Empirical comparison of two methods for the Bayesian update of the parameters of probability distributions in a two-level hybrid probabilistic-possibilistic uncertainty framework for risk assessment

Nicola Pedroni, Enrico Zio, A Pasanisi, M Couplet

Empirical comparison of two methods for the Bayesian update of the parameters of probability distributions in a two-level hybrid probabilistic-possibilistic uncertainty framework for risk assessment

N. Pedroni[1], E. Zio[2], A. Pasanisi[3], M. Couplet[4]

Abstract

In this paper, we address the issue of updating in a Bayesian framework, the possibilistic representation of the epistemically-uncertain parameters of (aleatory) probability distributions, as new information (e.g., data) becomes available. Two approaches are considered: the first is based on a purely possibilistic counterpart of the classical, well-grounded probabilistic Bayes' theorem; the second relies on the hybrid combination of (i) Fuzzy Interval Analysis (FIA) to process the uncertainty described by possibility distributions and (ii) repeated Bayesian updating of the uncertainty represented by probability distributions.

The feasibility of the two methods is shown on a literature case study involving the risk-based design of a flood protection dike.

Keywords: hierarchical two-level uncertainty, Bayesian update, possibility distributions, fuzzy interval analysis, possibilistic Bayes theorem, flood protection dike.

[1] Corresponding author. Assistant professor at the Chair of system science and the energetic challenge-Fondation EdF at CentraleSupélec, 3 Rue Joliot Curie – Plateau de Moulon, 91190, Gif-Sur-Yvette, France. E-mail: nicola.pedroni@supelec.fr.

[2] Full professor at the Chair of system science and the energetic challenge-Fondation EdF at CentraleSupélec, Grande Voie des Vignes, 92290, Chatenay-Malabry, France. E-mail: enrico.zio@ecp.fr. Also: Energy Department, Politecnico di Milano, Via Ponzio, 34/3 – 20133 Milano, Italy. E-mail: enrico.zio@polimi.it.

[3] Project Manager at EDF R&D - EIFER, Emmy-Noether-Str. 11. 76131 Karlsruhe, Germany. E-mail address: alberto.pasanisi@edf.fr.

[4] Research Engineer at Electricité de France, R&D, 6 Quai Watier, 78400, Chatou, France. E-mail address: mathieu.couplet@edf.fr.

Introduction

We consider a framework of uncertainty representation with two hierarchical levels [Limbourg and de Rocquigny, 2010], in which risk analysis models of aleatory (i.e., random) events (e.g., failures) contain parameters (e.g., probabilities, failure rates, …) that are epistemically-uncertain, i.e., known with poor precision due to lack of knowledge and information. Traditionally, both types of uncertainty are represented by probability distributions [Apostolakis and Kaplan, 1981; Apostolakis, 1990; NUREG-CR-6850, 2005; USNRC, 2002 and 2009; NASA, 2010] and Bayes' rule is useful for updating the (probabilistic) epistemic uncertainty representation as new information (e.g., data) becomes available [Bernardo and Smith, 1994; Siu and Kelly, 1998; Lindley, 2000 and 2006; Bedford and Cooke, 2001; Atwood et al., 2003; Kelly and Smith, 2009 and 2011; Pasanisi et al., 2012].

However, in some situations, insufficient knowledge, information and data impair a probabilistic representation of epistemic uncertainty. A number of alternative representation frameworks have been proposed for such cases [Aven, 2010 and 2011; Aven and Steen, 2010; Aven and Zio, 2011; Flage et al., 2009; Beer et al., 2013b and 2014b; Zhang et al., 2013], e.g., e.g., fuzzy set theory [Klir and Yuan, 1995], fuzzy probabilities [Buckley, 2005; Beer, 2009b; Pannier et al., 2013], random set theory [Molchanov, 2005], evidence theory [Ferson et al., 2003 and 2004; Helton et al., 2007 and 2008; Sentz and Ferson, 2002; Le Duy et al., 2013; Sallak et al., 2013], possibility theory (that can be considered a special case of evidence theory) [Baudrit and Dubois, 2006; Baudrit et al., 2006 and 2008; Dubois, 2006; Dubois and Prade, 1988], probability bound analysis using probability boxes (p-boxes) [Ferson and Ginzburg, 1996; Crespo et al., 2013; Mehl, 2013], interval analysis [Ferson and Hajagos, 2004; Ferson and Tucker, 2006; Ferson et al., 2007 and 2010; Jalal-Kamali and Kreinovich, 2013; Muscolino and Sofi, 2013; Zhang et al., 2013] and interval probabilities [Weichselberger, 2000]; notice that most of these theories can be included within the general common framework of *imprecise probabilities* [Kuznetsov, 1991; Walley, 1991; Kozine and

47 Filimonov, 2000; Kozine and Utkin, 2002; Coolen and Utkin, 2007; Beer and Ferson, 2013; Beer et
48 al., 2013a; Blockley, 2013; Reid, 2013; Sankararaman and Mahadevan, 2013].
49 In this paper, we adopt possibility distributions to describe epistemic uncertainty and address the
50 issue of updating, in a Bayesian framework, the possibilistic representation of the epistemically-
51 uncertain parameters of (aleatory) probability distributions. We take two approaches of literature.
52 The first is based on a purely possibilistic counterpart of the classical, well-grounded probabilistic
53 Bayes' theorem: it requires the construction of a possibilistic likelihood function which is used to
54 revise the prior possibility distributions of the uncertain parameters (determined on the basis of a
55 priori subjective knowledge and/or data) [Dubois and Prade, 1997; Lapointe and Bobee, 2000]. This
56 approach has been already applied by the authors for updating possibility distributions in [Pedroni
57 et al., 2014]. The second is a hybrid probabilistic-possibilistic method that relies on the use of
58 Fuzzy Probability Density Functions (FPDFs), i.e., PDFs with possibilistic (fuzzy) parameters. It is
59 based on the combination of (i) Fuzzy Interval Analysis (FIA) to process the uncertainty described
60 by possibility distributions and (ii) repeated Bayesian updating of the uncertainty represented by
61 probability distributions [Beer, 2009a; Stein and Beer, 2011; Stein et al., 2013; Beer et al., 2014a].
62 The objective (and the main contribution of the paper) is to compare the effectiveness of the two
63 methods. To the best of the authors' knowledge, this is the first time that the above mentioned
64 techniques are systematically compared with reference to risk assessment problems where hybrid
65 uncertainty is separated into two hierarchical levels. To keep the analysis simple and retain a clear
66 view of each step, the investigations are carried out with respect to a simple literature case study
67 involving the risk-based design of a flood protection dike [Pasanisi et al., 2009; Limbourg and de
68 Rocquigny, 2010]. In addition, different numerical indicators (e.g., cumulative distributions,
69 exceedance probabilities, percentiles, …) are considered to perform a fair, quantitative comparison
70 between the methods and evaluate their rationale and appropriateness in relation to risk analysis.
71 Other methods have been proposed in the literature to revise, in a Bayesian framework, non-
72 probabilistic representations of epistemic uncertainty [Ferson, 2005]. In [Viertl, 1996, 1997, 1999,

2008a,b and 2011; Viertl and Hareter, 2004a,b; Viertl and Hule, 1991] a modification of the Bayes' theorem is presented to account for the presence of fuzzy data and fuzzy prior PDFs: the approach is similar to that employed by [Beer, 2009a; Stein and Beer, 2011; Stein et al., 2013] and considered in this paper. In [Smets, 1993] a Generalized Bayes Theorem (GBT) is proposed within the framework of evidence theory: this approach is applied by [Le-Duy et al., 2011] to update the estimates of the failure rates of mechanical components in the context of nuclear Probabilistic Risk Assessment (PRA). Finally, in [Walley, 1996; Bernard, 2005; Masegosa and Moral, 2014] Imprecise Dirichlet Models (IDMs) are proposed for objective statistical inference from multinomial data. In the IDM, prior or posterior uncertainty about a parameter is described by a set of Dirichlet distributions, and inferences about events are summarized by lower and upper probabilities. This model has been extended by [Quaeghebeur and de Cooman, 2005] to generalized Bayesian inference from canonical exponential families and by [Walter and Augustin, 2009] with the aim of handling prior-data conflicts.

The remainder of the paper is organized as follows. First, the representation of aleatory (probabilistic) and epistemic (possibilistic) uncertainties in a "two-level" framework is provided; then, the two methods employed in this paper for the Bayesian update of the possibilistic parameters of aleatory probability distributions are described in details; after that, the case study concerning the risk-based design of a flood protection dike is presented; in the following Section, the methods described are applied to the case study: the results obtained are discussed and the two methods are synthetically compared; finally, some conclusions are drawn in the last Section.

Representation of aleatory and epistemic uncertainties in a two-level framework: fuzzy random variables

In all generality, we consider an uncertain variable Y, whose uncertainty is described by the Probability Distribution Function (PDF) $p^Y(y|\boldsymbol{\theta})$, where $\boldsymbol{\theta} = \{\theta_1, \theta_2, ..., \theta_m, ..., \theta_P\}$ is the vector of the corresponding internal parameters. In a two-level framework, the parameters $\boldsymbol{\theta}$ are themselves

affected by epistemic uncertainty [Limbourg and de Rocquigny, 2010]. In the present work, we describe these uncertainties by the (generally joint) possibility distribution $\pi^{\theta}(\boldsymbol{\theta})$ (it is straightforward to notice that in case the internal parameters $\{\theta_1, \theta_2, ..., \theta_m, ..., \theta_P\}$ are independent, then $\pi^{\theta}(\boldsymbol{\theta})$ is simply represented by a group of *P* separate, marginal and independent (i.e., non-interactive) possibility distributions, i.e., $\pi^{\theta}(\boldsymbol{\theta}) = \{\pi^{\theta_1}(\theta_1), \pi^{\theta_2}(\theta_2), ..., \pi^{\theta_m}(\theta_m), ..., \pi^{\theta_P}(\theta_P)\}$). A random variable Y with possibilistic parameters $\boldsymbol{\theta}$ is a particular case of a Fuzzy Random Variable (FRV), i.e., of a random variable whose values are not real, but rather fuzzy numbers [Féron, 1976; Kwakernaak, 1978; Puri and Ralescu, 1986; Baudrit et al., 2008; Couso and Sanchez, 2008]. The corresponding Fuzzy Probability Distribution Function (FPDF) is here indicated as $\tilde{p}^{Y}(y|\boldsymbol{\theta})$.

For clarification by way of example, we may consider the generic uncertain variable Y described by a Gumbel PDF, i.e., $Y \sim p^{Y}(y|\boldsymbol{\theta}) = Gum(\boldsymbol{\theta}) = Gum(\theta_1, \theta_2) = Gum(\gamma, \delta) = p^{Y}(y|\gamma, \delta)$. Parameter $\delta = \theta_2$ (i.e., the scale paramter) is a fixed point-wise value ($\delta = \theta_2$ = 100), whereas parameter $\gamma = \theta_1$ (i.e., the location parameter) is epistemically-uncertain. By hypothesis, the only information available on $\gamma = \theta_1$ is that it is defined on interval [a_γ, b_γ] = [900, 1300] and its most likely value is c_γ = 1100. Notice that such information is *not sufficient* for assigning a *single specific* probability distribution to describe the epistemic uncertainty in parameter $\gamma = \theta_1$. In facts, such scarce information is actually *compatible* with a *variety* of probability distributions (e.g., truncated normal, lognormal, triangular, …). To address this issue, this limited state of knowledge about $\gamma = \theta_1$ is here described by a triangular possibility distribution $\pi^{\gamma}(\gamma)$ with core c_γ = 1100 and support [a_γ, b_γ] = [900, 1300] (Fig. 1, top left column) [Baudrit and Dubois, 2006]. Indeed, this representation is *coherent* with the *information available*, as it can be demonstrated that such possibility distribution "encodes" the *family* of *all* the probability distributions with mode c_γ = 1100 and support [a_γ, b_γ] = [900, 1300] (obviously, this does *not* mean that the triangular possibility distribution is the *only one* with these characteristics, i.e., the only one able to encode such a

probability family). In other words, *one single* possibility distribution generates in practice a "*bundle*" of probability distributions with mode c_γ = 1100 and support $[a_\gamma, b_\gamma]$ = [900, 1300]. The reader is referred to [Baudrit and Dubois, 2006; Couso et al., 2001; Dubois et al., 2004] for further technical details and a formal proof of these statements.

In order to provide an additional *practical* interpretation of the possibility distribution $\pi^\gamma(\gamma)$ of $\gamma = \theta_1$, we can define its α-cut sets $A_\alpha^\gamma = \{\gamma:\ \pi^\gamma(\gamma) \geq \alpha\}$, with $0 \leq \alpha \leq 1$. For example, $A_{0.5}^\gamma$ = [1000, 1200] is the set of γ values for which the possibility function is greater than or equal to 0.5 (dashed segment in Fig. 1, top left column). Notice that the α-cut set A_α^γ of parameter γ can be interpreted also as the $(1 - \alpha)\cdot 100\%$ Confidence Interval (CI) for γ, i.e., the interval such that $P[\gamma \in A_\alpha^\gamma] \geq 1-\alpha$. For example, A_0^γ = [900, 1300] is the $(1 - 0)\cdot 100\% = 100\%$ CI for γ, i.e., the interval that contains the "true" value of γ with certainty (solid segment in Fig. 1, top left column); $A_{0.5}^\gamma$ = [1000, 1200] ($\subset A_0^\gamma$) is the $(1 - 0.5)\cdot 100\% = 50\%$ CI (dashed segment in Fig. 1, top left column); $A_{0.8}^\gamma$ = [1050, 1150] ($\subset A_{0.5}^\gamma \subset A_0^\gamma$) is the $(1 - 0.8)\cdot 100\% = 20\%$ CI, and so on. In this view, the possibility distribution $\pi^\gamma(\gamma)$ can be interpreted as a set of *nested* CIs for parameter γ [Baudrit and Dubois, 2006; Couso et al., 2001; Dubois et al., 2004].

For each possibility (resp., confidence) level α (resp., $1 - \alpha$) in [0, 1], a *family* of PDFs for *Y*, namely $\{p^Y(y \mid \gamma, \delta)\}_\alpha$, can be generated by letting the epistemically-uncertain parameter γ range within the corresponding α-cut set A_α^γ, i.e., $\{p^Y(y \mid \gamma, \delta)\}_\alpha = \{p^Y(y \mid \gamma, \delta): \gamma \in A_\alpha^\gamma, \delta = 100\}$. By way of example, Fig. 1, top right column, shows four PDFs belonging to the family $\{p^Y(y \mid \gamma, \delta)\}_{\alpha=0}$ (solid lines) and four PDFs belonging to the family $\{p^Y(y \mid \gamma, \delta)\}_{\alpha=0.5}$ (dashed lines).

In the same way, a *bundle* of Cumulative Distribution Functions (CDFs) for *Y*, namely $\{F^Y(y \mid \gamma, \delta)\}_\alpha$, can be constructed by letting γ range within A_α^γ, i.e., $\{F^Y(y \mid \gamma, \delta)\}_\alpha = \{F^Y(y \mid \gamma, \delta): \gamma \in A_\alpha^\gamma, \sigma = 100\}$. This *family* of CDFs (of level α) is bounded above and below by the

upper and lower CDFs, $\overline{F}_{\alpha}^{Y}(y)$ and $\underline{F}_{\alpha}^{Y}(y)$, defined as $\overline{F}_{\alpha}^{Y}(y)=\sup_{\gamma\in A_{\alpha}^{\gamma}}\left\{F^{Y}(y\,|\,\gamma,\delta=100)\right\}$ and $\underline{F}_{\alpha}^{Y}(y)=\inf_{\gamma\in A_{\alpha}^{\gamma}}\left\{F^{Y}(y\,|\,\gamma,\delta=100)\right\}$, respectively. Since $\pi^{\gamma}(\gamma)$ can be interpreted as a set of nested CIs for parameter γ (see above), it can be argued that the α-cuts of $\pi^{\gamma}(\gamma)$ induce also a *set* of *nested pairs* of CDFs $\left\{\left(\underline{F}_{\alpha}^{Y}(y),\overline{F}_{\alpha}^{Y}(y)\right):0\le\alpha\le1\right\}$ which bound the "true" CDF $F^{Y}(y)$ of *Y* with confidence larger than or equal to (1 – α), i.e., $P[\underline{F}_{\alpha}^{Y}(y)\le F^{Y}(y)\le\overline{F}_{\alpha}^{Y}(y)]\ge1-\alpha$, with $0 \le \alpha \le 1$ [Baudrit et al., 2007 and 2008]. In passing, notice that the upper and lower CDFs (of level α), $\overline{F}_{\alpha}^{Y}(y)$ and $\underline{F}_{\alpha}^{Y}(y)$, can be referred to as the *plausibility* and *belief* functions (of level α) of the set $Z = (-\infty, y]$, i.e., $\overline{F}_{\alpha}^{Y}(y)=Pl_{\alpha}^{Y}(Z)$ and $\underline{F}_{\alpha}^{Y}(y)=Bel_{\alpha}^{Y}(Z)$, respectively. For illustration purposes, Fig. 1, bottom row, shows the bounding upper and lower CDFs of *Y*, $\overline{F}_{\alpha}^{Y}(y)=Pl_{\alpha}^{Y}(Z)$ and $\underline{F}_{\alpha}^{Y}(y)=Bel_{\alpha}^{Y}(Z)$, built in correspondence of the α-cuts of level $\alpha = 0$ (solid lines), 0.5 (dashed lines) and 1 (dot-dashed line) of the possibility distribution $\pi^{\gamma}(\gamma)$ of parameter γ (Fig. 1, top left column).

Finally, the *set* of *nested pairs* of CDFs $\left\{\left(\underline{F}_{\alpha}^{Y}(y),\overline{F}_{\alpha}^{Y}(y)\right):0\le\alpha\le1\right\} = \left\{\left(Bel_{\alpha}^{Y}(Z),Pl_{\alpha}^{Y}(Z)\right):0\le\alpha\le1\right\}$, $Z = (-\infty, y]$, can be synthesized into a *single pair* of plausibility and belief functions as $Pl^{Y}(Z)=\int_{0}^{1}Pl_{\alpha}^{Y}(Z)d\alpha$ and $Bel^{Y}(Z)=\int_{0}^{1}Bel_{\alpha}^{Y}(Z)d\alpha$, respectively (dotted lines in Fig. 1, bottom row): in other words, $Pl^{Y}(Z)$ and $Bel^{Y}(Z)$ are obtained by *averaging* the different nested plausibility and belief functions (i.e., $\left\{\left(Bel_{\alpha}^{Y}(Z),Pl_{\alpha}^{Y}(Z)\right):0\le\alpha\le1\right\}$) generated at different possibility levels $\alpha \in [0, 1]$ (i.e., by averaging the different contributions to the plausibility and belief functions produced by different α-cuts of the epistemic parameter γ). The plausibility and belief functions $Pl^{Y}(Z)$ and $Bel^{Y}(Z)$, $Z = (-\infty, y]$, are shown to represent the "*best bounds*" for the "true" CDF $F^{Y}(y)$ of the uncertain variable *Y* [Ralescu, 2002; Baudrit et al., 2007 and 2008; Couso et al., 2004; Couso and Dubois, 2009; Couso and Sanchez, 2011].

Further details about FRVs are not given here for the sake of brevity: the interested reader is referred to the cited references.

Fig. 1

Bayesian update of the possibilistic parameters of aleatory probability distributions

In this Section, we present the methods employed in this study for updating, in a Bayesian framework, the possibilistic representation of the epistemically-uncertain parameters of (aleatory) probability distributions, as *new* information/evidence (e.g., data) becomes available. In this view, let $\pi^{\theta}(\boldsymbol{\theta})$ be the (joint) *prior* possibility distribution for the parameters $\boldsymbol{\theta} = [\theta_1, \theta_2, ..., \theta_m, ..., \theta_P]$ of the PDF $p^Y(y|\boldsymbol{\theta})$ of variable Y (built on the basis of a priori subjective engineering knowledge and/or data). For example, in the risk assessment context of this paper Y may represent the yearly maximal water flow of a river described by a Gumbel distribution: thus, $Y \sim p^Y(y|\boldsymbol{\theta}) = Gum(\boldsymbol{\theta}) = Gum(\theta_1, \theta_2) = Gum(\gamma, \delta) = p^Y(y|\gamma,\delta)$ and $\pi^{\theta}(\boldsymbol{\theta}) = \pi^{\gamma,\delta}(\gamma,\delta)$. Moreover, let $\boldsymbol{y} = [y_1, y_2, ..., y_k, ..., y_D]$ be a vector of D *observed* pieces of data representing the *new* information/evidence *available* for the analysis: referring to the example above, $\boldsymbol{y}$ may represent a vector of D values collected over a long period time (e.g., many years) of the yearly maximal water flow of the river under analysis. The objective of the Bayesian analysis is to update the a priori representation $\pi^{\theta}(\boldsymbol{\theta}) = \pi^{\gamma,\delta}(\gamma,\delta)$ of $\boldsymbol{\theta} = [\gamma, \delta]$ on the basis of the new evidence acquired, i.e., to calculate the *posterior* possibility distribution $\pi^{\theta}(\boldsymbol{\theta}|\boldsymbol{y})$ (i.e., $\pi^{\gamma,\delta}(\gamma,\delta|\boldsymbol{y})$) of $\boldsymbol{\theta}$ after $\boldsymbol{y}$ is obtained.

In the present paper, two methods are considered to this aim: the purely possibilistic method and the hybrid probabilistic and possibilistic approach.

Purely possibilistic approach

The purely possibilistic method (hereafter also referred to as 'Approach A' for brevity) is based on a purely possibilistic counterpart of the classical, probabilistic Bayes' theorem [Dubois and Prade, 1997; Lapointe and Bobée, 2000]:

$$\pi^{\theta}(\boldsymbol{\theta} \mid \boldsymbol{y}) = \frac{\pi_L^{\theta}(\boldsymbol{\theta} \mid \boldsymbol{y}) \cdot \pi^{\theta}(\boldsymbol{\theta})}{\sup_{\boldsymbol{\theta}} \left\{ \pi_L^{\theta}(\boldsymbol{\theta} \mid \boldsymbol{y}) \cdot \pi^{\theta}(\boldsymbol{\theta}) \right\}}, \tag{1}$$

where $\pi_L^{\theta}(\boldsymbol{\theta} \mid \boldsymbol{y})$ is the possibilistic likelihood of the parameter vector $\boldsymbol{\theta}$ given the newly observed data $\boldsymbol{y}$, and quantities $\pi^{\theta}(\boldsymbol{\theta} \mid \boldsymbol{y})$ and $\pi^{\theta}(\boldsymbol{\theta})$ are defined above. Notice that $\sup_{\boldsymbol{\theta}} \left\{ \pi_L^{\theta}(\boldsymbol{\theta} \mid \boldsymbol{y}) \cdot \pi^{\theta}(\boldsymbol{\theta}) \right\}$ is a normalization factor such that $\sup_{\boldsymbol{\theta}} \left\{ \pi^{\theta}(\boldsymbol{\theta} \mid \boldsymbol{y}) \right\} = 1$, as required by possibility theory [Dubois, 2006].

It is worth mentioning that forms of the possibilistic Bayes' theorem alternative to (1) can be constructed as a result of other definitions of the operation of 'conditioning' with possibility distributions: the reader is referred to [Dubois and Prade, 1997; Lapointe and Bobée, 2000] for technical details. In this paper, expression (1) has been chosen because "it satisfies desirable properties of the revision process and lead to continuous posterior distributions" [Lapointe and Bobée, 2000].

The possibilistic likelihood $\pi_L^{\theta}(\boldsymbol{\theta} \mid \boldsymbol{y})$ is here obtained by *transforming* the classical probabilistic likelihood function $L^{\theta}(\boldsymbol{\theta} \mid \boldsymbol{y})$ through *normalization*, i.e., $\pi_L^{\theta}(\boldsymbol{\theta} \mid \boldsymbol{y}) = \frac{L^{\theta}(\boldsymbol{\theta} \mid \boldsymbol{y})}{\sup_{\boldsymbol{\theta}} \left\{ L^{\theta}(\boldsymbol{\theta} \mid \boldsymbol{y}) \right\}}$ [Anoop et al., 2006] (obviously, $L^{\theta}(\boldsymbol{\theta} \mid \boldsymbol{y}) = \prod_{k=1}^{D} p^{Y}(y_k \mid \boldsymbol{\theta})$ in the case the observations $\{y_k: k = 1, 2, \ldots, D\}$ are independent and identically distributed). This choice has been made for the following main reasons:

i. the transformation is simple and can be straightforwardly applied to any distribution [Anoop et al., 2006];

ii. the resulting possibilistic likelihood is *very closely related* to the classical, purely probabilistic one (which is theoretically well-grounded) by means of the simple and direct

operation of normalization that preserves the "*original structure*" of the experimental evidence;

iii. it can be easily verified that the resulting possibilistic likelihood keeps the *sequential* nature of the updating procedure typical of the standard Bayes' theorem;

iv. the operation of likelihood normalization finds also *theoretical* justifications in some recent works of literature (see the brief discussion below) [Denoeux, 2014; Moral, 2014].

However, two considerations are in order with respect to this choice. First, it has to be admitted that the resulting possibility distributions do not in general adhere to the probability-possibility consistency principle [Dubois and Prade, 1980]. Second, it has to be remembered that the probabilistic likelihood function $L^{\theta}(\boldsymbol{\theta} \mid \boldsymbol{y})$ is *not* a probability distribution: in this view, from a rigorous mathematical viewpoint, speaking of probability/possibility transformation for it would be wrong. On the other side, from the practical engineering viewpoint of interest to the present paper, an operation of normalization can be performed (i.e., $L^{\theta}(\boldsymbol{\theta} \mid \boldsymbol{y}) \Big/ \int_{\theta} L^{\theta}(\boldsymbol{\theta} \mid \boldsymbol{y}) d\boldsymbol{\theta}$) in order to "technically" provide it with the "properties" of a probability distribution function.

It is worth noting that other techniques of transformation of probability density functions into possibility distributions exist, but the corresponding details are not given here for brevity sake: the interested reader is referred to [Dubois et al., 1993, 2004 and 2008; Flage et al., 2010 and 2013] for some proposed techniques, e.g., the principle of maximum specificity [Dubois et al., 1993] and the principle of minimal commitment [Dubois et al., 2008]. Also, it has to be noticed that techniques are also available to construct possibility distributions (and, thus, possibilistic likelihood functions) *directly* from rough experimental *data* (i.e., without resorting to probability-possibility transformations): see [Masson and Denoeux, 2006; Mauris, 2008; Hou and Yang, 2010; Serrurier and Prade, 2011] for more details. Finally, for a thorough *theoretical* justification of a "possibilistic vision" of the likelihood the reader is referred to: e.g., [Dubois et al., 1997], where possibility measures are considered as the *supremum* of a *family* of likelihood functions; [Denoeux, 2014],

where the evidence about a parameter (after observing a piece of data) is represented by a *consonant* "likelihood based" belief function, whose contour function equals the *normalized likelihood function* (see above): in the paper, this is also rigorously derived from three basic principles, i.e., the likelihood principle [Edwards, 1992], compatibility with Bayes' rule and the minimal commitment principle [Smets, 1993]; and finally [Moral, 2014], where the approach by [Denoeux, 2014] is discussed and the issue of representing likelihood information is taken from the point of view of imprecise probabilities.

Hybrid probabilistic and possibilistic approach

The hybrid probabilistic and possibilistic method (hereafter also called 'Approach B' for brevity) is based on the construction of a Fuzzy Probability Distribution Function (FPDF) to be used as a 'fictitious' prior for the epistemically-uncertain parameters $\boldsymbol{\theta}$ of the PDF $p^Y(y|\boldsymbol{\theta})$ of the uncertain variable *Y*: in other words, a fictitious (artificial) *probabilistic* function has to be 'superimposed' onto the *purely possibilistic* prior $\pi^{\theta}(\boldsymbol{\theta})$ that has to be updated. In more detail, let $\widetilde{p}^{\theta}(\boldsymbol{\theta}|\boldsymbol{\varphi})$ be the fictitious (prior) FPDF of $\boldsymbol{\theta}$, constructed by the superimposition of an (arbitrarily selected) fictitious PDF $p^{\theta}(\boldsymbol{\theta}|\boldsymbol{\varphi})$ and a vector of parameters $\boldsymbol{\varphi}$ described by a (properly selected) possibility distribution $\pi^{\varphi}(\boldsymbol{\varphi})$ (it is straightforward to notice that in the case parameters $\{\theta_m: m = 1, 2, \ldots, P\}$ are independent, then $p^{\theta}(\boldsymbol{\theta}|\boldsymbol{\varphi}) = \prod_{m=1}^{P} p^{\theta_m}(\theta_m|\boldsymbol{\varphi}_m)$, with $\boldsymbol{\varphi}_m = [\phi_{m1}, \phi_{m2}, \ldots, \phi_{mN_m}]$, $m = 1, 2, \ldots, P$; also, $\pi^{\varphi}(\boldsymbol{\varphi})$ is expressed as $\pi^{\varphi}(\boldsymbol{\varphi}) = \{\pi^{\varphi_1}(\boldsymbol{\varphi}_1), \pi^{\varphi_2}(\boldsymbol{\varphi}_2), ..., \pi^{\varphi_m}(\boldsymbol{\varphi}_m), ..., \pi^{\varphi_P}(\boldsymbol{\varphi}_P)\}$. In addition, if also the possibilistic parameters $\phi_{m1}, \phi_{m2}, \ldots, \phi_{mn_m}$ of the fictitious PDF $p^{\theta_m}(\theta_m|\boldsymbol{\varphi}_m)$, $m = 1, 2, \ldots, P$, are independent, then $\pi^{\varphi_m}(\boldsymbol{\varphi}_m) = \{\pi^{\varphi_{m1}}(\phi_{m1}), \pi^{\varphi_{m2}}(\phi_{m2}), ..., \pi^{\varphi_{mN_m}}(\phi_{mN_m})\}$). As before, let $Y \sim p^Y(y|\boldsymbol{\theta}) = Gum(\gamma, \delta) = p^Y(y|\gamma, \delta)$, with $\boldsymbol{\theta} = [\theta_1, \theta_2] = [\gamma, \delta]$ epistemically-uncertain. For simplicity, we consider $\theta_1 = \gamma$ and $\theta_2 = \delta$ independent, such that $\pi^{\theta}(\boldsymbol{\theta}) = \{\pi^{\gamma}(\gamma), \pi^{\delta}(\delta)\}$ and that

the (fictitious) PDF $p^{\theta}(\boldsymbol{\theta}|\boldsymbol{\varphi})$ can be written as $\prod_{m=1}^{P} p^{\theta_m}(\theta_m|\boldsymbol{\varphi}_m) = p^{\theta_1}(\theta_1|\boldsymbol{\varphi}_1)\cdot p^{\theta_2}(\theta_2|\boldsymbol{\varphi}_2) = p^{\gamma}(\gamma|\boldsymbol{\varphi}_\gamma)\cdot p^{\delta}(\delta|\boldsymbol{\varphi}_\delta)$. By hypothesis, the analyst *arbitrarily* selects Normal distributions to represent $p^{\gamma}(\gamma|\boldsymbol{\varphi}_\gamma)$ and $p^{\delta}(\delta|\boldsymbol{\varphi}_\delta)$, i.e., $p^{\gamma}(\gamma|\boldsymbol{\varphi}_\gamma) = N(\boldsymbol{\varphi}_\gamma) = N(\mu_\gamma,\sigma_\gamma)$ and $p^{\delta}(\delta|\boldsymbol{\varphi}_\delta) = N(\boldsymbol{\varphi}_\delta) = N(\mu_\delta,\sigma_\delta)$, respectively (Normal and Uniform distributions are indicated as good choices for $p^{\theta}(\boldsymbol{\theta}|\boldsymbol{\varphi})$ by [Beer, 2009a; Stein and Beer, 2011; Stein et al., 2013]). Then, parameters $\boldsymbol{\varphi}_\gamma = [\mu_\gamma,\sigma_\gamma]$ and $\boldsymbol{\varphi}_\delta = [\mu_\delta,\sigma_\delta]$ are represented by the possibility distributions $\pi^{\boldsymbol{\varphi}_\gamma} = \{\pi^{\mu_\gamma}(\mu_\gamma),\pi^{\sigma_\gamma}(\sigma_\gamma)\}$ and $\pi^{\boldsymbol{\varphi}_\delta} = \{\pi^{\mu_\delta}(\mu_\delta),\pi^{\sigma_\delta}(\sigma_\delta)\}$ (by so doing, the fictitious prior FPDF $\tilde{p}^{\theta}(\boldsymbol{\theta}|\boldsymbol{\varphi}) = \tilde{p}^{\theta_1}(\theta_1|\boldsymbol{\varphi}_1)\cdot\tilde{p}^{\theta_2}(\theta_2|\boldsymbol{\varphi}_2) = \tilde{p}^{\gamma}(\gamma|\boldsymbol{\varphi}_\gamma)\cdot\tilde{p}^{\delta}(\delta|\boldsymbol{\varphi}_\delta)$ is constructed). These possibility distributions should be *properly* selected by the analyst so as to reflect as closely as possible the structure of the '*real*' prior possibility distribution $\pi^{\theta}(\boldsymbol{\theta}) = \{\pi^{\gamma}(\gamma),\pi^{\delta}(\delta)\}$ that has to be updated: for example, $\pi^{\boldsymbol{\varphi}_\gamma}$ and $\pi^{\boldsymbol{\varphi}_\delta}$ could be identified by 'imposing' that the *expected value* of the FPDF $\tilde{p}^{\theta}(\boldsymbol{\theta}|\boldsymbol{\varphi})$ corresponds to the *real* prior possibility distribution $\pi^{\theta}(\boldsymbol{\theta})$, i.e., $E_{\boldsymbol{\varphi}}[\tilde{p}^{\theta}(\boldsymbol{\theta}|\boldsymbol{\varphi})] = \pi^{\theta}(\boldsymbol{\theta})$ or, in this case, $E_{\boldsymbol{\varphi}_\gamma}[\tilde{p}^{\gamma}(\gamma|\boldsymbol{\varphi}_\gamma)] = \pi^{\gamma}(\gamma)$ and $E_{\boldsymbol{\varphi}_\delta}[\tilde{p}^{\delta}(\delta|\boldsymbol{\varphi}_\delta)] = \pi^{\delta}(\delta)$.

In extreme synthesis, the method relies on the hybrid combination of (i) Fuzzy Interval Analysis (FIA) to process the uncertainty described by possibility distributions and (ii) *repeated* Bayesian updating of the uncertainty represented by probability distributions [Beer, 2009a; Stein and Beer, 2011; Stein et al., 2013]. In more details, the algorithm proceeds as follows:

1. set $\alpha = 0$;
2. select the α-cut $A_\alpha^{\boldsymbol{\varphi}}$ of the possibility distribution $\pi^{\boldsymbol{\varphi}}(\boldsymbol{\varphi})$ of vector $\boldsymbol{\varphi}$ of the parameters of the (fictitious) prior PDF $p^{\theta}(\boldsymbol{\theta}|\boldsymbol{\varphi})$;
3. letting the parameter vector $\boldsymbol{\varphi}$ range within the corresponding α-cut $A_\alpha^{\boldsymbol{\varphi}}$ identified at step 2. above, generate a family of (fictitious) prior PDFs $\left\{p^{\theta}(\boldsymbol{\theta}|\boldsymbol{\varphi})\right\}_\alpha = \left\{p^{\theta}(\boldsymbol{\theta}|\boldsymbol{\varphi}):\boldsymbol{\varphi}\in A_\alpha^{\boldsymbol{\varphi}}\right\}$. This is

empirically done by (i) randomly or deterministically selecting a *finite* number T (e.g., T = 100 in this paper) of parameter vectors $\boldsymbol{\varphi}_{l,\alpha}$, l = 1, 2, …, T, in $A_\alpha^{\boldsymbol{\varphi}}$ and (ii) evaluating $p^{\boldsymbol{\theta}}(\boldsymbol{\theta}\,|\,\boldsymbol{\varphi})$ in correspondence of these vectors, i.e., $\left\{p^{\boldsymbol{\theta}}(\boldsymbol{\theta}\,|\,\boldsymbol{\varphi})\right\}_\alpha \approx \left\{p^{\boldsymbol{\theta}}(\boldsymbol{\theta}\,|\,\boldsymbol{\varphi}_{l,\alpha}):l=1,2,...,T\right\}$;

4. apply the classical probabilistic Bayes theorem to *each* (fictitious) prior PDF $p^{\boldsymbol{\theta}}(\boldsymbol{\theta}\,|\,\boldsymbol{\varphi}_{l,\alpha})$ generated at step 3. above to get the corresponding posterior PDF $p_{l,\alpha}^{\boldsymbol{\theta}}(\boldsymbol{\theta}\,|\,\boldsymbol{y})$:

$$p_{l,\alpha}^{\boldsymbol{\theta}}(\boldsymbol{\theta}\,|\,\boldsymbol{y})=\frac{L^{\boldsymbol{\theta}}(\boldsymbol{\theta}\,|\,\boldsymbol{y})\cdot p(\boldsymbol{\theta}\,|\,\boldsymbol{\varphi}_{l,\alpha})}{\int_{\boldsymbol{\theta}} L^{\boldsymbol{\theta}}(\boldsymbol{\theta}\,|\,\boldsymbol{y})\cdot p(\boldsymbol{\theta}\,|\,\boldsymbol{\varphi}_{l,\alpha})d\boldsymbol{\theta}},\ l = 1, 2, \ldots, T \tag{2}$$

This is equivalent to generating a family $\left\{p^{\boldsymbol{\theta}}(\boldsymbol{\theta}\,|\,\boldsymbol{y})\right\}_\alpha$ of posterior PDFs for $\boldsymbol{\theta}$, i.e., $\left\{p^{\boldsymbol{\theta}}(\boldsymbol{\theta}\,|\,\boldsymbol{y})\right\}_\alpha \approx \{p_{l,\alpha}^{\boldsymbol{\theta}}(\boldsymbol{\theta}\,|\,\boldsymbol{y}):l=1,2,...,T\}$;

5. calculate the *expected value* of *each* posterior PDF $p_{l,\alpha}^{\boldsymbol{\theta}}(\boldsymbol{\theta}\,|\,\boldsymbol{y})$ generated at step 4. above to obtain a *point estimate* $\boldsymbol{\theta}_{l,\alpha}\,|\,\boldsymbol{y}$ for the epistemically-uncertain parameter vector $\boldsymbol{\theta}$ [Stein et al., 2013]:

$$\boldsymbol{\theta}_{l,\alpha}\,|\,\boldsymbol{y} = E_{\boldsymbol{\theta}}[\,p_{l,\alpha}^{\boldsymbol{\theta}}(\boldsymbol{\theta}\,|\,\boldsymbol{y})],\ l = 1, 2, \ldots, T \tag{3}$$

6. take the hull enveloping the T point estimates $\boldsymbol{\theta}_{l,\alpha}\,|\,\boldsymbol{y}$, l = 1, 2, …, T, as the (P-dimensional) α-cut $A_\alpha^{\boldsymbol{\theta}|\boldsymbol{y}}$ of the (joint P-dimensional) posterior possibility distribution $\pi^{\boldsymbol{\theta}}(\boldsymbol{\theta}\,|\,\boldsymbol{y})$ of $\boldsymbol{\theta}$. By way of example and for illustration purposes, Fig. 2 shows the identification of a (two-dimensional) α-cut $A_\alpha^{\boldsymbol{\theta}|\boldsymbol{y}} = A_\alpha^{\theta_1,\theta_2|\boldsymbol{y}}$ (solid line) as the contour enclosing T = 20 point estimates $\boldsymbol{\theta}_{l,\alpha}\,|\,\boldsymbol{y}$, l = 1, 2, …, 20 (dots), in the simplified case of P = 2 parameters θ_1 and θ_2;

7. if $\alpha < 1$, then set $\alpha = \alpha + \Delta\alpha$ (e.g., $\Delta\alpha$ = 0.05 in this paper) and return to step 2. above; otherwise, stop the algorithm.

The (joint *P*-dimensional) posterior possibility distribution $\pi^{\boldsymbol{\theta}}(\boldsymbol{\theta} \mid \boldsymbol{y})$ for the epistemically-uncertain parameter vector $\boldsymbol{\theta}$ is empirically constructed as the (discrete) collection of the α-cuts $A_{\alpha}^{\boldsymbol{\theta}|\boldsymbol{y}}$, α = 0, 0.05, …, 0.95, 1, found at step 6. above.

It is worth noting that the application of both approaches A and B always produces a *joint P*-dimensional posterior possibility distribution $\pi^{\boldsymbol{\theta}}(\boldsymbol{\theta} \mid \boldsymbol{y})$ (*whatever* the state of dependence between the priors), characterized by *P*-dimensional α-cuts $A_{\alpha}^{\boldsymbol{\theta}|\boldsymbol{y}}$, with $0 < \alpha < 1$: as a consequence, there is an *interactive dependence* between the values that parameters $\{\theta_m: m = 1, 2, \ldots, P\}$ can take when ranging within a given α-cut $A_{\alpha}^{\boldsymbol{\theta}|\boldsymbol{y}}$: for example, in Fig. 2 it is impossible that parameter θ_1 takes on low values *and* parameter θ_2 takes on high values at the same time.

From $\pi^{\boldsymbol{\theta}}(\boldsymbol{\theta} \mid \boldsymbol{y})$ it is straightforward to obtain the *marginal* posterior possibility distribution $\pi^{\theta_m}(\theta_m \mid \boldsymbol{y})$ for each parameter θ_m as $\pi^{\theta_m}(\theta_m \mid \boldsymbol{y}) = \max_{\theta_j \in \Re,\, j \neq m} \{\pi^{\boldsymbol{\theta}}(\boldsymbol{\theta} \mid \boldsymbol{y})\}$, $\forall \theta_m \in \Re$, m = 1, 2, …, *P* [Baudrit et al., 2006]: $\pi^{\theta_m}(\theta_m \mid \boldsymbol{y})$ is the projection of $\pi^{\boldsymbol{\theta}}(\boldsymbol{\theta} \mid \boldsymbol{y})$ onto the *m*-th axis. The (one-dimensional) α-cut $A_{\alpha}^{\theta_m|\boldsymbol{y}} = [\underline{\theta}_{m,\alpha} \mid y, \overline{\theta}_{m,\alpha} \mid y]$ of the marginal possibility distribution $\pi^{\theta_m}(\theta_m \mid \boldsymbol{y})$ is then related to the (*P*-dimensional) α-cut $A_{\alpha}^{\boldsymbol{\theta}|\boldsymbol{y}}$ of the joint possibility distribution $\pi^{\boldsymbol{\theta}}(\boldsymbol{\theta} \mid \boldsymbol{y})$ by the following straightforward relation, i.e., $A_{\alpha}^{\theta_m|\boldsymbol{y}} = [\underline{\theta}_{m,\alpha} \mid y, \overline{\theta}_{m,\alpha} \mid y] = [\min_{\boldsymbol{\theta} \in A_{\alpha}^{\boldsymbol{\theta}|\boldsymbol{y}}} \{\theta_m\}, \max_{\boldsymbol{\theta} \in A_{\alpha}^{\boldsymbol{\theta}|\boldsymbol{y}}} \{\theta_m\}]$. In this view, notice that the use of the *P*-dimensional α-cut $A_{\alpha,Cart}^{\boldsymbol{\theta}|\boldsymbol{y}}$ constructed by the Cartesian product of the (one-dimensional) α-cuts $A_{\alpha}^{\theta_m|\boldsymbol{y}}$ of the marginal distributions, m = 1, 2, …, *P* (i.e., $A_{\alpha,Cart}^{\boldsymbol{\theta}|\boldsymbol{y}} = A_{\alpha}^{\theta_1|\boldsymbol{y}}$ x $A_{\alpha}^{\theta_2|\boldsymbol{y}}$ x … x $A_{\alpha}^{\theta_m|\boldsymbol{y}}$ x … x $A_{\alpha}^{\theta_P|\boldsymbol{y}}$) would (incorrectly) imply *independence* between the posterior estimates of the parameters $\{\theta_m: m = 1, 2, \ldots, P\}$; however, since $A_{\alpha,Cart}^{\boldsymbol{\theta}|\boldsymbol{y}}$ completely *contains* $A_{\alpha}^{\boldsymbol{\theta}|\boldsymbol{y}}$ (i.e., by definition $A_{\alpha}^{\boldsymbol{\theta}|\boldsymbol{y}} \subset A_{\alpha,Cart}^{\boldsymbol{\theta}|\boldsymbol{y}}$), then conservatism would be still guaranteed [Stein et al., 2013]. For illustration purposes and with reference to the example above, Fig. 2 shows also the two-

dimensional α-cut $A_{\alpha,Cart}^{\theta_1,\theta_2|y}$ (dashed line) generated by the Cartesian product of the (one-dimensional) α-cuts $A_{\alpha}^{\theta_1|y}$ and $A_{\alpha}^{\theta_2|y}$, i.e., $A_{\alpha,Cart}^{\theta_1,\theta_2|y} = A_{\alpha}^{\theta_1|y}$ x $A_{\alpha}^{\theta_2|y}$.

Fig. 2

The characteristics of the two approaches are summarized in Table 1. Notice that both methods relies on *arbitrary assumptions* about either the prior or the likelihood functions: in the purely possibilistic approach (A) the "original" possibilistic prior is employed, but a possibilistic likelihood function has to be constructed (e.g., by probability-possibility transformations, directly from rough experimental data and/or by resorting to the guidelines provided by [Dubois et al., 1997; Denoeux, 2014]); instead, in the hybrid method (B) the original probabilistic likelihood function is used, but a "fictitious" prior Fuzzy Probability Distribution Function needs to be identified by superimposing an *arbitrarily selected* probabilistic PDF onto the "original" possibilistic prior that has to be updated.

Table 1.

A final consideration is in order with respect to the two approaches here outlined. In the hybrid probabilistic-possibilistic framework of interest to the present paper, the knowledge a priori available on the parameters $\boldsymbol{\theta} = [\theta_1, \theta_2, ..., \theta_m, ..., \theta_P]$ of a given (aleatory) probability model $p^Y(y|\boldsymbol{\theta})$ is described by the *prior* possibility distribution function $\pi^{\theta}(\boldsymbol{\theta})$. As detailed in the previous Section, the possibilistic approach is particularly suitable to address those situations where the information a priori available on $\boldsymbol{\theta}$ is scarce and imprecise, i.e., *not sufficient* for assigning a *single specific* probability distribution to (describe the epistemic uncertainty in) $\boldsymbol{\theta}$. Actually, the *possibilistic* function $\pi^{\theta}(\boldsymbol{\theta})$ is in practice "equivalent" to the *family* of *all* those *probability distributions* (of possibly *different shapes*) that are *coherent* with the scarce information available on $\boldsymbol{\theta}$. On the other hand, it is worth mentioning that in a classical *purely probabilistic* framework, imprecision in prior information about $\boldsymbol{\theta}$ can be also accounted for by means of the so-called

Hierarchical Bayes approach. Hierarchical Bayes is so-named because it utilizes *hierarchical* or *multistage* prior distributions [Gelman, 2006; Gelman et al., 2008; Congdon, 2010; Kelly and Smith, 2011; Chung et al., 2015; Shirley and Gelman, 2015]. To develop a hierarchical model for $\boldsymbol{\theta}$, we need to specify a *first-stage prior* (say, $p^{\theta}(\boldsymbol{\theta}|\boldsymbol{\varphi})$), which is often of a *particular functional form*, often a conjugate prior. However, analysts find it difficult to express their incertitude numerically at all, much less as particular probability distributions. Thus, a higher-dimensional model is defined to represent such (epistemic) uncertainty: in particular, we need to specify an "additional" prior distribution (say, $p^{\varphi}(\boldsymbol{\varphi}|\boldsymbol{\omega})$) on the first-stage parameters $\boldsymbol{\varphi}$. Distribution $p^{\varphi}(\boldsymbol{\varphi}|\boldsymbol{\omega})$ is called the *second-stage prior*, or *hyper-prior*. This way of proceeding amounts to generating a 'parametric' family $\left\{p^{\theta}(\boldsymbol{\theta}|\boldsymbol{\varphi})\right\}$ of (first-stage prior) probability distributions for $\boldsymbol{\theta}$, all obtained in correspondence of different possible values of the first-stage parameters $\boldsymbol{\varphi}$ (described by hyper-prior $p^{\varphi}(\boldsymbol{\varphi}|\boldsymbol{\omega})$). This method has been investigated in the field of social and behavioral sciences with the *main* aim of treating hierarchical data with different levels of variables in the same statistical model. For example, the hierarchical data for sociological survey analysis include measurements from individuals with different historical, geographic, or economic variables. To this end, the hierarchical modeling was proposed to account for the different grouping or times at which data are measured [Gill, 2002]. In addition, a common application of hierarchical Bayes analysis in the Probabilistic Risk Assessment (PRA) of nuclear power plants has been as a model of variability among data sources, for example variability in Emergency Diesel Generator (EDG) performance across different plants, or across time [Siu and Kelly, 1998; Atwood et al., 2003; Kelly and Smith, 2009]. Finally, similar analogy can be made for the collected measurements from a structure under different ambient and environmental conditions. This framework has been recently implemented for uncertainty quantification applications in structural dynamics [Behmanesh et al., 2015; Ballesteros et al., 2014]. Although hierarchical Bayes can address the issue of imprecise prior information by means of multi-level models, the following conceptual and practical considerations should be made

about its applicability: (i) in principle, one could define even higher dimensional models to represent the uncertainty. Some analysts have attempted three- and even four-level models [Jaworska and Aldenberg, 2000], but this hardly seems a workable solution when the complexity of the analysis is the primary problem; (ii) in addition, "if concern about the uncertainty in parameters can in principle drive the analysis to a higher level, one could fall into an insoluble infinite cascade" [Ferson, 2005]; (iii) finally, both first- and second-stage priors, $p^{\theta}(\boldsymbol{\theta}|\boldsymbol{\varphi})$ and $p^{\varphi}(\boldsymbol{\varphi}|\boldsymbol{\omega})$, respectively, are often of a *particular* functional form, which forces the analyst to make "overly-optimistic" and excessively precise statements about parameter uncertainties, even when the information and data available are scarce and/or vague, i.e., *not sufficient* for assigning a *single specific* probability distribution. In such cases, non-probabilistic methods (e.g., intervals, possibility distributions or Dempster-Shafer structures from evidence theory) have been shown to provide reliable and robust results [Ferson et al., 2003; Beer et al., 2013a]. On the basis of considerations (i)-(iii) above and given that the objective of the paper is the comparison of methods for the update of *possibilistic* parameters (of aleatory probability distributions), Hierarchical Bayes approaches are not considered in the present work.

Case study: flood protection risk-based design

The case study deals with the design of a protection dike in a residential area closely located to a river with potential risk of floods. Two issues of concern are: (i) high construction and annual maintenance costs of the dike; (ii) uncertainty in the natural phenomenon of flooding. Then, the different design options must be evaluated within a flooding risk analysis framework accounting for uncertainty.

The model

The maximal water level of the river (i.e., the output variable of the model, Z_c) is given as a function of several (and some uncertain) parameters (i.e., the inputs to the model) [Limbourg and de Rocquigny, 2010]:

$$Z_c = f(Q, Z_m, Z_v, K_s, B, L) = f(Y_1, Y_2, Y_3, Y_4, Y_5, Y_6) \quad (4)$$

where: $Y_1 = Q$ is the yearly maximal water discharge [m^3/s]; $Y_2 = Z_m$ and $Y_3 = Z_v$ are the riverbed levels [m asl] at the upstream and downstream parts of the river under investigation, respectively; $Y_4 = K_s$ is the Strickler friction coefficient; $Y_5 = B$ and $Y_6 = L$ are the width and length of the river part [m], respectively. Quantities $Y_5 = B$ (= 300m) and $Y_6 = L$ (= 5000m) are *constant* parameters, whereas quantities $Y_1 = Q$, $Y_2 = Z_m$, $Y_3 = Z_v$, $Y_4 = K_s$ are *uncertain* variables.

The input variables: physical description and representation of the associated uncertainty

The $n = 4$ input variables Y_i, $i = 1, 2, 3, 4$, are affected by aleatory and epistemic uncertainties. The aleatory part of the uncertainty is described by probability distributions of defined shape. The parameters of the probability distributions describing the aleatory uncertainty are themselves affected by epistemic uncertainty and represented in terms of possibility distributions.

The yearly maximal water flow, $Y_1 = Q$

The aleatory uncertainty in the yearly maximal water flow $Y_1 = Q$ is well described by a Gumbel probability distribution $p^Q(q|\gamma,\delta) = Gum(\gamma,\delta) = \frac{1}{\delta}\exp\left[-\exp\left(\frac{\gamma-q}{\delta}\right)\right]\exp\left[\frac{\gamma-q}{\delta}\right]$ [Limbourg and de Rocquigny, 2010]. The extreme physical bounds on variable Q are [Limbourg and de Rocquigny, 2010]:

- $Q_{\min} = 10\,m^3/s$, which is a typical drought flow level (irrelevant within a flood study);
- $Q_{\max}$ = 10000 m^3/s, which is three times larger than the maximal flood ever occurred.

When Approach A is used, the prior possibility distributions $\pi^\gamma(\gamma)$ and $\pi^\delta(\delta)$ for the epistemically-uncertain parameters γ and δ are subjectively chosen as triangular functions $TR(a_\gamma, c_\gamma, b_\gamma)$ and $TR(a_\delta, c_\delta, b_\delta)$, respectively, with cores (i.e., preferred or most likely values) c_γ = 955m^3/s and c_δ = 600m^3/s, and supports $[a_\gamma, b_\gamma]$ = [869, 1157] m^3/s and $[a_\delta, b_\delta]$ = [455, 660] m^3/s,

421 respectively. When Approach B is employed, the fictitious prior FPDFs $\tilde{p}^{\gamma}(\gamma)$ and $\tilde{p}^{\delta}(\delta)$ are
422 subjectively chosen as normal probability distributions $p^{\gamma}(\gamma) = N(\mu_{\gamma}, \sigma_{\gamma})$ and $p^{\delta}(\delta) =$
423 $N(\mu_{\delta}, \sigma_{\delta})$, respectively, with possibilistic parameters $\mu_{\gamma}, \sigma_{\gamma}, \mu_{\delta}$ and σ_{δ}: the characteristics of the
424 corresponding triangular possibility distributions $\pi^{\mu_{\gamma}}(\mu_{\gamma})$, $\pi^{\sigma_{\gamma}}(\sigma_{\gamma})$, $\pi^{\mu_{\delta}}(\mu_{\delta})$ and $\pi^{\sigma_{\delta}}(\sigma_{\delta})$ are
425 summarized in Table 2 for brevity. The Bayesian update of these uncertainty representations (based
426 on prior subjective knowledge) is realized with the aid of a vector $\boldsymbol{y}_1 = [y_{11}, y_{12}, \ldots, y_{1k}, \ldots, y_{1D_1}]$ of
427 $D_1 = 149$ (independent and identically distributed – iid) values of the annual maximal flow of the
428 river, i.e., $\boldsymbol{y}_1 = \boldsymbol{q} = [q_1, q_2, \ldots, q_k, \ldots, q_{149}]$. The point estimates for γ and δ obtained by the
429 classical, purely probabilistic Maximum Likelihood Estimation (MLE) method are $\hat{\gamma}^{MLE} = 1013.21$
430 m^3/s and $\hat{\delta}^{MLE} = 558.21\ m^3/s$, respectively.

431 **The upstream riverbed level, $Y_2 = Z_m$**

432 The aleatory part of the uncertainty in the upstream riverbed level $Y_2 = Z_m$ is represented by a
433 normal distribution, i.e., $Z_m \sim p^{Z_m}(z_m \mid \mu_{Zm}, \sigma_{Zm}) = N(\mu_{Zm}, \sigma_{Zm})$ [Limbourg and de Rocquigny,
434 2010]. This distribution is truncated at the minimum and maximum physical bounds on Z_m, i.e.,
435 $Z_{m,\min} = 53.5\ m$ (given by plausible lower geomorphologic limits to erosion) and $Z_{m,\max} = 57\ m$
436 (given by plausible upper geomorphologic limits to sedimentation), respectively. In Approach A,
437 the prior possibility distributions $\pi^{\mu_{Zm}}(\mu_{Zm})$ and $\pi^{\sigma_{Zm}}(\sigma_{Zm})$ for μ_{Zm} and σ_{Zm} are chosen as
438 triangular functions (Table 2). In Approach B, the fictitious prior FPDFs $\tilde{p}^{\mu_{Zm}}(\mu_{Zm})$ and $\tilde{p}^{\sigma_{Zm}}(\sigma_{Zm})$
439 are chosen as normal PDFs $p^{\mu_{Zm}}(\mu_{Zm}) = N(\mu_{\mu_{Zm}}, \sigma_{\mu_{Zm}})$ and $p^{\sigma_{Zm}}(\sigma_{Zm}) = N(\mu_{\sigma_{Zm}}, \sigma_{\sigma_{Zm}})$ with
440 parameters described by the triangular possibilistic functions $\pi^{\mu_{\mu_{Zm}}}(\mu_{\mu_{Zm}})$, $\pi^{\sigma_{\mu_{Zm}}}(\sigma_{\mu_{Zm}})$,
441 $\pi^{\mu_{\sigma_{Zm}}}(\mu_{\sigma_{Zm}})$ and $\pi^{\sigma_{\sigma_{Zm}}}(\sigma_{\sigma_{Zm}})$ reported in Table 2. The Bayesian update of these uncertainty
442 representations is carried out using a vector $\boldsymbol{y}_2 = [y_{21}, y_{22}, \ldots, y_{2k}, \ldots, y_{2D_2}]$ of $D_2 = 29$ (iid) values

of the upstream riverbed level, i.e., $\boldsymbol{y}_2 = \boldsymbol{z_m} = [z_{m,1}, z_{m,2}, \ldots, z_{m,k}, \ldots, z_{m,29}]$. The point estimates $\hat{\mu}_{Zm}^{MLE}$ and $\hat{\sigma}_{Zm}^{MLE}$ for μ_{Zm} and σ_{Zm} obtained by the MLE method are $\hat{\mu}_{Zm}^{MLE}$ = 50.19 m and $\hat{\sigma}_{Zm}^{MLE}$ = 0.38 m, respectively.

The downstream riverbed level, $Y_3 = Z_v$

As for $Y_2 = Z_m$, the aleatory part of the uncertainty in the downstream riverbed level $Y_3 = Z_v$ is represented by a normal distribution, i.e., $Z_v \sim p^{Z_v}(z_v \mid \mu_{Zv}, \sigma_{Zv}) = N(\mu_{Zv}, \sigma_{Zv})$, truncated at $Z_{v,\min} = 48\,\text{m}$ and $Z_{v,\max} = 51\,\text{m}$. As before, the prior possibility distributions $\pi^{\mu_{Zv}}(\mu_{Zv})$ and $\pi^{\sigma_{Zv}}(\sigma_{Zv})$ used in Approach A are triangular functions (Table 2) and the FPDFs $\tilde{p}^{\mu_{Zv}}(\mu_{Zv})$ and $\tilde{p}^{\sigma_{Zv}}(\sigma_{Zv})$ employed in Approach B are normal PDFs $p^{\mu_{Zv}}(\mu_{Zv}) = N(\mu_{\mu_{Zm}}, \sigma_{\mu_{Zm}})$ and $p^{\sigma_{Zv}}(\sigma_{Zv}) = N(\mu_{\sigma_{Zv}}, \sigma_{\sigma_{Zv}})$ with parameters described by the triangular possibility distributions $\pi^{\mu_{\mu_{Zv}}}(\mu_{\mu_{Zv}})$, $\pi^{\sigma_{\mu_{Zv}}}(\sigma_{\mu_{Zv}})$, $\pi^{\mu_{\sigma_{Zv}}}(\mu_{\sigma_{Zv}})$ and $\pi^{\sigma_{\sigma_{Zv}}}(\sigma_{\sigma_{Zv}})$ of Table 2. These representations are updated by means of a vector $\boldsymbol{y}_3 = [y_{31}, y_{32}, \ldots, y_{3k}, \ldots, y_{3D_3}]$ of $D_3 = 29$ (iid) values of the downstream riverbed level, i.e., $\boldsymbol{y}_3 = \boldsymbol{z_v} = [z_{v,1}, z_{v,2}, \ldots, z_{v,k}, \ldots, z_{v,29}]$. The MLE estimates of the parameters are $\hat{\mu}_{Zv}^{MLE}$ = 55.03 m and $\hat{\sigma}_{Zv}^{MLE}$ = 0.45 m, respectively.

The Strickler friction coefficient, $Y_4 = K_s$

The Strickler friction coefficient $Y_4 = K_s$ is the most critical source of uncertainty because it is usually a simplification of a complex hydraulic model. The absolute physical limits of K_s are 5 and 60, respectively [Limbourg and de Rocquigny, 2010]:

- $K_s < 5$ corresponds to an "extremely sinuous shape of the canal, with large dents and strong vegetation";
- $K_s = 60$ corresponds to a "canal with smoothest earth surface, rectilinear, without any vegetation".

The friction coefficient K_s is affected by random events modifying the river status (e.g., erosion, sedimentation, …): the corresponding variability is typically described by a normal distribution, i.e., $K_s \sim p^{K_s}(k_s \mid \mu_{Ks}, \sigma_{Ks}) = N(\mu_{Ks}, \sigma_{Ks})$ [Limbourg and de Rocquigny, 2010]. However, the parameters of this normal distribution are difficult to estimate because data can only be obtained through "indirect calibration characterized by significant uncertainty" [Limbourg and de Rocquigny, 2010]: the uncertainty in these parameters is described by possibility distributions. Details about the possibilistic functions $\pi^{\mu_{Ks}}(\mu_{K_s})$ and $\pi^{\sigma_{Ks}}(\sigma_{Ks})$ and the FPDFs $\tilde{p}^{\mu_{Ks}}(\mu_{Ks})$ and $\tilde{p}^{\sigma_{Ks}}(\sigma_{Ks})$ used to represent the a priori knowledge on μ_{Ks} and σ_{Ks} in Approaches A and B, respectively, are reported in Table 2. The Bayesian revision of these a priori representations is performed by means of a vector $\boldsymbol{y}_4 = [y_{41}, y_{42}, \ldots, y_{4k}, \ldots, y_{4D_4}]$ of $D_4 = 5$ (iid) values of the Strickler friction coefficient, i.e., $\boldsymbol{y}_4 = \boldsymbol{k_s} = [k_{s,1}, k_{s,2}, \ldots, k_{s,k}, \ldots, k_{s,5}]$. The MLE estimates of the parameters are $\hat{\mu}_{Ks}^{MLE} = 27.8$ and $\hat{\sigma}_{Ks}^{MLE} = 5.26$, respectively.

Table 2.

Results

In order to simplify the notation, in what follows let θ be one of the uncertain parameters of the PDFs of $Y_1 = Q$, $Y_2 = Z_m$, $Y_3 = Z_v$ and $Y_4 = K_s$, i.e., $\theta = \gamma$, δ, μ_{Zm}, σ_{Zm}, μ_{Zv}, σ_{Zv}, μ_{K_s} or σ_{Ks}. Fig. 3 illustrates the possibility distributions of the epistemically-uncertain parameters of the aleatory PDFs $p^{Q}(q|\gamma,\delta)$ (top row), $p^{Z_m}(z_m \mid \mu_{Zm}, \sigma_{Zm})$ (middle-top row), $p^{Z_v}(z_v \mid \mu_{Zv}, \sigma_{Zv})$ (middle-bottom row) and $p^{K_s}(k_s \mid \mu_{Ks}, \sigma_{Ks})$ (bottom row) of the uncertain input variables $Y_1 = Q$, $Y_2 = Z_m$, $Y_3 = Z_v$ and $Y_4 = K_s$, respectively, of the model of the previous Section: in particular, the prior possibility distributions $\pi^{\theta}(\theta)$ (= $\pi^{\gamma}(\gamma)$, $\pi^{\delta}(\delta)$, $\pi^{\mu_{Zm}}(\mu_{Zm})$, $\pi^{\sigma_{Zm}}(\sigma_{Zm})$, $\pi^{\mu_{Zv}}(\mu_{Zv})$, $\pi^{\sigma_{Zv}}(\sigma_{Zv})$, $\pi^{\mu_{K_s}}(\mu_{K_s})$ and $\pi^{\sigma_{Ks}}(\sigma_{Ks})$) are shown as solid lines, whereas the *marginal* posterior possibility distributions

$\pi^{\theta}(\theta|\boldsymbol{y})$ (= $\pi^{\gamma}(\gamma|\boldsymbol{q})$, $\pi^{\delta}(\delta|\boldsymbol{q})$, $\pi^{\mu_{Zm}}(\mu_{Zm}|\boldsymbol{z_m})$, $\pi^{\sigma_{Zm}}(\sigma_{Zm}|\boldsymbol{z_m})$, $\pi^{\mu_{Zv}}(\mu_{Zv}|\boldsymbol{z_v})$, $\pi^{\sigma_{Zv}}(\sigma_{Zv}|\boldsymbol{z_v})$, $\pi^{\mu_{Ks}}(\mu_{K_s}|\boldsymbol{k_s})$ and $\pi^{\sigma_{Ks}}(\sigma_{Ks}|\boldsymbol{k_s})$) obtained by Approaches A and B using D_1 = 149, D_2 = 29, D_3 = 29 and D_4 = 5 pieces of data are shown in dashed and dot-dashed lines, respectively; the point estimates $\hat{\theta}^{MLE}$ (= $\hat{\gamma}^{MLE}$, $\hat{\delta}^{MLE}$, $\hat{\mu}_{Zm}^{MLE}$, $\hat{\sigma}_{Zm}^{MLE}$, $\hat{\mu}_{Zv}^{MLE}$, $\hat{\sigma}_{Zv}^{MLE}$, $\hat{\mu}_{K_s}^{MLE}$ and $\hat{\sigma}_{Ks}^{MLE}$) produced by the classical MLE method are also shown for comparison (dots).

Fig. 3.

From a mere visual and qualitative inspection of Fig. 3 it can be seen that both approaches are suitable for revising the prior possibility distributions (based on a priori purely subjective knowledge) by means of empirical data. In particular, it is evident that: (i) the most likely (i.e., preferred) values c_θ of the epistemically-uncertain parameters (i.e., those values in correspondence of which the possibility function equals 1) are moved towards the MLE estimates $\hat{\theta}^{MLE}$ in all the cases considered; (ii) the area S_θ underlying the corresponding possibility distributions is significantly reduced: noting that this area is related to the imprecision in the knowledge of the possibilistic parameter (i.e., the larger the area, the higher the imprecision), it can be concluded that both approaches succeed in reducing the epistemic uncertainty. With respect to that, Table 3 reports the most likely values c_θ and the areas S_θ underlying the (marginal) possibility distributions of the uncertain parameter θ (= γ, δ, μ_{Zm}, σ_{Zm}, μ_{Zv}, σ_{Zv}, μ_{K_s} and σ_{Ks}) before and after the Bayesian update performed by Approaches A and B; the point estimates $\hat{\theta}^{MLE}$ obtained by the classical MLE method are also reported for completeness. In addition, in order to perform a quantitative comparison between the methods, two indicators are defined:

i. the relative absolute distance d_θ^{MLE} between the (posterior) most likely value c_θ of parameter θ and the corresponding MLE estimate $\hat{\theta}^{MLE}$, i.e.:

$$d_\theta^{MLE} = \frac{\left| c_\theta - \hat{\theta}^{MLE} \right|}{\hat{\theta}^{MLE}}. \quad (5)$$

Obviously, the lower is d_θ^{MLE} the closer is the most likely value c_θ to the MLE estimate $\hat{\theta}^{MLE}$, i.e., the higher is the *strength* of the approach in updating the prior possibilistic distribution on the basis of newly available experimental evidence;

ii. the percentage relative difference R_θ between the areas underlying the possibility distribution of parameter θ before and after the Bayesian update, namely S_θ^{Prior} and $S_\theta^{Posterior}$, respectively:

$$R_\theta = \frac{S_\theta^{Prior} - S_\theta^{Posterior}}{S_\theta^{Prior}} \cdot 100. \quad (6)$$

In this case, the higher is R_θ, the higher is the reduction in the area (i.e., in the epistemic uncertainty) and, thus, the higher is the "updating strength" of the approach.

Table 3.

It is evident that the *strength* of Approach A in moving the most likely values c_θ towards the corresponding MLE estimates $\hat{\theta}^{MLE}$ is *always higher* than that of Approach B. Actually, the values of d_θ^{MLE} (5) produced by Approach A are 1.58-5.24 times *lower* than those generated by Approach B for all the parameters.

From the analysis of quantitative indicator R_θ (6) it can be seen that both methods succeed in reducing the area underlying the possibility distributions of the uncertain parameters: in particular, the percentage reduction R_θ ranges between 8.33% and 30.49% for approach A and between 3.89% and 33.01% for Approach B. In addition, as expected, the strength of *both* approaches in reducing epistemic uncertainty *decreases* with the *size* of the data set used to perform the Bayesian update. For example, the area S_γ underlying the possibility distribution of γ (Fig. 3, top left column) is reduced by 30.49–33.01% with the aid of a *large* data set of size $D_1 = 149$; on the contrary, the area

537 $S_{\sigma_{Ks}}$ underlying the possibility distribution of σ_{Ks} (Fig. 3, bottom right column) is reduced only by
538 3.89–8.76% by means of D_4 = 5 pieces of data. Moreover, it is interesting to note that the strength
539 of Approach B in reducing epistemic uncertainty is *slightly higher* than that of Approach A *only*
540 when the amount of available data is quite *large* (i.e., in the revision of the possibility distributions
541 of parameters γ and δ of the PDF of $Y_1 = Q$, by means of D_1 = 149 pieces of data): actually, the
542 values of R_γ and R_δ range within 25.56-30.49% and 28.74-33.01% for Approaches A and B,
543 respectively. In all the other cases, the power of Approach A in reducing epistemic uncertainty is
544 *higher* than that of Approach B and this difference becomes more and *more evident* as the size of
545 the data set *decreases*: in particular, for medium-sized data sets (i.e., in the revision of the
546 possibility distributions of the parameters μ_{Zm}, σ_{Zm}, μ_{Zv}, σ_{Zv} of the PDFs of $Y_2 = Z_m$ and $Y_3 = Z_v$
547 with $D_2 = D_3$ = 29 pieces of data) the values of R_θ (6) produced by Approaches A and B range
548 within 8.33-24.38% and 4.17-22.00%, respectively; instead, for small-sized data sets (i.e., in the
549 revision of the possibility distributions of the parameters μ_{K_s} and σ_{Ks} of the PDF of $Y_4 = K_s$ with
550 D_4 = 5 pieces of data) the values of R_θ produced by Approaches A and B range within 8.76-16.28%
551 and 3.89-10.08%, respectively. This is particularly evident in the estimation of the standard
552 deviation σ_{Ks} of K_s (Fig. 3, bottom right column): on one side, the posterior distribution produced by
553 the hybrid approach (B) seems not to be influenced by the revision process (actually, the most likely
554 value of the parameter, $c_{\sigma_{Ks}}$ = 6.72, and the area underlying the corresponding posterior possibility
555 distribution, $S_{\sigma_{Ks}}$ = 3.95, are quite close to those of the prior, i.e., 6.89 and 4.11, respectively); on
556 the other side, the posterior distribution generated by the purely possibilistic approach (A) is almost
557 centered on the point estimates obtained by the MLE method and the corresponding area is reduced
558 by about 9%.
559
560 Finally, it has to be remarked that for the sake of simplicity the quantitative analyses above have
561 been performed only on the *marginal* posterior possibility distributions of the uncertain parameters

562 ($\pi^{\gamma}(\gamma|\boldsymbol{q})$, $\pi^{\delta}(\delta|\boldsymbol{q})$, $\pi^{\mu_{Zm}}(\mu_{Zm}|\boldsymbol{z_m})$, $\pi^{\sigma_{Zm}}(\sigma_{Zm}|\boldsymbol{z_m})$, $\pi^{\mu_{Zv}}(\mu_{Zv}|\boldsymbol{z_v})$, $\pi^{\sigma_{Zv}}(\sigma_{Zv}|\boldsymbol{z_v})$, $\pi^{\mu_{Ks}}(\mu_{K_s}|\boldsymbol{k_s})$

563 and $\pi^{\sigma_{Ks}}(\sigma_{Ks}|\boldsymbol{k_s})$); in fact, as highlighted before, the posterior possibility distributions of the

564 uncertain parameters are *multi-dimensional* functions (and the posterior estimates of the parameters

565 are *dependent*). Only for illustration purposes, Fig. 4 shows the α-cuts of level α = 0.05, 0.50 and

566 0.95 of the joint posterior possibility distributions $\pi^{\gamma,\delta}(\gamma,\delta|\boldsymbol{q})$, $\pi^{\mu_{Zm},\sigma_{Zm}}(\mu_{Zm},\sigma_{Zm}|\boldsymbol{z_m})$,

567 $\pi^{\mu_{Zv},\sigma_{Zv}}(\mu_{Zv},\sigma_{Zv}|\boldsymbol{z_v})$ and $\pi^{\mu_{K_s},\sigma_{K_s}}(\mu_{K_s},\sigma_{K_s}|\boldsymbol{k_s})$ of the parameters of the PDFs of $Y_1 = Q$ (top right

568 column), $Y_2 = Z_m$ (top left column), $Y_3 = Z_v$ (bottom left column) and $Y_4 = K_s$ (bottom right column),

569 respectively, produced by Approaches A (solid lines) and B (dashed lines).

570

571 **Fig. 4.**

572

573 The visual inspection of Fig. 4 confirms the results obtained by the quantitative analysis carried out

574 on the marginal distributions. For example, it can be seen that the areas of the α-cuts of level α =

575 0.95 of $\pi^{\mu_{Zv},\sigma_{Zv}}(\mu_{Zv},\sigma_{Zv}|\boldsymbol{z_v})$ and $\pi^{\mu_{K_s},\sigma_{K_s}}(\mu_{K_s},\sigma_{K_s}|\boldsymbol{k_s})$ produced by Approach A are consistently

576 smaller than those generated by Approach B; the α-cuts of $\pi^{\mu_{Zm},\sigma_{Zm}}(\mu_{Zm},\sigma_{Zm}|\boldsymbol{z_m})$ produced by

577 Approaches A and B are comparable in size, whereas the area of the α-cut of $\pi^{\gamma,\delta}(\gamma,\delta|\boldsymbol{q})$ obtained

578 by Approach B is slightly lower than the one produced by Approach A.

579

580 In order to show the effect that the reduction of the epistemic uncertainty in the distribution

581 parameters has on the uncertain input variables $Y_1 = Q$, $Y_2 = Z_m$, $Y_3 = Z_v$ and $Y_4 = K_s$, Fig. 5 reports

582 the upper and lower CDFs, $\overline{F}^{Y_j}(y_j)$ and $\underline{F}^{Y_j}(y_j)$ (i.e., the plausibility and belief functions,

583 $Pl^{Y_j}((-\infty, y_j])$ and $Bel^{Y_j}((-\infty, y_j])$, respectively), j = 1, 2, 3, 4, of $Y_1 = Q$ (top left column), $Y_2 = Z_m$

584 (top right column), $Y_3 = Z_v$ (bottom left column) and $Y_4 = K_s$ (bottom right column) before (solid

585 lines) and after the Bayesian update performed by Approaches A (dashed lines) and B (dot-dashed

lines). The calculation of CDFs is of great importance in risk assessment since they summarize the uncertainty "contained" in the variables of interest.

Fig. 5

Obviously, the gap between the plausibility and belief functions $Pl^{Y_j}\left(\left(-\infty, y_j\right]\right) = \overline{F}^{Y_j}\left(y_j\right)$ and $Bel^{Y_j}\left(\left(-\infty, y_j\right]\right) = \underline{F}^{Y_j}\left(y_j\right)$, j = 1, 2, 3, 4, is larger before the Bayesian update in all the cases considered: in particular, the 'prior' CDFs (solid lines) completely envelop the 'posterior' ones (dashed and dot-dashed lines). This larger gap is explained by the larger area contained under the possibility distribution functions of the corresponding epistemically-uncertain parameters (actually, as highlighted before, the larger the area, the higher the imprecision in the knowledge of the possibilistic parameters).

Then, in order to provide a fair and quantitative comparison between the two approaches adopted, the intervals $\left[\left[\overline{F}^{Y_j}\right]^{-1}(\beta), \left[\underline{F}^{Y_j}\right]^{-1}(\beta)\right]$ for the $\beta\cdot 100$-th percentiles Y_j^{β}, β = 0.05, 0.50 and 0.95, of the variables Y_j, j = 1, 2, 3, 4, are computed (Table 4). For example, analyzing Q^{β}, β = 0.05, 0.50 and 0.95, it can be seen that the width of the intervals is reduced by 30.54–40.34% and 29.88–43.29% by Approaches A and B, respectively: coherently with the results reported in Table 3, when a *large* data set is available (i.e., D_1 = 149 in this case) the strength of Approach B in reducing the epistemic uncertainty is *slightly* (i.e., 2-3%) *higher* than that of Approach A. On the contrary, analyzing K_s^{β}, β = 0.05, 0.50 and 0.95, it is evident that the width of the intervals is reduced of 25.30-30.80% and 5.74-19.45% by Approaches A and B, respectively: as highlighted before (Table 3), when a *small* data set is available (i.e., D_4 = 5 in this case) the power of Approach A in reducing epistemic uncertainty is *consistently* (i.e., 15-20%) *higher* than that of Approach B. This is confirmed *also* by the analysis of the quantiles of $Y_2 = Z_m$ and $Y_3 = Z_v$ (see Table 4).

Table 4.

Comparable conclusions can be drawn by the analysis of the upper and lower CDFs $\overline{F}^{Z_c}(z_c)$ and $\underline{F}^{Z_c}(z_c)$ (i.e., the plausibility and belief functions, $Pl^{Z_c}((-\infty, z_c])$ and $Bel^{Z_c}((-\infty, z_c])$, respectively), of the maximal water level of the river Z_c (i.e., the model output) obtained before (Fig. 6, solid lines) and after the Bayesian update performed by Approaches A (Fig. 6, dashed lines) and B (Fig. 6, dot-dashed lines). Notice that the distributions for Z_c have been obtained by propagating the two-level mixed probabilistic and possibilistic uncertainty through the mathematical model by means of a hybrid Monte Carlo (MC) and Fuzzy Interval Analysis (FIA) approach. This method combines the MC technique [Kalos and Withlock, 1986] with the extension principle of fuzzy set theory [Guyonnet et al., 2003; Kentel and Aral, 2004 and 2007] in two hierarchical, repeated steps [Baudrit et al., 2008; Kentel and Aral, 2005; Moller and Beer, 2004 and 2008; Moller et al., 2003 and 2006; Pedroni and Zio, 2012; Pedroni et al., 2013]: the reader is referred to the cited references for details. Again, in order to provide a fair comparison between the two approaches employed, proper quantitative indicators are computed. The final goal of the case study presented in the previous Section 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 issue (i) above, the quantity of interest that is most relevant to the decision maker is the $\beta \cdot 100$%-th quantile of Z_c (i.e., Z_c^{β}): this corresponds to the yearly maximal water level with a $\beta \cdot 100$-year return period. With respect to issue (ii) above, the quantity of interest that is most 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., $P[Z_c \geq z_c^*]$: in the present paper, z_c^* = 55.5 m (Table 5). Analyzing the intervals $[\left[\overline{F}^{Z_c}\right]^{-1}(\beta), \left[\underline{F}^{Z_c}\right]^{-1}(\beta)]$ for the percentiles Z_c^{β}, β = 0.05, 0.50 and 0.95, and the intervals $[1-\overline{F}^{Z_c}(z_c^*), 1-\underline{F}^{Z_c}(z_c^*)]$ for the exceedance probability $P[Z_c > z_c^*]$, it can be seen that their width is reduced of 28.14–38.63% and 16.88–29.23% by Approaches A and B, respectively (i.e., the

strength of Approach A in reducing epistemic uncertainty is about 4-11% *higher* than that of Approach B).

Fig. 6.

Table 5.

Several considerations are in order with respect to the results obtained. When the Bayesian update is performed based on a data set of *large* size (e.g., > 100 in this case), the difference in the behavior of the two approaches is quite *low*. This demonstrates that although the two methods are conceptually and algorithmically quite different, in presence of a "*strong*" experimental evidence they produce "*coherent*" results (i.e., posterior possibility distributions that bear the *same* overall "uncertainty content"): this is a fair outcome since the results provided by the two methods are expected to be more and more *similar* (i.e., more and more coherent with the experimental evidence) as the size of the data set *increases* (experts and practitioners may find a similarity and parallelism between these results and those obtained in purely probabilistic, graphical Bayesian models [Gelman et al., 2004]; in particular, see references concerning approaches used to "borrow strength" in (Hierarchical) Bayesian analyses [Atwood et al., 2003; Kelly and Smith, 2009 and 2011]). Instead, when the Bayesian update is performed on a data set of *small-medium* size (e.g., ≈ 5-30 in this case), the *strength* of Approach A in reducing epistemic uncertainty is *significantly higher* than that of Approach B. This can be explained as follows. In Approach A the *purely possibilistic* likelihood (i.e., the function "containing" the experimental evidence available) has a *direct* and strong influence on the *purely possibilistic* prior (actually, they are directly multiplied in (1)); on the contrary, in Approach B the *purely probabilistic* likelihood has a direct influence *only* on the (fictitious) *probabilistic* distributions that are superimposed onto the prior possibilistic parameters: this "artificial" procedure may in practice "soften" (i.e., reduce) the effect of the newly available information (i.e., of the data) in the revision of the possibilistic priors (obviously, this

663 effect is expected to be more evident if the amount of data, i.e., the *strength* of the experimental
664 evidence, is *small*). In such cases, embracing one method instead of the other may significantly
665 change the outcome of a decision making process in a risk assessment problem involving
666 uncertainties: this is of paramount importance in systems that are critical from the safety view point,
667 e.g., in the nuclear, aerospace, chemical and environmental fields.
668 Finally, it is absolutely important to acknowledge that even if the strength of one method (i.e., A in
669 this case) in reducing epistemic uncertainty is higher than that of the other one (i.e., B in this case),
670 this does *not necessarily* imply that one method is "*better*" or "*more effective*" than the other
671 *overall*. Actually, if on one side a consistent reduction in the epistemic uncertainty is in general
672 *desirable* in decision making processes related to risk assessment problems (since it significantly
673 *increases* the analyst's *confidence* in the decisions), on the other side this reduction must be
674 *coherent* with the amount of *information available*. In this view, an objection may arise in the
675 present case: is the *remarkable* strength of Approach A in reducing epistemic uncertainty (with very
676 few pieces of data) *fully justified* by such a small amount of data? In other words, is this
677 considerable reduction of epistemic uncertainty coherent with the strength of the experimental
678 evidence or is it too optimistic? These issues will be thoroughly discussed in the following
679 dedicated Section.
680
681 Finally, in addition to the strength of the approaches in revising the (prior) possibilistic description
682 of the uncertain parameters of aleatory variables, also the *computational time* associated to the
683 methods has to be taken into account. Table 6 reports the computational time t_{comp} required by the
684 Bayesian update of all the parameters of the PDFs of $Y_1 = Q$, $Y_2 = Z_m$, $Y_3 = Z_v$ and $Y_4 = K_s$ performed
685 by Approach A and by Approach B (with $T = 100$ repetitions of the purely probabilistic Bayes'
686 theorem for each of the $N_\alpha = 21$ α-cuts analyzed). Obviously, the computational time required by
687 Approach B is approximately $T \cdot N_\alpha$ (i.e., $T \cdot N_\alpha = 100 \cdot 21 = 2100$ in this case) times larger than that of
688 Approach A. On the other hand, notice that since Approach B is based on several *repetitions* of the

purely probabilistic Bayes' theorem, if possible, *parallelization* could be in principle employed to reduce the associated computational cost.

Table 6.

Discussion of the results and comparison of the approaches

The results of the comparisons performed in the previous Sections can be summarized as follows:

- both methods succeed in updating the possibilistic description of the epistemically-uncertain parameters of (aleatory) probability distributions by means of data. This is highlighted by the fact that in most cases the posterior possibility distributions produced by the two approaches are significantly different from the corresponding priors. In particular:
  - the most likely values of the parameters (i.e., those values in correspondence of which the possibility function equals 1) are moved towards the point estimates of the parameters obtained by the classical, purely probabilistic MLE method;
  - the area underlying the posterior possibility distributions is consistently lower than that of the priors: since this area is related to the imprecision in the knowledge of the possibilistic parameter (i.e., the larger the area, the higher the imprecision), it can be concluded that both approaches succeed in reducing the epistemic uncertainty in the possibilistic parameters of the aleatory probability distributions. This is also confirmed by the reduction of the gap between the upper and lower CDFs (i.e., the plausibility and belief functions) of the corresponding aleatory variables;
- when the Bayesian update is performed using a data set of *large* size (e.g., > 100), the *strength* of the two approaches in reducing the epistemic uncertainty is quite *similar*. By way of example, in the case study considered Approaches A and B reduce the areas underlying the possibility distributions of the uncertain parameters by 25.56-30.49% and 28.74-33.01%, respectively. This demonstrates that although the two methods are

conceptually and algorithmically quite different, in presence of a "strong" experimental evidence they produce "coherent" results (i.e., possibility distributions that bear almost the *same* overall "uncertainty content"): this is a fair outcome since the it is "desired" that both methods provide results that become more and more similar (i.e., that become more and more coherent with the experimental evidence) as the size of the data set increases;

- the strength of the purely possibilistic approach (A) in reducing epistemic uncertainty is *consistently higher* than that of the hybrid one (B) in presence of *medium*- and *small*-sized data sets (e.g., ≈ 5-30) (which is often the case in the risk analysis of complex safety-critical systems). For example, the width of the intervals for the quantiles of the variables of interest is reduced by 25.30-30.80% and 5.74-19.45% by Approaches A and B, respectively. This significantly different behavior is explained by the fact that in Approach A the (purely possibilistic) likelihood has an immediate and strong influence on the (purely possibilistic) prior (i.e., they are directly multiplied); on the contrary, in Approach B the (purely probabilistic) likelihood has a direct influence *only* on the (fictitious) probabilistic function that is superimposed onto the possibilistic parameter (subject to the Bayesian update): this "artificial" procedure could in practice weaken the effect of the newly available information (i.e., the data) in the revision of the possibilistic prior (this effect is expected to be more evident if the amount of data, i.e., the *strength* of the experimental evidence, is small).

As highlighted above, the fact that the power of Approach A in reducing epistemic uncertainty is higher than that of Approach B does *not necessarily* imply that one method is "*better*" or "*more effective*" than the other *overall*. For example, by hypothesis the remarkable strength of Approach A in reducing epistemic uncertainty by means of very few pieces of data (e.g., 5-30 in this case) might *not* be coherent with the "real" strength of such a scarce experimental evidence and could be accidentally due to some bias in the procedure. With respect to that, it has to be admitted that the uncertainty reduction power of the purely possibilistic approach (A) is strongly dependent on the shape of a *constructed* possibilistic

likelihood that could in principle *bias* the analysis. However: (i) in the present paper, this possibilistic function is *very closely related* to the classical, purely probabilistic one (which is theoretically well-grounded) by a simple and direct operation of *normalization* that preserves the "original structure" of the experimental evidence; (ii) the operation of normalization of the probabilistic likelihood finds also some *theoretical* justification in the work by [Denoeux, 2014]; (iii) in general, a probability-to-possibility transformation (properly performed according to the rules of possibility theory) *always* introduces *additional artificial* epistemic uncertainty into the analysis, i.e., it does not artificially reduce it (because it replaces a *single* probabilistic distribution by a *family* of distributions) [Dubois et al., 1993, 2004 and 2008; Flage et al., 2010 and 2013]. On the basis of considerations (i)-(iii) above, it seems unlikely that the purely possibilistic approach (A) may produce results that are dangerously over-optimistic with respect to those of the hybrid one (B). On the other hand, future research should be devoted to the study and development of rigorous, generalized methods for Bayesian model comparison and validation in a *purely possibilistic* framework, in order to complement and strengthen the conclusions drawn by means of the metrics originally introduced in the present manuscript: in this light, techniques from the classical, purely probabilistic field may serve as inspiring references [Gelman et al., 1996; Bayarri and Berger, 1999 and 2000; Johnson Valen, 2004];

- the computational time required by the hybrid approach (B) is *consistently higher* than that associated to the purely possibilistic one (A): this is explained by the necessity of repeatedly applying many times the purely probabilistic Bayes' theorem for *each* α-cut analyzed. More precisely, the application of method A just requires *one single* evaluation of the purely possibilistic Bayes' formula; on the other hand, approach B entails repeating $T \cdot N_\alpha$ times the classical probabilistic Bayes' theorem. In this respect, notice that N_α (i.e., the number of α-cuts processed) is typically of the order of 10 [Baudrit et al., 2006], whereas T (i.e., the number of parameter values selected to explore *each* α-cut) cannot be in principle prescribed

767 a priori: however, it needs to be large enough to thoroughly explore *each dimension* of the
768 epistemic parameter space (*at least* 10 values should be selected for each parameter);
769 • both methods relies on *assumptions* about either the prior or the likelihood functions: in the
770 purely possibilistic approach (A) the "original" possibilistic prior is employed, but a
771 possibilistic likelihood function has to be constructed (e.g., by probability-possibility
772 transformations or directly from rough experimental data); instead, in the hybrid method (B)
773 the original probabilistic likelihood function is used, but a "fictitious" prior Fuzzy
774 Probability Distribution Function needs to be identified by superimposing an *arbitrarily*
775 *selected* probabilistic PDF onto the "original" possibilistic prior that has to be updated.
776
777 Based on these findings, the advantages and drawbacks of the two approaches are summarized in
778 the following Table 7.
779
780 **Table 7.**
781

782 # Conclusions

783 In this paper, we have considered two methods for the Bayesian update of the possibilistic
784 parameters of aleatory probability distributions, with exemplification on a case study concerning the
785 risk-based design of a flood protection dike. The first method considered is based on a purely
786 possibilistic counterpart of the classical probabilistic Bayes' theorem; the second is a hybrid
787 (probabilistic and possibilistic) method combining Fuzzy Interval Analysis and the classical
788 probabilistic Bayes' theorem.
789 The findings of the work show that in general adopting different methods may generate different
790 results and possibly different decisions in risk problems involving uncertainties: this is of
791 paramount importance in systems that are critical from the safety viewpoint, e.g., in the nuclear,
792 aerospace, chemical and environmental fields.

In particular, on the basis of the results obtained, it seems advisable to suggest the use of the purely possibilistic approach (instead of the hybrid one) for the following reasons: (i) its strength in reducing epistemic uncertainty is significantly higher, in particular when the amount of available data is small: this is important in decision making processes since reducing epistemic uncertainty significantly increases the analyst confidence in the decisions; (ii) the computational time required is consistently lower.

However, it has to be remarked that the construction of a possibilistic likelihood required by the purely possibilistic method, although recently tacked in the literature, still represents an issue to be *further* investigated from both the theoretical and practical viewpoint in order to avoid introducing biases in the analysis and to suggest the application of the approach for real risk assessment problems: with respect to that, future research will be devoted to the investigation of additional methods developed to this aim. Also, future studies will be aimed at developing generalized methods for Bayesian model comparison and validation in a *purely possibilistic* framework, in order to complement and strengthen the conclusions drawn by means of the metrics originally introduced in the present manuscript.

References

Anoop MB, Balaji Rao K, Gopalakrishnan S, 2006. Conversion of probabilistic information into fuzzy sets for engineering decision analysis. Computers and Structures, 84, pp. 141-155.

Apostolakis, G.E., 1990. The concept of probability in safety assessment of technological systems. Science, 250, 1359.

Apostolakis, G.E., S. Kaplan, 1981. Pitfalls in risk calculations. Reliability Engineering, Vol. 2, Issue 2, pp. 135-145.

Atwood, C.L., JL. LaChance, H.F. Martz, D.L Anderson, M. Englehardte, D. Whitehead, T. Wheeler, 2003. Handbook of Parameter Estimation for Probabilistic Risk Assessment. NUREG/CR-6823, US Nuclear Regulatory Commission, Washington, DC; SAND2003-3348P, Sandia National Laboratories, Albuquerque, New Mexico.

Aven, T., 2010. On the Need for Restricting the Probabilistic Analysis in Risk Assessments to Variability. Risk Analysis, Vol. 30, Issue 3, pp. 354-360.

Aven, T., 2011. Interpretations of alternative uncertainty representations in a reliability and risk analysis context. Reliability Engineering & System Safety, Volume 96, Issue 3, pp. 353-360.

Aven, E. Zio, 2011. Some considerations on the treatment of uncertainties in risk assessment for practical decision making, Reliability Engineering and System Safety, Vol. 96, Issue 1, pp. 64-74.

Ballesteros, G., Angelikopoulos, P., Papadimitriou, C., Koumoutsakos, P., 2014. Bayesian Hierarchical Models for Uncertainty Quantification in Structural Dynamics. Vulnerability, Uncertainty, and Risk: Quantification, Mitigation, and Management, ASCE, pp. 1615-1624.

Baudrit, D. Dubois, 2006. Practical Representations of Incomplete Probabilistic Knowledge, Computational Statistics & Data Analysis, Vol. 51, Issue 1, pp. 86-108.

833 Baudrit, I. Couso, D. Dubois, 2007. Joint propagation of probability and possibility in risk analysis:
834 toward a formal framework, Internat. J. Approx. Reasoning, Vol. 45, Issue 1, pp. 82–105.
835 Baudrit, D. Dubois, D. Guyonnet, 2006. Joint Propagation and Exploitation of Probabilistic and
836 Possibilistic Information in Risk Assessment. IEEE Transactions on Fuzzy Systems, Vol. 14,
837 Issue 5, pp. 593-608.
838 Baudrit, D. Dubois, N. Perrot, 2008. Representing parametric probabilistic models tainted with
839 imprecision, Fuzzy Sets and System, Vol. 159, Issue 15, pp.1913-1928.
840 Bayarri MJ, Berger, JO, 1999. Quantifying surprise in the data and model verification. In: Bernardo,
841 JM, et al., editors. Bayesian statistics 6. Oxford: Oxford University Press. p.53–82.
842 Bayarri, MJ, Berger, JO, 2000. P-values for composite null models. J Am Stat Assoc; 95: 1127–42.
843 Bedford, T., Cooke, R., 2001. Probabilistic Risk Analysis. Foundations and Methods. Cambridge:
844 Cambridge University Publishing Ltd.
845 Beer, M., 2009a. Engineering quantification of inconsistent information, Int J Reliability and
846 Safety, 3(1/2/3): 174–197.
847 Beer, M., 2009b. Fuzzy probability theory, in: Encyclopedia of Complexity and Systems Science,
848 vol.6, Springer, NewYork, pp.4047–4059.
849 Beer, M., and Ferson, S., 2013. Special issue of Mechanical Systems and Signal Processing
850 "Imprecise probabilities - What can they add to engineering analyses?". Mechanical Systems
851 and Signal Processing, 37(1-2), 1-3.
852 Beer, M., Ferson, S., and Kreinovich, V., 2013a. Imprecise probabilities in engineering analyses.
853 Mechanical Systems and Signal Processing, 37(1-2), 4-29.
854 Beer, M., Zhang, Y., Quek, S.T., Phoon, K.K., 2013b. Reliability analysis with scarce information:
855 Comparing alternative approaches in a geotechnical engineering context, Structural Safety,
856 Vol. 41, pp. 1–10.

857 Beer, M., DiazDelaO, F.A., Patelli, E., Au, S.K., 2014a. Conceptual comparison of Bayesian
858 approaches and imprecise probabilities. In: Topping, B.H.V.; Ivanyi, P. (eds.), Computational
859 Technology Reviews 9, 2014, 1-29, Saxe-Coburg Publications.
860 Beer, M., Kougioumtzoglou, I.A., Patelli, E., 2014b. Emerging Concepts and Approaches for
861 Efficient and Realistic Uncertainty Quantification. In: Frangopol, D.M.; Tsompanakis, Y.
862 (eds.), Maintenance and Safety of Aging Infrastructure, Book Series "Structures &
863 Infrastructures", Vol 10, Chapter 5, 121–154, CRC Press, Taylor & Francis Group, Boca
864 Raton, London, New York, Leiden.
865 Behmanesh, I., Moaveni, B., Lombaert, G., Papadimitriou, C., 2015. Hierarchical Bayesian Model
866 Updating for Probabilistic Damage Identification. In: Atamturktur, HS, Moaveni, B.,
867 Papadimitriou, C., Schoenherr, T. Model validation and uncertainty quantification, Volume 3,
868 Conference Proceedings of the Society for Experimental Mechanics Series, pp. 55-66.
869 Bernard, J.M., 2005. An introduction to the imprecise Dirichlet model for multinomial data, Int. J.
870 Approx. Reason. 39, 123–150.
871 Bernardo, J M and Smith, A F M, 1994. Bayesian Theory. Chichester: Wiley.
872 Blockley, D., 2013. Analysing uncertainties: Towards comparing Bayesian and interval
873 probabilities. Mechanical Systems and Signal Processing, 37(1-2), 30-42.
874 Buckley, J.J., 2005. Fuzzy probabilities — new approach and applications, Studies in Fuzziness and
875 Soft Computing, vol.115, Springer, Berlin, Heidelberg.
876 Chung, Y., Gelman, A., Rabe-Hesketh, S., Liu, J., Dorie, V., 2015. Weakly Informative Prior for
877 Point Estimation of Covariance Matrices in Hierarchical Models. Journal of Educational and
878 Behavioral Statistics, Vol. 40, No. 2, pp. 136–157.
879 Congdon, P.D., 2010. Applied Bayesian Hierarchical Methods, CRC Press.
880 Coolen, F. P. A., and Utkin, L. V., 2007. Imprecise probability: a concise overview." Risk,
881 reliability and societal safety: proceedings of the European safety and reliability conference
882 (ESREL), Stavanger, Norway, 1959-66.

883 Couso, I., Dubois, D., 2009. On the Variability of the Concept of Variance for Fuzzy Random
884 Variables. Fuzzy Systems, IEEE Transactions on, Volume: 17, Issue: 5, pp. 1070 – 1080.
885 Couso, I., Miranda, E., de Cooman, G., 2004. A Possibilistic Interpretation of the Expectation of a
886 Fuzzy Random Variable. In: Soft Methodology and Random Information Systems, Series in
887 Advances in Soft Computing, Volume 26, pp. 133-140, Springer Berlin Heidelberg.
888 Couso, I., S. Montes, P. Gil, 2001. The necessity of the strong alpha-cuts of a fuzzy set,
889 International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 9, 249-262.
890 Couso, I., L. Sánchez, 2008. Higher order models for fuzzy random variables, Fuzzy Sets and
891 Systems 159, 237-258.
892 Couso, I., L. Sánchez, 2011. Upper and lower probabilities induced by a fuzzy random variable.
893 Fuzzy Sets and Systems, Volume 165, Issue 1, 16 February 2011, Pages 1–23.
894 Crespo, L. G., Kenny, S. P., and Giesy, D. P., 2013. Reliability analysis of polynomial systems
895 subject to p-box uncertainties. Mechanical Systems and Signal Processing, 37(1-2), 121-136.
896 Denoeux, T., 2014. Likelihood-based belief function: Justification and some extensions to low
897 quality data. International Journal of Approximate Reasoning, Volume 55, Issue 7, Pages 1535–
898 1547.
899 Dubois, D., 2006. Possibility Theory and Statistical Reasoning. Computational Statistics and Data
900 Analysis, Vol. 51, pp. 47-69.
901 Dubois, D., L. Foulloy, G. Mauris, H. Prade, 2004. Probability-Possibility Transformation,
902 Triangular Fuzzy Sets, and Probabilistic Inequalities. Reliable Computing, vol. 10, pp. 273-
903 297.
904 Dubois, D., H. Prade, 1980. Fuzzy Sets and Systems: Theory and Applications, Academic Press,
905 New York, NY.
906 Dubois, D., Prade, H., 1988. Possibility Theory: An Approach to Computerized Processing of
907 Uncertainty, New York, Plenum Press.

908 Dubois, D., Prade, H., 1997. Bayesian conditioning in possibility theory. Fuzzy Sets and Systems,
909 92, pp. 223-240.
910 Dubois, D, Prade, H & Sandri, S, 1993. On possibility/probability transformations. In: Lowen, R &
911 Roubens, M (eds). Fuzzy Logic: State of the Art. Dordrecht: Kluwer Academic Publishers.
912 pp. 103112.
913 Dubois D., Prade H., Smets P., 2008. A definition of subjective possibility. International Journal of
914 Approximate Reasoning, 2008, Vol. 48, pp. 352-364.
915 Dubois D., Moral, S., Prade H., 1997. A Semantics for Possibility Theory Based on Likelihoods.
916 Journal of mathematical analysis and applications, 205, pp. 359-380.
917 Edwards, A.W.F., 1992. Likelihood, expanded edition. The John Hopkins University Press,
918 Baltimore, USA.
919 Féron, R., 1976. Ensembles aléatoires flous, CR. Acad. Sc. paris - Série A, Vol 282, 903-906.
920 Ferson, S., 2005. Bayesian methods in risk assessment. Technical Report:
921 www.ramas.com/bayes.pdf.
922 Ferson, S. and Hajagos, J.G., 2004. Arithmetic with uncertain numbers: rigorous and (often) best
923 possible answers. Reliability Engineering and System Safety, 85, pp. 135-152.
924 Ferson, V. Kreinovich, L. Ginzburg, K. Sentz, D.S. Myers, 2003. Constructing probability boxes
925 and Dempster-Shafer structures, Sandia National Laboratories, Technical Report SAND2002-
926 4015, Albuquerque, New Mexico.
927 Ferson, V. Kreinovich, J. Hajagos, W. Oberkampf and L. Ginzburg, 2007. Experimental
928 Uncertainty Estimation and Statistics for Data Having Interval Uncertainty, Setauket, New
929 York 11733, Technical Report SAND2007-0939.
930 Ferson, R.B. Nelsen, J. Hajagos, D.J. Berleant, J. Zhang, W.T. Tucker, L.R. Ginzburg, W.L.
931 Oberkampf, 2004. Dependence in probabilistic modeling, Dempster-Shafer theory, and
932 probability bounds analysis, Technical Report SAND2004-3072, Albuquerque, New Mexico.

933 Ferson, W.T. Tucker, 2006. Sensitivity in risk analyses with uncertain numbers, Setauket, New
934 York 11733, Technical Report SAND2006-2801.
935 Ferson, P. Van den Brink, T.L. Estes, K. Gallagher, R. O'Connor, F. Verdonck, 2010. Bounding
936 uncertainty analyses. In: Application of uncertainty analysis to ecological risks of pesticides /
937 Warren-Hicks, W.J., A. Hart. Pensacola and Boca Raton, FL.: SETAC and CRC Press, 2010 -
938 ISBN 9781439807347.
939 Flage R., Aven T, Zio E., 2009. Alternative representations of uncertainty in reliability and risk
940 analysis – review and discussion. In: Martorell S, Guedes Soares C, Barnett J, editors.
941 Valencia, Spain, 22–25 September 2008 Safety, reliability and risk analysis. Theory, methods
942 and applications. Proceedings of the European safety and reliability conference 2009 (ESREL
943 2009). London: CRC Press, p. 2081–91.
944 Flage, P. Baraldi, E. Zio, T. Aven, 2010. Possibility-probability transformation in comparing
945 different approaches to the treatment of epistemic uncertainties in a fault tree analysis, in: B.
946 Ale, I.A. Papazoglu, E. Zio (Eds.), Reliability, Risk and Safety - Proceedings of the European
947 Safety and Reliability (ESREL) 2010 Conference, Rhodes, Greece, 5-9 September 2010, pp.
948 714-721, Taylor & Francis Group, London, United Kingdom, 2010, ISBN 978-0-415-60427-
949 7.
950 Flage, P. Baraldi, E. Zio, T. Aven, 2013. Probability and possibility-based representations of
951 uncertainty in fault tree analysis. Risk Analysis, 33(1), pp. 121-33.
952 Gelman A., et al., 2004. Bayesian data analysis. Second Ed. Chapman & Hall, CRC.
953 Gelman, A., 2006. Prior distributions for variance parameters in hierarchical models. Bayesian
954 Analysis, 1(3), pp. 515–533.
955 Gelman, A., Meng, XL, Stern, H., 1996. Posterior predictive assessment of model fitness via
956 realized discrepancies. Stat Sin: 733–806.

957 Gelman, A., van Dyk, D., Huang, Z., Boscardin, W.J., 2008. Using Redundant Parameterizations to
958 Fit Hierarchical Models. Journal of Computational and Graphical Statistics, Volume 17,
959 Number 1, Pages 95–122.
960 Gill, J., 2002. Bayesian methods: a social and behavioral sciences approach, Chapman & Hall/CRC,
961 Boca Raton, FL.
962 Guyonnet, D., Bourgine, B., Dubois, D., Fargier, H., Côme, B., Chilès, J.P., 2003. Hybrid approach
963 for addressing uncertainty in risk assessments. Journal of the Environmental Engineering
964 Division, ASCE 129, pp. 68–78.
965 Helton, J.D. Johnson, W.L. Oberkampf, C.J. Sallaberry, 2008. Representation of Analysis Results
966 Involving Aleatory and Epistemic Uncertainty, Sandia National Laboratories, Technical
967 Report SAND2008-4379, Albuquerque, New Mexico.
968 Helton, J.D. Johnson, W.L. Oberkampf, C.B. Storlie, 2007. A sampling-based computational
969 strategy for the representation of epistemic uncertainty in model predictions with evidence
970 theory, Computer Methods in Applied Mechanics and Engineering, 196: 3980–98.
971 Hou, Y. and Yang, B., 2010. Probability-Possibility Transformation for Small Sample Size Data.
972 2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery (FSKD
973 2010), pp. 1720-1724.
974 Jalal-Kamali, A., and Kreinovich, V., 2013. Estimating correlation under interval uncertainty.
975 Mechanical Systems and Signal Processing, 37(1-2), 43-53.
976 Jaworska, J.S., Aldenberg, T., 2000. Estimation of HC5 taking into account uncertainties of
977 individual dose response curves and species sensitivity distribution. Presentation at the 2000
978 Society for Risk Analysis annual meeting, Arlingotn, Virginia,
979 http://www.riskworld.com/Abstract/2000/SRAam00/ab0ac164.htm.
980 Johnson Valen, E., 2004. A Bayesian χ^2 test for goodness-of-fit. Ann Stat, 32: pp. 2361–84.
981 Kalos, M.H., P.A. Whitlock, 1986. Monte Carlo methods. Volume I: Basics, Wiley, New York,
982 NY.

983 Kelly, D.L., Smith, C.L, 2009. Bayesian inference in probabilistic risk assessment — The current
984 state of the art. Reliability Engineering and System Safety, 94, pp. 628–643.
985 Kelly, D.L., Smith, C.L, 2011. Bayesian Inference for Probabilistic Risk Assessment: A
986 Practitioner's Guidebook. Springer-Verlag: London, UK.
987 Kentel, E., M.M. Aral, 2004. Probabilistic-fuzzy health risk modeling. Stoch. Envir. Res. and Risk
988 Ass., Vol. 18: pp. 324–338.
989 Kentel, E., M.M. Aral, 2005. 2D Monte Carlo versus 2D Fuzzy Monte Carlo Health Risk
990 Assessment, Internat. J. Stochastic Environ. Res. Risk Assess., 19, pp. 86–96.
991 Kentel, E., M.M. Aral, 2007. Risk tolerance measure for decision-making in fuzzy analysis: a health
992 risk assessment perspective. Stoch. Environ Res. Ris. Assess., Vol. 21: pp. 405–417.
993 Klir, B. Yuan, 1995. Fuzzy Sets and Fuzzy Logic: Theory and Applications, Prentice-Hall, Upper
994 Saddle River, NJ.
995 Kozine, I., Filimonov, Y., 2000. Imprecise reliabilities: experiences and advances, Reliab. Eng.
996 Syst. Saf., 67, 75–83.
997 Kozine, I. O., and Utkin, L. V., 2002. Processing unreliable judgements with an imprecise
998 hierarchical model. Risk, Decision and Policy, 7(03), 325-339.
999 Kuznetsov, V. P., 1991. Interval statistical models (in Russian), Radio i Svyaz, Moscow.
1000 Kwakernaak, H., 1978. Fuzzy Random Variables-I. Definitions and Theorems, Information
1001 Sciences 15, 1-29.
1002 Lapointe S. and Bobeè B., 2000. Revision of possibility distributions: A Bayesian inference pattern.
1003 Fuzzy Sets and Systems, 116, pp. 119-140.
1004 Le Duy, T.D., Vasseur, D., Couplet, M., Dieulle, L., Bérenguer, C., 2011. A study on updating
1005 belief functions for pa-rameter uncertainty representation in Nuclear Probabilistic Risk
1006 Assessment. In F. Coolin, G. De Cooman, T. Fetz, M. Oberguggenberger (Eds.), Proceedings
1007 of the 7th Interna-tional Symposium on Imprecise Probability: Theories and Applications, 25-
1008 28 July 2011, Innsbruck, Austria: 247-256. Innsbruck, Austria: SIPTA.

1009 Limbourg, P., E. de Rocquigny, 2010. Uncertainty analysis using evidence theory – confronting
1010 level-1 and level-2 approaches with data availability and computational constraints,
1011 Reliability Engineering and System Safety, Vol. 95, Issue 5, pp. 550-564.
1012 Lindley, D.V., 2000. The philosophy of statistics. The Statistician, 49(3): 293-337.
1013 Lindley, D. V., 2006. Understanding uncertainty, Wiley, Hoboken, NJ.
1014 Masegosa, A., S. Moral, 2014. Imprecise probability models for learning multinomial distributions
1015 from data. Applications to learning credal networks. International Journal of Approximate
1016 Reasoning, Volume 55, Issue 7, Pages 1548–1569.
1017 Masson, M.H. and Denoeux, T., 2006. Inferring a possibility distribution from empirical data.
1018 Fuzzy Sets and Systems, 157, pp. 319–340.
1019 Mauris, G., 2008. Inferring a Possibility Distribution from Very Few Measurements. In: D. Dubois
1020 et al. (Eds.), Soft Methods for Hand. Var. and Imprecision, ASC 48, pp. 92–99, Springer-
1021 Verlag, Berlin, Heidelberg.
1022 Mehl, C. H., 2013. P-boxes for cost uncertainty analysis. Mechanical Systems and Signal
1023 Processing, 37(1-2), 253-263.
1024 Molchanov, I., 2005. Theory of Random Sets, Springer, New York.
1025 Möller, B., M. Beer, 2004. Fuzzy Randomness: Uncertainty in Civil Engineering and
1026 Computational Mechanics, Springer, Berlin.
1027 Möller, B., M. Beer, 2008. Engineering computation under uncertainty – Capabilities of non-
1028 traditional models. Computers and Structures, Vol. 86, pp. 1024–1041.
1029 Moller, B., W. Graf, M. Beer, 2003. Safety assessment of structures in view of fuzzy randomness.
1030 Computers and Structures, Vol. 81, pp. 1567–1582.
1031 Moller, B., M. Beer, W. Graf, J.U. Sickert, 2006. Time-dependent reliability of textile-strengthened
1032 RC structures under consideration of fuzzy randomness. Computers and Structures, Vol. 84,
1033 pp. 585–603.

Moral, S., 2014. Comments on “Likelihood-based belief function: Justification and some extensions to low-quality data” by Thierry Denoeux. International Journal of Approximate Reasoning, Volume 55, Issue 7, Pages 1591–1593.

Muscolino, G., and Sofi, A., 2013. Bounds for the stationary stochastic response of truss structures with uncertain-but-bounded parameters. Mechanical Systems and Signal Processing, 37(1-2), 163-181.

NUREG-CR-6850, 2005. EPRI/NRC-RES Fire PRA methodology for nuclear power facilities, Volume 2: detailed methodology. US Nuclear Regulatory Commission.

NASA, 2010. Risk-Informed Decision Making Handbook. NASA/SP-2010-576 – Version 1.0, April 2010.

Pannier, S., Waurick, M., Graf, W., and Kaliske, M., 2013. Solutions to problems with imprecise data” - An engineering perspective to generalized uncertainty models. Mechanical Systems and Signal Processing, 37(1-2), 105-120.

Pasanisi, A., E. de Rocquigny, N. Bousquet, E. Parent, 2009. Some useful features of the Bayesian setting while dealing with uncertainties in industrial practice, Proceedings of the ESREL 2009 Conference, Prague, Czech Republic, 7-10 September 2009, pp. 1795-1802.

Pasanisi, A., Keller, M., Parent, M., 2012. Estimation of a quantity of interest inuncertainty analysis: Some help from Bayesian decision theory. Reliability Engineering and System Safety, 100, pp. 93–101.

Pedroni, N., Zio, E., 2012. Empirical comparison of methods for the hierarchical propagation of hybrid uncertainty in risk assessment, in presence of dependences. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, Vol. 20, Issue 4, 2012, pp. 509-557.

Pedroni, N., E. Zio, E. Ferrario, A. Pasanisi, M. Couplet, 2013. Hierarchical propagation of probabilistic and non-probabilistic uncertainty in the parameters of a risk model. Computers and Structures, Volume 126, pp. 199–213.

Pedroni, N., E. Zio, A. Pasanisi, M. Couplet, 2014. Bayesian update of the parameters of probability distributions for risk assessment in a two-level hybrid probabilistic-possibilistic uncertainty framework. In: R.D.J.M. Steenbergen, P.H.A.J.M. van Gelder, S. Miraglia and A. C.W.M. Ton. Vrouwenvelder (Eds.), *Safety, Reliability and Risk Analysis, Beyond the Horizon, Proceedings of the European Safety and RELiability Conference (ESREL) 2013*, Amsterdam, The Netherlands, 29 September-2 October 2013, Taylor and Francis Group, London, UK, pp. 3295–3302.

Puri, M.L., D.A. Ralescu, 1986. Fuzzy random variables, Journal of Mathematical Analysis and Applications 114, 409-422.

Quaeghebeur and de Cooman, 2005. Imprecise probability models for inference in exponential families. Electronic Proceedings of the 4th International Symposium on Imprecise Probabilities and Their Applications, Pittsburgh, Pennsylvania, 2005.

Ralescu, D., 2002. Average level of a fuzzy set, in: Statistical Modeling, Analysis and Management of Fuzzy Data, eds. C. Bertoluzza, M.A. Gil, D.A. Ralescu (Springer, Heidelberg) pp. 119–126.

Reid, S. G., 2013. Probabilistic confidence for decisions based on uncertain reliability estimates. Mechanical Systems and Signal Processing, 37(1-2), 229-239.

Sankararaman, S., and Mahadevan, S., 2013. Distribution type uncertainty due to sparse and imprecise data. Mechanical Systems and Signal Processing, 37(1-2), 182-198.

Sentz, S. Ferson, 2002. Combination of Evidence in Dempster-Shafer Theory, Sandia National Laboratories, Technical Report SAND 2002-0835, Albuquerque, New Mexico.

Serrurier, M., and Prade, H., 2011. Maximum-Likelihood Principle For Possibility Distributions Viewed As Families Of Probabilities. 2011 IEEE International Conference on Fuzzy Systems June 27-30, 2011, Taipei, Taiwan, pp. 2987-2993.

Shirley, K.E., Gelman, A., 2015. Hierarchical models for estimating state and demographic trends in US death penalty public opinion. J. R. Statist. Soc. A, 178(1), pp. 1-28.

1085 Siu, N., Kelly, D., 1998. Bayesian parameter estimation in probabilistic risk assessment. Reliability
1086 Engineering and System Safety, 62, pp. 89–116.
1087 Smets, P., 1993. Belief Functions: The Disjunctive Rule of Combination and the Generalized
1088 Bayesian Theorem. International Journal of Approximate Reasoning, 9, pp. 1-35.
1089 Stein M. & M. Beer, 2011. Bayesian quantification of inconsistent information. Applications of
1090 Statistics and Probability in Civil Engineering – Faber, Köhler & Nishijima (eds), Taylor &
1091 Francis Group, London, ISBN 978-0-415-66986-3.
1092 Stein M., M. Beer, V. Kreinovich, 2013. Bayesian Approach for Inconsistent Information.
1093 Information Sciences: http://dx.doi.org/10.1016/j.ins.2013.02.024.
1094 USNRC, 2002. An approach for using probabilistic risk assessment in risk-informed decisions on
1095 plant-specific changes to the licensing basis. NUREG-1.174 – Revision 1, US Nuclear
1096 Regulatory Commission, Washington, DC.
1097 USNRC, 2009. Guidance on the Treatment of Uncertainties Associated with PRAs in Risk-
1098 Informed Decision Making. NUREG-1855, US Nuclear Regulatory Commission,
1099 Washington, DC.
1100 Viertl, R., 1996. Statistical Methods for Non-Precise Data. CRC Press, Boca Raton, New York,
1101 London, Tokyo.
1102 Viertl, R., 1997. On statistical inference for non precise data. Environmetrics, 8, pp. 541-568.
1103 Viertl, R., 1999. Statistics and integration of fuzzy functions. Environmetrics, 10, pp. 487-491.
1104 Viertl, R., 2008a. Foundations of Fuzzy Bayesian Inference. Journal of Uncertain Systems, 2(3), pp.
1105 187–191.
1106 Viertl, R., 2008b. Fuzzy Bayesian Inference. In: D. Dubois et al. (Eds.), Soft Methods for Handling
1107 Variability and Imprecision, ASC 48, pp. 10–15.
1108 Viertl, R., 2011. Statistical Methods for Fuzzy Data, Wiley, Chichester.
1109 Viertl, R. & Hareter, D., 2004a. Fuzzy information and imprecise probability. ZAMM, Z. Angew.
1110 Math. Mech., 84 (10–11), pp. 731 – 739.

Viertl, R. & Hareter, D., 2004b. Generalized Bayes-theorem for non-precise a-priori distribution, Metrika, 59(3), pp. 263–273.

Viertl R. and H. Hule, 1991. On Bayes' theorem for fuzzy data. Stat. Pap. 32.

Walley, P., 1991. Statistical reasoning with imprecise probabilities, Chapman and Hall, New York.

Walley, P., 1996. Inferences from multinomial data: learning about a bag of marbles (with discussion), J. R. Stat. Soc. B 58, 3–57.

Walter, G., and Augustin, T., 2009. Imprecision and Prior-data Conflict in Generalized Bayesian Inference. Journal of Statistical Theory and Practice, Volume 3, No.1, pp. 255-271.

Weichselberger, K., 2000. The theory of interval-probability as a unifying concept for uncertainty, Int. J. Approx. Reasoning, 24(2–3), 149–170.

Zhang, H., Dai, H., Beer, M., and Wang, W., 2013. Structural reliability analysis on the basis of small samples: An interval quasi-Monte Carlo method. Mechanical Systems and Signal Processing, 37(1-2), 137-151.

Figure captions list

Fig. 1. ***Top left****: triangular possibility distribution* $\pi^{\gamma}(\gamma)$ *of the epistemically-uncertain parameter* γ *of the Gumbel probability distribution of Y; in evidence the* α*-cuts of level* $\alpha = 0, 0.5$ *and 1.* ***Top right****: four PDFs belonging to the family* $\left\{p^{Y}(y \mid \gamma, \delta)\right\}_{\alpha=0}$ *and four PDFs belonging to the family* $\left\{p^{Y}(y \mid \gamma, \delta)\right\}_{\alpha=0.5}$. ***Bottom****: bounding upper and lower CDFs of Y,* $\overline{F}_{\alpha}^{Y}(y) = Pl_{\alpha}^{Y}(Z)$ *and* $\underline{F}_{\alpha}^{Y}(y) = Bel_{\alpha}^{Y}(Z)$, $Z = (-\infty, y]$, *built in correspondence of the* α*-cuts of level* $\alpha = 0, 0.5$ *and 1 of* $\pi^{\gamma}(\gamma)$*; the plausibility and belief functions* $Pl^{Y}(Z)$ *and* $Bel^{Y}(Z)$, $Z = (-\infty, y]$, *are also shown*

Fig. 2. *Identification of the P-dimensional* α*-cut* $A_{\alpha}^{\boldsymbol{\theta}|\boldsymbol{y}}$ *of the posterior possibility distribution of* $\boldsymbol{\theta}$ *as the hull enveloping T = 20 point estimates* $\boldsymbol{\theta}_{l,\alpha} \mid \boldsymbol{y}$, $l = 1, 2, \ldots, 20$ *(dots), generated by the repeated application of the Bayes' theorem in the hybrid probabilistic-possibilistic approach (B), in the particular case of P = 2 parameters* θ_1 *and* θ_2*. The corresponding* α*-cut* $A_{\alpha,Cart}^{\theta_1,\theta_2|\boldsymbol{y}}$ *generated by the Cartesian product of the (one-dimensional)* α*-cuts* $A_{\alpha}^{\theta_1|\boldsymbol{y}}$ *and* $A_{\alpha}^{\theta_2|\boldsymbol{y}}$ *of the marginal possibility distributions is also shown*

Fig. 3. *Prior and posterior possibility distributions of the epistemically-uncertain parameters of the aleatory PDFs of* $Y_1 = Q$ (***top row, left and right column***), $Y_2 = Z_m$ (***middle-top row, left and right column***), $Y_3 = Z_v$ (***middle-bottom row, left and right column***) *and* $Y_4 = K_s$ (***bottom row, left and right column***)*. The point estimates of the parameters obtained by the classical MLE method are also shown for comparison. Adapted and extended from [Pedroni et al., 2014]*

***Fig. 4**. Exemplary α-cuts of level α = 0.05, 0.50 and 0.95 of the joint posterior possibility distributions of the parameters of the PDFs of $Y_1 = Q$ (**top left**), $Y_2 = Z_m$ (**top right**), $Y_3 = Z_v$ (**bottom left**) and $Y_4 = K_s$ (**bottom right**), produced by Approaches A (solid lines) and B (dashed lines)*

***Fig. 5**. Upper and lower CDFs, $\overline{F}^{Y_j}(y_j)$ and $\underline{F}^{Y_j}(y_j)$, j = 1, 2, 3, 4, of the uncertain input variables $Y_1 = Q$ (**top left**), $Y_2 = Z_m$ (**top right**), $Y_3 = Z_v$ (**bottom left**) and $Y_4 = K_s$ (**bottom right**) before and after the Bayesian update performed by Approaches A and B*

***Fig. 6**. Plausibility and belief functions, $Pl^{Z_c}((-\infty, z_c])$ and $Bel^{Z_c}((-\infty, z_c])$, of the maximal water level of the river Z_c before and after the Bayesian update performed by Approaches A and B. Adapted and extended from [Pedroni et al., 2014]*

Figures and corresponding associated captions

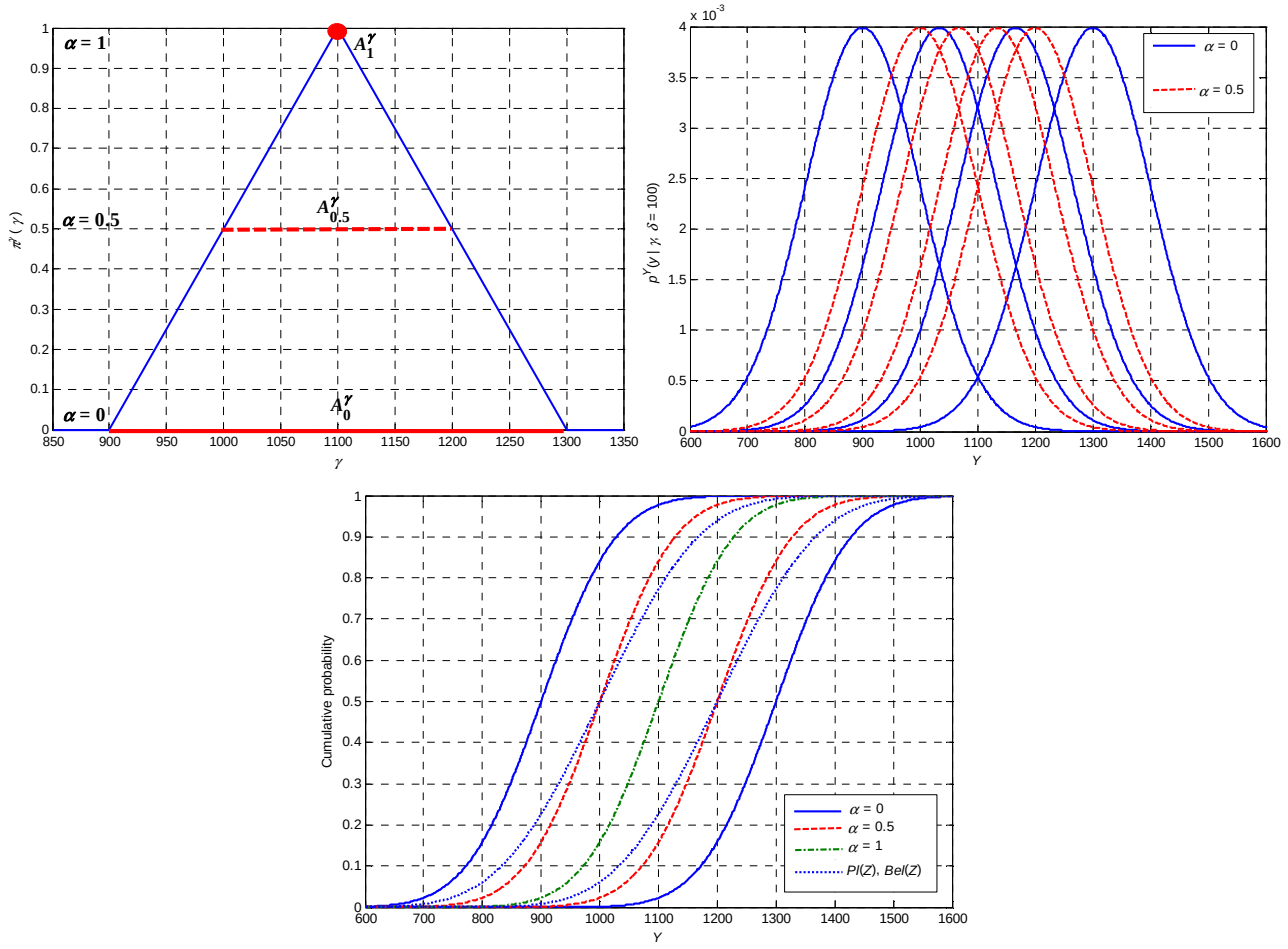


Fig. 1. ***Top left****: triangular possibility distribution* $\pi^\gamma(\gamma)$ *of the epistemically-uncertain parameter* γ *of the Gumbel probability distribution of Y; in evidence the α-cuts of level α = 0, 0.5 and 1.* ***Top right****: four PDFs belonging to the family* $\{p^Y(y \mid \gamma, \delta)\}_{\alpha=0}$ *and four PDFs belonging to the family* $\{p^Y(y \mid \gamma, \delta)\}_{\alpha=0.5}$. ***Bottom****: bounding upper and lower CDFs of Y,* $\overline{F}_\alpha^Y(y) = Pl_\alpha^Y(Z)$ *and* $\underline{F}_\alpha^Y(y) = Bel_\alpha^Y(Z)$, *Z = (–∞, y], built in correspondence of the α-cuts of level α = 0, 0.5 and 1 of* $\pi^\gamma(\gamma)$; *the plausibility and belief functions* $Pl^Y(Z)$ *and* $Bel^Y(Z)$, *Z = (–∞, y], are also shown*

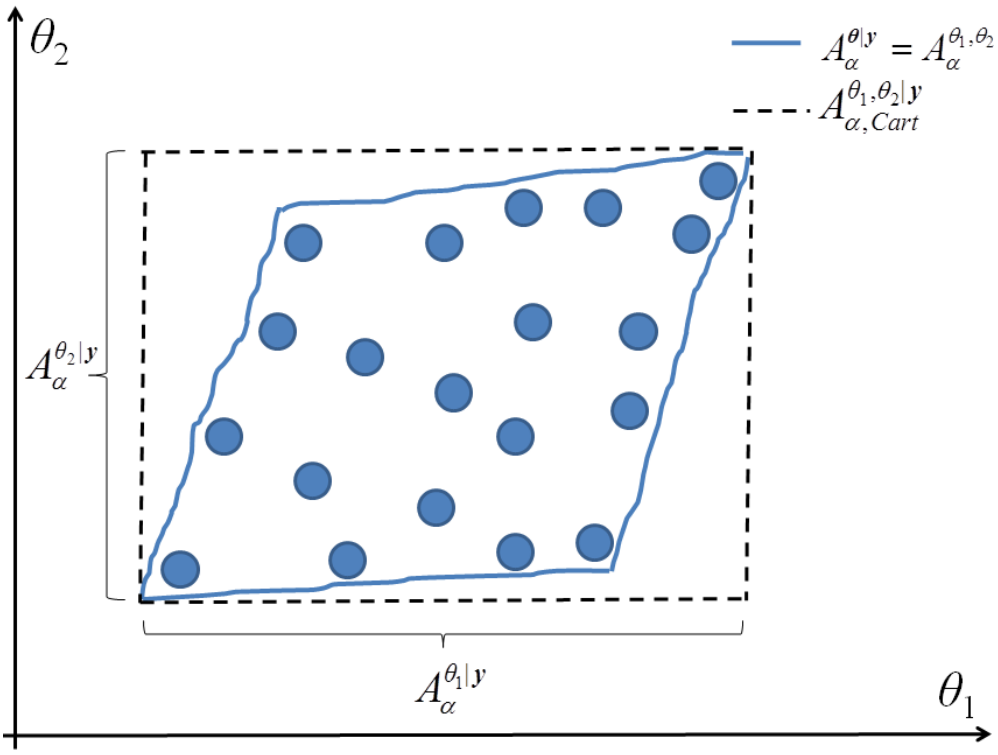
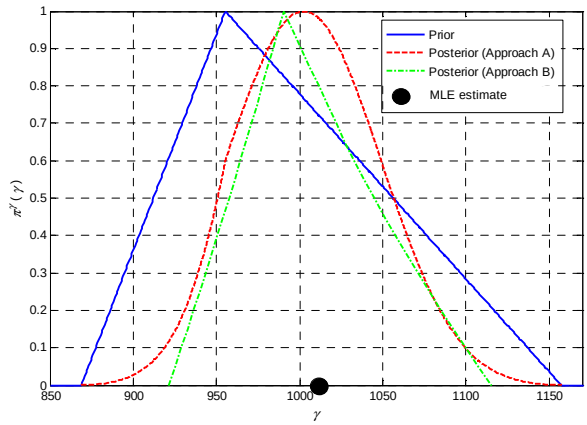
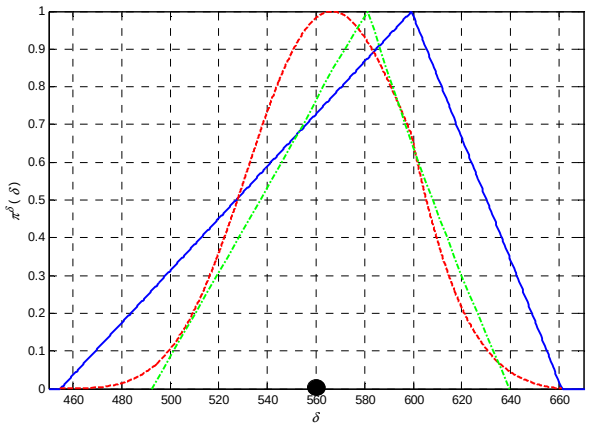


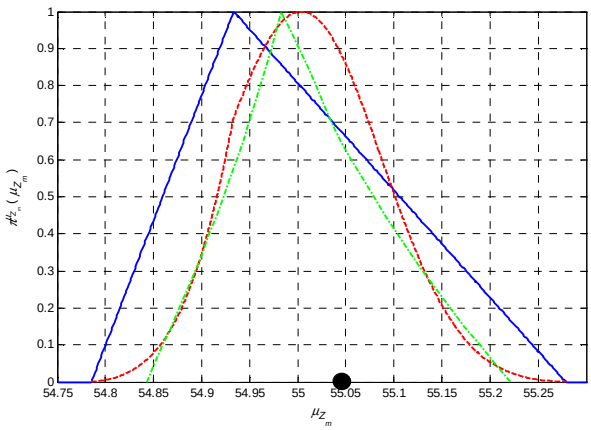
Fig. 2. *Identification of the P-dimensional α-cut* $A_\alpha^{\boldsymbol{\theta}|\boldsymbol{y}}$ *of the posterior possibility distribution of* $\boldsymbol{\theta}$ *as the hull enveloping T = 20 point estimates* $\boldsymbol{\theta}_{l,\alpha} \mid \boldsymbol{y}$, *l = 1, 2, …, 20 (dots), generated by the repeated application of the Bayes' theorem in the hybrid probabilistic-possibilistic approach (B), in the particular case of P = 2 parameters* θ_1 *and* θ_2. *The corresponding α-cut* $A_{\alpha,Cart}^{\theta_1,\theta_2|\boldsymbol{y}}$ *generated by the Cartesian product of the (one-dimensional) α-cuts* $A_\alpha^{\theta_1|\boldsymbol{y}}$ *and* $A_\alpha^{\theta_2|\boldsymbol{y}}$ *of the marginal possibility distributions is also shown*

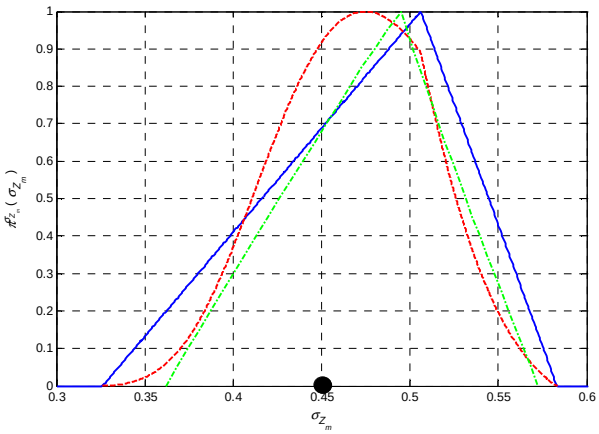
1184



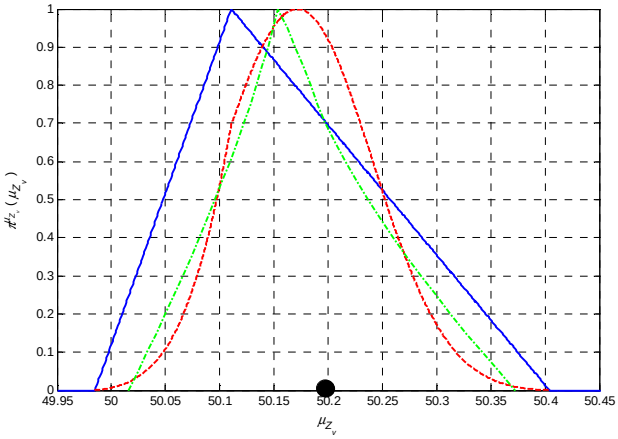


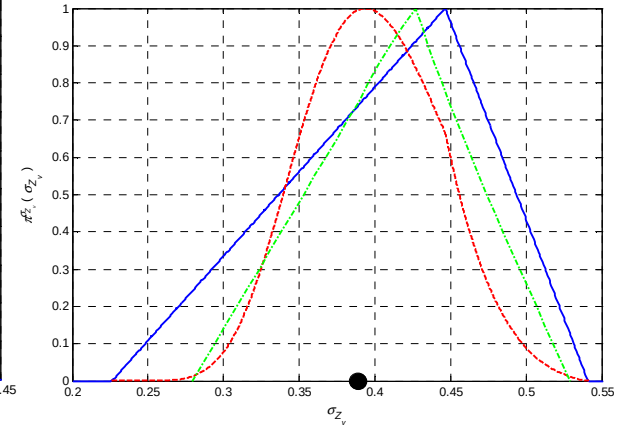
1185



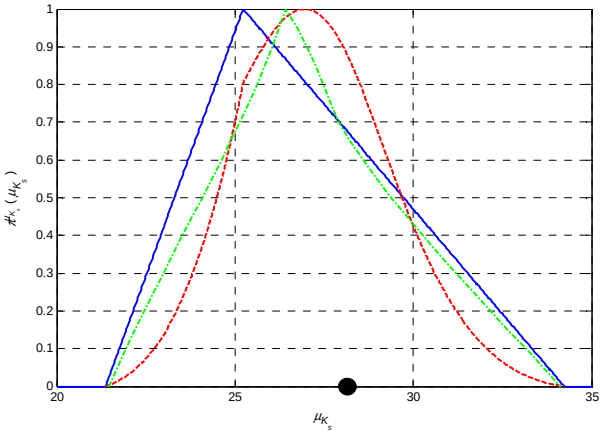


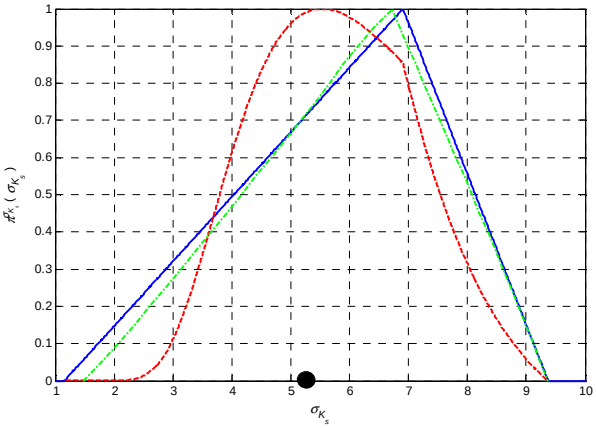
1186





1187





1188

Fig. 3. *Prior and posterior possibility distributions of the epistemically-uncertain parameters of the aleatory PDFs of* $Y_1 = Q$ ***(top row, left and right column)****,* $Y_2 = Z_m$ ***(middle-top row, left and right column)****,* $Y_3 = Z_v$ ***(middle-bottom row, left and right column)*** *and* $Y_4 = K_s$ ***(bottom row, left and right column)****. The point estimates of the parameters obtained by the classical MLE method are also shown for comparison. Adapted and extended from [Pedroni et al., 2014]*

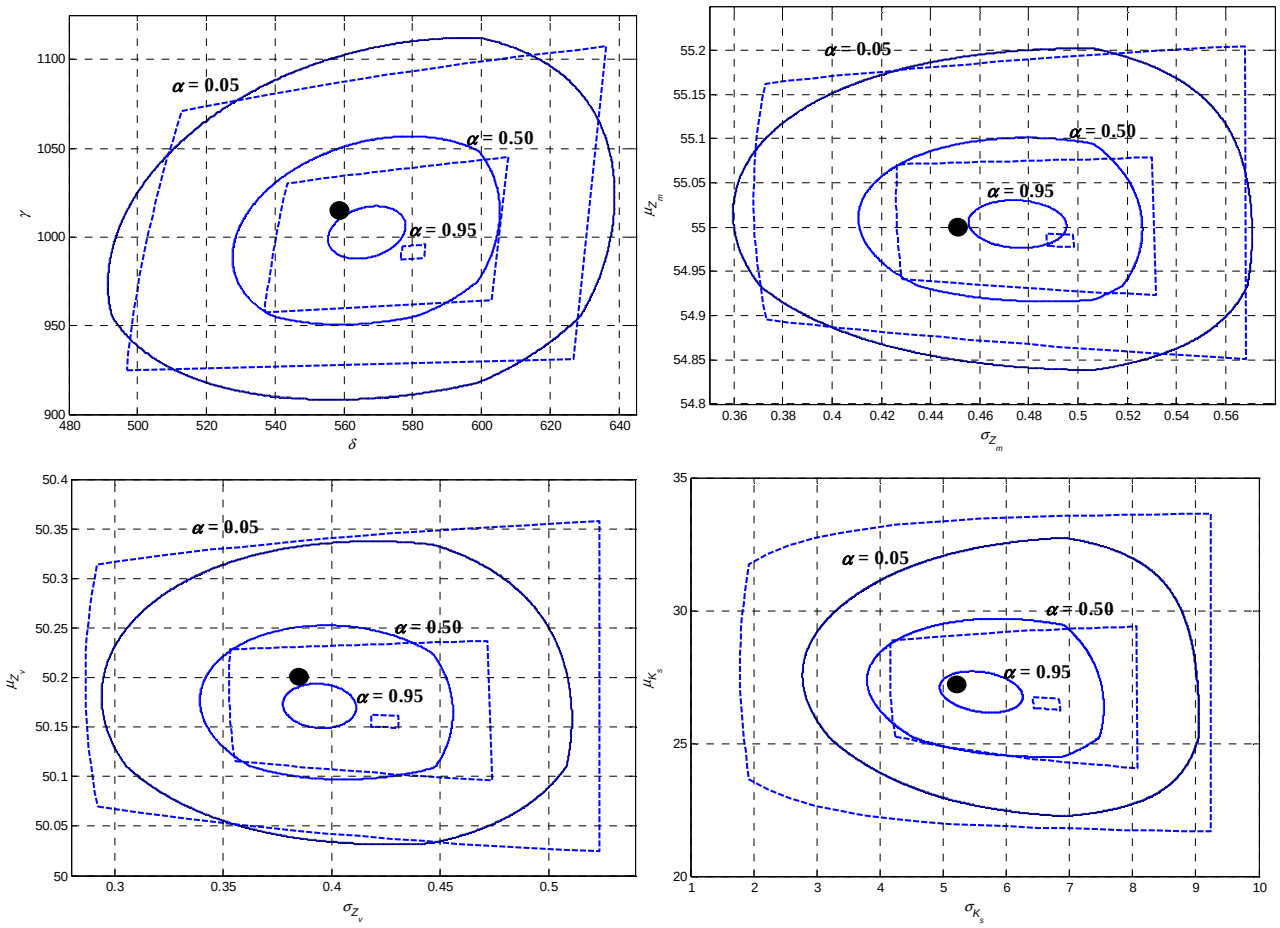


Fig. 4. *Exemplary α-cuts of level α = 0.05, 0.50 and 0.95 of the joint posterior possibility distributions of the parameters of the PDFs of $Y_1 = Q$ (**top left**), $Y_2 = Z_m$ (**top right**), $Y_3 = Z_v$ (**bottom left**) and $Y_4 = K_s$ (**bottom right**), produced by Approaches A (solid lines) and B (dashed lines)*

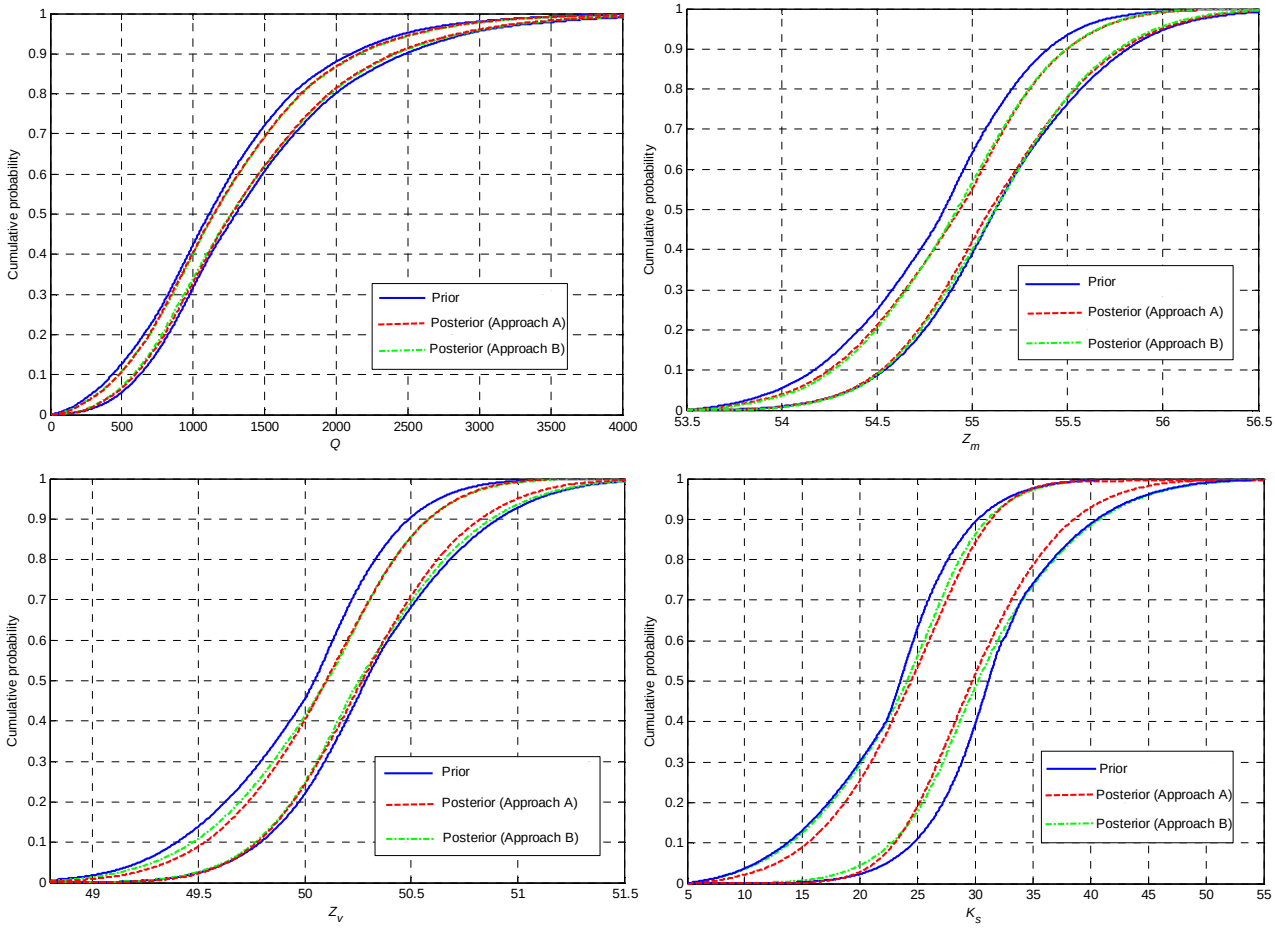


Fig. 5*. Upper and lower CDFs, $\overline{F}^{Y_j}(y_j)$ and $\underline{F}^{Y_j}(y_j)$, j = 1, 2, 3, 4, of the uncertain input variables $Y_1 = Q$ (**top left**), $Y_2 = Z_m$ (**top right**), $Y_3 = Z_v$ (**bottom left**) and $Y_4 = K_s$ (**bottom right**) before and after the Bayesian update performed by Approaches A and B*

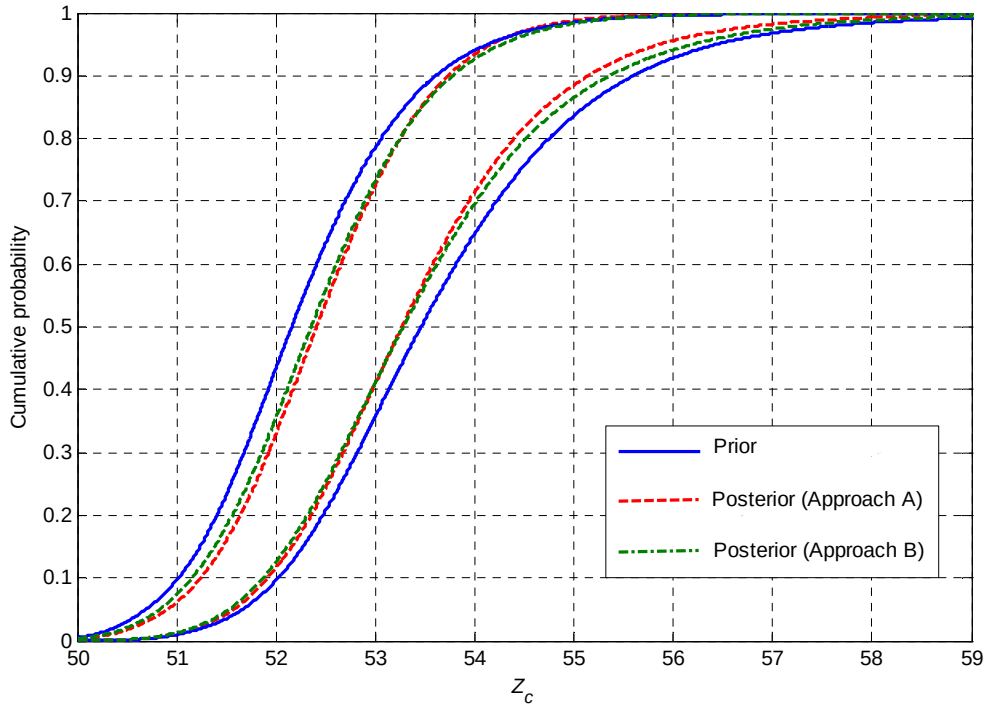


***Fig. 6**. Plausibility and belief functions, $Pl^{Z_c}((-\infty, z_c])$ and $Bel^{Z_c}((-\infty, z_c])$, of the maximal water level of the river Z_c before and after the Bayesian update performed by Approaches A and B. Adapted and extended from [Pedroni et al., 2014]*

Tables

	Purely possibilistic method (A)	Hybrid probabilistic-possibilistic method (B)
Prior	Possibility distribution	Fuzzy Probability Distribution Function
Likelihood	Possibility distribution: - Probability-possibility transformation in this paper (e.g., normalization: see also [Denoeux, 2014])	Probabilistic function
Posterior	Possibility distribution	Fuzzy Probability Distribution Function: - Expected value: possibility distribution

Table 1*. Characteristics of the purely possibilistic (A) and the hybrid probabilistic and possibilistic (B) approaches*

		Purely possibilistic method (Approach A)	
	Aleatory uncertainty	**Epistemic uncertainty (Purely possibilistic priors)**	
Q	$Gum(\gamma,\delta)$	$\pi^{\gamma}(\gamma) = TR(a_{\gamma}, c_{\gamma}, b_{\gamma}) = TR(869, 955, 1157)$	
		$\pi^{\delta}(\delta) = TR(a_{\delta}, c_{\delta}, b_{\delta}) = TR(455, 600, 660)$	
$\boldsymbol{Z_m}$	$N(\mu_{Zm},\sigma_{Zm})$	$\pi^{\mu_{Zm}}(\mu_{Zm}) = TR(a_{\mu_{Zm}}, c_{\mu_{Zm}}, b_{\mu_{Zm}}) = TR(54.78, 54.93, 55.28)$	
		$\pi^{\sigma_{Zm}}(\sigma_{Zm}) = TR(a_{\sigma_{Zm}}, c_{\sigma_{Zm}}, b_{\sigma_{Zm}}) = TR(0.33, 0.51, 0.58)$	
$\boldsymbol{Z_v}$	$N(\mu_{Zv},\sigma_{Zv})$	$\pi^{\mu_{Zv}}(\mu_{Zv}) = TR(a_{\mu_{Zv}}, c_{\mu_{Zv}}, b_{\mu_{Zv}}) = TR(49.98, 50.11, 50.40)$	
		$\pi^{\sigma_{Zv}}(\sigma_{Zv}) = TR(a_{\sigma_{Zv}}, c_{\sigma_{Zv}}, b_{\sigma_{Zv}}) = TR(0.23, 0.45, 0.54)$	
$\boldsymbol{K_s}$	$N(\mu_{Ks},\sigma_{Ks})$	$\pi^{K_s}(\mu_{K_s}) = TR(a_{\mu_{Ks}}, c_{\mu_{Ks}}, b_{\mu_{Ks}}) = TR(21.37, 25.23, 34.23)$	
		$\pi^{\sigma_{Ks}}(\sigma_{Ks}) = TR(a_{\sigma_{Ks}}, c_{\sigma_{Ks}}, b_{\sigma_{Ks}}) = TR(1.16, 6.91, 9.37)$	
		Hybrid probabilistic-possibilistic method (Approach B)	
	Aleatory uncertainty	**Epistemic uncertainty (Fuzzy Probability Density Functions - FPDFs)**	
		Probabilistic part	**Possibilistic part**
Q	$Gum(\gamma,\delta)$	$p^{\gamma}(\gamma) = N(\mu_{\gamma},\sigma_{\gamma})$	$\pi^{\mu_{\gamma}}(\mu_{\gamma}) = TR(a_{\mu_{\gamma}}, c_{\mu_{\gamma}}, b_{\mu_{\gamma}}) = TR(869, 955, 1157),$
			$\pi^{\sigma_{\gamma}}(\sigma_{\gamma}) = TR(a_{\sigma_{\gamma}}, c_{\sigma_{\gamma}}, b_{\sigma_{\gamma}}) = TR(39, 52, 57)$
		$p^{\delta}(\delta) = N(\mu_{\delta},\sigma_{\delta})$	$\pi^{\mu_{\delta}}(\mu_{\delta}) = TR(a_{\mu_{\delta}}, c_{\mu_{\delta}}, b_{\mu_{\delta}}) = TR(455, 600, 660)$
			$\pi^{\sigma_{\delta}}(\sigma_{\delta}) = TR(a_{\sigma_{\delta}}, c_{\sigma_{\delta}}, b_{\sigma_{\delta}}) = TR(25, 31, 43)$
$\boldsymbol{Z_m}$	$N(\mu_{Zm},\sigma_{Zm})$	$p^{\mu_{Zm}}(\mu_{Zm}) = N(\mu_{\mu_{Zm}},\sigma_{\mu_{Zm}})$	$\pi^{\mu_{\mu_{Zm}}}(\mu_{\mu_{Zm}}) = TR(a_{\mu_{\mu_{Zm}}}, c_{\mu_{\mu_{Zm}}}, b_{\mu_{\mu_{Zm}}}) = TR(54.78, 54.93, 55.28)$
			$\pi^{\sigma_{\mu_{Zm}}}(\sigma_{\mu_{Zm}}) = TR(a_{\sigma_{\mu_{Zm}}}, c_{\sigma_{\mu_{Zm}}}, b_{\sigma_{\mu_{Zm}}}) = TR(0.06, 0.09, 0.11)$
		$p^{\sigma_{Zm}}(\sigma_{Zm}) = N(\mu_{\sigma_{Zm}},\sigma_{\sigma_{Zm}})$	$\pi^{\mu_{\sigma_{Zm}}}(\mu_{\sigma_{Zm}}) = TR(a_{\mu_{\sigma_{Zm}}}, c_{\mu_{\sigma_{Zm}}}, b_{\mu_{\sigma_{Zm}}}) = TR(0.33, 0.51, 0.58)$
			$\pi^{\sigma_{\sigma_{Zm}}}(\sigma_{\sigma_{Zm}}) = TR(a_{\sigma_{\sigma_{Zm}}}, c_{\sigma_{\sigma_{Zm}}}, b_{\sigma_{\sigma_{Zm}}}) = TR(0.03, 0.04, 0.06).$
$\boldsymbol{Z_v}$	$N(\mu_{Zv},\sigma_{Zv})$	$p^{\mu_{Zv}}(\mu_{Zv}) = N(\mu_{\mu_{Zv}},\sigma_{\mu_{Zv}})$	$\pi^{\mu_{\mu_{Zv}}}(\mu_{\mu_{Zv}}) = TR(a_{\mu_{\mu_{Zv}}}, c_{\mu_{\mu_{Zv}}}, b_{\mu_{\mu_{Zv}}}) = TR(49.98, 50.11, 50.40)$
			$\pi^{\sigma_{\mu_{Zv}}}(\sigma_{\mu_{Zv}}) = TR(a_{\sigma_{\mu_{Zv}}}, c_{\sigma_{\mu_{Zv}}}, b_{\sigma_{\mu_{Zv}}}) = TR(0.04, 0.08, 0.10),$
		$p^{\sigma_{Zv}}(\sigma_{Zv}) = N(\mu_{\sigma_{Zv}},\sigma_{\sigma_{Zv}})$	$\pi^{\mu_{\sigma_{Zv}}}(\mu_{\sigma_{Zv}}) = TR(a_{\mu_{\sigma_{Zv}}}, c_{\mu_{\sigma_{Zv}}}, b_{\mu_{\sigma_{Zv}}}) = TR(0.23, 0.45, 0.54)$
			$\pi^{\sigma_{\sigma_{Zv}}}(\sigma_{\sigma_{Zv}}) = TR(a_{\sigma_{\sigma_{Zv}}}, c_{\sigma_{\sigma_{Zv}}}, b_{\sigma_{\sigma_{Zv}}}) = TR(0.02, 0.04, 0.08)$
$\boldsymbol{K_s}$	$N(\mu_{Ks},\sigma_{Ks})$	$p^{\mu_{Ks}}(\mu_{Ks}) = N(\mu_{\mu_{Ks}},\sigma_{\mu_{Ks}})$	$\pi^{\mu_{\mu_{Ks}}}(\mu_{\mu_{Ks}}) = TR(a_{\mu_{\mu_{Ks}}}, c_{\mu_{\mu_{Ks}}}, b_{\mu_{\mu_{Ks}}}) = TR(21.37, 25.23, 34.23)$
			$\pi^{\sigma_{\mu_{Ks}}}(\sigma_{\mu_{Ks}}) = TR(a_{\sigma_{\mu_{Ks}}}, c_{\sigma_{\mu_{Ks}}}, b_{\sigma_{\mu_{Ks}}}) = TR(0.50, 2.80, 3.79)$
		$p^{\sigma_{Ks}}(\sigma_{Ks}) = N(\mu_{\sigma_{Ks}},\sigma_{\sigma_{Ks}})$	$\pi^{\mu_{\sigma_{Ks}}}(\mu_{\sigma_{Ks}}) = TR(a_{\mu_{\sigma_{Ks}}}, c_{\mu_{\sigma_{Ks}}}, b_{\mu_{\sigma_{Ks}}}) = TR(1.16, 6.91, 9.37)$
			$\pi^{\sigma_{\sigma_{Ks}}}(\sigma_{\sigma_{Ks}}) = TR(a_{\sigma_{\sigma_{Ks}}}, c_{\sigma_{\sigma_{Ks}}}, b_{\sigma_{\sigma_{Ks}}}) = TR(0.10, 0.86, 2.64)$

Table 2*. Characteristics and parameters of the prior distributions of the uncertain variables Y_1 = Q, $Y_2 = Z_m$, $Y_3 = Z_v$ and $Y_4 = K_s$ used in Approaches A and B, respectively*

			Marginal possibility distributions update					
			Most likely value, c_θ (d_θ^{MLE})			Area, S_θ (R_θ)		
Variable	θ	**MLE** $\hat{\theta}^{MLE}$	**Prior**	**Posterior A**	**Posterior B**	**Prior**	**Posterior A**	**Posterior B**
Q [m^3/s]	γ	1013.21	955.55	1002.70 (0.010)	990.42 (0.023)	144.65	100.55 (30.49)	96.90 (33.01)
	δ	558.48	599.15	566.35 (0.014)	581.40 (0.041)	103.35	76.94 (25.56)	73.65 (28.74)
Z_m [m]	μ_{Zm}	55.03	54.93	55.00 (5.45e-4)	54.98 (9.09e-4)	0.25	0.190 (24.00)	0.195 (22.00)
	σ_{Zm}	0.45	0.51	0.47 (0.044)	0.50 (0.111)	0.12	0.110 (8.33)	0.115 (4.17)
Z_v [m]	μ_{Zv}	50.19	50.11	50.17 (3.98e-4)	50.15 (7.97e-4)	0.21	0.165 (21.43)	0.170 (19.05)
	σ_{Zv}	0.38	0.45	0.39 (0.026)	0.43 (0.132)	0.16	0.121 (24.38)	0.125 (21.88)
K_s [/]	μ_{Ks}	27.80	25.24	26.95 (0.031)	26.43 (0.049)	6.45	5.40 (16.28)	5.80 (10.08)
	σ_{Ks}	5.26	6.89	5.54 (0.053)	6.72 (0.278)	4.11	3.75 (8.76)	3.95 (3.89)

Table 3. *Most likely values c_θ of the parameters $\theta = \gamma$, δ, μ_{Zm}, σ_{Zm}, μ_{Zv}, σ_{Zv}, μ_{K_s} and σ_{Ks} of the aleatory PDFs of the input variables $Y_1 = Q$, $Y_2 = Z_m$, $Y_3 = Z_v$ and $Y_4 = K_s$ and areas S_θ underlying the corresponding (marginal) possibility distributions before and after the Bayesian update performed by Approaches A and B. The point estimates $\hat{\theta}^{MLE}$ obtained by the classical MLE method are also shown for comparison together with the values of the quantitative indicators d_θ^{MLE} (5) and R_θ (6) (in parentheses). Adapted and extended from [Pedroni et al., 2014]*

INPUT VARIABLES				
Variable	**Indicator**	**Prior**	**Purely possibilistic approach (A) (% width reduction)**	**Hybrid approach (B) (% width reduction)**
$Y_1 = Q$	$Q^{0.05}$	[284.8, 480.9]	[329.3, 446.3] (0.4033)	[322.9, 434.1] (0.4329)
	$Q^{0.50}$	[1110.5, 1295.5]	[1143.6, 1269.1] (0.3216)	[1145.5, 1263.7] (0.3610)
	$Q^{0.95}$	[2478.2, 2917.3]	[2553.4, 2857.9] (0.3065)	[2571.1, 2879.0] (0.2987)
$Y_2 = Z_m$	$Z_m^{0.05}$	[53.98, 54.36]	[54.09, 54.37] (0.2631)	[54.06, 54.36] (0.2105)
	$Z_m^{0.50}$	[54.85, 55.12]	[54.92, 55.09] (0.4074)	[54.93, 55.12] (0.2592)
	$Z_m^{0.95}$	[55.56, 56.02]	[55.66, 55.96] (0.3478)	[55.67, 55.99] (0.3043)
$Y_3 = Z_v$	$Z_v^{0.05}$	[49.22, 49.64]	[49.35, 49.64] (0.3095)	[49.29, 49.62] (0.2142)
	$Z_v^{0.50}$	[50.04, 50.29]	[50.10, 50.25] (0.4000)	[50.10, 50.27] (0.3199)
	$Z_v^{0.95}$	[50.63, 51.09]	[50.72, 51.00] (0.3913)	[50.72, 51.06] (0.2608)
$Y_4 = K_s$	$K_s^{0.05}$	[10.95, 22.33]	[12.75, 21.25] (0.2530)	[11.15, 20.38] (0.1889)
	$K_s^{0.50}$	[23.49, 31.15]	[24.43, 29.73] (0.3080)	[24.05, 30.22] (0.1945)
	$K_s^{0.95}$	[32.66, 44.14]	[33.01, 41.40] (0.2691)	[33.32, 44.14] (0.0574)

Table 4*. Intervals* $\left[\left[\overline{F}^{Y_j}\right]^{-1}(\beta),\ \left[\underline{F}^{Y_j}\right]^{-1}(\beta)\right]$ *for the β·100-th percentiles* Y_j^{β}*, β = 0.05, 0.50 and 0.95, of the input variables* Y_j*, j = 1, 2, 3, 4, before and after the Bayesian update performed by Approaches A and B. The percentage reduction in the width of the intervals produced by Approaches A and B is also shown in parentheses for comparison purposes*

OUTPUT VARIABLE Z_c				
Variable	**Indicator**	**Prior**	**Purely possibilistic approach (A) (% width reduction)**	**Hybrid approach (B) (% width reduction)**
Z_c	$Z_c^{0.05}$	[50.70, 51.67]	[50.90, 51.56] (0.3196)	[50.84, 51.53] (0.2887)
	$Z_c^{0.50}$	[52.16, 53.46]	[52.38, 53.23] (0.3462)	[52.36, 53.28] (0.2923)
	$Z_c^{0.95}$	[54.13, 56.44]	[54.21, 55.87] (0.2814)	[54.28, 56.20] (0.1688)
	$P[Z_c > z_c^*]$	[0.8908, 0.9946]	[0.9284, 0.9921] (0.3863)	[0.9103, 0.9946] (0.1879)

***Table 5**. Intervals $\left[\left[Pl^{Z_c}\right]^{-1}(\beta),\ \left[Bel^{Z_c}\right]^{-1}(\beta)\right]$ for the β·100-th percentiles Z_c^{β}, β = 0.05, 0.50 and 0.95, of the maximal water level of the river Z_c (i.e., the model output) and intervals $[1-\overline{F}^{Z_c}(z_c^*), 1-\underline{F}^{Z_c}(z_c^*)]$ for the exceedance probability $P[Z_c > z_c^* = 55.5m]$, before and after the Bayesian update performed by Approaches A and B. The percentage reduction in the width of the intervals produced by Approaches A and B is also shown in parentheses for comparison purposes. Adapted and extended from [Pedroni et al., 2014]*

	Approach A	Approach B (T = 100, N_α = 21)
Computational time t_{comp} [s] on a Intel® Core™2 Duo CPU E7600 @ 3.06 and 3.07 GHz	0.73	1536.65

***Table 6**. Computational time t_{comp} required by the Bayesian update of all the parameters of the PDFs of $Y_1 = Q$, $Y_2 = Z_m$, $Y_3 = Z_v$ and $Y_4 = K_s$ performed by Approach A and by Approach B with $T = 100$ repetitions of the purely probabilistic Bayes' theorem for each of the $N_\alpha = 21$ α-cuts analyzed*

	Purely possibilistic approach (A)	Hybrid probabilistic-possibilistic approach (B)
Advantages	- High strength in updating the possibilistic parameters independently on the size of the data set available - Computationally cheap	- Well-established and rigorous theoretical framework (i.e., probabilistic Bayes' theorem)
Drawbacks	- Necessity to "build" a possibilistic likelihood (e.g., by probability-possibility transformations, normalization or directly from rough experimental data)	- Necessity to "build" a (fictitious) Fuzzy Probability Distribution Functions as a prior - High strength in updating the possibilistic parameters only when the size of the data set available is large (e.g., > 100) - Computationally burdensome

Table 7*. Advantages and drawbacks of the purely possibilistic (A) and the hybrid probabilistic and possibilistic (B) approaches*