

# **Estimation de fonctions de masse de croyance, utilisation et confrontation dans le cadre d'expertises en Entomologie Médico-Légale**

## **Estimation of belief functions, use and confrontation in the context of forensic entomology expertises**

Alexandre Veremme<sup>1,2</sup>

Eric Lefèvre<sup>2</sup>

Daniel Jolly<sup>2</sup>

<sup>1</sup> ERASM - HEI

<sup>2</sup> LGI2A EA 3926

ERASM HEI - 13 rue de Toul 59046 Lille - France, {prenom.nom}@hei.fr

LGI2A - Université d'Artois - FSA - Technoparc Futura 62400 Béthune - France, {prenom.nom}@univ-artois.fr

### **Résumé :**

Dans le cadre de la théorie des fonctions de croyance, l'estimation des fonctions de masse demeure un problème non négligeable et pour lequel l'universalité n'existe pas. Pour la réalisation d'une application d'aide à la décision dédiée à l'entomologie médico-légale, nous sommes confrontés à cette problématique. Dans cet article, deux méthodes d'estimation sont présentées et confrontées. Nous portons notre attention sur une méthode d'estimation pour laquelle nous utilisons des facteurs de fiabilité estimés à partir de l'entropie de Shannon.

### **Mots-clés :**

théorie des fonctions de croyance, estimation de fonction de masse, facteur de fiabilité, SIAD, entomologie médico-légale

### **Abstract:**

Under the theory of belief functions, estimation of mass functions is a significant problem where the universality does not exist. For the achievement of a decision support system dedicated to forensic entomology, we are facing this problem. We propose in this article two methods to estimate belief masses and compare them. We are interested in a method of estimation for which we useifiability factors estimated thanks to Shannon entropy.

### **Keywords:**

belief function theory, belief function estimation, fiability factor, decision support system, forensic entomology

## **1 Introduction**

Représenter et gérer des connaissances au sein de systèmes complexes, où la quantité et la disparité des informations sont importantes, obligent les concepteurs à mettre en place un

formalisme particulier de traitement de l'information. L'objectif est alors d'appréhender au mieux les imperfections des connaissances (imprécision, incertitude, incomplétude etc.) afin de les manipuler et aboutir à une décision finale. Parmi les théories existantes, celle des fonctions de croyance (ou théorie de Dempster-Shafer) [1] se révèle de bon augure pour prendre en compte ces différentes caractéristiques.

Le *Modèle de Croyances Transférables* (MCT) de Philippe Smets [2, 3], vu comme une interprétation subjective et non probabiliste de la théorie des fonctions de croyance, est un instrument suffisamment souple et performant pour être appliqué dans de nombreux domaines [4] comme l'aide à la décision, le diagnostic etc. Ce modèle est d'autant plus intéressant qu'il dispose d'une palette développée d'outils de manipulations des fonctions de croyance (affaiblissement, conditionnement etc.). Mais, malgré l'importance des fondements du MCT, un problème non négligeable demeure : l'estimation des fonctions de masse de croyance. Cette estimation est en effet un problème difficile où l'universalité n'existe pas [5] et généralement dépendant de l'application et de la vision qu'on en a.

Dans le cadre de nos travaux, consistant en la réalisation d'un Système Informatique d'Aide à la Décision (SIAD) destiné à la médecine légale, nous sommes confrontés à ce problème d'estimation. Après un préambule sur la médecine légale (section 2), nous présentons dans la suite de l'article l'application d'aide à la décision, nommée ForenSeek [6, 7] (section 3). Nous revenons sur les bases du MCT (section 4) et détaillons ensuite les problématiques liées à la fusion d'informations ainsi que celles associées à la modélisation des fonctions de croyance dans le cadre de l'application. Nous confrontons nos méthodes d'estimation étudiées (section 5) et présentons les premiers résultats (section 6). Nous terminons sur les conclusions et perspectives (section 7).

## 2 L'Entomologie Médico-Légale et ses problématiques

Lors d'une enquête criminelle, il est indispensable d'obtenir un maximum d'informations sur les conditions de l'homicide. De nombreuses méthodes pour exploiter les indices présents sur la scène de crime sont connues des polices judiciaires mais, pour des intervalles *post-mortem* (IPM) supérieurs à 48 heures, une seule est utilisable en pratique : l'Entomologie Médico-Légale. Elle consiste à étudier l'entomofaune nécrophage retrouvée sur le corps afin d'estimer la date de la mort d'une victime. L'objectif est alors de dater les premières pontes en calculant l'âge des insectes prélevés à la découverte du corps.

Les méthodes modernes s'appuient principalement sur l'étude du développement des *diptères* (mouches) nécrophages, première espèce à coloniser un cadavre. Sachant que le développement d'un *diptère* est principalement dépendant de la température, il est aisément prédictible, dans un environnement contrôlé, le temps nécessaire à son développement et de dater le moment  $t_1$  de sa ponte à partir du taux de développement accumulé  $\Delta a$  au moment  $t_2$  de la découverte du corps. La date de la mort

est directement liée à cette date de ponte.

$$\Delta a = \int_{t_1}^{t_2} f(T(t))dt \quad (1)$$

Où  $T(t)$  représente la température  $T$  ressentie par un individu en fonction du temps  $t$ , et  $f$  un modèle de développement. De nombreux modèles de développement existent, le lecteur intéressé pourra se référer à [8] pour plus de détails.

Dans un écosystème complexe, l'estimation de ce taux de développement est loin d'être aussi évidente et de nombreux paramètres (pluviométrie, température durant le mois précédent, prédation *etc.*), plus difficiles à appréhender, influent sur la présence des *diptères*. De même, faute de données, la température considérée correspond généralement à la température relevée à la station météorologique la plus proche. Or, cette température peut différer de plusieurs degrés par rapport à la température du corps (ayant lui-même une inertie thermique), réellement ressentie par les insectes.

De plus, les larves ont un comportement grégiaire qui crée des augmentations locales de température. Aucune méthode ne permet encore de prendre en compte ces augmentations, signalées dans de nombreux articles comme une source importante d'erreur dans l'estimation de l'IPM [9].

Enfin, diverses techniques existent pour estimer l'âge des insectes, ayant chacune leurs avantages et inconvénients, et utilisant des données de référence très variables. Et, l'expert, par choix personnel, par habitude (peut-être lié aux difficultés d'utilisation), et surtout par manque de temps, se restreint généralement dans le nombre de techniques utilisées.

C'est à partir de ces différents constats et pour réaliser des expertises plus cohérentes et objectives, qu'est né le projet ForenSeek.

### 3 Le projet ForenSeek

Le SIAD développé dans le cadre du projet ForenSeek peut être décomposé en trois composants logiciels indépendants : (1) un simulateur multi-agents permettant de tester des heures du décès d'une victime, (2) un système de raisonnement analysant les résultats issus du simulateur, (3) un système de fusion agrégeant les résultats issus de différents systèmes de raisonnement. Dans cette section, nous détaillons brièvement le fonctionnement et l'inter-relation de ces trois composants. Le lecteur désireux de connaître plus en détail la structure du système pourra se référer à [6].

#### 3.1 Le modèle multi-agents de décomposition

Pour prendre en compte l'ensemble des paramètres écosystémiques intervenant dans le développement des *diptères* et considérer, entre autres, les augmentations de température émergeant des interactions entre larves, il a été naturel de se tourner vers la modélisation multi-agents. Un Système Multi-Agents (SMA) est en effet un très bon modèle pour reproduire des systèmes complexes, comme celui que peut représenter un corps où plusieurs milliers d'insectes interagissent.

Chaque insecte, *acteur* du système, est alors modélisé sous la forme d'un agent, *i.e.* un système informatique autonome, situé dans un environnement et capable de communiquer, directement ou à travers l'environnement, avec d'autres agents.

Il est ainsi possible de simuler le comportement thermique et nutritionnel d'une masse de larves et son influence sur le développement de ses individus et, plus globalement, de modéliser la décomposition du cadavre. Notons que ce modèle est prédictif : une simulation démarre à la date présumée de la mort et s'arrête au moment de la découverte du corps par les enquêteurs.

De plus, lors d'une simulation, l'ensemble des

paramètres sont initialisés à partir d'un modèle entomologique  $Mod_j$  issu de la littérature. Une simulation et ses résultats diffèrent en fonction de ces données initiales et il est alors possible de simuler un même cas autant de fois différentes qu'il existe de modèles.

#### 3.2 Le système de raisonnement et d'évaluation des hypothèses

Le système de raisonnement constitue le deuxième élément logiciel de ForenSeek. Le SMA étant prédictif et le raisonnement rétrodictionnaire (nécessité de revenir sur les origines du décès), ce deuxième niveau implémente une méthode abductive de raisonnement. L'abduction consiste à déterminer la cause la plus probable d'une observation [10]. Concrètement, cela consiste en une méthode itérative de choix et de validation d'hypothèses (ou heures possibles de la mort). Dans le cadre de l'application, il s'agit de parcourir « intelligemment » l'espace des hypothèses borné par les dates de disparition et de découverte de la victime.

En parallèle à ce système de raisonnement, nous avons mis en place un système à base de règles permettant de déterminer efficacement si une hypothèse est compatible avec les observations entomologiques faites par l'expert lors de la découverte du cadavre. Il est alors possible d'affecter des mesures de cohérence pour les hypothèses de l'intervalle, indiquant leur niveau de compatibilité respectif avec les observations entomologiques. Le modèle étant stochastique, il est nécessaire de réaliser un nombre de simulations statistiquement représentatif afin d'évaluer correctement une hypothèse.

La mesure de cohérence  $c_{j,k}$  associée à une hypothèse  $\omega_k$  et à un modèle  $Mod_j$  est définie de la manière suivante :

$$c_{j,k} = \frac{s_{j,k}^c}{s_{j,k}^t} \quad (2)$$

Où  $s_{j,k}^c$  et  $s_{j,k}^t$  représentent respectivement le nombre de simulations compatibles avec les données observées et le nombre total de simu-

lations réalisées en utilisant le modèle  $Mod_j$  et l'hypothèse  $\omega_k$ .

### 3.3 Le système de fusion

Afin d'améliorer l'estimation de l'IPM, plusieurs modèles  $Mod_j$  sont considérés, chacun fournissant plusieurs distributions de cohérence pour chaque hypothèse testée de l'intervalle *post-mortem*. Une étape de fusion est réalisée dans le cadre du *Modèle des Croyances Transférables* [2, 3].

## 4 Fusion de données dans le cadre du MCT

Dans cette section, les bases du *Modèle de Croyances Transférables*, l'interprétation adoptée dans nos travaux, sont succinctement rappelées. Le MCT peut être partagé en deux niveaux : le niveau crédal (permettant la représentation et la manipulation des croyances) et le niveau pignistique (utilisé pour la prise de décision).

### 4.1 Représentation et manipulation des informations

Soit  $\Omega = \{\omega_1, \dots, \omega_k, \dots, \omega_K\}$ , un ensemble fini, appelé cadre de discernement, regroupant l'ensemble des heures possibles de la mort. Une fonction de croyance *bel* est une fonction définie de  $2^\Omega$  vers  $[0, 1]$  :

$$bel(A) = \sum_{\emptyset \neq B \subseteq A} m(B) \quad \forall A \subseteq \Omega \quad (3)$$

Où  $m$ , appelé fonction de masse, est une fonction de  $2^\Omega$  dans  $[0, 1]$  vérifiant la condition :  $\sum_{A \subseteq \Omega} m(A) = 1$ . Un sous-ensemble  $A \subseteq \Omega$  tel que  $m(A) > 0$  est appelé élément focal de  $m$ . La masse  $m(A)$  représente la part de croyance attribuée à la proposition  $A$  et à aucun autre sous-ensemble plus spécifique, compte tenu de l'état de la connaissance.

Remarquons que dans le MCT, une fonction de masse  $m$  peut être non normalisée : la quantité  $m(\emptyset)$  peut être strictement positive.

Les fonctions de croyance permettent ainsi de modéliser des informations incertaines, de l'ignorance totale ( $m(\Omega) = 1$ ) à la connaissance complète ( $m(\{\omega_k\}) = 1$ ).

En présence d'une source d'information jugée non totalement fiable, il est possible d'intégrer un facteur de fiabilité  $\beta$  permettant de transférer les masses vers  $m(\Omega)$ . Cette opération d'affaiblissement [1] se traduit par :

$$\begin{aligned} m_\beta(A) &= \beta \cdot m(A) \quad \forall A \subseteq \Omega, A \neq \Omega \\ m_\beta(\Omega) &= 1 - \beta + \beta \cdot m(\Omega) \end{aligned} \quad (4)$$

Une fois représentées, les connaissances peuvent alors être révisées. Pour cela, la règle de combinaison conjonctive permet de fusionner deux fonctions de masse,  $m_1$  et  $m_2$ , provenant de sources distinctes et fiables. Cette règle est définie de la manière suivante :

$$m_1 \odot m_2(A) = \sum_{B \cap C = A} m_1(B)m_2(C), \quad \forall A \subseteq \Omega \quad (5)$$

Elle est encore appelée *règle de Dempster non normalisée*. Le réel  $m(\emptyset)$  représente le degré de conflit entre les deux sources et est obtenu par :  $m_1 \odot m_2(\emptyset)$ .

D'autres opérateurs de combinaisons sont disponibles. Le lecteur désireux d'en savoir plus sur leurs différentes propriétés respectives pourra se référer à [11].

### 4.2 Prise de décision

Le niveau précédent permet d'obtenir un résumé exhaustif de l'information sous forme d'une fonction de masse  $m$ . Pour prendre la décision, Smets et Kennes [2] proposent de transformer la fonction de masse  $m$  en une fonction de probabilité *BetP* exprimée sur  $\Omega$ . Cette transformation, nommée *transformation pignistique*, se caractérise pour tout  $\omega_k \in \Omega$  par :

$$BetP(\omega_k) = \frac{1}{1 - m(\emptyset)} \sum_{A \ni \omega_k} \frac{m(A)}{|A|} \quad (6)$$

Où  $|A|$  représente la cardinalité de  $A \subseteq \Omega$ . Lors de cette transformation, la masse de croyance

$m(A)$  est uniformément distribuée parmi les éléments de  $A$ . La *probabilité pignistique* obtenue, il est possible d'utiliser les outils classiques de la théorie de la décision statistique.

Si on considère le problème du choix d'une action parmi un ensemble fini  $\mathcal{A} = \{\delta_1, \dots, \delta_K\}$ , l'action  $\delta_k$  correspond à l'action de choisir l'hypothèse  $\omega_k \in \Omega$ . La mise en œuvre d'une action  $\delta_i$  alors que la réalité est l'hypothèse  $\omega_k$  est supposée entraîner un coût noté  $\lambda(\delta_i|\omega_k)$ . Le risque conditionnel de décider  $\delta_i$  s'exprime alors de la façon suivante :

$$R_{Bet}(\delta_i) = \sum_{\omega_k \in \Omega} \lambda(\delta_i|\omega_k) BetP(\omega_k). \quad (7)$$

Il suffit ensuite de choisir l'action qui minimise ce risque. Dans le cas de coûts  $\{0, 1\}$ , cela revient à choisir l'hypothèse pour laquelle la probabilité pignistique est maximale.

## 5 Estimations des fonctions de masse

Comme énoncé en introduction, le *Modèle de Croyances Transférables* constitue un cadre intéressant, par sa richesse et sa flexibilité, pour la représentation et la manipulation d'informations incomplètes. Cependant, le problème de l'estimation des fonctions de masse demeure et le choix de la modélisation pesera sur les résultats et décisions finaux. Nous présentons et confrontons ici deux méthodes de modélisation de fonctions de croyance.

### 5.1 Méthode 1

Nous revenons dans cette sous-section sur une méthode de modélisation présentée précédemment dans [12], auquel le lecteur pourra se référer pour plus d'explications.

Pour un modèle  $Mod_j$ , nous créons autant de fonctions de masse de croyance qu'il y a d'hypothèses  $\omega_k$ .  $K$  fonctions de masse  $m_{j,k}$  sont ainsi obtenues :

$$\begin{cases} m_{j,k}(\{\omega_k\}) &= \beta_{j,k} \cdot c_{j,k} \\ m_{j,k}(\Omega) &= 1 - \beta_{j,k} \cdot c_{j,k} \end{cases} \quad (8)$$

où  $c_{j,k}$  correspond à la mesure de cohérence de l'hypothèse  $\omega_k$  du modèle  $Mod_j$  (équation (2)) et  $\beta_{j,k}$  à un coefficient de fiabilité associé au modèle pour l'hypothèse  $\omega_k$ . Ce coefficient peut être fixé ou estimé en fonction, par exemple, des résultats obtenus par le modèle  $Mod_j$  dans un certain contexte [13]. Afin de déterminer la fonction de masse de chaque modèle, on peut utiliser l'équation (5) de combinaison conjonctive non-normalisée afin de combiner les  $K$  fonctions de masse  $m_{j,k}$ . La masse résultante du modèle  $Mod_j$  s'écrit alors :

$$m_j = \bigcirc_{k=1}^K m_{j,k} \quad (9)$$

Enfin, de manière identique, la fonction de masse représentant l'ensemble des connaissances est obtenue de la combinaison des  $J$  modèles :

$$m = \bigcirc_{j=1}^J m_j \quad (10)$$

A partir de cette fonction de masse finale, le calcul de la probabilité pignistique se fait à l'aide de l'équation (6). L'hypothèse retenue est celle minimisant le risque pignistique (équation (7)).

### 5.2 Méthode 2

La deuxième méthode, à l'inverse de la précédente, vise à considérer une distribution de mesure de cohérence de manière plus globale. Ainsi, nous ne créons qu'une fonction de masse de croyance par modèle  $Mod_j$ . L'affectation des masses est obtenue de la façon suivante par normalisation de la mesure de cohérence :

$$\begin{cases} m_j(\{\omega_k\}) &= \beta_j \cdot \frac{c_{j,k}}{c_j} \quad \forall \omega_k \in \Omega \\ m_j(\Omega) &= 1 - (\beta_j \cdot \sum_{k=1}^K \frac{c_{j,k}}{c_j}) \end{cases} \quad (11)$$

Où  $c_{j,k}$  est la mesure de cohérence de l'hypothèse  $\omega_k$  du modèle  $Mod_j$  (équation (2)) et où  $c_j$  est obtenu de la manière suivante :

$$c_j = \sum_{k=1}^K c_{j,k} \quad (12)$$

Dans l'équation (11), le réel  $\beta_j$  correspond au coefficient de confiance associé au modèle  $Mod_j$ . L'équation (10) est utilisée pour obtenir la fonction de masse finale.

### 5.3 Problématiques et détermination du facteur de fiabilité : entropie de Shannon

La problématique essentielle des méthodes présentées ci-dessus reste la détermination du facteur de fiabilité  $\beta_j$ . Ce facteur ne peut être omis puisque, sans lui, un conflit total ( $m(\emptyset) = 1$ ) est susceptible d'apparaître lors de la combinaison conjonctive. Dans le cadre de l'entomologie, nous ne souhaitons pas que ce conflit total soit un obstacle à la prise de décision (choix des heures les plus probables). Avoir deux sources en conflit total signifiera qu'elles interprètent différemment la même réalité.

Il est possible de fixer la valeur du facteur ou d'en obtenir une approximation à l'aide des diverses méthodes d'estimation existantes [13, 14, 15], mais ces dernières nécessitent des ensembles d'apprentissage que nous ne possédons pas dans le cadre de ForenSeek. Donner une valeur arbitraire n'aurait pas de justification.

Pour cela, nous proposons une méthode d'estimation dynamique du facteur de fiabilité relative à l'entropie de Shannon de la distribution de mesures de cohérence. De manière générale, l'entropie représente le niveau de désordre d'un système. Plus spécifiquement, l'entropie a déjà été utilisée dans le cadre des fonctions de croyance [16, 13] pour mesurer l'incertitude des fonctions de masse. L'entropie de Shannon  $H_j$  d'une distribution de mesures de cohérence est définie par :

$$H_j = - \sum_{k=1}^K c_{j,k} \ln(c_{j,k}) \quad (13)$$

Où  $c_{j,k}$  est la mesure de cohérence de l'hypothèse  $\omega_k$  pour le modèle  $Mod_j$ .

Comme présenté dans [13], il est possible d'obtenir un facteur de certitude normalisé  $C_j$  :

$$C_j = 1 - \frac{H_j}{\max_{\Omega}(H)} \quad (14)$$

Où  $\max_{\Omega}(H)$  représente le maximum de la valeur de Shannon pour un  $\Omega$  donné. Ce maxi-

mum est atteint lorsque tous les singlets sont équiprobables.

Ainsi, pour un modèle  $Mod_j$ , nous obtenons un facteur de fiabilité  $\beta_j$  égal à  $C_j$ .

La section suivante présente les premiers résultats issus de ces travaux.

## 6 Résultats préliminaires

### 6.1 Exemple 1

La figure 1 représente les distributions de mesures de cohérence fournies par deux modèles entomologiques  $Mod_1$  et  $Mod_2$  sur un intervalle de 150 heures. Le deuxième modèle apparaît comme moins engagé que le premier et d'après l'équation (13), l'entropie de  $Mod_2$  est donc plus élevée que celle du modèle  $Mod_1$ . Pour l'exemple courant, nous obtenons ainsi les facteurs de fiabilité  $\beta_1 = 0.376$  et  $\beta_2 = 0.203$ .

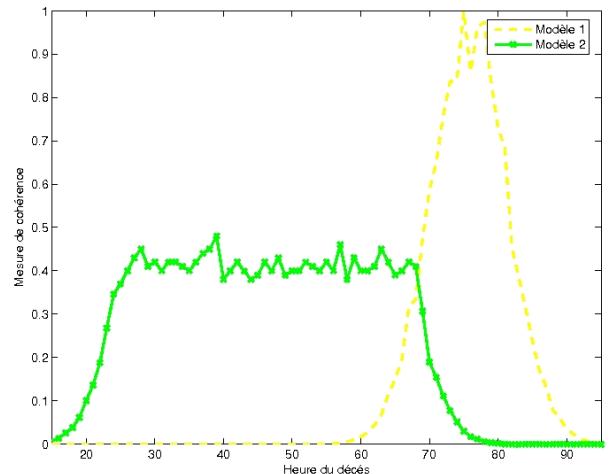


Figure 1 – Distributions de mesures de cohérence des modèles  $Mod_1$  et  $Mod_2$ .

Pour le même cas médico-légal, avec la méthode 2, nous comparons sur la figure 2 les probabilités pignistiques finales obtenues selon trois cas :

- 1. les facteurs de fiabilité sont estimés selon l'entropie de Shannon
- 2. les facteurs ne sont pas pris en compte ( $\beta_j = 1$ )

- 3. les facteurs sont arbitrairement fixés à 0.95 pour les deux modèles

!!!! NE PAS OUBLIER DE METTRE A JOUR L'IMAGE - inverser les valeurs des beta !!!!

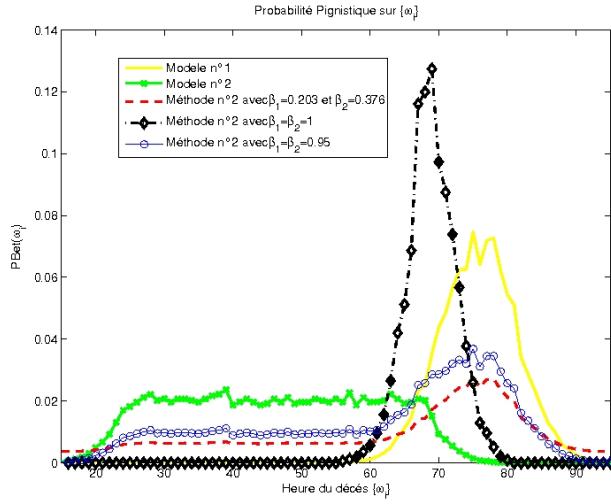


Figure 2 – Probabilités pignistiques des modèles  $Mod_1$  et  $Mod_2$  pris indépendamment et obtenues de la fusion suite à la méthode 2 selon les cas 1, 2 et 3 énoncés ci-dessus.

La non prise en compte des facteurs de fiabilité revient à une intersection des heures supportées par les deux modèles. Leur considération fait en revanche varier de manière importante le maximum de probabilité pignistique (décalage vers la droite dans l'exemple). Sans considération des  $\beta_j$  ( $\beta_j = 1$ ), nous choisissons l'hypothèse  $\omega_{69}$  alors qu'avec des facteurs fixés ou calculés selon l'entropie de Shannon, nous préférions l'hypothèse  $\omega_{75}$ , hypothèse qui semble la plus plausible au vu du faible niveau d'expressivité du modèle  $Mod_2$ .

Dans le cas de la prise en compte des facteurs de fiabilité, l'intérêt réside principalement dans le calcul dynamique de celui-ci, au contraire de la valeur arbitraire 0.95.

## 6.2 Exemple 2, confrontation des deux méthodes

La deuxième expérimentation permet de mettre en avant les résultats de la prise en compte de

l'entropie de Shannon lorsque six modèles entomologiques sont confrontés. Cette expérience considère un ensemble de 141 hypothèses, *i.e.* 141 heures entre la date de disparition de la victime et la découverte du cadavre.

Sur la figure 3 apparaissent les six fonctions de masse de croyance obtenues à partir de la méthode 2 ainsi que les facteurs de fiabilité de leur modèle respectif.

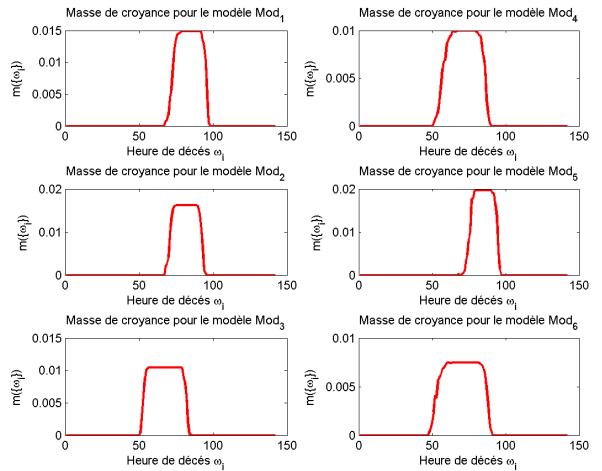


Figure 3 – Fonctions de masse de croyance obtenues à partir de la méthode 2 pour les six distributions de mesures de cohérence des différents modèles.  $\beta_1 = 0.3418$ ,  $\beta_2 = 0.3544$ ,  $\beta_3 = 0.3054$ ,  $\beta_4 = 0.2903$ ,  $\beta_5 = 0.3753$ ,  $\beta_6 = 0.2609$ .

La méthode 1 (figure 4) permet d'obtenir une probabilité pignistique très engagée ( $BetP = 0.46$ ). Ce fort engagement peut sembler intéressant mais le facteur de fiabilité de 0.95 n'a pas de justification. De plus, son estimation par l'entropie de Shannon serait ici injustifiée au vue de la construction de la fonction de masse (équation (8)), issue de la combinaison de  $K$  fonctions de masse intermédiaires.

Par contre, de manière identique à l'exemple précédent, la méthode 2, lorsque les facteurs sont calculés selon l'entropie de Shannon, permet de choisir les heures les plus supportées par les six modèles. Dans ce cas (figure 4) l'intervalle des hypothèses  $\{\omega_{75}, \dots, \omega_{86}\}$  est mis en avant.

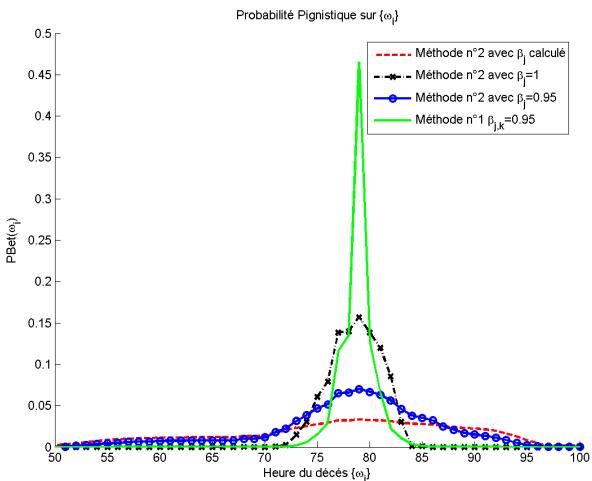


Figure 4 – Probabilités pignistiques obtenues à l’issue des deux méthodes.

## 7 Conclusion et perspectives

Nous avons confronté deux méthodes d'estimation de fonctions de masse de croyance. Dans le cas d'absence totale de connaissances sur la fiabilité de sources à combiner, nous proposons d'estimer les facteurs de fiabilité en fonction de l'entropie de Shannon des données fournies. Cette estimation permet d'obtenir une fiabilité instantanée et permet, dans le cadre de l'application d'aide à la décision ForenSeek, d'obtenir des premiers résultats intéressants. Dans l'avenir, nous aurons à prendre en considération diverses contraintes, telle que la prise en compte de l'incertitude liée à l'heuristique abductive (décrise en section 3). De plus, une entropie absolue (prise en compte de  $\max_{\Omega}(H)$  dans l'équation (13)) a été considérée, il est envisagé par la suite de tester le comportement de la méthode de Shannon utilisée de manière relative, où le maximum de la valeur de Shannon correspondrait au maximum de l'entropie des sources à combiner.

### Remerciements :

Ce travail est financé par le Ministère de la Recherche et la Fondation Norbert Ségard. Il est réalisé dans le cadre d'une collaboration entre différents laboratoires. Les auteurs tiennent ainsi à remercier David Mercier (LGI2A) pour sa participation et ses conseils. Les auteurs remercient également Daniel Dupont (Erasm - HEI Lille), Gildas Morvan (LGI2A), Philippe Kubiak (LAGIS - École

Centrale de Lille), Gilles Goncalves, Benoît Bourel et Damien Charabidze (Laboratoire d'Entomologie - Institut de Médecine Légale de Lille) pour leur soutien.

## Références

- [1] G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, 1976.
- [2] P. Smets and R. Kennes, “The transferable belief model,” *Artificial Intelligence*, vol. 66(2), pp. 191–234, 1995.
- [3] P. Smets, *The Transferable Belief Model for Quantified Belief Representation*, Kluwer Academic Publishers, 267–301.
- [4] P. Vannoorenberghe, “Un état de l’art sur les fonctions de croyance appliquées au traitement de l’information,” *Revue I3*, vol. 3, no. 2, pp. 9–45, 2003.
- [5] I. Bloch, “Fusion d’informations numériques : panorama méthodologique,” in *Journées Nationales de la Recherche en Robotique 2005*, Guidel, France, oct 2005, pp. 79–88.
- [6] Gildas Morvan, Daniel Jolly, Daniel Dupont, and Philippe Kubiak, “A decision support system for forensic entomology,” in *Proceedings of the 6<sup>th</sup> EUROSIM congress*, 2007.
- [7] G. Morvan, D. Charabidze, and A. Veremme, “<http://www.forenseek.org>,” 2007.
- [8] R. E. Stinner, A. P. Gutierrez, and G. D. Butler Jr, “An algorithm for temperature-dependant growth rate simulation,” *The Canadian Entomologist*, vol. 106, pp. 519–524, 1974.
- [9] M. I. Marchenko, “Medicolegal relevance of cadaver entomofauna for the determination of the time of death,” *Forensic Science International*, vol. 120, pp. 89–120, 2001.
- [10] C. S. Pierce, *Collected Papers of Charles Sanders Pierce*, vol. 2, Havard University Press, 1932.
- [11] Philippe Smets, “Analyzing the combination of conflicting belief functions,” *Inf. Fusion*, vol. 8, no. 4, pp. 387–412, 2007.
- [12] Gildas Morvan, Alexandre Veremme, David Mercier, and Éric Lefèvre, “Application du modèle des croyances transférables dans le cadre d’expertises en entomologie médico-légale,” in *5<sup>ème</sup> atelier sur la Fouille de Données Complexes (FDC 08), 8<sup>ème</sup> journées d’Extraction et Gestion des Connaissances (EGC 08)*, 2008, pp. 49–60.
- [13] H. Guo, W. Shi, and Y. Deng, “Evaluating sensor reliability in classification problems based on Evidence Theory,” *IEEE Transactions on Systems, Man and Cybernetics - Part B*, vol. 36, no. 5, pp. 970–981, 2006.
- [14] A. Jousselme, D. Grenier, and E. Bossé, “A new distance between two bodies of evidence,” *Information Fusion*, vol. 2, pp. 91–101, 2001.
- [15] Z. Elouedi, K. Mellouli, and P. Smets, “Assessing sensor reliability for multisensor data fusion with the transferable belief model,” *IEEE Trans. SMC B* :, vol. 34, pp. 782–787, 2004.
- [16] R. R. Yager, “Uncertainty representation using fuzzy measures,” *IEEE Transaction on Systems, Man and Cybernetics*, vol. 32, pp. 13–20, 2002.