

Estimation conjointe de moments spectraux d'échos Doppler. Application au cyclone 'Georges'

Eric BOYER¹, Claude ADNET², Pascal LARZABAL^{1,3} et Monique PETITDIDIER⁴

¹ LESiR/ENS Cachan, UPRESA 8029, 61 avenue du président Wilson, 94235 Cachan Cedex France.

²THALES Air Défense, 7-9 rue des Mathurins, Bagneux France.

³ IUT de Cachan, C.R.I.I.P, Université Paris Sud, BP140, 9 avenue de la division Leclerc, 94234 Cachan Cedex, France.

⁴ CETP 10-12 Avenue de l'Europe, 78140 Vélizy France.

eric.boyer@wanadoo.fr

Résumé – L'objet de ce travail est l'estimation des moments spectraux d'échos Doppler. Dans le cas d'échos proches, les algorithmes classiques basés sur la transformée de Fourier donnent de mauvais résultats du fait de leur manque de résolution. Nous proposons et comparons l'utilisation des méthodes du maximum de vraisemblance stochastique et de sous-espaces (algorithme Weighted Pseudo-Subspace Fitting : WPSF) pour une estimation conjointe des moments spectraux. Les performances statistiques (théorique et empirique par simulations de Monte Carlo) des estimateurs sont comparées à la borne de Cramér-Rao. Les résultats des tests effectués sur les données réelles obtenues lors de la campagne d'étude d'orages, à Arecibo, en Septembre et Octobre 1998 confirment l'intérêt de l'approche proposée.

Abstract – The purpose of this work is the estimation of Doppler echoes spectral moments. In case of strong overlapping, Fourier like techniques provide poor results because of the lack of resolution. We propose the use of stochastic maximum likelihood and subspace-based methods (WPSF algorithm) for a joint estimation of spectral moments. The statistical performances (theoretical and empirical by Monte Carlo simulations) of estimators are compared with the Cramér-Rao lower bound. The results of tests performed on VHF times series obtained during Thunderstorm, Arecibo, PR during September and October 1998 confirm the interest of the proposed approach.

1. Introduction

L'objet de ce travail est l'estimation des moments spectraux d'échos Doppler proches. Avec des radars Doppler VHF, cela peut être le cas lorsque l'écho rétrodiffusé se superpose à un écho de sol ou à d'autres échos dus à des hydrométéores (fortes pluies...). La puissance de chaque écho (moment d'ordre zéro) donne une estimation de l'intensité de la turbulence ou des caractéristiques des hydrométéores (nombre et taille) présents dans le volume de résolution; la fréquence moyenne due à l'effet Doppler (moment d'ordre 1) renseigne sur la valeur et le sens du vent radial ou du mouvement des hydrométéores; l'écart type de chaque écho (moment d'ordre 2) mesure la dispersion des vitesses par rapport à la vitesse moyenne [1]. Dans des cas de fort chevauchement, les algorithmes classiques basés sur la transformée de Fourier donnent de mauvais résultats du fait de leur manque de résolution. Nous proposons l'utilisation des méthodes de sous-espaces vectoriels (algorithme WPSF [3]) pour une estimation conjointe des moments spectraux. Cet article étend les résultats de [4], en présentant notamment une comparaison des variances théorique et expérimentale de l'estimateur WPSF ainsi qu'une comparaison avec la borne de Cramér-Rao. Les performances statistiques de l'estimateur WPSF sont de plus comparées à celles de l'estimateur du maximum de vraisemblance stochastique (MVS). Nous présentons au paragraphe 2 une modélisation paramétrique originale de la matrice de covariance des signaux reçus. Au paragraphe 3, nous présentons les algorithmes utilisés. Le paragraphe 4 donne les résultats de l'étude statistique des estimateurs. Enfin, au paragraphe 5 l'algorithme WPSF est validé sur les données réelles obtenues lors de la campagne d'étude d'orages, à Arecibo, en Septembre et Octobre 1998.

2. Modélisation paramétrique de la matrice de covariance des signaux reçus

D'après [1] et [2], les échos contenus dans le spectre de puissance $P(f)$ des séries temporelles sont supposés gaussiens, soit pour un écho

$$P(f) = (S(f) + \sigma_b^2) \cdot |\beta|^2$$

$$\text{avec } S(f) = \frac{P_0}{\sqrt{2\pi} \cdot \sigma_0} \cdot \exp\left[-\frac{1}{2} \left(\frac{f - f_0}{\sigma_0}\right)^2\right] \quad (1)$$

où σ_b^2 est une puissance de bruit additive, où β est une variable aléatoire gaussienne, de variance unité et où P_0 , f_0 et σ_0 sont les trois moments spectraux de l'écho gaussien. f_0 est la fréquence Doppler, reliée au vent radial moyen par la relation $f_0 = -\frac{2v_r}{\lambda}$, avec $\lambda = c \cdot T_0$ la longueur d'onde de

l'onde émise par le radar. De l'éq. (1), l'auto-corrélation de la série temporelle $x(t)$ du signal s'écrit :

$$\gamma(t) = P_0 \cdot \exp\left[(-2\pi^2\sigma_0^2t^2 + j \cdot 2\pi \cdot f_0 t)\right] + \sigma_b^2 \cdot \delta(t) \quad (2)$$

Notons \mathbf{x} les snapshots, de dimension $m \times 1$, tirés directement de la série temporelle $\{x(t), 0 \leq t \leq \tau\}$:

$$\mathbf{x} = \mathbf{y} + \mathbf{n} \quad (3)$$

où \mathbf{y} est un vecteur aléatoire gaussien centré et \mathbf{n} est un bruit blanc complexe et gaussien indépendant de \mathbf{y} , vérifiant

$$E[\mathbf{n}_i \mathbf{n}_j^*] = \sigma_b^2 \mathbf{I}_{\delta_{i,j}} \quad (4)$$

où $(\cdot)^*$ représente l'opérateur transposé conjugué.

Avec l'hypothèse d'un spectre gaussien pour les séries temporelles, la matrice de covariance de \mathbf{y} peut s'écrire, à partir de (2)

$$\mathbf{R}_y(P_0, \omega_0, \sigma_0^2) = E[\mathbf{y}\mathbf{y}^*] = P_0 \cdot \mathbf{A}(\omega_0) \mathbf{B}(\sigma_0^2) \mathbf{A}^*(\omega_0) \quad (5)$$

où

$$\mathbf{A}(\omega_0) = \text{diag}(1, e^{j\omega_0 T_s}, \dots, e^{j(m-1)\omega_0 T_s}) \quad (6)$$

$$\mathbf{B}(k, l) = e^{-2\pi^2 \sigma_0^2 (k-l)^2 T_s^2} \quad (7)$$

et où T_s est la période d'échantillonnage du radar. De (3) et (4), la matrice de covariance de \mathbf{x} s'écrit

$$\mathbf{R}_x(P_0, \omega_0, \sigma_0^2, \sigma_n^2) = \mathbf{R}_y(P_0, \omega_0, \sigma_0^2) + \sigma_n^2 \mathbf{I} \quad (8)$$

$$\mathbf{x} \in \mathcal{N}[\mathbf{0}, \mathbf{R}_x]$$

Dans le cas de n échos gaussiens indépendants, (8) reste valable avec $\mathbf{R}_y = \sum_{i=1}^n \mathbf{R}_{y_i}(P_i, \omega_i, \sigma_i^2)$. De plus, le système est totalement identifiable.

3. Présentation des estimateurs utilisés

3.1 Estimation au sens du maximum de vraisemblance stochastique

Notons $\boldsymbol{\mu}$ le vecteur paramètre à estimer

$$\boldsymbol{\mu} = [\omega_1 \sigma_1^2 P_1 \dots \omega_n \sigma_n^2 P_n \sigma_b^2] \quad (9)$$

De la relation (3), la log-vraisemblance négative s'écrit [5]

$$L(\boldsymbol{\mu}) = \log(|\mathbf{R}_x|) + \text{trace}[\mathbf{R}_x^{-1} \hat{\mathbf{R}}_x] + cte \quad (10)$$

L'estimée $\hat{\boldsymbol{\mu}}$ au sens du maximum de vraisemblance du vecteur paramètre $\boldsymbol{\mu}$ s'écrit alors

$$\hat{\boldsymbol{\mu}} = \arg \min_{\boldsymbol{\mu}} (L(\mathbf{x})) \quad (11)$$

3.2 Algorithme WPSF

Cet algorithme [3] est fondé sur une orthogonalité entre sous-espaces vectoriels qui ne fait pas intervenir les P_i et σ_b^2 . Par conséquent, le jeu de paramètres se réduit à $\boldsymbol{\eta}$

$$\boldsymbol{\eta} = [\omega_1 \sigma_1^2 \dots \omega_n \sigma_n^2] \quad (12)$$

Les paramètres P_i et σ_b^2 seront estimés au sens des moindres carrés dans une étape suivante.

Une décomposition en éléments propres de \mathbf{R}_x s'écrit

$$\mathbf{R}_x = \sum_{k=1}^m \lambda_k \mathbf{e}_k \mathbf{e}_k^* \quad (13)$$

$$= \mathbf{E}_s(\boldsymbol{\eta}) \boldsymbol{\Lambda}_s(\boldsymbol{\eta}) \mathbf{E}_s^*(\boldsymbol{\eta}) + \mathbf{E}_n(\boldsymbol{\eta}) \boldsymbol{\Lambda}_n(\boldsymbol{\eta}) \mathbf{E}_n^*(\boldsymbol{\eta})$$

où $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$ et où \mathbf{E}_s est "l'espace signal" engendré par les vecteurs propres correspondant aux d plus grandes valeurs propres de \mathbf{R}_x . Dans la suite, la dimension d de \mathbf{E}_s est un paramètre variable estimé par l'utilisateur; d n'est pas nécessairement égal à n (mais $n \leq d < m$).

Une estimée de \mathbf{R}_x est donnée par

$$\hat{\mathbf{R}}_x = \frac{1}{K} \sum_{t=1}^K \mathbf{x}(t) \mathbf{x}(t)^* \quad (14)$$

$$= \sum_{k=1}^m \hat{\lambda}_k \hat{\mathbf{e}}_k \hat{\mathbf{e}}_k^* = \hat{\mathbf{E}}_s \hat{\boldsymbol{\Lambda}}_s \hat{\mathbf{E}}_s^* + \hat{\mathbf{E}}_n \hat{\boldsymbol{\Lambda}}_n \hat{\mathbf{E}}_n^*$$

Dans l'algorithme WPSF [3], le vecteur paramètre $\boldsymbol{\eta}$ est estimé comme la valeur $\hat{\boldsymbol{\eta}}$ de $\boldsymbol{\eta}$ qui rend $\hat{\mathbf{E}}_s$ le plus orthogonal possible à $\mathbf{E}_n(\boldsymbol{\eta})$, soit :

$$\hat{\boldsymbol{\eta}} = \arg \min_{\boldsymbol{\eta}} f_{WPSF}(\boldsymbol{\eta}) = \arg \min_{\boldsymbol{\eta}} \boldsymbol{\rho}^*(\boldsymbol{\eta}) \mathbf{W} \boldsymbol{\rho}(\boldsymbol{\eta}) \quad (15)$$

$$\boldsymbol{\rho}(\boldsymbol{\eta}) = \text{vec}[\mathbf{E}_n^*(\boldsymbol{\eta}) \hat{\mathbf{E}}_s] \quad (16)$$

et où \mathbf{W} est une matrice de pondération de dimension $d(m-d) \times d(m-d)$, hermitienne définie positive. Pour le choix de \mathbf{W}_{opt} se reporter à [3].

De (14), une estimée de σ_b^2 est donnée par [5]

$$\hat{\sigma}_b^2 = \frac{1}{m-d} \sum_{i=d+1}^m \hat{\lambda}_i \quad (17)$$

De (2), la fonction d'auto-corrélation peut se modéliser sous la forme

$$\gamma(t) = \sum_i P_i \alpha_i(t) \cdot e^{j\omega_i t} + \sigma_b^2 \cdot \delta(t) \quad (18)$$

$$\alpha_i(t) = \exp(-2 \cdot \pi^2 \sigma_i^2 t^2)$$

De (18) les snapshots $\boldsymbol{\gamma}$ de dimension $m' \times 1$ ($m' \neq m$ a priori) tirés de la fonction d'auto-corrélation peuvent s'écrire

$$\boldsymbol{\gamma} = \mathbf{A}(\boldsymbol{\omega}, \boldsymbol{\sigma}) \cdot \mathbf{P} + \mathbf{n} \quad (19)$$

où

$$\mathbf{P} = [P_1 \quad \dots \quad P_n]^T \quad (20)$$

$$\mathbf{A}(\boldsymbol{\omega}, \boldsymbol{\sigma}) = [\mathbf{a}(\omega_1, \sigma_1) \quad \dots \quad \mathbf{a}(\omega_n, \sigma_n)] \quad (21)$$

$$\mathbf{a}(\omega_i, \sigma_i) = \left[1 \quad \alpha_i(T_s) e^{j\omega_i T_s} \quad \dots \quad (\alpha_i(T_s) e^{j\omega_i T_s})^{m-1} \right]^T \quad (22)$$

$$\mathbf{n} = [1 \quad 0 \quad \dots \quad 0]^T \quad (23)$$

De (19), les estimées \hat{P}_i de P_i sont obtenues en insérant dans la relation suivante l'estimée $\hat{\boldsymbol{\eta}}$ de $\boldsymbol{\eta}$ [5] :

$$\hat{\mathbf{P}} = [\hat{P}_1 \quad \dots \quad \hat{P}_n] = \left[(\mathbf{A}^* \mathbf{A})^{-1} \mathbf{A}^* \right] \cdot \hat{\boldsymbol{\gamma}} \quad (24)$$

On ne prend pas dans $\hat{\boldsymbol{\gamma}}$ l'échantillon $t=0$.

4. Etude de performance

4.1 Borne de Cramér-Rao (BCR)

La BCR est calculée directement à partir de la série temporelle composée des N échantillons [7].

Le ij ème élément de la matrice d'information de Fisher (FIM) s'écrit alors (formule de Bang) :

$$\{FIM\}_{ij} = \text{trace} \left\{ \mathbf{R}_x^{-1} \mathbf{R}_{x_i} \mathbf{R}_x^{-1} \mathbf{R}_{x_j} \right\} \quad (25)$$

$$\mathbf{R}_{x_i} = \frac{\partial \mathbf{R}_x}{\partial \mu_i} \quad (26)$$

Les expressions des \mathbf{R}_{x_i} ne sont pas reportées ici mais il est important de noter qu'elles ne font appel à aucune approximation, contrairement à des résultats similaires [3].

4.2 Etude statistique des estimateurs

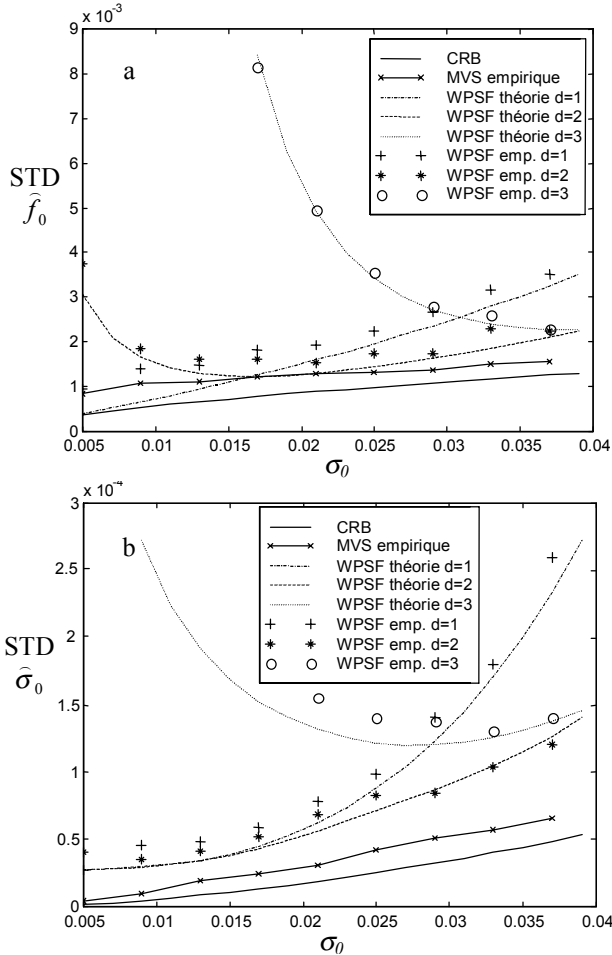


Fig. 1 : performances des estimateurs WPSF et MVS dans le cas d'un écho gaussien.

Les simulations ont été effectuées avec des snapshots de 8 échantillons et ne se recouvrant pas. Les fréquences et écarts-types donnés sont normalisés. La détermination du vecteur paramètre se fait par une descente de Newton initialisée à la

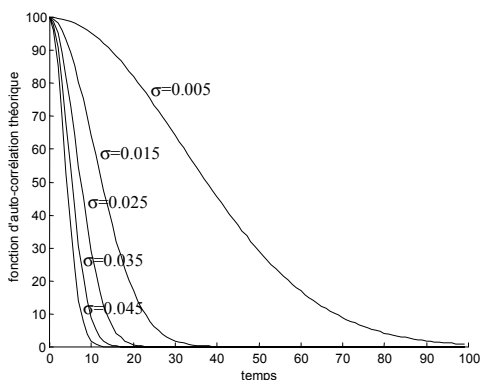


Fig. 2 : allure de la fonction d'auto-corrélation pour différentes valeurs de σ

vraie valeur du vecteur paramètre. La figure 1 présente les résultats des tirages de Monte Carlo (200 réalisations) dans le cas d'un écho Gaussien ($P_0=100 W, f_0=0$ et σ_0 variable). Les performances statistiques de l'estimateur MVS sont meilleures. Pour l'algorithme WPSF, le choix de d est primordial pour la qualité de l'estimation. Les expressions des variances théoriques de l'estimateur WPSF (Fig.1) sont données dans [3] dans le cas de snapshots indépendants. Comme le montre la figure 2, la fonction d'auto-corrélation (2) de l'écho gaussien ne décroît pas suffisamment rapidement pour les faibles valeurs de σ pour avoir des snapshots décorrelés. Cela explique l'écart observé entre les variances théorique et empirique de l'estimateur WPSF et la diminution de cet écart avec l'augmentation de σ .

La figure 3 présente les résultats des tirages de Monte Carlo (200 réalisations) dans le cas de deux échos gaussiens identiques ($\sigma=0.02$ et $P=100 W$) l'un en 0 et l'autre en Δf variable. Pour de faibles valeurs de Δf , la dégradation de l'estimation par la méthode WPSF est beaucoup plus rapide que pour l'algorithme MVS. WPSF ($d=2$) atteint MVS quand Δf augmente.

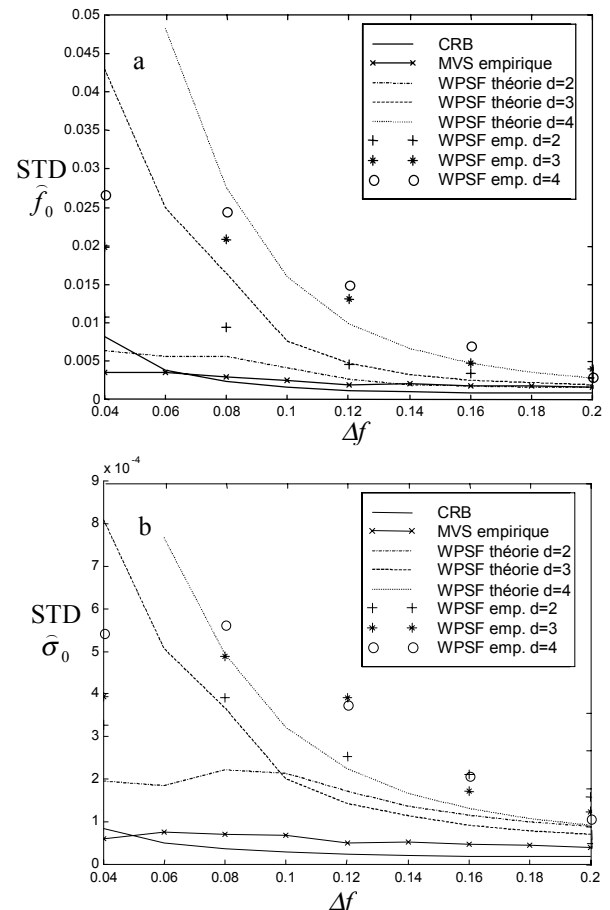


Fig. 3 : performances des estimateurs WPSF et MVS dans le cas de 2 échos gaussiens

5. Validation sur données réelles

Les résultats présentés figures 5 et 6 correspondent à l'application de la méthode WPSF sur les données réelles issues de la campagne d'orages, à Arecibo, en septembre et

octobre 1998. Les difficultés d'initialisation de l'algorithme MVS n'ont pas permis d'aboutir à des estimations concluantes de moments spectraux sur données réelles.



Fig. 4 : le radar d'Arecibo

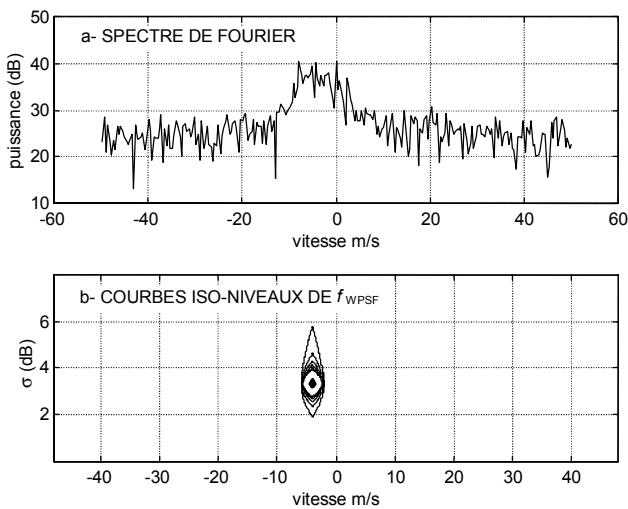


Fig.5 : validation de l'algorithme WPSF sur données réelles

L'estimation conjointe des moments spectraux d'ordre 1 et 2 présentés figure 5.b concordent avec le spectre réel de la figure 5.a. Sur la figure 6, la reconstruction du spectre par la méthode WPSF fait apparaître un troisième pic se superposant à l'écho de vent et probablement dû à une réflexion sur les hydrométéores.

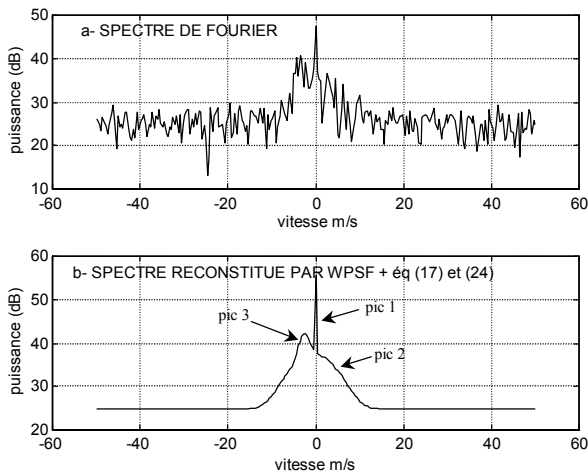


Fig. 6 : exemple d'estimation WPSF sur données réelles dans le cas de plusieurs échos

L'estimation concorde avec l'allure du spectre et permet de confirmer la présence d'un écho supplémentaire (pic 3 difficilement décelable avec Fourier).

Cependant, l'initialisation du vecteur paramètre reste une étape délicate dans la mise en œuvre opérationnelle de l'algorithme. De plus, il est nécessaire de choisir correctement la dimension d de l'espace signal.

6. Conclusions

Nous avons développé deux algorithmes d'estimation reposant sur une modélisation originale de la matrice de covariance des échos atmosphériques. Il s'agit tout d'abord de la méthode du MVS, qui bien que peu utilisée en traitement du signal se révèle intéressante dans notre problème, puis d'une généralisation de WSPF initialement introduit pour pallier aux difficultés de mise en œuvre du MVS tout en conservant les performances asymptotiques. Dans notre modèle, la matrice de covariance non bruitée est de rang plein, expliquant les meilleures performances statistiques (à priori inattendues) du MVS par rapport à WPSF. Les résultats empiriques obtenus par tirages de Monte Carlo viennent confirmer les développements théoriques. En raison de la difficulté d'initialisation du MVS, seul l'algorithme WPSF a été testé sur signaux réels. Dans des cas simples, le bon comportement de la méthode a été validé par la méthode de Fourier. Dans le cas de forts chevauchements, nous avons mis en évidence les améliorations introduites sur l'estimation des moments spectraux.

Remerciements

The National Astronomy and Ionospheric center is operated by Cornell University under contract with the National Science foundation.

Références

- [1] R.J. Doviak et D.S. Zrnic. *Doppler Weather Signal Radar Observations*. Academic Press, Inc., 1984.
- [2] A. Papoulis. *Probability, Random Variables and Stochastic Processes*. New York, MacGraw-Hill, 1965.
- [3] M. Bengtsson et B. Ottersten. *Antenna Array Signal Processing for High Rank Data Models*, PHD Thesis, KTH Sweden, 1999.
- [4] E. Boyer, M. Petitdidier, C. Adnet et P. Larzabal. *Subspace based spectral analysis for VHF and UHF radar signals*. PSIP, 2001.
- [5] S. Haykin, J. Litva et T.J. Shepherd. *Radar Array Processing*. Springer and Verlag, 1993.
- [6] R.O. Schmidt, *A signal subspace approach to multiple emitter location and spectral estimation*, IEEE Trans. Ant. Propag. AP-34, 276, 1986.
- [7] P. Stoica et B. Nehorai. *MUSIC, Maximum Likelihood and Cramér-Rao Bound*. IEEE Trans. on acoustics, speech and signal processing, vol. 37. Mai 1989.