



# **A comparison between probabilistic and Dempster-Shafer Theory approaches to Model Uncertainty Analysis in the Performance Assessment of Radioactive Waste Repositories**

Piero Baraldi, Enrico Zio

## **► To cite this version:**

Piero Baraldi, Enrico Zio. A comparison between probabilistic and Dempster-Shafer Theory approaches to Model Uncertainty Analysis in the Performance Assessment of Radioactive Waste Repositories. Risk Analysis, 2010, 30 (7), pp.1139-1156. <10.1111/j.1539-6924.2010.01416.x>. <hal-00610490>

**HAL Id: hal-00610490**

**<https://centralesupelec.hal.science/hal-00610490v1>**

Submitted on 25 Jul 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

# **A comparison between probabilistic and Dempster-Shafer Theory approaches to Model Uncertainty Analysis in the Performance Assessment of Radioactive Waste Repositories**

Piero Baraldi, Enrico Zio

*Energy Department, Polytechnic of Milan, Via Ponzio 34/3 20133 Milan, Italy*

## ABSTRACT

Model uncertainty is a primary source of uncertainty in the assessment of the performance of repositories for the disposal of nuclear wastes, due to the complexity of the system and the large spatial and temporal scales involved.

This work considers multiple assumptions on the system behaviour and corresponding alternative plausible modelling hypotheses. To characterize the uncertainty in the correctness of the different hypotheses, the opinions of different experts are treated probabilistically or, in alternative, by the belief and plausibility functions of the Dempster-Shafer theory.

A comparison is made with reference to a flow model for the evaluation of the hydraulic head distributions present at a radioactive waste repository site. Three experts are assumed available for the evaluation of the uncertainties associated to the hydro-geological properties of the repository and the groundwater flow mechanisms.

**KEYWORDS:** MODEL UNCERTAINTY, DEMPSTER-SHAFER THEORY, RADIOACTIVE WASTE REPOSITORIES

## Acronyms

|     |                               |
|-----|-------------------------------|
| DS  | Dempster-Shafer               |
| DST | Dempster-Shafer Theory        |
| bpa | basic probability assignment  |
| UOD | Universe Of Discourse         |
| PRA | Probabilistic Risk Assessment |

## Notation

|          |                                                                         |
|----------|-------------------------------------------------------------------------|
| $M_l$    | generic $l$ -th model developed within the model set expansion approach |
| $n$      | total number of possible models                                         |
| $y$      | quantity to be predicted by the model                                   |
| $x$      | input of the model (vector sign omitted for ease of representation)     |
| $S_l$    | structure of the generic $l$ -th model                                  |
| $F_l(y)$ | cumulative distribution of $y$ predicted by the model $M_l$             |
| $F(y)$   | Bayesian estimator of the cumulative distribution of $y$                |
| $P(M_l)$ | probability that the model $M_l$ is the correct one                     |
| $p_l$    | parameter of the categorical distribution indicating $P(M_l)$           |
| $f(p_l)$ | probability density function of $p_l$                                   |
| $k$      | number of assumptions on which the model $M$ is based                   |
| $H^i$    | $i$ -th assumption on which a model is based                            |
| $X^i$    | Universe Of Discourse of the assumption $H^i$                           |
| $n_i$    | number of available alternative hypotheses for the $i$ -th assumption   |
| $H_j^i$  | $j$ -th alternative hypothesis for the $i$ -th assumption               |

$p(H_j^i)$  probability that the  $j$ -th alternative hypothesis for the  $i$ -th assumption is the correct one

$z$  generic uncertain variable

$Z$  Universe of Discourse of  $z$

$P(\cdot)$  power set

$A$  generic element of  $P(\cdot)$

## 1. INTRODUCTION

Model uncertainty is considered a primary source of uncertainty in the assessment of the performance of repositories for the disposal of nuclear wastes. By definition, a model is a representation of a real system and uncertainties inevitably arise whenever there are possible alternative interpretations of the system and its phenomena, which are all plausible in light of the current knowledge of the system.

Methods to deal with model uncertainty include prediction expansion, model set expansion and formal Bayesian approaches. In prediction expansion <sup>(1)</sup>, a single model is chosen as the best one to represent the system. However, it is recognized that this model has drawbacks and may represent only some characteristics of the system better than others. Sensitivity studies are performed on the various assumptions to analyze the effects on the model output. The associated uncertainty is dealt with by applying a random adjustment factor to the model results. The adjustment factor may be multiplicative or additive, or both. However, it is not clear if the technical exercise of quantifying the gap between reality and the model (i.e., the quantification of the adjustment factor distribution) is feasible in practical analyses. Furthermore, a question arises on the reasoning behind this approach: if information exists that leads the expert to say something about the error in the prediction of a model, its correction can be simply considered as a modification of the original model <sup>(2,3)</sup>.

In model set expansion, the characteristics of the system under consideration are analyzed and models are created in an attempt to emulate the system based on goodness-of-fit criteria <sup>(4,5)</sup>. The models may use different assumptions and require different inputs. Each model has its own advantages and disadvantages, including limitations on and ranges of applicability. These models are then combined to produce a meta-model of the system.

Several methods have been proposed regarding the construction of this meta-model. They include mixture <sup>(3)</sup>, the NUREG-1150 approach <sup>(6)</sup>, the joint US/EC Probabilistic Accident Consequence Uncertainty Analysis (PACUA) approach <sup>(7)</sup>, and the Technical Facilitator-Integrator approach <sup>(8)</sup>. All of these methods rely on expert opinion. This is particularly inevitable in the case of interest here of the performance assessment of radioactive waste repositories, for which the predictive capabilities of the models cannot be verified over the time frames and spatial scales which they are required to apply to. As a consequence, the analyst cannot obtain empirical confirmation of the validity of a model from observations, so that the evaluation of the model must rely exclusively on the subjective interpretation of the information available at the time of the analysis. This leads to the conclusion that any attempt to address the issue of model uncertainty in a quantitative manner will rely on the use of expert judgment. In the mixture approach, the set of plausible models and their probabilities of being correct are agreed upon by the experts. The output distributions of the models are then linearly combined, with weights corresponding to the probabilities of correctness. The model distributions should be presented to the analysts before they are combined, allowing a transparent look at the range of variability that are combined into the meta-model. In some works, the model set expansion approach has been framed within a Bayesian approach to the estimation of the required parameters and combination of the individual model outputs <sup>(9-11)</sup>. These approaches are theoretically very attractive due to their mathematical rigor and ability to incorporate both objective and subjective information in a probabilistic representation.

On the other hand, it has been argued that since a model is a simplification of reality, it is possible to conclude that “every model is definitely false” <sup>(12)</sup>; therefore, it seems inappropriate to evaluate the probability that a model produces outputs that replicate perfectly the reality, because this is impossible by definition: the probability of a model being more or less correct can only be interpreted in relative terms, i.e. as the probability that the model produces outputs more or less close to reality, in comparison with other models. In this view, the alternative models can be included into a meta-model parameterized with an index parameter  $\varphi$  whose values  $(1,2,\dots,n)$  are associated to the different plausible models ( $M_1, M_2, \dots, M_n$ ) and the uncertainty about the model can be converted into uncertainty about the value of  $\varphi$ , represented by a given probability distribution. In practice, probability distributions can be assessed over the appropriate model structure and reinterpreted as associated to the model index parameter; this also allows comparing the impact of the uncertainty about the model with other uncertainties.

An alternative to this probabilistic representation of the uncertainty on the model is offered by a representation based on the belief and plausibility functions of the Dempster-Shafer Theory (DST). This approach requires an expert to provide for each plausible model an interval of values, limited by a lower bound, called belief, representing the amount of belief that directly supports a given model and an upper bound, called plausibility and measuring the fact that the model could possibly be true “up to that value” because there is only so much evidence that contradicts it. The Dempster rule of combination can then be used for the aggregation of the uncertainties of different experts on the correctness of the models. Actually, given that the models emerge from multiple modelling assumptions on the various aspects of system behaviour, the uncertainty analysis is made at the hypothesis level and then propagated into the uncertainty characterizing the entire model.

Finally, more powerful, fully formal Bayesian approaches have also been proposed for the treatment of model uncertainty. For example, in <sup>(13)</sup> the available models are treated as sources of information that can be used to predict the true value of the model output through the application of Bayes theorem. The proposed methodology allows for utilizing various types of information, including experimental data and subjective evidence and for updating the prior distribution of the model output when new information becomes available <sup>(11)</sup>.

In the present work, the model set expansion approach to model uncertainty has been embraced to handle a situation in which multiple assumptions are made on the behaviour of the system, giving rise to corresponding alternative plausible modelling hypotheses whose correctness is judged by different experts; a comparison of the classical probabilistic and Dempster-Shafer theory of Evidence frameworks of uncertainty representation is performed. Formal Bayesian methods are not considered in the comparison since the focus is on the proper setting of the Dempster-Shafer representation and it is assumed that all the information on the correctness of the hypotheses is available at the moment of the analysis, with no need of subsequently updating the probability distribution of the model output.

With respect to the application of the Dempster-Shafer theory of Evidence framework within the model set expansion approach, the belief and plausibility functions of the Dempster-Shafer theory of Evidence are used for the representation of the uncertainty in the correctness of the different hypotheses; then, the Dempster rule of combination is used for *i)* the aggregation of the uncertainties of different experts on the correctness of the hypotheses and *ii)* the propagation of the uncertainties regarding the individual assumptions into the uncertainty characterizing the entire model.

The proposed approach is applied to a groundwater flow model for the evaluation of the hydraulic head distributions present in the far field region of a radioactive waste repository. Two classes of model assumptions are required, regarding the hydro-geological properties of

the media in which the repository is located and the groundwater flow mechanisms. Three fictitious experts evaluate the associated uncertainties: the first one gives an evaluation of the correctness of the plausible alternative hypotheses for the first assumption on the geological properties, the second one for the second assumption on the groundwater flow mechanisms and the third one gives directly an opinion on the correctness of the overall groundwater flow models developed.

The paper is organized as follows. In Section 2, a brief overview of the model set expansion is given. Section 3 discusses the representation of model uncertainty by the probabilistic and the Dempster-Shafer (DS) frameworks. In Section 4, the problems of aggregating the uncertainties of different experts and of propagating the uncertainties in the model assumptions onto the uncertainty characterizing the entire model are addressed by a probabilistic and a DS approach. In Section 5 the proposed uncertainty representation, aggregation and propagation techniques are applied to a groundwater flow model for the evaluation of the hydraulic head distributions present at a radioactive waste repository. Some conclusions on the advantage and limitation of the proposed method are drawn in the last Section.

## 2. MODEL SET EXPANSION APPROACH

Consider the objective of expressing the uncertainty associated with the model prediction of the uncertain scalar output  $y$  of a system, due to the existence of a set of  $n$  models  $\{M_l\}$ , whose alternative structures  $\{S_l\}$ ,  $l=1,...,n$  are based on different hypotheses all plausible in light of the existing information on the system behaviour. The distribution  $F_l(y) = F(y|M_l)$  of  $y$  provided by the generic model,  $M_l$ , is conditional on the structure of the model as well as on the values of its internal parameters.

The family of distributions  $\{F_l(y)\}$  represents the uncertainty in the unknown  $y$  due to uncertainty in the models' structure; it can be probabilistically combined in a summary measure by means of a standard Bayesian approach <sup>(2,14)</sup>. Indicating with  $p(M_l)$  the epistemic probability which expresses the analyst's confidence in the set of assumptions underpinning the  $l$ -th model (here the term epistemic is used to characterize the uncertainty due to incomplete state of knowledge), the unconditional aleatory distribution  $F(y)$  is given by the standard Bayesian estimator:

$$F(y) = \sum_{l=1}^n F_l(y) p(M_l) \quad (1)$$

This average value has often encountered objections when employed for decision-making. The argument is that the decision maker should be aware of the full epistemic uncertainties  $\{p(M_l)\}$ . Their average value (1) can lead to erroneous decisions, particularly when the epistemic uncertainty is very large, since the average can be greatly affected by high values of the variable even though they may be very unlikely. In general, the entire distribution of the uncertainties should be presented to the decision-maker who may then choose his/her own criteria upon which to base the decisions.

Furthermore, the model set expansion approach is based on two fundamental assumptions: mutual exclusiveness and collective exhaustiveness of the set of models. While the first assumption can be often accepted in practice, excluding the case in which one of the model is a special case of another model, the second is often not met in practice because it requires that a perfect model not only exists but that it also be one of the  $n$  models considered. In general, the complexity of the phenomena is such that the list of plausible models considered is necessarily incomplete. Moreover, progress in understanding the physical laws underpinning

the process under analysis and the increasing computational capabilities are such that models evolve in time.

### 3. REPRESENTATION OF MODEL UNCERTAINTY

In this work, the probabilistic framework for the representation of the uncertainty regarding the probability of model correctness is compared with the Dempster-Shafer framework.

#### 3.1. The probabilistic approach to model uncertainty representation

Within a classical probabilistic approach, the uncertainty on the correctness of the alternative models  $M_l$  is modelled by the distributions of values  $p_l = P(M_l)$ , each one representing the probability that the  $l$ -th alternative is the correct one,  $l = 1, \dots, n$ . This representation can be mathematically formalized through the introduction of a categorical random variable  $M_l$ , whose distribution is completely given by the probabilities associated with each model alternative  $\Pr(M_l = \text{correct}) = P(M_l) = p_l$ , with  $\sum_{l=1}^n P(M_l) = \sum_{l=1}^n p_l = 1$ , according to the assumptions of exclusivity and exhaustivity of the set of models. The values of the  $n-1$  probabilities  $p_l$  defining the distribution of  $M_l$  need to be elicited from the experts; for example, in the simplest case of two alternative models,  $M_1, M_2$ , the experts would need to quantify  $p_1 = P(M_1)$ , i.e. the probability that  $M_1$  is the correct model; the value of  $p_2 = P(M_2)$  is then simply given by  $P(M_2) = 1 - p_1$ .

In practical cases, an expert may not have sufficient information and knowledge to provide fixed, certain values of the probabilities  $p_l$  and may prefer to express uncertainty on such values in terms of a subjective probability distribution  $f(p_1, p_2, \dots, p_{n-1})$ ; in the case of two possible alternatives  $M_1, M_2$ , the expert instead of assigning a fixed value of  $p_1$ , may provide

a probability density function expressing his/her epistemic uncertainty; if for example the expert believes that the value of  $p_1$  is located somewhere between 0.3 and 0.6, with no preference for any value within the range, this uncertainty could be represented by a uniform probability density function in  $[0.3, 0.6]$ . The uncertainty regarding the probabilities  $p_l$  described by the distribution  $f$  is often referred to as second-order uncertainty<sup>(15)</sup>.

The model set expansion approach of Section 2 can be modified to include this new source of epistemic uncertainty. In particular, the standard Bayesian estimator of the output distribution (eq.1) becomes:

$$F(y) = \int \left( \sum_{l=1}^n F_l(y) p_l \right) f(p_1, p_2, \dots, p_{n-1}) dp_1 dp_2 \dots dp_{n-1} \quad (2)$$

### 3.2. The Dempster-Shafer approach to model uncertainty representation

In this Section the DS evidence theory <sup>(16)</sup> is considered for the representation of the epistemic uncertainty affecting the expert knowledge of the probability  $P(M_l)$  that the alternative model  $M_l$ ,  $l=1,...,n$  be correct. In the DS framework, a lower and an upper bound are introduced for representing the uncertainty associated to  $P(M_l)$ . The lower bound, called belief,  $Bel(M_l)$ , represents the amounts of belief that directly supports  $M_l$  at least in part, whereas the upper bound, called plausibility,  $Pl(M_l)$ , measures the fact that  $M_l$  could be the correct model “up to that value” because there is only so much evidence that contradicts it.

From a general point of view, contrary to probability theory which assigns the probability mass to individual elementary events, the theory of evidence makes basic probability assignments (bpa)  $m(A)$  on sets  $A$  (the focal sets) of the power set  $P(Z)$  of the event space  $Z$ , i.e. on sets of outcomes rather than on single elementary events.

In more details,  $m(A)$  expresses the degree of belief that a specific element  $x$  belongs to the set  $A$  only, and not to any subset of  $A$ . The bpa satisfies the following requirements:

$$m : P(Z) \rightarrow [0, 1], \quad m(\emptyset) = 0; \quad \sum_{A \in P(Z)} m(A) = 1 \quad (3)$$

Also, probability theory imposes more restrictive conditions on the specification of the likelihood of events as a result of the requirement that the probabilities of the occurrence and non-occurrence of an event must sum to one.

As a result, while in probability theory a single probability distribution function is introduced to define the probabilities of any event or proposition, represented as a subset of the sample space, in evidence theory there are two measures of likelihood, belief and plausibility: the belief about events and propositions is represented as intervals, bounded by two values, belief and plausibility. The belief in a proposition (set)  $A$  is quantified as the sum of the probability masses assigned to all sets enclosed by it:

$$Bel(A) = \sum_{B \subseteq A} m(B) \quad (4)$$

Hence, it is a lower bound representing the amount of belief that directly supports the proposition at least in part. Plausibility is the sum of the probability masses assigned to all sets whose intersection with the proposition is not empty:

$$Pl(A) = \sum_{B \cap A \neq \emptyset} m(B) \quad (5)$$

Hence, it is an upper bound on the possibility that the proposition could be verified, i.e. it measures the fact that the proposition could possibly be true “up to that value” because there is only so much evidence that contradicts it.

#### **4. THE MULTIPLE-ASSUMPTIONS MULTIPLE-EXPERTS MODEL UNCERTAINTY PROBLEM**

Let us consider a model  $M$  that consists of a set of  $k$  assumptions  $H^i$ ,  $i=1, \dots, k$  where each assumption can assume  $n_i$  alternative plausible hypotheses  $H_j^i$ ,  $j=1, \dots, n_i$ . Combining all the

$n_i$  possible alternative hypotheses for each  $i$ -th assumption,  $n = \prod_{i=1}^k n_i$  model structures

$S_l, l=1, \dots, n$  in principle arise. For example, for a model for computing the neutron flux distribution in a nuclear reactor, there can be an assumption  $H^1$  on the spatial discretization of the reactor core in homogeneous regions, with hypothesis  $H_1^1$  corresponding to the homogeneous core model of only 1 region and  $H_2^1$  corresponding to  $N$  regions, and an assumption  $H^2$  on the number of groups of precursors and delayed neutrons, with  $H_1^2 = 1$

and  $H_2^2 = 5$ . This gives rise to  $n = \prod_{i=1}^{k=2} m_i = 2 \cdot 2 = 4$  model structures:

$$S_1 = (H_1^1, H_1^2), S_2 = (H_2^1, H_1^2), S_3 = (H_1^1, H_2^2), S_4 = (H_2^1, H_2^2).$$

Suppose that  $Q^i$  different experts  $E_q^i, q=1, \dots, Q^i$  give their opinions on the correctness of the alternative hypotheses for each  $i$ -th assumption,  $i=1, \dots, k$ . The objective of this Section is to compare the probabilistic and the Dempster-Shafer approaches in the estimation of the probability of model correctness based on the expert opinions.

A systematic framework is proposed for the representation of the uncertainty associated to the alternative model structures, as evaluated by the model experts with reference to the multiple alternative hypotheses. The framework consists in the following two steps:

- i) the aggregation of the opinions of the different experts on the correctness of a given hypothesis of assumption;
- ii) the propagation of the uncertainties arising from the expert evaluation of the individual assumptions into the uncertainty characterizing the entire model structure.

#### 4.1. Aggregation of expert opinions and propagation of uncertainties in the probabilistic framework

Assume that the  $Q^i$  experts have evaluated the probabilities of correctness of the  $n_i$  alternative hypotheses regarding assumption  $H^i$  in the model. Different approaches exist for the combination of multiple expert opinions. Axiomatic approaches such as the linear opinion pool <sup>(17)</sup> and the logarithmic opinion pool <sup>(18)</sup> are intuitive, easy to understand and calculate; for these reasons, they are often preferred in practice when there is no clear evidence justifying the adoption of a different pooling.

In this work the logarithmic opinion pool is adopted since, contrary to the linear opinion pool, it satisfies the principle of external Bayesianity.

According to the logarithmic opinion pool, the combined probability distribution  $p(H_j^i)$  is:

$$P(H_j^i) = k \prod_{q=1}^{Q^i} [P_q(H_j^i)]^{w_q} \quad j = 1, \dots, n_i \quad (6)$$

where  $P_q(H_j^i)$  is the expert  $q$ -th probability distribution for assumption  $H^i$ ,  $k$  a normalizing constant and  $w_q$  the weight used to represent the relative quality of the different experts. In the simplest case, the experts are viewed as equivalent and the combined distribution is proportional to the geometric mean of the individual distributions. Supposing that  $Q^i = 3$  experts are available, the results of the aggregation is the same whether formula (6) is directly applied to all 3 experts probabilities  $P_q(H_j^i)$  or it is firstly applied to the probabilities given by two experts and then the obtained probability is combined with the remaining expert probability.

Although axiomatic approaches have been largely used for aggregation of expert opinions, it has been pointed out that they cannot satisfy simultaneously a number of seemingly

compelling desiderata <sup>(19)</sup> and no foundationally-based method for determining the weights in their formula <sup>(20)</sup> have been proposed. For these reasons, in the last decades attention has shifted from the axiomatic approaches to the development of combination methods based on a clear Bayesian paradigm for aggregating information from experts <sup>(18)</sup>. However, Bayesian approaches are not considered in this work since the property of external Bayesianity of the logarithmic opinion pool aggregation method is considered sufficient for the purpose of a fair comparison with the DST aggregation; still, future work will be devoted to the analysis of Bayesian approaches for the aggregation of expert opinions in the context of model uncertainty here of interest, which concerns also the propagation of uncertainties on the individual assumptions onto the uncertainty characterizing the entire model structure.

Let us consider, first of all, the case in which there is no epistemic uncertainty on the probability distribution of the  $n_i$  alternative hypotheses  $H^i$  of the  $i$ -th assumption characterized by  $n_i$  discrete values parameters  $p_{ij} = P(H_j^i)$ ,  $j = 1, \dots, n_i$ . Assuming complete independence of the hypotheses, the probability  $P(M_l)$  of model  $S_l = (H_{l_1}^1, H_{l_2}^2, \dots, H_{l_k}^k)$  to be correct is given by:

$$P(M_l) = \prod_{i=1}^k P(H_{l_i}^i) = \prod_{i=1}^k p_{il_i} \quad (7)$$

In case the experts are uncertain on the exact values of the probabilities  $p_{ij}$  and thus provide the joint probability density functions  $f(p_{i1}, p_{i2}, \dots, p_{in_i-1})$ , then also the probability of correctness of model  $M_l$ ,  $p_l = P(M_l)$  is uncertain. From eq.(7), the joint probability density function  $f(p_1, p_2, \dots, p_{n-1})$  can be obtained numerically.

#### 4.2. Aggregation of expert opinions and propagation of uncertainties in the Dempster-Shafer framework

Within the DST, the assignment by the  $Q^i$  independent experts can be aggregated through the Dempster Rule of combination <sup>(16,21)</sup>.

For simplicity, let us consider two experts whose evidence on the value of an uncertain variable is expressed in terms of two sets of basic probability assignments (bpa's),  $m_1(A), m_2(A)$  on the sets  $A$  of the power set  $P(Z)$  of  $Z$ . Assuming that the bpa functions on the frame of discernment are based on independent arguments and bodies of evidence <sup>(16)</sup>, aggregation of this evidence into a joint bpa  $m_{12}(A)$  can be obtained by means of *Dempster rule* <sup>(23)</sup>:

$$\begin{aligned} m_{12}(A) &= \frac{\sum_{B \cap C = A} m_1(B) m_2(C)}{1 - K} \quad \forall A \neq 0 \\ m_{12}(0) &= 0 \end{aligned} \quad (8)$$

where the complementary normalization factor  $K$  is given by

$$K = \sum_{B \cap C = 0} m_1(B) m_2(C) \quad (9)$$

According to <sup>(4)</sup> and <sup>(5)</sup> above, the degree of evidence  $m_1(B)$  regarding focal set  $B \in P(Z)$ , from the first expert and the degree of evidence  $m_2(C)$  focused on focal set  $C \in P(Z)$ , from the second expert, are aggregated by taking their product  $m_1(B) \cdot m_2(C)$  focused on the intersection focal set  $B \cap C = A$ . This way of combining evidence sources is analogous to the way in which in probability theory joint probability density functions (pdfs) are calculated from two independent marginal pdfs, and is thus justified on the same grounds. However, some intersections  $B^i \cap C^i$  of different focal elements  $B^i$  and  $C^i$ , from the first and second

expert, may result in the same set  $A^i$  so that one must sum their product contribution to obtain  $m_{12}(A)$ . Furthermore, some of the intersections may be the empty set, for which  $m_{12}(0) = 0$ . Then, introducing  $K$  as the sum of products  $m_1(B) \cdot m_2(C)$  of all focal elements  $B$  of  $m_1$  and  $C$  of  $m_2$  such that  $B \cap C = 0$ , a normalized joint basic assignment  $m_{12}$  (as required by  $\sum_{A \in P(Z)} m(A) = 1$ ) is obtained by dividing by  $1 - K$ . As  $K$  is determined by the sum of products of the bpa's of all sets where the intersection is null <sup>(5)</sup>, it represents the basic probability mass associated with conflict; it is a measure of the amount of conflict between the two mass sets and the normalization factor  $1 - K$  has the effect of completely ignoring conflict by attributing any mass associated with conflict to the null set <sup>(21)</sup>. Consequently, counterintuitive results are obtained in the face of significant conflict among experts, which has raised serious criticism to the formula <sup>(22)</sup>. Other assignment rules have been proposed to more properly represent the conflict <sup>(22,24)</sup>; this, however, is out of the scope of the present work.

If  $Q^i > 2$  experts are available, then their opinion can be aggregated by applying  $Q^i - 1$  times the Dempster Rule (eq.(4)). Notice that, since the Dempster-rule satisfies the *associativity property*  $(A \times (B \times C) = (A \times B) \times C)$ , the order in which it is applied does not influence the final result <sup>(23)</sup>.

The problem of the propagation of the uncertainty in the assumptions onto the uncertainty in the model structure is here addressed by considering the general problem of aggregating uncertainties on independent variables. In this respect, let us introduce  $k$  independent uncertain variables  $z_i$  defined on their UODs  $Z^i, i = 1, \dots, k$  and let us suppose that there are  $k$  experts, the  $i$ -th of them providing the bpa  $m_i(A^i)$  for each  $A^i \subset P(Z^i)$ . According to the DST, the joint belief and plausibility of an element  $A = (A^1, A^2, \dots, A^k)$  of the power set of  $Z = Z^1 \times Z^2 \times \dots \times Z^k$  are defined by:

$$Bel(A) = Bel(A^1, A^2, \dots, A^k) = \prod_{i=1}^k Bel_i(A_i) \quad (10)$$

$$Pl(A) = Pl(A^1, A^2, \dots, A^k) = \prod_{i=1}^k Pl_i(A_i) \quad (11)$$

From the obtained belief and plausibility functions it is possible to obtain the basic probability assignment  $m(A) = m(A^1, A^2, \dots, A^k)$  by (25):

$$m(A) = \sum_{B \subseteq A} (-1)^{card(A-B)} Bel(B) \quad (12)$$

An alternative way that permits to directly compute the bpa of an element  $A = (A^1, A^2, \dots, A^k)$  of  $P(Z)$  is based on the application of the *Dempster rule* to the cylindrical extensions of the power sets elements <sup>(26)</sup>. The cylindrical extension  $A^{i\uparrow}$  of an element  $A^i$  of  $Z^i$  over  $Z$  is defined by considering for the variables  $x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_k$  their entire UODs:

$$A^{i\uparrow} = (Z^1, \dots, Z^{i-1}, A^i, Z^{i+1}, \dots, Z^k) \quad (13)$$

Given that the knowledge of the  $i$ -th expert is assumed to be only on the variable  $z_i$ , he/she is not able to assign specific masses to any subsets  $A_j$  of  $Z^i$ ,  $j \neq i$ . For this reason, it is intuitive for the  $i$ -th expert to associate to the cylindrical extension  $A^{i\uparrow}$  defined on  $Z$  the same mass given to  $A^i$  on the UOD  $Z^i$ :

$$m_i(A^{i\uparrow}) = m_i(A^i) \quad (14)$$

whereas all the other elements of  $P(Z)$  that are not cylindrical extensions of an element of  $Z^i$  have null mass:

$$m_i(A) = 0 \text{ if } \nexists A^i \in P(Z^i) \text{ such that } A^{i\uparrow} = A \quad (15)$$

Finally, for the aggregation of the masses given by different experts, it is possible to apply the *Dempster rule* to the cylindrical extensions. For example, considering two discrete independent variables  $z_1, z_2$  and supposing that Expert 1 assigns a bpa  $m_1(A^1)$  to all the sets  $A^1$  of  $P(Z^1)$  and Expert 2 assigns a bpa  $m_2(A^2)$  to all the sets  $A^2$  of  $P(Z^2)$ , the mass of all the sets  $(A^1, A^2)$  of the power set of  $Z = Z^1 \times Z^2$  can be obtained by:

$$m_{12}(A) = m_{12}(A^1, A^2) = \frac{\sum_{B \cap C = (A^1, A^2)} m_1(B) \cdot m_2(C)}{1 - \sum_{B \cap C = 0} m_1(B) \cdot m_2(C)} \quad (16)$$

that, due to eqs.(14) and (15), becomes:

$$m_{12}(A^1, A^2) = \frac{\sum_{B^{1\uparrow} \cap C^{2\uparrow} = (A^1, A^2)} m_1(B^{1\uparrow}) \cdot m_2(C^{2\uparrow})}{1 - \sum_{B^{1\uparrow} \cap C^{2\uparrow} = 0} m_1(B^{1\uparrow}) \cdot m_2(C^{2\uparrow})} \quad (17)$$

Furthermore, since in the case here discussed the set  $(A^1, A^2)$  can be obtained only by the intersection of  $A^{1\uparrow} \cap A^{2\uparrow}$ , and the intersection  $B^{1\uparrow} \cap C^{2\uparrow}$  is never empty, eq.(17) becomes:

$$m_{12}(A^1, A^2) = m_1(A^{1\uparrow}) \cdot m_2(A^{2\uparrow}) = m_1(A^1) \cdot m_2(A^2) \quad (18)$$

leading to the same results of the direct product of the belief and plausibility functions (eqs.(10) and (11)).

Notice, however, that eq.(17) can be applied also to the case in which the two experts assign a bpa on overlapping UODs whereas both its simplified version of eq.(18) and the method of eqs.(10) and (11) are not applicable to this case. For example, if Expert 1 assigns a bpa to the elements  $(A^1, A^2)$  of the power set of  $Z^1 \times Z^2$  and Expert 2 to the elements  $(A^2, A^3)$  of the power set of  $Z^2 \times Z^3$ , the application of eq.(12) to the cylindrical extensions  $(A^1, A^2)^\uparrow$  and  $(A^2, A^3)^\uparrow$  over  $Z^1 \times Z^2 \times Z^3$  permits to aggregate the uncertainty of the experts by assigning a mass to the elements  $(A^1, A^2, A^3)$ .

In the case of interest of this work, the  $i$ -th uncertain assumptions  $H^i$  has an UOD  $X^i$  defined by the set of the alternative hypotheses,  $\{H_1^i, \dots, H_{n_i}^i\}$ . The generic model  $S_l = (H_{l_1}^1, H_{l_2}^2, \dots, H_{l_k}^k)$  can then be seen as an element of the Cartesian product of the UODs  $X^i$ , i.e.  $X = X^1 \times X^2 \times \dots \times X^k = \{M_1, M_2, \dots, M_n\}$ . If we suppose that there are  $k$  experts, the  $i$ -th of them providing the bpa  $m_i(A^i)$  for each  $A^i \subset P(X^i)$ , it is then possible to compute the belief and plausibility of the model  $M_l$  with structure  $S_l = (H_{l_1}^1, H_{l_2}^2, \dots, H_{l_k}^k)$  by applying eqs.(10) and (11):

$$\begin{aligned} Bel(M_l) &= \prod_{i=1}^k Bel_i(H_{l_i}^i) \\ Pl(M_l) &= \prod_{i=1}^k Pl_i(H_{l_i}^i) \end{aligned} \quad (19)$$

or the mass by eq.(17):

$$m(M_l) = \prod_{i=1}^k m_i (H_{l_i}^i) \quad (20)$$

Notice that by using eqs.(10), (11) or (18) also the belief, plausibility and masses of the elements of the power set of  $X$  formed by more than one model can be directly obtained.

## 5. CASE STUDY

Various conceptual models have been proposed for describing the groundwater flow and solute transport at a radioactive waste repository site in unsaturated, fractured tuff formations. The analysis of the differences in the models has been mainly concerned with the fundamental assumptions underpinning their structures and their implications on the overall assessment by the models <sup>(27,28)</sup>. In the analysis, all of the models have been considered based on the same system geometry. A spatially uniformly distributed flux at the repository horizon equal to 0.1 mm/yr is assumed as boundary condition for all models. Steady-state groundwater flow through a one-dimensional system, occurring from the base of the repository to the water table, is also assumed. Only four radionuclide species, Tc-99, I-129, Cs-135 and Np-237, are assumed to be released from the repository and transported as solute. The transport model is based on the assumption of a single, dominant, non-branching transport path. Radionuclide retardation factors were based on the distribution coefficient  $K_d$  for each radionuclide. No other chemical reactions were included. Gaseous phase transport was not included either. The quantity considered as output of the model in this work is the hydraulic head distribution. In the analysis only the uncertainty due to the model structure is considered while the uncertainty on the model parameters is completely ignored.

With respect to the assessment of groundwater flow, the hypotheses made on two main, general assumptions drive the modeling:  $H^1$  = “hydraulic and geological properties” and  $H^2$  = “groundwater flow mechanisms”. Tables I and II report the alternative hypotheses considered for the two assumptions on groundwater flow <sup>(1)</sup>.

**Table I. Alternatives for assumption  $H^1$  “hydraulic and geological properties”**

**Table II. Alternatives for assumption  $H^2$  “groundwater flow mechanism”**

As a first case, consider that two experts have been asked to quantify the probabilities of correctness of the alternative hypotheses of the two assumptions  $H^1$  and  $H^2$ . In particular, Expert 1 gives a judgment only on the correctness of the alternative hypotheses of the first assumption  $H^1$  whereas Expert 2 on the probability of correctness of the alternative hypothesis for the second assumption  $H^2$ . Since the opinions elicited from the experts regard two different assumptions of the model, there is no overlapping between their knowledge and thus one needs only to tackle the problem of propagating the uncertainty in the assumptions onto the model. Let us also assume that the expert does not possess sufficiently refined knowledge of the system to express his/her belief on the correctness of the alternative hypotheses by discrete probability values, but that he/she feels more confident in giving an interval of possible values for the probabilities of correctness. For example, in evaluating the two different alternative hypotheses for the assumption  $H^1$ , Expert 1 judges that the probability  $P_1(H_1^1)$  of the correctness of hypothesis  $H_1^1$  is a value between 0.6 and 0.95. Expert 2 opinion on the three alternatives for assumption  $H^2$  is that  $P_2(H_1^2)$  is between 0.6 and 0.8. whereas  $P_2(H_2^2)$  is between 0.03 and 0.10.

### 5.1. Application of the probabilistic framework

For the assumption  $H^1$ , the categorical random variable is binomial with possible realizations  $H_1^1$  and  $H_2^1$  whose probabilities are:

$$\begin{aligned} P_1(H_1^1) &= p_{11} \\ P_1(H_2^1) &= p_{12} \end{aligned} \tag{21}$$

The expert assignment to the only parameter  $p_{11} = P_1(H_1^1)$  of the binomial distribution is affected by epistemic uncertainty; in particular, the expert suggests only the two limiting values of 0.6 and 0.95 for  $p_{11}$ , with no additional preference for values within the range, so that it seems reasonable to represent the uncertainty by a uniform probability density function  $f(p_{11})$  on the range defined by the limiting values (Figure 1).

**Figure 1. Probability density function representing the epistemic uncertainty on  $p_{11}$ .**

With respect to the assumption  $H^2$ , the probability, assigned by Expert 2, of the possible alternative hypotheses are:

$$\begin{aligned} P_2(H_1^2) &= p_{21} \sim U[0.60, 0.80] \\ P_2(H_2^2) &= p_{22} \sim U[0.03, 0.10] \\ P_2(H_3^2) &= 1 - p_{21} - p_{22} \end{aligned} \tag{22}$$

Notice that, since  $p_{21} + p_{22} \leq 1$ , the two parameters  $p_{21}$  and  $p_{22}$  of the multinomial distribution are not independent, i.e. a given value of  $p_{21}$  constrains  $p_{22}$  to be lower than  $1 - p_{21}$ . However, in the case here considered  $p_{21} + p_{22}$  is always less than 1 and thus the

two parameters are independent, i.e.  $f_{p_{21}, p_{22}}(p_{21}, p_{22}) = f_{p_1}(p_1)f_{p_2}(p_2)$ . In particular, uniform distributions are again assumed for  $p_{21}, p_{22}$  (Figure 2).

**Figure 2. Probability density functions representing the epistemic uncertainty on  $p_{21}$  and  $p_{22}$ .**

The uncertainty propagation from the alternative hypotheses onto the groundwater flow model structure leads to the introduction of a multinomial variable  $M$  whose possible realizations are:

$$M_1 = (H_1^1, H_1^2), M_2 = (H_2^1, H_1^2), M_3 = (H_1^1, H_2^2), M_4 = (H_2^1, H_2^2), M_5 = (H_1^1, H_3^2), M_6 = (H_2^1, H_3^2)$$

Assuming the independence of the two assumptions  $H^1$  and  $H^2$ , it is possible to compute the probability of correctness of the 6 models which emerge from the combination of the 2·3 alternative hypotheses, by simply multiplying the probability of correctness of the constituting alternative hypotheses:

$$\begin{aligned} P_{12}(M_1) &= P_1(H_1^1)P_2(H_1^2) = p_{11}p_{21} = p_1 \\ P_{12}(M_2) &= P_1(H_2^1)P_2(H_1^2) = (1 - p_{11})p_{21} = p_2 \\ P_{12}(M_3) &= P_1(H_1^1)P_2(H_2^2) = p_{11}p_{22} = p_3 \\ P_{12}(M_4) &= P_1(H_2^1)P_2(H_2^2) = (1 - p_{11})p_{22} = p_4 \\ P_{12}(M_5) &= P_1(H_1^1)P_2(H_3^2) = p_{11}(1 - p_{21} - p_{22}) = p_5 \\ P_{12}(M_6) &= P_1(H_2^1)P_2(H_3^2) = (1 - p_{11})(1 - p_{21} - p_{22}) = p_6 \end{aligned} \tag{23}$$

Note that the parameters  $p_1, p_2, p_3, p_4, p_5$  defining the multinomial distribution of  $M$ , are dependent given that they must satisfy the constraint  $p_1 + p_2 + p_3 + p_4 + p_5 \leq 1$  and thus they are characterized by the joint probability density function  $f(p_1, p_2, p_3, p_4, p_5)$ . Figure

3 reports the marginal probability density functions  $f_{p_l}(p_l)$  of the parameter  $p_l, l = 1, \dots, 5$  obtained numerically.

**Figure 3. Marginal probability density functions of the parameters  $p_1, p_2, p_3, p_4, p_5$  defining the multinomial distribution of  $M$ .**

From the obtained joint probability density function  $f(p_1, p_2, p_3, p_4, p_5)$ , it is then possible to estimate the probability of correctness of the models by:

$$P(M_i) = \int P(M_i | p_i) f(p_1, p_2, p_3, p_4, p_5) dp_1 dp_2 dp_3 dp_4 dp_5 = \int p_i f_{p_i}(p_i) dp_i \quad i = 1, \dots$$

$$P(M_6) = 1 - \sum_{i=1}^5 P(M_i) \quad (24)$$

which gives:

$$\begin{aligned} P(M_1) &= 0.54 \\ P(M_2) &= 0.16 \\ P(M_3) &= 0.05 \\ P(M_4) &= 0.02 \\ P(M_5) &= 0.18 \\ P(M_6) &= 0.05 \end{aligned} \quad (25)$$

## 5.2. Application of the DST framework

The UOD  $X^1$  of the assumption  $H^1$  is formed by two elements,  $H_1^1, H_2^1$ ; therefore, the definition of the corresponding bpa requires assigning a probability mass to all the three elements of the power set of  $X^1$ , i.e.  $P(X^1) = \{H_1^1, H_2^1, (H_1^1, H_2^1)\}$ . The opinion provided by the expert that the probability of the correctness of alternative  $H_1^1$  is a value between 0.6 and 0.95 leads to consider:

$$\begin{aligned}
Bel_1(H_1^1) &= 0.60 \\
Pl_1(H_1^1) &= 0.95
\end{aligned} \tag{26}$$

From this assignment, using the relation:

$$Bel(A) + Pl(\bar{A}) = 1 \tag{27}$$

the belief and plausibility of  $H_2^1$  are obtained:

$$\begin{aligned}
Bel_1(H_2^1) &= 0.05 \\
Pl_1(H_2^1) &= 0.95
\end{aligned} \tag{28}$$

From eq.(12) the following masses of the elements of the power set  $P(X^1)$  are then obtained:

$$\begin{aligned}
m_1(M_1^H) &= 0.60 \\
m_1(M_2^H) &= 0.05 \\
m_1(M_1^H, M_2^H) &= 0.35
\end{aligned} \tag{29}$$

The second assumption  $H^2$  is characterized by a UOD  $X^2$  formed by the three alternatives  $H_1^2, H_2^2, H_3^2$ , leading to a power set of 8 elements. Since the development of a full-fledged technique for the elicitation of the bpa from the expert knowledge is outside the scope of this work, here we limit ourselves to assigning directly the limiting values to the belief and plausibility of  $H_1^2$  and  $H_2^2$ :

$$\begin{aligned} Bel_2(H_1^1) &= 0.60, \quad Pl_2(H_1^1) = 0.80 \\ Bel_2(H_2^2) &= 0.03, \quad Pl_2(H_2^2) = 0.10 \end{aligned} \tag{30}$$

whereas in order to respect the constraints of a bpa, it is assumed that:

$$Bel_2(H_3^2) = 0.15, \quad Pl_2(H_3^2) = 0.35 \tag{31}$$

These values of belief and plausibility of the elements of the frame of discernment lead to the following bpa:

$$\begin{aligned} m_2(H_1^2) &= 0.60 \\ m_2(H_2^2) &= 0.03 \\ m_2(H_3^2) &= 0.15 \\ m_2(H_1^2, H_2^2) &= 0.02 \\ m_2(H_1^2, H_3^2) &= 0.15 \\ m_2(H_2^2, H_3^2) &= 0.02 \\ m_2(H_1^2, H_2^2, H_3^2) &= 0.03 \end{aligned} \tag{32}$$

The probability masses elicited from Experts 1 and 2 can be finally combined according to eq.(18). Table III reports bpa, belief and plausibility of the subsets of the frame of discernment  $\{M_1, M_2, M_3, M_4, M_5, M_6\}$  formed by a single element.

Notice that there is a considerable mass of  $1 - \sum_{i=1}^6 m(M_i) = 0.46$  assigned to the remaining subsets, i.e. a mass that cannot be attributed to a single model.

**Table III. Basic probability assignment, belief and plausibility of the six model hypotheses.**

It is also interesting to note that models  $M_1 = (H_1^1, H_1^2)$  and  $M_2 = (H_2^1, H_1^2)$  are characterized by a large difference between their belief and plausibility. This is caused by the fact that assumption  $H^1$  is characterized by a large amount of mass not allocated to a single hypothesis. For example, Expert 1 judges that the set  $(H_1^1, H_1^2)$  has a mass of 0.35; this means that there is a considerable evidence that one of these two models is correct but it is not possible to specify which one. For this reason, the strength of the belief in the models  $M_1$  and  $M_2$  is limited by the fact that it contains only the mass that can be attributed to the individual model, whereas the plausibility has an higher value because it includes all the mass that is not against such model and so it contains also the relatively large mass attributed to  $(H_1^1, H_1^2)$ . This effect of large uncertainty on the beliefs in the correctness of the models is less evident for the other models because the masses on the alternatives  $H_2^2$  and  $H_3^2$  are lower than  $m(H_1^2)$ .

Finally notice that  $Bel(M_3) > Bel(M_6)$  whereas  $Pl(M_3) < Pl(M_6)$ . Also this effect depends on the fact that there is a mass of 0.15 that Expert 2 assigns to  $(H_1^2, H_3^2)$ , which increases the plausibility of model 6 and not of model 3.

### 5.3. Prediction of the hydraulic head

In order to further compare the probabilistic and the DS frameworks, the estimation of the model output, i.e. the hydraulic head at  $d=90$  meters is here considered. The output values of models  $M_1, M_3, M_5, M_6$  have been taken from <sup>(1)</sup>, whereas, since models  $M_2$  and  $M_4$  have not actually been developed, their predictions have been arbitrarily set. The case in which some models, based on particular combinations of assumptions, are not developed will be further addressed in Section 5.4. Here, since by observing the predictions of models  $M_5 = (H_1^1, H_3^2)$  and  $M_6 = (H_2^1, H_3^2)$  one can argue that the effect of hypothesis  $H_2^1$  is to

decrease the prediction of the corresponding model with hypothesis  $H_1^1$ , the predictions of models  $M_2$  and  $M_4$ , characterized by hypothesis  $H_2^1$ , have been set, respectively, to a lower values than the predictions of models  $M_1$  and  $M_3$ , characterized by hypothesis  $H_1^1$ . Table IV reports all the model outputs considered in this Section.

**Table IV. Estimation of the hydraulic head at  $d=90$  meters**

According to the probabilistic framework, the probability  $p(z_i)$  that the hydraulic head  $z$  takes the value  $z_i$  is:

$$p(z_i) = \int p(z_i | p_1, p_2, p_3, p_4, p_5) f(p_1, p_2, p_3, p_4, p_5) dp_1 dp_2 dp_3 dp_4 dp_5 \quad (33)$$

and thus it is equal to the probability  $P(M_i)$  of the corresponding model  $M_i$  (eq.(25)). Figure 4 reports the discrete cumulative distribution representing the probability that the hydraulic head distribution at 90 metres is minor or equal to  $z$ .

**Figure 4. Discrete cumulative distribution of the hydraulic head distribution.**

Furthermore, the expected value of the hydraulic head can be computed by:

$$E[y] = \int \sum_{i=1}^6 y_i p(z_i | p_1, p_2, p_3, p_4, p_5) f(p_1, p_2, p_3, p_4, p_5) dp_1 dp_2 dp_3 dp_4 dp_5 = 817 \quad (34)$$

Within the DST framework, the basic probability assignment on the frame of discernment of the model output values  $\{z_1, z_2, z_3, z_4, z_5, z_6\}$  corresponds to the bpa defined on the frame of discernment of  $\{M_1, M_2, M_3, M_4, M_5, M_6\}$  (Table III). In particular, it is interesting to

compute, using eqs.(4) and (5), the belief and plausibility of the sets  $\{z \leq z_i\}, i = 1, \dots, 6$  which can be interpreted as limiting cumulative density distributions (Figure 5).

Notice that the probabilities  $P(z \leq z_i)$  obtained by applying the **probabilistic** framework are always within the interval of the corresponding belief and plausibility values. Thus, although the two approaches differ in the representation and propagation of the expert uncertainty on the probability of the model to be correct, in this case they lead to coherent results.

Finally, notice that synthetic measures of the expected value of the hydraulic head obtained from the belief and possibility distributions have been proposed in <sup>(29)</sup>, but they are not investigated in this work since decision makers usually demand to know the entire distribution of the uncertainties.

**Figure 5. Belief (circles), plausibility (squares) of the sets  $[z_1, \dots, z_i]$  obtained by the DST approach and cumulative density probability values (dots) obtained by the probabilistic approach.**

#### **5.4. Model uncertainty in presence of non developed models**

In this Section, the problem of aggregating the uncertainties of different experts on the correctness of the models is addressed on a case in which some possible models are not developed, although they have non zero probability of being correct for some experts. This case may arise when it is decided that some theoretically plausible models should not be developed due, for example, to the opinion of an expert or other practical limitations.

In particular, the previous case study is re-considered with the addition of the opinion of a third expert that judges models  $M_2$  and  $M_4$  surely not plausible. Furthermore, it is assumed that Expert 3 does not give any information to specify which of the remaining models  $M_1, M_3, M_5, M_6$  is the correct one.

#### 5.4.1. Probabilistic framework

According to the probabilistic framework, Expert 3 opinion is represented by introducing a multinomial variable with possible realization  $M_1, M_2, M_3, M_4, M_5, M_6$  whose probabilities are:

$$\begin{aligned}
 P_3(M_1) &= p_{31} \sim \mathcal{U}[0,1] \\
 P_3(M_2) &= p_{32} = 0 \\
 P_3(M_3) &= p_{33} \sim \mathcal{U}[0,1] \\
 P_3(M_4) &= p_{34} = 0 \\
 P_3(M_5) &= p_{35} \sim \mathcal{U}[0,1] \\
 P_3(M_6) &= p_{36} = 1 - p_{31} - p_{32} - p_{33} - p_{34} - p_{35}
 \end{aligned} \tag{35}$$

where  $p_{32}$  and  $p_{34}$  are equal to 0, whereas since Expert 3 does not give any suggestions on his/her beliefs on the remaining models, it is reasonable to characterize by uniform probability density functions the uncertainty on the parameters  $p_{31}, p_{33}$  and  $p_{35}$  of the multinomial distribution. Notice, however, that, due to the constraint  $p_{31} + p_{33} + p_{35} \leq 1$ , it is not possible to assume that the parameters of the multinomial distribution are independent and thus one should use a joint probability density function defined by:

$$\begin{cases} f(p_{31}, p_{33}, p_{35}) = 6 & \text{if } p_{31} + p_{33} + p_{35} \leq 1 \\ f(p_{31}, p_{33}, p_{35}) = 0 & \text{otherwise} \end{cases} \tag{36}$$

where the value of 6 is the normalization constant. From this a-priori distribution of the parameters of our multinomial distribution, the following estimation of the probabilities of the six models are obtained:

$$\begin{aligned}
P(M_1) &= 0.25 \\
P(M_2) &= 0.00 \\
P(M_3) &= 0.25 \\
P(M_4) &= 0.00 \\
P(M_5) &= 0.25 \\
P(M_6) &= 0.25
\end{aligned} \tag{37}$$

According to this interpretation, the total lack of information on which model between  $M_1, M_3, M_5, M_6$  is correct translates in the assignment of equal probabilities to the four models.

For the aggregation of the probabilities on the correctness of the 6 models given by Expert 3 with the probabilities obtained by combining the opinions of Expert 1 and Expert 2 reported in eq.(25), the logarithmic opinion pool method is used, assigning in eq.(6) equal weights to  $P_{12}(M_i)$  and  $P_3(M_i)$ . Thus,

$$P_{123}(M_i) = k \sqrt{P_{12}(M_i) P_3(M_i)} \tag{38}$$

leading to the results reported in the second column of Table V.

Notice that if the aggregation of  $P_3(M_i)$  with  $P_{12}(M_i)$  were performed by using the linear opinion pool, non zero probabilities would be obtained for  $P_{123}(M_2)$  and  $P_{123}(M_4)$ , leading to the paradox that a non-zero probability is assigned to a model that is actually not developed.

**Table V. Probabilities of correctness of the six models obtained by combining the three experts opinions (second column), only Expert 1 and 2 opinions (third column) and relative variation obtained by considering Expert 3 opinion (last column).**

Note that the effect of the aggregation of Expert 3 opinion with Experts 1 and 2 opinions is to set to zero the probabilities  $P_{123}$  of models  $M_2$  and  $M_4$ . Furthermore, due to the constraint that the sum of the probabilities of all the models must be one, the probability  $P_{12}(M_2) + P_{12}(M_4) = 0.17$  must be attributed in  $P_{123}$  to the remaining four models through a process of renormalization of the probabilities. Given that Expert 3 does not specify any preference between the remaining models  $M_1, M_3, M_5, M_6$ , one would expect that the probabilities  $P_{123}$  would all increase, with respect to the probabilities  $P_{12}$ , to consider the mass  $P_{12}(M_2) + P_{12}(M_4) = 0.17$ . However, using the logarithmic opinion pool method, it is obtained that  $P_{123}(M_1)$  decreases of 16.0%, whereas the other models probabilities increase, but not uniformly as it can be seen by observing the relative variations reported in the last column of Table V. The cause of this counterintuitive effect is that the probabilistic framework forces to assume that each of them has a probability of 0.25 and thus the aggregation of  $P_{12}(M_1) = 0.54$  with  $P_3(M_3) = 0.25$  through the geometric mean causes a decrease of  $P(M_1)$  that is not counterbalanced by the renormalization of the probabilities.

#### 5.4.2. *DS framework*

According to the DST framework, it is possible to represent Expert 3 judgment on the 6 models by the bpa:

$$\begin{aligned}
m_3(M_1, M_3, M_5, M_6) &= 1 \\
m_3(A) &= 0 \quad \forall A \neq (M_1, M_3, M_5, M_6) \in P(\{M_1, M_2, M_3, M_4, M_5, M_6\})
\end{aligned} \tag{39}$$

Notice that this representation of Expert 3 uncertainty on the models correctness is completely coherent with the information that he/she gives. In particular, through eqs.(4) and (5), it is possible to show that it is equivalent to assign belief zero and plausibility 1 to all the subsets containing one of the four models  $(M_1, M_3, M_5, M_6)$ .

By combining  $m_3$  with  $m_{12}$  using eq.(8), a new bpa  $m_{123}$  characterized by mass equal to 0 for all the subsets that contain  $M_2$  and/or  $M_4$  is obtained. Table VI represents bpa, belief and plausibility of the subsets formed by a single model while Table VII compares the bpas before and after the aggregation of Expert 3 opinion.

Notice that the DST combination of Expert 3 opinion with Expert 1 and 2 opinions results in a bpa with mass 0 for all the subsets that contain  $M_2$  or  $M_4$ . Thus, analogously to the probabilistic case of the previous Section, the four models  $M_1, M_3, M_5, M_6$  form an exhaustive set of models and the previous observation made on the limitation of the exhaustiveness hypothesis still arises. In this respect, it is interesting to point out the possibility, within the DST framework, to assume that the frame of discernment of an uncertain quantity is formed by all its known possible realizations and by an unknown element (30), (31). This approach, that seems very promising to address the hypothesis of exhaustiveness of the models, will be further investigated in future works. It is also interesting to consider that within the DST approach, contrary to the probabilistic approach, Expert 3 opinion causes an increase of 63% of the masses assigned to  $M_1$  by Experts 1 and 2. This is due to the fact that a considerable part of the mass  $m_{12}(M_1, M_2) = 0.35$ , that Experts 1 and 2

are not able to specify to which of the two models belongs, is assigned by eq.(8) to  $M_1$ , since  $m_3(M_2) = 0$ . Analogously, the mass assigned to the other single model subsets increases due to the fact that they collect all the mass assigned by Experts 1 and 2 to subsets containing  $M_2$  and/or  $M_4$ .

**Table VI. Mass, belief and plausibility of the subsets formed by a single model**

**Table VII. Comparison of the masses assigned to the single model subsets obtained by combining the three experts opinions (second column), only Expert 1 and 2 opinions (third column) and relative variation obtained by considering Expert 3 opinion (last column)**

Finally, Figure 6 reports the belief and plausibility values obtained within the DST approach for the hydraulic head  $z$  and the corresponding cumulative density values obtained by the probabilistic approach. Due to the different aggregation mechanism of Expert 3 opinion, the cumulative distribution is far below the DS limiting distribution. For example, the probability that  $z$  be lower than or equal to 810 is 0.46 according to the probabilistic framework and between 0.59 and 0.79 for the DST framework.

## 6. CONCLUSIONS

In this paper, the problem of representing and propagating model uncertainty within the model set expansion approach has been considered by the Dempster-Shafer Theory of Evidence compared to a classical probabilistic framework, with reference to a case study of a model of groundwater flow of interest for the performance assessment of radioactive waste

repositories. The methodological set up has considered the case in which multiple assumptions are made for the development of the model, giving rise to corresponding alternative plausible modelling hypotheses on whose correctness the opinions of different experts is sought.

Through the illustration of an example concerning the assumptions underpinning a model of groundwater flow, the DST has been shown a proper framework for effectively representing the uncertainty on the correctness of the different hypotheses by means of two limiting values, belief and plausibility, overtaking the difficulties that experts encounter in characterizing their uncertainty using single probability values or probability density functions. The added value of the approach is particularly evident in the case in which some possible models are not developed, for any reason, a situation which may lead to counterintuitive and non conservative results if addressed probabilistically.

Furthermore, within the DST, the Dempster rule of combination offers a natural basis for both the aggregation of the uncertainties of different experts on the correctness of the hypotheses and the propagation of the uncertainties regarding the individual assumptions onto the uncertainty characterizing the entire model. This is regardless of the limitations of the rule, which ignores conflicting expert opinions, which can be overcome by more elaborate aggregation rules not introduced here as the issue is out of the scope of the present work.

With respect to the future development of the work, notice that the model set expansion approach is based on the two assumptions of mutual exclusiveness and collective exhaustiveness of the set of models: while the first assumption may be acceptable, upon verification, the second is certainly not met in practice because it requires that the perfect model is one of the models taken into account, which is not realistic. This aspect deserves further study in the future; in this respect, it seems worth considering modifications to the

DST which assume that the frame of discernment of an uncertain quantity is formed by all its known possible realizations and by an unknown element which covers lack of completeness.

Another direction of development of the work deemed necessary regards the case in which new information on the model assumption correctness, including experimental data and subjective evidence, becomes available after the initial analysis and, thus, updating of the output distribution is required; in this respect, formal Bayesian approaches to model uncertainty seem very fit to the purpose due to their ability to update the distribution of the model output and should be compared to their counterparts within the DST framework of analysis.

Finally, the characterization of the model output uncertainty here analyzed, is expected to be used within the Probabilistic Risk Assessment (PRA) framework for the performance assessment of radioactive waste repositories: whereas the output probability distribution provided by a probabilistic approach to model uncertainty can be directly used in the PRA framework, the DST produces lower and upper cumulative density functions which can be used to bound the value of the probability that the output is below a given value, an informative metric that can be used in the PRA framework.

## **ACKNOWLEDGMENT**

This work has been funded by the Ente per le Nuove Tecnologie, L'Energia e l'Ambiente (ENEA) within the Framework Program Agreement with the Italian Ministry of Economic Development.

## REFERENCES

1. Zio E, Apostolakis GE. Two methods for the structured assessment of model uncertainty by experts in performance assessments of radioactive waste repositories. *Reliability Engineering and System Safety*, 1996; 54:225-241.
2. Apostolakis GE. A Commentary on Model Uncertainty. Washington, DC: NUREG/CP-0138, US Nuclear Regulatory Commission, Oct. 20-22, 1993.
3. Buslik A. A Bayesian Approach to Model Uncertainty. Washington, DC: NUREG/CP-0138, US Nuclear Regulatory Commission, Oct. 20-22, 1993.
4. Reinert JM, Apostolakis GE. Including model uncertainty in risk-informed decision making. *Annals of Nuclear Energy*, 2006; 33:354-369.
5. Winkler RL, Model Uncertainty: Probabilities for Models? Washington, DC: NUREG/CP-0138, US Nuclear Regulatory Commission, Oct. 20-22, 1993.
6. Keeney RL, von Winterfeldt D. Eliciting Probabilities from Experts in Complex Technical Problems. *IEEE Transactions on Engineering Management*, 1991; 38(3):191-201.
7. US Nuclear Regulatory Commission and the Commission of European Communities, Probabilistic Accident Consequence Uncertainty Analysis. Washington, DC: NUREG/CR-6555 and NUREG/CR-6523, 1997.
8. Budnitz RJ, Apostolakis GE, Boore DM, Cluft LS, , Coppersmith KJ, Cornell, Morris PA. Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on Uncertainty and Use of Experts. Washington, DC: NUREG/CR-6372, U.S. Nuclear Regulatory Commission, 1996.

9. Pulcini G. Accounting for model uncertainty in reliability analysis of mechanical repairable equipment. *Communications in Statistics-Theory and Methods*, 2002; 31(5).
10. Groen F, Mosleh A. Foundations of probabilistic inference with uncertain evidence. *Int. J. Approx. Reasoning*, 2005; 39(1):49-83.
11. Droguett EL, Mosleh A. Bayesian methodology for model uncertainty using model performance data. *Risk Analysis*, 2008; 28(5).
12. Morgan MG, Henrion M. *Uncertainty, A Guide to Dealing With Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge University Press, 1990.
13. Mosleh A. Model uncertainty and its implications in probabilistic risk assessment, 8th International Conference on Structural Mechanics in Reactor Technology. Brussels, Belgium, 1985.
14. Apostolakis GE. The Concept of Probability in Safety Assessments of Technological Systems. *Science*, 1990; 250:1359-1364.
15. De Rocquigny E, Devictor N, Tarantola S. *Uncertainty in Industrial Practice, a guide to quantitative uncertainty management*. John Wiley & Sons, 2008.
16. Shafer G. *A Mathematical Theory of Evidence*. Princeton, NJ: Princeton University Press, 1976.
17. Stone M. The opinion pool. *Annals of Mathematical Statistics*, 1961; 32:1339-1342.
18. Clemen RT, Winkler RL. Combining Probability Distributions From Experts in Risk Analysis. *Risk Analysis*, 1999; 19(2).
19. S. French. Group consensus probability distributions: A critical survey. In Bernardo JM, De Groot MH, Lindley DH, Smith AFM (Eds.), *Bayesian Statistics 2*. 1985.
20. Genest C, McConway KJ. Allocating the weights in the linear opinion pool. *Journal of Forecasting*, 1990; 9:53-73.

21. Yager R. On the Dempster-Shafer Framework and New Combination Rules. *Information Sciences*, 1987; 41:93-137.
22. Yager R. Quasi-Associative Operations in the Combination of Evidence. *Kybernetes*, 1987; 16:37-41.
23. Dempster AP. Upper and Lower Probabilities Induced by a Multi-valued Mapping. *Annals of Mathematical Statistics*, 1967; 38:325–339.
24. Ferson S, Kreinovich V. Representation, Propagation and Aggregation of Uncertainty. SANDIA Report, 2002.
25. Klir GJ. Principles of uncertainty: What are they? Why do we need them?. *Fuzzy sets and systems*, 1995; 74(1).
26. Haenni R, Lehmann N. Implementing belief function computations. *Journal of Intelligent Systems*, 2003; 18(1):31-49.
27. Gallegos DP, Pohl PI, Updegraff CD. Preliminary Assessment of the Impact of Conceptual Model Uncertainty on Site Performance, High Level Radioactive Waste Management. Proceedings of the Second International Conference, Las Vegas, Nevada, April 29-May 3, 1991.
28. Parsons M, Olague NE, Gallegos DP. Conceptualization of a Hypothetical High-Level Nuclear Waste Repository Site in Unsaturated Fractured Tuff. Washington, DC: NUREG/CR-5495, US Nuclear Regulatory Commission, 1991.
29. Smets P. Constructing the pignistic probability function in a context of uncertainty. In: Henrion M, Shachter RD, Kanal LN, Lemmer JF (Eds.). *Uncertainty in Artificial Intelligence*. Amsterdam, Holland: North Holland, 1990.
30. Smets P. The nature of the unnormalized beliefs encountered in the transferable belief model. In: Dubois D, Wellman MP, D'Ambrosio B, Smets P (eds.), *Uncertainty in Artificial Intelligence*. San Mateo, CA: Morgan Kaufman, 1992.

31. Zeng C, Zhao B. Support Function and Objecting Function, a New Viewpoint to the Dempster-Shafer Theory of Evidence. ICSP'O4 Proceedings, 2004.

**Table I. Alternatives for assumption  $H^1$  “hydraulic and geological properties”**

|         |                                       |
|---------|---------------------------------------|
| $H_1^1$ | “15 hydrogeological vertical layers ” |
| $H_2^1$ | “4 hydrogeological vertical layers”   |

**Table II. Alternatives for assumption  $H^2$  “groundwater flow mechanism”**

|         |                                                          |
|---------|----------------------------------------------------------|
| $H_1^2$ | “Groundwater flow in porous matrix only”                 |
| $H_2^2$ | “Groundwater flow in fractures only”                     |
| $H_3^2$ | “Groundwater flow in matrix and fracture simultaneously” |

**Table III. Basic probability assignment, belief and plausibility of the six model hypotheses.**

| Model | bpa  | belief | plausibility |
|-------|------|--------|--------------|
| $M_1$ | 0.36 | 0.36   | 0.76         |
| $M_2$ | 0.03 | 0.03   | 0.32         |
| $M_3$ | 0.02 | 0.02   | 0.10         |
| $M_4$ | 0.00 | 0.00   | 0.04         |
| $M_5$ | 0.09 | 0.09   | 0.33         |
| $M_6$ | 0.01 | 0.01   | 0.14         |

**Table IV. Estimation of the hydraulic head at  $d=90$  meters**

| Model                  | $z = \text{Head}(90\text{m})$ |
|------------------------|-------------------------------|
| $M_1 = (H_1^1, H_1^2)$ | 810m                          |
| $M_2 = (H_2^1, H_1^2)$ | 762m                          |
| $M_3 = (H_1^1, H_2^2)$ | 900m                          |
| $M_4 = (H_2^1, H_2^2)$ | 852m                          |
| $M_5 = (H_1^1, H_3^2)$ | 860m                          |
| $M_6 = (H_2^1, H_3^2)$ | 812m                          |

**Table V. Probabilities of correctness of the six models obtained by combining the three experts opinions (second column), only Expert 1 and 2 opinions (third column) and relative variation obtained by considering Expert 3 opinion (last column).**

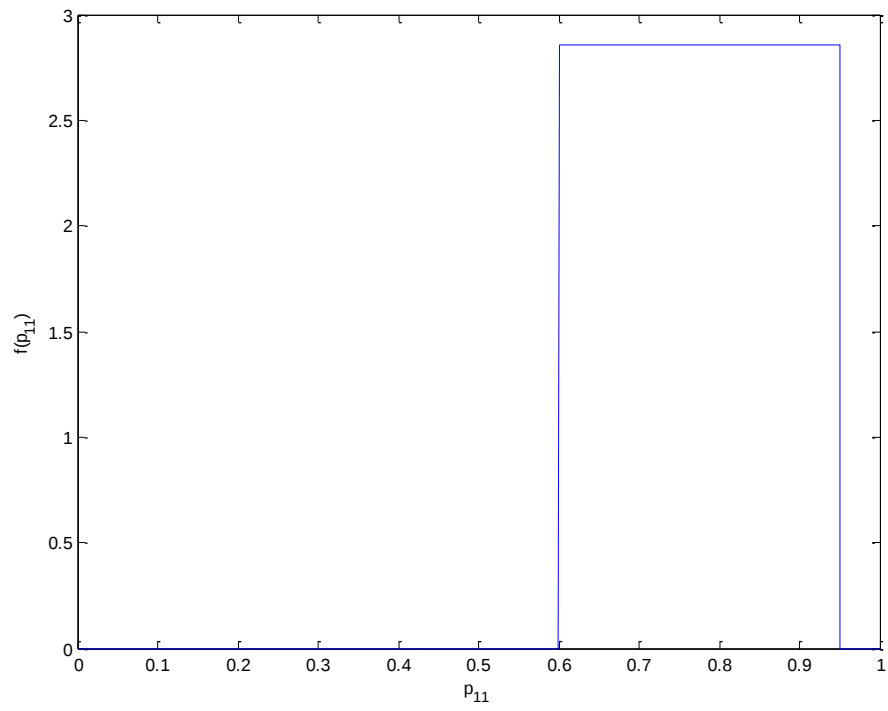
| Model | $P_{123}$ | $P_{12}$ | $(P_{123} - P_{12}) / P_{12}$ |
|-------|-----------|----------|-------------------------------|
| $M_1$ | 0.46      | 0.54     | -0.16                         |
| $M_2$ | 0.00      | 0.16     | -1.00                         |
| $M_3$ | 0.14      | 0.05     | 1.77                          |
| $M_4$ | 0.00      | 0.02     | -1.00                         |
| $M_5$ | 0.26      | 0.18     | 0.45                          |
| $M_6$ | 0.14      | 0.05     | 1.69                          |

**Table VI. Mass, belief and plausibility of the subsets formed by a single model**

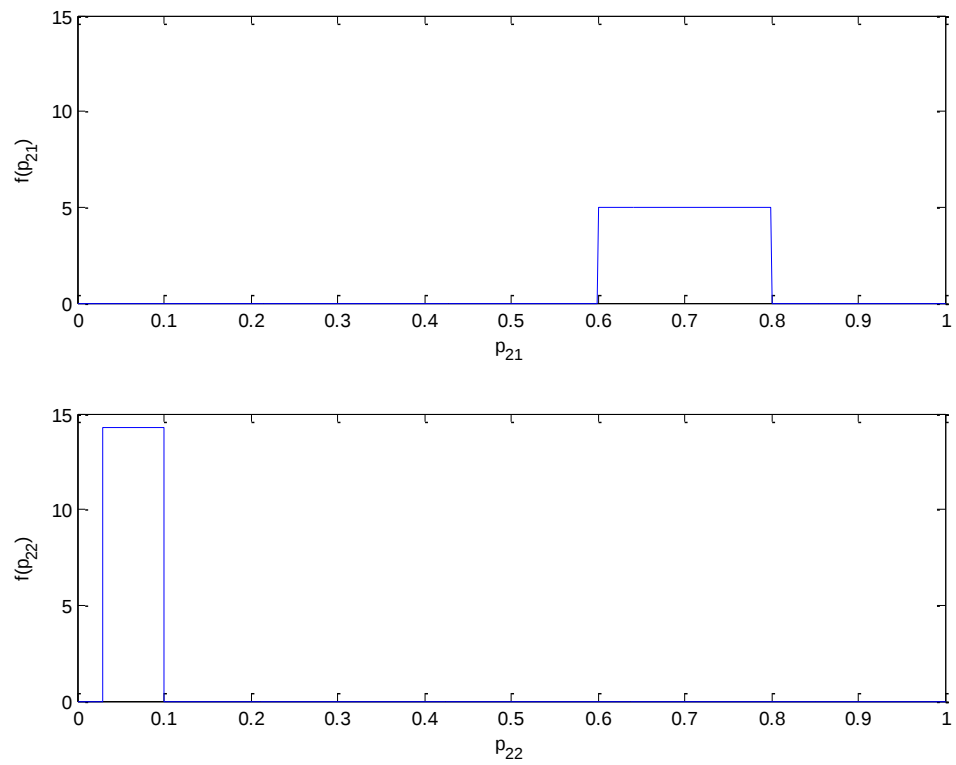
| <b>Subset</b> | <b>Mass</b> | <b>Belief</b> | <b>Plausibility</b> |
|---------------|-------------|---------------|---------------------|
| $M_1$         | 0.59        | 0.59          | 0.79                |
| $M_2$         | 0.00        | 0.00          | 0.00                |
| $M_3$         | 0.03        | 0.03          | 0.10                |
| $M_4$         | 0.00        | 0.00          | 0.00                |
| $M_5$         | 0.09        | 0.09          | 0.34                |
| $M_6$         | 0.02        | 0.02          | 0.14                |

**Table VII. Comparison of the masses assigned to the single model subsets obtained by combining the three experts opinions (second column), only Expert 1 and 2 opinions (third column) and relative variation obtained by considering Expert 3 opinion (last column)**

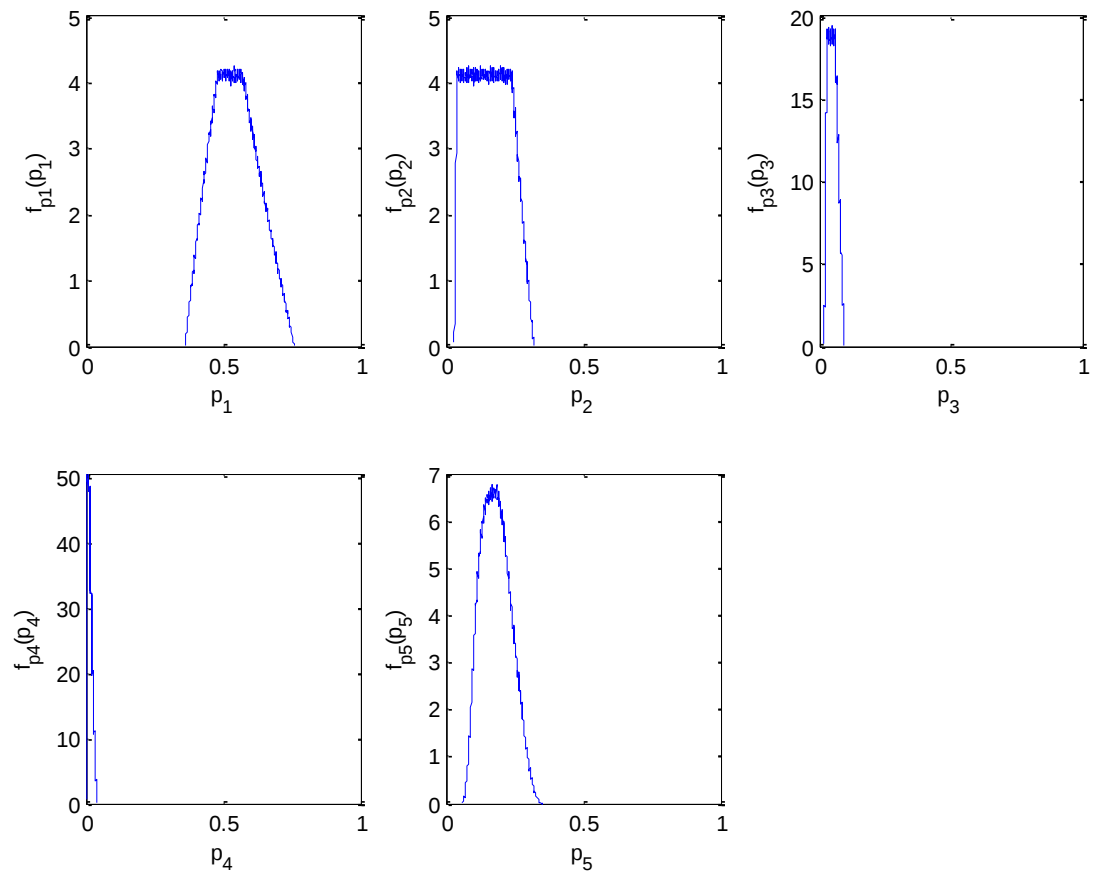
| <b>Subset</b>                      | $m_{123}$ | $m_{12}$ | $(m_{123} - m_{12}) / m_{12}$ |
|------------------------------------|-----------|----------|-------------------------------|
| $M_1$                              | 0.59      | 0.36     | 0.64                          |
| $M_2$                              | 0.00      | 0.03     | -1.00                         |
| $M_3$                              | 0.03      | 0.02     | 0.67                          |
| $M_4$                              | 0.00      | 0.00     | -1.00                         |
| $M_5$                              | 0.10      | 0.09     | 0.03                          |
| $M_6$                              | 0.018     | 0.01     | 1.57                          |
| <b>Sum of all the other masses</b> | 0.270     | 0.49     | -0.45                         |



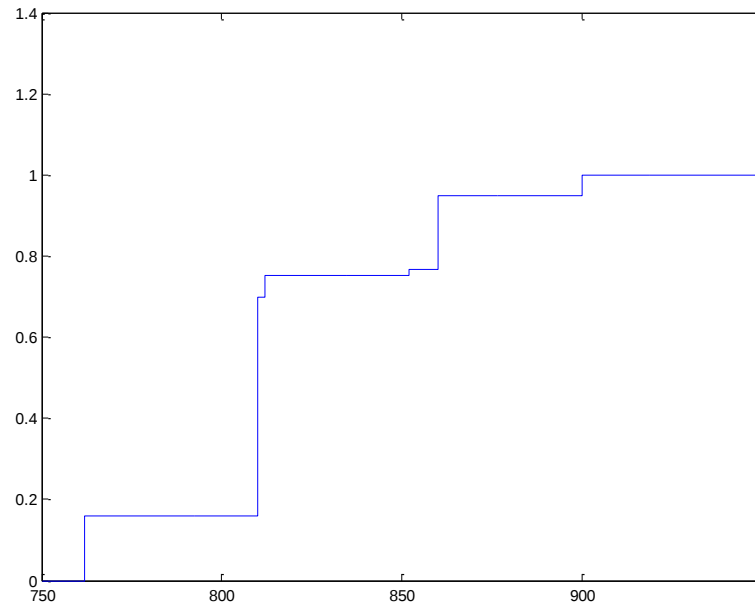
**Figure 1. Probability density function representing the epistemic uncertainty on  $p_{11}$ .**



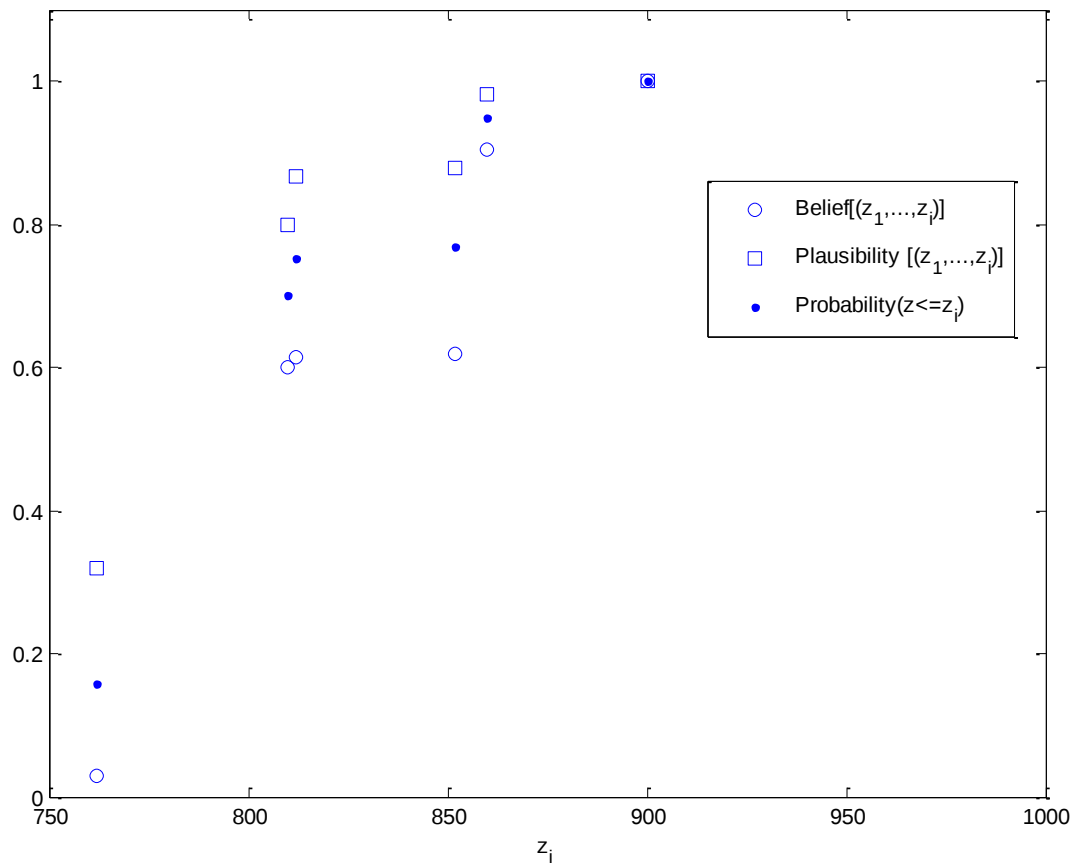
**Figure 2. Probability density functions representing the epistemic uncertainty on  $p_{21}$  and  $p_{22}$ .**



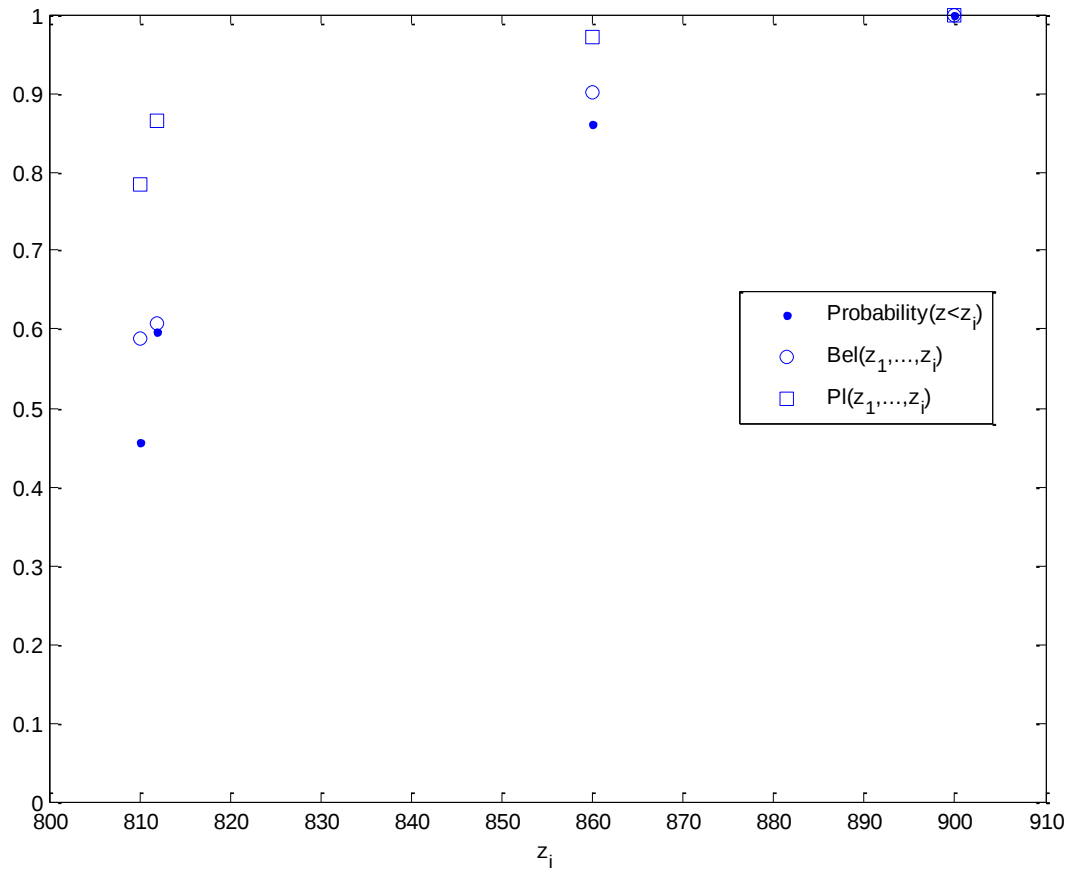
**Figure 3. Marginal probability density functions of the parameters  $p_1, p_2, p_3, p_4, p_5$  defining the multinomial distribution of  $M$ .**



**Figure4. Discrete cumulative distribution of the hydraulic head distribution.**



**Figure 5. Belief (circles), plausibility (squares) of the sets  $[z_1, \dots, z_i]$  obtained by the DST approach and cumulative density probability values (dots) obtained by the Bayesian approach.**



**Figure 6. Belief (circles), plausibility (squares) of the sets  $[z_1, \dots, z_i]$  obtained by the DST approach and cumulative density function values (dots) obtained by the Bayesian approach.**