

Diagnostic de la pollution atmosphérique par une approche RDF utilisant les fonctions de croyance

Patrick VANNOORENBERGHE et Thierry DENŒUX

Laboratoire HEUDIASYC, UMR 6599 CNRS

Université de Technologie de Compiègne

BP 20529, 60205 Compiègne cedex

Tel : (+33) 3.44.23.49.53, Fax : (+33) 3.44.23.44.77

Patrick.Vannoorenberghe@hds.utc.fr

Thierry.Denoeux@hds.utc.fr

Résumé

Dans cet article, nous proposons une approche du type « Reconnaissance des Formes » utilisant la théorie des croyances de Dempster-Shafer au problème de diagnostic de la pollution atmosphérique. A partir d'un formalisme théorique initialement introduit par Dempster [1] lors de ses travaux sur les bornes inférieure et supérieure d'une famille de distributions de probabilités, Shafer [2] a mis en évidence l'intérêt des fonctions de croyance pour la modélisation de connaissances incertaines. Cet outil mathématique est de nos jours reconnu pour sa grande capacité à modéliser les incertitudes via notamment le « Modèle des Croyances Transférables » introduit par Smets [3]. L'emploi de ce cadre théorique pour le diagnostic environnemental est motivé par la forte spécificité des données qui se caractérisent généralement par un nombre élevé de valeurs manquantes, de grandes imprécisions sur les mesures et une fiabilité limitée des capteurs de pollution. Cet article se focalise sur les méthodes de construction des fonctions de croyance ainsi que sur l'optimisation des paramètres d'apprentissage. L'utilisation de ces techniques est illustrée sur des données issues du réseau de surveillance de l'agglomération lyonnaise pour la détection de pic de pollution atmosphérique à l'ozone.

1 Position du problème

La place croissante occupée par les problèmes environnementaux dans notre société (gestion des ressources en eau, loi sur l'Air) et le poids économique important de l'industrie de l'environnement (élimination des déchets) ouvrent aujourd'hui un nouveau champ d'application aux techniques de diagnostic. Surveiller l'état de l'environnement pose des problèmes spécifiques liés aux difficultés rencontrées dans l'acquisition des informations (les informations délivrées par les capteurs étant imprécises, incertaines et incomplètes) et la modélisation des phénomènes [4]. En effet, les modèles analytiques sont souvent sujets à des hypothèses simplificatrices qui ne permettent pas de prendre en compte la complexité des phénomènes physiques (recombinaison de polluants dans l'atmosphère par exemple). Ces différents problèmes tendent à favoriser une approche du type « Reconnaissance De Formes » (RDF) au problème de la surveillance et du diagnostic [5]. L'approche RDF du diagnostic consiste à construire une fonction de décision par l'apprentissage d'un historique de mesures \mathbf{x}^i , $i = 1, \dots, n$ effectuées sur le système. Ces méthodes, dites "sans modèles", sont basées sur des techniques d'analyse de données et d'apprentissage statistique. Elles peuvent s'adapter aux données imparfaites et permettre ainsi de :

- modéliser les informations imprécises et incertaines sur les capteurs (par les fonctions de croyance [6] ou les distributions de possibilité par exemple [7]),
- rendre compte de l'évolution du contexte dans le cadre de l'apprentissage.

Dans le cadre de la surveillance de l'environnement, une telle prise en compte de ces informations permet dans un premier temps d'augmenter les performances du système de diagnostic [8, 9, 10] mais aussi de fournir une information plus riche sur la prédiction (intervalles de confiance, possibilité de rejet, ...).

Dans cet article, on propose d'utiliser une approche du type RDF utilisant la théorie des fonctions de croyance pour le problème de diagnostic de la pollution atmosphérique. La première section de ce papier permet d'appréhender les aspects spécifiques liés aux données dans le domaine de l'environnement (section 2) notamment dans le cadre de l'application envisagée. Quelques notions théoriques sur la théorie de l'évidence sont ensuite introduites à la section 3. La section 4 présente la méthodologie proposée ainsi que la description des différents modèles qui permettent de modéliser les informations disponibles sous forme de fonctions de croyance. L'optimisation des paramètres liés à ces fonctions est également abordée. Enfin, on présente une réalisation concrète sur des données réelles issues du réseau de surveillance de l'agglomération lyonnaise, visant la prédiction du pic d'ozone (cf. section 5).

2 Les problèmes liés aux données environnementales

Le problème envisagé concerne le diagnostic précoce de l'évolution de la pollution atmosphérique vers une situation correspondant à un événement anormal. On entend par « événement » une situation qui peut être profondément différente selon l'application envisagée. Par exemple, l'apparition d'un pic d'ozone peut survenir avec un niveau de seuil particulier [10]. Dans ce cadre, nous nous sommes focalisés sur le réseau de surveillance de la ville de Lyon. COPARLY (Comité de Coordination pour le Contrôle de la Pollution atmosphérique en région Lyonnaise), le nom donné au réseau de surveillance de la pollution sur l'agglomération lyonnaise, couvre une zone d'environ 30×30 km. Le réseau est constitué de 49 sites de mesures permettant un échantillonnage de différentes concentrations de polluants. Plus précisément, 17 capteurs permettent une mesure horaire du monoxyde d'azote (NO) et du dioxyde d'azote (NO₂) et 7 capteurs mesurant la concentration en ozone (O₃). La température est également mesurée chaque heure et est prédite à un horizon de 24 heures. Enfin, une prévision journalière des conditions atmosphériques sur 36 heures est réalisée par l'intermédiaire d'un indice climatologique, appelé géopotential, proposé par P. Benichou à Météo-France [11]. Nous disposons d'un ensemble de 897 jours de données réparties sur les périodes du 1 Avril au 30 Septembre des années 1994 à 1998.

Les informations délivrées par les capteurs de pollution sont de nature imparfaite c'est-à-dire imprécises, incertaines et incomplètes. En effet, une première imperfection réside dans l'incomplétude des bases de données. A titre d'exemple, sur les mesures de concentration en ozone, 24% des données sont manquantes! De surcroît, la pollution étant un phénomène difficilement modélisable, il s'avère que le contexte du problème posé évolue en fonction du temps. A titre d'exemple, on a représenté sur la figure 1 l'évolution de la concentration moyenne journalière en ozone sur un capteur particulier durant l'année 1995. Des statistiques annuelles sur ce capteur sont représentées sur la partie droite de la figure 1 sous forme de boîte à moustaches. On s'aperçoit que ces moyennes varient d'une année sur l'autre rendant le problème de l'apprentissage d'autant plus délicat à appréhender. De plus, le problème étudié concerne essentiellement la détection d'événements rares. En effet sur les 897 jours de données observés, seulement 58 jours sont considérés comme des situations anormales représentant uniquement 6.3% des données. Ces constatations tendent à favoriser l'emploi de techniques appropriées aux traitements de ce type de données comme la théorie des fonctions de croyance dont nous rappelons quelques fondements théoriques dans la section suivante.

3 Aspects théoriques relatifs aux fonctions de croyance

Dans cette partie, un certain nombre de concepts concernant la théorie de l'évidence de Demspter-Shafer [2] sont rappelés et permettent ainsi de définir l'ensemble des notations utilisées dans cet article. Soit $\Omega = \{\omega_q, q =$

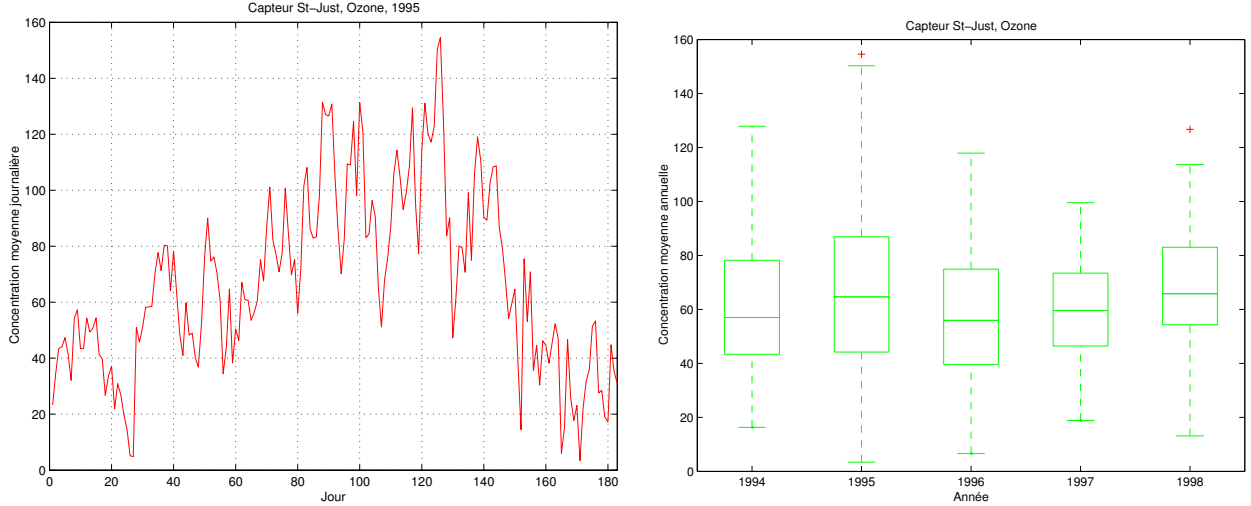


FIG. 1 – Concentration moyenne journalière en ozone (gauche) pour l'année 1995 et concentration moyenne annuelle (droite) sur le capteur St-Just

$1, \dots, Q\}$ un ensemble fini de valeurs possibles d'une variable d'intérêt y . Une masse de croyance m sur Ω est une fonction :

$$m : 2^\Omega \rightarrow [0, 1] \quad (1)$$

qui vérifie les propriétés suivantes : $m(\emptyset) = 0$ et $\sum_{A \subseteq \Omega} m(A) = 1$. Cette masse de croyance se différencie d'une probabilité au sens classique du terme par le fait que la totalité de la masse de croyance est répartie non seulement entre les hypothèses singletons ω_q mais aussi entre les hypothèses combinées $A \subseteq \Omega$. Chaque sous-ensemble $A \subseteq \Omega$ tel que $m(A) > 0$ est appelé élément focal de m . A partir de m , une fonction de croyance bel et une fonction de plausibilité pl peuvent être définies respectivement par $bel(A) = \sum_{B \subseteq A} m(B)$ et $pl(A) = \sum_{A \cap B \neq \emptyset} m(B)$. La quantité $bel(A)$ peut être interprétée comme un degré de croyance en l'hypothèse A . La plausibilité $pl(A)$ est vue comme la quantité totale de croyance qui peut être potentiellement placée sur A . Une fonction de croyance affaiblie $m_\alpha(\cdot)$ peut être obtenue à partir de m par l'équation suivante :

$$m_\alpha(A) = \alpha m(A) \quad \forall A \subseteq \Omega, A \neq \Omega \quad (2)$$

$$m_\alpha(\Omega) = 1 - \alpha + \alpha m(\Omega) \quad (3)$$

avec $0 \leq \alpha \leq 1$. Cette opération d'affaiblissement est généralement utilisée quand la source d'information n'est pas totalement fiable. Dans ce cas, le coefficient d'affaiblissement α représente une sorte de méta-connaissance sur la fiabilité du capteur qui n'est pas prise en compte directement par la fonction de croyance m .

Dans le cadre de la théorie de l'évidence de Dempster-Shafer, la fusion des informations issues de sources distinctes est réalisée en utilisant la *loi de combinaison de Dempster*. Deux corpus d'évidence m_1 et m_2 peuvent ainsi être fusionnés avec cette règle (somme orthogonale \oplus), conduisant ainsi à une fonction de croyance unique définie comme :

$$m(A) = \frac{\sum_{B \cap C = A} m_1(B) m_2(C)}{\sum_{B \cap C \neq \emptyset} m_1(B) m_2(C)} \quad \forall A \subseteq \Omega. \quad (4)$$

Pour le problème de la prise de décision, on suppose que nous avons une fonction de croyance m sur Ω qui résume l'ensemble des informations apportées sur la valeur de la variable y . La décision consiste à choisir une action a parmi un ensemble fini d'actions \mathcal{A} . Une fonction de perte $\lambda : \mathcal{A} \times \Omega \rightarrow \mathbb{R}$ est supposée définie de telle manière que $\lambda(a, \omega)$ représente la perte encourue si l'on choisit l'action a lorsque $y = \omega$. A partir d'arguments de rationalité, Smets [3] propose de transformer m en une fonction de probabilité p_m sur Ω (appelée fonction

de probabilité pignistique) définie pour tout $\omega \in \Omega$ comme : $p_m(\omega) = \sum_{A \ni \omega} \frac{m(A)}{|A|}$, où $|A|$ est la cardinalité de $A \subseteq \Omega$. Dans cette transformation, la masse de croyance $m(A)$ est uniformément distribuée parmi les éléments de A . A partir de cette probabilité, on peut associer à chaque action $a \in \mathcal{A}$ un risque défini comme le coût espéré relatif à p_m si on choisit a :

$$R(a) = \sum_{\omega \in \Omega} \lambda(a, \omega) p_m(\omega). \quad (5)$$

L'idée consiste ensuite à choisir l'action qui minimise ce risque, généralement appelé risque pignistique. En reconnaissance des formes, $\Omega = \{\omega_1, \dots, \omega_Q\}$ est l'ensemble des classes et les éléments de \mathcal{A} sont généralement les actions a_q qui consistent à assigner le vecteur inconnu à la classe ω_q . Avec des coût 0-1 ($\lambda(a_q, \omega_r) = 1 - \delta_{q,r}$ pour $q, r \in \{1, \dots, Q\}$), il est démontré [12] que la minimisation du risque pignistique R conduit à choisir la classe ω_0 de plus grande probabilité pignistique. Si une action supplémentaire de rejet a_0 de coût constant λ_0 est possible, alors le vecteur est rejeté si $p_m(\omega_0) < 1 - \lambda_0$.

4 Approche proposée

De manière à assurer une gestion des incertitudes liées aux données environnementales, nous proposons une approche basée sur l'utilisation des fonctions de croyance [2]. Cet outil théorique est désormais reconnu comme un cadre formel permettant de représenter et manipuler les incertitudes tout en trouvant des champs d'application divers et variés (diagnostic médical, traitement d'images, fusion multi-capteurs, ...). L'approche proposée dans ce papier adopte le point de vue du « Modèle des Croyances Transférables » qui permet de distinguer deux niveaux de représentation de l'information : le niveau credal et le niveau pignistique [3]. Pour obtenir les fonctions de croyance à partir des données d'apprentissage, deux familles de techniques sont généralement utilisées : les méthodes basées sur la vraisemblance qui utilisent l'estimation des densités et une méthode basée sur la distance dans laquelle les jeux de masses sont construits directement à partir des distances aux vecteurs d'apprentissage. Dans cet article, nous avons choisi de comparer ces techniques en se focalisant sur l'optimisation des coefficients de confiance qui permettent de quantifier la fiabilité des capteurs. Nous présentons brièvement les deux approches de construction des fonctions de croyance ainsi que l'optimisation des paramètres. Enfin, la méthodologie employée dans le cadre de notre application est présentée au paragraphe 4.4.

Dans la suite cet article, on suppose que le problème général de la surveillance de la pollution peut être posé sous la forme d'un problème de discrimination (RDF) où les états à diagnostiquer sont inventoriés dans l'ensemble Ω . Dans le cadre de la prédiction du pic d'ozone, l'événement à diagnostiquer est produit sous forme binaire comme la prédiction de l'absence ou de la présence d'un pic à un pas de temps donné. Dans l'application envisagée, on est donc en présence qu'un cadre de discernement à deux hypothèses ($Q = 2$) où les classes ω_1 et ω_2 représentent respectivement la prédiction au jour J de l'absence ou la présence d'un pic d'ozone au jour $J+1$.

4.1 Méthodes basées sur la vraisemblance

On suppose $f(\mathbf{x}|\omega_q)$ connues les densités de probabilités conditionnellement aux classes. En ayant observé \mathbf{x} , la fonction de vraisemblance $L(\omega_q|\mathbf{x})$ est une fonction de Ω dans $[0, +\infty)$ définie par $L(\omega_q|\mathbf{x}) = \mathbf{f}(\mathbf{x}|\omega_q)$, pour tout $q \in \{1, \dots, Q\}$. A partir de L , Shafer [2, p.238] a proposé de construire une fonction de croyance sur Ω définie par sa fonction de plausibilité comme :

$$pl(A) = \frac{\max_{\omega_q \in A} [L(\omega_q|\mathbf{x})]}{\max_q [L(\omega_q|\mathbf{x})]} \quad \forall A \subseteq \Omega. \quad (6)$$

En reconnaissances des formes, une application de cette méthode a été utilisée dans [13]. Remarquons que la fonction pl définie à l'équation (6) est consonante, c'est-à-dire que ses éléments focaux sont emboîtés. Pour cette raison, ce premier modèle serait appelée « Modèle Consonant ».

A partir de considérations axiomatiques, Appriou [14] a proposé une autre méthode basée sur la construction de Q fonctions de croyance $m_q(\cdot)$. L'idée consiste à prendre en compte de manière séparée chaque classe et à évaluer le degré de croyance accordée à chacune d'entre elles. Dans ce cas, les éléments focaux de chacune des fonctions de croyance m_q sont les singletons $\{\omega_q\}$, les sous-ensembles complémentaires $\overline{\omega_q}$ et Ω . Appriou obtient ainsi deux modèles différents :

Modèle 1 :

$$m_q(\{\omega_q\}) = \alpha_q \frac{R \cdot L(\omega_q | \mathbf{x})}{1 + R \cdot L(\omega_q | \mathbf{x})} \quad (7)$$

$$m_q(\overline{\omega_q}) = \alpha_q \frac{1}{1 + R \cdot L(\omega_q | \mathbf{x})} \quad (8)$$

$$m_q(\Omega) = 1 - \alpha_q, \quad (9)$$

Modèle 2 :

$$m_q(\{\omega_q\}) = 0 \quad (10)$$

$$m_q(\overline{\omega_q}) = \alpha_q (1 - R \cdot L(\omega_q | \mathbf{x})) \quad (11)$$

$$m_q(\Omega) = 1 - \alpha_q (1 - R \cdot L(\omega_q | \mathbf{x})). \quad (12)$$

Dans ces équations, α_q est un coefficient qui peut être utilisé pour modéliser une information complémentaire (comme par exemple la fiabilité d'un capteur), et R est une constante de normalisation qui est choisie dans la plage $]0, (\max_q(L(\omega_q | \mathbf{x})))^{-1}[$. En pratique, les performances de ces deux modèles semblent être quasi-équivalentes [15]. Cependant, Appriou (communication personnelle) recommande l'utilisation du modèle 2 qui a l'avantage d'être consistant avec le « théorème de Bayes généralisé » proposé par Smets [16]. Ce modèle, que nous appellerons « Modèle Séparable » dans la suite de ce papier, a été utilisé dans nos simulations dans sa forme la moins spécifique c'est-à-dire avec la plus grande valeur autorisée pour la constante R . A partir de ces Q fonctions de croyance et en utilisant la règle de combinaison de Dempster, une fonction de croyance unique m est obtenue $m = \bigoplus_q m_q$.

4.2 Méthode basée sur la distance

Une approche radicalement différente a été introduite par Denoeux [17]. Dans cette méthode, une fonction de croyance est directement construite par le biais des vecteurs d'apprentissage situés au voisinage du vecteur à classer. Si les k plus proches voisins (au sens de la distance euclidienne par exemple) sont considérés, nous obtenons ainsi k fonctions de croyance qui sont alors combinées par l'opérateur de Dempster. La méthode initiale a ensuite été améliorée en optimisant les paramètres [18], et une version neuronale a été récemment introduite [19]. Cette version utilise un ensemble de prototypes qui sont déterminés par minimisation d'une fonction d'erreur. Chaque prototype est alors considéré comme un corpus d'évidence qui influence la croyance dans l'appartenance de \mathbf{x} à une des classes. Une fonction de croyance m^i associée à chaque prototype i est ainsi définie par :

$$\forall q \in \{1, \dots, Q\} \quad m^i(\{\omega_q\}) = \alpha^i \phi^i(d^i) \quad (13)$$

$$m^i(\Omega) = 1 - \alpha^i \phi^i(d^i) \quad (14)$$

$$m^i(A) = 0 \quad \forall A \in 2^\Omega \setminus \{\{\omega_q\}, \Omega\} \quad (15)$$

où d^i est la distance Euclidienne au i -ème prototype, α^i est un paramètre associé à chaque prototype i et $\phi^i(\cdot)$ est une fonction décroissante définie par $\phi^i(d^i) = \exp[-\gamma^i (d^i)^2]$. Dans cette expression, γ^i est un paramètre positif associé à chaque prototype i . Les fonctions de croyance m^i pour chaque prototype sont ensuite combinées avec

la règle de Dempster. Cette technique de construction de fonctions de croyances sera appelée dans la suite « Modèle Distance ».

4.3 Optimisation des paramètres

Dans la pratique, l'application des méthodes basées sur la vraisemblance comporte une première difficulté qui concerne l'estimation des fonctions de densité associées à chaque classe. Plusieurs techniques d'estimation peuvent être utilisées, parmi lesquelles on retrouve les méthodes paramétriques basées sur un modèle gaussien par exemple, ou les méthodes de noyau non paramétriques. Notre choix, dans le cadre de cette étude, s'est porté sur l'utilisation d'un modèle de mélange gaussien associé à une technique d'estimation des paramètres via l'algorithme EM [20].

Il n'existe pas à notre connaissance de technique générale pour l'évaluation des coefficients d'affaiblissement α_q relatifs au « Modèle Séparable » (ceci est d'ailleurs remarqué par Bastière dans [21]). Dans cet article, nous proposons d'utiliser la même approche que celle introduite par Denceux [19], en minimisant le critère d'erreur suivant :

$$E(\alpha) = \sum_{i=1}^n \sum_{q=1}^Q (p^i(\omega_q) - u_q^i)^2 \quad (16)$$

où u_q^i est l'indicateur d'appartenance du vecteur \mathbf{x}^i ($u_q^i = 1$ si $\omega^i = \omega_q$), et $p^i(\omega_q)$ est la probabilité pignistique de ω_q pour le vecteur \mathbf{x}^i . L'application de ces différentes techniques au problème de l'environnement est présentée dans la section suivante.

4.4 Méthodologie proposée

Comme nous l'avons remarqué précédemment, on adopte dans cette étude le point de vue du Modèle des Croyances Transférables. Ainsi, le premier niveau (« credal level ») s'attache à modéliser l'information apportée par chaque capteur (ou source d'information) S_j sous la forme d'une fonction de croyances m_j . L'ensemble des sources d'informations ainsi modélisées est ensuite agrégé sous la forme d'une fonction de croyances m afin d'obtenir un résumé exhaustif de l'information. Le second niveau (« pignistic level ») concerne la prise de décision à partir de la distribution de probabilité pignistique p déduite de la fonction de croyance m . La figure 2 illustre ces deux niveaux de représentation de l'information.

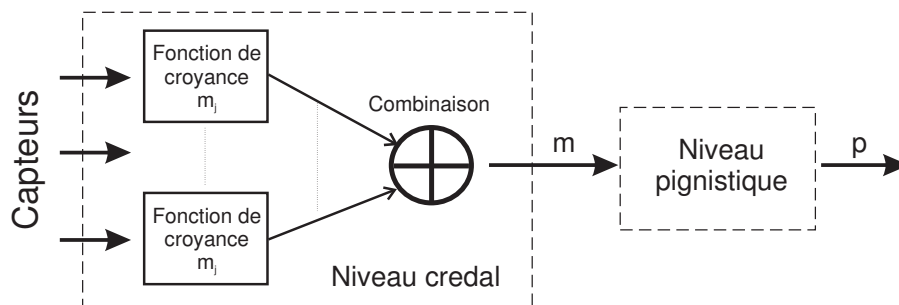


FIG. 2 – Méthodologie proposée

5 Application à la pollution atmosphérique

Même si la méthodologie proposée est suffisamment générique pour être appliquée à n'importe quel problème d'environnement, nous nous sommes focalisés sur le réseau de surveillance de la ville de Lyon. La constitution

des bases de données a été faite de manière chronologique. Les années 1994 à 1997 (488 jours) ont été choisies comme base d'apprentissage. Les bases de validation et de test correspondent respectivement aux années 1997 et 1998 représentant 245 (229 non pollués et 16 pollués) et 164 jours (152 non pollués et 12 pollués). La base d'apprentissage sert à la construction des fonctions de croyance tandis que la base de validation est utilisée pour l'optimisation des coefficients d'affaiblissement. En ce qui concerne les données, nous avons choisi de travailler par classe de géopotential en utilisant la température maximale mesurée au jour J ainsi que 3 capteurs d'ozone (concentrations maximales calculées au jour J).

Aggréger des informations provenant de plusieurs capteurs permet généralement un gain en fiabilité par l'utilisation de la redondance et la complémentarité des informations. Cette technique permet en outre de s'affranchir des données manquantes. La robustesse des modèles vis-à-vis de mesures aberrantes doit être étudiée en détail. La figure 3 propose les jeux de masses obtenus pour les trois modèles pour deux capteurs particuliers. Pour chaque capteur et pour chaque modèle, on a représenté les masses $m(\omega_1)$, $m(\omega_2)$ et $m(\Omega)$. On peut remarquer que les jeux de masses obtenus sont relativement différents. Ainsi, en cas de mesure aberrante ($400\mu g/m^3$ par exemple) pour le modèle basé sur la distance, la majorité de la masse est affectée à l'ensemble Ω ce qui n'est pas le cas pour les modèles basées sur les vraisemblances. En effet, on peut remarquer que pour ces deux modèles, la masse est affectée principalement à la classe ω_2 . Or, la masse $m(\Omega)$ reflète le manque partiel d'information pour la prise de décision. Elle peut être utilisée pour rejeter le vecteur à classer si l'incertitude sur la prédiction est trop grande. Cette propriété est d'un intérêt majeur et permet de conférer au « Modèle Distance » une plus grande robustesse vis-à-vis de mesures aberrantes.

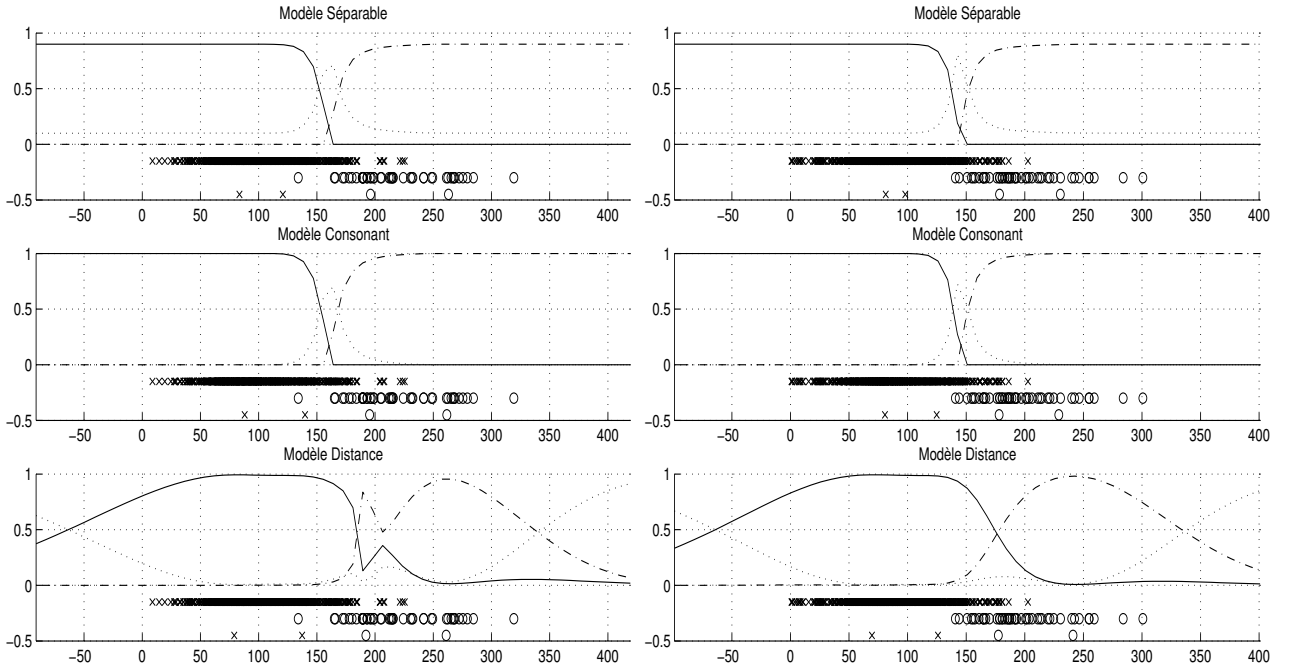


FIG. 3 – Jeux de masses obtenus pour les trois modèles pour deux capteurs différents; $m(\omega_1)$ en trait plein, $m(\omega_2)$ en trait mixte et $m(\Omega)$ en trait pointillé; x : classe ω_1 , o : classe ω_2

Le tableau 1 présente les résultats obtenus par le modèle basé sur les distances sur la base de test. Notons au passage que les paramètres de la méthode sont optimisés par le biais du critère proposé à l'équation (16) et que le nombre de prototypes est optimisé par validation croisée. Au vu de cette matrice de confusion, on peut noter que pour un taux de fausse alarme nul, le taux de non-détection (4 non-détections) reste admissible.

Un des atouts majeurs de la méthodologie développée consiste à étoffer la décision d'un indice de confiance accessible par l'intermédiaire de la fonction de croyance. Cet avantage peut se révéler très utile pour l'exploitation

TAB. 1 – Résultats sur la base de test pour le Modèle Distance

Observés	Prévus	
	Non pollué	Pollué
Non pollué	152	0
Pollué	4	8

du résultat. En effet, l’incertitude sur la prédiction est directement quantifiée par le degré de croyance $m(\Omega)$ accordée à l’ensemble Ω . Cet avantage consiste à pouvoir rejeter le vecteur si l’information disponible n’est pas suffisamment précise [12]. En effet, en supposant que l’action de rejet de coût λ_0 est possible, alors le vecteur \mathbf{x} sera rejeté si $p(\omega_d) < 1 - \lambda_0$ où ω_d est la classe de plus grande probabilité pignistique. La figure 4 représente les probabilités pignistiques pour les trois modèles, pour l’ensemble des vecteurs de la base de test. Sur ce graphique, les erreurs effectuées par le système sont encadrées. On peut y apercevoir également la zone de rejet. En faisant varier le coût λ_0 , il est possible de rejeter un nombre plus ou moins grand de vecteurs de la base de

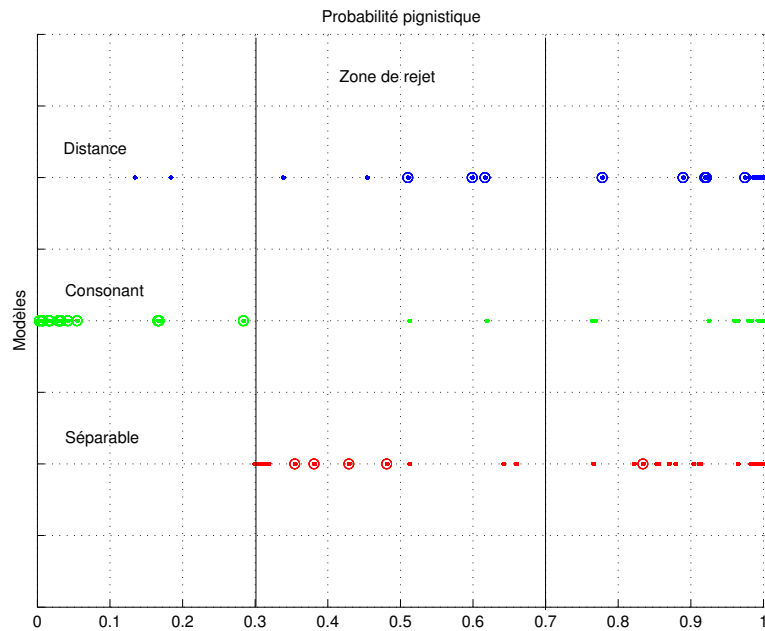


FIG. 4 – Probabilité pignistique pour les trois modèles pour l’ensemble des vecteurs de la base de test

test, et d’obtenir ainsi une décision plus fiable. A titre d’exemple, la figure 5 présente les performances en terme d’erreur en fonction du taux de rejet pour les trois modèles cités.

6 Conclusion

Dans cet article, nous avons proposé une application de la théorie des fonctions de croyance à la détection de situations anormales liées à la surveillance de l’environnement. L’approche présentée consiste à fusionner les informations (malheureusement imparfaites) issues des capteurs de mesures de pollution de manière à prédire l’apparition d’un pic à l’ozone. Ces informations ont été modélisées par des fonctions de croyance permettant ainsi la gestion de leur imprécision et leur incertitude. A ce titre, trois modèles de construction de fonctions de croyance ont été introduits et comparés. L’optimisation des paramètres de ces modèles a également été abordée. La méthodologie précédemment décrite a été appliqué au problème de diagnostic de pollution au pic d’ozone

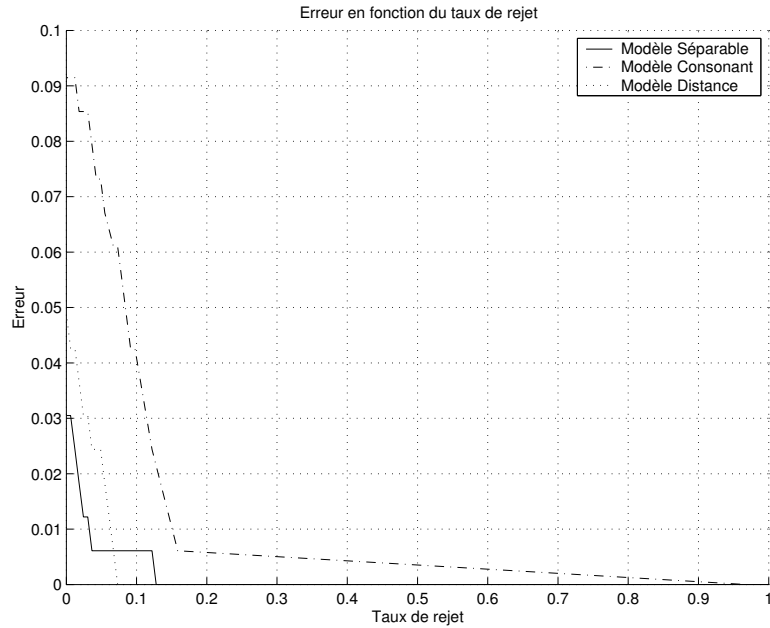


FIG. 5 – Taux d’erreur en fonction du taux de rejet pour les trois modèles

sur la ville de Lyon et permet d’obtenir des résultats satisfaisants de prédiction. Les avantages de la méthode résident principalement dans :

- une gestion des imprécisions et incertitudes relatives aux capteurs de pollution,
- une gestion directe des valeurs manquantes,
- une robustesse vis-à-vis des mesures aberrantes
- ainsi que la richesse de l’information disponible en sortie.

En outre, la méthodologie est suffisamment générique pour être appliquée à n’importe quel problème de diagnostic. Les travaux futurs portent sur une sélection optimale des capteurs comme sources d’informations pour la fusion.

Remerciements

Les auteurs tiennent à remercier l’association COPARLY pour avoir aimablement mis à disposition les informations concernant les mesures de pollution et les données météorologiques. Qu’il soit permis de citer l’Association ECRIN (Echange et Coordination Recherche-INDustrie) qui par l’intermédiaire de son groupe de travail "Prévention des pics de pollution" du Club Ingénierie du Traitement de l’Information propose un lieu d’échange et de rencontres entre différentes communautés (<http://www.ecrin.asso.fr/>). Il est à noter que ces travaux sont co-financés par l’intermédiaire d’un projet de recherche (<http://www.hds.utc.fr/~pvannoo/mese.html>) soutenu par le GDR-PRC ISIS du Centre National de la Recherche Scientifique (CNRS).

Références

- [1] A. Dempster, “Upper and lower probabilities induced by multivalued mapping,” *Annals of Mathematical Statistics*, vol. AMS-38, pp. 325–339, 1967.
- [2] G. Shafer, *A Mathematical Theory of Evidence*. Princeton University Press, 1976.

- [3] P. Smets and R. Kennes, "The Transferable Belief Model," *Artificial Intelligence*, vol. 66, no. 2, pp. 191–234, 1994.
- [4] A. Ionescu, *Contribution à la compréhension de la dispersion des polluants atmosphériques en zone urbaine*. PhD thesis, Université de Paris XII - Val de Marne, 1998.
- [5] T. Denoeux, M. Masson, and B. Dubuisson, "Advanced pattern recognition techniques for system monitoring and diagnosis : A survey," *Journal européen des systèmes automatisés*, vol. 31, no. 9-10, pp. 1509–1539, 1998.
- [6] S. Petit-Renaud, *Application des systèmes flous et de la théorie des croyances à l'estimation fonctionnelle en présence d'informations incertaines ou imprécises*. PhD thesis, Université de Technologie de Compiègne, 1999.
- [7] P. Vannoorenberghe, R. Lherbier, G. Delmaire, and G. Roussel, "Possibilistic identification of pollutant source by multisensor data fusion," in *Air Pollution Modelling and Simulation, APMS'98*, vol. 2, (Champs-Sur-Marne), pp. 767–768, 1998.
- [8] M. Masson, S. Canu, Y. Grandvalet, and A. Lynggaard-Jensen, "Software sensor design based on empirical data," *Ecological Modelling*, vol. 120, pp. 131–139, 1999.
- [9] S. Petit-Renaud and T. Denoeux, "Application de la théorie des fonctions de croyance en régression," in *Rencontres francophones sur la Logique Floue et ses Applications*, (Valenciennes), pp. 169–176, Cépadués, 1999.
- [10] N. Peton, G. Dray, D. Pearson, M. Mesbah, and B. Vuillot, "Modelling and analysis of ozone episodes," in *APMS'98, Air Pollution and Modelling Simulation*, (Champs-sur-Marne), pp. 75–86, 1998.
- [11] P. Benichou, "Classification automatique de configurations météorologiques sur l'europe occidentale," Tech. Rep. Monographie 8, METEO-FRANCE, 1995.
- [12] T. Denoeux, "Analysis of evidence-theoretic decision rules for pattern classification," *Pattern Recognition*, vol. 30, no. 7, pp. 1095–1107, 1997.
- [13] H. Kim and P. Swain, "Evidential reasoning approach to multisource-data classification in remote sensing," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 25, no. 8, pp. 1257–1265, 1995.
- [14] A. Appriou, "Uncertain data aggregation in classification and tracking processes," in *Aggregation and Fusion of imperfect information* (B. Bouchon-Meunier, ed.), pp. 231–260, Heidelberg : Physica-Verlag, 1998.
- [15] S. Fabre, A. Appriou, and X. Briottet, "Presentation and description of two classification methods using data fusion based on sensor management," *Information Fusion*, vol. 2, pp. 49–71, 2001.
- [16] P. Smets, "Belief functions : The disjunctive rule of combination and the Generalized Bayesian Theorem," *International Journal of Approximate Reasoning*, vol. 9, pp. 1–35, 1993.
- [17] T. Denoeux, "A k-nearest neighbour classification rule based on Dempster-Shafer theory," *IEEE Transactions on Systems Man and Cybernetics*, vol. 25, no. 5, pp. 804–813, 1995.
- [18] L. Zouhal and T. Denoeux, "An evidence-theoretic k-nn rule with parameter optimization," *IEEE Transactions on Systems, Man and Cybernetics C*, vol. 28, pp. 263–271, May 1998.
- [19] T. Denoeux, "A neural network classifier based on Dempster-Shafer theory," *IEEE Transactions on Systems, Man and Cybernetics A*, vol. 30, no. 2, pp. 131–150, 2000.
- [20] G. J. M. Lachlan and T. Krishnan, *The EM Algorithm and Extensions*. New-York : Wiley, 1997.
- [21] A. Bastière, "Methods for multisensor classification of airborne targets integrating evidence theory," *Aerospace Science and Technology*, no. 6, pp. 401–411, 1998.