Joint Treatment of Imprecision and Randomness in Uncertainty Propagation

Cédric Baudrit[1]   **Didier Dubois**[1]   **Dominique Guyonnet**[2]   **Hélène Fargier**[1]

[1]Institut de Recherche en Informatique de Toulouse
Université Paul Sabatier, 118 route de Narbonne 31062 Toulouse Cedex 4, France.
e-mail: baudrit@irit.fr, dubois@irit.fr, fargier@irit.fr
[2]Service Environnement et Procédés,
BRGM - 3, avenue Claude-Guillemin - BP 6009 - 45060 Orlans Cedex 2 - France
e-mail: d.guyonnet@brgm.fr

Abstract

This paper presents and studies in detail a hybrid method of uncertainty propagation for the case where knowledge regarding some parameters of a physical model is represented by probability measures, while others are represented by possibility measures or belief functions.

Keywords: (Random) Fuzzy Numbers, Probability, Possibility, Belief functions, dependence.

1 Introduction

Currently, decisions regarding the management of polluted sites very often rely on an evaluation of risks for man and the environment. Such an evaluation is typically performed with the help of models that simulate the transfer of pollutants from a source to a vulnerable target, for different scenarii of exposure. Due to time and financial constraints, information regarding model parameters is often incomplete. So the knowledge really available on model parameters is often vague. This leads to uncertainty that needs to accounted for in the decision-making process.

Uncertainty regarding model parameters may have essentially two origins [12]. It may arise from randomness due to natural variability resulting from heterogeneity or stochasticity. Or it may be caused by imprecision due to lack of information resulting, for example, from systematic measurement error or expert opinions. Imprecision and randomness are often confused in risk analysis. It may occur in practice, that some parameters of empirical models can be represented by probability distributions (variability) while others are better represented by possibility distributions (partial ignorance), or by belief functions of Shafer (variability and partial ignorance). Most researchers typically use either one or the other of these modes of uncertainty representation. But fewer are interested following in the combination of these different modes of representation (probability, possibility, belief function) in the same computation of risk.

Let $T:\mathbb{R}^n \to \mathbb{R}$ be a function (model) of n arguments x_i ($x = (x_1,...,x_n)$). The knowledge on parameters x_i can be represented by means of a probability, a possibility distribution or a mass function. The main issue is thus to carry the uncertainty attached to the variables over to $T(x)$ with the least possible loss of initial information. This is uncertainty propagation. Generally, in the evaluation of risks for man and the environment, one tries to estimate $P_{T(X)}(]-\infty,e])$, the probability that some pollutant concentrations be less than an absorbed dose limit e for example.

One can distinguish between three important issues [10] [12]: the first is how to represent the available information faithfully [1], the second is how to account for dependencies, correlations between the parameters in the propagation process (linear, non linear monotone dependency, interaction ...). For example the assumption of stochastic independence between parameters can generate too optimistic results [7] [8]. The last issue is the choice of the propagation technique [2] [9].

Kaufmann and Gupta [16] introduced hybrid numbers which simultaneously express imprecision (fuzzy number) and randomness (probabil-

ity). Ferson and Ginzburg [13] extended the approach of Kaufmann and used hybrid arithmetic to treat risk analysis [14]. We can consider hybrid numbers as random fuzzy numbers, which in fact can be encoded by means of belief functions (see further on).

In Section 2, we explain the hybrid method [13] [15] in detail. In Section 3, we study the links between the hybrid method and the random set approach [2] using the belief functions of Dempster-Shafer [2] to propagate uncertainties. In Section 4, we discuss the postprocessing step proposed by Guyonnet et al. [15] and propose alternative postprocessings for the hybrid method. We can see the "postprocessing" as a way to estimate the probability of events such that $]-\infty, e]$ from the results of the random fuzzy computation. We also present the postprocessing method of Ferson [13] [14] and compare it with our approach. We will compare results of the hybrid method with the random set approach on a synthetic example in Section 5.

2 Joint propagation of fuzzy numbers and probabilities

2.1 Methodology and discussion

Let us assume $k < n$ random variables $(X_1, ..., X_k)$ taking values $(x_1, ..., x_k)$ and $n - k$ possibilistic variables $(X_{k+1}, ..., X_n)$ taking values $(x_{k+1}, ..., x_n)$.

We explain in this section how to propagate the uncertainties generated by $(X_i)_{i=1...n}$ through T with the hybrid method. There are two steps (see Fig.1 [15]) that combine a Monte Carlo technique with the extension principe [6]. We first perform a Monte Carlo sampling of the random variables, thus processing variability (probability). The values thus obtained are then fixed $(X_1 = x_1, ..., X_k = x_k)$ and fuzzy interval analysis is used to estimate T. The knowledge on the value of $T(X)$ becomes a fuzzy subset. Random sampling is resumed and the process is performed in an iterative fashion to obtain a sample $(F_1, ..., F_m)$ of fuzzy subsets. $T(X)$ becomes a fuzzy random variable [17]. The Monte Carlo method has been criticized by Ferson [9] because it presupposes the independence of random variables. In the case where we know that random variables are independent, the Monte
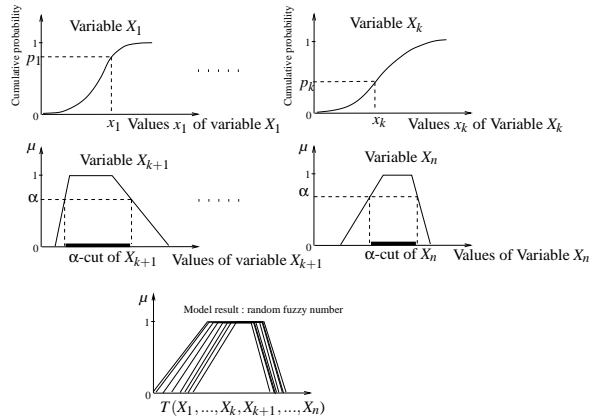

Figure 1: Schematic of the "hybrid" method

Carlo method is correct. It is worthwile noticing that within a Monte Carlo approach the rank correlation (non linear monotone dependency) between the random variables [4] can be taken into consideration (if known). Even if we can account for some dependencies between random variables with Monte-Carlo, it is necessary to be aware that the Monte Carlo method cannot account for all forms of dependency.

We must be careful with the extension principle because it also underlies an assumption on possibilistic variables. In fact the presence of imprecision on $X_{K+1}, ..., X_n$ generates two levels of dependencies. The first one is a dependency between information sources attached to variables and the second one is a dependency between variables themselves. The extension principe [6]:$\forall u \in \mathbb{R}$

$$\pi_T(u) = \sup_{x_{k+1}, ..., x_n, T(x_1, ..., x_n) = u} \min(\pi_1(x_{k+1}), ..., \pi_n(x_n))$$

first assumes strong dependence between information sources pertaining to possibilistic variables, i.e. on the choice of confidence levels or α-cuts induced by these confidence levels. For instance, one expert gives confidence degrees (sets) to two possibilistic variables A and B. That is if we know that A has a small imprecision then B will have a small imprecision. However, this form of dependency does not suppose any dependence between possibilistic variables. The use of "min" assumes the non-interaction of $X_{k+1}, ..., X_n$, which expresses a lack of knowledge about the links between the values of $X_{k+1}, ..., X_n$ and a lack of commitment as to whether $X_{k+1}, ..., X_n$ are linked or

not. The hybrid method also supposes independence between the group of probabilistic and possibilistic variables. It raises the following question for this method: how to take into consideration dependencies between the possibilistic variables, and between the random and possibilistic variables if such dependencies exist and are known?

3 Uncertainty propagation in the setting of the random sets

Belief functions [18] encompass possibility and probability theory. We will consider two variables and a continuous function T to illustrate the links between the propagation results obtained with the random set approach and the hybrid approach. Let X be a discrete random variable with $\Omega_X = \{x_1, ..., x_m\}$ and $p_i = P(X = x_i)$, Y a possibilistic variable. We denote π the possibility distribution of Y and π_α the α-cuts of π. Focal elements for X are singletons $(\{x_i\})_{i=1...m}$ and the mass distribution is equal to $(p_i)_{i=1..m}$ because X is discrete. We choose a discrete probability for the sake of clarity. Note however that if X has a probabilistic assignment function, by discretizing it, focal elements can be disjoint intervals and p_i can be equal to the surface under the probability distribution function for each disjoint interval. Focal elements for Y are denoted $(\pi_{\alpha_j})_{j=1...q}$ with $\alpha_q > 0$ and are nested. We denote $(\nu_j = \alpha_j - \alpha_{j+1})_{j=1...q}$ the mass distribution associated to $(\pi_{\alpha_j})_{j=1...q}$. We thus encode probabilistic and possibilistic variables as belief functions. With the hybrid method, $T(X,Y)$ is a discrete random fuzzy subset. That is, we obtain m fuzzy numbers $(\pi_T^i)_{i=1...m}$ with probability $(p_i)_{i=1..m}$. With the random set approach [2] we obtain mq focal elements (intervals) with mass distribution $(p_i\nu_j)_{i=1...m, j=1...q}$ and focal elements $T_{ij} = T(x_i, \pi_{\alpha_j})$. We have the following result: $\forall A$ measurable set

$$Pl_T(A) = \sum_{i=1}^{m} p_i \Pi_T^i(A)$$

$$Bel_T(A) = \sum_{i=1}^{m} p_i N_T^i(A)$$

Where Π_T^i (N_T^i) are the possibility (necessity) measures associated to fuzzy numbers π_T^i. Pl_T

(Bel_T) is the plausibility (belief) measure associated to focal elements T_{ij} and mass distribution $(p_i\nu_j)_{i=1...m, j=1...q}$. The calculation of Pl_T reads as follow:

$$Pl_T(A) = \sum_{(i,j),\ A \cap T_{ij} \neq \emptyset} p_i \nu_j$$

$$Pl_T(A) = \sum_{i=1}^{m} p_i \sum_{j=1...q,\ A \cap T_{ij} \neq \emptyset} \nu_j$$

$$\text{so,} \quad Pl_T(A) = \sum_{i=1}^{m} p_i Pl_T^i(A)$$

For each i varying from 1 to n, we have $T_{ij} \subseteq ... \subseteq T_{ik}\ \forall j \geq k$. Thus $Pl_T^i(A) = \Pi_T^i(A)$.

These results still hold when several probabilistic variables are involved because the random set approach presupposes the independence between focal elements. These results do not apply with more than one possibilistic variable. Indeed recall that fuzzy arithmetic presupposes a total dependency between α-cuts. Let us see when the two approaches are equivalent in the case where there are several possibilistic and probabilistic variables. Consider X, Y, two possibilistic variables encoded as belief functions by their focal elements $(\pi_{\alpha_i}^X)_{i=1...q}$, $(\pi_{\alpha_j}^Y)_{j=1...q}$ and the mass distributions $(\nu_i^X)_{i=1...q}$, $(\nu_j^Y)_{j=1...q}$. Let two discrete probabilistic variables Z, W be encoded by their focal elements $(\{z_k\})_{k=1...m}$, $(\{w_l\})_{l=1...m}$ and the mass distributions $(p_k^Z)_{k=1...m}$, $(p_l^W)_{l=1...m}$. With the random set approach, we define the mass distribution (noted ν_{ijkl}), associated to focal elements $T_{ijkl} = T(\pi_{\alpha_i}^X, \pi_{\alpha_j}^Y, \{z_k\}, \{w_l\})$ of $T(X,Y,Z,W)$, by:

$$\forall i,j,k,l \quad \nu_{ijkl} = \nu_i^X \nu_j^Y p_k^Z p_l^W$$

We thus assume independence of focal elements. Put now

$$\forall i,j,k,l;\ i=j \quad \nu_{ijkl} = \nu_i^{X,Y} p_k^Z p_l^W$$
$$\forall i,j,k,l;\ i \neq j \quad \nu_{ijkl} = 0$$

The possibility distribution $\pi_{X,Y}$ is characterized by $min(\pi_X, \pi_Y)$ what correspond to nested cartesian products of α-cuts and $\nu_i^{X,Y}$ is the mass associated to the cartesian product $\pi_{\alpha_i}^X \times \pi_{\alpha_i}^Y$. We thus assume a total dependency between focal elements associated to possibilistic variables. Hence,

if we want to estimate $Pl_T(A)$ $\forall A$ measurable set, using the last definition of v_{ijkl}, we still have:

$$Pl_T(A) = \sum_{i,k,l;\; A\cap T_{iikl}\neq\emptyset} v_i^{X,Y} p_k^Z p_l^W = \sum_{k,l} p_k^Z p_l^W \Pi_T^{kl}(A)$$

where Π_T^{kl} are the possibility measures associated to the fuzzy numbers π_{kl} obtained by the hybrid method.

4 Postprocessing of the hybrid method

4.1 The proposal of Guyonnet et al. [15]

Guyonnet et al. [15] propose to synthetise the random fuzzy result into a single fuzzy subset denoted F_d. For each α-cut of the random fuzzy subset obtained, Guyonnet et al. separately rearrange the left side and the right side of sets in increasing order. The set $[Finf_d^\alpha, Fsup_d^\alpha]$ is considered such that $P(leftside \leq Finf_d^\alpha) = 1 - d\%$ and $P(rightside \leq Fsup_d^\alpha) = d\%$. Varying $\alpha \in [0,1]$, a fuzzy interval F_d is thus built. The standard value $d = 95$ is chosen. That is they eliminate 5% on the left and on the right side and perform the union of the rest.

Starting from this F_d, we now can try to estimate the probability of events such that: $]-\infty, e]$, $]e, +\infty]$, $]e_1, e_2]$. However, there are caveats with this postprocessing. We get false estimates of $P_{T(X)}([e_1, e_2])$ with it. Indeed we treat independently *leftside* and *rightside* whereas *rightside* is entirely determined by *leftside* and conversely since any α-cut is generated as a whole. The more important problem is they confuse variability and imprecision. It does not account for the probabilities generated by the random variables and it thus forgets this knowledge (frequency of each fuzzy number). It may put excessive weights on randomly generated fuzzy numbers located on the extreme right and left parts of the result F_d. Indeed they may obtain the same fuzzy number F_d whether the F_i's have large imprecision and small variability as in Fig.2a, or are more precise with a great variability as in Fig.2b.

We can illustrate these problems more clearly when combining intervals and probability. For instance, let A, B be two independent random variables and C be an interval such that: $P(A = 1) = P(A = 2) = 0.5$, $P(B = 4) = 1/3$, $P(B = 6) = 2/3$

2a. Small variability of the sample
Large imprecision of each fuzzy number Fi

2b. Great variability of the sample
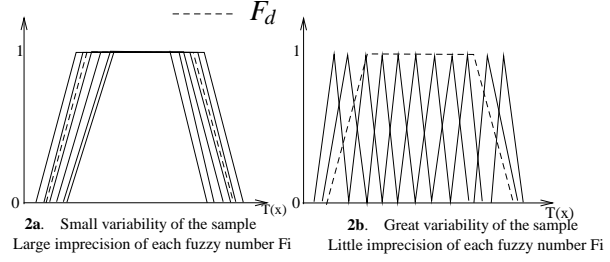Little imprecision of each fuzzy number Fi

Figure 2: A same result F_d obtained by the postprocessing proposed by Guyonnet et al. on two different possible results obtained by the hybrid method.

and $C = [1, 2]$. We compute $T = (A + B)/C$. With the hybrid method, we obtain a random interval: $T_1 = [2.5, 5]$ with probability 1/6, $T_2 = [3.5, 7]$ with probability 2/6; $T_3 = [3, 6]$ with probability 1/6 and $T_4 = [4, 8]$ with probability 2/6. Putting $d = 20\%$, with this postprocessing we obtain $T_d = [3, 7]$, and we assign to it a mass equal to 1, which is debatable. Indeed we eliminate the knowledge (frequency) brought by A and B, i.e. variability. As will be observed on some examples, the use of F_d may put excessive weight on outlier fuzzy values, in the case of precise results with high variability, even if some outliers have been deleted by the threshold d. The postprocessing proposed by Guyonnet et al. is thus debatable. Better alternative postprocessings can treat variability and imprecision in the hybrid method.

4.2 Processing variability and imprecision separately

We may wish to process variability and imprecision separely. To evaluate the average imprecision of T, we can estimate an average fuzzy number F_d^{mean} with:

$$F_d^{mean} = \frac{1}{m}\sum_{i=1}^{m} F_i$$

and compute the area under F_d^{mean}.

To estimate the variability of T, we can work with a representative value v_i^r of each fuzzy number F_i. Then we can estimate a variance V of the form:

$$V = \frac{1}{m}\sum_{i=1}^{m} v_i^{r2} - \frac{2}{m(m-1)}\sum_{j<i} v_j^r v_i^r$$

where v_i^r is a representative value of F_i. If V is small, the variability is small as in Fig.2a for

example. As the representative value v_i^r we can choose the middle point of the mean interval [19] which also is equal to the average of the pignistic probability associated to F_i :

$$v_i^r = \int_0^1 \frac{(supF_i^\alpha + infF_i^\alpha)}{2}d\alpha$$

V appears only as an indicator of result variability. We could also to estimate a fuzzy variance $F_d^{variance}$. Let f be the function which estimate the variance:

$$f : (x_1, ..., x_m) \mapsto \frac{1}{m}\sum_{i=1}^m x_i^2 - \frac{2}{m(m-1)}\sum_{j<i} x_j x_i$$

To obtain $F_d^{variance}$, we work by α-cut, we build nested intevals $F_{d,\alpha}^{variance} = [\underline{F_{d,\alpha}^{variance}}, \overline{F_{d,\alpha}^{variance}}]$ while solving:

$$\underline{F_{d,\alpha}^{variance}} = \inf_{x_i \in F_{i,\alpha}} f(x_1, ..., x_m)$$

$$\overline{F_{d,\alpha}^{variance}} = \sup_{x_i \in F_{i,\alpha}} f(x_1, ..., x_m)$$

This fuzzy variance describes a potential variability, because it scans all possible values of V compatible with the fuzzy data. Ferson et al. [11] propose an algorithm, which is of quadratic complexity, for computing the exact lower bound $\underline{F_{d,\alpha}^{variance}}$ of the sample variance for interval valued data. However, they show that computing the exact upper bound $\overline{F_{d,\alpha}^{variance}}$ is NP-hard. There exists an algorithm that computes $\overline{F_{d,\alpha}^{variance}}$ but it is exponential in the sample size. They propose an algorithm which is of quadratic complexity, but it presupposes all the interval midpoints are definitely different from each other. Computing a fuzzy variance doesn't appear straightforward, as we must apply these algorithms for all α-cuts.

4.3 Computing upper and lower cumulative distributions.

The use of belief functions [18] enables variability and imprecision to be processed in a common framework. Let $(\pi_i)_{i=1...m}$ be the sample of random fuzzy numbers resulting from the hybrid method. Let p_i be the probability associated at fuzzy number π_i. We encode each π_i with focal elements $(\pi_i^\alpha)_{i\alpha}$ and the mass distribution

associated $(\nu_\alpha p_i)_{i\alpha}$. We obtain a random sampling of intervals $(\pi_i^\alpha)_{i,\alpha}$ with a mass distribution $(\nu_\alpha p_i)_{i,\alpha}$. Then, we can estimate, $\forall A$ measurable set, $Pl_T(A)$ and $Bel_T(A)$ such that:

$$Pl_T(A) = \sum_{(i,\alpha);\ \pi_i^\alpha \cap A \neq \emptyset} \nu_\alpha p_i$$

$$Bel_T(A) = \sum_{(i,\alpha);\ \pi_i^\alpha \subseteq A} \nu_\alpha p_i$$

This homogeneous technique yields:

$$Pl_T(A) = \sum_i p_i \Pi_i(A)$$

$$Bel_T(A) = \sum_i p_i N_i(A)$$

If the Monte Carlo method yields distinct fuzzy numbers, then $p_i = (1/m)$ in this case. Let us compare the Guyonnet et al. [15] postprocessing with $Pl_T(] - \infty, e])$ and $Bel(] - \infty, e])$). For the sake of clarity, in Fig.3 three possibilistic distributions are shown with their associated probabilities to help us in the comparison. Value p_i represents the probability to ob-
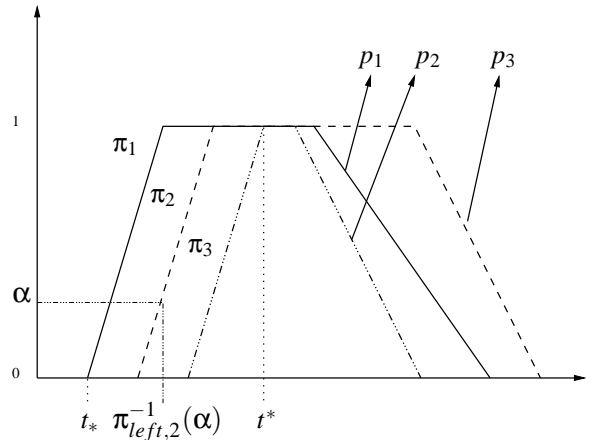

Figure 3:

tain the fuzzy number π_i. With the homogeneous postprocessing, we get $Pl_T(] - \infty, t^*]) = 1$ if and only if $\forall i = 1...m\ \Pi_i(] - \infty, t^*]) = 1$. That is $t^* = \max_i\{inf(core(\pi_i))\}$. We have also $Pl_T(] - \infty, t_*]) = 0$ if and only if $\forall i = 1...m\ \Pi_i(] - \infty, t_*]) = 0$. That is $t_* = \min_i\{inf(support(\pi_i))\}$. With the Guyonnet et al. method, we build for each α-cut the following intervals: $F_{d,\alpha} = [Finf_d^\alpha, Fsup_d^\alpha]$ such that $Proba(leftside \leq Finf_d^\alpha) = 1 - d\%$ and $P(rightside \leq Fsup_d^\alpha) = d\%$ (see Sec.4.1). We

study only the left part, that is upper probability. Note $Finf_d^\alpha = F_{left,\alpha}^{-1}(1-d\%)$ where $F_{left,\alpha}$ is the cumulative distribution function encoded by the probabilities p_i. We have: $F_{left,\alpha}^{-1}(1-d\%) = \sum_{i=1}^{m} \pi_{left,i}^{-1}(\alpha)\delta_{[p_0+...+p_{i-1},p_0+...+p_i[}(1-d\%)$. If $d \neq 100\%$, the construction of $F_{left,\alpha}^{-1}$ will necessarily imply:

$$for \quad \alpha = 0, \quad F_{left,0}^{-1}(1-d\%) \geq t_*$$

$$for \quad \alpha = 1, \quad F_{left,1}^{-1}(1-d\%) \leq t^*$$

$\Pi_T : e \mapsto]-\infty,e]$ obtained by Guyonnet et al. and $Pl_T : e \mapsto]-\infty,e]$ obtained by homogeneous postprocessing are increasing. We can conclude for $d \neq 100\%$ it exists a value t_c such that $\Pi_T(]-\infty,t_c]) = Pl_T(]-\infty,t_c])$ (see Fig.4). With a similar reasoning, we can also conclude it exits a value t_r such that $Bel_T(]-\infty,t_r]) = N_T(]-\infty,t_r])$. Thus we can say that with hybrid postprocessing according to [15], we will be more conservative on the right side of t_c and less conservative on the left side of t_c than our homogeneous postprocessing.
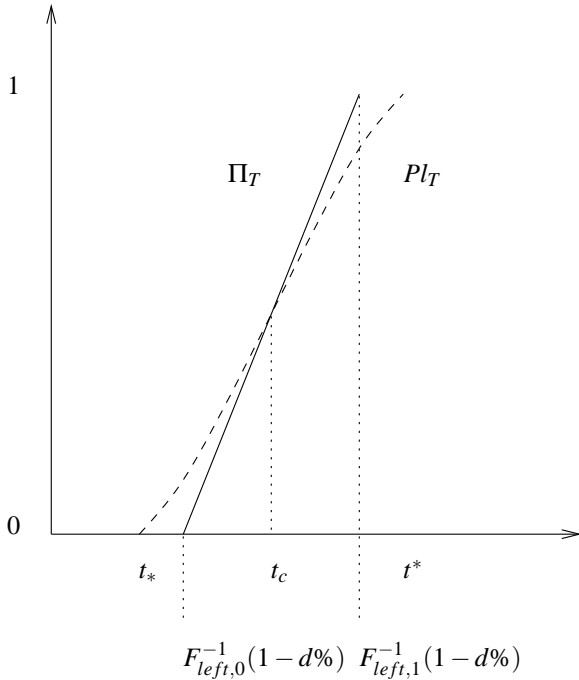

t_* t_c t^*

$$F_{left,0}^{-1}(1-d\%) \quad F_{left,1}^{-1}(1-d\%)$$

Figure 4:

4.4 Comparison with Ferson postprocessing method

Ferson [13] [14] treats variability and imprecision separately with random fuzzy numbers. In fact,

he fixes a degree of confidence (an α-cut) and is interested in the upper and lower cumulative probability distributions at this α-cut encoded by $(p_i)_i$. Thus if we are optimistic, we use the upper $\overline{P_1}$ and lower $\underline{P_1}$ cumulative distributions at $\alpha = 1$, but if we are pessimistic one uses the upper $\overline{P_0}$ and lower $\underline{P_0}$ cumulative distributions at $\alpha = 0$. We thus obtain two pairs of upper and lower cumulative distributions (see Fig.5). The gap between $\overline{P_\alpha}$ and $\underline{P_\alpha}$ represents the imprecision due to possibilistic variables.

We use the same notations as in Section 4.3 and compare this postprocessing with the previous one. With the homogeneous postprocessing, we have random intervals $(\pi_i^\alpha)_{i\alpha}$ associated to a mass distribution $(\nu_\alpha p_i)_{i\alpha}$. Note that $\pi_i^\alpha = [\underline{a_{i\alpha}},\overline{a_{i\alpha}}]$. For instance $Pl_T(]-\infty,e]) =$
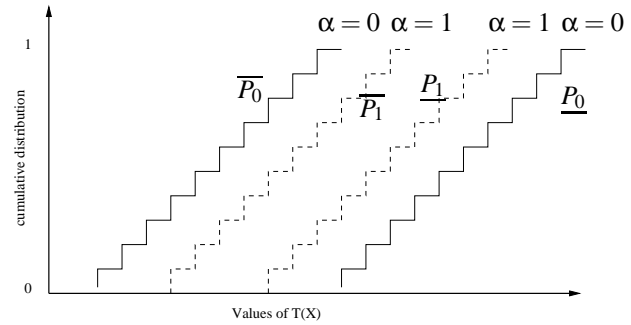

Figure 5: Postprocessing of Ferson with a sample of random fuzzy numbers equal to 10.

$\sum_{i,\alpha} \nu_\alpha p_i I_{\underline{a_{i\alpha}} \leq e}$ where $I_{\underline{a_{i\alpha}} \leq e} = 1$ if $\underline{a_{i\alpha}} \leq e$, otherwise 0. We work with finite sums, that is $i = 1...m$ and $\alpha = 1...q$. At i fixed, we have $\pi_{i,\alpha_1} \subseteq \pi_{i,\alpha_2}$ $\forall \alpha_1 \geq \alpha_2$. Thus, we have $\forall i,\alpha; \quad \forall e$:

$$I_{\underline{a_{iq}} \leq e} \quad \geq \quad I_{\underline{a_{i\alpha}} \leq e}$$

$$I_{\underline{a_{i1}} \leq e} \quad \leq \quad I_{\underline{a_{i\alpha}} \leq e}$$

Thus we have:

$$\sum_{i,\alpha} \nu_\alpha p_i I_{\underline{a_{iq}} \leq e} \quad \geq \quad \sum_{i,\alpha} \nu_\alpha p_i I_{\underline{a_{i\alpha}} \leq e}$$

$$\sum_{i,\alpha} \nu_\alpha p_i I_{\underline{a_{i1}} \leq e} \quad \leq \quad \sum_{i,\alpha} \nu_\alpha p_i I_{\underline{a_{i\alpha}} \leq e}$$

Thus

$$\sum_{\alpha} \nu_\alpha \sum_i p_i I_{\underline{a_{i1}} \leq e} \leq \sum_{i,\alpha} \nu_\alpha p_i I_{\underline{a_{i\alpha}} \leq e} \leq \sum_{\alpha} \nu_\alpha \sum_i p_i I_{\underline{a_{iq}} \leq e}$$

Then

$$\sum_i p_i I_{\underline{a_{i1} \leq e}} \ \leq \ \sum_{i,\alpha} \nu_\alpha p_i I_{\underline{a_{i\alpha} \leq e}} \ \leq \ \sum_i p_i I_{\underline{a_{iq} \leq e}}$$

That is

$$\overline{P_1} \leq Pl_T(]-\infty, e]) \leq \overline{P_0}$$

With a similar reasoning, we also show that:

$$\underline{P_0} \leq Bel_T(]-\infty, e]) \leq \underline{P_1}$$

5 Numerical example

Consider two possibilistic variables A, B (fuzzy numbers such that support(A)=[1,5], core(A)=[2,3]; support(B)=[3,8], core(B)=[4,6]), and one random variable C=Norm(7.5,1). We try to estimate $D = (A+B)/C$. This expression D commonly appears in risk assessment.

ag replacements
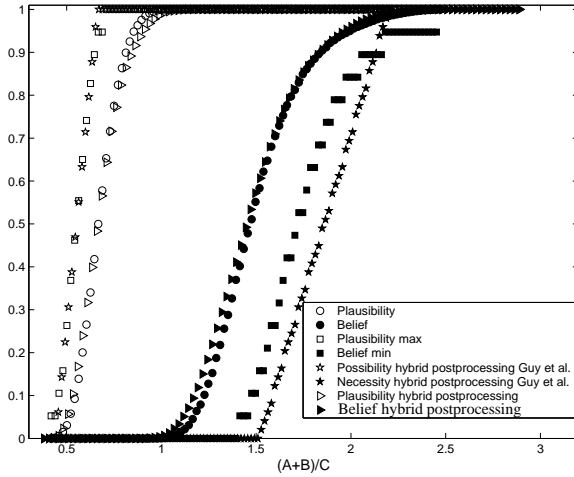
rid postprocessing Guy. et al.

Figure 6: Comparison of upper and lower cumulative distributions obtained by the random set approach and the postprocessing hybrid methods.

We used Guyonnet (stars on Fig.6) Ferson ($*,\diamond$ on Fig.7) and our homogeneous postprocessing methods ($\triangleright, \blacktriangleright$ on Fig.7 and Fig.6). We can compare these postprocessings on Fig.6 with the Dempster-Shafer (random set) approach of Sec.3 ($\circ, \bullet$ on Fig.6) and a conservative Dempster-Shafer approach ($\square, \blacksquare$ on Fig.6) [2] [3] to estimate upper and lower cumulative distributions of $(A+B)/C$. The conservative approach performs propagation without assuming knowledge on dependencies between variables. Dempster-Shafer conservative approach produces an enveloppe (Bel_{min} $\blacksquare$, Pl_{max} $\square$) for all the results ex-

cept for the Guyonnet et al. postprocessing. Indeed, the latter can sometimes be too conservative. It is due to the confusion between variability and imprecision plus the deletion of the probabilities (see sec.4.1). Bel_{min} and Pl_{max} remain the most credible conservative bounds on the cumulative distribution of $(A+B)/C$ in the case where we have no knowledge on dependencies. However, we can see that our postprocessing of the hybrid method (sec.4.3) is bracketed by Bel_{min} and Pl_{max}. We thus have for instance with our homogeneous postprocessing: $0.6 \leq P(\frac{A+B}{C} \leq 1.5) \leq 1$. We have no information on $P(\frac{A+B}{C} \leq 1)$ because we only know: $0 \leq P(\frac{A+B}{C} \leq 1) \leq 1$.
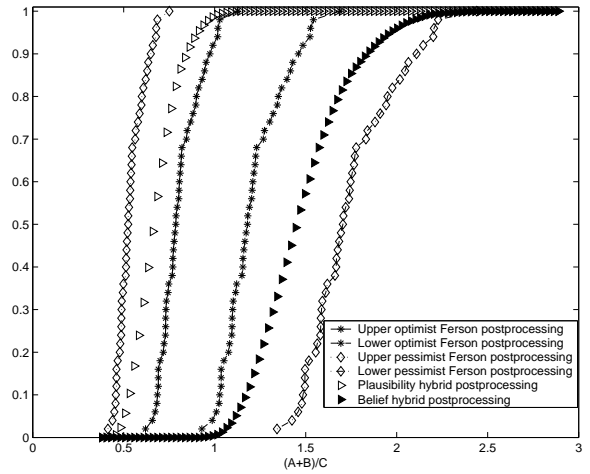
Figure 7: Comparison between Ferson and our homogeneous postprocessing results.

6 Conclusion

This paper proposes an approach to jointly propagate probabilistic and possibilistic uncertainty in deterministic mathematical models, and a postprocessing technique based on belief functions that computes imprecision and variability of the results separately, and extracts upper and lower cumulative distributions. Our proposal improves over the postprocessing method of Guyonnet et al. [15] which confuse variability and imprecision and neglect probabilities. Our method described in Section 4.3 seems to be an adequate approach to treat in a common framework the imprecision and the variability of the hybrid results knowing that probabilistic variables are independent; information sources on possibilistic variables are totally dependent and finally, probabilistic and pos-

sibilistic variables are independent of each other. Accounting for dependence is a very hard problem in the propagation process. Using ideas of [4], we may try to find a method to take into consideration some links or dependencies which could exist between possibilistic variables.

References

[1] C. Baudrit, D. Dubois, H. Fargier. Reprsentation de la connaissance probabiliste incomplte. *accepted to LFA'03*, 2003.

[2] C. Baudrit, D. Dubois, H.Fargier. Propagation of Uncertainty involving Imprecision and Randomness. *THIRD EUSFLAT'03*, 653-658, 2003.

[3] D. Berleant, C. Goodman-Strauss. Bounding results of Arithmetic Operations on Random Variables of Unknown Dependencies using Interval Arithmetic. *Reliable Computation*, 4, 147-165, 1998.

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

[5] D. Dubois, H. Prade. Random sets and Fuzzy Interval Analysis. *Fuzzy Sets and Systems*, 42, 87-101, 1991.

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

[7] M.A. Burman, S. Ferson. Correlation, dependency bounds and extinction risks. *Biological Conservation*, 73, 101-105, 1995.

[8] S. Ferson, T.F. Long. Conservative uncertainty propagation in environmental risk assessments. *Environmental Toxicology and Risk Assessment*. ASTM STP 1218, Ed J.S Hughes et al. Am. Soc. for Testing and Materials, Philadelphia, 97-110, 1994

[9] S. Ferson. What Monte Carlo methods cannot do. *Human and Ecology Risk Assessment*, 2, 990-1007, 1996.

[10] S.Ferson, L.Ginzburg, R.Akcakaya. Whereof one cannot speak: when input distributions are unknown. *Accepted for publication in Risk Analysis*, 2003.

[11] S. Ferson and L. Ginzburg and V. Kreinovich and L. Longpre and M. Aviles Computing Variance for Interval Data is NP-Hard. *ACM SIGACT News*, 33(2), 108-118, 2002.

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

[13] S. Ferson, L.R. Ginzburg. Hybrid Arithmetic. *Proceedings of ISUMA-NAFIPS'95, IEEE Computer Society Press, Los Alamitos, California*, 619-623, 1995.

[14] S. Ferson. Using fuzzy arithmetic in Monte Carlo simulation of fishery populations. *Management of Exploited Fish, T. Quinn (ed.), Proceedings of the ISMSEFP, Anchorage*, 595-608, 1992.

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

[16] A. Kaufmann, M.M. Gupta. Introduction to Fuzzy Arithmetic: Theory and Applications. *Van Nostrand Reinhold, New York*, 1985.

[17] R. Kruse, K.D Meyer. Statistics with Vague Data. *Reidel, Dordrecht, Netherlands*, 1987.

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

[19] R.R Yager. A procedure for ordering fuzzy subsets of the unit interval. *Information science*, 24, 143-161, 1981.