

# Théorie des fonctions de croyance pour la fusion et l'évaluation de la pertinence des sources d'informations: application à un bioprocédé fermentaire

Use of Evidence theory for the fusion and the estimation  
of relevance of data sources :  
application to an alcoholic bioprocess

**Sébastien Régis<sup>1</sup>, Andrei Doncescu<sup>2</sup> et Jacky Desachy<sup>1</sup>**

<sup>1</sup>Groupe de Recherche en Informatique et Mathématiques Appliquées des Antilles et de la Guyane, Université des Antilles et de la Guyane  
97159 Pointe-à-Pitre Guadeloupe, France  
sregis@univ-g.fr, jdesachy@univ-ag.fr

<sup>2</sup>Laboratoire d'Analyse et d'Architecture des Systèmes, 31077 cedex 04 Toulouse, France,  
adoncesc@laas.fr

Manuscrit reçu le 2 mai 2006



Résumé et mots clés

Dans cet article, nous présentons une application de la théorie des fonctions de croyance pour la classification d'états physiologiques dans un bioprocédé. Nous nous intéressons surtout à la pertinence des sources d'informations qui sont ici des paramètres biochimiques mesurés durant le procédé. La théorie des fonctions de croyance, et plus particulièrement la notion de conflit est utilisée pour évaluer la pertinence de chaque source d'information. Une autre mesure du conflit, basée sur une distance, est utilisée comme alternative, et fournit dans certains cas, des résultats plus cohérents qu'avec le conflit défini dans la théorie de Demspter et Shafer. Les résultats concernant deux types de bioprocédés (procédé batch correspondant à une classification supervisée, et procédé fed-batch correspondant à une classification non supervisée) sont présentés.

**Théorie des fonctions de croyance, classification, pertinence, bioprocédé**

Abstract and key words

In this paper, we present an application of the evidence theory for the classification of physiological states in a bioprocess. We are particularly interested by the relevance of the data sources which are here biochemical parameters measured during the bioprocess. The evidence theory, and more particularly the notion of conflict is used to evaluate the relevance of each data source. An other measure of conflict, based on a distance, is also used, and provides in some cases, better results than the classical notion of conflict of the evidence theory. Results are presented for two kinds of bioprocesses: batch process (which corresponds to a supervised classification) and fed-batch process (which corresponds to an unsupervised classification).

Evidence theory, classification, relevance, bioprocesses

# 1. Introduction

La théorie des fonctions de croyance connaît un succès croissant aujourd'hui car elle fournit un cadre conceptuel et applicatif très intéressant pour la gestion des notions de vague, d'incertain et pour la fusion d'information dans de nombreux contextes : la classification, la reconnaissance de formes, l'intelligence artificielle ou le traitement d'images par exemple.

Par ailleurs, les bioprocédés et les technologies de l'environnement utilisent de plus en plus des outils issus de l'informatique et des mathématiques : traitement du signal, fouille de données etc. Ainsi, l'utilisation des méthodes de classification représente à la fois une alternative et un complément aux méthodes de modélisation mathématique utilisées dans les bioprocédés. Ces outils de classification fournissent un modèle comportemental utile pour la compréhension et la reconnaissance de la physiologie des micro-organismes. Plus précisément, les méthodes de classification sont appliquées sur les paramètres biochimiques mesurés en ligne pendant le bioprocédé : les paramètres biochimiques représentent donc les sources d'information pour la classification. Le regroupement des paramètres en « paquets » permet de caractériser les états physiologiques des micro-organismes par une ou plusieurs classes. Ces méthodes de classification ont fourni des résultats intéressants [58] [24] [40] et le recours à ces approches dans le domaine des bioprocédés se multiplie. Une autre étape dans le processus est d'évaluer la pertinence des sources d'information. En effet, *a priori* tous les paramètres biochimiques (*ie* toutes les sources d'informations disponibles) sont utiles et nécessaires pour une bonne caractérisation des états physiologiques. Mais dans la pratique, lorsqu'un expert en microbiologie cherche à caractériser « manuellement » les états physiologiques à partir des paramètres biochimiques, il n'utilise qu'une partie de ces paramètres et laisse de côté les autres paramètres. Nous proposons donc d'utiliser la théorie des fonctions de croyance pour :

- évaluer la pertinence des sources d'information à partir du conflit.
- fusionner l'information pour la classification.

Nous proposons également d'utiliser une mesure du conflit intuitive basée sur une distance pour éviter certains manques de cohérence générés par le conflit de la théorie des fonctions de croyance. Cet article se présente comme suit. Dans le paragraphe 2, nous présentons la problématique des bioprocédés et de la classification. Dans le paragraphe 3, nous présentons la notion de pertinence, des méthodes existantes pour traiter la notion de pertinence, et l'approche basée sur la théorie des fonctions de croyance que nous proposons. Dans le paragraphe 4 nous présentons les résultats expérimentaux avant de conclure au paragraphe 5.

# 2. Bioprocédé et classification

Les bioprocédés utilisent des nombreux outils issus de l'informatique et des mathématiques. Si l'on se réfère uniquement au niveau macroscopique qui nous intéresse<sup>1</sup> on peut citer :

- les modèles mathématiques. Il s'agit de trouver un modèle susceptible de reconstituer les phases de croissance des micro-organismes [43]. La difficulté d'utilisation de ces modèles vient de la complexité du vivant qui nécessite de multiplier les paramètres dans l'équation qui définit le modèle. Par ailleurs, les modèles varient en fonction des micro-organismes utilisés (nous avons dénombré plus de 130 modèles).
- les techniques issues de l'intelligence artificielle. Elles cherchent à modéliser de façon explicite les connaissances des experts (voir par exemple [57]). Cependant le nombre de règles d'experts peut augmenter de façon quasi exponentielle et la modélisation de ces connaissances sous forme de règles pour un système expert est un travail long et fastidieux.
- les techniques issues de la classification. On cherche à regrouper les mesures effectuées en ligne dans des classes de telle sorte que ces classes soient bien différentes les unes des autres tout en ayant pour chacune d'elles, la plus grande homogénéité possible. Ces techniques peuvent faire appel ou non aux connaissances des experts.

Nous nous intéressons surtout aux méthodes de classification. Dans la pratique, on cherche à classer de façon automatique les paramètres biochimiques mesurés durant l'expérience. Il s'agit de trouver des classes qui correspondent aux états physiologiques des micro-organismes du bioprocédé. La reconnaissance de ces états permet de contrôler et d'optimiser le bioprocédé. Les paramètres biochimiques se présentent sous la forme de séries temporelles. La classification consiste donc à segmenter les séries temporelles de telle sorte qu'une classe ou un groupe de classes consécutives correspondent à un état physiologique donné. En fait la classification peut être réalisée « manuellement » par un expert en microbiologie (voir figure 1). On cherche donc à se rapprocher le plus possible du travail réalisé manuellement par l'expert en utilisant des méthodes automatiques. Comme nous l'avons dit plus haut, les experts en microbiologie cherchent non seulement à mieux connaître les états physiologiques mais aussi à caractériser la pertinence des paramètres biochimiques. Rappelons que la pertinence des paramètres est basée sur des connaissances *subjectives* et *empiriques*. Bien sûr ces connaissances sont essentiellement fondées sur des connaissances biologiques et biophysiques, mais le nombre des connaissances mises en jeu lors d'un bioprocédé

1. Sans tenir compte du niveau microscopique (voies métaboliques des micro-organismes) qui utilise de nombreux outils des équations différentielles ou du calcul matriciel, ni du niveau moléculaire (analyse génomique des micro-organismes) qui utilise des techniques d'analyse d'images ou de bio-informatique.

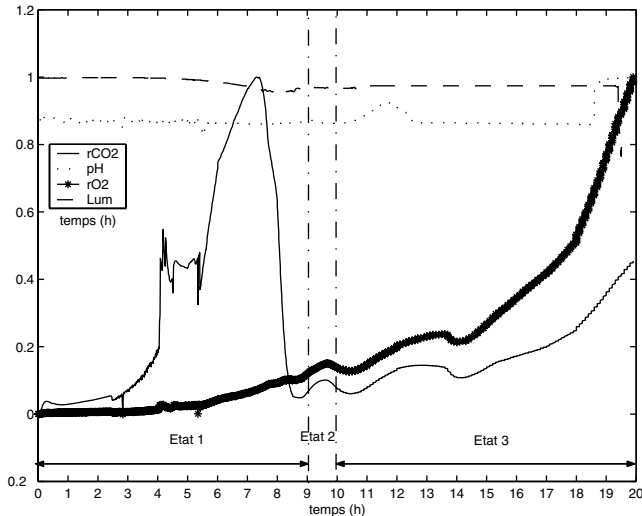


Figure 1. États physiologiques fournis par les experts sur un bioprocédé de type batch. L'axe des abscisses représente le temps exprimé en heures, l'axe des ordonnées représente l'amplitude des paramètres biochimiques (les valeurs ont été normalisées). 4 paramètres ont été utilisés : le pH, la vitesse de consommation de l'oxygène ( $rO_2$ ), la vitesse de production de dioxyde de carbone ( $rCO_2$ ) et la luminance ( $Lum$ ) qui traduit la production de biomasse. On distingue 3 états : la fermentation (état 1), la diauxie (état 2), l'oxydation (état 3).

étant très important, seule l'expérience du microbiologiste lui permet de faire le tri de celles-ci pour comprendre les phénomènes expérimentaux, évaluer la pertinence des paramètres lors de ces expériences, et éventuellement prendre une décision si cela est nécessaire. De plus les connaissances de ces experts concernent surtout 5 ou 6 paramètres alors qu'il peut y en avoir beaucoup plus (une vingtaine, voire plus). De ce fait, une partie des informations fournies par les paramètres peut être soit inexploitée, soit redondante, ou encore erronée. En fait on cherche à répondre aux questions suivantes : faut-il utiliser tous les paramètres biochimiques pour effectuer la classification ?

La réponse n'est pas évidente, d'autant plus que certains travaux (voir [38], cité dans [25]) ont montré que dans la pratique, certaines applications utilisent réellement moins de la moitié des paramètres ou variables disponibles.

À cette question, les microbiologistes seraient tentés de répondre immédiatement par l'affirmative : tous les signaux expriment, d'un certain point de vue, un phénomène biologique ou biophysique et sont donc indispensables. Cependant dans la pratique, on l'a vu, ces experts n'utilisent que quelques-uns de ces signaux.

Ainsi, en plus d'avoir une classification automatique des données qui leur fournit les états physiologiques, les microbiologistes cherchent une méthode automatique d'évaluation de la pertinence des paramètres biochimiques ayant une base *objective* ou du moins suffisamment *théorique*.

## 3. Pertinence des sources d'informations et théorie des fonctions de croyance

### 3.1. Notion de pertinence

La notion de pertinence reste une notion vague et il n'y a pas de définition générique de la pertinence. Pourtant la notion de pertinence (mais aussi et surtout celle de « non pertinence ») est très utilisée dans de nombreux domaines de l'informatique : classification [31] [10], détection de fautes [6], intelligence des moteurs de recherche [61]. Il faut noter que la notion de pertinence existe également dans des disciplines connexes à l'informatique [26] [11] : science de l'information, sciences cognitives, logique, épistémologie, etc. La définition de cette notion de pertinence varie sensiblement en fonction des disciplines et de l'application. Pour Lazo-Cortès et Ruiz-Schulcloper [31], la pertinence d'une variable dépend de sa capacité à discriminer différentes classes. Pour Zadeh [61], la pertinence globale d'un paramètre est fonction des informations fournies par les autres paramètres disponibles. Blum et Langley [10] proposent au moins 5 définitions différentes pour la pertinence. Greisdorf [26] souligne qu'en sciences cognitives, la vision de la pertinence est soit binaire (pertinent/non pertinent) soit graduelle. Il faut souligner que dans la plupart de ces disciplines, la tendance actuelle est de considérer qu'il existe différents niveaux dans la notion de pertinence (vision graduelle) [23] [11] [61]. D'un point de vue général la pertinence peut se définir comme étant le caractère de ce qui est plus ou moins approprié, qui s'inscrit dans la ligne de l'objectif poursuivi. Cette définition s'inspire de la définition encyclopédique de la pertinence [2] [1] [3].

Cette définition montre que la pertinence dépend d'un objectif. Or notre objectif ici est d'obtenir une classification qui soit correcte. Nous proposons donc la définition suivante pour la pertinence des signaux.

Une source est pertinente par rapport à la classification si :

- cette source n'induit pas de résultat aberrant.
- cette source fournit une information significative pour la classification.
- cette source génère des décisions en accord avec la plupart des autres signaux.

La première caractéristique est liée à la présence d'artefacts dans les sources. La deuxième traduit le fait que l'objectif lié à la pertinence est d'obtenir une classification correcte et montre qu'un signal pertinent a une influence sur la classification.

La dernière repose sur l'hypothèse que la majorité des signaux traduit la vérité au moins dans une certaine mesure (nous reviendrons ultérieurement sur cette hypothèse). Le fait de

considérer que la majorité des signaux traduit la vérité peut apparaître comme une hypothèse « radicale » mais il faut se rappeler que dans notre application les signaux sont issus de capteurs qui observent et traduisent le même phénomène biologique.

### 3.2. Méthodes d'évaluation de la pertinence

Il existe de plus en plus de méthodes « intelligentes » de classification tenant compte implicitement ou explicitement de la pertinence des variables disponibles. Blum et Langley [10] proposent de les classer en 4 groupes :

- les méthodes « embarquées ».

Il s'agit de méthodes de classification récursives où la recherche de la pertinence des paramètres se fait en même temps que la caractérisation des classes. Par exemple, on cherche tous les paramètres qui sont pertinents pour la caractérisation d'une classe donnée. Une fois que cette opération est terminée on passe à une nouvelle classe. La plupart de ces méthodes sont des méthodes utilisant des règles logiques [10] ou des propriétés statistiques [14].

- les méthodes « filtre ».

Il ne s'agit pas de méthodes utilisant un filtrage au sens mathématique du terme. La notion de filtrage vient du fait que l'évaluation de la pertinence des variables se fait en amont et indépendamment de la classification, comme un prétraitement de ces paramètres. Les méthodes comme l'ACP (Analyse en Composante Principale), l'ACI (Analyse en Composante Indépendante) appartiennent à cette catégorie. Dans ce groupe on trouve également les méthodes qui se basent sur les propriétés statistiques ou probabilistes des classes générées par la classification. Ces méthodes sont considérées comme des méthodes de type filtre car elles ne se basent pas sur le pourcentage d'erreur de la classification mais sur les propriétés intrinsèques des classes générées. Par exemple, l'utilisation de la théorie de l'information pour évaluer la pertinence appartient à cette catégorie.

- les méthodes « enveloppe ».

Cette classe de méthodes se différencie des deux premières dans la mesure où ce type de méthode utilise des résultats liés à la classification pour évaluer la pertinence des sources. L'approche proposée par Dubois et Prade [21] ainsi que la méthode que nous proposons appartiennent à cette catégorie. L'approche de Dubois et Prade repose sur la théorie des possibilités et sur la fusion par vote en tenant compte du degré d'accord entre les sources par rapport aux distributions de possibilités (voir aussi [22] [44]).

- les méthodes de pondération.

Ce groupe de méthodes utilise une pondération de la pertinence des différents paramètres. Cette pondération peut varier au cours d'une phase d'apprentissage. Les méthodes utilisant le perceptron appartiennent à cette catégorie.

Il faut aussi noter que dans la littérature des méthodes de sélection d'attributs (ou sélection de caractéristiques), il existe de nombreuses approches permettant d'évaluer la pertinence d'une source d'informations. On peut citer par exemple la méthode *Branch & Bound* [36] [13], les méthodes dites flottantes [39] ou les méthodes génétiques [52] [28]. Toutes ces méthodes de sélection d'attributs sélectionnent les paramètres pertinents en se basant sur un critère de minimisation (distance, information mutuelle, etc.). Si le critère de minimisation est basé sur les propriétés statistiques des classes générées, alors la méthode de sélection d'attributs appartient au groupe des méthodes « filtre ». Si le critère est basé sur le taux d'erreur de la classification, alors la méthode appartient au groupe des méthodes « enveloppe ».

On peut également citer les méthodes issues de la psychologie cognitive utilisées dans le cadre de travaux sur l'aide à la décision: les travaux et concept développés par Simon [53], Shanteau [51] [50], Montgomery [35] ou Barthélemy et Mullet [7] peuvent également être utilisés pour évaluer la pertinence. Ces concepts ont été utilisés dans des procédés industriels [45]. Enfin nous citerons des travaux issus de l'intelligence artificielle qui peuvent aussi être utilisés pour gérer la notion de pertinence: ainsi par exemple la fusion par vote proposée par Dubois et Prade [21] [22] tient compte implicitement de la pertinence des sources d'informations tandis que la méthode proposée par Felix [23] prend en compte de façon plus explicite la notion de degré de pertinence.

Les méthodes que nous présentons ne caractérisent pas toujours *explicitement* la pertinence des signaux. En effet, la pertinence d'une source n'est pas définie de façon absolue (nous avons proposé une définition possible basée sur la définition encyclopédique de la pertinence) et sa définition varie en fonction de la discipline ou du champ d'étude. D'un point de vue mathématique et informatique, il n'existe pas à proprement parler de méthode d'analyse de la pertinence d'une source. Cependant les techniques analysant les relations entre sources (dépendance ou indépendance, conflit, etc...) sont les approches les plus adaptées pour l'évaluation de la pertinence d'une source d'information. Ces techniques sont souvent utilisées pour rechercher et le cas échéant réduire les redondances d'informations ou les conflits entre diverses sources (voir par exemple [34]) mais elles servent également à évaluer l'apport réel de ces sources.

Toutes ces méthodes d'évaluation de la pertinence proposent une évaluation globale de la pertinence des sources. En effet un paramètre est considéré comme pertinent ou comme non pertinent pendant toute la durée de la classification. À notre connaissance, aucune méthode ne propose une évaluation locale qui pourrait évoluer dans le temps. La notion de pertinence dynamique est pourtant plus adaptée pour le bioprocédé, puisque la pertinence des paramètres peut varier en fonction de l'évolution de ce bioprocédé. Nous proposons donc d'utiliser la théorie des fonctions de croyance pour évaluer la pertinence des paramètres biochimiques.

### 3.3. Théorie des fonctions de croyance et pertinence

#### 3.3.1. Rappel sur les fonctions de croyance

La théorie des fonctions de croyance est une généralisation de la théorie bayésienne qui tient compte des notions d'incertitude et d'imprécision de l'information. Elle a été introduite par Dempster [15] puis a été formalisée mathématiquement par Shafer [49]. Elle a été utilisée récemment avec succès dans le domaine des bioprocédés [30] [29].

Considérons l'ensemble de toutes les évènements possibles (on parle d'ensemble de toutes les hypothèses); cet ensemble est appelé *ensemble de discernement* et est noté  $\Theta$ . Toutes ces hypothèses sont mutuellement exclusives et sont nommées *singletons*. La théorie des fonctions de croyance porte sur l'ensemble des sous-ensembles  $A$  de  $\Theta$ . Cet ensemble de sous-ensembles de  $\Theta$  est noté  $2^\Theta$ .  $A$  peut être composée d'un singleton ou d'une union de plusieurs singletons. Une fonction de masse  $m(\cdot)$  peut être alors définie de  $2^\Theta$  vers  $[0,1]$  avec les propriétés suivantes :

$$\sum_{A \subset \Theta} m(A) = 1 \quad (1)$$

$$m(\emptyset) = 0$$

$m(A)$  est la masse d'évidence associée à  $A$ .

Les fonctions de *plausibilité* ( $Pl(\cdot)$ ) et de *croyance* ( $Bel(\cdot)$ ) sont définies de  $2^\Theta$  vers  $[0,1]$  comme suit :

$$pl(A) = \sum_{B \cap A \neq \emptyset} m(B) \quad (2)$$

$$bel(A) = \sum_{B \subset A} m(B)$$

Pour obtenir une fusion de l'information de deux sources différentes 1 et 2, il existe une combinaison de leur masses d'évidence appelée règle de Dempster :

$$(m_1 \oplus m_2)(A) = m_{1,2}(A) = \frac{1}{1-K} \sum_{B \cap C = A} m_1(B).m_2(C) \quad A, B, C \subset 2^\Theta \quad (3)$$

où  $K$  est défini comme suit :

$$K = \sum_{B \cap C = \emptyset} m_1(B).m_2(C) \quad (4)$$

Le dénominateur  $1 - K$  est un facteur de normalisation. Plus précisément  $K$  représente la mesure du conflit entre les sources 1 et 2. Plus  $K$  est important, plus les sources sont en conflit et moins la fusion a de sens. Si  $K = 1$  alors le conflit est total et la fusion n'a pas de sens. On peut généraliser la règle de Dempster à  $n$  sources :

$$(\oplus m_i)_{i=1, \dots, n}(A) = \frac{1}{1-K} \sum_{X_1 \cap \dots \cap X_n = A} (\prod_{i=1}^n m_i(X_i)) \quad (5)$$

$$A, X_i \subset 2^\Theta$$

$$K = \sum_{X_1 \cap \dots \cap X_n = \emptyset} (\prod_{i=1}^n m_i(X_i))$$

Si les sources sont en conflit fort ( $K$  est grand) alors la règle de Dempster peut conduire à des résultats erronés [60]. La raison de ce comportement de la règle de Dempster provient du fait que la masse d'évidence affectée à l'ensemble vide est nulle. Cette contrainte  $m(\emptyset) = 0$  implique que l'intersection entre deux hypothèses est vide. Partant de cette contrainte, deux points de vue sont alors possibles :

- soit l'on travaille dans un *monde fermé*. On considère que les hypothèses décrivent totalement le problème à résoudre. Dans ce cas la solution au problème de classification se trouve forcément parmi les hypothèses données. C'est le point de vue classique de la théorie des fonctions de croyance.

- soit l'on travaille dans un *monde ouvert*, et dans ce cas les hypothèses modélisent partiellement le problème à résoudre. Soit la solution au problème de classification se trouve parmi les hypothèses données, soit il s'agit d'une nouvelle hypothèse omise ou du moins inconnue. Ce point de vue semble *en général* plus réaliste par rapport aux applications pratiques.

Quoiqu'il en soit, plusieurs alternatives ont été proposées pour pallier le manque de cohérence des résultats de la théorie des fonctions de croyance :

- la théorie développée principalement par P. Smets qui suppose que  $m(\emptyset) > 0$ . Cette supposition repose sur l'hypothèse du monde ouvert. Cette approche a été surtout développée et utilisée dans le cadre des probabilités pignistiques [54] [56].

- la combinaison proposée par Yager [59] qui affecte la masse de tous les conflits à la masse de l'ensemble  $\Omega$ . Cette approche repose sur l'hypothèse d'un monde fermé.

- La théorie développée par J. Dezert et F. Smarandache (théorie de Dezert-Smarandache) qui suppose que l'ensemble de discernement n'est pas formé uniquement de singletons mutuellement exclusifs mais également d'unions et d'intersections non vides de ces singletons ([18] [20] [19]). Cette théorie repose sur l'hypothèse du monde fermé.

- la combinaison proposée par Lefevre *et al.* [32] qui consiste à pondérer les différentes masses d'évidence. Cette approche peut être utilisée quel que soit l'hypothèse retenue (monde fermé ou monde ouvert).

- le déconditionnement qui peut être vu comme intermédiaire entre les deux premières alternatives puisqu'il utilise à la fois la notion de monde ouvert et de monde fermé mais à des échelles différentes (voir [12]). En effet chaque source est considérée comme un cadre local et non exhaustif (donc comme un monde ouvert) mais l'ensemble  $\Omega$  est lui considéré comme un cadre global exhaustif (donc comme un monde fermé). Le déconditionnement consiste à exprimer les masses d'évidence données par rapport à un cadre local, dans le cadre global. Il nécessite cependant de donner des informations *a priori* (sous forme de



probabilités par exemple) sur les événements par rapport au cadre global (pour plus de détails voir [12]).

Il existe encore d'autres alternatives pour pallier ce problème (on pourra trouver une liste plus exhaustive dans [32] et [30]). Il faut noter d'ailleurs que l'utilisation de nouvelles combinaisons dans la théorie des fonctions de croyance ne fait pas l'unanimité au sein de la communauté scientifique comme en témoigne la correspondance entre Lefevre *et al.* d'une part, et R. Haenni d'autre part [27] [33].

Pour notre application nous considérons que nous travaillons dans un monde fermé, s'il y a conflit entre les sources d'informations, on suppose qu'au moins une des sources est erronée ou non pertinente. Les sources doivent donc être suffisamment en accord pour que la fusion ait un sens. Ce choix du monde fermé pour notre application se justifie à la fois par rapport aux caractéristiques de l'application elle-même et par rapport aux techniques de classification que nous utilisons pour calculer les masses d'évidence. En effet comme nous le verrons en détails dans le sous-paragraphe 4.1 consacré aux résultats expérimentaux, nous nous intéressons à deux types de bioprocédés :

- le bioprocédé batch. Dans ce type de bioprocédé, tous les états physiologiques sont bien connus et l'on peut donc effectuer une classification supervisée avec des échantillons étiquetés. Comme tous les états sont connus, et comme les classes correspondant à ces différents états sont définies par apprentissage supervisé, on peut considérer que toutes les hypothèses possibles (c'est-à-dire dire ici les états physiologiques) sont connus : on est donc bien dans un monde fermé.

- le bioprocédé fed-batch. Dans ce type de bioprocédé, tous les états physiologiques ne sont pas connus à l'avance : il s'agit donc *a priori* d'un monde ouvert. Mais étant donné que nous utilisons pour ce type de bioprocédé, une méthode de clustering non supervisé, dans laquelle le nombre et la forme des classes ne sont pas fournis en entrée mais directement générés et mis à jour lors de la classification, on peut supposer que cette méthode de classification fournira des classes qui représenteront toutes les situations possibles (donc tous les états physiologiques) apparaissant lors du bioprocédé. Pour cette raison, nous considérons que dans le cas des bioprocédé fed-batch nous sommes également dans un monde fermé.

Une fois que la fusion de l'information est réalisée il est possible d'utiliser différentes règles de décision pour effectuer la classification. Le choix du maximum de plausibilité correspond à un choix optimiste tandis que le choix du maximum de croyance correspond à une décision pessimiste. Le choix d'une règle de décision dépend à la fois de l'application et du comportement souhaité. Le choix des probabilités pignistiques [55] est un choix intermédiaire.

### 3.3.2. Conflit dans la théorie des fonctions de croyance et pertinence

Nous proposons d'utiliser la notion de conflit pour caractériser la pertinence des paramètres biochimiques. L'utilisation du conflit pour comparer les sources d'information avait déjà été proposée par Schubert [46].

En calculant la pertinence deux à deux entre les paramètres, il est possible de savoir quels sont les paramètres en accord et quels sont ceux qui sont en conflit. Si un paramètre est en conflit avec la majorité des autres paramètres (par exemple avec plus de la moitié des autres paramètres), il est considéré comme non pertinent, sinon il est considéré comme pertinent. Il est ainsi possible de caractériser la pertinence d'un paramètre avec une certaine flexibilité. Cette caractérisation est faite pour chaque échantillon (mesuré à chaque instant donné  $t$ ) ; la pertinence est donc évaluée de façon locale. Cette caractérisation locale est *a priori* plus significative qu'une caractérisation globale qui ne tiendrait pas compte des évolutions possibles au cours du temps. Il est alors possible à un instant  $t$  d'enlever localement de la classification de Dempster-Shafer les sources qui ne sont pas pertinentes.

La notion de pertinence d'une source est basée sur son accord avec la majorité des autres sources. Cette hypothèse est fondée dans la mesure où les signaux mesurés traduisent tous le ou les mêmes phénomènes biologiques ou biophysiques. Par rapport à la définition de la pertinence que nous avons proposée, on notera que la notion d'accord avec les autres paramètres vérifie les 3 conditions posées dans la définition :

- la condition sur l'accord avec les autres est respectée par définition.

- la condition concernant les résultats aberrants est respectée dans la mesure où une valeur aberrante du paramètre en un instant donné  $t$  produira une classification *a priori* différente de celles fournies par les autres paramètres en ce point. Le paramètre sera donc en conflit avec les autres paramètres en ce point donné.

- la condition sur les informations significatives fournies par ce paramètre découle aussi de l'hypothèse selon laquelle la majorité des paramètres fournit les informations les plus significatives puisqu'il s'agit de la traduction par différents paramètres biochimiques d'un même phénomène.

Il faut néanmoins reconnaître que l'approche que nous proposons s'appuie essentiellement sur la troisième condition (cohérence avec la majorité des autres sources). Cependant, concernant la première condition (présence de résultats aberrants), et après discussion avec les experts en microbiologie, il semble peu probable que la fiabilité des paramètres biochimiques mesurés soit remise en cause, compte tenu du système physique de mesure mis en place (qui semble robuste et précis). Quant à la deuxième condition (l'apport d'informations) par rapport à l'application et d'après les experts, tous les paramètres biochimiques apportent de l'information. En effet il s'agit d'une application réelle et les sources d'informations sont des paramètres réels (il ne s'agit pas de capteurs logiciels) qui ont un sens physique ou biochimique pour les experts et leur apportent ainsi une information. On peut tout de même montrer que ces paramètres biochimiques apportent de l'information pour la classification et ce, en utilisant plusieurs points de vue. Ainsi par exemple du point de vue du traitement du signal, on peut montrer que la plupart des paramètres biochimiques ne sont pas des signaux stationnaires et fournissent donc *a priori* des informations pour

la classification (voir [40]). D'un point de vue statistique, on peut montrer que les paramètres biochimiques ne sont pas tous corrélés les uns aux autres<sup>2</sup> et qu'ils peuvent ainsi apporter des informations (au moins au niveau statistique) sensiblement différentes (et éventuellement complémentaires) les unes des autres. Il serait également possible d'utiliser par exemple la théorie de l'information pour montrer que les sources ont un apport d'informations mais l'utilisation de la théorie de l'information n'entre pas dans le cadre de cet article. Ainsi dans l'approche que nous proposons la pertinence d'une source d'information s'appuie essentiellement sur la troisième condition (cohérence avec la majorité des autres sources).

D'autre part, étant donné qu'il n'existe aucun facteur de pondération qui donnerait plus d'importance à telle source plutôt qu'à telle autre, l'hypothèse qui stipule que la majorité a toujours raison et l'emporte, semble être la plus naturelle. On notera également que la notion de majorité est elle-même sujette à interprétation puisque l'on peut décider d'être plus ou moins restrictif (la majorité peut représenter par exemple les 2/3 des sources). La notion de pertinence ainsi définie présente au moins deux avantages :

- elle ne dépend pas directement des propriétés statistiques des sources mais des classifications obtenues pour chacune d'elles. En effet, le conflit est calculé en utilisant les masses d'évidence qui dépendent de la classification fournie à partir de chaque paramètre (ie chaque source).

- en raison de la valeur numérique du conflit et de la flexibilité de la notion de majorité, il est possible de définir un certain degré de pertinence (vision graduelle de la pertinence).

La notion de pertinence ainsi définie est locale et non globale car elle est évaluée pour chaque échantillon. Cette notion de pertinence locale se rapproche de la notion de pertinence dynamique utilisée dans les méthodes proposées par Baluja et Pomerleau [6] d'une part et Greiner *et al.* [25] d'autre part. Néanmoins dans la méthode de Baluja et Pomerleau la pertinence d'une variable à l'instant  $t$  dépend directement de la pertinence de cette variable à l'instant  $t - 1$  ce qui n'est pas le cas dans la méthode que nous proposons. L'approche proposée par Greiner, elle, s'appuie explicitement sur la connaissance de l'expert pour modifier dynamiquement la pertinence. Puisque la pertinence des paramètres biochimiques varie en fonction de l'évolution du bioprocédé, cette notion de pertinence dynamique est donc bien adaptée à notre application.

Par ailleurs, comme nous l'avons dit plus haut, cette méthode s'appuyant sur le conflit de la théorie des fonctions de croyance, se rapproche de celle proposée par Schubert [46-48]. On notera également que Ayoun et Smets [5] ont proposé une méthode utilisant les probabilités pignistiques pour associer des sources

cohérentes : cette méthode s'appuyait sur l'approche du conflit des sources proposée par Schubert. Dans cette approche proposée par Schubert, le conflit est utilisé pour définir des sous-ensembles regroupant les sources d'information travaillant sur les mêmes éléments focaux. Schubert cherche en effet à former des groupes de sources d'information pour éviter de combiner des sources d'informations ne travaillant pas sur les mêmes éléments focaux, mais il ne s'intéresse pas à la notion de pertinence de ces sources. La méthode proposée dans cet article est similaire à celle de Schubert dans la mesure où la recherche des paramètres pertinents revient à chercher dans les regroupements des sources de Schubert, le groupe contenant le plus grand nombre de paramètres. Les paramètres appartenant à ce sous-ensemble seront considérés comme pertinents puisque ce sous-ensemble représente le plus grand nombre de paramètres.

Dans l'approche proposée dans cet article, les sources considérées comme non pertinentes ne sont pas intégrées dans la classification. Il existe d'autres alternatives pour gérer la fusion d'information de sources. En effet, la gestion d'une source non pertinente peut varier en fonction de l'objectif à atteindre, des informations exogènes disponibles et de l'application en question et de la méthode utilisée. Bloch, Hunter *et al.* [9]<sup>3</sup> propose au moins quatre possibilités pour gérer les conflits :

- (1) garder les sources pertinentes et ignorer les informations fournies par les sources non pertinentes (solution unilatérale)
- (2) réaliser un compromis c'est-à-dire fusionner dans une certaine mesure les informations fournies par les sources pertinentes et par les sources non pertinentes (solution du compromis)
- (3) créer un « nouveau monde » : on considère dans ce cas que les sources pertinentes n'observent pas le même objet ou le même monde que les sources non pertinentes.
- (4) reporter la classification pour attendre des informations complémentaires (solution du délai).

La première solution (qui est la plus classique) nous semble être la plus appropriée pour gérer les conflits entre sources dans cette application. Ainsi dans la méthode que nous proposons, les paramètres considérés comme non pertinents ne seront pas pris en compte lors de la classification.

Pour notre application biotechnologique, la seconde solution ne semble pas la mieux adaptée car on cherche justement à ne pas mélanger les informations fournies par les paramètres pertinents et les paramètres non pertinents puisqu'il est difficile d'évaluer *a priori* l'influence de ces paramètres non pertinents dans la classification. En outre, les techniques utilisées actuellement pour cette application fusionnent les informations sans tenir compte de la pertinence des sources, et de ce fait, utilise cette seconde solution. Cependant, si l'on réalise une fusion adaptative qui tient compte de la pertinence locale (*ie.* à l'instant  $t$ ), le compromis peut être intéressant. Par exemple, il est aussi possible d'utiliser la technique d'affaiblissement qui consiste à pondérer les sources en fonction de leur fiabilité (dans notre

2. Il est même possible que les corrélations des paramètres biochimiques évoluent dans le temps et ce, en raison notamment des diverses voies métaboliques mis en jeu par les micro-organismes et de l'influence de certains paramètres régulés par les experts, mais nous n'insisterons pas plus sur cet aspect (voir [42]).

3. À noter que Lefevre *et al.* évoquent également la gestion des conflits [32].

approche, la notion de fiabilité sera remplacée par celle de pertinence). Nous présenterons donc une variante de notre approche qui utilise l'affaiblissement.

La troisième solution ne nous semble pas non plus adéquate car les paramètres mesurés traduisent tous le même phénomène biochimique.

La dernière solution n'est pas adaptée car nous ne disposons pas d'informations (exogènes ou endogènes) supplémentaires en temps réel.

Ainsi l'approche que nous proposons pour évaluer la pertinence et effectuer la classification peut se résumer comme suit :

- Pour chaque instant donné  $t$  :

(1) Caractérisation de la pertinence

- Pour chaque paramètre P

- calcul du conflit deux à deux entre le paramètre P et les autres paramètres

- si le paramètre est en accord avec la majorité des autres paramètres, alors il est pertinent

- sinon, il est « non pertinent »

(2) Fusion de l'information

- fusion de tous les paramètres pertinents par la combinaison de Dempster

- si aucun paramètre n'est pertinent, fusion d'un nombre limité de paramètres choisis par l'expert et considérés comme pertinents

Le rajout d'une condition imposant un nombre de paramètres choisis par l'expert en microbiologie permet d'éviter le cas où aucun paramètre ne serait considéré comme pertinent car alors il serait impossible d'effectuer la classification. Dans ce cas de figure, les paramètres choisis par l'expert en microbiologie sont ceux sur lequel il s'appuie pour trouver « manuellement » les états physiologiques. Pour cette application, la méthode ne gère donc pas le cas où aucun paramètre n'est pertinent puisque l'expert impose dans ce cas le choix arbitraire de certains paramètres. Mais pour d'autres applications l'approche proposée pourrait gérer cette situation soit en assouplissant la valeur seuil du conflit, soit en assouplissant le nombre minimum de paramètres en accord (c'est-à-dire la « majorité ») pour qu'un paramètre soit considéré comme pertinent.

Nous présentons une variante de cette approche qui s'inspire cette fois de la méthode d'affaiblissement (voir [49] [8] [4]). Dans cette variante les paramètres considérés comme non pertinents sont intégrés dans la fusion mais avec un facteur de pondération moindre. Nous avons choisi d'utiliser un poids de 0.5 pour ces paramètres non pertinents alors que le facteur des paramètres pertinents reste à 1. Cette variante peut se résumer comme suit :

- Pour chaque instant donné  $t$  :

(1) Caractérisation de la pertinence

- Pour chaque paramètre P

- calcul du conflit deux à deux entre le paramètre P et les autres paramètres

- si le paramètre est en accord avec la majorité des autres paramètres, alors il est pertinent

- sinon, il est « non pertinent »

(2) Fusion de l'information par la combinaison de Dempster de tous les paramètres avec :

- un poids égal à 1 pour tous les paramètres pertinents

- un poids égal à 0.5 pour tous les paramètres non pertinents

- pour l'ensemble  $\Theta$ , la masse d'évidence  $m(\Theta)$  reste inchangée lors de la fusion pour les paramètres pertinents, tandis que pour les paramètres non pertinents  $m(\Theta)$  est mise à jour avec  $m(\Theta) = 0.5 \times (1 - m(\Theta))$

### 3.3.3. Mesure du conflit

Comme on l'a vu, la valeur  $K$  permet de mesurer le conflit entre sources. Néanmoins, le conflit doit être suffisamment faible pour que la fusion ait un sens. Pourtant en utilisant la mesure de conflit  $K$  on peut arriver à des résultats erronés. Ainsi pour deux paramètres fournissant exactement les mêmes masses d'évidence, on peut trouver un conflit  $K$  non nul entre les deux sources.

Par exemple supposons que nous avons 3 paramètres P1, P2 et P3 et deux classes  $C_1$  et  $C_2$  (on a donc  $\Theta = \{C_1, C_2\}$ ). Les masses d'évidence supposées connues pour les 2 classes sont respectivement pour P1, P2 et P3 (voir aussi tableau 1) :  $m_1(C_1) = \frac{1}{3}$ ,  $m_1(C_2) = \frac{1}{3}$ ,  $m_1(\Theta) = \frac{1}{3}$ ;  $m_2(C_1) = \frac{1}{3}$ ,  $m_2(C_2) = \frac{1}{3}$ ,  $m_2(\Theta) = \frac{1}{3}$ ;  $m_3(C_1) = 0$ ,  $m_3(C_2) = \frac{1}{3}$  et  $m_3(\Theta) = \frac{2}{3}$ .

Tableau 1. Les masses d'évidences pour P1, P2 et P3.

	P1	P2	P3
$m(C_1)$	$\frac{1}{3}$	$\frac{1}{3}$	0
$m(C_2)$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
$m(\Theta)$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{2}{3}$

Nous pouvons calculer le conflit  $K_{1,2}$  entre P1 et P2, et  $K_{1,3}$  entre P1 et P3.

Le calcul du conflit se simplifie et revient dans ce cas à la formule suivante :

$$K_{i,j,i \neq j} = \sum_{k,l,k \neq l} m_i(C_k) \cdot m_j(C_l) \quad (6)$$

Le conflit  $K_{1,2}$  est égal à 0.222 alors que les paramètres P1 et P2 sont parfaitement en accord. De plus le conflit  $K_{1,3}$  est égal à 0.111. Ainsi le conflit entre deux paramètres fournissant la même distribution de masses d'évidence est non nul et même supérieur au conflit entre deux paramètres donnant des résultats différents. Bien que diverses alternatives aient été proposées pour pallier ce problème [32], nous proposons une approche intuitive pour calculer le conflit entre les sources d'information. Cette approche conduit à la définition d'une nouvelle mesure de conflit basée sur la norme 1 (qu'on appelle aussi norme L1). Nous rappelons la définition de la norme 1 pour  $y \in R_n$  :



$$\|y\|_1 = \sum_{i=1}^n |y_i| \quad (7)$$

La nouvelle mesure du conflit que nous proposons est définie comme suit :

$$D = \frac{1}{2} \sum_i |m_1(C_i) - m_2(C_i)| \quad (8)$$

où  $m_1(C_i)$  et  $m_2(C_i)$  représente les masses d'évidence pour des paramètres 1 et 2 pour la classe  $i$ . Le facteur  $\frac{1}{2}$  est un facteur de normalisation.

On peut noter que la norme est mathématiquement équivalente à la norme Euclidienne et à la norme Sup sur  $R^n$  (avec  $n$  fini); il est donc possible de définir une mesure de conflit basée sur l'une de ces deux autres normes. Par ailleurs, il n'est pas nécessaire de définir une généralisation du calcul de la mesure du conflit basée sur la norme 1 pour  $N$  paramètres, puisque l'on calcule ce conflit deux à deux entre les paramètres. Il faut tout de même noter que cette généralisation n'est pas possible (contrairement à la notion classique du conflit de Dempster) étant donné que ce calcul du conflit n'est pas associatif.

Si nous reprenons l'exemple précédent, le calcul du conflit devient plus cohérent. Ainsi le conflit  $D_{1,2}$  entre P1 et P2 est égal à 0 et le conflit  $D_{1,3}$  entre P1 et P3 est égal à 0.666. Ainsi l'utilisation de cette nouvelle mesure du conflit donne des résultats plus cohérents que le conflit défini dans la méthode de Dempster-Shafer.

Notons tout de même une différence conceptuelle entre la mesure de conflit  $K$  de la théorie des fonctions de croyance, et la mesure  $D$  proposée dans cet article. En effet, le conflit  $K$  permet de mesurer, pour 2 sources, la différence de répartitions des masses d'évidence sur des éléments focaux exclusifs, alors que la mesure  $D$  mesure, pour 2 sources, la différence de répartition de l'ensemble des masses d'évidence, à condition que les 2 sources travaillent sur les mêmes éléments focaux. Cependant, autant que nous le sachions, la *plupart* des méthodes *automatiques* de calcul des masses d'évidence, fournissent pour toutes les sources une répartition de masses sur les mêmes éléments focaux. Ainsi, quand on utilise ce type de méthode automatique de calcul des masses d'évidence, la mesure  $D$  semble mieux adaptée pour calculer le conflit entre sources.

Toutefois, il est clair que ce conflit  $D$  peut être utilisé uniquement dans le cas où les sources d'information travaillent sur les mêmes éléments focaux de  $2^\Theta$ .

*Il faut cependant se rappeler que l'idée de base de cette approche est moins l'utilisation d'une nouvelle mesure du conflit que l'utilisation de la théorie des fonctions de croyance pour caractériser la pertinence des sources d'informations.*

## 4. Résultats expérimentaux

### 4.1. Bioprocédé batch et fed-batch : classifications supervisée et non supervisée

Les bioprocédés que nous étudions sont des bioprocédés de fermentation utilisant les micro-organismes appelés *Saccharomyces Cerevisiae*. Nous avons étudié deux types de bioprocédés :

- la fermentation batch. Dans ce type de fermentation, le substrat est placé au début du procédé puis aucune intervention extérieure n'est faite jusqu'à la fin de l'expérience. Pour ce bioprocédé, l'expert en microbiologie connaît parfaitement les différents états physiologiques (voir figure 1). Il est donc possible d'effectuer une classification supervisée en utilisant des échantillons étiquetés. On pourra également évaluer le pourcentage de classification correcte.

- la fermentation fed-batch. Cette fois, des interventions sont réalisées au cours de toute l'expérience. En particulier, les micro-organismes sont alimentés en substrat tout au long du cycle de la fermentation. Pour ce bioprocédé, l'expert ne connaît pas tous les états physiologiques. On effectuera une classification non supervisée et il n'est pas possible à l'heure actuelle d'évaluer le pourcentage de classification correcte.

Pour le bioprocédé batch, l'expérience dure environ 20 heures et correspond à 1012 points de mesures des paramètres biochimiques. On considère le début du bioprocédé comme étant  $t = 0$  heure (0h). On rappelle que l'on cherche à détecter trois états physiologiques principaux :

- l'état 1 : la fermentation (production d'éthanol). Elle va de 0h jusqu'à environ 9h ce qui représente un total de 590 points mesurés.

- l'état 2 : la diauxie. Cet état commence à environ 9h et se termine à 9h46 ce qui représente environ 33 points de mesure. C'est le plus petit état physiologique (en temps et en quantité de données) parmi les 3 et le plus difficile à caractériser

- l'état 3 : l'oxydation (production de biomasse). Elle commence à 9h46 et se termine en même temps que la fin de l'expérience à 20h ce qui représente 389 points de mesure.

Il y a 22 paramètres biochimiques et chacun d'eux a donc 1012 éléments.

Pour le bioprocédé fed-batch, l'expérience dure 21.60 heures. On considère le début du bioprocédé comme étant  $t = 0$  heure (0h). Il y a 12 paramètres ayant chacun 1400 points de mesures. Les états physiologiques ne sont pas connus; ce que l'on sait c'est qu'il existe deux types de phénomènes durant ce bioprocédé :

- les phénomènes dits biologiques. Ce sont des phénomènes qui sont dus uniquement à la biologie du système c'est-à-dire qu'ils sont générés par les micro-organismes et ne sont pas le résultats d'une action externe.

- les phénomènes dits biophysiques (ou physico-chimiques). Ils sont le résultat d'une réaction des micro-organismes suite à une action externe. Une action externe est la modification d'un paramètre régulé par l'expert en microbiologie (ex : changement de température, introduction d'un produit acide ou basique dans le milieu...).

#### 4.2. Méthodes de calcul des masses d'évidence

Nous avons utilisé une méthode de calcul des masses d'évidence distincte pour chaque expérience.

Pour le bioprocédé batch, puisque les classes étaient connues, nous avons utilisé la méthode de calcul des masses proposée par Denoeux [16] et basée sur la méthode des  $k$  plus proches voisins.

Pour le bioprocédé fed-batch, étant donné que nous ne connaissions pas les classes à chercher, nous avons utilisé une méthode de clustering floue, et non supervisée pour générer des classes à partir desquelles nous avons généré des masses d'évidence. Cette méthode est baptisée LAMDA (Learning Algorithm for Multivariate Data Analysis) [37] [41]. Les masses d'évidence sont calculées à partir des degrés d'appartenance aux classes fournies par LAMDA pour chaque paramètre : nous avons normalisé les degrés d'appartenance fournis par la classification LAMDA. Chaque masse d'évidence est donnée pour une classe et une seule : il s'agit donc de masses d'évidence calculées pour des singletons. Les masses d'évidence qui en résultent sont donc équivalentes à des probabilités.

Il faut noter que nous avons testé une autre méthode pour calculer les masses d'évidence à partir de ces degrés d'appartenance. Cette méthode développée par Desachy *et al.* [17] permet en effet d'affecter une masse d'évidence à l'union de toutes les classes ; cette masse d'évidence affectée à l'union de toutes les classes représente l'ignorance de la source d'information : les masses d'évidence ainsi calculées ne sont donc plus des probabilités. Cependant, cette méthode générait des difficultés au niveau des résultats numériques. En effet, quelque soit le paramètre biochimique, cette méthode affectait la majeure partie à l'ensemble  $\Theta$ , affectant ainsi un degré d'ignorance important à tous les paramètres biochimiques (le rapport entre la masse d'évidence de  $\Theta$  et les autres masses d'évidence est de l'ordre de  $10^3$ , voir [40]).

Par ailleurs, du fait que les deux méthodes utilisées pour le calcul des masses d'évidence travaillent soit sur  $\Theta$  et des singletons, soit sur des singletons, la plausibilité, la croyance ou les probabilités pignistiques pour la décision conduisent au même résultat.

#### 4.3. Résultats expérimentaux

Les premiers tests avaient été réalisés en utilisant la méthode LAMDA pour calculer les masses d'évidence sur un bioprocédé

de type batch : ces résultats sont disponibles dans [41]. La méthode proposée a cette fois été testée sur les deux bioprocédés de type batch et fed-batch avec les deux méthodes de calcul de masses d'évidence présentées ci-dessus.

Pour le bioprocédé batch, nous avons utilisé la méthode proposée par Denoeux [16]. 68 échantillons ont été testés (30 pour l'état 1, 7 pour l'état 2 et 31 pour l'état 3). Dans l'approche proposée par Denoeux, le calcul des masses d'évidence se fait ainsi : pour chacune des 3 classes  $C_1, C_2, C_3$  et pour l'ensemble  $\Theta$  la formule est donnée par les équations ci-dessous :

$$m(C_i) = \frac{m_i(C_i) \prod_{j \neq i} m_j(\Theta)}{K} \quad (9)$$

$$m(\Theta) = \frac{\prod_{i=1}^3 m_i(\Theta)}{K} \quad (10)$$

où  $K$  est le facteur de normalisation suivant :

$$K = \sum_{i=1}^3 \prod_{j \neq i} m_j(\Theta) + (1 - 3) \prod_{i=1}^3 m_i(\Theta) \quad (11)$$

Avec :

$$m_i(C_i) = 1 - \prod_{x_{ki} \in C_i} (1 - \alpha_0 e^{-d^{ki,l}}) \quad (12)$$

$$m_i(\Theta) = \prod_{x_{ki} \in C_i} (1 - \alpha_0 e^{-d^{ki,l}}) \quad (13)$$

où  $d^{ki,l}$  représente la distance métrique entre chaque mesure à classer  $x_l$  et chaque échantillon étiqueté  $x_{ki}$  de la classe  $C_i$  ( $i = \{1,2,3\}$ ) pour chaque paramètre biochimique ;  $e$  représente la fonction exponentielle et  $\alpha_0$  est une valeur fixée comprise entre 0 et 1 (pour plus de détails voir [16]). Pour l'application, nous avons choisi  $\alpha_0 = 0.95$ . Dans les tests que nous avons effectués nous avons retenu la valeur  $k = 7$ ,  $k$  étant le nombre de voisins pour la méthode des  $k$ -plus proches voisins (plusieurs valeurs de  $k$  ont été testées, voir figure 4). Un exemple de masse d'évidence pour quelques paramètres est donné dans le tableau 2. La classification a été réalisée en utilisant la théorie des fonctions de croyance avec et sans prise en compte de la pertinence. Un paramètre est considéré comme étant en accord avec un autre paramètre si leur conflit D est inférieure strictement à 0.5. Un paramètre est considéré comme pertinent s'il est en accord avec la majorité absolue des paramètres (*ie* la moitié des paramètres plus 1 soit au minimum 12 paramètres).

L'évaluation de la pertinence des différentes sources d'information a donné des résultats en accord avec la connaissance des experts. Tout d'abord, la majorité des paramètres sont considérés comme non pertinents vers la fin de l'expérience (approximativement de  $t = 18.5h$  à  $t = 20h$ ) comme on peut le voir sur la figure 2 (le détail des intervalles temporels de non pertinence est

donné dans le tableau 8 en annexe). Ce phénomène (déjà observé dans [41]) est connu des experts et s'explique par la décroissance et la mort des micro-organismes. Cette mort entraîne une situation chaotique qui explique le nombre élevé de paramètres non pertinents. Par ailleurs sur cette même figure 2 on constate également qu'aucun paramètre n'est pertinent pendant l'état 2 (diauxie) et au voisinage de cet état. Ceci peut s'expliquer principalement par le nombre peu élevé d'éléments de cet état (33 seulement) comparativement aux autres états, et de ce fait par le petit nombre d'échantillons pour l'apprentissage (7 échantillons). De plus on constate que certains paramètres ont du mal à discriminer d'un point de vue statistique cet état par rapport aux autres états : c'est le cas du pH ou de la luminance comme on peut le voir sur le tableau 3 (les statistiques pour chaque état et chaque paramètre est donné dans le tableau 7 en annexe).

Par ailleurs beaucoup de paramètres qui présentent des intervalles temporels de « non-pertinence » sont des paramètres régulés qui ne traduisent pas la biologie du système. Par exemple, le paramètre Température qui est un paramètre régulé reste pertinent de  $t = 0h$  jusqu'à  $t = 8.05h$  puis devient non pertinent juste avant l'apparition de la classe correspondant à l'état 2 (voir figure 3). Or en analysant ce paramètre on constate qu'il évolue peu jusqu'à  $t = 15h$  où il est modifié par l'expert. Or la pertinence revient à partir de  $t = 16.5h$  jusqu'à la mort des micro-organismes (vers  $t = 18.5h$ ) et correspond à la modification (le temps de retard de la pertinence peut correspondre à un temps de latence du système biologique). De même d'autres paramètres régulés comme le paramètre « Action froid », le paramètre « Ajout de base » deviennent non pertinents juste avant l'apparition de la classe correspondant à l'état 2 et redeviennent pertinents après qu'ils aient été régulés et ce jusqu'à la mort des micro-organismes. Un autre exemple parlant est celui du paramètre pression qui reste constant pendant tout le procédé. Comme on pouvait s'y attendre ce paramètre pression

Tableau 2. Une exemple de masse d'évidence pour 3 paramètres (pO<sub>2</sub>: pression partielle d'oxygène, CO<sub>2</sub>mes: dioxyde de carbone mesuré, et Q-resp: quotient respiratoire) pour les 4 éléments focaux: C1 (état 1), C2 (état 2), C3 (état 3) et l'ensemble des sous-ensemble  $\Theta$ . Le résultat de la fusion tient compte de tous les paramètres pertinents. Les valeurs ont été calculées pour  $t = 9.08h$  et  $k = 7$  (nombre de  $k$  plus proches voisins).

	pO <sub>2</sub>	CO <sub>2</sub> mes	Q-resp	Résultat de la fusion
m(C1)	0.948159	0.999817	0	0
m(C2)	0.051834	0.000182	0.000125	0.945084
m(C3)	0	0	0.999875	0
m( $\Theta$ )	0.000007	0.000001	0	0

devient non pertinent à  $t = 8h$  (juste avant l'état 2) et reste non pertinent jusqu'à la fin car il ne subit aucune régulation. Les résultats de l'évaluation de l'ensemble des paramètres sont toujours en cours d'analyse par les experts du domaine.

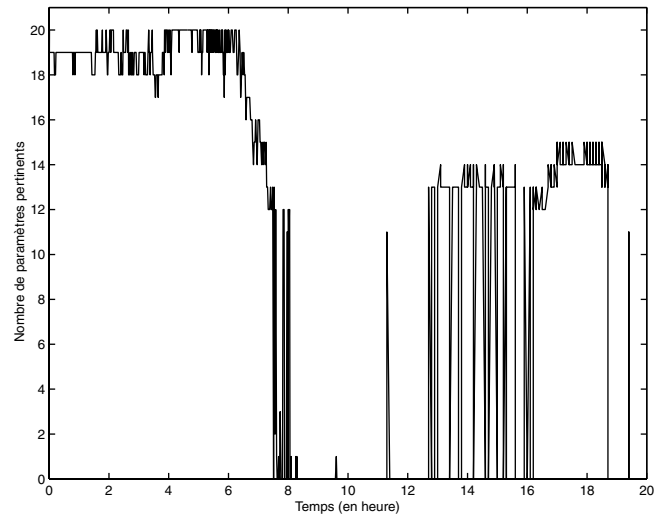


Figure 2. Évolution du nombre de paramètres pertinents en fonction du temps pour  $k = 7$  (l'axe des abscisses représente le temps mesuré en heures durant le bioprocédé, l'axe des ordonnées le nombre de paramètres pertinents).

Tableau 3. Propriétés statistiques des états physiologiques : moyenne  $m$  et écart-type  $\sigma$  pour quelques paramètres pH, rCO<sub>2</sub> (vitesse de production de dioxyde de carbone), rO<sub>2</sub> (vitesse de consommation du dioxygène), Lum (luminance).

	pH	rCO <sub>2</sub>	rO <sub>2</sub>	Lum
État 1	m=4.990 $\sigma=0.045$	m=22.241 $\sigma=20.641$	m=1.659 $\sigma=1.648$	m=69.109 $\sigma=0.996$
État 2	m=4.992 $\sigma=0.024$	m=6.278 $\sigma=0.736$	m=6.745 $\sigma=0.457$	m=67.806 $\sigma=0.098$
État 3	m=5.164 $\sigma=0.295$	m=13.857 $\sigma=7.687$	m=19.548 $\sigma=12.394$	m=67.180 $\sigma=3.447$

On voit donc que l'évaluation de la pertinence fournit des résultats cohérents par rapport à la gestion des paramètres régulés. Par ailleurs en ce qui concerne la classification proprement dite, la méthode d'évaluation de la pertinence fournit aussi des résultats satisfaisants. En effet, le fait d'évaluer la pertinence des paramètres et de ne pas inclure dans la fusion de données les paramètres non pertinents à chaque instant donné  $t$  permet d'améliorer les résultats de classification. Les résultats de la classification sont donnés dans le tableau 4. On rappelle que le nombre de  $k$  plus proches voisins est 7. Avec l'utilisation de la

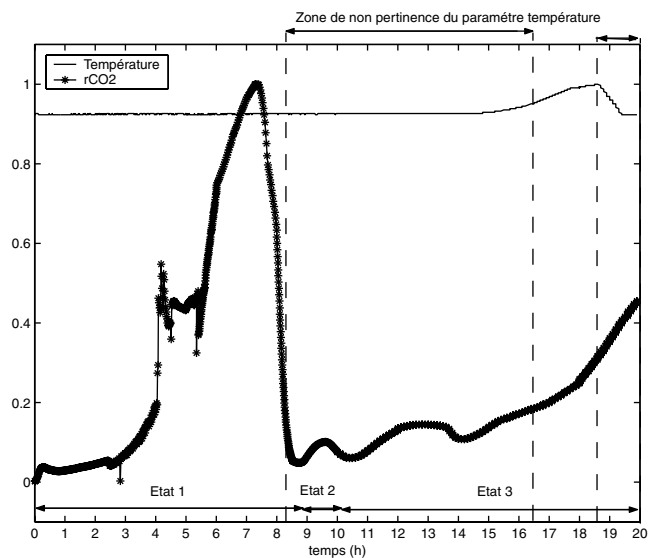


Figure 3. L'absence de pertinence du paramètre Température (qui est un paramètre régulé) apparaît juste avant l'arrivée de l'état 2 et disparaît une fois que ce paramètre Température est modifié. Le deuxième intervalle de non pertinence correspond à la mort des micro-organismes (sur ce schéma, les amplitudes ont été normalisées ; l'axe des abscisses représente le temps mesuré en heures durant le bioprocédé).

pertinence, les résultats sont améliorés de près de 20 points. Pour l'affaiblissement, on constate que pour  $k = 7$  l'approche donne des résultats au moins équivalents à ceux de l'approche classique sans prise en compte de pertinence. On voit bien que le fait d'inclure des sources d'information non pertinentes dans la classification biaise les résultats et que l'approche proposée dans cet article permet non seulement de donner des informations sur la pertinence des sources mais aussi d'améliorer les résultats.

Tableau 4. Résultats de la classification pour  $k = 7$ .

Méthode	% classification
fonctions de croyance	59.78 %
fonctions de croyance et pertinence	77.37 %
fonctions de croyance pertinence et affaiblissement	59.78 %

Des tests ont été réalisés en utilisant d'autres valeurs de  $k$  (voir figure 4). Globalement, les résultats de la classification par la théorie des fonctions de croyance qui prend en compte la pertinence sont meilleurs que ceux de la classification sans prise en compte de la pertinence. Ainsi le meilleur taux de classification (86.36 % de classification correcte pour  $k = 1$ ) est obtenu en tenant compte de la pertinence avec affaiblissement. Mais on notera que la méthode d'affaiblissement suit une trajectoire quasiment similaire à celle de l'approche classique à partir de  $k = 3$ . En moyenne c'est la méthode de prise en compte de la pertinence avec élimination qui obtient les meilleurs résultats (voir 5). Cependant cette méthode a ses limites : la classification sans prise en compte de la pertinence obtient de meilleurs résultats pour deux cas ( $k = 3$  et  $k = 2$ ). Cela est dû au fait que pour ces deux valeurs de  $k$ , lors de la classification, pour un nombre important d'instant  $t$ , aucun paramètre n'était considéré comme pertinent (on s'appuyait alors sur deux paramètres choisis par l'expert en microbiologie : le rCO2, soit la vitesse de production de dioxyde de carbone et rO2, soit la vitesse de consommation de l'oxygène) : ainsi pour 38 % des mesures effectuées durant le bioprocédé aucun paramètre n'était pertinent pour  $k = 3$ , et pour 39.92 % des mesures effectuées durant le bioprocédé aucun paramètre n'était pertinent pour  $k = 2$ .

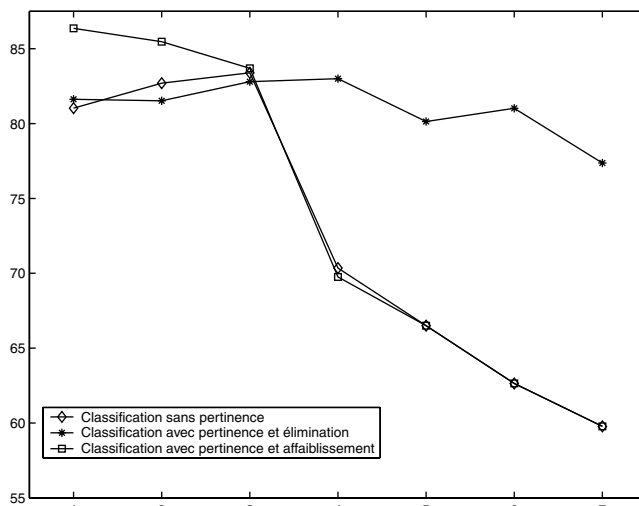


Figure 4. Pourcentage de classification correcte pour le bioprocédé batch, en fonction du nombre  $k$  (nombre de  $k$  plus proches voisins) l'axe des abscisses représente le nombre de  $k$  et l'axe des ordonnées le pourcentage de classification correcte. Globalement, la méthode tenant compte de la pertinence (avec élimination ou avec affaiblissement) donne les meilleurs résultats.

Pour le bioprocédé fed-batch, la méthode de clustering LAMDA a été utilisée. Les paramètres de la méthode ont été réglés avec l'accord de l'expert. Suite aux tests réalisés avec l'expert, un paramètre est considéré comme étant en accord avec un autre paramètre si leur conflit D est inférieur strictement à 0.3. Un paramètre est considéré comme pertinent s'il est accord avec la majorité absolue des paramètres (ie la moitié des paramètres plus 1 soit au minimum 7 paramètres).

Tableau 5. Moyenne des classifications pour toutes les valeurs de  $k$  (de  $k = 1$  à  $k = 7$ ).

Méthode	% classification moyenne
fonctions de croyance	72.34 %
fonctions de croyance et pertinence	81.07 %
fonctions de croyance pertinence et affaiblissement	73.45 %

Nous nous sommes surtout intéressés à la pertinence des paramètres biochimiques pour le bioprocédé fed-batch et plus particulièrement les paramètres régulés. Nous avons tout de même effectué la classification en utilisant la théorie des fonctions de croyance avec et sans la notion de pertinence, mais il n'a pas été possible d'évaluer numériquement le pourcentage de classification correcte.

Ainsi par exemple, l'agitation, qui est un paramètre régulé représenté par une fonction par paliers, est évaluée comme non pertinent par la méthode juste avant les deux premiers changements de palier comme cela est indiqué sur la figure 5, et de ce fait n'est pas inclus dans la fusion pour la classification (le détail des intervalles temporels de non pertinence est donné dans le tableau 9 en annexe). Une nouvelle classe apparaît en même temps que ce manque de pertinence (ce qui n'est pas le cas pour la classification par la théorie des fonctions de croyance qui ne tient pas compte de la pertinence). Or cela est confirmé par les experts puisque le changement de palier est réalisé manuellement par les experts en microbiologie qui considèrent qu'après un certain temps, la variable agitation ne sera plus en adéquation avec les autres paramètres de l'expérience, et ils effectuent ce changement pour anticiper ce phénomène. La non pertinence de ce paramètre régulé suggère ici une commande prédictive. Ainsi, en plus d'agir comme un système expert, la méthode développée agit donc comme un système prédictif, évaluant les signaux régulés comme non pertinents s'ils ne sont plus en adéquation avec le phénomène biologique. Par ailleurs, le premier et le dernier intervalle de non pertinence correspondent à des phénomènes biologiques [42], ce qui confirme le fait que l'agitation ne soit pas à ce moment là en adéquation avec le système biologique.

La méthode basée sur la pertinence à partir de la théorie des fonctions de croyance permet d'améliorer la connaissance sur les deux bioprocédés testés. D'une part, elle améliore les résul-

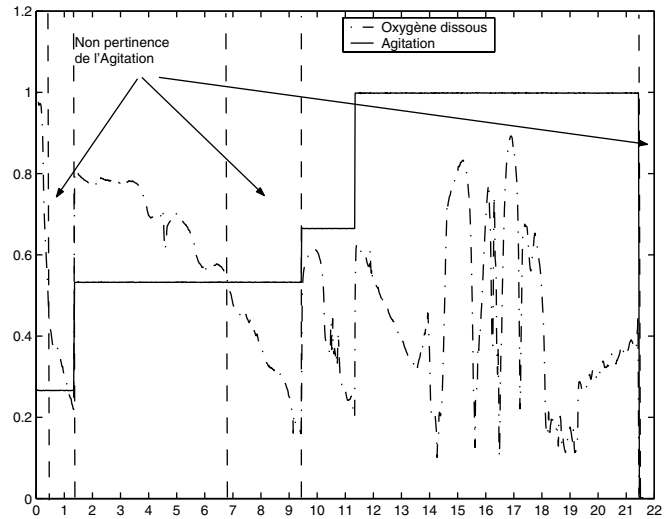


Figure 5. L'agitation est considérée comme non pertinente durant certains laps de temps dans le bioprocédé fed-batch (sur ce schéma, les amplitudes ont été normalisées; l'axe des abscisses représente le temps mesuré en heures durant le bioprocédé).

tats de la classification, d'autre part, elle permet un meilleur contrôle de certains paramètres régulés. Toutefois les résultats doivent être analysés plus profondément par les microbiologistes afin de mieux interpréter au niveau biologique et biophysique les phénomènes observés. Par ailleurs, le choix du niveau de conflit (0.5 et 0.3 respectivement pour le bioprocédé batch et fed-batch) ainsi que le choix de la majorité qui détermine le degré de pertinence restent empiriques et doivent être analysés plus profondément.

## 5. Conclusion

Nous avons présenté dans cet article une utilisation de la théorie des fonctions de croyance pour la classification dans des bioprocédés et l'évaluation de la pertinence des sources d'information. La définition de la pertinence que nous avons proposée repose sur la notion de conflit de la théorie des fonctions de croyance. Nous avons proposé une autre mesure du conflit basée sur la norme 1 et qui fournit des résultats plus cohérents. Cette mesure ne peut être utilisée que si toutes les sources travaillent sur les mêmes éléments focaux. Les résultats expérimentaux fournis par la pertinence définie dans cette méthode sont cohérents et confirmés par les observations des experts. Ces résultats montrent aussi que la notion de pertinence peut améliorer la classification. Plusieurs axes de travaux supplémentaires existent. L'analyse du conflit et le choix du seuil doivent faire l'objet d'études complémentaires. Les futurs travaux concerneront aussi l'analyse de la pertinence sur des intervalles de temps (et au lieu de l'analyse actuelle qui se fait point par



point). Une autre voie à explorer est d'utiliser une CAH (classification ascendante hiérarchique) non sur les données mais sur les paramètres pour choisir le nombre de paramètres pertinents fusionnés et analyser l'impact de cette fusion sur les résultats de la classification. L'approche proposée dans cet article pourrait aussi être appliquée à d'autres domaines comme la fusion d'informations issues d'images multispectrales.

Nous remercions l'équipe fermentation du LBB INSA pour avoir mis à notre disposition les données expérimentales et pour leur aide à la validation des résultats.

## Annexe

Dans cette annexe, nous donnons

- pour le bioprocédé batch: les statistiques des paramètres durant tout le bioprocédé, les statistiques de chaque paramètre

Tableau 6. Moyenne et écart-type des différents paramètres pour le procédé de type batch.

Paramètre/	moyenne	Écart-type
Température	30,407	0,624
pH	5,057	0,205
Agitation	561,069	122,261
pO2	77,510	25,809
Pression	0,000	0,000
Qair-entrée	13,546	3,87
Qair-sortie	13,558	3,893
CO2mes	0,039	0,005
O2mes	20,829	0,058
Action-chaud	1788,727	658,427
Action-froid	1797,421	3691,552
Ajout-base	385,190	349,283
U-volt(ppe)	4,396	0,627
Qair-sortie2	13,723	4,383
Lum	68,325	2,453
a	-1,652	0,134
b	7,938	0,903
delta	14,018	1,725
X-estim	0,974	0,810
rO2	8,701	11,615
rCO2	18,498	17,101
Q-resp	10,498	18,886

pour chaque état, et les intervalles temporels de non de pertinence pour chaque paramètre.

- pour le bioprocédé fed-batch: les intervalles temporels de non pertinence pour chaque paramètre.

Nous donnons également quelques explications pour certains des paramètres présentés.

L'action chaude permet de réguler la température du milieu. L'agitation correspond à la vitesse d'un rotor qui permet de garder l'homogénéité du milieu. L'ajout de base permet de réguler le pH. Le rCO2 et le rO2 correspondent respectivement à la vitesse de production de dioxyde de carbone, et à la vitesse de consommation de l'oxygène. Le CO2 mesuré et l'O2 mesuré correspondent à la mesure directe de dioxyde de carbone et d'oxygène présents dans le milieu. La luminance dépend de l'intensité lumineuse du milieu et mesure la concentration de la biomasse (la biomasse est la quantité de micro-organismes) dans le milieu. La tension électrique mesure l'activité électrochimique des micro-organismes.

Tableau 7. Moyenne (m) et écart-type (σ) des différentes classes (ou états) pour chaque paramètre pour le procédé de type batch.

État / Paramètre	État 1 : fermentation	État 2 : Diauxie	État 3 : Oxydation
Température	m= 30.155 σ=0.050	m= 30.191 σ=0.029	m=30.808 σ=0.865
pH	m=4.990 σ=0.045	m=4.992 σ=0.024	m=5.164 σ=0.295
Agitation	m=501.424 σ=126.196	m=656.606 σ=0.776	m=643.429 σ=42.183
pO2	m=90.764 σ=4.484	m=86.870 σ=0.742	m=56.613 σ=31.493
Pression	m=0 σ=0	m=0 σ=0	m=0 σ=0
Qair-entrée	m=12.495 σ=4.800	m=15.003 σ=0.017	m=15.017 σ=0.038
Qair-sortie	m=12.494 σ=4.825	m=15.030 σ=0.067	m=15.047 σ=0.059
CO2mes	m=0.042 σ=0.003	m=0.037 σ=0.003	m=0.034 σ=0.004
O2mes	m=20.827 σ=0.045	m=20.833 σ=0.047	m=20.830 σ=0.075
Action-chaud	m=1360.190 σ=542.968	m=2135.152 σ=23.500	m=2409.306 σ=82.511

Tableau 7. (suite)

État / Paramètre	État 1 : fermentation	État 2 : Diauxie	État 3 : Oxydation
Action-froid	m=204.897 $\sigma=0.929$	m=205.000 $\sigma=0$	m=4347.907 $\sigma=4988.585$
Ajout-base	m=150.330 $\sigma=150.138$	m=437.636 $\sigma=3.102$	m= 736.956 $\sigma=274.886$
U-volt(ppe)	m=4.500 $\sigma=0.000$	m=4.500 $\sigma=0.000$	m=4.231 $\sigma=0.989$
Qair-sortie2	m=12.884 $\sigma=5.581$	m=15.121 $\sigma=0.041$	m=14.877 $\sigma=0.398$
Lum	m=69.109 $\sigma=0.996$	m=67.806 $\sigma=0.098$	m=67.180 $\sigma=3.447$
a	m=-1.641 $\sigma=0.049$	m=-1.659 $\sigma=0.047$	m=-1.667 $\sigma=0.206$
b	m=7.634 $\sigma=0.428$	m=8.234 $\sigma=0.061$	m=8.373 $\sigma=1.227$
delta	m=14.583 $\sigma=0.694$	m=13.679 $\sigma=0.073$	m=13.191 $\sigma=2.415$
X-estim	m=0.727 $\sigma=0.276$	m=1.096 $\sigma=0.019$	m=1.338 $\sigma=1.167$
rO2	m=1.659 $\sigma=1.648$	m=6.745 $\sigma=0.457$	m=19.548 $\sigma=12.394$
rCO2	m=22.241 $\sigma=20.641$	m=6.278 $\sigma=0.736$	m=13.857 $\sigma=7.687$
Q-resp	m=17.463 $\sigma=22.259$	m=0.928 $\sigma=0.058$	m=0.746 $\sigma=0.093$

Tableau 8. Évaluation du manque de pertinence des paramètres biochimique durant le bioprocédé de type batch.

Paramètre	Intervalle de non pertinence (en heure)
Température	[8.05-16.5]; [18.7-20]
pH	[2.71-2.86]; [5.40-5.45] [7.04-7.94] [8-16.2]; [18.5-20]
Agitation	[5.58-5.71-5.75]; [5.83-5.85]; [6.03-6.06] [6.76-6.89]; [7.28-7.46]; [7.56-7.81] [8.11-13]; [15.6-16.20]; [18.7-20]
pO2	[3.48-3.78]; [5.09-5.14]; [7.61-8] [8.06-13]; [15.6-16.2]; [18.7-20]
Pression	[8.06-20]
Qair-entrée	[6.41-6.44]; [8-8.5]; [18.6-20]
Qair-sortie	[7.63-7.81]; [7.88-7.94]; [8.05-20]
CO2mes	[1.43-1.51]; [7.61-13]; [13.1-13.7] [15.6-16.2]; [18.7-20]
O2mes	[7.61-8]; [8.06-20]
Action-chaud	[7.35-13]; [15.6-15.9]; [18.7-20]
Action-froid	[7.61-8]; [8.06-16.6]; [18.7-20]
Ajout-base	[7.21-16.2]; [18.9-20]
U-volt(ppe)	[7.61-7.94]; [8.06-19.3]; [19.5-20]
Qair-sortie2	[1.95-1.98]; [4.98-5.03]; [7.61-7.94] [8.07-12.9]; [15.6-16.2]; [18.6-20]
Lum	[6.54-13]; [13.7-13.8]; [15.6-16.2] [18.7-20]
a	[0-2.06]; [2.33-2.56]; [2.66-2.73] [2.93-2.99]; [7.28-20]
b	[6.29-6.34]; [6.41-13]; [15.6-16.2]; [18.7-20]
X-estim	[6.59-13]; [15.6-16.2]; [18.7-20]
rO2	[7.61-7.66]; [7.76-7.78]; [7.86-12.9] [15.6-16.2]; [18.7-20]
rCO2	[2.15-2.45]; [2.71-2.79]; [2.83-3.83] [7.61-7.83]; [7.86-7.91]; [8.06-12.9] [15.6-16.9]; [18.7-20]
Q-resp	[7.61-7.81]; [7.86-7.91]; [8.05-13] [15.6-16.2]; [18.7-20]

Tableau 9. Évaluation du manque pertinence des paramètres biochimique durant le bioprocédé de type fed-batch

Paramètres	Intervalle de non pertinence (en heure)
T	[21,48-21,53]
pH	Aucun
Agitation	[0.21-1.33]; [6.90-9.43]; [21.45-21.60]
Antimousse	20.82; [20.99-21.60]
Conso. pompe 1	Aucun
Conso. Pompe sel	Aucun
% N <sub>2</sub>	Aucun
% O <sub>2</sub>	Aucun
%CO <sub>2</sub>	Aucun
Éthanol	[19.82-19.90]; [20.04-21.55]
Volume base	[0.99-1.33]; [3.98-4.21] [10.91-10.95]; [7.55-9.47] [21.50-21.52]; [21.57-21.58]
Oxygène dissous	Aucun



## Références

- [1] le Grand Dictionnaire Terminologique. <http://www.granddictionnaire.com>. Office Québécois de la langue française.
- [2] le Trésor de la Langue Française Informatisé. <http://atilf.atilf.fr/tlf.htm>. CNRS-ATILF.
- [3] Dictionnaire Alphabétique et Analogique de la Langue Française, 1969. Ed. Le Robert.
- [4] A. APPRIOU. *Décision et Reconnaissance des formes en signal*. Hermes Sciences, 2002.
- [5] A. AYOUN and P. SMETS. Data association in multi-target detection using the transferable belief model. *International Journal of Intelligent Systems*, 16:1167-1182, 2001.
- [6] S. BALUJA and D. POMERLEAU. Dynamic relevance: vision-based focus attention using artificial neural networks. *Artificial Intelligence*, 97:381-395, 1997.
- [7] J.-P. BARTHÉLEMY and E. MULLET. Choice basis: A model for multiattribute preferences. *British J. of Math. And Stat. Psychology*, 39:106-124, 1986.
- [8] I. BLOCH. Fusion d'informations numériques : panorama méthodologique. In *JNRR'05*, 2005.
- [9] I. BLOCH and A. HUNTER. Fusion: General concepts and characteristics. *International Journal of Intelligent Systems*, 16:1107-1134, Eds 2001.
- [10] A. BLUM and P. LANGLEY. Selection of relevant features and examples in machines learning. *Artificial Intelligence*, 97:245-271, 1997.
- [11] P. BORLUND. The concept of relevance in IR. *Journal of the American Society for Information Science and Technology*, 54(10):913-025, 2003.
- [12] H. BRACKER. *Utilisation de la théorie de Dempster-Shafer pour la classification d'images satellitaires à l'aide de données multi-sources et multi-temporelles*. Thèse de Doctorat, Ecole Nationale des Télécommunications de Bretagne, 1996.
- [13] X.-W. CHEN. An improved branch and bound algorithm for feature selection. *Pattern Recognition Letters*, 24:1925-1933, 2003.
- [14] A. CORD, C. AMBROISE, and J.-P. COCQUEREZ. Feature selection in robust clustering based on laplace mixture. *Pattern Recognition Letters*, 27:627-635, 2006.
- [15] A.P. DEMPSTER. A generalisation of bayesian inference. *Journal of the Royal Statistical Society*, 30:205-247, 1968.
- [16] T. DENOEU. A k-nearest neighbor classification rule based on Dempster-Shafer theory. *IEEE trans. on systems, man, and cybernetics*, 25(5):804-813, Mai 1995.
- [17] J. DESACHY, L. ROUX, and E.-H. ZAHZAH. Numeric and symbolic data fusion: a soft computing approach to remote sensing images analysis. *Pattern Recognition Letters*, 17:1361-1378, 1996.
- [18] J. DEZERT and F. SMARANDACHE. On the generation of hyper-powersets for the dezert-smarandache thoery. In *Fusion 2003*, pages 1118-1125, 2003.
- [19] J. DEZERT and F. SMARANDACHE. *Advances and Application in Dezert-Smarandache Theory*, chapter Combining Uncertain and Paradoxical Evidences in Dezert-Smarandache Thoery, 2004.
- [20] J. DEZERT and F. SMARANDACHE. The generalized pignistic transformation. In *Fusion 2004*, Stockholm Suède, 2004.
- [21] D. DUBOIS and H. PRADE. On the relevance of non-standard theories of uncertainty in modeling and pooling expert opinions. *Reliability Engineering and System Safety*, 36(2), 1992.
- [22] D. DUBOIS and H. PRADE. Possibility theory and data fusion in poorly informed environments. *Control Engineering Practice*, 2(5):811-823, 1994.
- [23] R. FELIX. Relationships between goals in multiple attribute decision making. *Fuzzy Sets and Systems*, 67:47-52, 1994.
- [24] G. GOMA, J.-L. URIBELARREA, V. GUILLOUET, and C. JOUVE. Tackling complexity in industrial microbiology for bioprocess. In *4th International Congress on Bioprocess in Food Industries*, Clermont-Ferrand, 11-13 Juillet 2004.
- [25] R. GREINER, A.J. GROVE, and A. KOGAN. Knowing what doesn't matter: exploiting the omission of irrelevant data. *Artificial Intelligence*, 97:345-380, 1997.
- [26] H. GREISDORF. Relevance: an interdisciplinary and information science perspective. *Informing Science*, 3(2):67-72, 2000.
- [27] R. HAENNI. Are alternatives to Dempster's rule of combination real alternatives? comments on "About belief function combination and conflict management problem"-Lefevre et al. *Information Fusion*, 3:237-239, 2002.
- [28] M. KUDO and J. SKLANSKY. Comparison of algorithms that select features for pattern classifiers. *Pattern Recognition*, 33:25-41, 2000.
- [29] L. LARDON. *Représentation et gestion des incertitudes pour le diagnostic par la théorie de Dempster-Shafer: application aux procédés biologiques*. Thèse de Doctorat, Ecole Nationale Supérieure Agronomique de Montpellier, Montpellier, 2004.
- [30] L. LARDON, A. PUNAL, and J.-P. STEYER. On-line diagnostic and uncertainty management using evidence theory-experimental illustration to anaerobic digestion processes. *Journal of Process Control*, 14:747-763, 2004.
- [31] M. LAZO-CORTÈS and J. RUIZ-SCHULCLOPER. Determining the feature relevance for non-classically described objects and a new algorithm to compute typical fuzzy testors. *Pattern Recognition Letters*, 16:1259-1265, 1995.
- [32] E. LEFEVRE, O. COLOT, and P. VANNOORENBERGHE. Belief function combination and conflict management. *Information Fusion*, 3:149-162, 2002.
- [33] E. LEFEVRE, O. COLOT, and P. VANNOORENBERGHE. Reply to the comments of R. Haenni on the paper "Belief function combination and conflict management". *Information Fusion*, 4:63-65, 2003.
- [34] L. MASCARILLA. *Apprentissage de connaissance pour l'interprétation d'images satellite*. Thèse de Doctorat, Univerisité P. Sabatier, Toulouse, 1996.
- [35] H. MONTGOMERY. *Analysing and Aiding Decision Processes*, chapter Decision rules and the search for a dominance structure :

- towards a process model of decision making, pages 343-369. North-Holland, 1983.
- [36] P. NARENDRA and K. FUKUNAGA. A branch and bound algorithm for feature subset selection. *IEEE Transactions on Computers*, 26(9):917-922, 1977.
- [37] N. PIERA-CARRETÉ, P. DESROCHES, and J. AGUILAR-MARTIN. Lamda: an incremental conceptual clustering system. Technical report, LAAS-CNRS, 1988.
- [38] B.W. PORTER, R. BAREISS, and R.C. HOLTE. Concept learning and heuristic classification in weak-theory domains. *Artificial Intelligence*, 45(1-2):229-263, 1990.
- [39] P. PUDIL, F. FERRI, J. NOVOVICOVA, and J. KITTLER. Floating search methods in feature selection. *Pattern Recognition Letters*, 15:1119-1125, 1994.
- [40] S. RÉGIS. *Segmentation, classification, et fusion d'informations de séries temporelles multi-sources: application à des signaux dans un bioprocédé*. Thèse de Doctorat, Université des Antilles et de la Guyane, Novembre 2004.
- [41] S. RÉGIS, A. DONCESCU, J.-P. ASSELIN DE BEAUVILLE, and J. DESACHY. Evaluation de la pertinence de paramètres biochimiques et classification pour la caractérisation des états physiologiques dans un bioprocédé par la théorie de l'évidence. *Revue des Nouvelles Technologies de l'Information*, C-1, 2004.
- [42] S. RÉGIS, L. FAURE, A. DONCESCU, J.-L. URIBELARREA, L. MANYRI, and J. AGUILAR-MARTIN. Adaptive physiological states classification in fed-batch fermentation process. In *IFAC CAB'9*, Nancy, France, Mars 2004.
- [43] J.A. ROELS. *Energetics and kinetics in biotechnology*, chapter Macroscopic theory and microbial growth and product formation, pages 23-73. Elsevier Biomedical Press, 1983.
- [44] L. ROUX. Auto-adaptive information-fusion for satellite images classification. In *Satellite Remote Sensing 2*, Paris, Septembre 1995.
- [45] E. LE SAUX. *Extraction et Implémentation de stratégies expertes. Application à des processus Industriels*. Thèse de Doctorat, EHES, 2000.
- [46] J. SCHUBERT. On non specific evidence. *International Journal of Intelligent Systems*, 8:711-725, 1993.
- [47] J. SCHUBERT. Finding a posterior domain probability distribution by specifying nonspecific evidence. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 3:163-185, 1995.
- [48] J. SCHUBERT. Specifying nonspecific evidence. *International Journal of Intelligent systems*, 11:525-563, 1996.
- [49] G. SHAFER. *A Mathematical Theory of Evidence*. Princeton University Press, New Jersey, 1976.
- [50] J. SHANTEAU. Competence in experts: The role of tasks characteristics. *Organizational behavior and human decision processes*, 1992.
- [51] J. SHANTEAU. How much information does an expert use? is it relevant? *Acta Psychologica*, pages 75-86, 1992.
- [52] J. SIEDLECKI and J. SKLANSKY. A note on genetic algorithm for large-scale feature selection. *Pattern Recognition Letters*, 10:335-347, 1989.
- [53] H. SIMON. *Models of Thought*. Yale University Press New Haven, 1979.
- [54] P. SMETS. *Non standard Logics for Automated Reasoning*, chapter Belief Functions, pages 29-39. Academic Press, 1988.
- [55] P. SMETS. The combination of evidence in the transferable belief model. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, (12):447-458, 1990.
- [56] P. SMETS and R. KENNES. The transferable belief model. *Artificial Intelligence*, 66:191-234, 1994.
- [57] J.P. STEYER. *Sur une approche qualitative des systèmes physiques : aide en temps réel à la conduite des procédés fermentaires*. Thèse de Doctorat, Université Paul Sabatier, Toulouse France, Décembre 1991.
- [58] J. WAISSMAN-VILANOVA. *Construction d'un modèle comportemental pour la supervision de procédés: application à une station de traitement des eaux*. Thèse de Doctorat, LAAS – CNRS, Novembre 2000.
- [59] R. YAGER. On the Dempster-Shafer framework and new combination rules. *Information Sciences*, 41:93-138, 1987.
- [60] L. ZADEH. A mathematical theory of evidence (book review). *AI magazine*, 5(3):81-83, 1984.
- [61] L. ZADEH. A note on web intelligence, world knowledge and fuzzy logic. *Data and Knowledge Engineering*, 50:291-304, 2004.





Sébastien Régis

Sébastien Régis a effectué des études de mathématiques à l'Université des Antilles et de la Guyane à Pointe-à-Pitre et à l'Université Joseph Fourier de Grenoble. Il obtient en 1999 un diplôme de Master en Ingénierie Mathématique de l'Université Joseph Fourier et devient docteur en Informatique de l'Université des Antilles et de la Guyane en 2004. Il est actuellement Attaché Temporaire à l'Enseignement et à la Recherche au sein de l'Université de Antilles et de la Guyane dans le laboratoire GRIMAAG. Ses travaux de recherche concernent le traitement du signal, la classification et la fusion d'information pour les séries temporelles issues des sciences expérimentales comme les biotechnologies ou la physique.



Jacky Desachy

Jacky Desachy est professeur des universités, directeur du laboratoire GRIMAAG au sein de l'Université des Antilles et de la Guyane. Ses axes de recherche concernent la vision par ordinateur, la télédétection, l'analyse et l'interprétation d'images, la reconnaissance des formes et l'intelligence artificielle. Il s'intéresse aussi à l'utilisation de connaissances, de données exogènes et de la logique floue pour l'interprétation d'images et à la fusion de sources d'information. Il est membre de divers comités scientifiques et d'organisations de conférences dans le domaine de l'image et de la télédétection et est relecteur de différentes conférences et revues dans le même domaine.



Andrei Doncescu

Andrei Doncescu a obtenu son diplôme d'ingénieur en électronique de l'Ecole Polytechnique de Bucarest en 1991 puis un DEA d'Informatique de l'Université de Poitiers en 1993. Docteur en Informatique de l'Université de La Rochelle depuis 1998, il est maître de conférences en informatique à l'Université Paul Sabatier de Toulouse et effectue ses travaux de recherche au sein du LAAS-CNRS de Toulouse. Ses axes de recherche sont le traitement du signal et des images, la fusion et la classification appliqués notamment aux biotechnologies et à la biologie.