HAL
open science

Vers une surveillance en temps réel des régions magnétosphériques à partir des radars cohérents HF SuperDARN

Guillaume Lointier

To cite this version:

Guillaume Lointier. Vers une surveillance en temps réel des régions magnétosphériques à partir des radars cohérents HF SuperDARN. Autre. Université d'Orléans, 2008. Français. NNT : . tel-00294447

HAL Id: tel-00294447

https://theses.hal.science/tel-00294447v1

Submitted on 9 Jul 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE PRESENTEE A L'UNIVERSITE D'ORLEANS
POUR OBTENIR LE GRADE DE

DOCTEUR DE L'UNIVERSITE D'ORLEANS

Discipline : physique des plasmas

par

Guillaume LOINTIER

Vers une surveillance en temps réel des régions magnétosphériques à partir des radars cohérents HF SuperDARN

Soutenue publiquement

le 31 mars 2008 à 13h30

Lieu: Amphithéâtre Charles Sadron
(Campus CNRS/ORLEANS)

Membres du jury :

R. André	Examinateur	Ingénieur, CEA/CESTA, Le Barp
A. Bourdillon	Président de Jury	Professeur à l'Université de Rennes 1, IETR
T. Dudok de Wit	Directeur de thèse	Professeur à l'Université d'Orléans, LPCE
C. Hanuise	Co-directeur de thèse	Directeur de recherche, LPCE, Orléans
C. Lathuillère	Rapporteur	Directeur de recherche, LPG, Grenoble
J.-P. Saint Maurice	Rapporteur	Professeur à l'Université de Saskatchewan, Saskatoon

Cette thèse à été réalisée avec le concours du Conseil Régional du Centre.

Remerciements

Je tiens tout d'abord à exprimer ma plus profonde gratitude à Thierry Dudok de Wit, directeur de thèse, qui a su se rendre disponible tout au long de ma thèse et ce malgré un emploi du temps digne d'un ministre. Merci pour m'avoir tout appris (parfois non sans malice), pour ton regard critique que tu as porté sur mon travail, pour m'avoir bousculé quand il le fallait et aussi pour ta franchise. Tout au long de ces trois années, j'ai eu le plaisir de partager ta rigueur dans le travail et le recul que tu portes sur la recherche en général. Thierry, je te dois beaucoup. Je remercie également chaleureusement Christian Hanuise, co-directeur de thèse, qui a su répondre présent, malgré la distance et ses nombreuses obligations à l'étranger, pour les moments décisifs de ma thèse et pour m'avoir prêter son bureau pendant ces trois dernières années (il m'a bien rendu service, maintenant je peux te le rendre !!!).

Je tiens également à remercier Pierre-Louis Blelly et Michel Tagger, directeurs du LPCE, pour m'avoir accueilli au sein de leur laboratoire et fait en sorte que ma thèse se déroule dans de bonnes conditions.

Je voudrais ensuite exprimer mes plus profonds remerciements à Chantal Lathuillère et à Jean-Pierre Saint Maurice pour avoir accepté de rapporter cette thèse, ainsi qu'à Raphaël André pour avoir participé à ce jury. Je vous suis très reconnaissant du temps que vous avez passé à relire, commenter, raturer et autres hiéroglyphes (il se reconnaîtra)... sur mon mémoire, mais également pour les nombreuses discussions qui ont suivies. Jean-Pierre, je me souviendrai longtemps de cette partie de beach soccer sur les plages de Chincoteague.

Et comment ne pas remercier les nombreuses personnes, jeunes et moins jeunes troublions du laboratoire avec lesquelles j'ai partagé d'excellents moments. Plus particulièrement Gwenaël avec qui j'ai pu longuement discuter de mon avenir et surtout

de choses plus ou moins sérieuses (il faut bien le dire...) et pour les parties de squash endiablées (tu ne perds rien pour attendre). J'ai gagné un véritable ami ; Xavier pour ton soutien, tes encouragements, ton aide et tes précieux conseils pour la rédaction de mon manuscrit (et il en avait besoin) ; Aurélie pour avoir partagé mon bureau. Matthieu pour ta gaieté et ton espièglerie car tu es arrivé au bon moment. Jean-Yves, tes conseils en informatique et nos longues discussions le soir vont me manquer. Et je ne peux certainement pas oublier ma copine de DEA, Farida. Je m'excuse pour toutes les personnes que je n'ai pas citées mais la liste est trop longue. Vous avez tous su rendre ces années tellement plus agréables, merci encore.

Je voudrais terminer par deux personnes qui me sont chères. Mon amie avec qui j'ai partagé toutes les joies d'une thèse. J'ai peut-être soutenu ma thèse mais tu m'as supporté, cela mérite au moins les félicitations du jury. Je ne saurai oublier Orphée qui m'a accompagné (et aidé) durant toutes ces longues nuits de rédaction.

Résumé

L'un des enjeux dans la compréhension des relations Soleil-Terre est l'étude de la dynamique des régions internes de la magnétosphère.

Dans ce contexte, l'objectif de mon travail a été de poser les bases d'un modèle opérationnel de détection et de suivi des régions frontières de la magnétosphère à partir du réseau de radars SuperDARN, qui sonde l'ionosphère à haute latitude. L'élaboration d'un tel modèle requiert une réduction de données. Pour cela, une méthode de décomposition en valeurs singulières (SVD) a été appliquée sur la mesure brute (une fonction d'autocorrélation) afin de définir trois nouveaux paramètres statistiques. L'interprétation de ces trois nouveaux paramètres montre que cette approche, bien qu'empirique, offre une description bien plus complète des échos radar que les modèles physiques habituellement utilisés. L'utilisation de ces paramètres avec une méthode de décision Bayésienne permet d'améliorer la détection de la frontière des lignes de champ géomagnétiques ouvertes/fermées. L'introduction d'un formalisme Bayésien comporte plusieurs avantages : il permet de valider le résultat en estimant une erreur sur la localisation, et de plus, il facilite l'introduction de nouvelles connaissances provenant de différents instruments. Ceci est loin d'être négligeable pour compléter les observations des radars SuperDARN. Par ailleurs, les propriétés de ce nouveau modèle ont permis l'élaboration d'un nouvel algorithme de prétraitement et d'analyse des mesures brutes.

Mots clefs : météorologie de l'espace, couplage magnétosphère-ionosphère, analyse statistique multivariée, classification Bayésienne, radars cohérents HF.

Abstract

One of the key issues in the study of the Sun-Earth connexion is the monitoring of the dynamics of the magnetospheric regions. In this context, the objective of my work is to lay the foundations for an operational model for monitoring the ionospheric projection of magnetospheric boundaries in near real-time. This model is based on the SuperDARN radar network, which was designed for the study of high latitude ionospheric convection. The elaboration of such a model requires a data reduction step. We use the Singular Value Decomposition (SVD) method to reduce the raw data (autocorrelation functions) to a set of three statistical parameters. The interpretation of these empirical parameters shows that they provide a more complete description of the radar echoes than the routinely employed physical parameters. By coupling these parameters to a Bayesian decision method, we show how the detection of the open/closed magnetic field line boundary can be improved. The Bayesian formalism offers several advantages : it provides confidence intervals for the location of the boundary and it can handle prior information, as well as constraints. In addition to that, a new strategy has been introduced for correcting outlier values that often plague the measurement of the autocorrelation function.

Keywords : space weather, magnetosphere-ionosphere coupling, multivariate statistical analysis, Bayesian classification, HF coherent radars.

Table des matières

Table des matières viii

Introduction 1

1 Contexte géophysique et expérimental 5

1.1 L'environnement terrestre 6

1.1.1 L'ionosphère 6

1.1.2 La magnétosphère 11

1.2 Couplage du système vent solaire-magnétosphère-ionosphère 14

1.2.1 Interaction vent solaire-magnétosphère 14

1.2.2 Dynamique du système magnétosphère-ionosphère 14

1.2.3 Projection ionosphérique des régions magnétosphériques 17

1.3 Surveillance des régions magnétosphériques : les moyens expérimentaux 19

1.3.1 Le réseau de radars cohérents HF SuperDARN 22

1.3.2 Principe de fonctionnement 23

1.3.2.1 Principe de mesure 23

1.3.2.2 Caractérisation de la fonction d'autocorrélation (ACF) 25

1.3.3 Les produits SuperDARN 28

1.3.3.1 Cartographie de la convection ionosphérique 28

1.3.3.2 Identifications des régions magnétosphériques : les prémices 28

2 Méthode d'analyse des mesures SuperDARN 33

2.1 Une approche statistique pour une description «temps réel» 34

2.1.1 Position du problème 34

2.1.2 Méthodologie 36

2.1.3 Décomposition en Valeurs Singulières 39

2.1.3.1 Définition algébrique 40
2.1.3.2 Définition variationnelle 42
2.1.3.3 Illustration géométrique de la SVD 43
2.2 Mise en oeuvre de la méthode de réduction statistique de l'ACF 45
2.2.1 Constitution d'une base de données 45
2.2.2 Réduction de l'ACF 49
2.3 Validation de la méthode 53
2.3.1 Influence de l'activité géomagnétique 53
2.3.2 Dépendance selon le secteur MLT 59
2.3.3 Parenthèse sur la représentativité des échos radar sélectionnés . . 62
2.4 Conclusion 64

3 Interprétation de la réduction statistique des ACFs **67**
3.1 Une interprétation physique des modes 68
3.1.1 Comparaison des paramètres $\{f_\alpha\}$ de la SVD avec les paramètres usuels issus de l'algorithme FitACF 69
3.1.1.1 Relation entre les paramètres $\{f_\alpha/f_1\}$ et la largeur spectrale 70
3.1.1.2 Relation entre les paramètres $\{f_\alpha/f_1\}$ et la forme des ACFs 73
3.1.2 Modélisation du comportement des paramètres SVD : $\{f_2/f_1, f_3/f_1\}$ 75
3.1.2.1 Choix d'un modèle analytique pour la simulation des ACFs 75
3.1.2.2 Méthode 76
3.1.2.3 Modélisation par mode SVD 78
3.2 Classification des échos radar : région E ou région F ? 81
3.3 Aspect géophysique de la réduction des ACFs 82
3.3.1 Comparaison des distributions polaires des paramètres issus de «FitACF» et de la SVD 84
3.3.2 Caractérisation de l'erreur des modèles 87
3.3.2.1 Distribution des écarts quadratiques relatifs des modèles 88
3.3.2.2 Comparaison des modèles 90
3.3.3 Interprétation de deux signatures particulières : le 12 février 2003 92
3.3.3.1 Evénement A 92
3.3.3.2 Evénement B 96
3.4 Détection de la frontière des lignes de champ ouvertes/fermées 99

3.4.1 Vers une Inférence Bayésienne de l'OCB 102
3.4.2 Etude de cas : 10 octobre 1999 103
3.4.2.1 Approximation de la vraisemblance 103
3.4.2.2 Projection de la frontière Bayésienne 106
3.4.3 Perspectives : aspect dynamique et amélioration 109
3.5 Conclusion 112

4 FitSVD : un nouvel algorithme d'analyse des mesures SuperDARN **115**
4.1 Schéma d'analyse des ACFs : FitSVD 116
4.2 Mise en place de la procédure 118
4.2.1 Détermination des seuils 118
4.2.2 Exemples d'application 120
4.3 Comparaison avec l'algorithme FitACF 122
4.3.1 Nature des retards aberrants détectés 122
4.3.2 Influence sur l'estimation des paramètres : $\{f_2/f_1, f_3/f_1\}$ et $\{\tau_e, n_\tau\}$ 125
4.4 Validation de la procédure 126
4.4.1 Validation d'un critère d'arrêt 127
4.4.2 Performance : détection des retards erronés 130
4.5 Conclusion 131

Conclusions et perspectives **133**

A Principe de fonctionnement des radars SuperDARN **139**
A.1 Radar multi-impulsions et fonctions d'autocorrélation 139
A.2 Identification des sources de bruits et d'interférences 141
A.2.1 Sources «naturelles» 141
A.2.2 Sources «instrumentales» 142
A.3 Algorithme «FitACF» : traitement des ACFs 143
A.3.1 Estimation des niveaux de fluctuations 144
A.3.2 Classification des «mauvais» retards (ou bad lags) 145

Actes de publication & communication **149**

Bibliographie **151**

Introduction

On peut considérer que la première étape vers une science des relations Soleil-Terre a été réalisée par Edmund Halley. Pour expliquer le développement spectaculaire d'une aurore en Europe en mars 1716, il suggère que des particules se mouvant le long des lignes de champ géomagnétique sont à l'origine de son existence. En 1747, Anders Celsius et Olav Hioters mettent en évidence une corrélation entre le moment de l'apparition des aurores et la variation d'une aiguille de boussole. En comparant les mesures simultanées du champ géomagnétique à Londres et à Uppsala, ils réalisent que le phénomène étudié est relié à un processus d'échelle planétaire. Richard Carrington démontre en 1859, pour la première fois, un lien entre le Soleil et la Terre, à court terme avec la corrélation temporelle entre les éruptions solaires et l'apparition des aurores, et à long terme, avec la variation de 11 années entre le nombre de taches solaires et de la fréquence de déclenchement des perturbations géomagnétiques et des aurores.

A partir des années 60, l'arrivée des mesures in-situ et d'autres moyens de sondage à distance de l'espace ont permis de mieux appréhender l'influence du Soleil sur l'environnement terrestre pour différentes échelles temporelles : 11 ans pour le cycle d'activité solaire, 27 jours pour la période de rotation du soleil, de l'ordre de la journée pour la propagation dans le milieu interplanétaire des éjections de masse coronale et de l'ordre de l'heure pour la propagation des particules relativistes. Les différents mécanismes d'échanges entre le Soleil et l'environnement terrestre, ainsi que le développement des perturbations observées au niveau de la Terre sont aujourd'hui bien connus. Par contre, il est encore difficile de prédire le déclenchement et l'amplitude de ces événements.

L'irruption toujours plus pressante des nouvelles technologies dans notre société rend celle-ci de plus en plus sensible aux perturbations de type électromagnétique, et au bombardement d'électrons et d'ions énergétiques. L'expansion spatiale et la magnitude de ces événements sont extrêmement variables. Ils peuvent affecter le fonctionnement

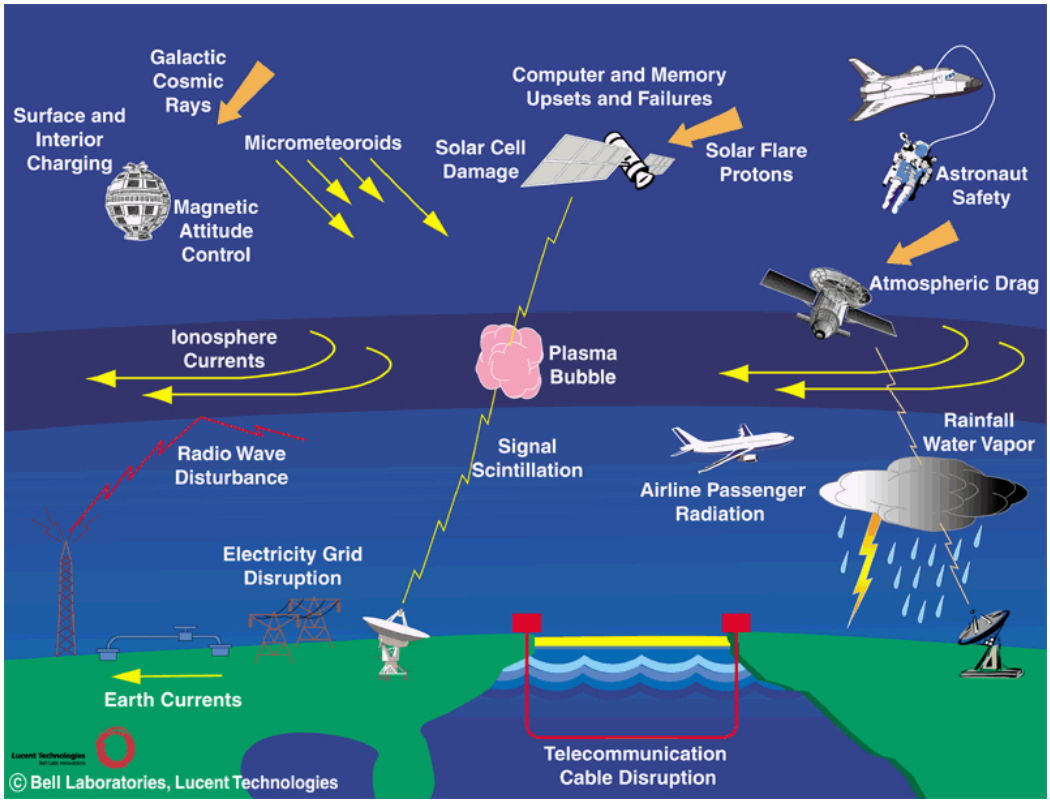


FIG. 1 – *Illustration de la chaîne des processus affectant l'environnement proche de la Terre (D'après L. J. Lanzerotti, cf Kamide and Chian, 2007, chap. 17).*

des satellites, perturber les systèmes de télé-communications au sol et dans l'espace, et les systèmes de navigation. Au sol, ils peuvent endommager les installations électriques, ou encore accélérer la corrosion des pipelines. L'exemple le plus marquant, et le plus couramment cité dans la littérature, est certainement l'événement du 13 mars 1989. Un orage magnétique particulièrement intense a causé une rupture du réseau électrique privant ainsi d'électricité une partie de Montréal et du Québec pendant près de 9 heures. La figure 1 illustre les différents effets de ces mécanismes perturbateurs sur l'environnement terrestre. Du besoin de prédire et de quantifier l'impact de ces perturbations sur l'environnement terrestre, sur les technologies et sur l'économie est née une nouvelle discipline la «météorologie de l'espace». Cette nouvelle discipline répond aux besoins de la communauté scientifique afin d'améliorer notre compréhension du système Soleil-Terre et répond aux besoins croissants de notre société en termes de prédiction.

Le système magnétosphère-ionosphère n'est pas statique et les réponses de l'ionosphère face aux perturbations subies par la magnétosphère sont loin d'être linéaires et globales. La compréhension de la dynamique des différentes régions constituant la magnétosphère nécessite cependant une approche globale. Les satellites n'apportent qu'une vision partielle tandis que les instruments au sol sondant l'ionosphère à hautes latitudes offrent une vision plus étendue du système magnétosphère-ionosphère. Notamment, le réseau de radars SuperDARN, avec sa large couverture spatiale des régions aurorales et polaires de l'ionosphère, offre une opportunité unique d'étudier le couplage magnétosphère-ionosphère à l'échelle planétaire et ce en temps réel.

Dans ce contexte, cette thèse propose les fondations nécessaires pour un nouveau schéma d'identification et de suivi en temps réel des régions magnétosphériques à partir du réseau de radars SuperDARN. Ce mémoire s'articule donc de la manière suivante. Dans le 1er chapitre, nous allons tout d'abord décrire les différents processus liés aux transferts d'énergie, de moment et de matière entre le vent solaire et la magnétosphère et nous aborderons les répercussions de ces mécanismes sur la dynamique de l'ionosphère. Nous terminerons ce chapitre par une description des radars SuperDARN. L'élaboration d'un modèle opérationnel exige que les quantités intéressantes pour notre objectif soient facilement accessibles. En effet, les modèles physiques ne permettent pas forcément d'y accéder directement et donc rapidement. Des méthodes de substitution doivent alors être employées. Dans le chapitre 2, nous montrerons comment une méthode de réduction de données appliquée sur les mesures brutes des radars SuperDARN permet de définir un nouveau jeu de paramètres statistiquement représentatif des conditions générales d'observation. Nous testerons la robustesse de cette méthode en fonction de différents facteurs (activité géomagnétique, dépendance en fonction du temps magnétique local). Le chapitre 3 se focalisera sur l'interprétation de ces nouveaux paramètres en termes de quantité physique et de leur signature géophysique. De cette interprétation, nous déduisons un nouveau jeu de paramètres empiriques qui améliore sensiblement la description de la mesure par rapport aux paramètres usuels. L'algorithme FitACF qui est l'algorithme officiel de traitement des mesures SuperDARN, est utilisé afin d'extraire des paramètres physiques de l'ionosphère. Notre objectif est différent de celui poursuivi

avec l'algorithme FitACF, puisque nous désirons élaborer un modèle opérationnel de suivi des régions magnétosphériques. Il est par conséquent nécessaire de mettre en place une nouvelle stratégie d'ajustement des mesures brutes. Le chapitre 4 sera consacré à l'élaboration d'un nouvel algorithme de traitement des données SuperDARN. Ce nouvel algorithme s'avère être une bonne alternative au traitement de la mesure brute par rapport à l'algorithme «FitACF», notamment pour la détection des mesures erronées.

1

Contexte géophysique et expérimental

1.1 L'environnement terrestre 6
1.1.1 L'ionosphère 6
1.1.2 La magnétosphère 11
1.2 Couplage du système vent solaire-magnétosphère-ionosphère 14
1.2.1 Interaction vent solaire-magnétosphère 14
1.2.2 Dynamique du système magnétosphère-ionosphère 14
1.2.3 Projection ionosphérique des régions magnétosphériques 17
1.3 Surveillance des régions magnétosphériques : les moyens expérimentaux 19
1.3.1 Le réseau de radars cohérents HF SuperDARN 22
1.3.2 Principe de fonctionnement 23
1.3.2.1 Principe de mesure 23
1.3.2.2 Caractérisation de la fonction d'autocorrélation (ACF) . 25
1.3.3 Les produits SuperDARN 28
1.3.3.1 Cartographie de la convection ionosphérique 28
1.3.3.2 Identifications des régions magnétosphériques : les prémices 28

La compréhension des relations Soleil-Terre nécessite l'étude de nombreux processus physiques qui affectent notamment la dynamique du système magnétosphère-ionosphère. La réponse de la magnétosphère induite par les sursauts d'activité solaire (éruptions solaires, éjections de masse coronale...) n'est pas globale et entraîne des perturbations de différentes natures qui peuvent être observées jusqu'à la surface de Terre. Afin de mieux comprendre l'impact sur notre société des mécanismes d'échange entre le milieu interplanétaire et l'environnement terrestre, il est crucial de développer des

outils capables de surveiller à l'échelle planétaire, les différentes régions qui constituent la magnétosphère.

Pour cela, nous nous sommes intéressés dans ce travail de thèse à leur signature dans l'ionosphère de haute latitude. L'ionosphère est un site d'observation privilégié car aux latitudes aurorale et polaires, cette région est particulièrement sensible au couplage entre le vent solaire et la magnétosphère. A cause notamment du caractère dipolaire du champ géomagnétique, chacune des régions de la magnétosphère se projette le long des lignes de force dans l'ionosphère. L'étude de la dynamique de l'ionosphère à haute latitude offre donc une opportunité unique de suivre les frontières magnétosphériques à une échelle quasi-globale.

Nous débuterons ce chapitre introductif par la description des différentes régions concernées par notre étude. Tout d'abord, nous décrirons la région ionosphérique car c'est à partir de cette dernière que le travail d'identification des régions magnétosphériques est effectué. Ensuite, nous discuterons de la structure interne de la magnétosphère. Nous aborderons l'influence du milieu interplanétaire sur la dynamique du système composé de la magnétosphère et de l'ionosphère, en introduisant les différents mécanismes d'échange d'énergie entre ces deux régions. Et nous terminerons ce chapitre par une description des radars cohérents HF SuperDARN et nous discuterons des possibilités qu'offre cet instrument pour identifier et suivre en temps (quasi) réel la projection dans l'ionosphère des régions magnétosphériques.

1.1 L'environnement terrestre

1.1.1 L'ionosphère

L'ionosphère est la portion ionisée de la couche supérieure de l'atmosphère. Elle forme la couche limite entre l'atmosphère et l'espace. Elle occupe une enveloppe sphérique qui s'étend de 60 km à environ 1000 km au-dessus de la surface de la Terre. La première mise en évidence de la présence d'une ionosphère remonte à la fin du XIXe siècle. C'est en 1882 que B. Stewart découvre une région conductrice autour de 100 km d'altitude. Il conclut que la haute atmosphère est le site le plus probable d'un courant électrique produisant une variation journalière du champ magnétique mesuré à la surface de la Terre. Ce n'est que quelques années plus tard, avec l'avènement des techniques de transmission radio, que la présence d'une ionosphère hautement conductrice a été confirmée. Tout d'abord en 1902, A. E. Kennelly et O. Heavyside ont supposé l'existence d'une telle région pour expliquer la première transmission radio transatlantique

par réflexion des ondes radio sur cette couche conductrice. Puis en 1925, E. V. Appleton démontre sa présence en étudiant le temps de propagation d'impulsions radio émises verticalement afin de déduire l'altitude des couches électriquement réfléchissantes. C'est à lui que l'on doit aujourd'hui la classification de l'ionosphère en différentes strates : les couches D, E et F illustrées à la figure 1.1.

La formation d'une ionosphère requiert deux ingrédients : une atmosphère neutre et des sources d'ionisation. La source principale d'ionisation est la photo-dissociation des espèces neutres par les radiations ultraviolet (UV), ultraviolet extrême (EUV) et X provenant du Soleil. Ce rayonnement excite les espèces neutres des couches supérieures de l'atmosphère libérant ainsi un électron et un ion positif

$$A + h\nu \Rightarrow A^+ + e^- \ , \tag{1.1}$$

où A^+ est l'ion positif associé à l'espèce chimique A, h la constante de Planck et ν la fréquence de la radiation. Les précipitations de particules énergétiques contribuent également au processus d'ionisation pour certaines régions de l'ionosphère. Ce processus peut être représenté par l'équation suivante :

$$A + e^{-*} \Rightarrow e^- + A^+ + e^- \ . \tag{1.2}$$

Où pour cet exemple, l'électron e^{-*} correspond à la particule pénétrante. Une partie de son énergie est absorbée lors de la collision avec l'espèce neutre A. Si l'énergie absorbée est suffisante, la collision libère un électron pour former un ion associé à l'atome ou à la molécule neutre. Ces particules ionisantes peuvent être soit des rayons cosmiques provenant du milieu interplanétaire, soit des particules énergétiques provenant du vent solaire, de la magnétosphère ou encore de l'ionosphère elle-même.

Bien que ces deux processus expliquent assez bien la production primaire de paire ion-électron, le processus d'ionisation ne s'arrête pas là et une cascade peut être observée. Les ions primaires subissent des réactions chimiques avec les neutres, se recombinent entre eux et avec les électrons, produisant notamment des ions secondaires. L'ensemble des différentes réactions chimiques (primaires et secondaires) avec les diverses espèces atmosphériques conduit à la présence majoritaire dans l'ionosphère des ions O_2^+, O^+ et NO^+.

Parmi les différents mécanismes d'échange entre les ions, les électrons et les neutres, trois types de réactions secondaires dominent :

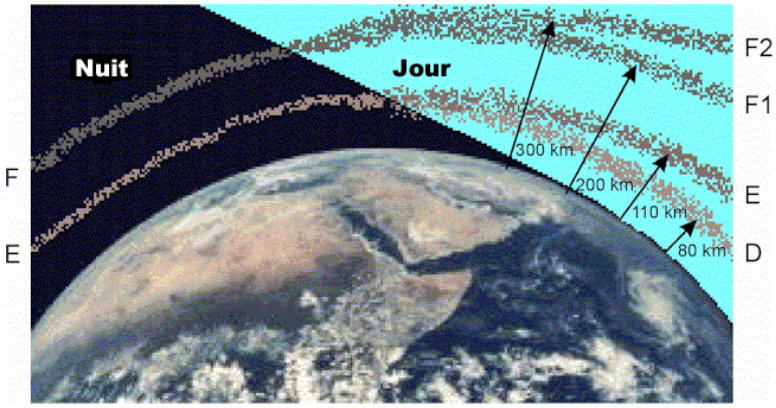


FIG. 1.1 – *Vue d'artiste des différentes couches constituant l'ionosphère.*

– par échange de charge entre un ions et un neutre

$$A^+ + B \Rightarrow A + B^+ \ , \tag{1.3}$$

– par échange de charge ionique

$$A^+ + BC \Rightarrow AB^+ + C \tag{1.4}$$

$$AB^+ + C \Rightarrow A + BC^+ \ , \tag{1.5}$$

– ou encore par recombinaison dissociative et radiative avec un électron, qui tendent à faire disparaître les espèces ionisées,

$$AB^+ + e^- \Rightarrow A + B \ , \tag{1.6}$$

$$A^+ + e^- \Rightarrow A + h\nu \ . \tag{1.7}$$

Les ions secondaires, qui sont principalement produits, concernent les espèces N^+, NO^+ et H^+. L'ensemble des ions (primaires et secondaires) sont transportés sous l'effet du vent neutre, soit vers les plus hautes altitudes, soit vers les plus basses altitudes. Les effets de diffusion et de transport sont fortement contraints par le champ magnétique intrinsèque à la Terre, qui peut être considéré comme dipolaire à l'altitude de l'ionosphère.

A cause des mécanismes d'ionisation, la production des ions est fortement contrôlée par la présence des espèces neutres dans l'atmosphère. Le champ gravitationnel de la

Terre qui est la source dominante de la diffusion des molécules de l'atmosphère, tend à séparer les différentes familles de composés atmosphériques en fonction de leur masse. Les espèces les plus légères diffusent vers les plus hautes altitudes. Chacune d'elles est en équilibre hydrostatique. Le profil en altitude de la concentration des espèces neutres suit une loi barométrique. Ainsi les molécules N_2 et O_2 sont les espèces majoritaires à basse altitude (< 200 km). Les espèces plus légères, comme par exemple l'oxygène atomique O, sont majoritaires entre 200 et 600 km. Finalement, l'espèce la plus légère, H, devient le composé prépondérant au-delà de 600 km d'altitude. Comme les réactions ne sont pas d'égales importances, l'ionosphère est stratifiée et est verticalement dépendante des processus chimiques qui dominent. Les couches D, E, F1 et F2 sont définies par leur altitude, composition, dynamique et par les densités électroniques. Les principales caractéristiques de ces régions sont résumées dans le tableau 1.1.

Tab. 1.1 – *Propriété de l'ionosphère côté jour.*

Région	Altitude (km)	Composition ionique	Densité électronique
D	$60 - 90$	O_2^-, NO^+, O_2^+, ions négatifs	10^8 - 10^{10} m^{-3}
E	$90 - 150$	O^+, O_2^+, NO^+	$\sim 10^{11}$ m^{-3}
F1	$150 - 250$	O^+, O_2^+, NO^+, N_2^+	10^{11} - 10^{12} m^{-3}
F2	$250 - 600$	O^+, O_2^+, H^+, He^+	$\sim 10^{12}$ m^{-3}

En région **D** ($60 - 90$ km), l'ionisation est faible car le rayonnement UV pénètre peu à ces altitudes. La densité électronique est de l'ordre de 10^8 - 10^{10} m^{-3}. C'est une région où l'atmosphère est suffisamment dense pour permettre la capture rapide des électrons par attachement électronique pour donner des ions négatifs, principalement des ions O_2^-,

$$O_2 + e^- + M \Rightarrow O_2^- + M \ , \tag{1.8}$$

où M désigne les espèces N_2 et O_2. D'autres processus chimiques plus complexes, impliquant les molécules CO_2 et NO, transforment les ions négatifs primaires en ions négatifs plus lourds NO_x^-, CO_x^-. La région D est une région où les ions positifs, négatifs et les électrons co-habitent à des niveaux de concentration suffisants pour assurer la quasi-neutralité. Côté nuit, les électrons tendent à disparaître (avec la source principale d'ionisation, le rayonnement UV, EUV et X) et la quasi-neutralité est assurée par la balance entre les ions positifs et négatifs. Ces processus de recombinaison sont rapides ce qui provoque la disparition de cette région.

La région **E** se situe entre 90 et 150 km, le plasma y est faiblement ionisé et les

collisions entre les particules chargées et neutres restent importantes. La densité électronique est plus élevée, de l'ordre de $\sim 10^{11}$ m^{-3}. Les ions O_2^+ et NO^+ sont majoritaires. L'ion O_2^+ est produit principalement par photo-ionisation de la molécule O_2 et par échange de charge avec N_2^+. L'ion NO^+ provient de l'échange de charge ionique avec les ions N_2^+ et O^+.

Les électrons deviennent magnétisés à partir de 90 km d'altitude, tandis que la fréquence de collision ion-neutre reste suffisamment élevée pour que les ions restent démagnétisés. Le comportement différentiel entre les ions et les électrons se traduit par la formation d'un courant électrique. L'électrodynamique de cette région dépend fortement de la latitude à cause de la topologie du champ magnétique terrestre. Ainsi à haute latitude, la circulation de courant horizontaux permet la fermeture des courants magnétosphériques et par conséquent favorise les échanges entre la magnétosphère et l'ionosphère.

En région **F** ($150-600$ km), les ions O_2^+, O^+ et NO^+ dominent. L'ion O_2^+ est produit principalement par photo-ionisation de son parent O_2 et par échange de charge avec N_2^+. NO^+ est ou sous-produit de la chimie. Il est principalement produit par échange de charge avec les ions O^+ et N_2^+. Quant à l'ion O^+, il devient prédominent à partir de 250 km, lorque la concentration de N_2 n'est plus suffisamment élevée pour contre-balancer la formation de l'ion O^+.

A partir de 150 km, les processus de transport deviennent importants. En effet, le plasma est faiblement collisionnel. Les ions et les électrons sont magnétisés et sont donc contrôlés par le champ électrique issu du couplage entre le vent solaire et le champ magnétique terrestre. Le maximum de densité électronique atteint $\sim 10^{12}$ m^{-3} autour de 300 km d'altitude. Généralement, on sépare la région **F** du côté jour, en deux sous-régions. La région **F1** et la région **F2**, dont l'interface est situé autour de 250 km, marquent la transition entre les ions atomiques O^+ et les ions moléculaires NO^+. La nuit, ces deux régions se fondent en une seule.

L'ionosphère est la résultante d'un équilibre entre les mécanismes d'ionisation et de recombinaison des différentes espèces chimiques et par conséquent est fortement liée au rayonnement solaire. Côté jour, les mécanismes d'ionisation des couches supérieures de l'atmosphère sont alors prépondérants. Cela se traduit par une expansion plus importante en altitude de l'ionosphère.

Dans des conditions géomagnétiques calmes, la dynamique à basses latitudes de l'ionosphère est principalement contrôlée par le mécanisme de chauffage par le Soleil du côté jour. Un vent neutre tend à souffler globalement du point sub-solaire vers les régions plus froides, côté nuit. A de plus hautes latitudes, cette dynamique est fortement

contrainte par la géométrie du champ magnétique terrestre qui réagit à l'activité du milieu interplanétaire et à la dynamique de la magnétosphère. Par exemple, on peut noter la formation de cellules de convection, dont la présence est la conséquence directe de l'écoulement du vent solaire autour de la magnétosphère (*cf* § 1.2.2).

L'étude des différents mécanismes, par exemple l'étude des précipitations de particules énergétiques ou de la dynamique de cette région, permet d'appréhender de manière globale les différentes réponses de l'environnement terrestre aux sursauts d'activités du Soleil.

1.1.2 La magnétosphère

La magnétosphère est une cavité née de l'interaction du vent solaire et du champ magnétique interplanétaire (IMF[1]) avec le champ magnétique intrinsèque de la Terre. Au voisinage de l'orbite terrestre, le vent solaire est un plasma fortement ionisé, non-collisionnel et gelé dans le champ magnétique (MHD idéale) possédant une vitesse d'écoulement supersonique et super-alfvénique. Lorsque le vent solaire aborde l'obstacle de l'environnement terrestre, l'écoulement du plasma est entravé par une *onde de choc*[2]. Le nez du choc est situé à environ $3-4$ R_T en amont de la magnétosphère (le rayon de la Terre $R_T = 6370$ km). En traversant ce choc, le plasma est freiné, comprimé et chauffé à des vitesses subsoniques dans une région appelée *magnétogaine*[3]. Le plasma contourne ensuite les frontières externes du champ géomagnétique, *i.e.* la magnétopause, avant de retrouver sa configuration initiale (*i.e* un écoulement supersonique et super-alfvénique) à partir d'une distance d'environ $100 - 200$ R_T du côté nuit. La région séparant le plasma magnétisé du vent solaire dans la magnétogaine et celui confiné dans le champ magnétique terrestre s'appelle la *magnétopause*. La magnétopause est généralement très fine (~ 100 km) et sa position dépend des conditions du milieu interplanétaire. La figure 1.2 illustre la topologie de la magnétosphère.

La morphologie de la magnétosphère résulte d'un équilibre (variable) entre la pression dynamique du vent solaire et celle du champ géomagnétique. Celui-ci est initialement un champ dipolaire d'une intensité d'environ ~ 50000 nT dans les régions polaires et de l'ordre de ~ 30000 nT à l'équateur. Sous la pression du vent solaire les lignes de force terrestres sont comprimées côté jour et s'étirent au-delà de l'orbite de la Lune du côté nuit, dans la direction anti-solaire. Pour une activité modérée du vent solaire, la magnétopause se situe à environ ~ 10 R_T au-dessus du point sub-solaire de la Terre et

[1]En anglais : *Interplanetary Magnetic Field*
[2]En anglais : *Bow shock*
[3]En anglais : *Magnetosheath*

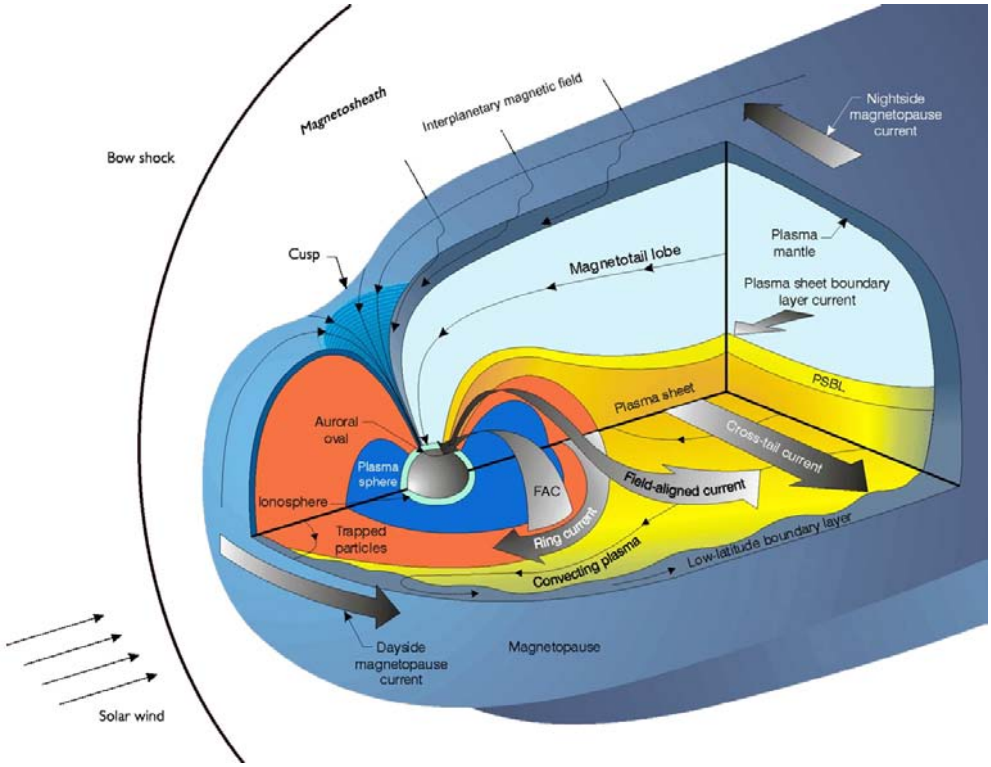


FIG. 1.2 – *Représentation schématique de la magnétosphère montrant les lignes de force magnétique, les systèmes de courant et les différentes régions plasmiques. Le vent solaire vient de la gauche.*

peut être repoussée à l'intérieur de l'orbite géostationnaire (6.6 R_T) lors d'événements intenses (Hanuise et al., 2006).

La structure interne de la magnétosphère est peuplée par un plasma thermique composé de particules énergétiques provenant à la fois du vent solaire et de l'atmosphère terrestre. Bien que l'essentiel du vent solaire soit dévié autour de la Terre, une certaine quantité de plasma peut traverser la magnétopause et entrer dans la magnétosphère (voir discussion § 1.2.1). Les *cornets polaires*[4] situés de part et d'autres du plan de l'écliptique, sont particulièrement sensibles aux sollicitations du milieu interplanétaire. Ils séparent les lignes de champ fermées du côté jour de celles qui s'étirent côté nuit. Ils se situent au-dessus des pôles magnétiques et constituent des régions faiblement magnétisées ouvertes sur la magnétogaine. A basse altitude ($\sim$ 300 km), les cornets polaires occupent une bande en latitude centrés autour de midi en heure magnétique locale (MLT[5]). Les particules du vent solaire peuvent, à travers cette bande, déposer

[4]En anglais : *Cusp*
[5]En anglais : *Magnetic Local Time*

leur énergie dans les couches supérieures de l'atmosphère.

Côté nuit, les *lobes* de la queue magnétosphérique sont les régions qui connectent magnétiquement la calotte polaire de l'ionosphère avec l'IMF. Le plasma y est peu dense ($\leqslant$ 0.1 cm^{-3}). Les lobes sont principalement alimentés par l'évaporation du plasma ionosphérique. Les lignes de force entre les lobes nord et sud de la queue sont quasi-parallèles. A l'interface, s'écoule une nappe de courant dans le sens matin-soir. Cette couche se nomme le *feuillet central de plasma* (CPS[6]). Au centre de cette couche, le champ magnétique est faible et le courant est hautement variable avec des sursauts d'écoulement rapide. Le feuillet central de plasma est le siège de reconnexion magnétique et de reconfiguration à grande échelle. L'accumulation du plasma au niveau des lobes semble être le facteur précurseur au déclenchement des sous-orages.

Le champ électrique, $\mathbf{E}$, qui traverse la magnétosphère ($\mathbf{E} = -\mathbf{V_{vs}} \wedge \mathbf{B_{IMF}}$, où $\mathbf{V_{vs}}$ est la vitesse du vent solaire et $\mathbf{B_{IMF}}$ le champ magnétique interplanétaire), contrôle le mouvement de convection du plasma magnétosphérique ($\mathbf{V_m} = \mathbf{E} \wedge \mathbf{B_m}$, où, $\mathbf{V_m}$ est la vitesse de dérive du plasma magnétosphérique et $\mathbf{B_m}$ le champ géomagnétique). Les particules énergétiques proches du feuillet plasma peuvent dériver en direction de la Terre et être piégées par les lignes de champ de plus basses altitudes pour former les ceintures de *Van Allen*. Les particules piégées se répartissent en deux régions, selon la gamme d'énergie. Dans la ceinture de radiation interne, les particules de très haute énergie (jusqu'au $\sim$ 1 MeV), majoritairement des ions, peuvent être piégées durant plusieurs mois. La ceinture de radiation externe se compose principalement d'électrons d'énergie allant du $\sim$ 1 keV à plusieurs MeV. Elle s'étend de 3 R_T jusqu'à l'orbite géostationnaire. Ces particules dérivent autour de la Terre formant le courant annulaire, localisé à $4 - 6$ R_T de distance de la Terre. Ce courant est hautement variable et est modulé par l'activité géomagnétique.

La structure magnétosphérique est maintenue par des courants électriques intenses parcourant ses régions frontières, à travers le feuillet plasma et parallèle aux lignes de force reliant l'ionosphère à haute latitude. Ces courants alignés aux champs (FAC[7]) interviennent dans le couplage entre la magnétosphère et l'ionosphère. Le feuillet plasma est magnétiquement connecté à l'ovale auroral, région encerclant les pôles magnétiques.

[6]En anglais : *Central Plasma Sheet*
[7]En anglais : *Field-aligned Current*

1.2 Couplage du système vent solaire-magnétosphère-ionosphère

Le vent solaire interagit en permanence avec le système magnétosphère-ionosphère se traduisant par un transfert d'énergie, de moment et de plasma vers les couches supérieures de l'atmosphère. Dans cette section, nous introduisons les différents concepts décrivant l'interaction entre le vent solaire et la magnétosphère, puis nous discuterons des conséquences de ce couplage sur la dynamique de l'ionosphère.

1.2.1 Interaction vent solaire-magnétosphère

Deux concepts d'interaction entre le vent solaire et la magnétosphère co-existent. Le premier suppose une magnétosphère magnétiquement hermétique au milieu interplanétaire (concept de la magnétosphère fermée). L'interaction est alors essentiellement visqueuse le long de la magnétopause. L'écoulement du vent solaire autour de la magnétosphère entraîne le plasma magnétosphérique des lignes géomagnétiques externes (internes) dans la direction anti-solaire (solaire). Deux cellules de convection sont formées de manière symétrique selon l'axe midi-minuit dans le plan équatorial de la magnétosphère. Une faible quantité d'énergie est déposée dans la magnétosphère par instabilité de Kelvin-Helmholtz ou encore par instabilité de type onde de surface.

Le transfert d'énergie est plus efficace lorsque la magnétosphère est ouverte sur le milieu interplanétaire (Dungey, 1961). C'est le concept de la magnétosphère ouverte. La description MHD idéale n'autorise pas d'échange de matière entre les champs magnétiques interplanétaire et terrestre. Lorsque la composante du champ magnétique interplanétaire est anti-parallèle au champ magnétique de la Terre, un des modes possibles d'instabilité est la reconnexion des lignes de champ terrestre et interplanétaire. Celle-ci rend aisé le transfert de plasma, d'énergie et de moment de la magnétogaine vers la magnétosphère. Les variations successives Nord-Sud du champ magnétique interplanétaire contrôle en grande partie le transfert d'énergie du vent solaire dans la magnétosphère, pouvant provoquer, à terme, le déclenchement des sous-orages. Les impulsions de pression du vent solaire sur la magnétosphère participent également à ce transfert.

1.2.2 Dynamique du système magnétosphère-ionosphère

La structure magnétosphérique est maintenue par des courants électriques intenses parcourant ses régions frontières, à travers le feuillet plasma et parallèles aux lignes

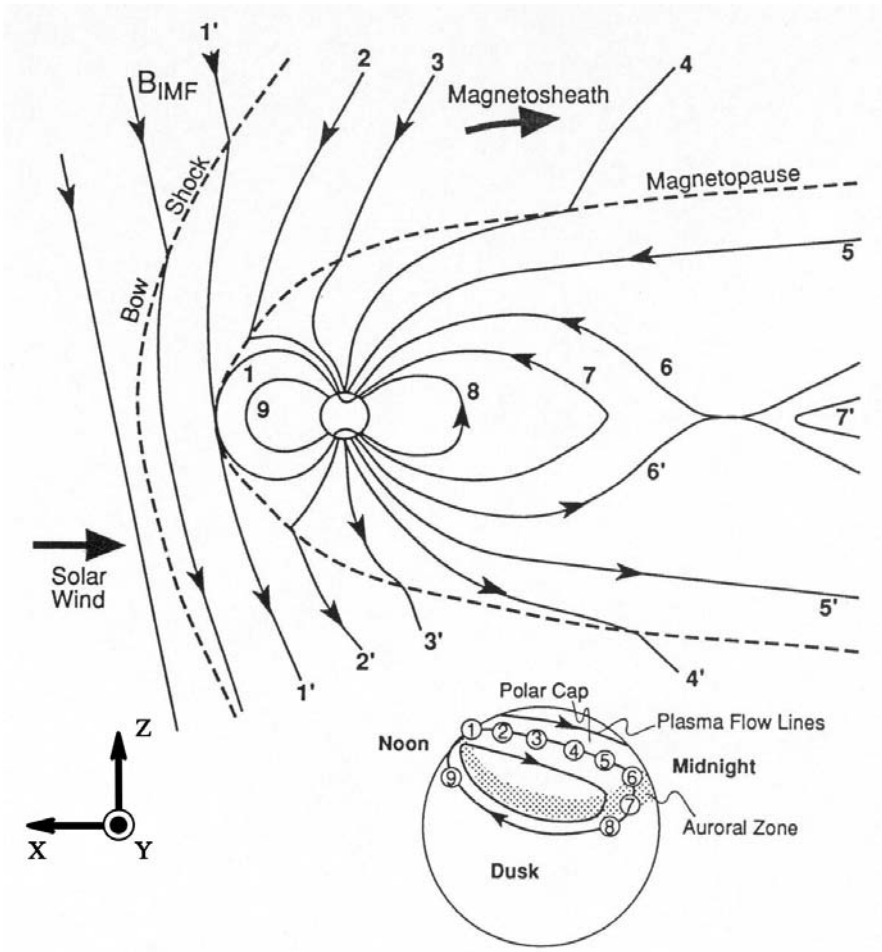


FIG. 1.3 – *Evolution temporelle des lignes de champ géomagnétique et de leur projection dans l'ionosphère, dans le système de coordonnées GSM. D'après Kivelson and Russel (1995).*

de force reliant l'ionosphère (à haute latitude). Les courants magnétosphériques se referment dans la région E de l'ionosphère, zone où la conductivité électrique perpendiculaire est forte. On parle alors de couplage électromagnétique.

Dans le référentiel de la Terre, le champ électrique, **E**, généré par le vent solaire qui est par exemple orienté dans le sens matin-soir lorsque le champ magnétique interplanétaire est orienté vers le sud, force la convection du plasma à grande échelle dans la magnétosphère. Les lignes de force étant considérées comme équi-potentielles (Dungey, 1961), cette convection impose dans l'ionosphère un écoulement du plasma dans la direction anti-solaire à travers la calotte polaire et un mouvement de retour (dans la direction solaire) sur les flancs à plus basses latitudes. La figure 1.3 illustre ce scénario dans le cas où le champ magnétique interplanétaire est dirigé vers le sud

($B_z < 0$). La ligne de champ magnétique (1) de la magnétopause se trouve dans une configuration anti-parallèle à celle du milieu interplanétaire (1'). Elle se reconnecte (2), ouvrant ainsi la magnétosphère au plasma solaire. Sous l'effet du vent solaire, la ligne de champ est entraînée vers la queue magnétosphérique (3 à 5) pour s'étirer jusqu'à ce qu'elle devienne quasi anti-parallèle avec le champ magnétique provenant du pôle opposé. Ces lignes de champ ouvertes sur le vent solaire se reconnectent ($6 - 6'$), libérant ainsi du plasma d'origine magnétosphérique dans la direction anti-solaire (7'). Dans la direction solaire, le plasma confiné dans les lignes nouvellement fermées, est accéléré et précipité dans l'ionosphère aux latitudes de l'ovale auroral (7 à 8). La projection dans l'ionosphère de la convection magnétosphérique avec la zone de précipitation est indiquée sur cette figure. Une cellule de convection (absente de la figure), symétrique par rapport à l'axe midi-minuit se forme également.

La convection du plasma en région F est donc contrôlée par le champ électrique projeté depuis la magnétosphère et par conséquent elle est imposée par l'orientation de l'IMF. Nous venons de voir que dans le cas de B_z négatif, la convection s'organise en deux cellules de sens opposé (Fig. 1.4 (b)), avec le plasma qui s'écoule dans la direction anti-solaire à travers la calotte polaire (imposé par l'écoulement du vent solaire) et la direction solaire à travers l'ovale auroral (imposé par la convection magnétosphérique). Lorsque que l'IMF est orienté vers le nord ($B_z > 0$), le schéma de convection est davantage confiné vers les plus hautes latitudes (Moses and Reiff, 1994) et devient plus complexe (Fig. 1.4 (e)). Deux cellules supplémentaires apparaissent avec un sens de circulation opposé au cas d'un $B_z < 0$. Ces deux cellules sont connectées aux lobes de la queue magnétosphérique.

La composante aube-crépuscule B_y de l'IMF rompt la symétrie de la convection magnétosphérique par rapport au méridien midi-minuit et déforme les cellules (Heppner and Maynard, 1987). Lorsque l'IMF possède une composante vers le côté matin ($B_y < 0$), la tension magnétique tire les lignes de champ nouvellement reconnectées vers le soir pour l'hémisphère Nord et vers le matin pour l'hémisphère Sud. Dans le cas ou $B_y > 0$, c'est l'inverse.

Les concepts présentés dans cette section supposent la stationnarité des processus de reconnexion à l'échelle planétaire. Des phénomènes de reconnexion transitoires et localisés à la magnétopause, côté jour, peuvent également se produire libérant ainsi localement du plasma dans la magnétosphère. Ces événements à transfert de flux (FTE[8]) induisent des perturbations qui se propagent le long des lignes de force de la magnétopause vers les plus basses altitudes, affectant localement le schéma de la convection ionosphérique (Cerisier et al., 2005).

[8]En anglais : *Flux Transfert Event*

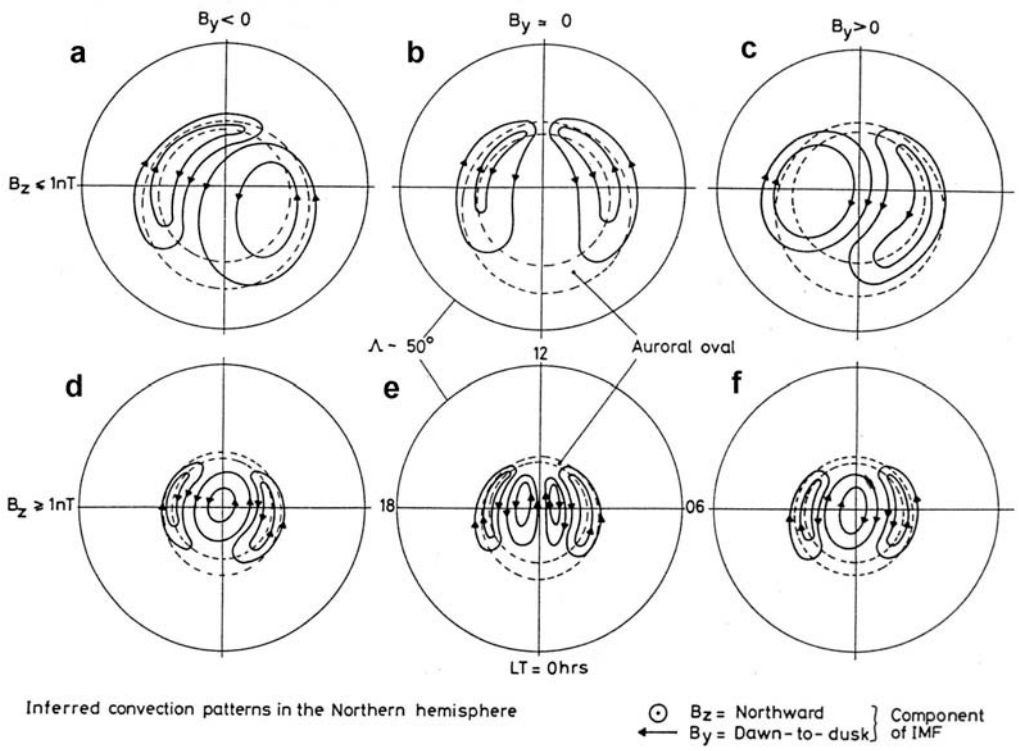


FIG. 1.4 – *Evolution attendue de la convection ionosphérique en coordonnées magnétiques en fonction de l'angle horaire du champ magnétique interplanétaire. Le Soleil est en haut. D'après Cowley and Lockwood (1992).*

1.2.3 Projection ionosphérique des régions magnétosphériques

Le plasma de la magnétosphère est non-collisionnel et magnétisé. Le mouvement des particules magnétosphériques peut être décrit comme le mouvement du centre guide. Le plasma dérive donc le long des lignes de champ magnétique. Le caractère dipolaire du champ géomagnétique implique que chacune des régions magnétosphériques possède sa projection dans l'ionosphère. On parle de couplage particulaire. Newell et al. (2004) ont étudié les précipitations des populations caractéristiques des régions magnétosphériques à partir d'une étude statistique basée sur 11 années de données enregistrées avec le satellite DMSP (*Defense Meteorological Satellite Program*). Les quatre panneaux de la figure 1.5 montrent la localisation statistique des différentes régions sources projetées sur une grille en latitude magnétique (MLAT) et en temps magnétique local (MLT) en fonction de l'orientation de l'IMF.

Depuis les basses latitudes, on rencontre tout d'abord le feuillet central de plasma

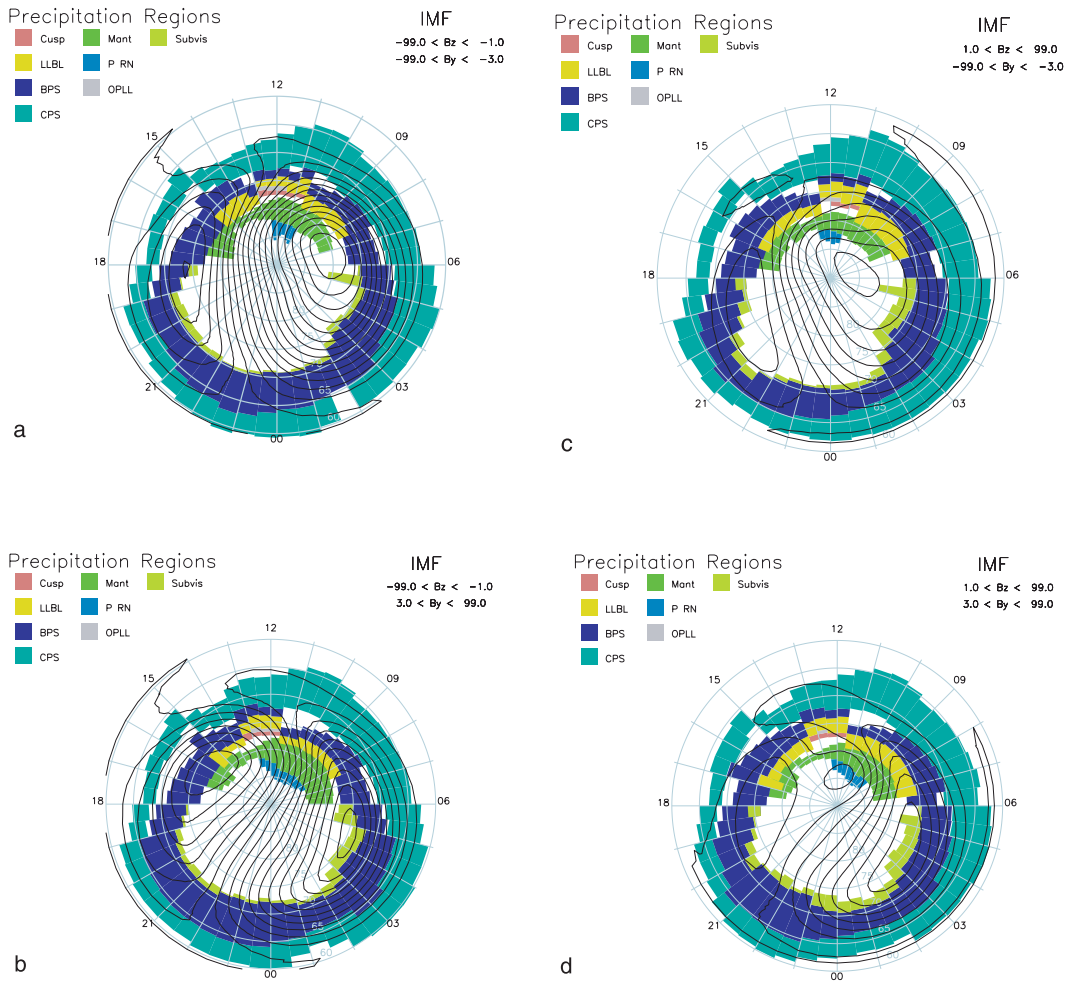


FIG. 1.5 – *Cartographie des précipitations ionosphériques classées selon les régions sources manétosphériques : (a) pour $B_z < 0$ et $B_y < 0$, (b) pour $B_z < 0$ et $B_y > 0$, (c) pour $B_z > 0$ et $B_y < 0$, (d) pour $B_z > 0$ et $B_y > 0$. D'après Newell et al. (2004)*

(CPS) qui épouse une forme annulaire. Ensuite la limite externe du feuillet plasma (BPS[9]). Centré autour de 12 : 00 MLT, le cornet polaire est la région la plus affectée par l'orientation de la composante B_y de l'IMF. Il varie entre $10^{\mathrm{h}}30$ et $13^{\mathrm{h}}30$ MLT. Notons que les régions magnétosphériques les plus externes se projettent vers les basses latitudes. L'anneau dessiné par la projection des régions magnétosphériques définit les frontières de l'ovale auroral et par conséquent la zone de précipitation dans l'ionosphère des particules provenant de la magnétosphère. Au-delà de l'ovale auroral se trouve la calotte polaire. C'est une région de faible précipitation correspondant aux lignes de champ connectées aux lobes de la queue magnétosphérique.

[9]En anglais : *Boundary Plasma Sheet*

A l'exception du cornet polaire, la projection ionosphérique des autres régions sources ne semble pas être réellement affectée par l'orientation de l'IMF. Cependant, nous notons que ces figures ne permettent pas d'appréhender les effets transitoires tels que les réponses des régions sources (en termes de dynamique) face aux variations brusques de l'IMF ou plus généralement du vent solaire pouvant amener une reconfiguration rapide sur de grandes échelles spatiales de la topologie interne de la magnétosphère. Ainsi, Newell et al. (1989) et Newell and Meng (1994) ont montré, à partir de mesures satellites basses altitudes, que le cornet polaire qui est magnétiquement connecté à la magnétopause (Smith and Lockwood, 1996) change de taille et de forme selon les conditions extérieures du vent solaire. La projection ionosphérique de la frontière BPS/CPS se déplace de plusieurs centaines de kilomètres (en latitude) en quelques minutes lors d'un retournement de la composante B_y (Dudeney et al., 1998).

1.3 Surveillance des régions magnétosphériques : les moyens expérimentaux

Les variations du champ géomagnétique enregistrées à la surface de la Terre peuvent être considérées comme la réponse non-linéaire du système complexe magnétosphère-ionosphère au forçage du milieu interplanétaire. En effet, la circulation du plasma magnétosphérique génère des courants dont les variations sont à l'origine des perturbations géomagnétiques enregistrées au sol. Les mesures du champ géomagnétique sont importantes pour l'élaboration d'indices géomagnétiques largement utilisés pour caractériser de manière quantitative les événements géomagnétiques. Le géophysicien allemand Julius Bartels a été le premier à introduire en 1949 l'indice géomagnétique planétaire Kp (Bartel, 1949). Ensuite, Mayaud (1968, 1973) élabore les indices am et aa. Les mesures continues effectuées depuis le milieu du XIXe siècle ont permis de remonter jusqu'à 1844 dans l'histoire de l'activité géomagnétique de notre planète Nevanlinna and Kataja (1993). Ces indices caractérisent les variations géomagnétiques, à hautes et moyennes latitudes, qui sont principalement reliées à l'électrojet auroral et par conséquent ils décrivent plutôt le niveau de perturbation global de la magnétosphère. D'autres indices ont été développés pour caractériser une partie de cette activité. A partir de mesures proches de l'équateur, l'indice Dst caractérise l'intensité du courant annulaire et du courant à la magnétopause (Chapman-Ferraro) (Sugiura, 1965). Enfin aux très hautes latitudes, l'indice AE introduit par Davis and Sugiura (1966), capturent la signature magnétique des phénomènes auroraux lors d'une activité géomagnétique intense. Ces différentes quantités empiriques sont calculées à partir des mesures de la composante horizontale du champ géomagnétique enregistrées par différents ré-

seaux de magnétomètres répartis sur l'ensemble du globe et sont librement accessibles sur Internet (http ://www.gfz-potsdam.de/pb2/pb23.GeoMag/niemegk/obs_eng.html ; http ://www.cetp.ipsl.fr/~isgi/lesdonne.htm ; http ://spidr.ngdc.noaa.gov/spidr/index.html). Ces indices géomagnétiques décrivent l'état énergétique de la magnétosphère à une l'échelle globale ou uniquement limitée à quelques régions frontières. Ils ne permettent pas de décrire sa topologie et donc de suivre la dynamique de la structure interne de la mgnétosphère.

Depuis les années 60, la magnétosphère est explorée par une multitude de satellites orbitant autour de la Terre, à diverses altitudes. Aujourd'hui, la flotte de satellites n'a jamais été aussi importante. Malgré cela, la taille de cette flotte reste insuffisante face aux vastes régions à couvrir. Une surveillance de l'ensemble de la magnétosphère à partir des mesures in-situ reste difficile puisque les missions spatiales ne proposent qu'une vision partielle. Néanmoins, les observations satellites permettent d'étudier localement les processus d'échange entre les régions magnétosphériques, leur structure, ainsi que leur dynamique. En effet, la magnétosphère est dominée par des structures dynamiques à trois dimensions. Les mesures provenant d'un seul satellite ne permet pas de séparer la plupart du temps les variations temporelles et spatiales de ces structures. Des missions composées de deux satellites, telles que ISEE-1 et 2, Interball ou encore Magion, ont permis une grande avancée en caractérisant une dimension spatiale, même si une seule dimension n'est pas suffisante au regard des trois dimensions des régions frontières à analyser. Par exemple, la comparaison des profils de densité mesurés entre les deux satellites ISEE-1 et 2, séparés par quelques centaines de kilomètres a permis d'étudier l'évolution temporelle de la magnétopause (Hubert et al., 1998). Il faut également citer les missions récentes de types multi-satellites telles que CLUSTER (Escoubet et al., 2001) et maintenant THEMIS (*Time History of Events and Macroscale Interactions during Substorms*). Ces deux missions sont composées respectivement de quatre et cinq sondes de conception identique. Leur configuration dans l'espace permet d'améliorer la résolution spatiale en trois dimensions des structures en multipliant les points de mesure. Des méthodes ont pu être mises en place pour observer le déplacement de certaines régions frontières de la magnétosphère le long de la trajectoire de la constellation de satellites, notamment avec la mission CLUSTER (Darrouzet, 2006).

De manière indirecte, l'étude de l'ionosphère à haute latitude, que ce soit depuis l'espace ou depuis le sol, permet de restituer une image quasi complète des processus d'échange entre la magnétosphère et l'ionosphère. Certains satellites ont à leur bord des caméras embarquées. Ces caméras qui travaillent en particulier dans la bande UV, permettent d'étudier l'ionosphère aurorale. Les premiers satellites à avoir embarqué une caméra UV sont les satellites POLAR puis IMAGE (*the Imager for Magnetopause-*

to-Aurora Global Exploration). Son imageur d'atomes neutres a permis de restituer une image globale de l'état des régions internes à la magnétosphère. Il est basé sur les processus d'échanges entre les espèces neutres de l'ionosphère aurorale et les particules chargées du courant annulaire qui sont suffisamment intenses pour permettre leur observation. (Burch, 2000; Burch et al., 2001). On peut également citer comme exemple l'instrument GUVI (*Global UltraViolet Imager*) à bord du satellite TIMED, qui permettait de déterminer l'abondance des composés principaux de la thermosphère à partir du rayonnement UV lointain de la Terre (Christensen et al., 2003). Quant à la mission POES, elle restitue une image de l'activité aurorale par extrapolation du flux de puissance déposé par les protons et les électrons qui produisent les aurores dans l'atmosphère. Les cartes d'activité aurorale sont accessibles en temps réel sur le site de la NOAA (http ://www.swpc.noaa.gov/pmap/index.html). Aujourd'hui, seul le satellite POLAR possède une caméra UV.

La combinaison entre les données satellites et les mesures radar ionosphériques effectuées au sol permettent d'analyser sous différents angles, un même événement ou une structure identique. Par exemple, Cerisier et al. (2005) ont mené une étude conjuguée entre les missions CLUSTER, IMAGE et les radars SuperDARN pour observer l'injection de plasma depuis le cornet polaire vers l'ionosphère. Le couplage des observations entre les mesures satellites et les mesures sol ont permis l'élaboration de modèles opérationnels de surveillance de la magnétosphère. Par exemple, le modèle OVATION qui est dédié à la modélisation en temps réel de l'ovale auroral (Newell et al., 2002). Ce modèle combine les mesures particules issues des satellites DMSP, de l'imageur UVI embarqué sur le satellite POLAR et des observations de la convection ionosphérique vue par les radars SuperDARN.

Les instruments au sol, comme par exemple le réseau de radars cohérents HF SuperDARN[10], ou encore les radars à diffusion incohérente (par exemple : EISCAT, Millstone Hill...), sans parler des caméras pleins ciels (associées à la mission THEMIS), offrent une opportunité unique grâce à leur couverture spatiale et à leur résolution temporelle, de suivre les régions internes de la magnétosphère en temps réel. Les modèles opérationnels précédemment cités sont essentiellement basés sur les mesures in-situ de la magnétosphère. Les observations radars contraignent la localisation notamment des frontières de l'ovale auroral. Pour ce travail de thèse, nous avons choisi une approche différente. Nous nous sommes basés sur les propriétés statistiques des échos radar mesurés par SuperDARN afin d'opérer une classification de ces échos en fonction de leur localisation afin d'associer ces classes à la projection de certaines régions frontières de la magnétosphère. Dans ce nouveau schéma d'identification, la projection des me-

[10]En anglais : *Super Dual Auroral Radar Network*

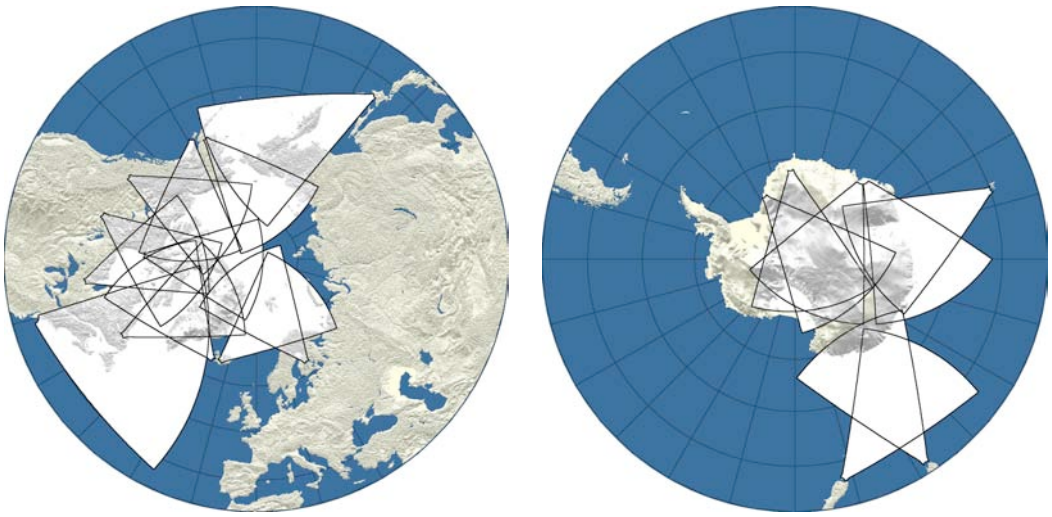
sures satellites pourra apporter dans le futur une amélioration sur la localisation des frontières.

1.3.1 Le réseau de radars cohérents HF SuperDARN

SuperDARN (pour *Super Dual Auroral Radar Network*) est un réseau international de radars cohérents Hautes Fréquences (HF) dédié à l'étude des processus dynamiques de l'ionosphère en régions aurorales et polaires (Greenwald et al., 1995). Grâce à la rétrodiffusion des ondes électromagnétiques par les fluctuations de densité électronique, les radars fournissent systématiquement des paramètres du milieu ; ces paramètres donnent accès à la vitesse de dérive du plasma déduite de la vitesse Doppler des échos radar et donnent également une information sur la turbulence du milieu. La superposition des champs de vue des radars permet de cartographier la convection ionosphérique globale à hautes latitudes. Ces cartes 2D de la convection ionosphérique sont actualisées toutes les 2 minutes avec une résolution spatiale de 100×100 km^2.

SuperDARN est constitué de deux chaînes de radars appariés situées dans les hémisphères nord et sud. Actuellement, 12 radars couvrent plus de 180 ° de longitude de l'hémisphère nord et 7 radars sont magnétiquement conjugués dans l'hémisphère sud. L'exploitation et la maintenance de chacun des 19 radars est sous la responsabilité de plusieurs pays. Ainsi, la France est engagée dans ce programme depuis 1994 avec l'exploitation d'un premier radar à Stokkseyri en Islande et d'un second radar en collaboration avec l'Italie, aux îles Kerguelen (2000). La figure 1.6 montre l'extension du champ de vue de l'ensemble du réseau dans les deux hémisphères.

Depuis une dizaine d'années, SuperDARN a prouvé avec succès qu'il est tout à fait adapté à l'étude des processus dynamiques du système magnétosphère-ionosphère à grande et à moyenne échelle et à l'étude du couplage entre le milieu interplanétaire et la magnétosphère (Chisham et al., 2007). A l'échelle planétaire, les radars SuperDARN contribuent à l'étude de la structure et de la dynamique de la convection globale de l'ionosphère en réponse aux variations du champ magnétique interplanétaire ou encore aux impulsions de pression du vent solaire. L'implantation des radars aux deux pôles permet de mener des études inter-hémisphériques de conjugaison magnétique. A moyenne échelle ($\sim$ 100 km), des études sont également menées sur les signatures ionosphériques des courants alignés (FAC) et des événements à transfert de flux (FTE), liés aux transferts d'énergie entre la magnétosphère et l'ionosphère. Et à petite échelle ($\lesssim$ 100 m), les radars SuperDARN permettent d'analyser les mécanismes de formation des irrégularités de densité par les instabilités de plasma.



(a) (b)

FIG. 1.6 – *Champ de vue des radars SuperDARN pour les hémisphères nord (a) et sud (b) (coordonnées géographiques).*

Avec sa haute résolution spatiale et temporelle, SuperDARN complète parfaitement la vision proposée par les mesures satellites. En effet, il restitue en temps réel une image globale de l'état du système magnétosphère-ionosphère, alors que les expériences spatiales produisent certes des mesures in-situ de la magnétosphère mais qui sont uniquement ponctuelles. Le fonctionnement continu et l'accessibilité en temps réel (au moins pour l'hémisphère nord) des mesures SuperDARN présentent, en plus des objectifs scientifiques mentionnés précédemment, un réel potentiel applicatif dans le domaine de la météorologie de l'espace (Greenwald, 1997; Rodger, 2000; Ruohoniemi et al., 2001). Par sa conception, SuperDARN offre une opportunité unique de pouvoir suivre en temps (quasi-) réel la dynamique de la magnétosphère et de ses régions internes, pour une activité géomagnétique faible à modérée. Pour un niveau d'activité trop élevé ou encore lors d'événements intenses, le manque d'écho radar ne permet plus d'effectuer cette surveillance (Hanuise et al., 2006).

1.3.2 Principe de fonctionnement

1.3.2.1 Principe de mesure

Les radars HF de la chaîne SuperDARN utilisent le principe de rétrodiffusion des ondes électromagnétiques par les irrégularités de densité électronique. En région F, les

irrégularités sont principalement créées par instabilité de *dérive de gradient* (ou dérive $\mathbf{E} \wedge \mathbf{B}$). Ces irrégularités possèdent un vecteur d'onde, $\mathbf{k_{irr}}$, perpendiculaire au champ magnétique local, puisqu'en région E et F, la conductivité parallèle du plasma (par rapport au champ magnétique) est très élevée ($\sigma_{\parallel}/\sigma_p \gg 1$) (Kelley, 1989).

La diffusion cohérente d'une onde électromagnétique résulte de l'excitation des ondes au-dessus du niveau naturel par les mécanismes d'instabilité. La section efficace de diffusion est alors proportionnelle, au sens de Bragg, à la valeur du spectre de puissance des irrégularités (Booker, 1956; Bowles and Balsley, 1963), telle que

$$\lambda_{\mathrm{irr}} = \frac{\lambda_{\mathrm{o}}}{2\sin(\theta/2)} \ , \tag{1.9}$$

où λ_{o} représente la longueur de l'onde incidente et θ l'angle entre le vecteur d'onde incident et diffusé. Dans la cas de la rétrodiffusion, $\theta = 180°$, l'expression (1.9) se réduit à

$$\lambda_{\mathrm{irr}} = \frac{\lambda_{\mathrm{o}}}{2} \ . \tag{1.10}$$

La rétrodiffusion d'une onde électromagnétique, $\mathbf{k_o}$, par les irrégularités de densité nécessite que la condition de Bragg soit satisfaite, *i.e.* lorsque $\mathbf{k_{irr}} = 2\mathbf{k_o}$. L'alignement le long du champ magnétique des irrégularités de densité impose donc la perpendicularité de l'onde incidente avec le champ magnétique ($\mathbf{k_o} \perp \mathbf{B}$). Cette condition est facilement réalisable aux latitudes équatoriales. Pour les régions aurorales et polaires, il faut utiliser les propriétés de réfraction des ondes dans l'ionosphère. Dans le domaine des hautes fréquences (3 - 30 MHz), les radars SuperDARN n'utilisent qu'une gamme réduite entre 9 et 16 MHz de cette bande afin de satisfaire la condition de perpendicularité à toutes les altitudes. Pour des raisons pratiques, lorsque la fréquence d'émission est trop faible ($<$ 9 MHz), les ondes électromagnétiques sont absorbées par les basses couches de l'ionosphère avant d'atteindre les cibles rétrodiffusantes. Tandis que pour de trop grandes valeurs ($>$ 16 MHz), la réfraction des ondes radios n'est pas suffisante pour atteindre la perpendicularité entre l'onde incidente et le champ géomagnétique. Dans cette bande, les radars détectent des irrégularités d'échelle caractéristique de l'ordre de 9 à 17 m ($\lambda_{\mathrm{irr}} = \lambda_{\mathrm{o}}/2$). L'onde radio qui est émise depuis le sol est réfractée tout le long de son parcours par l'ionosphère puis rétrodiffusée par les irrégularités électroniques lorsqu'il y a perpendicularité de l'onde incidente avec le champ magnétique. L'onde rétrodiffusée suit le même chemin de propagation que l'onde incidente mais en sens inverse, ce qui permet d'être détectée par le même radar.

L'utilisation d'une séquence de plusieurs impulsions permet par corrélation deux à deux des signaux reçus aux différents retards de reconstruire la fonction d'autocorrélation complexe (ACF[11]) du signal rétrodiffusé (Greenwald et al., 1985). Le principe de fonctionnement des radars, la recontruction et le traitement des ACFs seront abordés plus en détails dans l'annexe A. L'analyse de l'ACF donne accès aux paramètres physiques de l'ionosphère, tels que la vitesse Doppler radiale de dérive des irrégularités, la largeur spectrale du volume sondé et la puissance rétrodiffusée. La largeur spectrale est liée d'une part à la distribution des vitesses des irrégularités dans le volume sondé et d'autre part à la durée de vie de ces irrégularités. La puissance rétrodiffusée est liée à la présence des irrégularités dans le volume sondé, mais dépend également des processus physiques affectant la propagation de l'onde (absorption, diffusion par les irrégularités de moyennes échelles...).

1.3.2.2 Caractérisation de la fonction d'autocorrélation (ACF)

La fonction d'autocorrélation est une fonction complexe dont on extrait les paramètres physiques. La partie réelle est décrite par un *cosinus* amorti et la partie imaginaire par un *sinus* amorti. La caractérisation du spectre de rétrodiffusion peut se faire de deux manières différentes : soit par une transformée de Fourier rapide (FFT) (Baker et al., 1986), soit par un ajustement, au sens des moindres carrés d'une fonction paramétrique (Hanuise et al., 1985; Villain et al., 1987). Historiquement, la deuxième solution a été préférée car elle offre l'avantage de caractériser directement les paramètres de rétrodiffusion.

Pour chaque cellule radar, trois quantités sont systématiquement estimées par l'algorithme "FitACF". Le processus est résumé avec la figure 1.7. Le panneau 1.7-a montre la partie réelle (décroissance en cosinus) et imaginaire (décroissance en sinus) de l'ACF avec à droite le spectre de puissance correspondant (Fig. 1.7-b). Le spectre Doppler est centré sur la vitesse Doppler et la largeur spectrale est donnée par la largeur à mi-hauteur de la composante principale. Deux modèles de spectres sont utilisés pour ajuster analytiquement la largeur spectrale : un spectre lorentzien et un spectre gaussien.

Les deux formes utilisées pour ajuster les spectres de puissance correspondent en réalité à des extremums. La majorité des spectres observés par les radars SuperDARN varie entre ces deux formes. Moorcroft (2004) montre ainsi que d'un point de vue statistique, les spectres sont de forme intermédiaire.

[11] En anglais : *Auto-Correlation Function*

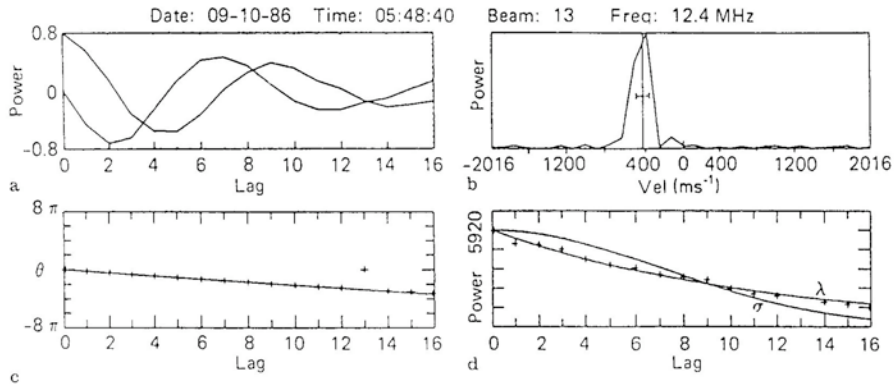


FIG. 1.7 – *Exemple de fonction d'autocorrélation complexe mesurée par les radars SuperDARN. (d'après Villain et al. (1996))*

L'ambiguïté sur la forme des spectres est liée à la physique des plasmas, mais également au principe de la mesure. Le processus physique qui est mesuré, est un processus non-stationnaire. Les caractéristiques du spectre, en termes de forme et de largeur, dépendent à la fois de la distribution moyenne des vitesses de dérive des irrégularités dans le volume sondé et de leur durée de vie par rapport à la durée de la mesure. De plus, le volume sondé, qui est typiquement de l'ordre 50×50 km dans le plan perpendiculaire au champ magnétique, possède une taille bien plus grande que celle des irrégularités observées ($\sim$ 10 m). Le spectre qui est mesuré n'est autre que la superposition de spectres «élémentaires» de forme différente associés à chacune des structures présentes dans un même volume. Dans le cas où dans un même volume, des régimes de vitesses bien distincts co-existent, ces spectres élémentaires peuvent être indirectement observés. Cela se traduit par une modulation de la décroissance du module de l'ACF et par conséquent se traduit au niveau du spectre de puissance, par l'apparition de plusieurs composantes, *i.e.* plusieurs pics. Il faut également noter que la présence de bruit dans la mesure peut entraîner l'apparition de plusieurs pics.

La largeur spectrale (Fig. 1.7-d) est estimée à partir de la décorrélation de la puissance de l'ACF ($|\mathrm{R}(\tau)| = \sqrt{\Re[\mathrm{R}(\tau)]^2 + \Im[\mathrm{R}(\tau)]^2}$). En supposant une distribution monomodale des vitesses (spectre à une composante), la largeur spectrale est calculée analytiquement en ajustant deux formes : une exponentielle décroissante ($|\mathrm{R}| \sim e^{-\lambda\tau}$) et une fonction gaussienne ($|\mathrm{R}| \sim e^{-\sigma^2\tau^2}$) au module de l'ACF. La puissance rétrodiffusée est définie par $|\mathrm{R}(0)|$. La vitesse Doppler radiale des irrégularités (Fig. 1.7-c) est calculée à partir d'un ajustement linéaire au sens des moindres carrés de la variation temporelle de la phase ($\varphi(\tau) = \arctan \Im[\mathrm{R}(\tau)]/\Re[\mathrm{R}(\tau)]$).

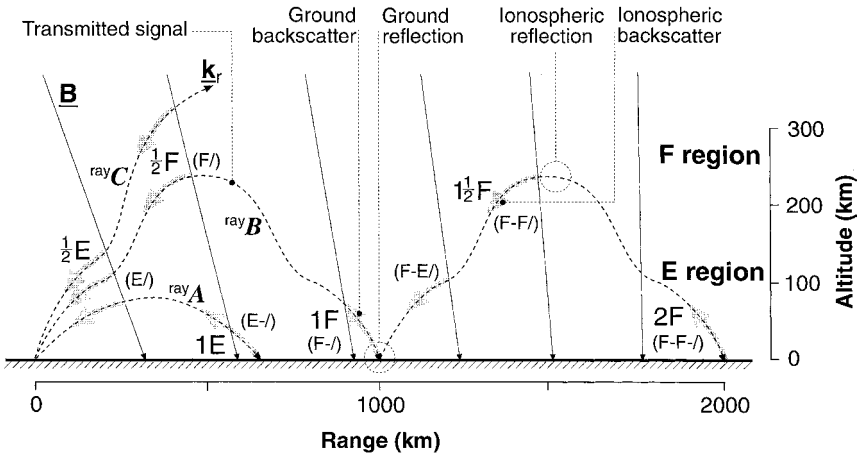


FIG. 1.8 – *Illustration des différents chemins de propagation des ondes radio à travers l'ionosphère. D'après Milan et al. (1997).*

La vitesse Doppler mesurée par les radars SuperDARN est la vitesse des irrégularités en région E et F. En Région F, la dérive des irrégularités est associée à la composante $\mathbf{E} \wedge \mathbf{B}$ de la dérive ionosphérique (Villain et al., 1985; Ruohoniemi et al., 1987). En région E, la description de la dérive du plasma est plus délicate car les ions sont démagnétisés et donc découplés du mouvement d'ensemble des électrons. Au-delà de la vitesse acoustique des ions (~ 400 m.s^{-1}), la vitesse des irrégularités électroniques sature et n'est plus significative de la dérive plasmique $\mathbf{E} \wedge \mathbf{B}$ (Villain et al., 1987). C'est pourquoi, en restreignant la mesure aux échos de région F, SuperDARN fournit un diagnostic très fiable de la convection du plasma.

La localisation des échos radar (en altitude, en distance ou en azimut) est délicate car les conditions du milieu peuvent mener à des chemins de propagation différents (Milan et al., 1997). La figure 1.8 illustre les différents trajets que peuvent emprunter les ondes radio à travers l'ionosphère. Hanuise and Crochet (1980) montrent un exemple où une onde est rétrodiffusée puis réfléchie par la couche F avant de revenir au radar. Pour cet exemple, l'hypothèse d'un chemin de propagation en ligne droite est donc erronée. La cible apparaît plus loin en distance qu'elle ne l'est en réalité. De plus, l'angle entre l'onde incidente et l'onde rétrodiffusée fausse l'estimation des paramètres ($\lambda_{\mathrm{irr}} = \lambda_{\mathrm{o}}/[2\sin(\theta/2)]$, avec $\theta \neq 180°$ l'angle entre l'onde incidente et l'onde rétrodiffusée). Selon André et al. (1997), l'incertitude sur la localisation est de l'ordre de 50 km en considérant un chemin de propagation direct. Dans le cas d'un chemin de propagation plus complexe, l'incertitude peut augmenter rapidement. Un autre effet des multiples chemins de propagation est le mélange d'échos radar provenant de régions différentes.

D'une manière générale, deux catégories peuvent être distinguées, les échos proches (< 900 km) associés à la région E et les échos lointains (> 1200 km) associés à la région F (Hanuise et al., 1991).

1.3.3 Les produits SuperDARN

1.3.3.1 Cartographie de la convection ionosphérique

L'assimilation des vitesses radiales à un modèle statistique de convection tenant compte des conditions du milieu interplanétaire permet de dessiner la carte globale de la convection du plasma à travers la calotte polaire (Ruohoniemi and Baker, 1998). La figure 1.9 donne un exemple de carte de convection pour le 25 octobre 2005 à 01h40 temps universel, à travers l'hémisphère nord. Les points situent l'origine des vecteurs, le trait la direction de la convection. La couleur code l'intensité de la vitesse. Ces cartes sont réalisées par le JHU/APL[12] et sont accessibles en temps réel sur leur page Internet (http ://superdarn.jhuapl.edu/). Pour cet exemple, on peut noter la présence de deux cellules de convection de part et d'autre du méridien midi-minuit caractéristique d'un schéma de convection dans le cas d'un $B_z < 0$ et d'un $B_y > 0$.

La cartographie 2D de la convection ionosphérique est l'information principalement utilisée pour étudier la dynamique du système magnétosphère-ionosphère. Cependant, de nombreuses études se sont portées sur l'utilisation de la largeur spectrale pour identifier la signature ionosphérique de certaines régions magnétosphériques comme par exemple, le cornet polaire ou encore la frontière des lignes de champ ouvertes/fermées.

1.3.3.2 Identifications des régions magnétosphériques : les prémices

De nombreuses études ont montré que les propriétés spectrales des échos radar étaient fonction de la latitude magnétique et du secteur MLT. Baker et al. (1995) et Rodger et al. (1995) montrent un accroissement de la largeur spectrale co-localisée avec la projection dans l'ionosphère du cornet polaire. Tandis qu'un gradient en latitude de la largeur spectrale marquant la séparation entre les petites et grandes largeurs spectrales est fréquemment observé tout autour de l'ionosphère polaire. La relation entre la frontière de la largeur spectrale (SWB[13]) et les frontières géophysiques a été étudiée en détail (Dudeney et al., 1998; Parkinson et al., 2002; Woodfield et al., 2002a,b;

[12] The John Hopkins University/Applied Physics Laboratory
[13] En anglais : *Spectral Width Boundary*

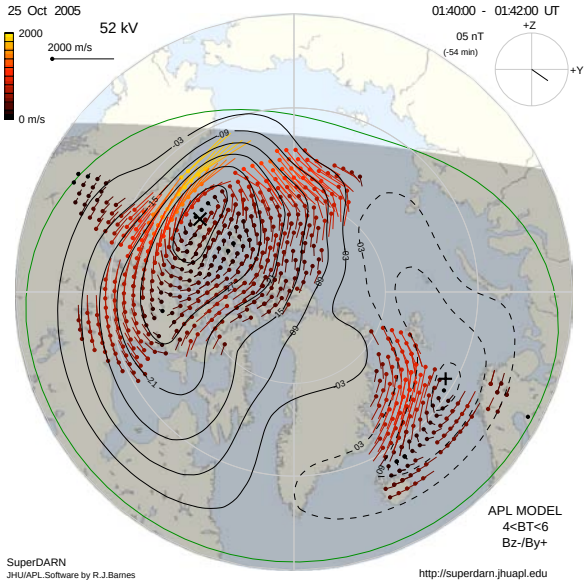


FIG. 1.9 – *Exemple de carte de convection (source JHU/APL).*

Chisham and Freeman, 2004; Chisham et al., 2005b,a). Pour une large gamme de secteur MLT, la SWB est notamment co-localisée avec la frontière entre les lignes de champ géomagnétiques ouvertes et fermées.

En parallèle, André et al. (2002) ont étudié les propriétés statistiques, sur plusieurs années, des spectres Doppler. Pour cela, ils ont réalisé une classification des échos radar selon leur complexité et leur largeur spectrale. Les auteurs séparent les spectres à une composante de vitesse des spectres à plusieurs composantes de vitesse. La figure 1.10 montre la probabilité d'observer différentes classes spectrales dans les zones aurorales et polaires de l'ionosphère. La classe «S», correspondant aux spectres étroits à une composante, se manifeste principalement à des latitudes magnétiques inférieures à 75°, entre $10^{h}00$ et $18^{h}00$ MLT. Cette classe est principalement enregistrée au niveau des lignes de champ fermées co-localisées avec le feuillet plasma central (CPS) et la couche limite du feuillet plasma (PSBL[14]) dans le secteur après-midi à basse latitude ($< 70°\Lambda$). Elle peut être associée à l'accroissement de l'ionisation transportée par la convection à grande échelle. La classe «T», assimilée aux spectres larges à une composante est associée à une région de précipitation intense de particules faiblement énergétiques ($\leqslant$ 1 keV) dans les secteurs nuit et matin, co-localisée à la calotte polaire. La classe «m» représente la catégorie des spectres complexes. Elle est observable dans une région

[14]En anglais : *Plasma Sheet Boundary Layer.*

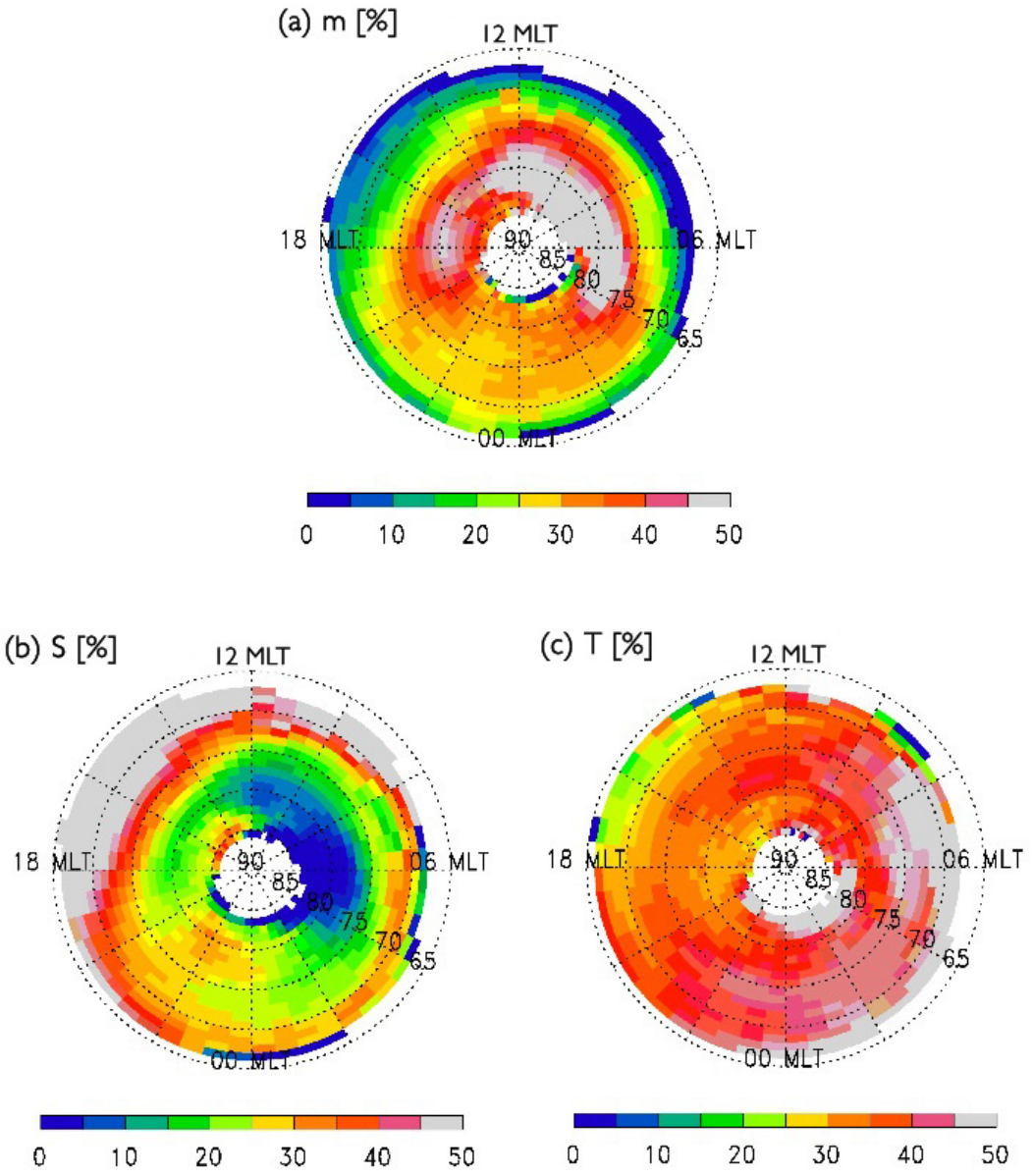


FIG. 1.10 – *Probabilité d'observer un écho radar rétrodiffusé de classe m (a), de classe S (b) et de classe T (c), en coordonnées MLT/Λ (cf texte pour la définition des classes). D'après André et al. (2002).*

en forme de «u» qui se superpose à la région du cornet polaire. Certaines activités d'ondes ultra-basses fréquences ($\lesssim$1 Hz) y sont géneralement détectées.

A partir d'une description statistique, les résultats de André et al. (2002) montrent qu'il est possible de mettre en oeuvre une procédure de classification des échos radar

observés par SuperDARN. La comparaison de ces classes avec des données satellites permet leur association avec la projection dans l'ionosphère de certaines régions magnétosphériques.

Ces nombreuses études suggèrent que certains paramètres mesurés par SuperDARN peuvent être utilisés comme indicateur[15] de certaines frontières magnétospériques (Rodger, 2000).

Un étape supplémentaire vers une modélisation opérationnelle a été réalisée par André and Dudok de Wit (2003). A partir de la même statistique que celle employée par André et al. (2002), ils ont reconstruit les fonctions de densité de probabilité de ces mêmes paramètres en fonction de la latitude et du secteur MLT. A l'aide d'une méthode statistique de décomposition en valeurs singulières, les auteurs ont montré qu'il était possible d'identifier ces régions de manière objective, sans employer de seuil ad-hoc.

Les études qui se sont intéressées aux propriétés des ACFs en fonction de la latitude magnétique et du secteur MLT ont montré qu'il est possible d'identifier la projection ionosphérique de certaines régions de la magnétosphère. Le principe même de fonctionnement des radars SuperDARN permet en plus d'envisager cette identification en temps (quasi)réel. Dans cette optique qui est de développer un outil de surveillance opérationnel, il est souvent préférable de travailler directement sur la mesure brute, *i.e* l'ACF, afin d'extraire des quantités appropriées pour réaliser cette tâche. Dans le chapitre suivant, nous introduisons une méthode empirique basée sur les propriétés statistiques des ACFs brutes qui permet de définir de nouveaux paramètres pour l'identification des régions magnétosphériques.

[15] Un indicateur empirique permet de quantifier indirectement un processus physique. Par exemple, le début de la période des vendanges est un traceur du climat passé (Chuine et al., 2004).

2

Méthode d'analyse des mesures SuperDARN

2.1 Une approche statistique pour une description «temps réel» 34
2.1.1 Position du problème . . . 34
2.1.2 Méthodologie . . . 36
2.1.3 Décomposition en Valeurs Singulières . . . 39
2.1.3.1 Définition algébrique . . . 40
2.1.3.2 Définition variationnelle . . . 42
2.1.3.3 Illustration géométrique de la SVD . . . 43
2.2 Mise en oeuvre de la méthode de réduction statistique de l'ACF . . . 45
2.2.1 Constitution d'une base de données . . . 45
2.2.2 Réduction de l'ACF . . . 49
2.3 Validation de la méthode . . . 53
2.3.1 Influence de l'activité géomagnétique . . . 53
2.3.2 Dépendance selon le secteur MLT . . . 59
2.3.3 Parenthèse sur la représentativité des échos radar sélectionnés . . . 62
2.4 Conclusion . . . 64

La majorité des études abordant le problème de détection des régions de la magnétosphère, impliquant SuperDARN, se sont essentiellement basées sur les propriétés statistiques déduites de l'analyse du spectre de puissance, *i.e* en ajustant un modèle analytique au module des ACFs et en étudiant la distribution de ses paramètres et éventuellement celle des écarts au modèle. En effet, l'étude de la vitesse Doppler, qui est estimée à partir de la variation de la phase selon le retard τ, permet de localiser la ligne neutre de renversement de convection (CRB[1]). Cette frontière n'est pas reliée,

[1]Cross Reversal Boundary

comme on pouvait le penser, à une frontière magnétosphérique (Lester et al., 2001). De plus, la vitesse Doppler dépend de l'orientation des radars, ce qui complique ce genre d'approche statistique. Pour ces différentes raisons, nous nous intéressons dans cette étude uniquement au module des ACFs et par conséquent au spectre de puissance lui-même.

La méthode que nous développons s'applique directement aux données brutes afin d'en extraire directement l'information pertinente. Ce faisant, nous arrivons à une méthode qui présente les avantages suivants :

- Gain de temps (travail sur les données brutes).

- En cherchant à déterminer les frontières à partir des données brutes et non à partir des résultats des modèles physiques, nous nous affranchissons du biais que ceux-ci peuvent introduire.

- Mise en place d'une méthode de réduction statistique des données qui facilite l'appplication de modèles physiques.

Une des méthodes les mieux adaptées à la réduction de données est la Décomposition en Valeurs Singulières (SVD). En l'appliquant au module de l'ACF, nous montrerons dans ce chapitre comment on peut reproduire les propriétés globales du spectre rétrodiffusé à partir d'un faible nombre de paramètres statistiques. Une seconde étape, qui sera abordée dans le chapitre suivant, consistera à interpréter ces paramètres en termes de classes et à les associer à la projection dans l'ionosphère de certaines régions magnétosphériques.

2.1 Une approche statistique pour une description «temps réel»

2.1.1 Position du problème

L'analyse des signaux radars donne accès aux paramètres physiques de l'ionosphère par l'intermédiaire des 18 retards τ de la fonction d'autocorrélation complexe $R(\mathbf{r}, t, \tau)$, localisée en un point $\mathbf{r}$ de l'ionosphère à un instant t. De nombreux travaux ont exploité la dépendance en latitude magnétique et en MLT des propriétés spectrales des échos radar afin d'identifier les signatures ionosphériques de régions ou de frontières internes à la magnétosphère (Baker et al., 1995; André and Dudok de Wit, 2003; Chisham and Freeman, 2003; Woodfield et al., 2002a,b). Que ce soit pour des études statistiques

ou pour des études de cas, ces travaux se sont plus particulièrement focalisés sur la distribution spatiale de la largeur spectrale (*cf* § 1.3.2.2).

Bien que l'interprétation physique de la largeur spectrale soit sujette à caution, notamment pour les grandes valeurs, ce paramètre semble être un bon candidat pour identifier certaines régions ou frontières magnétosphériques (Baker et al., 1986; Milan et al., 1999; Pinnock et al., 1995; Rodger et al., 1995). Par exemple, Baker et al. (1995) et Rodger et al. (1995) ont montré un accroissement de la largeur spectrale au niveau de l'empreinte ionosphérique du cornet polaire. Une autre propriété remarquable de la largeur spectrale est un gradient latitudinal observé régulièrement autour de l'ionosphère polaire. Ce gradient, lorsque la variation est brusque, marque la limite entre le régime des petites et des grandes largeurs spectrales. La limite entre ces deux régimes de largeur spectrale est nommée la frontière de la largeur spectrale (SWB[2]). La relation entre la SWB et les frontières géophysiques a été étudiée en détail et a permis de mettre en évidence une bonne corrélation avec la projection de la frontière des lignes de champ ouvertes/fermées (OCB[3]) pour une large gamme de secteurs MLT (Dudeney et al., 1998; Lester et al., 2001; Lointier et al., 2008; Parkinson et al., 2002; Woodfield et al., 2002a,b; Chisham and Freeman, 2004; Chisham et al., 2005c,a). L'étude de la vitesse d'écoulement traversant d'OCB, par exemple le long de la frontière côté jour, permet d'estimer le taux de reconnexion à la magnétopause (Baker et al., 1997; Pinnock et al., 1999; Hubert et al., 2006). Dans ce qui suit, nous nous intéresserons au suivi de l'OCB.

Cependant, la largeur spectrale ne suffit pas pour la détection des régions frontières magnétosphériques, car son estimation souffre de quelques faiblesses. En effet, les modèles utilisés pour ajuster l'ACF ne sont pas adéquats pour reproduire l'ensemble des formes observées (*cf* § 1.3.2.2). D'après André et al. (2002), la détection des régions frontières magnétosphériques peut être améliorée en incorporant comme paramètres additionnels les écarts entre les modèles ajustés et les observations. Les auteurs ont montré que les erreurs résiduelles sur l'estimation de la phase et de la puissance au retard 0 ne sont pas aléatoirement réparties à travers la calotte polaire mais sont au contraire spatialement organisées. A partir de seuils ad-hocs appliqués sur ces trois paramètres (largeur spectrale, erreur sur la phase et sur la puissance), ils ont identifié trois classes d'ACFs : spectre large à une composante, spectre étroit à une composante et spectre multi-composantes. D'un point de vue statistique, chacune de ces classes apparaît préférentiellement selon certaines heures locales. Les auteurs associent ces classes à des processus physiques classiquement observés plus haut en altitude, chacun de ces

[2]En anglais : *Spectral Width Boundary.*
[3]En anglais : *Open/Closed field line Boundary.*

processus étant lié à une région spécifique de la magnétosphère (par exemple : activité d'ondes Pc1, précipitation de particules). Les résultats de André et al. (2002) montrent comment l'erreur sur l'estimation de la phase et de la puissance sont complémentaires pour la description des régions ionosphériques. Ceci pose le problème de la capacité des modèles à reproduire la mesure et donc à estimer la largeur spectrale. En réalisant une analyse statistique multivariée pour laquelle aucune classe *a priori* n'a été définie, André and Dudok de Wit (2003) confirment ce résultat.

Tous ces résultats obtenus avec les radars SuperDARN ne posent que les premières fondations d'une surveillance automatisée et en temps réel des régions magnétosphériques. Cependant, le développement d'un tel outil requiert une méthodologie spécifique. Par exemple, nous pouvons citer le travail de Wing et al. (2003) qui ont développé un réseau de neurones destiné à séparer les ACFs pour ne conserver que celles qui permettent une estimation fiable des paramètres. L'objectif premier était de démontrer que les réseaux de neurones pouvaient opérer à un niveau de performance suffisant pour automatiser les tâches de classification des échos radar. Ces auteurs notent également qu'une telle approche peut être d'une grande aide pour des applications en météorologie de l'espace, en ce sens qu'il serait aisé d'identifier différents types d'ACFs comme par exemple les fonctions modulées (spectre multi-composantes).

Toute l'information dont nous disposons est contenue dans les ACFs. La voie classique consiste à ajuster un modèle physique ou empirique, et à étudier les moments statistiques des différents paramètres issus de l'ajustement. Le passage par un modèle suppose des hypothèses simplificatrices qui peuvent biaiser la statistique. Pour contourner ce problème, nous nous sommes donc focalisés sur les propriétés statistiques des échos radar rétrodiffusés. En utilisant sur un échantillon représentatif d'ACFs une méthode de décomposition en valeurs singulières comme outil de réduction de données, nous caractérisons la mesure avec un nombre réduit de paramètres statistiques.

2.1.2 Méthodologie

Le terme ACF, sauf mention contraire, désignera dorénavant le module de l'ACF. Nous écartons volontairement l'information sur la phase car elle ne semble pas pertinente pour notre objectif. L'information sur la phase décrit la convection ionosphérique qui est une conséquence directe du couplage entre le vent solaire et l'environnement terrestre. Cette quantité, qui contient dans une certaine mesure la trace indirecte des régions magnétosphériques, reste difficilement exploitable pour une approche statistique comme celle employée dans ce travail.

Les ACFs possèdent toutes une variation selon le retard τ qui varie relativement peu d'un cas à l'autre. Dans un échantillon statistique, cette reproductibilité implique une redondance d'information. Il devrait donc être aisé de reproduire la variabilité des ACFs à partir de quelques paramètres. L'objectif est alors de décrire chaque ACF par une somme de fonctions de base (ou modes) caractéristiques des propriétés générales de la rétrodiffusion à hautes latitudes. Contrairement aux nombreuses études citées jusqu'ici, ces fonctions de base ne seront pas imposées a priori à partir d'un modèle, mais déduites des observations. Nous perdons donc en interprétation physique, mais gagnons en capacité du modèle à reproduire les observations.

Afin de simplifier la notation, nous parlerons d'une ACF $R(t, \tau)$ enregistrée à un instant t, ce qui signifiera implicitement une ACF localisée en un point $\mathbf{r}$ de l'ionosphère, enregistrée à instant t.

Le problème peut se poser de la manière suivante : comment réduire une forme bi-variée dépendante de t et de τ par une somme finie de termes séparables de la forme

$$R(t, \tau) \simeq \sum_{k=1}^{K} a_k(t) \cdot \Phi_k(\tau) \ , \tag{2.1}$$

où $a_k(t)$ est une fonction temporelle et $\Phi_k(\tau)$ est une fonction de base ? Dans notre contexte Φ_k est appelé mode ; il est indépendant de t et de dimension identique à l'ACF, *i.e* défini sur les 18 retards. Cette représentation n'est certainement pas unique, il est donc nécessaire de poser des contraintes. Il nous faut déterminer une famille de fonctions de base Φ_k et y associer un ensemble de fonctions temporelles $a_k(t)$ afin d'obtenir la meilleure approximation possible.

De façon à rendre la décomposition unique, nous imposons la contrainte selon laquelle l'ensemble des modes $\Phi_k(\tau)$ forment une base orthonormale, telle que

$$\langle \Phi_{k_1}(\tau), \Phi_{k_2}(\tau) \rangle = \begin{cases} 0 & \text{si } k_1 \neq k_2 \ , \\ 1 & \text{si } k_1 = k_2 \ , \end{cases} \tag{2.2}$$

où $\langle \cdot, \cdot \rangle$ représente le produit scalaire. Ainsi une fonction temporelle $a_k(t)$ quelconque est uniquement liée au k-ième mode et peut être calculée sans ambiguïté à partir de l'expression (2.1) de la manière suivante

$$a_k(t) = \langle R(t, \tau), \Phi_k(\tau) \rangle \ . \tag{2.3}$$

De la même façon, on contraint les termes $a_k(t)$ à être orthogonaux :

$$\langle a_k(t), a_l(t) \rangle = 0 \quad \text{ssi} \quad k \neq l \ . \tag{2.4}$$

Par conséquent, à chacune des fonctions de base $\Phi_k(\tau)$, on associe une fonction temporelle $a_k(t)$ unique d'ordre k. En d'autres termes, l'ensemble des coefficients $a_k(t)$ forme les paramètres du modèle exprimant la contribution de chacun des modes $\Phi_k(\tau)$. Dans un certain sens, une telle décomposition est analogue à une décomposition en modes de Fourier. Dans le cas d'une projection d'une ACF (ou plus généralement d'une série temporelle quelconque) sur des modes de Fourier, les fonctions de bases $(\Phi_k(\tau))$ sont des fonctions connues, *i.e.* des fonctions cosinus et sinus et les coefficients pondérateurs associés à chacune de ses fonctions de base ne sont autres que les paramètres $a_k(t)$. La différence majeure réside uniquement dans le fait que l'allure des modes de Fourier est donnée par les hypothèses de départ tandis que pour notre problème la forme et le nombre nécessaire de modes $\Phi_k(\tau)$ résultera des propriétés statistiques des ACFs. Nous reviendrons sur cet aspect dans les sections suivantes.

Afin de réaliser cette réduction, il est plus commode de ranger les mesures dans une matrice de dimension $N_t \times N_\tau$:

$$\mathrm{R} = \begin{pmatrix} R(t_1,\tau_1) & R(t_1,\tau_2) & \cdots & R(t_1,\tau_{N_\tau-1}) & R(t_1,\tau_{N_\tau}) \\ R(t_2,\tau_1) & R(t_2,\tau_2) & \cdots & R(t_2,\tau_{N_\tau-1}) & R(t_2,\tau_{N_\tau}) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ R(t_{N_t-1},\tau_1) & R(t_{N_t-1},\tau_2) & \cdots & R(t_{N_t-1},\tau_{N_\tau-1}) & R(t_{N_t-1},\tau_{N_\tau}) \\ R(t_{N_t},\tau_1) & R(t_{N_t},\tau_2) & \cdots & R(t_{N_t},\tau_{N_\tau-1}) & R(t_{N_t},\tau_{N_\tau}) \end{pmatrix} \in \mathbb{R}^{N_t \times N_\tau} \tag{2.5}$$

où la ligne i définit l'ACF enregistrée au temps t_i, et la colonne j l'ensemble des ACFs observées pour un retard τ_j.

A partir de la matrice R, nous allons chercher à déterminer une famille de fonctions orthonormales $\{\Phi_k\}_{k=1}^{K}$ caractéristiques des propriétés statistiques des ACFs telles que l'approximation (2.1) soit la meilleure au sens des moindres carrés. Cette approximation

est équivalente au problème de minimisation au sens de la norme Euclidienne[4] suivant :

$$\min \left\{ \sum_{i=1}^{N_\tau} \left\| R(t,\tau_i) - \sum_{k=1}^{K} a_k(t) \cdot \Phi_k(\tau_i) \right\|_2^2 \right\} . \tag{2.6}$$

Nous allons montrer dans la section 2.1.3 que la solution de ce problème de minimisation est directement donnée par la Décomposition en Valeurs Singulières de la matrice R.

2.1.3 Décomposition en Valeurs Singulières

La Décomposition en Valeurs Singulières (SVD[5]) est un théorème fondamental de l'algèbre linéaire. Basée sur la reconnaissance des traits dominants d'un ensemble de données, la SVD est une méthode d'analyse efficace qui permet l'approximation d'un système de dimension élevée par un autre de dimension généralement plus faible. Elle autorise une transformation d'un ensemble de paramètres corrélés dans un autre ensemble de paramètres non-corrélés.

A l'aube de l'algèbre linéaire, cinq mathématiciens entre la fin XIXe et le début du XXe siècle ont largement contribué au développement de la méthode SVD, sous des formes différentes (Stewart, 1993) : les mathématiciens Eugenio Beltrami (1835-1899), Camille Jordan (1838-1921), James Joseph Sylverster (1814-1897), Erhard Schmidt (1876-1959) et Hermann Weyl (1885-1955). De nos jours, la SVD existe sous différentes déclinaisons selon le champ disciplinaire dans lequel elle est employée. Ainsi en météorologie, on utilise la méthode de Décomposition en Fonctions Orthogonales Empiriques (EOF[6]) afin d'identifier des modèles de circulation atmosphérique globaux (Preisendorfer, 1988). En traitement d'image, on préfère le terme de Décomposition de Karhunen-Loève (KLT[7]). Pour la simplification des modèles d'écoulement turbulent (Aubry et al., 1988; Cordier and Bergmann, 2003), on parle plutôt de Décomposition Orthogonale aux Valeurs Propres (POD[8]). On retrouve également dans la littérature les méthodes de Décomposition BiOrthogonales ou encore d'Analyse en Composantes Principales (PCA[9]) (Chatfield and Collins, 1995) dans le domaine d'analyse de données multidimensionnelles.

[4]La norme 2 ou Euclidienne d'un vecteur $\mathbf{x} = [x_1, x_2, \cdots, x_n]^T \in \mathbb{R}^n$ est définie par $\| \mathbf{x} \|_2 = \sqrt{\sum_{i=1}^{p} |x_i|^2}$.
[5]En anglais : *Singular Value Decomposition.*
[6]En anglais : *Empirical Orthogonal Functions.*
[7]En anglais : *Karhunen-Loeve Transform*
[8]En anglais : *Proper Othogonal Decomposition.*
[9]En anglais : *Principal Component Analysis.*

Il est nécessaire de préciser que ces méthodes diffèrent selon le domaine d'application. Cependant, elles restent toutes basées sur le même principe, celle de la décomposition en valeurs/vecteurs propres de la matrice de covariance. Dans ce contexte, la méthode décomposition en valeurs singulières peut être vue comme un outil permettant de réaliser une analyse en PCA ou encore en KLT (Gerbrand, 1981). Le point commun de ces applications résident dans la recherche des axes naturels de corrélation linéaire (*i.e.* orthogonaux), permettant d'accéder ainsi à une description plus compacte du jeu de données initial. Ce qui est important de noter ici, c'est que l'interprétation de ces méthodes dépend avant tout du nombre de réalisations par rapport à la taille des variables.

La multiplication des moyens d'observation dans le domaine des relations Soleil-Terre implique une quantité toujours plus importante de données de sources diverses, dont il faut faire la synthèse. Il est alors naturel d'avoir recours à des méthodes spécifiques afin d'extraire l'information pertinente des jeux de données multivariées. Par exemple, Dudok de Wit and Auchère (2007) utilisent cette méthode de décomposition afin de caractériser les différentes structures du Soleil à partir des observations à différentes longueurs d'ondes du spectre EUV. Santolík et al. (2006) ont développé un modèle empirique afin de localiser les sources des ondes électromagnétiques multi-composantes observées à bord du satellite DEMETER. Menvielle et al. (2007) utilisent la SVD comme un outil de différenciation afin de modéliser la variation de la densité thermosphérique. André and Dudok de Wit (2003) ont également montré que la SVD pouvait être utilisée afin de réaliser une classification des régions magnétosphériques à partir des fonctions de densité de probabilité des paramètres usuels mesurés par SuperDARN et ce, sans imposer de modèle a priori.

2.1.3.1 Définition algébrique

Pour des raisons de simplicité, nous ne définirons la méthode SVD que pour le cas où R est à valeur dans $\mathbb{R}$. A titre d'exemple, une extension au cas complexe de la SVD est présentée dans Cordier and Bergmann (2003). Dans cette étude, nous nous restreignons à caractériser les propriétés du spectre de puissance en appliquant la méthode SVD au module de l'ACF.

Considérons le cas général d'un ensemble de séries temporelles rangées en colonne dans une matrice réelle définie positive $(\mathrm{B})_{ij} = B(t_i, \mathbf{x}_j)$ de dimension $N_t \times N_{\mathbf{x}}$. La Décomposition en Valeurs Singulières n'est autre que la factorisation de B en trois

matrices (Golub and Van Loan, 1993), telles que :

$$\mathrm{B} = \mathrm{F}\,\Lambda\,\mathrm{G}^T \tag{2.7}$$

Où Λ est une matrice diagonale de dimension $N_t \times N_{\mathbf{x}}$, contenant les valeurs $\lambda_1, \cdots, \lambda_p \in \mathbb{R}$, appelées valeurs singulières ou poids de B, avec $p = min\{N_t, N_{\mathbf{x}}\}$. Par convention, les valeurs singulières sont rangées par ordre décroissant, telles que $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_r = \lambda_{r+1} = \cdots = \lambda_p = 0$. Le rang r représente le nombre de valeurs singulières non-nulles de la matrice B. Le symbole $\cdot^T$ définit la transposée. Les matrices F et G de dimension respective $N_t \times N_t$ et $N_{\mathbf{x}} \times N_{\mathbf{x}}$ sont orthogonales et unitaires,

$$\begin{cases} \mathrm{F}\,\mathrm{F}^T = \mathrm{I}_{\mathrm{N_t}}\,, \\ \mathrm{G}\,\mathrm{G}^T = \mathrm{I}_{\mathrm{N_x}}\,, \end{cases} \tag{2.8}$$

telles que $\mathrm{I}_{\mathrm{N_t}} \in \mathbb{R}^{N_t \times N_t}$ et $\mathrm{I}_{\mathrm{N_x}} \in \mathbb{R}^{N_{\mathbf{x}} \times N_{\mathbf{x}}}$ sont les matrices identités. Il est facilement démontrable que les r premières colonnes de $\mathrm{F} = [f_1, \cdots, f_{N_t}]$ et $\mathrm{G} = [g_1 \cdots, g_{N_{\mathbf{x}}}]$ forment les vecteurs singuliers gauches et droits "non-nuls" de la matrice de corrélation ($\mathrm{B}\,\mathrm{B}^T \in \mathbb{R}^{N_t \times N_t}$) et de covariance ($\mathrm{B}^T\,\mathrm{B} \in \mathbb{R}^{N_{\mathbf{x}} \times N_{\mathbf{x}}}$)

$$\begin{cases} (\mathrm{B}\,\mathrm{B}^T) f_\alpha = {\lambda_\alpha}^2\, f_\alpha\,, \\ (\mathrm{B}^T\,\mathrm{B}) g_\alpha = {\lambda_\alpha}^2\, g_\alpha\,, \end{cases} \tag{2.9}$$

Par ailleurs, les matrices F et G étant orthogonales et unitaires, les vecteurs singuliers dont elles sont constituées forment une base orthonormée :

$$\langle f_\alpha, f_\beta \rangle = \langle g_\alpha, g_\beta \rangle = \begin{cases} 0 & \text{si } \alpha \neq \beta \\ 1 & \text{si } \alpha = \beta\,. \end{cases} \tag{2.10}$$

Cette contrainte assure l'unicité, au sens de la norme Euclidienne, de la décomposition. Basée sur la matrice de covariance (et donc limitée à l'étude statistique à l'ordre 2 des données), la SVD permet donc de définir une nouvelle base dont les vecteurs propres sont décorrélés. Il existe cependant une contrainte plus forte qui consiste à supposer l'indépendance statistique de signaux sources. En effet, il faut rappeler que deux modes orthogonaux ne sont pas forcément indépendants tandis que l'inverse reste vrai (Hyvärinen et al., 2001). La méthode d'analyse en composantes indépendantes (ICA[10]) (Hyvärinen et al., 2001) permet de faire cela. Cette technique peut être pré-

[10]En anglais : *Independant Component Analysis*

sentée comme une extension de la SVD. Elle exploite les propriétés statistiques des données d'ordre supérieur à 2 (et également inférieur), afin de rechercher les directions génératrices dans l'espace des données qui sont statistiquement indépendantes.

L'équation (2.7) peut se réécrire :

$$\mathrm{B} = \lambda_1 \cdot f_1 \cdot {g_1}^T + \lambda_2 \cdot f_2 \cdot {g_2}^T + \cdots + \lambda_r \cdot f_r \cdot {g_r}^T \ , \tag{2.11}$$

où le produit dyadique $f_k \cdot {g_k}^T$ est une matrice $N_t \times N_{\mathbf{x}}$, puisque c'est le produit d'un vecteur $N_t \times 1$ par un vecteur $1 \times N_{\mathbf{x}}$. L'expression (2.11) montre que la méthode SVD constitue le développement limité de la matrice B par une somme de matrice séparable de rang 1. Elle démontre notamment que les premiers modes, du fait de l'ordonnement décroissant des poids (λ_α), pèsent davantage sur l'approximation de B. Lorsque la suite des valeurs singulières λ_α décroît rapidement alors les propriétés sous-jacentes de la matrice B peuvent donc être approchées avec une certaine précision à partir de l'expression réduite

$$\widehat{\mathrm{B}}_m = \lambda_1 \cdot f_1 \cdot {g_1}^T + \lambda_2 \cdot f_2 \cdot {g_2}^T + \cdots + \lambda_m \cdot f_m \cdot {g_m}^T \ , \tag{2.12}$$

où $m \leqslant r$. On dit que $\widehat{\mathrm{B}}_m$ est l'approximation de rang m de la matrice B. On peut alors montrer que la norme 2 de l'erreur de troncature vaut

$$\left\|\mathrm{B} - \widehat{\mathrm{B}}_m\right\|_2^2 = \sum_{k=m+1}^{p} {\lambda_k}^2 \ . \tag{2.13}$$

La norme de la matrice est définie comme $\left\|A\right\|_2^2 = \mathrm{tr}\left(A^T A\right)$ et l'opérateur tr signifie la trace du produit $A^T A$. Par conséquent, si la distribution des valeurs singulières diminue rapidement, alors nous pouvons espérer trouver une approximation de B possédant un rang faible. La SVD est en ce sens une puissante méthode de compression avec perte.

Il faut rappeler que la décomposition par SVD, sous la contrainte d'orthogonalité des modes est unique au signe près, à l'exception du cas où deux modes possèdent le même poids. Dans ce cas, on peut faire une rotation de leur sous-espace. On parle alors de dégénérescence des deux modes en question.

2.1.3.2 Définition variationnelle

Il existe également dans la littérature une autre manière d'introduire la décomposition en valeurs singulières. Pour cela, il faut considérer les éléments de la matrice B comme une image échantillonnée de la forme bilinéaire $B(t, \mathbf{x})$ telle que $(\mathrm{B})_{ij} = B(t_i, x_j)$. Dans le contexte variationnel, la méthode SVD permet de décomposer une forme bilinéaire $B(t, \mathbf{x})$ en une somme de produits de fonctions à variables séparées, comme suit :

$$B(t_i, x_j) = \sum_{\alpha=1}^{\min\{N_t, N_{\mathbf{x}}\}} \lambda_\alpha \cdot f_\alpha(t_i) \cdot g_\alpha(x_j), \qquad \text{avec } i = 1, \ldots, N_t \text{ et } j = 1, \ldots, N_{\mathbf{x}}. \tag{2.14}$$

où les fonctions $f_\alpha(t)$ et $g_\alpha(\mathbf{x})$ définissent donc respectivement le mode temporel et spatial d'ordre α de la fonction initiale $B(t, \mathbf{x})$. Et λ_α correspond au poids de ce mode. Bien que la terminologie soit différente, les modes portent la même signification que les vecteurs propre de la définition matricielle de la SVD et possèdent donc les mêmes propriétés.

La définition variationnelle de la méthode SVD est donc basée sur le problème d'approximation posé à l'équation (2.6). Par identification entre les équations (2.1) et (2.14), nous montrons aisément les équivalences suivantes : $a_\alpha(t) \equiv \lambda_\alpha \cdot f_\alpha(t)$ et $\Phi_\alpha(\tau) \equiv g_\alpha(\tau)$. Dans la suite de ce manuscrit, nous adoptons la notation de la définition variationnelle (2.14). Ainsi, nous distinguons les poids (λ_α) des modes temporels ($f_\alpha(t)$).

Cette méthode permet donc de séparer les variables d'une application spatio-temporelle à partir d'une représentation statistique sans imposer un a priori sur la forme des modes spatiaux et temporels et de concentrer l'information pertinente sur les premiers.

2.1.3.3 Illustration géométrique de la SVD

Une interprétation géométrique peut être donnée à la SVD. Pour cela, nous étudions un exemple concernant deux indices indirects de l'activité solaire : l'indice décimétrique *f10.7* et le nombre de neutrons enregistré par la station Climax au Colorado. L'indice décimétrique est la mesure du flux radio émis par le Soleil entier à une longueur d'onde de 10.7 cm. *f10.7* est un proxy du flux solaire UV qui affecte l'ionosphère et la thermosphère terrestre. Quant au flux de neutrons, il résulte de l'interaction des

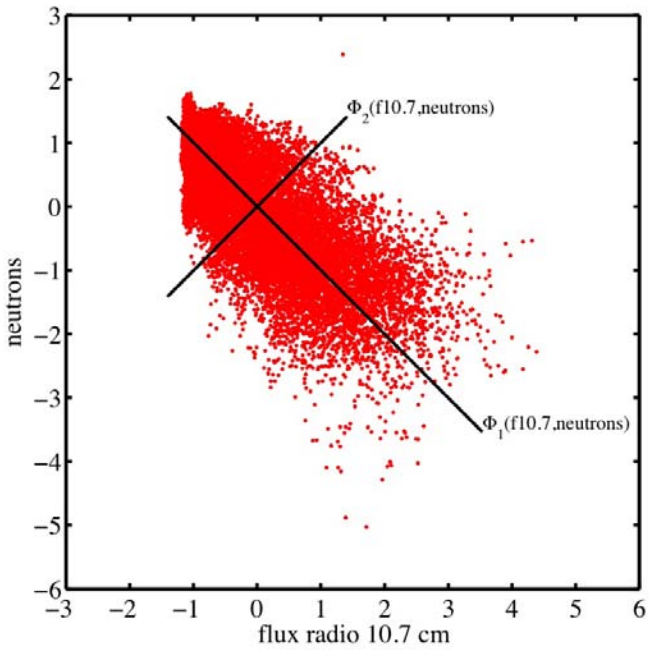


FIG. 2.1 – *Concept de la SVD : les points représentent la corrélation entre le flux de neutrons et le flux radio à 10.7 cm dans l'espace des phases bi-dimensionel des données. Les lignes noires montrent le nouveau système de coordonnées* $\{\Phi_1, \Phi_2\}$ *obtenu par SVD.*

rayons cosmiques avec les particules des couches supérieures de l'atmosphère terrestre. Le rayonnement cosmique est sensible au niveau de turbulence du milieu interplanétaire et au champ magnétique héliosphérique, qui sont eux-mêmes affectés par l'activité solaire. Il en résulte une anti-corrélation entre le flux de neutrons et l'indice *f10.7*. De nombreux auteurs ont déjà étudié cette anti-corrélation (Svensmark, 2000). Nous nous contenterons d'utiliser ici la SVD pour illustrer géométriquement la méthode.

Sur une période allant du 2 janvier 1957 au 30 novembre 2006, les données journalières du flux de neutrons et de l'indice *f10.7* ont été ordonnées dans un tableau de dimension $17\,835 \times 2$, comme une liste de coordonnées de 17 835 points de l'espace des phases $\{\mathrm{f10.7}, \mathrm{neutrons}\}$. La projection des données dans cet espace est indiquée à la figure 2.1. Les données ont été centrées et réduites afin de faciliter la comparaison (*i.e* $\hat{x}_i = \frac{x_i - \langle x \rangle}{\sigma_x}$).

L'utilisation de la SVD revient à chercher une nouvelle paire de variables qui sont à la fois orthogonales et qui extrémalisent la dispersion. Par conséquent, la SVD peut être interprétée comme une rotation de l'espace des phases $\{\mathrm{f10.7}, \mathrm{neutrons}\}$, du système de coordonnées initial à un nouveau système de coordonnées dont les axes orthogonaux coïncident avec les axes d'inerties des données. La figure 2.1 montre comment le pre-

mier axe (Φ_1) du nouveau système de coordonnées ajuste la direction de la variabilité maximale des données. Le second Φ_2, par définition orthogonale et défini au signe près, décrit plutôt la dispersion des données selon une direction perpendiculaire à l'axe Φ_1.

La figure 2.1 illustre géométriquement le concept de réduction de données. Initialement, nous disposions de deux paramètres décrivant l'activité solaire. La SVD a permis de découpler la variabilité selon deux directions privilégiées dans l'espace des phases et ainsi de réduire la variabilité principale à un unique paramètre. En d'autres termes, le premier axe décrit dans une certaine mesure la moyenne des deux paramètres initiaux. Plus la dimension de l'espace de phases est importante (pour les ACFs de SuperDARN, elle sera de dimension 18), plus la projection sur quelques axes s'avère appropriée pour décrire des données multivariées corrélées.

Faisons le lien avec la définition algébrique de la SVD (*cf* § 2.1.3.1). L'équation (2.7) peut s'écrire sous la forme $\mathrm{BG} = \mathrm{F}\Lambda$ puisque la matrice Λ est orthogonale. Cette expression n'est autre que l'application linéaire de B permettant le passage du système de coordonnées initial au système de coordonnées réduit. Dans ce cadre, la direction des axes principaux est donnée par les colonnes de F. La matrice G peut être interprétée comme une matrice de passage entre les deux systèmes de coordonnées. La norme des axes est donnée par leur poids respectif, qui n'est autre que la valeur λ_α, des données le long de cet axe.

2.2 Mise en oeuvre de la méthode de réduction statistique de l'ACF

2.2.1 Constitution d'une base de données

Nous voulons mener une analyse des ACFs afin de définir un nouveau jeu de paramètres (ou de modes) qui soit représentatif des propriétés statistiques des ACFs mesurées dans toute l'ionosphère. Pour cela, nous devons constituer un échantillon qui soit à la fois représentatif et dépourvu de mesures erronées. Il est donc nécessaire de distinguer les ACFs représentatives de l'activité ionosphérique de celles qui sont affectées par des perturbations. En effet, de nombreux mécanismes peuvent dégrader la qualité des mesures : effets de propagation, d'absorption des ondes HF, par les irrégularités électroniques ou encore par la présence d'interférences... L'intensité de ces différentes perturbations dépendent à la fois de la distance des cibles rétrodiffusantes par rapport au radar, mais également de leur localisation à travers l'ionosphère. Par exemple, la

présence d'une activité d'ondes au voisinage du cornet polaire semble être propice à la formation d'ACFs de forme irrégulière (André et al., 2000a,b, 2002). La détection de ces mesures anormales ou «mauvais retards» est délicate. Il est également difficile de déterminer précisément leur nature, *i.e.* s'ils sont dus à des artefacts instrumentaux ou bien s'il sont causés par des «perturbations extérieures» (*e.g.* augmentation de la turbulence, précipitation de particules, variation du champ électrique...). Il est donc nécessaire d'opérer une sélection rigoureuse afin de construire une statistique d'ACFs qui soit à la fois complète et représentative des différents types d'échos radar provenant de la région F. Dans un premier temps, nous nous efforcerons d'éliminer au maximum les échos radar affectés par du bruit ou par des interférences, même si nous omettons certaines informations susceptibles d'être utiles pour la détection des frontières. Ces échos radars se traduisent au niveau des ACFs par des retards dont la puissance mesurée s'écarte de plusieurs ordres de grandeurs de celle escomptée. La détection de ces derniers seront traités plus en détail dans le chapitre 4.

Pour cela, nous avons retenu 4 jours d'observations SuperDARN répartis sur l'ensemble de l'année 2003 : le 12 février 2003, le 24 mai 2003, le 1er septembre 2003 et le 21 décembre 2003. Ces 4 jours ont été choisis pour leur occurrence élevée en termes d'échos radar et pour leur bonne couverture spatio-temporelle de la convection ionosphérique. Tous les échos radar de l'hémisphère nord ont ainsi été regroupés. Nous nous intéressons dans cette étude uniquement aux radars de l'hémisphère nord car le réseau y est plus développé que dans l'hémisphère sud. Son extension spatiale permet de couvrir pendant un peu plus de 12 heures un même secteur. La taille de cet échantillon peut paraître faible par rapport à d'autres études regroupant des saisons entières d'observations SuperDARN (Ruohoniemi and Greenwald, 1997; André et al., 2002; Villain et al., 2002), néanmoins, nous le considérons suffisamment représentatif pour estimer les quantités qui nous intéressent. Nous justifions ceci par le fait que la décomposition des ACFs en valeurs singulières a été testée sur différents échantillons. Pour chacun de ces échantillons (de taille différente, conditions géophysiques de détection des échos différentes), le résultat de la décomposition varie peu (voir § 2.3).

Pour ces 4 jours, l'activité géomagnétique est modérée (Kp < 4) avec dans le vent solaire une rotation régulière de la composante verticale B_z du champ magnétique du Nord vers le Sud. Ce forçage régulier sur la magnétosphère pourrait expliquer l'occurrence élevée des échos radar. Aucun événement géomagnétique intense (choc interplanétaire, sous-orage géomagnétique...) n'est détecté pendant la période. L'absence d'événements intenses peut être interprétée comme une faiblesse pour cette étude, en ce sens que l'échantillon ne décrit pas toutes les conditions que l'on peut observer par les radars SuperDARN. Cependant nous montrerons au § 2.3.1 que la réduction des

ACFs (au sens de la SVD) est peu affectée par le niveau de l'activité géomagnétique.

Afin de constituer notre base de données, il nous faut sélectionner uniquement des ACFs «propres». Pour cela, nous appliquons un certain nombre de critères basés sur la distance des échos radar et le diagnostic de l'algorithme FitACF. Nous conservons les données provenant seulement des radars fonctionnant dans le mode opératoire commun afin d'éviter la contamination par des mesures issues de modes de fonctionnement spécifiques. Une restriction sur les échos radar de région F est nécessaire car aux altitudes associées à cette région, le plasma est directement contrôlé par l'activité magnétosphérique. Il n'y a pas, comme en région E, des effets de collision qui rendent l'interprétation plus délicate (Villain et al., 1987). Pour éliminer une majorité d'échos de région E, nous sélectionnons les ACFs associées à des distances d'au moins 900 km du radar (André et al., 2002; Villain et al., 2002) (se reporter à la discussion au § 1.3.2.2). Un moyen sûr d'évaluer l'altitude des cibles rétrodiffusantes serait de faire une étude de tracé de rayon afin d'étudier le chemin de propagation de chacun des faisceaux. Ceci est difficilement envisageable dans le cadre d'une étude statistique regroupant plusieurs dizaines de milliers d'échos radar car il faudrait obtenir un profil en altitude de densité électronique pour chacun des cas et en déduire les conditions de propagation des ondes.

A partir de l'algorithme FitACF, la largeur spectrale est calculée à partir d'un ajustement de l'ACF, au sens des moindres carrés, avec un modèle gaussien, s, ($|\mathrm{R}| \sim e^{-\sigma^2\tau^2}$) et avec un modèle exponentiel, l, ($|\mathrm{R}| \sim e^{-\lambda\tau}$). Pour chaque modèle, l'erreur quadratique relative, $\langle\epsilon_{l,s}\rangle$:

$$\langle\epsilon_{l,s}(t)\rangle = \sqrt{\frac{1}{N-2}\sum_{i=1}^{N}\left(\frac{R(t,\tau_i)-\widehat{R}_{l,s}(t,\tau_i)}{R(t,\tau_i)}\right)^2}, \tag{2.15}$$

issue de l'ajustement donne une indication sur l'adéquation avec chacun des modèles. Nous conservons, empiriquement, les données pour lesquelles les deux erreurs relatives normalisées sont inférieures à 1. En utilisant un seuil plus restrictif, nous rejetons non seulement des mesures dégradées mais également des mesures qui sont mal reproduites à la fois par les modèles gaussien et exponentiel. Après avoir testé plusieurs valeurs, le seuil de 1 correspond à un bon compromis entre la présence de bruit dans l'échantillon et la représentativité des ACFs.

Les ACFs sont souvent entachées de valeurs aberrantes, qui se traduisent par des écarts importants pour certains retards entre la puissance mesurée et la décroissance escomptée de l'ACF. La présence de ces mauvais retards oblige leur élimination lors de l'ajustement des modèles. Or, la méthode SVD n'autorise pas des valeurs manquantes.

Il est donc nécessaire de construire une base de données dépourvue de ces ACFs erronées. Ceci peut laisser craindre que notre statistique est biaisée par une sélection trop drastique. Dans un premier temps, cette contrainte est nécessaire afin de définir des modes réguliers. Pour cela, nous conservons uniquement les ACFs pour lesquelles l'algorithme FitACF (avec ce qu'il comporte comme défauts, *cf* Annexe A) ne détecte aucun mauvais retard. La rétrodiffusion au sol, caractérisée par des petites vitesses et par des faibles largeurs spectrales, est également éliminée de l'échantillon. Pour cela, nous utilisons le critère défini par Baker (2003) :

$$G(V_D, W) = |V_D| - V_D^{max}\left(1 - \frac{W}{W^{max}}\right), \tag{2.16}$$

où V_D est la vitesse Doppler ($V_D^{max} = 200$ m.s^{-1}) et où W est la largeur issue du modèle exponentiel ($W^{max} = 90$ m.s^{-1}). Les échos radar avec $G \leqslant 0$ sont considérés comme issus de la rétrodiffusion au sol.

Nous utilisons enfin les ACFs avec un rapport signal/bruit (S/B) supérieur à 4 dB et une largeur spectrale inférieure à 600 m.s^{-1}. Au-delà de 600 m.s^{-1}, l'observation des ACFs montre que le temps de décorrélation (même s'il demeure réaliste d'un point de vue de la physique) est trop rapide pour que la largeur spectrale puissent être correctement ajustée par un des deux modèles précités.

Après application des différentes contraintes (résumées dans le tableau 2.1), la base de données contient 21 972 ACFs, soit 20% des observations enregistrées pour l'hémisphère nord. Le 12 février 2003 contribue pour 32%, le 24 mai 2003 pour 6%, le 1er septembre 2003 pour 32% et le 21 décembre 2003 pour 27% des échos. Cet échantillon est suffisamment grand et représentatif pour permettre une estimation correcte des quantités statistiques pour les conditions géomagnétiques spécifiées. Nous avons vérifié que le modèle statistique que nous obtenons à partir de ces quatre journées ne change pas de manière significative si d'autres jours sont ajoutés (*cf* § 2.3).

Tab. 2.1 – *Critères de sélection des ACFs*

rapport S/B	$\geqslant 4$ dB
Nombre de mauvais retards par ACF	0
Erreur quadratique relative, $\langle \epsilon_{l,s} \rangle$	$\leqslant 1$
Largeur spectrale	$\leqslant 600$ m.s^{-1}
Distance radiale au radar	$[900, 3500]$ km
$G(V_D, W)$	$\geqslant 0$

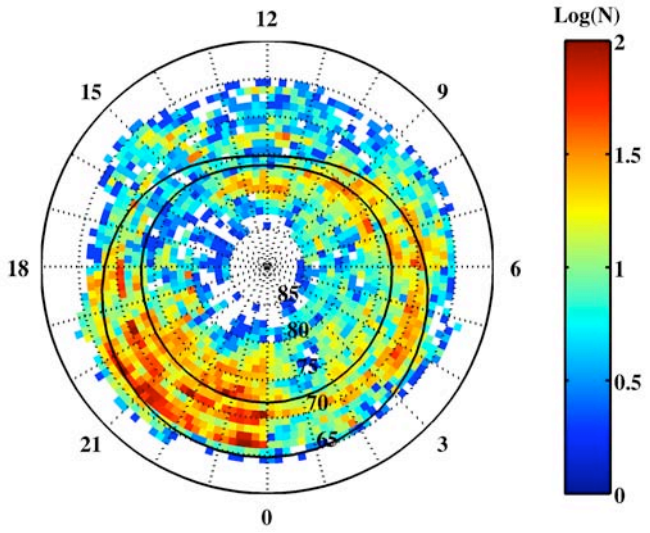


FIG. 2.2 – *Occurrence des échos radar en MLT/Λ (latitude invariante magnétique), basée sur le système de coordonnées AACGM (Baker and Wing, 1989). Le logarithme de l'effectif pour chaque cellule est codé avec une échelle de couleur. L'ovale auroral, défini pour une activité géomagnétique modérée ($3 < Kp < 4^-$) est superposée (Holzworth and Meng, 1975).*

La figure 2.2 montre la distribution spatiale des échos radar en fonction du temps magnétique local (MLT) et de la latitude invariante magnétique (Λ) basée sur le système de coordonnéess $AACGM$[11] (Baker and Wing, 1989). Dans ce système de coordonnées, les données sont ramenées aux pieds des lignes de champ, à une altitude de 300 km. Cette figure montre que notre statistique est cohérente avec celles obtenues sur des échantillons plus grands (Ruohoniemi and Greenwald, 1997; Villain et al., 2002; André et al., 2002; André and Dudok de Wit, 2003). La distribution est loin d'être uniforme, ce qui révèle l'impact du couplage magnétosphère-ionospère sur le niveau de turbulence et la présence d'irrégularités dans l'ionosphère. L'accroissement du taux de rétrodiffusion côté nuit est relativement bien co-localisé avec l'ovale auroral défini par Holzworth and Meng (1975), pour une activité géomagnétique modérée (Kp < 4).

2.2.2 Réduction de l'ACF

Notre échantillon d'ACFs se présente sous la forme d'un tableau rectangulaire R de dimensions $21\,972 \times 18$, dans lequel chaque ligne représente le module d'une ACF enregistrée à un instant t ; les colonnes correspondent aux retards τ. Par SVD, la dé-

[11] En anglais : *Altitude Adjusted Corrected GeoMagnetic.*

composition en $n = 18$ modes d'une ACF enregistrée au temps t, s'écrit

$$R(t,\tau) = \sum_{\alpha=1}^{n} \lambda_\alpha \cdot f_\alpha(t) \cdot g_\alpha(\tau) \ , \tag{2.17}$$

où chaque mode $g_\alpha(\tau)$ représente la α'ième fonction de base orthonormée de l'ACF, $f_\alpha(t)$ décrit la contribution du mode g_α à la reconstruction de l'ACF au temps t et ${\lambda_\alpha}^2$ donne l'énergie du α-ième mode. La condition d'orthogonalité des modes entraîne l'unicité de la décomposition.

Dans ce type d'analyse statistique, la préparation des données est importante. Ici, la SVD est appliquée aux ACFs brutes sans normalisation. Cela implique donc que la statistique est plus fortement pondérée par des événements intenses et par les régions générant le plus d'échos radar. Sur cet échantillon, la puissance rétrodiffusée définie par le retard zéro de l'ACF ($R(t,0)$) peut varier de 1 000 à 100 000 en unités mesurées. Une approche différente consiste, par exemple, par une normalisation préalable de chaque ACF par la puissance au retard zéro ou par la puissance moyenne de l'ACF. Une telle normalisation a l'avantage de donner le même poids à toutes les ACFs mais elle a l'inconvénient d'augmenter la contribution des échos avec un plus faible rapport signal/bruit ; la présence accrue d'ACFs bruitées affectera les modes de la SVD. En particulier, la séparation est moins nette entre les modes reproductibles caractérisés par une forte variance et ceux qui sont noyés dans le bruit (et donc moins reproductibles). La figure 2.3 illustre justement la distribution de l'énergie (λ_α^2) captée par chacun des 18 modes pour ces trois configurations : pour le cas où la méthode SVD a été appliquée sur les ACFs brutes (points bleus), sur les ACFs qui ont été normalisées par leur puissance au retard nul (diamants verts) et dans le cas où chaque ACF a été normalisée par sa puissance moyenne, *i.e.* moyennée sur les 18 retards (triangles rouges). Pour faciliter la comparaison, l'énergie de chaque mode a été normalisée par l'énergie totale ($\sum_{k=1}^{n} {\lambda_k}^2$). On peut notamment constater que les modes sont ordonnés par ordre décroissant d'importance. La superposition des trois distributions résultantes confirme bien que la séparation entre les modes reproductibles est plus évidente lorsque la méthode SVD est appliquée directement sur les ACFs brutes, sans préparation particulière (au préalable) de l'échantillon. Dans ce qui suit, nous nous concentrerons donc sur les modes obtenus, dans le cas où la méthode SVD appliquée directement sur les ACFs.

On peut également évaluer une ACF moyenne de l'échantillon et retrancher celle-ci à chaque ACF avant d'appliquer la SVD. Cela permet de se focaliser sur les variations dans la forme de l'ACF. Ceci est envisageable dans le mesure où l'on cherche à caractériser un écart par rapport à un équilibre bien déterminé. Or avec les données

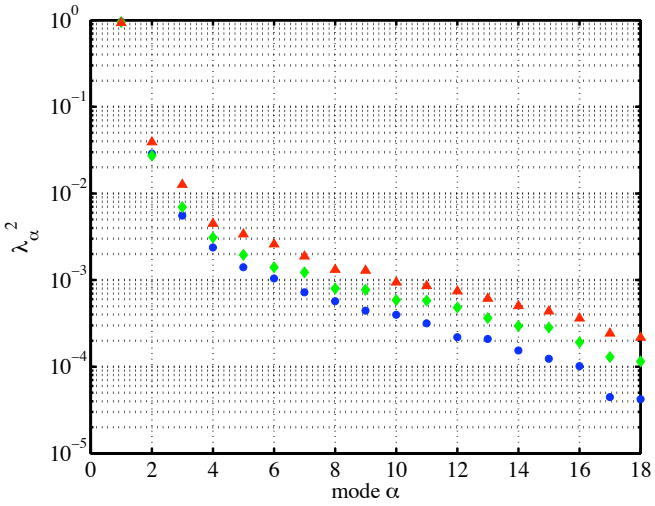


FIG. 2.3 – *Distribution des énergies captées par chacun des 18 modes, dans les cas ou la méthode SVD est appliquée aux ACFs brutes : sans normalisation (points bleus), avec normalisation par la puissance au retard nul (diamants verts) et avec une normalisation par la puissance moyenne de l'ACF (triangles rouges). Chaque énergie est exprimée comme une fraction de l'énergie totale ($\sum_{k=1}^{n} \lambda_k^2$). La variabilité des données est concentrée dans les premiers ordres.*

SuperDARN, l'ACF moyenne fait elle aussi partie des paramètres intéressants, qu'il convient donc de conserver.

La clef de l'interprétation de la SVD réside dans l'étude de la distribution des poids λ_α (ou de l'énergie ${\lambda_\alpha}^2$). Les poids issus de la décomposition donnent en effet une indication sur le degré de cohérence spatiale et temporelle des données. La figure 2.3 montre bien que la distribution de l'énergie est piquée, avec uniquement les quelques premiers modes qui se détachent du reste de la distribution. Ceci suggère qu'un faible nombre de modes captent l'essentiel des caractéristiques des données sélectionnées. En effet, ${\lambda_\alpha}^2$ décrit l'énergie des données selon la α-ième direction de l'espace de la mesure. Si chaque ACF était une fonction aléatoire de t et de τ, une distribution quasi-plate serait alors observée, puisqu'aucune direction ne serait privilégiée. La forme piquée de la distribution est donc une conséquence directe du degré de redondance de l'information contenue dans chacune des ACFs. En projetant les données sur un seul mode, 95% de l'énergie totale est décrite, sur deux modes 98.5%, sur trois modes 99.2% ...

Il n'existe pas de critère pertinent permettant l'estimation du nombre de modes significatifs. En général, le point d'inflexion du spectre des valeurs singulières est utilisé. Ce critère peut être justifié par les travaux de Broomhead and King (1986) et de Elsner and Tsonis (1996). Les auteurs expliquent la forme coudée du spectre par la présence de bruit additif dans le jeu de données. Prenons le cas d'un jeu de données bruité et

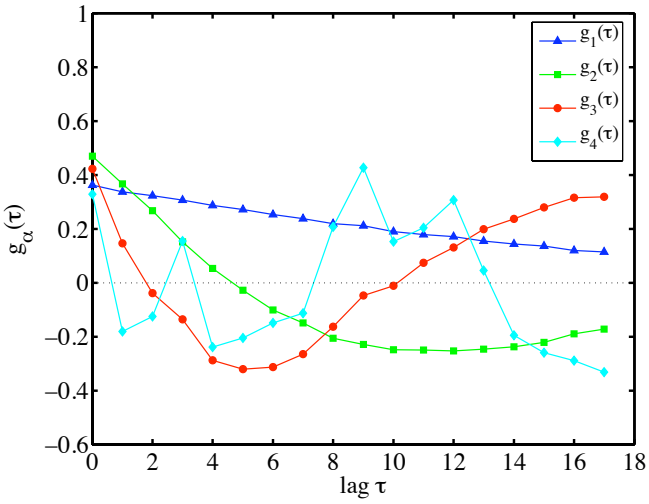


FIG. 2.4 – *Les quatre premiers modes $g_\alpha(\tau)$, avec $\alpha = [1, 2, 3, 4]$. Seul les trois premiers possèdent une allure régulière et reproductible. Ainsi, ces trois modes sont utilisés pour reconstruire toutes les ACFs.*

caractérisons ce bruit par une loi normale $N(0, \sigma^2)$ de moyenne nulle et de variance, σ^2. Il est alors raisonnable de supposer que la partie cohérente du signal (dans notre cas la décroissance escomptée de l'ACF) soit décorrélée du bruit. La distribution des poids, ${\lambda_\alpha}^2$, qui décrit le degré d'information capté par chacun des modes, peut alors se décomposer en une somme de deux termes :

$$ {\lambda_\alpha}^2 = \overline{\lambda_\alpha}^2 + M\sigma^2 \tag{2.18} $$

où $\overline{\lambda_\alpha}^2$ est l'énergie du mode α issu du jeu de données non bruité et M la dimension la plus élevée de la matrice de données. L'équation 2.18 révèle l'existence d'un plateau dans la distribution des poids λ_α. L'ordination des poids, $\overline{\lambda_\alpha}^2$, qui tend vers zéro avec l'ordre des modes, montrent que les modes d'ordre élevé captent des structures avec des échelles de plus en plus petite. L'apparition d'un plateau suggère donc que les modes d'ordre élevé sont dominés par le bruit, ce qui limite la taille des petites structures qui peut être caractérisée par la méthode SVD.

Dans notre étude, les fluctuations observées dans les mesures SuperDARN ne peuvent pas être assimilées à un simple bruit blanc. Le plateau s'incline alors, et ce d'autant plus que le bruit est corrélé dans l'espace ou dans le temps. Le point d'inflexion reste cependant un critère qualitatif pour séparer les modes significatifs de ceux qui sont dominés par du bruit.

Les quatre premières fonctions de base $g_\alpha(\tau)$, avec $\alpha = [1, 2, 3, 4]$, sont exposées à la figure 2.4. On notera une similitude avec la décomposition de Fourier, dans laquelle on projette aussi sur une base de modes dont le nombre de noeuds va en augmentant. Seules les trois premières fonctions connaissent une variation régulière selon τ. Les modes d'ordre 4 et plus ont une allure plus irrégulière. Contrairement aux trois premiers modes, qui sont relativement reproductibles, la forme des modes pour $\alpha \geqslant 4$ varie en fonction des critères de sélection. Le premier mode $g_1(\tau)$ est simplement la moyenne pondérée de l'ensemble des ACFs contenues dans l'échantillon et ressemble à une exponentielle décroissante. Les modes $g_2(\tau)$ et $g_3(\tau)$ peuvent être interprétés comme des corrections d'ordre supérieur nécessaires pour décrire toute la gamme possible des formes. Le point d'inflexion dans la figure 2.3 suggère que les 3 - 4 premiers modes sont les plus significatifs. Etant donné le faible poids des modes 4 et supérieurs, ainsi que leur manque de reproductibilité, nous les négligerons par la suite. Une reconstruction précise des ACFs est donc possible à partir du développement tronqué :

$$\hat{R}(t, \tau) = \lambda_1 f_1(t) g_1(\tau) + \lambda_2 f_2(t) g_2(\tau) + \lambda_3 f_3(t) g_3(\tau) \tag{2.19}$$

Nous remarquons que les modes d'ordre supérieur à 1 possèdent des valeurs négatives. Ceci est dû à la condition d'orthogonalité des modes de la méthode SVD. Sur le principe, on peut reprocher à la SVD de produire des modes non réalistes (dans notre cas avec des valeurs négatives). Il est possible d'y pallier, en utilisant des méthodes plus sophistiquées de classification, telles que la séparation Bayésienne en source positive (Cardoso, 1998). Or le résultat de ces méthodes n'est pas forcément unique, et l'estimation des modes est nettement plus coûteuse en temps de calcul. Nous n'envisagerons donc pas ce genre de solution dans la suite.

L'équation (2.19) suggère que chaque ACF peut être caractérisée par les trois quantités $\{f_1(t), f_2(t), f_3(t)\}$ uniquement. Il est important de noter que cette description compacte des ACFs est issue d'une description statistique par opposition aux méthodes classiques qui imposent un modèle physique ou empirique. La réduction de données, elle, impose peu de contraintes sur la forme des ACFs. La figure 2.5 illustre comment une variété d'ACFs de forme différente peut être ajustée avec trois modes seulement. Afin d'apprécier le caractère général de cette description, même si l'échantillon utilisé pour la SVD ne représente que 20% du nombre total d'échos radar détectés, il est important de tester maintenant cette représentation compacte selon différentes conditions ionosphériques.

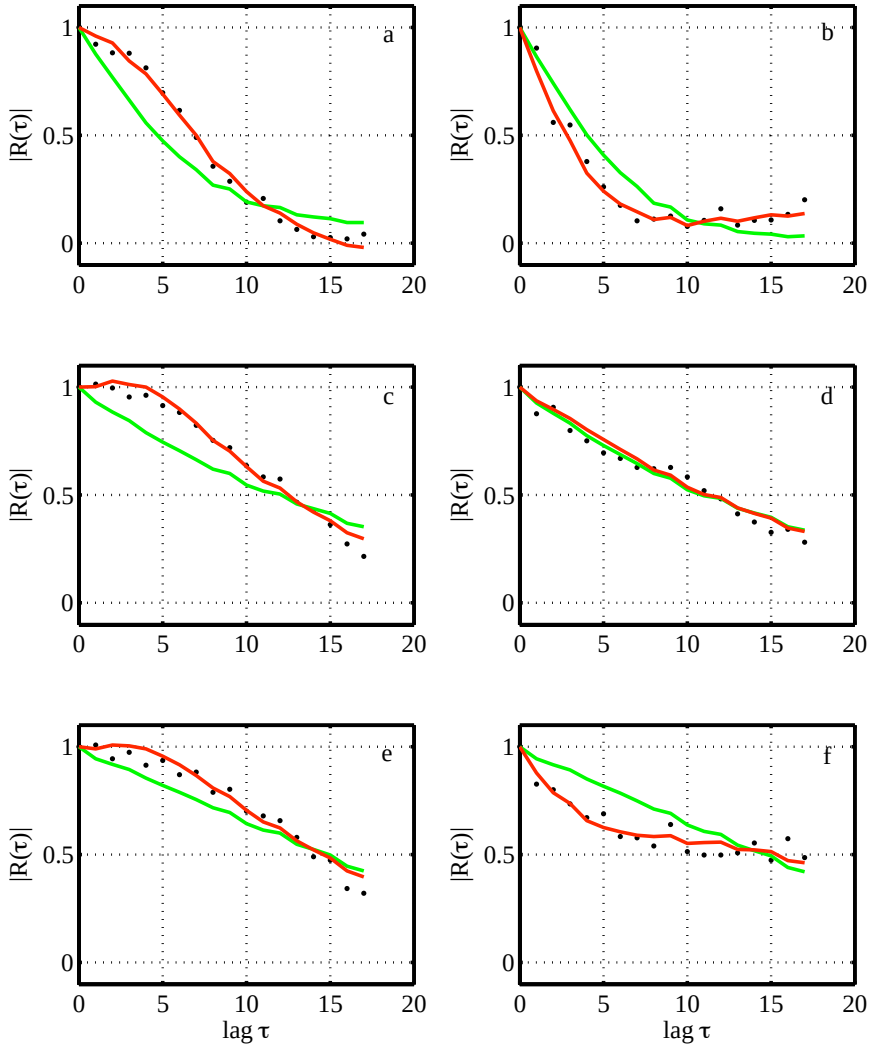


FIG. 2.5 – *Illustration de six formes typiques d'ACFs reconstruites par le modèle de la SVD. Chacune des ACFs (points noirs) a été reconstruite à partir des deux premiers modes (trait vert) et à partir des trois premiers modes (trait rouge). Les panneaux (de a à f) se réfèrent à la position des différentes ACFs dans l'espace des paramètres définis à la figure 3.2.*

2.3 Validation de la méthode

Dans cette section, nous allons éprouver la robustesse de la réduction statistique des ACFs. Pour cela, nous allons étudions l'incidence de l'activité géomagnétique et de la dépendance zonale (secteur horaire magnétique local) sur la définition des modes SVD, *i.e.* sur le nombre de modes significatifs, mais également sur leur allure respective.

2.3.1 Influence de l'activité géomagnétique

La présence d'irrégularités électroniques en région F, capables de rétrodiffuser les ondes radio, est fortement reliée au niveau de perturbation géomagnétique (via les précipitations de particules, la génération de turbulence ...). Ruohoniemi and Greenwald (1997) ont étudié en détail l'influence de cette activité sur la dépendance spatiale de l'occurrence des échos radar durant un maximum d'activité du cycle solaire. La question ici est de savoir si la dépendance spatiale des propriétés de la rétrodiffusion en fonction du niveau d'activité géomagnétique peut avoir un impact sur la représentativité des modes statistiques $g_\alpha(\tau)$ définis à la section 2.2.2.

Pour cela, nous opérons une sélection d'événements en fonction du niveau de perturbation géomagnétique afin de comparer le résultat de la réduction des ACFs par la méthode SVD. A partir de la valeur journalière de l'indice géomagnétique Kp, nous classons toutes les journées d'observations de l'année 2003 en quatre catégories différentes. Chacune de ces catégories caractérise différents niveaux de l'activité magnétosphérique, d'une activité calme, modérée à une activité très intense $\{0 < \mathrm{Kp} \leqslant 2,\ 2 < \mathrm{Kp} \leqslant 4,\ 4 < \mathrm{Kp} \leqslant 6,\ 6 < \mathrm{Kp} \leqslant 8\}$.

D'autre part, l'accessibilité des cibles rétrodiffusantes en région F par les ondes radios HF dépend fortement de l'état de ionisation des régions E et D (via les mécanismes de réfraction et d'absorption), en région F (de la décorrélation du front d'onde)

Tab. 2.2 – *Sélection des événements*

Intervalles	$0 < \mathrm{Kp} \leqslant 2$	$2 < \mathrm{Kp} \leqslant 4$	$4 < \mathrm{Kp} \leqslant 6$	$6 < \mathrm{Kp} \leqslant 8$
Evénements	06/01/2003 07/01/2003	20/12/2003 21/12/2003	16/11/2003 17/11/2003 05/12/2003	18/08/2003 29/10/2003 30/10/2003 31/10/2003 20/11/2003
Nombre d'échos radar	22 829	17 158	18 452	18 529

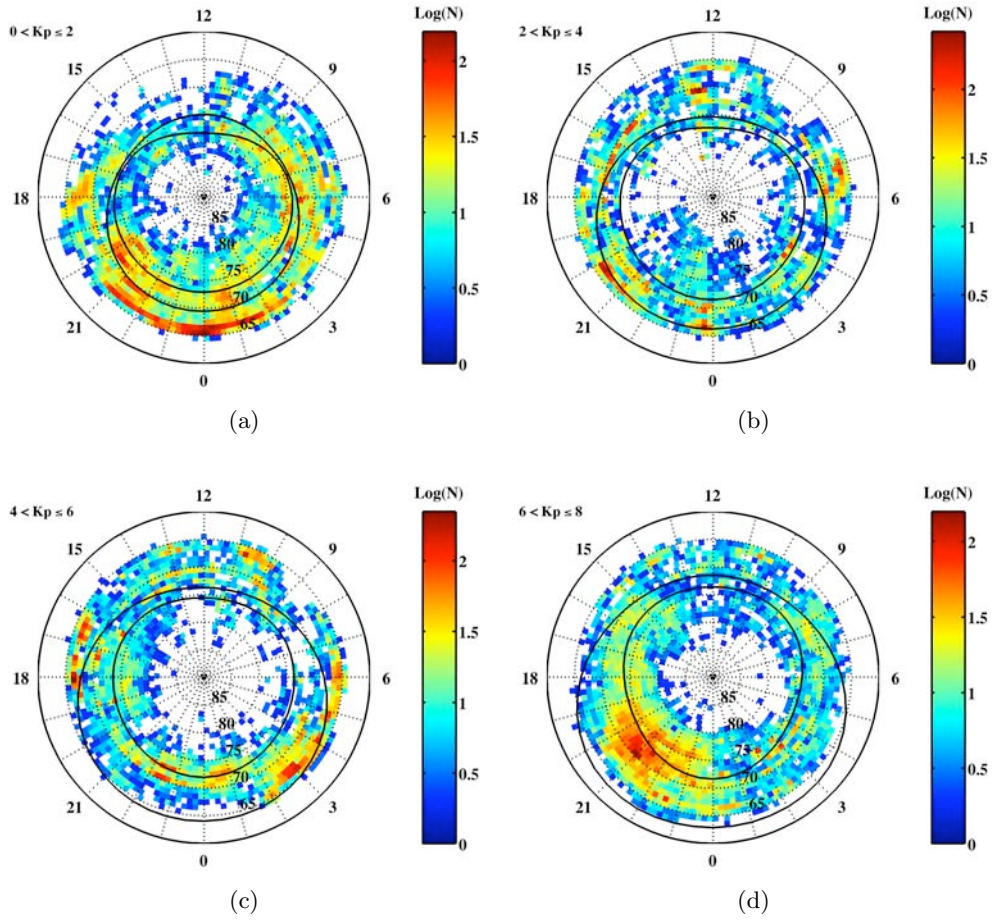


FIG. 2.6 – *Distribution en MLT/Λ (comme dans la Fig. 2.2) des échos radar selon différents niveaux d'activité géomagnétique définie à partir de l'indice Kp :* $0 < \mathrm{Kp} \leqslant 2$ *(a),* $2 < \mathrm{Kp} \leqslant 4$ *(b),* $4 < \mathrm{Kp} \leqslant 6$ *(c),* $6 < \mathrm{Kp} \leqslant 8$ *(d). L'occurrence est codée avec une échelle de couleur logarithmique. Un modèle d'ovale auroral, calculé à partir de la valeur médiane de chacun des intervalles de* Kp*, est superposé (Holzworth and Meng, 1975).*

et par conséquent est fonction du niveau de perturbations Ruohoniemi and Greenwald (1997). En d'autres termes, lorsque l'activité magnétosphérique devient élevée, le couplage entre magnétosphère et l'ionosphère est modifié, se traduisant par une expansion vers les basses latitudes de l'ovale auroral, modifiant les conditions de propagation et d'absorption des ondes HF. Les régions rétrodiffusantes se déplacent alors vers les plus basses latitudes, en deçà du champ de vue des radars SuperDARN. Ceci se traduit par une diminution de la production des échos radar lorsque l'indice Kp augmente (Milan et al., 1997). Notre objectif est de construire quatre échantillons de taille comparable. Nous avons donc regroupé des journées pour chaque classe de Kp.Tous les échos radar de l'hémisphère nord ont été regroupés et classés avec les mêmes critères de sélection

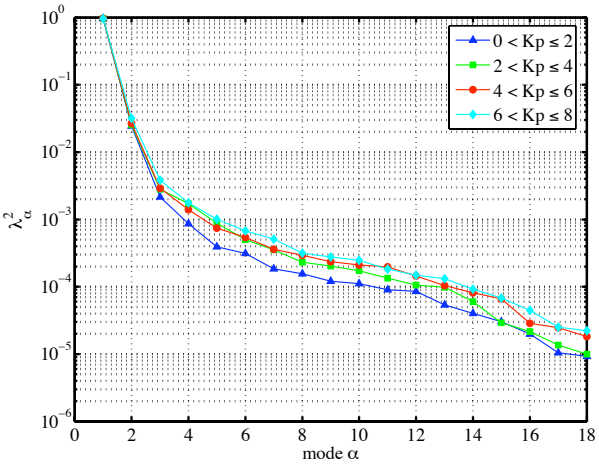


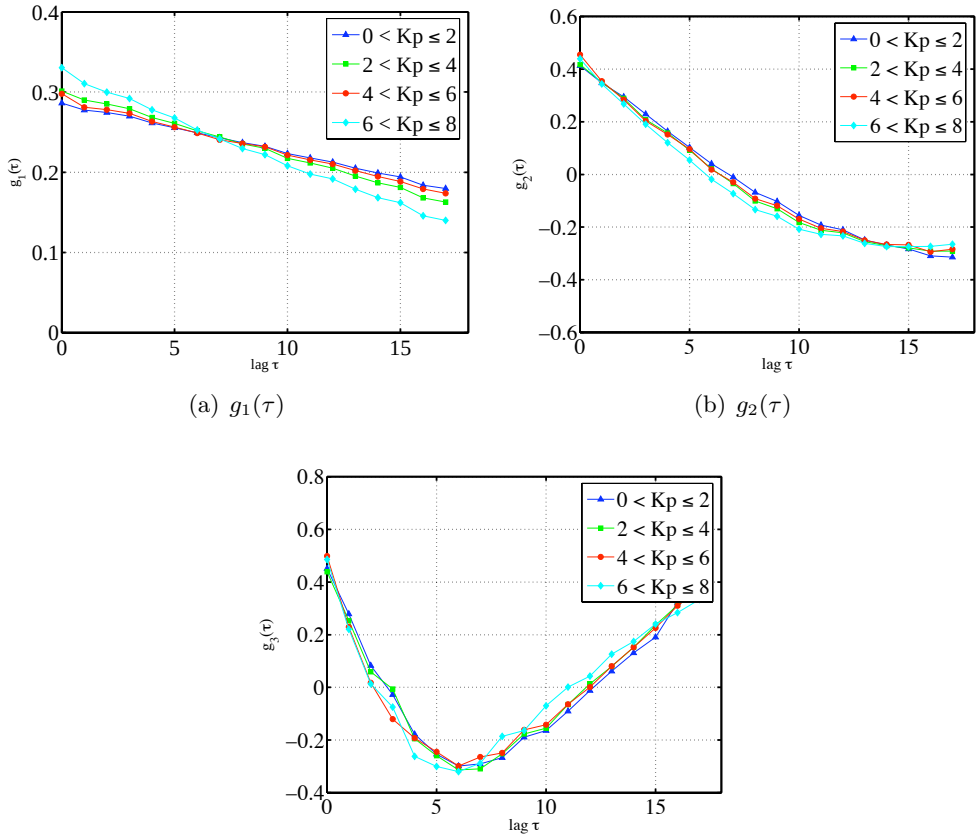
FIG. 2.7 – *Comparaison de la distribution des énergies normalisées par l'énergie totale* ($\sum_{k=1}^{18} {\lambda_k}^2$) *pour les quatre intervalles de l'activité géomagnétique*

qu'à la section 2.1. La taille respective de chacun des échantillons est résumée dans le tableau 2.2, avec les événements qui y contribuent.

La figure 2.6 montre la répartition en MLT/Λ (système de coordonnées AACGM) des échos radar pour chacune des quatre catégories. Bien qu'il soit difficile de comparer cette statistique avec celle présentée par Ruohoniemi and Greenwald (1997, planche 3), des similitudes existent. Pour les quatre intervalles, le côté nuit apparaît plus productif et est associé aux frontières de l'ovale auroral, tandis que la probabilité d'observer la région du cornet polaire côté jour reste faible.

Nous appliquons la méthode SVD indépendamment sur les quatre échantillons d'ACFs et ainsi obtenons quatre ensembles de fonctions de base avec leur distribution d'énergie associée (*cf* Eq. (2.17)). La figure 2.7 compare le spectre d'énergie relatif aux quatre intervalles d'activités géomagnétiques. La distribution des poids ne révèle pas de différence majeure. Pour chacun des échantillons, les trois premiers modes surclassent nettement les modes d'ordre supérieur et sont de même amplitude relative. L'augmentation des modes d'ordre élevé avec le Kp correspond à une augmentation du niveau de fluctuations aléatoires.

La forme et l'amplitude des modes, $g_\alpha(\tau)$, varient peu d'un niveau de Kp à un autre (Fig. 2.8). Pour l'ensemble de la gamme d'activités géomagnétiques, le premier mode décroît de manière monotone. Cependant, nous pouvons remarquer une différence : la fonction $g_1(\tau)$ définie pour une activité intense ($6 < \mathrm{Kp} \leqslant 8$) possède une décroissance plus forte selon τ que dans le cas d'une activité plus faible. Ceci peut être vu à travers la



(a) $g_1(\tau)$ (b) $g_2(\tau)$

(c) $g_3(\tau)$

FIG. 2.8 – *Comparaison des trois premiers modes $g_1(\tau)$ (a), $g_2(\tau)$ (b) et $g_3(\tau)$ (c), pour différents niveaux d'activité géomagnétique.*

distribution de la largeur spectrale qui diffère d'un échantillon à un autre. Notamment la valeur médiane de la largeur spectrale qui est de l'ordre de 116 m.s^{-1} pour l'intervalle $6 < \mathrm{Kp} \leqslant 8$, tandis que pour les trois autres sous-échantillons, la valeur médiane varie uniquement autour de $\sim$ 65 m.s^{-1}. Un tel décalage de la valeur médiane traduit la présence plus importante d'ACFs possédant une décorrélation rapide dans le cas d'une activité élevée. Ceci se manifeste au niveau de la réduction des ACFs par une augmentation de la pente du premier mode. Quant aux modes $g_2(\tau)$ et $g_3(\tau)$, ils conservent leur allure propre par rapport à celle définie au § 2.2.2, apportant une correction d'ordre supérieur nécessaire à la description des différentes formes d'ACF.

Nous pouvons en conclure que les modes définis dans ce travail pour une faible valeur de Kp restent représentatifs de la signature radar (à l'échelle de l'ionosphère aurorale et polaire) quelles que soient les conditions géomagnétiques.

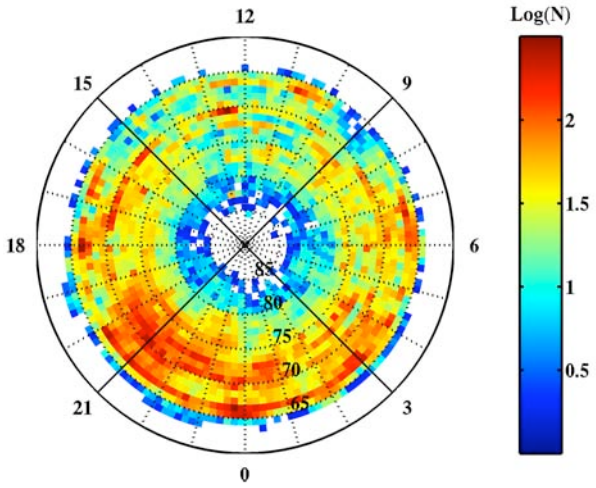


FIG. 2.9 – *Distribution en MLT/Λ (ACCGM) des échos radar. L'occurence est codée avec une échelle de couleur logarithmique. Les quatre sous-échantillons sont délimités par les traits pleins noirs radiaux.*

2.3.2 Dépendance selon le secteur MLT

Dans cette partie, nous nous intéressons à la dépendance des modes en fonction de l'heure locale magnétique. Pour cela, nous opérons une séparation en quatre secteurs MLT des échos radar précédemment sélectionnés (*cf* § 2.3.1).

Ainsi, nous regroupons 27 767 ACFs côté nuit $\{21^{\mathrm{h}}00 < \mathrm{MLT} \leqslant 03^{\mathrm{h}}00\}$, 16 460 ACFs côté matin $\{03^{\mathrm{h}}00 < \mathrm{MLT} \leqslant 09^{\mathrm{h}}00\}$, 11 273 ACFs l'après-midi $\{09^{\mathrm{h}}00 < \mathrm{MLT} \leqslant 15^{\mathrm{h}}00\}$ et 21 528 ACFs côté soir $\{15^{\mathrm{h}}00 < \mathrm{MLT} \leqslant 21^{\mathrm{h}}00\}$. Ce choix des secteurs MLT est motivé par leur dynamique propre face aux variations du milieu interplanétaire. La figure 2.9 montre la répartition spatiale MLT/Λ des échos radar sélectionnés. Les traits pleins de couleur noire (radiaux) délimitent les quatre sous-echantillons en fonction du secteur MLT.

Pour chacun des 4 secteurs MLT, nous appliquons la SVD. La figure 2.10 présente la distribution des poids calculée pour chacun des quatre secteurs. Cette figure suggère là encore que les trois premiers modes semblent être suffisants pour une description précise de la mesure sur la totatilé des secteurs MLT. On note toutefois que l'énergie des trois premiers modes diffère selon le secteur MLT. La différence la plus significative est à noter pour les modes d'ordre supérieur à 3 . Bien qu'en général, le troisième mode reste très faible devant les modes 1 et 2, le mode 3 défini à partir du secteur $\{09^{\mathrm{h}}00 < \mathrm{MLT} \leqslant 15^{\mathrm{h}}00\}$ surclasse celui des autres secteurs d'un facteur variant entre 2

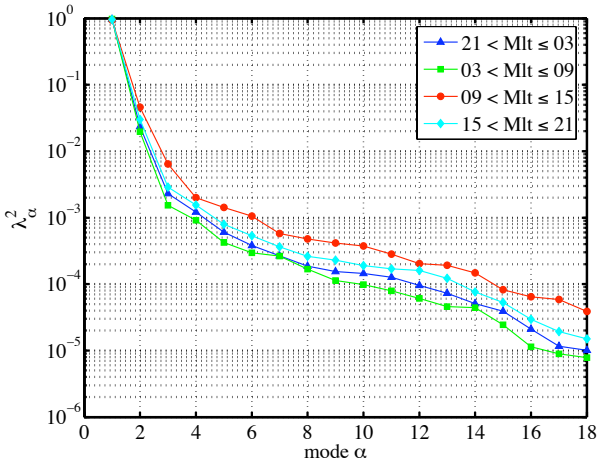


FIG. 2.10 – *Même représentation que dans la figure 2.7, mais en fonction du secteur MLT.*

et 7 (*cf* tableau 2.3). Ce résultat montre que le troisième mode est davantage nécessaire pour une représentation précise de la mesure dans le secteur midi que sur les flancs ou encore dans le côté nuit. Ceci peut s'expliquer par la présence d'ACFs de formes très piquées ou encore modulées dues à des précipitations structurées d'électrons ($\sim$ 100eV) dans la région du cornet polaire (Rodger et al., 1995; Dudeney et al., 1998). Comme ces ACFs ont des formes plus complexes, la distribution des énergies, ${\lambda_\alpha}^2$, décroît moins rapidement que pour le côté nuit, où la reproductibilité entre les ACFs est meilleure et l'ajustement plus aisé.

Notons que la distribution des poids pour l'intervalle $\{09^{\mathrm{h}}00 < \mathrm{MLT} \leqslant 15^{\mathrm{h}}00\}$ possède un plateau qui est supérieur aux trois autres. Ce décalage en amplitude indique que le niveau de bruit est différent selon le secteur MLT. En effet, la distribution du rapport signal/bruit montre que le niveau de bruit moyen est plus élevé pour ce secteur (*cf* Fig. 2.11). La présence plus importante d'échos radar de faible puissance se répercute normalement sur l'augmentation des poids des modes d'ordre supérieur à 3 pour ce

TAB. 2.3 – *Dépendance de l'énergie relative du mode α en fonction sur secteur* MLT *(exprimée en %).*

	${\lambda_1}^2$	${\lambda_2}^2$	${\lambda_3}^2$	${\lambda_4}^2$
$21:00 < \mathrm{MLT} \leqslant 03:00$	97.08	2.35	0.23	0.12
$03:00 < \mathrm{MLT} \leqslant 09:00$	97.63	1.15	0.09	0.04
$09:00 < \mathrm{MLT} \leqslant 15:00$	94.03	4.59	0.63	0.20
$15:00 < \mathrm{MLT} \leqslant 21:00$	96.26	2.99	0.29	0.15

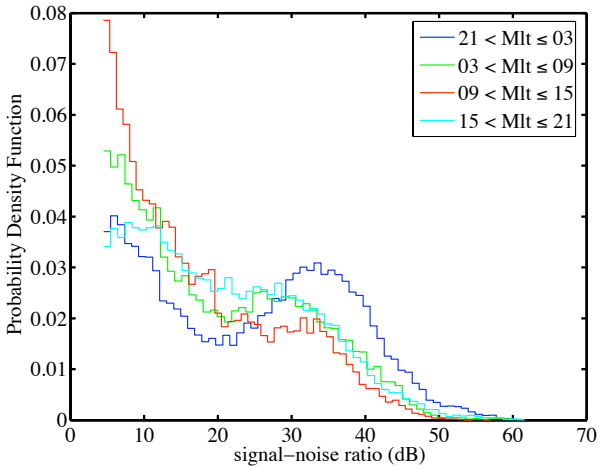


FIG. 2.11 – *Evolution de la fonction de densité de probabilité du rapport signal/bruit (estimé à partir du programme FitACF) en fonction du secteur MLT.*

même secteur MLT. L'augmentation du niveau de bruit peut être interprétée comme la conséquence du chauffage de l'ionosphère par photo-ionisation des couches supérieures de l'atmosphère neutre. L'augmentation de la densité de l'ionosphère entraîne des perturbations sur la transmission des ondes HF jusqu'aux cibles rétrodiffusantes (Hanuise et al., 2006).

Les panneaux (a), (b), (c) et (d) de la figure 2.12 comparent respectivement l'allure des modes $g_1(\tau)$, $g_2(\tau)$, $g_3(\tau)$ et $g_4(\tau)$ en fonction du secteur horaire MLT. La décomposition des ACFs en fonction du secteur MLT confirme les résultats présentés précédemment. Le secteur horaire MLT n'influence pas de manière significative la définition des premiers modes. Pour l'ensemble des secteurs MLT, la décroissance selon τ du mode $g_1(\tau)$ est quasi-linéaire. Seul le secteur $\{09^{\mathrm{h}}00 < \mathrm{MLT} \leqslant 15^{\mathrm{h}}00\}$ semble se démarquer légèrement des trois autres. On remarque notamment que la décroissance des trois premiers modes est plus rapide. Ceci est la conséquence de la distribution de la largeur spectrale qui est décalée vers de plus grandes valeurs dans le secteur midi que pour les flancs ou pour le côté nuit. Cette différence entre les quatre secteurs MLT est davantage observée avec le mode $g_3(\tau)$, en accord avec la distribution des poids (*cf* 2.10). Il faut également noter que la présence d'ACFs multi-composantes localisées aux pieds d'ancrage ionosphérique du cornet polaire ne sont pas suffisamment représentées, du moins dans notre sélection, pour avoir une certaine influence sur la forme des modes. Bien qu'il soit difficile de comptabiliser la nombre de spectres multi-composantes enregistrés, Ponomarenko and Waters (2006) ont montré que la contribution de ces échos radar, sur la journée du 10 décembre 1999, par rapport à l'ensemble des échos radar détecté

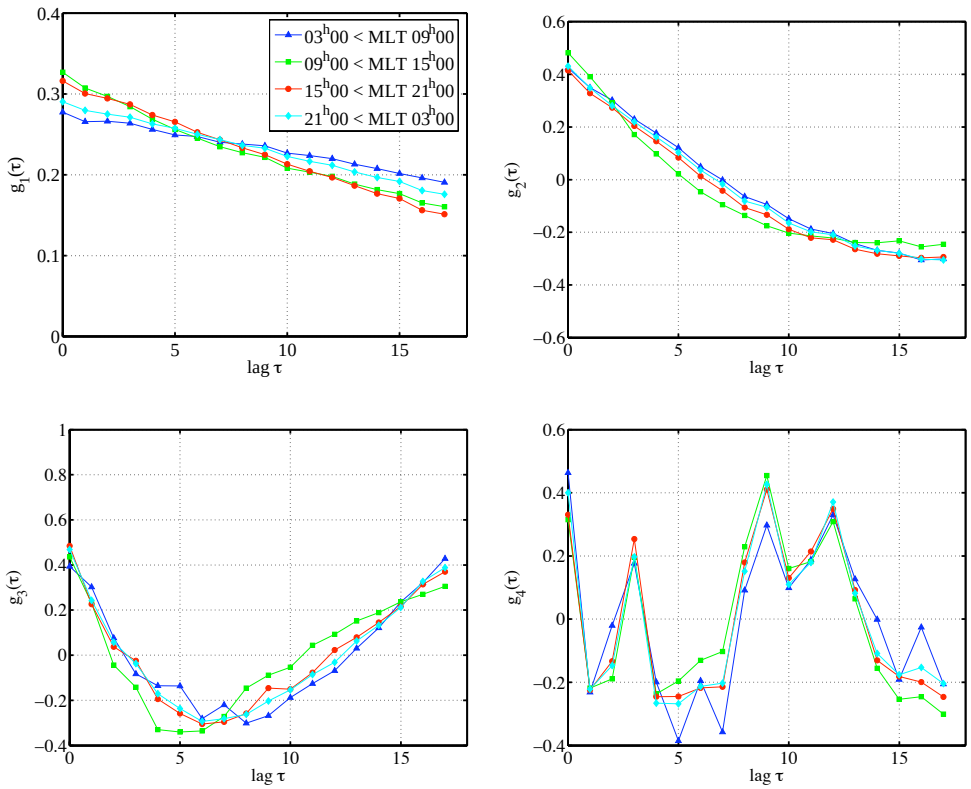


FIG. 2.12 – *Caractérisation de la décomposition en valeurs singulières des ACFs en fonction du secteur MLT. les quatre premiers modes $g_\alpha(\tau)$, avec $\alpha = [1, 2, 3, 4]$ sont présentés.*

par le radar TIGER représente seulement $\sim 6\,\%$.

Nous remarquons également que le quatrième mode, bien qu'il soit trop irrégulier pour être utilisé lors de la reconstruction des ACFs, conserve une certaine reproductibilité d'un échantillon à un autre. Ceci peut s'expliquer par la présence répétitive dans les ACFs de certains mauvais retards non détectés par l'algorithme FitACF (*cf* annexe A pour une description des différentes sources possibles d'erreur de mesure). Cet effet est également observé pour les modes d'ordre supérieur à 4.

L'investigation de la dépendance en secteur MLT de la décomposition des ACFs ne révèle pas de changement notable sur les modes $g_\alpha(\tau)$. A partir de ces quatre échantillons définis sur un intervalle de $06^{\mathrm{h}}00$ MLT, nous notons que les propriétés physiques des régions magnétosphériques pouvant entraîner une signature spécifique et différente de la rétrodiffusion ne pèsent pas suffisamment sur les échantillons pour affecter de manière significative le résultat de la décomposition.

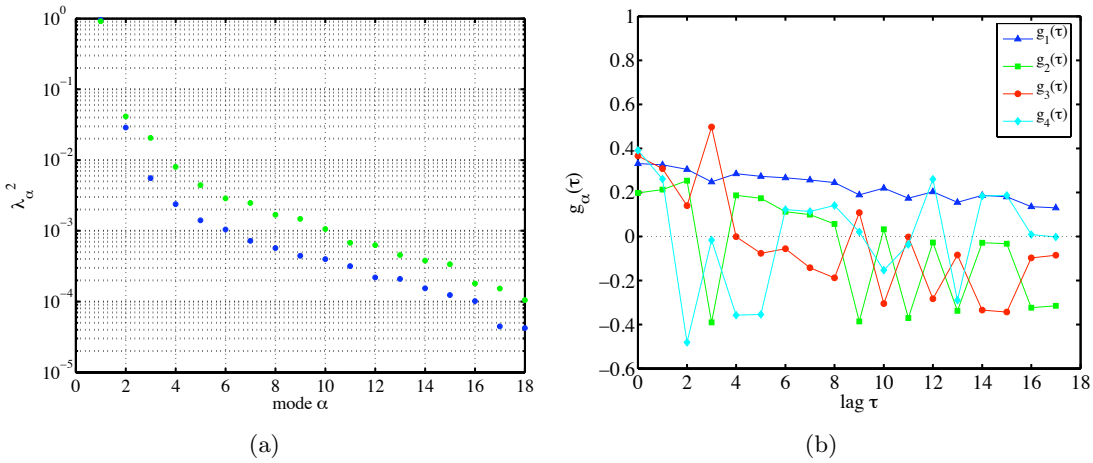


FIG. 2.13 – *Influence des mauvais retards sur la décomposition en valeurs singulières des ACFs : (a) distribution des énergies normalisée pour des ACFs avec (courbe verte) et sans retards erronés (courbe bleue). (b) Les quatre premiers modes $g_\alpha(\tau)$ qui sont associés aux spectres d'énergie indiqués en vert, sont également représentés.*

2.3.3 Parenthèse sur la représentativité des échos radar sélectionnés

Dans ce travail de thèse, nous avons défini des échantillons statistiques dans le but de définir une base la plus complète possible afin de reproduire le maximum de variabilité des échos radars observés. Pour cela, un certain nombre de critères ont été utilisés pour sélectionner les ACFs. Le choix de ces critères peut être critiqué puisqu'au final, nous avons sélectionné uniquement $\sim 20\,\%$ des ACFs observées pour chaque échantillon. On peut effectivement se demander si la base ainsi définie est suffisamment complète pour reproduire l'ensemble des ACFs observées, comme par exemple celles qui possèdent une allure modulée. Parmi ces critères, le plus restrictif est le nombre de mauvais retards (contenus dans chaque ACF), car il contribue à la hauteur de $\sim 90\,\%$ de la totalité des ACFs éliminées. Dans cette première étape, nous avons volontairement conservé les ACFs pour lesquelles aucun mauvais retard n'a été détecté par l'algorithme FitACF. Néanmoins, la détection systématique de ces valeurs erronées est une tâche délicate. Il peut arriver que certaine valeurs soient identifiées comme mauvaises (à cause de la présence de bruit ou d'interférence), même si elles véhiculent des informations sur l'origine micro-physique des échos radar (*cf* chapitre 4 et l'annexe A pour une discussion sur l'origine des différentes erreurs). L'élimination de toutes les ACFs dégradées par de mauvais retards suppose que l'on écarte des informations susceptibles d'affecter la représentativité du modèle SVD. Ce choix, comme le prouve la figure 2.13, est malgré tout nécessaire pour pouvoir définir un ensemble de fonctions empiriques de forme

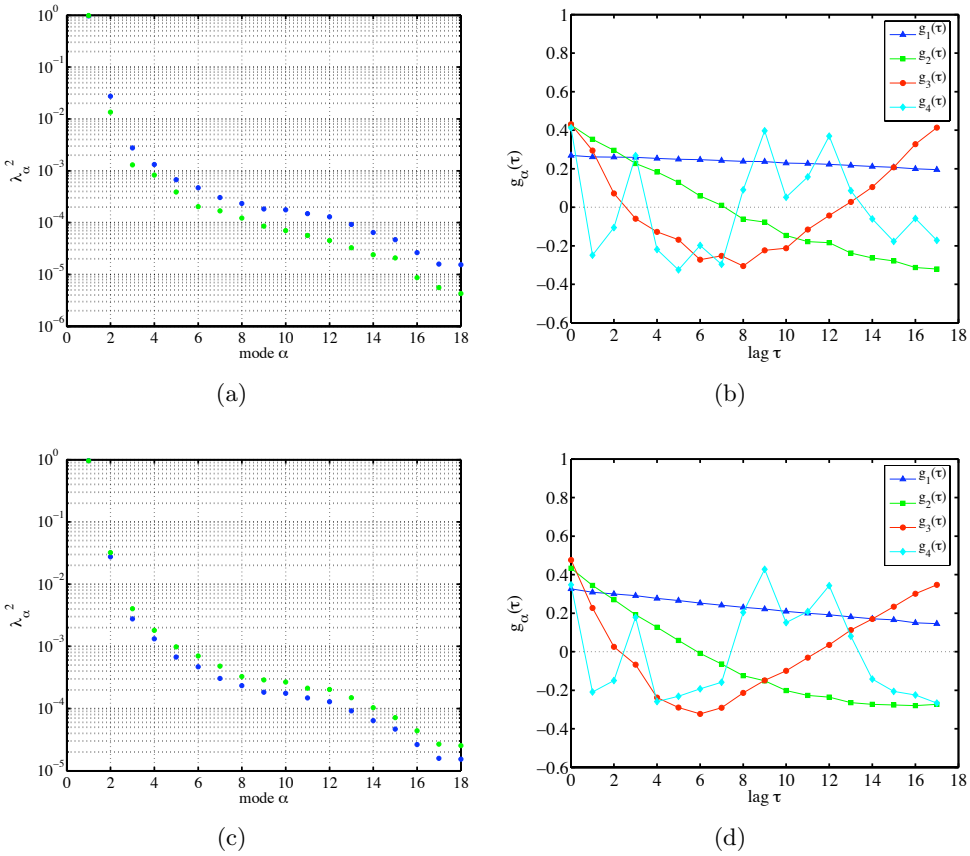


FIG. 2.14 – *Influence de l'orientation des champs de vue des radars sur la décomposition en valeurs singulières des ACFs.* ***En haut*** *: résultat obtenu pour des radars méridionaux.* ***En bas*** *: celui obtenu pour des radars zonaux. Pour chacun des deux cas, le spectre d'énergie normalisée est indiqué (point vert), ainsi que les quatre premiers modes $g_\alpha(\tau)$, avec $\alpha = [1, 2, 3, 4]$. Le spectre d'énergie indiqué en bleu montre le résultat de la décomposition lorsque les deux orientations sont regroupées dans un même échantillon.*

suffisamment régulière. Dans le cas où l'on conserve les ACFs avec mauvais retards, le nombre trop faible de modes significatifs (1 ou 2 selon la Fig. 2.13(a)), ainsi que leur forme respective (Fig. 2.13(b)), ne permettent pas de reproduire et de caractériser correctement les observations. Ceci nous oblige donc à omettre volontairement une certaine quantité d'information pour définir un ensemble de fonctions empiriques suffisamment complet pour caractériser un maximum de la variabilité des échos radar. Nous devons également faire remarquer que les seuils appliqués sur les erreurs quadratiques moyennes des modèles exponentiel et gaussien permettent (dans une certaine mesure) de conserver les ACFs dont l'allure s'éloigne de ces modèles et par conséquent d'élargir la sélection des échos radar.

Nous avons enfin mené un dernier test sur l'orientation du champ de vue des radars (zonal vs méridional). En effet, selon l'orientation du radar, la signature de la rétrodiffusion diffère quelque peu. Par exemple, l'interprétation géophysique de la frontière de la largeur spectrale (SWB) est différente. La SWB observée par les radars zonaux est plutôt reliée à la ligne neutre de renversement de la convection (CRB) tandis que les radars méridionaux observent plutôt la frontière entre les lignes de champ géomagnétique ouvertes et fermées (OCB) (Chisham and Freeman, 2004). Nous avons vérifié que de telles différences selon l'orientation des radars ne se répercutent pas sur la définition des modes statistiques, comme indiqué sur la figure 2.14.

2.4 Conclusion

Dans ce chapitre, nous avons mis en oeuvre une méthode de réduction des ACFs brutes observées par les radars SuperDARN dans l'ionosphère des hautes latitudes. Notre objectif était de définir un nouveau jeu de paramètres compact caractérisant l'ACF.

Pour cela, nous avons utilisé une méthode de décomposition en valeurs singulières comme un outil de réduction de données. A partir d'un échantillon représentatif des observations SuperDARN, nous avons montré que l'ensemble des ACFs observées peut être reconstruit à partir d'une combinaison linéaire de trois modes seulement. Avec ceux-ci, on reproduit en moyenne 99.2 % de l'énergie des ACFs sélectionnées.

La description est robuste et consistante, en ce sens que la variation de l'activité géomagnétique ou encore la localisation des échos radar selon le secteur MLT n'affecte pas la forme des trois modes significatifs. De plus, nous devons noter que la décomposition est indépendante de l'orientation des faisceaux radars.

Nous utiliserons donc dans ce qui suit la projection des ACFs sur ces trois modes pour définir un nouveau jeu de paramètres empiriques qui caractérise les différents types d'ACFs observées. Le chapitre suivant étudiera les modes statistiques sous forme de quantité et interprétera ces résultats d'un point de vue géophysique.

3

Interprétation de la réduction statistique des ACFs

3.1 Une interprétation physique des modes 68
3.1.1 Comparaison des paramètres $\{f_\alpha\}$ de la SVD avec les paramètres usuels issus de l'algorithme FitACF 69
3.1.1.1 Relation entre les paramètres $\{f_\alpha/f_1\}$ et la largeur spectrale 70
3.1.1.2 Relation entre les paramètres $\{f_\alpha/f_1\}$ et la forme des ACFs 73
3.1.2 Modélisation du comportement des paramètres SVD : $\{f_2/f_1, f_3/f_1\}$ 75
3.1.2.1 Choix d'un modèle analytique pour la simulation des ACFs 75
3.1.2.2 Méthode 76
3.1.2.3 Modélisation par mode SVD 78
3.2 Classification des échos radar : région E ou région F ? 81
3.3 Aspect géophysique de la réduction des ACFs 82
3.3.1 Comparaison des distributions polaires des paramètres issus de «FitACF» et de la SVD 84
3.3.2 Caractérisation de l'erreur des modèles 87
3.3.2.1 Distribution des écarts quadratiques relatifs des modèles 88
3.3.2.2 Comparaison des modèles 90
3.3.3 Interprétation de deux signatures particulières : le 12 février 2003 92
3.3.3.1 Evénement A 92
3.3.3.2 Evénement B 96
3.4 Détection de la frontière des lignes de champ ouvertes/fermées 99
3.4.1 Vers une Inférence Bayésienne de l'OCB 102
3.4.2 Etude de cas : 10 octobre 1999 103
3.4.2.1 Approximation de la vraisemblance 103
3.4.2.2 Projection de la frontière Bayésienne 106
3.4.3 Perspectives : aspect dynamique et amélioration 109
3.5 Conclusion 112

Nous avons montré dans le chapitre précédent que le module de toutes les ACFs observées avec les radars SuperDARN peut être reproduit par une combinaison linéaire de trois fonctions orthonormales de base ($g_\alpha(\tau)$) donnant ainsi accès à trois paramètres statistiques $\{f_1(t), f_2(t), f_3(t)\}$. La détermination de ces trois modes par une méthode de décomposition en valeurs singulières est robuste en ce sens que la forme des modes $g_\alpha(\tau)$ est reproductible : ils sont insensibles à l'activité géomagnétique (Kp) et au secteur MLT.

Avant de passer à l'étape suivante, *i.e.* celle de classification des régions géophysiques, une phase intermédiaire d'interprétation des paramètres statistiques est nécessaire. Cette phase d'interprétation est essentielle pour nous guider dans l'élaboration de nouvelles quantités appropriées à notre objectif.

Ce chapitre se décompose alors en deux parties. Une première aborde le problème de la signification physique des modes. La seconde, plus appliquée, est dédiée à la détection de la frontière entre les lignes de champ ouvertes et fermées.

3.1 Une interprétation physique des modes

L'approche exploratoire permet une description compacte de la mesure à partir de ses propriétés statistiques. L'inconvénient des méthodes de réduction telles que la SVD par rapport à des modèles physiques ou empiriques est que la description est sans lien direct avec la physique sous-jacente. En effet, un mode peut décrire plusieurs processus physiques et réciproquement, un processus physique peut être réparti sur plusieurs modes.

Une phase d'interprétation est donc primordiale afin d'extraire les quantités pertinentes. Pour cela, nous comparons tout d'abord les paramètres $\{f_1(t), f_2(t), f_3(t)\}$ avec les paramètres usuels extraits des échos radar. Puis, nous modélisons le comportement de ces coefficients avec des données simulées afin de définir un nouveau jeu de paramètres. Ce chapitre sera donc entièrement consacré à des comparaisons entre les paramètres obtenus par SVD et ceux issus du modèle FitACF, même si par la suite, nous utiliserons uniquement les premiers pour localiser les frontières.

3.1.1 Comparaison des paramètres $\{f_\alpha\}$ de la SVD avec les paramètres usuels issus de l'algorithme FitACF

Les modèles gaussien ($|R(\tau)| \propto e^{-\sigma^2\tau^2}$) et exponentiel ($|R(\tau)| \propto e^{-\lambda\tau}$) qui sont implémentés dans l'algorithme FitACF permettent d'extraire de l'ACF les paramètres physiques tels que la largeur spectrale et la puissance du signal rétrodiffusé. En comparant ces paramètres usuels avec les paramètres $\{f_1(t), f_2(t), f_3(t)\}$ de chacune des ACFs, nous allons explorer la physique qui est sous-jacente à chacun des trois modes. Il faut noter, par ailleurs, que l'utilisation de l'algorithme FitACF intervient ici uniquement pour la sélection des ACFs et la comparaison des modes avec les paramètres usuels. La projection des ACFs sur les trois modes ($g_\alpha(\tau)$, avec $\alpha = [1, 2, 3]$) est par ailleurs suffisamment robuste pour développer un schéma d'interpolation automatisé permettant une estimation fiable des $\{f_1(t), f_2(t), f_3(t)\}$. Cet aspect sera abordé dans le chapitre 4.

Nous cherchons donc à comparer les paramètres $\{f_1(t), f_2(t), f_3(t)\}$ aux paramètres physiques classiques, qui sont la largeur spectrale et la puissance du signal rétrodiffusé. Pour cela, nous utilisons l'échantillon d'ACFs constitué au § 2.2.1, à partir duquel nous appliquons à la fois l'algorithme FitACF et la méthode SVD. Pour chacune des mesures, nous estimons donc à la fois les paramètres $\{f_1(t), f_2(t), f_3(t)\}$, la largeur spectrale évaluée à partir des modèles gaussien (W_s) et exponentiel (W_l) ainsi que la puissance rétrodiffusée. Les facteurs physiques à l'origine de la signature de l'amortissement et de la forme du module des ACFs mesurées par SuperDARN ne sont pas encore entièrement compris. On les associe soit à la distribution des vitesses de dérives ionosphériques soit à la durée de vie des irrégularités électroniques (Hanuise et al., 1993; Villain et al., 1996; Ponomarenko and Waters, 2006). Le plus rapide des deux processus détermine le temps de décorrélation de l'ACF et par conséquent sa forme (exponentielle, gaussienne ou intermédiaire) et sa largeur spectrale. La largeur spectrale est donc liée, entre autres, au schéma de convection ionosphérique et aux particules énergétiques qui se précipitent dans l'ionosphère. La puissance du signal rétrodiffusé dépend essentiellement de la présence d'irrégularités électroniques (Ballatore et al., 2001; Danskin et al., 2002). Danskin et al. (2002) montrent notamment que la puissance des échos est plus élevée lorsque le champ électrique est plus intense ($\sim$ 10 mV/m). Or ces irrégularités se produisent sur des échelles de temps plus petites que le temps nécessaire au balayage radar. Les mesures intègrent donc un processus non-stationnaire, ce qui complique l'interprétation de l'ACF. Cette non-stationarité signifie qu'il est difficile d'interpréter une ACF en termess de processus physiques bien déterminés et qu'il faut davantage les considérer comme la résultante d'un mélange de processus.

La question est de savoir quelle information nous désirons extraire des modes statistiques. f_1 est une moyenne pondérée sur les dix-huit retards de l'ACF et peut donc être interprété comme une puissance moyenne de l'ACF. Cependant, la puissance rétrodiffusée n'est pas un bon candidat pour l'identification des régions ionosphériques : elle dépend de l'état de la cible rétrodiffusante mais également des perturbations subies par l'onde radio tout au long de son trajet (Fejer and Kelley, 1980; Hanuise, 1983; Greenwald et al., 1995). La caractérisation de la forme de l'ACF semble être plus intéressante pour une classification. Cette information devrait se situer dans les contributions de f_2 et de f_3 relatives à f_1. En effet, chaque mode, $g_\alpha(\tau)$, possède une décroissance selon τ qui est figée. C'est donc la contribution relative de ces trois modes qui permet d'épouser la décroissance de l'ACF. Par conséquent, nous allons nous concentrer sur les rapports adimensionnels f_2/f_1 et f_3/f_1. Dans la mesure où f_α n'est jamais nul à cause de la normalisation imposée par la méthode SVD, ces rapports sont bien définis et ceux quelles que soient les conditions. L'étude des f_α seuls est dénuée d'intérêt dans la mesure où ils sont fortement corrélés avec la puissance rétrodiffusée, qui varie sur plusieurs ordres de grandeur. Toute dépendance sera donc dominée par la puissance.

3.1.1.1 Relation entre les paramètres $\{f_\alpha/f_1\}$ et la largeur spectrale

Chaque mode possède une forme propre et décroît différemment avec les retards. C'est la contribution relative de chacun des trois modes qui permet de reproduire l'amortissement des ACFs. La contribution des trois modes à l'ajustement des ACFs est exprimée par l'intermédiaire des coefficients f_α. Les figures 3.1(a) et 3.1(b) montrent de manière convaincante une corrélation des rapports f_2/f_1 et f_3/f_1 avec la largeur spectrale (W) exprimée en Hertz. Pour réaliser ces figures, nous avons conservé la valeur obtenue à partir du modèle (gaussien ou exponentiel) qui ajuste le mieux les données. Le lien entre la largeur spectrale et le rapport f_2/f_1 est plus marqué que pour f_3/f_1 en ce sens que la figure 3.1(b) révèle une dispersion plus grande de cette corrélation. Nous exprimons la largeur spectrale en Hz et non en m.s^{-1} car la fréquence radar n'intervient pas dans le modèle statistique. Les valeurs supérieures à 35 Hz (soit W $\geqslant$ 525 m.s^{-1} pour une fréquence radar de 10 MHz) correspondent à des ACFs de forme très piquée et sont donc difficiles à caractériser. Seuls quelques premiers retards dépassent alors du niveau de bruit. Nous devons également noter que la dispersion de la distribution est exacerbée par l'échelle logarithmique. Une simulation a été réalisée afin d'étudier l'influence du bruit sur l'estimation de ces rapports. Pour cela nous avons utilisé le modèle d'Obhukov introduit par (Hanuise et al., 1993) et par Villain et al. (1996) pour décrire le caractère turbulent des irrégularités électroniques. Bien que son interprétation soit controversée (à cause des valeurs obtenues qui diffèrent de la théorie), ce modèle

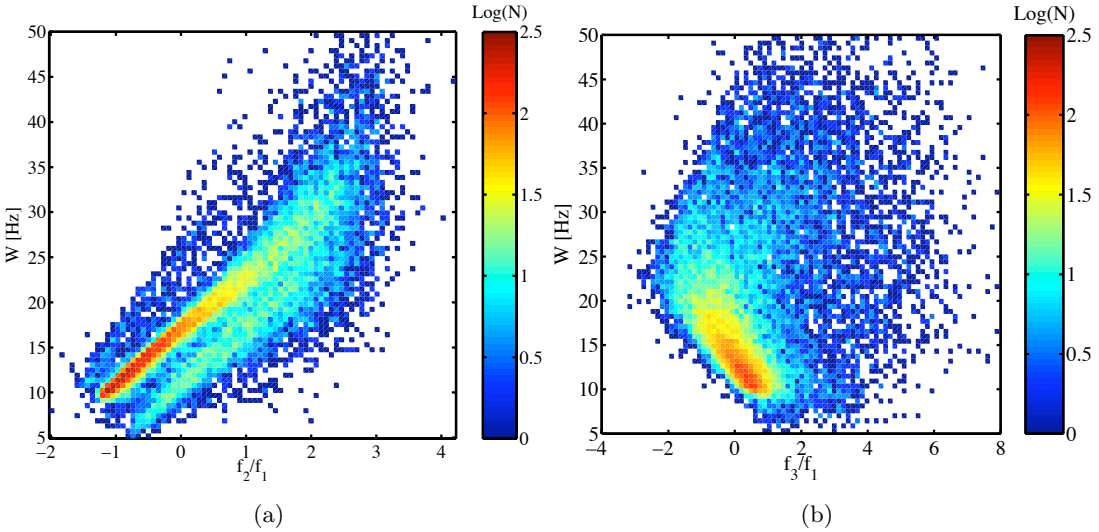


FIG. 3.1 – *Distribution des échos radar en fonction du rapport adimensionnel f_2/f_1 et de la largeur spectrale W (en Hz) (a) et en fonction du rapport f_3/f_1 et de la largeur spectrale W (en Hz) (b). Le nombre d'échos radar est représenté avec une échelle de couleur logarithmique. Cet histogramme a été construit à partir d'une grille de 80^2 intervalles équi-répartis.*

a l'avantage de reproduire l'ensemble des formes observées d'ACFs. Nous avons donc reconstruit à partir de ce modèle, un jeu d'ACFs synthétiques que nous avons volontairement dégradé en ajoutant un bruit blanc. L'ajustement de ces ACFs à partir des trois modes SVD, a permis de montrer que, pour un même niveau de bruit, l'estimation des paramètres est moins bonne pour celles qui sont davantage piquées.

Notons la présence sur la figure 3.1(a) de deux arêtes qui révèlent deux populations différentes. Une investigation montre qu'elles correspondent aux valeurs calculées à partir de l'algorithme FitACF en utilisant respectivement un modèle gaussien (arête supérieure) et exponentiel (arête inférieure). Ce genre de choix entraîne une différence systématique de 5 Hz, soit une largeur spectrale de 300 m.s^{-1} pour une fréquence radar de 10 MHz, ce qui est loin d'être négligeable. Ce décalage illustre clairement comment le modèle sous-jacent peut introduire un biais sur la caractérisation de l'ACF. Un des avantages majeurs de la SVD est de ne pas imposer d'hypothèses sur le modèle pouvant biaiser l'estimation des paramètres. Ces deux populations n'apparaissent pas sur la figure 3.1(b).

Comme l'information sur la forme des ACFs est principalement contenue dans les rapports f_2/f_1 et f_3/f_1, il est intéressant de représenter les différents types d'ACFs dans un espace dont les axes sont ces deux rapports. La figure 3.2 montre la valeur

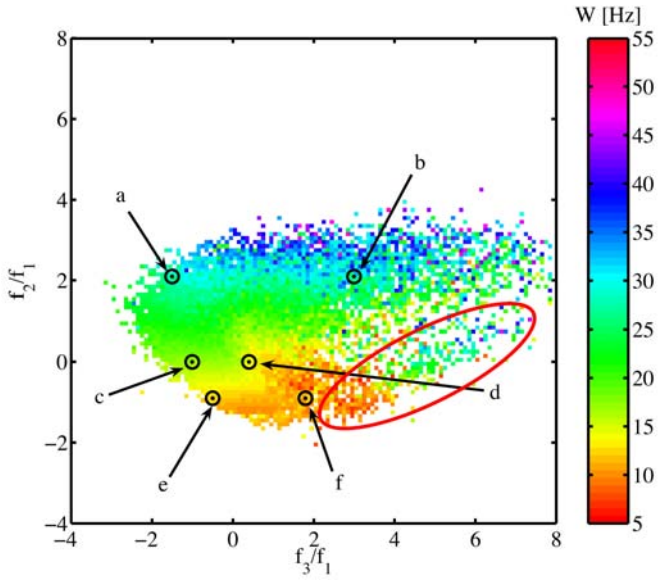


FIG. 3.2 – *Distribution de la largeur spectrale W en fonction des rapports adimensionnels f_2/f_1 et f_3/f_1. Cette distribution a été construite à partir de 120^2 intervalles équi-répartis. Chaque lettre définit une région correspondant à une forme spécifique d'ACF (voir Fig. 2.5).*

de la largeur spectrale en fonction de ces deux rapports. La valeur moyenne de la largeur spectrale, codée à l'aide d'une échelle de couleur, augmente graduellement avec le rapport f_2/f_1. Nous notons que son influence sur le rapport f_3/f_1 reste relativement faible. Ceci est normal, dans la mesure où le mode f_3 ne décrit qu'une faible variance des observations. Le rapport f_2/f_1 caractérise la largeur spectrale, tandis que le rapport f_3/f_1 apparaît comme une correction d'ordre supérieur, dont le lien avec la largeur spectrale est bien établi, quoique plus complexe que pour f_2/f_1 (*cf* les figures 3.1(a) et 3.1(b)). Chaque panneau de la figure 2.5 (*cf* Chap. 2) montre une forme typique d'ACF localisée dans le plan $\{f_2/f_1, f_3/f_1\}$ (indiqué par les lettres a à f dans la figure 3.2). Bien que le troisième mode possède un poids (λ_3) par rapport aux deux autres modes (λ_1 et λ_2) relativement plus faible, il est définitivement nécessaire pour ajuster correctement la décroissance de l'ACF. Comme nous le montrerons dans la section suivante, le rapport f_3/f_1 véhicule une information supplémentaire intéressante en donnant une indication sur l'allure (*i.e.* la nature exponentielle ou gaussienne de la transition) des ACFs, laquelle permet d'identifier les processus prédominants qui affectent la rétrodiffussion des ondes radio (Hanuise et al., 1993; Villain et al., 1996; Ponomarenko and Waters, 2006). La caractérisation de la forme des ACFs est délicate et dépend de la qualité de la mesure. En effet, la reconstruction des ACFs est souvent entachée d'erreurs qui affectent seulement certains retards (*cf* Annexe A pour une discussion des sources

potentielles d'erreurs). Il est alors nécessaire de les détecter et de les éliminer avant d'ajuster les modèles. Nous montrerons dans le chapitre 4 qu'au-delà de 10 mauvais retards, les paramètres f_1, f_2 et f_3 ne permettent plus de caractériser correctement les l'ACFs.

3.1.1.2 Relation entre les paramètres $\{f_\alpha/f_1\}$ et la forme des ACFs

Afin d'étudier l'influence de la forme des ACFs sur les rapports f_2/f_1 et f_3/f_1, nous évaluons les erreurs quadratiques relatives pour le modèle exponentiel et le modèle gaussien :

$$\langle \epsilon_{l,s} \rangle = \sqrt{\frac{1}{N-2} \sum_{i=1}^{N} \left(\frac{|R(\tau_i)| - |\widehat{R}_{l,s}(\tau_i)|}{|R(\tau_i)|} \right)^2} \tag{3.1}$$

où $|R(\tau_i)|$ est la puissance mesurée, $|\widehat{R}_{l,s}(\tau_i)|$ la puissance donnée par les modèles exponentiel (l) et gaussien (s) au retard τ_i. Le rapport de ces deux quantités (ϵ_l/ϵ_s) permet de caractériser la forme de l'ACF : $\epsilon_l/\epsilon_s > 1$ identifie une décroissance de type préférentiellement gaussien de l'ACF et $\epsilon_l/\epsilon_s < 1$ une décroissance de type plutôt exponentiel. Lorsque ce rapport est proche de $\epsilon_l/\epsilon_s \sim 1$, la forme est intermédiaire. Cela signifie qu'en termes de moindres carrés, les deux modèles sont équivalents. Il faut cependant noter qu'un rapport proche de 1 peut résulter à la fois d'un bon ajustement des deux modèles mais peut aussi provenir d'ACFs avec des valeurs si aberrantes qu'aucun des deux modèles ne convient.

L'influence de la forme des ACFs sur les rapports sans dimension f_2/f_1 et f_3/f_1 est illustrée avec la figure 3.3. Elle montre la projection du rapport des erreurs quadratiques moyennes (*cf* Eq. (3.1)) du modèle exponentiel sur le modèle gaussien dans le plan $\{f_2/_1, f_3/f_1\}$. Dans cette projection, le rapport ϵ_l/ϵ_s n'est pas aléatoirement distribué. Une région de couleur jaune est clairement identifiable pour laquelle les échos radar sont typiquement gaussiens. A droite de cette région, il existe une zone de transition de couleur orange qui caractérise des ACFs de forme intermédiaire ou mal définie par les deux modèles analytiques ($\epsilon_l/\epsilon_s \sim 1$). Puis une région de couleur noire localise des ACFs de forme typiquement exponentielle.

Cette figure illustre clairement le rôle de f_3/f_1 et de f_2/f_1 sur la description de la forme de la décroissance des ACFs. On voit très clairement que c'est le rapport f_3/f_1 qui contrôle la forme de l'ACF (gaussienne, exponentielle) car la variation se fait

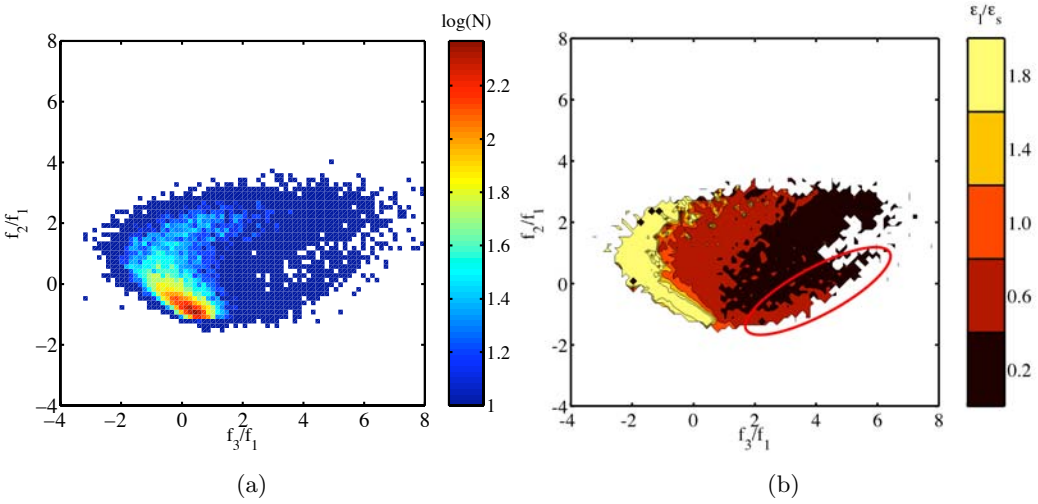


FIG. 3.3 – *Histogramme bi-varié des échos radar en fonction des rapports adimensionnels f_2/f_1 et f_3/f_1 (a). Cet histogramme a été reconstruit à partir d'une grille de 80^2 intervalles équi-répartis et l'échelle de couleur logarithmique indique le nombre d'échos radar. Iso-contours du rapport ϵ_l/ϵ_s des erreurs quadratiques moyennes des modèles exponentiel (ϵ_l) et gaussien (ϵ_s) projetés dans le même plan (b).*

principalement selon cet axe. Cette représentation met également en avant l'influence de la largeur spectrale sur la forme du spectre de rétrodiffusion. En effet, selon la valeur de la largeur spectrale et donc du rapport f_2/f_1, les formes observées du spectre de rétrodiffusion diffèrent. Par exemple, pour un rapport de $f_2/f_1 \lesssim -1$, le modèle exponentiel ne semble pas décrire la forme des ACFs et ce pour tout l'intervalle de définition de f_3/f_1. Tandis que pour des ACFs plus étroites, *i.e.* pour f_2/f_1 plus grand, toutes les formes semblent être représentées. A la différence du lien évident entre la largeur spectrale et le rapport f_2/f_1, la variabilité de la forme de la décroissance des ACFs est plus complexe à décrire et nécessite l'utilisation des deux quantités f_2/f_1 et f_3/f_1.

A partir de cette représentation, on observe notamment que 32 % des ACFs suivent un modèle gaussien ($\epsilon_l/\epsilon_s \geqslant 1.05$), 28 % un modèle exponentiel ($\epsilon_l/\epsilon_s \leqslant 0.95$). En accord avec les résultats de Moorcroft (2004), 40 % des ACFs sont de forme intermédiaire. Selon Villain et al. (1987), et même si les auteurs n'ont pas précisé les critères qui ont été employés pour faire cette distinction, ont estimé à 20 % la proportion d'ACFs qui est mieux représentée par un amortissement de type gaussien et à 70 % celle qui est mieux représentée par une exponentielle décroissante. Même si le rapport ϵ_l/ϵ_s ne donne qu'une première approximation sur la forme des ACFs, ce résultat indique que le développement du programme FitACF ne permet pas de caractériser correctement

près de 40 % des observations radars.

Notons la présence sur les figures 3.2 et 3.3 d'une population (marquée par une ellipse rouge) qui se détache du reste de l'échantillon et pour laquelle la largeur spectrale et le rapport ϵ_l/ϵ_s se comportent différemment. Une étude approfondie a permis de mettre en évidence qu'une quantité importante de ces échos radar possèdent des propriétés proches de ceux issus de la région E ou encore de ceux issus de la rétrodiffusion au sol. On trouve également des ACFs modulées ou encore des ACFs qui possèdent une décroissance très rapide (avec uniquement 1 ou 2 retards au dessus du niveau de bruit). Ce résultat montre bien en autres toute l'ambiguïté sur la localisation en altitude des échos radar puisque le critère unique, qui consiste à utiliser la distance comme séparateur des échos radar, ne permet pas de lever complètement l'incertitude. Nous reviendrons plus en détail sur ce problème de classification à la section § 3.2.

La réduction statistique des ACFs à partir de la méthode SVD a permis d'isoler les caractéristiques du signal rétrodiffusé à partir de deux paramètres statistiques. Le rapport f_2/f_1 offre une description unique de la largeur spectrale et donc sur l'amortissement. Le rapport f_3/f_1 couplé à f_2/f_1 donne une information sur la forme de la décroissance. Ils sont par conséquent reliés aux processus micro-physiques (prédominant) qui génèrent l'écho radar.

3.1.2 Modélisation du comportement des paramètres SVD : $\{f_2/f_1, f_3/f_1\}$

Afin de compléter les résultats du paragraphe précédent, nous désirons modéliser le comportement des modes statistiques afin de poursuivre l'interprétation physique des rapports f_2/f_1 et f_3/f_1. Pour cela, il faut utiliser un modèle analytique capable de reproduire l'ensemble de la variabilité des ACFs observées, dans le but d'en déduire un nouveau jeu de paramètres empiriques.

3.1.2.1 Choix d'un modèle analytique pour la simulation des ACFs

Plusieurs auteurs (Hanuise et al., 1993; Villain et al., 1996; Ponomarenko and Waters, 2006) se sont basés sur la théorie de la diffusion collective des ondes HF pour examiner plus en détail la forme des ACFs observées afin de caractériser les processus micro-physiques liés à la mesure. A partir d'un modèle unique, le profil d'Obukhov :

$$|R(\tau)| = R_o e^{-\mathbf{k}^2 D T_L \left(\frac{\tau}{T_L} + e^{-\left(\frac{\tau}{T_L}\right)} - 1 \right)} , \tag{3.2}$$

où $|\mathbf{k}|$ est la norme du vecteur d'onde d'analyse (défini par $\mathbf{k} = \mathbf{k_d} - \mathbf{k_i}$, où $\mathbf{k_i}$ est le vecteur de l'onde incidente et $\mathbf{k_d}$ celui de l'onde diffusée), D le coefficient de diffusion ambipolaire et T_L le temps de corrélation des irrégularités. Ce modèle permet la description de la décroissance des ACFs, depuis une forme purement gaussienne à une forme purement exponentielle. Il relie donc le spectre de rétrodiffusion au rapport entre la taille des échelles spatiales caractéristiques de la turbulence et la longueur d'onde des fluctuations de densité détectées par le radar ($\lambda_{irr} = \lambda_o/2$). Villain et al. (1996) argumentent notamment que les ACFs intermédiaires sont assimilables à des échelles de turbulence du même ordre de grandeur que la taille caractéristique des irrégularités de plasma. Bien que ce modèle fournit les paramètres de la turbulence de l'ionosphère, leur interprétation reste délicate et controversée (Ponomarenko and Waters, 2006). De plus, d'un point de vue numérique, l'estimation simultanée des paramètres de ce modèle peut rapidement conduire à des problèmes mal conditionnés. Nous lui préférons donc un modèle paramétrique plus fonctionnel, celui utilisé par Jackel (2000) et par Moorcroft (2004) pour étudier la forme des spectres auroraux observés par les radars incohérents UHF de Millstone Hill (440 MHz) et de EISCAT et COSCAT (933 MHz) :

$$|R(\tau)| = R_o e^{-\left(\tau/\tau_e\right)^{n_\tau}} , \tag{3.3}$$

où R_o est la puissance au retard zéro, le paramètre τ_e fixe la largeur de l'ACF (ou du spectre) et n_τ est un coefficient adimensionnel qui contrôle le degré de gaussiannité de l'ACF ; lorsque $n_\tau = 1$, l'ACF suit une décroissance exponentielle et quand $n_\tau = 2$, l'ACF est gaussienne.

Le choix du modèle (3.3) est motivé par : (1) les ACFs observées n'appartiennent que rarement à l'une des deux classes (exponentielle ou gaussienne) suggérées par la théorie et se trouvent fréquemment entre les deux, (2) ce modèle est moins sensible aux fluctuations que le modèle (3.2). Chacun des trois paramètres $\{R_o, \tau_e, n_\tau\}$ permet de caractériser des propriétés différentes de l'ACF.

3.1.2.2 Méthode

A partir de l'échantillon choisi pour identifier les modes principaux de la SVD (§ 2.2.1), nous avons ajusté le modèle (3.3) à l'ensemble des ACFs afin de définir la distribution des paramètres $\{R_o, \tau_e, n_\tau\}$. Le paramètre τ_e varie typiquement entre 2 et 200 s^{-1} avec une valeur médiane de $\sim 15\ \mathrm{s}^{-1}$; n_τ est quant à lui borné entre 0.5 et 2.5. La distribution de n_τ est très étroite et centrée sur 1.5. La majorité des ACFs possède

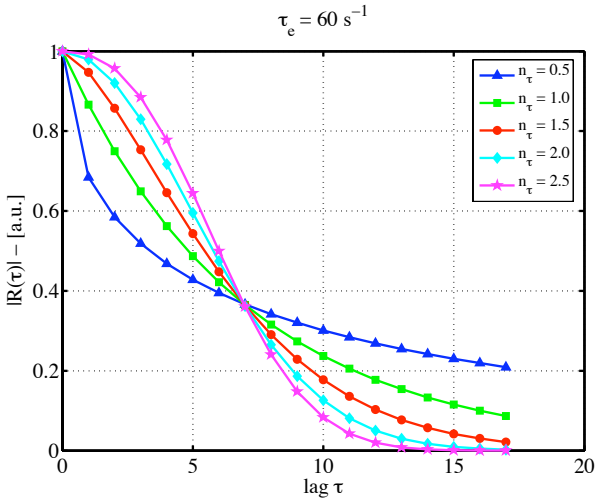


FIG. 3.4 – *Exemples d'ACFs reproduites à partir de l'équation* (3.3)*, pour* $\tau_e = 60$ s^{-1} *et* $n_\tau = [0.5, 1.0, 1.5, 2, 2.5]$.

donc une forme intermédiaire entre exponentielle ($n_\tau = 1$) et gaussienne ($n_\tau = 2$), ce qui confirme le résultat de Moorcroft (2004). A partir du domaine de variation ainsi défini, nous avons reconstruit des ACFs synthétiques en faisant varier indépendamment les paramètres de décroissance (τ_e) et de forme (n_τ) pour ensuite estimer les paramètres f_2/f_1 et f_3/f_1. Notons que nous ne prenons pas en compte la puissance au retard 0 comme paramètre libre car nous nous intéressons uniquement à l'allure des ACFs et à leur influence sur les modes. Les ACFs reproduites sont toutes normalisées à $|R(0)| = 1$.

La figure 3.4 donne quelques exemples d'ACFs modélisées à partir de l'équation (3.3), avec un taux de décroissance unique $\tau_e = 60$ s^{-1} et pour des valeurs du facteur de forme n_τ allant de 0.5 à 2.5. Pour un taux de décroissance identique, ces exemples mettent en relief l'influence du facteur de forme sur l'ensemble des retards de l'ACF. On voit notamment comment la transition de la forme gaussienne à une exponentielle décroissante relève la queue de l'ACF. La queue que l'on observe dans les mesures SuperDARN a une origine différente. Elle est causée par la présence d'un bruit de fond dans les données. Celui-ci peut perturber la reconstruction de l'ACF par les modèles. Notons également l'apparition d'une discontinuité dans la dérivée $d|R|/d\tau$ à l'origine dans le cas $n_\tau < 1$. Il n'y a pas de lien entre la discontinuité et les ACFs à décorrélation rapide. Ce lien est uniquement une conséquence de ce modèle. On pourrait très bien avoir $|R(\tau)| = R_o e^{-(\tau/\tau_e)^{n_\tau}} + Cste$ pour prendre en compte le niveau de bruit. Cette hypothèse suggère que le bruit affecte l'ensemble des retard avec la même amplitude, ce qui est difficilement vérifiable. Or comme l'ont montré Ponomarenko and Waters (2006),

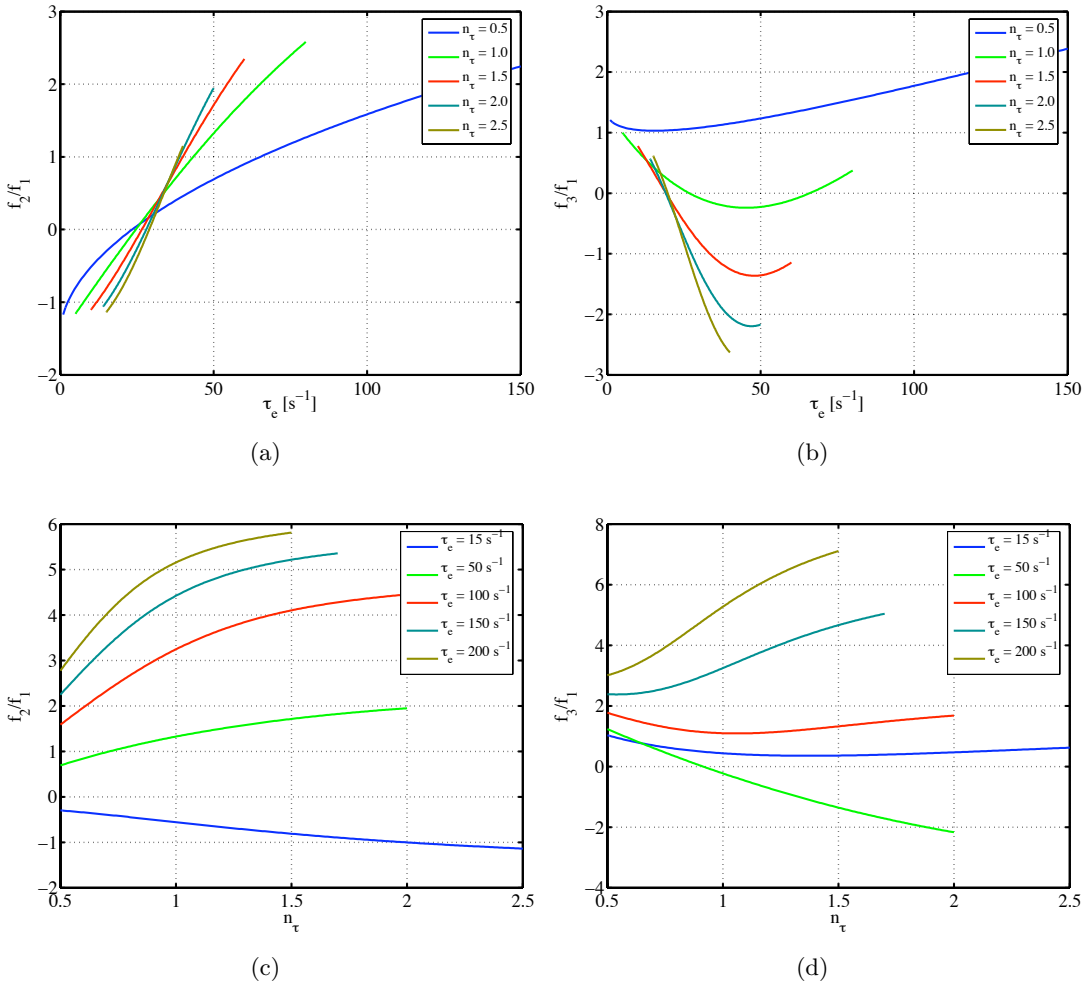


FIG. 3.5 – *Variation des paramètres statistiques f_2/f_1 et f_3/f_1 en fonction des paramètres τ_e et n_τ.*

la soustraction d'un bruit constant comme c'est le cas dans le programme FitACF, a pour effet de sur-évaluer systématiquement l'estimation de la largeur spectrale. Ce dernier point montre l'influence du bruit sur la mesure des grands retards et donc l'ajustement par les modèles des ACFs, puisque la présence d'une longue queue biaise l'estimation de la largeur spectrale et la caractérisation de l'allure des ACFs.

3.1.2.3 Modélisation par mode SVD

Pour mieux comprendre la signification des rapports adimensionnels f_2/f_1 et f_3/f_1, nous étudions maintenant leur dépendance envers les paramètres $\{\tau_e, n_\tau\}$ du modèle

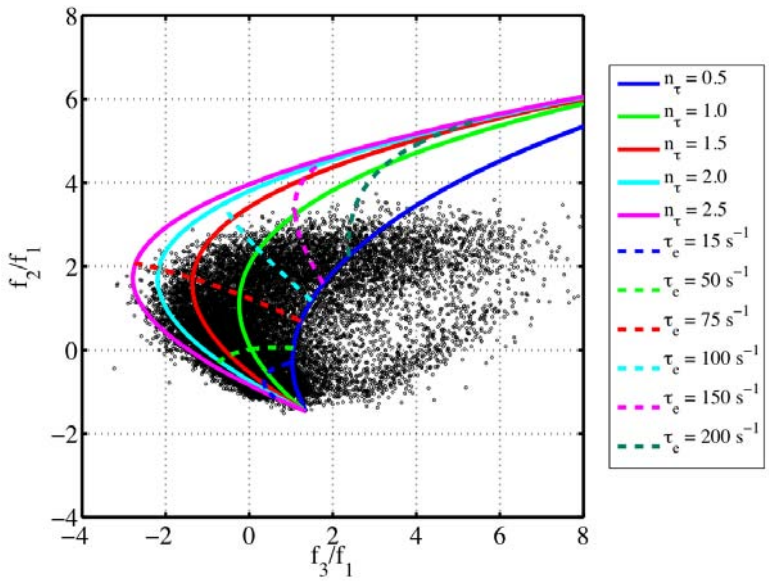


FIG. 3.6 – *Projection des courbes iso-formes (trait plein) et iso-largeurs (tiret) projetées dans le plan* $\{f_2/f_1, f_3/f_1\}$. *Le nuage de points représente la distribution des ACFs utilisées pour reconstruire les figures 3.2 et 3.3.*

analytique de l'équation (3.3).

Les quatre panneaux de la figure 3.5 montrent l'influence de ces deux paramètres sur les rapports adimensionnels f_2/f_1 et f_3/f_1. Les panneaux 3.5(a) et 3.5(b) présentent l'influence du facteur de forme n_τ sur les modes en fonction du taux de décroissance. Le rapport f_2/f_1 augmente avec τ_e de manière monotone ce qui confirme notre interprétation du rapport f_2/f_1 comme une signature de la largeur spectrale. A l'exception du cas $n_\tau < 1$, la variation n'est que très peu affectée par la forme de l'ACF, ce qui est particulièrement intéressant. L'évolution de f_3/f_1 dépend davantage du facteur de forme n_τ (*cf* Fig. 3.5(b)). Pour une valeur constante de $\tau_e = 40$ s^{-1}, le rapport f_3/f_1 varie de -2 à 1.1, tandis que pour $\tau_e = 20$ s^{-1}, ce rapport varie seulement entre 0 et 1. Les figures 3.5(c) et 3.5(d) représentent respectivement les iso-largeurs, τ_e, en fonction de f_2/f_1 et de n_τ et en fonction de f_3/f_1 et de n_τ. Le rapport f_2/f_1 augmente lorsque l'ACF se rapproche d'une forme gaussienne (*cf* 3.5(c)). L'interprétation de f_3/f_1 est plus délicate puisque le signe de la pente dépend du taux de décroissance. De manière générale, les ACFs très piquées ($n_\tau < 1$) influencent différemment les rapports f_2/f_1 et f_3/f_1 (panneaux (a), (b) et (d) de la figure 3.5). Leur comportement se détache par rapport aux formes d'ACF plus régulières. Ce cas correspond à des ACFs difficilement ajustables que nous éliminons de notre étude. On peut déduire de f_2/f_1 et de f_3/f_1 d'autres paramètres, de manière empirique, afin de mieux caractériser la largeur

spectrale et la forme des ACFs.

La figure 3.6 montre la projection des courbes iso-formes et iso-largeurs dans le plan $\{f_2/f_1, f_3/f_1\}$. La superposition des courbes iso- sur la projection des ACFs montrent clairement la dépendance des propriétés des ACFs en termes de décroissance et de forme dans cet espace. Le rapport f_2/f_1 décrit principalement la largeur de l'ACF. On note également que les deux rapports f_2/f_1 et f_3/f_1 décrivent la forme. Ceci veut dire que la forme et la largeur des ACFs sont non-linéairement corrélés puisque la SVD ne permet pas de découpler ces deux quantités. Pour cette échantillon, nous avons regroupé des ACFs qui sont à la fois bruitées et définies avec la même résolution temporelle (*i.e.* de retard élémentaire identique). Il est donc difficile de caractériser, selon la largeur spectrale mesurée, la forme réelle du spectre de rétrodiffusion. La séparation entre les différentes formes de spectres observés est ainsi maximisée pour des valeurs de f_2/f_1 comprises entre -0.5 et 2.5, correspondant à $\tau_e \in [15, 200]$ s^{-1}. Pour des valeurs de f_2/f_1 supérieures à 2.5, les ACFs deviennent trop piquées, *i.e.* que la résolution temporelle n'est pas suffisante pour que l'on puisse distinguer les formes. Remarquons qu'un grand nombre d'échos radar qui se détachent de l'essaim principal ne sont pas correctement décrits par ce modèle, pourtant très flexible. Cette population qui est identique à celle indiquée sur les figures 3.2 et 3.3, concerne des ACFs difficilement reproductibles par les deux modèles analytiques gaussien et exponentiel puisque $\epsilon_l/\epsilon_s \sim 1$, et qui plus est ne sont pas décrits par le modèle de l'équation (3.3) puisque les courbes iso- ne recouvrent pas cette zone de l'espace des paramètres.

Ce petit groupe d'ACFs qui se détache du reste la population semble être composé en partie par des ACFs extrêmement piquées. On y trouve également quelques ACFs modulées (spectres multi-composantes). Ces ACFs possèdent donc des formes difficiles à ajuster que ce soit pour le modèle analytique de l'équation (3.3) ou par les trois modes issus de l'analyse par SVD. Elles sont notamment observées au voisinage de l'empreinte ionosphérique du cornet polaire (Baker et al., 1995; Moen et al., 2001).

Notons le bon accord entre le facteur de forme, n_τ et l'information véhiculée par le rapport des erreurs du modèle exponentiel sur le modèle gaussien. En rapprochant les figures 3.3 et 3.6, on remarque que les iso-contours formés par le rapport ϵ_l/ϵ_s épousent la forme des courbes iso-formes, $n_\tau = Cste$.

La comparaison des trois modes déduits de la réduction statistique des ACFs avec les paramètres usuels a permis d'isoler l'information capturée par chacun d'eux en termes de quantité physique. Le rapport f_2/f_1 apparaît comme un bon estimateur de la largeur spectrale puisqu'il n'est que faiblement affecté par l'allure des ACFs. Tandis que f_3/f_1 est principalement sensible à l'allure piquée ou non. Néanmoins sa

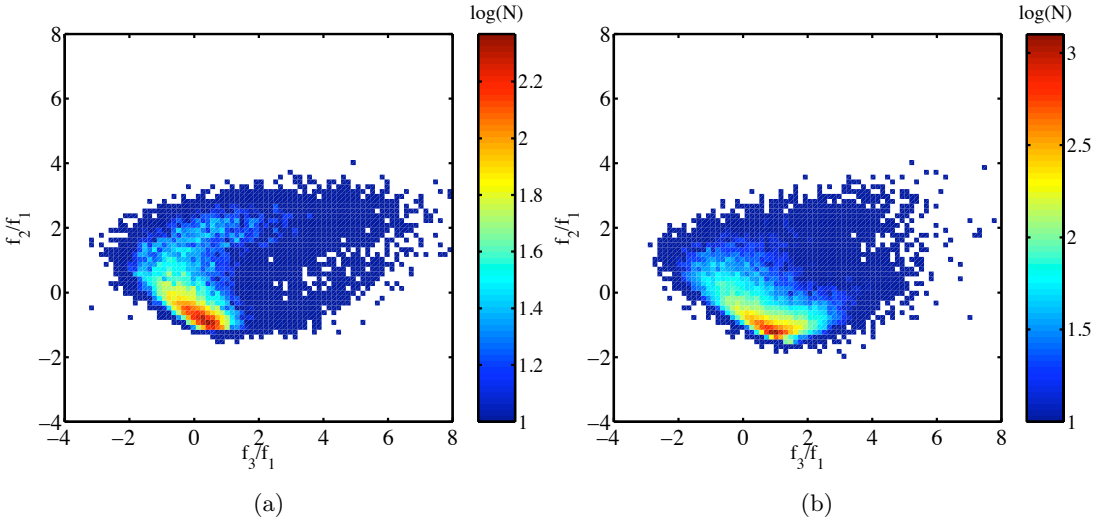


FIG. 3.7 – *Comparaison des histogrammes bi-variés des échos radar issus de la région F (a) et de la région E (b), en fonction des rapports adimensionnels f_2/f_1 et f_3/f_1. Ces histogrammes ont été reconstruits à partir d'une grille de 80^2 intervalles équi-répartis et l'échelle de couleur logarithmique indique le nombre d'échos radar.*

dépendance vis à vis de la forme des spectres est loin d'être linéaire et dépend également de f_2/f_1. Le rapport f_3/f_1 seul livre également une information additionnelle sur la qualité de la mesure. En première approximation, on peut l'utiliser couplé au rapport signal/bruit pour quantifier la qualité de la mesure, *i.e.* pour déterminer si la résolution temporelle utilisée est adaptée ou non par rapport au temps de décorrélation ou si le signal rétrodiffusé est trop bruité. En considérant la dépendance des modes SVD avec les paramètres τ_e et n_τ, il est possible d'en déduire un nouveau jeu de paramètres empiriques, décrivant uniquement la largeur spectrale et la forme des ACFs.

3.2 Classification des échos radar : région E ou région F ?

Nous avons remarqué au § 3.1.1, et ce malgré les précautions prises pour définir un échantillon représentatif des observations radar issues de la région F, qu'il reste néanmoins une faible quantité de données dont l'origine est ambiguë (*cf* figures 3.2 et 3.3).

Une étude détaillée des mesures révèle en effet que ces échos radar possèdent une vitesse de dérive en moyenne inférieure à 100 m.s^{-1} et un rapport signal-bruit qui varie entre entre 4 à 60 dB. Ces caractéristiques sont typiques d'échos radar provenant

soit de la région E, soit issus de la rétrodiffusion par le sol (après réflection par les couches ionosphériques). En projetant uniquement les échos de région E (*i.e.* associés à des distances $\leqslant$ 900 km, *cf* § 1.3.2.2), nous montrons que ces derniers (par rapport aux échos dits de région F) sont ordonnés différemment dans l'espace des paramètres $\{f_2/f_1, f_3/f_1\}$. La comparaison entre les figures 3.7(a) et 3.7(b) indique un décalage entre les deux maximums de l'ordre ~ 0.5 selon f_2/f_1 et selon f_3/f_1. Ceci montre que les critères utilisés pour l'identification des échos radar de région E (basés uniquement sur la distance estimée au radar) ou encore des échos de sol (à partir de la fonction empirique $G(V_D, W)$) sont loin d'être suffisants. En effet, les échos radar de région E peuvent apparaître à de grandes distances, par propagation des rayons après réflection au sol. Et inversement, des échos proches possédant une faible vitesse de dérive et une petite largeur spectrale peuvent provenir de la région F. Quant aux échos de sol, ils peuvent être caractérisés par de grandes largeurs spectrales et par de fortes dérives Doppler (supérieures aux limites fixées, *cf* Eq. (2.16)). Notons qu'il est extrêmement difficile de connaître la proportion d'échos radar provenant de la région E ou du sol. Pour cela, il faudrait étudier le chemin de propagation de chacun des échos radar qui constituent cette région de l'espace des phases en prenant en compte les conditions géophysiques de l'ionosphère (*e.g.* profil vertical de densité électronique), ce qui est difficilement envisageable. Une autre solution consisterait à utiliser le réseau d'interférométrie des radars SuperDARN, ce qui permettrait de déterminer l'angle d'élévation des échos radar. Nous pourrions approcher l'altitude des irrégularités cibles moyennant une hypothèse sur le chemin de propagation.

La figure 3.3(b) montre que la SVD offre des perspectives intéressantes dans la classification rapide des échos radar. Une analyse systématique des échos radar provenant des régions E, F et même des échos provenant du sol, pourrait aboutir sur des nouveaux critères améliorant l'identification des différentes sources d'échos radar. Par exemple, la construction des cartes de convection pourrait dès lors être améliorée puisque la vitesse Doppler des irrégularités provenant de la région E sature à la vitesse iono-acoustique, qui peut être inférieure à celle de la dérive plasmique (Villain et al., 1987).

3.3 Aspect géophysique de la réduction des ACFs

La question qui se pose maintenant est celle de l'apport des paramètres $\{f_2/f_1, f_3/f_1\}$ par rapport à la largeur spectrale estimée par FitACF, dans la description des régions ionosphériques. Pour cela, nous étudions les propriétés spatiales des paramètres issus du modèle SVD afin de les confronter à ceux donnés par FitACF.

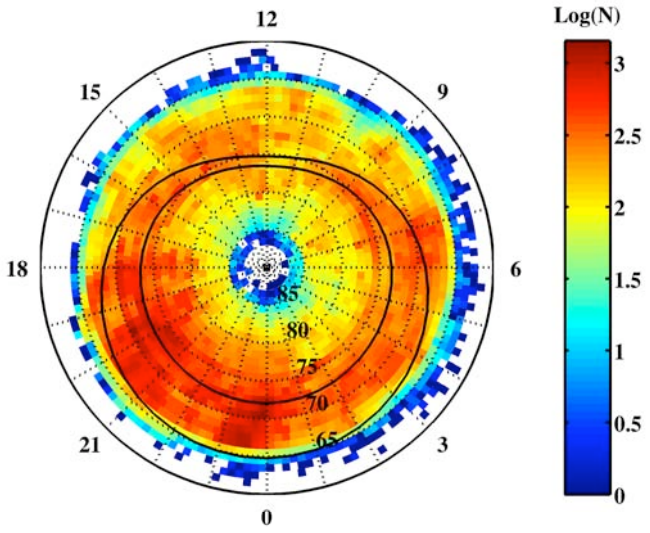


FIG. 3.8 – *Nombre d'échos radar projetés sur une grille MLT/Λ. Le nombre d'échos radar est codé avec une échelle de couleur logarithmique. L'ovale auroral superposé est défini pour une activité géomagnétique modérée,* $3 < \mathrm{Kp} < 4^-$ *(Holzworth and Meng, 1975).*

Notre comparaison est basée sur les quatre journées étudiées au § 2.2. Nous avons regroupé toutes les ACFs de l'hémisphère nord répondant aux critères de sélection définis au paragraphe § 2.2.1, à une différence près. Le nombre maximum de mauvais retards autorisé dans chaque ACFs a été étendu à 10 au lieu de 0 initialement. En effet, au paragraphe § 2.2.2, il s'agissait de définir les modes ($g_\alpha(\tau)$). Maintenant, nous utilisons ces modes pour caractériser l'ensemble des ACFs observées. L'algorithme FitACF fournit un diagnostic des mauvais retards. Il est dès lors aisé d'ajuster les ACFs réduites uniquement aux «bons» retards. Nous verrons dans le chapitre 4 comment utiliser la SVD pour détecter et éliminer les mauvais retards et nous montrerons que la reconstruction d'une ACF n'est plus possible si celle-ci comprend plus de 10 mauvais retards. Nous devons également ajouter que l'augmentation du nombre maximum de mauvais retards par ACF permet d'améliorer la couverture spatiale de l'ionosphère puisque nous augmentons le nombre de points. Comme le suggère la figure 3.8, l'augmentation de la taille de l'échantillon, de 21 972 à 479 336 ACFs, améliore la description des régions ionosphériques peu sondées. La répartition des échos radar est plus homogène que celle présentée à la figure 2.2. On note que la probabilité d'observer la rétrodiffusion côté nuit est plus grande que côté jour (Ruohoniemi and Greenwald, 1997; Villain et al., 2002). Un maximum d'échos radar se produit le long des frontières de l'ovale auroral côté nuit, modélisé ici pour un Kp modéré ($3 < \mathrm{Kp} < 4^-$).

Pour les sous-sections suivantes, nous ajustons les ACFs sélectionnées avec les trois

modèles (exponentiel, gaussien et SVD) et nous ne conservons que les mesures dont l'écart quadratique relatif pour l'un des trois modèles est inférieur à 0.8 (*cf* éq. (3.4) et (3.5)). Ainsi, nous obtenons trois jeux de paramètres différents comprenant chacun 479 336 événements.

3.3.1 Comparaison des distributions polaires des paramètres issus de «FitACF» et de la SVD

Nous confrontons ici la projection spatiale des paramètres issus de l'algorithme FitACF et du modèle SVD. Pour cela, nous projetons ces paramètres sur une grille MLT/Λ en fonction de leur position (Ruohoniemi and Baker, 1998). Pour faciliter la visualisation, les valeurs ont été moyennées au sein de chaque cellule. Les figures 3.9(a), 3.9(b) et 3.9(c) montrent respectivement la distribution moyenne du paramètre σ du modèle gaussien $\left(R \propto e^{-\sigma^2\tau^2}\right)$, du paramètre λ issu du modèle exponentiel $\left(R \propto e^{-\lambda\tau}\right)$ à partir desquels la largeur spectrale est estimée et de l'estimateur empirique de la largeur spectrale issu du modèle SVD, f_2/f_1. Nous ne projetons pas ici la largeur spectrale exprimée en m.s^{-1} qui est déduite des modèles gaussien et exponentiel afin de faciliter la comparaison entre les trois modèles. Par abus de langage, nous utiliserons dans la suite le terme largeur spectrale pour désigner les paramètres σ, λ et f_2/f_1. La figure 3.9(d) présente la projection moyenne du paramètre empirique f_3/f_1. Seules les cellules contenant au moins dix échos radar sont indiquées par une couleur.

La distribution de la largeur spectrale estimée à partir des trois modèles révèle des structures identiques. Ces structures sont nettes, quel que soit le modèle employé. Une bande correspondant à des spectres larges est observée tout le long de la frontière polaire de l'ovale auroral. Le bord équatorial de cette bande forme la frontière de la largeur spectrale qui est généralement associée à la frontière des lignes de champ magnétique ouvertes/fermées (Lester et al., 2001). Une investigation sur chacune des cellules révèle que le coefficient de corrélation moyen entre σ et f_2/f_1 est de $\rho_{\{\sigma,f_2/f_1\}} = 0.90 \pm 0.06$ et celui entre λ et f_2/f_1 est $\rho_{\{\lambda,f_2/f_1\}} = 0.94 \pm 0.05$, ce qui est remarquable. La différence réside plutôt au niveau de la valeur de la largeur spectrale donnée par chacun des modèles, comme c'est le cas avec les régions définies par des spectres larges (de couleur vert - jaune, voir rouge selon les modèles). On remarque notamment que les régions associées à de fortes largeurs spectrales (selon les modèles gaussien et exponentiel) sont identifiées par le modèle SVD avec des valeurs positives de f_2/f_1. Nous notons également que ces régions sont beaucoup plus étendues avec le modèle gaussien qu'avec le modèle exponentiel.

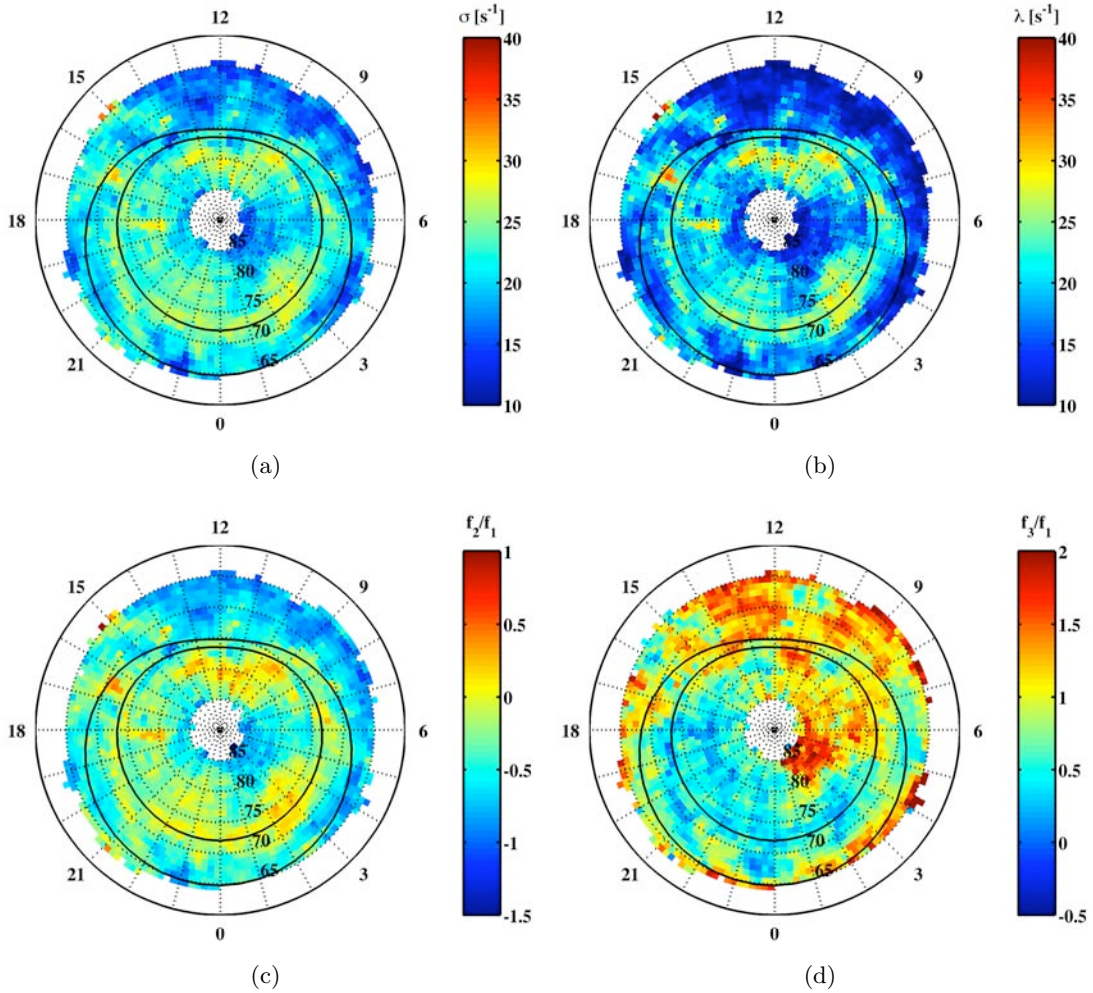


FIG. 3.9 – *Projection en MLT/Λ de (a) σ (temps caractérisque issu du modèle gaussien), (b) λ (temps caractéristique issu du modèle exponentiel), (c) f_2/f_1 et (d) f_3/f_1. Les latitudes invariantes magnétiques (Λ) varient de 60° à 90°. L'ovale auroral superposé est défini pour une activité géomagnétique modérée (3 < Kp < 4) (Holzworth and Meng, 1975).*

Quant à la distribution moyenne du paramètre f_3/f_1, exposée à la figure 3.9(d), elle permet de mettre en évidence une asymétrie jour/nuit. Une large région apparaît entre $03^{h}00$ et $15^{h}00$ MLT pour laquelle le rapport $f_3/f_1 \geqslant 1.5$. Cette région caractérise des ACFs possédant une forme très piquée (*cf* discussion § 3.1.2.2). Une seconde région est identifiable sur l'intervalle $18^{h}00$ - $24^{h}00$ MLT pour laquelle le spectre de rétrodiffusion est caractérisé (d'après la figure 3.6) par une forme intermédiaire (gaussien - exponentiel).

La comparaison des figures 3.9(c) et 3.9(d) permet de mettre en avant deux comportements distincts de la rétrodiffusion. Un premier où les paramètres f_2/f_1 et f_3/f_1

sont liés et un second où ils semblent être indépendants l'un de l'autre. Notamment dans le secteur MLT centré sur $12^{h}00$, où une large région avec un champ électrique turbulent intense (Golovchanskaya et al., 2006) se traduit par un accroissement de la largeur spectrale. Le paramètre f_3/f_1 devient fort sur une petite zone centrée sur 11h00 MLT qui est co-localisée sur le gradient de la largeur spectrale. La tache que l'on peut observer à partir de la projection de f_3/f_1 (Fig. 3.9(d)), localisée entre $23^{h}00$ - $24^{h}00$ MLT et entre 65 - 70°Λ, semble être indépendante de la valeur de f_2/f_1 (Fig. 3.9(c)). D'après sa position en latitude, on peut suspecter une reconnexion des lignes de force localisée dans la queue magnétosphérique qui peut par transfert de flux magnétique entraîner des modifications localisées du régime de rétrodiffusion (Cerisier et al., 2005). On peut également noter une transition entre les largeurs spectrales étroites et larges, localisée à 70°Λ pour le secteur $02^{h}00$ - $08^{h}00$ MLT. Sur cet intervalle, la frontière de la largeur spectrale est co-localisée avec une transition du rapport f_3/f_1. De part et d'autre de cette frontière, les ACFs présentent une forme plutôt exponentielle au-dessus de cette frontière et une forme plutôt intermédiaire vers les plus basses latitudes (*cf* abaque de la figure 3.6). La forme plutôt exponentielle des ACFs est la conséquence directe d'un faible rapport signal/bruit qui a pour effet de relever la queue des ACFs (*cf* Fig. 3.10(d)). A partir de $08^{h}00$, la transition du rapport f_3/f_1 disparaît, tandis que la frontière de la largeur spectrale dérive vers les hautes latitudes, jusqu'à 75°Λ à $12^{h}00$ MLT. La frontière de la largeur spectrale (caractérisée par une transition en latitude de l'échelle de couleur du bleu vers le jaune - rouge) semble suivre la frontière entre la couche limite du feuillet plasma (PSBL[1]) et le feuillet plasma central (CPS[2]) sur l'ensemble de l'intervalle $02^{h}00$ - $12^{h}00$ MLT (*cf* figure 1.5). Woodfield et al. (2002b) ont justement montré que la frontière de la largeur spectrale pour l'intervalle $03^{h}00$ - $08^{h}00$ MLT se positionne plutôt au niveau des lignes de champ dipolaires localisées avec la frontière CPS/PSBL tandis que pour le secteur pré-midi MLT, la frontière de la largeur spectrale est plutôt associée à la séparatrice entre les lignes de champ fermées et ouvertes (Chisham et al., 2005c).

En marge du comportement global des paramètres de la rétrodiffusion pour lequel nous retrouvons des propriétés mises en évidence par des travaux faisant intervenir des statistiques plus importantes (Villain et al., 2002; Woodfield et al., 2002b; Chisham et al., 2005c), nous notons la présence de structures plus spécifiques, dues à la taille limitée de l'échantillon. Notamment dans le secteur soir où certaines signatures semblent être reliées à des événements particuliers On remarque sur les figures 3.9(a), 3.9(b) et 3.9(c), une tache orange définie par l'intervalle $16^{h}00$ - $17^{h}00$ MLT et 69 - 74°Λ, indiquant une augmentation localisée de la largeur spectrale ou encore une traînée

[1]En anglais : *Plasma Sheet Boundary Layer.*
[2]En anglais : *Central Plasma Sheet.*

entre $18^{\mathrm{h}}00$ et $22^{\mathrm{h}}00$ MLT. Ces cas particuliers mettent en évidence toute la difficulté à distinguer les signatures potentielles des régions frontières magnétosphériques des événements singuliers indépendants. Nous reviendrons sur ces deux exemples à la sous-section § 3.3.3 où une interprétation sur l'origine de ces signatures sera donnée.

L'identification des régions magnétosphériques à partir des seules observations SuperDARN demeure ambiguë. La largeur spectrale seule n'est pas suffisante pour identifier les signatures des régions magnétosphériques. L'ajout de paramètres supplémentaires, tels que le paramètre f_3/f_1 bien qu'il soit principalement anti-corrélé avec la largeur spectrale, permet d'améliorer cette description. Comme nous l'avons mentionné précédemment, le comportement f_3/f_1 est différent de celui de la largeur spectrale dans certain secteur MLT/Λ.

3.3.2 Caractérisation de l'erreur des modèles

L'objectif de cette section est de mettre en évidence l'influence du secteur horaire magnétique et de la latitude magnétique sur la précision des modèles. Les mesures SuperDARN reflètent l'état de turbulence ionosphérique conditionnée principalement par les gradients de densité (Fejer and Kelley, 1980; Greenwald et al., 1995; Ballatore et al., 2001). Ces gradients de densité sont en parti contrôlés par les flux de particules (plus ou moins) énergétiques qui se précipitent depuis la magnétosphère (Danskin et al., 2002). Leurs caractéristiques varient donc selon le secteur horaire magnétique et la latitude magnétique : vitesse Doppler des irrégularités, puissance des signaux rétrodiffusés, largeur des spectres, la forme des spectres. Par exemple, on observe des spectres à une composante de type gaussien ou lorenzien, ou des spectres plus complexes multi-pics par la présence de régimes distincts de vitesse à l'intérieur du volume diffuseur. Cette dernière catégorie est associée à des ACFs dont l'amortissement est modulé. Nous remarquons que cette dernière catégorie de spectre ne peut pas être entièrement décrite par les modèles de l'algorithme FitACF, ni par les trois modes SVD. Pour ces cas, seule la composante principale, *i.e.* celle qui est associée à la modulation la plus grande, est décrite puisque les modèles caractérisent uniquement l'amortissement moyen de l'ACF. Néanmoins, une autre utilisation de la méthode SVD permettrait une analyse plus complète de ces ACFs. Pour cela, nous pourrions utiliser le nombre de modes nécessaires pour reproduire ces ACFs comme paramètre afin de caractériser qualitativement la complexité des ACFs.

Dans ce qui suit, nous profitons donc du fait que les modèles gaussien, exponentiel et SVD, de par les hypothèses de départ, ajustent préférentiellement des ACFs de forme

différente pour étudier la projection des écarts entre la mesure et ces modèles et ainsi faire ressortir les différents régimes de la rétrodiffusion. Ce genre de comparaison est toujours délicat en ce sens que l'ACF intègre les erreurs liées aux principes de mesure, à l'hypothèse de stationnarité des processus physiques et à la propagation...

3.3.2.1 Distribution des écarts quadratiques relatifs des modèles

Afin d'étudier les performances des modèles selon les caractéristiques de la rétrodiffusion, nous étudions la projection des écarts entre les modèles et les observations. Pour cela, nous calculons l'écart quadratique relatif de chaque modèle :

$$\langle \epsilon_{l,s} \rangle = \sqrt{\frac{1}{N-2} \sum_{i=1}^{N} \left(\frac{|R(\tau_i)| - |\widehat{R}_{l,s}(\tau_i)|}{|R(\tau_i)|} \right)^2} \tag{3.4}$$

$$\langle \epsilon_{SVD} \rangle = \sqrt{\frac{1}{N-3} \sum_{i=1}^{N} \left(\frac{|R(\tau_i)| - |\widehat{R}_{SVD}(\tau_i)|}{|R(\tau_i)|} \right)^2} \tag{3.5}$$

Notons que le nombre de degrés de liberté n'est pas le même dans le modèle SVD ($\nu = 3$) et dans les modèle gaussien et exponentiel ($\nu = 2$).

Les panneaux (a), (b) et (c) de la figure 3.10 montrent respectivement la distribution moyenne de l'écart quadratique relatif commis sur l'ajustement des ACFs à partir des modèles gaussien, exponentiel et de celui issu de la SVD. Pour chacune des distributions, les erreurs se comportent globalement de manière identique. Cependant, une différence nette de l'amplitude des erreurs entre le modèle statistique et les modèles issus de FitACF est à souligner. Le modèle SVD permet un ajustement des ACFs avec une erreur inférieure d'un facteur 1.5 par rapport aux deux autres modèles. Ceci s'explique par le fait que ce dernier est mieux adapté aux données.

Côté nuit, nous pouvons voir, pour les trois modèles, une frontière entre les grandes erreurs (vers le pôle) et celles plus faibles d'un facteur 2 (vers l'équateur). Cette frontière est co-localisée avec la frontière polaire de l'ovale auroral autour de $\sim 72°\Lambda$ et qui coïncide avec la frontière de la largeur spectrale pour le secteur $21^{\mathrm{h}}00$ - $08^{\mathrm{h}}00$ MLT. A partir de $06^{\mathrm{h}}00$ jusqu'à $09^{\mathrm{h}}00$ MLT, la frontière entre les grandes et les petites erreurs suit la frontière observée par le rapport n_{SVD}. Côté jour, à partir de $09^{\mathrm{h}}00$, le gradient en latitude de l'erreur s'estompe. Une région plus spécifique, localisée sur le secteur $09^{\mathrm{h}}00$ - $13^{\mathrm{h}}00$ MLT entre 78° et 84°Λ de latitude, ou l'erreur est plus forte peut être

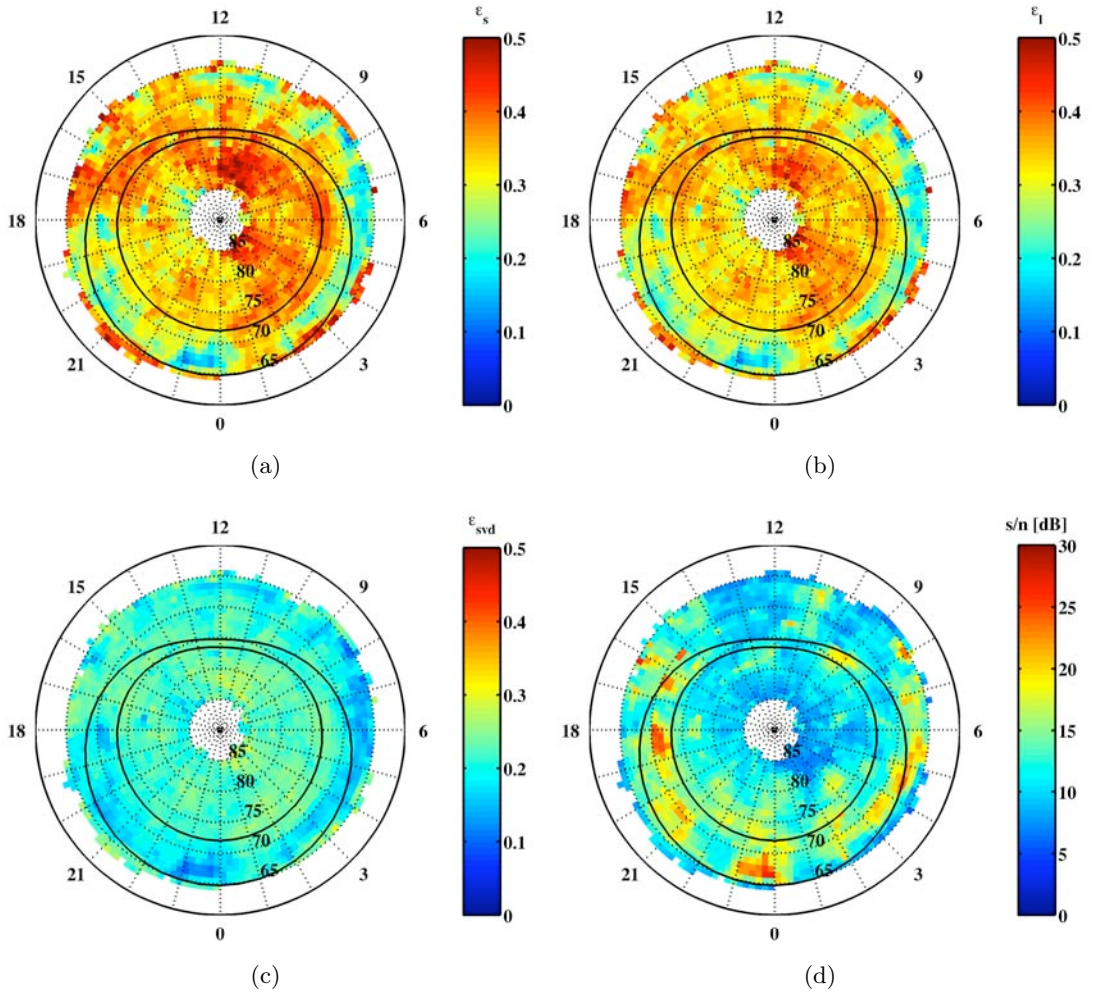


FIG. 3.10 – *Comparaison de la distribution MLT/Λ de l'écart quadratique relatif aux modèles gaussien* $\langle \epsilon_s \rangle$ *(a), exponentiel* $\langle \epsilon_l \rangle$ *(b), celui issu de la description SVD* $\langle \epsilon_{svd} \rangle$ *(c). La projection du rapport signal/bruit est indiquée à la figure (d). L'ovale auroral superposé est défini pour une activité géomagnétique modérée,* $3 < \mathrm{Kp} < 4$ *(Holzworth and Meng, 1975).*

associée au cornet polaire. Cette région pour laquelle la signature radar des irrégularités électroniques est moins bonne correspond à une région où le niveau de turbulence est plus élevé. Nous remarquons également que la bande à basse latitude entre $02^{\mathrm{h}}00$ et $08^{\mathrm{h}}00$ MLT observée à partir de la largeur spectrale et du facteur de forme coïncide également avec une région où les trois modèles reproduisent mieux les mesures. Cette bande est associée à un régime ou le rapport signal/bruit est plus élevé (*cf* Fig. 3.10(d)). On note un comportement similaire pour les régions suivantes :

- 18^h00 - 19^h00 MLT, 70° - 73°Λ
- 19^h30 - 21^h00 MLT, 65° - 70°Λ
- 23^h00 - 24^h00 MLT, 66° - 69°Λ

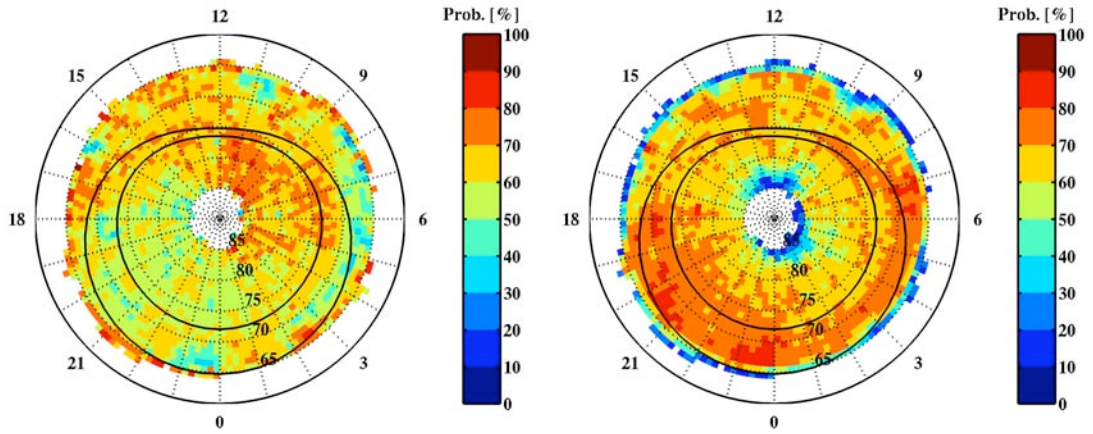
La comparaison de la distribution des erreurs avec celle du rapport signal/bruit (*cf* Fig. 3.10(d)) permet de mettre en avant un comportement différent entre le côté jour et le côté nuit. Côté nuit, les régions de plus faible erreur, pour les trois modèles, sont associées à une augmentation du rapport signal/bruit. Côté jour, nous notons globalement des erreurs plus fortes. Ces erreurs sont liées principalement à la reproductibilité des ACFs et dans une moindre mesure au niveau de bruit. A titre d'illustration, notons les comportements différents de deux régions :

– 09^h00 MLT autour de 73°Λ : le rapport signal/bruit est localement élevé (18 dB) et les erreurs sur les modèles y sont (légèrement) plus faibles. Ceci indique que les écarts aux modèles sont influencés par la qualité des données.
– 16^h00 et 17^h00 MLT centré sur l'ovale auroral : Le rapport signal/bruit est localement élevé ($\sim$ 23 dB), mais les erreurs sur les modèles restent fortes. Dans cette région, les erreurs semblent être dominées par le manque de reproductibilité des ACFs.

La projection des erreurs permet de mettre en relief deux régions bien distinctes. Une région à basse latitude côté nuit où l'erreur est plus faible pour les trois modèles peut être associée aux lignes de champ géomagnétique fermées. La seconde région associée à la calotte polaire est caractérisée par des erreurs plus importantes. Ces erreurs sont la conséquence des irrégularités formées sur les lignes de champ ouvertes sur le milieu interplanétaire. Côté jour, la distribution des erreurs ne laisse pas apparaître de structures évidentes à cause du niveau de turbulence plus important à l'approche du cornet polaire. Ceci est également dû à la taille de l'échantillon qui moyenne des cas très différents. La projection des erreurs pour un jour individuel permet de dessiner des frontières plus nettes. Notamment, la frontière entre les grandes et petites erreurs que l'on distingue difficilement sur les panneaux (a), (b) et (c) de la figure 3.10, apparaît dans ces cas plus distinctement.

3.3.2.2 Comparaison des modèles

Nous étudions maintenant la cohérence entre les modèles gaussien, exponentiel et SVD et les observations radar. Notre objectif est de mettre les modèles en compétition. Dans chaque cellule de la grille MLT/Λ et pour chaque ACF, nous comparons



(a) (b)

FIG. 3.11 – *Projection en MLT/Λ de la probabilité $P(\epsilon_l \leqslant \epsilon_s)$ (a) et de la probabilité que $\epsilon_{SVD} \leqslant \min\{\epsilon_s, \epsilon_l\}$ (b). L'ovale auroral superposé est défini pour une activité géomagnétique modérée,* $3 < \mathrm{Kp} < 4$ *(Holzworth and Meng, 1975).*

les erreurs respectives des trois modèles (*cf* Eq. (3.4) et (3.5)). Nous déduisons les probabilités suivantes : la probabilité que le modèle exponentiel soit plus performant que le modèle gaussien ($P(\epsilon_l \leqslant \epsilon_s)$), ainsi que la probabilité que le modèle SVD soit plus performant que les modèles gaussien et exponentiel ($P(\epsilon_{SVD} \leqslant \min\{\epsilon_s, \epsilon_l\})$). Par exemple, la probabilité $P(\epsilon_l \leqslant \epsilon_s)$ est définie par l'équation (3.6), où $N_{(\epsilon_l \leqslant \epsilon_s)}$ est le nombre d'ACFs pour lesquelles l'erreur du modèle gaussien est inférieure à l'erreur du modèle exponentiel et N_{tot} est le nombre total d'ACFs dans la cellule considérée :

$$P(\epsilon_l \leqslant \epsilon_s) = 100 \times \frac{N_{(\epsilon_l \leqslant \epsilon_s)}}{N_{tot}} \tag{3.6}$$

Les panneaux (a) et (b) de la figure 3.11 indiquent les probabilités respectives, $P(\epsilon_l \leqslant \epsilon_s)$ et $P(\epsilon_{SVD} \leqslant \min\{\epsilon_s, \epsilon_l\})$. Les probabilités d'observer le modèle exponentiel (et par le complément le modèle gaussien) sont indépendantes de la distribution de la largeur spectrale (*cf* Fig. 3.9). Côté nuit, au-dessus de 70°Λ, ces deux modèles sont équivalents. Côté jour, le modèle exponentiel par rapport au modèle gaussien possède la plus forte probabilité de décrire les signaux rétrodiffusés. La répartition des deux modèles est cohérente avec la projection du facteur de forme n_{SVD} (*cf* Fig. 3.9(d)). Les régions pour lesquelles la probabilité que le modèle exponentiel soit plus performant se superposent aux régions décrites par un facteur de forme inférieur à 1, et réciproquement.

Nous remarquons également que les régions les moins bruitées, côté nuit à basse

latitude (*i.e.* celles correspondant à un rapport signal/bruit plus élevé) voient augmenter la probabilité que le modèle gaussien décrive mieux les échos radar que le modèle exponentiel. Cette dépendance des modèles face à la qualité des mesures, nous indique que la présence de bruit (qui a pour effet de relever la queue des ACFs), privilégie le modèle exponentiel. Ceci remet en cause la capacité des modèles utilisés par le programme FitACF pour reproduire les ACFs mesurées. De plus, le nombre limité de retards ainsi que la faible résolution temporelle des ACFs ne facilite pas l'évaluation de leur forme, même visuellement. Pire encore, le nombre de points significatifs d'une ACF est bien souvent inférieur à 18 à cause de la présence de mauvais retards, de la durée de vie des irrégularités par rapport aux échelles de la turbulence (décroissance plus ou moins rapide des ACFs) et du niveau de bruit (Ponomarenko and Waters, 2006).

3.3.3 Interprétation de deux signatures particulières : le 12 février 2003

Dans la section précédente, nous avons mis en évidence l'influence du secteur MLT et de la latitude magnétique sur les paramètres issus des trois modèles gaussien, exponentiel et SVD et ainsi que de leur erreur respective. A l'échelle globale de l'ionosphère, nous avons notamment identifié une signature qui semble coïncider avec la frontière CPS/PSBL côté matin. A des échelles spatiales plus petites, des signatures plus singulières apparaissent. Ces signatures sont dues à la taille relativement faible de l'échantillon (quatre jours uniquement), et semblent être davantage liées à des événements de type précipitation qu'à l'empreinte ionosphérique des régions magnétosphériques.

Afin de mettre en évidence toute la difficulté de séparer les signatures des régions frontières de la magnétosphère de celles d'événements singuliers, nous allons dans cette section interpréter deux événements particuliers. Ces événements sont définis à partir des quatre panneaux de la figure 3.9 par :

- Evénement A : localisé entre $16^{h}00$ - $17^{h}00$ MLT et centré sur 73°Λ de latitude magnétique, cet événement se révèle particulièrement intense selon la description du modèle SVD (panneau (c)) par rapport à la description donnée par les modèles gaussien et exponentiel (panneau (a) et (b)).
- Evénement B : caractérisé par une longue zone délimitée par une augmentation de la largeur spectrale qui dérive vers les basses latitudes à partir de $18^{h}00$ MLT jusqu'à $22^{h}00$ MLT. Cet événement, au contraire, est davantage visible sur les panneaux (a) et (b) que sur le panneau (c).

3.3.3.1 Evénement A

L'analyse de l'échantillon montre que la signature de cet événement est majoritairement associé aux données des radars d'Hankasalmi, de Þykkvibaer et de Stokkseyri pour des temps allant de $12^{h}00$ à $17^{h}30$ temps universel (TU) le 12 février 2003. C'est le radar de Þykkvibaer (basé en Islande) qui contribue le plus à cet événement (en termes d'échos radar).

La figure 3.12 montre les conditions du milieu interplanétaire enregistrées au point de Lagrange (L1) par le satellite ACE. Pendant cet intervalle de temps les conditions du milieu interplanétaire sont plutôt stables. La vitesse du vent solaire notamment varie peu et oscille autour de ~ 370 km.s^{-1}. Afin de positionner les observations SuperDARN dans le contexte géophysique, il est nécessaire d'ajouter un délai aux observations du satellite ACE, ce qui permet de prendre en compte la propagation du vent solaire entre le point de Lagrange et l'ionosphère. Dans ce cas, on parlera alors de «temps terrestre». En supposant une vitesse d'écoulement moyenne du vent solaire (~ 370 km.s^{-1}), nous ajoutons ~ 50 min au temps de mesure du satellite ACE. On remarque donc pour l'intervalle de temps qui nous intéresse *i.e.* $12^{h}00$ à $17^{h}30$ TU terrestre, que la composante verticale (B_z) du champ magnétique interplanétaire s'oriente vers le sud à partir de $13^{h}10$. Elle se stabilise autour de -5 nT jusqu'à $15^{h}20$ TU, puis elle devient quasi-nulle. En début de période, la composante B_y est orientée selon la direction matin - soir ($B_y \geqslant 0$) puis se retourne (orientée vers le matin) lorsque B_z devient négatif. Ensuite, elle augmente graduellement jusqu'à la fin de l'après-midi. La composante B_x reste globalement constante et plutôt négative. Aucune variation importante de la pression dynamique n'est à noter.

La figure 3.13 montre l'évolution temporelle des paramètres de la rétrodifussion sur l'intervalle $15^{h}00$ - $15^{h}16$ TU (temps terrestre) à partir du radar de Þykkvibaer. Un balayage complet sur deux est montré (une image toutes les quatre minutes). Les colonnes 2, 3 et 4 montrent respectivement l'évolution de la vitesse Doppler radiale au radar (en m.s^{-1}), de la largeur spectrale (en m.s^{-1}) et de la puissance rétrodiffusée (en dB). Les vitesses négatives identifient des irrégularités qui s'éloignent du radar.

Les cartes successives montrent la formation d'une frontière de renversement de convection (CRB[3]), marquée sur la figure (colonne 2) par une transition du code de couleur bleu/rouge, qui est co-localisée avec l'événement (Fig. 3.9). Associée à cette ligne de renversement, on note une augmentation au cours du temps de la largeur spectrale qui englobe cette zone. L'accroissement de la puissance rétrodiffusée indique

[3]En anglais : *Convection reversal boundary*

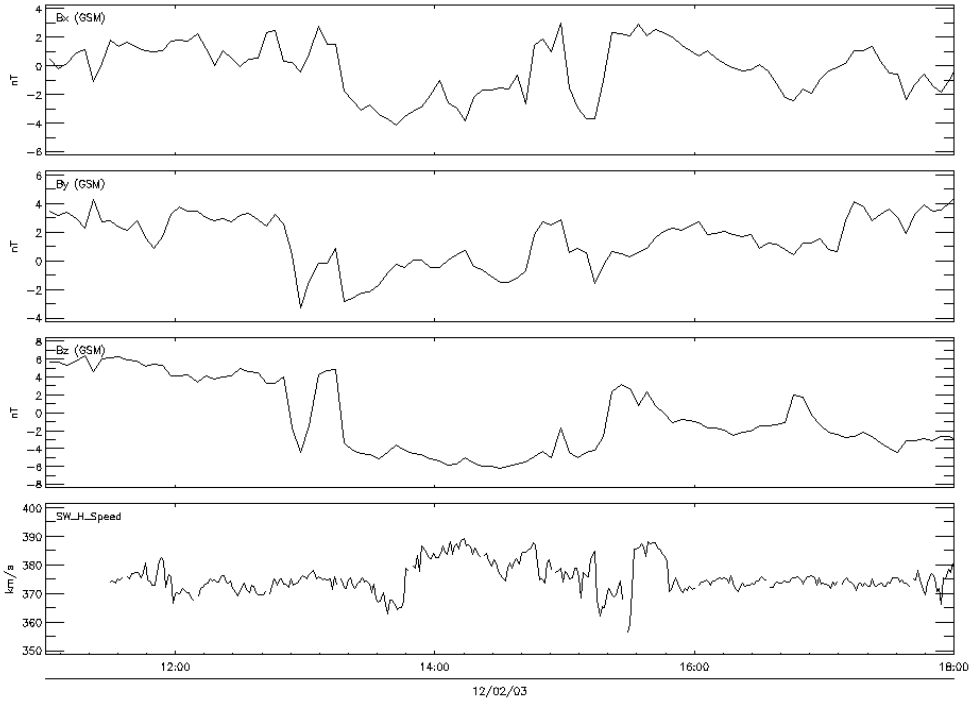


FIG. 3.12 – *Evolution des conditions du milieu interplanétaire pendant l'intervalle* $12^{h}00$ - $18^{h}00$ *TU, enregistrées par le satellite ACE au point de Lagrange (L1). De haut en bas, les composantes* B_x, B_y, B_z *du champ magnétique interplanétaire, ainsi que la vitesse du vent solaire sont indiquées.*

que les observations radars sont de bonne qualité et par conséquent valide la signature spatiale de la largeur spectrale et de la vitesse.

La valeur de la largeur spectrale résulte en partie de la distribution moyenne des vitesses des irrégularités dans le volume sondé. La présence d'un cisaillement de vitesse accroît localement la largeur spectrale. La CRB, qui est la conséquence directe de l'écoulement du vent solaire autour de l'environnement terrestre en entraînant les lignes de champ géomagnétiques dans un mouvement de convection (par frottement visqueux et par reconnexion), ne correspond pas forcément à la signature d'une frontière magnétosphérique. Notamment, (Lester et al., 2001) ont montré que la CRB n'est pas forcément co-localisée avec la frontière entre les lignes de champ géomagnétiques ouvertes et celles fermées. De nombreux auteurs se sont davantage intéressés à étudier la position en latitude d'un gradient de la largeur spectrale (qui sépare deux régions, celles des fortes largeurs spectrales et celles des petites largeurs spectrales) pour localiser cette frontière (Dudeney et al., 1998; Lester et al., 2001; Parkinson et al., 2002; Woodfield et al., 2002a,b,c; Chisham et al., 2005a,c). Nous aborderons plus en détail les

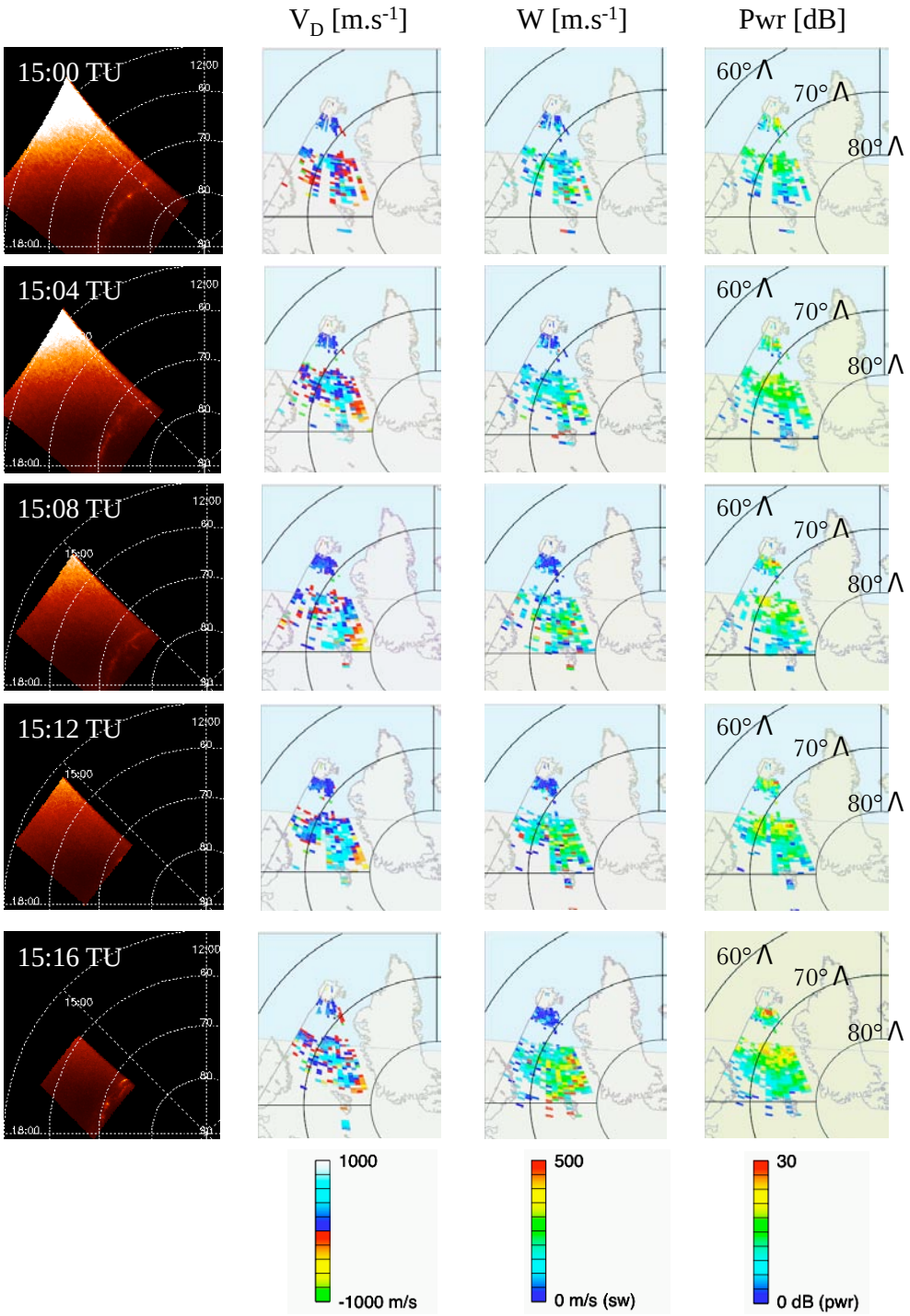


FIG. 3.13 – *Evolution de* 15^h00 *à* 15^h16 *TU de l'événement A.* **1ère colonne :** *images UV de l'ovale auroral vu par l'intrument WIC du satellite IMAGE.* **Colonnes 2, 3 et 4 :** *signature radar observée par le radar de* Þykkvibaer. *De gauche à droite, les paramètres vitesse Doppler, largeur spectrale et puissance rétrodiffusée sont exposés. Les données sont projetées en coordonnées* MLT/Λ *(12 MLT en haut de la figure).*

caractéristiques particulières de la frontière de la largeur spectrale et nous proposerons une nouvelle méthode pour notamment en améliorer sa détection à la section 3.4.

La comparaison des données SuperDARN avec l'instrument WIC/IMAGE (*Wideband Imager Camera*) qui propose une imagerie ultraviolette (140 - 160 nm) de l'ovale auroral révèle la présence d'un arc auroral au-dessus de 70°Λ et qui est décalé de quelques degrés vers l'équateur par rapport à la CRB (*cf* 1ère colonne de la figure 3.13). La succession des images révèle également un scintillement de l'arc auroral, ainsi qu'une lente dérive vers les basses latitudes. La présence d'un arc auroral est associé à une intensification localisée et structurée de la précipitation de particules faiblement énergétiques ($<$ 500 keV) dans l'ionosphère. Dans le cas d'une précipitation modérée comme ici, la scintillation de la luminosité suggère que le flux de particules varie. Cette variation contribue à une intensification des courants de fermeture dans l'ionosphère, et par conséquent à une augmentation du champ électrique. Ceci se traduit par une augmentation de la dérive $\mathbf{V} \propto \mathbf{E} \wedge \mathbf{B}$. Les observations SuperDARN enregistre une augmentation de la convection (couleur bleu ciel) à partir de 15^{h}08 TU terrestre. La présence d'un champ électrique turbulent fort explique également la présence de spectres larges plus bas en latitude par rapport à la CRB, ainsi que l'augmentation de la puissance rétrodiffusée. En effet, l'existence d'une région localisée dans laquelle la précipitation de particules énergétiques est exacerbée et structurée, participe à une augmentation locale de la densité électronique et donc des gradients électroniques associés ce qui favorise le développement d'instabilités (Danskin et al., 2002).

Les conditions du milieu interplanétaire étant stables, le radar de Stokkseyri qui pointe dans la direction opposée de celle de Þykkvibaer observe autour de 17^{h}00 TU terrestre, la même signature localisée dans la même zone.

L'interprétation géophysique des fortes largeurs spectrales mesurées par SuperDARN est délicate. Dans cet exemple, la région associée à de fortes largeurs spectrales se juxtapose à une CRB et à un arc auroral. Ce dernier est une structure fine (de $\sim$ 100 km en latitude qui peut être observé tout autour de la calotte polaire) qui couple l'ionosphère à la magnétosphère par l'intermédiaire de courant aligné au champ circulant le long de la couche limite du feuillet plasma (PSBL). Le PSBL est la couche supérieure du feuillet plasma central dont le bord externe se situe à la frontière des lignes de champ ouvertes/fermées. L'étude de cet événement illustre les résultats obtenus par Chisham et al. (2005a). Ces auteurs ont démontré que dans le secteur après-midi (12^{h}00 - 18^{h}00 MLT), l'association du gradient de la largeur spectrale comme traceur de la frontière des lignes de champ ouvertes/fermées n'est pas suffisamment robuste. La position de cette transition est localisée plus bas en latitude. Sur la figure 3.13, la frontière de la largeur spectrale n'est pas observée, malgré un **Bz** négatif. L'influence de ces phéno-

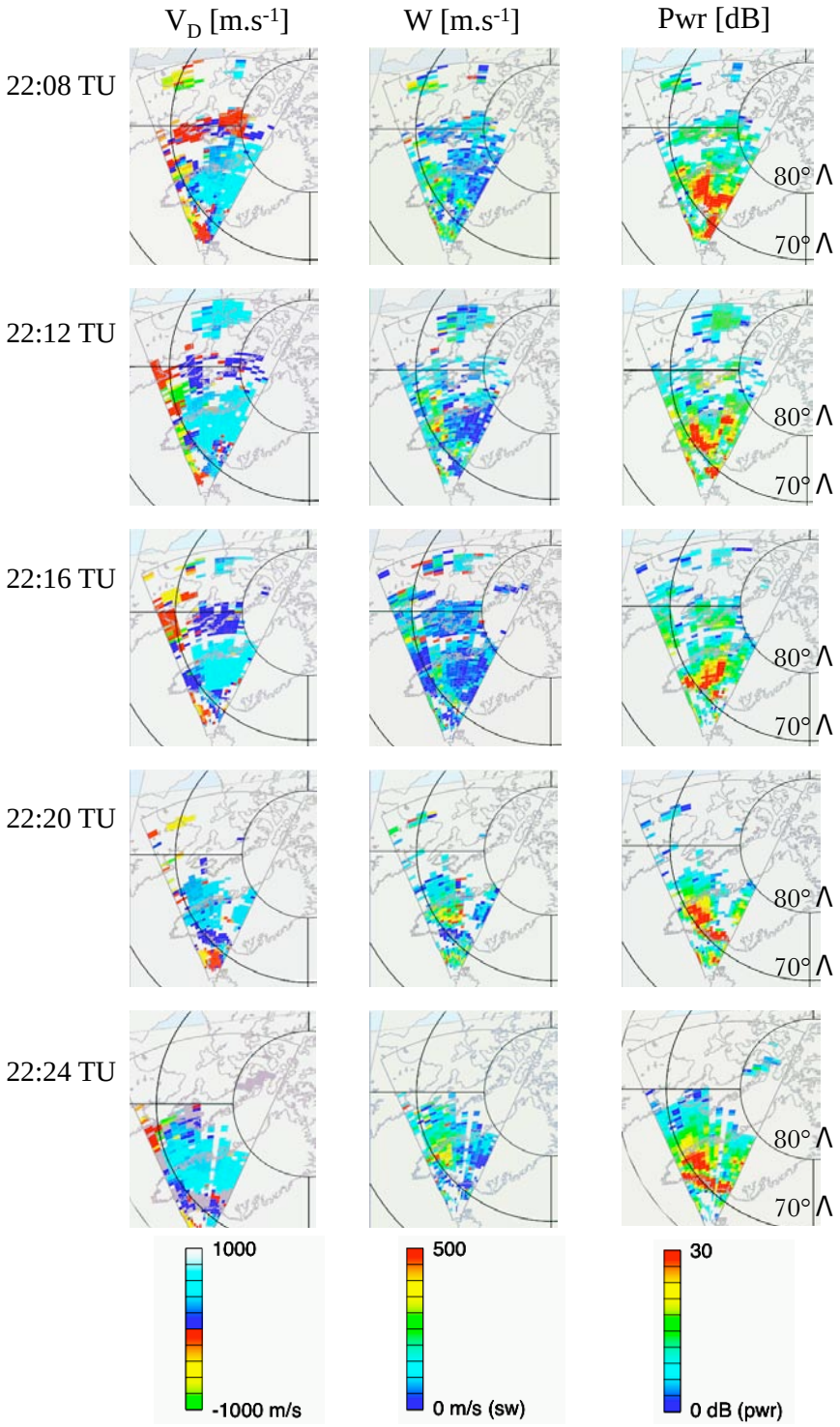


FIG. 3.14 – *Evolution de l'événement B entre* $22^{h}08$ *et* 22h24. *Les colonnes, de gauche vers la droite, montrent respectivement la projection de la vitesse Doppler* (V_D)*, de la largeur spectrale* (W) *et de la puissance rétrodiffusée* (Pwr)

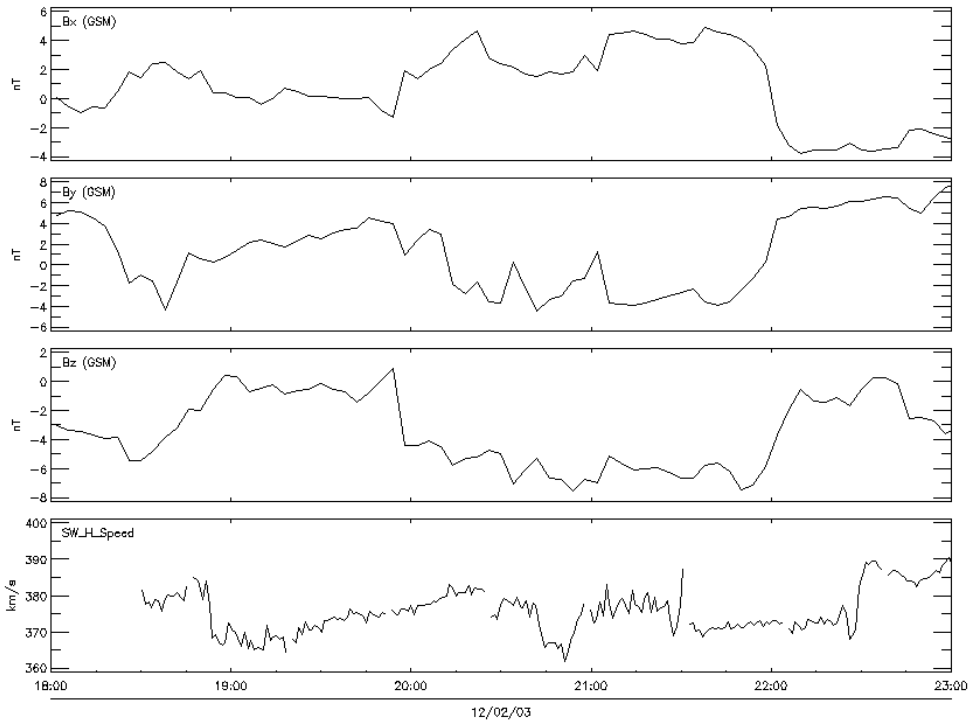


FIG. 3.15 – *Evolution des conditions du milieu interplanétaire sur la période* 18^h00 - 23^h00 *TU. De haut en bas, les composantes* B_x, B_y, B_z *du champ magnétique interplanétaire et la vitesse du vent solaire sont indiquées.*

mènes sur la signature largeur spectrale complique son utilisation comme un traceur éventuel de certaines régions frontières de la magnétosphère.

3.3.3.2 Evénement B

Nous nous intéressons maintenant à l'interprétation de l'événement B. Il correspond à une augmentation de la largeur spectrale et forme une traînée débutant vers 18^h00 MLT à 70°Λ de latitude et qui dérive lentement en direction du soir vers les basses latitudes et atteint 65°Λ à 22^h00 MLT (*cf* Fig. 3.9). Cette singularité correspond également à un cisaillement de vitesse observé par les radars d'Hankasalmi et de Stokkseyri, le 12 février 2003.

Le radar d'Hankasalmi (basé en Finlande) qui pointe en direction du Nord observe en premier le cisaillement durant l'intervalle 17^h00 - 20^h00 TU. Ensuite, c'est le radar de Stokkseyri, qui en direction de l'Ouest, observe le même cisaillement de vitesse

vers $22^{\mathrm{h}}10$ TU, plus bas en latitude $\leqslant 70°\Lambda$, entre $19^{\mathrm{h}}00$ - $22^{\mathrm{h}}00$ MLT. La figure 3.14 montre uniquement l'évolution temporelle des paramètres de la rétrodiffusion mesurés par le radar de Stokkseyri, les données du satellite IMAGE n'étant pas disponibles pour cet intervalle. L'accroissement de la largeur spectrale semble être la conséquence de la présence d'un cisaillement de vitesse quoique légèrement décalé vers les plus hautes latitudes. On note également l'augmentation du rapport signal/bruit sur cette même zone.

Les conditions du milieu interplanétaire au point de Lagrange (L1) sont exposées à la figure 3.15. Pour cette période, la vitesse du vent solaire augmente graduellement avant de chuter à partir de $20^{\mathrm{h}}30$ TU. En moyenne la vitesse oscille autour de $\sim$ 380 m.s^{-1}, ce qui donne un délai d'environ une heure pour que le vent solaire depuis le point de Lagrange atteigne la magnétopause. Sur cette période qui est comprise entre $17^{\mathrm{h}}00$ et $22^{\mathrm{h}}00$ TU terrestre, la composante B_z du champ magnétique interplanétaire est négative et la composante B_x est quasi nulle. La composante B_y est positive jusqu'à $21^{\mathrm{h}}00$ TU terrestre. Après, elle fait une incursion négative à $B_y \approx -4$ nT pendant environ une heure. La rotation de la composante B_y correspondant à une augmentation de la composante $B_x \approx 5$ nT et à une diminution brusque de la vitesse du vent solaire. La dérive de la CRB vers les plus basses latitudes observée entre les deux radars semble suivre les variations de la composante aube-crépuscule (B_y) de l'IMF.

Les deux événements singuliers que nous venons d'étudier révèlent toute l'ambiguïté au sujet de l'interprétation de la largeur spectrale. De nombreuses sources peuvent être à l'origine de son accroissement : cisaillement de vitesse, arc auroral, augmentation du champ électrique turbulent... Ces différentes sources ne sont pas forcément la manifestation d'une région ou d'une frontière particulière de la magnétosphère. L'utilisation des données UV provenant du satellite IMAGE apporte un diagnostic important dans l'interprétation des observations SuperDARN. Il contribue à lever l'ambiguïté sur l'origine géophysique de la signature particulière de la largeur spectrale.

L'utilisation des paramètres f_2/f_1 et f_3/f_1 issus de la méthode SVD doit permettre d'améliorer la description des régions frontières de la magnétosphère observées par les radars SuperDARN puisqu'ils permettent de décrire la largeur spectrale mais également la forme des spectres. Dans la section suivante, nous les utilisons pour identifier une frontière particulière : celle de la largeur spectrale qui est généralement associée à la frontière entre les lignes de champ géomagnétique fermées et celles ouvertes sur le milieu interplanétaire.

3.4 Détection de la frontière des lignes de champ ouvertes/fermées

> *« Si une personne a une espérance dépendant de l'arrivée d'un événement, la probabilité de cet événement est à la probabilité qu'il fasse défaut comme la perte en cas de non-occurence est au gain résultant de son arrivée.»*
>
> Enoncé du théorème de Bayes - Révérend Thomas Bayes (1702 – 1761)

Dans cette section, nous nous intéressons plus particulièrement à la détection de la frontière des lignes de champ ouvertes/fermées (OCB[4]). En effet, le réseau de radars SuperDARN est un instrument bien adapté à ce problème (Rodger, 2000).

L'OCB est la projection dans l'ionosphère de la frontière entre les lignes de champ géomagnétiques ouvertes et fermées. L'OCB est définie comme le bord polaire de l'émission de la raie rouge (630 nm) (Blanchard et al., 1995; Lester et al., 2001; Lockwood et al., 1993). en termes de précipitation d'électrons, l'OCB est équivalente à la frontière entre la pluie polaire et les précipitations provenant du feuillet plasma. La détection de l'OCB est importante pour réaliser un bilan des échanges d'énergie entre le vent solaire et l'environnement terrestre par l'intermédaire du taux de reconnexion (Hubert et al., 2006). Par exemple, le suivi de cette frontière permet d'étudier les différentes phases de l'évolution des sous-orages géomagnétiques, de la phase de déclenchement jusqu'à la phase de détente en passant par la phase d'expansion. La prédiction des sous-orages est un des enjeux pour la météorologie de l'espace. Par exemple, les perturbations du champ géomagnétique engendrées par les sous-orages peuvent affecter les dispositifs technologiques sur Terre et dans l'espace.

De nombreux travaux ont utilisé les propriétés spatiales des échos radar observés par SuperDARN pour localiser la frontière OCB. Par exemple, Newell et al. (2002) ont développé le modèle OVATION qui est dédié à la description de l'ovale auroral. Ce modèle utilise principalement l'imageur UVI embarqué à bord du satellite POLAR et les radars SuperDARN. Ces derniers interviennent notamment pour contraindre la localisation de la frontière polaire de l'ovale auroral et donc celle de l'OCB à partir de la ligne neutre de renversement de la convection (CRB[5]), même si, selon Lester et al. (2001), la CRB et l'OCB ne coïncident pas forcément. D'autres études se sont plutôt appuyées sur les propriétés spatiales de la largeur spectrale. Elles ont démontré que cette dernière est un sérieux candidat pour l'identification des régions frontières magnétosphériques (Baker et al., 1995; Rodger, 2000). La présence d'une transition graduelle de la largeur spectrale en latitude est fréquemment observée tout le long de la frontière équatoriale de la calotte polaire. Cette transition marquant la séparation

[4]En anglais : *Open/Closed field line Boundary*
[5]En anglais : *Convection Reversal Boundary*

entre les spectres étroits et larges (respectivement au Sud et au Nord de cette limite) est appelée frontière de la largeur spectrale (SWB[6]). Les résultats montrent une bonne corrélation avec l'OCB pour une large gamme de secteurs MLT (Baker et al., 1986; Milan et al., 1999; Rodger et al., 1995). De nombreuses études ont tenté de déterminer les processus physiques à l'origine des spectres larges observés à haute latitude. La présence d'un champ électrique turbulent plus important pourrait en être la cause (Rodger et al., 1995). Dudeney et al. (1998) évoquent les précipitations structurées d'électrons de faible énergie (< 100 eV). André et al. (2000a) met en avant la présence d'activité d'ondes électrostatiques basse fréquence (Pc1) dans les régions se projetant dans l'ionosphère à des latitudes coïncidant avec le régime des grandes largeurs spectrales.

L'association de la SWB avec les frontières géophysiques est une tâche délicate. En effet, différentes études ont montré que la SWB peut être l'empreinte ionosphérique de différentes régions frontières magnétosphériques selon le secteur MLT (Dudeney et al., 1998; Lester et al., 2001; Parkinson et al., 2002; Woodfield et al., 2002a,b; Chisham et al., 2004, 2005c,a). Pour les secteurs MLT $18^{h}00$ - $02^{h}00$ et $08^{h}00$ - $12^{h}00$, la localisation de la SWB coïncide avec celle de l'OCB, tandis que pour les secteurs après-midi et post-minuit, la SWB est localisée au Sud de l'OCB, sur des lignes de champ fermées qui sont co-localisées avec la séparation entre le feuillet plasma central et la couche limite du feuillet plasma. La présence de deux régimes distincts de largeur spectrale de part et d'autre de la SWB suggère la possibilité de discriminer deux familles de spectres et donc deux régions géophysiques. Les quantités telles que le rapport signal/bruit ou encore le facteur de forme semblent être moins pertinentes pour la localisation de l'OCB. Le premier dépend surtout de la présence d'irrégularités, tout le long du chemin de propagation du faisceau. Le second, comme nous venons de le discuter au § 3.3.1, porte une signature qui ne semble pas être connectée à l'OCB, côté matin. De plus, la présence du bruit rend son utilisation délicate.

La détection de la SWB est en soi un problème délicat puisqu'il n'existe pas de critère évident permettant de discriminer deux classes de spectres rétrodiffusés. La méthode communément employée est celle d'un seuillage appliqué sur la distribution spatiale de la largeur spectrale. La difficulté principale réside dans la définition d'une valeur seuil sur la largeur spectrale pour séparer les deux régions situées de part et d'autre de la frontière. Les études mentionnées ci-dessus ont utilisé des valeurs ad hoc pour étudier la dynamique de la SWB, allant de 150 à 300 m.s^{-1}. Cet éventail de valeurs proposées atteste bien de la difficulté de définir une valeur limite universelle. De plus, Chisham and Freeman (2004) notent que la nature de la transition (en termes de gradient et d'amplitude) varie selon le secteur MLT et qu'elle est bien définie jus-

[6]En anglais : *Spectral Width Boundary*

tement pour les secteurs MLT où l'OCB et la SWB correspondent le mieux. Chisham and Freeman (2003) montrent également que l'utilisation d'un filtre médian appliqué spatialement et temporellement aux données avant le seuillage améliore la corrélation entre la SWB et l'OCB.

Les différentes études se sont contentées de localiser la SWB et de la comparer empiriquement avec les régions frontières identifiées par satellite sans tenter d'établir un critère permettant de calculer une erreur sur sa localisation. De plus, comme nous l'avons remarqué précédemment, la largeur spectrale estimée par le biais de l'algorithme FitACF ne permet pas une description complète de l'ensemble des spectres. Ces différentes considérations nous amènent donc à utiliser un critère de décision Bayésien afin de discriminer les deux classes de spectres à partir du rapport adimensionnel f_2/f_1. Comme nous l'avons démontré dans le chapitre précédent, le paramètre f_2/f_1 décrit la largeur spectrale de l'ACF de façon quasi-indépendante de sa forme, ce qui n'est pas le cas de la largeur spectrale calculée avec l'algorithme FitACF à partir d'un modèle gaussien et exponentiel. L'approche Bayésienne, en plus de fournir un critère décisionnel robuste, permet également de quantifier l'erreur commise sur la localisation de la SWB. Cette erreur est généralement omise dans la littérature, alors qu'elle est capitale pour valider l'emplacement de l'OCB.

3.4.1 Vers une Inférence Bayésienne de l'OCB

L'approche Bayésienne est largement utilisée en théorie de la décision et en classification (Duda et al., 2001). C'est une approche probabiliste basée sur le degré de confiance qu'un événement se produise. Elle permet une description complète de la mesure.

Soit X l'espace des observations. Nous considérons la mesure x, dans notre cas la largeur spectrale f_2/f_1, comme une variable aléatoire continue. Posons $\omega = \omega_1$ et $\omega = \omega_2$ les deux classes possibles de la largeur spectrale situées de part et d'autre de la frontière ($\Omega = \{\omega_1, \omega_2\}$). Ces classes sont supposées disjointes. L'ensemble Ω représente les classes de départ du phénomène à étudier. Le problème revient à déterminer l'appartenance à une des deux classes de la mesure x, ainsi que le niveau de confiance. On appelle inférence la capacité de l'approche Bayésienne de déduire la cause à partir de la conséquence. A partir des définitions de la probabilité conditionnelle[7], la formule de Bayes énonce la probabilité d'observer la classe ω_j connaissant la valeur de la mesure x, comme suit :

[7]Soit A et B deux événements, la probabilité conjointe de A et B s'écrit : $P(A, B) = P(A|B)P(B) = P(B|A)P(A)$

$$P(\omega_j|x,\mathrm{I}) = \frac{p(x|\omega_j,\mathrm{I})P(\omega_j,\mathrm{I})}{p(x,\mathrm{I})} \quad ; \quad j = 1,2 \tag{3.7}$$

$$\text{avec} \quad p(x,\mathrm{I}) = \sum_{j=1}^{2} p(x|\omega_j,\mathrm{I})P(\omega_j,\mathrm{I}) \tag{3.8}$$

où $P(\omega_j,\mathrm{I})$ est la probabilité[8] a priori d'observer la classe ω_j. Le terme $p(x|\omega_j,\mathrm{I})$ est la vraisemblance de x et exprime la densité de probabilité conditionnelle de la mesure x connaissant la classe ω_j. Nous avons explicitement inclus le terme I dans la conditionnelle pour rappeler que n'importe quelle relation entre probabilités, selon Bayes, nécessite la connaissance du phénomène étudié et donc en dépend. Cette notion est importante dans la description Bayésienne des probabilités (*cf* § 3.4.3).

La quantité qui nous intéresse ici est la probabilité a posteriori $P(\omega_j|x,\mathrm{I})$. La formule de Bayes montre qu'en observant une mesure x, nous pouvons déduire de la probabilité a priori $P(\omega_j,\mathrm{I})$ une probabilité a posteriori $P(\omega_j|x,\mathrm{I})$. En d'autres termes, la formule de Bayes permet d'évaluer l'origine ω_j de la mesure x. L'évidence $p(x,\mathrm{I})$ est la densité de probabilité de x ; elle peut être interprétée comme un facteur de normalisation qui permet de conserver la notion de probabilité ($P(\omega_1|x,\mathrm{I}) + P(\omega_2|x,\mathrm{I}) = 1$). En effet, une observation appartient forcément à une classe et à une seule, par conséquent la somme des probabilités d'appartenance est égale à 1.

L'analyse Bayésienne s'appuie donc sur deux grandeurs : l'information a priori et la vraisemblance. L'a priori représente notre connaissance sur la phénoménologie étudiée et la vraisemblance permet d'assimiler de nouvelles observations avec notre connaissance.

3.4.2 Etude de cas : 10 octobre 1999

Afin d'illustrer la méthode, nous nous focaliserons sur l'événement du 10 octobre 1999, qui a été étudié en détail par Hosokawa et al. (2003). Les auteurs ont menés une étude conjuguée entre les deux hémisphères afin de comparer la dynamique de la SWB du côté jour, lors d'un sous-orage magnétique.

[8] la majuscule $P(\cdot)$ définit une probabilité et la minuscule $p(\cdot)$, une densité de probabilité

3.4.2.1 Approximation de la vraisemblance

Avant d'appliquer le théorème de Bayes, la première étape consiste à définir la vraisemblance des deux classes ω_1 et ω_2 de la largeur spectrale. Pour cela, nous devons construire la fonction de densité de probabilité du paramètre f_2/f_1, autrement dit l'évidence de f_2/f_1. A partir des données du 10 octobre 1999, nous avons regroupé l'ensemble des échos radar de l'hémisphère nord répondant aux critères de sélection définis au § 2.2.1. Toutefois, le nombre maximum de mauvais retards contenus par ACF a été étendu à 10 au lieu de 0 initialement afin d'augmenter la taille de l'échantillon. Comme nous le montrerons dans le chapitre suivant, la corrélation entre la largeur spectrale estimée par l'algorithme FitACF et le rapport adimensionnel f_2/f_1 est conservée pour des ACFs ne possédant pas plus de 10 mauvais retards. Au-delà, la corrélation chute dramatiquement. Cet assouplissement permet d'augmenter la taille de l'échantillon (N= 366 168) et ainsi de palier au manque de statistique.

Le panneau supérieur de la figure 3.16 montre la fonction de densité de probabilité de f_2/f_1 (courbe bleue). Cette courbe bimodale suggère que la distribution de f_2/f_1 peut être approchée par une somme de deux distributions. En première approximation, nous supposons que les deux vraisemblances $p(x|\omega_1, \mathrm{I})$ et $p(x|\omega_2, \mathrm{I})$, suivent une loi normale et que les deux classes suivent une loi a priori uniforme telle que $P(\omega_1, \mathrm{I}) = P(\omega_2, \mathrm{I}) = \frac{1}{2}$. Les courbes verte et rouge du panneau supérieur de la figure 3.16 montrent l'approximation du terme $p(x|\omega_j, \mathrm{I})P(\omega_j, \mathrm{I})$ de l'équation (3.8) pour chacune des deux classes. Les deux lois normales ont été ajustées à partir de leur moyenne, variance et amplitude. L'évidence résultante, indiquée par les tirets noirs, ajuste relativement bien la distribution observée. Notons cependant la forme tronquée de $p(x, \mathrm{I})$ pour $x < -1$, qui peut être expliquée par la limite basse appliquée à la largeur spectrale pour éliminer les échos de sol. L'équation (3.8) permet alors une estimation directe des deux vraisemblances. Une troisième catégorie aurait pu être ajoutée pour améliorer l'ajustement de $p(x, \mathrm{I})$ pour $x < -1$, mais cela n'entraîne pas une amélioration significative de la détection de la SWB. De plus, la nature «abrupte» de la transition entre les deux régimes de largeur spectrale ne permet pas de décrire cette transition comme une classe supplémentaire.

En supposant l'équiprobabilité des deux a priori $P(\omega_1, \mathrm{I})$ et $P(\omega_2, \mathrm{I})$, nous ne spécifions aucune information particulière (a priori non-informatif) quant à l'état de connaissance du système. La seule information dont nous disposons réellement est l'existence de deux régimes de largeur spectrale. Le choix de la vraisemblance est un problème délicat. Les seules données SuperDARN ne permettent pas son estimation. D'autres instruments pourraient permettre son estimation. Par exemple, Baker et al. (1997) et Chisham and Freeman (2003) ont utilisé les données satellites pour identifier les échos

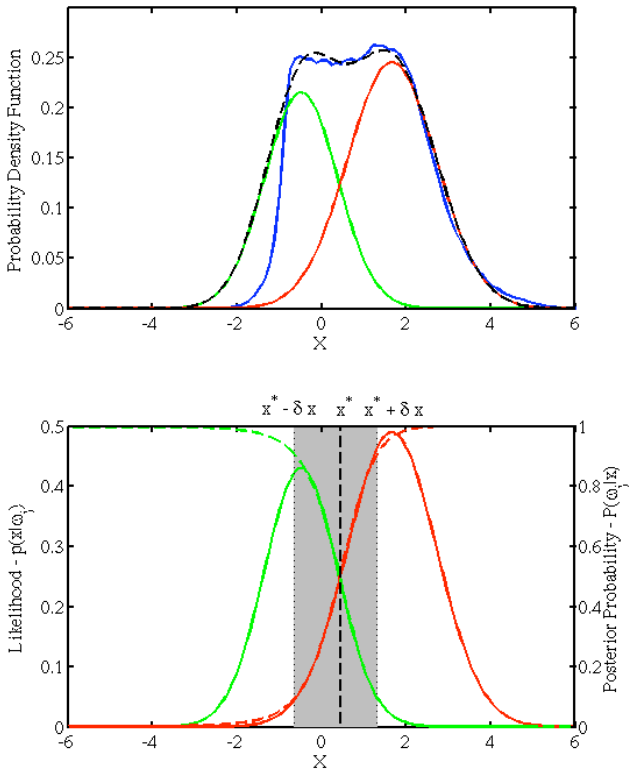


FIG. 3.16 – **Panneau supérieur :** *les fonctions de densité de probabilité pour le 10 octobre 1999 reconstruites à partir de l'ensemble des radars de l'hémisphère nord : l'évidence de f_2/f_1, i.e. $p(x, \mathrm{I})$ (courbe bleue) ; l'évidence de f_2/f_1 sachant la classe ω_1, i.e. $p(x|\omega_1, \mathrm{I})P(\omega_1, \mathrm{I})$ (courbe verte) ; l'évidence de f_2/f_1 sachant la classe ω_2, i.e. $p(x|\omega_2, \mathrm{I})P(\omega_2, \mathrm{I})$ (courbe rouge). Les tirets noirs montrent l'approximation de $p(x, \mathrm{I})$.* **Panneau inférieur :** *les courbes verte et rouge montrent les vraisemblances et les probabilités a posteriori respectives de ω_1 et ω_2. La zone grisée indique l'intervalle de confiance définie entre $P(\omega_1|x) \geqslant 10\%$ et $P(\omega_2|x) \geqslant 10\%$.*

radar répartis de part et d'autre de l'OCB afin d'étudier la distribution de la largeur spectrale indépendamment. Cependant, la distribution de la largeur spectrale étant large, le recouvrement des deux distributions est important, ce qui révèle la difficulté majeure dans le problème de la définition d'un seuil universel. Les différentes valeurs de seuil trouvées dans la littérature de 150 à 300 m.s^{-1}, en attestent. En choisissant un a priori non-informatif, nous minimisons notre impact sur l'estimation de ce seuil. Il existe également d'autres lois non-informatives, moins naturelles qu'une loi uniforme, comme par exemple les lois de Jeffreys (Jeffreys, 1961).

Le panneau inférieur de la figure 3.16 montre la vraisemblance $p(x|\omega_j, \mathrm{I})$ des classes

ω_1 et ω_2 (courbe verte et rouge). Les tirets, de même couleur, indiquent la probabilité a posteriori $P(\omega_j|x, \mathrm{I})$ estimée à partir de l'équation (3.7). L'égalité $P(\omega_1|x^*, \mathrm{I}) = P(\omega_2|x^*, \mathrm{I}) = \frac{1}{2}$ des deux probabilités a posteriori fixe la limite de décision entre les deux classes à $x^* = 0.52$. Par conséquent, cette limite localise la frontière de la SWB dans le domaine de définition de f_2/f_1, puisque :

- si $x < x^*, P(\omega_1|x, \mathrm{I}) > P(\omega_2|x, \mathrm{I})$ et donc ω_1 est la classe la plus probable.
- si $x > x^*, P(\omega_1|x, \mathrm{I}) < P(\omega_2|x, \mathrm{I})$ et donc ω_2 est la classe la plus probable.

L'erreur de décision sur ω_j après avoir observé x est donnée directement par

$$P(err|x, \mathrm{I}) = min[P(\omega_1|x, \mathrm{I}), P(\omega_2|x, \mathrm{I})] \ . \tag{3.9}$$

Notons que la limite entre les deux classes est estimée en maximisant l'incertitude sur l'origine de chaque mesure, là où la superposition est maximisée. Un des avantages de ce formalisme est la possibilité de définir un intervalle de confiance. Cet intervalle est défini par la plage des valeurs de x telle que $P(\omega_2|x, \mathrm{I}) \geqslant 10\%$ et $P(\omega_1|x, \mathrm{I}) \geqslant 10\%$. L'intervalle résultant est indiqué par la zone grisée du panneau inférieur de la figure 3.16.

La méthode que nous venons d'introduire permet la définition d'un seuil pour la classification de la largeur spectrale. Ce seuil correspond à peu près au minimum local du pic de l'évidence. Cependant, par rapport aux études précédemment citées, elle présente l'avantage d'être explicitée dans un cadre probabiliste général, ce qui permet d'auto-calibrer la valeur de ce seuil mais également de déduire des quantités aussi primordiales que les intervalles de confiance. De plus, la méthode intègre les a priori et les vraisemblances, ce qui en fait une méthode générale et facilement adaptable : l'a priori ou la vraisemblance peuvent s'adapter à des objectifs différents.

Connaissant maintenant la vraisemblance des classes ω_1 et ω_2, nous pouvons déterminer la probabilité a posteriori pour chacune des valeurs de f_2/f_1 d'appartenir au régime des petites ou des grandes largeurs spectrales.

3.4.2.2 Projection de la frontière Bayésienne

Comparons les performances de l'approche Bayésienne appliquée au paramètre f_2/f_1 avec celle plus classique de la SWB déterminée à partir de la largeur spectrale. Afin de faciliter cette comparaison, nous définissons une seule et unique fonction de

discrimination $-1 \leqslant G(x, \mathrm{I}) \leqslant 1$

$$G(x, \mathrm{I}) = P(\omega_2|x, \mathrm{I}) - P(\omega_1|x, \mathrm{I}) \ , \tag{3.10}$$

au lieu des deux fonctions a priori, telles que $G(x)$ est négative (respectivement positive) lorsque la classe dominante est ω_1 (ω_2).

Nous prenons comme exemple le faisceau 6 du radar de Þykkvibaer sur l'intervalle $09^{\mathrm{h}}30$ - $13^{\mathrm{h}}30$ TU du 10 octobre 1999. La figure 3.17 montre l'évolution temporelle en fonction de la latitude magnétique invariante de la largeur spectrale (panneau supérieur), de G(x) (au centre) et de la vitesse Doppler (panneau inférieur). Cet événement est identique à celui qui est présenté dans Hosokawa et al. (2003). Pendant cette période, le champ de vue du radar couvre depuis les latitudes moyennes jusqu'à la calotte polaire les régions du feuillet plasma central, de la couche limite basse latitude et du cornet polaire.

A la figure 3.17(a), nous employons une stratégie identique à celle utilisée dans Hosokawa et al. (2003) pour identifier la SWB. Nous fixons la frontière à partir d'un seuil défini à $W^* = 150$ m.s^{-1}. Ce seuil a été fixé ainsi car il semble s'adapter au mieux aux données étudiées. A chaque intervalle de temps, nous cherchons en direction du pôle la première paire de portes successives pour lesquelles la largeur spectrale excède la valeur seuil (W^*). La première de ces deux portes fixe la position de la SWB. Dans la figure 3.17(a), le trait noir gras indique la position de la SWB. Dans le panneau inférieur (Fig. 3.17(b)), nous utilisons le critère Bayesien et nous fixons la frontière à $G(x, \mathrm{I}) = 0$. $G(x, \mathrm{I}) > 0$ est associé à la classe ω_2, *i.e.* aux grandes largeurs spectrales et $G(x, \mathrm{I}) < 0$ est associé aux largeurs spectrales étroites (ω_1). La frontière Bayésienne est identifiée avec la même stratégie que celle utilisée pour localiser la SWB. L'intervalle marqué en vert indique l'extension en latitude de l'intervalle de confiance défini auparavant.

La comparaison des figures 3.17(a) et 3.17(b) montre un excellent accord entre les deux méthodes, avec quelques différences, notamment à $11^{\mathrm{h}}35$, $12^{\mathrm{h}}00$ et à $13^{\mathrm{h}}10$, où la position de la SWB s'étend au-delà de l'intervalle de confiance détecté avec la méthode Bayésienne. Il n'existe pas de calibre absolu pour déterminer ici la position exacte de la SWB. D'un point de vue géophysique, il est difficile de déterminer quelle frontière est la meilleure puisque l'association de la SWB avec les régions magnétosphériques requiert une comparaison détaillée avec les mesures satellites. De plus, différents effets (par exemple : champ électrique, turbulence, distribution des vitesses des inhomogénéités, effet de propagation...) peuvent affecter la localisation de la SWB (Baker et al., 1995; Moen et al., 2001; André et al., 2000a; Vallières et al., 2004). L'un d'eux est l'orienta-

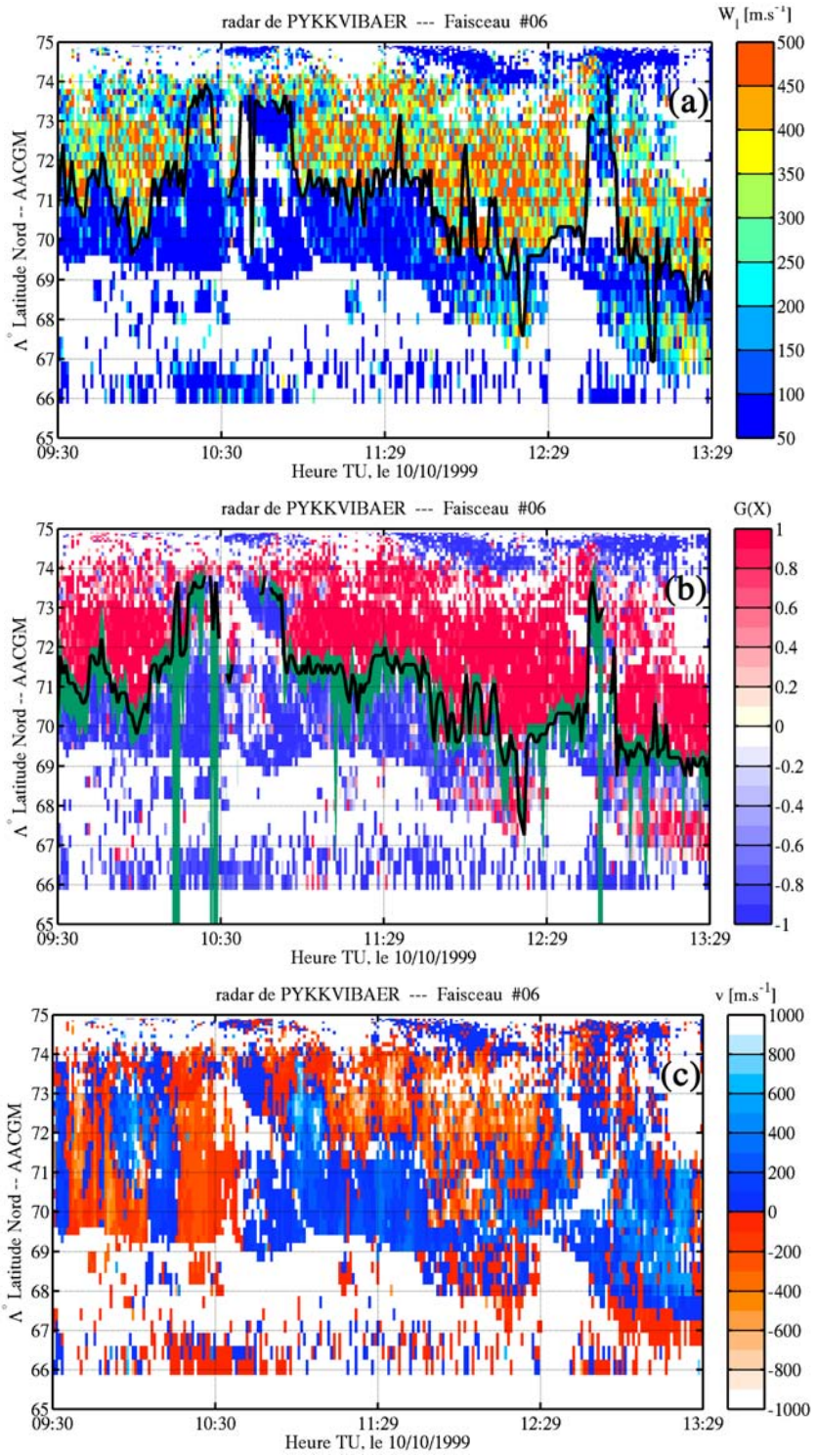
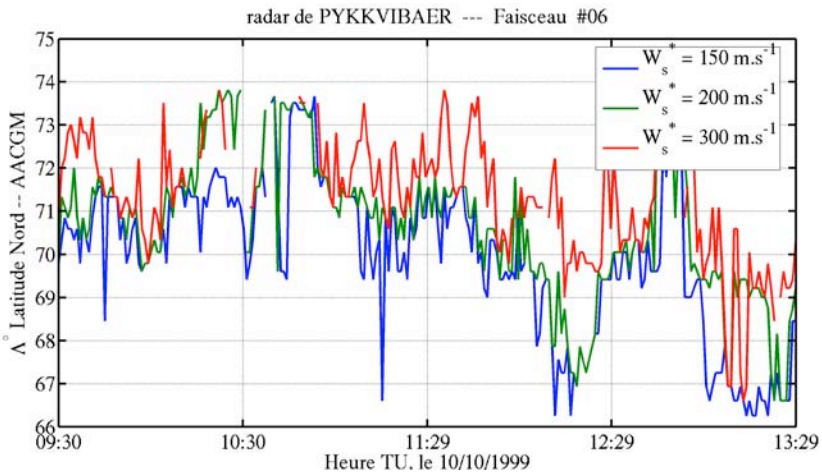


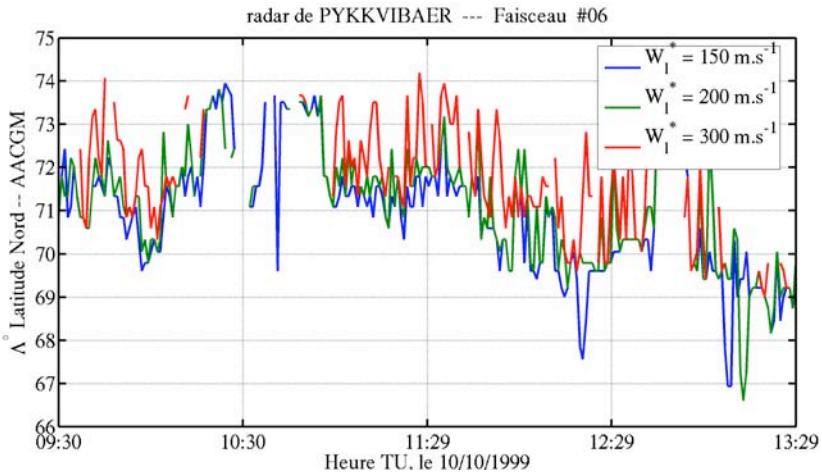
FIG. 3.17 – *Détermination de la SWB pour le 10 octobre 1999 entre* $09^{h}30$ - $13^{h}30$ *TU, observée par le faisceau #06 du radar de Þykkvibaer. Le* **panneau supérieur** *(a) montre la SWB (trait noir) détectée à partir d'une valeur seuil de la largeur spectrale lorenzienne* ($W^* = 150$ m.s^{-1})*. Le code de couleur se réfère à la valeur de la largeur spectrale. Le* **panneau central** *(b) identifie la SWB à partir de la fonction discriminante* $G(x)$*. Son intervalle de confiance est indiqué en vert. Le* **panneau inférieur** *(c) expose la projection de la vitesse Doppler.*

tion des champs de vue des radars. Les faisceaux non-méridiaux peuvent observer un accroissement de la largeur spectrale induit pas la CRB, qui n'est pas forcément relié à l'OCB (Lester et al., 2001; Villain et al., 2002; Chisham et al., 2005b). Le faisceau 6 exposé à la figure 3.17 est justement zonal. La projection de la vitesse Doppler sur ce même intervalle de temps, indique clairement que pour la période $11^{h}00$ - $12^{h}30$ TU, la SWB et la frontière Bayésienne sont co-localisées avec la ligne neutre de la cellule de convection. Avant $11^{h}00$ TU, la SWB et la frontière Bayésienne sont associées à la limite basse-latitude du cornet polaire. C'est un cas typique où deux régions distinctes peuvent amener une signature identique sur la largeur spectrale et de f_2/f_1. Pour les raisons que nous avons déjà évoqué, la méthode Bayésienne est préférable. De plus, l'introduction de nouvelles connaissances dans l'a priori permettrait de faire la distinction entre ces deux frontières. Néanmoins, nous pouvons citer les avantages de notre approche :

- Nous détectons la frontière à partir d'un traceur de la largeur spectrale, qui ne requiert pas la modélisation de l'ACF soit par une forme exponentielle soit par une forme gaussienne. En effet, l'utilisation de ces deux modèles peut entraîner une divergence sur la localisation de la SWB (*cf* les intervalles de temps TU : $09^{h}30$ - $09^{h}45$, $12^{h}00$ - $12^{h}30$ et $13^{h}00$ - $13^{h}30$ entre les panneaux (a) et (b) de la figure 3.18).
- L'approche Bayésienne court-circuite la tâche subjective de définir un seuil universel. Les panneaux (a) et (b) de la figure 3.18 montrent l'influence du seuil sur la position en latitude de la SWB pour les modèles respectifs gaussien et exponentiel. On remarque notamment que la SWB dérive vers le pôle lorsque l'on augmente le seuil. La variation du seuil peut entraîner une erreur de plusieurs degrés en latitude soit de plusieurs centaines de kilomètres. Cet effet est davantage marqué avec le modèle gaussien. A partir du critère Bayésien, une valeur unique est définie $x^* = 0.52$.
- La frontière Bayésienne souffre moins des fluctuations en latitude que la SWB puisque le critère Bayesien maximise la séparation entre les deux classes et en ce sens est mieux adapté pour traquer la frontière avec des données éparses. Même avec des données discontinues, la méthode quantifie la probabilité d'appartenance à chaque classe.
- Un sous-produit important est l'intervalle de confiance. Il permet la validation du résultat. Cet intervalle est relié aux échelles spatiales du gradient des largeurs spectrales. A cause de la résolution spatiale de SuperDARN et du bruit ambiant, la nature de la transition (forme et amplitude) entraîne une incertitude, d'où la nécessité de localiser une bande en latitude dans laquelle se projette l'OCB et non la SWB uniquement.



(a) modèle gaussien



(b) modèle exponentiel

FIG. 3.18 – *Identification de la SWB à partir des modèles gaussien (a) et exponentiel (b). Trois valeurs seuils ont été appliquées aux deux modèles :* $W^* = 150$ m.s^{-1} *(courbe bleue),* $W^* = 200$ m.s^{-1} *(courbe verte) et* $W^* = 300$ m.s^{-1} *(courbe rouge).*

D'autres événements (~ 10) ont été étudiés de la même façon. Les résultats montrent une excellente corrélation des deux approches.

3.4.3 Perspectives : aspect dynamique et amélioration

Nous nous sommes jusqu'ici contentés d'utiliser le formalisme Bayésien sans injecter d'information supplémentaire dans le système. Autrement dit, nous avons utilisé ici

un a priori non-informatif. En première intention, cette hypothèse est nécessaire pour approcher les vraisemblances $p(x|\omega_j)$, car il n'y a pas de raison de donner plus de poids à l'une ou à l'autre des deux classes.

Un des intérêts majeurs de l'approche probabiliste est sa capacité de fusionner différentes sources d'informations. Par exemple, l'a priori se décompose comme le produit

$$P(\omega_j, \mathrm{I}) = \prod_{k=1}^{N} P(\omega_j, \mathrm{I}_k) \ , \tag{3.11}$$

où I_k représente la k-ième source d'information. Une notation plus juste devrait être $\mathrm{I}_k(\Lambda, \Phi, t)$, en ce sens que notre connaissance de la probabilité a priori doit évoluer avec l'état du système magnétosphère-ionosphère et doit également dépendre des variables d'espace : la latitude invariante magnétique (Λ) et le secteur horaire magnétique local (Φ). On peut dès lors envisager un schéma d'intégration à plusieurs entrées, permettant d'introduire des informations supplémentaires pour actualiser notre connaissance a priori des deux classes. Voici quelques pistes :

- La fréquence d'apparition de la SWB côté jour est fortement liée avec l'orientation vers le sud du champ magnétique. Lorsque que la composante B_z du champ magnétique est orientée vers le nord, la quantité d'énergie (en termes de précipitation de particules) entrant dans la magnétosphère n'est pas suffisante pour que la SWB puisse être observée.
- La prise en compte de l'évolution temporelle de la distribution spatiale des probabilités a posteriori doit être également envisagée. Par exemple, dans le cas où la variation en latitude de la SWB sur deux des intervalles successifs est trop rapide la probabilité d'observer une frontière géophysique (OCB) à partir de la SWB devrait être plus faible. Cette probabilité d'observer la signature ionosphérique d'une région magnétosphérique dépend également du secteur MLT.
- L'introduction d'un modèle de l'ovale auroral peut contraindre le modèle. La probabilité d'observer l'OCB diminue lorsque la distance entre la frontière polaire de l'ovale auroral et la SWB augmente.

D'autres améliorations sont possibles, notamment avec le rapport f_3/f_1. Nous avons jusqu'ici basé la localisation de la frontière sur la variation de la largeur spectrale et sur le lien direct de celle-ci avec le rapport f_2/f_1. Or, on peut maintenant envisager de s'affranchir de cette étape intermédiaire, pour chercher à localiser la frontière directement à partir des rapports f_2/f_1 et f_3/f_1. En effet, la combinaison entre le rapport f_3/f_1 le rapport f_2/f_1 permet de déduire empiriquement un facteur de forme, ce qui

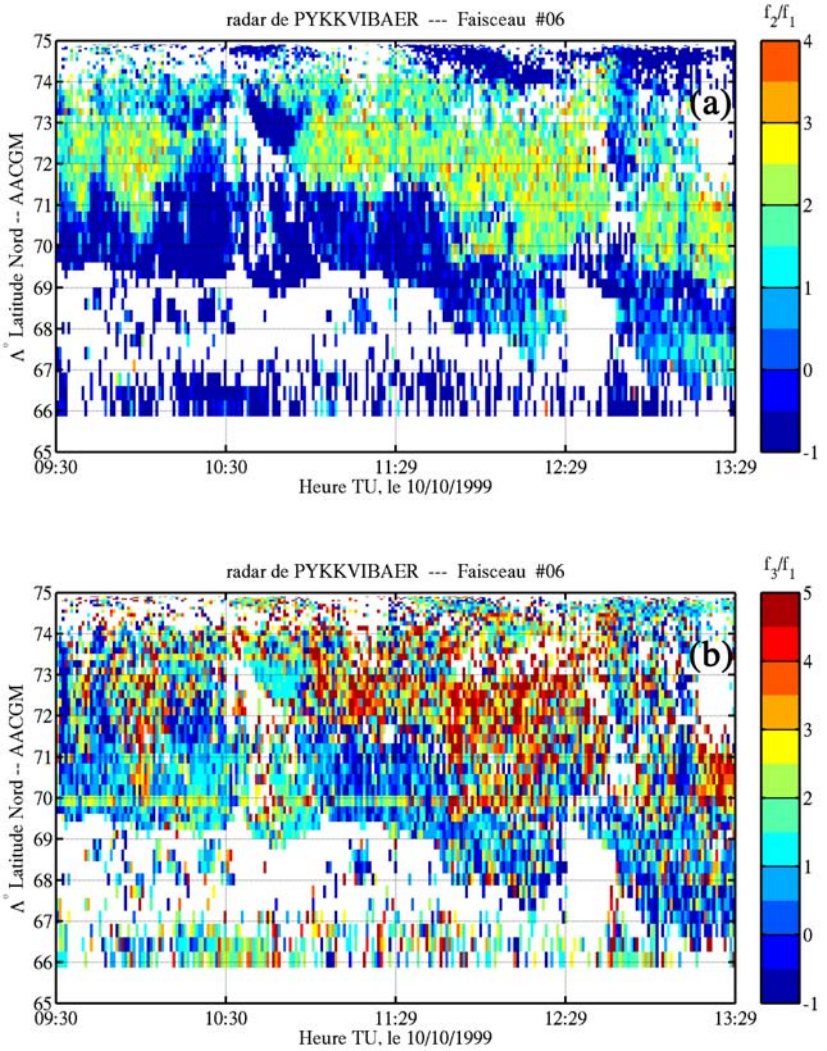


FIG. 3.19 – *Evolution temporelle des rapports adimensionnels f_2/f_1 (a), f_3/f_1 (b) obtenus par le faisceau #6 du radar de Þykkvibaer, entre* 09^h30 *et* 13^h30 *TU.*

contribue à une meilleure description des ACFs. Son étude pourrait améliorer la description des deux classes de la largeur spectrale. Comme le suggère la comparaison des figures 3.19(a) et 3.19(b), on peut noter un comportement différent du rapport f_3/f_1 de part et d'autre de la frontière. Au Nord de la SWB, on distingue un régime spécifique du rapport f_3/f_1 pour lequel, bien qu'il fluctue davantage que le rapport f_2/f_1, est principalement supérieur à 3.5. Au Sud de la SWB, deux régimes spécifiques peuvent être identifiés. Avant 11^h00, on remarque que f_3/f_1 est principalement positif, tandis qu'après 11^h00, là où les vitesses Doppler détectées sont essentiellement positives, f_3/f_1 est plutôt négatif. On note également la présence, sur la projection du paramètres

f_3/f_1, de plusieurs bandes (70°, 72° et 73.5° Λ) qui n'apparaissent pas sur la projection des autres paramètres : f_2/f_1, largeur spectrale (W_l) et la vitesse Doppler (V_D). Ces bandes peuvent être associées à des ACFs ou résident une ambiguïté sur l'estimation de la distance.

L'intégration de nouvelles informations (mesures satellites ou sol, modèles, etc ...) dans l'a priori nécessite la mise en place d'une statistique importante regroupant plusieurs années d'observations. Nous avons évoqué la nécessité de prendre en compte l'évolution temporelle du système magnétosphère-ionosphère. L'aspect dynamique peut être pris en compte dans la description Bayésienne puisqu'on peut estimer par récurrence les probabilités a posteriori. Pour cela, il n'est pas nécessaire de recalculer les fonctions de densité de probabilité à chaque mise à jour des connaissances. On peut déduire de la probabilité a posteriori au temps t, la probabilité a posteriori au temps $t + \delta t$ en ajoutant de nouvelles mesures. Le fonctionnement par récurrence de la méthode Bayésienne est tout à fait adapté à une modélisation en temps réel.

3.5 Conclusion

Dans ce chapitre, nous nous sommes attachés à décrire l'information sous-jacente aux trois modes SVD en termes de quantité physique. Tout d'abord en comparant les rapports adimensionnels f_2/f_1 et f_3/f_1 avec les paramètres calculés par l'algorithme FitACF, nous avons démontré que l'un d'eux, le rapport f_2/f_1 s'avère être un bon estimateur de la largeur spectrale. Le second rapport f_3/f_1 décrit plutôt la forme du spectre, même si la relation est loin d'être linéaire. Une confrontation des modes avec un modèle analytique, en ce sens qu'il reproduit toutes les formes d'ACFs observées, valide ces résultats. Le rapport f_2/f_1 offre une description unique de la largeur spectrale, *i.e.* f_2/f_1 dépend peu de la forme des ACFs, ce qui n'est pas le cas de la largeur spectrale estimée avec l'algorithme FitACF.

L'étude géophysique a permis de confirmer la bonne corrélation de la largeur spectrale entre le modèle SVD et les modèles gaussien et exponentiel. L'association des paramètres f_2/f_1 avec f_3/f_1 apparaît comme une valeur ajoutée pour la description des échos radar. Il permet une description supplémentaire sur la forme du spectre rétrodiffusé. Nous avons remarqué à partir de ces deux rapports une transition du côté matin entre les formes exponentielles et intermédiaires associée à la frontière de la largeur spectrale. Cependant, la difficulté principale réside dans la discrimination entre les signatures caractéristiques des régions frontières magnétosphériques et celles d'événements singuliers. L'application d'un critère Bayésien basé sur le traceur de la largeur

spectrale fournit une approche pertinente pour la détection de ces frontières. En effet, la largeur spectrale seule ne permet pas d'identifier correctement les régions frontières de la magnétosphère, dans le sens où différentes sources possibles amènent une même signature de la largeur spectrale. L'approche Bayésienne offre une base formelle tout à fait adaptée à ce problème.

Cette étude constitue donc la première phase dans l'élaboration d'une détection opérationnelle des régions magnétosphériques. Pour cela, nous sommes restés le plus proche possible de la mesure brute, afin de contourner les modèles physiques, qui peuvent introduire un biais (Ponomarenko and Waters, 2006). La difficulté de définir un seuil pour l'identification de la SWB est contournée par l'introduction d'un critère Bayésien comme un classifieur auto-calibré, de plus mieux adapté à l'analyse en temps réel.

L'aspect dynamique de l'approche Bayésienne est important. L'apprentissage d'un a priori sur une grille MLT/Λ via l'ajout de nouvelles connaissances, venant par exemple de différents instruments (satellites, radars incohérents, caméra plein ciel, etc...) ou encore de modèles est à envisager. Cela permettrait de faire évoluer notre a priori du système et par conséquent de contraindre la détection des régions frontières magnétosphériques.

4

FitSVD : un nouvel algorithme d'analyse des mesures SuperDARN

4.1 Schéma d'analyse des ACFs : FitSVD 116
4.2 Mise en place de la procédure 118
4.2.1 Détermination des seuils 118
4.2.2 Exemples d'application 120
4.3 Comparaison avec l'algorithme FitACF 122
4.3.1 Nature des retards aberrants détectés 122
4.3.2 Influence sur l'estimation des paramètres : $\{f_2/f_1, f_3/f_1\}$ et $\{\tau_e, n_\tau\}$ 125
4.4 Validation de la procédure 126
4.4.1 Validation d'un critère d'arrêt 127
4.4.2 Performance : détection des retards erronés 130
4.5 Conclusion 131

L'analyse d'une mesure, d'un signal, requiert un pré-traitement afin d'en extraire la partie cohérente des sources de bruit susceptibles d'affecter les paramètres du modèle. L'élaboration d'un algorithme opérationnel nécessite l'automatisation de ces tâches de pré-traitement.

Dans le cas des radars SuperDARN, la mesure brute est une ACF définie sur 18 retards. Lors de la mesure, différentes sources d'erreurs, comme la présence dans l'ionosphère d'un bruit fond ou encore d'interférence, viennent perturber la reconstruction des ACFs. Ces effets se traduisent par des retards qui peuvent parfois s'écarter d'un ordre de grandeur par rapport à la valeur escomptée. La présence dans l'ACF de retards

erronés perturbe l'ajustement des modèles et ainsi biaise l'estimation de ces paramètres. Il est donc nécessaire de les détecter et de les éliminer au préalable. Cette tâche de détection et d'ajustement des ACFs est réalisée directement sur chacun des sites radars par l'intermédiaire de l'algorithme FitACF. Ce dernier utilise des critères empiriques afin d'identifier et d'éliminer ces mauvais retards avant d'estimer la vitesse Doppler, la largeur spectrale et la puissance rétrodiffusée. Le principe de reconstruction des ACFs ainsi qu'une description de l'algorithme FitACF sont abordés dans l'Annexe A. Notons cependant le caractère ambigu d'une telle détection. Si certaines causes instrumentales sont bien identifiées, il est probable qu'un nombre non négligeable de valeurs dites aberrantes soient dues à des spectres multi-composantes et revêtent dans ce sens un véritable sens physique.

Dans ce chapitre, nous allons étudier les nouvelles perspectives qu'offre la description SVD en termes de détection de retards erronés. Nous mettrons en place une nouvelle procédure «FitSVD» permettant d'éliminer les mesures erronées et d'ajuster fiablement les paramètres statistiques, f_α, en tirant profit du fait que la variabilité des ACFs peut être reproduite par une combinaison linéaire des trois modes $g_\alpha(\tau)$ décrits dans les chapitres précédents. Nous validerons cette approche en confrontant ses performances avec celles de l'algorithme FitACF.

4.1 Schéma d'analyse des ACFs : FitSVD

Dans l'optique de développer un modèle opérationnel, il est nécessaire de mettre en oeuvre une procédure adaptée et robuste, permettant une estimation fiable des paramètres f_α même en présence de «mauvais retards» ou mieux encore, d'identifier et de corriger les mauvais retards. Pour cela, une étape supplémentaire d'analyse de l'ACF est nécessaire pour valider les f_α.

En comparaison avec l'algorithme FitACF, nous abordons le problème de la détection des mauvais retards d'une manière différente. En effet, l'algorithme FitACF évalue des critères empiriques afin de prédire l'impact des différentes sources de bruit sur les ACFs et ainsi de détecter les retards affectés. L'évaluation des niveaux de fluctuations est en soi une tâche délicate, ce qui rend faillible l'algorithme FitACF en matière de détection. Comme nous l'avons montré dans le chapitre 2, la combinaison des trois modes statistiques, $g_\alpha(\tau)$, permet de reproduire la décroissance moyenne du module des ACFs. Les retards qui s'écartent au-delà d'une certaine valeur seuil du modèle SVD peuvent donc être facilement détectés. Notre hypothèse de base est la suivante : toute valeur qui s'écarte de manière significative de l'ACF, telle qu'elle est reconstruite à partir des trois

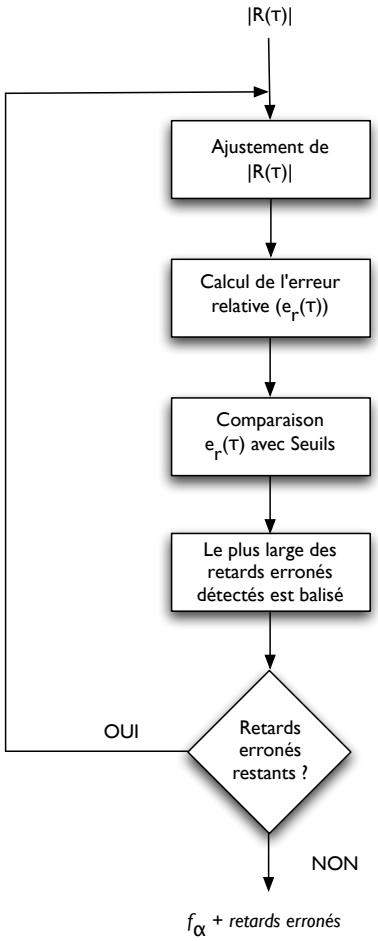


FIG. 4.1 – *Synopsis de l'algorithme FitSVD.*

modes, est considérée comme erronée. Cela revient donc à supposer que toute ACF non erronée est une fonction régulière de τ, qui se laisse bien décrire par une combinaison linéaire des trois modes. Nous excluons donc toutes les ACFs multimodales, quelle que soit leur origine (instrumentale, physique,...), même si leurs propriétés peuvent être pertinentes pour étudier notamment la microphysique. Ces ACFs sont en effet difficilement ajustables par un modèle simple. Il reste alors à définir le seuil ainsi qu'à spécifier la procédure utilisée pour reconstruire l'ACF à partir des données bruitées. Les propriétés du modèle SVD permet donc de procéder par itération (voir Fig. 4.1) :

1. Nous ajustons l'ensemble de l'ACF brute à partir des trois modes de la SVD.
2. Nous comparons l'écart entre le modèle SVD et la puissance mesurée pour chacun des 18 retards.
3. Les retards qui s'écartent au-delà d'une valeur seuil sont balisés comme aberrants.

Parmi ces retards aberrants, celui qui possède l'écart le plus important est éliminé.

4. L'ACF est ré-ajustée à partir des 17 retards restants.
5. La procédure de détection est à nouveau relancée avec la nouvelle combinaison de modes obtenue.

L'opération est réitérée jusqu'à ce qu'il n'y ait plus de retard aberrant dans l'ACF. Pour cela, nous devons définir un critère d'arrêt.

Dans un souci de différencier les retards erronés détectés par l'algorithme FitACF de ceux détectés par la méthode présentée ci-dessus, nous désignerons par le terme «aberrants» ces derniers et nous conserverons le terme «mauvais retards» pour ceux identifiés par l'algorithme FitACF.

La définition des valeurs seuils et du critère d'arrêt seront l'objet des paragraphes suivants.

4.2 Mise en place de la procédure

4.2.1 Détermination des seuils

L'objectif est de définir une valeur seuil pour détecter les retards aberrants qui biaisent le plus l'estimation des paramètres. La présence de bruit et de différentes sources de mauvais retards (FitACF) affecte différemment la puissance les retards. Ceci implique que la valeur de ces seuils va varier avec τ. Les différentes sources d'erreur et leur influence sur la résolution des ACFs sont discutées en détail au § A.2 de l'annexe. Pour évaluer ces seuils, nous étudions le comportement des erreurs indépendamment sur chacun des 18 retards. L'algorithme FitACF, malgré ses défauts inhérents, est employé afin de différencier les erreurs commises par le modèle SVD entre les bons et les mauvais retards.

Nous sélectionnons 2.10^6 ACFs à partir des observations du 20 décembre 2003 sur l'ensemble de l'hémisphère nord, sans aucune contrainte particulière. A partir des trois modes SVD, nous les avons ajustées afin de reconstruire la distribution des erreurs relatives commises sur chacun des 18 retards :

$$\epsilon_r(\tau) = \frac{|R(\tau)| - |\hat{R}(\tau)|}{|R(\tau)|} \ , \tag{4.1}$$

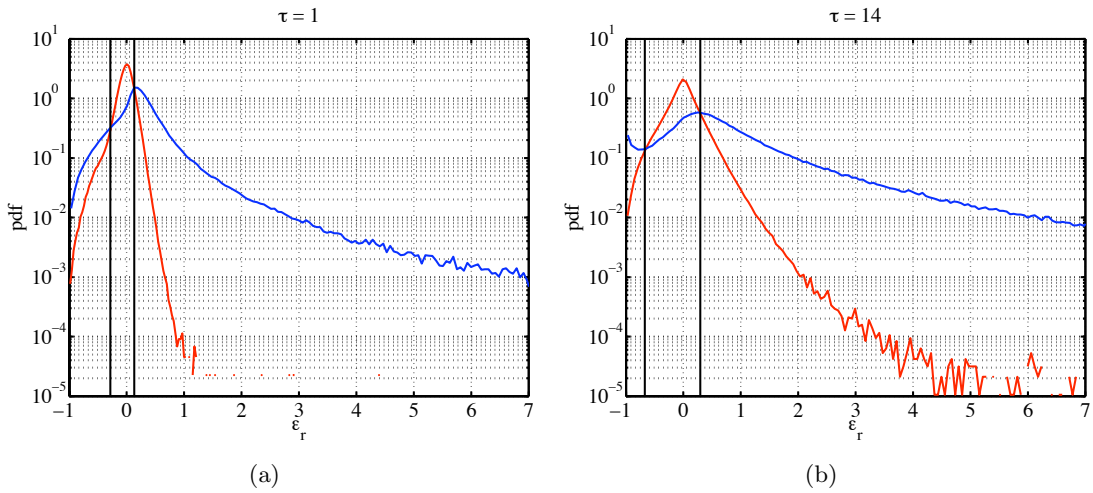


FIG. 4.2 – *Fonction de densité de probabilité des écarts relatifs ϵ_r au modèle issu de la SVD pour les retards $\tau = 1$ (a) et $\tau = 14$ (b). La courbe bleue montre l'erreur relative correspondant aux mesures balisées comme «mauvaises» par l'algorithme FitACF et en rouge, les bonnes mesures. Les lignes noires verticales fixent les limites séparant les valeurs acceptables des autres.*

où $|R(\tau)|$ et $|\hat{R}(\tau)|$ désignent respectivement le module de l'ACF mesuré et modélisé par SVD au retard τ.

Les panneaux (a) et (b) de la figure 4.2 montrent les densités de probabilité (PDF[1]) des erreurs pour les retards $\tau = 1$ et $\tau = 14$. Les courbes rouges désignent la distribution des écarts relatifs au modèle pour les bons retards et celles en bleu la même distribution pour les retards identifiés comme mauvais par l'algorithme FitACF. Ces deux exemples confirment que les mauvais retards donnent lieu à des erreurs bien plus élevées. Nous notons également que la queue des distributions augmente graduellement avec la valeur du retard τ. Ceci résulte de la présence de bruits et d'interférences qui affectent davantage les grands retards. La superposition des deux courbes suggère qu'un nombre important de retards pour lesquels l'écart aux modèles est important ne sont pas signalés comme mauvais par l'algorithme FitACF. Ce dernier constitue ce qu'il y a de mieux pour identifier les valeurs erronées, mais ne saurait être considéré comme une référence absolue pour les détecter.

Pour chaque retard, nous prenons l'intersection entre les deux densités de probabilité, de part et d'autre du maximum de la courbe rouge, pour définir deux valeurs seuils de l'erreur. Nous aurions pu choisir un autre critère, par exemple, en prenant la probabilité que l'erreur sur les bons retards soit égale à la moitié de la probabilité de

[1]Probability Density Function

l'erreur sur les mauvais retards. Ce choix permettrait d'être plus conservatif en termes de sélection des retards valides, mais il aurait pour conséquence d'augmenter les erreurs lors de l'ajustement des paramètres. Les seuils sont indiqués par un trait vertical noir sur la figure 4.2. En définissant les seuils par l'intersection des deux distributions (*cf* traits verticaux noirs de la figure 4.2), nous supposons implicitement qu'en deçà de cet intervalle, la contribution des sources de bruit et d'interférences sont statistiquement négligeables et donc n'influence que faiblement l'estimation des paramètres. Autrement dit, les deux valeurs seuils définies pour chacun des 18 retards, fixent les limites inférieure et supérieure entre les mesures valides et celles erronées. Notons également sur cette figure, que la valeur de ces seuils augmentent avec le retard τ.

En rouge, la largeur de la distribution des erreurs sur les retards valides donnés par FitACF (qui peut se tromper) caractérise les fluctuations aléatoires des centres diffuseurs dans les volumes sondés, et par conséquent, elle est reliée au niveau de bruit. La définition de ces seuils dépend alors du niveau de bruit contenu dans l'échantillon. Nous avons étudié la dépendance de ces seuils selon différents niveaux de bruit, de 4 à 60 dB. Notre choix s'est porté sur les seuils les plus larges, *i.e.* ceux qui ont été obtenus à partir de l'échantillon d'ACFs pour lequel le rapport signal/bruit était le plus faible. Nous avons opté pour cette solution afin d'être le moins contraignant possible au niveau de la sélection des retards valides.

Les retards éliminés par l'algorithme FitACF ou par le modèle SVD sont différents par la nature des deux approches. Nous éliminons ainsi qualitativement les puissances qui affectent le plus le modèle SVD, sans nous préoccuper des sources éventuelles d'interférences.

4.2.2 Exemples d'application

La figure 4.3 montre six exemples de détection de retards aberrants. Les retards aberrants détectés par la méthode FitSVD sont indiqués par des points rouges. Les mauvais retards détectés par l'algorithme FitACF sont superposés (diamants verts). Comme le suggèrent ces exemples, la technique du seuillage permet uniquement d'éliminer les retards qui s'éloignent le plus de la décroissance moyenne de l'ACF. Ainsi, les mauvais retards identifiés par FitACF mais qui sont proches de l'ACF brute, ne sont pas détectés avec notre technique. Notons l'exemple de gauche, dernière ligne, ou les deux méthodes sont complètement en désaccord. C'est un cas typique où l'ACF brute est trop dégradée et ne permet pas une détection fiable (même de manière subjective) des puissances erronées. Notons également que l'ACF ajustée dans la figure à droite au

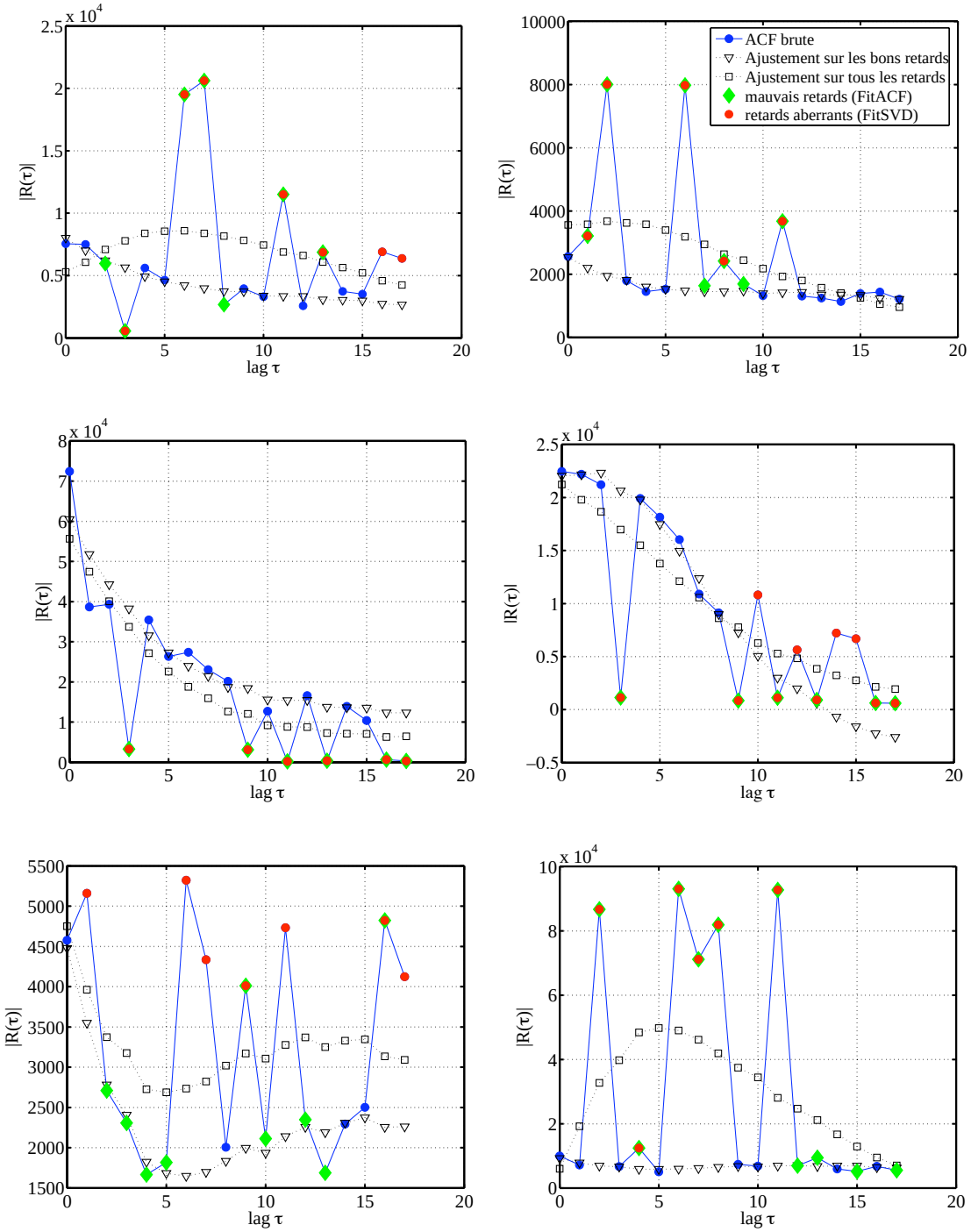


FIG. 4.3 – *Six exemples d'identification de retards anormaux. Les points bleus montrent l'ACF brute. Les diamants verts indiquent la position des mauvais retards détectés avec l'algorithme FitACF et les retards marqués d'un point rouge montrent ceux identifiés par la procédure FitSVD. Les carré montre la première itération de la procédure et les triangles inversés montrent le résultat final des ACFs.*

milieu est négative. Ceci est un inconvénient de la méthode SVD dont les fonctions ne sont pas définies positives.

A la vue de la différence entre la première et la dernière itération, il est évident que l'ajustement des ACFs à partir des trois modes de la SVD nécessite un pré-traitement. L'ajustement des trois modes aux données brutes est largement influencé par la présence de puissances erronées et par conséquent l'estimation des paramètres f_α l'est également. Pour les exemples de la figure 4.3, les rapports f_2/f_1 et f_3/f_1 peuvent varier d'un facteur 2 entre la première itération et la dernière, ce qui est loin d'être négligeable. Il faut donc trouver un compromis entre la qualité de l'ajustement des ACFs et celle de l'estimation des paramètres et ce de manière robuste. Il est donc nécessaire d'introduire un critère d'arrêt. L'élaboration de ce critère d'arrêt sera discuté au § 4.4.1.

4.3 Comparaison avec l'algorithme FitACF

Nous comparons dans cette section, les performances de la stratégie employée par l'algorithme FitACF avec celle basée sur les propriétés de la SVD. L'objectif est d'identifier les catégories de mauvais retards détectées par la méthode FitSVD et d'en étudier l'impact sur l'estimation des modes $\{f_2/f_1, f_3/f_1\}$ et des paramètres $\{\tau_e, n_\tau\}$ du modèle analytique (*cf* Eq. 3.3). Pour cela, nous nous basons sur le même échantillon d'ACF que celui de la section § 4.2.1.

4.3.1 Nature des retards aberrants détectés

Pour l'échantillon de la section 4.2.1, l'algorithme FitACF détecte 3 599 044 retards aberrants contre 3 030 061 avec la procédure FitSVD (*cf* § 4.2.2). Parmi tous les mauvais retards identifiés, 46.5 % sont détectés par les deux méthodes, 33 % le sont uniquement par FitACF et 20.5 % par la méthode SVD (*cf* figure 4.4).

Sur l'ensemble des 45.6 % retards détectés par les deux techniques, la méthode SVD identifie les différentes catégories définies dans l'algorithme FitACF (*cf* § A.3.2 de l'annexe A) avec plus ou moins de réussite. Le tableau 4.1 récapitule pour chacune des catégories, le nombre de retards détectés par les deux méthodes, ainsi que le taux de réussite de la méthode SVD. Ce taux de réussite, qui est exprimé en pourcentage, est calculé à partir du rapport entre le nombre de retards détectés avec l'algorithme FitSVD sur le nombre de retards détectés avec l'algorithme FitACF. Nous notons ainsi que la technique de seuillage détecte préférentiellement les retards affectés par une

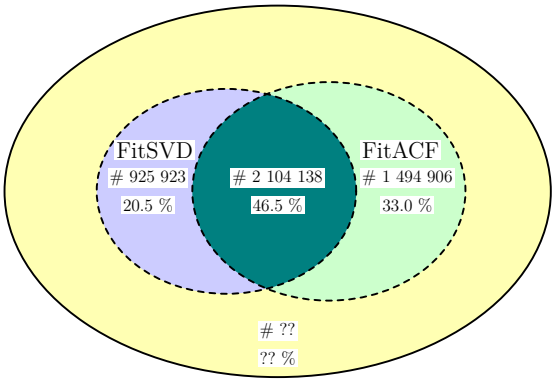


FIG. 4.4 – *Bilan des retards erronés détectés à partir de l'algorithme FitACF (zone verte) et à partir de la méthode SVD (zone bleue). La région en jaune définit l'ensemble des retards affectés par du bruit qui sont pas détectés par aucune des deux méthodes.*

superposition de signaux provenant de différentes portes (interférences inter-portes ou CRIs[2]), catégorie 1, et ceux dont la puissance après soustraction du bruit est négative, catégorie 3 (*cf* tableau A.1), avec un taux de réussite respectif de 75 % et de 61 %. Le taux de réussite baisse à 35% et 25% pour les retards assimilés selon l'algorithme FitACF, à des pics positifs (catégorie 5) ou négatifs (catégorie 9). Selon les critères définis à partir de la méthode SVD, nous montrons statistiquement que les CRIs et les retards dont la puissance est négative (cat. 1 et 3) impactent d'une manière non négligeable sur l'ajustement des modèles. Ce résultat peut être interprété d'un point de vue historique. Depuis vingt ans, l'algorithme a subi de nombreuses améliorations. Les catégories 5, 7 et 9 ont été définies ultérieurement, elles correspondent à des corrections d'ordre secondaire.

[2]Cross-Range Interference

TAB. 4.1 – *Répartition des retards erronés par la méthode SVD.*

	Catégories des mauvais retards (FitACF)				
	1	3	5	7	9
Nombre de retards détectés par FitACF	2 035 847	186 888	1 228 509	52 319	95 481
Nombre de retards détectés par SVD	1 516 681	113 110	426 054	24 182	24 111
Pourcentage détecté par la méthode SVD (%)	75	61	35	46	25

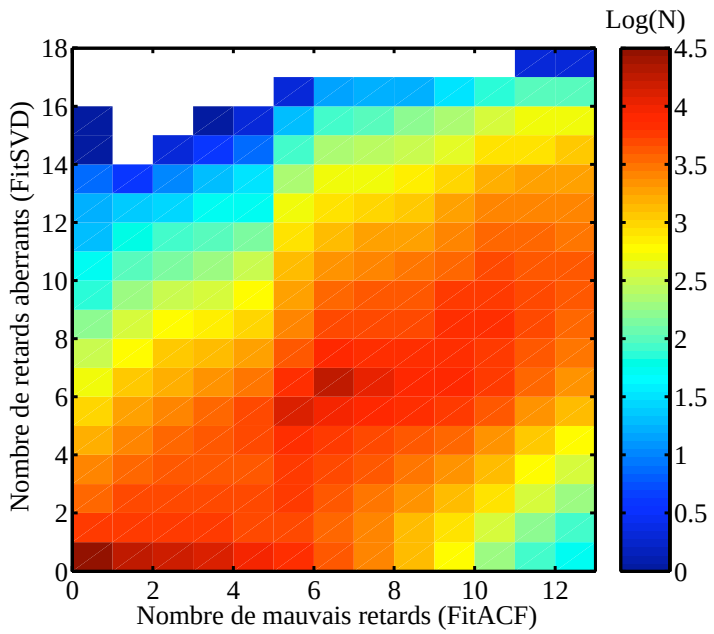


FIG. 4.5 – *Histogramme bi-varié du nombre d'ACFs en fonction du nombre de retards aberrants détectés par ACF avec l'algorithme FitACF et ceux détectés avec la technique du seuillage appliquée au modèle SVD.*

La figure 4.5 montre la distribution des échos radar en fonction du nombre de mauvais retards par ACF (identifiés avec FitACF) et en fonction du nombre de retards aberrants par ACF (identifiés avec la méthode SVD). Bien que l'approche des deux méthodes soit foncièrement différente, le nombre de retards aberrants détectés par ACF est bien corrélé. En accord avec les remarques précédentes, on remarque que la méthode SVD par rapport à l'algorithme FitACF sous-estime le nombre de retards aberrants contenus dans chacune des ACFs. Les retards qui sont uniquement détectés par l'algorithme FitACF correspondent à des retards qui, visuellement, n'ont pas l'air aberrants.

L'utilisation de la méthode SVD couplée avec des seuils *ad-hoc* s'appuie sur un critère heuristique. Elle ne permet pas de détecter les retards possédant une erreur faible. Ainsi 33 % des mauvais retards détectés par l'algorithme FitACF ne le sont pas avec cette technique, tandis que 20.5 % des retards aberrants possédant une erreur importante ne sont pas identifiés par l'algorithme FitACF. Ceci nous amène naturellement à étudier l'impact des différences qui existent entre les deux approches sur l'estimation des paramètres.

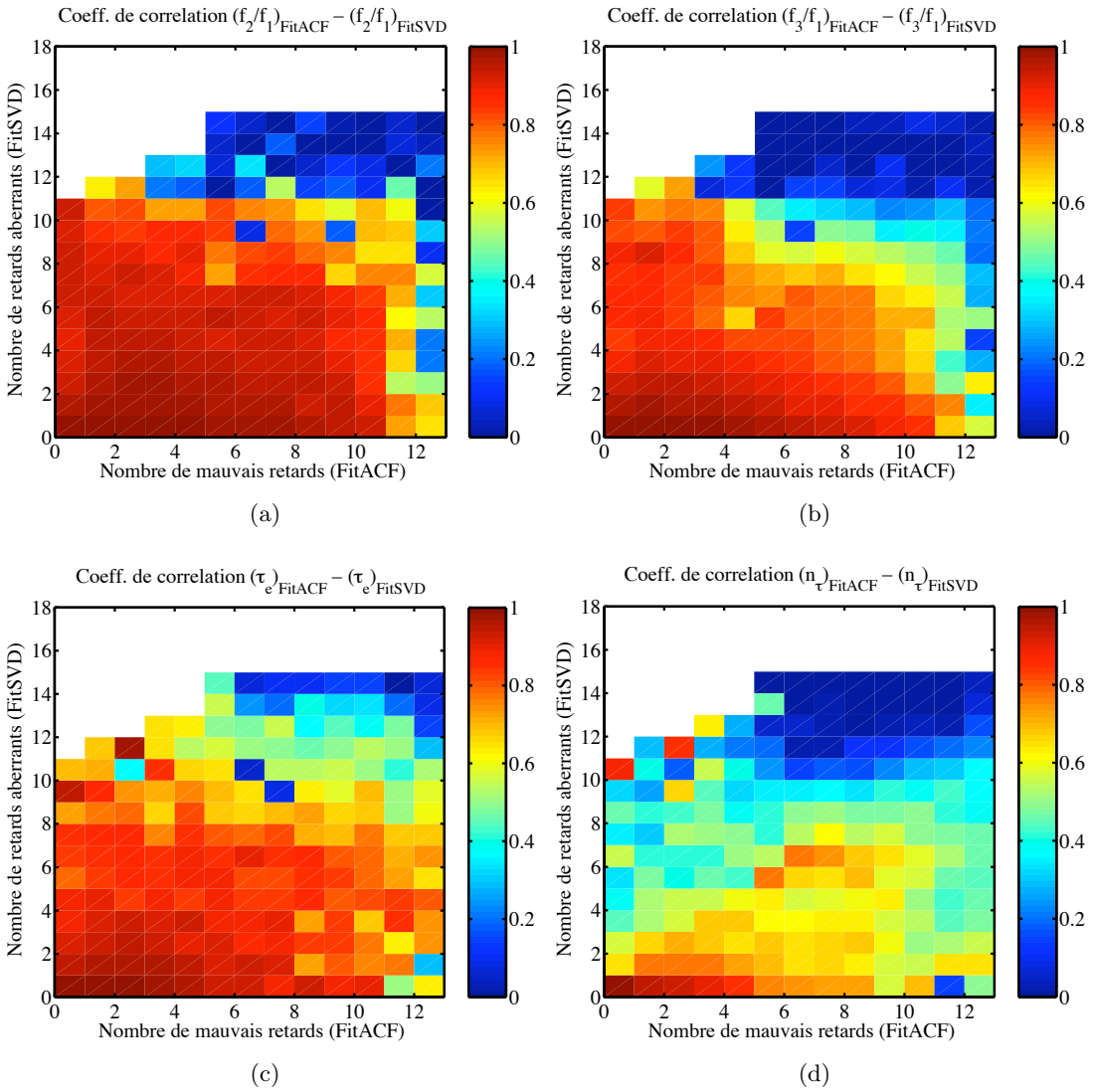


FIG. 4.6 – *Projection dans le plan défini par le nombre de mauvais retards détectés, par ACF, avec l'algorithme FitACF (bad lag) et avec la procédure FitSVD (outlier), du coefficient de corrélation des paramètres f_2/f_1 (a), f_3/f_1 (b), τ_e (c) et de n_τ (d).*

4.3.2 Influence sur l'estimation des paramètres : $\{f_2/f_1, f_3/f_1\}$ et $\{\tau_e, n_\tau\}$

Nous étudions l'influence de ce nouveau schéma d'identification des retards aberrants sur l'estimation des paramètres $\{f_2/f_1, f_3/f_1\}$ et sur les paramètres $\{\tau_e, n_\tau\}$ issus du modèle analytique utilisé à la section § 3.1.2 (*cf* Eq. (3.3)). Pour cela, nous ajustons les ACFs avec ces deux modèles à partir de la procédure FitSVD et ensuite à partir de l'algorithme FitACF. Ainsi pour chacune des ACFs, nous obtenons, pour chaque

modèle, deux jeux de paramètres. Nous classons ensuite l'ensemble des paramètres calculés selon le nombre de retards aberrants (FitSVD) et de mauvais retards (FitACF) détectés par ACF.

La figure 4.6 montre l'évolution du coefficient de corrélation des paramètres f_2/f_1 (a), f_3/f_1 (b), τ_e (c) et n_τ (d) estimés à partir des ACFs «nettoyées» par les procédures FitACF et FitSVD. Nous traçons cette corrélation en fonction du nombre de retards erronés détectés par les deux méthodes. Cette représentation permet de voir le domaine commun aux deux approches. Pour le paramètre f_2/f_1, on remarque que la corrélation est supérieure à 0.8 pour un nombre de retards aberrants inférieur à 9 pour FitSVD et inférieur à 10 pour FitACF. Au-delà de 10 retards erronés par ACF détectés avec les deux méthodes, la corrélation du rapport f_2/f_1 chute dramatiquement. Le rapport f_3/f_1 est plus sensible à la méthode d'identification. A partir de 4 mauvais retards, la correlation est plus faible lorsque la méthode FitSVD détecte davantage de valeurs erronées que l'algorithme FitACF.

Le modèle analytique, ajusté par une méthode de gradient, est plus sensible aux deux schémas d'identification. Le paramètre τ_e évolue de la même manière que les rapports f_2/f_1 et f_3/f_1, mais avec une corrélation qui est globalement plus faible. Le paramètre n_τ est le plus affecté. La corrélation devient nulle systématiquement lorsque la méthode FitSVD détecte plus de mesures erronées que l'algorithme FitACF. L'estimation du facteur de forme, n_τ, et du temps de corrélation, τ_e, est donc particulièrement sensible à la méthode de détection. Ceci confirme la robustesse de la méthode SVD et nous incite une fois de plus à la préférer pour caractériser les ACFs.

Ces résultats permettent de mettre en évidence que la méthode de détection employée, pour peu que le nombre de valeurs erronées présent dans l'ACF reste faible ($\leqslant 10$), n'a pas de réel impact sur l'estimation des paramètres f_2/f_1 et f_3/f_1, ce qui n'est pas le cas du modèle analytique (3.3).

4.4 Validation de la procédure

Comme nous venons de le montrer, la corrélation des paramètres $\{f_2/f_1, f_3/f_1\}$ entre deux pré-traitements des ACFs chute dramatiquement pour un nombre de retards aberrants supérieur à 10. Il faut alors trouver un compromis entre le nombre suffisant de retards à éliminer pour ne pas être biaisé par les valeurs anormales, et le nombre suffisant de retards à conserver pour pouvoir estimer correctement les paramètres de l'ACF. Afin d'améliorer la robustesse de la méthode, il est nécessaire d'introduire un critère d'arrêt.

4.4.1 Validation d'un critère d'arrêt

Pour définir un critère d'arrêt, nous avons reconstruit 1 000 ACFs synthétiques à partir des modes de la SVD que nous avons volontairement dégradées en introduisant un bruit blanc gaussien, en faisant varier le niveau de fluctuation entre 10 dB et 30 dB. Nous avons ensuite introduit manuellement des mauvais retards, entre 4 et 12 retards erronés par ACF. La position de ces retards est choisie de façon aléatoire sur les 18 retards et leur amplitude est simulée à partir des distributions des écarts relatifs des mauvais retards identifiés par FitACF (*cf* Fig. 4.2). Nous contrôlons uniquement le niveau de fluctuations et le nombre de retards aberrants. Notons que cette procédure ne permet pas de reproduire exactement le bruit détecté par les radars SuperDARN. Pour un rapport signal/bruit donné, nous surévaluons l'effet du bruit par rapport aux vraies mesures. La façon dont nous générons le bruit peut amener des puissances négatives dans l'ACF. Dans ce cas, ces retards sont mis à zéro.

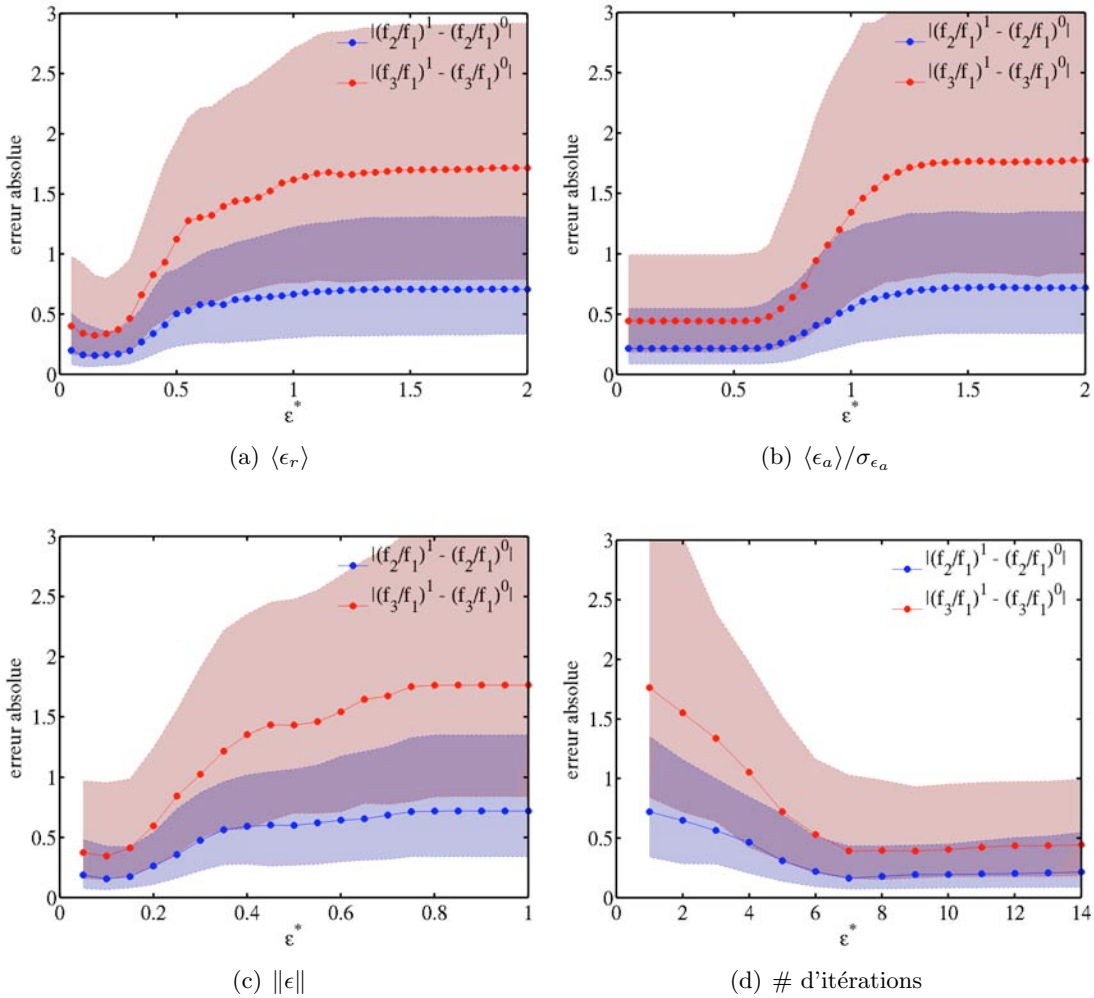
A partir de ce jeu d'ACFs bruitées, nous appliquons l'algorithme FitSVD en suivant, pour chaque itération, l'évolution de différentes quantités. Notre objectif est de définir une valeur seuil, ϵ^*, à partir de ces quantités permettant d'ajuster de manière robuste les rapports adimensionnels f_2/f_1 et f_3/f_1, *i.e.* une valeur ϵ^* qui minimise au mieux l'écart entre les paramètres d'origines $\{(f_2/f_1)^0, (f_3/f_1)^0\}$ et ceux ajustés par la procédure $\{(f_2/f_1)^1, (f_3/f_1)^1\}$. Nous testons donc différentes quantités couramment utilisées dans la pratique pour quantifier l'erreur relative : l'erreur quadratique relative, le rapport de la moyenne de l'erreur absolue sur son écart type (qui peut être interprété comme l'erreur absolue réduite), le rapport de la norme des erreurs absolues sur la norme de la puissance et ainsi que le nombre d'itérations :

$$\langle \epsilon_r \rangle = \sqrt{\frac{1}{n-3} \sum_{i=1}^{n} \left(\frac{|\mathrm{R}(\tau_i)| - |\widehat{\mathrm{R}}(\tau_i)|}{|\mathrm{R}(\tau_i)|} \right)^2} , \tag{4.2}$$

$$\frac{\langle \epsilon_a \rangle}{\sigma_{\epsilon_a}} \quad \text{avec } \epsilon_a(\tau) = \big| |\mathrm{R}(\tau)| - |\widehat{\mathrm{R}}(\tau)| \big| , \tag{4.3}$$

$$\|\epsilon\| = \frac{\|\epsilon_a\|}{\|\mathrm{R}\|} \quad \text{avec } \epsilon_a(\tau) = \big| |\mathrm{R}(\tau)| - |\widehat{\mathrm{R}}(\tau)| \big| , \tag{4.4}$$

où $|\mathrm{R}(\tau)|$ est la mesure brute de la puissance au retard τ et $|\hat{\mathrm{R}}(\tau)|$ la valeur approchée par le modèle. Le symbole $\langle \cdot \rangle$ exprime une moyenne et le symbole $\| \cdot \|$ exprime la norme. Ces différentes quantités ne véhiculent pas tout à fait la même information



(a) $\langle \epsilon_r \rangle$

(b) $\langle \epsilon_a \rangle / \sigma_{\epsilon_a}$

(c) $\|\epsilon\|$

(d) # d'itérations

FIG. 4.7 – *Evolution de la distribution de l'erreur absolue entre les paramètres* $\{(f_2/f_1)^0, (f_3/f_1)^0\}$ *et les paramètres calculés* $\{(f_2/f_1)^1, (f_3/f_1)^1\}$ *en fonction d'un seuil* ϵ^* *appliqué à :* $\langle \epsilon_r \rangle$ *(a),* $\langle \epsilon_a \rangle / \sigma_{\epsilon_a}$ *(b),* $\|\epsilon\|$ *(c) et du nombre d'itérations (d). En bleu, c'est l'écart sur l'estimation de* f_2/f_1 *et en rouge celui sur l'estimation de* f_3/f_1*. La courbe indique la valeur médiane et la zone colorée définit l'intervalle entre le premier et le troisième quartile.*

et se comportent différemment. L'erreur quadratique relative (4.2) diminue lorsque les premiers retards aberrants sont éliminés, puis diverge lorsque le nombre de retards restants est égal à 3. Le rapport de la moyenne sur l'écart type de l'erreur absolue (4.3) se comporte sensiblement de la même manière. Le rapport de la norme des erreurs absolues sur la norme de la puissance (4.4) diminue à chaque itération et un décrochement devrait être observé lorsque l'ensemble des retards invalides est éliminé. Ce dernier paraît plus robuste en ce sens que l'erreur absolue est pondérée par la norme de la puissance, qui fluctue moins entre chaque itération.

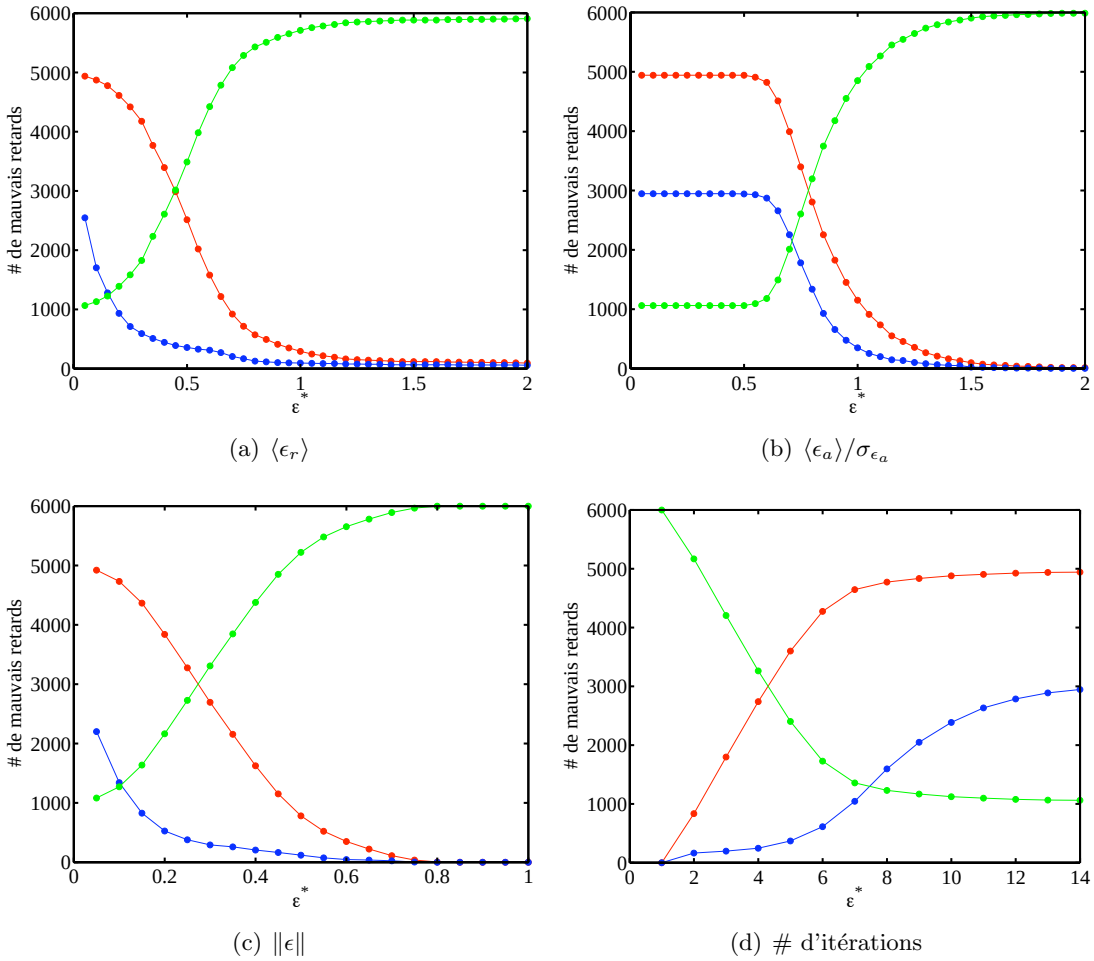


FIG. 4.8 – *Nombre de retards détectés en fonction du seuil ϵ^*. En rouge, le nombre de retards erronés détectés par FitSVD. En vert, le nombre de retards erronés non détectés par FitSVD. La courbe bleue représente le nombre de retards détectés par FitSVD affecté par le bruit.*

Pour chacune de ces quantités, nous ajustons l'échantillon d'ACFs à partir de l'algorithme FitSVD. Nous arrêtons volontairement la procédure lorsque la quantité atteint une valeur seuil, ϵ^*. Nous réitérons l'opération en faisant varier ce seuil afin de déterminer la valeur qui permet de réduire au mieux l'écart entre les paramètres d'origine et ceux calculés. Le résultat de cette simulation est illustré avec les figures 4.7(a) 4.7(b) 4.7(c) et 4.7(d) dans le cas d'un échantillon d'ACFs bruitées avec un niveau de 20 dB et par 6 retards aberrants. Ces figures montrent la distribution des écarts absolus sur l'estimation des rapports f_2/f_1 (en bleu) et f_3/f_1 (en rouge) en fonction de ϵ^*, pour les quatre critères d'arrêt : erreur quadratique relative 4.7(a), rapport de la moyenne sur l'écart type de l'erreur absolue 4.7(b), rapport de la norme des erreurs absolues sur la norme de la puissance 4.7(c) et le nombre d'itérations 4.7(d). Le bandeau de couleur

définit l'intervalle entre le premier et le troisième quartile et la courbe matérialise la valeur médiane de la distribution.

Comme le suggèrent ces 4 figures, l'erreur commise sur l'évaluation du rapport f_3/f_1 est systématiquement plus importante. Ceci s'explique par le poids (λ_3) relativement faible du mode 3 par rapport au deux premiers modes. Ces figures montrent que pour chacun des critères, une valeur seuil peut être définie. Lorsque la contrainte augmente, l'écart sur les paramètres diminue pour atteindre des valeurs qui sont sensiblement identiques selon les critères d'arrêt. Notons que lorsque le seuil appliqué sur $\langle\epsilon_a\rangle/\sigma_{\epsilon_a}$ devient trop restrictif (*cf* Fig. 4.7(b)), la valeur seuil n'est plus atteinte. L'algorithme s'arrête donc une fois avoir identifié tous les retards erronés. Cette remarque est également vraie pour le nombre d'itérations (*cf* Fig. 4.7(d)), puisque dans cet exemple nous n'avons simulé que 6 retards aberrants par ACF. Nous devons préciser qu'en simulant des cas plus défavorables, l'écart entre les paramètres d'origine et ceux ajustés augmente à nouveau lorsque $\epsilon^* = 10$ itérations.

Le critère basé sur l'erreur quadratique relative 4.7(a) semble être le plus performant par rapport aux trois autres, en ce sens que l'intervalle entre le premier et le troisième quartile est autour de $\epsilon^* \approx 0.2$ le plus petit. Nous avons réitéré cette simulation pour différents niveaux de bruit de (10 à 30 dB) et pour différents nombres de retards aberrants (2 à 13). On note un élargissement de la distribution des erreurs lorsque que l'on dégrade davantage les ACFs. En revanche, la valeur seuil ϵ^* n'est que très faiblement affectée. Les résultats obtenus confirment ceux présentés ci-dessus et valident l'erreur quadratique relative comme le critère d'arrêt le plus performant.

4.4.2 Performance : détection des retards erronés

La figure 4.8 présente l'évolution en fonction du seuil, ϵ^*, du nombre de retards erronés détectés (rouge) et non détectés (vert) par l'algorithme FitSVD pour les quatres critères d'arrêt : $\langle\epsilon_r\rangle$ (a), $\langle\epsilon_a\rangle/\sigma_{\epsilon_a}$ (b), $\|\epsilon\|$ (c) et du nombre d'itérations (d). La courbe bleue indique le nombre de retards dégradés sous l'action du bruit, qui sont identifiés comme erronés par FitSVD. Le nombre de retards détectés par la méthode augmente avec la contrainte appliquée au seuil. On note également que le nombre de retards considérés comme erronés par la méthode (FitSVD) tandis qu'ils sont dégradés sous l'action du bruit augmente avec la contrainte. Ceci n'est pas gênant dans la mesure où l'on considère tout retard s'écartant trop de la forme moyenne des ACFs comme erroné et ce sans distinguer la source. La valeur seuil pour laquelle nous obtenons la meilleure estimation possible des paramètres correspond également au maximum de

retards erronés détectés par la méthode (*cf* Fig. 4.8). Nous montrons ainsi que ce nouveau schéma de détection permet un ajustement robuste des paramètres f_1, f_2, f_3 de la SVD.

4.5 Conclusion

Dans ce chapitre, nous avons défini un jeu de seuils permettant de détecter les retards erronés contenus dans les ACFs, et ce de manière automatisée. A partir de ces seuils, un schéma d'interpolation opérationnel a été développé dont le but est de fournir un diagnostique fiable des paramètres issus de la rétrodiffusion.

Pour développer ce modèle, nous avons étudié les écarts relatifs entre les ACFs brutes et les approximations calculées à partir des trois modes statistiques $g_\alpha(\tau)$. Bien que l'algorithme FitACF n'identifie pas tous les mauvais retards, il nous a néanmoins permis de différencier les distributions des erreurs relatives entre les bons et les mauvais retards. Nous avons ainsi défini des limites plus contraignantes au-delà desquelles nous considérons tous retards comme non-valides. Nous avons validé cette méthode par une simulation d'ACFs synthétiques dégradées avec un bruit réaliste ainsi que des retards erronés. Nous avons également montré que cette nouvelle stratégie est nécessaire pour améliorer de manière significative l'erreur sur l'ajustement des ACFs brutes. Bien que cette méthode soit itérative, elle a l'énorme avantage de s'auto-limiter en ce sens que le nombre d'itérations dépend uniquement du nombre de retards erronés contenus dans l'ACF.

La méthode SVD fournit un critère explicite au sens où la valeur qu'elle donne à un retard τ, représente la partie reproductible du signal. La différence entre la mesure et la valeur ajustée représente la contribution due à ce que nous interprétons comme erreur (bruit, interférence, etc ...). Par conséquent, lorsque cette différence devient plus importante qu'une valeur seuil, nous considérons le retard correspondant comme aberrant et nous l'éliminons lors de l'ajustement de l'ACF. Une amélioration est cependant possible. Au lieu d'éliminer ces retards aberrants, nous pouvons envisager un ajustement pondéré au sens des moindres carrés. L'algorithme FitACF nous permet de connaître la distribution des erreurs commises par le modèle SVD sur les bons comme sur les mauvais retards. Nous pourrions envisager d'introduire le formalisme Bayésien afin de déterminer la probabilité qu'un retard τ soit correct et utiliser celui-ci comme coefficient pondérateur. Ce coefficient permettrait de donner plus ou moins de poids à ce retard lors de l'ajustement de l'ACF.

L'algorithme FitSVD offre une bonne alternative à la détection des retards erronés telle qu'elle est faite dans FitACF. En simplifiant de façon non négligeable le schéma d'identification, il pourrait dès lors être directement implémenté dans un évolution future de l'algorithme FitACF. La meilleur solution consisterai à utiliser l'information physique, telle qu'elle est utilisée dans FitACF (*e.g.* pour les CRIs) et à combiner ces critères-là avec ceux de FitSVD.

Conclusions et perspectives

Le travail présenté dans ce mémoire est consacré à la caractérisation des régions frontières de la magnétosphère à partir de leur projection dans l'ionosphère à haute latitude, en utilisant les mesures de la chaîne de radars SuperDARN. Ce travail a permis de poser les premiers jalons d'un schéma opérationnel pour la détection et le suivi, idéalement en temps réel, de ces régions.

L'environnement spatial de la Terre, dont les frontières sont dessinées par son champ magnétique, subit les assauts continuels de l'activité solaire. La magnétosphère terrestre résulte d'un équilibre dynamique entre le champ magnétique terrestre et celui du milieu interplanétaire. Cet équilibre est régi par des transferts d'énergie, de matière et de moment. Le suivi des régions frontières de la magnétosphère est donc un enjeu crucial dans l'étude de la dynamique de notre environnement et notamment dans la compréhension de l'impact de l'activité solaire.

L'élaboration d'un modèle opérationnel requiert une méthodologie spécifique. En effet, en météorologie de l'espace, il est souvent préférable d'avoir accès à des paramètres empiriques robustes et faciles à interpréter, qu'à des quantités physiques qui sont parfois plus difficiles à estimer. C'est dans cette optique que notre travail se focalise sur la caractérisation statistique des échos radar mesurés par SuperDARN, en se basant pour cela sur les données brutes.

L'approche habituelle consiste à ajuster à partir de l'algorithme FitACF, un modèle physique (gaussien ou exponentiel) aux fonctions d'autocorrélation (ACF) produites par les radars. A partir des paramètres de ces modèles, on déduit des quantités qui possèdent une signification physique mais qui sont également controversées : largeur spectrale, puissance rétrodiffusée, retard de phase. Nous avons opté pour une approche statistique (et en ce sens plus empirique). Sachant que le module $|R(\tau)|$ des ACFs possède une forme reproductible, il devient possible d'approcher son allure à partir d'une combinaison linéaire d'un nombre limité de fonctions de base $g_k(\tau)$ communes à toutes les régions géophysiques, appelées «modes» :

$$|R(t,\tau)| = \sum_k \lambda_k \cdot f_k(t) \cdot g_k(\tau) \ .$$

La méthode qui permet de reproduire, au sens des moindres carrés, toutes les ACFs à partir d'un nombre minimum de modes statistiques, est la Décomposition en Valeurs Singulières (SVD), que nous avons donc utilisée.

Dans le chapitre 2, nous montrons que le module des ACFs peut se décomposer uniquement en une série de trois modes (ou de fonctions de base) $g_k(\tau)$. Une combinaison linéaire de ces modes permet de reproduire 99.2 % de l'énergie des ACFs observées. Nous montrons également que cette décomposition est robuste puisque ces modes ne sont guère affectés par la taille de l'échantillon statistique (pour autant que celui-ci couvre les différents types d'ACF que mesure SuperDARN), par le niveau d'activité géomagnétique, ou encore, ne dépendent pas du secteur horaire magnétique.

L'ajustement du module de l'ACF par les trois modes donne accès à trois coefficients, f_k. Ces coefficients expriment la contribution de chacun des trois modes à la description de l'ACF et reproduisent donc à eux seuls l'ensemble des types d'ACF mesurées par SuperDARN. L'inconvénient d'une telle approche est que ces modes proposent une description statistique de la mesure, dénuée de sens physique. La comparaison avec les paramètres physiques issus de l'algorithme FitACF révèle que l'information pertinente est contenue dans les rapports adimensionnés f_2/f_1 et f_3/f_1. Le rapport f_2/f_1 est fortement corrélé avec la largeur spectrale, et peut donc être utilisé comme traceur de celle-ci. Le rapport f_3/f_1 caractérise la décroissance et la forme des ACFs et apporte une correction d'ordre supérieur à l'ajustement de la mesure. L'intérêt de cette description des ACFs réside dans le fait que les paramètres obtenus ne sont pas liés au choix du modèle physique en particulier. En effet, le choix d'un modèle gaussien ou exponentiel dans FitACF entraîne un écart systématique de 5 Hz (soit $\sim$ 300 m.s^{-1} à 10 MHz) dans la valeur de la largeur spectrale, alors que les observations montrent qu'il est souvent difficile de préférer un modèle à un autre. Cet écart est une conséquence de l'inadéquation du modèle physique et justifie d'autant plus le recours à un modèle empirique.

Notre méthode a été validée de trois façons :

1. D'abord, nous montrons que l'ajustement des ACFs est systématiquement meilleur avec le modèle empirique SVD qu'avec les modèles physiques, et ce quelle que soit la région couverte.
2. Ensuite, nous avons vérifié que les trois modes dominants du modèle SVD sont robustes, en ce sens qu'il ne dépendent pas de façon sensible de la taille de l'échantillon analysé, ni de la région couverte. Notons cependant que nous avons délibérément exclu de cet échantillon les ACFs possédant des retards aberrants afin de définir un jeu de fonctions de base de forme régulière.

3. Enfin, nous avons comparé la distribution spatiale des paramètres obtenus par SVD avec les paramètres du modèle physique, pour montrer que les résultats sont cohérents.

L'objectif de ce modèle empirique s'éloigne de l'objectif premier de SuperDARN qui est d'étudier la convection. Nous avons donc choisi de mettre en place une nouvelle procédure «FitSVD» de traitement et d'analyse des ACFs brutes à partir du modèle SVD. Les mesures sont fréquemment corrompues par des valeurs erronées qu'il convient de détecter puis d'éliminer avant de pouvoir estimer correctement les quantités physiques. Sachant que ces valeurs corrompues se démarquent de celles fortement reproductibles, décrites avec les modes de la SVD, nous obtenons un critère simple pour les éliminer :

- Les trois paramètres f_k sont estimés par un ajustement au sens des moindres carrés des trois fonctions de bases $g_k(\tau)$.
- L'ACF est reconstruite à partir de ces trois modes seulement. On détermine ensuite l'écart entre la mesure et le modèle.
- Les écarts qui dépassent une valeur seuil sont éliminés un à un. A chaque étape, les paramètres f_k sont estimés de nouveau. On procède par itérations jusqu'à ce qu'il y ait un bon accord entre l'ACF reconstruite et celle observée.

Dans le chapitre 4, nous avons montré comment sélectionner les valeurs seuil de cet algorithme et défini sa plage de validité. Nous avons montré également qu'au-delà de 10 valeurs corrompues (sur 18) dans une ACF, il n'est plus possible d'estimer correctement les paramètres f_k. L'algorithme «FitSVD» permet de simplifier de manière non négligeable la procédure de détection des retards erronés, avec de bonnes performances.

A partir de ces résultats, nous nous sommes concentrés sur la distribution spatiale de la largeur spectrale en tant que traceur de la frontière des lignes de champ ouvertes/fermées (OCB). De nombreuses études ont en effet montré que cette frontière coïncide avec une brusque variation en latitude de la largeur spectrale (SWB), et ce dans une large gamme de valeur MLT. Toutefois, au lieu d'utiliser la largeur spectrale, comme il est habituel de le faire, nous utilisons le rapport f_2/f_1, qui est plus robuste et n'est pas affecté par le choix du modèle physique. Nous introduisons à cette occasion un critère de décision Bayésien dans le seul but de localiser la position de la frontière.

L'approche Bayésienne s'appuie sur une conception des probabilités en termes de degré de confiance qu'un événement survienne. Dans notre étude, le point de départ est la présence de deux classes de valeurs distinctes de la largeur spectrale de part et d'autre de la SWB. La transition entre ces deux régimes est parfois progressive selon les conditions locales du milieu, ce qui rend d'autant plus difficile la définition d'une

valeur seuil délimitant les deux régions. L'approche Bayésienne permet donc d'évaluer le degré d'appartenance de l'ACF à l'une des deux classes à partir de sa valeur f_2/f_1. On associe la classe des grandes largeurs spectrales à la région des lignes de champ ouvertes sur le milieu interplanétaire et celle des petites largeurs spectrales aux lignes de champ fermées.

Nous montrons que le formalisme Bayésien est bien adapté aux données éparses. Non seulement, il permet de déterminer la valeur seuil optimale qui différencie au mieux les deux régions situées de part et d'autre de la frontière, mais en plus, il fournit un intervalle de confiance, ce qui est précieux pour statuer sur les cas fréquents où la frontière est mal déterminée. Il permet ainsi de localiser la position de la SWB avec un degré de confiance donné, ce que ne permet pas la méthode classique.

Le formalisme Bayésien offre des perspectives intéressantes. Comme nous l'avons illustré dans ce mémoire, l'association de la SWB avec les frontières géomagnétiques est ambiguë. Le caractère générique des méthodes Bayésiennes est ici un atout majeur : nous pouvons en effet introduire de nouvelles connaissances dans le modèle et ensuite les mettre à jour pour mieux réconcilier l'emplacement de la SWB avec l'OCB. L'étape suivante consiste alors à travailler sur l'écart entre ces deux frontières. En confrontant la position de la SWB avec la projection de l'OCB identifiée par satellite, nous pourrions étudier l'influence des conditions du milieu interplanétaire (orientation du champ magnétique interplanétaire, pression dynamique du vent solaire...), de l'activité géomagnétique (Kp), de la position de l'ovale auroral sur cet écart, ou encore des ondes (variations temporelles). L'intégration de ces lois comportementales dans le modèle permet de déduire la position de l'OCB de celle de la SWB. Une extension de cette approche est envisageable à d'autres frontières.

Le problème qui se pose ici est cependant celui de l'assimilation de données dans le modèle Bayésien. En effet, ce modèle d'identification et de suivi des régions frontières magnétosphériques est destiné à opérer en temps réel. Ceci implique implicitement que les nouvelles connaissances que l'on désire intégrer à ce modèle le soient également. C'est un des facteurs à prendre en compte en plus de leur pouvoir descriptif du système magnétosphère-ionosphère. Ce dernier point est peut-être le plus délicat pour l'élaboration d'un modèle opérationnel car les structures informatisées existantes ne sont pas forcément adaptées pour cela. La météorologie de l'espace souffre de cela. En effet, cette nouvelle discipline nécessite la mise en place de nouvelles structures donnant accès à de nombreuses données de diverses disciplines et ce en temps réel, ce qui fait défaut aujourd'hui. Des projets comme par exemple «SWENET» (ESA) vont dans ce sens.

En parallèle de l'objectif premier de ce travail qui est de détecter et de suivre les

régions frontières de la magnétosphère, la méthode SVD ouvre de nouvelles possibilités pour l'analyse des mesures SuperDARN :

- La définition d'un nouveau jeu de paramètres empiriques permet d'améliorer la description des spectres Doppler par rapport aux modèles utilisés dans l'algorithme FitACF. En effet, le paramètre f_2/f_1 fournit une valeur unique de la largeur spectrale, mais la combinaison de f_2/f_1 avec f_3/f_1 permet en plus de caractériser la forme des spectres. Cette nouvelle information est importante pour étudier les processus micro-physiques dont dépendent les mesures radars. Nous pouvons dès lors utiliser ce modèle empirique dans une prochaine évolution de l'algorithme FitACF, afin d'étudier ces mécanismes en fonction des conditions du milieu.
- L'interprétation physique des paramètres, f_k/f_1, a permis d'identifier un certain nombre d'échos radar dont le comportement diffère de l'ensemble de l'échantillon utilisé. Ces échos radar possèdent des caractéristiques proches des échos issus de la région E (vitesse Doppler faible et petite largeur spectrale). Cette étude pourrait donc être étendue à la région E afin d'établir de nouveaux critères plus robustes (que ceux utilisés dans ce mémoire) afin de différencier les régions E et F. La catégorie des échos de sol pourrait également être étudiée.
- Nous pourrions également envisager d'utiliser la méthode SVD non plus comme un outil réduction mais comme un outil de différenciation. A partir d'échantillons restreints à certaines régions spécifiques de l'ionosphère, comme par exemple celle qui est associée à la projection du cornet polaire, nous pouvons appliquer la méthode SVD afin d'étudier le degré de complexité du spectre de rétrodiffusion. Nous pourrions dès lors caractériser des spectres plus complexes, car pour le moment seuls les spectres à une composante peuvent être caractérisés.

Les méthodes d'analyse statistique comme celles employées dans ce travail peuvent être également étendues à d'autres expériences. Ainsi, nous pourrions appliquer la méthode SVD aux radars incohérents (EISCAT), qui mesurent en fonction de l'altitude les paramètres du plasma (densité ionique, température ionique et électronique, vitesse du plasma). En établissant à partir de la méthode SVD des classes statistiques de ces paramètres, nous pourrions mettre en avant des modes spatio-temporels de plasma caractéristiques de différentes régions selon l'altitude, ou encore identifier les échos cohérents.

A

Principe de fonctionnement des radars SuperDARN

A.1 Radar multi-impulsions et fonctions d'autocorrélation 139
A.2 Identification des sources de bruits et d'interférences 141
A.2.1 Sources «naturelles» 141
A.2.2 Sources «instrumentales» 142
A.3 Algorithme «FitACF» : traitement des ACFs 143
A.3.1 Estimation des niveaux de fluctuations 144
A.3.2 Classification des «mauvais» retards (ou bad lags) 145

Nous abordons dans cette annexe, le principe de fonctionnement des radars SuperDARN. Nous détaillerons le processus de reconstruction des fonctions d'autocorrélation, ainsi que de leur traitement à partir de l'algorithme FitACF.

A.1 Radar multi-impulsions et fonctions d'autocorrélation

Les caractéristiques géophysiques de l'ionosphère des hautes latitudes sont telles que SuperDARN doit détecter des dérives Doppler jusqu'à 2 km.s^{-1}, pour des distances allant au-delà de 3330 km. Ceci n'est pas possible avec l'utilisation d'un schéma d'impulsions régulier. En effet, la fréquence de répétition induite par le temps de propagation (aller et retour) des ondes rétrodiffusées par les irrégularités les plus éloignées est de $f_r \leqslant 1/\tau_{max} = 45$ Hz (τ_{max} le temps mis par l'onde pour parcourir un aller et retour entre le radar et les irrégularités les plus éloignées), ce qui limite la mesure de

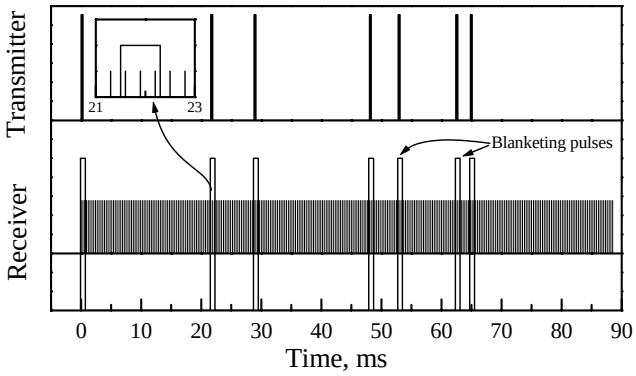


FIG. A.1 – *Schéma de 7 impulsions (d'après Ponomarenko and Waters (2006)). L'encart, situé en haut à gauche de la figure, montre la superposition de l'échantillonnage du récepteur avec l'émission d'une impulsion.*

dérive à $f_n \leqslant 20$ Hz (la fréquence de Nyquist). Or les paramètres physiques du milieu sont tels que la convection ionosphérique peut produire des dérives Doppler jusqu'à des valeurs de l'ordre de $f_D = 2V_D/\lambda_{rad} \approx 100 - 200$ Hz (avec f_D la fréquence Doppler, V_D la dérive Doppler et λ_{rad} la longueur d'onde d'émission du radar).

Les schémas multi-impulsions permettent de résoudre les ambiguïtés en distance ou en vitesse (Fig. A.1). Cette technique, introduite par Farley (1972) a d'abord été expérimentée sur les radars à diffusion incohérente avant d'être adaptée avec succès sur les radars coéhrents HF (Greenwald et al., 1985). Elle consiste à émettre une série d'impulsions de manière à ce que le délai entre deux impulsions consécutives soit un multiple d'un intervalle élémentaire (par exemple en mode commun $\tau_o = 2.4$ ms, soit $f_n = 200$ Hz). Un train de sept impulsions permet d'obtenir 18 retards (de $0\tau_o$ à $17\tau_o$) de l'ACF. La figure A.2 illustre, dans le cas d'un schéma à deux impulsions, la manière dont les retards de l'ACF sont reconstruits et comment on estime la distance des irrégularités associées. La durée des impulsions fixe la résolution radiale du volume sondé. En mode commun, la longueur des impulsions est de 300 μs ; soit une résolution en distance de 45 km. L'échantillonnage donne accès à 70 portes (ou cellules radars) distantes de 180 à 3330 km du radar.

L'émission et la réception utilisent le même réseau d'antennes. L'échantillonnage du récepteur se superpose à l'émission des impulsions (Fig. A.1). Le récepteur est échantillonné à une fréquence de ~ 3 kHz et reste ouvert le temps nécessaire pour recevoir le dernier signal rétrodiffusé par les irrégularités distantes d'au plus de 3330 km, soit $\Delta t_{max} \approx 90$ ms. Pour chaque écho radar reçu, le récepteur numérise une double valeur :

une composante en phase (I) et une seconde déphasée de 90° (Q), permettant ainsi la définition d'une amplitude complexe unique ($I + jQ$) du signal reçu. La re-combinaison par paire des signaux enregistrés permet de définir une amplitude complexe pour chacun des 18 retards de l'ACF. L'étude du temps de parcours entre l'émission et la réception de l'onde permet, moyennant une hypothèse sur le trajet de l'onde, d'estimer la distance radiale des irrégularités rétrodiffusantes associées à l'ACF. Quant à l'allure du module de l'ACF, elle permet de caractériser les propriétés spectrales du volume rétrodiffusant associé (*cf* § 1.3.2.2).

Pour chaque faisceau, l'opération ci-dessus est répétée pendant le temps d'intégration. En mode commun, le temps d'intégration est de $t_i \approx 7$ s. Ce temps fixe le nombre d'ACFs à moyenner ($N_{avg} = t_i/\Delta t_{max}$), soit $N_{avg} \approx 75$. Les N_{avg} ACFs définies pour chacune des portes sont moyennées afin d'atténuer l'effet des fluctuations statistiques. Ce sont ces données que nous analysons, ce qui veut dire que nous caractérisons une mesure qui est la moyenne d'un processus non stationnaire.

A.2 Identification des sources de bruits et d'interférences

En plus des mécanismes physiques tels que les effets liés à la propagation, l'absorption des ondes HF ou encore les effets liés à la turbulence du milieu, de nombreux facteurs viennent perturber les signaux rétrodiffusés. Ces effets affectent différemment l'allure générale de l'ACF et peuvent entraîner une déviation de la puissance mesurée sur certains retards de plusieurs ordres de grandeur par rapport à celle attendue. Ces valeurs aberrantes sont donc une source non négligeable d'erreur sur l'estimation des paramètres. Une détection des retards affectés est nécessaire afin de ne pas les prendre en compte lors de l'ajustement de l'ACF par un modèle.

Les sources de bruits ou d'interférences peuvent être classées en deux catégories : les sources dites «naturelles» présentes dans l'ionosphère et les sources dites «instrumentales» causées par le principe de fonctionnement même des radars.

A.2.1 Sources «naturelles»

Parmi les facteurs, nous notons la présence dans l'ionosphère d'un bruit de fond incohérent que l'on associe généralement au bruit galactique. Dans la bande HF, le bruit galactique est supérieur au bruit électronique. La non-stationnarité des processus physiques pendant la durée de la mesure est également une source de fluctuations. La

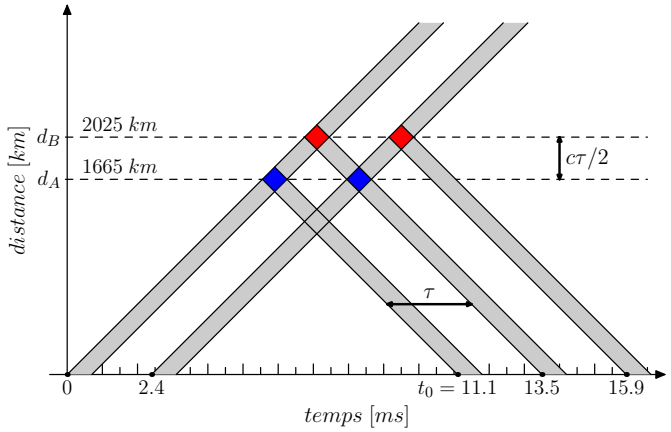


FIG. A.2 – *Exemple de CRI dans le cas d'un schéma à 2 impulsions. La longueur des impulsions (en gris) n'est pas à l'échelle.*

répétition de la séquence de mesure permet d'en atténuer l'effet. Une troisième source de bruit, cohérente celle-ci, est liée à l'activité humaine. Les interférences radio (de magnitude supérieure au niveau de bruit galactique) peuvent recouvrir la signature radar des irrégularités. Le problème du bruit cohérent est partiellement résolu. Un sondage en fréquence est réalisé afin de sélectionner une bande de fréquence libre de toutes interférences avant chaque balayage en azimut. Cette sélection de la fréquence radar s'effectue toutes les 2 min. Cette précaution n'exclut pas l'interférence de cette fréquence sélectionnée par une source radio cohérente pendant le balayage.

L'identification de la nature des sources de fluctuations et l'estimation de leur niveau (ou de leur variance) sont un problème en soi qui n'est pas encore résolu et fait l'objet d'un débat animé au sein de la communauté SuperDARN.

A.2.2 Sources «instrumentales»

La conception même des radars et leur mode de fonctionnement sont également une source d'erreur supplémentaire. En plus du bruit électronique, il arrive parfois, de par la nature mono-statique des antennes, que l'émission d'une impulsion se superpose à l'échantillonnage du récepteur. L'encadré de la figure A.1 illustre ce cas. Il montre la superposition de l'échantillonnage du récepteur pendant la durée d'une impulsion qui est émise. Prévisible, le signal mesuré est filtré et se traduit par la mise à zéro du ou des retards correspondants de l'ACF.

Le schéma multi-impulsions qui est nécessaire pour mesurer les larges dérives Dop-

pler est également une source d'interférences inter-distance (CRI[1]). Le CRI se traduit par le retour simultané de plusieurs échos radar provenant de différentes cibles. Une illustration de la formation d'un CRI est donnée à la figure A.2, dans le cas d'un schéma à deux impulsions. Les cibles A et B sont séparées de $c\tau/2$ l'une de l'autre. On retrouve dans la littérature de nombreuses discussions au sujet des CRIs. Barthes et al. (1998) abordent le problème de manière formelle. Yukimatu and Tsutsumi (2002) présentent un exemple de CRI et montrent comment les signaux reçus se superposent.

Les irrégularités de petites échelles étendues sur des distances de l'ordre de 1000 km forment une cible continue fournissant des conditions favorables à la formation des CRIs. Avec l'exemple donné à la figure A.2, il est aisé de comprendre comment les CRIs peuvent affecter la puissance des ACFs. Les retards de l'ACF affectés par un CRI dépendent évidemment de la distance respective entre celui-ci et le volume rétrodiffusant qui est associé à l'ACF. Le déphasage entre les 2 signaux rétrodiffusés par les deux volumes sondés peut être considéré comme aléatoire. La répétition de la séquence d'émission durant le temps d'intégration permet de diminuer l'effet de ces termes interférentiels. En effet, le moyennage sur un nombre infini de séquence tend à les annuler. Ce n'est plus le cas pour un nombre fini de séquence où les termes interférentiels peuvent produire une fluctuation aléatoire supplémentaire. Les CRIs peuvent ainsi contribuer à la diminution ou à l'augmentation de la puissance de l'ACF au retard τ. A partir d'une séquence d'impulsions, il est facile de calculer quand et pour quels retards de l'ACF le CRI peut se produire. Par contre, il est beaucoup plus compliqué de savoir si le CRI est significatif à une distance donnée.

La présence de ces différentes sources de bruit, d'interférence dans les mesures influence l'ajustement par les modèles de l'ACF et donc biaise l'estimation des paramètres issus de la rétrodiffusion. L'algorithme FitACF, qui est le logiciel standard de traitement et d'analyse des mesures SuperDARN, utilise un certain nombre de critères afin d'améliorer l'ajustement des ACFs. Dans la suite, nous allons décrire en détail le fonctionnement de cet algorithme pour mieux confronter les performances et les différences respectives entre FitACF et le nouveau schéma de détection développé à partir du modèle SVD.

A.3 Algorithme «FitACF» : traitement des ACFs

L'algorithme FitACF est la procédure standard implémentée sur chacun des sites radars du réseau SuperDARN. Cette algorithme est le fruit de vingt années de recherche

[1]En anglais : *Cross-Range Interference*

pour améliorer l'analyse des échos radar, Il est constitué d'un ensemble de programmes dédiés : (1) au pré-traitement des ACFs (*i.e.* qu'il teste la validité de chacun des 18 retards), (2) il ajuste ensuite des modèles permettant d'extraire des quantités physiques, telles que la vitesse Doppler, la largeur spectrale et la puissance rétrodiffusée.

Une ACF est enregistrée pour chaque cellule radar formant un faisceau. Pour chaque faisceau, l'algorithme FitACF procède comme suit. Les ACFs brutes sont d'abord traitées afin d'exclure les retards affectés par les CRIs et par la superposition du récepteur par l'émission d'une impulsion. Ensuite, le niveau de bruit de fond est estimé. On fait l'hypothèse d'un bruit additif sur le module de l'ACF. Ce bruit est corrigé en soustrayant pour chaque retard τ de l'ACF, la même valeur. Les retards possédant une puissance négative sont également éliminés. Les ACFs dites «propres» sont ensuite ajustées par un modèle gaussien $\left(R(\tau) \propto e^{-\sigma^2\tau^2}\right)$ et par un modèle exponentiel $\left(R(\tau) \propto e^{-\lambda\tau}\right)$ d'où l'on déduit la puissance rétrodiffusée et la largeur spectrale. La vitesse Doppler est estimée en ajustant une droite sur la variation de la phase avec les retards.

A.3.1 Estimation des niveaux de fluctuations

Afin de détecter les retards affectés par ces sources d'erreurs, l'algorithme FitACF calcule avant l'ajustement de l'ACF par les modèles, trois quantités : le niveau des fluctuations statistiques, le bruit au retard zéro et le bruit aux retards non-nuls ($\tau \neq 0$). Ces trois quantités sont réactualisées pour chaque faisceau.

Le niveau (ou la valeur efficace) des fluctuations statistiques, σ_R, est déduite du module de l'ACF, en calculant le rapport de la puissance moyenne au retard 0 sur la racine carrée du nombre d'ACFs utilisées dans le calcul de la moyenne, $\sigma_R = |\mathrm{R}(0)|/\sqrt{N_{avg}}$.

Le bruit au retard nul, $R_n(0)$, correspond à la valeur la plus élevée entre le bruit de fond mesuré en réception passive, *i.e.* lorsque l'émetteur est éteint, et la moyenne des 10 échos radar les plus faibles le long d'un faisceau. Ce paramètre est utilisé pour éliminer les ACFs avec un rapport signal sur bruit (en dB) trop faible.

Le bruit aux retards non-nuls, $\langle R_n(\tau \neq 0)\rangle$, est obtenu en identifiant toutes les ACFs du faisceau qui satisfont à la condition $|\mathrm{R}(0)| < 1.6R_n(0) + \sigma_R$ et en moyennant la puissance de tous les retards de toutes ces ACFs. C'est l'expression $|\mathrm{R}(\tau)| - (\langle R_n(\tau \neq 0)\rangle + \sigma_R)$ qui est ajustée pour déterminer la largeur spectrale, ce qui signifie que σ_R est plutôt considéré comme un biais. Ponomarenko and Waters (2006) montrent en outre que la soustraction de σ_R à la puissance de chaque retard de l'ACF augmente artificiellement la largeur spectrale. L'algorithme FitACF fournit des valeurs de la largeur

spectrale dont les plus élevées atteignent $\sim$ 500 m.s^{-1} ; tandis que l'ajustement des ACFs sans opérer de soustraction du niveau de fluctuations confine les largeurs spectrales à des valeurs de l'ordre de 250 m.s^{-1}. Ces nouvelles valeurs se rapprochent de celles prédites par la théorie de la diffusion. Les auteurs concluent que dans la grande majorité des publications la largeur spectrale est certainement largement sur-estimée. Ils préconisent notamment dans la prochaine version de l'algorithme FitACF de ne plus soustraire cette quantité lors de l'estimation de la largeur spectrale. Dans le cas de la SVD, nous avons opté de travailler l'ACF brute, *i.e.* sans extraire le niveau. En effet, il est extrêmement difficile de vérifier si ce bruit peut être considéré comme constant sur les 18 retards et d'un point de vue pratique, il est difficilement envisageable d'appliquer la méthode SVD sur un échantillon possédant des trous.

A.3.2 Classification des «mauvais» retards (ou bad lags)

La présence de bruits et d'interférences dans les mesures se manifeste donc de différentes manières sur l'allure générale de l'ACF. A partir des trois quantités, fluctuations statistiques, bruit au retard nul et non-nuls, l'algorithme FitACF effectue une série de tests empiriques afin d'éliminer les retards affectés par des sources potentielles de bruits ou d'interférences. Les retards qui s'écartent trop de la forme attendue de l'ACF sont signalés comme «**mauvais retards**» (ou bad lags) et sont ignorés lors de l'ajustement par les modèles de l'ACF.

En pratique, l'algorithme FitACF effectue une classification des retards selon qu'ils soient affectés ou non par une source de bruit potentielle. A chaque retard de chacune des ACFs est attribué un identifiant. Cet identifiant indique si le retard peut être

Tab. A.1 – *Classification des mauvais retards selon l'algorithme FitACF.*

Identifiant	Origine des «mauvais retards»
0	Si le retard τ est valide.
1	Si le retard τ est affecté par un CRI ou par une superposition avec une impulsion.
3	Si la puissance du retard τ est négative après extraction du bruit.
5	Si le retard τ est assimilé à un pic positif.
7	Si le retard τ se situe après deux retards consécutifs dont leur puissance est négative.
9	Si le retard τ est assimilé à pic négatif.

considéré comme valide et par conséquent utilisé lors de l'ajustement de l'ACF par les modèles. Ainsi, seuls les retards considérés comme représentatifs du volume rétrodiffuseur sont validés. Le tableau A.1 résume la classification des retards effectués par l'algorithme FitACF. Une description détaillée des différents tests effectués à partir de l'algorithme FitACF est fournie par Ponomarenko and Waters (2006).

La figure A.3 montre cinq ACFs brutes avec le résultat de la détection des mauvais retards selon l'algorithme FitACF. Le chiffre à côté des retards erronés (point rouge) indique sa catégorie (résumé dans le tableau A.1). D'un point de vue qualitatif, nous pouvons voir qu'un certain nombre de retards qui ne sont pas signalés comme erronés (par exemple à la figure. A.3(b)) semblent s'écarter significativement de la décroissance moyenne de l'ACF, tandis que d'autres signalés comme erronés ne semblent affecter que faiblement l'allure moyenne des ACFs. Ces exemples amènent naturellement à se demander si la stratégie employée avec l'algorithme FitACF est suffisante puisqu'elle ne permet pas d'éliminer automatiquement les retards qui influencent le plus l'ajustement des ACFs. D'une façon plus générale et ce à cause du caractère bruité de la mesure, on ne peut pas dire si l'ACF comprend uniquement des bonnes valeurs ou seulement des valeurs erronées.

Ponomarenko and Waters (2006) ont étudié en détail les performances de l'algorithme FitACF et son influence sur l'estimation des paramètres. Ils sont arrivés à la conclusion que tous les mauvais retards ne sont pas détectés. Les auteurs montrent en effet que la non-détection des retards affectés par des CRIs a pour effet de surévaluer, systématiquement, le niveau de fluctuations statistiques σ_R, d'un facteur ~ 5. Ainsi, beaucoup de retards avec une puissance s'écartant au-delà des limites de fluctuations sont conservés.

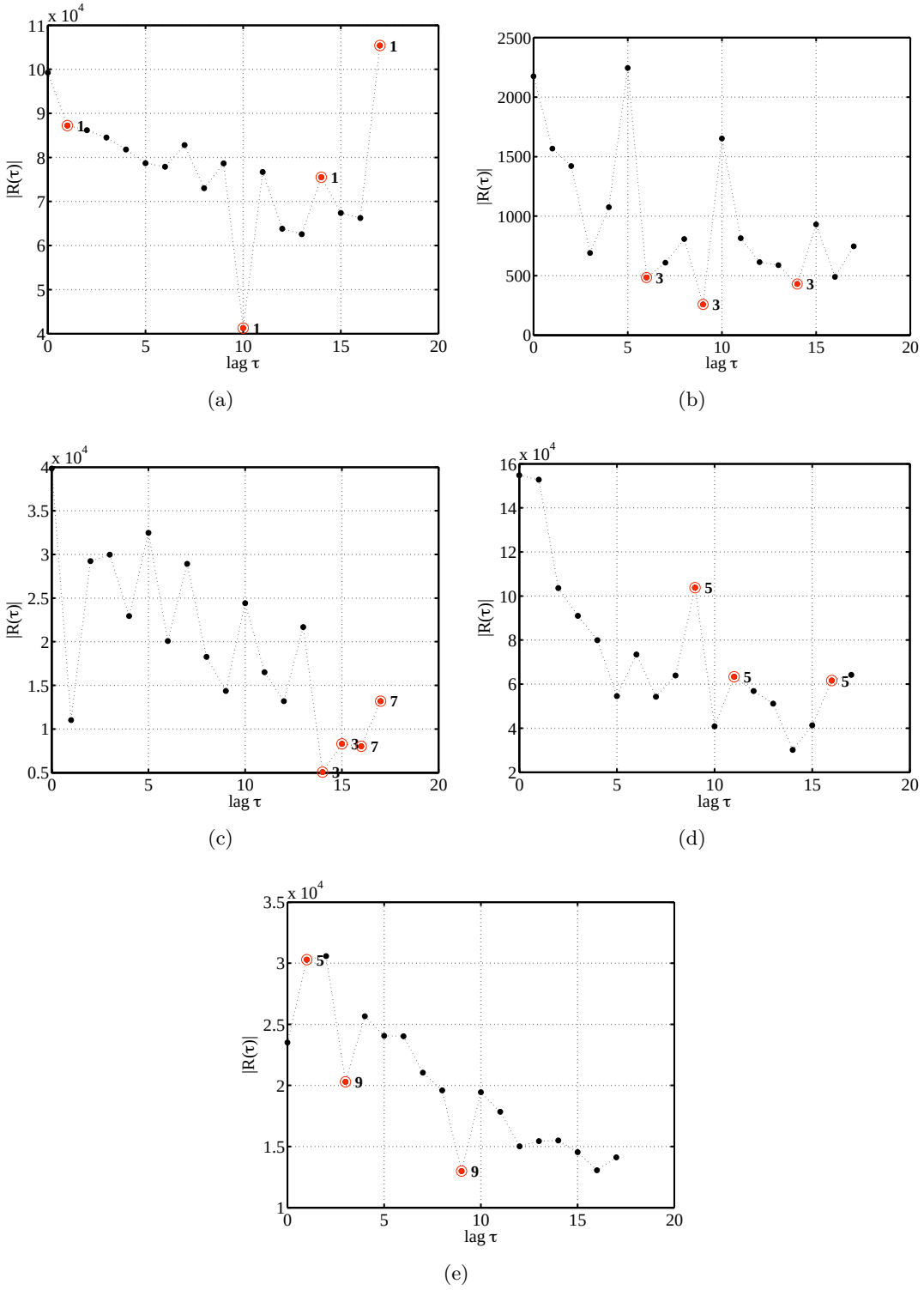


FIG. A.3 – *Cinq exemples d'ACF. Les mauvais retards détectés par l'algorithme FitACF sont indiqués par un point rouge et le chiffre indique l'identifiant correspondant (cf tableau A.1).*

Actes de publication & communication

Publication (revue à comité de lecture) :

- **A statistical approach for identifying ionospheric footprint of magnetospheric boundaries from SuperDARN observations**
G. Lointier, T. Dudok de Wit, C. Hanuise, X. Vallières, J.-P. Villain, Annales Geophysicae, 26, 305-314, www.ann-geophys.net/26/305/2008/, 2008.

Communications :

Présentations orales :

- **A statistical characterization for quasi real-time tracking of magnetospheric regions with SuperDARN**
G. Lointier, C. Hanuise, T. Dudok de Wit, X. Vallières, J.-P. Villain
4-8th June 2007, WorkShop SuperDARN, Abashiri, Hokkaido, Japan

- **Monitoring ionospheric signatures of magnetospheric boundaries with SuperDARN : a statistical approach**
G. Lointier, T. Dudok de Wit, C. Hanuise, J.-P. Villain
5-9th June 2006, WorkShop SuperDARN, Chincoteague, United State

Présentations Posters :

- **Vers une modélisation en temps réel des régions magnétosphériques à partir du réseau de radars cohérents HF SuperDARN**
G. Lointier, T. Dudok de Wit, C. Hanuise, X. Vallières
24-28 mars 2008, Atelier PNST 2008, Obernai, France

- A statistical approach for characterizing the SuperDARN auto-correlation function
G. Lointier, T. Dudok de Wit, C. Hanuise, X. Vallières, J.-P. Villain
5-9th June 2007, WorkShop SuperDARN, Abashiri, Hokkaido, Japan

- Statistical reduction of the SuperDARN data to identify the ionospheric footprint of the magnetospheric regions
G. Lointier, T. Dudok de Wit, C. Hanuise, X. Vallières, J.-P. Villain
13-17th November 2006, 3rd European Space Weather Week (ESWW), Brussels, Belgium

- Identification des régions frontières magnétosphériques par le réseau de radars HF cohérents SuperDARN
G. Lointier, T. Dudok de Wit, C. Hanuise, J.-P. Villain
26-30 juin 2006, Semaine Française de l'Astrophysique (SF2A), Université de Jussieu, Paris, France

- Monitoring Magnetospheric boundaries using the statistical properties of the SuperDARN radars echoes
G. Lointier, T. Dudok de Wit, C. Hanuise, J.-P. Villain
14-18th November 2005, 2nd European Space Weather Week (ESWW), ESTEC/Noordwijk, Netherlands

Bibliographie

André, R. and Dudok de Wit, T. : Identification of the ionospheric footprint of magnetospheric boundaries using SuperDARN coherent HF radars, Planetary and space Science, 51, 813–820, 2003.

André, R., Hanuise, C., Villain, J.-P., and Cerisier, J.-C. : HF radars : Multifrequency study of refraction effects and localization of scattering, Radio Science, 32, 153–168, 1997.

André, R., Pinnock, M., and Rodger, A. S. : On the factor conditioning the Doppler spectral width determined from SuperDARN HF radars, International Journal of Geomagnetism and Aeronomy, 2, 77–86, 2000a.

André, R., Pinnock, M., and Rodger, A. S. : Indentification of the low-altitude cusp by Super Dual Auroral Radar Network radars : A physical explanation for the empirically signature, Journal of Geophysical Research, 105, 27 081–27 093, 2000b.

André, R., Pinnock, M., Villain, J.-P., and Hanuise, C. : Influence of magnetospheric processes on winter radar spectra characteristics, Annales Geophysicae, 20, 1783–1793, 2002.

Aubry, N., Holmes, P., Lumley, J. L., and Stone, E. : The dynamics of coherent structures in the wall region of a turbulent boundary layer., Journal of Fluid Mechanics, 192, 115–173, 1988.

Baker, K. B. : White paper on FITACF (update on FITACF for Sept. 2003), Private communication, 2003.

Baker, K. B. and Wing, S. : A new magnetic coordinate system for conjugate studies at high latitudes, Journal of Geophysical Research, 94, 9139–9143, 1989.

Baker, K. B., Greenwald, R. A., Walker, A. D. M., Bristow, P. F., and Zanetti, L. J. : A case study of plasma processes in the dayside cleft, Journal of Geophysical Research, 91, 3130–3144, 1986.

Baker, K. B., Dudeney, J. R., Greenwald, R. A., Pinnock, M., Newell, P. T., Rodger, A. S., Mattin, N., and Meng, C.-I. : HF radar signatures of the cusp and low-latitude boundary layer, Journal of Geophysical Research, 100, 7671–7695, 1995.

Baker, K. B., Rodger, A. S., and Lu, G. : HF-radar observations of the dayside magnetic merging rate : A Geospace Environment Modeling boundary layer, Journal of Geophysical Research, 102, 9603–9617, 1997.

Ballatore, P., Villain, J.-P., Vilmer, N., and Pick, M. : The influence of the interplanetary medium on SuperDARN radar scattering occurrence, Annales Geophysicae, 18, 1576–1583, 2001.

Bartel, J. : The standardized index, Ks, and the planetary index, Kp, IATIME Bulletin, 12b, 1949.

Barthes, L., André, R., Cerisier, J.-C., and Villain, J.-P. : Separation of multiple echoes using a high-resolution spectral analysis for SuperDARN HF radars, Radio Science, 33, 1005–1017, 1998.

Blanchard, G. T., Lyons, L. R., Samson, J. C., and Rich, F. J. : Locating the polar cap boundary from observations of 6300 A auroral emission, Journal of Geophysical Research, 100, 7855–7862, 1995.

Booker, H. G. : A theory of scattering by no-istropic irregularities with application to radar reflections from the aurora, Journal of Atmospheric and Terrestrial Physics, 8, 204–221, 1956.

Bowles, K. L. and Balsley, B. B. : Radio waves scattering in the ionosphere, Advances in Electronics and Electron Physics, 19, 55–176, 1963.

Broomhead, D. S. and King, G. P. : extracting qualitive dynamics from experimental data, Physica D, 20, 217–236, 1986.

Burch, J. L. : IMAGE mission overview, Space Science Reviews, 2000.

Burch, J. L., Mende, S. B., Mitchell, D. G., Moore, T. E., Pollock, C. J., Reinisch, B. W., Sandel, B. R., Fuselier, S. A., Gallagher, D. L., Green, J. L., Perez, J. D., and Reiff, P. H. : Views of Earth's Magnetosphere with the IMAGE Satellite, Science, 291, 619–624, 2001.

Cardoso, J. F. : Blind signal separation : statistical principles, Proceedings of the IEEE, 9, 2009–2025, 1998.

Cerisier, J.-C., Marchaudon, A., Bosqued, J.-M., McWilliams, K. A., Frey, H. U., Bouhram, M., Laakso, H., Dunlop, M., Föster, M., and Fazakerley, A. : Ionospheric signatures of plasma injections in the cusp triggered by solar wind pressure pulses, Journal of Geophysical Research, 110, A08 204, doi :10.1029/2004JA010962, 2005.

Chatfield, K. and Collins, A. J. : introduction to Multivariate Statistics., Chapman & Hall, London, 1995.

Chisham, G. and Freeman, M. P. : A technique for accurately determining the cusp-region polar cap boundary using SuperDARN HF radar measurements, Annales Geophysicae, 21, 983–996, 2003.

Chisham, G. and Freeman, M. P. : An investigation of latitudinal transitions in the SuperDARN Doppler spectral width parameter at different magnetic local time, Annales Geophysicae, 22, 1187–1202, 2004.

Chisham, G., Freeman, M. P., and Sotirelis, T. : A statistical comparison of SuperDARN spectral width boundaries and DMSP particle precipitation boundaries in the nightside ionosphere, Geophysical Research Letters, 31, L02 804, doi : 10.1029/2003GL019074, 2004.

Chisham, G., Freeman, M. P., Lam, M. M., Abel, G. A., Sotirelis, T., Greenwald, R. A., and Lester, M. : A statistical comparison of SuperDARN spectral width boundaries and DMSP particle precipitation boundaries in the afternoon sector ionosphere, Annales Geophysicae, 23, 3645–3654, 2005a.

Chisham, G., Freeman, M. P., Sotirelis, T., and Greenwald, R. A. : The accuracy of using the spectral width boundary measured in off-meridional SuperDARN HF radar beams as a proxy for the open-closed field line boundary, Annales Geophysicae, 23, 2599–2604, 2005b.

Chisham, G., Freeman, M. P., Sotirelis, T., Greenwald, R. A., Lester, M., and Villain, J.-P. : A statistical comparison of SuperDARN spectral width boundaries and DMSP particle precipitation boundaries in the morning sector ionosphere, Annales Geophysicae, 23, 733–743, 2005c.

Chisham, G., Lester, M., Milan, S. E., Freeman, M. P., Bristow, W. A., Grocott, A., McWilliams, K. A., Ruohoniemi, J. M., Yeoman, T. K., Dyson, P. L., Greenwald, R. A., Kikuchi, T., Pinnock, M., Rash, J. P. S., Sato, N., Sofko, G. J., Villain, J.-P., and Walker, A. D. M. : A decade of the Super Dual Auroral Radar Network (SuperDARN) : scientific achievements, new techniques and future directions, Surveys in Geophysics (Springer Ed.), 28, 33–109, doi :10.1007/s10712-007-9017-8, 2007.

Christensen, A. B., Paxton, L. J., Avery, S., Craven, J., Crowley, G., Humm, D. C., Kil, H., Meier, R. R., Meng, C.-I., Morrison, D., Ogorzalek, B. S., Straus, P., Strickland, D. J., Swenson, R. M., Walterscheid, R. L., Wolven, B., and Zhang, Y. : Initial observations with the Global Ultraviolet Imager (GUVI) in the NASA TIMED satellite mission, Journal of Geophysical Research, 108, 1451, doi :10.1029/2003JA009918, 2003.

Chuine, I., Yiou, P., Viovy, N., Seguin, B., Daux, V., and Le Roy Ladurie, E. : Grape ripening as a past climate indicator, Nature, 432, 289–290, 2004.

Cordier, L. and Bergmann, M. : Proper Orthogonal Decomposition : an overview, in : on post-processing of experimental and numerical data, Von Karman Institute for Fluid Dynamics, 2003.

Cowley, S. W. H. and Lockwood, M. : Excitation and decay of solar wind-driven flows in the magnetosphere-ionosphere system, Annales Geophysicae, 10, 103–105, 1992.

Danskin, D. W., Koustov, A. V., Ogawa, T., Nishitani, N., Nozawa, S., Milan, S. E., Lester, M., and Andre, D. : On the factor controlling occurence of F-region coherent echoes, Annales Geophysicae, 20, 1385–1397, 2002.

Darrouzet, F. : Etude de la magnétosphère terrestre par analyse multipoint des données de la mission CLUSTER. Contributions à la caractérisation des frontières et de la magnétosphère interne, Ph.D. thesis, Université d'Orléans, 2006.

Davis, T. N. and Sugiura, M. : Auroral electrojet activity index AE and its universal time variation, Journal of Geophysical Research, 71, 785–801, 1966.

Duda, R., Hart, P., and Stork, D. : Pattern Classification, A Wiley-Interscience Publication, 2 edn., 2001.

Dudeney, J. R., Rodger, A. S., Freeman, M. P., Picket, J., Scudder, J., Sofko, G., and Lester, M. : The nightside ionospheric response to IMF By changes, Geophysical Research Letters, 25, 2601–2604, 1998.

Dudok de Wit, T. and Auchère, F. : Multispectral analys of solar EUV images : linking temperature to morphology, Astronomy and Astrophysics, 466, 347–355, doi :10.1051/0004-6361:20066764, 2007.

Dungey, J. W. : Interplanetary magnetic field and auroral zones, Physical Review Letter, 6, 47–48, 1961.

Elsner, J. B. and Tsonis, A. A. : Singular spectrum analysis, a new tool in time series analysis, Plenum Press, New York, 1996.

Escoubet, C. P., Fehringer, M., and Goldstein, M. : Introduction The Cluster mission, Annales Geophysicae, 19, 1197–1200, provided by the SAO/NASA Astrophysics Data System, 2001.

Farley, D. T. : Multiple-pulse incoherent-scatter correlation function measurements, Radio Science, 7, 661–666, 1972.

Fejer, B. G. and Kelley, M. C. : Ionospheric irregularities, Review of Geophysics, 18, 1980, 1980.

Gerbrand, J. J. : On the relationships between SVD, KLT and PCA, Pattern Recognition, 14, 375–381, 1981.

Golovchanskaya, I. V., Ostapenko, A. A., and Koselov, B. V. : Relationship between the high-latitude electric and magnetic turbulence and the Birkeland field-algined currents, Journal of Geophysical Research, 111, A12 301, doi :10.1029/2006JA011835, 2006.

Golub, G. and Van Loan, C. : Matrix Computations, The John Hopkins University Press, Baltimore, 4 edn., 1993.

Greenwald, R. A. : Space weather, SuperDARN and Tasmanian Tiger, Australian Journal of Physics, 50, 773–792, 1997.

Greenwald, R. A., Baker, K. B., and Hanuise, C. : An HF phased-array radar for studying small-scale structure in the high-latitude ionosphere, Radio Science, 20, 63–79, 1985.

Greenwald, R. A., Baker, K. B., Dudeney, J. R., Pinnock, M., Jones, T. B., Thomas, E. C., Villain, J.-P., Cerisier, J.-C., Senior, C., Hanuise, C., Hunsucker, R. D., Sofko, G., Koehler, J., Nielsen, E., Pellinen, R., Walker, A. D. M., Sato, N., and Yamagishi, H. : DARN/SUPERDARN : A Global View of the Dynamic of High-Latitude Convection, Space Science Reviews, 71, 761–796, 1995.

Hanuise, C. : High-latitude ionospheric irregularities : a review of recent results, Radio Science, 18, 1983.

Hanuise, C. and Crochet, M. : On new scattering modes for studying plasma instabilities with a HF radar system, Geophysical Research Letters, 7, 556–558, 1980.

Hanuise, C., Greenwald, R. A., and Baker, K. B. : Drift motions of small-scale irregularities in the high-latitude F region : an experimental comparison with plamas drift motions, Journal of Geophysical Research, 90, 9717–9725, 1985.

Hanuise, C., Villain, J.-P., Cerisier, J.-C., Senior, C., Ruohoniemi, J. M., Greenwald, R. A., and Baker, K. B. : Statistical study of high-latitude E-region Doppler spectra obtained with SHERPA HF radar, Annales Geophysicae, 9, 273–285, 1991.

Hanuise, C., Villain, J.-P., Gresillon, D., Cabrit, B., and Baker, K. B. : Interpretation of HF radar ionopsheric Doppler spectra by collective wave scattering theory, Annales Geophysicae, 11, 29–39, 1993.

Hanuise, C., Cerisier, J.-C., Auchère, F., Bocchialini, K., Bruinsma, S., Cornilleau-Wehrlin, N., Jakowski, N., Lathuillère, C., Menvielle, M., Valette, J.-J., Vilmer, N., Watermann, J., and Yaya, P. : From the Sun to the Earth : impact of the 27-28 May 2003 solar events on the magnetosphere, ionosphere and thermosphere, Annales Geophysicae, 24, 129–151, 2006.

Heppner, J. P. and Maynard, N. C. : Empirical high-latitude electric fiel models, Journal of Geophysical Research, 92, 4467–4489, 1987.

Holzworth, R. and Meng, C.-I. : Mathematical representation of the auroral oval, Geophysical Research Letters, pp. 377–380, 1975.

Hosokawa, K., Woodfield, E. E., Lester, M., Milan, S. E., Sato, N., Yukimatu, A. S., and Iyemori, T. : Interhemispheric comparison of spectral width boundary as observed by SuperDARN radars, Annales Geophysicae, 21, 1553–1565, 2003.

Hubert, B., Milan, S. E., Grocott, A., Blockx, C., Cowley, S. W. H., and Gérard, J.-C. : Dayside and nightside reconnection rates inferred from IMAGE FUV and Super Dual Auroral Radar Network data, Journal of Geophysical Research, 111, A03 217, doi :10.1029/2005JA011140, 2006.

Hubert, D., Harvey, C. C., Roth, M., and De Keyser, J. : Electron density at the subsolar magnétopause for high magnetic shear : ISEE 1 and 2 observations, Journal of Geophysical Research, 103, 6685–6692, 1998.

Hyvärinen, A., Karhunen, J., and Oja, E. : Independant component analysis, Wiley - InerScience, 2001.

Jackel, B. J. : Characterization of auroral power spectra and autocorrelation functions, Radio Science, 35, 1009–1023, 2000.

Jeffreys, H. : Theory of probability, 3rd, Oxford University Press, 1961.

Kamide, Y. and Chian, A. C.-L., eds. : Handbook of the Solar-Terrestrial Environment, vol. XIV, Springer, 2007.

Kelley, M. C. : The earth's ionosphere, Academic Press, London, 1989.

Kivelson, M. G. and Russel, C. T. : Introduction to Space Physics, Cambridge University Press, New York, 1995.

Lester, M., Milan, S. E., Besser, V., and Simth, R. : A case study of HF radar spectra and 630 nm auroral emission in the pre-midnight sector, Annales Geophysicae, 19, 327–339, 2001.

Lockwood, M., Carlson, H. C., and Sandholt, P. E. : Implications of the altitude of transient 630-nm dayside auroral emissions, Journal of Geophysical Research, 98, 15 571–15 587, 1993.

Lointier, G., Dudok de Wit, T., Hanuise, C., and Villain, J.-P. : A statistical approach for identifying the ionosphéric footprint of magnetospheric boundaries from SuperDARN observation, Annales Geophysicae, 26, 305–314, www.ann-geophys.net/26/305/2008/, 2008.

Mayaud, P. N. : Indices Kn, Ks, Km, Centre National de la Recherche Scientifique, Paris, pp. 1964–1967, 1968.

Mayaud, P. N. : A handred year series of geomagnetic data, 1868-1967 : indices aa, storm sudden commencements, IUGG Publ. Office, Paris, p. 256, 1973.

Menvielle, M., Lathuillère, C., Bruinsma, S., and Viereck, R. : A new method for studying the thermospheric density variability derived form CHAMP/STAR accelerometer data for magnetically active conditions, Annales Geophysicae, 25, 1949–1958, http://direct.sref.org/1432-0576/angeo/2007-25-1949, 2007.

Milan, S. E., Yeoman, T. K., Lester, M., Thomas, E. C., and Jones, T. B. : Initial backscatter occurence statistics from the CUTLASS HF radars, Annales Geophysicae, 15, 703–718, 1997.

Milan, S. E., Lester, M., Cowley, S. W. H., Moen, J., Sandholt, P. E., and Owen, C. J. : Meridian-scanning photometer, coherent HF radar, and magnetometer observations of the cusp : a case study, Annales Geophysicae, 17, 159–172, 1999.

Moen, J., Carlson, H. C., Milan, S. E., Shumilov, N., Lybekk, K., Sandholt, P. E., and Lester, M. : On the collocation between dayside auroral actvity and coherent HF radar backscatter, Annales Geophysicae, 18, 1531–1549, 2001.

Moorcroft, D. R. : The shape of the auroral backscatter spectra, Geophysical Research Letters, 31, L09 802, doi :10.10129/2003GL019340, 2004.

Moses, J. J. and Reiff, P. H. : Empirical convection models for northward IMF, Journal of Atmospheric and Terrestrial Physics, 56, 195–207, 1994.

Nevanlinna, H. and Kataja, E. : An extension of the geomagnetic activity index series aa for two solar circles, Geophysical Research Letters, 20, 2703–2706, 1993.

Newell, P. T. and Meng, C.-I. : Ionospheric projections of magnetospheric regions under low and high solar wind pressure condtions, Journal of Geophysical Research, 99, 273–286, 1994.

Newell, P. T., Meng, C.-I., Sibeck, D. G., and Lepping, R. : Some low-altitude cusp dependencies on the interplanetary magnetic field, Journal of Geophysical Research, 99, 8921–8927, 1989.

Newell, P. T., Sotirelis, T., Ruohoniemi, J. M., Carbary, J. F., Liou, K., Skura, J. P., Meng, C.-I., Deehr, C., Wilkinson, D., and Rich, F. J. : OVATION : Oval variation, assesment, tracking, intensity, and online nowcasting, Annales Geophysicae, 20, 1039–1047, 2002.

Newell, P. T., Ruohoniemi, J. M., and Meng, C.-I. : Maps of precipitation by source region, binned by IMF, with inertial convection streamlines, Journal of Geophysical Research, 109, A10 206, doi :10.1029/2004JA010499, 2004.

Parkinson, M. L., Dyson, P. L., Pinnock, M., Devlin, J. C., Hairston, M. R., Yizengaw, E., and Wilkinson, P. J. : Signatures of the midnight open-closed magnetic field line boundary during balanced dayside and nightside reconnection, Annales Geophysicae, 20, 1616–1630, 2002.

Pinnock, M., Rodger, A. S., Dudeney, J. R., Rich, F., and Baker, K. B. : High spatial end temporal resolution observations of the ionospheric cusp, Annales Geophysicae, 13, 919–925, 1995.

Pinnock, M., Rodger, A. S., Baker, K. B., Lu, G., and Hairston, M. : Conjugate observations of the day-side reconnection electric field : A GEM boundary layer campaign, Annales Geophysicae, 17, 443–454, 1999.

Ponomarenko, P. V. and Waters, C. L. : Spectral width of SuperDARN echoes : measurement, use and physical interpretation, Annales Geophysicae, 24, 115–128, 2006.

Preisendorfer, R. W. : Principal Component Analyses in Meteorology and Oceanography, Elsevier Science Publishing, 1988.

Rodger, A. S. : Ground-based imaging of magnetospheric boundaries, Advances in Space Research, 25, 1461–1470, 2000.

Rodger, A. S., Mende, S. B., Rosenberg, T. J., and Baker, K. B. : Simultaneous optical and HF radar observations of the ionospheric cusp, Geophysical Research Letters, 22, 2045–2048, 1995.

Ruohoniemi, J. M. and Baker, K. B. : Large-scale imaging of high-latitude convection with Super Dual Auroral Radar Network HF radar observations, Journal of Geophysical Research, 103, 20 797–20 811, 1998.

Ruohoniemi, J.-M. and Greenwald, R. A. : Rates of scattering ocurrence in routine HF radar observations during solar cycle maximum, Journal of Geophysical Research, 32, 1051–1070, 1997.

Ruohoniemi, J. M., Greenwald, R. A., Baker, K. B., Villain, J.-P., and McCready, M. A. : Drift motions of small-scale irregularities in the high-latitude F region : an experimental comparison with plamas drift motions, Journal of Geophysical Research, 92, 4553–4564, 1987.

Ruohoniemi, J. M., Barnes, R. J., Greenwald, R. A., and Shepherd, S. G. : The response of the high-latitude ionosphere to the coronal mass ejection event of April 6, 2000 : A practical demonstration of space weather nowcasting with the Super Dual Auroral Radar Network HF radars, Journal of Geophysical Research, 106, 30,085–30,097, 2001.

Santolík, O., Nĕmec, F., Parrot, M., Lagoutte, D., and Madrias, L. end Berthelier, J.-J. : Analysis methods for multi-component wave measurements on board the DEMETER spacecraft, Planetary and space Science, 54, 512–527, 2006.

Smith, M. F. and Lockwood, M. : Earth's magnetospheric cusps, Review of Geophysics, 34, 233–260, 1996.

Stewart, G. W. : On the early history of the singular value decomposition, SIAM Rev., 35, 551–566, doi :http://dx.doi.org/10.1137/1035134, 1993.

Sugiura, M. : Hourly values of equatorial Dst for the IGY, Annales of International Geophysical Year, 35, 9–45, 1965.

Svensmark, H. : Cosmic Rays and Earth's Climate, Space Science Reviews, 93, 175–185, doi :10.1023/A:1026592411634, 2000.

Vallières, X., Villain, J.-P., Hanuise, C., and André, R. : Ionospheric propagation effects on spectral widths measured by SuperDARN HF radars, Annales Geophysicae, 22, 2023–2031, 2004.

Villain, J.-P., Caudal, G., and Hanuise, C. : A SAFARI-EISCAT comparison between the velocity of F region small-scale irregularities and the ion drift, Journal of Geophysical Research, 90, 8433–8443, 1985.

Villain, J.-P., Greenwald, R. A., Baker, K. B., and Ruohoniemi, J. M. : HF Radar observations of E Region Plasma Irregularities Produced by Oblique Electron Streaming, Journal of Geophysical Research, 92, 12,327–12,342, 1987.

Villain, J.-P., André, R., Hanuise, C., and Grésillon, D. : Observations of the high latitude ionosphere by HF radars : interpretation in terms of collective wave scattering and characterization of turbulence, Journal of Atmospheric and Terrestrial Physics, 58, 943–958, 1996.

Villain, J.-P., André, R., Pinnock, M., Greenwald, R. A., and Hanuise, C. : A Statistical study of the Doppler spectral width of high-latitude ionospheric F-region echeos recorded with SuperDARN coherent HF radars, Annales Geophysicae, 20, 1769–1781, 2002.

Wing, S., Greenwald, R. A., Meng, C.-I., Sigillito, V. G., and Hutton, L. V. : Neural networks for automated classification of ionospheric irregularities in HF radar backscattered signals, Radio Science, 38, 1063, doi :10.1029/2003RS002869, 2003.

Woodfield, E. E., Davies, J. A., Eglitis, P., and Lester, M. : A case study of HF radar spectral width in the post midnight magnetic local time sector and its relationship to the polar cap boundary, Annales Geophysicae, 20, 501–509, 2002a.

Woodfield, E. E., Davies, J. A., Lester, M., Yeoman, T. K., Eglitis, P., and Lockwood, M. : Nightside studies of coherent HF Radar spectral width behaviour, Annales Geophysicae, 20, 1399–1413, 2002b.

Woodfield, E. E., Hosokawa, K., Milan, S. E., and Lester, M. : An inter-hemispheric, statistical study of nightside spectral width distributions from HF scatter radars, Annales Geophysicae, 20, 1921–1934, 2002c.

Yukimatu, A. S. and Tsutsumi, M. : A new SuperDARN meteor wind measurement : Raw time series analysis method and its application to mesopause region dynamics, Geophysical Research Letters, 29, 42–1, doi :10.1029/2002GL015210, 2002.

Vers une surveillance en temps réel des régions magnétosphériques à partir des radars cohérents HF SuperDARN

Résumé

L'un des enjeux dans la compréhension des relations Soleil-Terre est l'étude de la dynamique des régions internes de la magnétosphère.

Dans ce contexte, l'objectif de mon travail a été de poser les bases d'un modèle opérationnel de détection et de suivi des régions frontières de la magnétosphère à partir du réseau de radars SuperDARN, qui sonde l'ionosphère à haute latitude. L'élaboration d'un tel modèle requiert une réduction de données. Pour cela, une méthode de décomposition en valeurs singulières (SVD) a été appliquée sur la mesure brute (une fonction d'autocorrélation) afin de définir trois nouveaux paramètres statistiques. L'interprétation de ces trois nouveaux paramètres montre que cette approche, bien qu'empirique, offre une description bien plus complète des échos radar que les modèles physiques habituellement utilisés. L'utilisation de ces paramètres avec une méthode de décision Bayésienne permet d'améliorer la détection de la frontière des lignes de champ géomagnétiques ouvertes/fermées. L'introduction d'un formalisme Bayésien comporte plusieurs avantages : il permet de valider le résultat en estimant une erreur sur la localisation, et de plus, il facilite l'introduction de nouvelles connaissances provenant de différents instruments. Ceci est loin d'être négligeable pour compléter les observations des radars SuperDARN. Par ailleurs, les propriétés de ce nouveau modèle ont permis l'élaboration d'un nouvel algorithme de prétraitement et d'analyse des mesures brutes.

Mots-clefs: météorologie de l'espace, couplage magnétosphère-ionosphère, analyse statistique multivariée, classification Bayésienne, radars cohérents HF.

Toward a real time monitoring of magnetospheric regions using the SuperDARN coherent HF radar network

Abstract

One of the key issues in the study of the Sun-Earth connexion is the monitoring of the dynamics of the magnetospheric regions. In this context, the objective of my work is to lay the foundations for an operational model for monitoring the ionospheric projection of magnetospheric boundaries in near real-time. This model is based on the SuperDARN radar network, which was designed for the study of high latitude ionospheric convection. The elaboration of such a model requires a data reduction step. We use the Singular Value Decomposition (SVD) method to reduce the raw data (autocorrelation functions) to a set of three statistical parameters. The interpretation of these empirical parameters shows that they provide a more complete description of the radar echoes than the routinely employed physical parameters. By coupling these parameters to a Bayesian decision method, we show how the detection of the open/closed magnetic field line boundary can be improved. The Bayesian formalism offers several advantages : it provides confidence intervals for the location of the boundary and it can handle prior information, as well as constraints. In addition to that, a new strategy has been introduced for correcting outlier values that often plague the measurement of the autocorrelation function.

Keywords: space weather, magnetosphere-ionosphere coupling, multivariate statistical analysis, Bayesian classification, HF coherent radars.

Discipline - spécialité: physique des plasmas

Intitulé et adresse du laboratoire
Laboratoire de Physique et Chimie de l'Environnement (LPCE) / CNRS
3A, avenue de la Recherche Scientifique
45 071 ORLÉANS CEDEX 02
FRANCE