Belief assignment on compound hypotheses within the framework of the Transferable Belief Model

Alexandre Veremme, Daniel Dupont, Eric Lefevre, David Mercier

Belief assignment on compound hypotheses within the framework of the Transferable Belief Model

Alexandre Veremme[1,2], Daniel Dupont[1], Éric Lefevre[2], David Mercier[2]
Université Lille Nord de France, F-59000 Lille France
[1] HEI, Département Ingénierie et Sciences du Vivant
13 rue de Toul, F-59800 Lille, firstname.name@hei.fr
[2] U-Artois, LGI2A EA 3926
Technoparc Futura, F-62400 Béthune, firstname.name@univ-artois.fr

Abstract – *Within the framework of evidence theory, the construction of mass functions remains an important problem which can considerably influence final results. Several propositions of construction are available in literature but generally, these methods do not consider the belief masses on compound hypotheses and do not exploit all the strengths of the theory. In this article, we propose two methods of grouping hypotheses allowing to reduce the number of focal elements during the construction of mass functions. The first one is based on a well-known approach in cartography and geoinformation fields: Jenks' method. The second one is based on information criteria. We present and confront both methods. Results in a forensic entomology application are then presented.*

Keywords: mass function construction, compound hypotheses, belief theory, evidence theory, Jenks' method, information criteria, forensic entomology.

1 Introduction

Representing and managing knowledge in complex systems, where quantity and disparity of information are important, generally requires the system designer to set up a particular formalism of information processing. The goal is to better reason with imperfect information (imprecision, uncertainty, incompleteness *etc.*) and to handle these imperfections to make a final decision. Among the existing theories, the belief function theory [1] could be well adapted to take into account these different characteristics.

The transferable belief model (TBM) of Philippe Smets [2] can be seen as a subjective and non probabilistic interpretation of evidence theory. It is a flexible and efficient tool to be applied in many areas such as decision support, diagnosis *etc.* The TBM is all the more interesting as it has a range of tools to handle belief functions (discounting, conditioning operations *etc.*). Despite the importance of the model foundations, a significant problem remains: the construction (or estimation) of basic belief assignments.

Construction is a difficult step that is not universal and which generally depends on the application and the vision we have on it. We usually distinguish the following methods of estimation: those based on the calculations of distances [3, 4] and those based on likelihood inspirations [5, 6]. However, these two types of estimation present various inconveniences: the first one can supply very engaged mass functions but a solution can always be obtained even if the frame of discernment is large. The second method respects more the Minimal Commitment Principle and seems better to make a careful decision. But we are then confronted with a computing problem when the frame of discernment increases. This last method becomes thus unusable because of the number of focal elements, close to two to the power of the cardinality of the frame of discernment. Different works deal with the assignment of belief masses on the disjunctions (*compound hypotheses*) to better use the power of evidence theory and so to reduce the space complexity. Affectations on the disjunctions can be classified in two axes: those where a knowledge on the hypotheses is *a priori* known [7] and those consisting in a learning of the compound focal elements [8]. More recently, [9] also introduces imprecision on mass functions by using dual fuzzy morphological operators. Interested readers can find a good range of methods of estimation in [10].

Within the framework of our works, consisting in realizing a decision support system intended for forensic medicine, we are confronted with this problem of construction. According to experts' observations, we try to determine the time of death of a person with a possible large frame of discernment (several hundred hours). In this paper, we present two methods allowing to reduce the number of focal elements and to affect our belief masses on disjunctions.

After a display of the transferable belief model (section 2), we propose two methods allowing to estimate the belief masses on disjunctions of hypotheses (section 3). The first one is based on a well-known method

in cartography and geoinformation fields: Jenks' natural breaks classification [11]. The second one is based on an approach of information criteria and originally comes from the study of histograms [12]. We are going to detail these two methods of estimation, present and compare their advantages and inconveniences in the framework of belief functions. We apply them to our forensic medicine problem and end on the interests of such methods in the framework of the TBM.

2 Transferable belief model

The evidence theory was initially introduced by Dempster [13] in relation to his work on lower and upper bounds of a distribution probability family. Using this mathematical formalism, Shafer [1] shows the benefits of belief functions to model uncertain knowledge. The usefulness of belief functions, as an alternative to subjective probabilities, was subsequently proved in an axiomatic way by Smets [2] in the transferable belief model thus giving a clear and coherent interpretation of the subjacent concept of this theory. In this section, the essential mathematical elements of the TBM will be introduced.

2.1 Basic concepts

The transferable belief model is based on the assumption that, from the beginning, a set Ω called the frame of discernment is known and which is defined as follows:

$$\Omega = \{\omega_1, \omega_2, \dots \omega_N\} \tag{1}$$

This set is composed of N exclusive hypotheses. From this frame of discernment, a power set noted 2^Ω can be built, including the 2^N proposals A of Ω:

$$2^\Omega = \{A | A \ \subseteq \Omega\} \tag{2}$$

A belief function can be mathematically defined by a mass function (or allocation [1]), noted m defined by 2^Ω in $[0, 1]$, and that verifies:

$$\sum_{A \subseteq \Omega} m(A) = 1 \tag{3}$$

Each subset $A \subseteq \Omega$ such that $m(A) \neq 0$ is called a focal element of Ω. Thus, mass $m(A)$ represents the degree of belief allocated to the proposal A and that cannot, in the present state of knowledge, be attributed to a more specific subset than A. The belief function for which $m(\emptyset) = 0$ is called normal. In the transferable belief model, the condition $\sum_{\emptyset \neq A \subseteq \Omega} m(A) = 1$ is not imposed and it is possible to have $m(\emptyset) \neq 0$. This can introduce the notion of *open world* while assuming that the belief cannot be attributed to a subset of Ω. In this case, $\emptyset$ can be interpreted as a proposal which is not in the frame of discernment Ω and that it is likely to be the solution to the problem as opposed to the *closed world* where the set Ω is assumed to be exhaustive.

[1]The term *bba* for *basic belief assignment is often found in literature*

2.2 Discounting

When the information that results in the belief function is not totally reliable, it may be useful to discount this belief. In order to do that, a coefficient α is used, which represents knowledge of the source reliability. This coefficient will allow the transfer of the belief to the set Ω when the information is not totally reliable. The discounting belief function m^α, defined by a reliability coefficient α can then be deduced from m by means of the following expression:

$$\begin{cases} m^\alpha(A) = \alpha m(A) \\ m^\alpha(\Omega) = 1 - \alpha + \alpha m(\Omega). \end{cases} \tag{4}$$

In literature several methods have been developed to compute the discounting factor [14, 15].

2.3 Fusion of belief functions

In the transferable belief model, data from distinct sources are fused using the *conjunctive rule of combination*. This rule, which is commutative and associative, is defined by:

$$\forall\, A \in 2^\Omega \qquad m(A) = m_1(A) \textcircled{\cap} \dots \textcircled{\cap} m_Q(A) \tag{5}$$

where $\textcircled{\cap}$ represents the combination operator. In cases of two sources noted S_i and S_j, giving respectively belief functions noted m_i et m_j, the combination can be written as follows:

$$m(C) = \sum_{A \cap B = C} m_i(A).m_j(B) \tag{6}$$

The value $m(\emptyset) = m_1 \textcircled{\cap} m_2(\emptyset)$ reflects the existing conflict between two sources S_i and S_j. When this factor equals 1, the sources are in total conflict. Various works deal anyway with the problem of the appearance of conflict with this combination rule [16]. In a more general way, within the framework of the evidence theory, different combination rules, not defined in the TBM, were proposed. Interested readers could refer to [17, 18, 19] for more details about some of these rules.

2.4 Pignistic transformation

Within the framework of the TBM, Smets [2] recommends to construct a probability function in order to take an optimal decision allowing to make decision on individual hypotheses. The transformation from a mass function to a probability function is achieved by the *pignistic* transformation and is given by:

$$\forall\, \omega_n \in \Omega \qquad BetP(\omega_n) = \frac{1}{1 - m(\emptyset)} \sum_{A \ni \omega_n} \frac{m(A)}{|A|} \tag{7}$$

where $|A|$ represents the cardinality of $A \subseteq \Omega$.

Once the *pignistic* probability obtained, it is possible to use classic tools of statistical decision theory.

Readers could find justifications and details of this transformation in [20].

When comparing it to the probability theory, the transferable belief model shows some advantages, the most important of which is the possibility of expressing the degree of uncertainty. In this model, the mass assignment on a subset does not mean that the remainder automatically goes to the complement. After this introduction to the basic concepts of the TBM, the next section deals with the construction of mass functions.

3 Construction of mass functions

As mentioned in the introduction, the wealth and flexibility of the TBM constitute an interesting frame to represent and manage incomplete information. However, the problem of the construction of basic belief assignments persists and the modelling choice can have important consequences on the final results and decisions. In the next sections (3.2 and 3.3), we will present two methods of automatic grouping of hypotheses to construct mass functions (section 3.4).

3.1 Context and objectives

Let $S = \{S_1, \ldots, S_j, \ldots, S_J\}$ be a set of sources expressing themselves on the validity of the hypotheses $\{\omega_1, \ldots, \omega_n, \ldots, \omega_N\} = \Omega$ with regard to a given question Q. In a general way, each source S_j can answer: "Is the hypothesis ω_n a possible solution of the question Q?". The validity degree $\delta^Q_{j,n}$ associated to this answer varies between the values 0 and 1. It takes the value 0 when the hypothesis is not a solution of the question Q and the value 1 when the hypothesis is exactly a possible solution. So a source S_j supplies a distribution of validity degrees defined by $\Delta^Q_{j,\Omega} : \Omega \mapsto [0,1]$. The final objective in general is to find the hypothesis ω_n (or the set of hypotheses Z_λ) according to the validity degree(s) (function of a given criterion relating to the application and the objectives for example).

When the number N is important ($N > 15$), it is generally difficult to work by means of belief functions because of the number of focal elements. And it is not easy to estimate correctly the *bbas* with a good report commitment/time calculations. So, the purpose is to obtain an estimation $\widehat{\Delta^Q_{j,Z}}$ of the distribution $\Delta^Q_{j,\Omega}$ to reduce these focal elements afterwards. This estimate distribution is defined by $\widehat{\Delta^Q_{j,Z}} : Z \mapsto [0,1]$ where Z corresponds to the set of the disjunctions of hypotheses having been grouped together in an "intelligent" way according to the validity degrees: $Z = \{z_1, ..., z_\lambda, ..., z_\Lambda\} \subseteq \Omega$, with $\Lambda < N$.

In this article, two methods allowing to get an estimation of the distribution $\Delta^Q_{j,\Omega}$ are presented. This estimation will then be at the basis of the assignment of belief masses. The first method is based on a well-known method in cartography and geoinformation fields: Jenks' natural breaks classification [21, 11]. The last one is based on an approach of information criterion and originaly comes from the optimization of histograms. Both methods also allow to reduce complexity by keeping a satisfactory expressiveness to make decisions. In the next sections, we present these two methods and go on the comparison in the framework of the forensic application.

3.2 Grouping of hypotheses based on Jenks' approach

Jenks' algorithm is a much used empirical approach within the geographical information systems (GIS) [22, 21]. It allows an optimal classification [11] of a population in a certain number of classes (or partitions) and minimizes the sum of absolute deviations about class means. It starts with an arbitrary set of classes, calculates an error (also named GVF for Goodness of Variance Fit index), and attempts to maximise it by moving observations between adjacent classes. The optimization so minimizes within-class variance and maximize between-class variance in an iterative series of calculations.

In short, Jenks' algorithm can be broken down into separate sequences:
1). the determination of the class number,
2). an iterative heuristics allowing to maximize the GVF index.

3.2.1 Jenks' optimized breaks algorithm

We consider the source S_j, the set of hypotheses Ω and the relative ditribution of validity degrees $\Delta^Q_{j,\Omega}$ with regard to a certain question Q (see the previous section).

For a fixed number Λ of classes and a number N of hypotheses, there is a number Γ, equaled to $\frac{N!}{\Lambda!.(N-\Lambda)!}$, of possible combinations. Let us denote $Z_{\Lambda,N} = \{Z_1, \ldots, Z_\gamma, \ldots, Z_\Gamma\}$ the set of these possible combinations. We try to get the best combination Z which maximizes the GVF index. The GVF index is given by the following equation:

$$GVF(Z_\gamma) = \frac{SDAM - SDCM(Z_\gamma)}{SDAM} \tag{8}$$

where the SDAM index is the Squared Deviations from the Array Mean. It is obtained by:

$$SDAM = \sum_{\omega_n \in \Omega} (\delta^Q_{j,n} - Mean(\Omega))^2 \tag{9}$$

with $Mean(\Omega)$, the mean of the validity degrees $\delta^Q_{j,n}$. In the equation (8), the SDCM value is the Squared Deviations from the Class Means:

$$SDCM(Z_\gamma) = \sum_{z_\lambda \in Z_\gamma} \sum_{\omega_n \in z_\lambda} (\delta^Q_{j,n} - Mean(z_\lambda))^2 \tag{10}$$

In (10), $Mean(z_\lambda)$ is the mean of the validity degrees of the class z_λ with for the combination Z_γ.

We notice that the GVF index so varies between 0 and 1 (its value is 1 when $\Lambda = N$). This index can be viewed as an entropic index that we maximize for a fixed number of classes. A heuristic has to be developed to cross the set of $Z_{\Lambda,N}$ possible combinations. We do not detail it in this article.

3.3 Grouping of hypotheses based on the approach of information criteria

The main idea is to sum up the information given by each source S_j by means of an optimum histogram in the sense of the maximum likelihood and of a mean square cost. As previously, sources could supply degrees of validity $\delta^Q_{j,n}$ on hypotheses $\omega_n \in \Omega$. First, we get the distribution of validity degrees given by each source and then make the optimum histogram building which is led by the use of an information criterion. We will see that different information criteria initially designed for model selection can be used.

3.3.1 Histogram approximation

We consider the source S_j, the set Ω of N hypotheses and the relative distribution of validity degrees $\Delta^Q_{j,\Omega}$ with regard to a question Q of interest (see the section 3.1). The aim is to approximate $\Delta^Q_{j,\Omega}$ with a histogram $\widehat{\Delta^Q_{j,Z}}$ built on a subpartition $Z = \{z_1, ..., z_\lambda, ..., z_\Lambda\} \subseteq \Omega$ with Λ bins such as $\Lambda < N$. The distribution $\widehat{\Delta^Q_{j,Z}}$ built with Z is an optimum estimation of $\Delta^Q_{j,\Omega}$ according to a cost function to define. Z results from an information criterion called IC issued from the basic Akaike's information criterion (AIC) [12], AIC^* or ϕ^* [23] which are respectively Hannan-Quinn's criterion and Rissanen's criterion. These criteria have the following form:

$$IC(\Lambda) = g(\Lambda) - 2 \sum_{z_\lambda \in Z} \widehat{\Delta^Q_{j,Z}}(z_\lambda) \ln \frac{\widehat{\Delta^Q_{j,Z}}(z_\lambda)}{Card(z_\lambda)} \quad (11)$$

where $g(\Lambda)$ is a penalty term which differs from one criterion to another one. $Card(z_\lambda)$ is the number of hypotheses $\omega_k \in z_\lambda$. $\widehat{\Delta^Q_{j,Z}}(z_\lambda)$ is the maximum likelihood estimator for the partition z_λ which is given by the following equation:

$$\widehat{\Delta^Q_{j,Z}}(z_\lambda) = \frac{1}{\sum_{\omega_n \in \Omega} \delta^Q_{j,n}} \sum_{\omega_k \in z_\lambda} \delta^Q_{j,k} \quad (12)$$

An optimum histogram $\widehat{\Delta^Q_{j,Z}}$ to approximate the distribution $\Delta^Q_{j,\Omega}$ is obtained in two steps. The first one consists in merging two contiguous bins in a histogram with Λ bins among the $(\Lambda - 1)$ possible fusions of two bins. This is made by minimizing the IC criterion. The second one consists in finding the "best" histogram with Λ bins. The optimum histogram with $\Lambda = \Lambda_{opt}$ bins is the one which minimizes IC.

3.3.2 Selection of the bin number of a histogram

Obtaining the optimum histogram is based on the use of an information criterion IC which gives the optimal number of bins thanks to a cost function based on Kullback's contrast or Hellinger's distance. We do not detail these cost functions but if the cost function is expressed according to KullBack's contrast [24], we obtain the criterion such as :

$$AIC^*(\Lambda) = \frac{\Lambda\ (1 + \ln P)}{P} - 2 \sum_{z_\lambda \in Z} \widehat{\Delta^Q_{j,Z}}(z_\lambda) \ln \frac{\widehat{\Delta^Q_{j,Z}}(z_\lambda)}{Card(z_\lambda)} \quad (13)$$

where $P = \sum_{\omega_n \in \Omega} \delta^Q_{j,n}$.

This criterion can be used to select the optimum histogram with Λ bins to approximate the distribution of validity degrees. Other criteria exist (*e.g.* AIC or ϕ^* criteria) but we do not detail them in this article, simply because the criterion AIC^* gave us better results in our application in term of time of calculations and final decision obtained. Detailed demonstrations for this section are available in [12, 23].

3.3.3 Optimum histogram building process

At first, an initial histogram with N bins is built thanks to the distribution $\Delta^Q_{j,\Omega}$. Then, a partition with $(N - 1)$ bins is considered. For each possible fusion of two contiguous bins among $(N-1)$ the criterion $IC(N-1)$ is computed. The choice of the best fusion is made according to the minimization of $IC(N - 1)$. When it is done, we look for the best partition with $(N - 2)$ bins according to the same rule. Finally, the histogram with Λ bins that minimize $IC(\Lambda)$ for $\Lambda \in \{1, \ldots, N\}$ is retained.

3.4 Mass function construction by affecting masses on disjunctions

Having detailed the methods of grouping hypotheses, we present in this part the method of mass assignments. Whatever the previous method used, we obtain the set $Z = \{z_1, ..., z_\lambda, ..., z_\Lambda\} \subseteq \Omega$ of compound hypotheses. For a source S_j, we propose to make the average of the validity degrees of each z_λ-class to affect the masses. We create as many basic belief assignments as disjunctions z_λ:

$$\begin{cases} m_{j,\lambda}(z_\lambda) &= 0 \\ m_{j,\lambda}(\overline{z_\lambda}) &= \alpha_j \ (1 - \frac{Mean(z_\lambda)}{\sum_{z_\lambda \in Z} Mean(z_\lambda)}) \\ m_{j,\lambda}(\Omega) &= 1 - \alpha_j + \alpha_j \ . \ \frac{Mean(z_\lambda)}{\sum_{z_\lambda \in Z} Mean(z_\lambda)} \end{cases} \tag{14}$$

In (14), $Mean(z_\lambda)$ is the mean of validity degrees of the class z_λ and α_j is the reliability coefficient of the source S_j. This method of construction means that the more the source S_j believes in a disjunction z_λ, the less it will allocate a belief mass to its opposite $\overline{z_\lambda}$.

When we have the Λ created *bbas* $m_{j,\lambda}$, we can combine them by means of the Dempster combination. The *bba* m_j from the source S_j is obtained by:

$$m_j = \textcircled{\cap}_{\lambda=1}^{\Lambda} m_{j,\lambda} \tag{15}$$

As in the method presented in [5], it is a careful way to generate mass functions. We are going to study its results in the following sections.

4 Applications

4.1 Confrontation of both methods

In this section, we confront mass functions given by both methods. Figure 1 presents the validity degree distribution given by a source S_j for a frame of discernment $\Omega = \{\omega_0, ..., \omega_{150}\}$. As described in the section 3.2, in order to use Jenks' method, it is necessary to know *a priori* the number of possible disjunctions. Within the framework of this pratical case, we subjectively consider that an expert would have classed hypotheses in three classes. So, the belief function creating process is detailed for Jenks' method in tables 1 and 2. Figure 2 allows to compare the results obtained by both methods. We can see that with Jenks' method (with only three classes), the two principal peaks are eliminated. These peaks are always present with the AIC^*-criterion based method.

4.2 Mass function construction in the forensic entomology project

During a criminal investigation, it is essential to obtain a maximum of information on the conditions of the manslaughter. Many methods to exploit the indications on the murder scene are known but, for large post-mortem intervals (PMI), only one of these techniques is useful in practice: forensic entomology. It consists in studying the insects found on a cadaver to estimate the time of its death. The objective is then to date the first layings by calculating the insects' age. Modern PMI estimation methods are based on insect development models. These models consider that insect development speed is temperature-dependent and

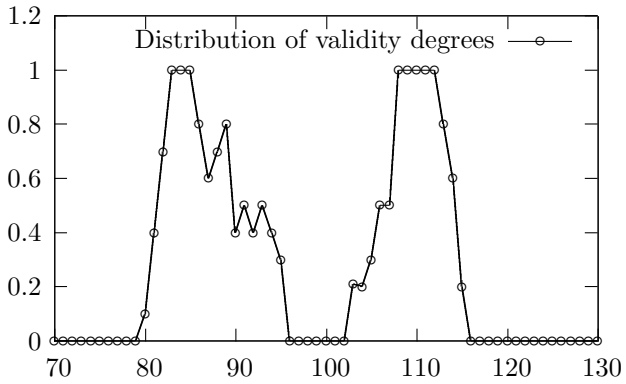

Figure 1: Validity degree distribution given by a source S_j. Between the hypotheses ω_0 and ω_{70} and between the hypotheses ω_{130} and ω_{150}, the validity degrees are null.

Table 1: The three mass functions obtained with the grouping of hypotheses based on Jenks' method from data presented in figure 1.

	$m_{j,1}$
$m_{j,1}(\{\overline{\omega_0, \ldots, \omega_{80}}\})$	$0,997$
$m_{j,1}(\Omega)$	$0,003$

	$m_{j,2}$
$m_{j,2}(\{\overline{\omega_{81}, \ldots, \omega_{114}}\})$	$0,013$
$m_{j,2}(\Omega)$	$0,987$

	$m_{j,3}$
$m_{j,3}(\{\overline{\omega_{115}, \ldots, \omega_{150}}\})$	$0,990$
$m_{j,3}(\Omega)$	$0,010$

has the next general form:

$$\Delta a = \int_{t_1}^{t_2} f(T(t))dt \tag{16}$$

where $T(t)$ represents the temperature T felt by an insect at the time t, f is a development model and Δa is the accumulated rate of development of the insect. The interested reader can refer to [25] for more details.

There are a lot of models in literature but they do not take into account the entire ecosystemic context (thermic inertia of the body, thermic behavior of the larvae *etc.*). No current method takes into account these parameters, reported in numerous articles as an important source of error in the estimating of the PMI [26]. Moreover, these models are based on different original biological data that were not still verified in all the contexts, but are used by the experts because of the lack of better information. Thus, estimations performed using these methods are often overestimated and not as precise as they could be. To improve the decision-making

Table 2: The final mass function m_j obtained from the combination (eq. 15) of the mass functions $m_{j,1}$, $m_{j,2}$ and $m_{j,3}$.

	m_j
$m_j(\{\overline{\omega_0, \ldots, \omega_{80}}\})$	0.010
$m_j(\{\overline{\omega_0, \ldots, \omega_{114}}\})$	1.4×10^{-4}
$m_j(\{\overline{\omega_{81}, \ldots, \omega_{114}}\})$	3.2×10^{-7}
$m_j(\{\overline{\omega_{81}, \ldots, \omega_{150}}\})$	3.0×10^{-5}
$m_j(\{\overline{\omega_{115}, \ldots, \omega_{150}}\})$	0.002
$m_j(\{\omega_{81}, \ldots, \omega_{114}\})$	0.986
$m_j(\Omega)$	2.5×10^{-5}

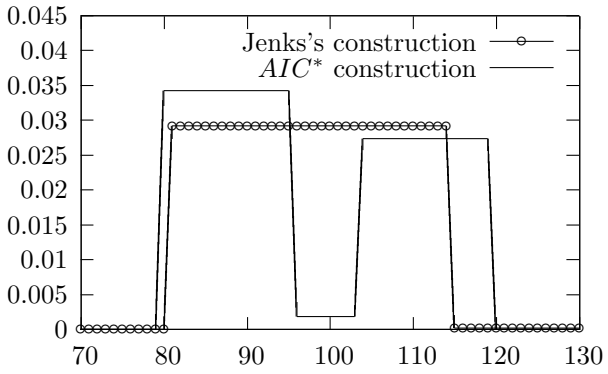

Figure 2: *Pignistic* probability distributions for both methods. The Jenks' curve is obtained from the *pignistic* transformation of the mass function m_j detailed in table 2.

and assist the forensic scientists, a decision support system has been developed. This project is based on a predictive multiagent model of insect development and cadaver decomposition in a complex ecosystem. It is used to determine if a hypothesis - a possible time of death - is coherent with the observations available on the ecosystem of the crime scene and the entomofauna found on the victim. More information about this model and the validation process can be found in [27].

To sum up the notations of the section 3.1, there are several sources (entomological models) $S_j \in S$ supplying distributions of validity degrees $\Delta^Q_{j,\Omega} : \Omega \mapsto [0, 1]$, Ω is the set of the possible hours. Each distribution indicates how the model thinks that the hypotheses of the time interval are in accordance with the reality Q observed by the forensic scientists.

In the figure 3, we present the distributions of validity degrees supplied by four entomological models on an interval of 150 hours. For Jenks' based construction, the entomological expert recommends to use 3 disjunctions/classes. This value corresponds to a "normal" expected shape of a distribution supplied by the models. For the method based on information criteria, we used

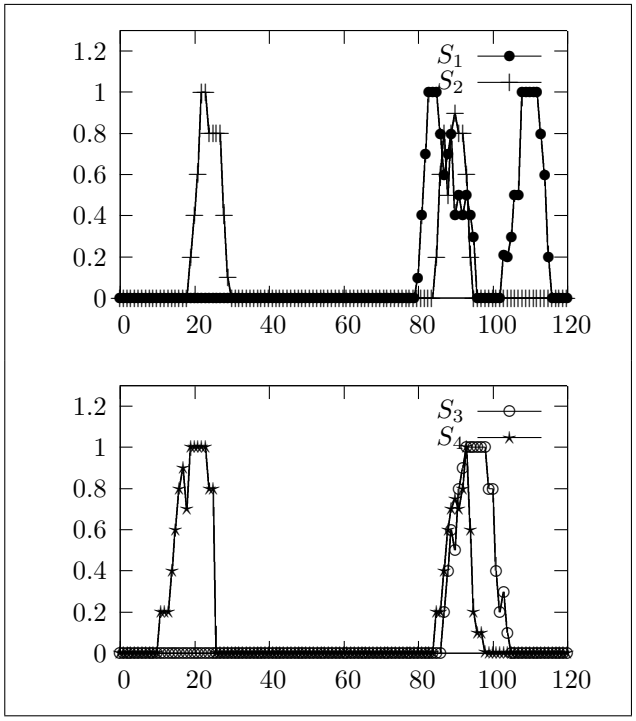

Figure 3: Validity degree distributions given by four entomological models. To make it simpler, we do not print hypotheses between ω_{120} and ω_{150} where validity degrees are null.

the AIC^* criterion allowing to obtain a good number of disjunctions.

We notice in figures 4 and 5 that the method based on the AIC^*-criterion retranscribes in a better way the reality of the information than Jenks' method. An important loss of information is indeed observed during the use of this last method: mass functions of S_2 and S_4 contain no more information of the second observable peak in figure 3 around the hypothesis ω_{90}. This loss of information related to Jenks' method can be annoying during the final decision-making but could also be handled by means of the value of the Jenks index (see section 3.2). For the models S_2 and S_4, the Jenks' calculated index is very low and a treatment *a posteriori* could be made for, either to discount the mass functions proportionally in the index value, or to make a treatment by considering a bigger number of classes.

However, fixing the number of classes to a weaker value for Jenks' method allows us to reduce the calculation time during the mass function combinations even if we lose in precision, as we can see in figure 6.

5 Discussion

Methods of mass function construction allowing to affect beliefs on compound hypotheses try to answer problems linked to methods of estimation based on the calculation of distances and those based on the calcula-

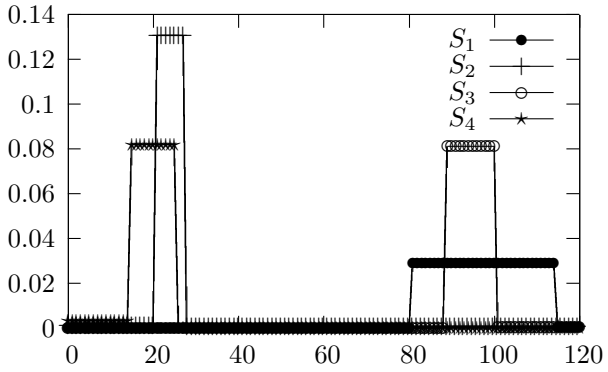

Figure 4: The four *pignistic* probability distributions obtained with Jenks' method.

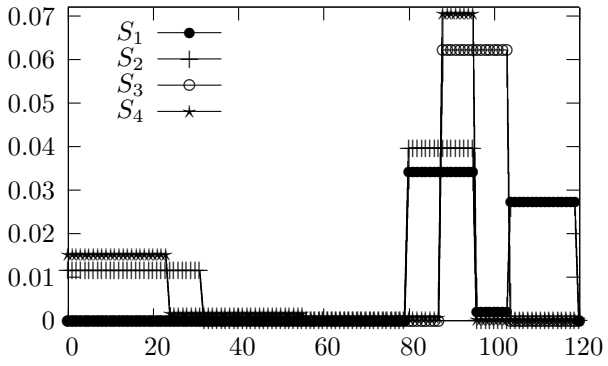

Figure 5: The four *pignistic* probability distributions obtained with the AIC^*-criterion based method.

tion of likelihoods. They allow a compromise between "being less engaged" and "being calculable". Both methods can be used in two different contexts: Jenks' method can be used when the modeller knows *a priori* the shape of the distributions $\Delta^Q_{j,\Omega}$. It is also possible to use the value of the GVF index to verify that the number of classes was correct and to modify it if necessary to generate other *bbas* with more disjunctions. The method based on information criteria can lead to a higher number of classes and thus obtain a *bba* closer to raw information, while being less involved with methods based on the calculation of distances. It seems interesting, but can lead to too important a number of disjunctions and too expensive a computation time.

6 Conclusion

When the number of hypotheses of the frame of discernment is too large, it is generally difficult to manage belief functions because of the number of potential focal elements. In this article, we propose two automatic methods of grouping hypotheses to reason on

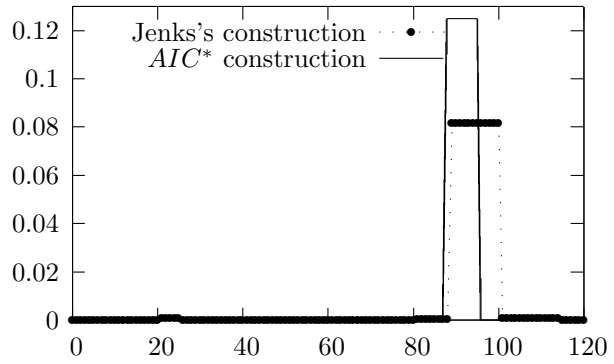

Figure 6: Final *pignistic* probability distributions obtained from the combination of the four models for both methods.

compound hypotheses. Jenks' method is particularly useful if the modeller has the *a priori* knowledge of the information provided by the sources to combine. The second one, based on information criteria, is interesting to group the closest hypotheses in a totally automatic way. Both methods allow to reduce the calculation time and to improve the quality of our decisions in comparison with the other existing construction methods. Afterwards, we shall work on the complexity of the presented methods and possibly analyse previous works from possibility distribution construction.

The results obtained within the framework of the forensic medicine application are promising and allow us to obtain less committed results while being explicit enough to choose one or several hours of death. In the future, we shall work on an iterative and interactive method allowing to approach gradually hypotheses supported by the majority of the models.

Acknowledgments.

This work is financed by the Norbert Ségard foundation. The authors thank Daniel Jolly, Gildas Morvan and Gilles Goncalves (LGI2A), Philippe Kubiak (LAGIS-École Centrale of Lille), Benoît Bourel and Damien Charabidze (Institute of forensic medicine of Lille) for their support.

References

[1] Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press, Princeton, New Jersey (1976)

[2] Smets, P., Kennes, R.: The transferable belief model. Artificial Intelligence **66**(2) (1994) 191–234

[3] Schubert, J.: On nonspecific evidence. International Journal of Intelligent Systems **8** (1993) 711

[4] Denœux, T.: A k-nearest neighbour classification rule based on dempster-shafer theory. IEEE Transactions on Systems, Man and Cybernetics **25**(5) (1995) 804–813

[5] Appriou, A.: Multisensor signal processing in the framework of the theory of evidence. In: Application of Mathematical Signal Processing Techniques to Mission Systems, Research and Technology Organization (Lecture Series 216) (November 1999) (5–1)(5–31)

[6] Smets, P.: Belief functions: The disjonctive rule of combination and the generalized bayesian theorem. International Journal of Approximate Reasoning **9** (1993) 1–35

[7] Bloch, I.: Some aspects of dempster-shafer evidence theory for classification of multi-modality medical images taking partial volume effect into account. Pattern recognition Letters **17** (1996) 905–919

[8] Milisavljevic, N., Bloch, I.: A two-level approach for modeling and fusion of humanitarian mine detection sensors within the belief function framework. In: Applied Stochastic Models and Data Analysis (ASMDA'2001). (June 2001) 743–748

[9] Bloch, I.: Defining belief functions using mathematical morphology application to image fusion under imprecision. International Journal of Intelligent Systems **48** (2007) 437–465

[10] Bloch, I., Maitre, H.: Fusion en traitement d'images : spécificités et approches numériques. Revue Techniques de l'Ingénieur **TEB1 - TE5230** (2002) 1–23

[11] Slocum, T.A., McMaster, R.B., Kessler, F.C., Howard, H.H.: Thematic Cartography and Geographic Visualization (2nd Edition). Prentice Hall Series in Geographic Information Science (2005)

[12] Colot, O.: Apprentissage et détection automatique de changements de modèles - Application aux signaux électro-encéphalographiques. PhD thesis, Université de Rouen (1993)

[13] Dempster, A.: Upper and lower probabilities induced by multivalued mapping. Annals of Mathematical Statistics **AMS-38** (1967) 325–339

[14] Guo, H., Shi, W., Deng, Y.: Evaluating sensor reliability in classification problems based on Evidence Theory. IEEE Transactions on Sytems, Man and Cybernetics - Part B **36**(5) (2006) 970–981

[15] Mercier, D., Quost, B., Denoeux, T.: Refined modeling of sensor reliability in the belief function framework using contextuel discounting. Information Fusion **9**(2) (2008) 246–258

[16] Smets, P.: Analyzing the combination of conflicting belief functions. Information Fusion **8**(4) (2007) 387–412

[17] Dubois, D., Prade, H.: A set-theoric view of belief functions: Logical operations and approximations by fuzzy sets. International Journal of General Systems **12** (1986) 193–226

[18] Florea, M.C., Jousselme, A.L., Bossé, ı., Grenier, D.: Robust combination rules for evidence theory. Information Fusion **10**(2) (2009) 183–197

[19] Lefevre, E., Colot, O., Vannoorenberghe, P.: Belief function combination and conflict management. Information Fusion **3**(2) (2002) 149–162

[20] Smets, P.: Decision making in a context where uncertainty is represented by the belief functions. Belief functions in Business Decisions (2002) 17–61

[21] Jenks, G.: Optimal data classification for choropleth maps. Occasional paper No. 2, department of geography, University of Kansas (1977)

[22] Fisher, W.D.: On grouping for maximum homogeneity. J. Amer. Statist. Assoc. **53** (1958) 789–798

[23] Colot, O., Olivier, C., Courtellemont, P., El-Matouat, A., de Brucq, D.: Information criteria and abrupt changes in probability laws. In Holt, M., Cowan, C., Grant, P., Sandham, W., eds.: Signal Processing VII: Theories and Applications, EUSIPCO'94 (September 1994) 1855–1858

[24] El-Matouat, A., Olivier, C.: Sélection du nombre de classe d'un histogramme et contraste de kullback. In: 25ème Journées Internationales de Statistiques, ASU (1993) 193–196

[25] Stinner, R.E., Gutierrez, A.P., Butler Jr, G.D.: An algorithm for temperature-dependant growth rate simulation. The Canadian Entomologist **106** (1974) 519–524

[26] Marchenko, M.I.: Medicolegal relevance of cadaver entomofauna for the determination of the time of death. Forensic Science International **120** (2001) 89–120

[27] Morvan, G., Jolly, D., Dupont, D., Kubiak, P.: A decision support system for forensic entomology. In: Proceedings of the 6^{th} EUROSIM congress. (2007)