

Estimation d'un modèle stationnaire de dispersion et localisation de source. Application à la surveillance de la pollution

Scattering model estimation and source localization. Application to pollution monitoring

par G. ROUSSEL, E. TERNISIEN, M. BENJELLOUN

Laboratoire d'analyse des systèmes du Littoral 50 rue Ferdinand Buisson – BP 699 62228 CALAIS Cedex e-mail : Gilles. Roussel@lasl.univ-littoral.fr

résumé et mots clés

Cet article aborde le problème de la localisation en deux étapes d'une source responsable de l'apparition d'un signal bruité sur des capteurs mesurant la concentration d'un composant chimique se dispersant sur un domaine spatial à surveiller. Après avoir modélisé le problème d'identification aveugle des paramètres physiques inconnus conditionnellement à une position quelconque de source, on détermine ensuite la position réelle de cette dernière en minimisant un critère basé sur la dissemblance spectrale partielle de la source estimée à partir des dsp des observations. Quelques simulations illustrent l'efficacité et l'exactitude de la méthode.

Localisation de source, dissemblance de spectre, identification aveugle

abstract and key words

This paper deals with a two-step source localization problem, which aims to identify which source is responsible for a significant change in the measured pollutant concentration in a monitored area. We first develop a blind identification procedure for unknown physical parameters conditional to the source coordinates. The true source coordinates are then obtained by minimizing a source partial spectral density distance, based on the power spectral densities of the observations. Simulations examples attest to the effectiveness of the algorithm.

Source localization, spectral difference, blind identification

1. introduction

La surveillance de la pollution atmosphérique ou aquatique connaît un intérêt particulier à l'heure où il convient d'évaluer l'impact des rejets produits par l'industrie ou le transport sur l'homme et son environnement. Il s'agit dans ce cas de déterminer les coordonnées de la source responsable d'une émission dont le débit inhabituel présente un risque de pollution pour la région placée en aval du courant horizontal de transport. C'est un problème de diagnostic faisant appel à des méthodes de résolution de problèmes inverses. Outre la difficulté mathématique inhérente à toute résolution de problèmes inverses (fermeture, stabilité,...), la qualité de la solution dépendra toujours de la précision du modèle direct pris comme référence au comportement réel du phénomène physique et de la connaissance de certaines grandeurs exogènes au modèle. Par exemple, concernant le modèle général de transport-diffusion qui nous intéresse ici, la connaissance précise des champs de vent $U(x, y, z, t)$ considérés stables et homogènes spatialement constitue une hypothèse forte. Le vent peut être mesuré et estimé *in situ* (par exemple par des méthodes d'estimation de données directionnelles [BPRB01] lorsque le domaine surveillé n'est pas trop étendu. S'il s'agit d'une pollution aquatique ou maritime, cela sous entend alors la connaissance du courant de transport. D'autres paramètres internes du modèle, comme les coefficients de diffusion par exemple, sont souvent choisis en référence à des études effectuées dans des conditions similaires à la situation courante. Dans cet article nous opterons plutôt pour une identification aveugle et en ligne de ces coefficients. Pour ce qui est de l'applicabilité du modèle, le cas du gaz passif en milieu atmosphérique (celui qui n'influence pas les mouvements de l'air qui l'entoure) est plus simple à traiter par le fait que seuls le vent moyen et la turbulence naturelle de l'atmosphère interviennent dans le processus de dispersion, et on peut alors négliger les recombinaisons chimiques avec les espèces environnantes. Pour des épisodes de dispersion où les conditions d'advection sont relativement stables en direction et en vitesse, le vecteur $\vec{U} = [U_x, U_y, U_z]^T$ peut alors être considéré approximativement constant et égal au vent (ou courant) moyen. Le comportement réel est alors assez voisin du modèle stationnaire issu de l'équation fondamentale (équation panache) de transport diffusion. Ce modèle, sous sa forme générale 3D, s'écrit :

$$\begin{aligned} & \frac{\partial X(x, y, z, t)}{\partial t} + U_x \frac{\partial X(x, y, z, t)}{\partial x} + U_y \frac{\partial X(x, y, z, t)}{\partial y} \\ & + U_z \frac{\partial X(x, y, z, t)}{\partial z} \\ & = K_x \frac{\partial^2 X(x, y, z, t)}{\partial x^2} + K_y \frac{\partial^2 X(x, y, z, t)}{\partial y^2} \\ & + K_z \frac{\partial^2 X(x, y, z, t)}{\partial z^2} + S(x_s, y_s, z_s, t) \end{aligned} \quad (1)$$

où $X(x, y, z, t)$ représente la concentration (en g/m^3) au point (x, y, z) à l'instant t , et $S(x_s, y_s, z_s, t)$ le débit (en g/s) de la source supposée ponctuelle aux coordonnées (x_s, y_s, z_s) .

Ce modèle, dans sa version statique ($\frac{\partial X(x, y, z, t)}{\partial t} = 0$), est très largement exploité dans la littérature [Con90, PS82, Erm77, RDTL00] mais ne rend pas compte de l'évolution temporelle. La version dépendante du temps est plus complexe à résoudre sous sa forme analytique. Les solutions directes et *a fortiori* inverses obtenues par le calcul analytique ou intégral sont, de toute façon, assez difficiles à expliciter lorsqu'un modèle fait intervenir différentes lois de la mécanique des fluides (conservation de masse, quantité de mouvement, équation de transport,...). Un certain nombre de travaux issus de diverses communautés scientifiques s'intéressent au problème sous-jacent de la localisation de sources. Pour un modèle dynamique parabolique de diffusion, on pourra lire dans [Kha94] une méthode de localisation d'une source parmi un ensemble connu de P sources potentielles. Malheureusement, la solution suppose de connaître l'intensité de l'émission, et suggère par ailleurs un problème linéaire d'estimation des intensités de sources (dans le cas des sources multiples) avec la connaissance de l'intensité totale. Une autre approche mathématique consiste à utiliser les modèles maillés dont les avantages sont liés essentiellement à l'absence de restriction dans les conditions aux limites qu'ils peuvent accepter, les hypothèses simplificatrices applicables aux équations, la prise en compte de l'hétérogénéité du milieu de dispersion et la description de tout le domaine spatial d'intérêt. Ainsi, la connaissance est à la fois spatiale et temporelle. Faisant largement appel aux calculs numériques, les coûts sont plus élevés mais diminuent avec les progrès informatiques. Ils constituent en cela une voie prometteuse. Dans [HI00], les auteurs s'appuient sur le théorème de réciprocité pour déterminer les rétropanaches issus du modèle adjoint de dispersion de polluants radioactifs.

Dans la communauté du traitement du signal, la localisation d'une source émettant de l'énergie dans un milieu de propagation est abordée sous l'angle du traitement d'antenne. Impulsés par l'apparition du goniomètre adaptatif pour localiser les émissions radio, de nombreux travaux (largement recensés dans [1Dd99]) sur la localisation par la recherche des directions d'arrivée de sources (DOA) sont fondés sur les propriétés haute résolution de la matrice de corrélation spatiale [BK83] ou l'utilisation des statistiques d'ordres supérieurs [LAC97]. La résolution asymptotiquement illimitée des méthodes HR en font une technique de choix pour bon nombre d'applications (télécommunication, géophysique, biomédical, sonar et radar). Nous proposons ici, d'approcher le problème de la localisation de source de pollution dans un milieu dispersif non isotropique par décomposition en sous-espace signal et bruit, technique largement utilisée en méthodes HR. Cependant, le choix de la modélisation par modèle maillé nous incite ici à abandonner la recherche des directions d'arrivée, mais plutôt à sélectionner de façon opti-

male un nœud source dans un repère cartésien au sens d'un critère de vraisemblance. Nous ferons cette étude sous l'hypothèse d'une source unique délivrant un signal large bande, telle que la nature spectralement riche des émissions polluantes le laisse supposer.

2. représentation dans l'espace d'état

La localisation de sources sur la base des modèles maillés pose donc une problématique qui d'ailleurs dépasse largement le cadre de la pollution atmosphérique ou aquatique. En effet, on peut aussi s'intéresser, par exemple, à la localisation de points chauds en diagnostic de systèmes mécaniques, etc ... Pour la surveillance de pollution, nous supposons donc que le modèle de connaissance (1) de la dispersion dans l'air du polluant émis par la source est une bonne approximation du phénomène réel. Cette émission va subir sur son trajet vers les capteurs qui le mesurent une atténuation complexe provoquée par un phénomène de diffusion (variable selon les directions) et un retard lié au temps de transport. Les observations du signal de l'unique source convolué au modèle de propagation sont ici délivrées par un ensemble de n_c capteurs spatialement répartis. Lorsque le maillage est suffisamment fin, nous avons montré dans [TRB00a] que la discrétisation de (1) peut s'effectuer en adoptant les différences finies comme approximation au premier ordre des dérivées. Plus spécifiquement, nous utilisons la méthode décentrée en espace dite méthode « upwind » [Euv94], plus appropriée lorsque la présence d'un terme d'advection suggère d'adopter une direction privilégiée dans le calcul des dérivées. Dans ce cas, la dérivée spatiale selon une direction $\vec{\alpha}$ s'écrit :

$$\frac{\partial X}{\partial \alpha} = \begin{cases} \frac{1}{h}(x_i^n - x_{i-1}^n) & \text{si } \vec{U} \cdot \vec{\alpha} > 0 \\ \frac{1}{h}(x_{i+1}^n - x_i^n) & \text{si } \vec{U} \cdot \vec{\alpha} < 0 \end{cases} \quad (2)$$

Le calcul de la dérivée est ici effectué sur une région en amont du point i par rapport à la direction du vent. Le modèle discrétisé, dans le cas 2D, est alors :

$$x_{n+1}^{i,j} = m_1 x_n^{i+1,j} + m_2 x_n^{i,j} + m_3 x_n^{i-1,j} + m_4 x_n^{i,j+1} + m_5 x_n^{i,j-1} + m_6 S_n^{i,j} \quad (3)$$

$x_n^{i,j}$ étant la représentation discrète à l'instant n et à la position (i, j) de la concentration $X(x, y, t)$. Les paramètres m_q ($q = 1 \dots 6$) sont fonctions des composantes de l'advection U_x et U_y , de la diffusion K_x et K_y , des pas du maillage spatial h et k selon les axes x et y et du pas temporel p défini par l'échantillonnage des observations. Plus précisément, et en prenant en

compte les changements de direction du vent par les fonctions $signe(U_\alpha)$, on obtient les relations suivantes :

$$\begin{aligned} m_1 &= \frac{pK_x}{h^2} + \frac{1}{2} [1 - signe(U_x)] \frac{p|U_x|}{h} \\ m_2 &= 1 - \frac{p|U_x|}{h} - \frac{p|U_y|}{k} - \frac{2pK_x}{h^2} - \frac{2pK_y}{k^2} \\ m_3 &= \frac{pK_x}{h^2} + \frac{1}{2} [1 + signe(U_x)] \frac{p|U_x|}{h} \\ m_4 &= \frac{pK_y}{k^2} + \frac{1}{2} [1 - signe(U_y)] \frac{p|U_y|}{k} \\ m_5 &= \frac{pK_y}{k^2} + \frac{1}{2} [1 + signe(U_y)] \frac{p|U_y|}{k} \\ m_6 &= p \end{aligned}$$

La condition de stabilité du schéma numérique impose cependant que pour la période d'échantillonnage p on ait :

$$p \leq \frac{h^2 k^2}{2k^2 K_x + 2h^2 K_y + hk^2 |U_x| + kh^2 |U_y|}$$

Ici, la source est ponctuelle. Sa contribution dans l'équation est nulle sauf au point $(i, j) = (x_s, y_s)$ où elle vaut S_n .

Il est alors possible de déduire un modèle d'état discret (4) dans lequel les composantes du vecteur X_n représente la grille des concentrations $x_n^{i,j}$ de l'espèce chimique surveillée aux différents points (i, j) du maillage et à l'instant d'échantillonnage n . Dans (5), X_n est construit en concaténant les L lignes contenant chacune l nœuds du maillage rectangulaire de taille $(L \times l)$. X_n est donc un vecteur de taille $(L \times l) \times 1$.

$$\begin{cases} X_{n+1} = AX_n + BS_n + w_n \\ y_n = CX_n + b_n \end{cases} \quad (4)$$

avec

$$X_n = \begin{bmatrix} [x_n^{1,1}, \dots, x_n^{1,j}, \dots, x_n^{1,l}]^T, \dots, \\ [x_n^{i,1}, \dots, x_n^{i,j}, \dots, x_n^{i,l}]^T, \dots, [x_n^{L,1}, \dots, x_n^{L,j}, \dots, x_n^{L,l}]^T \end{bmatrix}^T \quad (5)$$

La forme d'état (4) sépare le terme d'évolution AX_n du terme source BS_n . La matrice d'état A modélise la réponse libre du système de dispersion selon la définition analytique de l'équation différentielle (1) dans le cas stationnaire. Dans le cas d'un espace ouvert (2D), sans condition connue aux limites horizontales sur la frontière $\partial\Omega$ du domaine d'étude Ω , les concentrations aux points de $\partial\Omega$ sont considérés nulles. Cette hypothèse, un peu naïve du point de vue théorique, est en réalité une approximation qui facilite la dispersion aux bords et évite les accumulations. Ce codage conduit à une forme tri-diagonale par blocs de la matrice d'état A telle que :

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} & 0 & \cdots & \cdots & 0 \\ A_{2,1} & A_{2,2} & A_{2,3} & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & 0 \\ \vdots & & & & & \vdots \\ 0 & \cdots & \cdots & A_{L-1,L-2} & A_{L-1,L-1} & A_{L-1,L} \\ & & & 0 & A_{L-1,L} & A_{L,L} \end{pmatrix}$$

Chaque bloc $A_{k,k}$, $A_{k,k+1}$ et $A_{k-1,k}$ dépend des paramètres m_q constants obtenus à partir de (3) tels que :

$$A_{k,k} = \begin{bmatrix} m_2 & m_1 & & & & \\ m_3 & m_2 & m_1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & m_3 & m_2 & m_1 & \\ & & & m_3 & m_2 & \end{bmatrix},$$

$$A_{k,k+1} = \begin{bmatrix} m_4 & & & & \\ & \ddots & & & \\ & & m_4 & & \end{bmatrix} \text{ et } A_{k-1,k} = \begin{bmatrix} m_5 & & & & \\ & \ddots & & & \\ & & m_5 & & \\ & & & \ddots & \\ & & & & m_5 \end{bmatrix}.$$

Les concentrations initiales X_0 seront considérées nulles, de par l'effet évanescent des valeurs propres stabilisantes de la matrice A qui rendent imperceptibles, avec le temps écoulé, le dernier épisode de pollution. Par contre, on pourra tenir compte de w_n qui modélise dans (4) la concentration de fond considérée comme un vecteur de variables aléatoires de d.d.p. à support positif, plus particulièrement de loi poissonnienne. Conséquence d'une émission lointaine ou très diffuse, sa variance est peu significative en comparaison de celle du signal associé à la source à détecter, au point que, son effet est quasi imperceptible sur les mesures pendant l'épisode de pollution et nous l'ignorons dans la suite.

La matrice d'application de la commande B représente ici la configuration spatiale de la source. Lorsque cette dernière est ponctuelle (on peut étendre le problème au cas d'une source diffuse à flux spatial homogène sans perte de généralité) la matrice B est représentée par un vecteur de taille $(L \times l) \times 1$ contenant $Ll - 1$ zéros et un terme non nul m_6 , à l'indice correspondant au nœud source dans le maillage. La matrice C de l'équation d'observation caractérise la configuration du réseau de capteurs sur l'espace surveillé. De taille $n_c \times (L \times l)$, elle dépend de la position des capteurs sur la grille, et reprend le même codage que la matrice B . La présence d'un capteur d'étiquette k aux coordonnées (i, j) dans la grille est alors modélisé par $[C_{k,(j-1)l+i}] = 1$. Enfin, les observations y_n sont bruitées par le vecteur $b_n = [b_n^1, \dots, b_n^{n_c}]^T$ considéré ici comme la réalisation de n_c variables aléatoires *iid* gaussiennes, dont la variance σ^2 est déterminée par les caractéristiques du capteur.

Le modèle d'état couvre ainsi toute la zone d'observation et chaque transfert source-capteur est alors modélisable par un filtre (figure 1), *a priori* ARMA, de réponse impulsionnelle *a priori* infinie.

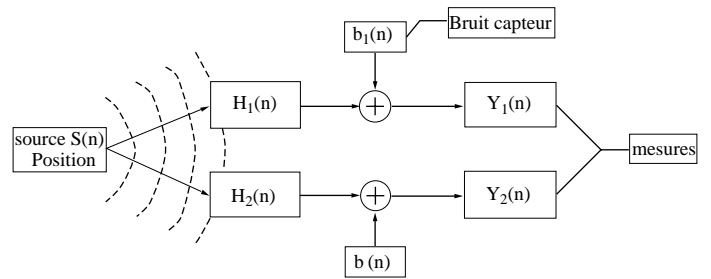


Figure 1. – Interprétation des canaux source-capteurs.

En tronquant les réponses impulsionnelles à l'ordre M maximal supposé connu du canal source-capteur pour le capteur le plus éloigné, on peut alors considérer que la source émet le signal S_n dans le canal de réponse impulsionnelle finie

$$h^i = [C^{(i)}B, \dots, C^{(i)}A^k B, \dots, C^{(i)}A^M B]^T \\ = [h_0^i, \dots, h_k^i, \dots, h_M^i]^T$$

et produit les observations y_n^i au capteur i selon la convolution :

$$y_n^i = \sum_{\Delta=0}^M h_n^i S_{n-\Delta T} + b_n^i \quad (6)$$

En multicapteur, les h^i ($1 \leq i \leq n_c$) sont les canaux vus par chacun des capteurs. Si N échantillons de y_n^i sont reçus, la forme matricielle du filtrage (6) s'écrit :

$$Y_n^i = \mathcal{H}_N^i S_n + \beta_n^i \quad (7)$$

avec $Y_n^i = [y_n^i, \dots, y_{n-N+1}^i]^T$, $\beta_n^i = [b_n^i, \dots, b_{n-N+1}^i]^T$ et \mathcal{H}_N^i la matrice de filtrage (dite de Sylvester) de taille $N \times (M + N)$ associée au canal h^i .

$$\mathcal{H}_N^i = \begin{pmatrix} h_0^i & \cdots & h_M^i & 0 & \cdots & 0 \\ 0 & \ddots & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & 0 \\ 0 & \cdots & 0 & h_0^i & \cdots & h_M^i \end{pmatrix}$$

En regroupant les n_c sorties, l'ensemble des observations devient :

$$Y_n = \mathcal{H}_N S_n + \beta_n \quad (8)$$

avec

$$\mathcal{H}_N = (\mathcal{H}_N^1 \cdots \mathcal{H}_N^{n_c})^T \quad (9)$$

La figure 2 montre une simulation du processus de dispersion directe. Le signal source, ici synthétique, simule un épisode de rejet discontinu mais d'amplitude constante (signal carré). Les conditions suivantes ont été imposées :

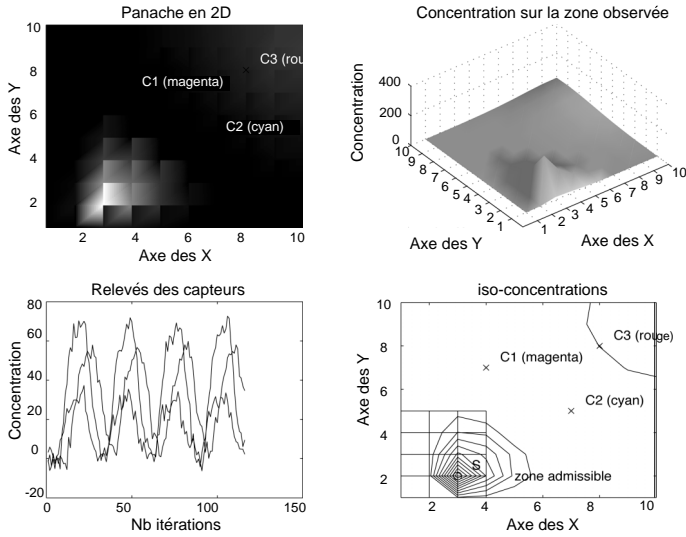


Figure 2. – Simulations directes. a) : panache 2D à l'instant final, b) : concentrations à l'instant final, c) : séquence d'observation, d) : courbes iso-concentrations à l'instant final.

\vec{U} : (10 m/s, 45°).

zone d'observation carrée : (10 km, 10 km), pas spatial 1 km.

coordonnée source : (3 km, 2 km) par rapport à l'origine.

coordonnées capteurs ($n_c = 3$) : (4 km, 7 km), (7 km, 5 km), (8 km, 8 km).

Période d'échantillonnage $p = 50$ s.

$K_x = 1$ m²/s, $K_y = 2$ m²/s.

Amplitude du signal source 200 g/s, période $T = 1500$ s.

bruit : $\sigma^2 = 10$.

3. identification aveugle

Dans bon nombre de cas pratiques, même si l'on dispose de quelques paramètres et de la connaissance des composantes du vent (U_x, U_y), on a souvent beaucoup d'incertitudes sur les paramètres qui caractérisent la diffusion. Certes, un certain nombre de travaux portant sur le cas du modèle statique [Tur94] de (1) ont permis d'établir des déviations standard de diffusion calculables en fonction des conditions météorologiques. Notre approche consiste ici à considérer que ces paramètres de diffusion $\theta = [K_x, K_y]$ sont inconnus et qu'il est plus fiable de les identifier en ligne. La matrice d'évolution $A = A(\theta)$ étant de fait partiellement connue, une estimation $\hat{\theta}$ est indispensable. Nous nous plaçons donc dans le contexte où la résolution du problème de localisation (détermination de B) est subordonnée à l'identification « aveugle » conjointe des paramètres θ du modèle directe de propagation $A(\theta)$ à l'aide des observations.

Nous avons développé dans [TRB00a] et [TRB00b] le principe utilisé pour résoudre le problème précédent. Notre méthode est

basée sur la décomposition en sous-espace bruit et sous-espace signal à partir des valeurs propres et vecteurs propres de la matrice d'autocorrélation R_y présentée dans (10).

$$R_y = E[Y_n Y_n^T] = \mathcal{H}_N R_s \mathcal{H}_N^T + R_b \\ = U \text{diag}(\lambda_0, \lambda_1, \dots, \lambda_{M+N-1}) U^H + \sigma^2 V V^H \quad (10)$$

Dans [MDCM95], Moulines *et al.* montrent que sous réserve que les conditions d'identifiabilité suggérées par [TXK91] (diversité des canaux impliquant que la matrice \mathcal{H}_N soit de rang colonne plein, $N > M$) soient vérifiées, ainsi que l'indépendance statistique de la source et du bruit, le classement des $n_c N$ valeurs propres $\lambda_0 \geq \lambda_1 \geq \dots \geq \lambda_{M+N-1} \geq \dots \geq \lambda_{n_c N-1}$ et des vecteurs propres associés de l'EVD de la matrice d'autocorrélation des sorties (10) permet d'isoler deux sous-espaces orthogonaux :

– l'espace source engendré par les colonnes de $U = [U_0, \dots, U_{M+N-1}]$ dont les valeurs propres associées sont supérieures à σ^2 . Cet espace est le même que celui engendré par les colonnes de \mathcal{H}_N .

– l'espace bruit engendré par les colonnes de $V = [V_0, \dots, V_{n_c N - M - N - 1}]$ dont les valeurs propres associées sont théoriquement égales à σ^2 .

Les deux sous-espaces étant orthogonaux, alors $\mathcal{H}_N^H V_i = 0$ pour tout V_i de V .

D'un point de vue calcul, il est plus pratique d'exprimer la condition d'orthogonalité en la réécrivant sous la forme scalaire équivalente :

$$\|\mathcal{H}_N^H V_i\| = V_i^H \mathcal{H}_N \mathcal{H}_N^H V_i = 0, \quad V_i \in V \quad (11)$$

Remarquons que si la matrice de filtrage a la forme (9), chaque vecteur propre V_i peut être partitionné en n_c sous-vecteurs tel que : $V_i = [V_i^{(1)}, \dots, V_i^{(n_c)}]^T$. Il est montré dans [MDCM95], que moyennant la définition de la matrice de filtrage multidimensionnelle $\mathcal{V}_i = [\mathcal{V}_i^{(1)}, \dots, \mathcal{V}_i^{(n_c)}]$ construite avec les éléments de $V_i^{(j)}$ telle que :

$$\mathcal{V}_i^{(j)} = \begin{bmatrix} v_{i,0}^{(j)} & v_{i,1}^{(j)} & \dots & v_{i,N-1}^{(j)} & 0 & \dots & \dots & 0 \\ 0 & v_{i,0}^{(j)} & v_{i,1}^{(j)} & \dots & v_{i,N-1}^{(j)} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & v_{i,0}^{(j)} & v_{i,1}^{(j)} & \dots & \dots & v_{i,N-1}^{(j)} \end{bmatrix} \quad (12)$$

alors,

$$V_i^H \mathcal{H}_N \mathcal{H}_N^H V_i = H^H \mathcal{V}_i \mathcal{V}_i^H H$$

où $H = [h^{(1)T}, \dots, h^{(n_c)T}]^T$ est défini comme le vecteur contenant les n_c réponses impulsionnelles inconnues. On peut alors redéfinir la condition d'orthogonalité (11) par :

$$H^H \mathcal{V}_i \mathcal{V}_i^H H = 0, \quad i = 0, 1, \dots, n_c N - M - N - 1 \quad (13)$$

soit encore, en généralisant à l'ensemble des vecteurs propres du sous-espace bruit :

$$H^T Q H = 0 \quad (14)$$

avec $Q = \sum_{i=0}^{n_c N - M - N - 1} \mathcal{V}_i \mathcal{V}_i^T$.

En minimisant le critère quadratique d'orthogonalisation $J(H) = H^T Q H$ par rapport aux coefficients de la forme vectorielle de la réponse impulsionnelle $H = [h^{(1)T}, \dots, h^{(n_c)T}]^T$, on détermine ainsi les réponses impulsionnelles recherchées.

Naturellement, l'application de la méthode est sous-tendue à la connaissance de l'ordre M du système. M peut être déduit de la taille du sous-espace signal soit parce que la variance σ^2 du bruit est connue, soit parce que la rupture des deux sous espaces est détectable par l'analyse de la distribution des valeurs propres de la matrice de covariance R_y [GLC96]. Il est clair que la précision de l'estimation va dépendre de l'estimation préalable de M . La méthode étant relativement robuste à une sous-estimation de l'ordre, nous accepterons une légère diminution de M , se traduisant par une troncature un peu plus forte de la plus grande réponse impulsionnelle. Cette minoration doit demeurer faible (approximation de la RII par une RIF), étant donné l'amplitude supposée élevée (pouvant être proche de l'unité) des valeurs propres de la matrice $A(\theta)$, et donc du nombre élevé de termes significatifs des réponses impulsionnelles.

Nous avons illustré dans [TRB00a] que la fermeture du problème de détermination des réponses impulsionnelles (réalisée dans [MDCM95] par une contrainte de normalisation des R.I.), était dans notre cas, assurée grâce aux relations introduites par la paramétrisation de la matrice $A(\theta)$. Nous ne cherchons pas ici à déterminer les réponses impulsionnelles de tous les canaux sources-capteurs, mais plutôt le vecteur de paramètres θ unifiant l'ensemble des canaux directs. Lorsque la source est localisée (B connue) l'identification de θ est donc sous-tendue à la résolution du problème non linéaire en θ :

$$\hat{\theta} = \arg \min_{\theta} J(\theta) \quad \text{avec} \quad J(\theta) = H^T(\theta) Q H(\theta) \quad (15)$$

L'optimisation du critère $J(\theta)$ sera menée par exemple par un algorithme de Levenberg-Marquart en initialisant la recherche à partir de la suggestion θ_0 non nul. On remarquera que le caractère unimodal (figure 3) de la fonction $J(\theta)$ au voisinage des valeurs pratiques du paramètre θ , nous assure la convergence vers la solution optimale.

La solution du programme (15) fournit dans $\hat{\theta}$ une estimation des paramètres physiques du modèle conditionnellement à une position de source.

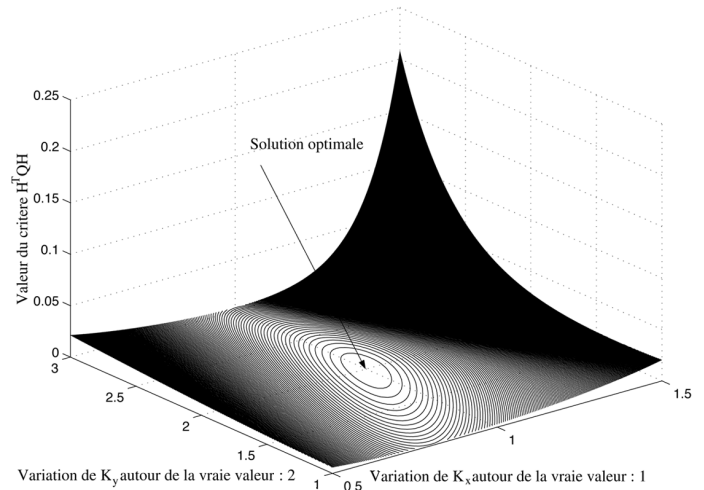


Figure 3. – Allure de la fonction $J(\theta)$ en iso-contours.

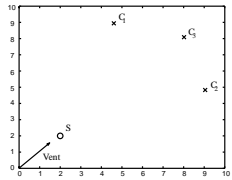
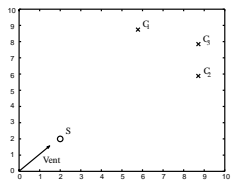
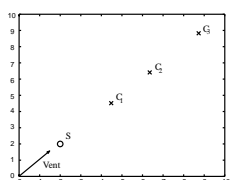
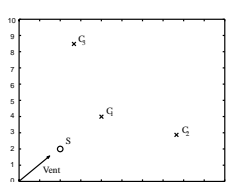
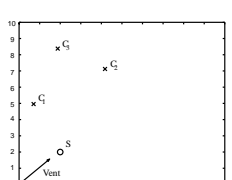
Le tableau Tab.1 renseigne sur la qualité de l'estimation du vecteur de paramètres $\theta = [K_x, K_y]$ en fonction de quelques situations typiques de placement. Il est clair que le biais est globalement sensible au rapport signal à bruit, qui pour un bruit de variance donnée, va dépendre de la puissance du signal reçu et donc de son placement par rapport aux paramètres de dispersion (\vec{U} et \vec{K}). Notons que le biais sur $\hat{\theta}$ est de plus fortement dépendant du biais sur l'estimation de la matrice d'autocorrélation R_y .

Le tableau Tab.2 montre, pour quelques exemples de signaux sources, le biais d'estimation sur θ . On constate que si le signal source est de faible puissance (exemple de l'impulsion isolée), le rapport signal/bruit au niveau des capteurs s'affaiblit trop (inférieur à 10 ou 15 dB) et l'erreur relative d'estimation croît. Résultat en tant que tel si la source est déjà localisée, le modèle ainsi identifié permet ensuite une déconvolution du signal source et la détermination des flux émis par la source de pollution. Si la source doit être localisée, l'identification aveugle constitue l'étape d'estimation qui précède l'étape de décision basée sur un critère évaluant la vraisemblance de la position de source parmi un ensemble possible de positions.

4. localisation

La localisation pose donc le problème de la détermination de la matrice B . Les premières approches testées dans [TRB00c, TRB00a, TRB00d] procèdent par une approche jointe de décision basée sur la minimisation d'un critère caractérisant la vraisemblance de la position testée par rapport aux autres positions potentielles. Il convient dans ces méthodes de définir un terme

Tableau 1. – Qualité de l'estimation en fonction de la position des capteurs

Placement des capteurs	$100 \times \frac{\ \hat{\theta} - \theta\ }{\ \theta\ }$
positions dispersées, rapport signal/bruit homogène : 	22 %
positions peu dispersées, rapport signal/bruit homogène : 	25 %
positions alignées : 	32 %
positions dispersées, rapport signal/bruit faibles sur deux capteurs : 	50 %
positions peu dispersées, aucun rapport signal/bruit élevé : 	127 %

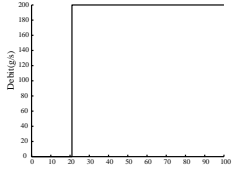
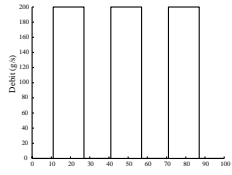
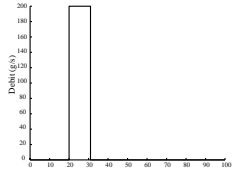
$J'(B)$ dépendant de B permettant de retrouver l'information retard qu'il n'est pas possible d'estimer par les techniques aveugles du second ordre.

Dans [TRB00a], le terme additif $J'(B)$ complète le critère général par une mesure de la distance entre l'indice du maximum de la séquence d'intercorrélation des observations $\tau_{x^i y^j}$ d'une part, et l'indice du maximum de la séquence d'intercorrél-

ation des réponses impulsionnelles $\tau_{r_{h^i h^j}}$ d'autre part. La solution optimale B^* est obtenue lorsque, pour chaque nœud du maillage où l'on estime la matrice A , le minimum de $J'' = J(\theta(B)) + J'(B)$ est trouvé : $B^* = \arg \min(J''(B))$.

Cette technique est coûteuse car, en itérant B sur l'ensemble des nœuds du maillage, elle devient dépendante de la finesse de ce dernier.

Tableau 2. – Qualité de l'estimation pour quelques signaux source

Signaux sources (cas de placement de la figure 2 avec les échelles en m)	$100 \times \frac{\ \hat{\theta} - \theta\ }{\ \theta\ }$
<p>Échelon (amplitude 200 g/s) :</p> 	10 %
<p>Signal rectangulaire périodique (amplitude 200 g/s, Largeur 15 T, Période 30 T) :</p> 	11 %
<p>Impulsion (amplitude 200 g/s, Largeur 12 T) :</p> 	25 %

Dans [TRB00d], on complète aussi l'information par un raisonnement d'approximation des phases. En considérant que si le spectre du signal d'émission se situe à l'intérieur de la bande passante de chaque canal source-capteur, alors les déphasages mesurables entre deux signaux sont essentiellement dus à la différence de temps liée au phénomène de transport de la matière. Dans ce cas, l'estimation $\tau_{r_{y_i y_j}}$ peut alors être considérée équivalente à la différence des retards purs présents dans les réponses impulsionnelles : $\tau_{h^i} - \tau_{h^j} \approx \arg \max_{\tau} (r_{y_i y_j}(\tau))$ où le terme τ_{h^i} est tel que $h^i(0) = \dots = h^i(\tau_{h^i}) = 0$. Dans ce cas, la position optimale est obtenue par la résolution du problème :

$$\begin{aligned}
 B^* &= \arg \min_{B \in \mathbf{B}_{admissible}} (J'(B)) \\
 &= \sum_{i,j}^{n_c} (\tau_{h^i}(B) - \tau_{h^j}(B)) - \arg \max_{\tau} (r_{y_i y_j}(\tau))
 \end{aligned}$$

$\mathbf{B}_{admissible}$ correspond à l'ensemble des points potentiels possibles, qui peut être réduit pour des raisons de complexité numérique par une méthode de prélocalisation présentée dans [TRB00c].

Enfin, dans [TRB00b], on vise à faire coïncider les déphasages observés entre chaque paire de capteurs et ceux introduits par les fonctions de transfert intercapteurs estimées pour chaque position admissible. Soient h^i et h^j les réponses impulsionnelles entre la source et deux capteurs i et j . Il est clair que la réponse impulsionnelle équivalente de la fonction de transfert intercapteurs h^{ij} du capteur i vers le capteur j peut prendre la forme $h^{ij} = \frac{h^j}{h^i}$. La matrice de filtrage \mathcal{H}^{ij} associée s'obtient en calculant la pseudo inverse $\mathcal{H}_N^{i \dagger}$ telle que :

$$\mathcal{H}^{ij} = \mathcal{H}_N^j \mathcal{H}_N^{i \dagger} \tag{16}$$

Pour chaque couple de capteurs, il est alors possible de calculer une erreur quadratique entre les estimations $\hat{y}^{(j)} = \mathcal{H}^{ij}(B)y^{(i)}$ au capteur d'arrivée j et l'observation à ce même capteur. Ainsi, la norme du vecteur d'erreur $\varepsilon_{ij} = y^{(j)} - \hat{y}^{(j)}$ tend vers 0 lorsque la réponse impulsionnelle intercapteur coïncide au bruit près à celle du système à identifier. De façon classique, on cherchera à optimiser le programme de localisation :

$$B^* = \arg \min_{B \in \mathbf{B}_{admissible}} (J'(B) = \left\| \sum_{\substack{i \\ \forall j > i}}^{n_c-1} y_j - \mathcal{H}^{ij}(B)y_i \right\|) \quad (17)$$

Le critère est considéré valide dans des cas où le RSB n'est pas trop faible, typiquement supérieur à 10 dB. En deçà, on peut aboutir à une erreur (faible) de localisation en raison de l'influence du bruit sur la mesure de coïncidence des réponses impulsionnelles.

5. dissemblance spectrale de source

Le critère précédent, bien que simple à mettre en œuvre, possède de l'inconvénient d'effectuer une somme dans le domaine temporel de signaux dont les rapports signaux à bruit sont parfois très différents. En effet, pour une position de source donnée et en fonction de la position des capteurs (figure 2.d), les réponses en fréquences de chaque filtre source-capteur varient, faisant apparaître une atténuation croissante avec l'éloignement de l'axe source-capteur par rapport à la direction du flux (figure 4). Ce phénomène est mécaniquement lié à la combinaison advection (déphasage linéaire) et diffusion (déphasage non linéaire, atténuation anisotrope et non monotone en fréquence).

En supposant constante la variance du bruit des différents capteurs, la variance du signal utile diminuera fortement avec la distance, augmentant ainsi fortement la variance de tout estimateur de ressemblance des signatures temporelles. Pour remédier à ce problème, on suggère, dans cette méthode, de comparer les

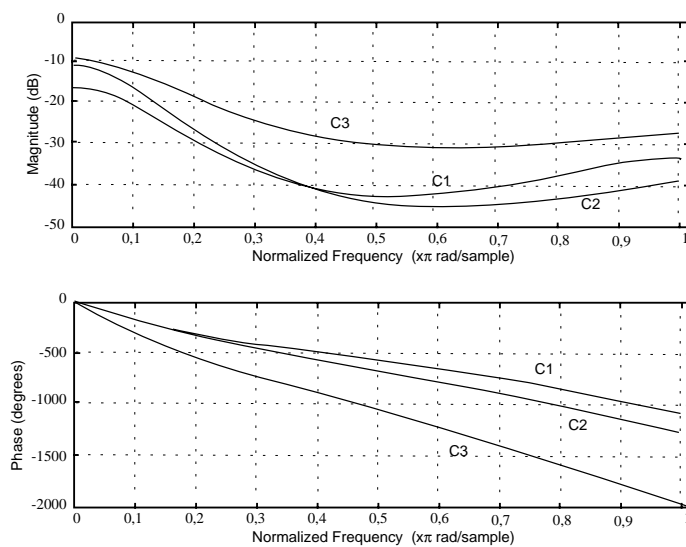


Figure 4. – Réponses fréquentielles des 3 canaux source-capteurs

spectres de puissance de tous les signaux capteurs ramenés au point source testé : la similitude des spectres de source devant être obtenue lorsque la position testée (parmi les positions de l'ensemble $\mathbf{B}_{admissible}$) correspond à la position effective.

Soit $\mathcal{S}_s(f) = \frac{\frac{1}{N} \sum_{l=1}^N w_l S^i(l) e^{-2\pi j f t}}{\frac{1}{N} \sum_{l=1}^N |w_l|^2}$ la densité spectrale de

puissance (dsp) de la source S_n , où la séquence $[w_1 \dots w_N]$ correspond aux coefficients d'une fenêtre de pondération.

Il est utile de rappeler ici que $\mathcal{S}_s(f)$ et $\mathcal{S}_{Y_{u_n}^i Y_{u_n}^i}(f)$, c'est-à-dire la dsp du signal utile $Y_{u_n}^i = (y_n^i - b_n^i)$ au capteur i , sont liées par le carré du module de la fonction de transfert harmonique $h^i(f)$ par :

$$\mathcal{S}_{Y_{u_n}^i Y_{u_n}^i}(f) = |h^i(f)|^2 \mathcal{S}_s(f) \quad \forall i \in [1..n_c] \quad (18)$$

$h^i(f)$ est la transformée de Fourier de la réponse impulsionnelle $h^i(t)$ obtenue par $h^i(f) = \int_{-\infty}^{+\infty} h^i(t) e^{-j2\pi f t} dt$.

Pour la bonne position de source B^* , on peut exprimer l'égalité des densités spectrales de la source correspondantes à chaque canal source-capteur telle que :

$$\mathcal{S}_s(f) = \frac{\mathcal{S}_{Y_{u_n}^1 Y_{u_n}^1}(f)}{|h^1(f, B^*)|^2} = \dots = \frac{\mathcal{S}_{Y_{u_n}^{n_c} Y_{u_n}^{n_c}}(f)}{|h^{n_c}(f, B^*)|^2} \quad (19)$$

Bien sûr, la relation (19) n'est vérifiée qu'aux fréquences basses où la dsp $\mathcal{S}_{Y_{u_n}^i Y_{u_n}^i}(f)$ du signal utile $Y_{u_n}^i$ ramenée à la source possède un niveau de puissance bien supérieur au bruit capteur (supposé blanc) ramené à la source par filtrage inverse. En effet, si $|h^i(f)|$ est décroissante en f (canal passe-bas), alors il existe une fréquence seuil f_{seuil} telle que pour $f > f_{seuil}$:

$$\frac{\mathcal{S}_{b_n^i b_n^i}(f)}{|h^i(f, B^*)|^2} > \frac{\mathcal{S}_{Y_{u_n}^i Y_{u_n}^i}(f)}{|h^i(f, B^*)|^2} \quad \forall f > f_{seuil} \quad (20)$$

En conséquence, à des fréquences très basses le bruit harmonique est faible en comparaison des signaux utiles, on peut réécrire (19) en substituant les dsp des signaux utiles par l'estimation des dsp $\widehat{\mathcal{S}}_{y_n^i y_n^i}(f)$ des observations y_n^i sur un même support temporel $[0, T_0]$:

$$\widehat{\mathcal{S}}_s(f) \simeq \frac{\widehat{\mathcal{S}}_{y_n^1 y_n^1}(f, T_0)}{|h^1(f, B^*)|^2} \simeq \dots \simeq \frac{\widehat{\mathcal{S}}_{y_n^{n_c} y_n^{n_c}}(f, T_0)}{|h^{n_c}(f, B^*)|^2} \quad (21)$$

$$\forall f < f_{max} \leq f_{seuil}$$

À la position optimale B^* , ces relations relatives à chaque paire de canaux conduisent, sous leurs formes logarithmiques, aux relations :

$$20 \log \frac{\widehat{S}_{y_n^i y_n^i}(f)}{|h^i(f, B^*)|^2} - 20 \log \frac{\widehat{S}_{y_n^j y_n^j}(f)}{|h^j(f, B^*)|^2} \simeq 0$$

On voit, que pour les cas où $h^i(f, B^*) \neq h^i(f, B)$ (ce qui est quasiment sûr si $B \neq B^*$), il est possible de définir une mesure d'écart spectral par la différence :

$$\text{diffSpec}_{ij}(f, B) = 20 \log \frac{\widehat{S}_{y_n^i y_n^i}(f)}{|h^i(f, B)|^2} - 20 \log \frac{\widehat{S}_{y_n^j y_n^j}(f)}{|h^j(f, B)|^2}$$

en dB (23)

La décision de localisation peut alors s'effectuer au sens de la minimisation de la *dissemblance spectrale de source* sur le support fréquentiel $[0..f_{max}]$:

$$B^* = \arg \min_B (\text{disSpec}(B)) \tag{24}$$

avec

$$\text{disSpec}(B) = \frac{2}{(k_{max} + 1) n_c (n_c - 1)} \sum_{i=1}^{n_c-1} \sum_{k=0}^{k_{max}} \text{diffSpec}_{ij}(f = \frac{k f_e}{2N}, B) \tag{25}$$

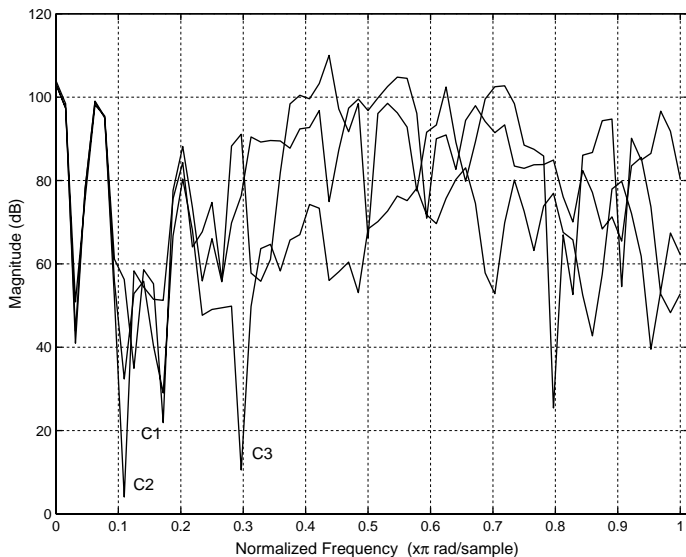


Figure 5. – Densités spectrales de puissance des trois observations capteurs à la bonne position de source. Elles se séparent nettement pour les fréquences normalisées > 0.1 rad/sample.

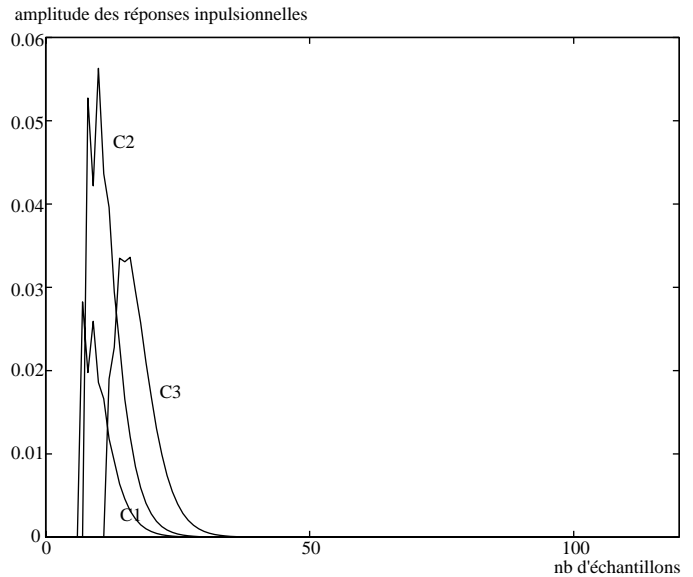


Figure 6. – Réponses impulsionnelles de trois canaux source-capteurs à la bonne position de source.

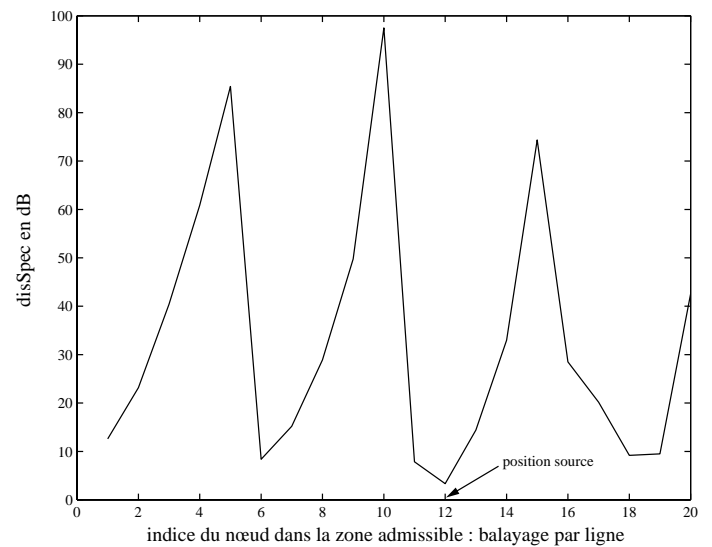


Figure 7. – Les valeurs de *disSpec* dans la zone admissible. *disSpec* est minimal pour B^* .

$\forall j > i, k_{max} = f_{max}/f_e$, et $f_e = 1/T_e$ la fréquence d'échantillonnage. À noter que k est un entier couvrant l'ensemble des points de fréquence du support fréquentiel $[0..k_{max}]$. k_{max} constitue la limite de bande basse fréquence que l'on prend de façon empirique égale, à 1/10 de la largeur du spectre total, ici $k_{max} = 5$.

Pour la simulation réalisée, le programme de décision (24) revient numériquement à calculer préalablement les estimations des dsp $\widehat{S}_{y_n^i y_n^i}(f)$ (figure 5) des signaux capteurs y_n^i (figure 2c),

$1 \leq i \leq n_c = 3$ (par la méthode du périodogramme sur le support temporel $[0..T_0]$ avec $T_0 = NT_e$). Pour l'ensemble des positions admissibles, l'identification aveugle de la matrice $A(B)$ conditionnellement à B permet ensuite de calculer les transformées de Fourier de chaque réponse impulsionnelle $h^i(t)$ identifiée (figure 6). Le calcul de la dissemblance s'obtient facilement en moyennant la séquence concaténée des k_{\max} premières densités des $n_c(n_c - 1)/2$ écarts spectraux. Enfin, la décision se fait en sélectionnant la valeur minimale critère (25). Celle-ci fournit la position du nœud contenant la source (figure 7).

6. conclusion

Après avoir présenté l'approche de modélisation du problème de localisation aveugle dans le cas d'un système de dispersion basé sur l'équation d'advection-diffusion, nous avons résumé quelques techniques abordées dans des travaux antérieurs pour fournir une solution de la matrice d'application de la commande du modèle d'état du système. Dans cet article, nous proposons, sans remettre en cause l'étape préalable d'identification aveugle des paramètres inconnus de la matrice d'état A , d'assurer l'étape de localisation par un critère de décision basé sur les écarts spectraux des paires d'observations ramenées à l'entrée des canaux source-capteurs. Ce critère, appelé dissemblance de spectre donne une mesure globale de la coïncidence du signal source déconvolué des différents signaux capteurs en s'affranchissant du bruit haute fréquence. Une simulation montre à travers un exemple de système, la validité de la méthode.

BIBLIOGRAPHIE

- [BK83] G. Bienvenu and L. Kopp. Optimality of high resolution array processing using the eigensystem approach. *IEEE trans. on ASSP*, 31(5) : 1235-1248, Oct 1983.
- [BPRB01] D. Brige, L. Pruvost, S. Reboul, and M. Benjelloun. MAP estimate for directional data : Application to the atmospheric turbulence. In *Second International Symposium PSIP'2001, Physics in Signal and Image Processing*, Marseille, France, Janvier 2001.
- [Con90] TRC Environmental Consultants. FDM : Fugitive dust model. Technical report, 21907 64th Avenue, Suite 230, Mountlake Terrace, WA 98043, 1990.
- [Erm77] D.L. Ermak. An analytical model for air pollutant transport and deposition from point sources. *Atmospheric Environment*, vol 11 : 231-237, 1977.
- [Euv94] D. Euvrard. *Résolution numérique des équations aux dérivées partielles*. MASSON, 1994.
- [GLC96] J. Gouffaud, P. Larzabal, and H. Clergeot. Some properties of ordered eigenvalues of a wishart matrix : Application in detection test and model order selection. In *Proceedings of ICASSP*, volume 5, pages 2465-2468, 1996.
- [HI00] F. Hourdin and J.-P. Issartel. Sub-surface nuclear test monitoring through the CTBT xenon network. *Geophysical research letters*, 27(15) : 2245-2248, August 2000.
- [Kha94] A.Y. Khapalov. Localization of unknown sources for parabolic systems on the basis of available observations. *INT. J. Systems Sci.*, 25(8) : 1305-1322, 1994.
- [LAC97] J.L. Lacoume, P.O. Amblard, and P. Comon. *Statistiques D'ordre Supérieur Pour Le Traitement Du Signal*. Masson, 1997.
- [IDd99] S. Marcos (Sous la Direction de). *Méthodes Haute Résolution : Traitement D'antenne et Analyse Spectrale*. HERMES, 1999.
- [MDCM95] E. Moulines, P. Duhamel, J-F. Cardoso, and S. Mayrargue. Subspace methods for the blind identification of multichannel FIR filters. *IEEE Trans on Signal Processing*, 43 : 516-525, 1995.
- [PS82] F. Pasquill and F.B. Smith. *Atmospheric Diffusion, third edition*. ELLIS HORWOOD LIMITED, 1982.
- [RDTL00] G. Roussel, G. Delmaire, E. Ternisien, and R. Lherbier. Separation problem of industrial particles using a stationary scattering model. *Journ. Environment Modelling and software*, vol. 15/6-7 : pp 653-661, december 2000.
- [TRB00a] E. Ternisien, G. Roussel, and M. Benjelloun. Blind identification and source localization of dispersive SIMO model by subspace method. In *Sysid'2000*, Santa Barbara, California, June 2000. IFAC.
- [TRB00b] E. Ternisien, G. Roussel, and M. Benjelloun. Critère de localisation de source et d'identification pour un modèle de dispersion SIMO. In *CIFA 2000*, Lille, France, 5-8 juillet 2000. CESA.
- [TRB00c] E. Ternisien, G. Roussel, and M. Benjelloun. Identification aveugle du milieu de dispersion et localisation de source de pollution par observation multicapteurs. In *Journées de l'automatique et Environnement*, Nancy, 9-10 Mars 2000. CRAN.
- [TRB00d] E. Ternisien, G. Roussel, and M. Benjelloun. Subspace method for blind characterisation of atmospheric scattering model. In *Scientific Computation, Applied Mathematics and Simulation*, Lausanne, Suisse, 09 2000. 16th IMACS.
- [Tur94] D. Bruce Turner. *Atmospheric Dispersion Estimates. An Introduction to Dispersion Modeling*. CRC Press, Inc, 2000 Corporate Blvd, Broca Raton, Florida 33431, 1994.
- [TXK91] L. Tong, G. Xu, and T Kailath. A new approach to blind identification and equalization of multipath channels. In *Proc. 25th Asilomar Conf.*, pages 856-860, 1991.

Manuscrit reçu le 18 janvier 2001

LES AUTEURS

Gilles ROUSSEL



Gilles ROUSSEL est né en 1963 à Saint-Omer (France). Il est titulaire d'un DEA et d'un Doctorat de 3^{ème} cycle de l'Université de Lille I depuis 1994, en Automatique et Informatique Industrielle. Il a rejoint l'Université du Littoral Côte d'Opale comme maître de Conférences en 1994, et effectue ses recherches dans le Laboratoire d'Analyse des Systèmes du Littoral (LASL). Il a travaillé de 1989 à 1996 sur les problèmes d'optimisation combinatoire (optimisation de placement géométrique, optimisation par inférence, ..). Depuis 1996, il s'intéresse aux problèmes d'estimation et de décision appliqués à la caractérisation des systèmes et des signaux sources par la perception multidimensionnelle.

Eric TERNISIEN



Eric TERNISIEN est né en 1973 à Boulogne-sur-mer (France). Il est titulaire d'un DEA d'Automatique et d'Informatique Industrielle et d'un Doctorat de 3^{ème} cycle en Génie Informatique, Automatique et Traitement du Signal. Ses activités de recherche se concentrent sur les problèmes d'estimation et de décision appliqués à la caractérisation des systèmes physiques et à la localisation de source. Il enseigne actuellement à l'IUP Génie des Systèmes Industriels du Littoral dans le cadre de son contrat d'Attaché Temporaire d'Enseignement et de Recherche.

Mohammed BENJELLOUN



Mohammed BENJELLOUN est docteur de 3^{ème} cycle de l'Université de Lille I en Génie Electrique et habilité à diriger des recherches de l'Université d'Evry Val d'Essonne. Il est actuellement Professeur à l'Université du Littoral Côte d'Opale dans le Laboratoire d'Analyse des Systèmes du Littoral (LASL, UPRES EA 2600), dont il est directeur depuis 1997. Son domaine d'intérêt en recherche est l'analyse des signaux multidimensionnels et la fusion d'informations.