Monte Carlo and fuzzy interval propagation of hybrid uncertainties on a risk model for the design of a flood protection dike

Piero Baraldi, Nicola Pedroni, Enrico Zio, Elisa Ferrario, Alberto Pasanisi, Mathieu Couplet

Monte Carlo and fuzzy interval propagation of hybrid uncertainties on a risk model for the design of a flood protection dike

P. Baraldi, N. Pedroni, E. Zio & E. Ferrario
Politecnico di Milano, Milano, Italy

A. Pasanisi, M. Couplet
Electricité de France, Chatou, France

ABSTRACT: A risk model may contain uncertainties that may be best represented by probability distributions and others by possibility distributions. In this paper, a computational framework that jointly propagates probabilistic and possibilistic uncertainties is compared with a pure probabilistic uncertainty propagation. The comparison is carried out with reference to a risk model concerning the design of a flood protection dike.

1 INTRODUCTION

Uncertainty analysis is a fundamental part of the risk analysis of complex systems such as nuclear power plants, aerospace systems and others.

In risk analysis, uncertainty is conveniently considered of two different types: randomness due to inherent variability in the system behavior and imprecision due to lack of knowledge and information on the system. The former type of uncertainty is often referred to as objective, aleatory, stochastic whereas the latter is often referred to as subjective, epistemic, state-of-knowledge (Helton 2004).

In the context of risk analysis, the aleatory uncertainty is related to the occurrence of the events which define the various possible accident scenarios, whereas epistemic uncertainty arises from a lack of knowledge of fixed but poorly known parameter values entering the evaluation of the probabilities and consequences of the accident scenarios.

In the current risk assessment practice, both types of uncertainties are represented by means of probability distributions. However, resorting to a single probabilistic representation of epistemic uncertainty may not be possible when sufficient data is not available for statistical analysis, even if one adopts expert elicitation procedures to incorporate diffuse information into the corresponding probability distributions, within a subjective view of probability. Indeed, an expert may not have sufficiently refined knowledge or opinion to characterize the relevant epistemic uncertainty in terms of probability distributions (Helton 2004).

As a result of the potential limitations associated to a probabilistic representation of epistemic uncertainty under limited information, a number of alternative representation frameworks have been proposed, e.g., fuzzy set theory, evidence theory, possibility theory and interval analysis (Klir & Yuan 1995; Aven & Zio 2011).

Possibility theory is attractive for risk assessment, because of its representation power and its relative mathematical simplicity. It offers two measures of likelihood, namely possibility and necessity measures, that may be interpreted as lower and upper probabilities in the representation of imprecision in the experts' probability assignments.

The possibilistic representation of uncertainty can both be combined with and transformed into the traditional probabilistic representation. In this respect, an integrated ("hybrid") computational framework has been proposed for jointly propagating probabilistic and possibilistic representations through a model (Baudrit et al. 2006). This framework has been applied to propagate uncertainties in event trees (Baraldi & Zio, 2008) and fault trees (Flage et al. 2010; Flage et al. 2011).

In the present paper, the integrated framework of propagation is tested on a flood risk model (Pasanisi et al. 2009; Limbourg & de Rocquigny, 2010) considered a realistic benchmark for uncertainty modeling.

The reminder of the paper is organized as follows. In Section 2, some basic concepts about possibility theory are summarized; in Section 3, the details about the integrated propagation framework are given; in Section 4, approaches for constructing possibility distributions are discussed; in Section 5, the flood model considered for the uncertainty propagation task is presented; in Section 6, the results of the joint propagation of aleatory and epistemic uncertainties through the flood model are reported and commented; finally, in Section 7, conclusions and direction for future work are provided.

2 BASICS OF POSSIBILITY THEORY

In possibility theory, uncertainty is represented by a possibility function $\pi(y)$. For each y in a set Ω, $\pi(y)$ expresses the degree of possibility of y. When $\pi(y)=0$ for some y, it means that the outcome y is considered an impossible situation. When $\pi(y)=1$ for some y, it means that the outcome y is possible, i.e., is just unsurprising, normal, usual (Dubois 2006). This is a much weaker statement than when probability is 1.

The possibility function gives rise to probability bounds, upper and lower probabilities, referred to as necessity and possibility measures (N, Π). They are defined as follows.

The possibility of an event A, $\Pi(A)$, is defined by

$$\Pi(A)=\sup_{y\in A}\{\pi(y)\}, \tag{1}$$

and the necessity measure $N(A)$ is defined by

$$N(A)=1-\Pi(notA)=1-\sup_{y\notin A}\{\pi(y)\}.$$

Let $\mathcal{P}(\pi)$ be a family of probability distributions such that for all events A, $N(A)\le P(A)\le \Pi(A)$. Then,

$$N(A)=\inf P(A) \text{ and } \Pi(A)=\sup P(A) \tag{2}$$

where *inf* and *sup* are with respect to all probability measures in $\mathcal{P}$. Hence the necessity measure is interpreted as a lower level for the probability and the possibility measure is interpreted as an upper limit. Referring to subjective probabilities, the bounds reflect that the analyst is not able or willing to precisely assign his/her probability, and the bounds are the best he/she can do given the information available; in other words, he or she can only describe a subset of $\mathcal{P}$ which contains his/her probability (Dubois 2006).

3 PROPAGATION OF EPISTEMIC AND ALEATORY UNCERTAINTIES THROUGH A MODEL

Let us consider a model whose output is a function $Z=f(Y_1,Y_2,...,Y_n)$ of n uncertain variables Y_j, $j=1,\cdots,n$, ordered in such a way that the first k are random (aleatory uncertainty), with probability distributions $p_{Y_i}(y)$ and the last $n-k$ are possibilistic (epistemic uncertainty), represented by possibility distributions $(\pi^{Y_{k+1}},\cdots,\pi^{Y_n})$. The propagation of this hybrid uncertainty information can be performed by combining the Monte Carlo technique with the extension principle of fuzzy set theory (Baudrit et al. 2006).

The operative steps of the procedure are:

1 set $i = 0$
2 set $i = i + 1$
3 sample the $i-th$ realization $(y_1^i,\cdots,y_k^i)$ of the random variable vector $(Y_1,\cdots,Y_k)$
4 set the possibility value α to 0
5 select the corresponding α-cuts of the possibility distributions $(\pi^{Y_{k+1}},\cdots,\pi^{Y_n})$ as intervals of possible values of the possibilistic variables $(Y_{k+1},\cdots,Y_n)$
6 calculate the smallest and largest values of $f(y_1^i,\cdots,y_k^i,Y_{k+1},\cdots,Y_n)$, denoted by $\underline{f}_\alpha^i$ and $\overline{f}_\alpha^i$ respectively, considering the fixed values $(y_1^i,\cdots,y_k^i)$ sampled for the random variables $(Y_1,\cdots,Y_k)$ and all values of the possibilistic variables $(Y_{k+1},\cdots,Y_n)$ in the $\alpha-cuts$ of their possibility distributions $(\pi^{Y_{k+1}},\cdots,\pi^{Y_n})$; such extreme values $\underline{f}_\alpha^i$ and $\overline{f}_\alpha^i$ are the lower and upper limits, respectively, of the $\alpha-cut$ interval $[\overline{f}_\alpha^i,\underline{f}_\alpha^i]$ of $f(y_1^i,\cdots,y_k^i,Y_{k+1},\cdots,Y_n)$
7 if $\alpha<1$ then set $\alpha=\alpha+\Delta\alpha$ (e.g., $\Delta\alpha=0.05$) and return to step 5. above; otherwise obtain the fuzzy random realization (fuzzy interval) π_i^f of $Z=f(Y)$ as the collection of the values $\underline{f}_\alpha^i$ and $\overline{f}_\alpha^i$ for each $\alpha-cut$ (notice that since $\Delta\alpha=0.05$ then $N_\alpha = 1/\Delta\alpha + 1 = 1/0.05 + 1 = 21$ values of α are considered in the procedure, i.e., $N_\alpha = 21$ α-cuts of the possibility distributions $(\pi^{Y_{k+1}},\cdots,\pi^{Y_n})$ are selected; thus, the fuzzy random realization π_i^f of $Z=f(Y)$ is constructed as the collection of its $N_\alpha = 1/\Delta\alpha + 1 = 1/0.05 + 1 = 21$ α-cut intervals $[\overline{f}_\alpha^i,\underline{f}_\alpha^i]$)
8 if $i < m$ (e.g., $m = 10000$) then return to step 2. to generate a new realization of the random variables; otherwise, stop the algorithm

At the end of the procedure an ensemble of m random realizations of fuzzy intervals is obtained, i.e., $(\pi_1^f,\cdots,\pi_m^f)$.

Two considerations are in order with respect to the choices of N_α and m. A small number N_α of α-cuts (e.g., $N_\alpha = 5$) leads to a rough and imprecise characterization of the fuzzy random realizations π_i^f of Z; on the other hand, a large number N_α of α-cuts (e.g., $N_\alpha = 100$) causes a remarkable increase in the computational time. Thus, the choice of the number N_α of α-cuts is driven by the trade–off between estimation accuracy and computational cost. Similarly, the number m of realizations of the random variables has to be large enough to guarantee an accurate and precise propagation of the corresponding aleatory uncertainty: in practice, m is usually of the order of thousands.

For each set A contained in the universe of discourse U_Z of the output variable Z, it is possible to obtain the possibility measure $\Pi_i^f(A)$ and the ne-

cessity measure $N_i^f(A)$ from the corresponding possibility distribution $\pi_i^f(z)$, by:

$$\Pi_i^f(A) = \max_{z \in A}\{\pi_i^f(z)\},\ \forall A \subseteq U_Z \quad (3)$$

$$N_i^f(A) = \inf_{z \notin A}\{1 - \pi_i^f(z)\} = 1 - \Pi_i^f(\overline{A}),\ \forall A \subseteq U_Z \quad (4)$$

The *m* different realizations of possibility and necessity can then be combined to obtain the belief $Bel(A)$ and the plausibility $Pl(A)$ for any set A, respectively (Baudrit et al. 2006):

$$Bel(A) = \sum_{i=1}^{m} p_i N_i^f(A) \quad (5)$$

$$Pl(A) = \sum_{i=1}^{m} p_i \Pi_i^f(A) \quad (6)$$

where p_i is the probability of sampling the $i-th$ realization $(y_1^i, \cdots, y_k^i)$ of the random variable vector $(Y_1, \cdots, Y_k)$. For each set A, this technique thus computes the probability-weighted average of the possibility measures associated with each output fuzzy interval.

The likelihood of the value $f(Y)$ passing a given threshold z can then be computed by considering the belief and the plausibility of the set $A = (-\infty, z]$; in this respect, $Bel(f(Y) \in (-\infty, z])$ and $Pl(f(Y) \in (-\infty, z])$ can be interpreted as bounding, average cumulative distributions $\underline{F}(z) = Bel(f(Y) \in (-\infty, z])$, $\overline{F}(z) = Pl(f(Y) \in (-\infty, z])$ (Baudrit et al. 2006).

4 APPROACHES FOR CONSTRUCTING POSSIBILITY DISTRIBUTIONS

In this Section, a number of approaches for constructing possibility distributions of the variables subject to epistemic uncertainty are briefly described. In Section 4.1, triangular possibility distributions are considered; in Section 4.2, the use of Chebyshev inequality is illustrated; finally, in Section 4.3, two methods for transforming a probability distribution into a possibility distribution are described based on the principle of maximum specificity (Section 4.3.1) and on the normalization of the probability density function (Section 4.3.2).

4.1 *Triangular function*

Let us suppose that the analyst knows that an uncertain variable can take values in a given range [*a*, *b*] and the most likely value is *c*. To represent this information a possibility distribution can be taken as a triangle with base determined by the range [*a*, *b*] (i.e., the absolute physical limits of the variable) and with vertex taken in correspondence of the most likely value *c*: in other words, the possibility distribution equals 0 in correspondence of the extreme values *a* and *b* of the physically allowable range and 1 in correspondence of the most likely value *c*. It has been shown that the family of probability distributions defined by a triangular possibility distribution with range [*a*,*b*] and vertex c contains all the probability distributions with support [*a*,*b*] and mode *c* (Baudrit & Dubois 2006).

4.2 *Chebyshev inequality*

If the analyst knows the mean μ and the standard deviation σ of the uncertain variable of interest, then the Chebyshev inequality (Baudrit & Dubois 2006) can be used to construct a possibility distribution. Actually, the use of continuous possibility distributions for representing probability families heavily relies on probabilistic inequalities. Such inequalities provide probability bounds for intervals forming a continuous nested family around a typical value. This nestedness property leads to interpreting the corresponding family as being induced by a possibility measure. These bounds are usually used for proving convergence properties but, in this context, they can be used for representing knowledge. This is the case of the Chebyshev inequality, for instance. As pointed, for instance by Baudrit & Dubois (2006), the classical Chebyshev inequality defines a bracketing approximation on the confidence intervals around the known mean μ of a random variable *Y*, knowing its standard deviation σ. The Chebyshev inequality can be written as follows:

$$P(|Y - \mu| \le k\sigma) \ge 1 - \frac{1}{k^2} \text{ for } k \ge 1 \quad (7)$$

Chebyshev inequality defines a possibility distribution that dominates any density with given mean and variance: it allows to define a possibility distribution π by considering intervals $[\mu - k\sigma, \mu + k\sigma]$ as α-cuts of π and letting $\pi(\mu - k\sigma) = \pi(\mu + k\sigma) = 1/k^2$. Notice that from (7) $P(Y \in [\mu - k\sigma, \mu + k\sigma]) \ge 1 - 1/k^2$, for $k \ge 1$; moreover, $P(Y \in [\mu - k\sigma, \mu + k\sigma]) \ge 1 - \alpha$ by definition of α-cut interval. As a consequence, $\alpha = 1/k^2$ by construction. Since in this paper α = 0.05, 0.1, …, 0.95, 1 (see Section 3), then *k* = 4.4721, 3.1623, …, 1.0260, 1. The resulting possibility distribution defines a probability family $\mathcal{P}^{\mu,\sigma}(\pi)$ which has been proven to contain all distributions with mean μ and standard deviation σ, whether the unknown probability distribution function is symmetric or not, unimodal or not (Baudrit & Dubois 2006).

4.3 *Probability – possibility transformations*

In this Section, we consider transformations from probability distributions to possibility distributions. It is worth noting that in the transformation procedure (i.e., going from probability to possibility) "some information is lost because there is a conversion from pointed-valued probabilities to interval-valued ones" (Dubois et al. 1993).

Given the interpretation of possibility and necessity measures as upper and lower probabilities, a possibility distribution π induces a family $\mathcal{P}(\pi)$ of probability measures. There is not a one-to-one relation between possibility and probability, and transformations from a probability measure P into a possibility distribution π can only ensure that

- $\mathcal{P}(\pi)$ includes P; and
- $\mathcal{P}(\pi)$ is selected according to some principle (rationale); e.g., "minimize loss of information", in some sense.

The following should be basic principles for such transformations (Dubois et al. 1993):

1 The probability-possibility consistency principle.
The family $\mathcal{P}(\pi)$ is formally defined as $\mathcal{P}(\pi) = \{P : \forall A \subseteq Y, P(A) \le \Pi(A)\}$.
It seems natural to require a transformation to select P from $\mathcal{P}(\pi)$ (Dubois et al. 1993). This is referred to as the probability-possibility consistency principle, formulated as $P(A) \le \Pi(A), \forall A \subseteq Y$

2 Preference preservation.
A possibility distribution π induces a preference ordering on Y, such that $\pi(y) > \pi(y')$ means that the outcome y is preferred to y'. A transformation should therefore satisfy $\pi(y) > \pi(y') \Leftrightarrow p(y) > p(y')$.

In the following, two methods for transforming a probability distribution into a possibility distribution are considered: the first one is based on the principle of maximum specificity (Section 4.3.1); the second one on the normalization of the probability density function (Section 4.3.2).

4.3.1 *The principle of maximum specificity*

The most specific possibility distribution π^t, or rather the minimum area under π, that dominates a given probability density p is given by:

$$\pi^t(y) = \pi^t(h(y)) = \int_{-\infty}^{y} p(x)dx + \int_{h(y)}^{+\infty} p(x)dx =$$

$$= F(y) + \overline{F}(h(y)) \tag{8}$$

where $\overline{F}(\cdot) = 1 - F(\cdot)$ and $h(y) = \max\{x : p(x) \ge p(y)\}$.

It is interesting to observe that for this transformation: $N([y, h(y)]) = P([y, h(y)]) \le \Pi([y, h(y)]) = 1$, i.e., the transformation prescribes equality between the necessity of a given α-cut and the probability of the same α-cut.

The transformation applies to unimodal, continuous and support bounded probability densities p. Moreover, this criterion is not necessarily adapted to the transformation of a subjective probability distribution reflecting an expert opinion (Dubois et al. 2004).

4.3.2 *Normalization of probability density*

The possibility distribution resulting from the transformation is given by the normalization of probability density, i.e., $\mu_p = p(y) / \sup p(y)$. Note that the distribution resulting from this normalization (when taken to be a possibility distribution) does not in general adhere to the probability-possibility consistency principle (Dubois and Prade 1980).

5 CASE STUDY: FLOOD PROTECTION DESIGN

The case study deals with the design of a protection dike. The problem concerns a residential area that is closely located to a river; due to this location, there is a potential risk of flood. As prevention and mitigation measures, a dike has to be built to protect the area. Different design options must be considered taking into account that the construction of a dike involves high building costs and annual maintenance costs and the natural phenomenon of flooding is subject to a large amount of uncertainties. Thus, the analyst has to evaluate different design options while covering flood uncertainty.

The system model analyzed in this study is an analytical approximation for calculating the maximal water level of the river (i.e., the output variable of the model, Z_c) as a function of a number of parameters (i.e., the input variables of the model):

$$Z_c = Z_v + \left(\frac{Q}{K_s * B * \sqrt{(Z_m - Z_v)/L}} \right)^{3/5} \tag{9}$$

where: Q is the yearly maximal water discharge (m3/s); Z_m and Z_v are the riverbed levels (m asl) at the upstream and downstream part of the river under investigation, respectively; K_s is the Strickler friction coefficient; B and L are the width and length of the river part (m), respectively.

The input variables are classified as follows (Pasanisi et al. 2009):

- Constants:
  - $B = 300$ m;
  - $L = 5000$ m.
- Aleatory variables:
  - The maximal water flow Q is the variable with the largest amount of data available. A large set of water flow data is available to perform Bayesian inference (149 annual

maximal flow values) on the parameters of the distribution. The Gumbel distribution $Gum(q|\alpha,\beta)$ is a well-established probabilistic model for maximal flows:

$$Gum(q|\alpha,\beta)= \\ =\frac{1}{\beta}\exp\left[-\exp\left(\frac{q-\alpha}{\beta}\right)\right]\exp\left[\frac{\alpha-q}{\beta}\right] \qquad (10)$$

The Bayesian estimates (posterior means) of the parameters of the distribution are $\alpha=1014.0$ and $\beta=565.4$ (Pasanisi et al. 2009).

- The uncertainty in the upstream and downstream levels, Z_m and Z_v respectively, are quantified by a bivariate normal distribution $N(\boldsymbol{\mu},\boldsymbol{\Sigma})$. Indeed, as the upstream and downstream sections are quite close it seems reasonable to model them as possibly dependent variables. A total of 29 pairs of data $(Z_m^{(i)},Z_v^{(i)})$ have been used to perform Bayesian inference and setting the posterior distributions for $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. The point values μ^* and Σ^* used in this paper for the distribution parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are $\boldsymbol{\mu}^*=[\mu_m^*,\mu_v^*]=[55.0,\ 50.2]$ and $\boldsymbol{\Sigma}^*=[\sigma_m^*,\sigma_{mv}^*,\sigma_{vm}^*,\sigma_v^*]$ = [0.46, 0.3388; 0.3388, 0.39]: these are the posterior means of the probability density functions of the parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$.

  An additional consideration is in order with respect to the fact that variable Z_v appears *twice* in model function (9). Actually, the presence of *repeated* variables causes analytical difficulties in several uncertainty propagation methods, including interval analysis and fuzzy arithmetic (Ferson et al. 2003): in particular, it typically leads to an *overestimation* of the uncertainty. However, this is not the case for MC simulation (here employed to propagate aleatory uncertainty) because it (automatically) escapes the problem of repeated variables by instantiating *each* occurrence to the *same* randomly sampled value.

– Epistemic variables:

- The Strickler friction coefficient K_s is perhaps the most critical source of uncertainty. It is affected by epistemic uncertainty, since it is a simplification of a much more complex hydraulic model. In addition, assessing the uncertainty of K_s is difficult because, in practice, even if this coefficient is strongly related to the morphology of the river, it cannot be measured. As a consequence, data may only be retrieved through indirect calibration noised by significant observational uncertainty: this is reflected in the availability of only a very small series of 5 data sets with ± 15% noise (Limbourg & de Rocquigny 2010). The absolute physical limits of K_s are [*a*, *b*] = [5, 60], but the real value is expected to vary in a smaller range.

  In (Pasanisi et al. 2009), this epistemic variable is treated within a probabilistic framework: it is considered that the probability distribution of K_s is normal with mean μ and standard deviation σ equal to 30 and 7.5, respectively. In this paper, the epistemic uncertainty associated to K_s is represented by means of possibility distributions; the four methods described in the previous Section 4 are used. For the method of Section 4.1 (i.e., triangular possibility distribution), the base of the triangle is [5, 60] (i.e., the absolute physical limits of K_s) and the most likely value is 30 (i.e., the mean μ of the normal probability density function of K_s used in (Pasanisi et al. 2009)); for the methods of Section 4.2, 4.3, 4.4 (i.e., Chebyshev inequality and probability-possibility transformations), the mean μ and the standard deviation σ used are 30 and 7.5 (i.e., the mean and the standard deviation of the probability density function of K_s used in (Pasanisi et al. 2009)).

  The possibility distributions for K_s resulting from the application of the methods in Section 4.1-4.4 are shown in Figure 4.

It is worth noting that the area underlying the possibility distribution is related to the imprecision in the knowledge of the possibilistic variable: the larger the area, the higher the imprecision. In Figure 4, a direct visual comparison of the areas underlying the four possibility distributions considered is shown.

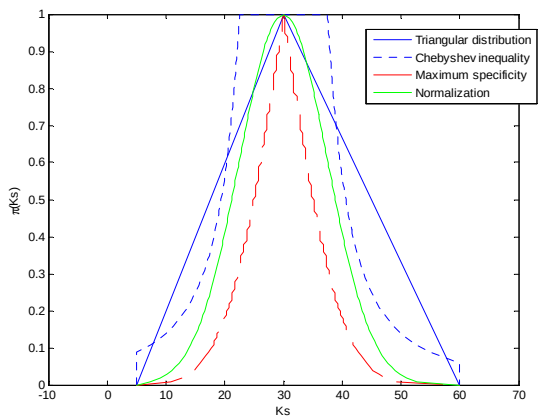

Figure 4. Comparison of the four different possibility distributions used for K_s: triangular function (Section 4.1), Chebyshev inequality (Section 4.2), principle of maximum specificity (Section 4.3.1), normalization of the probability density (Section 4.3.2).

It can be noticed that the larger areas are those underlying the possibility distributions built using the triangular function and the Chebyshev inequality. In fact, the information available to the analyst for

building these two possibility distributions is quite scarce: in the first case, only the physical limits and the most likely value of the variable are known; in the second case, only the mean value and the standard deviation are considered. On the contrary, the smaller areas underlying the possibility distributions constructed by the transformation methods are explained by the larger amount of information available to the analyst on the epistemic variable of interest, i.e., the probability distribution function itself.

For each of the four cases considered (i.e., for each of the possibility distributions built), the integrated uncertainty propagation procedure of Section 2 is run with m = 10000 realizations of the probabilistic variables; for each realization of the probabilistic variables, 21 values of α (0, 0.05, 0.1, …, 1) are considered to process the epistemic uncertainty associated to K_s. The results are compared to those obtained with a pure probabilistic approach, as in (Pasanisi et al. 2009), by sampling 10000 times a joint vector (Q, Z_m, Z_v, K_s) by the joint pdf $p(Q) \cdot p(Z_m, Z_v) \cdot p(K_s)$. With respect to the epistemic uncertain variable K_s, this approach consists in marginalizing the joint pdf of (Z_c, K_s): $p(Z_c \mid K_s) \cdot p(K_s)$, over all possible values of K_s.

6 RESULTS

Figure 5 (a-d) shows the comparison of the cumulative distribution functions of the maximal water level of the river (i.e., the output variable of the model, Z_c) obtained by the probabilistic uncertainty propagation approach (solid lines) with the belief (lower curves) and plausibility (upper curves) functions obtained by the integrated framework of uncertainty propagation where the possibility distributions for K_s are constructed with the methods of Sections 4.1-4.3.

It can be seen that:

- the integrated framework explicitly propagates the uncertainty by separating the contributions coming from the probabilistic and possibilistic variables; this separation is visible in the output distributions of the maximal water level of the river where the separation between the belief and plausibility functions reflects the imprecision in the knowledge of the possibilistic variable K_s and the slope pictures the variability of the probabilistic variables Q, Z_m, Z_v;
- the separation between the belief and plausibility functions is larger for the cases in Figures 5a and 5b (where the possibility distributions are those of Figure 4, built using the triangular function and Chebyschev inequality, respectively) than for those in Figures 5c and 5d (where the possibility distributions are those of Figure 4, built using the probability-possibility transformations); the larger gap between the belief and plausibility functions in Figures 5a and 5b than in Figures 5c and 5d is explained by the larger area contained under the corresponding possibility distribution functions (actually, the larger the area, the higher the imprecision in the knowledge of the possibilistic variable).
- the uncertainty in the output distribution of the pure probabilistic approach is given only by the slope of the cumulative distribution;
- as expected, the cumulative distribution of the maximal water level of the river obtained by the pure probabilistic method is within the belief and plausibility functions obtained by the hybrid approach

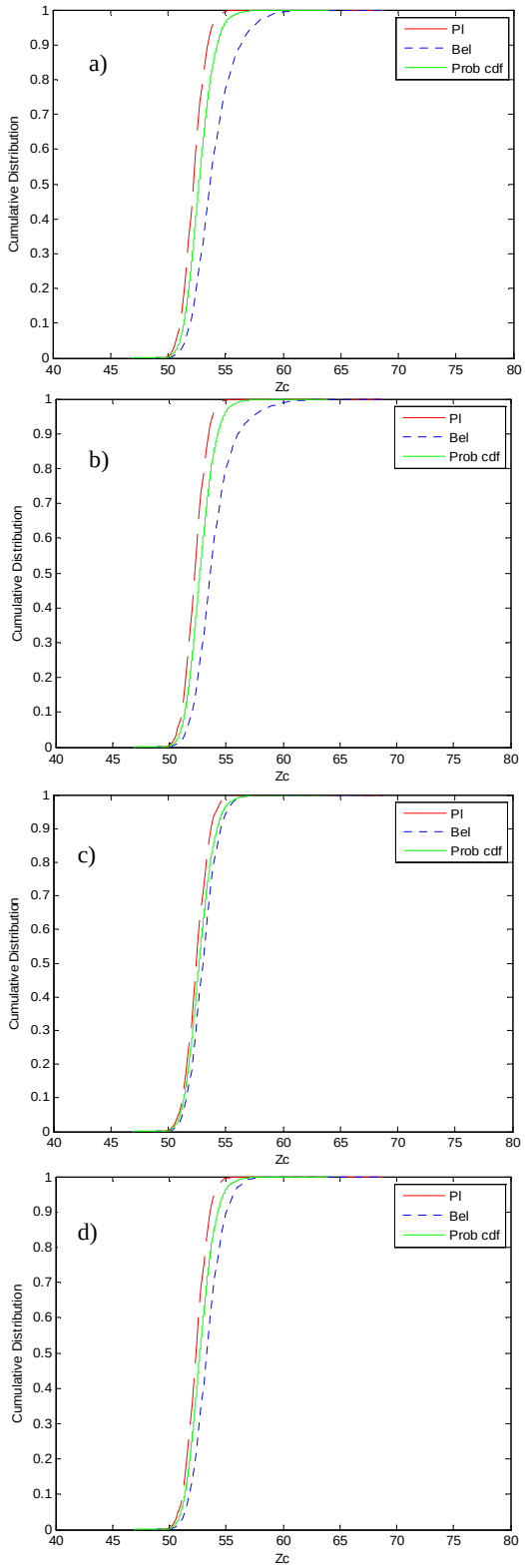

Figure 5. Comparison of the cumulative distribution functions of the maximal water level of the river Z_c obtained by the probabilistic uncertainty propagation approach (solid line) with the belief (lower dashed curve) and plausibility (upper dashed curve) functions obtained by the hybrid approach with the possibility distribution of K_s built using a) a triangular function, b) the Chebyshev inequality, c) the principle of maximum specificity d) the normalization of probability density (see Section 4 and Figure 4).

The final goal of the risk model assessment is to determine i) the dike level necessary to guarantee a given flood return period or ii) the flood risk for a given dike level.

With respect to item i) above, a reasonable quantity of interest is the 99% quantile of Z_c, i.e., $Z_c^{0.99}$, taken as the annual maximal flood level. This corresponds to the level of a "centennial" flood, the yearly maximal water level with a 100 year-return period. With respect to item ii) above, the quantity of interest that is mostly relevant to the decision maker is the probability that the maximal water level of the river Z_c exceeds a given threshold z^*, i.e., $P(Z_c > z^*)$; in the present paper, z^* = 55.5 m as in (Limbourg & de Rocquigny 2010). Table 1 reports the lower ($Z_{c,lower}^{0.99}$) and upper ($Z_{c,upper}^{0.99}$) 99th percentiles obtained from the two limiting cumulative distributions by using the four different possibility distributions proposed in Section 4 (i.e., triangular function, Chebyshev inequality, principle of maximum specificity and normalization of the probability density function) and the corresponding $Bel(Z_c > z^*)$ and $Pl(Z_c > z^*)$. In addition, as synthetic mathematical indicators of the imprecision in the knowledge of Z_c (i.e., of the separation between the belief and plausibility functions), the percentage widths:

- $W_{Zc} = (Z_{c,upper}^{0.99} - Z_{c,lower}^{0.99}) / Z_{c,prob}^{0.99}$ of the interval $\left[Z_{c,lower}^{0.99}, Z_{c,upper}^{0.99}\right]$ with respect to the percentile $Z_{c,prob}^{0.99}$ obtained by the pure probabilistic approach
- $W^* = (Pl(Z_c > z^*) - Bel(Z_c > z^*)) / P(Z_c > z^*)_{prob}$ of the interval $\left[Bel(Z_c > z^*), Pl(Z_c > z^*)\right]$ have been reported.

The numerical results in Table 1 confirm the similarities between the cumulative distributions obtained by using the triangular function and the Chebyshev inequality for the possibilistic representation of the uncertainty on K_s, and between the cumulative distributions obtained by the two different transformations from probability to possibility distributions.

Table 1. Lower and upper values of the Z_c percentiles and the threshold exceedance probability, and calculation of the indicator W about the width of the confidence interval.

Possibility distribution	$Z_c^{0.99}$ **(Pure probabilistic value = 56.10)**		$P[Z_c \geq 55.5]$ **(Pure probabilistic value = 0.0191)**	
	$\left[Z_{c,lower}^{0.99}, Z_{c,upper}^{0.99}\right]$	$W_{Zc}[\%]$	$[Bel, Pl]$	$W^*[\%]$
Triangular function	[54.57, 59.29]	8	[0.0015, 0.1682]	873
Chebyshev inequality	[54.40, 60.00]	10	[0.0014, 0.1631]	847
Transformation probability to possibility (principle of maximum specificity)	[54.83, 55.99]	2	[0.0043, 0.0344]	157
Transformation probability to possibility (normalization)	[54.60, 56.69]	4	[0.0028, 0.0705]	355

7 CONCLUSIONS

We have applied to a risk flood model a computational framework for the joint propagation of probabilistic and possibilistic uncertainties. Aleatory and epistemic uncertainties have been kept separate in the model, i.e., some of the variables are purely probabilistic (aleatory uncertainty) and some are purely possibilistic (epistemic uncertainty).

The following analyses have been carried out:

1 A comparison between the "hybrid" and the "pure probabilistic" approach, highlighting that:
   - the "global" uncertainty of the output within the pure probabilistic approach is only given by the cumulative distribution: the contributions of aleatory and epistemic uncertainties are here merged;
   - the hybrid approach explicitly propagates the uncertainty by separating the contributions coming from the probabilistic and possibilistic variables;
   - the larger gap between the belief and plausibility functions is explained by the larger area contained under the corresponding possibility distribution functions;
   - as expected, the cumulative distribution of the model output obtained by the pure probabilistic method is within the belief and plau-

sibility functions obtained by the hybrid approach.

2 A comparison of four methods for constructing the possibility distributions of the variables subject to epistemic uncertainty, showing that:
   - the choice of the possibility distribution depends on the information available about the variable: when the physical limits and the most likely value are available, a triangular possibility distribution can be constructed; when the mean and the standard deviation can be computed, e.g., by means of empirical data, the Chebyshev inequality can be used; when a probability distribution is available, the methods for transforming probability into possibility distributions can be employed;
   - there are similarities between the results obtained by using:
     - the triangular function and the Chebyshev inequality;
     - the two transformations from probability to possibility distributions (i.e., those based on the principle of maximum specificity and on the normalization of the probability density function).

   These similarities are explained by the same "uncertainty content" borne by the corresponding possibility distributions (as demonstrated by the similar area limited by the possibility distribution functions).

Future research will be devoted to:
- the study and development of methods for the joint propagation of probabilistic and possibilistic uncertainty in cases when the uncertain variables are not purely aleatory or purely epistemic (for example, the uncertainty of a variable might be described by a probability distribution whose parameters are themselves poorly known and represented by a possibility distribution);
- the comparison of the integrated probabilistic – possibilistic approach to other computational frameworks for the joint propagation of aleatory and epistemic uncertainties, e.g., the double Monte Carlo method (Baudrit et al. 2008) and the Dempster – Shafer theory of evidence (Ferson et al. 2003);
- the treatment of dependencies between probabilistic and possibilistic variables;
- the use of advanced integrated simulation methods within the framework of uncertainty propagation for reducing the associated computational cost.

REFERENCES

Aven, T. & Zio, E. 2010. Some considerations on the treatment of uncertainties in risk assessment for practical decision making. *Reliability Engineering and System Safety*, 96 (1): 64-74.

Baraldi, P. & Zio, E. 2008. A Combined Monte Carlo and Possibilistic Approach to Uncertainty Propagation in Event Tree Analysis. *Risk Analysis*, 28 (5): 1309-1326.

Baudrit, C. & Dubois, D. 2006. Practical Representations of Incomplete Probabilistic Knowledge. *Computational Statistics & Data Analysis*, 51 (1): 86-108.

Baudrit, C., Dubois, D., & Guyonnet, D. 2006. Joint Propagation of Probabilistic and Possibilistic Information in Risk Assessment. *IEEE Transactions on Fuzzy Systems*, 14 (5): 593-608.

Baudrit, C. Dubois, D. & Perrot, N. 2008. Representing parametric probabilistic models tainted with imprecision. *Fuzzy sets and system*, 159: 1913-1928.

Dubois, D. 2006. Possibility theory and statistical reasoning. *Computational Statistics & Data Analysis*, 51 (1): 47-69

Dubois, D. & Prade, H. 1980. *Fuzzy Sets and Systems: Theory and Applications*, New York: Academic Press.

Dubois, D., Prade, H., & Sandri, S. 1993. On possibility/probability transformations. In R. Lowen & M. Roubens (Eds), *Fuzzy Logic: State of the Art*: 103–112. Dordrecht: Kluwer Academic Publishers.

Dubois, D., Foulloy, L., Mauris, G., & Prade, H. 2004. Probability-possibility transformations, triangular fuzzy sets, and probabilistic inequalities. *Reliable Computing* 10 (4): 273-297.

Ferson, S., Ginzburg, L., Kreinovich, V., Myers, D.M. & Sentz, K. 2003. Construction Probability Boxes and Dempster-Shafer Structures, Sandia National Laboratories.

Flage, R., Baraldi, P., Zio, E., Aven, T. 2010. Possibility-probability transformation in comparing different approaches to the treatment of epistemic uncertainties in a fault tree analysis. In B. Ale, I.A. Papazoglu, E. Zio (Eds.), Reliability, Risk and Safety - Proceedings of the ESREL 2010 Conference, Rhodes, Greece: 714-721,

Flage, R., Baraldi, P., Zio, E., Aven, T. 2011. Probabilistic and possibilistic treatment of epistemic uncertainties in fault tree analysis. Submitted to *Risk analysis*.

Helton, J.C. 2004. Alternative Representations of Epistemic Uncertainty. *Reliability Engineering and System Safety*, 85 (1-3): 1-10.

Klir, G.J. & Yuan, B. 1995. *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Upper Saddle River, NJ: Prentice-Hall.

Limbourg, P. & de Rocquigny, E. 2010. Uncertainty analysis using evidence theory – confronting level-1 and level-2 approaches with data availability and computational constraints. *Reliability Engineering and System Safety* 95 (5): 550-564.

Pasanisi, A., de Rocquigny, E., Bousquet, N. & Parent, E. 2009. Some useful features of the Bayesian setting while dealing with uncertainties in industrial practice. Proceedings of the ESREL 2009 Conference, Prague, Czech Republic: 1795-1802