega$, we have $\underline{A} \subseteq A \subseteq \overline{A}$ where $\underline{A} = \{\omega \in \Omega, \Gamma(\omega) \subseteq A\}$ and $\overline{A} = \{\omega \in \Omega, \Gamma(\omega) \cap A \neq \emptyset\}$. By defining the mass distribution ν_Γ on Ω by $\nu_\Gamma(E) = P(\{\omega, \Gamma(\omega) = E\})$. We thus retrieve belief and plausibility functions as follows:

$$\underline{P}(A) = P(\underline{A}) = Pl_\Gamma(A) = \sum_{E \cap A \neq \emptyset} \nu_\Gamma(E)$$

$$\overline{P}(A) = P(\overline{A}) = Pl_\Gamma(A) = \sum_{E \subseteq A} \nu_\Gamma(E)$$

In the continuous case, when $S = \mathbb{R}$, continuous belief functions can be defined letting $(\Omega, \mathcal{A}, P)$ be the unit interval equipped with the Lebesgue measure, and $\Gamma(\omega)$ be a Borel-measurable subset of reals (e.g. an interval) (see Smets[37]). Then again, $Pl(A) = P(\Gamma(\omega) \cap A \neq \emptyset)$ and $Bel(A) = P(\Gamma(\omega) \subseteq A)$.We may extract upper $\overline{F}_X$ and lower $\underline{F}_X$ cumulative distribution functions such that, $\forall x \in \mathbb{R} \quad \underline{F}_X(x) \leq F(x) \leq \overline{F}_X(x)$ with :

$$\overline{F}_X(x) = Pl(X \in (-\infty, x]) \quad \text{and} \quad \underline{F}_X(x) = Bel(X \in (-\infty, x]).$$

This is a particular p-box. But this p-box contains many more probability functions than $\mathcal{P}(\nu)$. Interestingly, a p-box is a special case of continuous belief function (see for instance [14]) with focal sets in the form of intervals $([\overline{F}_X^{-1}(\alpha), \underline{F}_X^{-1}(\alpha)]), \alpha \in (0,1]$. $\nu([\overline{F}_X^{-1}(p_i), \underline{F}_X^{-1}(p_i)]) = p_i - p_{i-1}$

3 Fuzzy random variables

A fuzzy random variable associates a fuzzy set to each possible result of a random experiment. In the literature, fuzzy random variables can be interpreted in different ways depending on the context of the study. Originally, a fuzzy random variable $\tilde{T}$ assigns a (precise) probability to each possible (fuzzy) image of $\tilde{T}$ (for instance the membership function of a linguistic term in a term scale), by considering it as a “classical” measurable mapping [34]. There are variants according to the type of metrics chosen on the set of measurable membership functions [30]. Below, we are going to briefly describe two other interpretations, The first one views a fuzzy random variable as a possibility distribution over classical random variables (named the 2d order model [6,8]). In the other one, a fuzzy random variable corresponds to a family of probability measures constructed from a probability space and a fuzzy relation between this space and another space, and the fuzzy relation is interpreted as a family of conditional probabilities. The fuzzy random variable is then equated to a set of probability functions [1]. In both views, the membership function μ_F of a fuzzy set F is interpreted as a possibility distribution π associated to some unknown quantity.

Let $(\Omega, \mathcal{A}, P)$ be a probability space. Let $\tilde{\mathcal{F}}(\mathbb{R})$ be the set of measurable fuzzy subsets of $\mathbb{R}$. Here, we briefly recall how a second-order possibility measure on the set of classical random variables, induced by a fuzzy random variable, is constructed [6,8]. For instance, consider the random variable $T = f(X, Y)$, where $f : \mathbb{R}^2 \to \mathbb{R}$ is a known mapping, X a random variable $\Omega \to \mathbb{R}$ and Y is another imprecisely known random variable described by a fuzzy set associated to the membership function $\mu_{\tilde{Y}} : \mathbb{R} \to [0, 1]$. Hence, it defines a constant mapping, $\tilde{Y} : \Omega \to \tilde{\mathcal{F}}(\mathbb{R})$ that assigns, to every element of Ω, the same fuzzy set $\tilde{Y}$. That means for each $\omega \in \Omega$ and each $y \in \mathbb{R}$, $\mu_{\tilde{Y}}(y)$ represents the possibility grade that $Y(\omega)$ coincides with y. Then, $\tilde{T} = f(X, \tilde{Y})$ is a fuzzy random variable, defined by the extension principle, by

$$\mu_{\tilde{T}(\omega)}(t) = \sup_{t=f(X(\omega),y)} \mu_{\tilde{Y}}(y). \tag{1}$$

Let $[\tilde{T}(\omega)]_\alpha = f(X(\omega), Y_\alpha) = \{f(X(\omega), y) : y \in \tilde{Y}_\alpha\}$. The fuzzy random variable $\tilde{T} : \Omega \to \tilde{\mathcal{F}}(\mathbb{R})$ represents the following imprecise information about the random variable $T : \Omega \to \mathbb{R}$: for each $\alpha > 0$, the probability $P(T(\omega) \in [\tilde{T}(\omega)]_\alpha)$ is greater than or equal to $1-\alpha$. This is in agreement with the fact that the possibility distribution associated with $\tilde{T}$ is equivalent to stating that for each cut $[\tilde{T}(\omega)]_\alpha$, the degree of necessity has a lower bound: $N([\tilde{T}(\omega)]_\alpha) \geq 1 - \alpha$. Under this interpretation we can say that, for each confidence level $1 - \alpha$, the probability distribution associated with T belongs to the set $\mathcal{P}_{\tilde{T}_\alpha} = \{P_T, T \in s(\tilde{T}_\alpha)\}$, where $s(\tilde{T}_\alpha)$ is the set of selections from the random set $\tilde{T}_\alpha$, *i.e* $s(\tilde{T}_\alpha) = \{T : \Omega \to \mathbb{R}, T(\omega) \in [\tilde{T}(\omega)]_\alpha\}$. Thus, given an arbitrary event A of the final space, the probability $P_T(A)$ belongs to the set

$$\mathcal{P}_{\tilde{T}_\alpha}(A) = \{P_T(A), T \in s(\tilde{T}_\alpha)\} \tag{2}$$

with confidence level $1 - \alpha$. In [6] the fuzzy set $\tilde{P}_{\tilde{T}}$ of probability functions, with membership function given by the equation:

$$\mu_{\tilde{P}_{\tilde{T}}}(Q) = \sup\{\alpha \in [0, 1],\ Q \in \mathcal{P}_{\tilde{T}_\alpha}\},\ \forall\, Q$$

is viewed as an imprecise representation of the probability measure P_T. In fact, $\mu_{\tilde{P}_{\tilde{T}}}$ is a possibility distribution on the space of probability functions. According to the available information, the quantity $\mu_{\tilde{P}_{\tilde{T}}}(Q)$ represents the possibility degree that Q coincides with the true probability measure associated with T, P_T. On the other hand, for each event A, the fuzzy subset of the unit interval $\tilde{P}_{\tilde{T}}(A)$, defined as

$$\mu_{\tilde{P}_{\tilde{T}}(A)}(p) = \sup\{\alpha \in [0, 1]/p \in \mathcal{P}_{\tilde{T}_\alpha}(A)\},\ \forall\, p \in [0, 1],$$

represents our imprecise information about the quantity $P_T(A) = P(T \in A)$. Thus, the value $\mu_{\tilde{P}_{\tilde{T}}(A)}(p)$ represents the degree of possibility that the “true” degree of probability $P_T(A)$ is p. De Cooman recently proposed a behavioral interpretation of such fuzzy probabilities [10].

The possibility measure $\tilde{P}_{\tilde{T}}$ is a "second order possibility measure". We use this term because it is a possibility distribution defined over a set of probability measures [40]. A second order possibility measure is associated with a set of (meta) probability measures, each of them defined, as well, over a set of probability measures. Thus, a second order possibility measure allows us to state sentences like "the probability that the true probability of the event A is 0.5 ranges between 0.4 and 0.7". On the other hand, it is easily checked that the set of probability functions considered in equation (2) yields a plausibility function :

$$Pl_{\alpha}(A) = P(\{\omega \in \Omega, [\tilde{T}(\omega)]_{\alpha} \cap A \neq \emptyset\})$$

and lower bounded by a belief function:

$$Bel_{\alpha}(A) = P(\{\omega \in \Omega, [\tilde{T}(\omega)]_{\alpha} \subseteq A\}).$$

The interval $[Bel_{\alpha}(A), Pl_{\alpha}(A)]$ is the α-cut of the fuzzy probability $\tilde{P}_{\tilde{T}}(A)$.

3.2 The order 1 model of fuzzy random variables

The second order model (of the last subsection) associates, to each event (crisp subset of the final space), a fuzzy set in the unit interval. Here we assume a probability space $(\Omega, \mathcal{A}, P)$ and an ill-known conditional probability function $P(t|\omega)$, relating spaces Ω and the range $\mathbb{R}$ of a variable T. It is supposed that the knowledge about $P(t|\omega)$ consists of a conditional possibility distribution $\pi(t|\omega)$, such that, when ω is fixed $\Pi(A|\omega) \geq P(A|\omega)$. It induces the probability family $\mathcal{P}_T$ on the output space, defined by ([1,5]):

$$\mathcal{P}_T = \{P_T, P_T(A) = \int_{\Omega} P(A|\omega)dP(\omega), \Pi(A|\omega) \geq P(A|\omega)\}$$

For instance, consider again the random variable $T = f(X, Y)$, where $f : \mathbb{R}^2 \to \mathbb{R}$ is a known mapping, X a random variable $\Omega \to \mathbb{R}$. Now, $Y \in \mathbb{R}$ is another, imprecisely known, quantity that deterministically affects the relation between X and T via the function f. The information about Y is given by means of a possibility distribution $\mu_{\tilde{Y}}$. $\tilde{T}(\omega) = f(X(\omega), \tilde{Y})$ defined by equation (1) is now interpreted as a conditional possibility distribution $\pi(t|\omega)$. According to Section 2.2, for each $\omega \in \Omega$ the set of probability measures $\{P(\cdot|\omega), P([\tilde{T}(\omega)]_{\alpha}|\omega) \geq 1 - \alpha, \forall\alpha > 0\}$ coincides with the set of probability measures dominated by the possibility measure $\Pi(\cdot|\omega)$. This view comes down to considering the fuzzy random variable assigning to each realization ω the fuzzy set $\pi(\cdot|\omega)$, as a standard random set that assigns to each pair (ω, α) the set $[\tilde{T}(\omega)]_{\alpha}$ with a mass density $d\alpha \times dP(\omega)$ (it is a continuous belief function in the spirit of Smets [37]). The plausibility measure of a measurable set (e.g. an interval)

A, describing our information about T is thus of the form :

$$Pl_T(A) = \sup\{P_T(A), P_T \in \mathcal{P}_T\} = \int_\Omega \int_{A\cap[\tilde{T}(\omega)]_\alpha \neq \emptyset} d\alpha dP(\omega) = \int_\Omega \Pi(A|\omega)dP(\omega). \quad (3)$$

Similarly for the lower bound:

$$Bel_T(A) = \inf\{P_T(A), P_T \in \mathcal{P}_T\} = \int_\Omega \int_{[\tilde{T}(\omega)]_\alpha \subseteq A} d\alpha dP(\omega) = \int_\Omega N(A|\omega)dP(\omega). \quad (4)$$

We can interpret the values $Pl_T(A)$ and $Bel_T(A)$ as the most precise bounds (the least upper one for $Pl_T(A)$, the greatest lower one for $Bel_T(A)$) for the "true" probability of A, according to the available information [6]. There exists a strong relationship between these plausibility and belief functions and the fuzzy set $\tilde{P}_{\tilde{T}}(A)$ with cuts $[Bel_\alpha(A), Pl_\alpha(A)]$ defined in Section 3.1. Equating the fuzzy sets $\tilde{T}(\omega)$ with possibility distributions $\pi(\cdot|\omega)$, the following result holds:

$$[Bel_T(A), Pl_T(A)] = [\int_0^1 Bel_\alpha(A)d\alpha, \int_0^1 Pl_\alpha(A)d\alpha],$$

In other words, the interval $[Bel_T(A), Pl_T(A)]$ coincides with the "mean value" [19] and the average level [35] of the fuzzy set $\tilde{P}_{\tilde{T}}(A)$ [1].

3.3 Discretized encoding of probability, possibility, p-boxes and fuzzy random variables as random sets

Belief functions [12,36] encompass possibility, probability, probability-boxes theories and the previous subsection shows it may as well account of a special view of fuzzy random variables. Hence, we can encode probability distributions p, p-boxes $[\underline{F}_X, \overline{F}_X]$, possibility distributions π and fuzzy random variables $\tilde{X}$. Continuous representations on the real line will be approximated in a discrete framework, by using mass distribution ν, for making practical computations.

- Let X be a real random variable. In the discrete case, focal elements are singletons $(\{x_i\})_i$ and the mass distribution ν is defined by $\nu(\{x_i\}) = P(X = x_i)$. In the continuous case, we define focal intervals $((x_i, x_{i+1}])_i$ by discretizing probability density into m intervals and a mass distribution ν is defined by $\nu((x_i, x_{i+1}]) = P(X \in (x_i, x_{i+1}]), \forall\, i = 1 \ldots m$.
- Let X be an ill-known random variable described by a possibility distribution π. Focal sets correspond to the α-cuts

$$E_j = \{x|\pi(x) \geq \alpha_j\},\ \forall\ j = 1...q$$

of possibility distribution π associated with X such that $\alpha_1 = 1 \geq \alpha_j \geq \alpha_{j+1} \geq \alpha_q > 0$ and $E_j \subseteq E_{j+1}$. Mass distribution ν is defined by $\nu(E_j) = \alpha_j - \alpha_{j+1}$, $\forall\ j = 1 \ldots q$ where $\alpha_{q+1} = 0$.

- Let X be an ill-defined random variable represented by a p-box $[\underline{F}_X, \overline{F}_X]$. By putting

$$\underline{F}_X^{-1}(p) = \min\{x | \underline{F}_X(x) \geq p\}, \ \forall\ p \in [0,1] \quad (5)$$

$$\overline{F}_X^{-1}(p) = \min\{x | \overline{F}_X(x) \geq p\}, \ \forall\ p \in [0,1] \quad (6)$$

we can choose focal sets of the form $([\overline{F}_X^{-1}(p_i), \underline{F}_X^{-1}(p_i)])_i$ and the mass distribution ν such that $\nu([\overline{F}_X^{-1}(p_i), \underline{F}_X^{-1}(p_i)]) = p_i - p_{i-1}$ where $1 \geq p_i > p_{i-1} > 0$. Kriegler et al. [31] have shown that this p-box is representable by a belief function, so that $\mathcal{P}(\underline{F}_X \leq \overline{F}_X) = \mathcal{P}(\nu)$. This result is generalised to cumulative distributions on any finite ordered set by Destercke et al. [13].
- Let $\tilde{X}$ be a fuzzy random variable described by a finite set of possibility distributions $(\pi^1, \dots, \pi^n)$ with respective probability $(p_1, \dots, p_n)$. Using the order 1 view developed in the previous sections, focal sets of possibility distributions $(\pi^i)_{i=1,\dots,n}$ correspond to the α-cuts

$$E_j^i = \{x | \pi^i(x) \geq \alpha_j\}, \ j = 1, \dots, q, \ i = 1, \dots, n$$

with $\alpha_1 = 1 \geq \alpha_j \geq \alpha_{j+1} \geq \alpha_q > 0$. Mass distribution ν is defined by $\nu(\pi^i_{\alpha_j}) = (\alpha_j - \alpha_{j+1}) \times p_i$, for all $j = 1, \dots, q$ and $i = 1, \dots, n$ where $\alpha_{q+1} = 0$. Besides, we can observe that the induced plausibility (resp. belief) of a measurable set A coincides with the arithmetic mean of the possibility (resp. necessity) measures Π^i (weighted by the probabilities of the different values of X,) i.e. [2]:

$$Pl(A) = \sum_{(i,j):A \cap E_j^i \neq \emptyset} \nu_{ij} = \sum_{i=1}^{m} \sum_{j:A \cap E_j^i \neq \emptyset} p_i\, \nu_j = \sum_{i=1}^{m} p_i \Pi^i(A), \quad (7)$$

$$Bel(A) = \sum_{(i,j):E_j^i \subseteq A} \nu_{ij} = \sum_{i=1}^{m} \sum_{j:E_j^i \subseteq A} p_i\, \nu_j = \sum_{i=1}^{m} p_i N^i(A). \quad (8)$$

4 Representation of parametric probabilistic models tainted with imprecision

In uncertainty analysis, it is usual to represent knowledge pertaining to uncertain quantities by parametric probabilistic models P_θ. But one is not always able to specify the values of parameters $\theta \in \Theta$ precisely. Indeed, based on their experience or their knowledge about random phenomenon, experts can provide a class of distributions having, for instance, the same shape but differing in central tendency. They can also provide a class of distributions being from the same distribution family (e.g. normal distribution $\mathcal{N}(\mu, \sigma)$) with the same mean (e.g. μ) but different variances (e.g. $\sigma \in [\underline{\sigma}, \overline{\sigma}]$). That means that experts may be able to provide an interval regarding possible values for a given parameter (e.g. $\sigma \in [\underline{\sigma}, \overline{\sigma}]$), but also to express preferences within this interval by defining confidence intervals (e.g. $[\underline{\sigma}, \overline{\sigma}]_\alpha$

is a confidence interval of σ with a level 1-α). Frequentists use two-dimensional Monte-Carlo (2MC) simulation to account for uncertainty (imprecision) associated with the parameters of probability model. Because of its mathematical simplicity, this approach is now widely used and recommended [27,39]. It seems clear that the two-dimensional Monte-Carlo method faces the same difficulties, in particular regarding the choice of the meta-probability function that represents knowledge about the parameter of probability model.

4.1 The classical approach: the two-dimensional Monte-Carlo method

A two-dimensional Monte-Carlo simulation [9] is a nesting of two ordinary Monte-Carlo simulations [25]. This subsection presents its basic steps and discusses the underlying assumptions.

4.1.1 Presentation

Let $\vec{X} : \Omega \rightarrow \mathbb{R}^n$ be a random vector and consider the random variable $T = f(\vec{X})$, where $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is a known mapping. Assume that random variables $(X_1, \ldots, X_n)$ are represented by parametric probabilistic models $(P^{X_i}_{\vec{\theta}_i})_{i=1,\ldots n}$. Moreover, the vector parameters $(\vec{\theta}_1, \ldots, \vec{\theta}_n)$ of the probability models $P^{X_1}_{\vec{\theta}_1}, \ldots, P^{X_n}_{\vec{\theta}_n}$ are themselves represented by single probability distributions $P_1, \ldots, P_n$. For instance, $\vec{X} = (X_1, X_2) \hookrightarrow \mathcal{N}(\mu_1, \sigma_1) \cdot \mathcal{N}(\mu_2, \sigma_2)$. Then $\vec{\theta}_1 = (\mu_1, \sigma_1)$ has distribution $P_1 = \mathcal{U}([a_1, a_2]) \cdot \mathcal{U}([a_3, a_4])$, $\vec{\theta}_2 = (\mu_2, \sigma_2)$ has distribution $P_2 = \mathcal{U}([a_5, a_6]) \cdot \mathcal{U}([a_7, a_8])$, where all a_i are precise values, $\mathcal{U}$ means "uniform distribution". The 2MC method is summarized as follows [9]:

(1) Generate an n-realization $(\vec{\theta}_1, \ldots, \vec{\theta}_n)$ according to the probability distributions $P_1, \ldots, P_n$ and in accordance with dependencies (if known).
(2) Generate m realizations $(x^j_1(\vec{\theta}_1), \ldots, x^j_n(\vec{\theta}_n))_{j=1,\ldots,m}$ according to the probabilities $P^{X_1}_{\vec{\theta}_1}, \ldots, P^{X_n}_{\vec{\theta}_n}$, respecting dependencies (if known), based on the parameter selection $\vec{\theta}_1, \ldots, \vec{\theta}_n$.
(3) Compute m realizations $(t_j)_{j=1,\ldots,m} = (f(x^j_1(\vec{\theta}_1), \ldots, x^j_n(\vec{\theta}_n))_{j=1,\ldots,m}$ for the random variable T.
(4) Return to step 1 until a collection of n possible probability distribution functions (each corresponding to a choice of parameters) is obtained (see Figure 1).

Typically, the first step of the simulation represents the expert's uncertainty about the parameters that should be used to specify the probabilities about $\vec{X}$ for step 2. The second step of the simulation represents natural variability of the underlying physical and biological processes. The two-dimensional Monte-Carlo method

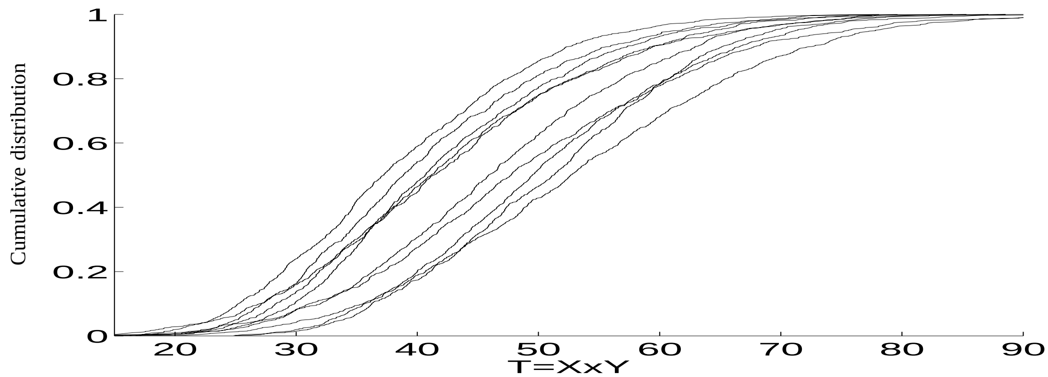

Fig. 1. Sample of 10 cumulative probability distributions resulting from 2MC simulation where $T = X \times Y$ with X=Triangular(3,M,8), $Y = \mathcal{N}(\mu, \sigma)$ and $(M, \mu, \sigma) = (\mathcal{U}([4, 6]), \mathcal{U}([7.5, 10]), \mathcal{U}([1, 2]))$ by assuming independence between X and Y, between M and (μ, σ) and strong independence between μ and σ.

hence provides a probability measure on the space of probability functions $P^T_{\vec{\theta}}$ called a "meta-distribution".

4.1.2 Comments on the two-dimensional Monte-Carlo method

There are some limitations to the use of the 2MC approach [20]:

(1) The two-dimensional Monte-Carlo simulation requires an expert to specify one probability distribution function (it is often just postulated) for each uncertain parameter of probability models and potential inter-parameter dependencies as well. Analysts already face real difficulties to characterize the probability distribution function pertaining to the underlying physical and/or biological process, and one may wonder to what extent they can justify relevant probability distributions regarding the parameters of probability models. Not being able to correctly specify probability distributions about parameters, certain analysts tried three- and even four-dimensional Monte-Carlo simulations. They need to provide higher-order probabilities modeling their state of knowledge about the parameters of a mathematical model, but the higher the order of the distribution, the less useful and meaningful such information is for decision-making.

(2) The two-dimensional Monte-Carlo simulation provides a sample of cumulative distribution functions (see Figure 1). It thus appears difficult to explain this kind of outputs to managers and decision makers. Faced with these results, analysts conceal the complexity of the meta-distribution by representing the median or the mean distribution and the $100\alpha^{th}$ and $100(1-\alpha)^{th}$ percentiles of the distributions considered as an envelope of the meta-distribution (see Figure 2 with α=0.05). This postprocessing may involve a great loss of

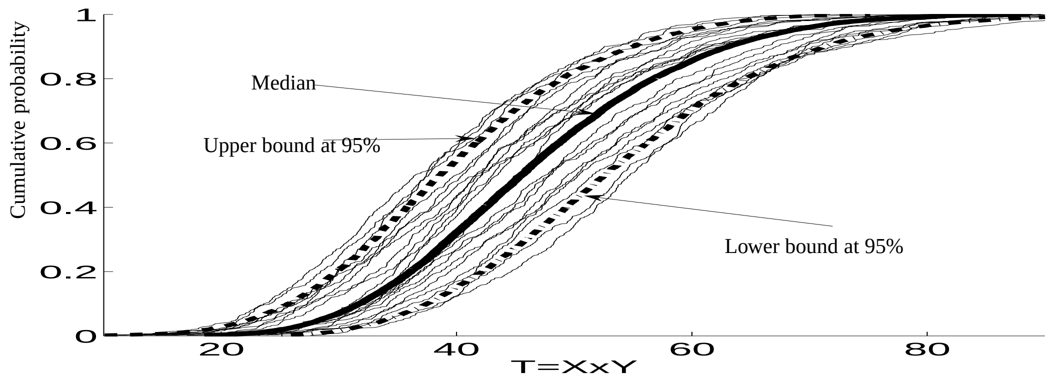

Fig. 2. Median, lower and upper percentiles at 95% resulting from the post-processing of Monte-Carlo 2D simulation.

information and a misinterpretation of the results. Indeed, they might believe that 100(1-2α)% of possible distributions represented by the meta-distribution are within the envelope defined by the $100\alpha^{th}$ and $100(1-\alpha)^{th}$ percentiles of the distributions. This interpretation would be totally incorrect. For instance, consider five possible cumulative distributions resulting from the 2MC simulation, from which 20% on the left and on the right side are eliminated, and perform the pointwise union of the remaining cumulative distributions (see the 20^{th} and 80^{th} percentiles of distributions in dotted line on Figure 3). According to Figure 3, we can observe that neither F_T^1 nor F_T^2 (two of five possible cumulative distributions) lie within the envelope defined by the 20^{th} and 80^{th} percentiles. It is thus not true that 60% of the possible distributions lie inside bound limits. The same problem occurs, for instance, with the estimated aver-

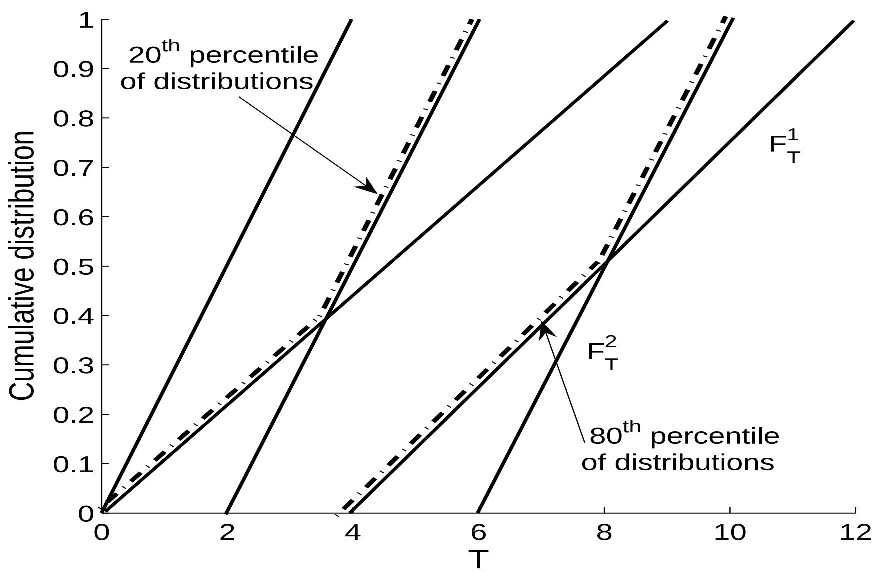

Fig. 3. Envelope defined by the 20^{th} and 80^{th} percentiles of distributions in the meta-distribution.

age distribution because an ad hoc distribution might be obtained which might differ from any of the distributions of the meta-distribution.

(3) There are also difficulties about parameter dependencies in the probability model. Due to a lack of knowledge, analysts often assume independence between such parameters which can create impossible mathematical structures. For example, consider an ill-known random variable X represented by a uniform distribution $\mathcal{U}([a,b])$ where $(a,b) \in [\underline{a},\overline{a}] \times [\underline{b},\overline{b}]$. Assume $\underline{a} < \underline{b} < \overline{a} < \overline{b}$, it is then possible during step 1 of the two dimensional Monte Carlo simulation, under independence assumptions, to obtain $(a_1, b_1) \in [\underline{a},\overline{a}] \times [\underline{b},\overline{b}]$ such that $\underline{a} < \underline{b} < b_1 < a_1 < \overline{a} < \overline{b}$ which is meaningless. According to this example, taking into account of dependencies between parameters is necessary, but not so obvious.

Even if the 2MC approach purposedly tries to separate variability from imprecision, its first main problem is that it treats partial ignorance in the same way it treats variability. Faced with imprecise information the two-dimensional Monte Carlo simulation does not allow to handle this kind of knowledge more correctly than a classical Monte-Carlo method. distinction between imprecision and variability. The second main problem is the treatment of uncertainty about parameter dependencies. The same difficulties appear as in classical Monte Carlo methods [21,23].

4.2 The fuzzy random variable approach

In this subsection, we propose a practical uncertainty propagation model expressed in terms of fuzzy random variables in order to represent a parametric probabilistic model tainted with imprecision faithfully. Consider ill-known random variables $(X_1, \ldots, X_n)$ represented by a class of probability measures $(P^{X_1}_{\vec{\theta}_1}, \ldots, P^{X_n}_{\vec{\theta}_n})$ and $T = f(X_1, \ldots, X_n)$, where $f : \mathbb{R}^n \to \mathbb{R}$ is a known mapping. For the sake of clarity in notations, we consider θ_i as a vector for all i (*i.e.* for instance $\theta_1 = (\mu, \sigma)$ where μ is the mean and σ the standard deviation). Assume that expert gives vector parameters $(\theta_1, \ldots, \theta_n) \in [\underline{\theta}_1, \overline{\theta}_1] \times \ldots \times [\underline{\theta}_n, \overline{\theta}_n]$ or confidence intervals $[\underline{\theta}_i, \overline{\theta}_i]_\alpha$ for all $i = 1 \ldots n$ with confidence level $1 - \alpha$. Let $Q^{X_i}_{\theta_i} :]0,1[\to \mathbb{R}$ be a possible quantile function of X_i such that

$$\forall u \in]0,1[,\ Q^{X_i}_{\theta_i}(u) = \inf\{x | F^{X_i}_{\theta_i}(x) \geq u\}$$

where $F^{X_i}_{\theta_i}$ defines a possible cumulative distribution function of X_i. We decide that $Q^{X_i}_{\theta_i}(0)$ is the smallest possible value for X_i and $Q^{X_i}_{\theta_i}(1)$ the greatest. The function $Q^{X_i}_{\theta_i}$ can be interpreted as the quasi-inverse function of $F^{X_i}_{\theta_i}$ and if a random variable U is associated with uniform distribution on $[0,1]$, $Q^{X_i}_{\theta_i}(U)$ then has cumulative distribution function $F^{X_i}_{\theta_i}$.

4.2.1 Interval case

Assume that random variables $U_1,\ldots,U_n$ follow uniform distributions on $[0,1]$ and that expert provides parameters $(\theta_1,\ldots,\theta_n) \in [\underline{\theta}_1,\overline{\theta}_1] \times \ldots \times [\underline{\theta}_n,\overline{\theta}_n]$. By combining random sampling with interval analysis, the distribution function becomes a random set $\Gamma_T : \Omega \to \mathcal{P}(\mathbb{R})$ defined as follows:

$$\Gamma_T(\omega) = [\inf_{i,\theta_i\in[\underline{\theta}_i,\overline{\theta}_i]} f(Q^{X_1}_{\theta_1}(U_1(\omega)),\ldots,Q^{X_n}_{\theta_n}(U_n(\omega))), \sup_{i,\theta_i\in[\underline{\theta}_i,\overline{\theta}_i]} f(Q^{X_1}_{\theta_1}(U_1(\omega)),\ldots,Q^{X_n}_{\theta_n}(U_n(\omega)))],\ \forall\omega\in\Omega.$$

If we perform a random sampling

$$\begin{pmatrix} u_1^1 & \ldots & u_1^m \\ \vdots & & \vdots \\ u_n^1 & \ldots & u_n^m \end{pmatrix}$$

of size m from a uniform distribution on $[0,1]$ according to dependencies (if known), the mass distribution ν can be defined by

$$\nu([\inf_{i,\theta_i\in[\underline{\theta}_i,\overline{\theta}_i]} f(Q^{X_1}_{\theta_1}(u_1^j),\ldots,Q^{X_n}_{\theta_n}(u_n^j)), \sup_{i,\theta_i\in[\underline{\theta}_i,\overline{\theta}_i]} f(Q^{X_1}_{\theta_1}(u_1^j),\ldots,Q^{X_n}_{\theta_n}(u_n^j))]) = \frac{1}{m},\ \forall j=1\ldots m.$$

By construction, we have $\forall\omega\in\Omega$, $T(\omega)\in\Gamma_T(\omega)$ and upper and lower bound probabilities can be estimated by means of plausibility and belief functions as defined in Section 2.3. For example, consider the trivial case $T = f(X)$ where $f : x \mapsto x$ and X is associated with the normal distribution $\mathcal{N}(\mu,\sigma)$ with $(\mu,\sigma)\in[\underline{\mu},\overline{\mu}]\times[\underline{\sigma},\overline{\sigma}]$. We thus define a random set $\Gamma_T : \Omega \to \mathcal{P}(\mathbb{R})$ such that

$$\Gamma_T(\omega) = [\inf_{(\mu,\sigma)\in[\underline{\mu},\overline{\mu}]\times[\underline{\sigma},\overline{\sigma}]} Q^X_{\mu,\sigma}(U(\omega)), \sup_{(\mu,\sigma)\in[\underline{\mu},\overline{\mu}]\times[\underline{\sigma},\overline{\sigma}]} Q^X_{\mu,\sigma}(U(\omega))]$$

where $Q^X_{\mu,\sigma} : u \mapsto \mu + \sigma \times Q_{0,1}(u)$ and $Q_{0,1}$ is a numerical approximation of the inverse normal distribution $\mathcal{N}(0,1)$. That means we have

$$\Gamma_T(\omega) = [min(\underline{\mu}+\underline{\sigma}\times Z(\omega),\underline{\mu}+\overline{\sigma}\times Z(\omega)), max(\overline{\mu}+\underline{\sigma}\times Z(\omega),\overline{\mu}+\overline{\sigma}\times Z(\omega))]$$

with $Z(\omega) = Q_{0,1}(U(\omega))$ (*i.e.* $Z = \mathcal{N}(0,1)$).

4.2.2 Fuzzy interval case

In this case, experts provide nested intervals $[\underline{\theta}_i,\overline{\theta}_i]_\alpha$ for all $i = 1\ldots n$ with certainty levels $1-\alpha$. Then, for each value of α, $f(X_1,\ldots,X_n)$ becomes a random set Γ^T_α :

$\Omega \to \mathcal{P}(\mathbb{R})$ defined for each confidence level $1-\alpha$ by

$$\Gamma_\alpha^T(\omega) = [\inf_{i,\theta_i \in [\underline{\theta}_i, \overline{\theta}_i]_\alpha} f(Q_{\theta_1}^{X_1}(U_1(\omega)), \ldots, Q_{\theta_n}^{X_n}(U_n(\omega))), \\ \sup_{i,\theta_i \in [\underline{\theta}_i, \overline{\theta}_i]_\alpha} f(Q_{\theta_1}^{X_1}(U_1(\omega)), \ldots, Q_{\theta_n}^{X_n}(U_n(\omega)))],\ \forall \omega \in \Omega.$$

This approach assumes a strong dependence between information sources pertaining to parameters $(\theta_1, \ldots, \theta_n)$, *i.e.* on the choice of the confidence level. This suggests that if the source informing θ_1 is rather precise then the one informing on other parameters is also precise. We thus define a fuzzy random variable $\tilde{T} : \Omega \to \tilde{\mathcal{F}}(\mathbb{R})$ from the above described random simulation process:

$$\mu_{\tilde{T}(\omega)}(t) = \sup\{\alpha \in [0,1] | t \in \Gamma_\alpha^T(\omega)\},\ \forall \omega \in \Omega$$

where the α-cuts of possibility distribution $\tilde{T}(\omega)$ correspond to the random set $\Gamma_\alpha^T(\omega)$. According to Section 3.1, the fuzzy random variable $\tilde{T}$ induces a fuzzy set $\tilde{P}_{\tilde{T}}$ of probability functions that in turn induces fuzzy probabilities $\tilde{P}_{\tilde{T}}(A)$ of events A, following the definitions in Section 3.1. As previously, if we perform a random sampling $(u_1, \ldots, u_m)$

$$\begin{pmatrix} u_1^1 & \ldots & u_1^m \\ \vdots & & \vdots \\ u_n^1 & \ldots & u_n^m \end{pmatrix}$$

the fuzzy random variable $\tilde{T}$ takes the possibility distribution values $\tilde{T}_1, \ldots, \tilde{T}_m$ with corresponding probabilities $1/m$. According to Section 3.2, we can estimate the lower and upper probabilities $[Bel, Pl]$ for all measurable events $A \subset \mathbb{R}$ such that:

$$Pl(T \in A) = \sum_{i=1}^{m} \frac{i}{m} \times \sup_{t \in A} \mu_{\tilde{T}_i}(t) \text{ and } Bel(T \in A) = \sum_{i=1}^{m} \frac{i}{m} \times \inf_{t \notin A} (1 - \mu_{\tilde{T}_i}(t))$$

For example, consider the random variable $T = f(X)$ where $f : x \mapsto x$ and X is represented by a normal distribution $\mathcal{N}(\mu, \sigma)$ where μ (resp. σ) is represented by the triangular possibility distribution with core $\{8.75\}$ and support $[7.5, 10]$ (resp. with core $\{1.5\}$ and support $[1, 2]$). Figure 4 shows lower and upper cumulative distributions $[Bel((-\infty, .]), Pl((-\infty, .])$ (see Section 3.2) and allows to display the order 1 model induced from the fuzzy random $\widetilde{T}$. Figure 5 presents lower and upper cumulative distributions $[\underline{F}_\alpha, \overline{F}_\alpha]$=$[Bel_\alpha((-\infty, .]), Pl_\alpha((-\infty, .])]$ for $\alpha \in \{0, 0.5, 1\}$ (see Section 3.1) and allows to display the second order possibility model induced from $\widetilde{T}$. We can see that $Bel((-\infty, .]) \simeq 1/3(\underline{F}_0(.) + \underline{F}_{0.5}(.) + \underline{F}_1(.))$ and $Pl((-\infty, .]) \simeq 1/3(\overline{F}_0(.) + \overline{F}_{0.5}(.) + \overline{F}_1(.))$.

Contrary to the previously shown postprocessing of the meta-distribution, obtained by the two dimensional Monte Carlo simulation, the new model presents the advantage of being able to estimate all measurable events. Compared to the method by

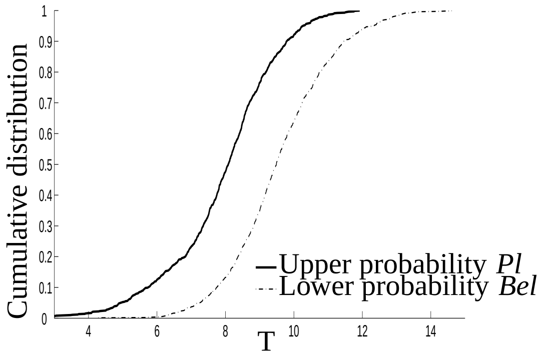

Fig. 4. First order model induced from a fuzzy random variable sampling

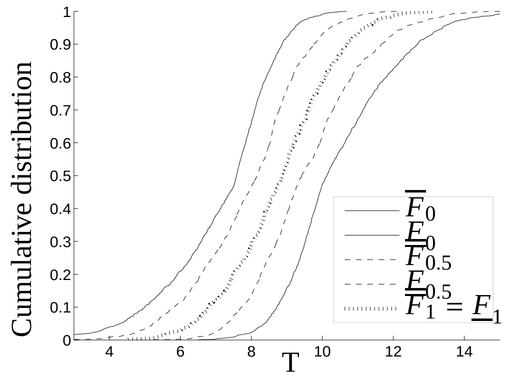

Fig. 5. Second order model induced from a fuzzy random variable sampling

Möller and Beer [32] (section 4.1.1 p. 110 et seq.), the proposed approach propagates imprecise and uncertain information exactly (up to discretization). While these authors compute a fuzzy probability density and a fuzzy p-box via fuzzy arithmetics, our method may compute the probability of any measurable event directly. Moreover, this model might be combined with other kinds of knowledge (see Section 5).

5 Illustrative example combining heterogeneous knowledge

Consider the previous mapping $f : (x, y) \mapsto x \times y$ and the ill-known random variables X and Y. In a first step, an expert provides predictive intervals about the quantity x defining a triangular possibility distribution π_X with core $[4, 6]$ and support $[3, 8]$. The quantity y is represented by a normal distribution $\mathcal{N}(\mu, \sigma)$ where $(\mu, \sigma) \in [7.5, 10] \times [1, 2]$ with the central values $(8.75, 1.5)$ being estimated more likely.
A random set $\Gamma_X : \Omega \to \mathcal{P}(\mathbb{R})$ is then defined such that $X(\omega) \in \Gamma_X(\omega)$ where $\Gamma_X(\omega)$ corresponds to an α-cut of π_X (*i.e.* $\Gamma(\omega) = \pi_X^{-1}(U(\omega))$ where $\pi_X^{-1}(z) = \{x \in \mathbb{R} / \pi_X(x) \geq z\}$ and $U = \mathcal{U}([0, 1])$). A triangular possibility distribution π^μ with core $\{8.75\}$ and support $[7.5, 10]$ (resp. π^σ with core $\{1.5\}$ and support $[1, 2]$) is proposed to represent the knowledge relative to μ (resp. σ). The fuzzy random variable $\widetilde{Y} : \Omega \to \widetilde{\mathcal{F}}(\mathbb{R})$ is thus defined such that

$$\mu_{\widetilde{Y}(\omega)}(z) = \sup\{\alpha \in [0, 1] | \; z \in [\inf_{(\mu,\sigma) \in \pi^\mu_\alpha \times \pi^\sigma_\alpha} Q^Y_{\mu,\sigma}(U(\omega)), \sup_{(\mu,\sigma) \in \pi^\mu_\alpha \times \pi^\sigma_\alpha} Q^Y_{\mu,\sigma}(U(\omega))]\}$$

where $U = \mathcal{U}([0, 1])$ and $Q^Y_{\mu,\sigma}$ is the inverse normal distribution $\mathcal{N}(\mu, \sigma)$ deduced by the numerical approximation of inverse normal distribution $\mathcal{N}(0, 1)$. Assuming stochastic independence between X and Y, we obtain a fuzzy random variable $\widetilde{T} : \Omega \to \widetilde{\mathcal{F}}(\mathbb{R})$, to represent the ill-known quantity $t = f(x, y)$, which verifies

$$\mu_{\widetilde{T}(\omega)}(z) = \sup\{\alpha \in [0, 1] | \; z \in f \circ (\Gamma_X(\omega), [\widetilde{Y}(\omega)]_\alpha)\}.$$

Figure 6 shows the lower and upper cumulative distributions (*Bel* $- - -$, *Pl* • • •)

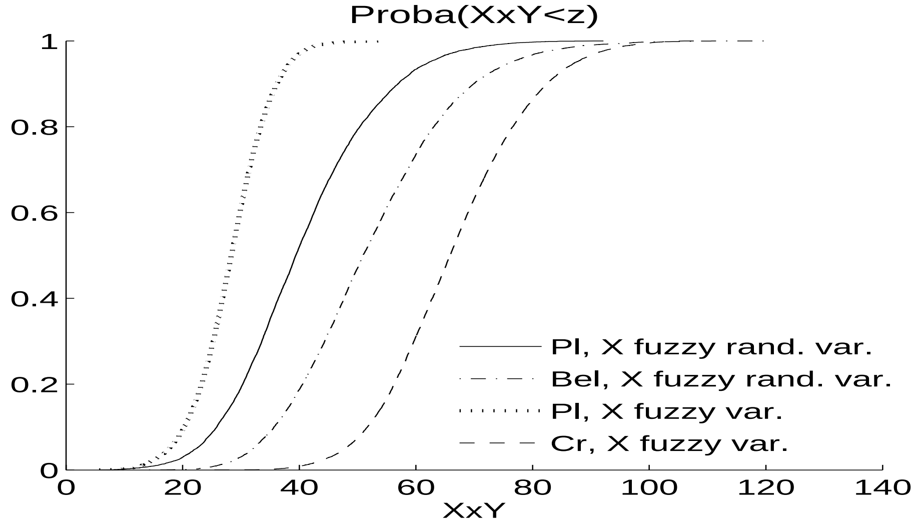

Fig. 6. Upper and lower cumulative probabilities of $X \times Y$ where X is either described by a possibility or by a parametric probabilistic model.

induced from the fuzzy random variable $\widetilde{T}$.

Assume now that expert upgrades his/her original knowledge pertaining to quantity x and provides a class of distributions for X namely a triangular distribution of support $[3, 8]$ and mode $M \in [4, 6]$ with a central value $\{5\}$ being estimated more likely. In order to represent the knowledge associated with the mode M, we propose a triangular possibility distribution π^M with core $\{5\}$ and support $[4, 6]$. One thus obtains the following class of possibility distributions:

$$\mu_{\widetilde{T}(\omega)}(z) = \sup\{\alpha \in [0, 1]/\ z \in f \circ ([\widetilde{X}(\omega)]_\alpha, [\widetilde{Y}(\omega)]_\alpha)\}$$

where $\widetilde{X} : \Omega \to \widetilde{\mathcal{F}}(\mathbb{R})$ is defined by

$$\mu_{\widetilde{X}(\omega)}(z) = \sup\{\alpha \in [0, 1]/\ z \in [\inf_{M \in \pi^M_\alpha} Q^X_M(U(\omega)), \sup_{M \in \pi^M_\alpha} Q^X_M(U(\omega))]\}$$

with Q^X_M is the inverse triangular cumulative distribution of support $[3, 8]$ and mode M. Figure 6 displays lower (*Bel* $. - . - .$) and upper (*Pl* $-$) probabilities induced from the new representation of X namely a fuzzy random variable. The total uncertainty about T can be characterized by the interval $[\overline{F}^{-1}(0.05), \underline{F}^{-1}(0.95)]$ corresponding to the lower 5% and the upper 95% percentiles of credibility and plausibility measures. According to Figure 6 we thus obtain the interval $[14, 87.5]$ with the first representation of X and $[17.5, 76]$ in the second one, that is a reduction of 20% of total uncertainty pertaining to T.

Compare now these results with the two-dimensional Monte Carlo approach. Figure 7 presents the cumulative distribution F_1 of $X \times Y$, (resp F_2) obtained with

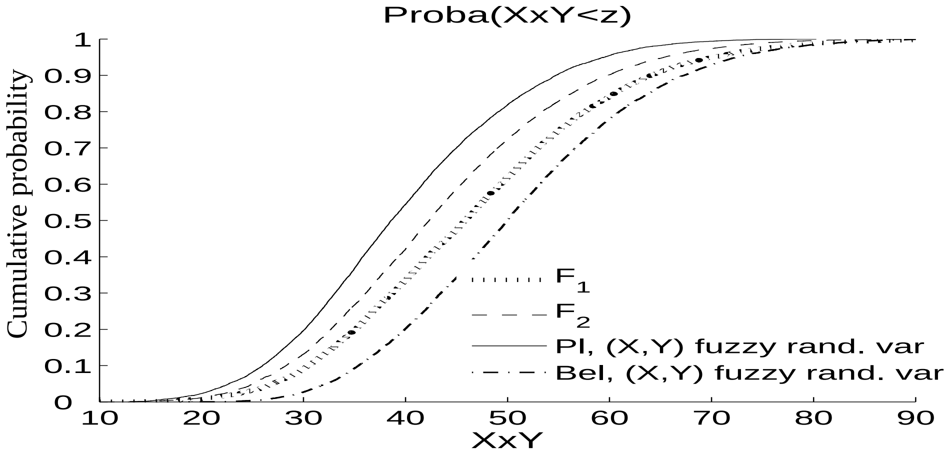

Fig. 7. First order probability and possibility measures

(M,μ,σ)=(Triang(4, 5, 6),$\mathcal{U}$([7.5,8.75]), Triang(1,1.5,2)) (resp (M,μ,σ)=($\mathcal{U}$([4,6]), $\mathcal{U}$([7.5,10]), $\mathcal{U}$([1,2]) assuming independence). To estimate F_1 (resp F_2), we use the Theorem of Total Probability namely: $F(t) = P(X \times Y \leq t) = \sum_i P((X,Y) = (x_i,y_i)) \times P(X \times Y \leq t|(X,Y) = (x_i,y_i))$. According to Figure 7, we can conclude that $F_1(40) = 0.3 < F_2(40) = 0.4$ but these results must be used with caution. Indeed, according to the nature of knowledge pertaining to (M,μ,σ), the estimated probabilities F_1 et F_2 are subjective. If we want to remain faithful to available information, we can only assert that the probability $P(X \times Y \leq 40)$ can potentially reach $Pl(X \times Y \leq 40) = 0.5$ and we are certain that it is not lower than $Bel(X \times Y \leq 40) = 0.2$.

6 Conclusion

In uncertainty analysis, imprecise knowledge pertaining to uncertain quantities is often modeled by parametric probabilistic models. When data are lacking, it is hard to specify the value of their parameters precisely. Because of its mathematical and computational simplicity, many analysts routinely use and recommend the two-dimensional Monte-Carlo simulation as a convenient approach to distinguish imprecision from variability in uncertainty analysis. This paper has recalled the main disadvantages and limits of the 2MC simulation, which may significantly underestimate over-estimate uncertainty about the results, and can thus be misleading. This paper suggests that parametric probabilistic models tainted with imprecision can be processed within the framework of fuzzy random variables and we propose a practical method based on combining Monte-Carlo simulation and interval analysis to represent them. When information about model parameter values is scarce, our approach looks more faithful to available knowledge than the two-dimensional Monte-Carlo method. There certainly might be situations, where statistical evidence is richer, where the two-dimensional Monte-Carlo method could be more appropriate. Nevertheless, it is not very easy to represent classes of probability distributions in general cases. It is thus interesting to further investigate the potential of triangular or trapezoidal possibility distributions in order to better control the probability families they encompass. The aim is to eliminate distributions that are not in conformity with the natural process under study, and thus to reduce imprecision. Because the presence of imprecision potentially generates two levels of dependency [7], further research is also needed for representing knowledge about dependence. Indeed, both the fuzzy random variable approach and the two-dimensional Monte-Carlo simulation cannot easily account for dependence between variables and parameters. Borrowing from results on rank correlations [4], copulas [33] and the general framework of upper and lower probabilities introduced by Couso et al. [7], we may try to take into consideration some links or dependencies which could exist between model parameters, and between the group of parameters and the group of variables.

References

[1] Baudrit, C., Couso, I., Dubois, D. Joint Propagation of Probability and Possibility in Risk Analysis : toward a formal framework. *International Journal of Approximate Reasoning*. In press.

[2] Baudrit, C., Dubois, D., Guyonnet, D. Joint Propagation and Exploitation of Probabilistic and Possibilistic Information in Risk Assessment Models. *IEEE Transaction on Fuzzy Systems*, vol 14, Issue 5, pp. 593-608, 2006.

[3] Baudrit, C., Dubois, D. Practical Representation of Incomplete Probabilistic Knowledge. *Computational Statistic and Data Analysis*, Vol 51, Issue 1, pp. 86-108, Nov 2006.

[4] Connover, W.J., Iman, R.L. A Distribution-Free Approach to Inducing Rank Correlation Among Input Variables. *Technometric*, 3, 311-334, 1982.

[5] Couso, I., Dubois, D., Montes, S., Sánchez, L. On various definitions of the variance of a fuzzy random variable. Proc. International Symposium on Imprecise Probability: Theories and Applications (ISIPTA 2007), Prague (Czech Republic), p. 135-144, 2007.

[6] Couso, I., Montes, S., Gil, P. Second order possibility measure induced by a fuzzy random variable. In C. Bertoluzza, M. A. Gill, and D. A. Ralescu (Eds.), *Statistical modeling, analysis and management of fuzzy data*, 127-144. Springer, Heidelberg, 2002.

[7] Couso, I., Moral, S., and Walley, P. A survey of concepts of independence for imprecise probabilities. *Risk Decision and Policy*, 5, 165-181, 2000.

[8] Couso, I., Sánchez, L. Higher order models for fuzzy random variables Fuzzy Sets and Systems, 159(3), 2008 (to appear).

[9] Cullen, A.C., Frey, H.C. *Probabilistic Techniques in Exposure Assessment: A Handbook for Dealing with variability and uncertainty in Models and Inputs*, Lenum press: New york, 1999.

[10] G. De Cooman, A behavioural model for vague probability assessments, *Fuzzy Sets and Systems*, Vol. 154, Issue 3, Pages 305-358, 2005.

[11] De Cooman, G., Aeyels, D. Supremum-preserving upper probabilities. *Information Sciences*, 118, 173-212, 1999.

[12] Dempster, A.P. Upper and Lower Probabilities Induced by a Multivalued Mapping. *Annals of Mathematical Statistics*, 38, 325-339, 1967.

[13] Destercke, S., Dubois, D. A unified view of some representations of imprecise probabilities. International Conference on Soft Methods in Probability and Statistics (SMPS 2006), J. Lawry , E. Miranda, A. Bugarin, S. Li (Eds.), Springer, Advances in Soft Computing, Volume 37, p. 249-257, 2006.

[14] Destercke, S., Dubois, D. Chojnacki, E., Relating practical representations of imprecise probabilities. Proc. International Symposium on Imprecise Probability: Theories and Applications (ISIPTA 2007), Prague (Czech Republic), p. 155-164, 2007.

[15] Dubois, D., Kerre, E., Mesiar, R, and Prade, H. Fuzzy interval analysis.*Fundamentals of Fuzzy Sets*, D. Dubois, H. Prade, Eds. Boston MA: Kluwer, 483-581, 2000.

[16] Dubois, D., Nguyen, H.T., Prade, H. Possibility theory, probability and fuzzy sets: misunderstandings, bridges and gaps. *Fundamentals of Fuzzy Sets*, Dubois,D. Prade,H., Eds: Kluwer , Boston, Mass, 343-438 , 2000.

[17] Dubois, D., Prade, H., Sandri, S.A. On possibility/probability transformations.*Fuzzy Logic: State of the Art* (R. Lowen, M. Lowen, eds.), Kluwer Academic Publ., 103-112, 1993.

[18] Dubois, D., Prade, H. When upper probabilities are possibility measures. *Fuzzy Sets and Systems*, 49, 65-74, 1992.

[19] Dubois D., and Prade, H. The mean value of a fuzzy number, *Fuzzy Sets and Systems*, 24, 279-300, 1987.

[20] Ferson, S. Bayesian methods in risk assessment. *Technical Report on using Bayes methods in risk analysis* http://www.ramas.com/bayes.pdf

[21] Ferson, S., R.B. Nelsen, J. Hajgos, D.J. Berleant, J. Zhang, W.T. Tucker, L.R Ginzburg, and W.L. Oberkampf. 2004. Dependence in Probabilistic Modeling, Dempster-Shafer Theory, and Probability Bounds Analysis. *SAND2004-3072, Sandia National Laboratories*, Albuquerque, 2004. http://www.ramas.com/depend.pdf

[22] Ferson, S., Ginzburg, L., Kreinovich, V., Myers, D.M., Sentz, K.Construction Probability Boxes and Dempster-Shafer structures. *Sandia National Laboratories, Technical report SANDD2002-4015*, 2003. URL:www.sandia.gov/epistemic/Reports/SAND2002-4015.pdf.

[23] Ferson, S., Ginzburg, L.R. Different methods are needed to propagate ignorance and variability. *Reliability Engineering and Systems Safety*, 54, 133-144, 1996.

[24] Guyonnet, D., Bourgine, B., Dubois, D., Fargier, H., Côme, B., Chilès, J.P. Hybrid approach for addressing uncertainty in risk assessments. *Journal of Environmental Engineering*, 126, 68-78, 2003.

[25] Gentle, J.E. Monte-Carlo methods. In S. Kotz and N.L. Johnson eds, editors, *Encyclopedia of Statistics*, 5, 612-617. John Wiley and Sons, New York, 1985.

[26] Helton, J.C. Treatment of uncertainty in performance assessments for complex systems. *Risk Analysis*, 14, 483-511, 1994.

[27] Hoffman, F.O. Hammonds, J.S. Propagation of uncertainty in risk assessment: the need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability. *Risk Analysis*, Vol. 14, pp. 707-712, 1994.

[28] Jaffray, J.Y., Bayesian updating and belief functions. *IEEE Trans. Systems, Man Cybern*, 22, 1144-1152.

[29] Kentel, E. and Aral, M.M. 2D Monte Carlo versus 2D Fuzzy Monte Carlo Health Risk Assessment. *International Journal for Stochastic Environmental Research and Risk Assessment*, 19, 86-96, 2005.

[30] Krätschmer, V. A unified approach to fuzzy random variables. *Fuzzy Sets and Systems*, 123, 1-9, 2001.

[31] Kriegler, E., Hermann, H. Utilizing belief functions for the estimation of future climate change. *International journal of Approximate Reasoning*, 39, 185-209, 2005.

[32] Möller, B., Beer, M. *Fuzzy Randomness: Uncertainty in Civil Engineering and Computational Mechanics*, Springer Verlag, 2004.

[33] Nelsen, R.B. An Introduction to Copulas. *Lecture Notes in Statistics*, Springer-Verlag, New York, v.139, 1999.

[34] Puri, M.L., Ralescu, D. Fuzzy random variables. *J. Math. Anal. Appl.*, 114, 409-422, 1986.

[35] Ralescu, D. Average level of a fuzzy set. In C. Bertoluzza, M. A. Gil, and D. A. Ralescu, editors (2002) Statistical modeling, analysis and management of fuzzy data, pages 119–126. Springer, Heidelberg.

[36] Shafer, G. *A Mathematical Theory of Evidence*. Princeton University Press, 1976.

[37] Smets P., Belief functions on real numbers. Int. J. of Approximate Reasoning, 40:181-223, 2005.

[38] Smets P. and Kennes R. (1994). The transferable belief model. *Artificial Intelligence*, 66, 191-234.

[39] U.S. EPA (2001) RAGS Vol. 3 Part A : Process for conducting Probabilistic Risk Assessment. Chapter 1, december 31, 2001.

[40] Walley, *P. Statistical Reasoning with Imprecise Probabilities*, Chapman and Hall, 1991.