

Reconnaissance automatique de fréquences de résonance à partir des données satellitaires de la mission spatiale CLUSTER

X. SURAUD¹, J.-L. RAUCH¹, R. LEDEE², P. DECREAU¹

¹ Laboratoire de Physique et Chimie de l'Environnement / Centre National de la Recherche Scientifique (LPCE/CNRS), 3A, avenue de la recherche scientifique 45071 Orléans cedex 2

² Laboratoire d'Electronique, Signaux, Images (LESI) EA 1715 / Université d'Orléans Polytech'Orléans - Site Galilée - 12 rue de Blois, BP 6744 45067 Orléans cedex 2

Xavier.Suraud@cnsr-orleans.fr; Jean-Louis.Rauch@cnsr-orleans.fr; Roger.Ledee@univ-orleans.fr;

Pierrette.Decreau@cnsr-orleans.fr

Résumé – Dans cette communication, nous présentons une méthode automatique pour l'identification de la fréquence de résonance d'un plasma spatial. Les données satellitaires sont issues de deux instruments de mesure. Elles se présentent sous la forme d'un spectrogramme bruité et d'un signal. Ce signal permet de procéder automatiquement à un découpage adapté en temps des spectrogrammes et à une analyse en fréquence de chaque bande temporelle. La méthode mise en place exploite la distance de Fréchet comme critère de ressemblance. Une courbe de l'évolution de la fréquence plasma au cours du temps est ensuite reconstruite. Le logiciel développé permet l'étude systématique, dans certaines régions de la magnétosphère terrestre, des spectrogrammes de l'expérience WHISPER de la mission spatiale CLUSTER.

Abstract – This paper presents a new method in order to automatically identify the spacial plasma frequency. Data stem from two onboard satellite instruments. These are a two dimension spectrogram and a one dimension signal. From this last signal, the spectrogram is cutted out along the temporal axis in a adapted way, and a frequency analysis is performed. The method uses the Frechet distance as a form similarity criterium. Plasma frequency time evolution is then extracted from the spectrogram. A software can treat spectrograms produced by the WHISPER instrument of the CLUSTER mission in several Earth magnetosphere regions.

1. Introduction

La mission scientifique CLUSTER est dédiée à l'étude des frontières des régions clefs de la magnétosphère terrestre. Cette mission est constituée de quatre satellites identiques évoluant en constellation tétraédrique de forme et de dimension ajustable en fonction des objectifs scientifiques. Chaque satellite a embarqué 11 instruments scientifiques pour mesurer les paramètres du milieu : champ magnétique, ondes électromagnétiques, flux, température, espèces de particules, etc. L'originalité de ce projet [1] réside dans la possibilité de réaliser des mesures en trois dimensions et ainsi de séparer les variations temporelles des variations spatiales.

La mesure de la concentration électronique du plasma est un paramètre clef pour déterminer les propriétés des régions actives de l'environnement terrestre ainsi que les conditions de propagation des ondes qui rayonnent dans cet environnement. L'expérience Wave of High frequency and Sounder for Probing Electron density by Relaxation (WHISPER) mesure le champ électrique sur la gamme 2-80 kHz à l'aide d'antennes double-sphères de 88 m, puis en fait l'analyse fréquentielle en temps réel, par transformée de Fourier rapide. Un émetteur travaillant sur la gamme 4 – 80 kHz complète cet ensemble pour former un sondeur à relaxation, et ainsi obtenir, par identification des résonances propres du plasma, une mesure précise et fiable de la densité électronique totale du milieu [2]. Cette densité peut être déduite facilement de la fréquence de résonance « plasma » à l'aide de la formule suivante :

$$Fp_e = 9\sqrt{N_e} \quad (1)$$

où Fp_e est la fréquence « plasma » en kHz et N_e la densité électronique en cm^{-3} .

Des informations complémentaires concernant la mission CLUSTER sont disponibles sur le site de l'ESA (European Spatial Agency) [3] et les données de l'instrument WHISPER sur le site [4].

A chaque instant, l'expérience WHISPER construit un spectre correspondant à la réponse du plasma faisant suite à une excitation. Ces spectres sont juxtaposés et l'amplitude du signal pour chaque fréquence est codée en couleurs. Cet ensemble forme ainsi une image temps-fréquence appelée « spectrogramme ». Une série de points plus ou moins intenses correspondant aux fréquences de résonances du milieu apparaîtra et cette représentation nous permet de mettre en évidence les variations temporelles de la densité du plasma. L'enjeu de ce travail est donc d'être capable d'identifier et de reconnaître cette série de points associés à la résonance plasma, et ce, même dans les régions où la réponse du plasma est faible. Malheureusement, l'évolution temporelle de cette résonance plasma n'est pas forcément continue et peut présenter de brusques variations quand le satellite traverse des régions frontières comme l'illustre la figure 1. De plus, le spectre peut présenter des résonances d'autres natures perturbant la localisation de celle associée à la fréquence plasma. Pour palier au problème de la discontinuité et de l'aspect lacunaire des données nous sommes intéressés aux mesures de l'expérience EFW (Electric Field and Waves)

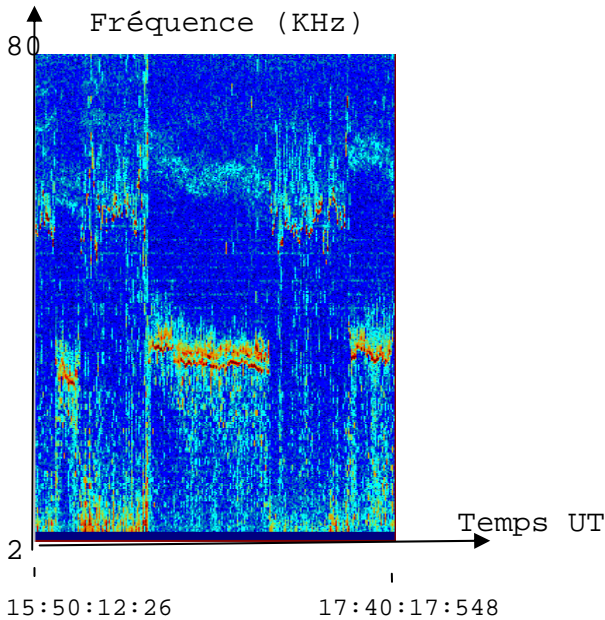


FIG. 1 : spectrogramme du 2002/05/13

qui mesure la différence de potentiel entre le corps du satellite, chargé positivement, et un ensemble de capteurs placés à ~ 45 m du corps du satellite en liaison électrique avec le plasma. Cette différence de potentiel, d'autant plus élevée que la densité électronique dans le milieu est faible, permet de suivre qualitativement les variations temporelles de la densité, sans pour autant en donner une estimation précise. En effet, s'il existe une relation empirique bi-univoque entre le « potentiel EFW » et la fréquence plasma, celle-ci dépend clairement du régime de plasma concerné et donc de la région étudiée [5]. Les données auxiliaires de l'expérience EFW nous aideront par conséquent à identifier la résonance « plasma » présente dans les spectrogrammes sondeur comme un « fil d'Ariane » nous renseignant sur le sens des fluctuations de la densité. Les mesures précises de cette quantité seront obtenues au mieux à la résolution temporelle de l'instrument WHISPER.

Notre problème se résume donc à extraire d'une image une série de points formée de maxima locaux dont les variations sont « semblables » à celle d'une courbe auxiliaire de référence.

Quelques travaux relatifs à la recherche de signatures dans ce type de données existent. On peut citer ceux de Galkin [6] qui identifie les fréquences caractéristiques du plasma dans la bande 3 KHz et 3 MHz à l'aide de l'instrument RPI (Radio Plasma Imager) embarqué sur le satellite IMAGE. Il met en œuvre un seuillage adaptatif sur le gradient de l'image et exploite les réseaux de neurones pour l'identification des fréquences caractéristiques. Mais, l'utilisation des réseaux de neurones nécessite une connaissance *a priori* de la signature temporelle. Les travaux de Trotignon [7] ont concerné uniquement un traitement spectre à spectre sans intégrer l'information de continuité temporelle et qui nécessitait une intervention manuelle constante. En raison du nombre considérable de données à traiter qui présentent des variations imprévisibles, il s'avère nécessaire de mettre au point une procédure de localisation de la fréquence plasma de manière automatique adaptée à une succession de régions bien particulières de l'espace (vent solaire, magnétogaine).

Dans ce contexte, nous proposons une nouvelle stratégie qui consiste à découper notre spectrogramme en intervalles temporels de manière à isoler les régions de l'espace dont les frontières sont reconnaissables par des variations brusques du potentiel. Ensuite, chacun de ces intervalles est ensuite redécoupé suivant la variable fréquentielle pour former des sous-images. Ce double découpage exploite l'information provenant de l'instrument EFW. Bien que ces mesures ne soient pas calibrées, elles nous donnent accès à une information précieuse sur les variations de la densité. Une mesure basée sur la distance de Fréchet entre le signal EFW et la détection des fréquences de résonance dans chaque sous-image permet la sélection automatique de la bande fréquentielle qui contient la fréquence plasma. L'évolution temporelle continue de cette fréquence est ensuite retrouvée par une mise en correspondance non rigide de l'information EFW sur nos données WHISPER.

Nous présentons ci-après le principe du partitionnement en sous-images des spectrogrammes. Nous rappellerons ensuite le principe de la distance de Fréchet et définirons les critères de sélection des bandes fréquentielles. Le principe de déformation de l'information EFW pour une adaptation à nos données sera ensuite explicité. L'ensemble sera complété par une analyse statistique des résultats obtenus sur une période de cinq mois.

2. Méthodes

Dans le spectrogramme, l'extraction de la courbe associée à l'évolution temporelle de la fréquence plasma se réalise en quatre étapes. La première est le découpage temps-fréquence de l'image basé sur les caractéristiques du signal EFW. Dans chaque bloc temps-fréquence, il est pratiqué une détection de la fréquence d'amplitude maximale au cours du temps. Il n'est conservé que les points significatifs au sens du rapport signal à bruit. Au terme de la deuxième phase, chaque bande temporelle est associée à une famille de courbes de moyenne nulle (il y a une courbe par bloc). Ensuite, à l'aide d'un critère de ressemblance, chaque courbe est comparée au signal EFW également recentré en zéro. La distance de Fréchet s'est avérée bien adaptée à notre problématique ; la série temporelle de points retenue est celle associée à la plus petite distance mesurée. Enfin, dans une quatrième étape, les points retenus sont utilisés pour calculer les deux paramètres définissant un modèle physique de calibration des données EFW, définissant ainsi une nouvelle courbe EFW localisée au sens des moindres carrés sur nos fréquences choisies. Nous définissons ensuite une bande centrée autour de ce nouveau signal de manière à récupérer toutes les fréquences maximales qui forment ainsi la courbe d'évolution temporelle de la fréquence plasma recherchée.

2.1 Partitionnement en sous-images

Le partitionnement des spectrogrammes repose sur les variations en amplitude du signal EFW. Il est recherché un découpage temporel qui définit des intervalles dans lesquels les variations en amplitude sont modérées. Le signal EFW correspondant aux données présentées en figure 1 est reproduit en figure 2.

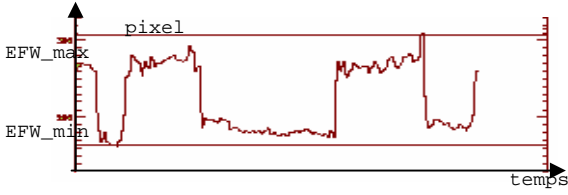


FIG. 2 : signal EFW

Sur ce signal, il est calculé une valeur continue comme suit :

$$EFW_c = \frac{EFW_{\max} + EFW_{\min}}{2}$$

Sur tout l'espace temporel considéré, les points du signal sont comparés à cette valeur continue. Il est ainsi défini une partition composée d'intervalles sur chacun desquels une suite de points a une amplitude supérieure ou inférieure à EFW_c . Néanmoins, il est introduit un nombre minimum de points consécutifs de manière à rendre significatif le critère de ressemblance entre ce signal et des fréquences de résonance de l'instrument WHISPER. De ce fait, l'algorithme fait la distinction entre un passage de frontière et une brusque variation du potentiel induit par l'apparition d'un phénomène physique de courte durée.

Le partitionnement fréquentiel se fait tout aussi simplement : la bande d'étude est recouverte par des sous-bandes fréquentielles de largeur $EFW_{\max} - EFW_{\min}$ et se chevauchant de moitié. En supposant que l'amplitude du potentiel sera de l'ordre de celle de la courbe de fréquence plasma sur chaque intervalle temporel, ces choix impliquent l'unicité de la bande fréquentielle qui contient cette dernière.

2.2 Distance de Fréchet

La distance Euclidienne permet de comparer deux objets ponctuels. Par contre, pour comparer deux objets évoluant suivant une variable comme le temps par exemple, il existe plusieurs types de distances dont les plus usitées sont la distance moyenne de déplacement ou la distance de Hausdorff [8]. Ces deux mesures nécessitent la prise en compte de l'ensemble des points de la courbe considérée. Notre problématique étant la mesure de ressemblance entre l'évolution temporelle des points détectés sur les spectrogrammes et le signal EFW, la distance de Fréchet est la mieux adaptée. En effet, elle représente une distance max entre deux lignes orientées [9]. Elle se définit de la manière suivante :

Toute polyligne orientée est équivalente à une application continue $f : [a, b] \rightarrow \mathcal{R}^2$ où $a, b \in \mathcal{R}$, $a < b$. La distance de Fréchet (d_f) est la suivante : Soit $f : [a, a'] \rightarrow \mathcal{R}^2$ et $g : [b, b'] \rightarrow \mathcal{R}^2$ deux polygones et $\| \cdot \|$ la norme euclidienne,

$$d_f(f, g) = \inf_{\substack{\alpha : [0,1] \rightarrow [a,a'] \\ \beta : [0,1] \rightarrow [b,b']}} \max_{t \in [0,1]} \|f(\alpha(t)) - g(\beta(t))\|$$

Une illustration intuitive de la distance de Fréchet est la suivante : un maître et son chien suivent deux chemins. Ils avancent ou s'arrêtent à volonté, indépendamment l'un de

l'autre, mais ils ne peuvent pas revenir sur leur pas. La distance de Fréchet entre ces deux chemins est la longueur minimale de la laisse qui permet de réaliser une progression de conserve satisfaisant ces conditions.

Par contre, le calcul de cette distance est complexe. Néanmoins, les travaux d'Eiter et Mannila [10] donnent un algorithme simple pour son approximation dont l'erreur maximale est majorée. Nous avons utilisé cet algorithme.

2.3 Principe de déformation

L'instrument EFW mesure le potentiel du milieu en Volts. Plusieurs études théoriques [5] ont contribué à définir un modèle reliant le potentiel et la fréquence plasma. Il s'exprime comme suit :

$$\alpha * \text{potentiel}^\beta \text{ avec } \beta < 0 \quad (2)$$

où α et β ajustent expérimentalement la densité au potentiel.

Les coefficients usuellement adoptés pour définir une courbe moyenne dans cette région sont :

$$\alpha = 200 \text{ et } \beta = -1.85.$$

3. Résultats expérimentaux

L'algorithme décrit précédemment a fait l'objet d'une implantation dans un logiciel de visualisation nommé WHIMAGE spécialement conçu pour faciliter le dépouillement des données WHISPER. Nous présenterons les résultats obtenus avec les données des figures 1 et 2. Une analyse des résultats sur cinq mois de données complète notre étude.

Dans le cas des données du 13 mai 2005, l'orbite du satellite SC1 traverse différentes régions. Le passage des frontières se voit tout aussi bien sur le spectrogramme de la figure 1 que sur le potentiel EFW de la figure 2.

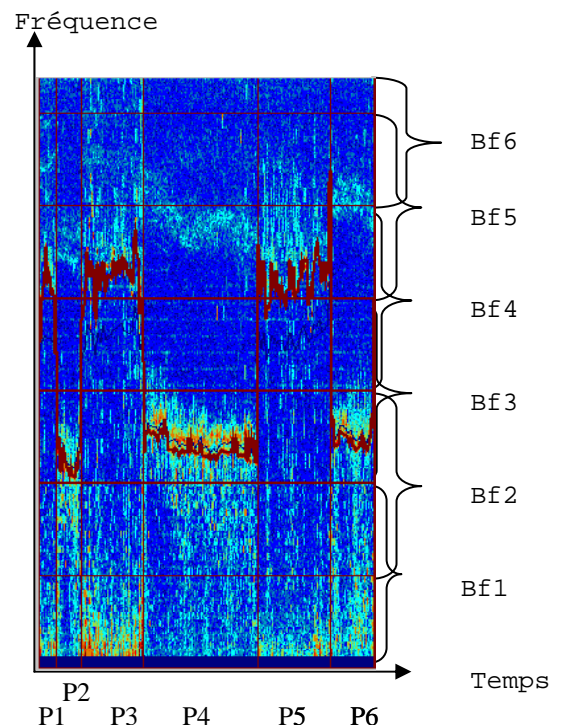


FIG. 3 : grille d'étude

La méthode simple exposée en 2.1 donne sans ambiguïté 36 sous-images ($P_i * BF_j$) représentées sur la figure 3 et dont les distances de Fréchet sont reportées dans le tableau 1.

TAB. 1 : distances de Fréchet

	P1	P2	P3	P4	P5	P6
BF6	12.5	106.1	45.8	76.7	47.0	94.6
BF5	7.4	69.8	40.1	30.7	16.0	86.9
BF4	8.0	59.8	21.0*	104.7	21.1	78.7
BF3	46.4	23.0	76.2	8.1	74.9	61.4
BF2	29.2	97.2	36.0	111.6	52.8	113.9
BF1	25.0	113.0	19.2	98.3	78.8	114.9

Le choix de la bande de fréquence est effectué sur les valeurs minimales de la distance de Fréchet mises en rouge dans le tableau 1 et la courbe résultat, corrigée en *, est affichée en rouge sur la figure 3.

Nous avons procédé au lancement en batch du logiciel sur cinq mois de données entre le 01/12/2003 et le 31/04/2004.

Un examen visuel des résultats a été effectué pour évaluer l'efficacité de notre méthode. Le logiciel a traité des périodes contenant 500 spectres. Dans 85% du temps de ces cinq mois de données la fréquence plasma a été reconnue correctement. Les échecs de détection sont attribués aux deux raisons principales suivantes :

(i) La méthode utilisée basée sur les variations de EFW amène dans certains cas un découpage mal adapté. Des variations brusques de la fréquence plasma peuvent être parfaitement visible sur les données WHISPER alors qu'elles n'apparaissent pas clairement sur la courbe de référence EFW. Ces différences localement significatives peuvent augmenter considérablement la distance de Fréchet et provoquer un choix erroné de la bande fréquentielle.

(ii) Dans certains cas, les mesures contiennent trop de points de bruit et provoquent une augmentation anormale de la distance de Fréchet bien que visuellement les courbes se ressemblent globalement. Dans cette situation, le choix de la distance de Fréchet n'est pas pertinent et une mesure de distance plus globale serait préférable.

Afin d'améliorer les résultats, un deuxième test a été fait en imposant les contraintes suivantes : on effectue une recherche sur un nombre limité de bandes fréquentielles autour de la courbe EFW par défaut. Cela permet de donner en cas d'échec un résultat pas trop éloigné de la courbe de résonance. De plus, cela permet d'intégrer en quelque sorte une dépendance entre les sous intervalles de temps jusqu'alors traités de manière indépendante. Nous ajoutons également un poids à chaque courbe qui est une fonction du nombre de maxima globaux qu'elle contient. Ces nouvelles propriétés; proximité de la courbe EFW et nombre de maxima globaux, attribuées aux courbes va nous permettre de définir des critères de sélection finale quand les distances de Fréchet sont voisines. Ces nouvelles contraintes ont permis d'améliorer significativement la reconnaissance de la

fréquence plasma. Nous estimons que sur 90 % du temps d'étude l'intervalle fréquentiel a été correctement choisi.

4. Conclusion et perspectives

En conclusion, la méthode donne de bons résultats. Une attention toute particulière est portée sur le découpage en temps et sur l'interprétation de la distance de Fréchet. La logique d'extraction proposée dans le paragraphe 2 bien que simple nous semble bien adaptée à nos données dans des régions où les frontières sont bien définies. L'outil développé permet actuellement une automatisation du dépouillement de données des quatre satellites CLUSTER et de fournir à la communauté scientifique une évaluation correcte de la densité électronique. Cette automatisation permettra d'effectuer des statistiques sur de nombreux passages dans des régions clés et de localiser plus finement les zones frontières et leurs fluctuations temporelles et spatiales. Néanmoins, d'autres méthodes sont en cours de développement pour les régions où la fréquence plasma se trouve mélangée en permanence avec d'autres résonances. En effet, le choix d'un maximum local dans une bande fréquentielle n'est plus valide. Des techniques de tout autre ordre sont actuellement à l'étude.

Références

- [1] C.T. Russel et R. Schmidt. *The Cluster and Phoenix Missions*. Edité par C.P. Escoubet, Ed. Kluwer Academic Publishers, 1997.
- [2] P. Décréau et al. and Whisper investigators. *WHISPER, a resonance sounder and wave analyser: performances and perspectives for the Cluster Mission*. Space Science Reviews 79: 157-193, 1997.
- [3] www.esa.int/science/cluster
- [4] www.whisper.cnrs-orleans.fr
- [5] A. Pedersen et al. *Four-point high time resolution information on electron densities by the electric field experiments (EFW) on Cluster*. Annales Geophysicae 19:1483-1489, 2001.
- [6] I.A. Galkin. *A pre-attentive vision model for automated data exploration applied to space plasma remote sensing data*. PhD thesis in computer science, University of Massachusetts Lowell, 2004.
- [7] J.G. Trotignon, P. Décréau, J.L. Rauch et al. *How to determine the thermal electron density and the magnetic field strength from the Cluster/Whisper observations around the Earth*. Annales Geophysicae, 19:1711-1720, 2001.
- [8] T. Devogele, Y. Riffault. *Alarime de précision pour les lignes de côte des ECDIS*, CoastGIS'99, Editeur Populus and Loubersac, Ed. IFREMER, Actes de colloques n°25, p 159-169, 2000.
- [9] M. Fréchet. *Sur quelques points du calcul fonctionnel*. Rendiconti del Circolo Mathematico di Palermo. 22 :1-74, 1906.
- [10] T. Eiter and K. Mannila. *Computing Discrete Fréchet Distance*. Rapport technique Christian Doppler Labor für Expertensysteme, Université Technique de Vienne, n° CD-TR 94/64.