

# Decision-Making with Belief Interval Distance

Jean Dezert<sup>1</sup>, Deqiang Han<sup>2</sup>, Jean-Marc Tacnet<sup>3</sup>,  
Simon Carladous<sup>3,4,5</sup>, and Yi Yang<sup>6</sup>

<sup>1</sup> The French Aerospace Lab, F-91761 Palaiseau, France. [jean.dezert@onera.fr](mailto:jean.dezert@onera.fr)

<sup>2</sup> CIESR, Xi'an Jiaotong Univ, Xi'an, China 710049. [deqhan@gmail.com](mailto:deqhan@gmail.com)

<sup>3</sup> UGA, Irstea, UR ETGR, 2 rue de la Papeterie-BP 76, F-38402 St-Martin-d'Hères, France. [{simon.carladous,jean-marc.tacnet}@irstea.fr](mailto:{simon.carladous,jean-marc.tacnet}@irstea.fr)

<sup>4</sup> AgroParisTech, 19 avenue du Maine, 75732 Paris, France.

<sup>5</sup> ENSMSE - DEMO, 29, rue Ponchardier 42100 Saint-Etienne, France.

<sup>6</sup> SKLSVMS, School of Aerospace, Xi'an Jiaotong Univ, Xi'an, China 710049. [jiafeiyy@mail.xjtu.edu.cn](mailto:jiafeiyy@mail.xjtu.edu.cn)

**Abstract.** In this paper we propose a new general method for decision-making under uncertainty based on the belief interval distance. We show through several simple illustrative examples how this method works and its ability to provide reasonable results.

**Keywords:** belief functions, decision-making, distance between BBAs.

## 1 Introduction

Dempster-Shafer Theory (DST), also known as the Mathematical Theory of Evidence or the Theory of Belief Functions (BF), was introduced by Shafer in 1976 [1] based on Dempster's previous works [2]. This theory offers an elegant theoretical framework for modeling uncertainty, and provides a method for combining distinct bodies of evidence collected from different sources. In the past more than three decades, DST has been used in many applications, in fields including information fusion, pattern recognition, and decision making [3]. Although belief functions are very appealing for modeling epistemic uncertainty, the two main important questions related to them remain still open:

1. How to combine efficiently several independent belief functions?

This open question is out of the scope of this paper and it has been widely disputed by many experts [4–14]. In this short paper, we focus on the second question below.

2. How to take a final decision from a belief function?

This second question is also very crucial in many problems involving epistemic uncertainty where the final step (after beliefs elicitation, and beliefs combination) is to make a decision.

In the sequel, we assume that the reader is familiar with Dempster-Shafer Theory of belief functions [1] and its notations. Due to space restriction, we will not recall the definitions of basic belief assignment  $m(\cdot)$ , belief  $Bel(\cdot)$  (also called

credibility by some authors), and plausibility functions  $Pl(\cdot)$  functions defined over a given finite discrete frame of discernment (FoD)  $\Theta$ . For any focal element  $X$  of the powerset of  $\Theta$ , denoted by  $2^\Theta$ , the interval  $BI(X) \triangleq [Bel(X), Pl(X)]$  is called the belief interval of  $X$ . Its length  $Pl(X) - Bel(X)$  characterizes the uncertainty on  $X$  (also called ambiguity in [15]). This paper is organized as follows. In section 2, we recall the common decision-making techniques used so far to make a decision from belief functions. In section 3 we recall the new distance measure based on Belief interval, and we present a new general method for decision-making with belief functions. Finally, examples of this new approach are given in section 4, with concluding remarks in section 5.

## 2 Classical decision-making methods with belief functions

We assume a given FoD  $\Theta = \{\theta_1, \dots, \theta_n\}$  and a given BBA  $m(\cdot)$  defined on  $2^\Theta$ . We want to make a decision from  $m(\cdot)$ . It consists in choosing a particular element of the FoD that solves the problem under consideration, which is represented by the set of potential solutions (choices)  $\theta_i, i = 1, \dots, n$ . How to do this in an effective manner is the fundamental question of decision-making under epistemic uncertainty. Many decision-making criteria have been proposed in the literature. Some advanced techniques developed in the 1990s [16–19, 15] have not been widely used so far in the BF community, probably because of their complexity of implementation. In this section, we only present briefly the simplest ones frequently used.

### 1. Decision based on maximum of credibility:

This decision-making scheme is the so-called prudent (or pessimistic) scheme. It consists in choosing the element of the FoD  $\Theta$  that has the maximum of credibility. In other terms, one will decide  $\hat{\theta} = \theta_{i^*}$  with<sup>1</sup>

$$\theta_{i^*} = \arg \max_i Bel(\theta_i) \quad (1)$$

### 2. Decision based on maximum of plausibility:

On the contrary, if we prefer to adopt a more optimistic decision-making (less prudent) attitude, one will choose the element of the FoD  $\Theta$  that has the maximum of plausibility. In other terms, one will decide  $\hat{\theta} = \theta_{i^*}$  with

$$\theta_{i^*} = \arg \max_i Pl(\theta_i) \quad (2)$$

### 3. Decision based on maximum of probability:

Usually decision-makers prefer to adopt a more balanced decisional attitude making a compromise between the aforementioned pessimistic and optimistic attitudes. For this, the BBA  $m(\cdot)$  is transformed into a subjective probability measure  $P(\cdot)$  compatible with the belief interval  $[Bel(\cdot), Pl(\cdot)]$ , and one will

---

<sup>1</sup> The notation with hat indicates the decision taken. Here  $\hat{\theta}$  specifies that the decision taken is only a singleton of  $\Theta$ .

choose the element of the FoD  $\Theta$  that has the maximum of probability. In other terms, one will decide  $\hat{\theta} = \theta_{i^*}$  with

$$\theta_{i^*} = \arg \max_i P(\theta_i) \quad (3)$$

In practice, many probabilistic transformations are available to approximate (or transform) a BBA  $m(\cdot)$  in a probability measure  $P(\cdot)$ . By example, the pignistic transformation [20], the plausibility transformation [21], the DSmP transformation and other ones presented in [22], etc.

Of course, in case of multiple maximum values, no decision can be clearly drawn. Usually if only one decision must be made, a random sample between elements  $\theta_i$  generating the maximal decision-making criterion value is used to make a unique final decision  $\hat{\theta}$ . Another more prudent decision scheme is to use the disjunction of all elements generating the maximal decision-making criterion value, to provide a less specific final decision (if it is allowed for the problem under concern).

Our main criticism about using these decision-making schemes is that they do not use the whole information contained in the original BBA, which is in fact expressed by the whole belief interval. The pessimistic attitude uses only the credibility values, whereas the optimistic attitude uses only the plausibility values. The prudent attitude based on the criteria (3) requires a particular choice of probabilistic transformation which is often disputed by users. Making a decision from the  $P(\cdot)$  measure is theoretically not satisfactory at all because the transformation is lossy since we cannot retrieve  $m(\cdot)$  from  $P(\cdot)$  when some focal elements of  $m(\cdot)$  are not singletons. In the next section, we propose a better justified decision scheme based on the belief interval distance [23, 24].

### 3 Decision-making method using belief interval distance

In our previous works [23, 24], we have defined a Euclidean belief interval distance between two BBAs  $m_1(\cdot)$  and  $m_2(\cdot)$  defined on the powerset of a given FoD  $\Theta = \{\theta_1, \dots, \theta_n\}$  as follows

$$d_{BI}(m_1, m_2) \triangleq \sqrt{N_c \cdot \sum_{X \in 2^\Theta} d_W^2(BI_1(X), BI_2(X))} \quad (4)$$

where  $N_c = 1/2^{n-1}$  is a normalization factor to have  $d_{BI}(m_1, m_2) \in [0, 1]$ , and  $d_W(BI_1(X), BI_2(X))$  is the Wasserstein's distance [25] between belief intervals  $BI_1(X) \triangleq [Bel_1(X), Pl_1(X)] = [a_1, b_1]$  and  $BI_2(X) \triangleq [Bel_2(X), Pl_2(X)] = [a_2, b_2]$ . More specifically,

$$d_W([a_1, b_1], [a_2, b_2]) \triangleq \sqrt{\left[ \frac{a_1 + b_1}{2} - \frac{a_2 + b_2}{2} \right]^2 + \frac{1}{3} \left[ \frac{b_1 - a_1}{2} - \frac{b_2 - a_2}{2} \right]^2} \quad (5)$$

In [23], we have proved that  $d_{BI}(x, y)$  is a true distance metric because it satisfies the properties of non-negativity ( $d(x, y) \geq 0$ ), non-degeneracy ( $d(x, y) = 0 \Leftrightarrow x = y$ ), symmetry ( $d(x, y) = d(y, x)$ ), and the triangle inequality ( $d(x, y) + d(y, z) \geq d(x, z)$ ), for any BBAs  $x, y$  and  $z$  defined on  $2^\Theta$ . The choice of Wasserstein's distance in  $d_{BI}$  definition is justified by the fact that Wasserstein's distance is a true distance metric and it fits well with our needs because we have to compute a distance between  $[Bel_1(X), Pl_1(X)]$  and  $[Bel_2(X), Pl_2(X)]$ .

For notation convenience, we denote  $m_X$  the categorical BBA having only  $X$  as focal element, where  $X \neq \emptyset$  is an element of the powerset of  $\Theta$ . More precisely,  $m_X$  is the particular (categorical) BBA defined by  $m_X(X) = 1$  and  $m_X(Y) = 0$  for any  $Y \neq X$ . Such basic BBA plays an important role in our new decision scheme because its corresponding belief interval reduces to the degenerate interval  $[1, 1]$  which represents the certainty on  $X$ . The basic principle of the new decision scheme we propose is very simple and intuitively makes sense. It consists in selecting as the final decision (denoted by  $\hat{X}$ ) the element of the powerset for which the belief interval distance between the BBA  $m(\cdot)$  and  $m_X$ ,  $X \in 2^\Theta \setminus \{\emptyset\}$  is the smallest one<sup>7</sup>. Therefore, take as the final decision  $\hat{X}$  given by

$$\hat{X} = \arg \min_{X \in 2^\Theta \setminus \{\emptyset\}} d_{BI}(m, m_X) \quad (6)$$

where  $d_{BI}(m, m_X)$  is computed according to (4).  $m(\cdot)$  is the BBA under test and  $m_X(\cdot)$  the categorical BBA focused on  $X$  defined above.

This decision scheme is very general in the sense that the decision making can be done on any type of element<sup>2</sup> of the power-set  $2^\Theta$ , and not necessarily only on the elements (singletons) of the FoD (see examples in the next section). This method not only provides the final decision  $\hat{X}$  to make, but also it evaluates how good this decision is with respect to its alternatives if we define the quality indicator  $q(\hat{X})$  as follows

$$q(\hat{X}) \triangleq 1 - \frac{d_{BI}(m, m_{\hat{X}})}{\sum_{X \in 2^\Theta \setminus \{\emptyset\}} d_{BI}(m, m_X)} \quad (7)$$

One sees that the quality indicator  $q(\hat{X})$  of the decision  $\hat{X}$  made will become maximum (equal to one) when the distance between the BBA  $m(\cdot)$  and  $m_{\hat{X}}$  is zero, which means that the BBA  $m(\cdot)$  is focused in fact only on the element  $\hat{X}$ . The higher  $q(\hat{X})$  is, the more confident in the decision  $\hat{X}$  we should be.

Of course, if a decision must be made with some extra constraint<sup>3</sup> defined by a (or several) condition(s), denoted  $c(X)$ , then we must take into account  $c(X)$  in Eq. (6), that is  $\hat{X} = \arg \min_{X \in 2^\Theta \setminus \{\emptyset\} \text{ s.t. } c(X)} d_{BI}(m, m_X)$ , and also in the derivation of quality indicator by taking  $\sum_{X \in 2^\Theta \setminus \{\emptyset\} \text{ s.t. } c(X)} d_{BI}(m, m_X)$  as

<sup>7</sup> This simple principle has also been proposed by Essaid et al. [26] using Jousselme's distance.

<sup>2</sup> empty set excluded.

<sup>3</sup> for instance, making a choice only among the singletons of  $2^\Theta$ .

denominator in (7). Theoretically any other strict distance metric, for instance Jousselme's distance [27–29], could be used instead of  $d_{BI}(\cdot, \cdot)$ . We have chosen  $d_{BI}$  distance because of its ability to provide good and reasonable behavior [23] as will be shown. When there exists a tie between multiple decisions  $\{\hat{X}_j, j > 1\}$ , then the prudent decision corresponding to their disjunction  $\hat{X} = \cup_j \hat{X}_j$  should be preferred (if allowed), otherwise the final decision  $\hat{X}$  is made by a random selection of elements  $\hat{X}_j$ .

## 4 Examples and comparisons

In this section we present several examples when the cardinality of the FoD  $|\Theta|$  is only 2 and 3 because it is easier to see whether the decision-making results make sense or not. We compare and discuss decisions only made with the belief interval distance  $d_{BI}$  and Jousselme's distance  $d_J$  because the other lossy decision schemes do not exploit both credibility and plausibility values. The examples corresponding to cases where the BBA  $m(\cdot)$  is focused on a single element  $X$  of  $2^\Theta$  are not presented because one trivially gets  $\hat{X} = X$  using either  $d_{BI}$  or  $d_J$  distances. The next tables present several BBAs from which a decision has to be made. By convention, and since we work with normal BBAs satisfying  $m(\emptyset) = 0$ , the empty set is not included in the tables. The rows for  $d_{BI}^{\min}(m_i, m_X)$  and for  $d_J^{\min}(m_i, m_X)$  list the minimal values obtained for  $d_{BI}(m_i, m_X)$  and  $d_J(m_i, m_X)$ . The rows for  $\hat{X}^{d_{BI}}$  and for  $\hat{X}^{d_J}$  list the decision(s)  $\hat{X}$  made when using  $d_{BI}(m_i, m_X)$  and  $d_J(m_i, m_X)$  respectively. The rows for  $q(\hat{X}^{d_{BI}})$  and  $q(\hat{X}^{d_J})$  list the quality indicators of decision(s) made using  $d_{BI}(m_i, m_X)$  and  $d_J(m_i, m_X)$  respectively. Depending on the BBA, it is possible to have multiple decisions  $\{\hat{X}_j\}$  in case of a tie. If a tie occurs either a random sampling of  $\{\hat{X}_j\}$  must be drawn, or (if allowed) the disjunction of decisions  $\hat{X}_j$  is preferred. In the next subsections, we present results in free-constraint case (i.e.  $c(X) = \emptyset$ ), as well as when the decisions are restricted to be singletons (i.e.  $c(X) \equiv "|X| = 1"$ ).

### 4.1 Examples with $\Theta = \{A, B\}$

Table 1 shows the decisions made when there is no constraint on the cardinality of the decision  $\hat{X}$ .

One sees that methods based on min of  $d_{BI}(m, m_X)$  and on min of  $d_J(m, m_X)$  yield the same reasonable decisions in almost all cases. With  $m_2$ , one has multiple decisions  $\hat{X}^{d_J} = \{A, B, A \cup B\}$  with quality 0.6667 when using  $d_J$ , which is a bit surprising in our opinion because there is a real tie between  $A$  and  $B$ . Consequently, the decision  $A \cup B$  should be preferred when there is no constraint on the cardinality of decisions. For this  $m_2$  case, one gets a unique decision  $\hat{X}^{d_{BI}} = A \cup B$  with a better quality 0.776 which seems more reasonable. We see also that all minimal distance values obtained with  $d_{BI}$  are less (or equal in case  $m_1$ ) to the minimal values obtained with  $d_J$ . In fact, when the mass function is distributed symmetrically, it is naturally expected that no real decision can be easily taken (as illustrated for BBA's  $m_2(\cdot)$  and  $m_5(\cdot)$  in Table 1). Here, the

**Table 1.** Examples of several BBA's and decisions made (no constraint case).

$X \in 2^\Theta$	$m_1(\cdot)$	$m_2(\cdot)$	$m_3(\cdot)$	$m_4(\cdot)$	$m_5(\cdot)$	$m_6(\cdot)$	$m_7(\cdot)$
$A$	0.9	0.5	0.8	0.1	0.4	0.9	0.1
$B$	0.1	0.5	0.1	0.1	0.4	0	0
$A \cup B$	0	0	0.1	0.8	0.2	0.1	0.9
$d_{BI}^{\min}(m_i, m_X)$	0.1000	0.2887	0.1528	0.0577	0.2309	0.0577	0.0577
$q(\hat{X}^{d_{BI}})$	0.9330	0.7760	0.8939	0.9502	0.8134	0.9622	0.9513
$\hat{X}^{d_{BI}}$	$A$	$A \cup B$	$A$	$A \cup B$	$A \cup B$	$A$	$A \cup B$
$d_J^{\min}(m_i, m_X)$	0.1000	0.5000	0.1581	0.1000	0.4000	0.0707	0.0707
$q(\hat{X}^{d_J})$	0.9390	0.6667	0.8999	0.9276	0.6409	0.9574	0.9501
$\hat{X}^{d_J}$	$A$	$A, B, A \cup B$	$A$	$A \cup B$	$A \cup B$	$A$	$A \cup B$

decision  $A \cup B$  for BBA's  $m_2(\cdot)$  and  $m_5(\cdot)$  can be interpreted as a no proper decision, in the sense that  $A \cup B$  is the whole universe of discourse, hence we are merely selecting anything (and discarding nothing). Such kind of no proper decision may however be very helpful in some fusion systems because it warns that input information is not rich enough, and that one needs more information to take a proper decision (by including more sensors or more experts reports in the system for instance). For symmetrical mass function, the decision drawn from the new proposed decision rule is consistent with what we can reasonably get because. To make a proper decision we will always need to introduce some possibly arbitrary additional constraints.

Table 2 shows the decisions made for same examples when we force the decision to be a singleton, that is when the constraint is  $c(X) \equiv |X| = 1$ . One sees that the decisions restricted to the set of singletons using  $d_{BI}(m, m_X)$  or  $d_J(m, m_X)$  are the same but the quality indicators are a bit better when using  $d_{BI}(m, m_X)$  with respect to  $d_J(m, m_X)$ . The values of the quality indicators in Table 2 are different to those of Table 1 which is normal because we use the constraint  $c(X)$  in the denominator of the formula (7).

**Table 2.** Examples of several BBA's and decisions made (restricted to singletons).

$X \in 2^\Theta$	$m_1(\cdot)$	$m_2(\cdot)$	$m_3(\cdot)$	$m_4(\cdot)$	$m_5(\cdot)$	$m_6(\cdot)$	$m_7(\cdot)$
$A$	0.9	0.5	0.8	0.1	0.4	0.9	0.1
$B$	0.1	0.5	0.1	0.1	0.4	0	0
$A \cup B$	0	0	0.1	0.8	0.2	0.1	0.9
$d_{BI}^{\min}(m_i, m_X)$	0.1000	0.5000	0.1528	0.5508	0.5033	0.0577	0.5196
$q(\hat{X}^{d_{BI}})$	0.9000	0.5000	0.8477	0.5000	0.5000	0.9427	0.5393
$\hat{X}^{d_{BI}}$	$A$	$A, B$	$A$	$A, B$	$A, B$	$A$	$A$
$d_J^{\min}(m_i, m_X)$	0.1000	0.5000	0.1581	0.6403	0.5099	0.0707	0.6364
$q(\hat{X}^{d_J})$	0.9000	0.5000	0.8434	0.5000	0.5000	0.9308	0.5276
$\hat{X}^{d_J}$	$A$	$A, B$	$A$	$A, B$	$A, B$	$A$	$A$

#### 4.2 Examples with $\Theta = \{A, B, C\}$

Table 3 shows the decisions made when there is no constraint on the cardinality of the decision  $\hat{X}$ , whereas Table 4 shows the results for the same examples when the decisions made are restricted to singletons. As shown in the tables all minimal distance values obtained with  $d_{BI}$  are less (or equal) to the minimal values obtained with  $d_J$  and the quality indicator decisions is better when computed with  $d_{BI}$  (except in case  $m_1$  of Table 3). The decisions results obtained with  $d_J$  are mostly consistent with those obtained with  $d_{BI}$  (except in case  $m_2$  and  $m_3$  of Table 3) where a larger set of decisions (tie) is obtained using  $d_J$ . If the decisions are restricted to singletons (see Table 4), then the decision-making based on  $d_{BI}$  and on  $d_J$  provides the same results with a better quality of decisions using  $d_{BI}$ .

**Table 3.** Examples of several BBA's and decisions made (no constraint case).

$X \in 2^\Theta$	$m_1(\cdot)$	$m_2(\cdot)$	$m_3(\cdot)$	$m_4(\cdot)$	$m_5(\cdot)$	$m_6(\cdot)$	$m_7(\cdot)$	$m_8(\cdot)$
$A$	0.9	0.5	1/3	0.5	0	0	0	0.2
$B$	0.1	0.5	1/3	0	0	0	0	0.1
$A \cup B$	0	0	0	0.5	0.5	2/3	1/3	0.05
$C$	0	0	1/3	0	0	0	0	0.05
$A \cup C$	0	0	0	0	0	0	1/3	0.1
$B \cup C$	0	0	0	0	0.5	1/3	1/3	0.2
$A \cup B \cup C$	0	0	0	0	0	0	0	0.3
$d_{BI}^{\min}(m_i, m_X)$	0.1000	0.2887	0.4082	0.2887	0.2887	0.1925	0.2357	0.2227
$q(\hat{X}^{d_{BI}})$	0.9776	0.9242	0.8787	0.9271	0.9120	0.9421	0.9241	0.9280
$\hat{X}^{d_{BI}}$	$A$	$A \cup B$	$2^\Theta \setminus \{\emptyset, A, B, C\}$	$A, A \cup B$	$A \cup B, B \cup C, \Theta$	$A \cup B$	$A \cup B \cup C$	$A \cup B \cup C$
$d_J^{\min}(m_i, m_X)$	0.1000	0.5000	0.5774	0.3536	0.4082	0.2722	0.3333	0.3149
$q(\hat{X}^{d_J})$	0.9798	0.8870	0.8571	0.9225	0.8989	0.9337	0.9111	0.9152
$\hat{X}^{d_J}$	$A$	$A, B, A \cup B$	$2^\Theta \setminus \{\emptyset\}$	$A, A \cup B$	$A \cup B, B \cup C, \Theta$	$A \cup B$	$A \cup B \cup C$	$A \cup B \cup C$

**Table 4.** Examples of several BBA's and decisions made (restricted to singletons).

$X \in 2^\Theta$	$m_1(\cdot)$	$m_2(\cdot)$	$m_3(\cdot)$	$m_4(\cdot)$	$m_5(\cdot)$	$m_6(\cdot)$	$m_7(\cdot)$	$m_8(\cdot)$
$A$	0.9	0.5	1/3	0.5	0	0	0	0.2
$B$	0.1	0.5	1/3	0	0	0	0	0.1
$A \cup B$	0	0	0	0.5	0.5	2/3	1/3	0.05
$C$	0	0	1/3	0	0	0	0	0.05
$A \cup C$	0	0	0	0	0	0	1/3	0.1
$B \cup C$	0	0	0	0	0.5	1/3	1/3	0.2
$A \cup B \cup C$	0	0	0	0	0	0	0	0.3
$d_{BI}^{\min}(m_i, m_X)$	0.1000	0.5000	0.5774	0.2887	0.5000	0.5092	0.6236	0.5770
$q(\hat{X}^{d_{BI}})$	0.9488	0.7321	0.6667	0.8531	0.7388	0.7364	0.6667	0.6855
$\hat{X}^{d_{BI}}$	$A$	$A, B$	$A, B, C$	$A$	$B$	$B$	$A, B, C$	$A$
$d_J^{\min}(m_i, m_X)$	0.1000	0.5000	0.5774	0.3536	0.5774	0.5932	0.6667	0.6117
$q(\hat{X}^{d_J})$	0.9488	0.7321	0.6667	0.8300	0.7257	0.7229	0.6667	0.6836
$\hat{X}^{d_J}$	$A$	$A, B$	$A, B, C$	$A$	$B$	$B$	$A, B, C$	$A$

## 5 Conclusions

We have presented a new method for decision-making with belief functions which truly exploits the belief interval value of each focal element of a BBA. It is easy to implement and can be applied with any strict distance metric between two BBAs. We have considered and compared the well-known Jousselme's distance and the recent belief interval distance. This method is general because the decision can be made not only on singletons, but also on any other compound focal elements (if needed and allowed). It also provides a quality indicator of the decision made.

## References

1. Shafer, G.: *A Mathematical Theory of Evidence*. Princeton University Press (1976)
2. Dempster, A.: Upper and lower probabilities induced by a multivalued mapping. *Ann. Math. Statist.*, Vol. 38, 325–339 (1967)
3. Smets, P.: Practical uses of belief functions. In: K. B. Laskey and H. Prade, Editors, *Uncertainty in Artificial Intelligence 15*, pp. 612–621, Stockholm, Sweden (1999)
4. Zadeh, L.A.: On the validity of Dempster's rule of combination. Memo M79/24, Univ. of California, Berkeley, U.S.A. (1979)
5. Lemmer, J.: Confidence factors, empiricism and the Dempster-Shafer theory of evidence. In: Proc. of 1st Conf. on Uncertainty in Artificial Intelligence (UAI-85), pp. 160–176 (1985)
6. Pearl, J.: Reasoning with belief functions: An analysis of compatibility. *IJAR*, Vol. 4, pp. 363–389, (1990), with Rejoinder in *IJAR*, Vol. 6, pp. 425–443, (1992)
7. Voorbraak, F.: On the justification of Dempster's rule of combination. Dept. of Philosophy, Univ. of Utrecht, The Netherlands, Logic Group Preprint Series, No. 42 (1988)
8. Wang, P.: A defect in Dempster-Shafer theory. In: Proc. of 10th Conf. on Uncertainty in AI, pp. 560–566 (1994)
9. Gelman, A.: The boxer, the wrestler, and the coin flip: a paradox of robust Bayesian inference and belief functions. *American Statistician*, Vol. 60, No. 2, pp. 146–150 (2006)
10. Dezert, J., Wang, P., Tchamova, A.: On The Validity of Dempster-Shafer Theory. In: Proc. of Fusion 2012, Singapore (2012)
11. Tchamova, A., Dezert, J.: On the Behavior of Dempster's Rule of Combination and the Foundations of Dempster-Shafer Theory. In: Proc. of IEEE IS2012 Conf., Sofia, Bulgaria (2012)
12. Dezert, J., Tchamova, A., Han, D., Tacnet, J.-M.: Why Dempster's rule doesn't behave as Bayes rule with informative priors. In: Proc. of 2013 IEEE INISTA 2013, Albena, Bulgaria (2013)
13. Dezert, J., Tchamova, A.: On the validity of Dempster's fusion rule and its interpretation as a generalization of Bayesian fusion rule. *IJIS*, Vol. 29, Issue 3, pp. 223–252 (2014)
14. Smarandache, F., Dezert, J.: Advances and applications of DSmT for information fusion. Vols. 1–4, ARP, USA (2004–2015) <http://www.onera.fr/fr/staff/jean-dezert>
15. Srivastava, R.P.: Decision Making Under Ambiguity: A Belief-Function Perspective. *Archives of Control Sciences*, Vol. 6 (XLII), No. 1-2, pp. 5–27 (1997)

16. Nguyen, H.T., Walker, E.A.: On Decision Making Using Belief Functions, in Advances in the Dempster-Shafer Theory of Evidence. Edited by R.R. Yager, M. Fedrizzi, and J. Kacprzyk, John Wiley & Sons, New York, NY, USA (1994)
17. Jaffray, J-Y.: Utility Theory for Belief Functions. *Oper. Res. Letters*, Vol. 8, pp. 107-112 (1989)
18. Strat, T.M.: Decision Analysis Using Belief Functions. *IJAR*, Vol. 4, No. 5, p. 6 (1990)
19. Yager, R.R.: Decision making under Dempster-Shafer uncertainties. *Int. J. general systems*, Vol. 20, pp. 233–245 (1992)
20. Smets, P., Kennes, R.: The transferable belief model. *Artif. Int.*, Vol. 66, pp. 191–234 (1994)
21. Cobb, B.R., Shenoy, P.P.: On the plausibility transformation method for translating belief function models to probability models. *IJAR*, Vol. 41, No. 3, pp. 314–330 (2006)
22. Dezert, J., Smarandache, F.: A new probabilistic transformation of belief mass assignment. In: Proc. of Fusion 2008, Cologne, Germany (2008)
23. Han, D., Dezert, J., Yang, Y.: New Distance Measures of Evidence based on Belief Intervals. In: Proc. of Belief 2014, Oxford, UK (2014)
24. Han, D., Dezert, J., Yang, Y.: Belief Interval Based Distance Measures in the Theory of Belief Functions. (submitted to Journal Jan. 2016).
25. Irpino, A., Verde, R.: Dynamic Clustering of Interval Data Using a Wasserstein-based Distance. *Pattern Recognition Letters*, Vol. 29, pp. 1648–1658 (2008)
26. Essaid, A., Martin, A., Smits, G., Ben Yaghane, B.: A distance-based decision in the credal level. *Int. Conf. AISC 2014*, Sevilla, Spain, (2014).
27. Jousselme, A.-L., Grenier, D., Bossé É.: A new distance between two bodies of evidence. *Information Fusion*, Vol. 2, No. 2, pp. 91–101 (2001)
28. Jousselme, A.-L., Maupin, P.: Distances in evidence theory: Comprehensive survey and generalizations. *IJAR*, Vol. 53, No. 2, pp. 118–145 (2012)
29. Bouchard M., Jousselme A.-L., Doré P.-E.: A proof for the positive definiteness of the Jaccard index matrix. *IJAR*, Vol. 54, pp. 615–626 (2013)