Comparing methods for joint objective and subjective uncertainty propagation with an example in a risk assessment

Cédric Baudrit **Didier Dubois**
Institut de Recherche en Informatique de Toulouse
Université Paul Sabatier, 118 route de Narbonne 31062 Toulouse Cedex 4, France
baudrit@irit.fr dubois@irit.fr

Abstract

Probability-boxes, numerical possibility theory and belief functions have been suggested as useful tools to represent imprecise, vague or incomplete information. They are particularly appropriate in environmental risk assessment where information is typically tainted with imprecision or incompleteness. Based on these notions, we present and compare four different methods to propagate objective and subjective uncertainties through multivariate functions. Lastly, we use these different techniques on an environmental real case of soil contamination by lead on an ironworks brownfield.

Keywords. Imprecise Probabilities, Possibility, Belief functions, Probability-Boxes, Dependency Bounds.

1 Introduction

In risk analysis, uncertainty regarding model parameters has two origins. It may arise from randomness (often referred to as "objective uncertainty") due to natural variability of observations. Or it may be caused by imprecision (often referred to as "subjective uncertainty") due to a lack of 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 encoded either by probability-boxes (upper & lower cumulative distribution functions [18, 19]) or possibility distributions (also called fuzzy intervals) [12] or yet by random intervals using belief functions of Shafer [27].

Let T: $\mathbb{R}^n \to \mathbb{R}$ be a function (model) of n parameters x_i ($x = (x_1, ..., x_n)$). The knowledge on parameters x_i can be represented using a distribution of probability, of possibility, a mass function or a probability-box. Generally, in the evaluation of risks for man and the environment, one tries to estimate the probability $P_{T(X)}((-\infty, e])$ that the output remains under a threshold e, an absorbed pollutant dose limit for example. The current interest for the problem of propagating heterogeneous uncertainties through mathematical models of phenomena of interest is witnessed by recent publications of special issues of risk analysis journals, like in [25].

In this paper, we present and compare four different methods to estimate $P_{T(X)}((-\infty, e])$. The first method, dubbed "*Hybrid*", processes variability and imprecision separately by combining a Monte-Carlo technique with the extension principle of fuzzy set theory [10]. The second method, dubbed "*Independent Random Sets*", processes variability and imprecision in the common framework of belief functions and assumes independence between focal sets [2]. The third method, dubbed "*Conservative Random Sets*", uses the same idea as the "*Independent Random Sets*" approach except that we now do not assume anything on the dependency between variables. It determines the maximal (resp. minimal) plausibility of events of interest (resp. belief) by solving a linear optimization problem [2]. The last method, dubbed "*Dependency Bounds Convolution*" proposed by Williamson and Downs [29], gives extreme bounds resulting from convolutions of two random variables under unknown dependency. This method propagates probability-boxes (upper & lower cumulative distribution functions) through mathematical models and thus only provides probability-boxes as results. Each method models (in)dependency between variables in a different way. Hence, for instance, the "*Hybrid*" approach, contrary to the "*Independent Random Sets*" approach, is not a counterpart to the calculus of probabilistic variables under stochastic independence. We notice that the "*Dependency Bounds Convolution*" approach does not give the same results as the "*Conservative Random Sets*" approach when there are more than two variables (whereas they are equivalent when we are faced with two variables as shown in [17]). According to mathematical models used to represent the knowledge of x_i, we will see that the "*Conservative Random Sets*" approach, contrary to the "*Dependency Bounds Convolution*" approach, provides the extent to which a criterion $T(X) \in A$ is satisfied where A is any measurable set.

In Section 2, we recall basic notions of imprecise probabilities and present three special cases that are easier to encode than general probability families: p-boxes, possibility distributions and finite random intervals (which induce belief functions). In Section 3, we present and compare four numerical methods for propagating objective (variability) and subjective (imprecision) information through multivariate function T so to estimate probability bounds for event of the form $P(T(X) \leq e)$ that refer to the likelihood of a threshold violation, commonly found in environmental studies. Lastly, in Section 4, we use these different methods to process uncertainty on a real case of soil contamination by lead on an ironworks brownfield.

2 Concise Representations of Imprecise Probability

Consider a probability space $(\Omega,\mathcal{A},P)$. Let $\mathcal{P}$ be a probability family on the referential Ω and X be a random variable associated with probability measure P. For all $A \subseteq \Omega$ measurable, we define its upper probability $\overline{P}(A) = \sup_{P\in\mathcal{P}} P(A)$ and its lower probability $\underline{P}(A) = \inf_{P\in\mathcal{P}} P(A)$. It is clear that representing and reasoning with a family of probabilities may be very complex. In the following we consider three frameworks for representing special sets of probability functions, which are more convenient for a practical handling.

2.1 Probability boxes

A natural model of an ill-known probability measure is obtained by considering a pair $(\underline{F},\overline{F})$ of non-intersecting cumulative distribution functions, generalising an interval. The interval $[\underline{F},\overline{F}]$ such that $\underline{F}(x) \leq F(x) \leq \overline{F}(x)\ \forall x \in \mathbb{R}$ is called a probability box (p-box) [18] [19]. In the probability box $[\underline{F},\overline{F}]$, the gap between $\underline{F}$ and $\overline{F}$ reflects the incomplete nature of the knowledge, thus picturing the extent of what is ignored. A p-box encodes the set of probability measures

$$\mathcal{P}(\underline{F} \leq \overline{F}) = \{P, \forall x, \underline{F}(x) \leq F(x) \leq \overline{F}(x)\}$$

whose cumulative distribution functions F are restricted by the bounding pair $(\underline{F}, \overline{F})$.

Suppose $(\underline{F},\overline{F})$ is induced from a probability family $\mathcal{P}$ where $\underline{F}(x) = \underline{P}((-\infty, x])$ and $\overline{F}(x) = \overline{P}((-\infty, x])\ \forall x \in \mathbb{R}$. Clearly, $\mathcal{P}(\underline{F} \leq \overline{F})$ strictly contains the set $\mathcal{P}$ it is built from. This is already true if the family $\mathcal{P}$ is represented by the upper and lower probabilities $\overline{P}(A)$ and $\underline{P}(A)$ for all measurable sets A. One may not be able to reconstruct $\mathcal{P}$ from these projections, only a superset of it can be recovered. A fortiori, the p-box is even a looser approximation of $\mathcal{P}$. As we shall see, the p-box representation method can be very imprecise, if $\underline{F}$ and $\overline{F}$ are not close to each other. Nevertheless, it is clearly a very convenient representation.

2.2 Possibility Theory

Possibility theory [12] is relevant to represent consonant imprecise knowledge. The basic tool is the possibility distribution, a mapping π from X to the unit interval such that $\max_{x\in X} \pi(x) = 1$. If X is the real line π is taken as the membership function of a fuzzy interval, that is π is upper semi-continuous, and its α-cuts $\{x, \pi(x) \geq \alpha\}$ are closed intervals [10]. A possibility distribution induces a pair of functions $[N, \Pi]$ such that

$$\Pi(A) = \sup_{x\in A} \pi(x)$$

and

$$N(A) = 1 - \Pi(\overline{A}).$$

The following noticeable properties of possibility and necessity measures are

$$\Pi(A \cup B) = \max(\Pi(A), \Pi(B));$$

and

$$N(A \cap B) = \min(N(A), N(B))$$

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 [8, 14]) hold. In other words, the family $\mathcal{P}(\pi)$ is entirely determined by the probability intervals it generates.

Possibility distributions can be obtained by extracting prediction intervals from probability measures ([11]), or by linear approximation between a core and a support provided by some expert. More generally the expert may supply pairs (measurable set A_i, necessity weight λ_i) interpreted as stating that the probability $P(A_i)$ is at least equal to λ_i where A_i is a measurable set (an interval containing the value of interest). We define the corresponding probability family as follows: $\mathcal{P} = \{P, \forall A_i, \lambda_i \leq P(A_i)\}$. If the sets A_i are nested ($A_1 \subset A_2 \subset \cdots \subset A_n$, as can be expected for a family of confidence intervals), then there is a possibility distribution (namely $\pi(x) = 1$ if $x \in A_1$ and $1 - \lambda_i$ if $x \in A_{i+1} \setminus A_i, \forall i > 1$) such that $\overline{P} = \Pi$ and $\underline{P} = N$ (see [14], and in the infinite case [8]).

It is tempting to define a particular p-box $[\underline{F},\overline{F}]$ from π such that $\overline{F}(x) = \Pi((-\infty, x])$ and $\underline{F}(x) = N((-\infty, x])$. But this p-box contains many more probability functions than $\mathcal{P}(\pi)$ (see [1])

2.3 Random sets

A discrete random set is defined by a mass distribution $\nu : \mathcal{F} \to [0, 1]$ assigning positive weights to measurable subsets of Ω in a finite family $\mathcal{F}$ It is such that

$\sum_{E\in\mathcal{F}} \nu(E) = 1$. A set $E \in \mathcal{F}$ is called a focal set. The theory of evidence introduced by Shafer [27] and elaborated further by Smets [28] introduces belief *Bel* and Plausibility *Pl* measures defined as $Bel(A) = \sum_{E,E\subseteq A} \nu(E)$ and $Pl(A) = \sum_{E,E\cap A\neq\emptyset} \nu(E) = 1 - Bel(\overline{A})$.
This framework allows imprecision and variability to be treated separately within a single framework. As in possibility theory, any pair of dual functions belief/plausibility $[Bel, Pl]$ can be interpreted as lower and upper probabilities. Indeed a mass distribution ν encodes the probability family $\mathcal{P}(\nu) = \{P, \forall A$ measurable, $Bel(A) \leq P(A)\} = \{P, \forall A$ measurable, $P(A) \leq Pl(A)\}$. In this case we have: $\overline{P} = Pl$ and $\underline{P} = Bel$.This view of belief functions is at odds with the theory of evidence of Shafer and and the transferable belief model of Smets (who never consider this probability family as part of their framework) but was originally proposed by Dempster [9]. He considered the image of a probability space via a set-valued mapping. In this view, $Bel(A)$ is the minimal amount of probability that must be assigned to A by sharing the probability weights defined by the mass function among single values in the focal sets. $Pl(A)$ is the maximal amount of probability that can be likewise assigned to A. The random set framework encompasses probability theory (when focal sets are singletons) and possibility theory (when focal sets are nested). In the former case $Bel = Pl = P$, and in the latter case $Bel = N$, $Pl = \Pi$.
We may define an upper $\overline{F}$ and a lower $\underline{F}$ cumulative distribution function (a particular p-box) such that

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

But this p-box contains again many more probability functions than $\mathcal{P}(\nu)$.

2.4 Discretized encoding of probability, possibility and p-box as random sets

Random sets as used by Shafer and Smets [27] encompass possibility and probability theories in the discrete case, not in the continuous case (even if continuous belief functions can be envisaged). Hence, we must build a discretized version of continuous probability distributions p, and possibility distributions π by means of some mass distribution ν. The discrete representation will be approximate but it allows for computations. Note that statistical data and poor probabilistic data is often obtained in a discrete, finite format. Continuous distributions are thus idealisations of the actual data. Hence one may argue that discrete representations are often closer to the way information is actually obtained. The step consisting of discretising continuous distributions may sometimes be bypassed, because the data can sometimes be directly modelled as a discrete random set or interval. This is clearly one of the ideas pervading Shafer's theory of evidence.

- Let X be a real random variable. In the discrete case, focal elements are singletons $(\{x_i\})_i$ and the mass distribution ν is just defined by $\nu(\{x_i\}) = P(X = x_i)$. In the continuous case, we can define focal intervals $(]x_i, x_{i+1}])_i$ by discretizing the support of a probability density into m intervals and a mass distribution ν is defined by $\nu(]x_i, x_{i+1}]) = P(X \in]x_i, x_{i+1}])$ $\forall\ i = 1 \ldots m$. This discretication achieves a bracketing approximation of the continuous probability measure, in the sense that $Bel(A) \leq P(A) \leq Pl(A), \forall A$ measurable, as noticed by Dubois and Prade [13].

- Let X be a possibilistic variable. Focal sets correspond to a selection of α-cuts

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

  of possibility distribution π associated with X such that $\alpha_1 = 1 > \alpha_j > \alpha_{j+1} > \alpha_q > 0$, which ensures $\pi_{\alpha_j} \subseteq \pi_{\alpha_{j+1}}$). Mass distribution ν is defined by $\nu(\pi_{\alpha_j}) = \alpha_j - \alpha_{j+1}\ \forall\ j = 1 \ldots q$ with $\alpha_{q+1} = 0$. Then the corresponding discrete possibility distribution is a lower approximation (more precise) of the continuous one. Alternatively, an upper approximation is obtained by letting $\nu(\pi_{\alpha_{j+1}}) = \alpha_j - \alpha_{j+1}$, by convention $\pi_{\alpha_{q+1}}$ being the support of π.

- Let X be an ill-defined random variable represented by a p-box $[\underline{F}_X, \overline{F}_X]$. There is no unique way of representing a p-box by a mass assignment. If $\underline{F}_X^{-1}(0) < \overline{F}_X^{-1}(1)$ where

$$\underline{F}_X^{-1}(p) = \min\{x|\underline{F}_X(x) \geq p\},\ \forall\ p \in [0, 1] \quad (2)$$

$$\overline{F}_X^{-1}(p) = \min\{x|\overline{F}_X(x) \geq p\},\ \forall\ p \in [0, 1] \quad (3)$$

  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 $1 \geq p_i > p_{i-1} > 0$. However, if $\underline{F}_X^{-1}(0) > \overline{F}_X^{-1}(1)$, we can choose focal sets of the form $[\overline{F}_X^{-1}(p_i), \underline{F}_X^{-1}(1 - p_i)]$ with the same last mass distribution. How to determine the least specific mass distribution ν associated with focal sets from a p-box (or most reasonable in some sense) is an open problem.

3 Approaches to the joint propagation of imprecision and variability

Let us assume $k < n$ random variables $(X_1, ..., X_k)$ taking values $(x_1, ..., x_k)$ and $n - k$ ill-known quantities $(X_{k+1}, ..., X_n)$ taking values $(x_{k+1}, ..., x_n)$ represented by possibility distributions $(\pi^{X_{k+1}}, ..., \pi^{X_n})$. This section presents four methods to propagate such heterogeneous uncertainties pervading the parameters $(X_i)_{i=1...n}$ through a multivariate function T. Typically, $(X_1, ..., X_k)$ are supposed to be variable quantities that can be properly observed via sufficiently rich statistical experiments. On the contrary, $(X_i)_{i=1...n}$ may be quantities on which no significant statistical data is available, but that can be informed via expert opinions in the form of "confidence" intervals. These quantities may be purely deterministic (not subject to variability) but in any case ill-known.

3.1 "*Hybrid*" possibility-probability approach

The "*Hybrid*" propagation method, a first version of which was proposed in [24] involves two main steps (see Figure 1). It combines a Monte-Carlo technique (Random Sampling) with the extension principle of fuzzy set theory [10] (interval analysis by α-cuts). We first perform a random sampling of the random variables ($X_1 = x_1, ..., X_k = x_k$) by taking into consideration known dependencies (as non linear monotone dependence [5] for instance) and fuzzy interval analysis is used to estimate T. Even if we can account for some dependencies between random variables with the Monte-Carlo method, it is necessary to be aware that it cannot account for all forms of dependence [20]. The knowledge on the value of $T(X)$ becomes a fuzzy subset, for each k-tuple. Random sampling is resumed and the process is performed in an iterative fashion in order to obtain a sample $(\pi_1^T, ..., \pi_m^T)$ of fuzzy subsets where m is the number of samples of the k random variables. $T(X)$ then becomes a fuzzy random variable (or a family of random possibility distributions) in the sense of [23].

We must emphasize the fact that the extension principle also underlies a dependence assumption on possibilistic variables. In fact the presence of imprecision on $X_{k+1}, ..., X_n$ potentially generates two levels of dependencies. The first one is a dependence between information sources attached to variables and the second one is a dependence between variables themselves. The extension principle [10]; $\forall u \in \mathbb{R}$:

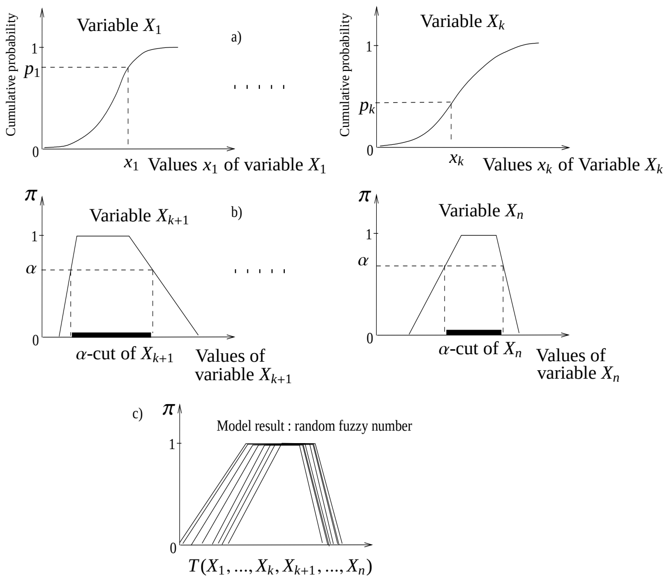

Figure 1: Outline of the "*Hybrid*" method [24]

$$\pi^T(u) = \sup_{x_{k+1},...,x_n, T(x_1,...,x_n)=u} \min(\pi^{X_{k+1}}(x_{k+1}), ..., \pi^{X_n}(x_n))$$

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 [13]. However, this form of dependence does not presuppose any genuine functional (objective) dependence between possibilistic variables inside the domain $\pi_\alpha^{X_{k+1}} \times \cdots \times \pi_\alpha^{X_n}$. The use of "minimum" assumes non-interaction between $X_{k+1}, ..., X_n$, 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. Indeed, the least specific joint possibility distribution whose projections on the $X_{k+1}, \ldots, X_n$ axes are $\pi^{X_{k+1}}, \ldots, \pi^{X_n}$, is precisely $\pi^{X_{k+1},...,X_n} = \min(\pi^{X_{k+1}}, \ldots, \pi^{X_n})$. As a consequence of the dependence on the choice of confidence levels, one cannot interpret the calculus of possibilistic variables as a conservative counterpart to the calculus of probabilistic variables under stochastic independence.

Now, from the sample $(\pi_i^T)_{i=1...m}$ of random fuzzy subsets $T(X)$, we encode each π_i^T as a belief function with focal sets corresponding to α-cuts $(\pi_{i\alpha}^T)_\alpha$ and the associated mass distribution is $(\nu_\alpha p_i)_\alpha$ (see Section 2.3). We obtain a weighted random sampling of intervals defining a belief function. Then, we can estimate, for all measurable sets A, $Pl^T(A)$ and $Bel^T(A)$ such that [3]:

$$Pl^T(A) = \sum_{(i,\alpha)\ \pi_{i\alpha} \cap A \neq \emptyset} \nu_\alpha p_i \qquad Bel^T(A) = \sum_{(i,\alpha)\ \pi_{i\alpha} \subseteq A} \nu_\alpha p_i$$

These evaluations are of the form:

$$Pl^T(A) = \sum_i p_i \Pi_i^T(A) \qquad Bel^T(A) = \sum_i p_i N_i^T(A)$$

where Π_i^T (resp. N_i^T) is the possibility measure (resp. necessity measure) associated with the possibility distribution π_i^T. This technique thus computes the eventwise weighted average of the possibility measures associated with each output fuzzy interval, and applies to any event.

A more refined representation of our knowledge of the probability of an event A induced by such a random fuzzy variable takes the form of a second order possibility distribution $\widetilde{\pi}(A)$ [6] on the unit interval. Indeed for each confidence level α, we obtain a random set $(\pi_{i\alpha}^T)_i$, we thus can compute the plausibility $Pl_\alpha(A)$ and the belief $Bel_\alpha(A)$ for each value α. Nested intervals of the form $([Pl_\alpha(A), Bel_\alpha(A)])_\alpha$ define the α-cuts of $\widetilde{\pi}(A)$ and it is easy to see that $Pl^T(A)$ and $Bel^T(A)$ are upper and lower bounds of the mean interval of $\widetilde{\pi}(A)$ [15].

$$Pl^T(A) = \int_0^1 Pl_\alpha(A) d\alpha \qquad Bel^T(A) = \int_0^1 Bel_\alpha(A) d\alpha$$

The proof of this result is given in [3].

3.2 The random set approach

In this Section, we exploit the fact that belief functions [27] encompass possibility and probability theory. We present two different ways to use belief functions to propagate heterogeneous information (variability + imprecision) in a homogeneous framework by assuming first independence among all focal sets, and then unknown dependence among focal sets.

3.2.1 "*Independent Random Sets*" approach

By using Section 2.4, consider $X_{k+1}, \ldots, X_n$, possibilistic variables encoded as belief functions by their focal sets $(\pi^{X_{k+1}}_{\alpha_{k+1}})_{\alpha_{k+1}}, \ldots, (\pi^{X_n}_{\alpha_n})_{\alpha_n}$ and the mass distributions $(\nu^{X_{k+1}}_{\alpha_{k+1}})_{\alpha_{k+1}}, \ldots, (\nu^{X_n}_{\alpha_n})_{\alpha_n}$. For the sake of clarity, we suppose random variables $X_1, \ldots, X_k$ are discrete. Let discrete probabilistic variables $X_1, \ldots, X_k$ be encoded by their focal singletons $(\{x_1^{\beta_1}\})_{\beta_1}, \ldots, (\{x_k^{\beta_k}\})_{\beta_k}$ and the mass distributions $(p^{X_1}_{\beta_1})_{\beta_1}, \ldots, (p^{X_k}_{\beta_k})_{\beta_k}$. With the "*Independent Random Sets*" approach, we define the mass distribution (denoted by $\nu^T_{\beta_1..\beta_k\alpha_{k+1}...\alpha_n}$), associated with focal sets $\pi^T_{\beta_1..\beta_k\alpha_{k+1}...\alpha_n} = T(x_1^{\beta_1}, \ldots, x_k^{\beta_k}, \pi^{X_{k+1}}_{\alpha_{k+1}}, \ldots, \pi^{X_n}_{\alpha_n})$ of $T(X)$, by $\forall\, \beta_1, \ldots, \beta_k, \alpha_{k+1}, \ldots, \alpha_n$:

$$\nu^T_{\beta_1..\beta_k\alpha_{k+1}...\alpha_n} = p^{X_1}_{\beta_1} \times \cdots \times p^{X_k}_{\beta_k} \times \nu^{X_{k+1}}_{\alpha_{k+1}} \times \cdots \times \nu^{X_n}_{\alpha_n}$$

Then, we can estimate, for all measurable sets A, $Pl^T(A)$ and $Bel^T(A)$ as follows:

$$Pl^T(A) = \sum_{(\beta_1,\ldots,\beta_k,\alpha_{k+1},\ldots,\alpha_n);\pi^T_{\beta_1..\beta_k\alpha_{k+1}...\alpha_n} \cap A \neq \emptyset} \nu^T_{\beta_1..\beta_k\alpha_{k+1}...\alpha_n}$$

$$Bel^T(A) = \sum_{(\beta_1,\ldots,\beta_k,\alpha_{k+1},\ldots,\alpha_n);\pi^T_{\beta_1..\beta_k\alpha_{k+1}...\alpha_n} \subseteq A \neq \emptyset} \nu^T_{\beta_1..\beta_k\alpha_{k+1}...\alpha_n}$$

It corresponds to applying a Monte-Carlo method to all variables. For each possibility distribution, an α-cut is independently selected. The "*Independent Random Sets*" approach is a conservative counterpart to the calculus of probabilistic variables under stochastic independence [7, 22].

3.2.2 Casting the "hybrid" approach in the random set setting

Suppose now the same value of α is selected in the Monte-Carlo simulation, for all possibilistic variables. Then: $\forall\, \beta_1, \ldots, \beta_k, \alpha_{k+1}, \ldots, \alpha_n$

$$\begin{array}{ll} \alpha_{k+1} = \cdots = \alpha_n & \nu^T_{\beta_1..\beta_k\alpha_{k+1}...\alpha_n} = \nu^{X_{k+1},\ldots,X_n}_{\alpha} \times p^{X_1}_{\beta_1} \times \cdots \times p^{X_k}_{\beta_k} \\ \alpha_{k+1} \neq \cdots \neq \alpha_n & \nu^T_{\beta_1..\beta_k\alpha_{k+1}...\alpha_n} = 0 \end{array}$$

The possibility distribution $\pi^{X_{k+1},\ldots,X_n}$ is characterized by $min(\pi^{X_{k+1}}, \ldots, \pi^{X_n})$ which corresponds to nested Cartesian products of α-cuts and $\nu^{X_{k+1},\ldots,X_n}_{\alpha_{k+1}}$ is the mass associated with the Cartesian product $\pi^{X_{k+1}}_{\alpha_{k+1}} \times \cdots \times \pi^{X_n}_{\alpha_n}$. Like in 3.1 we thus assume total dependence between focal sets associated with possibilistic variables. Hence, if we want to estimate $Pl^T(A)$, for all measurable sets A, using the last definition of $\nu^T_{\beta_1..\beta_k\alpha_{k+1}...\alpha_n}$, we deduce that:

$$Pl^T(A) = \sum_{\beta_1,\ldots,\beta_k} p^{X_1}_{\beta_1} \times \cdots \times p^{X_k}_{\beta_k} \times \Pi^T_{\beta_1..\beta_k}(A)$$

where $\Pi^T_{\beta_1..\beta_k}$ are the possibility measures associated with the joint non-interactive possibility distribution $\pi^T_{\beta_1..\beta_k}$ obtained by the "*Hybrid*" method (see Section 3.1). That means the "*Hybrid*" approach and "*Independent Random Sets*" approaches are equivalent when we combine only one possibilistic variable with other probabilistic variables .

3.2.3 "*Conservative Random Sets*" approach

Borrowing from [4], the idea of the "*Conservative Random Sets*" approach [2] is to compute extreme upper plausibility Pl^T_{max} and lower belief functions Bel^T_{min} without assuming any knowledge about dependencies. It yields a linear optimization problem whose unknown is the joint mass function. It yields the loosest possible bracketing $[Bel^T_{min}(A), Pl^T_{max}(A)]$ of $P(T(X) \in A)$, which can be attained using some feasible joint mass distribution, for any measurable set A of interest.
For the sake of clarity, consider an example involving three parameters x, y, z which can be represented by probability distribution, p-box or possibility distribution and the function $T : (x,y,z) \mapsto T(x,y,z)$. Let $(\nu^x_i)_i$, $(\nu^y_j)_j$ and $(\nu^z_k)_k$ be the mass distributions associated with focal sets $([\underline{x}_i, \overline{x}_i])_i$, $([\underline{y}_j, \overline{y}_j])_j$ and $([\underline{z}_k, \overline{z}_k])_k$. In the "*Conservative Random Sets*" approach, contrary to the "*Independent Random Sets*" approach where $h_{ijk} = \nu^x_i \times \nu^y_j \times \nu^z_k$, we must find the mass distribution $(h_{ijk})_{ijk}$ such that $Pl^T(A)$ is maximal and $Bel^T(A)$ is minimal. That means we obtain $Pl^T_{max}(A)$ by solving the following maximization problem:

$$\max \sum_{T([\underline{x}_i,\overline{x}_i],[\underline{y}_j,\overline{y}_j],[\underline{z}_k,\overline{z}_k]) \cap A \neq \emptyset} h_{ijk}$$

$$\begin{array}{c} \sum_{jk} h_{ijk} = \nu^x_i \quad \forall i \\ \sum_{ik} h_{ijk} = \nu^y_j \quad \forall j \\ \sum_{ij} h_{ijk} = \nu^z_k \quad \forall k \\ \sum_{i,j,k} h_{ijk} = 1 \end{array}$$

Similary, we obtain $Bel^T_{min}(A)$ by minimizing

$$\sum_{T([\underline{x}_i,\overline{x}_i],[\underline{y}_j,\overline{y}_j],[\underline{z}_k,\overline{z}_k]) \subseteq A} h_{ijk}$$

under the same constraints. The "*Conservative Random Sets*" approach is a rigorous method to obtain a conservative, but attainable bracketing of the ill-known probability $P(A)$ for all measurable sets A when nothing about dependencies between variables is assumed.

3.3 "*Dependency Bounds Convolution*" approach

This section describes a method that can be used to compute extreme upper and lower cumulative distribution functions on results of probabilistic model no matter what correlations or statistical dependencies exist among the variables. Williamson and Downs [29] gave a numerical method for computing these bounds by using p-boxes $[\underline{F}_X, \overline{F}_X]$, $[\underline{F}_Y, \overline{F}_Y]$ representing two ill-known random variables X and Y, without using any information about their joint distribution for arithmetic operations $\{+, -, \times, \div\}$. The idea is to use the Fréchet bounds :

$$max(F_X(x)+F_Y(y)-1, 0) \leq F_{(X,Y)}(x, y) \leq min(F_X(x), F_Y(y)) \quad (4)$$

based on the theory of copulas [26]. An important result due to Sklar proves the existence, for any joint probability distribution $F_{(X,Y)}$, of a function C, called “Copula”, from the unit square to the unit interval, such that $F_{(X,Y)}$ is completely determined by its marginals F_X, F_Y via C, that is, $F_{(X,Y)} = C(F(x), F(y)) = P(X \leq x, Y \leq y)$. It means that a copula C contains all information related to dependence among random variables X and Y. Formally, a two-dimensional copula C is a mapping $C : [0, 1] \times [0, 1] \to [0, 1]$ such that

1. $C(0, u) = C(u, 0) = 0 \; \forall \; u \in [0, 1]$
2. $C(u, 1) = C(1, u) = u \; \forall \; u \in [0, 1]$
3. $C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0$ $\forall \; (u_1, u_2, v_1, v_2) \in [0, 1]^4$ such that $u_1 \leq u_2$ and $v_1 \leq v_2$

All copulas verify, $\forall \; (u_1, u_2) \in [0, 1]^2$

$$max(u_1 + u_2 - 1, 0) \leq C(u_1, u_2) \leq min(u_1, u_2) \quad (5)$$

corresponding to Fréchet bounds (4). Copula $C : (F_X(x), F_Y(y)) \mapsto \min(F_X(x), F_Y(y))$ means Y is almost surely an increasing function of X. Copula $C : (F_X(x), F_Y(y)) \mapsto \max(F_X(x) + F_Y(y) - 1, 0)$ means Y is almost surely a decreasing function of X.
Hence, from Fréchet bounds, extreme upper and lower cumulative distribution functions of the addition of two ill-defined random variables, for instance, are given by [29]:

$$\underline{F}_{X+Y}(z) = \sup_{x+y=z} \{\max(\underline{F}_X(x) + \underline{F}_Y(y) - 1, 0)\} \quad (6)$$

$$\overline{F}_{X+Y}(z) = \inf_{x+y=z} \{\min(\overline{F}_X(x) + \overline{F}_Y(y), 1)\}. \quad (7)$$

These results come from the following inequalities [29]

$$\sup_{L(x,y)=z} \max(F_X(x) + F_Y(y) - 1, 0) \leq F_{L(X,Y)}(z)$$

$$F_{L(X,Y)}(z) \leq \inf_{L(x,y)=z} \min(F_X(x) + F_Y(y), 1)$$

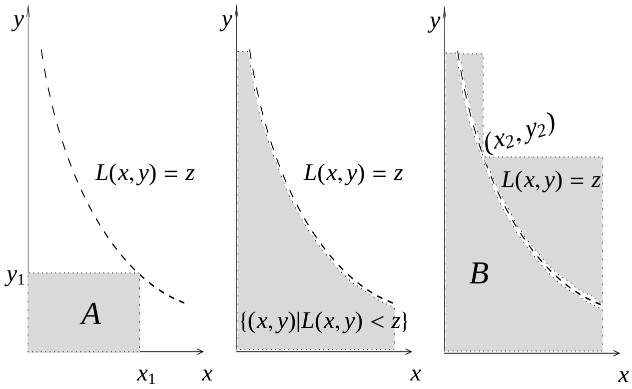

Figure 2: The joint probability bounds of cumulative distribution function $F_{L(x,y)}$

where $L \in \{+, -, \div, \times\}$. Indeed (see Figure 2), $\forall \; (x_1, y_1)$, $(x_2, y_2) \in \{(x, y)|L(x, y) = z\}$

$$\begin{aligned} &\max(F_X(x_1) + F_Y(y_1) - 1, 0) \\ &\quad \leq C(F_X(x_1), F_Y(y_1)) \\ &\quad = \int\int_A dC(F_X(x), F_Y(y)) \\ &\quad \leq \int\int_{\{(x,y)|L(x,y)<z\}} dC(F_X(x), F_Y(y)) \\ &\quad \leq \int\int_B dC(F_X(x), F_Y(y)) \\ &\quad = F_X(x_2) + F_Y(y_2) - C(F_X(x_2), F_Y(y_2)) \\ &\quad \leq \min(F_X(x_2) + F_Y(y_2), 1) \end{aligned}$$

Hence, $\underline{F}_{L(X,Y)}(z)$ is the greatest value of $\max(F_X(x) + F_Y(y) - 1, 0)$ where $(x_1, y_1) \in \{(x, y)|L(x, y) = z\}$ and $\overline{F}_{L(X,Y)}(z)$ is the smallest value of $\min(F_X(x_2) + F_Y(y_2), 1)$ where $(x_2, y_2) \in \{(x, y)|L(x, y) = z\}$. Then, it is easy to deduce $\underline{F}_{X+Y}$ and $\overline{F}_{X+Y}$ (see equations 6 and 7) from $\underline{F}_X$, $\underline{F}_Y$, $\overline{F}_X$ and $\overline{F}_Y$.
Ferson [21] extends these results to other operators like min, max, log, exp and power. As in fuzzy arithmetic (interval analysis by α-cuts), the quasi-inverses of the resulting lower and upper distribution function bounds can be calculated in terms of quasi-inverses of the upper and lower cumulative distribution functions. They employ lower and upper discrete approximations to the quantile function. That is, they discretize upper and lower cumulative distributions $\underline{F}_X$, $\underline{F}_Y$, $\overline{F}_X$ and $\overline{F}_Y$ into $m + 1$ elements and they obtain bounds on the quantile functions $\underline{F}^{-1}_{X+Y}$ and $\overline{F}^{-1}_{X+Y}$ as follows (see Figure 3):

- $\underline{F}^{-1}_{X+Y}(\frac{i}{m}) = \min_{j=i...m}\{\underline{F}^{-1}_X(\frac{j}{m}) + \underline{F}^{-1}_Y(\frac{i-j+m}{m})\}$
- $\overline{F}^{-1}_{X+Y}(\frac{i}{m}) = \max_{j=0...i}\{\overline{F}^{-1}_X(\frac{j}{m}) + \overline{F}^{-1}_Y(\frac{i-j}{m})\}$

where i varies between 0 and m. The quantile bounds are then inverted to obtain bounds on the cumulative distribution functions.
The first obvious disadvantage of this method, although it does not assume anything about dependencies, is that

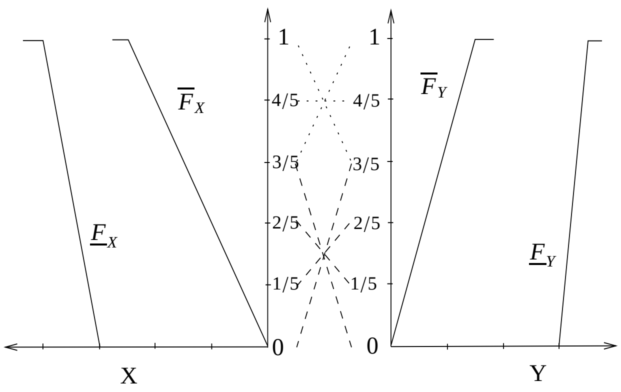

Figure 3: Practical computation to obtain $[\underline{F}_{X+Y}(3/5), \overline{F}_{X+Y}(3/5)]$.

it does not allow to estimate probability for all measurable sets A (contrary to the "*Conservative Random Sets*" approach) but only measurable sets as $(-\infty, e]$ or $[e, \infty)$. Moreover, contrary to the other propagation methods, this process cannot be applied to any multivariate functions like semi-analytical models for instance. The next subsection shows another problem with this approach.

3.4 "*Conservative Random Sets*" versus "*Dependency Bounds Convolution*"

Ferson et al. [17] show that the "*Conservative Random Sets*" approach and the "*Dependency Bounds Convolution*" method to propagate uncertainty converge to the same result when the domain is restricted to the positive real line, and the output is expressed in the cumulative distributional form. But, for more than two variables, we show with an example that the equivalence is no longer true. When one is faced with more than two variables, the "*Dependency Bounds Convolution* " method combines the first two variables. and then combines the previous results with other variables and so on. This process looks sound because the weakest and the strongest copulas are associative (they are triangular norms), but it is is questionable because the weakest copula extended to 3 arguments is no longer a copula [26] even if it still provides a bracketing of the ill-known joint probability. As a consequence, this method models impossible dependencies among random variables. Indeed, consider three ill-defined random variables X, Y and Z; such a combination of copulas then provides the alleged joint probability distribution

$$F_{(X,Y,Z)}(x, y, z) = \max(F_X(x) + F_Y(y) + F_Z(z) - 2, 0).$$

To see that the joint probability $F_{(X,Y,Z)}$ takes into account impossible dependence structures first obtain by projection

$$F_{(X,Y)}(x, y) = \lim_{z \to +\infty} \max(F_X(x) + F_Y(y) + F_Z(z) - 2, 0)$$

That is

$$F_{(X,Y)}(x, y) = \max(F_X(x) + F_Y(y) - 1, 0) \quad (8)$$

Similarly we get:

$$F_{(Y,Z)}(y, z) = \max(F_Y(y) + F_Z(z) - 1, 0) \quad (9)$$

$$F_{(Z,X)}(z, x) = \max(F_X(x) + F_Z(z) - 1, 0) \quad (10)$$

which means Y is almost surely a decreasing function of X, Z is almost surely a decreasing function of Y and Z is almost surely a decreasing function of X. However, conditions (8) (9) imply that Z must be almost surely an increasing function of X, which is contradictory with the last condition (10).

Example. Let X (resp. Y and Z) be an ill-known random variable represented by a belief function such that $\nu_X([3, 4]) = 0.5$, $\nu_X([2, 5]) = 0.5$ (resp. $\nu_Y([3, 5]) = 0.5$, $\nu_Y([2, 6]) = 0.5$ and $\nu_Z([4, 5]) = 0.5$, $\nu_Z([3, 6]) = 0.5$) and $T(x, y, z) = (x + y) \times z$. We try to estimate a bracketing of $P(T(X, Y, Z) \leq e)$ using the "*Dependency Bounds Convolution*" method and the "*Conservative Random Sets*" approach. Figure 4 represents cumulative distributions deduced from the mass distributions of variables X, Y and Z. Figure 5 represents the cumulative distribution of T by applying either the "*Dependency Bounds Convolution*" method or the "*Conservative Random Sets*" approach. We

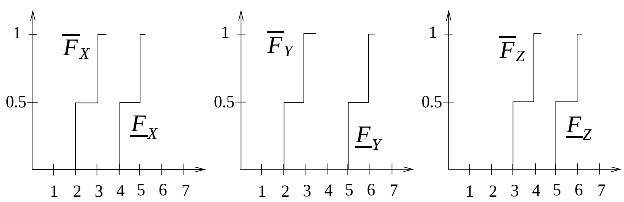

Figure 4: p-box of X, Y and Z

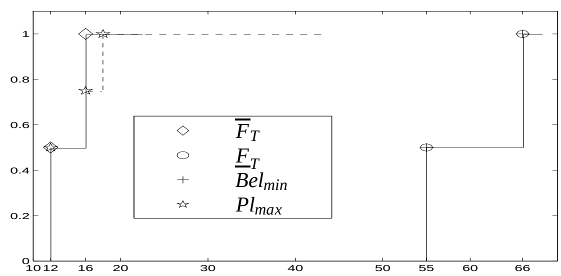

Figure 5: Upper and Lower cumulative dustributions of T

can observe that $\overline{F}_T(16) = 1$ whereas $Pl_{max}((-\infty, 16]) = 0.75$. That means there does not exist any mass distribution ν_T such that $Pl((-\infty, 16]) = 1$. Indeed, suppose that such a mass distribution exists. There are two focal elements for each variable; that means that there are eight focal elements for T:

X	Y	Z	$(X+Y)\times Z$	ν_T
$[3,4]$	$[3,5]$	$[4,5]$	$[45,100]$	ν_{111}
$[3,4]$	$[3,5]$	$[3,6]$	$[18,54]$	ν_{112}
$[3,4]$	$[2,6]$	$[4,5]$	$[20,50]$	ν_{121}
$[3,4]$	$[2,6]$	$[3,6]$	$[15,60]$	ν_{122}
$[2,5]$	$[3,5]$	$[4,5]$	$[20,50]$	ν_{211}
$[2,5]$	$[3,5]$	$[3,6]$	$[15,60]$	ν_{212}
$[2,5]$	$[2,6]$	$[4,5]$	$[16,55]$	ν_{221}
$[2,5]$	$[2,6]$	$[3,6]$	$[12,66]$	ν_{222}

We compute $Pl_{max}((-\infty,16])$ by solving the following problem:

$$\begin{array}{ll}
\max \nu_{122}+\nu_{212}+\nu_{221}+\nu_{222} & \\
\nu_{111}+\nu_{112}+\nu_{121}+\nu_{122}=0.5 & C(1)\\
\nu_{211}+\nu_{212}+\nu_{221}+\nu_{222}=0.5 & C(2)\\
\nu_{111}+\nu_{112}+\nu_{211}+\nu_{212}=0.5 & C(3)\\
\nu_{121}+\nu_{122}+\nu_{221}+\nu_{222}=0.5 & C(4)\\
\nu_{111}+\nu_{121}+\nu_{211}+\nu_{221}=0.5 & C(5)\\
\nu_{112}+\nu_{122}+\nu_{212}+\nu_{222}=0.5 & C(6)\\
\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2}\nu_{ijk}=1 & C(7)
\end{array}$$

We supposed that $Pl_{max}((-\infty,16])=1$. It means that $\nu_{122}+\nu_{212}+\nu_{221}+\nu_{222}=1$. Hence, the last constraint C(7) implies:

$$\nu_{111}=0;\nu_{112}=0;\nu_{121}=0;\nu_{211}=0.$$

The constraints C(1) and C(5) thus imply: $\nu_{122}=0.5$ and $\nu_{221}=0.5$. Hence, $\nu_{111}+\nu_{221}=1$, while according to C(4), it should be equal to 0.5. So, the probability family $\mathcal{P}(\nu_T)$ obtained from the belief function approach with optimization is more precise than $\mathcal{P}(\underline{F}_T<\overline{F}_T)$ obtained by the "*Dependency Bounds Convolution*" method and p-boxes, and the latter is clearly over pessimistic.

4 Application to soil contamination

In this Section, we apply the previous uncertainty propagation methods to a real case of soil contamination by lead on an ironworks brownfield in the south of France. Following an on-site investigation revealing the presence of lead in the superficial soil at levels on the order of tens of grams per *kg* of dry soil, a cleanup objective of 300 mg/kg was established by a consulting company, based on a potential risk assessment, taking into account the most significant exposure pathway and the most sensitive target (direct soil ingestion by children). The mathematical model calculating the quantity D_{lead} absorbed by a child living on the site and exposed via soil ingestion is given by [16]:

$$D_{lead}=\frac{C_{soil}\times IR_{soil}\times(Fi_{inside}+Fi_{outside})\times Ef\times ED}{Bw\times AT\times 10^6}$$

where

D_{lead}	=	Absorbed lead dose related to the ingestion of soil ($mg/[kg.day]$)
C_{soil}	=	Lead concentration in Soil (mg/kg)
IR_{soil}	=	Ingestion Rate mg soil/day
Fi_{indoor}	=	Indoor Fraction of contaminated soil ingestion (unitless)
$Fi_{outdoor}$	=	Outdoor Fraction of contaminated soil ingestion (unitless)
Ef	=	Exposure Frequency ($days/year$)
ED	=	Exposure Duration ($years$)
Bw	=	Body Weight (kg)
AT	=	Averaging time (period over which exposure is averaged–days)

The World Health Organization prescribed the acceptable lead dose related to the ingestion of polluted soil to be equal to 3.5 $\mu g/[kg.day]$. That means that after the cleanup objective of 300 mg/kg on the site (ironworks brownfield), calculated doses D_{lead} should not be larger than 3.5 $\mu g/[kg.day]$. The situation is considered acceptable in terms of public health if calculated doses are not superior to 3.5 $\mu g/[kg.day]$.

Typically, the model parameters are tainted by objective and subjective uncertainty. The cleanup objective of C_{soil}=300 mg/kg of the consulting company cannot be achieved uniformly over the entire polluted site. So, this objective can be viewed as the mode of a distribution on C_{soil} after cleanup. Moreover, C_{soil} values greater than 500 mg/kg are not tolerable, while values lower than 40 mg/kg are unrealistic. To summarize, for lead concentration in soil (C_{soil}) after cleanup, we consider a modal value of 300 mg/kg and a support [40,500] mg/kg. A rigorous way to represent this knowledge is to use the triangular possibilistic distribution $\pi_{C_{sol}}$ which encodes a more precise probability family (see [1]) than the one defined by the p-box of Ferson et al. [18] [19] such that:

$$\underline{F}_{Csoil}(x)=\frac{x-300}{200}\quad \text{for } x\in[300,500] \text{ and 0 otherwise}$$

$$\overline{F}_{Csoil}(x)=\frac{x-40}{260}\quad \text{for } x\in[40,300] \text{ and 1 otherwise}$$

Concerning the ingestion rate IR_{soil} (resp. the indoor fraction ingested Fi_{indoor}), experts say that it is sure that the values of IR_{soil} are within $[20,300]$ (resp. of Fi_{indoor} are within $[0.2,0.9]$) but they say that the most likely values of IR_{soil} are within $[50,200]$ (resp. $[0.5,0.7]$ for Fi_{indoor}). This knowledge is typically represented by the possibility distribution $\pi_{IR_{soil}}$, $\pi_{Fi_{indoor}}$ defined by support($\pi_{IR_{soil}}$)=$[20,300]$, core($\pi_{IR_{soil}}$)=$[50,200]=\{x|\pi_{IR_{soil}}(x)=1\}$ and support($\pi_{Fi_{indoor}}$)=$[0.2,0.9]$, core($\pi_{Fi_{indoor}}$)=$[0.5,0.7]$.

For the body weight Bw, we have sufficient knowledge to represent it by the normal distribution with mean (resp. standard deviation) equal to 17.4 kg (resp. equal to 2.57

kg).
The child's time budget (related to exposure frequency) is divided into two components: 2 hours per day outdoors and 16 hours per day indoor (experts opinion).
The outdoor contaminated soil ingestion fraction $Fi_{outdoor}$ is taken as unity. Experts consider the exposure duration ED equal to 6 years and the exposure is averaged on 6 years so $AT = 2190$ days. The period over which exposure is averaged is taken as the exposure duration, as carcinogenic effects for lead are not proven.
From this imprecise and random data, we propagate information through the model (absorbed lead dose D_{lead} related to the ingestion of soil) by means of the four previous methods of propagation to estimate $P(D_{lead} \leq threshold)$. Figure 6 represents the upper

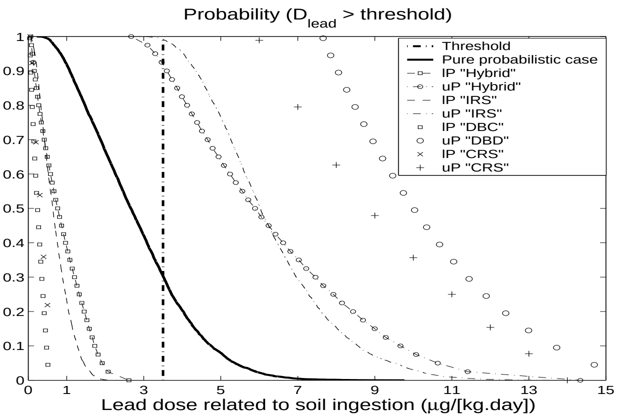

Figure 6: Upper and Lower cumulative distributions of D_{lead} according to the four methods of propagation. "lP"=Lower Probability, "uP"=Upper Probability, "DBC"="*Dependency Bounds Convolution*", "IRS"="*Independent Random Sets*", "CRS"= "*Conservative Random Sets*"

and lower cumulative distribution functions of D_{lead} resulting from these four methods. It enables the four methods to be compared with the purely probabilistic approach where one assumes all knowledge to be of a random nature and stochastic independence between variables (classical procedure in risk assessment). Hence we can observe that the probability $P(D_{lead} > 3.5)$ is equal to 30% in the pure probabilistic case but the other methods show that this probability can be much higher, leading us to reject the hypothesis of non violation of the tolerable threshold. The distance between upper and lower probabilities characterizes the imprecision of result D_{lead} brought by the imprecise knowledge of parameters and by the modeling of (in)dependence enforced by each propagation method. We also can see that the "*Dependency Bounds Convolution*" approach provides the most imprecise results, which are actually too pessimistic from a mathematical point of view since this method considers impossible dependencies among ill-defined random variables represented by p-boxes. These results show that it would be preferable to decrease the cleanup objective of 300 mg/kg in order to reduce the upper probability of exceeding the tolerable dose of 3.5 mg/kg.

5 Conclusion

This paper compares three techniques for the representation of imprecise subjective and objective probabilistic knowledge, and four practical propagation methods through multivariate functions. The p-box representation reflects only a small part of the information contained in a probability set. Using p-boxes can become very imprecise if the distance between the lower and upper distribution is large. A possibility distribution or a belief function is then more precise.
All propagation methods carry their own assumptions regarding (in)dependence between variables. The "*Hybrid*" approach, contrary to the "*Independent Random Sets*" approach, is not a conservative counterpart of the calculus of independent probabilistic variables. Indeed the extension principle applied to possibility distributions involves a dependence assumption between observers, but no dependence is assumed between observed variables.
The "*Dependency Bounds Convolution*" can only work with p-boxes and may produce erroneously overpessimistic bounds, due to a mathematical difficulty with copulas of order higher than two. On the contrary, the "*Conservative Random Sets*" approach can compute tight pessimistic bounds for any kind of measurable event without making any assumption about dependence, at the expense of solving one optimisation problem for each event and each probability bound.

Acknowledgements

We thank BRGM, INERIS and ADEME for providing the information relative to the case of lead contamination.

References

[1] Baudrit, C., Dubois, D., Fargier, H. Practical representation of incomplete probabilistic information. In M. Lopez-Diaz et al.: eds *Soft Methodology and Random Information Systems* , Advances in Soft Computing Series, Springer, 149-156, 2004.

[2] Baudrit, C., Dubois, D., Fargier, H. Propagation of Uncertainty involving Imprecision and Randomness. *Proc.*3^{d} *European Conference on Fuzzy Logic and Technology, Zittau (Germany)*, 653-658, 2003.

[3] Baudrit, C., Couso, I., Dubois D., Joint propagation of probability and possibility in risk analysis: towards a formal framework, submitted.

[4] Berleant D. , Goodman-Strauss, C. Bounding results of arithmetic operations on random variables of unknown dependencies using interval arithmetic. *Reliable Computing*, 4, 147-165, 1998.

[5] Connover, W.J., Iman, R.L. A distribution-free approach to inducing rank correlation among input variables. *Technometric*, 3, 311-334, 1982.

[6] Couso, I., Montes, S., Gil, P. Second order possibility measure induced by a fuzzy random variable. In C. Bertoluzza et al., eds.: *Statistical Modeling, Analysis and Management of Fuzzy Data*, Springer, Heidelberg, 2002.

[7] Couso, I., Moral, S., Walley, P. A survey of concepts of independence for imprecise probabilities. *Risk Decision and Policy*, 5, 165-180, 2000.

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

[9] Dempster, A.P. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38, 325-339, 1967.

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

[11] Dubois D., Foulloy L., Mauris G., Prade H., Probability-possibility transformations, triangular fuzzy sets, and probabilistic inequalities. *Reliable Computing* 2004. 10, 273-297

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

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

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

[15] Dubois, D., Prade, H. The mean value of a fuzzy number. *Fuzzy Sets and Systems*, 24(3), 279-300, 1987.

[16] EPA, U.S. (1989) - Risk Assessment Guidance for Superfund. *Volume I: Human Health Evaluation Manual (Part A). (EPA/540/1-89/002). Office of Emergency and Remedial Response. U.S EPA., Washington D.C.*

[17] Ferson, S., Berleant, D., Regan, H.M. Equivalence of methods for uncertainty propagation of real-valued random variables. *International Journal of Approximate Reasoning*, 36, 1-30, 2004.

[18] Ferson, S., Ginzburg, L., Kreinovich, V., Myers, D.M., Sentz, K. Construction of Probability Boxes and Dempster-Shafer Structures. *Sandia National Laboratories, Technical report SANDD2002-4015*, 2003.

[19] Ferson, S., Ginzburg, L., Akcakaya, R. Whereof one cannot speak: when input distributions are unknown. To appear in *Risk Analysis*.

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

[21] Ferson, S. RAMAS Risk Calc 4.0 Software: *Risk Assessment with Uncertain Numbers*. Lewis Publishers, Boca Raton, Florida.

[22] Fetz, Th., Oberguggenberger, M. Propagation of uncertainty through multivariate functions in the framework of sets of probability measures. *Reliability Engineering and System Safety*, 85, 73-87, 2004.

[23] Gil, M.A. (Ed.) Fuzzy Random Variables. Special issue of *Information Sciences*, 133, Nos. 1-2, 2001.

[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] Helton J.C. Oberkampf W.L., Eds. (2004) Alternative Representations of Uncertainty, Reliability Engineering and Systems Safety, vol. 85, Elsevier, 369 p.

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

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

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

[29] Williamson, R.C., Downs, T. Probabilistic arithmetic I: Numerical methods for calculating convolutions and dependency bounds. *International Journal of Approximate Reasoning*, 4, 89-158, 1990.