HAL
open science

Intégration des données d'observatoires magnétiques dans l'interprétation de sondages magnétotelluriques : acqusition, traitement, interprétation

Hugo Larnier

To cite this version:

Hugo Larnier. Intégration des données d'observatoires magnétiques dans l'interprétation de sondages magnétotelluriques : acqusition, traitement, interprétation. Géophysique [physics.geo-ph]. Université de Strasbourg, 2017. Français. NNT : 2017STRAH003 . tel-02268559

HAL Id: tel-02268559

https://theses.hal.science/tel-02268559

Submitted on 21 Aug 2019

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.

UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE 413 – SCIENCES DE LA TERRE ET DE L'ENVIRONNEMENT

[Institut de Physique du Globe de Strasbourg, CNRS UMR 7516]

THÈSE présentée par :

[Hugo LARNIER]

soutenue le : **07 Février 2017**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Géophysique

Intégration des données d'observatoires magnétiques dans l'interprétation de sondages magnétotelluriques :

Acquisition, Traitement, Interprétation

THÈSE dirigée par :

[M. SAILHAC Pascal]	Professeur, Université d'Orsay

RAPPORTEURS :

[M. GIBERT Dominique]	Physicien, Université de Rennes
[M. JUANJO Ledo]	Professeur, Université de Barcelone

AUTRES MEMBRES DU JURY :

[M. RIVERA Luis]	Professeur, Université de Strasbourg
[M. PERRIER Frédéric]	Professeur, Institut de Physique du Globe de Paris
[Mme. CHAMBODUT Aude]	Physicien Adjoint, Université de Strasbourg

Résumé Français

Dans ce manuscrit, nous développons des méthodologies de détection et caractérisation de sources géomagnétiques et atmosphériques en se basant sur la transformée en ondelettes continues. Les techniques introduites se basent sur les caractéristiques temps-fréquence des ondes observées dans les séries temporelles magnétotelluriques (MT). Nous détaillons en particulier leurs évolutions temporelles et spatiales et l'implication dans les procédures d'acquisitions de données MT. A partir de ces procédures de détection, nous détaillons l'implémentation d'une stratégie de détermination des fonctions de réponse MT basée sur les statistiques robustes, et du bootstrap hiérarchique pour le calcul des incertitudes sur les sondages MT. Nous discutons l'apport de ces méthodes sur l'amélioration des sondages MT.

Deux études de sondages MT sont également détaillées. La première étude MT concerne la caractérisation de la structure géoélectrique situé sous l'observatoire magnétique de Chambon-La-Forêt, France. La seconde étude concerne des mesures effectuées dans la vallée de Trisuli au Népal en mars 2016, faisant suite au séisme de Gorkha du 25 Avril 2015. L'objectif de cette campagne est la comparaison avec une étude effectuée en 1996. Nous discutons des effets topographiques sur les sondages MT et leurs implications sur l'ancien profil de conductivité obtenu. Nous présentons également une nouvelle interprétation de la distribution de conductivité dans le sous-sol de vallée de Trisuli.

Mots clés : Magnétotellurique, Géomagnétisme, Observatoires magnétiques, Transformée en ondelette continue, Statistiques robustes, Chambon-La-Forêt, Népal.

Résumé anglais

In this manuscript, we detail the application of continuous wavelet transform to processing schemes for the detection and the characterisation of geomagnetic and atmospheric sources. Presented techniques are based on time-frequency properties of electromagnetic (EM) waves observed in magnetotellurics (MT) time series. We particularly detail temporal and spatial variations of EM waves and their implication on the acquisition of MT data. We detail the application of these detection procedures in a MT processing scheme. To recover MT response functions, we use robust statistics and a hierarchical bootstrap approach for uncertainties determination. We discuss the pros and cons on the improvement of MT soundings.

Interpretation of two datasets are also presented. The first MT study deals with the caracterisation of the resistivity distribution below the French National magnetic observatory of Chambol-la-Forêt. The second study details the interpretation of new MT soundings acquired in March 2016 in the Trisuli valley, Nepal, following the Gorkha earthquake of April 25, 2015. The main objective of this campaign was to compare the new soundings with an old campaign in 1996. We discuss topography effects on MT soundings and their implication on the resistivity distribution. We also introduce a new interpretation of the resistivity distribution in Trisuli valley.

Keywords : Magnetotellurics, Geomagnetism, Magnetic Observatories, Continuous wavelet transform, Robust statistics, Chambon-La-Forêt, Népal.

Remerciements

Après trois années qui sont passées à la vitesse de l'éclair, il est temps de clôturer ce manuscrit.

Tout d'abord, un grand merci à mes deux encadrants de thèse, Pascal Sailhac et Aude Chambodut aussi bien pour leur accompagnement scientifique tout au long de ces trois années, que leur encouragements constants qui m'ont permis de mener ce travail jusqu'au bout.

Pascal, tu m'as toujours fait confiance au point d'attendre plusieurs mois pour voir mes résultats sans me mettre la pression. Mais toujours en remettant en cause ce que je te montrais pour en être parfaitement convaincu. Ces qualités me seront certainement utiles dans la suite ! Tu m'as aussi permis de rencontrer les bonnes personnes pour discuter et m'a toujours mis en avant pour que je me fasse connaitre. J'espère sincèrement que l'on pourra continuer à travailler ensemble par la suite. Merci pour ces trois ans et amuse-toi bien à l'Université d'Orsay !

Aude, tu m'as (presque) directement considéré comme un collègue plus que comme un doctorant dès le départ de cette thèse (et même avant en stage de recherche), ce qui m'a mis en confiance. Même si tu est toujours overbookée, tu as toujours aussi eu ta porte ouverte si il y avait une nécessité de discuter. A la fin, avec le stress de fin de thèse, ça a été critique pour moi. Tu m'as aussi permis de bien voir les rouages humains de la recherche scientifique au travers de tes remarques en congrès et sur mes corrections, ce qui je pense devrait me servir par la suite, aussi bien dans le domaine professionnel que personnel.

Un grand merci aux membres du jury qui se sont déplacés parfois d'assez loin pour évaluer mon travail de thèse. Malgré les 300 pages et le peu de temps disponible pour lire ce manuscrit (la prochaine fois que je fais une thèse, je ne rends pas les manuscrits en décembre promis), vous avez rendu les documents à temps, ce qui enlève un poids considérable au jeune thésard qui espère soutenir !

Je remercie aussi l'ensemble du couloir du troisième étage pour leur accueil tout au long de ces trois ans, Jean-Paul Boy, Maksim Bano, Séverine Rosat, Jean-Michel Marthelot, Matthias Zillmer ; Yves Rogister, "alors, t'as fini de rédiger ?" ; Patrice, Pierre-Daniel Matthey, Caroline Lazarus, Marc Schaming... Je ne vais pas citer l'ensemble du personnel du troisième étage, mais sachez que ça a été un plaisir de bosser ici parmi vous !
Parmi les non-troisième étage, un énorme merci à Magali Pierrat sans qui j'aurais probablement un procès pour non rendu de livres scientifiques en temps et en heure, Jean-Marc Brendle pour son soutien informatique (récuperer des données vieilles de vingt ans, c'est une prise de tête), Antoine Bur, Samantha Borsoni et Muriel Eichhorn pour leur soutien logistique dans l'organisation des cours, Dilek Karayigit pour ses explications sur les ordres de mission (clair comme de l'eau de roche).

Finissons par le meilleur, les personnes qui m'ont permis de finir sans partir en hurlant. Commençons par le meilleur bureau que j'ai jamais connu jusqu'à maintenant (prenez pas la grosse tête, c'est mon premier). Un gros BIG UP à ma colocataire de bureau depuis presque trois ans, Floriane Provost, qui a toujours été là dans les bons moments et les mauvais, ça m'embête de partir en te laissant, même si je sais que la relève du bureau 304 (et à côté) est assurée et te permettra de décompresser pendant ta dernière année. Aline Deprez, partie au milieu de ma thèse, que je comprends beaucoup mieux maintenant que j'ai vécu la période de rédaction. Toutes les deux, avec le craquage du vendredi, vous étiez terribles en co-bureaux. Merci aux nouveaux co-bureaux qui sont arrivés, Emilie Klein et Eric Henrion. C'était un vrai plaisir de partager ce bureau avec vous !
Non loin du bureau 304, merci à Yann Ziegler pour nos délires sans queue ni-tête d'après midi (de toute façon t'étais pas là le matin), Jean-Rémi Dujardin, Semih Turkaya.
Merci à Charlène Aldebert (check wesh) et à Edouard Alkange pour leur soutien indéfectible pendant ces trois ans et qui m'ont empêché de passer mes soirées au travail (merci le mal de crâne le lendemain). Merci aux irréductibles potes d'enfances, Damien Barranger, Nicolas Bel, Charly Martin, Fabien Mennessier, Tristan Prudhomme. Ca faisait du bien les pauses avec vous.
Le plus grand soutien (26 ans quand même) vient probablement de mes parents qui ont été là tout du long de ma scolarité, avec ses hauts et ses bas. Ca y est, c'est fini, vous pouvez souffler, le plus dur va commencer maintenant pour le fiston. Merci indéfiniment pour tout.

Enfin, finissons par la meilleure, même si je ne vais pas développer ici, *merci Camille*.

Table des matières

Résumé ii

Abstract iii

Remerciements iv

Sommaire vi

Table des Figures xi

Liste des Tableaux xxiii

Introduction 1

I Eléments de magnétotellurique 6

1 Induction dans un milieu conducteur par une onde monochromatique 7
1.1 Milieu homogène . . . 7
1.1.1 Régime général dans un milieu homogène . . . 7
1.1.2 Régime diffusif . . . 10
1.1.3 Diffusion dans un milieu homogène . . . 11
1.1.3.1 Diffusion . . . 11
1.1.3.2 Relation entre champ magnétique et électrique . . . 11
1.2 Diffusion dans un milieu unidimensionnel . . . 12
1.3 Milieux avec variations latérales de conductivité . . . 14
1.3.1 Tenseur d'impédance . . . 14
1.3.2 Fonction de transfert verticale . . . 15

II Traitement de données magnétotelluriques 18

2 Détection et caractérisation des sources géomagnétiques 19
2.1 Environnement électromagnétique . . . 19
2.1.1 Eléments du champ interne . . . 19
2.1.2 Eléments du champ externe . . . 21
2.1.2.1 Ionosphère . . . 21
2.1.2.2 Magnétosphère . . . 22
2.2 Ondes haute fréquence (> Hz) . . . 24

2.2.1 Origine . . . 24
2.2.2 Formes des ondes . . . 24
2.2.2.1 *Sferic* - Ondes VLF (*Very Low Frequency*) . . . 24
2.2.2.2 *Slow tails* - Ondes ELF (*Extremely Low Frequency*) . . . 25
2.2.2.3 *Whistlers* . . . 26
2.2.2.4 Autres ondes utilisables . . . 27
Résonances de Schumann . . . 27
Q-Bursts . . . 28
2.2.3 Intérêt en MT . . . 30
2.2.4 Article GJI - Détection et caractérisation des sources hautes fréquences. . . 30
2.2.5 Orages proches . . . 61
2.3 Ondes basses fréquences . . . 62
2.3.1 Pulsations géomagnétiques . . . 62
2.3.1.1 Mécanismes de génération . . . 64
2.3.1.2 Caractéristiques spatio-temporelles . . . 65
2.3.2 Détection - Potential paper . . . 68
2.3.3 Importance des observatoires magnétiques dans la détection . . . 84
2.4 Conclusion . . . 85

3 Calcul des fonctions de transfert 86
3.1 Introduction . . . 86
3.2 Statistiques robustes . . . 87
3.2.1 Segmentation des séries temporelles . . . 87
3.2.2 Régression linéaire et moindres carrés . . . 87
3.2.3 Statistiques robustes . . . 88
3.2.3.1 Espérance . . . 89
3.2.3.2 Incertitude sur la valeur des fonctions de réponse . . . 93
3.3 Utilisation d'une station de référence . . . 94
3.4 Utilisation de l'analyse en ondelette . . . 95
3.4.1 Inversion des coefficients . . . 95
3.4.2 Méthodologie . . . 96
3.4.2.1 Stockage des coefficients . . . 96
3.4.2.2 Utilisation de combinaisons d'événements . . . 96
3.4.2.3 Article EPS - Réduction du biais sur l'impédance en station simple 97
3.4.2.4 Transition vers le bootstrap . . . 108
3.4.2.5 Bootstrap et tirage aléatoire . . . 108
Présentation . . . 108
Application à la magnétotellurique : Problème de régression . . . 110
Bootstrap hiérarchique . . . 111
3.4.3 Applications . . . 112
3.4.3.1 Eclairs atmospherics . . . 112
3.4.3.2 Article GJI - Détermination de la réponse MT par bootstrap . . 112
3.4.3.3 Pulsations géomagnétiques . . . 144
Etude du cas des données MT de Schwabwiller - Données peu bruitées 144
3.5 Conclusions et perspectives . . . 151

III Interprétation et inversion de sondages magnétotelluriques 152

4 Chambon La Forêt 153
4.1 Situation géographique de l'observatoire . . . 153
4.1.1 Observatoire magnétique national de Chambon-La-Forêt (CLF) . . . 153
4.1.2 Géologie du bassin parisien . . . 155
4.1.3 Etudes géophysiques précédentes dans la région de CLF . . . 155
4.2 Présentation de la campagne . . . 157
4.2.1 Mesures utilisées . . . 157
4.2.1.1 Dipôles électriques permanents . . . 157
4.2.1.2 Mesures MT . . . 157
4.3 Bruit anthropique autour de l'observatoire . . . 158
4.4 Traitement et interprétation des données MT . . . 162
4.4.1 Traitement des données MT . . . 162
4.4.2 Dimensionnalité des sondages MT . . . 162
4.4.3 Modélisation . . . 166
4.4.3.1 1D . . . 166
4.4.3.2 3D . . . 169
Effet topographique . . . 169
Effet de l'océan . . . 171
4.4.3.3 2D . . . 172
4.5 Conclusion . . . 174

5 Interprétation de données MT au Népal 176
5.1 Introduction . . . 176
5.1.1 Contexte géologique . . . 176
5.1.1.1 Collision Inde-Eurasie . . . 176
5.1.1.2 Structure de la chaîne Himalayenne . . . 177
5.1.2 Le séisme de Gorkha . . . 179
5.1.3 Travaux antérieurs en magnétotellurique sur la chaîne Himalayenne . . . 180
5.2 Campagne magnétotellurique 2016 . . . 181
5.3 Traitement des données . . . 183
5.4 Analyse des sondages . . . 185
5.4.1 Comparaison avec les anciennes campagnes . . . 185
5.4.2 Distorsions du champ électrique et dimensionnalité . . . 189
5.5 Modélisation des effets topographiques . . . 195
5.5.1 Modélisation et inversion d'effets 2D . . . 195
5.5.2 Modélisation 3D . . . 200
5.6 Inversion du profil 2D . . . 202
5.6.1 Présentation de l'ancien profil . . . 202
5.6.2 Inversion des données de la campagne 2016 . . . 205
5.6.3 Tests de sensibilité . . . 207
5.6.3.1 Aux stations . . . 207
5.6.3.2 Variation de la résolution verticale et horizontale . . . 212
5.7 Interprétation . . . 213
5.7.1 Comparaison entre les différents modèles obtenus et les données . . . 213
5.7.2 Lien avec d'autres études géophysiques . . . 215

5.8 Conclusion . . . 216
5.9 Annexes . . . 218
5.9.1 Comparaison entre les campagnes de 1996 et 2016 . . . 218
5.9.2 Décompositions de Groom-Bailey . . . 221
5.9.2.1 Station simple, par période . . . 221
5.9.2.2 Station simple, par période, strike fixé à N20E . . . 222
5.9.2.3 Multi-station, pour les périodes > 1s . . . 223

IV Conclusion **224**

6 Conclusion générale **225**

Bibliographie **228**

A Données utilisées et traitements appliqués **238**
A.1 Données d'observatoires . . . 238
A.1.1 Présentation . . . 238
A.1.2 Matériel . . . 239
A.1.3 Type de données . . . 240
A.2 Données MT . . . 240
A.2.1 Matériel . . . 240
A.2.1.1 Mesure du champ électrique . . . 240
A.2.1.2 Mesure du champ magnétique . . . 241
A.2.1.3 Procédures d'acquisition . . . 241
A.3 Filtrage de données MT . . . 242

B Transformée en ondelettes **244**
B.1 Introduction sur le temps-fréquence . . . 244
B.1.1 Représentations classiques . . . 244
B.2 Transformée en ondelettes continues - 1D . . . 245
B.2.1 Fonctions d'ondelettes . . . 245
B.2.2 Ondelettes mère . . . 246
B.2.3 Ondelettes filles . . . 248
B.2.4 Calcul des coefficients . . . 249
B.3 Noyau d'ondelette et distance de corrélation . . . 250
B.4 Autres quantités définies à partir des coefficients . . . 250
B.4.1 Coefficients de corrélation croisée . . . 250
B.4.2 Spectre global d'ondelette . . . 251

C Polarisation **252**
C.1 Attributs de polarisation . . . 252
C.1.1 Domaine temporel . . . 252
C.1.2 Domaine de Fourier . . . 255
C.1.3 Coefficients de polarisation instantanée . . . 260

D Outils d'interprétation **263**
D.1 Etude des distortions . . . 263
D.1.1 Tenseur de phase . . . 264
D.1.2 Décomposition de Groom-Bailey . . . 265
D.2 Codes de modélisation et d'inversion . . . 267
D.2.1 ModEM . . . 268
D.2.2 Mare2DEM . . . 269

E Détail des données **271**
E.1 Mission Népal . . . 271
E.1.1 Calendrier d'acquisition . . . 271
E.1.2 Détail des mesures . . . 271
E.1.3 Détail des points mesurés . . . 274
E.2 Chambon-La-Forêt . . . 285
E.2.1 Détail des mesures . . . 285

Table des figures

1 Table de résistivité de différents matériaux géologiques. D'après la présentation d'EM donnée par Constable, S à la 2015 *IAGA Summer School*. 1

1 *Resistivity values for geological materials. From the EM course given by Steven Constable, at the 2015 IAGA Summer School.* 1

1.1 Représentation d'un milieu 1D dont la variation de résistivité est verticale. 13

1.1 *Representation of a 1D conductivity medium. Only vertical variation of resistivity is considered.* 13

1.2 Représentation classique des sondages magnétotelluriques. Seules les composantes anti-diagonales ont été représentées ici. Noir : Composante Z_{xy}. Gris : Composante Z_{yx}. Les barres verticales indiquent les incertitudes autour de l'estimation. 15

1.2 *Classical representation of MT soundings. Only off-diagonal components are shown. Black : Z_{xy}. Grey : Z_{yx}. Uncertainties are indicated by vertical bars.* 15

1.3 Haut : Le corps conducteur est indiqué par la lettre **C** et s'étend de façon infinie dans la direction horizontale perpendiculaire à Y. Bas : Évolution de la composante T_y le long du profil dans la direction Y. 16

1.3 *Top : The conductive body is the black box indicated by the letter **C** and is extending infinitely in the horizontal direction perpendicular to Y. Bottom : T_y variation along the profil in the Y direction.* 16

2.1 Spectre en puissance des variations naturelles du champ magnétique. D'après C. Constable (2016). 20

2.1 *Power spectrum of natural variations of the magnetic field. From C.* Constable *(2016)* 20

2.2 Profil en altitude de la densité électronique aux coordonnées 18°N, 67°W, enregistré à l'équinoxe de Septembre. D'après Richmond (2007). 21

2.2 *Vertical profil of electronic density recorded at September equinox at 18°N, 67°W. From* Richmond *(2007).* 21

2.3 Système de courant ionosphérique moyen pour l'année 1964–1965. Les isovaleurs sont indiquées en kA. D'après Richmond (2007). 22

2.3 *Annual mean of the ionospheric current system for year 1964. Values are indicated in kA. From* Winch *(2007).* 22

2.4 Schéma de la géométrie de la magnétosphère dans l'approximation du flux gelé. Les traits pleins indiquent les lignes de champ magnétique. Les pointillés les lignes de courant du plasma. D'après Cowley (1995). 23

2.4 *Sketch of magnetosphere geometry in frozen-flux approximation. Magnetic field lines are indicated by full lines. Dashed lines indicated plasma flux lines. From* Cowley *(1995)* 23

2.5 Schéma du cycle de Dungey. D'après Axford (1969). 24

2.5 *Sketch of Dungey cycle. From* Axford *(1969)* 24
2.6 Schéma représentant la propagation des ondes EM issues de l'occurrence d'un éclair atmosphérique. D'après Mackay et A. C. Fraser-Smith (2010). 25
2.6 *Sketch of atmospheric waves propagation from lightning strike occurence to measuring station. From* Mackay *et A. C.* Fraser-Smith *(2010).* 25
2.7 Haut : Onde VLF en domaine temporel pour un échantillonnage à 65536 Hz. Bas : Spectre en puissance des séries temporelles présentées au dessus. 26
2.7 *Top : Atmospheric wave in time domain for a sampling frequency of 65,536 Hz. Bottom : Power spectrums of the above time series.* 26
2.8 Haut : Onde ELF en domaine temporel pour un échantillonnage à 2048 Hz. Bas : Spectre de puissance des séries temporelles présentées au dessus. 27
2.8 *Top : Slow tail in time domain for a sampling frequency of 2,048 Hz. Bottom : Power spectrum of above time series.* 27
2.9 Haut : Onde *whistler* en domaine temporel sur le champ magnétique horizontal pour un échantillonnage à 65536 Hz. A noter la présence d'une large *slow tail* simultanée. Bas : Spectres de puissance des séries présentées au dessus. 28
2.9 *Top : Whistler in time domain on horizontal magnetic field for a sampling frequency of 65,536 Hz. Notice the presence of a large slow tail appearing simultaneously. Bottom : Power spectrums of above time series.* 28
2.10 Résonances de Schumann - Spectre du champ magnétique au Welschbruch. Les 6 premières résonances sont clairement visibles. 29
2.10 *Schumann resonances on magnetic field power spectrum at Welschbruch geophysical station. The first six harmonics are clearly visible.* 29
2.11 Haut : Onde *Q-Burst* en domaine temporel pour un échantillonnage à 64 Hz. Bas : Spectre en puissance des séries temporelles présentées au dessus. 29
2.11 *Top : Q-Burst in time domain for a sampling frequency of 65,536 Hz. Bottom : Power spectrum of above time series.* 29
2.12 Comparaison des angles de polarisation horizontale enregistrés pour une onde ELF enregistrée au Népal aux stations de Trisuli et Grang le 14 Mars 2016. . . . 61
2.12 *Comparison of horizontal polarisation angle recorded at Trisuli and Grang MT stations on 14 March 2016.* 61
2.13 Pulsations de type Pc3 observées à l'observatoire d'Uppsala (UPS, Suède) le 12 novembre 2013. 63
2.13 *Pc3 pulsation recorded at Uppsala (UPS, Sweden) observatory on 12 November 2013.* 63
2.14 Pulsations de type Pi1 observées à l'observation de Newport (NEW, Etats-Unis), le 05 janvier 2013. 63
2.14 *Pi1 pulsation recorded at Newport (NEW, United States of America) observatory on 05 January 2013.* 63
2.15 Variation de l'amplitude de trois pulsations Pc3 (A, B et C) avec la latitude géomagnétique, extrait de Takao Saito (1969) d'après les études de J. Jacobs et Sinno (1960), Kato (1962) et Kato et T Saito (1962). 64
2.15 *Variation of the amplitude of 3 Pc3 pulsations (A, B, and C) with geomagnetic latitude. From Takao* Saito *(1969), results from studies by J.* Jacobs *et* Sinno *(1960),* Kato *(1962) et* Kato *et T* Saito *(1962).* 64
2.16 Variation de la période de pulsations Pc3 et Pc4 avec le cycle solaire. Extrait de Takao Saito (1969), d'après les études de T Saito et Matsushita (1967) et Kawamura (1969b). 65

2.16 *Variation of Pc3 and Pc4 pulsations period with the solar cycle. From Takao* SAITO *(1969), results from studies by T* SAITO *et* MATSUSHITA *(1967) et* KAWAMURA *(1969b).* . . . 65
2.17 Variation de la probabilité d'apparition d'une pulsation Pc3 en fonction de l'heure du jour et du mois en 1961 et 1962. Extrait de Takao SAITO (1964) et Takao SAITO (1969). . . . 66
2.17 *Diurnal variation of Pc3 pulsation occurence from 1961 to 1962. From Takao* SAITO *(1964) et Takao* SAITO *(1969).* . . . 66
2.18 Haut : Variation de la fréquence d'apparition des pulsations Pc1 entre 1952 et 1966. Bas : Nombre de taches solaires pour la même période. Extrait de Takao SAITO (1969), d'après les études de BENIOFF (1960) et TROITSKAYA (1967). . . . 67
2.18 *Top : Variation of Pc1 pulsation occurence rate from 1952 to 1966. Bottom : Solar sunspot number for the same period. From Takao* SAITO *(1969), results from studies by* BENIOFF *(1960) et* TROITSKAYA *(1967).* . . . 67
2.19 Variation de la polarisation des pulsations Pc4 et Pc5 au Canada. Extrait de SAMSON, J. JACOBS et ROSTOKER (1971). . . . 68
2.19 *Variation of the polarisation of geomagnetic pulsations Pc4 and Pc4 in Canada. From* SAMSON, *J.* JACOBS *et* ROSTOKER *(1971).* . . . 68
2.20 Haut : Séries temporelles magnétiques filtrées enregistrées à Syafru-Bensi, Népal, le 09/03/2016. Bas : Séries temporelles magnétiques filtrées simultanées à l'observatoire d'Hyderabad (HYB, Inde). Rouge : H_x, Bleu : H_y. . . . 84
2.20 *Top : Filtered magnetic time series recorded at Syafru-Bensi, Nepal, on March 09, 2016. Bottom : Simultaneous filtered magnetic time series of Hyderabad observatory (HYB, India). Red : H_x. Blue : H_y.* . . . 84

3.1 Exemple de problème de régression linéaire avec des points dont les résidus ne vérifient pas l'hypothèse de distribution gaussienne. Les points verts vérifient : $e = 0.5b + \mathcal{N}(0, 0.1)$. Les points rouges suivent une distribution de Cauchy de paramètres de position 0 et d'échelle 0.1. La courbe en pointillés noirs est le modèle $e = 0.5b$, celle en pointillés bleus le modèle obtenu par les moindres carrés. 89
3.1 *Example of linear regression where residuals do not follow gaussian distribution. Green dots follow $e = 0.5b + \mathcal{N}(0, 0.1)$. Red dots follow Cauchy distribution with location and scale parameters at 0 and 0.1 respectively. Dotted black line is $e = 0.5b$ model, dashed blue line is model obtained through least-squares inversion.* . . . 89
3.2 Fonction de coût pour les modèles de moindres carrés (trait plein), norme L_1 (trait pointillé), et de Huber (trait discontinu). . . . 91
3.2 *Cost function for least-squares model (full line), L_1 norm (dotted line) and Huber model (dashed line).* . . . 91
3.3 Poids pour les modèles de moindres carrés (trait pointillé bleu), norme L_1 (trait pointillé), et de Huber (trait pointillé rouge). . . . 92
3.3 *Weights for least-squares model (full line), L_1 norm (dotted line) and Huber model (dashed line).* . . . 92
3.4 Exemple de problème de régression linéaire avec des points dont les résidus ne vérifient pas l'hypothèse de distribution gaussienne. Les points verts vérifient : $e = 0.5b + \mathcal{N}(0, 0.1)$. Les points rouges suivent une distribution de Cauchy de paramètres de position 0 et d'échelle 0.1. La courbe en pointillés noirs est le modèle $e = 0.5b$, celle en pointillés bleus le modèle obtenu par les moindres carrés, et celle en pointillés rouge est celle obtenue par la procédure de moindres carrés pondérés par les poids de Huber. . . . 93

3.4 *Example of linear regression where residuals do not follow gaussian distribution. Green dots follow $e = 0.5b + \mathcal{N}(0, 0.1)$. Red dots follow Cauchy distribution with location and scale parameters at 0 and 0.1 respectively. Dotted black line is $e = 0.5b$ model, dashed blue line is the model obtained through least-squares inversion. Red dashed line is the model obtained through application of iterative least squares regression with Huber weights.* . . . 93
3.5 Distribution des événements inversés par la méthode utilisant les combinaisons pour une fréquence de 107 Hz. Les distributions ont été ramenés à une médiane nulle et divisées par l'écart type de la distribution Bootstrap. Rouge : Combinaisons de 2 événements. Vert : Combinaisons de 3 événements. Bleu : Combinaisons de 4 événements. Magenta : Combinaisons de 5 événements. . . . 109
3.5 *Impedance tensor distribution at 107 Hz frequency by inverting combinations of events wavelet coefficients. Medians of each distribution have been set to 0 and distributions have been divided by the bootstrap distribution standard deviation. Red : Combinations of two events. Green : Combinations of three events. Blue : Combinations of four events. Pink : Combinations of five events.* . . . 109
3.6 Distribution des événements inversés par la méthode utilisant les combinaisons pour une fréquence de 107 Hz. Les distributions ont été ramenés à une médiane nulle et divisées par l'écart type de la distribution Bootstrap. Rouge : Combinaisons de 2 événements. Vert : Combinaisons de 3 événements. Bleu : Combinaisons de 4 événements. Magenta : Combinaisons de 5 événements. . . . 113
3.6 *Impedance tensor distribution at 107 Hz frequency by inverting combinations of events wavelet coefficients. Medians of each distribution have been set to 0 and distributions have been divided by the bootstrap distribution standard deviation. Red : Combinations of two events. Green : Combinations of three events. Blue : Combinations of four events. Pink : Combinations of five events. Black : Bootstrap inversion.* . . . 113
3.7 Distributions des inversions bootstrap pour une fréquence de 107 Hz. . . . 114
3.7 *Bootstrap distributions for 107 Hz frequency* . . . 114
3.8 Carte indiquant la position de la station MT et de l'observatoire magnétique de CLF. . . . 145
3.8 *Map of MT station and CLF magnetic observatory.* . . . 145
3.9 Séries temporelles à Schwabwiller et à l'observatoire magnétique de Chambon-La-Forêt pour le 03/12/2014. . . . 146
3.9 *MT (top) and observatory (bottom) time series on the 02/12/2014. Time series start at 18 :00 :00 PM.* . . . 146
3.10 Cohérence sur 24h entre les séries MT et d'observatoire. . . . 147
3.10 *24h coherency between MT and observatory time series.* . . . 147
3.11 Cohérence sur 1h entre les séries MT et d'observatoire. . . . 148
3.11 *1h coherency between MT and observatory time series.* . . . 148
3.12 Illustration de la connectivité des pixels et de l'assignation d'un label à chaque ensemble. . . . 149
3.12 *Illustration of the label assignement to groups of significant coefficients.* . . . 149
3.13 Indices des événements après application de la méthodologie de segmentation de He et al. (2009). . . . 149
3.13 *Events labels after application of segmentation methodology of* He *et al. (2009).* . . . 149
3.14 Sondages MT obtenus par application du code BIRRP (gauche) et du bootstrap (droite). . . . 150

3.14 *MT soundings obtained through BIRRP processing (left) and bootstrap processing (right).* . . . 150
3.15 Gauche : Résidu maximal obtenu en prenant toute la série temporelle ou seulement la sélection des coefficients. Carrés : Résidus sur E_x, Cercles : Résidus sur E_y. Droite : QQ-plots des résidus par rapport à la distribution de Rayleigh. Haut : Résidus sur E_x, Bas : Résidus sur E_y. . . . 150
3.15 *Left : maximum residual using all wavelet coefficients or only selected coefficients. Squares : E_x residus, Circles : E_y residus. Left : Residuals QQ-plots against Rayleigh distribution. Top : E_x residuals. Bottom : E_y residuals.* . . . 150

4.1 Carte de la France avec la localisation de la campagne Audio-MT, MT. La topographie est extraite du modèle ETOPO1 (Amante et Eakins, 2009). . . . 154
4.1 *Map of France with location of Audio-MT, MT campaign. Topography is from the ETOPO1 model (Amante et Eakins, 2009).* . . . 154
4.2 Coupe géologique Est-Ouest du bassin Parisien. L'échelle verticale est amplifiée. D'après Lorenz (1980). . . . 155
4.2 *East-West geological cross section of the Paris basin. Note the vertical exaggeration From Lorenz (1980)* . . . 155
4.3 Intensité de l'anomalie magnétique (non réduite aux pôles) obtenue en France. Le cercle rouge indique la position de l'observatoire magnétique. Modifié d'après Thébault, Mandea et Schott (2006). . . . 156
4.3 *Magnetic anomaly (without reduction to the pole) intensity over France. The red circle indicates the location of CLF magnetic observatory. Modified from Thébault, Mandea et Schott (2006)* . . . 156
4.4 Position des stations Audio-MT et MT installées en mars 2016 relativement à l'observatoire située en (0, 0). Projection UTM20N. . . . 158
4.4 *Position of Audio-MT and MT stations relative to the observatory located at position (0, 0). UTM 20N projection.* . . . 158
4.5 Comparaison des séries temporelles entre la station Observatoire et D4. L'échantillonnage des séries temporelles est à 32768 Hz. . . . 159
4.5 *Comparison between D4 and Observatory station time series. Sampling frequency is 32,768 Hz.* . . . 159
4.6 Comparaison des séries temporelles entre la station Observatoire et B6. L'échantillonnage des séries temporelles est à 32,768 Hz. . . . 160
4.6 *Comparison between B6 and Observatory station time series. Sampling frequency is 32,768 Hz.* . . . 160
4.7 Comparaison des séries temporelles entre la station Observatoire et C1. L'échantillonnage des séries temporelles est à 1024 Hz. . . . 161
4.7 *Comparison between B6 and Observatory station time series. Sampling frequency is 1,024 Hz.* . . . 161
4.8 Ellipses du tenseur de phase aux fréquences : a) 6 kHz b) 279 Hz c) 34 Hz d) 7 Hz e) 0.5 Hz f) 0.039062 Hz g) 0.007812 Hz h) 0.000521 Hz i) 6.5e-5 Hz. La couleur de l'ellipse indique la valeur du paramètre λ. . . . 163
4.8 *Phase tensor ellipses at frequencies : a) 6 kHz b) 279 Hz c) 34 Hz d) 7 Hz e) 0.5 Hz f) 0.039062 Hz g) 0.007812 Hz h) 0.000521 Hz i) 6.5e-5 Hz. Ellipse color indicates the value of λ parameter.* . . . 163
4.9 *Twist* et *Shear* de la décomposition de Groom-Bailey aux fréquences : a) 6 kHz b) 279 Hz c) 34 Hz d) 7 Hz e) 0.5 Hz f) 0.039062 Hz g) 0.007812 Hz h) 0.000521 Hz i) 6.5e-5 Hz. Rouge : *Twist*. Blue : *Shear* . . . 164

4.9 *Twist and shear of Groom-Bailey decompositions at frequencies : a) 6 kHz b) 279 Hz c) 34 Hz d) 7 Hz e) 0.5 Hz f) 0.039062 Hz g) 0.007812 Hz h) 0.000521 Hz i) 6.5e-5 Hz. Red : Twist, Blue : Shear.* . . . 164
4.10 Partie réelle des flèches d'induction en convention de Parkinson aux fréquences : a) 6 kHz b) 279 Hz c) 34 Hz d) 7 Hz e) 0.5 Hz f) 0.039062 Hz g) 0.007812 Hz h) 0.000521 Hz i) 6.5e-5 Hz. . . . 165
4.10 *Real part of inductions arrows in Parkinson convention at frequencies : a) 6 kHz b) 279 Hz c) 34 Hz d) 7 Hz e) 0.5 Hz f) 0.039062 Hz g) 0.007812 Hz h) 0.000521 Hz i) 6.5e-5 Hz.* . . . 165
4.11 Carte géologique reproduite à partir de la carte au 1/50 000 fournie par le BRGM. Les points rouges indiquent la position des forages proche-surface. . . . 166
4.11 *Géological map of the area reproduced from the 1/50 000 map available on BRGM website. Red dots indicates near-surface wells.* . . . 166
4.12 Strike obtenus par les différentes décompositions. Rouge : Sondage MT Observatoire. Vert : Sondage MT Thivet. Bleu : Sondage MT Ingrannes. Marron : Sondage dipôles d'observatoire. . . . 167
4.12 *Strike obtained through the different decompositions. Red : Observatory MT sounding. Green : Thivet MT sounding. Blue : Ingrannes MT sounding. Brown : Observatory electric dipoles sounding.* . . . 167
4.13 Haut : Position des sites Audio-MT (triangles blancs), des forages (points rouges) et de la coupe géologique (ligne rouge). Milieu : Exemple de sondages obtenus sur le site du BRGM. Gauche : 03642X0066. Droite : 03642X0085. Bas : Coupe géologique obtenue à partir de données de forage. D'après Gigot, 1984. . . . 168
4.13 *Top : Location of Audio-MT sites (white triangles), wells (red dots) and geological cross-section (red line). Middle : Example of wells, Left : 03642X0066, Right : 03642X0085. Bottom : Geological cross-section from drilling information. From* Gigot *(1984).* . . . 168
4.14 Gauche : Modèle 1D de résistivité utilisé dans les modélisations. Droite : Comparaison entre le sondage correspondant au modèle 1D utilisé à gauche dans les modélisations et le sondage MT obtenu à l'observatoire. . . . 169
4.14 *Left : 1D resistivity model used in the modeling. Right : Comparison between the modeling of the mean 1D model and the Observatory MT sounding.* . . . 169
4.15 Comparaison entre les modélisations de sondage MT pour des modèles avec ou sans topographie aux positions de la campagne. . . . 170
4.15 *Comparison between the modeling of MT soundings for 3D models with and without topography.* . . . 170
4.16 Comparaison entre les modélisations de sondage MT pour des modèles avec ou sans océans aux positions de la campagne. Haut : Fonctions de transfert verticales réelles et imaginaires. Bas : Direction des flèches d'induction réelles. Rouge : Données. Blanc : Modèle avec océan. Noir : Modèle sans océan. . . . 171
4.16 *Comparison between the modeling of MT soundings for 3D models with and without oceans. Top : Vertical transfer function. Bottom : Direction of induction arrows. Red : Data. White : Ocean model. Black : Land model.* . . . 171
4.17 Sondages MT obtenus par la décomposition de Groom-Bailey avec une direction de strike 45°. . . . 172
4.17 *Groom-Bailey MT soundings with a strike direction of 45°.* . . . 172
4.18 Modèles d'origine de l'anomalie magnétique du bassin parisien proposés par Autran et al. (1986). . . . 173

4.18 *Models for the origin of the Paris basin magnetic anomaly. From* AUTRAN *et al. (1986)* . . . 173
4.19 Modèles de conductivité 2D utilisés pour le calcul des réponses MT. Haut : Modèle avec une anomalie résistante. Bas : Modèle avec une anomalie conductrice. Les triangles blancs indiquent la position des stations MT. . . . 174
4.19 *2D conductivity models for the computation of MT response functions. White triangles indicates the location of MT stations.* . . . 174
4.20 Sondages modélisés aux stations situées à -50 km (gauche) et à 5.8 km (droite). . 175
4.20 *Response functions modelling for MT stations located at -50 km (left) and 5.8 km (right).* . . . 175

5.1 Mouvement de l'Inde par rapport à la plaque eurasienne. D'après ROYER et PATRIAT (2002). . . . 177
5.1 *Displacement of Indian plate relative to Eurasia plate. From* ROYER *et* PATRIAT *(2002).* . . . 177
5.2 Haut : Position de la coupe géologique traversant la chaîne Himalayenne. Bas : Schéma des principales formations géologiques de la chaîne Himalayenne. DKI : Dadeldhura klippe. JKI : Jajarkot klippe. KKI : Katmandhu klippe. D'après BOLLINGER et al. (2004). Voir le texte pour la description des formations géologiques. 178
5.2 *Top : Position of geological cross-section across the Himalayan chain. Bottom : Sketch of main geological formations of Himalayan chain. DKI : Dadeldhura klippe. JKI : Jajarkot klippe. KKI : Katmandhu klippe. From* BOLLINGER *et al. (2004). See text for a description of geological structures.* . . . 178
5.3 Carte des répliques du séisme de Gorka au niveau de l'Himalaya. KKI : Katmandhu klippe. D'après ADHIKARI et al. (2015). . . . 179
5.3 *Map of Gorka earthquake aftershocks. KKI : Katmandhu klippe. From* ADHIKARI *et al. (2015).* . . . 179
5.4 Modèle de conductivité électrique dans la vallée de Katmandou. D'après Carole LEMONNIER et al. (1999). . . . 180
5.4 *Electrical conductivity model in Katmandu valley. From Carole* LEMONNIER *et al. (1999).* . . . 180
5.5 Localisation de la zone étudiée. Modèle topographique extrait du modèle ETOPO1 (AMANTE et EAKINS, 2009). . . . 181
5.5 *Location of studied area. Topographic data is taken from ETOPO1 model (*AMANTE *et* EAKINS*, 2009).* . . . 181
5.6 Plan des stations magnétotelluriques installées. Topographie combinée des modèles ETOPO1 (AMANTE et EAKINS, 2009) et SRTM (FARR et al., 2007). . . . 182
5.6 *Map of magnetotelluric stations. Topography extracted from ETOPO1 (*AMANTE *et* EAKINS*, 2009) and SRTM (*FARR *et al., 2007) models.* . . . 182
5.7 Carte indiquant la position de la campagne MT et de l'observatoire magnétique utilisée comme station de référence. . . . 183
5.7 *Map of local MT stations in Nepal and HYB magnetic observatory India used as reference station.* . . . 183
5.8 Haut : Comparaison des composantes magnétiques Nord-Sud entre locale et *remote*. Milieu : Comparaison des composantes magnétiques Est-Ouest entre locale et remote. Bas : Cohérence entre les composantes locales et *remotes* pour les mêmes directions. . . . 184

5.8 *Top : Comparison between North-South magnetic components at local and remote station. Top : Comparison between East-West magnetic components at local and remote station. Bottom : Coherency between local and remote stations.* 184
5.9 Comparaison entre les sondages obtenus en 1996 (haut) et 2016 (bas) pour la station de Grang. 186
5.9 *Comparison between 1996 MT soundings (top) and 2016 MT soundings (bottom) for Grang station.* 186
5.10 Comparaison entre les sondages obtenus en 1996 (haut) et 2016 (bas) pour la station de Kalikastan. 187
5.10 *Comparison between 1996 MT soundings (top) and 2016 MT soundings (bottom) for Kalikastan station.* 187
5.11 Comparaison entre les sondages obtenus en 1996 (haut) et 2016 (bas) pour la station de Pipaltar. 188
5.11 *Comparison between 1996 MT soundings (top) and 2016 MT soundings (bottom) for Pipaltar station.* 188
5.12 Comparaison des vecteurs d'induction obtenus aux campagnes de 1996 et 2016. . 190
5.12 *Comparison of induction arrows between 1996 and 2016 MT campaigns.* 190
5.13 Indicateurs de dimensionnalité du tenseur de phase λ et β pour les 6 stations MT. 191
5.13 *Dimensionality indicators of phase tensor λ and β for the 6 MT stations.* 191
5.14 Strike du tenseur de phase $\alpha - \beta$ pour les 6 stations MT. 192
5.14 *Phase tensor strike $\alpha - \beta$ for the 6 MT stations.* 192
5.15 Comparaison des paramètres de distorsion pour une décomposition avec un strike de N20E et une décomposition multi-sites pour les 6 stations MT. 193
5.15 *Comparison of distortion parameters for single-site single frequency decomposition and multi-sites decomposition of the 6 MT stations.* 193
5.16 Test du χ^2 appliqué aux différentes décompositions. Multi-sites : Décomposition simultanée des sites du profil. Strike 10W : Décomposition en fixant le strike à N10W. SS,SF : Décomposition par site et par fréquence. 194
5.16 *χ^2 test applied to the different decompositions. Multi-sites : Simultaneous decomposition of all MT stations. Strike 10W : Decomposition by setting the strike at N10W. SS,SF : Single site, single frequency decomposition.* 194
5.17 Haut : Profils 2D utilisés. Bas : Dénivelé le long des profils et position des stations après projection orthogonale. 195
5.17 *Top : 2D profil used in the modeling. Bottom : Topography along profil with MT stations positions after orthogonal projection.* 195
5.18 Haut : Maillage du modèle en éléments finis pour le calcul de l'effet topographique. Bas : Maillage final pour le calcul du modèle direct. 196
5.18 *Top : Finite element mesh for topography effect computation. Bottom : Final mesh used in forward modeling.* 196
5.19 Modèle obtenu après inversion sans topographie des données modélisées avec topographie. 197
5.19 *Inversion without topography of synthetic soundings from model with topography.* . 197
5.20 Comparaison entre données et résultat de l'inversion sans topographie. 198
5.20 *Comparison between data and inversion result without topography included.* 198
5.21 Modèle de topographie 3D pour le site de Grang 199
5.21 *3D topography model for Grang MT site..* 199
5.22 Sondages 3D pour un modèle homogène de terre à 1000 $\Omega.m$. 201
5.22 *3D MT soundings for an homogenous earth with a resistivity of 1000 $\Omega.m$.* 201

5.23 Modèle de Carole LEMONNIER et al. (1999) non interpolé. 202
5.23 *Final resistivity model from Carole* LEMONNIER *et al. (1999) without interpolation.* 202
5.24 Données réelles et modélisées pour les stations de Simra, Pantale, Daman et Kakani d'après l'inversion de 1999. 203
5.24 *Real data and modeling of 1999 conductivity model for Simra, Pantale, Daman and Kakani MT stations.* 203
5.25 Données réelles et modélisées pour les stations de Pipaltar, Dhunche, Ramche, Syafru d'après l'inversion de 1999. 204
5.25 *Real data and modeling of 1999 conductivity model for Pipaltar, Dhunche, Ramche and Syafru MT stations.* 204
5.26 Modèle de résistivité issue de l'inversion à partir des données de 2016. 205
5.26 *Resistivity modeling from 2016 dataset.* 205
5.27 Comparaison entre données et résultat de l'inversion à partir des données de 2016. 206
5.27 *Comparison between data and inversion results from 2016 dataset.* 206
5.28 Modèle de résistivité issue de l'inversion du profil en retirant Syafru (haut), Dhunche (milieu) et Grang (bas). Les triangles blancs indiquent la position des stations utilisées dans l'inversion. 208
5.28 *Resistivity model from inversion without Syafru MT station (top), without Dhunche MT station (middle), and without Grang MT station (bottom). White triangles indicate MT stations used in inversion.* 208
5.29 Modèle de résistivité issue de l'inversion du profil en retirant Kalikastan (haut) et Trisuli (bas). Les triangles blancs indiquent la position des stations utilisées dans l'inversion. 209
5.29 *Resistivity model from inversion without Kalikastan MT station (top), without Trisuli MT station (bottom). White triangles indicate MT stations used in inversion.* 209
5.30 Sondages obtenus par le retrait des stations MT pour les stations de Syafru (top), Dhunche (middle) and Grang (bottom). 210
5.30 *Sondages obtained from forward modeling with removal of MT stations for Syafru (top), Dhunche (middle), Grang (bottom).* 210
5.31 Sondages obtenus par le retrait des stations MT pour les stations de Kalikastan (top), Trisuli (middle) and Pipaltar (bottom). 211
5.31 *Sondages obtained from forward modeling with removal of MT stations for Kalikastan (top), Trisuli (middle), Pipaltar (bottom).* 211
5.32 Modèles de résistivité obtenus en changeant le paramètre de régularisation réglant la résolution horizontale et verticale. Haut : Rapport horizontal/vertical 1/3. Bas : Rapport horizontal/vertical 1. Les triangles blancs indiquent la position des stations utilisées dans l'inversion. 212
5.32 *Resistivity model obtained by changing the regularization parameter between horizontal and vertical resolution. Top : Ratio horizontal/vertical 1/3. Bottom : Ratio horizontal/vertical 1. White triangles indicate MT stations used in inversion.* . . 212
5.33 Modèles de conductivité obtenus par l'inversion des données de 2016 (haut), de 1996 (milieu) et des campagnes de 1996, 1999, 2000 (bas). Les triangles blancs indiquent la position des stations utilisées dans l'inversion. 214
5.33 *Conductivity distributions from the inversion of 2016 data (top), 1996 data (middle) and 1996, 1999, 2000 data (bottom). White triangles indicate MT stations used in inversion.* 214
5.34 RMS mesurées entre les sondages MT 2016 et les sondages modélisés par les différentes distributions de conductivité. 215

5.34 *Difference between 2016 MT soundings and the modeling of conductivity distributions.* . . . 215
5.35 Différents modèles d'origine des flux de CO_2 et de radon à la source hydrothermale de Syabru-Bensi. D'après GIRAULT et Frédéric PERRIER (2014). . . . 216
5.35 *Sketch of possible origine of CO_2 and radon fluxes at the Syabru-Bensi hydrothermal system. From* GIRAULT *et Frédéric* PERRIER *(2014).* . . . 216
5.36 Coupe géologique de la vallée de Trisuli au Népal. D'après HUBBARD et al. (2016). 217
5.36 *Geological cross-section of Trisuli valley in Nepal. From* HUBBARD *et al. (2016).* 217
5.37 Comparaison entre les sondages obtenus en 1996 (haut) et 2016 (bas) pour la station de Syafru. . . . 218
5.37 *Comparison between 1996 MT soundings (top) and 2016 MT soundings (bottom) for Syafru station.* . . . 218
5.38 Comparaison entre les sondages obtenus en 1996 (haut) et 2016 (bas) pour la station de Dhunche. . . . 219
5.38 *Comparison between 1996 MT soundings (top) and 2016 MT soundings (bottom) for Dhunche station.* . . . 219
5.39 Comparaison entre les sondages obtenus en 1996 (haut) et 2016 (bas) pour la station de Trisuli. . . . 220
5.39 *Comparison between 1996 MT soundings (top) and 2016 MT soundings (bottom) for Trisuli station.* . . . 220
5.40 Décomposition de Groom-Bailey par fréquence et par site. . . . 221
5.40 *Single site, single frequency Groom-Bailey decomposition..* . . . 221
5.41 Décompositions de Groom-Bailey pour un strike fixé à N20E. . . . 222
5.41 *Groom-Bailey decomposition for a N20E strike.* . . . 222
5.42 Décompositions de Groom-Bailey multi-sites. . . . 223
5.42 *Multi-sites Groom-Bailey decomposition.* . . . 223

A.1 Carte des observatoires magnétiques dans le monde en 2013. La déclinaison magnétique est superposée à la carte. D'après LOVE et CHULLIAT, 2013. . . . 239
A.1 *Map of worldwide magnetic observatories in 2013. Magnetic declination is overlaid on the map. From* LOVE *et* CHULLIAT *(2013)* . . . 239
A.2 Fonctions de calibration des sondes MFS06 et MFS07 pour leurs deux modes de fonctionnement. . . . 242
A.2 *MFS06e and MFS07e calibration function for both acquisition modes.* . . . 242

B.1 Ondelette mère de Morlet avec le paramètre $\omega_0 = 5$. Gauche : Représentation temporelle. Trait plein : Partie réelle. Pointillé : Partie imaginaire. Droite : Module de la transformée de Fourier. . . . 247
B.1 *Morlet mother wavelet with parameter $\omega_0 = 5$. Left : Temporal representation. Full line : Real line. Dashed line : Imaginary line. Right : Amplitude of Fourier transform.* . . . 247
B.2 Comparaison entre deux ondelettes mère de Morlet. Gauche : Partie réelle en domaine temporel. Droite : Domaine fréquentiel. Trait plein noir : Ondelette mère avec $\omega_0 = 5$. Trait pointillé rouge : Ondelette mère avec $\omega_0 = 10$. . . . 247
B.2 *Comparison between two Morlet mother wavelet. Left : Real part of temporal representation. Right : Amplitude of Fourier transform. Full line : $\omega_0 = 5$. Red dashed line : $\omega_0 = 10$.* . . . 247

B.3 Ondelette mère de Paul d'ordre $m = 2$. Gauche : Représentation temporelle. Trait plein : Partie réelle. Pointillé : Partie imaginaire. Droite : Module de la transformée de Fourier. 248
B.3 *Paul mother wavelet of order m. Left : Temporal representation. Full line : Real line. Dashed line : Imaginary line. Right : Amplitude of Fourier transform.* 248
B.4 Comparaison entre l'ondelette mère et l'ondelette fille de Morlet par dilatation de facteur $a = 1/2$. Gauche : Partie réelle en domaine temporel. Droite : Domaine fréquentiel. Trait plein noir : Ondelette mère. Trait pointillé rouge : Ondelette fille. 249
B.4 *Comparison between mother and dilated Morlet wavelet. Left : Real part of temporal representation. Right : Amplitude of Fourier transform. Full line : Mother wavelet. Red dashed line : Dilated wavelet.* 249

C.1 Illustration des attributs de polarisation dans le cas général d'une polarisation elliptique. 253
C.1 *Illustration of polarisation attributes in the case of elliptic polarisation.* 253
C.2 Gauche : Signal monochromatique décrit équation C.4. Trait plein : f_x. Trait pointillé : f_y. Droite : Hodogramme des signaux. 254
C.2 *Left : Monochromatic signal from equation (C.4). Full line : f_x. Dashed line : f_y. Right : Signal hodogram.* 254
C.3 Gauche : Signal monochromatique décrit équation C.5. Trait plein : f_x. Trait pointillé : f_y. Droite : Hodogramme des signaux. 255
C.3 *Left : Monochromatic signal from equation (C.5). Full line : f_x. Dashed line : f_y. Right : Signal hodogram.* 255
C.4 Gauche : Signal monochromatique décrit équation C.6. Trait plein : f_x. Trait pointillé : f_y. Droite : Hodogramme des signaux. 255
C.4 *Left : Monochromatic signal from equation (C.6). Full line : f_x. Dashed line : f_y. Right : Signal hodogram.* 255
C.5 Haut, Gauche : Signal monochromatique décrit équation (C.4). Trait plein : f_x. Trait pointillé : f_y. Haut, Droite : Spectre d'amplitude du signal décrit équation (C.4). Bas : Attributs de polarisation en fonction de la fréquence. 256
C.5 *Top, Left : Monochromatic signal from equation (C.4). Top, Right : Amplitude spectra of the signal from equation (C.4). Bottom : Fourier polarisation attributes versus frequency.* 256
C.6 a) Signal monochromatique décrit équation (C.5). Trait plein : f_x. Trait pointillé : f_y. b) Spectre d'amplitude du signal décrit équation (C.5). c) d) e) Attributs de polarisation en fonction de la fréquence. 257
C.6 *a) Monochromatic signal from equation (C.5). b) Amplitude spectra of the signal from equation (C.5). c) d) e) Fourier polarisation attributes versus frequency.* . . 257
C.7 a) Signal monochromatique décrit équation (C.6). Trait plein : f_x. Trait pointillé : f_y. b) Spectre d'amplitude du signal décrit équation (C.6). c) d) e) Attributs de polarisation en fonction de la fréquence. 258
C.7 *a) Monochromatic signal from equation (C.6). b) Amplitude spectra of the signal from equation (C.6). c) d) e) Fourier polarisation attributes versus frequency.* . . 258
C.8 a) Signal monochromatique décrit équation (C.16). Trait plein : f_x. Trait pointillé : f_y. b) Spectre d'amplitude du signal décrit équation (C.16). c) d) e) Attributs de polarisation en fonction de la fréquence. 259
C.8 *a) Monochromatic signal from equation (C.16). b) Amplitude spectra of the signal from equation (C.16). c) d) e) Fourier polarisation attributes versus frequency.* . . 259

C.9 De haut en bas : Signal monochromatique décrit équation (C.16) ; Amplitude des coefficients d'ondelettes ; Coefficients d'ellipticité instantanée ; Coefficients d'angle de polarisation instantanée ; Coefficients de différence de phase instantanée. . . . 261
C.9 *From top to bottom : Monochromatic signal from equation (C.16) ; Wavelet coefficients amplitude ; Instantaneous ellipticity coefficients ; Instantaneous polarisation angle coefficients ; Instantaneous phase difference coefficients.* 261
C.10 Détermination des attributs de polarisation par la méthode d'analyse en ondelettes. Noir : Moyenne arithmétique des coefficients de polarisation instantanés. Rouge : Modèle imposé. 262
C.10 *Determination of polarisation attributes using continuous wavelet transform. Black : Arithmetic mean of instantanous polarisation coefficients. Red : Model.* 262

D.1 Représentation graphique du tenseur de phase. D'après Caldwell, Bibby et Brown (2004). 265
D.1 *Graphical representation of the phase tensor. From* Caldwell, Bibby *et* Brown *(2004).* 265
D.2 Illustration de l'effet des matrices de distortion de Groom-Bailey. D'après A. D. Chave et A. Jones (2012). 266
D.2 *Illustration of the effect of Groom-Bailey distortion matrix. From A. D.* Chave *et A.* Jones *(2012).* 266
D.3 Grille en différence finies pour le calcul des champs électromagnétiques. D'après Kelbert et al. (2014) 268
D.3 *Finite-difference grid for the computation of EM fields. From* Kelbert *et al. (2014).* 268

Liste des tableaux

2.1 Classification des pulsations géomagnétiques. D'après J. A. Jacobs et al. (1964). 62
2.1 *Geomagnetic pulsations classification. From J. A. Jacobs et al. (1964).* 62

4.1 Coordonnées de l'observatoire magnétique de Chambon-La-Forêt. Les coordonnées géomagnétiques corrigées sont calculées à partir du modèle IGRF 2010 pour l'année 2011. 153
4.1 *Geomagnetic coordinates of the magnetic observatory of Chambon-La-Forêt. Corrected geomagnetic coordinates are computed from the IGRF-10 model of year 2011.* 153

E.1 Détail des configurations AMT 275
E.2 Détail des configurations MT 285
E.3 Détail des configurations AMT 286
E.4 Détail des configurations AMT 287

Introduction

En géophysique, les méthodes électromagnétiques (EM) permettent la caractérisation de la conductivité électrique du sous-sol, elle-même directement reliée aux caractéristiques physiques des roches comme leur température, leur porosité ou encore leur perméabilité. La conductivité variant sur plusieurs ordres de grandeurs, les méthodes EM apportent un éclairage intéressant pour la caractérisation du sous-sol et complémentaire aux méthodes basées sur la densité des roches telles que la géophysique sismique ou la gravimétrie. En EM, la méthode magnétotellurique (MT) se base sur l'induction de courants électriques dans le sous-sol par des sources naturelles, telles que les orages atmosphériques ou encore l'activité magnétique (issue de l'action du vent solaire sur le champ magnétique terrestre). Cette diversité de sources utilisée en MT permet d'investiguer des profondeurs allant de la dizaine de mètres à la centaine de kilomètres et, donc, d'étudier la caractérisation d'aquifères comme les systèmes de collision continentales ou de subduction. La simplicité de mise en oeuvre sur le terrain de la méthode géophysique MT est également un atout dans les environnements difficiles d'accès. Le travail effectué au cours

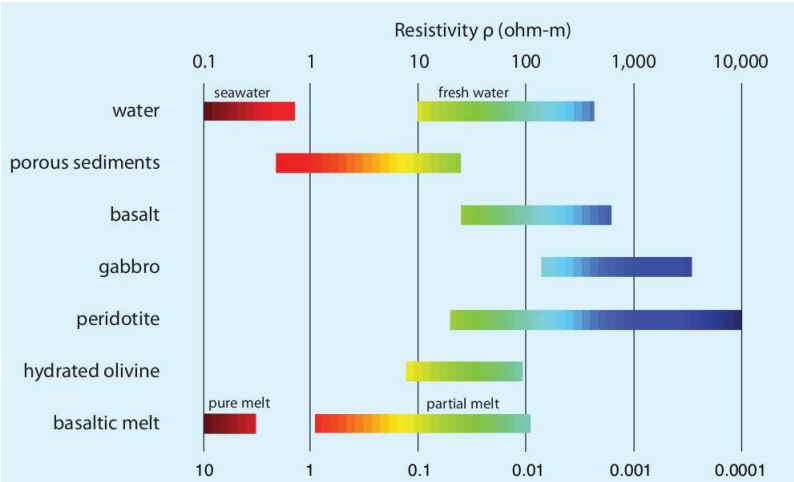


Figure 1: Table de résistivité de différents matériaux géologiques. D'après la présentation d'EM donnée par Constable, S à la 2015 *IAGA Summer School*.
FIGURE 1: *Resistivity values for geological materials. From the EM course given by Steven Constable, at the 2015 IAGA Summer School.*

de cette thèse a principalement porté sur l'acquisition de nouvelles données MT, le développement de nouvelles méthodologies de traitements des données obtenues et leur interprétation en adéquation avec chaque contexte géophysique.

Dans le présent manuscrit de thèse, nous avons séparé les travaux en 5 chapitres. Le premier chapitre rappelle les bases de la méthode magnétotellurique.

Les chapitres 2 et 3 concernent les développements méthodologiques effectués en traitement de données MT. Ce traitement consiste en la caractérisation des fonctions de transferts entre les composantes horizontales du champ électrique et du champ magnétique. Ces fonctions de transferts sont ensuite interprétées pour apporter des informations sur les structures géologiques sous-jacentes. Les procédures actuelles de traitements de données MT utilisent peu l'information sur les sources géomagnétiques. Dans le cas de données bruitées, ou bien dans le cas des intervalles de fréquences où le signal est naturellement faible - telles que l'Audio-MT *dead band* (1-5 kHz)-, cette information peut s'avérer critique afin de réduire l'incertitude sur les sondages MT.

Nous avons travaillé en ce sens pendant cette thèse en détectant et caractérisant les sources géomagnétiques par transformées en ondelettes continues avant le calcul des fonctions de réponses MT. Deux types de sources majeurs interviennent dans l'induction de courants électriques dans le sous-sol et nécessitent une approche différente. Les sources hautes fréquences (supérieures au Hz) proviennent principalement de l'activité orageuse globale par l'occurrence d'éclairs atmosphériques. Ceux-ci apparaissent comme des événements transitoires ayant un large contenu fréquentiel. A l'inverse, les sources basses fréquences résultent plutôt des oscillations de la cavité formée par la magnétosphère. Elles apparaissent bien localisées en fréquence, et contiennent des événements entretenus (de quelques minutes à plusieurs heures).

Nous avons montré que les différentes ondes électromagnétiques émises par l'occurrence d'un éclair atmosphérique ont des contenus fréquentiels et des formes différentes dans les spectrogrammes d'ondelettes. Trois types d'ondes ont été particulièrement étudiés, les *atmospherics*, *slow tails*, et les *whistlers*. Une méthodologie de détections automatiques et de caractérisations a été développée en se basant sur le large contenu fréquentiel et l'éventuelle dispersions de chacune de ces ondes électromagnétiques. A l'aide de cet outil, nous avons montré que les événements, ayant un contenu fréquentiel dans l'Audio-MT *dead band*, présentent une fréquence d'occurrence nettement inférieure à l'ensemble des événements. Ainsi, nous avons démontré tout le bénéfice de l'augmentation de la durée des mesures Audio-MT, afin d'obtenir un nombre suffisant d'événements géomagnétiques propres à permettre la caractérisation des réponses MT dans cette intervalle de fréquences. Notre méthodologie de détections a été utilisée pour le calcul des fonctions de réponse MT. Ces fonctions sont obtenues par régression robuste sur les coefficients d'ondelettes des événements géomagnétiques. Les incertitudes sont déterminées par bootstrap hiérarchique sur l'ensemble des événements utilisés. A l'aide de cette méthodologie, nous avons

montré une amélioration significative des réponses MT pour les fréquences supérieures au Hz, et permis la réduction drastique des incertitudes dans l'Audio-MT *dead band*.

Pour les sources dont les fréquences sont inférieures au Hertz, nous avons choisi d'étudier les pulsations géomagnétiques (0.001-1 Hz). Contrairement aux sources atmosphériques, les pulsations n'ont pas une signature caractéristique permettant de les détecter aisément dans le plan temps-fréquence. En revanche, les pulsations présentent une variabilité temporelle principalement diurne. Nous avons donc mis en place une stratégie basée sur la simulation de séries temporelles dont les caractéristiques sont déterminées par un spectre de puissance moyen calculé sur 24h. Les spectrogrammes d'ondelettes, croisés des séries temporelles analysées, sont ainsi comparés à ceux de la population simulée. Les coefficients significatifs correspondant à ceux des pulsations géomagnétiques sont ainsi plus aisément détectés. Si l'amplitude des signaux demeure ici le critère principal de détection, l'utilisation des données des observatoires magnétiques s'est, elle aussi, révélée critique. En effet, contrairement aux données MT, les données des observatoires magnétiques restent peu affectées par les bruits anthropiques. La méthodologie de détermination des fonctions de réponses MT, développées pour les sources atmosphériques, a ici été simplement adaptée et appliquée pour le cas des pulsations géomagnétiques.

Nous avons montré que, l'utilisation des événements géomagnétiques dans le calcul des fonctions de réponses MT, réduit significativement l'amplitude des résidus obtenus sur le champ électrique. De plus, nous obtenons une caractérisation plus précise des fonctions de réponses magnétotellurique lors des procédures de statistiques robustes. Enfin, l'utilisation des signaux à fort rapport signal sur bruit réduit significativement le biais observé sur l'impédance MT dans le cas d'un traitement sans station de référence.

Aux chapitres 4 et 5, nous présentons l'interprétation de deux ensembles de données magnétotelluriques acquis durant ces trois ans. L'interprétation consiste en l'étude des fonctions de réponses MT obtenues. Dans le chapitre 4, nous présentons les résultats d'une campagne d'acquisition de données menées en mars 2015 autour de l'observatoire magnétique de Chambon-La-Forêt en collaboration avec Prof. F. Perrier de l'Institut de Physique du Globe de Paris (IPGP). Cette mission avait deux objectifs principaux : (i) connaître et investiguer les sources de bruits et, (ii) étudier la distribution de résistivités dans le sous-sol autour de l'observatoire. Les sources de bruits autour de l'observatoire magnétique sont apparues comme beaucoup plus importantes que prévu. Les méthodes de traitements Audio-MT, développées pendant cette thèse, ont été cruciales pour réduire les incertitudes. L'interprétation des sondages MT indique une distribution quasi-verticale de la résistivité, n'occasionnant que très peu de distorsions du champ électrique, jusqu'à une profondeur d'environ 3 km. La modélisation 3D des effets électriques de la topographie, dans les environs de l'observatoire, montre une très faible influence de la topographie pour des fréquences inférieures à 500 Hz. Les modélisations 1D sont en accords avec d'anciens forages superficiels effectués dans la zone. Les mesures MT basses fréquences apportent des informations

plus profondes. Ainsi, à partir de 2.2 km de profondeur, la complexité du milieu augmente pour devenir un milieu 2D. La direction d'invariance du lieu à cette profondeur indique une structure pouvant correspondre à l'anomalie du bassin de Paris, de précédentes études ayant montré une source de profondeur équivalente. Le peu de données basses fréquences récoltées durant cette courte mission ne permettent en revanche pas de préciser l'extension ou la conductivité de cette anomalie par inversion de données.

Le chapitre 5 présente la seconde campagne d'acquisition, à laquelle j'ai activement contribué, qui s'est déroulée au Népal en mars 2016 dans la vallée de Katmandou, en collaboration avec : le Professeur Frédéric Perrier de l'IPGP, le National Seismological Center de Katmandou et l'Earth Observatory of Singapor. Cette mission avait pour objectif de répéter la campagne MT organisée en 1996 par l'IPGS. 9 sondages ont été acquis, 3 à de nouveaux emplacements et 6 de répétition sur l'ancien profil. L'analyse des sondages obtenus montre une complexité croissante des structures sous-jacentes de la vallée de Katmandou jusqu'au pied de la chaîne Himalayenne. La comparaison directe des flèches d'inductions, entre celles calculées suite à la campagne de 1996 et celle de 2016, montre peu de différences. Une zone conductrice profonde au nord du profil est mise en évidence par nos nouvelles mesures, et, confirme cette partie des résultats de la campagne précédente. La modélisation 3D de la topographie souligne son influence importante sur les composantes horizontales du champ électrique. En particulier, certaines composantes du tenseur d'impédance présentent une décroissance de la résistivité apparente avec la fréquence pouvant simuler la présence d'un corps conducteur. Les données de la campagne de 1996 avaient été inversées sans topographie. Le conducteur profond aurait donc pu, à l'époque, être créé artificiellement par l'algorithme d'inversion. Il n'en est rien puisque les données MT de 2016, inversées en prenant en compte la topographie, mettent également en évidence cette structure. Il est à noter que les données récentes sont également de meilleures qualités qu'en 1996 grâce à l'amélioration des procédures de traitements de données et du matériel d'acquisition utilisé. Par contre, les différents modèles de résistivités électriques obtenus ne sont pas en accord avec la campagne de 1996. En particulier, le conducteur profond, au niveau de la rampe du *Main Himalayan Thrust*, est totalement absent des nouveaux modèles qui positionnent les conducteurs au niveau de l'ancienne faille *Main Boundary Thrust* inactive. La principale caractéristique des nouveaux modèles est la présence d'un autre conducteur dans la partie sud du profil. Cette nouvelle structure conductrice tend à être connectée avec un conducteur superficiel présent dans la partie nord. Ce nouveau modèle est cohérent avec les modèles géologiques les plus récents et peut éventuellement être mis en relation avec des mesures de CO2 effectuées dans la région. Nous présentons dans cette thèse l'interprétation des données de cette campagne. Il est cependant à noter que les travaux n'en sont qu'aux prémices et demanderaient dans le futur de nouvelles campagnes de mesures sur une base plus régulière afin de définitivement valider les nouveaux résultats obtenus mais également appréhender l'évolution temporelle des caractéristiques physiques de cette zone active.

Le dernier chapitre de conclusion rappelle brièvement les achèvements majeurs de cette thèse, tant pour les développements méthodologiques que pour les études de cas avec acquisitions de nouvelles données MT. Ensuite, dans ce contexte, nous ouvrons et présentons les possibles prochains développements que je souhaiterais poursuivre dans le futur.

Première partie

Eléments de magnétotellurique

Chapitre 1

Induction dans un milieu conducteur par une onde monochromatique

La majeure partie des développements dans cette partie se retrouve dans les principales publications de magnétotellurique telles que Cagniard *(1953),* Rikitake *(1948), ou encore J. R.* Wait*, 1954. Ces auteurs utilisent des notations différentes de celles utilisées ci-dessous, mais les développements présentés ici sont extraits de ces publications.*

1.1 Milieu homogène

1.1.1 Régime général dans un milieu homogène

L'ensemble des méthodes utilisant les ondes électromagnétiques est régi par un ensemble d'équations appelées équations de Maxwell (Maxwell, 1865). Dans le vide, ces équations sont :

$$\nabla \cdot \mathbf{e} = \frac{\rho}{\varepsilon_0} \tag{1.1}$$

$$\nabla \cdot \mathbf{b} = 0 \tag{1.2}$$

$$\nabla \times \mathbf{e} = -\frac{\partial \mathbf{b}}{\partial t} \tag{1.3}$$

$$\nabla \times \mathbf{b} = \mu_0 \mathbf{j} + \mu_0 \varepsilon_0 \frac{\partial \mathbf{e}}{\partial t} \tag{1.4}$$

Avec :

$$\begin{aligned}
\mathbf{e} &= \text{Champ électrique } (V.m^{-1}) \\
\mathbf{b} &= \text{Induction magnétique } (T) \\
\mathbf{d} &= \text{Déplacement diélectrique } (C.m^{-2}) \\
\mathbf{h} &= \text{Champ magnétique } (A.m^{-1}) \\
\mathbf{j} &= \text{Courant électrique } (A.m^{-2})
\end{aligned}$$

L'équation (1.1) est appelée équation de *Maxwell-Gauss* et décrit la relation entre charges électriques et la création d'un champ électrique. L'équation (1.2) de *Maxwell-Flux* décrit le caractère conservatif du flux du champ magnétique impliquant l'absence de monopôles magnétiques. L'équation (1.3) décrit la création d'un champ électrique en présence d'une variation de champ magnétique. Enfin, l'équation (1.4) explique la création d'un champ magnétique en présence de courants ou de variations de champs électriques.
Dans un milieu de paramètres $\bar{\bar{\varepsilon}}$, $\bar{\bar{\mu}}$, $\bar{\bar{\sigma}}$, ces relations sont accompagnées de trois relations constitutives :

$$\begin{aligned}
\mathbf{d} &= \bar{\bar{\varepsilon}}\,\mathbf{e} \\
\mathbf{b} &= \bar{\bar{\mu}}\,\mathbf{h} \\
\mathbf{j} &= \bar{\bar{\sigma}}\,\mathbf{e} \qquad \text{appelée loi d'Ohm,}
\end{aligned}$$

où $\bar{\bar{\varepsilon}}$ est la *permittivité* du milieu $(F.m^{-1})$, $\bar{\bar{\mu}}$ est la *perméabilité* du milieu $(H.m^{-1})$ et $\bar{\bar{\sigma}}$ est la *conductivité* du milieu (en $S.m^{-1}$). En général, ces quantités sont considérées scalaires et μ est considérée égale à sa valeur attendue dans le vide ($\mu_0 = 4\pi.10^{-7} H.m$).

En magnétotellurique, le phénomène physique considéré est le phénomène d'*induction électromagnétique*, qui consiste en la création de courants électriques dans un matériau conducteur par la variation temporelle d'un champ magnétique source. Ce phénomène est particulièrement décrit par les équations (1.3) et (1.4).

A partir de ces relations, en appliquant l'opérateur $\nabla\times$ à (1.4), en utilisant l'identité : $\nabla \times \nabla \times \mathbf{u} = \nabla(\nabla \cdot \mathbf{u}) - \Delta\mathbf{u}$, et, en supposant que les paramètres du milieu μ, σ et ε ne varient

pas en fonction du temps,

$$\begin{aligned}
\nabla(\nabla\cdot\mathbf{e})-\Delta\mathbf{e} &= \nabla\times(-\frac{\partial\mathbf{B}}{\partial t}) && (1.5)\\
&= -\frac{\partial}{\partial t}(\nabla\times\mathbf{B}) && (1.6)\\
&= -\frac{\partial}{\partial t}\left(\mu\mathbf{j}+\mu\varepsilon\frac{\partial\mathbf{e}}{\partial t}\right) && (1.7)\\
&= -\mu\sigma\frac{\partial\mathbf{e}}{\partial t}+\mu\varepsilon\frac{\partial^2\mathbf{e}}{\partial t^2} && (1.8)
\end{aligned}$$

Dans un milieu homogène isotrope, aucune charge n'est présente, donc $\nabla\cdot\mathbf{e}=0$.

$$\Delta\mathbf{e}+\mu\varepsilon\frac{\partial^2\mathbf{e}}{\partial t^2}-\mu\sigma\frac{\partial\mathbf{e}}{\partial t}=\mathbf{0} \tag{1.9}$$

De la même façon, la relation équivalente pour $\mathbf{h}$ s'écrit :

$$\Delta\mathbf{h}+\mu\varepsilon\frac{\partial^2\mathbf{h}}{\partial t^2}-\mu\sigma\frac{\partial\mathbf{h}}{\partial t}=\mathbf{0} \tag{1.10}$$

Les relations (1.9) et (1.10) sont appelées *équations d'Helmholtz.*

Considérons maintenant le cas d'une onde monochromatique dont la variation temporelle est définie par $+i\omega t$, avec ω sa pulsation et $\mathbf{k}$ son vecteur d'onde. Supposons que φ, la phase à l'origine, est égale dans les trois directions, sa variation temporelle s'écrit :

$$\begin{aligned}
\mathbf{h}(\mathbf{x},t) &= \mathbf{h_0}\cos(\omega t+\varphi) && (1.11)\\
\mathbf{e}(\mathbf{x},t) &= \mathbf{e_0}\cos(\omega t+\varphi) && (1.12)
\end{aligned}$$

Cette onde s'écrit en notation complexe par :

$$\begin{aligned}
\mathbf{H}(\mathbf{x},t) &= \mathbf{h_0}e^{i\omega t+\varphi} && (1.13)\\
\mathbf{E}(\mathbf{x},t) &= \mathbf{e_0}e^{i\omega t+\varphi} && (1.14)
\end{aligned}$$

La notation complexe permet de réécrire les équations de Maxwell de façon beaucoup plus aisée. Pour obtenir les champs EM réels à partir de la notation complexe, il suffit de considérer :

$$\begin{aligned}
\mathbf{h}(\mathbf{x},t) &= \Re(\mathbf{H}(\mathbf{x},t)) && (1.15)\\
\mathbf{e}(\mathbf{x},t) &= \Re(\mathbf{E}(\mathbf{x},t)) && (1.16)
\end{aligned}$$

Ainsi, en notation complexe, les équations d'Helmoltz s'écrivent :

$$\Delta \mathbf{E} + k^2 \mathbf{E} = \mathbf{0} \tag{1.17}$$

$$\Delta \mathbf{H} + k^2 \mathbf{H} = \mathbf{0} \tag{1.18}$$

Avec :

$$k^2 = \omega^2 \mu \varepsilon - i\omega\mu\sigma \tag{1.19}$$

Le nombre d'onde k^2 contient deux processus physiques :

- $k_p^2 = \omega^2 \mu \varepsilon$ décrit les processus de propagation liés aux courants de convection $\mathbf{j_D}$.
- $k_d^2 = \omega\mu\sigma$ décrit les processus de diffusion liés aux courants de conduction $\mathbf{j_C}$.

1.1.2 Régime diffusif

Dans la majorité des méthodes EM utilisées en géophysique, les fréquences utilisées sont relativement basses, inférieures à la dizaine de kHz (Swift, 1967). Dans ce cas :

$$\omega^2 \mu \varepsilon \ll \omega\mu\sigma \tag{1.20}$$

Cette hypothèse, appelée approximation des régimes quasi-stationnaires, traduit le caractère négligeable des courants de déplacement devant les courants de diffusion. Les équations d'Helmoltz en domaine temporel deviennent donc :

$$\Delta \mathbf{e} = \mu\sigma \frac{\partial \mathbf{e}}{\partial t} \tag{1.21}$$

Cette équation traduit le même processus physique que la diffusion de la chaleur. Dans les méthodes EM utilisant les hautes fréquences, comme le géoradar, les courants de diffusion deviennent négligeables devant les courants de déplacement. Les équations obtenues sont les équations d'ondes telles qu'utilisées en géophysique sismique pour décrire la propagation d'une onde mécanique dans le sol.

Les paramètres étudiés varient aussi selon les deux types de méthodes électromagnétiques. Dans le cas des méthodes diffusives, le paramètre étudié est la résistivité électrique σ. Dans le cas des méthodes de propagation, le paramètre étudié est la permittivité électrique ε.

Les équations à résoudre sont donc :

$$\Delta \mathbf{E} - i\omega\mu\sigma \mathbf{E} = \mathbf{0} \tag{1.22}$$

$$\Delta \mathbf{H} - i\omega\mu\sigma \mathbf{H} = \mathbf{0} \tag{1.23}$$

1.1.3 Diffusion dans un milieu homogène

1.1.3.1 Diffusion

Considérons la diffusion d'une onde incidente à la surface du sous-sol, de conductivité constante σ. L'onde est incidente de vecteur d'onde $\mathbf{k} = k\mathbf{e_z}$. L'équation d'Helmoltz en milieu homogène devient alors :

$$\frac{\partial^2 H_x}{\partial z^2} - k^2 H_x = 0 \tag{1.24}$$

$$\frac{\partial^2 H_y}{\partial z^2} - k^2 H_y = 0 \tag{1.25}$$

Les solutions de ces équations sont :

$$H_x(z,t) = H_{x,0} e^{i\omega t} e^{-ikz} \tag{1.26}$$

$$H_y(z,t) = H_{y,0} e^{i\omega t} e^{-ikz} \tag{1.27}$$

La régime est diffusif : $k = \sqrt{\omega\mu\sigma/2} - i\sqrt{\omega\mu\sigma/2} = \alpha - i\beta$. Le champ magnétique est donc égal à :

$$H_x(z,t) = H_{x,0} e^{i\omega t} e^{-i\alpha z} e^{-\beta z} \tag{1.28}$$

$$H_y(z,t) = H_{y,0} e^{i\omega t} e^{-i\alpha z} e^{-\beta z} \tag{1.29}$$

Le terme en $e^{-\beta z}$ décrit une atténuation exponentielle dans la direction verticale. La profondeur δ à laquelle l'amplitude du champ est diminuée d'un facteur e est appelée *profondeur de peau*, égale à :

$$\delta = \frac{1}{\beta} = \sqrt{\frac{2}{\omega\mu\sigma}} \tag{1.30}$$

1.1.3.2 Relation entre champ magnétique et électrique

Dans le milieu homogène, l'équation de Maxwell-Ampère donne :

$$\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} = \sigma E_x \tag{1.31}$$

$$\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} = \sigma E_y \tag{1.32}$$

$$\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} = \sigma E_z \tag{1.33}$$

Par hypothèse, le milieu étant homogène, les dérivées latérales sont nulles. Ces relations se simplifient en :

$$\frac{\partial H_x}{\partial z} = \sigma E_y \tag{1.34}$$

$$\frac{\partial H_y}{\partial z} = \sigma E_x \tag{1.35}$$

Or, comme $\sigma = k^2/(-i\omega\mu)$, ces équations deviennent :

$$\frac{E_y}{H_x} = \frac{-ik}{k^2/(-i\omega\mu)} \tag{1.36}$$

$$\frac{E_y}{H_x} = \frac{-\omega\mu}{k}, \tag{1.37}$$

De la même façon :

$$\frac{E_x}{H_y} = \frac{\omega\mu}{k}. \tag{1.38}$$

Le rapport entre le champ électrique et le champ magnétique est défini comme l'*impédance électrique* Z. Deux paramètres physique sont déterminés à partir de cette impédance, la résistivité ρ ($\Omega.m$) et la phase φ (en radians) (Cagniard, 1953), tels que :

$$\rho = \frac{|Z|^2}{\omega\mu} = \frac{1}{\omega\mu}\left|\frac{E_x}{H_y}\right|^2 = \frac{1}{\omega\mu}\left|\frac{E_y}{H_x}\right|^2 \tag{1.39}$$

$$\varphi = \arctan\left(\frac{E_x}{H_y}\right) = \arctan\left(\frac{-E_y}{H_x}\right) = \frac{\pi}{4} \tag{1.40}$$

La phase de l'impédance pour un milieu homogène vaut donc $\pi/4$.

1.2 Diffusion dans un milieu unidimensionnel

Les premières études magnétotelluriques se sont basées sur l'interprétation de sondages MT en utilisant l'hypothèse de milieux tabulaires (Cagniard, 1953 ; J. R. Wait, 1954). Pour un milieu tabulaire dont la résistivité ne varie qu'avec la profondeur, Cagniard (1953) a montré que, comme dans le cas homogène, les champs électriques et magnétiques sont reliés par :

$$E_x = Z_{xy} H_y \tag{1.41}$$

$$E_y = -Z_{xy} H_x \tag{1.42}$$

Il a démontré les relations de calcul de l'impédance pour des milieux de 1 à 3 couches de résistivité différentes. J. R. Wait (1954) a poursuivi ces travaux en montrant que la réponse MT d'un milieu tabulaire se calcule par une formule de récurrence.

Soit un milieu tabulaire à N couches de conductivité σ_i et d'épaisseur h_i situé au dessus d'un demi-espace de conductivité σ_{n+1} (Figure 1.1). L'impédance de chaque couche Z_j est définie

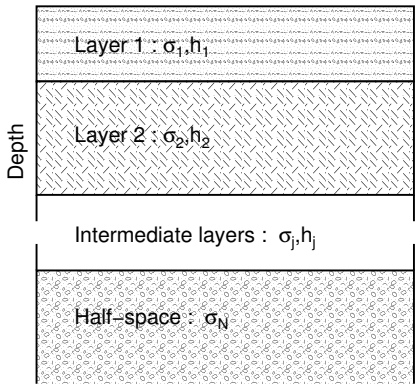


Figure 1.1: Représentation d'un milieu 1D dont la variation de résistivité est verticale.
FIGURE 1.1: *Representation of a 1D conductivity medium. Only vertical variation of resistivity is considered.*

par :

$$Z_j(\omega) = \frac{\omega\mu}{k_j} \qquad \text{avec } k_j = -i\omega\mu\sigma_j \tag{1.43}$$

Le calcul de l'impédance effective $\tilde{Z}$ en surface commence par celui de la couche n :

$$\tilde{Z}_n(\omega) = \frac{Z_{n+1}(\omega) + Z_n(\omega)\tanh(ik_nh_n)}{Z_n(\omega) + Z_{n+1}(\omega)\tanh(ik_nh_n)} \tag{1.44}$$

L'impédance effective de chaque couche $(j-1)$ se détermine grâce à l'impédance effective de la couche du dessous j par la relation :

$$\tilde{Z}_{j-1}(\omega) = \frac{\tilde{Z}_j(\omega) + Z_{j-1}(\omega)\tanh(ik_jh_j)}{Z_{j-1}(\omega) + \tilde{Z}_j(\omega)\tanh(ik_jh_j)} \tag{1.45}$$

Il est important de noter ici que l'impédance effective d'une couche ne dépend que de celles des couches situées en dessous. Finalement, l'impédance de surface se détermine en calculant par récurrence les impédances effectives de chaque couche jusqu'à remonter à la première couche de surface.

Finalement, deux paramètres peuvent se définir à partir de l'impédance effective à la surface, *la résistivité apparente* ρ_a $(\Omega.m)$ et la *phase* φ (rad), définis par :

$$\rho_a(\omega) = \frac{|\tilde{Z}(\omega)|^2}{\omega\mu} \tag{1.46}$$

$$\varphi_a(\omega) = \arctan\frac{\Im(\tilde{Z}(\omega))}{\Re(\tilde{Z}(\omega))} \tag{1.47}$$

Le processus physique sous-jacent à ces quantités étant un processus diffusif, ces paramètres sont dits "apparents". Ils prennent donc en compte l'ensemble des propriétés (conductivité, géométrie) du milieu dans lequel l'onde se diffuse. Une onde haute fréquence étant rapidement atténuée, la réponse électrique en surface ne sera déterminée que par les propriétés des couches proches de la surface. Plus l'onde est basse fréquence, plus l'atténuation se fait en profondeur. La réponse électrique en surface est donc déterminée par l'ensemble des couches traversées avant atténuation. Ce processus impose donc les paramètres apparents à ne présenter que peu de variations en fonction de la fréquence de l'onde. Si l'impédance, après traitement des données MT, présente des discontinuités entre deux fréquences proches, le sondage doit être interprété avec précaution (la réciproque n'étant pas vraie, un sondage lisse pouvant être faux).

La représentation en fonction de la fréquence n'étant pas suffisante pour décrire les milieux géologiques, différentes méthodes permettant d'interpréter les sondages magnétotelluriques effectués sur un milieu 1D ont été développées. Parmi-celles ci, les méthodes de Niblett et Bostick (Niblett et Sayn-Wittgenstein, 1960 ; F. X. Bostick, 1977) ; démontrées équivalentes par A. G. Jones (1983) ; permettent d'effectuer une transformation du sondage afin de passer d'une représentation résistivité apparente/fréquence à une représentation en résistivité/profondeur.

1.3 Milieux avec variations latérales de conductivité

1.3.1 Tenseur d'impédance

Pour des milieux dont la distribution de conductivité est plus complexe, la relation entre le champ électrique **E** et le champ magnétique **H** est exprimée sous forme tensorielle en domaine fréquentiel, par l'intermédiaire du *tenseur d'impédance* $\bar{\bar{Z}}$ (Cantwell, 1960), où $\omega = 2\pi f$:

$$\begin{pmatrix} E_x(\omega) \\ E_y(\omega) \end{pmatrix} = \begin{pmatrix} Z_{xx}(\omega) & Z_{xy}(\omega) \\ Z_{yx}(\omega) & Z_{yy}(\omega) \end{pmatrix} \begin{pmatrix} H_x(\omega) \\ H_y(\omega) \end{pmatrix}, \tag{1.48}$$

Ce tenseur, s'il est exprimé entre **E** et **B**, est appelé *tenseur de réponse*. Le tenseur d'impédance contient les informations sur les propriétés géo-électriques du milieu étudié. Chaque composante est reliée à deux quantités interprétables, la *résistivité apparente* ρ_a $(\Omega.m)$ et la *phase* φ (en

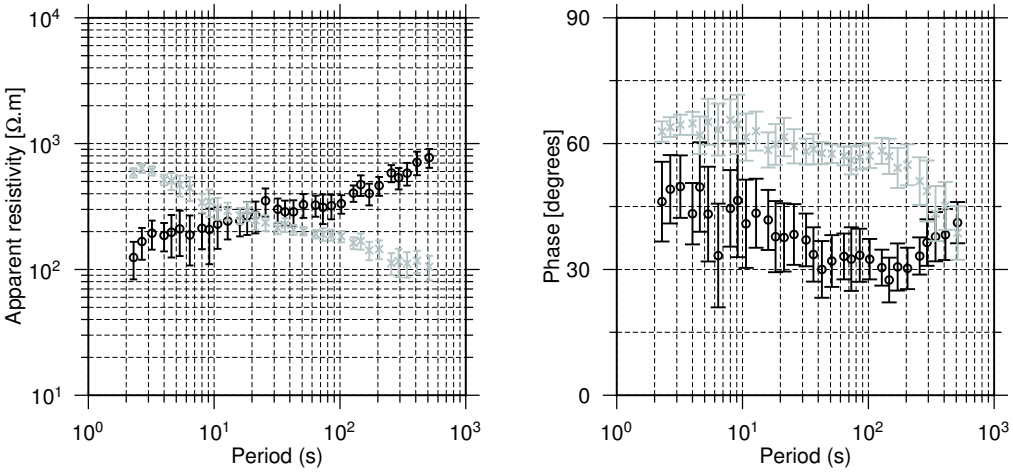


FIGURE 1.2: Représentation classique des sondages magnétotelluriques. Seules les composantes anti-diagonales ont été représentées ici. Noir : Composante Z_{xy}. Gris : Composante Z_{yx}. Les barres verticales indiquent les incertitudes autour de l'estimation.
FIGURE 1.2: *Classical representation of MT soundings. Only off-diagonal components are shown. Black : Z_{xy}. Grey : Z_{yx}. Uncertainties are indicated by vertical bars.*

radians) définis par :

$$\rho_{ij}(\omega) = \frac{|Z_{ij}(\omega)|^2}{\omega\mu} \text{ où l'indice } _a \text{ est retiré pour simplicité d'écriture.} \tag{1.49}$$

$$\varphi_{ij}(\omega) = \arctan \frac{\Im(Z_{ij}(\omega))}{\Re(Z_{ij}(\omega))} \tag{1.50}$$

Les tenseurs sont conventionnellement représentés à travers ces paramètres. Chaque composante est régulièrement échantillonnée en fréquence et accompagnée d'une estimation ainsi que d'une incertitude. La Figure 1.2 montre un exemple de sondage MT. La relation entre fréquence et profondeur de peau δ permet d'interpréter rapidement les sondages MT. Les hautes fréquences correspondent à des couches superficielles, tandis que les longues périodes correspondent aux structures géologiques profondes.

La géologie du sous-sol étudié influe sur l'amplitude et la phase de chaque élément du tenseur d'impédance. L'étude de l'ensemble du tenseur permet de déterminer un modèle de conductivité électrique du milieu étudié. Des outils existent pour effectuer cette analyse. Ils seront présentés à l'Annexe D.

1.3.2 Fonction de transfert verticale

Si le milieu présente des anomalies latérales de conductivité, l'induction du champ magnétique créé une composante magnétique verticale (Parkinson, 1959 ; Parkinson, 1962). Cette composante est reliée aux composantes magnétiques horizontales par les fonctions de transfert

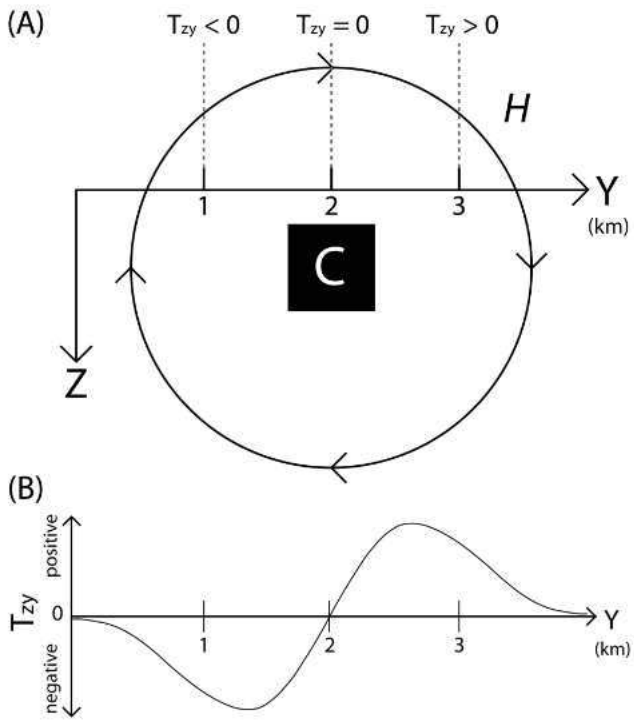


FIGURE 1.3: Haut : Le corps conducteur est indiqué par la lettre **C** et s'étend de façon infinie dans la direction horizontale perpendiculaire à Y. Bas : Évolution de la composante T_y le long du profil dans la direction Y.
FIGURE 1.3: *Top : The conductive body is the black box indicated by the letter **C** and is extending infinitely in the horizontal direction perpendicular to Y. Bottom : T_y variation along the profil in the Y direction.*

appelées *tipper*, et notées T telles que :

$$H_z(\omega) = T_x(\omega)H_x(\omega) + T_y(\omega)H_y(\omega) \tag{1.51}$$

Les fonctions de transfert verticales sont traditionnellement représentées sous la forme de vecteurs, appelés *flèches d'induction*. Deux conventions existent pour les représenter, la convention de Parkinson (Parkinson, 1962) et la convention de Wiese (Wiese, 1962). Dans la convention de Parkinson, le vecteur partie réelle est inversée et pointe vers les zones conductrices (ou loin des zones résistantes). Dans la convention de Wiese, ce vecteur pointe loin des zones conductrices (et vers les zones résistantes). Pour la convention de Parkinson, les vecteurs sont donc :

$$\mathbf{c}_{\Re} = -\Re(T_x)\mathbf{e_x} - \Re(T_y)\mathbf{e_y} \tag{1.52}$$

$$\mathbf{c}_{\Im} = \Im(T_x)\mathbf{e_x} + \Im(T_y)\mathbf{e_y} \tag{1.53}$$

La Figure 1.3 illustre le principe de la variation des fonctions de réponse verticales au dessus d'un corps conducteur. La composante T_y change de signe de chaque côté du corps conducteur

et s'annule au dessus, indiquant la position du corps.

Deuxième partie

Traitement de données magnétotelluriques

Chapitre 2

Détection et caractérisation des sources géomagnétiques

2.1 Environnement électromagnétique

La méthode magnétotellurique (MT) se base sur la présence de sources naturelles externes de courants produisant un champ magnétique. Comme présenté au Chapitre 1, la profondeur de pénétration dépend de la fréquence de l'onde induisant des courants dans le sous-sol. L'intérêt de la méthode MT provient de la variété des sources naturelles externes permettant une caractérisation à plusieurs échelles du sous-sol étudié. Un exemple de spectre de puissance du champ magnétique terrestre est représenté Figure 2.1 (C. Constable, 2016) pour des périodes allant de quelques dizaines de millions d'années à moins d'un dix-millième de seconde. Deux sources majeures ; internes et externes ; dont les origines sont radicalement différentes, interviennent dans les variations temporelles du champ dont nous décrirons les principales caractéristiques dans les sections suivantes.

2.1.1 Eléments du champ interne

La Terre peut se décomposer au premier ordre en une succession de couches de la surface jusqu'au centre. Les principales discontinuités (Dziewonski et Anderson, 1981) établissent l'existence de quatre couches majeures :

- Le noyau interne.
- Le noyau externe liquide.
- Le manteau.
- La lithosphère.

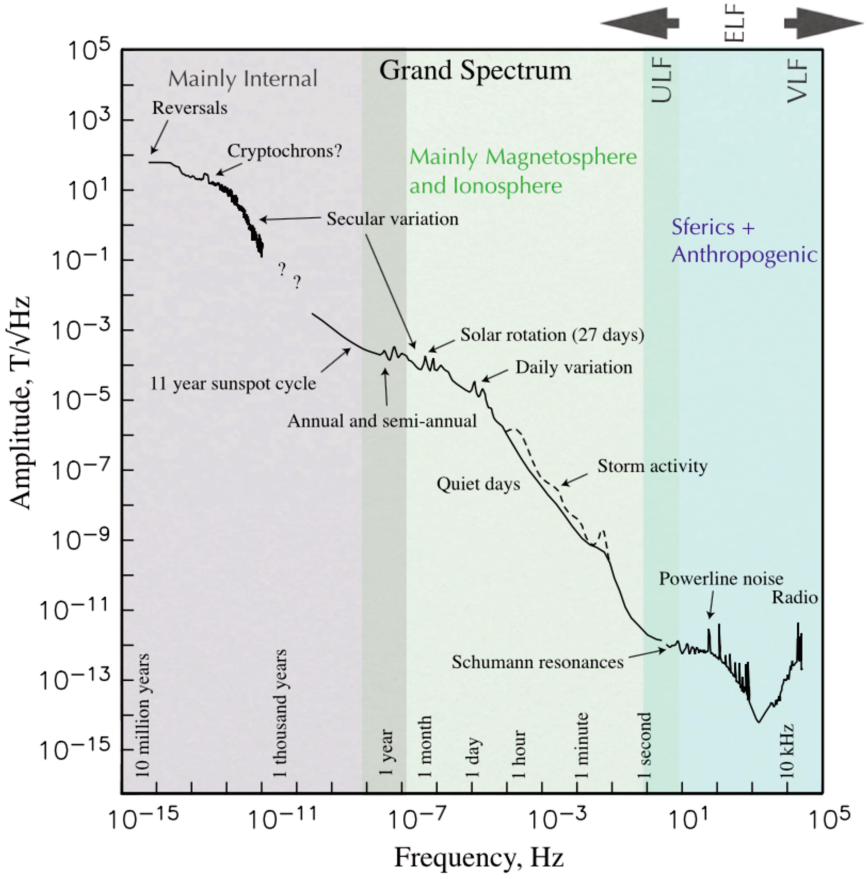


Figure 2.1: Spectre en puissance des variations naturelles du champ magnétique. D'après C. Constable (2016).

FIGURE 2.1: *Power spectrum of natural variations of the magnetic field. From C.* Constable *(2016)*

Le champ magnétique à la surface de la Terre est en première approximation (∼90 %) équivalent à celui d'un dipôle magnétique dont ses sources se situent dans le noyau externe liquide. Ce noyau est liquide et est composé principalement de fer et de nickel. Par convection, des mouvements de fluides dans le noyau liquide sont provoqués, qui, par induction, vont provoquer la création de courants électriques dans le noyau. Ce mécanisme est appelé géodynamo est est auto-entretenu depuis au moins 3 milliards d'années. La géodynamo va entraîner un nombre important de phénomènes dont les variations temporelles ont des périodes observables au delà de deux mois. Parmi les phénomènes observables, les inversions du champ sont les plus notables.

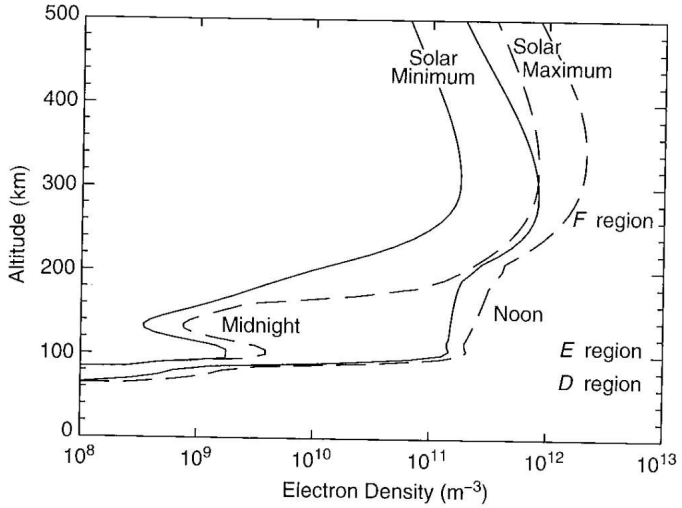


FIGURE 2.2: Profil en altitude de la densité électronique aux coordonnées 18°N, 67°W, enregistré à l'équinoxe de Septembre. D'après RICHMOND (2007).
FIGURE 2.2: *Vertical profil of electronic density recorded at September equinox at 18°N, 67°W. From* RICHMOND *(2007).*

2.1.2 Eléments du champ externe

2.1.2.1 Ionosphère

L'ionosphère est une couche chargée de l'atmosphère dont la limite inférieure se situe à moins de 100 km d'altitude. Cette couche est créée par l'ionisation des molécules de la haute atmosphère par les rayonnements ultraviolets et X provenant du soleil. Cette ionisation crée un plasma permettant la conduction de courants électriques et influençant la propagation des ondes radio.

L'ionosphère est traditionnellement divisée en un nombre fini de couches. Ces couches sont reliées au maxima locaux de densité d'électrons (Figure 2.2). La couche située au dessus de 150 km est appelée couche *F*. Cette couche contient le maximum global de densité d'électrons. La couche possédant le second maximum local est appelée la couche *E*, et est située aux alentours de 110 km. Au dessous de 90 km d'altitude se trouve la couche *D*. Dépendant de l'incidence des rayons électromagnétiques provenant du soleil, la densité d'électrons varie avec les cycles solaires et le cycle diurne.

L'ionosphère est le siège de nombreux phénomènes électromagnétiques. La variation diurne observable sur les magnétogrammes des observatoires magnétiques trouve ainsi sa source dans le phénomène de dynamo ionosphérique (Figure 2.3).

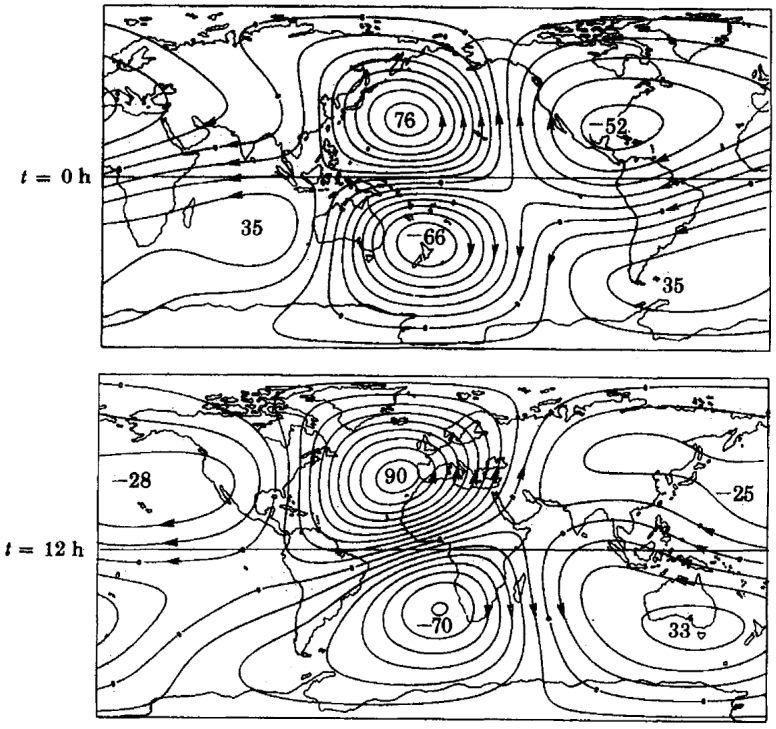


FIGURE 2.3: Système de courant ionosphérique moyen pour l'année 1964–1965. Les isovaleurs sont indiquées en kA. D'après RICHMOND (2007).
FIGURE 2.3: *Annual mean of the ionospheric current system for year 1964. Values are indicated in kA. From* WINCH *(2007).*

2.1.2.2 Magnétosphère

La magnétosphère est la région de l'espace où le mouvement des particules chargées (ions, électrons) est défini par le champ magnétique terrestre par opposition au champ magnétique interplanétaire (IMF, pour *Interplanetary Magnetic Field*). Cette région, commençant à environ 500 km au dessus de la surface terrestre (COWLEY, 2007), a une structure complexe qui est en constante évolution du fait son interaction avec le vent solaire (Figure 2.4).

Le soleil projette dans l'espace interplanétaire un plasma de protons et d'électrons, dont l'ensemble est appelé le vent solaire. La vitesse des particules émises est de l'ordre de 500 km/s en moyenne mais peut atteindre la vitesse de 1200-1300 km/s lors d'un orage magnétique.

Comme vu précédemment, au premier ordre, le champ magnétique terrestre peut être considéré comme dipolaire. Ce champ est plongé dans l'IMF qui va alors déterminer sa géométrie en le confinant dans une cavité appelée cavité de Chapman-Ferraro (CHAPMAN et FERRARO, 1931). La description de la magnétosphère terrestre se base sur le principe du "flux gelé" (COWLEY, 1995). Ce principe postule qu'une particule en giration autour d'une ligne de champ magnétique

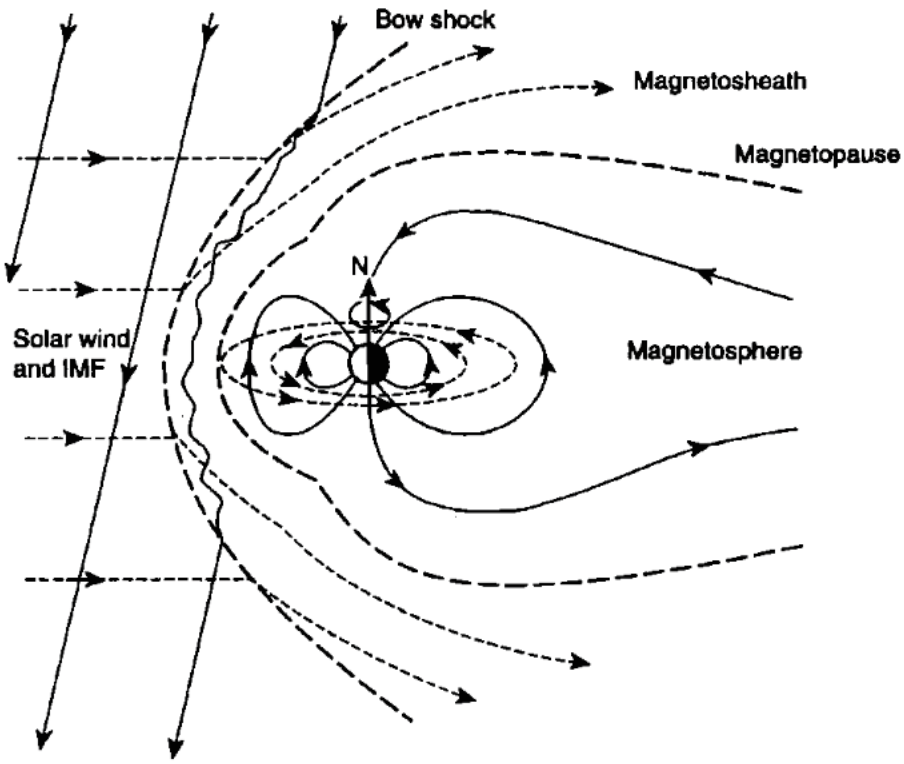


FIGURE 2.4: Schéma de la géométrie de la magnétosphère dans l'approximation du flux gelé. Les traits pleins indiquent les lignes de champ magnétique. Les pointillés les lignes de courant du plasma. D'après COWLEY (1995).
FIGURE 2.4: *Sketch of magnetosphere geometry in frozen-flux approximation. Magnetic field lines are indicated by full lines. Dashed lines indicated plasma flux lines. From* COWLEY *(1995)*

reste en giration autour de cette ligne de champ à tout instant. Ainsi, le mouvement du plasma dans la magnétosphère est conjugué à celui de l'organisation des lignes de champ terrestre.

En approchant du voisinage de la Terre, la vitesse du vent solaire va diminuer et celui-ci sera dévié à partir de la structure appelée onde de choc (ang. *bow shock*). L'action du vent solaire va alors compresser du côté diurne la magnétosphère, rapprochant les lignes de champ magnétique. La limite géographique où la pression du vent solaire et la pression magnétique de la magnétosphère se compense s'appelle la *magnétopause*. Entre les deux limites, l'espace est appelé magnétogaine (ang. *magnetosheath*). Dans cette région, le vent solaire est compressé et chauffé.

Du côté nuit, les lignes de champ vont s'étirer et former la *queue de la magnétosphère* sur plusieurs dizaines de rayons terrestres. Lors d'un orage magnétique, les lignes de champ situées au niveau de la magnétopause vont être étendues et ramenées au niveau de la queue de la magnétosphère, où des phénomènes de reconnexion magnétiques vont se produire (Figure 2.5). Ce cycle de reconnexion est appelée Cycle de Dungey (DUNGEY, 1961).

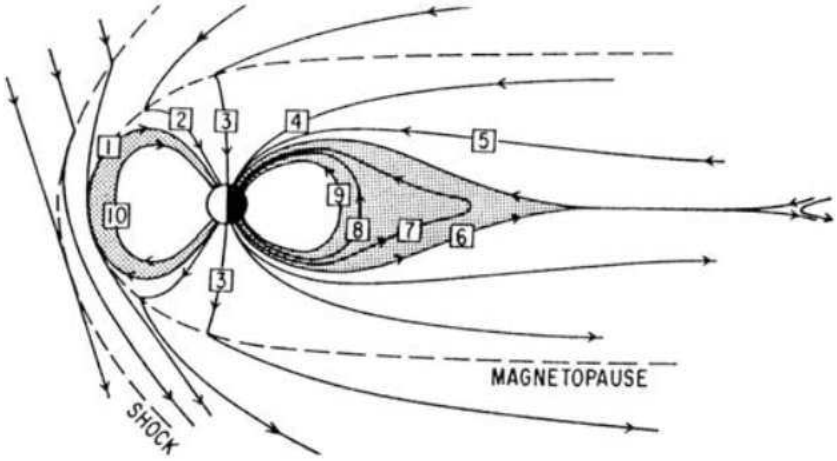


FIGURE 2.5: Schéma du cycle de Dungey. D'après AXFORD (1969).
FIGURE 2.5: *Sketch of Dungey cycle. From* AXFORD *(1969)*

2.2 Ondes haute fréquence (> Hz)

2.2.1 Origine

Les sources naturelles de fréquences supérieures au Hz utilisées en MT proviennent de l'activité orageuse dans le monde (C. CONSTABLE, 2016). Un éclair unique permet l'émission de plusieurs ondes électromagnétiques (EM) dont les mécanismes de propagation diffèrent suivant la nature de ces ondes. Deux principaux mécanismes de propagation sont observables : une propagation dans le guide d'onde formé par l'ionosphère et la surface de la Terre (Figure 2.6), et une propagation le long des lignes de champ du champ magnétique terrestre par transmission dans l'ionosphère.

Dans les sections suivantes, nous décrirons les trois grands types d'ondes, d'origine atmosphérique, étudiées dans cette thèse.

2.2.2 Formes des ondes

2.2.2.1 *Sferic* - Ondes VLF (*Very Low Frequency*)

Les *radio atmospherics* ou *sferics* sont les ondes ayant le contenu fréquentiel le plus élevé, entre 1 kHz et 50 kHz, recouvrant le domaine fréquentiel des émetteurs artificiels VLF. Ces ondes ont une forme oscillatoire amortie dépendant de la distance à la source. Dans les premiers kilomètres, leur forme peut être expliquée par des principes de réflexion et transmission aux interfaces ionosphère/surface de la Terre (RAKOV et UMAN, 2003). A longue distance, les sources

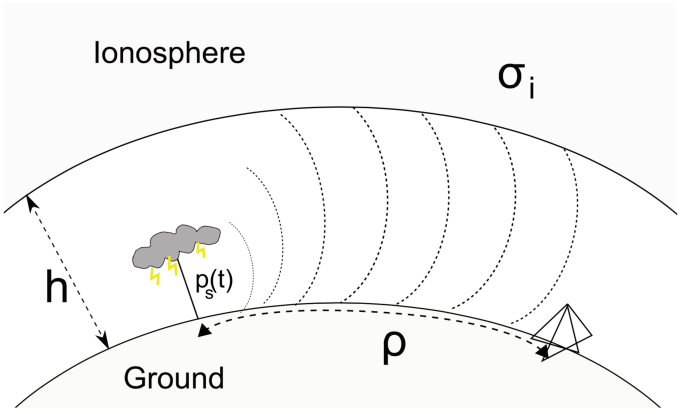


FIGURE 2.6: Schéma représentant la propagation des ondes EM issues de l'occurrence d'un éclair atmosphérique. D'après MACKAY et A. C. FRASER-SMITH (2010).
FIGURE 2.6: *Sketch of atmospheric waves propagation from lightning strike occurence to measuring station. From* MACKAY *et A. C.* FRASER-SMITH *(2010).*

ont une forme d'onde sinusoïdale. La Figure 2.7 représente un *sferic* dont le champ EM a été mesuré sur le dôme du volcan de la Soufrière, à la station de Guenon, le 16 novembre 2011 (P. SAILHAC et al., 2016). Le spectre associé présente une rapide diminution de l'amplitude pour des fréquences inférieures à 4 kHz. Cette diminution est liée à l'atténuation dans une bande de fréquence appelée l'Audio-MT *dead band*. Dans cette intervalle de fréquence, généralement compris entre 1 et 5 kHz, l'amplitude des ondes EM est fortement attenuée. La largeur fréquentielle de l'Audio-MT *dead band* varie en fonction de la hauteur de l'ionosphère, des conditions ionosphériques... (RAKOV et UMAN, 2003).

L'augmentation du spectre en puissance pour les fréquences inférieures à 1 kHz correspond à la présence d'une seconde onde appelée *slow tail* (voir section 2.2.2.2).

J. R WAIT (1960b) a notamment développé une théorie sur la propagation des *sferics* dans le guide d'onde formé par l'ionosphère et la surface de la Terre. Cette théorie est basée sur la propagation par modes propres et permet de prédire la dispersion des ondes sur le trajet entre la source et le site de mesure ainsi que l'atténuation des fréquences dans l'Audio-MT *dead band*.

2.2.2.2 *Slow tails* - Ondes ELF (*Extremely Low Frequency*)

En dessous d'une certaine fréquence, d'environ 1 kHz, dépendant de la hauteur de l'ionosphère, la théorie des modes propres prédit une atténuation des ondes VLF. Un autre type d'onde, appelée *slow tail*, peut alors se propager dans l'atmosphère, Ces ondes ont un contenu fréquentiel en général compris entre 10 Hz et 1 kHz. La forme des *slow tails* est impulsive avec des oscillations très rapidement amorties. La Figure 2.8 montre un exemple de *slow tail* mesurée

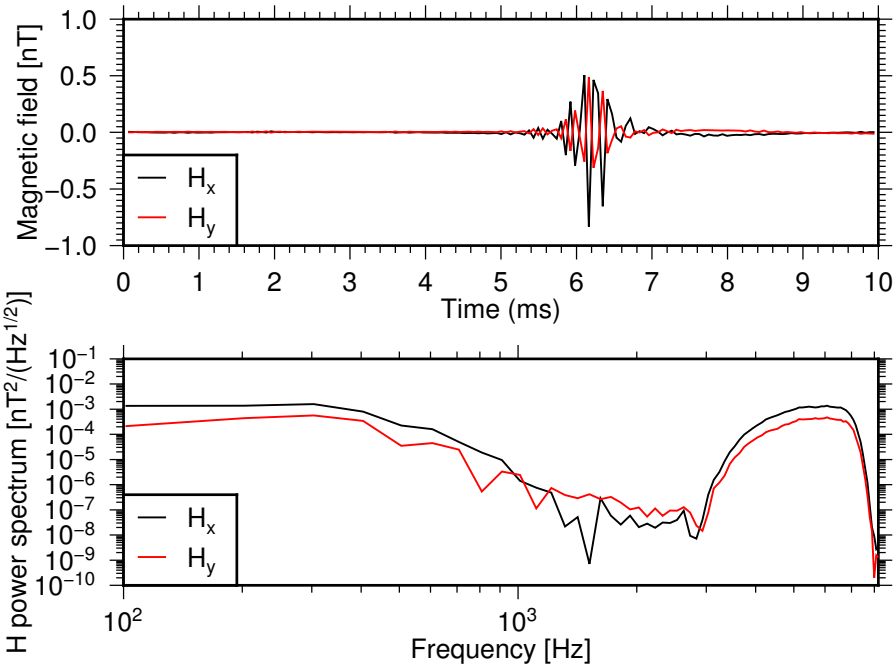


FIGURE 2.7: Haut : Onde VLF en domaine temporel pour un échantillonnage à 65536 Hz. Bas : Spectre en puissance des séries temporelles présentées au dessus.
FIGURE 2.7: *Top : Atmospheric wave in time domain for a sampling frequency of 65,536 Hz. Bottom : Power spectrums of the above time series.*

sur le dôme du volcan La Soufrière, à la station de Guenon. Le spectre d'amplitude décroît avec les fréquences croissantes.

J. R Wait (1960a), comme dans le cas des ondes VLF, a proposé une modélisation numérique de ces ondes basées sur la théorie des modes propres. Cette théorie permet notamment d'utiliser les ondes ELF dans la caractérisation de la distance aux éclairs dont la *slow tail* est suffisamment forte en amplitude (Mackay et A. C. Fraser-Smith, 2010).

2.2.2.3 *Whistlers*

Les *whistlers* sont des ondes EM qui se propagent dans la magnétosphère par transmission dans l'ionosphère. Leur nom provient du sifflement qui fut entendu dans les radios pendant la première guerre mondiale (Barkhausen, 1930). Ces ondes ont un contenu fréquentiel entre 100 Hz et 10 kHz (Figure 2.9), et se superposent aux deux ondes précédentes. Leur maximum d'amplitude est généralement observé autour de 5 kHz (Helliwell, 1965). Contrairement aux ondes se propageant uniquement dans l'atmosphère, ces ondes ne subissent pas l'atténuation dans l'Audio-MT *dead band*. La caractéristique principale des *whistlers* est la dispersion fréquentielle subie par cette onde lors de son passage dans l'ionosphère et la magnétosphère. Deux principaux types de mesure permettent d'étudier les *whistlers* : les mesures au sol et les mesures satellitaires.

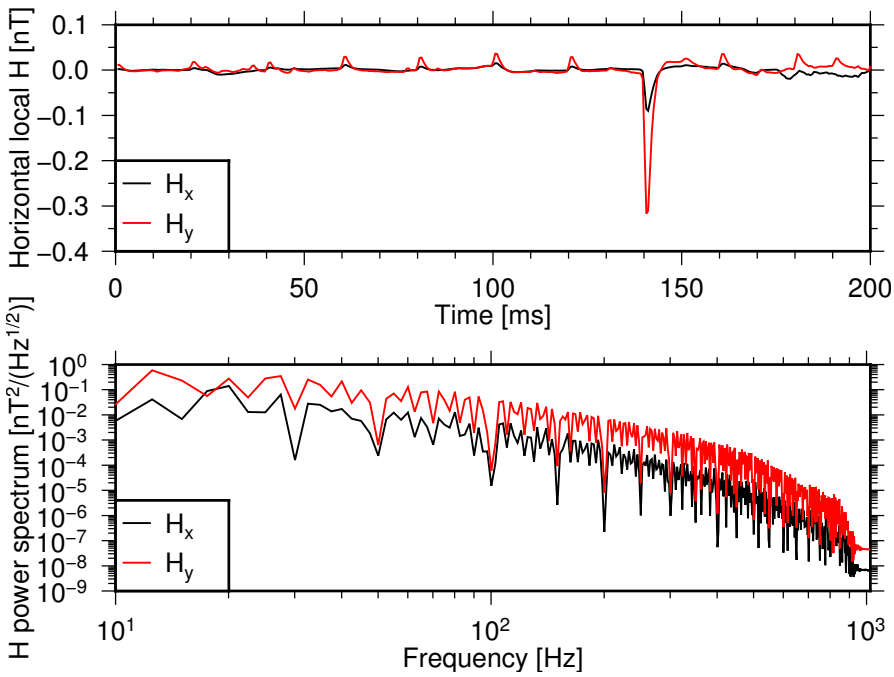


FIGURE 2.8: Haut : Onde ELF en domaine temporel pour un échantillonnage à 2048 Hz. Bas : Spectre de puissance des séries temporelles présentées au dessus.
FIGURE 2.8: *Top : Slow tail in time domain for a sampling frequency of 2,048 Hz. Bottom : Power spectrum of above time series.*

Ces dernières permettent de donner des informations sur la trajectoire des *whistlers* dans la magnétosphère (Helliwell, 1965). Ce type d'onde permet par exemple de renseigner sur la composition électronique des milieux traversés (Carpenter et al., 1981). Un exemple d'onde *whistler* enregistré sur le dôme du volcan La Soufrière, à la station de la Faille de la Ty, est présenté Figure 2.9.

2.2.2.4 Autres ondes utilisables

D'autres types d'ondes permettent d'étudier l'induction de courants électromagnétiques dans la Terre. Parmi celles-ci, les résonances de Schumann et les *Q-Bursts* sont notamment été utilisées pour la caractérisation de l'impédance MT.

Résonances de Schumann Les résonances de Schumann sont les témoignages de l'activité orageuse dans le monde. Si on considère une onde électromagnétique de vitesse 3.10^8 m/s, une circonférence de la Terre moyenne de $2\pi a$ avec $a = 6371$ km, le temps que met une onde EM pour faire le tour de la Terre est d'environ (1/7.5) s, ce qui correspond à une fréquence de 7.5 Hz. En réalité, à cause de la pression du vent solaire, de la composition électronique de l'ionosphère, les fréquences du fondamental à 7.5 Hz et des harmoniques varient. Les résonances de Schumann

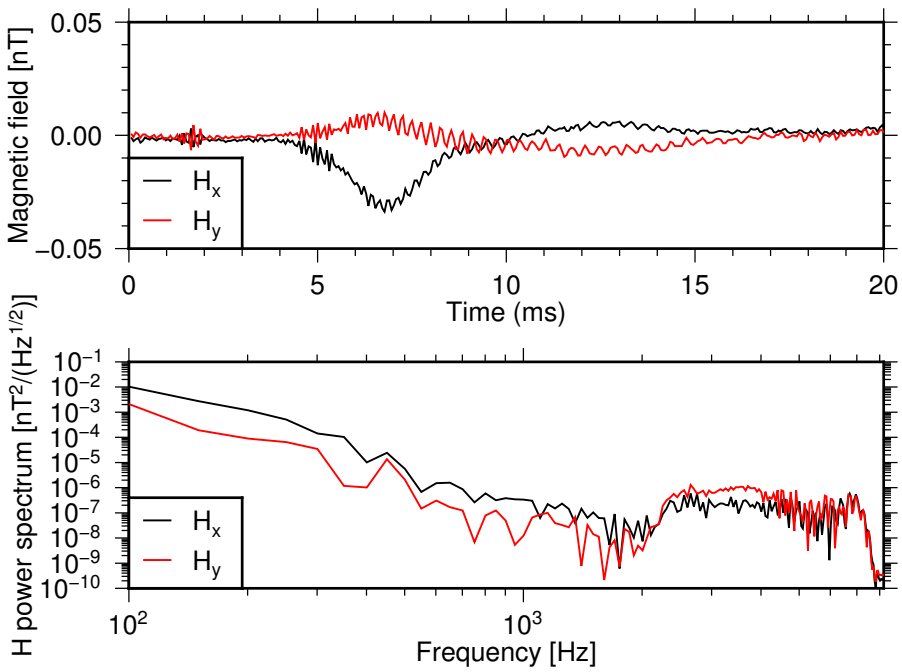


FIGURE 2.9: Haut : Onde *whistler* en domaine temporel sur le champ magnétique horizontal pour un échantillonnage à 65536 Hz. A noter la présence d'une large *slow tail* simultanée. Bas : Spectres de puissance des séries présentées au dessus.
FIGURE 2.9: *Top : Whistler in time domain on horizontal magnetic field for a sampling frequency of 65,536 Hz. Notice the presence of a large slow tail appearing simultaneously. Bottom : Power spectrums of above time series.*

correspondent aux résonances des ondes électromagnétiques émises par l'activité orageuse autour de la planète. Outre le fondamental d'environ 7.8 Hz, les harmoniques à 14.1, 20.3, 26.3 et 32.5 Hz sont clairement identifiables (T. Madden et Thompson, 1965) (Figure 2.10).
Ces résonances permettent de renseigner sur l'état électronique de l'ionosphère, ainsi que sur sa hauteur (Chand, Israil et Rai, 2009).

Q-Bursts Les *Q-bursts* sont des ondes EM identifiées par Ogawa, Tanaka et al. (1966). Ces ondes présentent des maximums d'amplitude en fréquence correspondant aux fréquences des résonances de Schumann (Ogawa et Komatsu, 2007) et sont reliées aux événements atmosphériques appelés *sprites* (Boccippio et al., 1995). Ces signaux présentent plusieurs formes d'ondes présentant des caractéristiques différentes (Ogawa et Komatsu, 2007). Un exemple d'onde *Q-Burst* enregistré à Chambon-La-Forêt, à la station Ingrannes est présenté Figure 2.11.

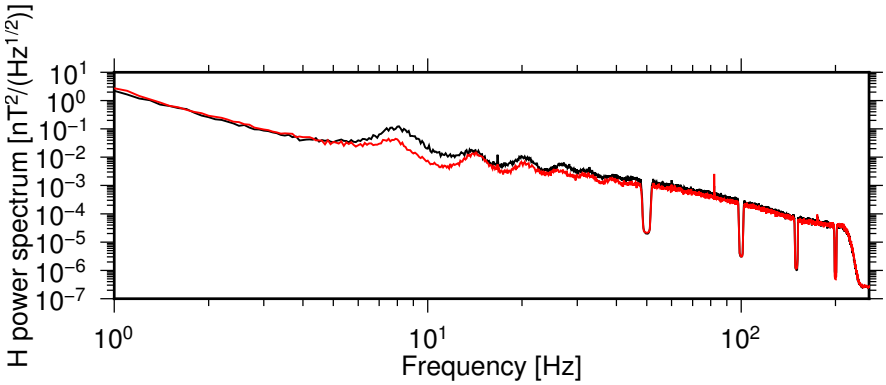


FIGURE 2.10: Résonances de Schumann - Spectre du champ magnétique au Welschbruch. Les 6 premières résonances sont clairement visibles.
FIGURE 2.10: *Schumann resonances on magnetic field power spectrum at Welschbruch geophysical station. The first six harmonics are clearly visible.*

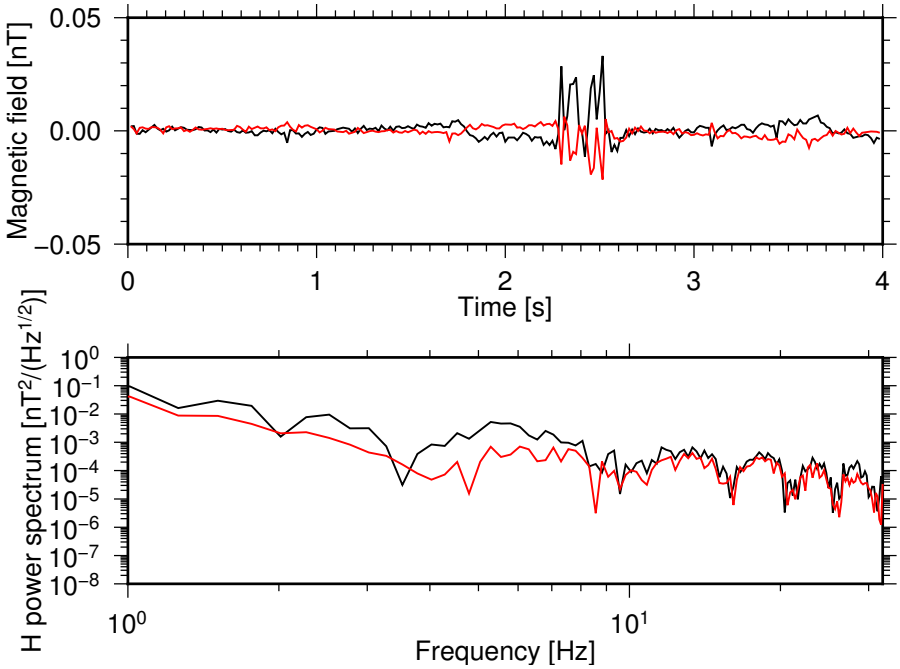


FIGURE 2.11: Haut : Onde *Q-Burst* en domaine temporel pour un échantillonnage à 64 Hz. Bas : Spectre en puissance des séries temporelles présentées au dessus.
FIGURE 2.11: *Top : Q-Burst in time domain for a sampling frequency of 65,536 Hz. Bottom : Power spectrum of above time series.*

2.2.3 Intérêt en MT

Pour la caractérisation des sondages MT, ces sources permettent d'obtenir les fonctions de transfert pour des fréquences entre 1 Hz et 10 kHz (Strangway, Swift Jr et Holmer, 1973). La principale caractéristique des sources géomagnétiques dans cette bande de fréquence est la présence de deux types de sources simultanées. En effet, le signal obtenu dans cette bande de fréquence est la somme d'un bruit atmosphérique (Reuveni et al., 2010), lui même étant la superposition de toutes les ondes émises par des éclairs dans le monde, et d'ondes transitoires avec un très fort ratio signal sur bruit et avec une polarisation linéaire. La variabilité de l'activité orageuse est multiple (Chrissan et A. Fraser-Smith, 1996). X Garcia et A. G. Jones (2002) ont montré une variation de l'énergie sur le champ magnétique entre 10 Hz et 32 kHz différente entre les hémisphères Nord et Sud. L'énergie magnétique étant plus importante en été dans l'hémisphère nord et plus importante en hiver dans l'hémisphère sud. Les mêmes auteurs ont montré de façon similaire une variabilité diurne du champ magnétique dans l'hémisphère nord. Le champ magnétique a une énergie supérieure de plusieurs ordres de grandeur la nuit aux alentours de minuit par rapport à l'amplitude en journée. Cette variabilité occasionne des problèmes dans la détermination des fonctions de transfert dans la bande de fréquences Audio-MT.
En effet, dans certaines zones industrialisées, le bruit anthropique peut être supérieur au bruit atmosphérique utilisé pour la caractérisation des fonctions de transfert MT. Le bruit anthropique ne respectant pas les hypothèses MT de source quasi-uniforme, il est difficile d'obtenir des fonctions de transfert satisfaisantes.

2.2.4 Article GJI - Détection et caractérisation des sources hautes fréquences.

Résumé en français Les ondes atmosphériques créées par l'activité orageuse mondiale apportent des informations a propos des courants électriques dans et à l ?extérieur de la Terre. Les événements avec un fort rapport signal-sur-bruit permettent de déterminer les propriétés des milieux traversés le long de leur propagation et sont ainsi particulièrement intéressants.

Nous introduisons dans ce papier une nouvelle méthodologie pour la détection et la caractérisation automatique des ondes atmosphériques en utilisant la décomposition temps-fréquence obtenue par application de la transformée en ondelette continue.

Nous nous intéressons particulièrement à trois types d'ondes, les atmosphérics, les slow tails et les whistlers, couvrant la totalité de la bande de fréquence 10 Hz - 10 kHz. Chaque type d'onde possède une signature caractéristique dans le plan temps-fréquence. Notre méthodologie permet la recherche automatique et de chaque type d'onde grâce à leur signature dans le plan temps-fréquence. La transformée en ondelette continue nous permet aussi de déterminer les attributs de polarisation horizontaux.

Cette procédure est tout d'abord appliquée sur des séries synthétiques de slow tails avec différents ratio signal-sur-bruit. Nous l'appliquons ensuite sur les séries temporelles de deux séries temporelles magnétiques acquises dans le cadre d'une campagne d'acquisition Audio-MT en Guadeloupe, DOM-TOM. Les atmospherics et les slow tails analysées présentent une polarisation linéaire, alors que les whistlers analysées présentent une polarisation elliptique.

La diversité de l'activité atmosphériques est enfin mise en relation avec le traitement de données Audio-MT, utilisé dans la caractérisation de la sub-surface. Nous montrons que les résultats des procédures de traitement Audio-MT dépendent principalement du contenu fréquentiel des ondes transitoires observées. Notre nouvelle méthodologie permet ainsi de donner des critères reliant la qualité des estimations MT à la diversité des ondes atmosphériques observées.

submitted to *Geophys. J. Int.*

Detection and characterisation of lightning-based sources using continuous wavelet transform: application to audio-magnetotellurics

H. Larnier[1], P. Sailhac[1,*], A. Chambodut[1]

[1] *Institut de Physique du Globe de Strasbourg, UMR7516, Université de Strasbourg/EOST, CNRS; Strasbourg, France*

[*] *now at Laboratoire GEOPS, UMR 8148, Université Paris Sud, Orsay, France*

SUMMARY

Atmospheric electromagnetic waves created by global lightning activity contain information about electric processes of the inner and the outer Earth. Large signal-to-noise ratio events are particularly interesting because they convey information about electromagnetic properties along their path. We introduce a new methodology to automatically detect and characterize lightning-based waves using a time-frequency decomposition obtained through the application of continuous wavelet transform. We focus specifically on three types of sources that cover the whole frequency range 10 Hz - 10 kHz, namely atmospherics, slow tails, and whistlers. Each wave has specific behavior in the time-frequency plane due to source shape and dispersion processes. Our methodology allows to automatically seek each type of event in the time-frequency plane thanks to their specific signature. The polarization attributes are also recovered in the time-frequency domain. This procedure is first applied on synthetic extremely-low-frequency synthetic time series with different signal-to-noise ratios. We then apply it on real data: two stations of audio- magnetotelluric data acquired in Guadeloupe, oversea French territories. Most of analysed atmospherics and slow tails can be considered as linearly polarized, whereas analysed whistlers are elliptically polarized. The diversity of lightning activity is finally analysed in an audio-

magnetotelluric data processing framework, as used in subsurface prospecting, through estimation of the impedance response functions. We show that audio-magnetotelluric processing results depends mainly on the frequency content of electromagnetic waves observed in processed time series.

Our new methodology based on the time-frequency signature of lightning-based electromagnetic waves allows to automatically detect and characterize electromagnetic events in audio-magnetotelluric time series, providing the means to assess quality of response functions obtained through processing.

Key words: Atmospherics, polarization attributes, Continuous Wavelet Transform.

1 INTRODUCTION

Natural geomagnetic activity in the frequency range 1 Hz to 10 kHz mainly originates from global lightning activity with about 100 strikes per second (Brook, 1925). Each single lightning strike gives rise to various phenomena. Among those phenomena, electromagnetic (EM) waves propagate through various paths. Most EM waves propagate in the waveguide formed by two conductive layers: the Earth's surface and the ionosphere (Laby *et al.*, 1940). Others, as whistlers, follow the magnetic field lines from one hemisphere to the other through the magnetosphere. Analysis of such waves' properties supplies information about electric and atmospheric processes of the Earth. Indeed, atmospherics characterization provides information about the ionospheric conductivity (Hughes & Pappert, 1975) and the distance to the lightning source (Heydt, 1982). Slow tails properties provide information about the current moment of the lightning strike (Cummer & Inan, 2000) and the distance from the source (Mackay & Fraser-Smith, 2010). Measurements of the Schumann resonances (Schumann & Koenig, 1954) provide insights about the height of the ionosphere, its electron density (Chand *et al.*, 2009), or the global ground temperature change (Williams, 1992).

EM waves emitted from a lightning strike act as induction sources by creating electric currents in the ground. If the source is far enough, the emitted waves can be considered as quasi-uniform over the area of measurements. The framework of audio-magnetotellurics (AMT) uses this assumption to give information about the electric structure of the subsurface (Strangway *et al.*, 1973). Processing AMT data requires state-of-the-art numerical codes assuming that EM time series are quasi-stationary to apply windowed fourier transform. Only a few studies actually detail the magnetic events in the processing of MT time series. Among those studies, a detailed analysis of the seasonal variation of atmospherics amplitude was conducted by Garcia & Jones (2002). They demonstrated that the optimal

season to acquire AMT data is summer in the northern hemisphere at mid-latitude. In the framework of AMT processing, Zhang & Paulson (1997) proposed a methodology using a moving threshold in time-frequency domain to detect atmospheric waves before processing. We previously extended Zhang & Paulson method to consider the choice of the wavelet depending on the frequency domain to be processed (Larnier *et al.*, 2016).

In this paper, we propose a new methodology: Automatic Detection of ElectroMagnetic waves (ADEM), to automatically detect EM waves in the frequency range 10 Hz – 10 kHz. ADEM is based on the time-frequency properties of sought waves using continuous wavelet transform (CWT). We first show in Section 2, AMT data acquired in Novembre 2012 in Guadeloupe, French Overseas Territories. They exhibit interesting features of typical EM waves. The ADEM method, including details on the computation of wavelet coefficients and polarization attributes, is given in Section 3. In section 4, we show the application of the ADEM to real data and discuss about polarization attributes and frequency content of atmospheric waves with consequences on the acquisition and processing of AMT data.

2 LIGHTNING-BASED WAVES

2.1 Sites locations and data

EM time series acquired during a magnetotelluric campaign on La Soufriere lava dome (Guadeloupe, France) in 2012 (Sailhac *et al.*, 2016) are analysed. All MT sites use ADU07 data logger as recording system from Metronix Geophysics. Let us now consider two stations (Figure 1).

The first one, named Guenon, is located on the lava dome. Magnetic field was recorded using MFS07e induction coils (using chopper off), and uses 41 and 31 meters electric dipoles for NS direction and EW direction respectively for electric measurements.

The second station, named Youketi, is used as a remote station for AMT processing. It is located 8 km off the lava dome and 600 meters off the Soufriere Volcano Observatory. This station was composed by 50 meters electric dipole for NS and EW measurements and MFS06 induction coils (using chopper off) to measure the magnetic field.

The sample time series analysed in this paper is one minute duration acquired at 16384 Hz sample frequency. This corresponds to 983 040 data samples for each channel (horizontal magnetic and electric fields).

In the one minute sample, many types of EM waves in the frequency band of 1 Hz up to 10 kHz coexist. In this paper, we focus on three types of EM waves, atmospherics, slow tails and whistlers. Sample time series from this campaign actually exhibit significative signals of these three types to be illustrated and tested in the ADEM.

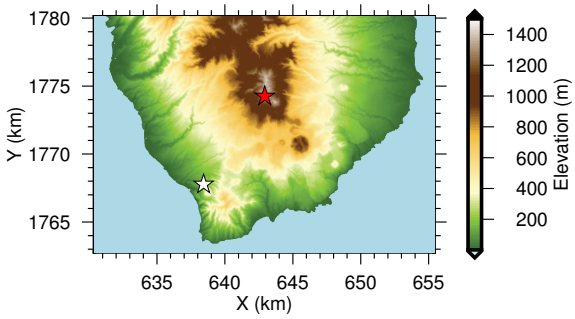


Figure 1. Location of magnetotelluric stations used in Guadeloupe, French Overseas Departments and Territories. UTM 20N projection. Red star: Guenon AMT station. White star: Youketi AMT station.

2.2 Atmospherics

Atmospherics appear on EM time series as damped oscillations (Figure 2a). Their shapes vary, depending on ionospheric conditions (e.g. night or day) and distance from the lightning strike (Rakov & Uman, 2003). According to the waveguide mode theory (Wait, 1962; Cummer, 2000), atmospherics undergo dispersion from the source to the measurement point where highest frequencies travel faster than lowest ones. Atmospherics have frequency content ranging from 2 kHz up to more than 30 kHz (Figure 2b). On Figure 2, we present the CWT analysis of a short portion of Guenon time series. The largest atmospherics at 5.727s exhibits large frequency content, from 3-4 kHz to more than 10 kHz (Figure 2c). In addition, other atmospherics appear later (around 5.730s). They also have large frequency content but with lower amplitude and with lower signal-to-noise ratio. An interesting feature is the dispersion effect on wavelet coefficients. Indeed, large coefficients at high frequencies around 10 kHz appear to be located a few milliseconds before large coefficients at lower frequencies around 4 kHz.

Polarization attributes (Figure 2d, 2e and 2f) are constant over the area of high signal-to-noise ratio. Atmospherics in this typical time sequence display quasi-null ellipticity, a phase difference close to phase opposition and a polarization angle close to 60 degrees. If such polarization attributes are observed with accuracy, we will consider atmospherics to have linear polarization.

2.3 Slow tails

Slow tails are emitted by a lightning strike and arrive slightly later than the atmospherics due to lower speed propagation. These waves only propagate in quasi-Transverse ElectroMagnetic (TEM) field mode. Characteristics of both atmospherics and slow tails waves have been used in several studies to determine the distance from the source lightning (Wait, 1960; Mackay & Fraser-Smith, 2010) or

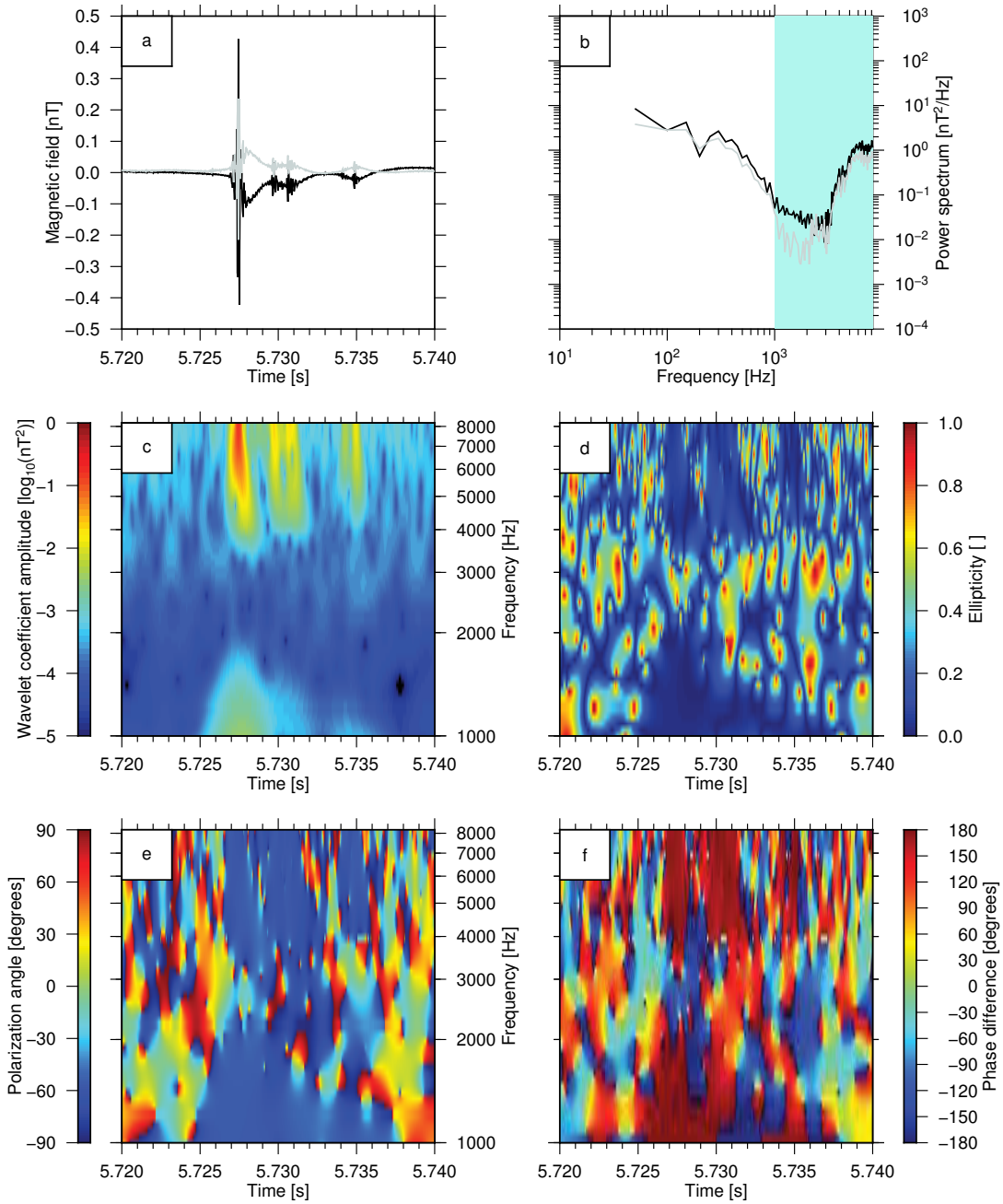


Figure 2. Atmospherics observed at Guenon station, on the 16th of November 2012. Time series start at 19h 45min 16s TU. a) Time series, black line: North-South magnetic field H_x, grey line: East-West magnetic field H_y. b) Power spectrum, black line: North-South magnetic field, grey line: East-West magnetic field. The blue area indicates the frequency domain covered by the wavelet analysis. c) Spectrogram of wavelet coefficients. d) Ellipticity coefficients. e) polarization angle coefficients. f) Phase difference coefficients.

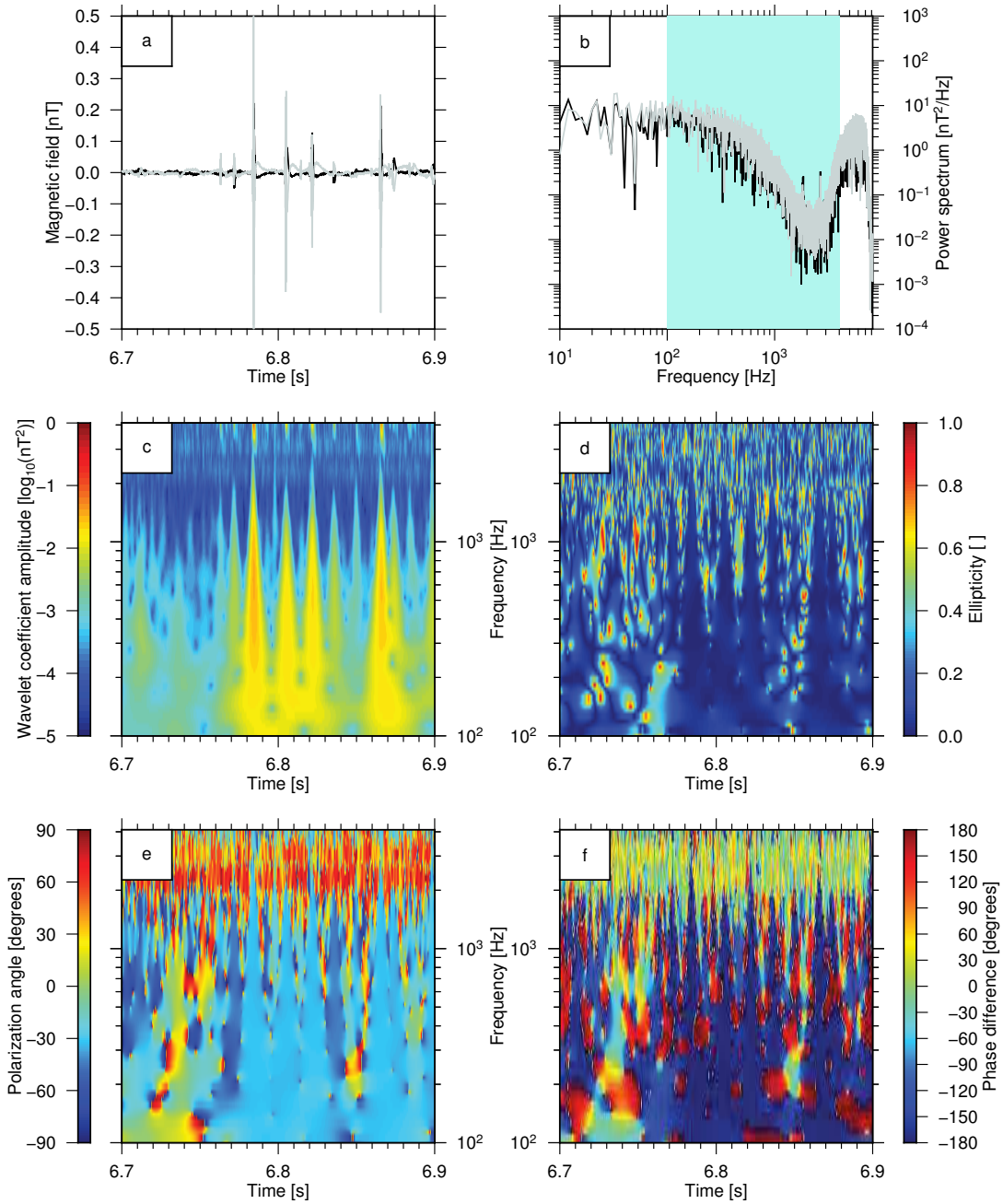


Figure 3. Slow tails observed at Guenon station, on the 16th of November 2012. Time series start at 19h 45min 16s TU. a) Time series, black line: North-South magnetic field H_x, grey line: East-West magnetic field H_y. b) Power spectrum, black line: North-South magnetic field, grey line: East-West magnetic field. The blue area indicates the frequency domain covered by the wavelet analysis. c) Spectrogram of wavelet coefficients. d) Ellipticity coefficients. e) Polarization angle coefficients. f) Phase difference coefficients.

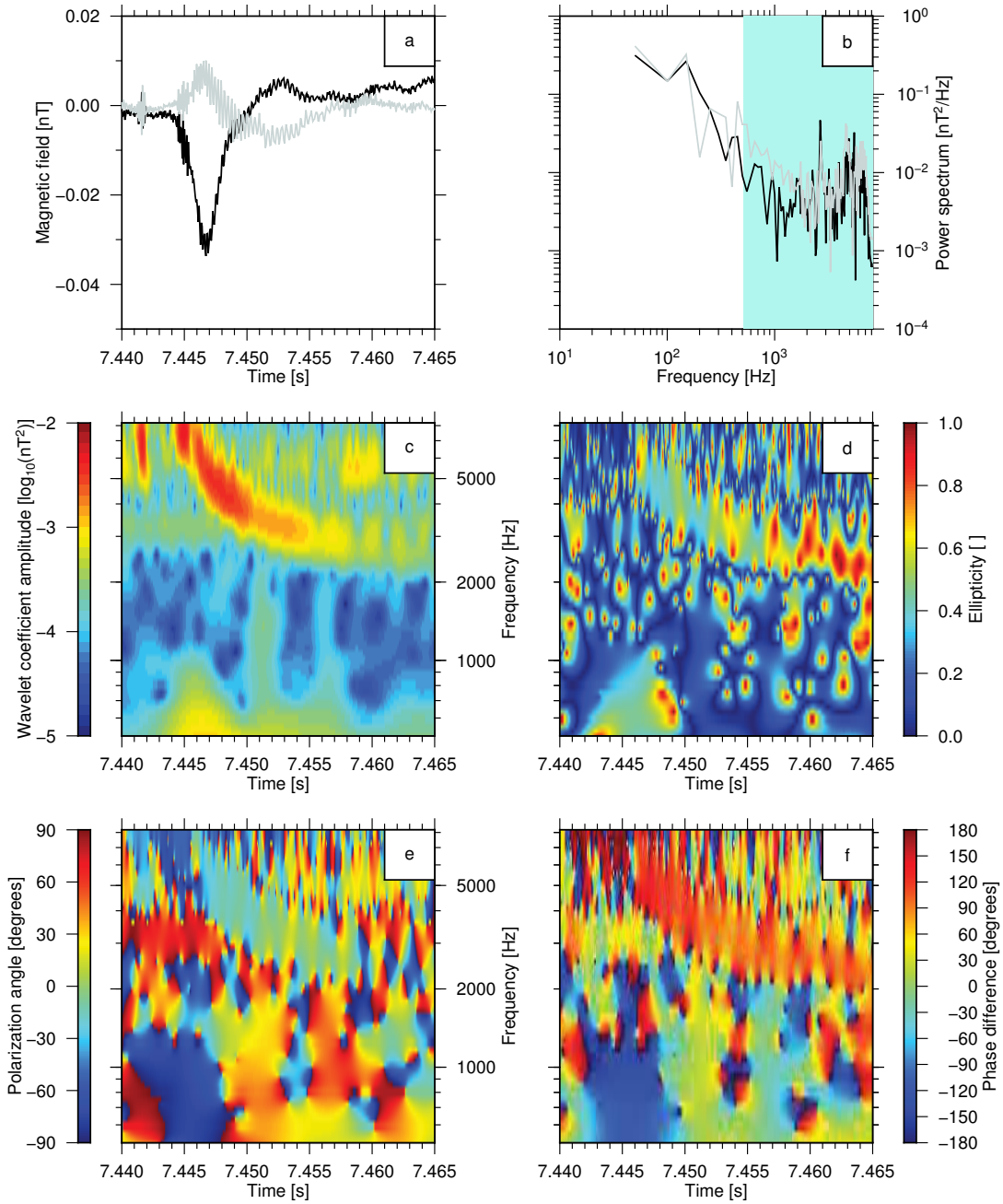


Figure 4. Whistlers observed at Guenon station, on the 16th of November 2012. Time series start at 19h 45min 16s TU. a) Time series, black line: North-South magnetic field H_x, grey line: East-West magnetic field H_y. b) Power spectrum, black line: North-South magnetic field, grey line: East-West magnetic field. The blue area indicates the frequency domain covered by the wavelet analysis. c) Spectrogram of wavelet coefficients. d) Ellipticity coefficients. e) polarization angle coefficients. f) Phase difference coefficients.

its current moment (Cummer & Inan, 2000). Their frequency content starts from less than 30 Hz up to 3 kHz (Figure 3a, 3b). Figure 3c show wavelet analysis of a short portion of Guenon time series showing several large slow tails. Like atmospherics, slow tails have large frequency content, from 1 kHz down to 50 Hz. Variability in amplitude of wavelet coefficients also affects the frequency content. By comparison, the slow tail located at 6.835 seconds has narrower frequency content, from less than 2 kHz down to 300 Hz, whereas the slow tail at 6.785 seconds has high amplitude wavelet coefficients from 2.5 kHz down to 100 Hz.

Polarizations attributes (Figure 3d, 3e, 3f) are constant over the region of large amplitude coefficients indicating constant polarizations attributes. The largest slow tail has ellipticity close to 0, a phase difference of ± 180 degrees and a polarization of about -30 degrees. Like atmospherics, we consider these attributes as signature of linear polarization.

2.4 Whistlers

Whistlers are also EM waves emitted from a lightning strike, but they do not propagate in the waveguide. They penetrate through the ionosphere and follow magnetic field lines to the other hemisphere where they can be measured on ground surface (Rakov & Uman, 2003). ”Ducted” whistlers are seen on ground surface, while ”unducted” are trapped in the magnetosphere and invisible in ground measurements. These waves are dispersed during their propagation through ionosphere and magnetosphere. Thus high frequencies arrive significantly sooner than low frequencies. A whistler may contain frequencies from 100 Hz up to 10 kHz (Figure 4a, 4b).

On Figure 4c, we present the wavelet analysis of a whistler observed in Guenon data. The most distinctive feature in the wavelet spectrogram is the dispersion effect. This dispersion affects the wavelet coefficient such that high amplitude coefficients at low frequencies (around 3 kHz) are located 10 ms after the high frequencies.

The polarizations attributes (Figure 4d, 4e, 4f) indicate elliptical polarization with ellipticity around 0.5, phase difference around 120 degrees positive, and polarization angle of -40 degrees.

3 METHODS

3.1 Wavelet analysis

We recall that the continuous wavelet transform (CWT) is a mathematical tool that uses a special class of functions called wavelets to decompose a signal s into time-frequency representation. The basis functions are derived from the mother wavelet that is dilated by a dimensionless factor a (corresponding to the inverse of frequency) and translated by a factor b (corresponding to time). The signal s is

then decomposed into coefficients by:

$$\mathcal{W}_\psi(a,b)[s] = \int_{-\infty}^{+\infty} dt \frac{1}{a} \psi^*_{a,b}(t) s(t), \tag{1}$$

where * represents the complex conjugate and $\psi_{a,b}(t)$ is the dilated and translated function given by equation (2).

$$\psi_{a,b}(t) = \psi\left(\frac{t-b}{a}\right). \tag{2}$$

For discrete time series, the function $\mathcal{W}_\psi(a,b)[s]$ is sampled on discrete values for a and b, constituting a matrix of wavelet coefficients.

First introduced in seismic (Goupillaud *et al.*, 1984), the CWT has already been widely used in acoustics and climate studies to analyse non stationary processes and proven to be an efficient tool to study magnetotelluric data (Zhang & Paulson, 1997; Garcia & Jones, 2008; Escalas *et al.*, 2013; Larnier *et al.*, 2016).

A large number of mother wavelets is available (Torrence & Compo, 1998; Holschneider, 1995). As Zhang & Paulson (1997), we use the Morlet wavelet for its properties of good accuracy both in time and frequency. It is defined in time domain by:

$$\psi_m(\eta) = \pi^{-1/4} e^{i\omega_0 \eta} e^{-\eta^2/2}, \tag{3}$$

and in frequency domain by:

$$\widehat{\psi}_m(a\omega) = \pi^{-1/4} H(\omega) e^{-(a\omega-\omega_0)^2/2}, \tag{4}$$

where H is the Heaviside function and ω_0 the wavelet parameter. ω_0 can be modified to adjust the time-frequency resolution of the analysing wavelet.

3.2 Detection of EM waves

Lightning EM waves display different time-frequency patterns according to their nature (Figures 2, 3, 4). Such waves display different time-frequency patterns according to their nature. Some have wavelet coefficients with large amplitude on a large frequency band whereas some display narrow frequency content. The detection of these waves in the time-frequency domain may be based on the characteristics shown by the CWT. We thus set two criteria:

- A minimum amplitude threshold in the time-frequency domain (signal-to-noise ratio).
- A specific pattern in time-frequency (e.g. dispersion).

3.2.1 *First criterion of significant amplitude*

The signals we seek are above the signal-to-noise ratio. In order to get all wavelet coefficients corresponding to this criterion, we are looking for a threshold at each scale. We follow the method of significant scales described in Torrence & Compo (1998) in a way similar to that used in Larnier *et al.* (2016).

We want to detect time-frequency patterns by comparing their amplitude to the background noise. Background noise power spectrum is estimated by the method of Torrence & Compo (1998). They have used Monte Carlo methods and considered stationary time series s of length N described by order one Auto-Regressive (AR-1) process or Gaussian white noise. They demonstrated that the square modulus wavelet distribution: $D(a_k) = \{|\mathcal{W}_\psi(a_k, b_i)[s]|^2,\ i = 1...N\}$ at each scale a_k, is related to the χ_2 distribution. $D(a_k)$ is distributed as $P_k\sigma^2\chi_2^2/2$ (with same variance and confidence intervals), where σ is the noise standard deviation and P_k the mean fourier background spectrum of signal s at frequency f_k corresponding to the scale a_k (such that $a_k = (\omega_0 + \sqrt{2+\omega_0^2})/4\pi f_k$). The subscript 2 on the χ_2^2 distribution corresponds to the number of degrees of freedom (2 for a complex mother wavelet, 1 for a real mother wavelet). By computing $\chi_2^2(\alpha)$, the variance σ and P_k can be used to define a threshold value $\eta(a_k)$ at each scale a_k with confidence level $(1-\alpha)$:

$$\eta(a_k, \alpha) = \frac{1}{2} P_k \sigma^2 \chi_2^2(\alpha). \tag{5}$$

Unfortunately, knowledge of the electromagnetic background spectrum is not trivial since background EM noise does not follow any classical shape such as white or red noise processes. For example, it presents deviations such as the MT and AMT dead bands (Chave & Jones, 2012), which have to be taken into consideration. We follow the suggestion of Torrence & Compo (1998), and use the global wavelet spectrum of the magnetic field as background spectrum. The global wavelet spectrum is defined as:

$$G_\psi^2(a)[s] = \int_{-\infty}^{+\infty} db |\mathcal{W}_\psi(a,b)[s]|^2. \tag{6}$$

Figure 5 presents an example of the comparison between the global wavelet spectrum and the fourier power spectrum H_k. The global wavelet spectrum appears as a smoothed fourier spectrum reducing the effects of abrupt variations in the fourier spectrum. Using the global wavelet spectrum as background spectrum P_k, the wavelet modulus distribution at each scale a_k is constrained by the following threshold at confidence level $1-\alpha$:

$$\eta(a_k, \alpha) = \frac{1}{2} G_\psi^2(a_k)[s] \chi_2^2(\alpha). \tag{7}$$

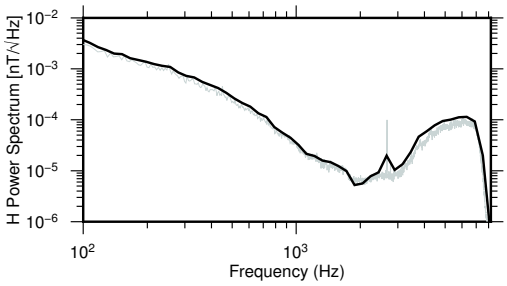


Figure 5. Comparison between fourier spectrum H_k (thin grey line) and global wavelet spectrum (thick black line) of North-South magnetic field for an extract of 262144 samples.

3.2.2 Second criterion, extent in time-frequency (dispersion pattern)

Another criterion we use to discriminate EM waves from noise, is supported by their frequency band and their dispersion behavior. Indeed, as illustrated on Figures 2, 3 and 4, different patterns in time-frequency domains are clearly identified. We use these characteristic patterns as a second criterion which also yields automatic segmentation. All sought EM waves have large frequency content. This is characterized by a large number of high wavelet coefficients along the scale axis (a values) within a very short time (b interval). Atmospherics and whistlers are affected by dispersion. The maximum in correlation with a large-scale wavelet (lower frequency) is then located later than those with a short-scale wavelet (higher frequency). Dispersion affects this series of high wavelet coefficients by a shift along the time axis (Δb of a few milliseconds as seen on Figure 2 and 4). Two main behavior are thus expected and are used to discriminate EM waves from noise:

- Large frequency content and no dispersion for slow tails, with vertical maxima chains.
- Large frequency content and dispersion effect for atmospherics and whistlers, with curved maxima chains.

From the wavelet coefficients matrix, we compute and chain local modulus maxima along the scale axis (see Appendix B and C). The chain rule is wave-dependent, but the general rule is to chain two local maxima along a dispersion curve controlled by a dispersion coefficient D. Flexibility is allowed and controlled by the correlation kernel of the analysing wavelet (Farge, 1992; Maraun *et al.*, 2007). The exact chain rule implemented in the algorithm is described in Appendix C.

3.2.3 Combinaison of criteria

Each criterion alone should not be used as an effective tool to discriminate EM waves from noise. As discussed by Torrence & Compo (1998) and detailed by Maraun *et al.* (2007), it can be shown that even pure white or red noise processes create local maxima of high magnitude in the time-frequency

domain that will be kept after thresholding. Vertical maxima chains must not be used alone either. If a wavelet with narrow time support such as Cauchy wavelet (Holschneider, 1995) is used, the coefficients' behavior will create many vertical maxima chains in the matrix of wavelet coefficients. The simultaneous use of both criteria helps to compensate the flaw of each independent criterion.

Due to the large diversity of EM waves intensity and polarizations, we decide to look for specific maxima chains with fixed extent in the time-frequency domain. For example, we may decide to look for atmospherics having significant content in the frequency range 4 kHz - 10 kHz, or, be more ambitious and only look for atmospherics having significant content in the frequency range 2 kHz - 16 kHz. The number of detected events will be highly dependent on this choice. Detecting more events is achieved by lowering the confidence level α or by reducing the expected frequency range in the maxima chains.

3.3 Polarization attributes

Monochromatic plane waves are characterized by elliptical polarization in the horizontal plane. The polarization represents the space evolution of the electric and magnetic components with time. This evolution describes an ellipse in the orthogonal plane which ellipse can be characterized by three attributes, e the ellipticity, θ the polarization angle and $\Delta\phi$ the phase difference between orthogonal components.

The ellipticity e represents the ratio of the minor r and the major R axis with: $e = (r/R) \in [0, 1]$. The polarization angle θ is the angle between the horizontal axis and the major axis with $\theta \in [-\pi/2, \pi/2]$. The phase difference is computed between the horizontal components, with $\Delta\phi \in [-\pi, \pi]$. Many methods exist to recover the polarization attributes by retrieving ellipses parameters on hodograms or in Fourier domain (Fowler *et al.*, 1967). We choose the wavelet based determination of polarization attributes described by Diallo *et al*, (2006). Let the signal c be defined by:

$$c(t) = h_y(t) + ih_x(t), \tag{8}$$

where h_x and h_y are the north-south and east-west component of the magnetic field respectively. The time-frequency polarization attributes are given by:

$$e(a,b) = \frac{\left||\mathcal{W}_\psi^+(a,b)[c]| - |\mathcal{W}_\psi^-(a,b)[c]|\right|}{|\mathcal{W}_\psi^+(a,b)[c]| - |\mathcal{W}_\psi^-(a,b)[c]|} \tag{9}$$

$$\theta(a,b) = \frac{1}{2}\arg(\mathcal{W}_\psi^+(a,b)[c] \otimes \mathcal{W}_\psi^-(a,b)[c]) \tag{10}$$

$$\Delta\phi(a,b) = \arg\left(\frac{\mathcal{W}_\psi^+(a,b)[c] + (\mathcal{W}_\psi^-(a,b)[c])^*}{\mathcal{W}_\psi^+(a,b)[c] - (\mathcal{W}_\psi^-(a,b)[c])^*}\right) + \frac{\pi}{2}. \tag{11}$$

where $\mathcal{W}_{\psi}^{+}(a,b)[c]$ and $\mathcal{W}_{\psi}^{-}(a,b)[c]$ are the coefficients obtained using a progressive and regressive wavelet respectively (Holschneider, 1995). Diallo *et al*, (2006) underline that these instantaneous coefficients must not be interpreted alone and have to be studied on a time-frequency window to make sure they are meaningful. This method has been previously used to describe cultural noise sources in Controlled-Source AMT experiments by Escalas *et al.* (2013) or to determine the polarisation attributes of geomagnetic pulsations by Kulesh *et al* (2007).

To recover the polarization attributes of each event, we use the maxima chain kept at the detection stage. For maxima chain i, at each scale a_j, let $b_{i,j}$ be the time position of the local maxima in the maxima chain. We define a distance d_j based on the correlation kernel of the mother wavelet (Appendix A). We then store wavelet polarization coefficients according to:

$$\begin{aligned} D_{e_{i,j}} &= \{e(a_j,b),\ |b-b_{i,j}|<d_j\} \\ D_{\theta_{i,j}} &= \{\theta(a_j,b),\ |b-b_{i,j}|<d_j\} \\ D_{\Delta\phi_{i,j}} &= \{\Delta\phi(a_j,b),\ |b-b_{i,j}|<d_j\}\,. \end{aligned} \tag{12}$$

For each event i, the polarization attributes e_i, θ_i and $\Delta\phi_i$ are determined by taking the median value of the distributions $D_{e_{i,j}}$, $D_{\theta_{i,j}}$, $D_{\Delta\phi_{i,j}}$ respectively for all scales. Note that careful attention is paid to the determination of circular quantities such as θ and $\Delta\phi$ using the methodology of directional statistics (Fisher, 1995).

3.4 Test on synthetic data

We first apply the ADEM algorithm on synthetic data created by following Wait, (1960), as described in Appendix C. Three types of slow tails time series are synthesized with different noise level (Figure 6). The lowest noise level is typical of most AMT data while the moderate noise level corresponds to noisy AMT data. The highest noise level is $\sigma = 0.02$ and is used to test the robustness of the method but is not representative of most AMT data. Higher values for σ result in time series where the noise level is too high and unrealistic fo field AMT data. This test was carried out to assess the efficiency of the code in the detection and characterization of non-monochromatic waves. The synthetic time series are 40 seconds long. For each event we attribute a random time index and random polarization attributes. Results are described in Table 1.

Let us now discuss about results without remote site first. For low noise level, only one true event is not detected and no false event is detected by the code. The characterization of polarization attributes is efficient with almost no difference between recovered attributes and true attributes. For moderate noise level, only one true event is not detected and no false event is detected. Error slightly increases but remains low. Large noise level leads to the recovery of only 70 % of true events. Even worse for

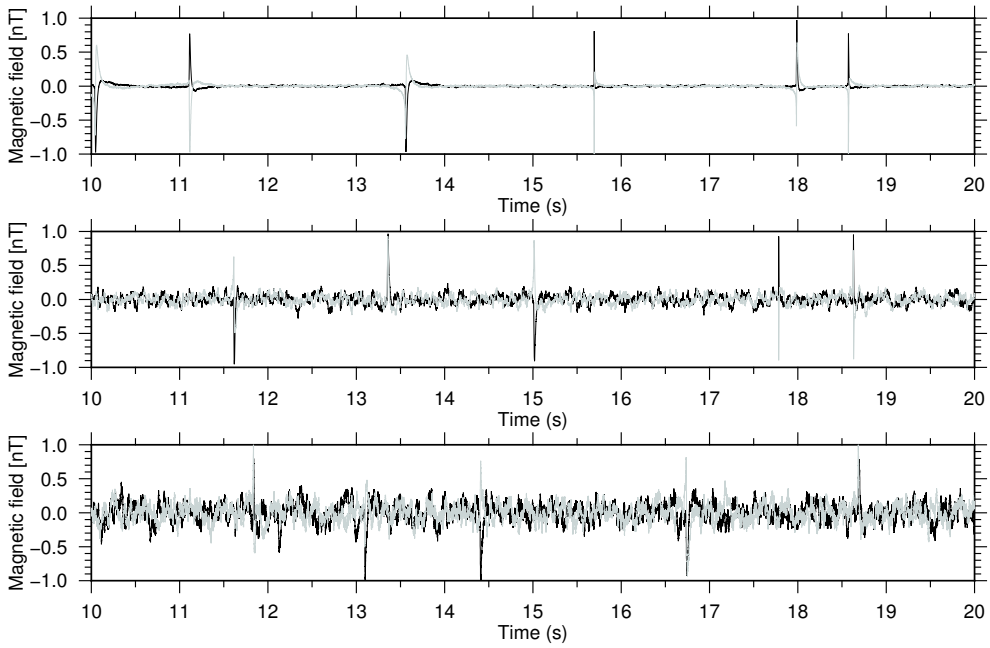


Figure 6. Synthetic time series. Top: Low level of noise. Middle: Medium noise level. Bottom: High noise level. Black: North-South magnetic field. Grey: East-West magnetic field.

automatic detection, false events are detected. In spite of the trouble with false events, polarization attributes are correct for the true events recovered by the process.

When a remote time series is used in the detection, it removes every false detection made by the ADEM algorithm. The remote station is therefore critical in the detection of atmospherics to prevent any false detection from the time series. The use of a remote station is standard process in AMT acquisition so does not act as a burden in the field acquisition scheme.

4 APPLICATION

4.1 Atmospherics

4.1.1 Characterisation

We first apply the ADEM method to the real datasets of Guenon and Youketi time series. The detection was set using a 90 % confidence interval. Sought maxima chains were expected to have significant content from 8 kHz down to 3 kHz. The correlation kernel was obtained using a critical value of 0.90. On this basis, 172 atmospherics are detected. To illustrate the detection process, we have represented on Figure 7 the two stages described previously on 0.2 seconds of the analysed time series. The distribution of polarization attributes is illustrated Figure 8. Most of detected waves have quasi-null ellipticity

Criteria	Low noise level ($\sigma = 0.001$)	Moderate noise level ($\sigma = 0.01$)	High noise level ($\sigma = 0.02$)
Number of detected true events (over a total 40).			
Single site	39	39	30
Remote reference	39	39	27
Number of detected false events			
Single site	0	0	3
Remote reference	0	0	0
$(e_m - e_d)$			
Single site	-0.002 $\pm$ 0.004	0.012 $\pm$ 0.045	0.046 $\pm$ 0.086
Remote reference	-0.002 $\pm$ 0.004	0.011 $\pm$ 0.045	0.050 $\pm$ 0.086
$(\theta_m - \theta_d)$ (degrees)			
Single site	0.022 $\pm$ 0.6	0.78 $\pm$ 12.9	-2.16 $\pm$ 9.8
Remote reference	0.066 $\pm$ 0.76	0.39 $\pm$ 12.7	-3.00 $\pm$ 10
$(\Delta\phi_m - \Delta\phi_d)$ (degrees)			
Single site	-0.187 $\pm$ 0.773	-0.659 $\pm$ 4.91	-2.32 $\pm$ 8.66
Remote reference	-0.210 $\pm$ 0.960	-0.962 $\pm$ 5.00	0.692 $\pm$ 8.45

Table 1. Results of synthetic tests. Subscript m indicates the parameters of the model. Subscript d indicates the results from the code.

and phase difference around 0 modulo π. This indicates linear polarization for all events. The polarization angle indicates a large distribution of wave polarization, with two principal directions of +60 and -40 degrees. Computation of the difference in polarization attributes for common events gives difference below 10 degrees in polarization angle, 15 degrees for phase difference and 0.1 for ellipticity (Figure 9).

4.1.2 Frequency content and occurence

The variability of frequency content in magnetic time series may be illustrated by the systematic use of our algorithm to detect atmospherics with variable frequency content. Several detections are performed: atmospherics with significant frequency content from 5 kHz to 8 kHz, from 4 kHz to 8 kHz, from 3 kHz to 8 kHz and finally from 2.5 kHz to 8 kHz. We apply the algorithm on 1 minute

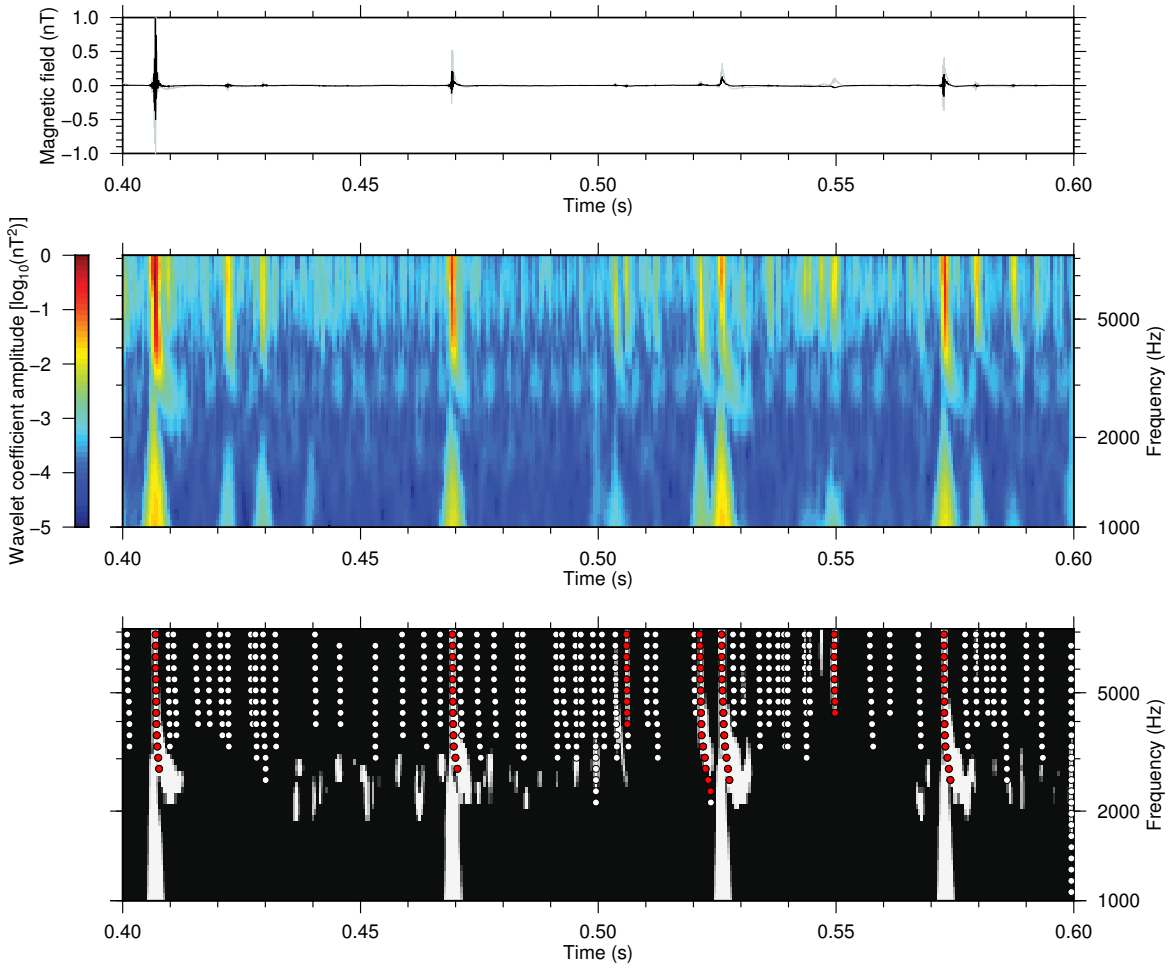


Figure 7. Illustration of detection processes on Guenon time series by CWT. Top: Orthogonal magnetic time series, black line: North-South magnetic field, grey line: East-West magnetic field. Middle: Sum of wavelet spectrograms of time series. Bottom: White: Significant coefficients kept after Torrence & Compo (1998) detection scheme. White and red chains: Maxima chains detected. Red chains: Maxima chains kept after application of both criteria.

time series at 16384 Hz. Only events detected at Guenon and Youketi simultaneously are kept. Time series of the flow of detected events (number of events detected by second) are illustrated Figure 10.

Every second, at least 15 atmospherics with frequency content from 5 kHz to 8 kHz are detected, with peaks to nearly 50 events, making an average of about 33 per second over the whole minute. Atmospherics from 4 kHz to 8 kHz occur less often, with an average of about 22 detections per second. The search for atmospherics with frequency content from 3 kHz to 8 kHz outputs even less

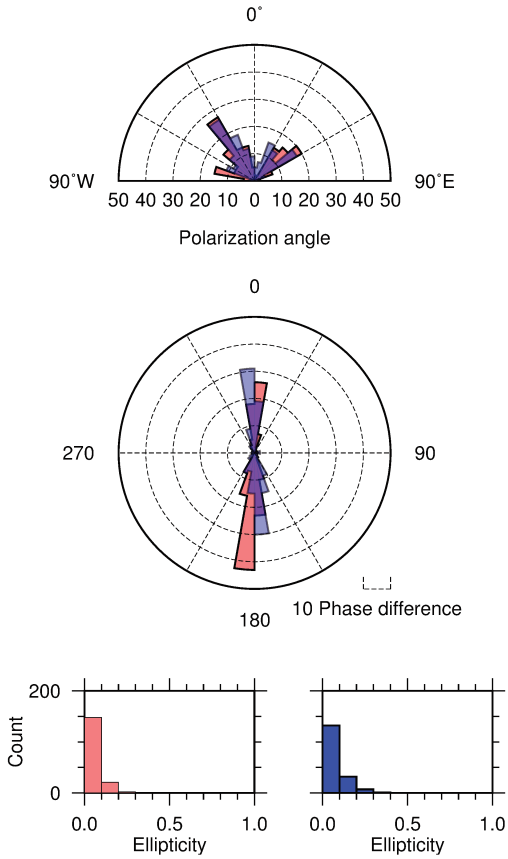


Figure 8. Distributions of atmospherics polarization attributes obtained independently at Guenon and Youketi AMT stations. Red: Guenon. Blue: Youketi

detected events: 3 on average per second, with periods of no occurrence at all. Finally, atmospherics with frequency content from 2.5 kHz to 8 kHz occur on average 0.2 times per second.

4.2 Slow tails

4.2.1 Characterisation

We repeat the methodology on Guenon and Youketi time series but this time to detect slow tails. The detection was set using 99 % confidence interval. Maxima chains with significant frequency content from 64 Hz up to 1 kHz are sought. The critical kernel was built using a critical value of 0.9. As a result, 287 events are detected (Figure 11). The principal direction of polarization is about -45 degrees west. As for atmospherics' case, the distribution of the ellipticity and phase difference indicates linear polarization for most of the slow tails with some departure that may be due to noise in the data. Common detected waves with Guenon and Youketi have similar features. Differences in polarization

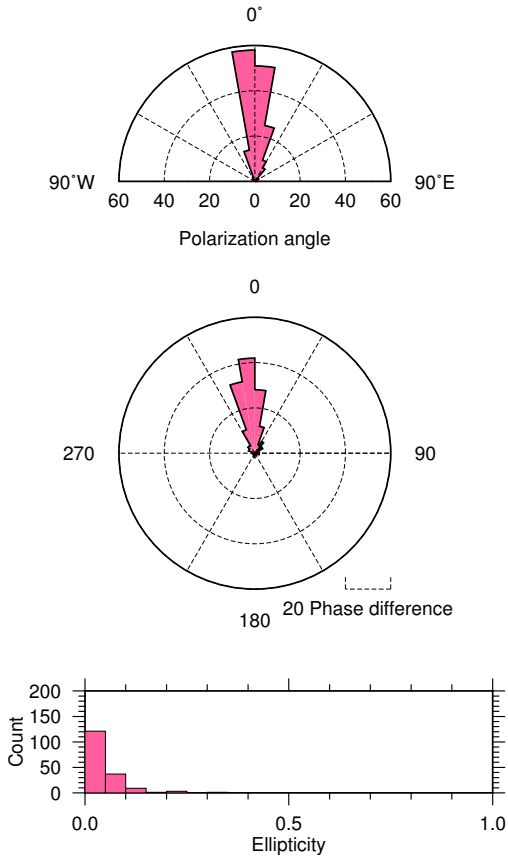


Figure 9. Difference of atmospherics polarisation attributes for common events obtained at Guenon and Youketi AMT stations.

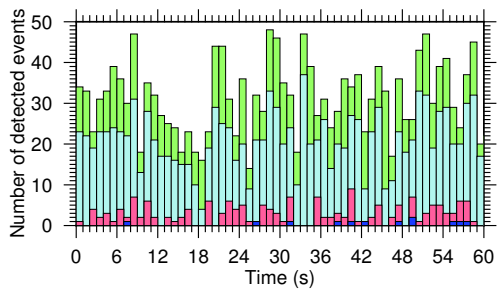


Figure 10. Flows of detected atmospherics at Guenon station depending on expected frequency content. Green: From 5 kHz to 8 kHz, Light blue: From 4 kHz to 8 kHz, Red: From 3 kHz to 8 kHz, Dark blue: From 2.5 kHz to 8 kHz.

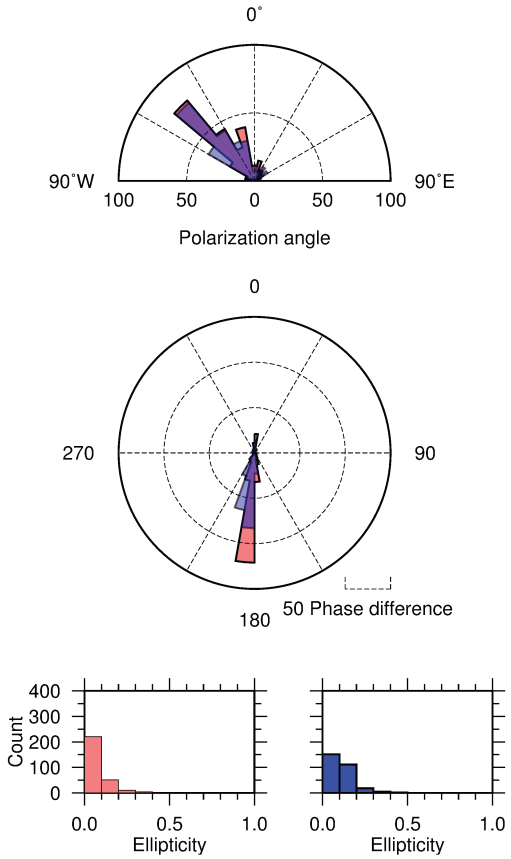


Figure 11. Distributions of slow tails polarization attributes obtained independently at Guenon and Youketi AMT stations. Red: Guenon. Blue: Youketi.

attributes for common events on both sites show values below 10 degrees for polarisation, 20 degrees for phase difference and 0.1 for ellipticity (Figure 12).

4.2.2 Frequency content and occurence

Similar to the process performed on atmospherics, we study the relationship between frequency content and occurrence. We use one minute of measurement at 16384 Hz. Detection of slow tails is performed with frequency content from 128 Hz to 500 Hz, from 128 Hz to 1 kHz, from 128 to 1.5 kHz, and finally from 128 Hz to 2.5 kHz. Time series for the flow of detected events (number of events detected by second) are illustrated Figure 13.

Concerning slow tails with significant frequency content from 128 Hz to 500 Hz, such events occur quite often, 7.1 events per second on average. Events with frequency content from 128 Hz to 1 kHz occur on average 5.15 per second and events with frequency content from 128 Hz to 1.5 kHz

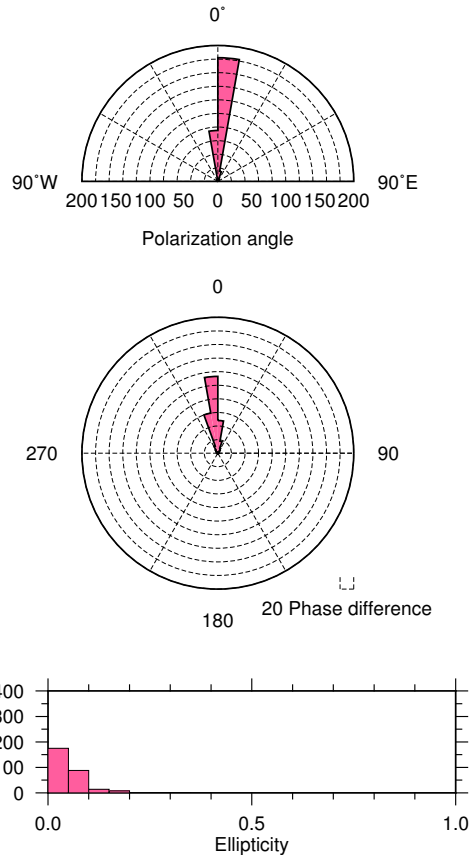


Figure 12. Difference of slow tails polarisation attributes for common events obtained at Guenon and Youketi AMT stations.

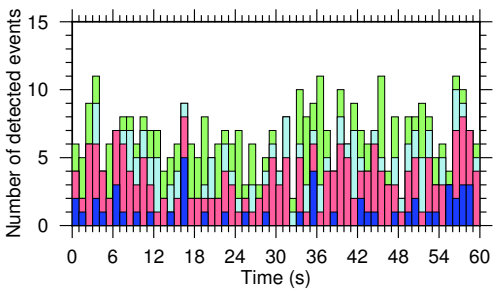


Figure 13. Flows of detected slow tails at Guenon station depending on expected frequency content. Green: From 128 Hz to 500 Hz, Light blue: From 128 Hz to 1 kHz, Red: From 128 Hz to 1.5 kHz, Dark blue: From 128 Hz to 2.5 kHz.

occur with an average of 3.7 per second. Finally, the largest events with frequency content from 128 Hz to 2.5 kHz do not occur often with an average of 0.83 per second, with periods of no occurrence.

4.3 Whistlers

Whistlers occur only a few times in our datasets. We are not able to provide extensive analysis over several tens of events such as in the case of atmospherics or slow tails. Nevertheless, the whistler shown Figure 4 is detected and analysed by the ADEM code. The event is characterised by an ellipticity of $e = 0.53$, a polarization angle of $\theta = 1.8$ in degrees and a phase difference $\Delta\phi = 73$ degrees.

4.4 AMT impedance tensor

Natural EM waves show large and low intensity events with large spectral content and event with narrow spectral content. Among many properties of EM waves, we are concerned with audio-magnetotellurics (AMT). We use the principle of electromagnetic induction to constrain electrical resistivity of the subsurface. Such EM waves induce electrical currents in the ground that can be measured using voltmeter and grounded electrodes. These currents are related to the inductive magnetic field through the so-called impedance tensor Z (Cagniard, 1953):

$$\begin{aligned} E_x(\omega) &= Z_{xx}(\omega)H_x(\omega) + Z_{xy}(\omega)H_y(\omega) \\ E_y(\omega) &= Z_{yx}(\omega)H_x(\omega) + Z_{yy}(\omega)H_y(\omega). \end{aligned} \tag{13}$$

One major advantage of EM waves such as slow tails and atmospherics is their wide frequency content. This wideness helps to constrain Z over a large frequency band at once. One critical concern in audio-magnetotellurics is the low signal-to-noise ratio in the so-called AMT dead band, from 1 kHz to 5 kHz. In this frequency band, standard processing procedures often fail to properly determine MT response functions Z (Garcia & Jones, 2002). We have previously shown that waves with significant frequency content in this AMT dead band still occur and should provide information to constrain the impedance tensor (section 4.1 and 4.2).

The ADEM algorithm is applied on Guenon EM data samples and is used to define periods of active and quiet lightning activity before processing using standard schemes. We use the algorithm to detect, slow tails presenting maxima chains from 128 Hz to 1024 Hz providing a catalog of significative geomagnetic events. Each segment of ”active” time series starts 0.1 second before the occurrence of a selected event. While the difference between two successive events is below 0.5 seconds, the segment is still considered as active. When no major event is found for 0.5 seconds, the active segment stops 0.1 seconds after the last major event in the active segment. We consider as ”quiet” the remaining part of the time series. Active and quiet segments are illustrated on Figure 14.

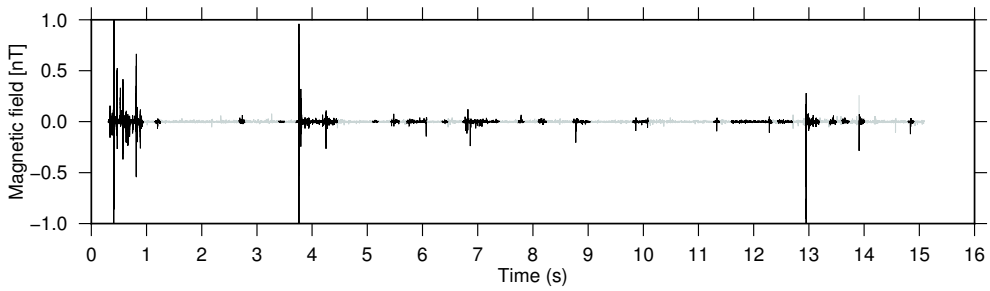


Figure 14. Illustration of the selection of active segments of Guenon time series. Black: Active time segments. Gray: Quiet time segments.

In the following, we distinguish three distincts datasets. The first one corresponds to active time periods, the second one to quiet time periods, and the third one gathering the whole time series. We process MT time series using the BIRRP (Bounded Influence Remote Reference Processing) code from the work of Chave & Thomson (2004).

On Figure 15, we present the output of the MT processing of the three datasets to be compared. The transfer functions corresponding to the quiet segment start to have wide confidence intervals for frequencies above 300 Hz. Active segments transfer functions remain interpretable for frequencies up to 1 kHz. Transfer functions corresponding to active segments show little difference with transfer functions corresponding to the processing of the whole time series despite its reduced length.

This indicates that processing results are mostly dependent on large signal-to-noise ratio events. Still, even though EM waves with significant frequency content over 1 kHz occurs, processing fail to recover the response functions over 1 kHz.

5 CONCLUSION

We develop a methodology for the automatic detection of lightning-based electromagnetic waves (ADEM). It is based on the characteristics of the EM waves in time-frequency plane. Two criteria are used: one relying on the amplitude of the expected waves; the second one lies in the behavior of local maxima in the time-frequency domain. Using both criteria, this procedure allows for direct detection, segmentation and characterisation of three types of lightning based EM waves, namely atmospherics, slow tails and whistlers. This scheme can be extended to events such as q-Bursts or tweeks. It allows either for a large detection where many events are recovered or a precise detection for events with large frequency content. Using this technique, we have shown that large signal-to-noise ratio events have significant impact on magnetotelluric (MT) data processing, specially for frequencies near the audio-MT dead band. Finally the use of a remote station is critical for a proper detection of lightning-based

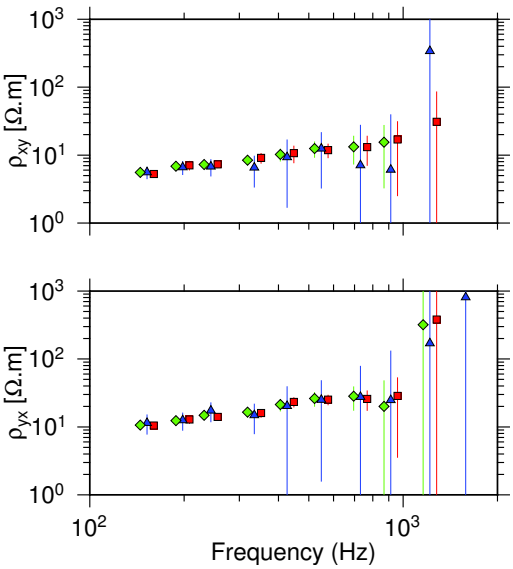


Figure 15. Results from MT processing of ELF waves using the whole data series (green), quiet times (blue) or active times (red).

EM waves. This does not change the framework of AMT acquisition where a remote station is almost always set up for processing. However, we have shown that large events with significant content in AMT dead band have low occurrence frequency. This suggests that longer time series may be necessary to acquire proper AMT data depending on the atmospheric activity. In our examples, 1 second of data may no be sufficient if no slow tails are present in the signal. Finally, this methodology allows to characterize the geomagnetic activity in events number, frequency content and polarisation of geomagnetic waves in AMT time series. This provides a way to control the quality of AMT response obtained through MT processing of time series.

REFERENCES

Brook, C.E.P., 1925, The distribution of thunderstorms over the globe, *Geophys. Mem.* (Air Ministry, Meteorological Office, London), **24**, 147–164.

Cagniard, L., 1953, Basic theory of the magnetotelluric method of geophysical prospecting, *Geophysics*, **18**, 605–635.

Chand, R., Israil, M., Rai, J., 2009, Schumann resonance frequency variations observed in magnetotelluric data recorded from Garhwal Himalayan region India, *Annales Geophysicae*, **27**:9, 3497–3507

Chave, A. D., Thomson, D. J., 2004, Bounded influence magnetotelluric response function estimation, *Geophys. J Int.*, **130**(2), 475–496.

Chave, A.D., Jones, A. G., 2012, The Magnetotelluric Method: Theory And Practice, *Cambridge University Press*, Cambridge

Cummer, S.A., Inan, U.S., 2000, Modeling ELF radio atmospheric propagation and extracting lightning currents from ELF observations, *Radio Science*, **35**, 1437–1444.

Cummer, S. A., 2000, Modeling electromagnetic propagation in the Earth-Ionosphere waveguide, *IEEE Transactions on antennas and propagation*, **48**, 1420–1429.

Diallo, M. S., Kulesh, M. Holschneider, M., Scherbaum, F., Adler, F., 2006, *Geophysics*, **71**(3), 67-77.

Escalas, M., P. Queralt, J. Ledo, A. Marcuello, Polarisation analysis of magnetotelluric time series using a wavelet-based scheme: A method for detection and characterisation of cultural noise sources, *Physics of the Earth and Planetary Interiors*, **218**, 31–50.

Farge M., 1992, Wavelet transforms and their applications to turbulence. *Annu. Rev. Fluid Mech.*, **24**, 395–457

Fisher, N. I., 1995, Statistical analysis of circular data, *Cambridge University Press*, Cambdridge, 277pp.

Fowler, R. A., Kotick, B. J., Elliott, R. D., 1967, Polarization analysis of natural and artificially induced geomagnetic micropulsations, *J. geophys. Res.*, **72**(11), 2871–2883.

Garcia, X., Jones, A. G., 2002, Atmospheric sources for audio-magnetotelluric (AMT) sounding, *Geophysics*, **67**(2), 448–458.

Garcia, X., Jones, A. G., 2008, Robust processing of magnetotelluric data in the AMT dead band using the continuous wavelet transform. *Geophysics*, , **73**(6), 223–F234.

Goupillaud, P., Grossmann, A., Morlet, J., 1984, Cycle-octave and related transforms in seismic signal analysis, *Geoexploration*, **23**, 85–102.

Helliwell, R. A., 1965. Whistlers and related ionospheric phenomena, *Stanford University Press*.

Holschneider, M., 1995, Wavelets, an analysis tool, *Oxford Science Publications*, Oxford.

Hughes, H.G., Papert, R.A., 1975, Propagation prediction model selection using VLF atmospherics, *Geophys. Res. Lett.*, **2**, 96–98.

Heydt, G., 1982, Instrumentation, in *Handbook of Atmospherics*, vol. 2, ed. H. Volland, Boca Raton, FLorida, CRC Press, 203–256

Kulesh, M., Nosé, M., Holschneider, M., Yumoto, K., 2007, Polarization analysis of a Pi2 pulsation using continuous wavelet transform, *Earth Planets Space*,**59**, 961–970.

Laby, T.H., McNeill, J.J., Nicholls, F.G., Nickson, A.F.B., 1940, Waveform, energy and reflexion by the ionosphere, of atmospherics, *Proc. Roy. Soc. A*, **174**, 145–163.

Larnier, H., Sailhac, P., Chambodut, A., 2016, New application of wavelets in magnetotelluric data processing: Reducing impedance bias, *Earth Planets Space*,**68**(70).

Mackay, C., Fraser-Smith, A. C., 2000, Lightning location using the slow tails of sferics, *Radio Science*, **45**

Maraun, D., Kurths, J., Holschneider, M., 2007. Nonstationary gaussian processes in wavelet domain: Synthesis, estimation, and significance testing, *Physical Review E*, **75**.

Rakov, V. A. & Uman, M. A., 2003, Lightning: Physics and Effects, *Cambridge University Press*, Cambridge.

Sailhac, P., Larnier, H., Warden, S., Nicollin, F., Zlotnicki, J., Bruère, F., 2016, Processing and modeling mag-

netotelluric data at La Soufrière of Guadeloupe lava dome (France), *EMIW 2016 Chiang Mai Extended Abstract*.

Schumann, W.O, Koenig, H., 1954, Über die beobachtung von atmospherics bei geringsten frequenzen, *Naturwissenschaften*, **41**, 183–184.

Sentman, D.D., 1990, Approximate Schumann resonance parameters for a two-scale-height ionosphere, *J. Atmos. Terr. Phys*, **52**, 35–46.

Strangway, D. W., Swift, C. M., Holmer, R. C., 1973, The application of audio frequency magnetotellurics (AMT) to mineral exploration, *Geophysics*, **38**, 1159–1175.

Torrence, C., Compo, G. P., 1998, A practical guide to wavelet analysis, *Bulletin of the American Meterological Society*, **79**:1, 61–78.

Wait, J. R., 1960, On the theory of the slow tail portion of atmospheric waveforms, *J. geophys. Res.*, **65**(7), 1939–1946.

Wait, J. R., 1962, Electromagnetic waves in stratified media, *Pergamon Press*, Oxford.

Williams, E. R., 1992, The Schumann Resonance: A Global Tropical Thermometer, *Science*, **256**, 1184–1187.

Zhang, Y., Paulson, K. V., 1997, Enhancement of signal-to-noise ratio in natural-source transient MT data with wavelet transform, *Pure and Applied Geophysics*, **149**, 405–419.

APPENDIX A: REPRODUCING KERNEL OF MORLET WAVELET

The reproducing kernel K_ψ of the wavelet ψ characterizes the correlation of the wavelet transform between two different points in the scale-time domain (Farge, 1992; Holschneider, 1995). It is defined by the correlation of the wavelet function dilated and translated using different parameters (a_1, b_1) and (a_2, b_2):

$$K_\psi(a_1, a_2, b_1, b_2) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \frac{1}{a_1} \psi\left(\frac{t - b_1}{a_1}\right) \frac{1}{a_2} \psi\left(\frac{t - b_2}{a_2}\right) dt \tag{A.1}$$

where C_ψ is the following normalizing factor defined from the Fourier transform of the wavelet ψ:

$$C_\psi = \int_{-\infty}^{\infty} |\hat{\psi}_0(\omega)| \frac{d\omega}{|\omega|} \tag{A.2}$$

By simple change of variables, one can write the reproducing kernel as the wavelet transform of the wavelet ψ:

$$K_\psi(a_1, a_2, b_1, b_2) = \frac{1}{C_\psi} \mathcal{W}_\psi\left(\frac{a_2}{a_1}, \frac{b_2 - b_1}{a_1}\right) \tag{A.3}$$

Calculation for the Morlet wavelet yields the following formula (Maraun *et al.*, 2007):

$$\begin{aligned} K_\psi(a_1, a_2, b_1, b_2) = & \frac{2a_1a_2}{a_1^2 + a_2^2} \exp\left(i\omega_0 \frac{a_1 + a_2}{a_1^2 + a_2^2}(b_2 - b_1)\right) \\ & \exp\left(-\frac{1}{2}\frac{(b_2 - b_1)^2 + \omega_0(a_2 - a_1)^2}{a_1^2 + a_2^2}\right) \end{aligned} \tag{A.4}$$

The correlation length of the Morlet reproducing kernel is used to define the resolution in time. By choosing a critical correlation level c, we define the resolution time limit at scale a by:

$$\begin{aligned} l_{\psi,c}(a_1) = & \max(b_2 | K_\psi(a_1, a_2, b_1, b_2) > c, \forall a_2, \forall b_2) \\ & - \min(b_2 | K_\psi(a_1, a_2, b_1, b_2) > c, \forall a_2, \forall b_2) \end{aligned} \tag{A.5}$$

APPENDIX B: DEFINITION AND RULES OF MAXIMA CHAINS

At scale a_j a local maximum at index i is defined by:

$$\begin{aligned} |\mathcal{W}_\psi(a_j, b_i)| > & |\mathcal{W}_\psi(a_j, b_{i+1})| \\ |\mathcal{W}_\psi(a_j, b_i)| > & |\mathcal{W}_\psi(a_j, b_{i-1})| \end{aligned} \tag{B.1}$$

Helliwell (1965) has shown that the group delay for whistlers (for frequencies below the nose frequency) can be approximated by:

$$\Delta t = D f^{-1/2}, \tag{B.2}$$

with D being a constant factor for every frequency called the dispersion (in $s^{1/2}$). This parameter is related to the density of electron on the path taken by the whistler wave (Helliwell, 1965). The time difference is characterized by being proportional to the inverse square root of the frequency. In the following, we use this dispersion assumption to chain local maxima in the time frequency plane for the three types of waves namely the atmospherics, the slow tails and the whistlers. The factor D is chosen as the parameter responsible for dispersion.

For slow tails, no dispersion is expected, so D is set at 0. For atmospherics and whistlers, the dispersion depends on the path taken from the source. The D is therefore variable and has to be assigned to each type of events. In the paper, we have used $D = 0.012\ t^{1/2}$ for atmospherics and $D = 0.12\ t^{1/2}$ for whistlers waves. Besides, because of noise in the time series, we allow a slight variation along the time axis around the function defined by equation B.2. This time interval where the local maxima is sought is defined as proportional to the half width $l_{\psi,c}(a)/2$ of the Morlet kernel K above a critical value c (as detailed in Appendix A).

In the following, we call M_i the set of all local maxima at scale a_i. The chain rule is:

(i) At starting scale a_α, we seek every local maxima and determine M_α, the set of all local maxima at scale a_α.

(ii) For every $b_{i_\alpha} \in M_\alpha$, we connect the local maxima (a_α, b_{i_α}) to a maximum on larger scale when:

$$\exists b_{i_{\alpha+1}} \begin{cases} b_{i_{\alpha+1}} \in M_{\alpha+1} \\ b_{i_{\alpha+1}} \in [b_{i_\alpha} + \Delta_t - l_{\psi,c}(a_{\alpha+1})/2 : b_{i_\alpha} + \Delta_t + l_{\psi,c}(a_{\alpha+1})/2] \end{cases} \tag{B.3}$$

(iii) To connect it to a maximum on smaller scale:

$$\exists b_{i_{\alpha-1}} \begin{cases} b_{i_{\alpha-1}} \in M_{\alpha-1} \\ b_{i_{\alpha-1}} \in [b_{i_\alpha} - \Delta_t - l_{\psi,c}(a_{\alpha-1})/2 : b_{i_\alpha} + \Delta_t + l_{\psi,c}(a_{\alpha-1})/2] \end{cases} \tag{B.4}$$

APPENDIX C: GENERATION OF SYNTHETIC ELF WAVES

To create synthetic magnetic field time series, we generate magnetic field responses of far lightning pulses following Wait (1960) and Wait (1962). The propagation medium is defined as the atmosphere comprised in a spherical shell from the ground at $r = a$ and the ionosphere at a distance $r = a + h$ where h the ionosphere height. Both layers act as waveguide boundary where the field propagates in modes. We assume that the lightning source is equivalent to a vertical electric dipole source on the Earth's surface. The current source is assumed to be a impulse described by a dirac δ function. The frequency domain response is then given by:

$$P_s(\omega) = |p_s| \tag{C.1}$$

The vertical electric field at the source location is given by:

$$E_0(\omega) = i\left(\frac{\eta\omega}{2\pi c\rho}\right)\left(e^{-i\omega\rho/c}\right)P_s(\omega) \tag{C.2}$$

where ρ is the distance source-receiver, c the speed of light, η the impedance of free space and $\omega = 2\pi f$ the wave pulsation.

Wait (1962) derived the expression of the EM field in frequency domain at receiver location. This field is given by a sum over all the waveguide modes:

$$\begin{aligned} H_r(\omega) \cong \left(\frac{\rho/R_E}{\sin(\rho/R_E)}\right)^{1/2} \frac{(2\pi c\rho)^{1/2}}{h\eta} \frac{E_0\omega}{(i\omega)^{1/2}} \\ \cdot \sum_{n=0}^{\infty} \Re(S_n(\omega))^{1/2} \delta_n e^{-i\omega\rho S_n(\omega)/c}, \end{aligned} \tag{C.3}$$

where R_E is the Earth radius, h the height of the ionosphere, η the intrinsic impedance of free space. δ_n is related to the excitation of each mode and depends on reflexion coefficients on the ground and the ionosphere. S_n is the sinus of the complex angle of the incident wave on the ionosphere. The magnetic

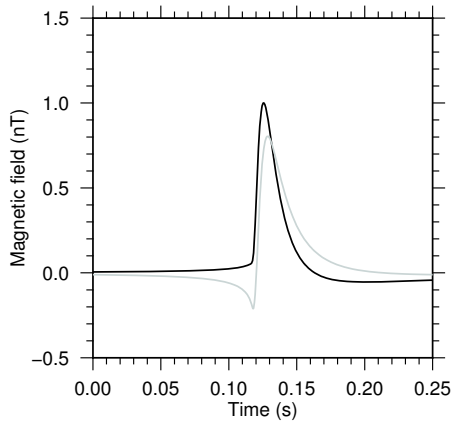


Figure A1. Synthetic slow tail with parameters $e = 0.262$, $\theta = 0.826$, $\Delta\phi = 150.5$.

field in time domain at receiver location is then given by:

$$h_r(t) = \int_{-\infty}^{+\infty} H_r(\omega) e^{i\omega t} d\omega \tag{C.4}$$

The slow tails propagate for frequencies below 2 kHz. The first cutoff frequency $n = 0$ is about 1.8 kHz (Mackay & Fraser-Smith, 2010), so only the 0th mode is kept in the sum.

Following these assumptions, the magnetic field equation becomes:

$$\begin{aligned} H_r(\omega) \cong \left(\frac{\rho/R_E}{\sin(\rho/R_E)}\right)^{1/2} \frac{(2\pi c\rho)^{1/2}}{h\eta} \frac{E_0\omega}{(i\omega)^{1/2}} \\ \cdot \Re(S_0(\omega))^{1/2} \delta_0 e^{-i\omega\rho S_0(\omega)/c}, \end{aligned} \tag{C.5}$$

The Earth's radius is set to $R_E = 6371$ km, the speed of light is $c = 3\ 10^5$ km/s, free space impedance is equal to 376.730 Ω, the ionospheric conductivity is $\sigma = 10^{-6}$ S/m, and the free space permittivity $\varepsilon_0 = 9.10^{-12}$ F/m. δ_0 is approximatively equal to 1 for $n = 0$ (Wait, 1960) and S_0 can be approximated to:

$$S_0(\omega) = \left(1 + \frac{c}{h(i\sigma_i\omega/\varepsilon_0)^{1/2}}\right)^{1/2}, \tag{C.6}$$

where σ_i is the ionosphere conductivity (assumed to be homogeneous here with a value of 10^{-6} S/m).

From this magnetic field, we build a polarized signal following the monochromatic definition of polarization attributes (Fowler *et al.*, 1967). For each event, we set $\Delta\phi$ the phase difference between horizontal components and the relative amplitude A_x and A_y of both orthogonal components. The

ellipticity e and the polarization angle θ are then given by:

$$\sin(2\arctan(e)) = \frac{2A_xA_y}{A_x^2+A_y^2}\sin(\Delta\phi) \tag{C.7}$$

$$\tan(2\theta) = \frac{2A_xA_y}{A_x^2-A_y^2}\cos(\Delta\phi) \tag{C.8}$$

On Figure A1, we illustrate an exemple of a polarized synthetic slow tail. Each event was then randomly assigned to a time location along the time series.

Finally, we add order 1 autoregressive noise to the time series using the following model:

$$x_{i+1} = c + \alpha x_i + \sigma\varepsilon_i, \tag{C.9}$$

with α the parameter of the AR-1 model, ε_i a standard white noise model, σ the variance applied to the white noise to increase its amount in the modeled time series and c a constant. c is set to zero in this experiment.

To create the associated remote time series, we assume that the wave is quasi-uniform on the area where both stations are set. The polarisation attributes and the time index of synthetic events are then the same in both local and remote time series. Only the added noise is different in the two sets of time series.

2.2.5 Orages proches

Le traitement des données MT sous-entend l'hypothèse de source lointaine pour respecter le principe de source quasi-uniforme (A. D. Chave et A. Jones, 2012) ce qui n'est pas nécessairement avéré pour toutes les ondes observées. La source doit donc avoir des propriétés constantes sur une grande distance, ainsi, les attributs de polarisation doivent donc aussi être constants pour deux stations proches. Cette propriété peut nous permettre d'éliminer une partie des ondes ne respectant pas la théorie MT en utilisant un simple critère sur la polarisation horizontale.

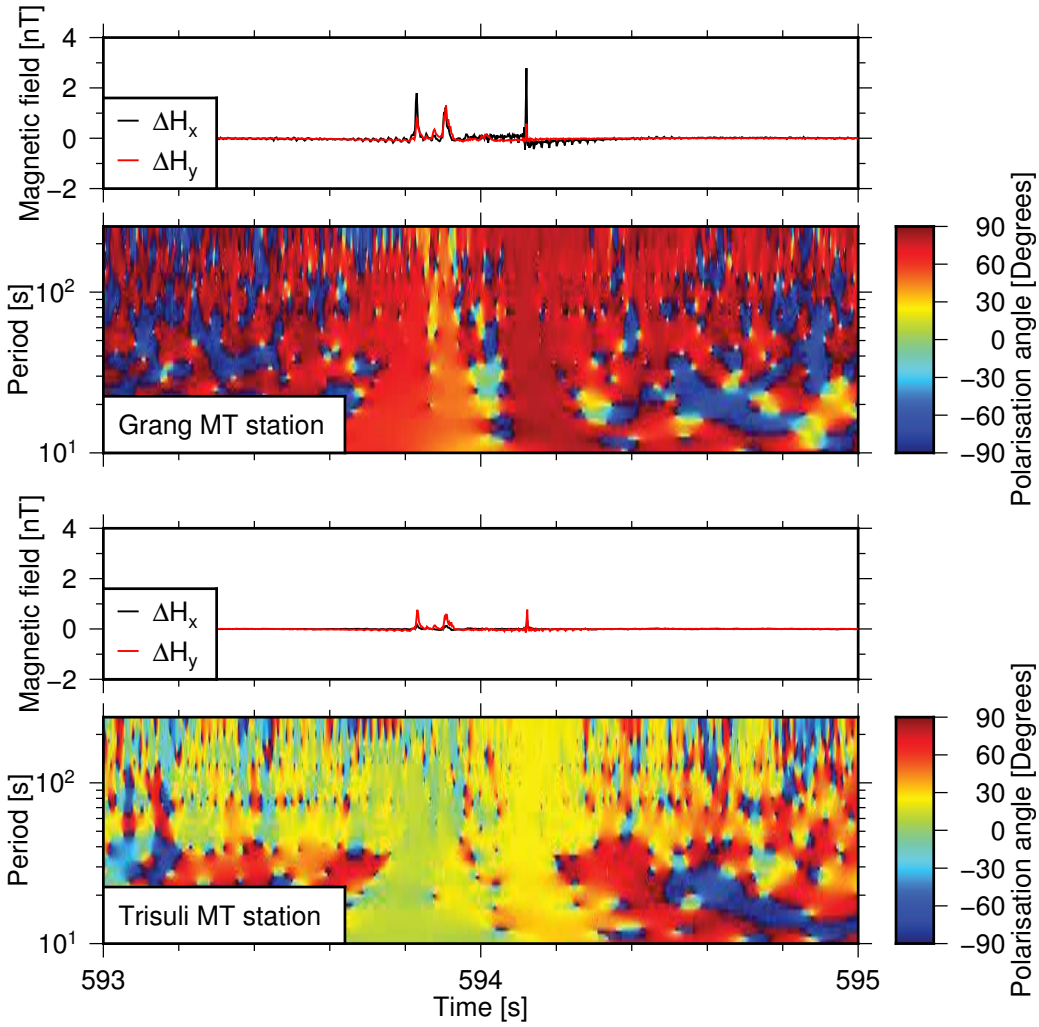


Figure 2.12: Comparaison des angles de polarisation horizontale enregistrés pour une onde ELF enregistrée au Népal aux stations de Trisuli et Grang le 14 Mars 2016.
FIGURE 2.12: *Comparison of horizontal polarisation angle recorded at Trisuli and Grang MT stations on 14 March 2016.*

Lors de la campagne 2016 au Népal (voir Chapitre 5), un violent orage a eu lieu le 14 Mars 2016. Les stations Trisuli et Grang étaient en fonctionnement à cette date et nous permettent donc d'étudier les attributs de polarisation pour des éclairs proches. La Figure 2.12 montre un

Type de pulsations		Périodes (s)
Pulsations continues	Pc1	0.2 - 5
	Pc2	5 - 10
	Pc3	10 - 45
	Pc4	45 - 150
	Pc5	150 - 600
Pulsations irrégulières	Pi1	1 - 40
	Pi2	40 - 150

TABLE 2.1: Classification des pulsations géomagnétiques. D'après J. A. JACOBS et al. (1964).
TABLE 2.1: *Geomagnetic pulsations classification. From J. A.* JACOBS *et al. (1964).*

exemple d'onde ELF enregistrée. Les attributs de polarisation calculés pour 2 stations espacées de seulement 15 km indiquent clairement une large différence de l'angle de polarisation. Il apparaît ici que l'utilisation de l'angle de polarisation horizontale permet de déterminer une partie des ondes ne respectant pas l'hypothèse de source quasi-uniforme.

2.3 Ondes basses fréquences

L'autre catégorie de sources étudiées dans cette thèse est l'ensemble des sources dont les fréquences d'oscillations sont inférieures au Hz. Elles sont principalement produites par l'interaction du soleil avec la magnétosphère terrestre (GUBBINS et HERRERO-BERVERA, 2007).

2.3.1 Pulsations géomagnétiques

Du 28 août au 7 septembre 1861, le scientifique écossais Balfour Stewart rapporte la présence d'aurores d'intensités exceptionnelles (*unprecedented magnificience*) tout autour du globe et enregistra des variations sinusoïdales du champ magnétique (STEWART, 1861). Par la suite, du fait de l'évolution technologique des instruments utilisés, les enregistrements de ce type d'oscillations ont augmenté. Le nom donné à ces oscillations est celui de micropulsations géomagnétiques.

En 1963, à la treizième assemblée générale de l'Union Internationale De Géodésie et de Géophysique (IUGG), un comité a travaillé sur la classification des pulsations. J. A. JACOBS et al. (1964) a proposé pour la première fois une classification (Table 2.1). Deux types de pulsations ont été décrits :

- Les pulsations continues sont des oscillations du champ magnétique pouvant durer plusieurs heures (Figure 2.13).
- Les pulsations irrégulières sont des oscillations dont la durée totale est courte (Figure 2.14).

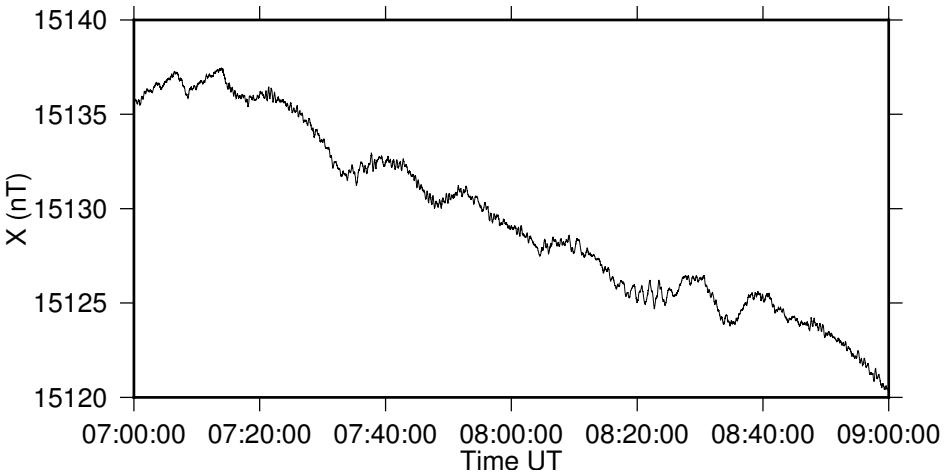


FIGURE 2.13: Pulsations de type Pc3 observées à l'observatoire d'Uppsala (UPS, Suède) le 12 novembre 2013.
FIGURE 2.13: *Pc3 pulsation recorded at Uppsala (UPS, Sweden) observatory on 12 November 2013.*

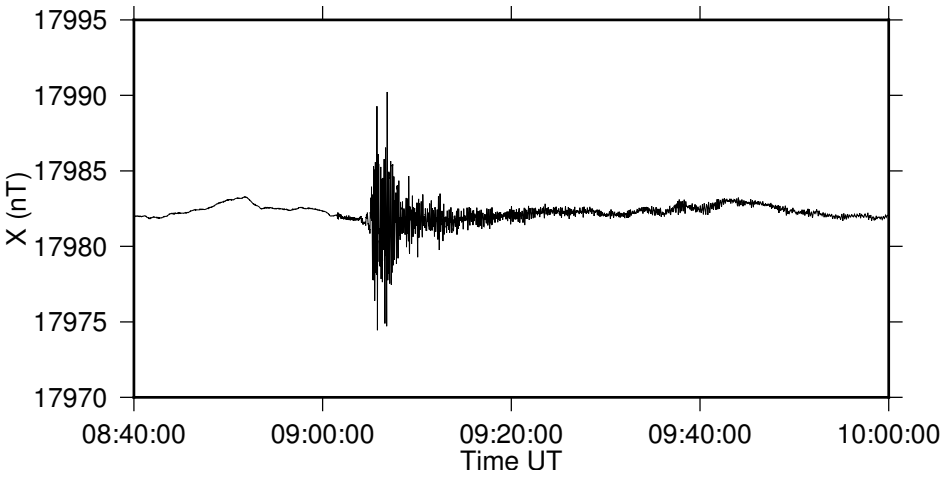


FIGURE 2.14: Pulsations de type Pi1 observées à l'observation de Newport (NEW, Etats-Unis), le 05 janvier 2013.
FIGURE 2.14: *Pi1 pulsation recorded at Newport (NEW, United States of America) observatory on 05 January 2013.*

Cette classification est cependant artificielle car son établissement s'est fait à une date où les processus physiques sous-jacents à la génération des pulsations géomagnétiques étaient encore relativement peu connus. Ainsi, certaines pulsations continues peuvent présenter une période principale, variant avec le temps, passant d'une classe à une autre.

Dans la suite de cette section, nous décrirons les principaux mécanismes de génération et de propagation des pulsations géomagnétiques en les illustrant sur des données d'observatoires magnétiques.

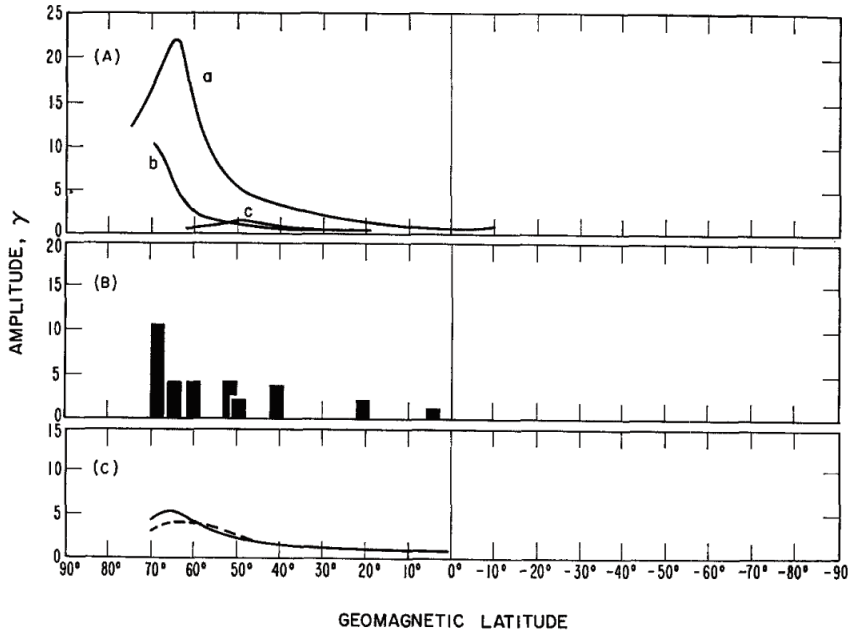


FIGURE 2.15: Variation de l'amplitude de trois pulsations Pc3 (A, B et C) avec la latitude géomagnétique, extrait de Takao SAITO (1969) d'après les études de J. JACOBS et SINNO (1960), KATO (1962) et KATO et T SAITO (1962).
FIGURE 2.15: *Variation of the amplitude of 3 Pc3 pulsations (A, B, and C) with geomagnetic latitude. From Takao* SAITO *(1969), results from studies by J.* JACOBS *et* SINNO *(1960),* KATO *(1962) et* KATO *et T* SAITO *(1962).*

2.3.1.1 Mécanismes de génération

La source des pulsations géomagnétiques se trouve essentiellement dans l'interaction entre le vent solaire et la magnétosphère terrestre (MCPHERSON, 2005). Il est généralement admis que les principaux mécanismes pour les ondes de périodes supérieures à 40 s concernent une excitation par un cisaillement entre le vent solaire et le plasma de la magnétosphère au niveau de la magnétopause. Les phénomènes d'instabilités de Kelvin-Helmoltz (FEJER, 1964 ; YUMOTO et Takao SAITO, 1980 ; HWANG et al., 2012) fournissent un premier modèle expliquant la présence d'ondes longues périodes. Ce phénomène se créé à une interface entre deux fluides de vitesses différentes et provoque l'apparition de tourbillons à l'interface entre les deux fluides. Un tel mécanisme a récemment été mis en doute quand à la génération de l'ensemble des pulsations longues périodes (PILIPENKO et al., 2010 ; KAVOSI et RAEDER, 2015). D'autres sources de génération d'ondes MagnétoHydroDynamiques (MHD) longues périodes identifiées sont les phénomènes de résonance de ligne de champ (CHEN et HASEGAWA, 1974 ; SOUTHWOOD et W. HUGHES, 1983 ; VELLANTE et al., 2004), de variations oscillatoires dans la pression du vent solaire (KEPKO, SPENCE et SINGER, 2002) ou de dérives des particules dans le champ magnétique terrestre (YEOMAN et WRIGHT, 2001 ; MCPHERSON, 2005). Les pulsations courtes périodes sont créés par d'autres mécanismes tels que les instabilités ion/cyclotron lors d'injections de particules chargées dans

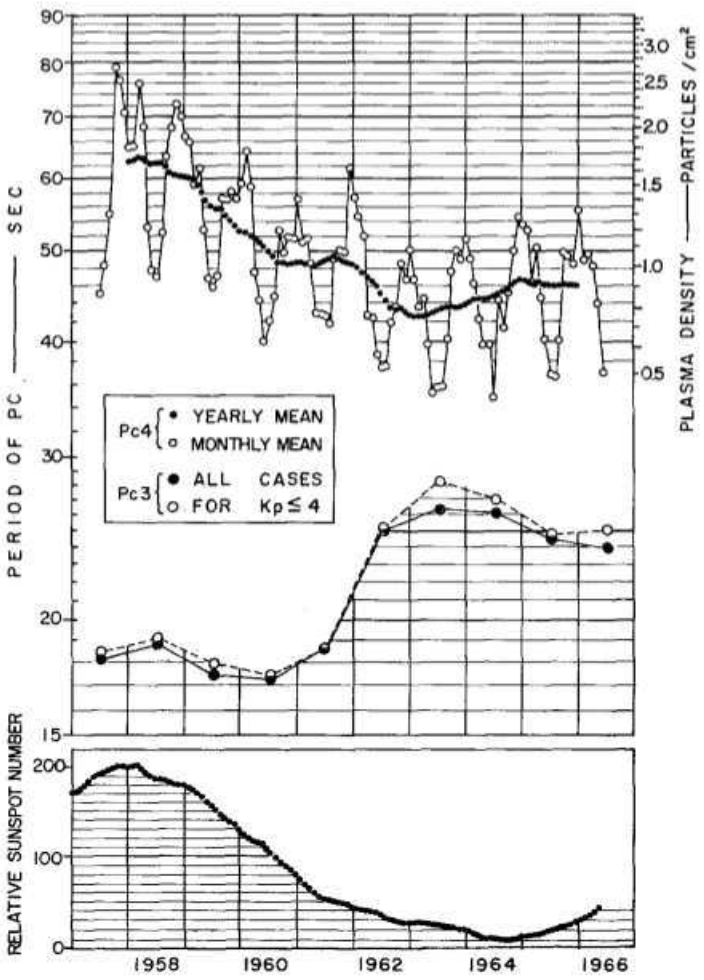


FIGURE 2.16: Variation de la période de pulsations Pc3 et Pc4 avec le cycle solaire. Extrait de Takao SAITO (1969), d'après les études de T SAITO et MATSUSHITA (1967) et KAWAMURA (1969b).
FIGURE 2.16: *Variation of Pc3 and Pc4 pulsations period with the solar cycle. From Takao* SAITO *(1969), results from studies by T* SAITO *et* MATSUSHITA *(1967) et* KAWAMURA *(1969b).*

la magnétosphère à l'occasion de sous-orages géomagnétiques (LEE et KWOK, 1984 ; KANGAS, GUGLIELMI et POKHOTELOV, 1998).

2.3.1.2 Caractéristiques spatio-temporelles

L'ensemble des pulsations présente une complexité liée à leurs variabilités à la fois spatiales et temporelles. Sans vouloir être exhaustif sur l'ensemble des propriétés recensées dans la littérature, nous allons présenter les principales caractéristiques des pulsations géomagnétiques. Pour aller plus loin, MATSUSHITA (1967), Takao SAITO (1969) et MCPHERSON, 2005 sont un bon point de départ et contiennent un grand nombre de références à ce sujet.

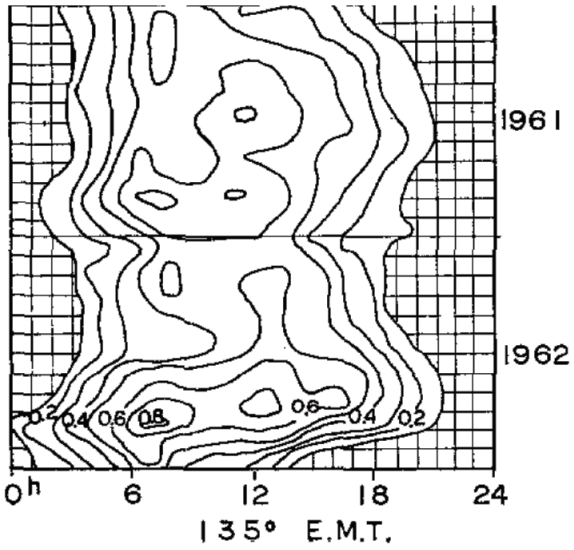


FIGURE 2.17: Variation de la probabilité d'apparition d'une pulsation Pc3 en fonction de l'heure du jour et du mois en 1961 et 1962. Extrait de Takao SAITO (1964) et Takao SAITO (1969).
FIGURE 2.17: *Diurnal variation of Pc3 pulsation occurence from 1961 to 1962. From Takao* SAITO *(1964) et Takao* SAITO *(1969).*

Amplitude des pulsations Ce paramètre est sans doute le caractère morphologique le plus aisément observable de façon directe. Ainsi, l'amplitude des pulsations Pc1 est généralement de l'ordre du dixième voir du centième de nT, et leur observation nécessite des appareils avec une grande sensibilité ou un environnement magnétique particulièrement calme. Les pulsations de type Pc2-3 ont une amplitude plus importante, pouvant atteindre le nT. Enfin, les pulsations de type Pc4-5 peuvent atteindre la dizaine de nT. Ces affirmations peuvent être tempérées puisqu'elles dépendent également de la latitude d'observation. Comme le montre la Figure 2.15, l'amplitude d'une même Pc3 varie avec la latitude géomagnétique.

Période des pulsations La seconde caractéristique visible des pulsations est l'évolution de leur période avec le temps. Les Pc4-5 présentent une variation saisonnière de la période avec une diminution de la période pendant l'été. Le cycle solaire présente aussi une influence sur la période des pulsations Pc4 avec une augmentation de la période pendant les minimums du cycle solaire. Les Pc1 et Pc2-3 présentent elles une variation de leur période avec l'heure du jour. De plus, la période moyenne des Pc3 varie aussi avec le cycle solaire (Figure 2.16).

Fréquence d'apparition La fréquence d'apparition des pulsations varie spatio-temporellement sur plusieurs échelles. L'analyse des données d'observatoires montre que les pulsations Pc2-3 apparaissent principalement du côté jour de la magnétosphère, avec un maximum de probabilité autour de 12h MLT (Figure 2.17). Cette variabilité peut être interprétée comme un manque de

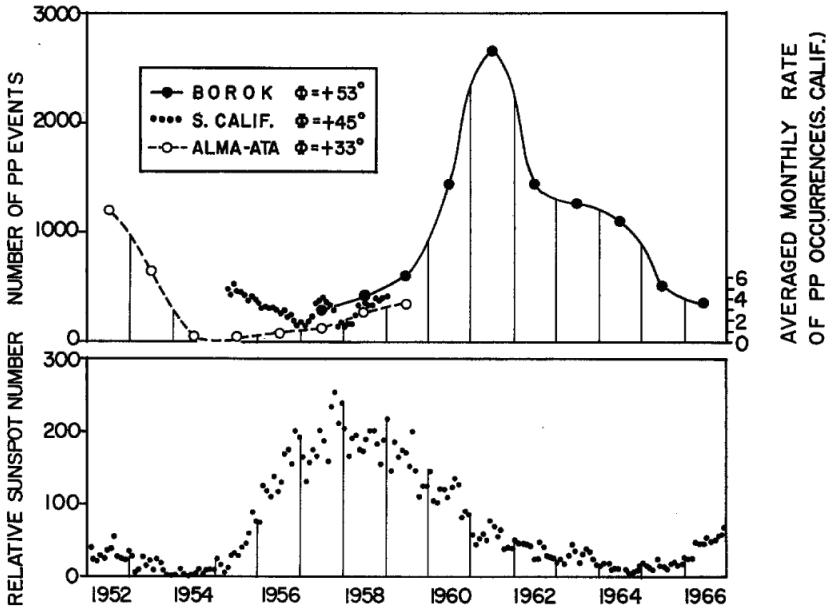


FIGURE 2.18: Haut : Variation de la fréquence d'apparition des pulsations Pc1 entre 1952 et 1966. Bas : Nombre de taches solaires pour la même période. Extrait de Takao SAITO (1969), d'après les études de BENIOFF (1960) et TROITSKAYA (1967).
FIGURE 2.18: *Top : Variation of Pc1 pulsation occurence rate from 1952 to 1966. Bottom : Solar sunspot number for the same period. From Takao* SAITO *(1969), results from studies by* BENIOFF *(1960) et* TROITSKAYA *(1967).*

sensibilité des capteurs à une faible amplitude des pulsations (MATSUSHITA, 1967) qui peuvent être présentes à tout moment. Un comportement opposé est observé sur les Pi2 qui apparaissent principalement côté nuit. Ces pulsations présentent donc une variation avec la longitude ou avec le temps local. Le cycle solaire a aussi une influence sur la fréquence d ?occurrence des pulsations. Les Pc1 montrent en particulier un pic d ?occurrence situé légèrement avant le maximum solaire (BENIOFF, 1960) (Figure 2.18).

Polarisation SAMSON, J. JACOBS et ROSTOKER (1971) ont décrit l'évolution de la polarisation des pulsations géomagnétiques pour 7 stations enregistrant les variation du champ magnétique pendant 3 jours. Ils ont montré que la polarisation des pulsations Pc4-5 variait à la fois avec la latitude géomagnétique et l'heure locale (Figure 2.19).

Toutes ces caractéristiques renseignent sur l'état de la magnétosphère et doivent être retrouvées par l'analyse des pulsations par un algorithme automatique.

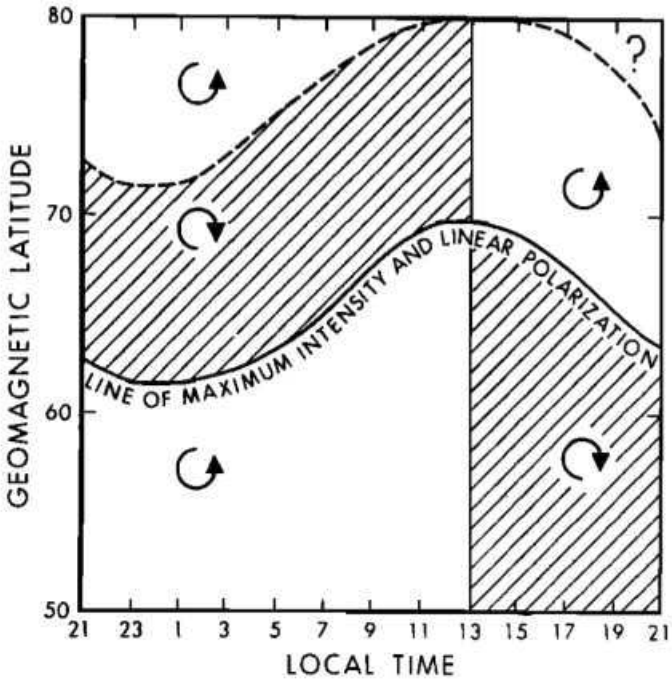


FIGURE 2.19: Variation de la polarisation des pulsations Pc4 et Pc5 au Canada. Extrait de SAMSON, J. JACOBS et ROSTOKER (1971).
FIGURE 2.19: *Variation of the polarisation of geomagnetic pulsations Pc4 and Pc4 in Canada. From* SAMSON, *J.* JACOBS *et* ROSTOKER *(1971).*

2.3.2 Détection - Potential paper

Résumé en français Cet article introduit une méthodologie pour la détection automatique des oscillations du champ magnétique de période 1-1000 secondes. Cette détection est basée sur la comparaison entre une population de *surrogates* de Fourier et les signaux analysés, permettant la sélection de coefficients significatifs dans le plan temps-fréquence. Les *surrogates* de Fourier sont calculés à partir d'un spectre en puissance moyen calculé sur 24h. Les pulsations étant visibles simultanément sur les trois directions de l'espace, les *surrogates* sont générés indépendamment dans les trois directions.

Cette procédure permet le traitement rapide de nombreuses données d'observatoires, ouvrant la porte à l'étude des pulsations à un niveau global sur un grand nombre d'observatoires magnétiques. Nous avons implémenté cette méthode sur un *cluster* de calcul et les résultats bruts de l'application sur des données seconde d'observatoires magnétiques sont disponibles. Deux applications de ce travail sont montrés. La première est une étude statistique des propriétés des pulsations Pc3 pendant 2 ans de mesures au niveau des observatoires magnétiques d'Uppsala et de Nurmijarvi. La seconde est une étude courte de la variation spatial des pulsations Pc5 au niveau de l'Amérique du Nord.

An automatic algorithm for the detection of geomagnetic pulsations based on continuous wavelet transform and Fourier surrogates

H. Larnier[1]*, A. Chambodut[1], P. Sailhac[1,2]

[1]Institut de Physique du Globe de Strasbourg; UMR7516, Université de Strasbourg/EOST, CNRS; 5 Rue René Descartes, F-67084 Strasbourg Cedex

[2]Université Paris Sud.

Key Points:

- Automatic detection of geomagnetic pulsations of the magnetic field.
- Available catalog for the geomagnetic community.
- Global studies available.

*Current adress:

Corresponding author: H. Larnier, `hlarnier@unistra.fr`

Abstract

This paper introduces a methodology to automatically detect oscillations of the magnetic field in the period range 1-1000 seconds. The detection scheme is based on the detection of large amplitude wavelet coefficients and cross-spectrogram coefficients between magnetic components and the use of Fourier surrogates. The reference Fourier spectrum for surrogate generation is chosen as a quiet daily magnetic spectrum. Fourier surrogates are generated independently on each channel so geomagnetic pulsations are tested against independence of channels. This procedure allows to efficiently process large datasets to perform global studies on several magnetic observatories. We implemented the method on a computing cluster and raw results for 1 Hz magnetic observatories data are made publicly available. Two applications of this work are shown. The first one is a statistical study of Pc3 pulsations properties over two years span at Uppsala and Nurmijarvi magnetic observatories. The other one is a short spatial study of Pc5 pulsations over North America.

1 Introduction

Geomagnetic pulsations are oscillations of the magnetic field with periods from 1 to 1000 seconds [*Saito*, 1969]. The origin of such oscillations lies in the interaction of the solar wind with the magnetosphere [*McPherron*, 2005]. Various mechanisms produce magnetohydrodynamic (MHD) waves in the magnetosphere whose manifestation on ground surface are the so-called geomagnetic pulsations. Various types of pulsations are observable. A first attempt of classification was made in 1964 by *Jacobs et al.* [1964]. Two main classes of pulsations were derived, continuous (Pc) and irregular (Pi). Pc pulsations are considered as quasi-stationary waves with varying amplitude during the day that depend mainly on the magnetic local time (MLT). Irregular pulsations are transient events with large amplitude. Such waves characteristics obtained through magnetic measurements, from satellites or magnetic observatories, provide valuable information about space weather parameters [*McPherron*, 2005]. Time occurrence of pulsations may also be of interest for magnetotelluric practitioners in data processing schemes [*Larnier et al.*, 2016].

Currently, more than 150 INTERMAGNET observatories are performing measurements of the Earth's magnetic field worldwide. A significant number of these observatories produce 1 Hz measurements that can be used to study processes such as geomagnetic pulsations. This rather large amount of data requires an automatic procedure to

study large scale processes. Several schemes have already been introduced to recover geomagnetic pulsations occurrence and properties. *Nosé et al.* [1998] developed a methodology based on wavelet analysis and empirical thresholding of the amplitude of wavelet coefficients. Manual thresholding has to be adapted to each observatory and each day of measurements so fast application of this methodology to a large dataset is a critical bottleneck for global studies. *Bortnik et al.* [2007] used Fourier transform and thresholding by comparison to a median spectrum. However, Fourier transform does not account for fast variations of signal's frequency or polarization properties [*Kulesh et al.*, 2007].

In this paper, we introduce a methodology based on the application of the continuous wavelet transform combined with surrogate simulation to perform automatic detection and characterization of geomagnetic pulsations. In section 2, we describe the method to recover their frequency and amplitude in the time-frequency domain as well as their polarization attributes. In section 3, we consider two examples of applications. The first one is a statistical study of temporal evolution of pulsations parameters on two observatories over three years. The second one is a spatial study over three observatories on North America for one day of measurements. Observatories considered in this study are:

- Uppsala (UPS, Sweden): $\theta = 59.903°$N, $\lambda = 17.353°$E.
  $\theta_m = 56.6°$N, longitude $\lambda_m = 95.43°$E.
- Nurmijarvi (NUR, Finland): $\theta = 60.508°$N, $\lambda = 24.655°$E.
  $\theta_m = 57.06°$N, $\lambda_m = 101.84°$E.
- Newport (NEW, United States of America): $\theta = 48.265°$N, $\lambda = 242.877°$E.
  $\theta_m = 54.71°$N, $\lambda_m = 304.75°$E.
- Fresno (FRN, United States of America): $\theta = 37.091°$N, $\lambda = 240.281°$E.
  $\theta_m = 42.91°$N, $\lambda_m = 304.64°$E.
- Meanook (MEA, Canada): $\theta = 54.616°$N, $\lambda = 246.653°$E.
  $\theta_m = 61.79°$N, $\lambda_m = 307.52°$E,

where θ and λ are respectively the geographic latitude and longitude, and θ_m and λ_m are respectively the corrected geomagnetic coordinates (CGM) using IGRF-11 for epoch 2010.00. X, Y and Z are the geographic north, east and vertical down magnetic field components respectively. Only variations with periods below 1000 seconds are consid-

ered in this work, so, all time series are high-pass filtered with a cut-off frequency corresponding to 2000 seconds. Filtered data are denoted $\Delta X, \Delta Y, \Delta Z$ in the following.

2 Methods

The methodology described in this section is illustrated on the magnetic time series of Uppsala geomagnetic observatory for the 5th of January 2013.

2.1 Wavelet analysis

To account for the large variability of geomagnetic activity in terms of frequency and time simultaneously, the continuous wavelet transform (CWT) is applied on magnetic time series using the Morlet wavelet with wavelet parameter $\omega_0 = 6$ [*Holschneider*, 1995]. The CWT allows to obtain a time-frequency representation of time series through scale dilatation a (corresponding to frequency) and time translation b of an analyzing wavelet ψ. Wavelet coefficients of a signal s are obtained by application of the convolution (1).

$$\mathcal{W}_\psi(a,b)[s] = \int_{-\infty}^{\infty} s(t)\frac{1}{a}\psi\left(\frac{t-b}{a}\right)dt \tag{1}$$

The CWT allows to describe the time-frequency evolution of temporal signals and also to obtain the horizontal polarization attributes in the wavelet domain [*Diallo et al.*, 2006; *Kulesh et al.*, 2007]. Geomagnetic pulsations often occur simultaneously on the three channels. To enhance their signal-to-noise ratio, we use the sum of the three cross-spectrograms between magnetic channels [*Bortnik et al.*, 2007]. The cross-spectrogram between two time series x and y is:

$$\mathcal{W}_\psi^{XY}(a,b) = \mathcal{W}_\psi(a,b)[x] \cdot \mathcal{W}_\psi^*(a,b)[y], \tag{2}$$

where * corresponds to the complex conjugate operator and $\cdot$ is the element-to-element product. The cross-spectrogram tested for significance is then given by:

$$\mathcal{W}_\psi^{XYZ}(a,b) = |\mathcal{W}_\psi^{XY}(a,b)| + |\mathcal{W}_\psi^{XZ}(a,b)| + |\mathcal{W}_\psi^{YZ}(a,b)| \tag{3}$$

Figure 1 presents $\mathcal{W}_\psi^{XYZ}$, the cross-spectrogram of the example time series for the 5th of January 2013. Coherent processes have enhanced amplitude while the amplitude of noise is reduced. Coherent wavelet coefficients with large amplitude correspond on all components to part of the signals with higher energy that can be related to geomagnetic pulsations.

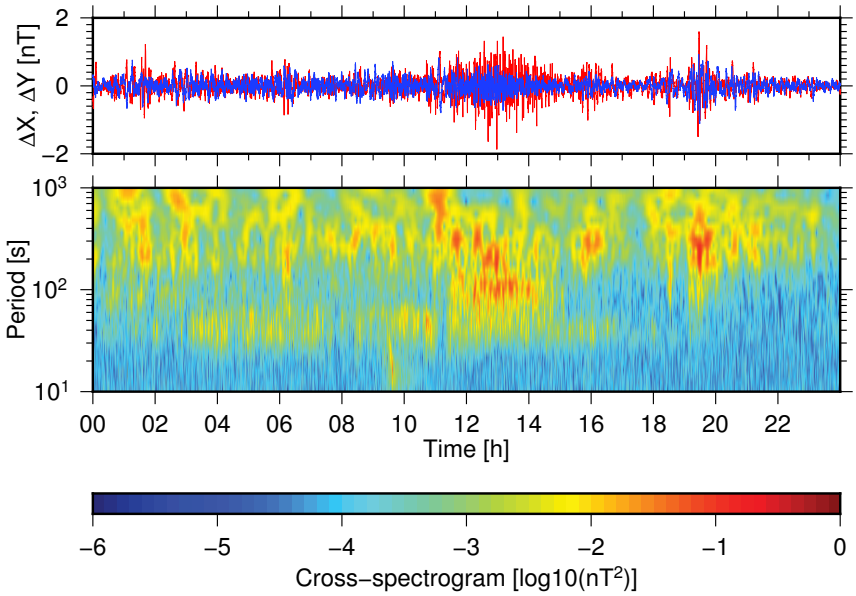


Figure 1. Top: Filtered time series of horizontal magnetic field at UPS magnetic observatory for the 05th of January 2013. Vertical magnetic field is not shown for visual convenience. Bottom: Cross-spectrogram $\mathcal{W}_{\psi}^{XYZ}(a,b)$.

2.2 Surrogate data testing

Geomagnetic pulsations have different diurnal variations for each pulsation type. Pc3 pulsations have larger occurrence rate in the morning in local time than in the evening whereas Pi2 pulsations have a maximum of occurrence located near the local midnight [*Saito*, 1969]. Moreover, noise level differs from one observatory to another. Such characteristics have to be taken into account by a data adaptive technique.We consider the surrogate technique that has been introduced to recover significant frequencies in non-stationary processes. Surrogates are realizations of time-series sharing specified properties with the analyzed time series.The analyzed time series is then tested against the whole surrogate population. Among the large range of surrogates available [*Cazelles et al.*, 2014], we choose the constrained Fourier surrogate. It relies on the null hypothesis: the original time series is considered as a stochastic process that can be characterized by the autocorrelation function of the original time series. Realizations of such surrogates are obtained by randomizing the phases of the Fourier spectrum of the original time series and returning back in time domain through inverse Fourier transform. If the Fourier trans-

form of the original time series s, is given by coefficients S_k at a frequency index k in Fourier domain, then, a Fourier surrogate $\hat{s}$ in time domain is given by:

$$\hat{s_n} = \sum_{k=1}^{N} |S_k| e^{i\beta_n} \exp\left(\frac{2ik\pi n}{N}\right), \tag{4}$$

where n is the time index and the set of phase differences β_n is drawn independently from the uniform distribution from $-\pi$ to π. Estimation of the amplitude of Fourier transform through simple Fast Fourier Transform estimation often leads to bias due to the finite length of the analyzed time series. To properly estimate its amplitude, we use the multi-taper technique, proposed by *Thomson* [1982] and *Wardinski & Mandea* [2006], to estimate the power spectrum P_k. Power spectrum and Fourier transform are then related by:

$$|S_k| = N(P_k)^{1/2} \tag{5}$$

Due to the variability of the geomagnetic pulsations with time, the power spectrum associated to the geomagnetic pulsations also varies with time. Figure 2 shows the hourly power spectra of the considered UPS dataset. The power spectrum associated to Pc3 pulsations varies on almost two orders of magnitude during the day. One hour is particularly characterized by large amplitude in the Pc2 pulsations frequency domain from 9 to 10 LT shown by gold line on Figure 2. As reference power spectrum for the surrogate time series simulation, we use the median spectrum $\bar{P}_k$ of hourly power spectra P_k^i at each frequency index k defined by equation (6.) Such example of average spectrum is shown Figure 2 by the thick red line.

$$\bar{P}_k = Med(P_k^i,\ i = 1, .., 24) \tag{6}$$

Like *Cazelles et al.* [2014], we test the original wavelet quantities against the distributions of wavelet quantities of surrogates functions. For each surrogate triplet (X_n, Y_n and Z_n), we compute the cross-spectrogram defined by equation (3). At each scale a_l in the wavelet analysis, the original cross-spectrogram is tested against the population of surrogates. The α percentile is extracted and any cross-spectrogram coefficients above this percentile is kept for further analysis.

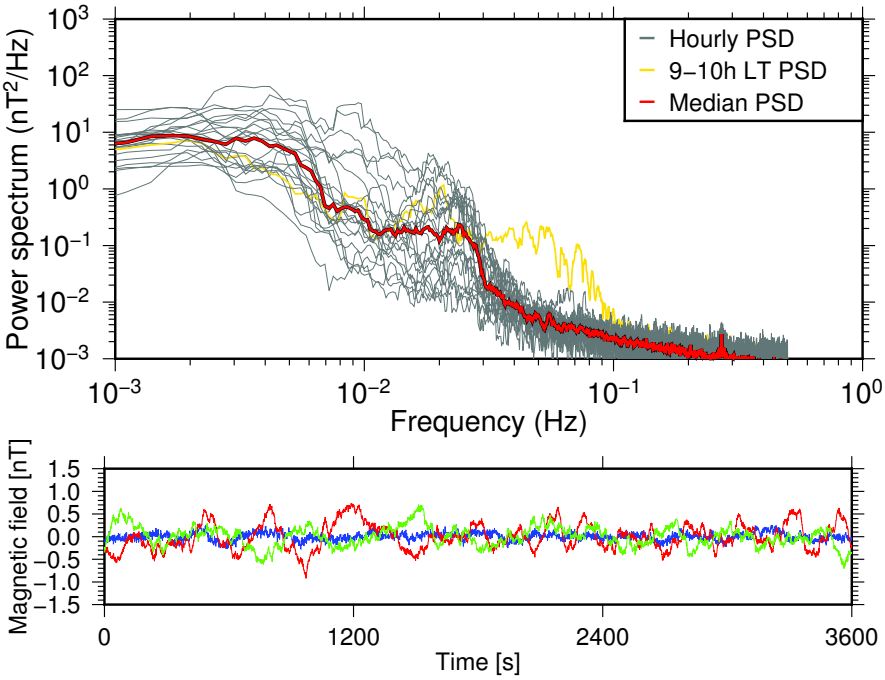


Figure 2. Top: Gray: Hourly power spectra at UPS magnetic observatory for the 05th of January 2013. Gold: Hourly power spectra for the time window 9-10h LT. Red: Median spectrum. Bottom: Example of surrogate time series. Red: X, Green: Y, Blue: Z.

2.3 Areawise thresholding

As stated by *Torrence & Compo* [1998] and later *Maraun et al.* [2007], pointwise extraction of wavelet coefficients is not enough to recover only the interesting processes. Wavelet coefficients obtained through CWT application are by definition correlated, so if a process is significant then a whole area of coefficients should also be significant. To account for this aspect, we use the areawise thresholding introduced by *Maraun et al.* [2007]. The Morlet wavelet kernel is computed at each scale a_l. The correlation area is then defined by the shape of the correlation kernel above a critical value c chosen as $c = 0.75$ in this study. Indeed, larger values result in smaller kernels and therefore meaningless areas can be kept. Lower values lead to the selection of only very large areas of large amplitude. Low amplitude continuous pulsations may therefore be missed in the detection process. From the set of thresholded wavelet coefficients, only areas where the whole correlation kernel fits in are kept for pulsation analysis (Figure 3).

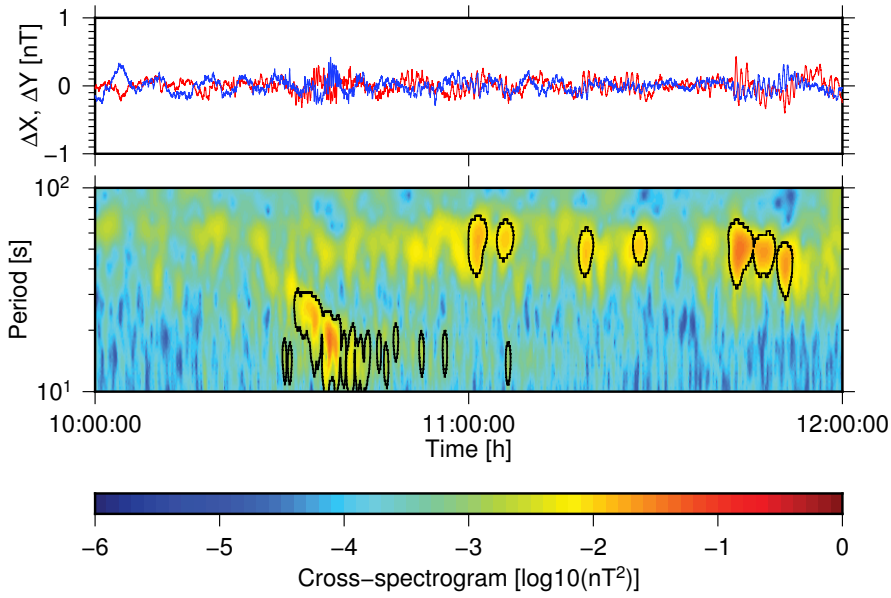


Figure 3. Top: Filtered time series of horizontal magnetic field at UPS magnetic observatory for the 05th of January 2013. Bottom: Cross-spectrogram $\mathcal{W}_{\psi}^{XYZ}(a,b)$. Areawise significant coefficients are indicated by the black contour.

2.4 Average properties

Wavelet coefficients must not be interpreted alone [*Diallo et al.*, 2006]. To describe the temporal evolution of significant coefficients, we divide the time axis into two minutes windows with an overlap of one minute between two adjacent windows. On each window, and for each frequency domain covered by Pc3, Pc4 and Pc5 pulsations, we seek the time-frequency position of the cross-spectrogram maximum. We apply on this maxima the wavelet kernel, and average the properties on significant coefficients belonging to the kernel. Circular quantities such as polarization angle and phase difference are accounted for using the methodology of circular statistics [*Fisher*, 1995]. Figure 4 presents the raw processing of the considered time series.

One minute data of pulsations frequency and polarisation attributes are then produced on this basis with a ASCII column format similar to IAGA-2002 format of magnetic observatory data. In the framework of this study, we choose not to perform any filtering on the event length or the continuity in event polarization prior to the delivering of raw data files.

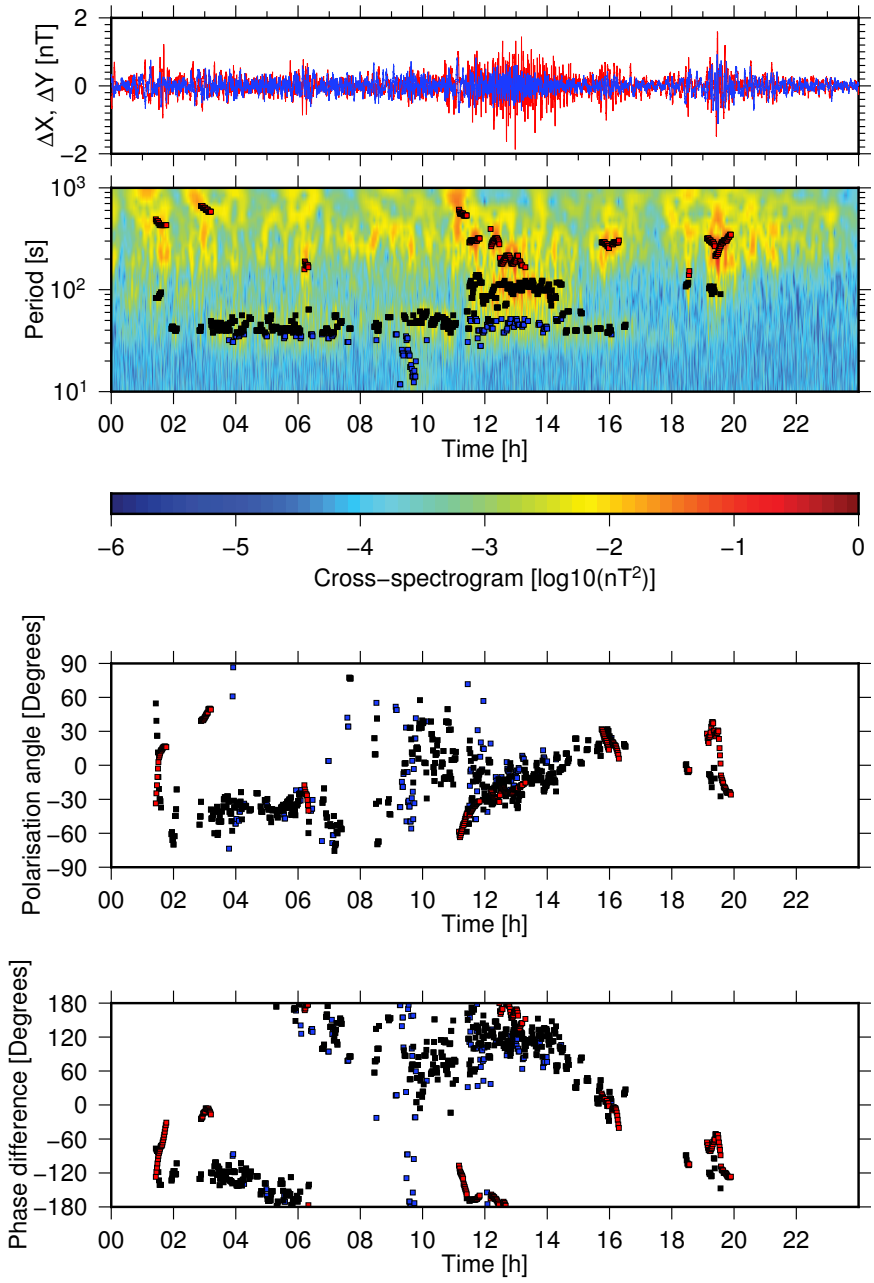


Figure 4. From top to bottom: Filtered time series of horizontal magnetic field at UPS magnetic observatory for the 05th of January 2013. Cross-spectrogram $\mathcal{W}_{\psi}^{XYZ}(a,b)$ with raw results for the detection of geomagnetic pulsations in the 10 to 1000 s range. Raw results for the horizontal polarization angle. Raw results for the phase difference.

3 Results and applications

Any automatic algorithm must be able to recover statistical properties of geomagnetic pulsations such as described in *Matsushita & Campbell* [1967] or *Saito* [1969]. In this section, we introduce two potential applications of the presented algorithm in the study of temporal and spatial variations of geomagnetic pulsations.

3.1 Temporal

We apply the algorithm on UPS and NUR magnetic observatory time series. UPS and NUR observatory are approximatively located at the same corrected geomagnetic latitude, with only a few degrees of difference in longitude. Properties of geomagnetic pulsations should then show little difference on those two observatories.

Three years of data are analyzed for the UPS observatory from January 2013 to December 2016. NUR observatory has available second data from September 2013, so nine months are missing in compared to UPS. Two analysis are performed. The first one is the calculation of the average period of Pc3 pulsations for each month along the three years. The second one is a six-months average of the pulsations characteristics from early July to the end of December 2014. No filtering has been performed on the results, so Pi pulsations are also taken into account. Results are shown Figure 5.

NUR and UPS analysis display similar patterns in the temporal variation of Pc3 characteristics. Pc3 period shows fast decrease from more than 35 seconds to less than 30 seconds after the maximum sunspot number where it stabilizes at about 30 seconds. Prior to the maximum sunspot number, Pc3 period is also relatively low compared to the 35 seconds period maximum. Hourly means of geomagnetic pulsations period also display the U shape previously described in the literature [*Saito*, 1969]. The maximum period of oscillation is at about 07:00 LT. Their occurrence is also greater in the morning than in the evening as well as the amplitude of the pulsations.

3.2 Spatial

Another application of this scheme is the monitoring of waves properties on a global scale. We apply the algorithm on NEW, FRN and MEA magnetic observatories time series. The three observatories are located at different corrected geomagnetic latitudes but

approximatively on the same corrected geomagnetic longitude. Raw results of the analysis on the 7th of January are shown Figure 6. The three observatories display similar patterns detected during the day. One known fact is the dependence of Pc5 amplitude with the geomagnetic latitude [*Matsushita & Campbell*, 1967]. Indeed, amplitude of the cross-spectrogram at MEA observatory is almost two orders of magnitude above the cross-spectrogram at FRN. Interestingly, the polarization of Pc5 pulsations is very similar between FRN and NEW observatories with increased difference with MEA. These similitudes can be detailed by filtering the raw files and performing averaging of parameters on the events to obtain an overview of the distribution of parameters.

4 Discussion and Conclusions

An automatic procedure for the detection of geomagnetic pulsations is described in this paper. The performance of this scheme is acknowledged on two examples on a temporal scale and a spatial scale demonstrating its ability to detect geomagnetis pulsationc and recover their statistical properties (typical periods, latitude, time of day and polarisation attributes).

Acknowledgments

The results presented in this paper rely on data collected at magnetic observatories. We thank the national institutes that support them and INTERMAGNET for promoting high standards of magnetic observatory practice (www.intermagnet.org). Sunspot number values were downloaded from the Sunspot Index and Long-term Solar Observations (SILSO) website (www.sidc.be/silso/home). We thank the Royal Observatory of Belgium for making the data available. All figures in this paper are plotted using GMT 5 by *Wessel et al.* [2013]. The raw results of this work are publicly available under a Creative Commons licence CCBY onto the Playground area of the ISGI Web site. ASCII Column files (in a format similar to magnetic observatory format IAGA2002) are available to all people that would be interested.

References

Bortnik, J., Cutler, J. W., Dunson, C., Bleier, T. E. (2007). An automatic wave detection algorithm applied to Pc1 pulsations. Journal of Geophysical Research: Space Physics, 112(A4).

Cazelles, B., Cazelles, K., Chavez, M. (2014). Wavelet analysis in ecology and epidemiology: impact of statistical tests. Journal of the Royal Society Interface, 11(91), 20130585.

Diallo, M. S., Kulesh, M. Holschneider, M., Scherbaum, F., Adler, F., 2006, Geophysics **71**(3), 67-77.

Fisher, N. I., 1995. Statistical analysis of circular data, Cambridge University Press, Cambridge, 277pp.

Goupillaud, P., Grossmann, A., Morlet, J., 1984, Cycle-octave and related transforms in seismic signal analysis, *Geoexploration*, **23**, 85–102.

Holschneider, M., 1995, Wavelets, an analysis tool, *Oxford Science Publications*, Oxford.

Jacobs, J. A., Kato, Y., Matsushita, S., Troitskaya, V. A. (1964). Classification of geomagnetic micropulsations. Journal of Geophysical Research, 69(1), 180-181.

Kulesh, M., Nosé, M., Holschneider, M., Yumoto, K., 2007, Polarization analysis of a Pi2 pulsation using continuous wavelet transform, Earth, Planets and Space **59**, 961–970.

Larnier, H., Sailhac, P., Chambodut, A., 2016, New application of wavelets in magnetotelluric data processing: Reducing impedance bias, Earth, Planets and Space **68**(70).

Maraun, D., Kurths, J., Holschneider, M., 2007. Nonstationary gaussian processes in wavelet domain: Synthesis, estimation, and significance testing, *Physical Review E*, **75**.

Matsushita, S., Campbell, W. H. (1967). Physics of Geomagnetic Phenomena, Volume 1. In Physics of Geomagnetic Phenomena (Vol. 1).

McPherron, R. L. (2005). Magnetic pulsations: their sources and relation to solar wind and geomagnetic activity. Surveys in Geophysics, 26(5), 545-592.

Nosé, M., Iyemori, T., Takeda, M., Kamei, T., Milling, D. K., Orr, D., Singer, H. J. Worthington, E. W., Sumitomo, N. (1998). Automated detection of Pi2 pulsations using wavelet analysis: 1. Method and an application for substorm monitoring.

Saito, T. (1969). Geomagnetic pulsations. Space Science Reviews, 10(3), 319-412.

Schreiber, T., Schmitz, A. (2000). Surrogate time series. Physica D: Nonlinear Phenomena, 142(3), 346-382.

Thomson, D. J. (1982). Spectrum estimation and harmonic analysis. Proceedings of the IEEE, 70(9), 1055-1096.

Torrence, C., Compo, G. P. (1998). A practical guide to wavelet analysis. Bulletin of the American Meteorological society, 79(1), 61-78.

Wardinski, I., & Mandea, M. (2006). Annual and semi-annual variations of the geomagnetic field components analysed by the multi-taper method. Earth, planets and space, 58(6), 785-791.

Wessel, P., Smith, W. H., Scharroo, R., Luis, J., Wobbe, F. (2013). Generic mapping tools: Improved version released. Eos, Transactions American Geophysical Union, 94(45), 409-410.

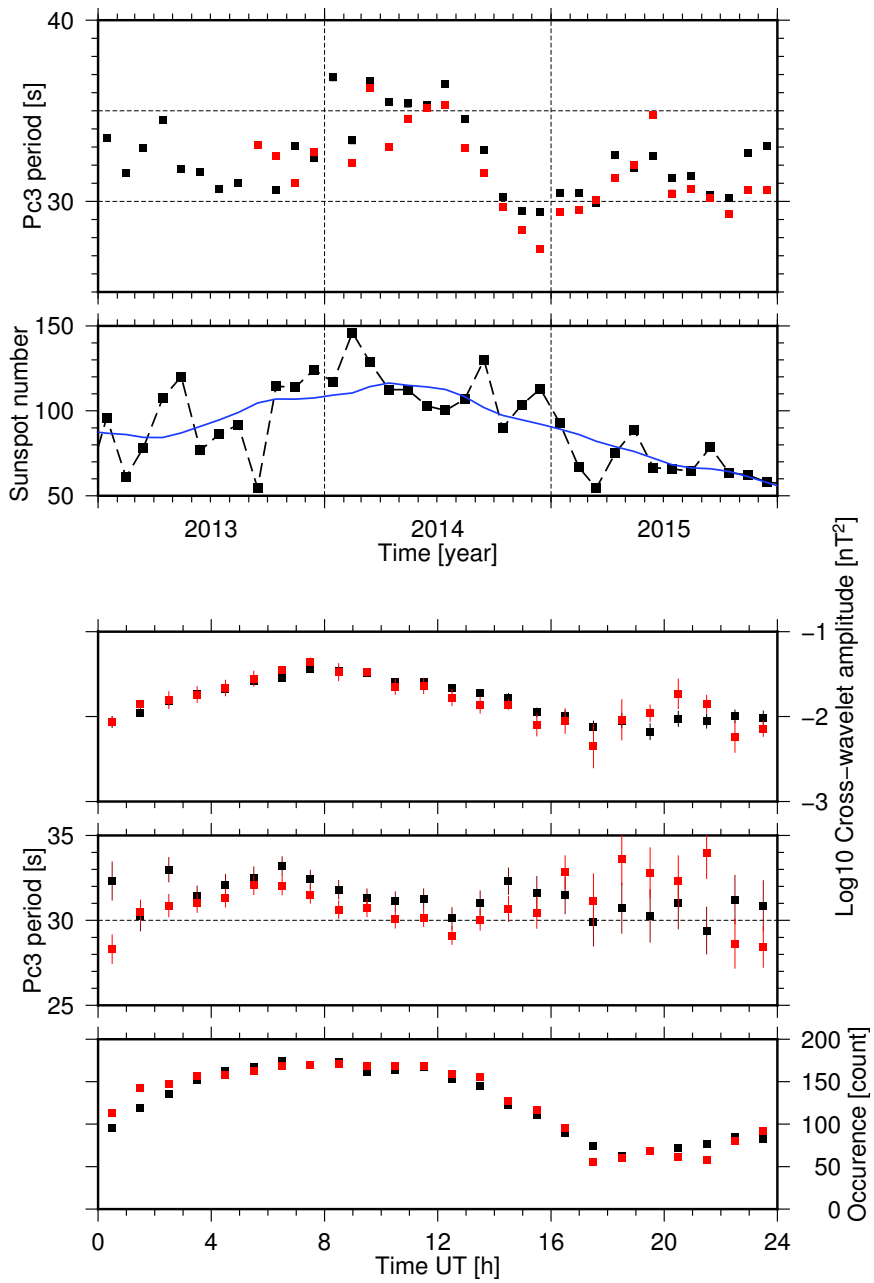


Figure 5. From top to bottom: Variation of monthly means of Pc3 period for UPS (black) and NUR (red) observatory. Variation of the monthly mean sunspot number (black) and smoothed sunspot number (blue line) (SILSO data, Royal Observatory of Belgium, Brussels). Hourly means of cross-spectrogram amplitude for UPS (black) and NUR (red) observatory. Hourly means of Pc3 periods over a six month period for UPS (black) and NUR (red) observatory. Hourly occurrence of Pc3 pulsations for UPS (black) and NUR (red) observatory.

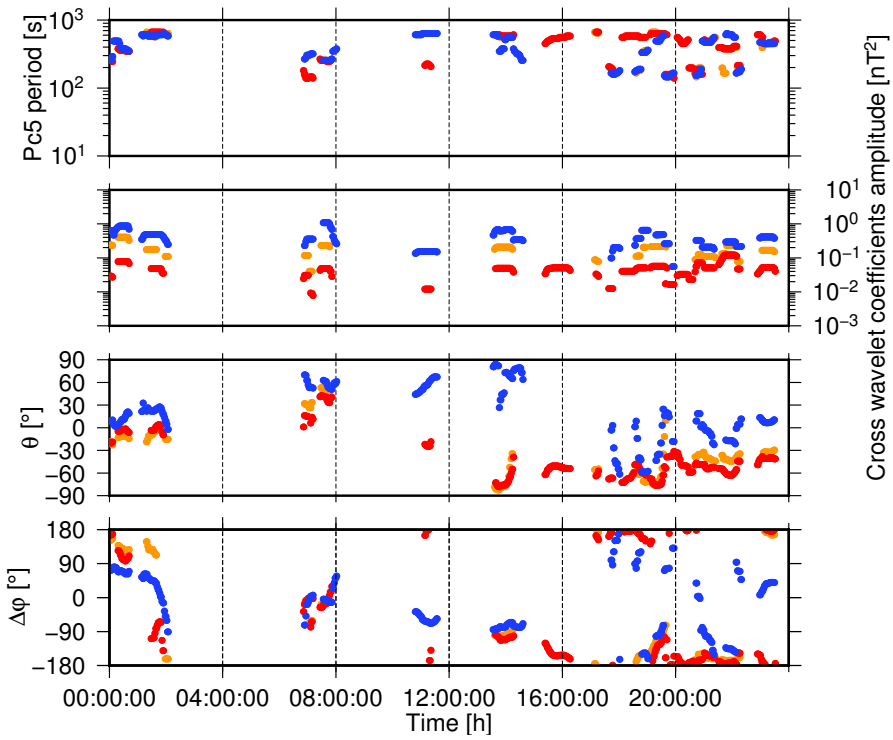


Figure 6. Raw results from the application of the algorithm on NEW (orange), FRN (red) and MEA (blue) observatory time series for the 7th of January 2013. From top to bottom: Period of detected Pc5 pulsations. Cross-spectrogram coefficients amplitude. Polarization angle in the horizontal plane. Phase difference between the horizontal components.

2.3.3 Importance des observatoires magnétiques dans la détection

La procédure de sélection d'un site d'observatoire magnétique impose une isolation la plus complète possible des signaux EM non naturels. En revanche, dans certains cas, les stations MT peuvent être installées aux abords de sources EM anthropiques pour des raisons logistiques. Ainsi, les séries temporelles sont susceptibles d'être contaminées par des signaux anthropiques ayant des manifestations diverses sur les champs EM.

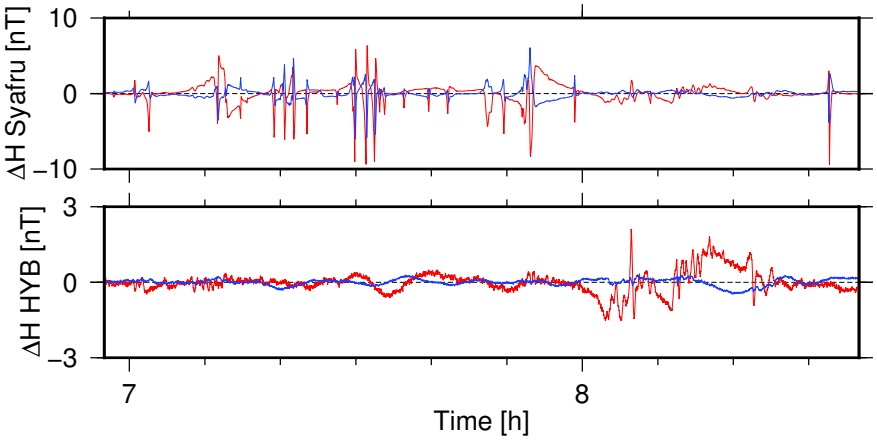


FIGURE 2.20: Haut : Séries temporelles magnétiques filtrées enregistrées à Syafru-Bensi, Népal, le 09/03/2016. Bas : Séries temporelles magnétiques filtrées simultanées à l'observatoire d'Hyderabad (HYB, Inde). Rouge : H_x, Bleu : H_y.
FIGURE 2.20: *Top : Filtered magnetic time series recorded at Syafru-Bensi, Nepal, on March 09, 2016. Bottom : Simultaneous filtered magnetic time series of Hyderabad observatory (HYB, India). Red :* H_x*. Blue :* H_y*.*

La Figure 2.20 présente des séries temporelles magnétiques enregistrées en mars 2016 (Appendice E), d'une station MT acquise à Syafru-Bensi, Népal, et de l'observatoire magnétique d'Hyderabad (HYD), Inde. Un filtre passe-haut avec une fréquence de coupure à 0.0005 Hz a été appliqué aux deux séries temporelles.

Pour des raisons logistiques, la station de Syafru-Bensi a été installée en bordure sud du village, à 50 mètres de la route principale permettant l'accès à la frontière chinoise. Le trafic routier de marchandises y était relativement important. De plus, une centrale hydroélectrique est située en contrebas du village, à 500 mètres de la position de la station MT. Les séries temporelles MT (Figure 2.20, haut) présentent donc un niveau de bruit important avec des transitoires dont l'amplitude peut atteindre des valeurs supérieures à 10 nT. Ces types de transitoires ont un contenu fréquentiel large et vont donc impacter largement le spectre de puissance des séries magnétiques. La procédure décrite Section 2.3.2, appliquée aux séries temporelles magnétiques MT, va donc être perturbée par ces signaux anthropiques. La détection correcte des signaux naturels est donc périlleuse avec les seules données de la station MT.

Sur la même figure, bas, nous avons représenté les séries temporelles de l'observatoire magnétique d'Hyderabad aux mêmes instants. Une analyse de cohérence entre les stations du Népal et cet observatoire indien est effectuée Section 5.3. Les signaux naturels sont mis en évidence sur les deux séries temporelles, mais seul le champ magnétique enregistré à l'observatoire HYD permet une analyse statistique fine des variations du champ magnétique naturel telle que présentée précédemment Section 2.3.2 pour permettre la détection des oscillations naturelles du champ magnétique nécessaire à l'interprétation MT.

Les observatoires magnétiques sont donc critiques dans l'analyse des sources MT basses fréquences.

2.4 Conclusion

Dans ce chapitre, nous avons étudié la variabilité des sources géomagnétiques et atmosphériques observées dans les séries MT aux périodes comprises entre 0.0001 et 1000 secondes. Ces sources sont caractérisées par un fort rapport signal sur bruit et peuvent être détectées en se basant sur leurs caractéristiques.

Les sources hautes fréquences, de par leur caractère transitoire, sont aisément détectables par de simples critères de rapport signal-sur-bruit, et l'analyse de plusieurs stations simultanément permet de discriminer les sources proches des sources lointaines.

Les sources basses fréquences présentent un caractère oscillatoire et entretenu qui ne permet pas leur détection par une simple analyse de rapport signal sur bruit. Pour certaines sources, telles les pulsations continues, la question reste ouverte de savoir si les oscillations sont présentes de façon permanente mais avec une amplitude inférieure à celles des capteurs actuels ou bien si elles présentent effectivement des périodes d'amplification et d'inhibition .

Dans cette optique, l'utilisation de séries temporelles magnétiques affranchies de tout signal anthropique est primordiale pour une analyse correcte des variations du champ magnétique naturel. Certaines séries temporelles MT, acquises lors de campagnes de mesures, présentent à posteriori lors du traitement en laboratoire de fortes perturbations anthropiques et ne permettent donc pas une caractérisation correcte de l'activité géomagnétique ou atmosphérique. Le traitement MT de tels signaux nécessite l'usage de stations de références.

Ces procédures de détection seront appliquées au chapitre suivant dans le calcul des fonctions de réponses MT.

Chapitre 3

Calcul des fonctions de transfert

3.1 Introduction

Dans cette partie, nous introduisons de nouveaux algorithmes de traitement de données MT. Nous comparerons les résultats obtenus par nos méthodes à celles obtenues par les traitements classiques. Nous nous attacherons à décrire les grands principes des algorithmes. Pour une description détaillée des principes de base, nous invitons le lecteur à se plonger dans le chapitre 5 de A. D. Chave et A. Jones (2012).

Le traitement de données MT cherche à déterminer les fonctions de réponse magnétotelluriques définies par :

$$\mathbf{E} = \mathbf{Z} \times \mathbf{B}. \tag{3.1}$$

L'impédance MT étant définie en domaine fréquentiel, c'est tout naturellement que les principaux algorithmes de traitement MT se basent sur la transformée de Fourier. L'amélioration des techniques de traitement a été graduelle des années 1960 jusqu'à aujourd'hui. Les années 60 et 70 ont consisté en la recherche de méthodes simples pour réduire les biais sur l'estimation MT (Sims, F. X. J. Bostick et H. W. Smith, 1971). La redécouverte de la technique de *remote reference* par Gamble, Goubeau et J. Clarke (1979) a permis une amélioration drastique des estimations MT par rapport aux estimations précédentes. Les années 80 ont permis l'émergence de l'application des méthodes utilisant les statistiques dites "robustes" dans l'estimation des réponses MT (A. G. Jones et al., 1989). Depuis, les améliorations ont été plus fines, avec l'utilisation des réseaux de stations MT dans l'analyse des sources (G. D. Egbert, 1997), dans l'amélioration des statistiques robustes (Alan D. Chave et David J. Thomson, 2004), ou enfin avec l'introduction d'un changement drastique sur le modèle statistique appliqué aux résidus du champ électrique (Alan D Chave, 2014).

Peu de traitements ont été développés utilisant l'information sur les sources MT afin de déterminer les réponses MT. Parmi ceux-ci, TZANIS et BEAMISH (1987) et Y. ZHANG et PAULSON (1997) ont développé des algorithmes utilisant les sources hautes fréquences.

Dans les sections 3.2 et 3.3, nous décrirons les principes généraux des statistiques robustes utilisées par les principaux algorithmes de traitement de données MT développés dans la communauté MT. Dans la section 3.4, nous développerons la nouvelle méthodologie implémentée dans ce travail qui utilise la détection des événements géomagnétiques présentée au chapitre précédent.

3.2 Statistiques robustes

Dans cette section, nous considérons que les séries temporelles du champ magnétique ont été acquises sans bruit. La nature des signaux est non stationnaire, mais il ne s'agit que d'ondes planes, sans perturbations par des ondes issues de sources proches.

3.2.1 Segmentation des séries temporelles

Dans le traitement MT, nous disposons donc de 4 séries temporelles (e_x, e_y, h_x, h_y) pour la station locale.

Dans les traitements conventionnels, la première étape du traitement consiste en la segmentation des séries temporelles en N segments avec chevauchement. La taille D du segment est variable selon le choix de l'utilisateur du code, mais est toujours proportionnelle à la période T considérée. Ainsi, les segments correspondant aux longues périodes contiennent plus d'échantillons que les segments correspondants aux courtes périodes. Pour chaque segment, un fenêtrage est appliquée afin de limiter le biais des basses fréquences sur les estimations des transformées de Fourier. Un grand nombre de fenêtrage est possible : les fenêtres de Hamming, Blackmann... Parmi ceux-ci, la classe la plus utilisée en MT est la famille des séquences de Slepian (SLEPIAN, 1978 ; David J THOMSON, 1982 ; Alan D. CHAVE et David J. THOMSON, 1987).

Après application du fenêtrage, une transformée de Fourier discrète est appliquée sur chaque section. Les données sont donc N valeurs de E et H à une fréquence particulière.

3.2.2 Régression linéaire et moindres carrés

Considérons que les signaux électriques et magnétiques ont été segmentées en N sections de taille égale et dont les effets d'aliasing et de biais basse fréquence ont été réduits par l'application d'une fenêtre d'apodisation. La transformée de Fourier est appliquée sur chaque segment et la

fréquence f_k est considérée par la suite. Le problème MT consiste en la recherche à la fréquence f_k d'une solution à la régression linéaire :

$$\mathbf{E} = \mathbf{B} \times \mathbf{Z} + \varepsilon, \tag{3.2}$$

avec $\mathbf{E}$ un vecteur de taille $(N, 1)$, $\mathbf{Z}$ une matrice de taille $(2, 1)$ et $\mathbf{B}$ un vecteur de taille $(N, 2)$. E contient les transformées de Fourier du champ électrique (E_x ou E_y), H les transformées de Fourier du champ magnétique (B_x et B_y), et Z les fonctions de transfert recherchées. ε est un vecteur de taille $(N, 1)$ représentant les erreurs de mesure sur E. $\times$ représente le produit matriciel.

Jusqu'à l'introduction des statistiques robustes, les méthodes utilisées pour déterminer Z se basaient sur la simple application des moindres carrés à la régression (3.2). Les moindres carrés permettent de déterminer le modèle minimisant les résidus $\mathbf{r}$ entre le champ électrique mesuré $\mathbf{E}$ et le champ électrique modélisé $\mathbf{B} \times \mathbf{Z}$:

$$\sum_{i=1}^{N} r_i^2 = \sum_{i=1}^{N} |E_i - B_i Z|^2 \tag{3.3}$$

Le modèle des moindres carrés est juste dans le cas où les erreurs ε sont indépendamment et identiquement distribuées selon un modèle gaussien (G. D. Egbert et J. R. Booker, 1986). En pratique, les erreurs suivent rarement ce modèle et les moindres carrés ne permettent pas de déterminer les fonctions de réponse magnétotelluriques (G. D. Egbert et J. R. Booker, 1986 ; Alan D. Chave et David J. Thomson, 1989). Pour donner un exemple, nous représentons à la Figure 3.1, une illustration de l'échec des moindres carrés à récupérer la tendance d'une droite linéaire définie par $e = 0.5b$. Les points ne suivant pas la loi $e = 0.5b$, appelés *outliers*, perturbent l'estimation des moindres carrés et empêchent de déterminer correctement la pente de la droite.

L'écart à un modèle gaussien peut venir de plusieurs origines, comme une erreur de mesure ou un bruit local ne correspondant pas au processus d'induction tel que défini par le modèle MT (Alan D. Chave et David J. Thomson, 1987). Dans ces conditions, le modèle des moindres carrés ne permet pas de déterminer avec précision Z.

3.2.3 Statistiques robustes

Pour accommoder ces écarts, les statistiques robustes ont été appliquées en magnétotellurique. Ces statistiques robustes sont définies comme permettant d'obtenir des estimations en présence de points aberrants appelés *outliers*.

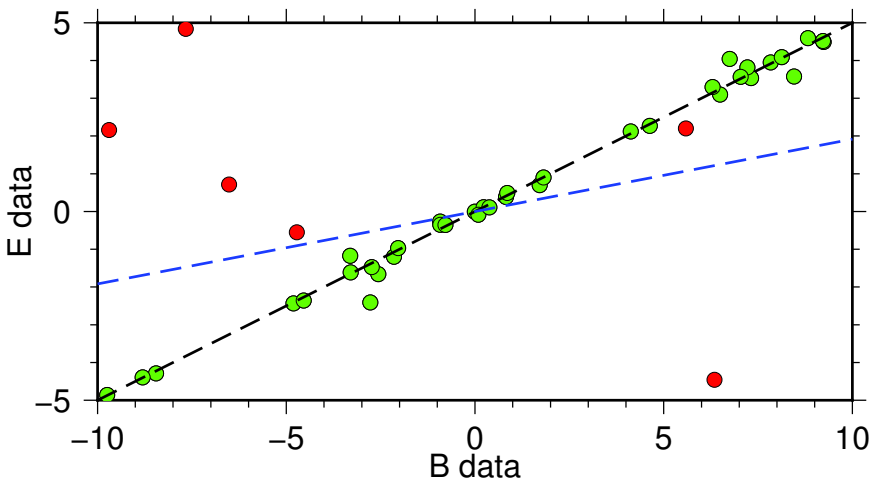


FIGURE 3.1: Exemple de problème de régression linéaire avec des points dont les résidus ne vérifient pas l'hypothèse de distribution gaussienne. Les points verts vérifient : $e = 0.5b + \mathcal{N}(0, 0.1)$. Les points rouges suivent une distribution de Cauchy de paramètres de position 0 et d'échelle 0.1. La courbe en pointillés noirs est le modèle $e = 0.5b$, celle en pointillés bleus le modèle obtenu par les moindres carrés.

FIGURE 3.1: *Example of linear regression where residuals do not follow gaussian distribution. Green dots follow* $e = 0.5b + \mathcal{N}(0, 0.1)$. *Red dots follow Cauchy distribution with location and scale parameters at 0 and 0.1 respectively. Dotted black line is* $e = 0.5b$ *model, dashed blue line is model obtained through least-squares inversion.*

3.2.3.1 Espérance

La robustesse d'un estimateur se détermine par rapport à la quantité maximale de données pouvant être modifiée dans un ensemble sans perturber l'espérance faite par l'estimateur. Deux estimateurs pouvant être utilisés pour calculer l'espérance d'un ensemble sont la médiane $\tilde{x}$ et la moyenne $\bar{x}$. La robustesse de la médiane est de $1/2 - 1/N$ si N est impair, et $1/2 - 1/2N$ si N est pair indiquant que l'espérance est perturbée fortement si la part d'outliers dépasse cette proportion dans l'échantillon. La moyenne a en revanche une robustesse de 0, il suffit d'un seul outlier peut modifier fortement l'espérance. Parmi les estimateurs, on trouve plusieurs classes, parmi celles-ci, les estimateurs de type L (pour *linear*) ou M (pour *Maximum Likelihood Type*).

Les estimateurs de type L se basent sur une combinaison linéaire des déciles. La médiane est un exemple d'estimateur de l'espérance de type L (5e décile), comme le *midhinge* (moyenne des 3eme et 7eme déciles). L'intérêt des estimateurs de type L est leur facilité de calcul, leur robustesse et leur capacité à être visualisés rapidement grâce à des graphes de types boites à moustache. En revanche, leur utilisation dans les problèmes de régression pose problème car ils ne permettent pas l'adaptation rapide à des jeux de données différents.

Les estimateurs de type M se basent sur la minimisation d'une fonction objectif ρ. Ils sont une généralisation de la méthode du maximum de vraisemblance (*Maximum Likelihood Estimation*,

MLE) (Huber, 1981). Soit θ le paramètre recherché d'une loi de probabilité et $(X_1, X_2, ..., X_n)$ un ensemble de variables aléatoires de densité de probabilité continue f. Soit $(x_1, x_2, ..., x_n)$ des échantillons tirés pour chaque variable aléatoire X_i. En supposant que les variables sont identiquement et indépendamment distribuées, la fonction de vraisemblance s'écrit :

$$L(\theta) = \prod_{i=1}^{n} f(x_i|\theta) \tag{3.4}$$

Au lieu de maximiser L, principe de la méthode MLE, les estimateurs de type M minimisent une fonction coût ρ ou annulent sa dérivée, appelée fonction d'influence ψ, définies par :

$$\begin{aligned} &\hat{\theta} : \min(\sum_{i=1}^{n} \rho(x_i|\theta)), \qquad \text{où } \rho(x_i|\theta) = -\log(f(x_i|\theta)) \\ &\psi(x) = \frac{\partial \rho}{\partial \theta}(x) \end{aligned} \tag{3.5}$$

Il est impossible de savoir à priori quel paramètre θ permet de minimiser la fonction ρ. Le principe des estimateurs de type M se base sur une procédure de régression (appelée *Iterative Reweighted Least Squares*) en plusieurs itérations permettant de déterminer le meilleur paramètre θ à partir de la population $(x_1, x_2, ..., x_n)$.

Cette méthode utilise une solution initiale θ^0, par exemple obtenue à partir de l'application des moindres carrés (e.g. équation 3.3), ou d'une solution de norme L_1 dans le cas de données extrêmement bruitées (Alan D. Chave et David J. Thomson, 2004). A partir de cette solution initiale, les résidus $\mathbf{r^0}$ sont calculés entre la population $(x_1, x_2, ..., x_n)$ et le modèle obtenu utilisant θ^0. Dans le cas du problème MT, les résidus $\mathbf{r^0}$ sont obtenus par $r^0 = e - \mathbf{hz^0}$.

Le problème de l'estimation se pose par la dispersion des données autour du paramètre recherché θ, à savoir déterminer quels résidus sont à considérer comme significatifs. θ est invariant par translation $(X_i \to X_i + a \Rightarrow \theta \to \theta + a)$, mais pas par changement d'échelle $(X_i \to bX_i \not\Rightarrow \theta \to b\theta)$. Les résidus doivent donc être pondérés par une échelle d qui va déterminer à partir de quelle proportion les échantillons sont considérés comme des points aberrants dans la régression :

$$x_i = \frac{r_i}{d} \tag{3.6}$$

L'échelle doit être choisie robuste pour que l'ensemble de l'algorithme soit robuste.

Le choix de la fonction ρ et de l'échelle est déterminée par le modèle désiré pour modéliser les résidus obtenus à partir des données à disposition. Parmi les modèles les plus couramment utilisés, les modèles pour les normes L_1 et L_2, les moindres carrés, utilisent les fonctions objectifs :

$$\begin{aligned} \rho(x) &= \frac{-x^2}{2} \\ \rho(x) &= |x| \end{aligned} \tag{3.7}$$

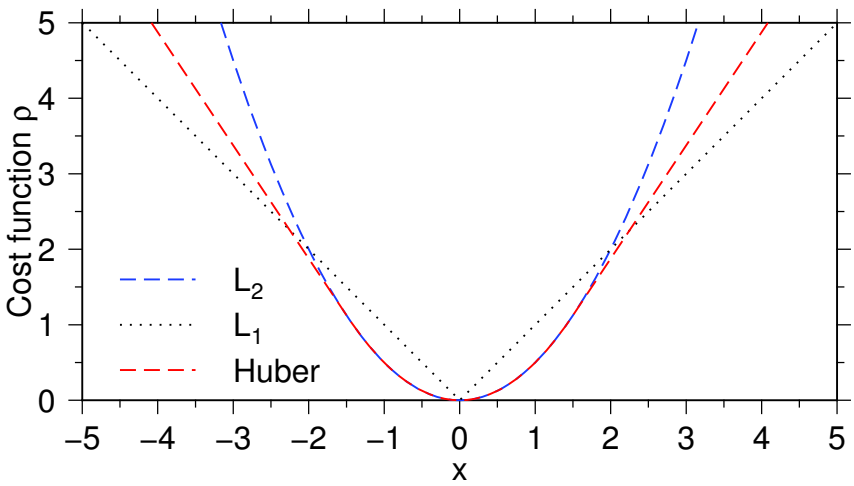


FIGURE 3.2: Fonction de coût pour les modèles de moindres carrés (trait plein), norme L_1 (trait pointillé), et de Huber (trait discontinu).
FIGURE 3.2: *Cost function for least-squares model (full line), L_1 norm (dotted line) and Huber model (dashed line).*

Le modèle utilisé en MT pour les résidus jusqu'à Alan D CHAVE (2014) est celui d'un noyau gaussien contaminé par des valeurs déviant de ce modèle. Pour accommoder cette distributions de résidus, un estimateur de type M adapté est celui basé sur les poids de HUBER (1981). Les fonctions objectifs sont définies par :

$$\rho(x) = \begin{cases} -x^2/2 & \text{if } |x| \leqslant \gamma \\ -\gamma|x| - \gamma^2/2 & \text{if } |x| > \gamma \end{cases} \tag{3.8}$$

Ainsi, le modèle sous-jacent contient un noyau de données gaussien défini par la fonction objectif associée aux moindres carrés, et un modèle en double exponentielle aux extrémités. La continuité est assurée par le paramètre $\gamma^2/2$. Le paramètre γ permet d'ajuster la robustesse de la fonction objective par rapport au problème étudié. De nombreuses références conseillent l'utilisation du paramètre γ égal à 1.5. Pour le choix de l'échelle déterminant à partir de quelle distance les résidus sont considérés importants, Alan D. CHAVE et David J. THOMSON (1989) et Alan D. CHAVE et David J. THOMSON (2004) ont choisi la statistique d'ordre MAD définie par :

$$MAD(X) = \text{median}(X_i - \text{median}(X_i)) \tag{3.9}$$

Le MAD est pondérée par une valeur numérique pour qu'elle soit cohérente dans le cas d'une distribution supposée. Cette distribution supposée est la distribution de Rayleigh (appropriée pour la distribution du module d'une quantité complexe). La valeur du MAD est donc pondérée par 0.44845 (Alan D. CHAVE et David J. THOMSON, 1989).

Une fois que le modèle décrivant les résidus est choisi, chaque élément x_i de la population se

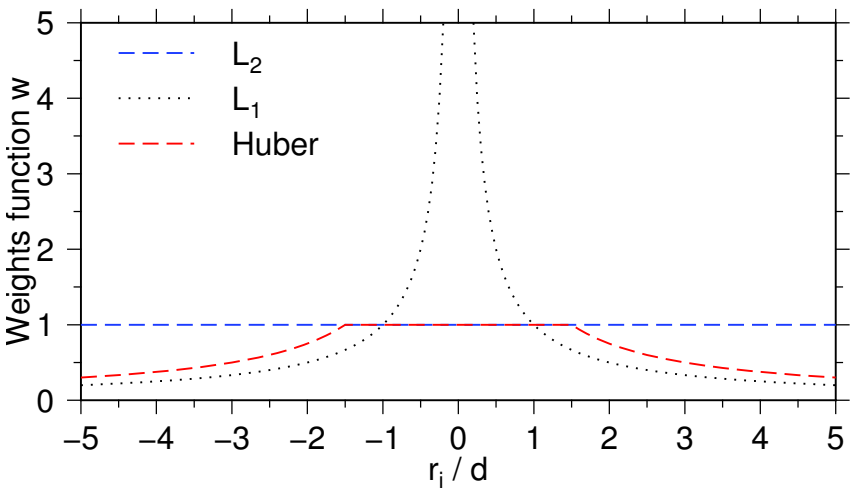


FIGURE 3.3: Poids pour les modèles de moindres carrés (trait pointillé bleu), norme L_1 (trait pointillé), et de Huber (trait pointillé rouge).
FIGURE 3.3: *Weights for least-squares model (full line), L_1 norm (dotted line) and Huber model (dashed line).*

voit attribuer un poids dont la forme dépend de la fonction objectif ρ. Dans le cas de la fonction objectif de Huber, les poids sont définis par :

$$w_i = \begin{cases} 1 & \text{if } |r_i/d| \leqslant \gamma \\ \gamma/|r_i/d| & \text{if } |r_i/d| > \gamma \end{cases} \tag{3.10}$$

A chaque itération k, la solution $\mathbf{z^k}$ et les poids attribués w^k sont modifiés. La solution finale est obtenue quand la quantité rw^kr ne change que peu entre deux itérations successives. La fonction objectif définie par HUBER (1981) étant convexe (Figure 3.2), la convergence est nécessairement atteinte au bout d'un certain nombre d'itérations. A la Figure 3.4 est représenté l'application de la méthode de régression robuste sur le même cas que présenté à la Figure 3.1.

Les poids définis par l'équation (3.10) ne peuvent être nuls (Figure 3.3) et ainsi éliminer les points aberrants qui passeraient quand même la première régression robuste, (Alan D. CHAVE et David J. THOMSON, 1989) ont appliqué une deuxième régression robuste dont les poids sont définis par :

$$v_i = \exp(-e^{-\xi(|x_i|-\xi)}) \tag{3.11}$$

Le paramètre ξ définit la taille des résidus à partir desquels les poids commencent à perdre de l'influence dans le calcul de l'estimation finale.

D'autres raffinements tels que les estimateurs *bounded influence* (Alan D. CHAVE et David J. THOMSON, 2004) sont utilisés dans la régression robuste comme l'ajout d'une régression pour éliminer les points extrêmes dans le champ magnétique. Alan D CHAVE (2014) a modifié le

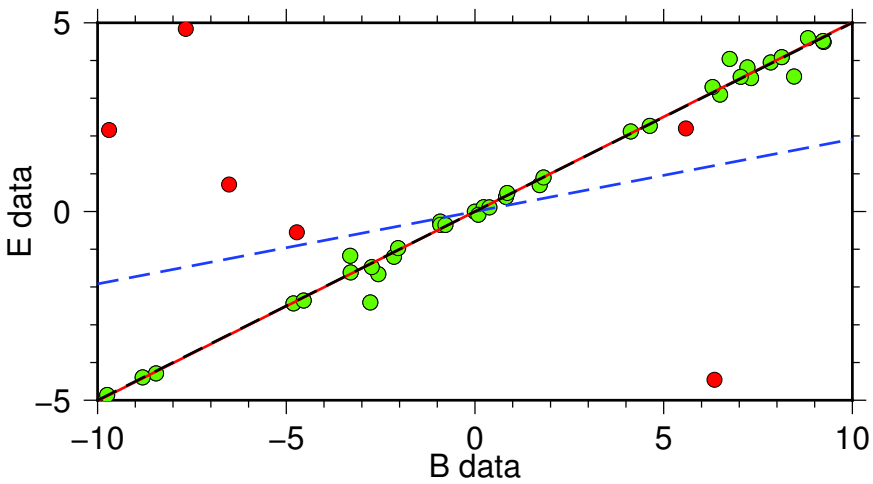


FIGURE 3.4: Exemple de problème de régression linéaire avec des points dont les résidus ne vérifient pas l'hypothèse de distribution gaussienne. Les points verts vérifient : $e = 0.5b + \mathcal{N}(0, 0.1)$. Les points rouges suivent une distribution de Cauchy de paramètres de position 0 et d'échelle 0.1. La courbe en pointillés noirs est le modèle $e = 0.5b$, celle en pointillés bleus le modèle obtenu par les moindres carrés, et celle en pointillés rouge est celle obtenue par la procédure de moindres carrés pondérés par les poids de Huber.
FIGURE 3.4: *Example of linear regression where residuals do not follow gaussian distribution. Green dots follow* $e = 0.5b + \mathcal{N}(0, 0.1)$. *Red dots follow Cauchy distribution with location and scale parameters at 0 and 0.1 respectively. Dotted black line is* $e = 0.5b$ *model, dashed blue line is the model obtained through least-squares inversion. Red dashed line is the model obtained through application of iterative least squares regression with Huber weights.*

modèle statistique appliqué aux résidus pour y intégrer les résidus importants dans un modèle de distribution Lévy stable. Cette modification permet d'expliquer l'ensemble des résidus sans avoir recours à un modèle de distribution gaussienne tronqué tels que défini par les poids de Huber.

3.2.3.2 Incertitude sur la valeur des fonctions de réponse

Dans cette section, nous présentons l'estimateur non-paramétrique *jackknife* auquel nous avons comparé nos résultats. Soit $\hat{\theta}$, l'impédance obtenue en appliquant l'estimateur de type M à l'ensemble des données $(x_1, x_2, ..., x_N)$. L'estimation *delete-one* $\theta_{\backslash n}$ est obtenue en appliquant l'estimateur aux données $(x_1, x_2, ..., x_{n-1}, x_{n+1}, ..., x_N)$. Pour éviter de recalculer les poids w_n, ceux obtenus pour l'estimation de z peuvent être réutilisés pour obtenir une approximation de la solution $\theta_{\backslash n}$. L'estimateur de la variance *jackknife* sur le paramètre θ est obtenu par :

$$\tilde{\theta} = \frac{1}{N} \sum_{i=1}^{N} (\theta_{\backslash i} - \bar{\theta})^2, \tag{3.12}$$

avec $\bar{\theta}$ la moyenne des estimations jackknife $\theta_{\backslash i}$.

Le calcul des estimations *jackknife* est rapide, même pour de grands volumes de données tels qu'à disposition dans le traitement de données MT. En revanche, le problème devient plus compliqué car z est un vecteur, la variance *jackknife* devient ainsi une matrice de covariance. Alan D. Chave et David J. Thomson (1989) ont aussi rappelé que la variance classique était biaisée car le problème de la régression MT utilise deux variables d'entrée, h_x et h_y. Ils ont ainsi proposé de pondérer l'estimation de la variance *jackknife* en MT par l'utilisation de la matrice de projection H définie par :

$$\mathbf{H} = \mathbf{b} \cdot (\mathbf{b}^{\dagger} \cdot \mathbf{b})^{-1} \cdot \mathbf{b}^{\dagger}. \tag{3.13}$$

L'estimation jackknife devient ainsi :

$$\sigma^2 = \frac{N}{N-2} \sum_{i=1}^{N} (1 - H_{ii})(\hat{\mathbf{z}} - \hat{\mathbf{z}}_{\backslash i})(\hat{\mathbf{z}} - \hat{\mathbf{z}}_{\backslash i})^{\dagger} \tag{3.14}$$

Pour le calcul des estimations de la variance *jackknife* sur la résistivité apparente et la phase, le calcul est plus délicat car la distribution de la résistivité apparente n'est pas gaussienne et la circularité de la phase doit être prise en compte. Pour la résistivité apparente, il est nécessaire de passer par une transformation permettant de stabiliser la distribution pour en calculer une variance (Rice, 2006). La variance *jackknife* doit être ainsi calculée sur le logarithme de la résistivité apparente , et repassée à la fin du calcul dans le domaine de résistivité. L'autre solution utilisée par Alan D. Chave et David J. Thomson (1989) est d'utiliser les variances obtenues sur les fonctions de réponse $\mathbf{z}$ pour directement les convertir par développement au premier ordre. Les erreurs obtenues sont :

$$\delta\rho_{ij} = 2\mu_0 |z_{ij}| \delta z_{ij} / \omega \tag{3.15}$$

$$\delta\phi_{ij} = sin^{-1}(\delta z_{ij} / |z_{ij}|) \tag{3.16}$$

Alan D Chave et Lezaeta (2007) ont montré que ces erreurs fournissaient une bonne approximation des erreurs *jackknife* sur la résistivité apparente et la phase et permettaient de s'affranchir des calculs *jackknifes* sur ceux-ci.

3.3 Utilisation d'une station de référence

Dans l'ensemble des méthodes présentées précédemment, la détermination des estimations sur les fonctions de réponse utilise le calcul de la quantité $(\mathbf{b}^{\dagger} \cdot \mathbf{b})^{-1}$. Cette quantité, en présence de bruit sur le champ magnétique, biaise l'impédance en la diminuant. En effet, en considérant $\mathbf{b} = \mathbf{b_0} + \varepsilon$, avec $\mathbf{b_0}$ le champ magnétique non bruité et ε le bruit sur le champ magnétique, on

observe :

$$\begin{aligned} \mathbf{b}^\dagger \cdot \mathbf{b} &= (\mathbf{b_0} + \varepsilon)^\dagger (\mathbf{b_0} + \varepsilon) && (3.17) \\ &= \mathbf{b_0}^\dagger \mathbf{b_0} + \varepsilon^\dagger \mathbf{b_0} + \mathbf{b_0}^\dagger \varepsilon + \varepsilon^\dagger \varepsilon && (3.18) \\ & && (3.19) \end{aligned}$$

En supposant que le bruit est non corrélé avec le champ magnétique, alors $\varepsilon^\dagger \mathbf{b_0} = \mathbf{b_0}^\dagger \varepsilon = 0$. Ainsi :

$$\mathbf{b}^\dagger \cdot \mathbf{b} = \mathbf{b_0}^\dagger \cdot \mathbf{b_0} + \varepsilon^\dagger \varepsilon \tag{3.20}$$

Ainsi, dans le calcul de l'impédance MT, le bruit sur le champ magnétique produit une sous-estimation de z. A l'ensemble des méthodes présentées précédemment, il est donc nécessaire d'inclure la station de référence pour réduire le biais sur l'impédance MT. Cette méthode a été introduite par Gamble, Goubeau et J. Clarke (1979) et a permis une réduction significative du biais sur z. Le calcul de la solution initiale par moindres carrés dans la régression robuste s'effectue avec :

$$\mathbf{z}_1^r = (\mathbf{b^r} \times \mathbf{b}^\dagger)^{-1} \times (\mathbf{b^r} \times \mathbf{e}) \tag{3.21}$$

Les résidus $\mathbf{r}$ sur le champ électrique sont quand à eux toujours calculés en utilisant le champ magnétique local.

3.4 Utilisation de l'analyse en ondelette

3.4.1 Inversion des coefficients

Y. Zhang et Paulson (1997) ont pour la première fois utilisé les coefficients en ondelettes pour la détermination des fonctions de réponse MT. Une des remarques importantes introduites par les auteurs est l'utilisation d'une ondelette dont la résolution fréquentielle est suffisamment précise pour permettre une détermination précise de Z. Comme précisé au chapitre 1, le processus physique sous-jacent est un processus de diffusion donc l'impédance ne varie que peu sur un intervalle de fréquence restreint. Si l'ondelette a une résolution fréquentielle faible, alors les coefficients d'ondelettes peuvent intégrer une variation de l'impédance. Si la résolution fréquentielle est suffisamment forte, l'impédance peut être considérée constante sur l'intervalle de fréquence définie par l'ondelette. Y. Zhang et Paulson (1997) ont ainsi postulé que l'énergie de l'ondelette doit être contenue dans un intervalle de fréquences défini par l'équation (3.22), où f^+ et f^- sont respectivement les bornes supérieures et inférieures de l'intervalle de fréquence.

$$\log_{10}(f^+) - \log_{10}(f^-) < 0.2 \tag{3.22}$$

Cette condition, entièrement empirique, permet d'écrire le système d'équations MT dans le domaine temps-fréquence :

$$\begin{aligned} e_x(a,b) &= Z_{xx}(a)h_x(a,b) + Z_{xy}(a)h_y(a,b) \\ e_y(a,b) &= Z_{yx}(a)h_x(a,b) + Z_{yy}(a)h_y(a,b) \end{aligned} \tag{3.23}$$

Sous cette forme, le système s'écrit comme un système linéaire et les méthodes de régression classiques comme les moindres carrés ou les régressions robustes décrites précédemment peuvent s'appliquer.

3.4.2 Méthodologie

Nous avons introduit deux méthodologies d'inversion des coefficients d'ondelettes afin de déterminer les fonctions de réponse MT. La première méthode développée est basée sur une inversion de coefficients correspondant à des combinaisons d'événements. La seconde méthode est basée sur le bootstrap.

3.4.2.1 Stockage des coefficients

De la même façon que pour le calcul des attributs de polarisation des événements géomagnétiques (voir chapitre précédent), nous utilisons le noyau de corrélation pour attribuer à chaque événement un ensemble de coefficients. Les notations pour la suite de ce chapitre sont les suivantes. Après la détection, nous disposons donc d'un ensemble d'événements, chacun caractérisé par un groupe de coefficients (vide ou non) à chaque échelle.

3.4.2.2 Utilisation de combinaisons d'événements

Le système MT, tel que défini par l'équation 3.1, est constitué de deux équations qui peuvent être résolues séparément (e_x et e_y). Ces équations contiennent deux inconnues donc sont des problèmes sous-dimensionnés par nature. Il est nécessaire de disposer d'au moins deux couples différents de h_x et h_y pour permettre de déterminer une solution. Notre première méthode d'estimation s'est donc basée sur le calcul simple d'estimations de la fonction de réponse MT par régression robuste de coefficients appartenant à des paires d'événements. La régression robuste appliquée utilise les poids de Huber tels que définis par Alan D. Chave et David J. Thomson (1989).

3.4.2.3 Article EPS - Réduction du biais sur l'impédance en station simple

Résumé en français Les données MT consistent en la somme de sources naturelles, transitoires ou quasi-périodiques, et de diverses sources de bruit (instrumental ou d'origine humaine), dont la nature doit être prise en compte dans les traitements MT. La majeure partie des procédures de traitement se basent sur la transformée de Fourier et sur les statistiques robustes pour déterminer les fonctions de réponses MT. La plupart n'utilisent pas l'information sur les sources MT. De plus, la procédure principale pour réduire l'influence du bruit sur le calcul de l'impédance MT est l'installation d'une station de référence est nécessaire. Cette deuxième acquisition, même installée avec soin, peut néanmoins échouer et ne pas permettre la réduction du biais MT.

Nous proposons ici d'utiliser la transformée en ondelettes continues appliquées aux séries temporelles MT pour réduire l'influence du bruit sur les estimations des fonctions de réponses MT sans station de référence. Nous utilisons deux ondelettes différentes, appelées Cauchy et Morlet, adaptées à la forme temporelle des événements géomagnétiques observés. Nous montrons, en utilisant les coefficients d'ondelettes d'événements localisés, que cette procédure permet de retrouver sans station de référence, les fonctions de réponse MT non biaisées obtenues avec station de référence.

LETTER

New wavelets application in magnetotelluric data processing: reducing impedance bias

Hugo Larnier*†, Pascal Sailhac† and Aude Chambodut†

Abstract

Magnetotelluric (MT) data consists of the sum of several types of natural sources whether they are transient or quasi-periodic and of various sources of noise (instrumental, anthropogenic) whose nature has to be taken into account in MT processing. Most processing techniques are based on Fourier transform of MT time series and robust statistics at a fixed frequency to compute the MT response functions. But only a few take into account the nature of the sources. Moreover, to reduce the influence of noise in the inversion of the response functions, one often sets up another MT station called remote station. However, even careful set-up of this remote station can not prevent its failure in some cases. We offer here to use the continuous wavelet transform on magnetotelluric time series to reduce the influence of noise even for single site processing. We use two different types of wavelets, Cauchy and Morlet, according to the shape of observed geomagnetic events. We show that using wavelet coefficients at clearly identified geomagnetic events, we are able to recover the unbiased response function obtained using robust remote processing algorithms. This makes it possible to process even single station sites and increase the confidence in its interpretation.

Keywords: Magnetotelluric bias; Continuous wavelet transform; Response Function; Sferics; Geomagnetic pulsations

Introduction

Magnetotelluric (MT) method is based on the induction of natural electromagnetic (EM) fields in the ground. These natural fields can be distinguished into two main classes. High frequency (HF) signals (with frequency over 1 Hz) are mainly due to the lightning activity around the world with EM waves conveyed in the waveguide formed by the ionosphere and the conductive earth. The other main classe of EM waves is due to the interaction of the solar wind with the Earth's magnetic field that produces magnetohydrodynamic waves transmitted in the atmosphere through the ionosphere from the magnetosphere (Saito, 1969 [1], McPherson, 2005 [2]).

The MT method is thus based on the quasi-uniform source (Chave & Jones, 2012, [3, Chapter 2]) assumption where sources are supposed to be far from the measurement point. In this approximation, the horizontal electric field $e = (e_x, e_y)$ is linked to the horizontal magnetic field $h = (h_x, h_y)$ by convolution products $*$ (in time domain) with impulse response functions (z_{xx}, z_{xy}, z_{yx} and z_{yy}) which are components of the impedance tensor z:

$$\begin{aligned} e_x(t) &= z_{xx}(t) * h_x(t) + z_{xy}(t) * h_y(t) \\ e_y(t) &= z_{yx}(t) * h_x(t) + z_{yy}(t) * h_y(t). \end{aligned} \tag{1}$$

By properties of the Fourier transform, this gives us the following relation:

$$\begin{pmatrix} E_x \\ E_y \end{pmatrix} = \begin{pmatrix} Z_{xx} & Z_{xy} \\ Z_{yx} & Z_{yy} \end{pmatrix} \begin{pmatrix} H_x \\ H_y \end{pmatrix} \tag{2}$$

where E and H are Fourier transform of e and h respectively, and Z_{ij} is a 2×2 complex tensor.

Z can be transformed into apparent resistivity ρ_a and phase ϕ with:

$$\begin{aligned} \rho_a &= |Z|^2/(\mu\omega) \\ \phi &= \arctan(\Im(Z)/\Re(Z)), \end{aligned}$$

$\mu = 4\pi 10^{-7}$ H/m being the magnetic permeability, and $\omega = 2\pi f$ with f the signal frequency. These quantities are then analyzed or inverted to infer geoelectrical properties of the subsurface.

Because of the nature of MT time series (e.g. noise heteroscedasticity) and the complexity of the source

*Correspondence: hlarnier@unistra.fr
Institut de Physique du Globe de Strasbourg, UMR 7516, 5 Rue Rene Descartes, 67084 CEDEX Strasbourg, France
Full list of author information is available at the end of the article
†Equal contributor

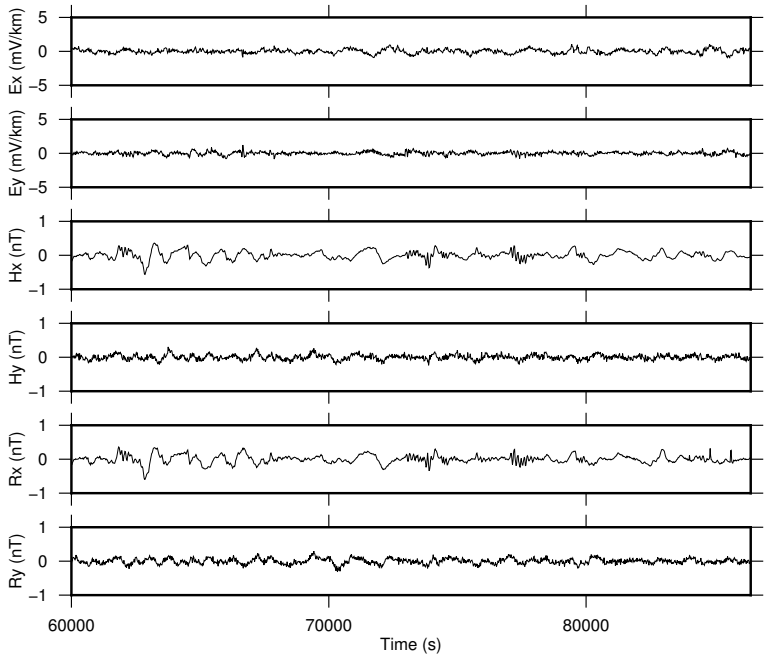


Figure 1 EM time series at Piton de la Fournaise MT stations on 28th May 1997. Sampling frequency : 0.05 Hz

field (transient nature and diversity of polarizations), Z computation is not straightforward. It has been the object of several methods published since the 70s. Addressing those issues is the object of this paper.

The first important issue to address is the effect of noise in MT time series. Indeed, Z estimations can be severely down weighted by noise on the magnetic field (Sims *et al.*, 1971, [4]). The usual way to get rid of this bias is to set up a second MT station called a remote station (Gamble *et al.*, 1979 [5]). This station is used in the processing assuming that sources of noise at both stations are not correlated so the bias would be drastically reduced.

The second issue we would like to address comes from the complexity of MT time series that we mentioned before. The first MT processing techniques (e.g. Sims *et al.*, 1971, [4]) were based on least-squares analysis which fails in the presence of outliers. The origin of outliers may come from several issues: non stationarity, source effect of MT signals, brief overwhelming noise. Robust statistics were then applied to MT processing (Chave *et al.*, 1987, [6], Egbert & Booker, 1986, [7]) to accommodate these outliers. Another approach was the use of wavelet techniques to look for MT sources or remove noise in the time series. Zhang & Paulson (1997) [8] applied the continuous wavelet transform (CWT) on high frequency data to select sferics events. They thus increase the signal-to-noise ratio in the response function determination in a straightforward way by the selection of useful signal parts. Trad and Travassos (2000) [9] used wavelets specifically to filter MT data in the time-frequency domain before processing using robust algorithms. Escalas *et al.* (2013) [10] used wavelet analysis to study the polarisations properties of cultural noise sources in MT time series.

We took into account both the issues of noise and of the transient nature of some MT sources by applying the CWT on two types of signals: geomagnetic pulsations and sferics. We demonstrate that the continuous wavelet transform may also be a powerful tool to reduce the bias in the computation of the impedance tensor, even in the case of single station processing.

Methodology

Continuous wavelet transform

The CWT is a mathematical technique used to decompose a signal on a time-frequency representation using a special class of functions called wavelets. A wavelet can be defined as a physical event (e.g. Green function) and has to fulfill two conditions: it has to be localized both in time and frequency domains and to be zero-mean (the so-called admissibility criteria) (Holschnei-

der, 1995, [11]). The wavelet transform allows to compute coefficients depending on two factors, dilatation $a > 0$ (corresponding to frequency) and translation τ (time) respectively. Each coefficient is computed following equation (3), where s is the analyzed signal (e.g. electric or magnetic field), ψ is the mother wavelet and * represents the complex conjugate:

$$\mathcal{W}_\psi[s](a,\tau) = \int_{-\infty}^{+\infty} dt \frac{1}{a}\psi^*\left(\frac{t-\tau}{a}\right)s(t) \quad \tau \in \mathbb{R} \quad (3)$$

By using properties of the Fast Fourier Transform (FFT) with the convolution, one is able to compute wavelet coefficients very quickly.

Wavelet's choice

In this work, we have chosen two different wavelets, each adapted to the shape of the considered geomagnetic event, the Cauchy wavelet and the Morlet wavelet. Both wavelets are progressive ones with real valued Fourier transform.

The Cauchy wavelet is given in frequency domain by:

$$\widehat{\psi}_c(a\omega) = \frac{2^m}{\sqrt{m\Gamma(2m)}}H(\omega)(a\omega)^m e^{-a\omega}, \quad (4)$$

where m is the wavelet's order (corresponding to the derivative order if integer), H the Heaviside step function, Γ the Gamma function and a the dilatation factor.

The Morlet wavelet is given in frequency domain by:

$$\widehat{\psi}_m(a\omega) = \pi^{-1/4}H(\omega)e^{-(a\omega-\omega_0)^2/2}, \quad (5)$$

where ω_0 is the central pulsation of the wavelet, H the Heaviside function and a the dilatation factor. The equation (5) is not representing an admissible wavelet if ω_0 is not high enough to fulfill the zero mean condition. To make it admissible, ω_0 must be above $\pi\sqrt{2/\log(2)}$. A highest ω_0 implies a highest number of oscillations in the wavelet functions.

The following sections illustrate the wavelet choice according to the observed geomagnetic events.

Pulsations

Geomagnetic pulsations are the result of the interactions between the solar wind and the magnetosphere and will appear on EM time series as short oscillations (Figure 1). The pulsations can be divided into two main classes, continuous (Pc pulsations) or irregular (Pi pulsations). Continuous pulsations are quasi-periodic oscillations whose occurrences vary during the day. For example, Pc1 pulsations are predominant in day-hours (in local time) at high latitudes and in night-hours at low latitudes (Saito, 1969, [1], Bortnik *et al.*, 2008, [12]). Moreover even in periods of occurence, they often appear as pearls on magnetograms with period of very low signal-to-noise ratio (Jun *et al.*, 2014, [13]). Irregular pulsations are characterized by a short time duration comparatively to Pc pulsations and can be classified as transient events in MT time series. Geomagnetic pulsations signal contains several oscillations so they are very localized in time and

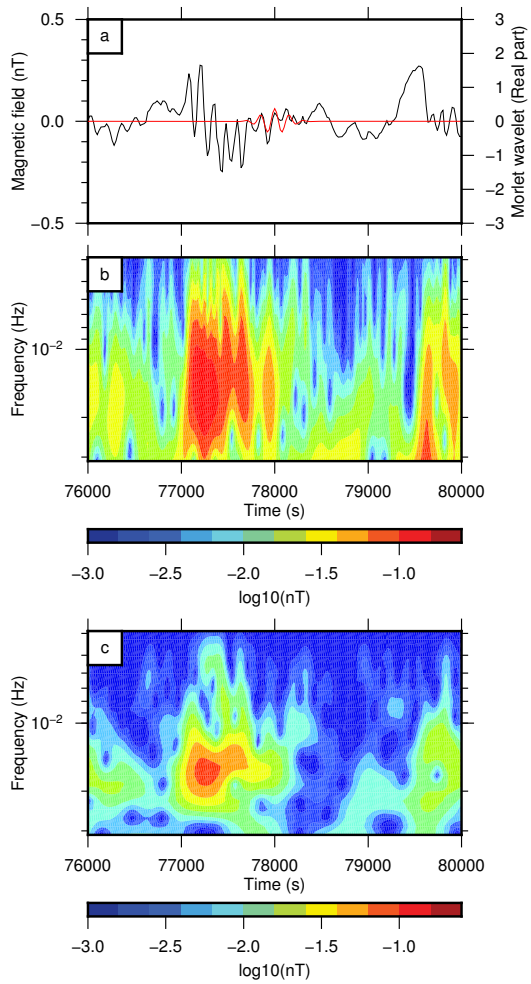


Figure 2 Geomagnetic pulsation on H_x and its representation in time-frequency space. MT site: Piton de la Fournaise (see Real data application section). The time axis is the same as Figure 1. a) Comparison between a pulsation (black line) and Morlet wavelet (real part, red line). b) Wavelet coefficients of the pulsation shown in a) using Cauchy wavelet. c) Wavelet coefficients of the pulsation shown in a) using Morlet wavelet.

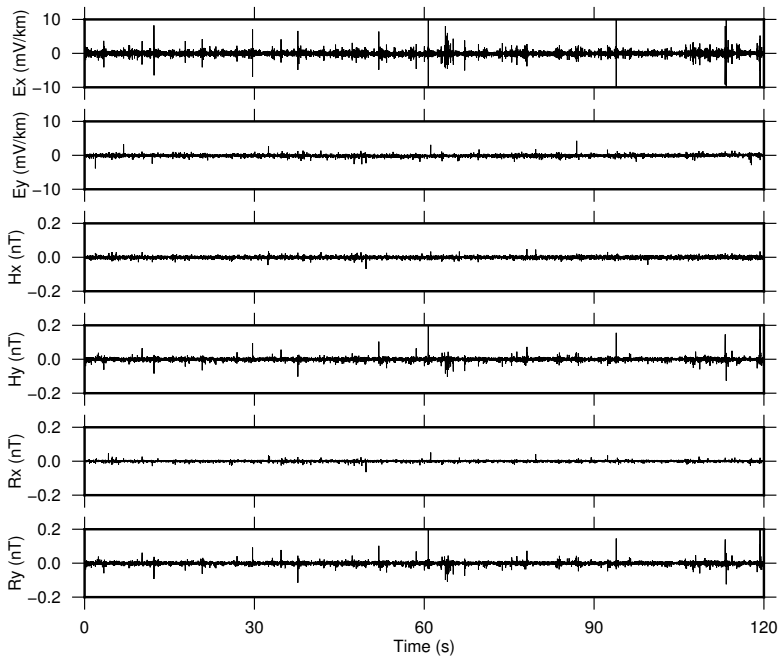


Figure 3 EM time series at Rittershoffen MT stations. Sampling frequency : 512 Hz. Time series start on the 2nd of June 2014 at 18h17 and 24 seconds (UT)

frequency simultaneously. Thus like previous authors (Zhang & Paulson, 1997, [8]; Garcia & Jones, 2008, [14] for atmospherics signals) we have chosen the Morlet wavelet which has the best compromise between time and frequency resolution. Using Cauchy wavelet would be hazardous because of its low frequency resolution. Both Morlet and Cauchy wavelet coefficients are shown Figure 2. As illustrated, highest coefficients corresponding to the geomagnetic pulsations are much more scattered in frequency using Cauchy wavelet than Morlet wavelet.

Sferics

Sferics are transient events occurring in the frequency band starting from 1 Hz to more than 10 kHz (Garcia & Jones, 2002, [15]). These signals are the most energetic part of the MT signal in these frequency bands (even if other phenomena occur such as the Schumann resonances between the wave guide made by ionosphere and Earth's surface).

These sferics are characterized by a sudden impulse in time domain and can be divided into two main frequency bands, above 1 kHz and below. In the low frequency band of Extremely Low Frequency (ELF) waves, sferics are constituted of a high amplitude (Figure 3) impulse called the slow tail (Mackay & Fraser-Smith, 2010 [16]) whereas in the high frequency band of Very Low Frequency (VLF) waves, sferics are mainly shaped as oscillations. To analyse ELF waves using the CWT, we have chosen the Cauchy wavelet. Most of the previous work on MT using wavelets was based on Morlet wavelet for its properties of good resolution both in time and frequency (Zhang & Paulson, 1997, [8], Garcia & Jones, 2008, [14]). However, the shape of the ELF wave (Figure 4, (a)) is very impulsive and therefore wide in frequency so it is closer to Cauchy's shape than to Morlet's one. Indeed, the ELF wave is very short in time and has very few oscillations so we need a wavelet which has similar properties. As illustrated on Figure 4 (b,c), the highest coefficients obtained with Morlet wavelet span a wide time length, longer than the actual length of the ELF wave. Due to their impulsive nature, sferics have a wide frequency content (Mackay & Fraser-Smith, 2010, [16]). In a time-frequency domain, sferics appear as a series of high valued coefficients along the frequency axis (very localized in time and spread in frequency).

Selection of wavelet coefficients

We manually picked geomagnetic events from our MT time series, and selected wavelet coefficients from the time-frequency plane using the following methodology:

In a way similar to wavelet based denoising techniques, amplitudes of the wavelet coefficients are used

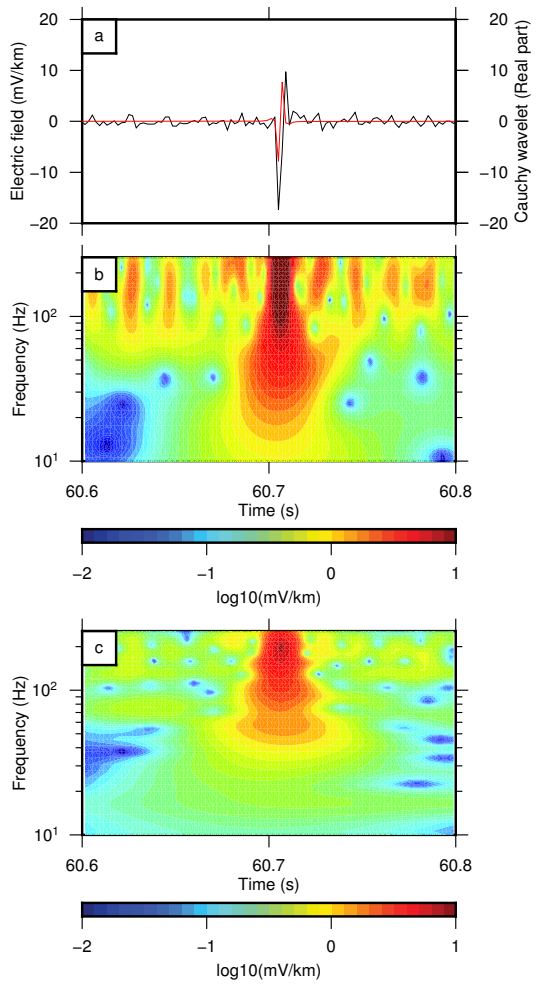


Figure 4 Sferic on E_x and its decomposition in time-frequency space. MT site: Rittershoffen (see Real data application section). Time axis the same as Figure 3. a) Comparison between a sferic (black line) and Cauchy wavelet (real part, red line). b) Wavelet coefficients of the sferic shown in a) using Cauchy wavelet. c) Wavelet coefficients of the sferic shown in a) using Morlet wavelet.

to define selection scheme. The MT system (1) can be solved in a univariate way for e_x and e_y separately. In both cases, one has to obtain good signal-to-noise ratio for both h_x and h_y. At each scale, for every channel s (output channel (e_x or e_y), and input channel (h_x, h_y)):

- We compute the median value α of the distribution of $|\mathcal{W}_\psi[s](a,\tau)|$ over a length $T_a = Nt_a$ around t_0 the time position of the previously picked event. N in this work is set to 30 and t_a is the time step corresponding to the analyzed scale.
- From this value, we define a threshold set as β times the median value α.
- All coefficients below this threshold level are discarded for the magnetotelluric response function computation.
- Coefficients are kept for final computation when the threshold was reached on (e_x, h_x and h_y) or (e_y, h_x and h_y).

This criteria is sufficient for geomagnetic pulsations (below the Hz). However for ELF waves, we add one more criteria to fulfill: selected coefficients must span a large frequency range (basically the whole frequency range of ELF waves without MT and AMT dead bands). In practice, we check the Fourier spectrum of EM time series which is a good indicator of dead band boundaries. Depending on the signal-to-noise ratio or the time of the day, this limit can be highly variable.

When a remote station is available, we add another criteria to discriminate noise from geomagnetic events. Indeed, the horizontal magnetic coefficients at the remote station have to fulfill the same properties. At the end of this stage, we obtain a group of wavelet coefficients for each scale and each selected geomagnetic event.

Inversion of wavelet coefficients

Zhang & Paulson (1997) [8] demonstrated that the equation (2) can be written as:

$$\begin{aligned}\mathcal{W}_\psi[e_x](a,\tau) &= Z_{xx}\mathcal{W}_\psi[h_x](a,\tau) + Z_{xy}\mathcal{W}_\psi[h_y](a,\tau)\\ \mathcal{W}_\psi[e_y](a,\tau) &= Z_{yx}\mathcal{W}_\psi[h_x](a,\tau) + Z_{yy}\mathcal{W}_\psi[h_y](a,\tau)\end{aligned} \tag{6}$$

a system of linear equations in the time-frequency domain described by the following equation:

$$d = Gm \tag{7}$$

where d contains the electric field wavelet coefficients ($\mathcal{W}_\psi[e_x](a,\tau)$, $\mathcal{W}_\psi[e_y](a,\tau)$), G contains the magnetic field coefficients ($\mathcal{W}_\psi[h_x](a,\tau)$, $\mathcal{W}_\psi[h_y](a,\tau)$), and m contains the MT response functions Z. The classical least-squares solution to this equation is:

$$Z = (G^T G)^{-1} G^T d. \tag{8}$$

This solution is inevitably downward-biased (Sims *et al.*, 1971, [4]). One solution is to include a remote station in the processing scheme (Gamble *et al.*, 1979, [5]). The solution then becomes:

$$Z = (G_r^T G)^{-1} G_r^T d \tag{9}$$

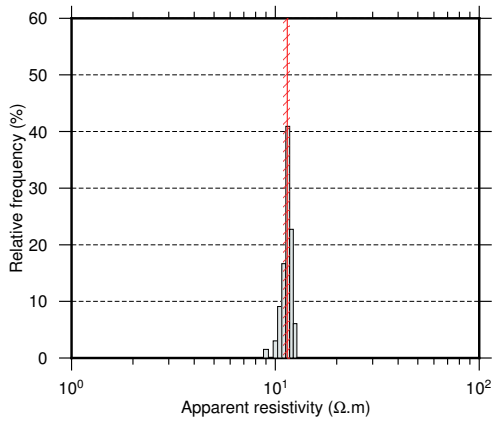


Figure 5 Relative frequency histogram of ρ_{xy} at a frequency of 45 Hz. Processing of Rittershoffen site in remote reference processing. Vertical red thick line : Median value of the distribution. Hatched domain : domain defined by the first and third quartile of the distribution.

where G_r contains the remote station magnetic field coefficients $\mathcal{W}_\psi[r_x](a,\tau)$ and $\mathcal{W}_\psi[r_y](a,\tau)$. Another way to reduce the bias is to actually make the predictor variable (here the h field) as noise-free as possible. In this alternate way our goal is to use wavelet coefficients of high signal-to-noise ratio geomagnetic events in the computation. By doing so we will reduce the bias introduced by noise on the magnetic field in the solution (8).

To reduce the influence of noise in the electric field, we also use robust statistics instead of classical least-squares. As explained in Chave & Jones (2012) [3, chapter 5], robust statistics were introduced in MT because of the properties of natural-source electromagnetic data. Among these reasons, they state "finite duration of many geomagnetic or cultural events" and "marked *non stationarity*". Moreover, the heteroscedasticity of the noise can hide natural events along the time series. The main argument for using wavelet transform is precisely to take into account these issues by getting rid of the non-useful part of the time series (e.g. no apparent geomagnetic signal) before computing the MT response function. In practice, in many time series the non-useful part are the largest. In consequence, Z computation is only based on parts of the time series where there is a significant induction occurring in the subsurface. Indeed, a significant part of the records can not be used because of the noise level of current state-of-the art magnetic sensors (Chave & Jones, 2012, [3, Chapter 9]). For example, in the AMT (Audio MT) dead band, the signal does not raise above the noise level of induction coils in daytime (Garcia & Jones, 2002, [15]). Robust statistics are still necessary, but we considerably increase the signal to noise ratio on all selected events.

Zhang & Paulson, 1997, [8] worked in single-station configuration and used conventional least-squares analysis to resolve equation (7). Instead of using all selected wavelet coefficients to recover the MT response functions, we estimate the response function on pairs of geomagnetic events. Doing so allows us to build a distribution for each component of the impedance tensor (e.g. Figure 5).

At each scale for each pair of geomagnetic events, the response function is estimated using the Huber M-estimator described by Chave *et al.*, (1987), [6] in a MT context. From the distribution of response functions at each scale estimated from all available pairs of geomagnetic events, we represent the final estimation by taking the median value (less sensitive to outliers). The confidence interval is represented using the interquartile range (IQR) L-estimator (Huber & Ronchetti, 2009, [17])of the distribution. The IQR has a breakdown point of 50 % which makes it robust to long tails in the distribution and is still a good measure of the dispersion or skewness in the distribution.

Real data application

Datasets

We applied the previously described technique on two real datasets to illustrate the potential of using wavelet coefficients on both geomagnetic pulsations and sferics.

The first dataset consists of MT data acquired with a sampling frequency of 512 Hz in the North East part of France near the town of Rittershoffen as part of a geothermal monitoring experiment (Abdelfettah *et al.*, 2014, [18]). The electric measurements were made with EPF06 electrodes, and the magnetic measurements with MFS07e coils from Metronix SA. Data in this area are badly contaminated by 50/3 Hz and 50 Hz noise and their harmonics. We filtered this noise using a notch filter with 2048 coefficients. The remote used with these measurements is located on the site of Welschbruch geophysical station in the Vosges mountains at about 60 kilometers away from the measurement site.

The second dataset consists of electromagnetic horizontal measurements acquired on La Fournaise volcano, before, during and after the 1998 eruption (Zlotnicki *et al.*, 2005, [19]). The data were acquired with induction coils similar to those made by Metronix or Phoenix Geophysics (Clerc, 1971, [20]), and $Pb-PbCl_2$ electrodes at a sampling frequency of 0.05 Hz. The remote station used here is also located on the volcano, a little less than 10 km away from the measurement station.

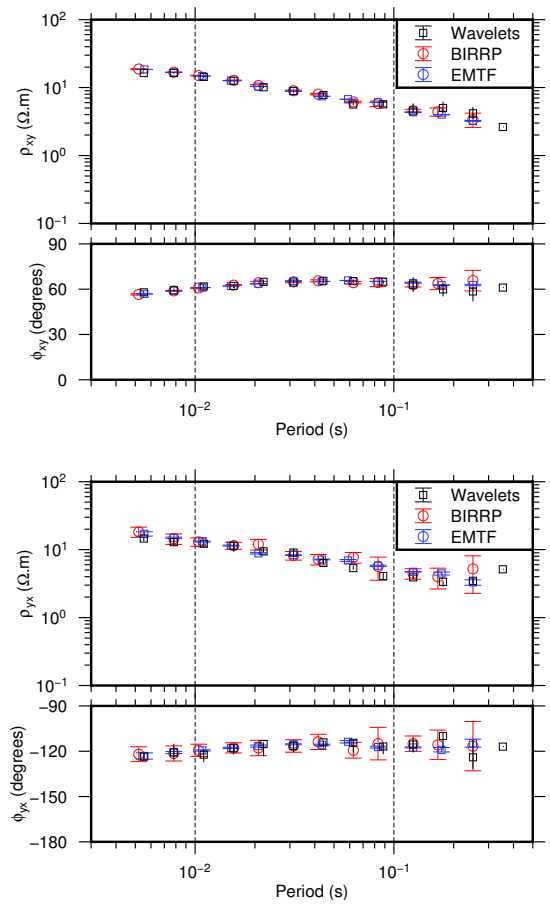


Figure 6 Comparison of wavelet based determination of the impedance tensor and robust algorithms MT site: Rittershoffen site with remote reference processing.

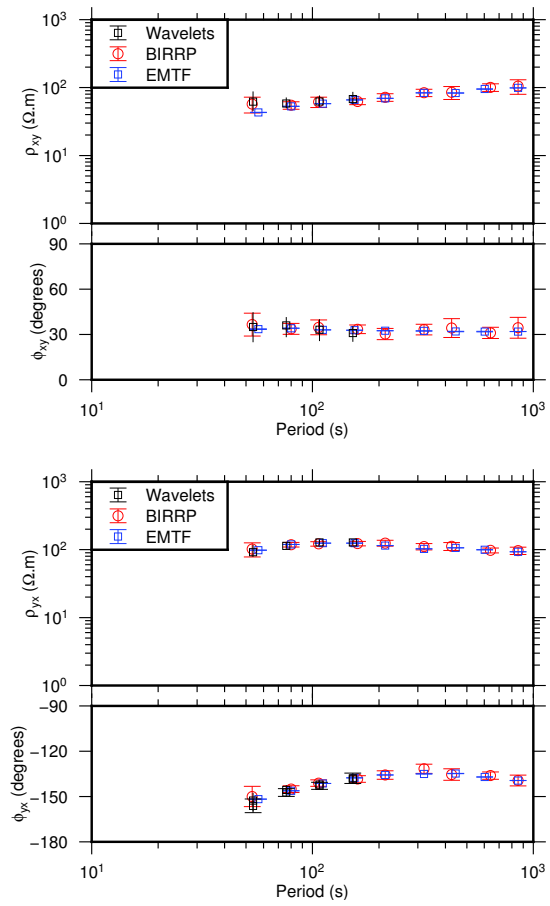


Figure 7 Comparison of wavelet based determination of the impedance tensor and robust algorithms MT site: Piton de la Fournaise site with remote reference processing.

To assess the validity of our method, we have compared our results with the ones obtained with robust processing codes of Alan D. Chave (BIRRP, Bounded Influence Remote Reference Processing, Chave & Thomson, 1987, [6], and 2004, [21]), and Gary Egbert (EMTF, Egbert, 1997, [22]). For simplicity, we consider error bars as they are calculated using robust processing codes, even if they can be discussed (Wawrzyniak *et al.*, 2013 [23], Chave, 2015, [24], Wawrzyniak *et al.*, 2015, [25]).

The threshold factor value β was set to 1 for sferic's application and 4 for pulsations' application but other values can be chosen depending on the signal-to-noise ratio of the processed time series.

Remote processing

We compared the MT response functions obtained using wavelet coefficients with the application of other processing codes on both datasets.

In both frequency bands (Figures 6 and 7), the response functions obtained using wavelet coefficients are comparable to both processing codes. For response functions in the frequency band covered by geomagnetic pulsations (low frequency band), wavelet inversion is also in agreement with robust processing even if the source is not as impulsive as sferics. Figure 5 also demonstrate that using robust statistics on the wavelet coefficients of only two events allows the recovery of the response functions with good accuracy.

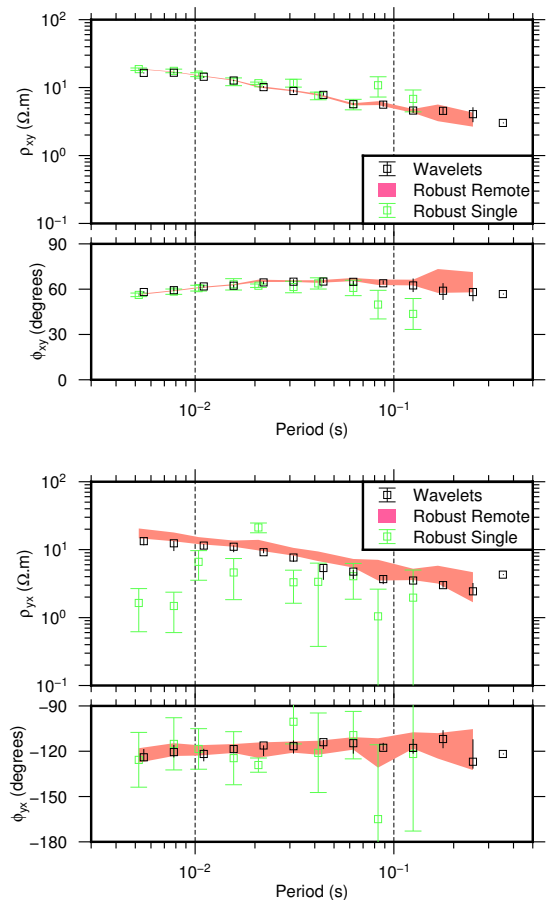


Figure 8 Comparison of wavelet based determination of the impedance tensor and robust algorithms MT site: Rittershoffen site in single site processing.

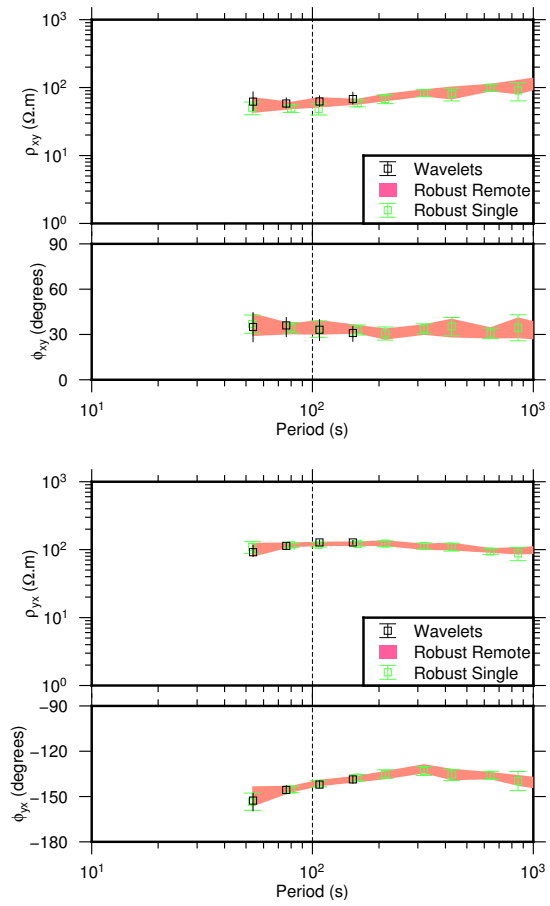


Figure 9 Comparison of wavelet based determination of the impedance tensor and robust algorithms MT site: Piton de la Fournaise site in single site processing.

Single site processing

For single site processing, we processed single site MT stations by removing the remote station for all processing codes. For visual comparison, only BIRRP results are shown, but conclusions remain the same for EMTF.

For high frequency results (Figure 8), ρ_{yx} is downward biased by noise over a large frequency bandwidth and ρ_{xy} is also biased around 0.08 Hz. Because of the position of the station near housing, it is difficult to assess where the noise source is (it could be electric fences or pipelines). In this case, the result given by robust processing in single site configuration would lead to a wrong interpretation of the MT response. Wavelet processing allows to drastically reduce the noise bias and recover the impedance tensor obtained using remote processing even near the 50 Hz frequency.

For low frequency processing (Figure 9), the single site ρ_{xy} is slightly downward biased by noise where the wavelet-based result remains within the confidence bounds of the remote processing. Particularly, the ρ_{xy} component was very sensitive to the coherence parameter. Some examples of interpretable MT responses for this component are illustrated Figure 10. If the coherence parameter is too low (below 0.5), ρ_{xy} is distincly biased by noise (up to a factor of three) for periods below 200 seconds.

Other transient applications

So far, we have illustrated only two types of transient signals, sferics in their lowest frequency range (e.g. slow tails), and geomagnetic pulsations. This approach can also be applied on wider MT bands than the ones presented in this paper.

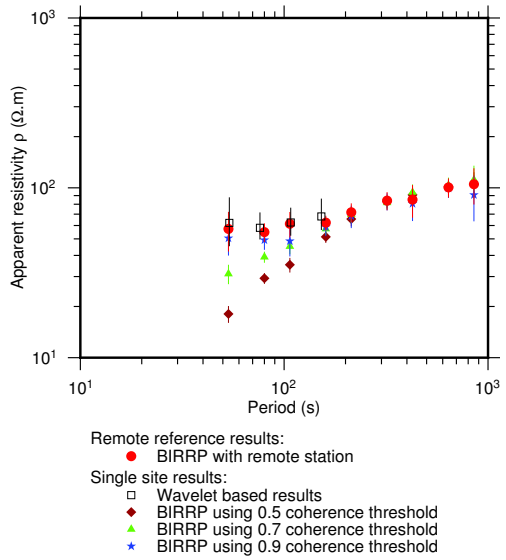


Figure 10 Dependance of ρ_{xy} robusts results on coherence parameter MT site: Piton de la Fournaise site in single site processing.

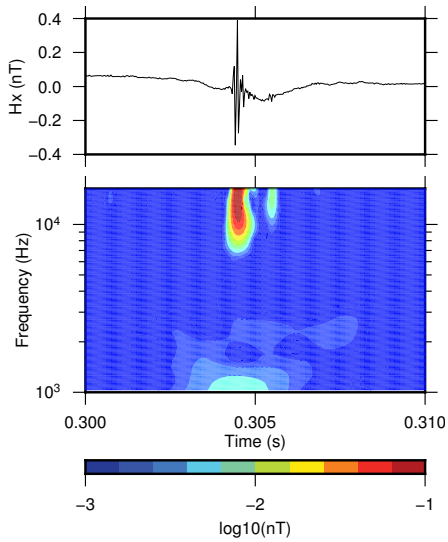


Figure 11 VLF wave on H_x time series at Rittershoffen. Sampling frequency: 32768 Hz.

For higher frequencies of application, the slow tails studied in this paper already contains frequency content above 256 Hz (up to 1 kHz). Other impulsive waves such as VLF waves have distinctive signature in the time-frequency plane and frequency content from 1 kHz up to more than 10 kHz (Rakov & Uman, 2003, [26]). On Figure 11, we show an example of such a VLF wave and its time-frequency analysis. These waves were already used in MT processing in Zhang & Paulson, (1997), [8] study.

The assortment of low frequency geomagnetic sources is large. Among transient events at lower periods than 200 seconds, there are Pc5 geomagnetic pulsations (Pc5) which have periods up to 500 - 600 seconds (Saito, 1969, [1]). Geomagnetic storms are also low frequency sources of induction. They have already been studied using wavelet analysis (e.g. Mendes Jr. *et al.*, 2005, [27]). We show for example on Figure 12 a signal of interest at low frequency recorded on Piton de La Fournaise volcano on 16th July 1997 which contains significant frequency content to more than 1000 seconds.

Conclusion

We have shown through these experiments that the continuous wavelet transform (CWT) is an easy and efficient way to characterize the MT response function for transient geomagnetic events. Using this technique, we have shown that most of the information contained in the sources wavelet coefficients is sufficient in datasets to characterize the MT impedance tensor. Two types of geomagnetic events were studied, geomagnetic pulsations and sferics, increasing the frequency band studied in previous publications of CWT application on MT processing. The mother wavelet has to be adapted to each transient event to accuratly recover the source information. Even irregular geomagnetic pulsations which are very localized in time and frequency simultaneously can be used in MT processing. The other transient sources of induction are currently being investigated and will make the subject of future paper on wavelet analysis of magnetotelluric time series.

We have also shown that we were able to drastically reduce the noise bias of the considered datasets on the MT response function in the case of single station configuration. This is achieved by using the high signal-to-noise ratio transient events in MT time series. By using robust statistics with only two geomagnetic events in the wavelet domain, we recover the accurate response functions.

Yet, we still have to develop a way to automatically detect geomagnetic events to analyse their properties (e.g. polarization) and their effect on MT impedance computation.

Competing interests

The authors declare that they have no competing interests.

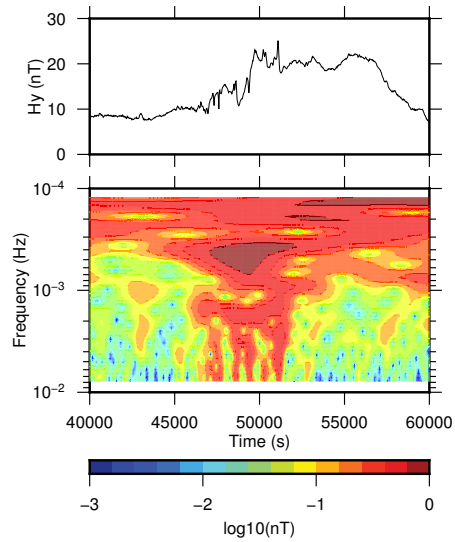


Figure 12 Low frequency variations of the magnetic field at Piton de La Fournaise MT site. Time series start on the 16th of July 1997.

Author's contributions

The manuscript was written by HL and reviewed by all authors. All authors read and approved the final manuscript.

Acknowledgements

We would like to thank the two anonymous reviewers who greatly helped to improve this manuscript. We thank the LABEX G-Eau-Thermie of the University of Strasbourg and the field crew (Y. Abdelfettah) for the acquisition of MT measurements at Rittershoffen. We also thank J. Zlotnicki for Le Piton de la Fournaise EM time series. Finally we also acknowledge discussions with various colleagues and students at early stages of this work.

References

1. Saito, T.: Geomagnetic pulsations. Space Science Reviews **10**, 319–412 (1969)
2. McPherson, R.L.: Magnetic pulsations: Their sources and relation to solar wind and geomagnetic activity. Surveys in Geophysics **26**(5), 545–592 (2005)
3. Chave, A.D., Jones, A.: The Magnetotelluric Method: Theory and Practice. Cambridge University Press, Cambridge (2012)
4. Sims, W.E., Bostick, F.X.J., Smith, H.W.: The estimation of magnetotelluric impedance tensor elements from measured data. Geophysics **36**(5), 938–942 (1971)
5. Gamble, T.D., Goubau, W.M., Clarke, J.: Magnetotellurics with a remote magnetic reference. Geophysics **44**(1), 53–68 (1979)
6. Chave, A.D., Thomson, D.J., Ander, M.E.: On the robust estimation of power spectra, coherences, and transfer functions. Journal of Geophysical Research **92**(B1), 633–648 (1987)
7. Egbert, G.D., Booker, J.R.: Robust estimation of geomagnetic transfer functions. Geophysical Journal of the Royal Astronomical Society **87**(1), 173–174 (1986)
8. Zhang, Y., Paulson, K.V.: Enhancement of signal-to-noise ratio in natural-source transient magnetotelluric data with wavelet transform. Pure and Applied Geophysics **149**, 405–419 (1997)
9. Trad, D.O., Travassos, J.M.: Wavelet filtering of magnetotelluric data. Geophysics **65**(2), 482–491 (2000)
10. Escalas, M., Queralt, P., Ledo, J., Marcuello, A.: Polarisation analysis of magnetotelluric time series using a wavelet-based scheme: A method for detection and characterisation of cultural noise sources. Physics of the Earth and Planetary Interiors **218**, 31–50 (2013)
11. Holschneider, M.: Wavelets: An Analysis Tool. Oxford Science Publications, Oxford (1995)
12. Bortnik, J., Cutler, J.W., Dunson, C., Bleier, T.E., McPherron, R.L.: Characteristics of low-latitude Pc1 pulsations during geomagnetic storms. Journal of Geophysical Research **113** (2008)
13. Jun, C., Shiokawa, K., Connors, M., Schofield, I., Poddelsky, I., Shevtsov, B.: Study of Pc1 pearl structures observed at multi-point ground stations in Russia, Japan, and Canada. Earth, Planets and Space **66**(140) (2014)
14. Garcia, X., Jones, A.G.: Robust processing of magnetotelluric data in the AMT dead band using the continuous wavelet transform. Geophysics **73**(6), 223–234 (2008)
15. Garcia, X., Jones, A.G.: Atmospheric sources for audio-magnetotelluric (AMT) sounding. Geophysics **67**(2), 448–458 (2002)
16. Mackay, C., Fraser-Smith, A.C.: Lightning location using the slow tails of sferics. Radio Science **45** (2010)
17. Huber, P.J., Ronchetti, E.M.: Robust Statistics, 2nd Edition. Wiley, New York (2009)
18. Abdelfettah, Y., Sailhac, P., Schill, E., Larnier, H.: Preliminary magnetotelluric monitoring results at Rittershoffen. 3rd European Geothermal Workshop, Karlsruhe, 15-16 October 2014 (2014)
19. Zlotnicki, J., Le Mouël, J.-L., Gvishiani, A., Agayan, S., Mikhailov, V., Bogoutdinov, S., Kanwar, R., Yvetot, P.: Automatic fuzzy-logic recognition of anomalous activity on long geophysical records: Application to electric signals associated with the volcanic activity of La Fournaise volcano (Reunion Island). Earth and Planetary Science Letters **234**, 261–278 (2005)
20. Clerc, G.: Contribution à l'optimisation des capteurs à induction destinés à la mesure des variation du champ magnétique terrestre (10^{-3} à 10^{4} hz). PhD thesis, Paris, France (1971)
21. Chave, A.D., Thomson, D.J.: Bounded influence magnetotelluric response function estimation. Geophysical Journal International **157**(3), 988–1006 (2004)
22. Egbert, G.D.: Robust multiple-station magnetotelluric data processing. Geophysical Journal International **130**(2), 475–496 (1997)
23. Wawrzyniak, P., Sailhac, P., Marquis, G.: Robust error on magnetotelluric impedance estimates. Geophysical Prospecting **61**, 533–546 (2013)
24. Chave, A.: Comment on 'Robust error on magnetotelluric impedance estimates' by P. Wawrzyniak, P. Sailhac and G. Marquis, Geophysical Prospecting, DOI: 10.1111/j.1365–2478.2012.01094.x. Geophysical Prospecting (2015)
25. Wawrzyniak, P., Sailhac, P., Marquis, G.: Reply of the authors to AD Chave's comment on Wawrzyniak P., Sailhac P. and Marquis G. 2013. Robust error on magnetotelluric impedance estimates. Geophysical Prospecting (2015)
26. Rakov, V.A., Uman, M.A.: Lightning: Physics and Effects. Cambridge University Press, Cambridge (2003)
27. Mendes Jr., O., Domingues, M.O., da Costa, M., Clúa de Gonzalez, A.: Wavelet analysis applied to magnetograms: Singularity detections related to geomagnetic storms. Journal of Atmospheric and Solar-Terrestrial Physics **67**, 1827–1836 (2005)

3.4.2.4 Transition vers le bootstrap

Lors du test de la méthode, nous nous sommes rendus compte que l'utilisation d'un grand nombre d'événements dans nos combinaisons permettait de diminuer la dispersion des distributions sur les fonctions de réponse MT. Pour montrer cet effet, nous avons calculé le tenseur d'impédance en utilisant des combinaisons de 2, 3 ou 4 événements parmi l'ensemble des événements disponibles. Pour obtenir une population représentative, nous avons permis le calcul de 1000 combinaisons pour chaque cas. La série temporelle utilisée est celle de la station Audio-MT Faille Ty acquise en Guadeloupe en Novembre 2012. La fréquence d'acquisition est de 16,384 Hz, la station de référence est installée sur le site de Youketi. La détection des événements est effectuée par le code décrit précédemment dans la section 2.2.4. Pour ce traitement, les ondes EM cherchées sont les *slow tails*.

Les résultats sont présentés à la Figure 3.5. Plus le nombre d'événements utilisé dans les combinaisons est grand, plus la dispersion des valeurs de Z obtenues est grande. Cette constatation nous a dirigé vers la technique du bootstrap qui permet d'utiliser l'ensemble des événements à disposition.

3.4.2.5 Bootstrap et tirage aléatoire

La technique bootstrap est une méthode permettant de donner des informations sur une distribution de probabilité inconnue par ré-échantillonnage de tous les échantillons à disposition. Contrairement à l'approche précédente où nous n'utilisions qu'une partie des événements pour chaque calcul des fonctions de réponse, nous allons ici utiliser l'ensemble des événements pour estimer ces fonctions.

Présentation Soit $\mathbf{x} = (x_1, x_2, ..., x_n)$ n échantillons identiquement et indépendamment distribués selon une loi de probabilité F dont on cherche à évaluer un paramètre $\theta = t(F)$. La façon dont les échantillons $\mathbf{x}$ sont distribués par F est inconnue. Le bootstrap propose d'estimer la fonction F en la remplaçant par la loi de répartition empirique $\hat{F}$ qui attribue à chaque échantillon un probabilité $1/n$ d'être tiré. Ainsi, l'estimation de θ est donné par $\hat{\theta} = t(\hat{F})$. Ainsi, si l'espérance et la variance de F sont notées μ_F et σ_F^2, alors l'estimateur de la variance de F est donné par :

$$\hat{\sigma} = \sigma_{\hat{F}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{3.24}$$

Soit s un estimateur $\theta = s(\mathbf{x})$, s pouvant être la moyenne ou l'écart type. La procédure bootstrap se base sur le calcul de B échantillons bootstrap $\mathbf{x^b}$ définis comme des échantillons de la même taille que $\mathbf{x}$ mais dont les valeurs $(x_1^b, x_2^b, ..., x_n^b)$ sont tirées aléatoirement dans la

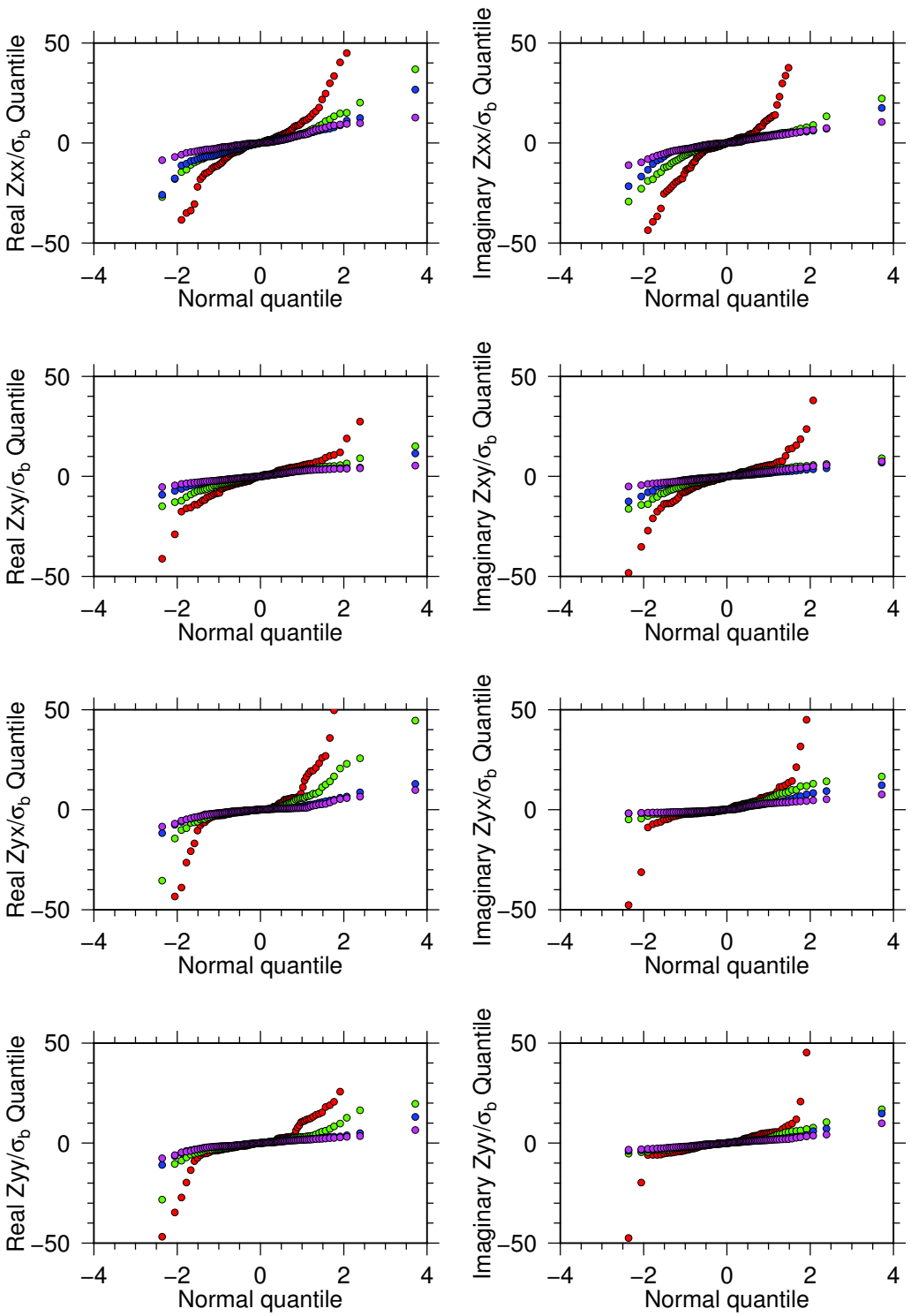


FIGURE 3.5: Distribution des événements inversés par la méthode utilisant les combinaisons pour une fréquence de 107 Hz. Les distributions ont été ramenés à une médiane nulle et divisées par l'écart type de la distribution Bootstrap. Rouge : Combinaisons de 2 événements. Vert : Combinaisons de 3 événements. Bleu : Combinaisons de 4 événements. Magenta : Combinaisons de 5 événements.

FIGURE 3.5: *Impedance tensor distribution at 107 Hz frequency by inverting combinations of events wavelet coefficients. Medians of each distribution have been set to 0 and distributions have been divided by the bootstrap distribution standard deviation. Red : Combinations of two events. Green : Combinations of three events. Blue : Combinations of four events. Pink : Combinations of five events.*

population de $\mathbf{x}$ avec retirage. Ainsi, un même élément peut se retrouver plusieurs fois dans l'échantillon bootstrap ou pas du tout.

Chaque échantillon bootstrap $\mathbf{x^b}$ permet de donner une estimation θ^b du paramètre θ par :

$$\theta^b = s(\mathbf{x^b}) \tag{3.25}$$

La moyenne bootstrap est définie par :

$$\hat{\theta}^* = \sum_{b=1}^{B} \theta^b / B \tag{3.26}$$

L'ensemble des estimations θ^b permet de déterminer l'estimation bootstrap de l'écart type par :

$$\hat{se}_B = \sqrt{\sum_{b=1}^{B} \frac{[\hat{\theta}^b - \hat{\theta}^*]^2}{B-1}} \tag{3.27}$$

Application à la magnétotellurique : Problème de régression Dans le cas de l'estimation des fonctions de réponses MT, chaque échantillon x_i est un couple $(e_i, \mathbf{h_i})$ estimé à une fréquence. Le problème revient à l'estimation du vecteur $\mathbf{z}$ défini par :

$$e_i = \sum_{j=1}^{2} z_{ij} h_{ij} + \varepsilon_i, \qquad i = 1, 2 \tag{3.28}$$

Soit $\hat{\mathbf{z}}$ une estimation de $\mathbf{z}$ obtenue par une méthode quelconque (moindres carrés ou régression robuste). Efron et Tibshirani, 1994 développe plusieurs méthodes pour l'application du bootstrap aux problèmes de régression linéaire (Chapitre 7 et 9). L'un des modèles se base sur le bootstrapping des paires $x_i = (e_i, \mathbf{h_i})$, l'autre sur le bootstrapping des résidus $x_i = (\hat{\mathbf{z}}\mathbf{h_i} + \varepsilon_i, \mathbf{h_i})$. Le second modèle se base sur l'hypothèse que l'erreur ε_i est indépendante de $\mathbf{h}_i$. Or en MT, les erreurs au modèle peuvent être indépendantes de la valeur du champ magnétique, comme dans le cas de bruit anthropique tels que des spikes dans les données, ou dépendantes comme dans le cas des effets de source proche.

Le jackknife, présenté au paragraphe 3.2.3.2, est une linéarisation du bootstrap (Efron et Tibshirani, 1994). Le bootstrap est de fait applicable dans le calcul des erreurs pour la méthodologie développée par Alan D. Chave et David J. Thomson (1987) et Alan D. Chave et David J. Thomson (2004), mais nécessite un plus long temps de calcul.

Dans la méthodologie par ondelettes continues développée dans cette thèse, une approximation est cependant violée par le principe même des ondelettes continues qui implique une corrélation entre les coefficients. Dans le problème de la détermination des fonctions de réponses d'ondelettes, chaque échantillon x_i est un couple $(\mathcal{W}_\psi[e_i](a), \mathcal{W}_\psi[\mathbf{h_i}])$. Le problème est toujours

celui d'une régression linéaire :

$$\mathcal{W}_\psi[e_i](a) = \sum_{j=1}^{2} z_{ij}\mathcal{W}_\psi[h_{ij}](a) + \mathcal{W}_\psi[\varepsilon_i], \qquad i = 1, 2 \tag{3.29}$$

Dans Larnier, Pascal Sailhac et Chambodut (2016), pour un même échantillon, nous sélectionnons plusieurs coefficients à chaque échelle pour le calcul des fonctions de réponse. Ces échantillons sont dont corrélés entre eux par le noyau de corrélation de l'ondelette utilisée. Ainsi, deux échantillons au sein même de $\mathbf{x}$ peuvent ne pas être indépendamment distribués.

Bootstrap hiérarchique Pour pouvoir appliquer la méthodologie du bootstrap, il est donc nécessaire de pouvoir assurer l'indépendance de nos échantillons dans notre population $\mathbf{x}$. Le problème est le suivant, nous disposons d'une population $\mathbf{x}$ dont plusieurs coefficients peuvent être corrélés entre eux. Les données peuvent donc être décrites sur deux niveaux : le premier correspond aux groupes de coefficients corrélés entre eux, le deuxième aux coefficients eux-mêmes. Le premier niveau correspond au niveau des événements géomagnétiques.

Ainsi, nous disposons donc d'une population de N groupes de coefficients $\mathbf{w} = (w_1, ..., w_N)$, avec $n \neq N$. Chaque groupe w_i contient n_i coefficients, avec :

$$n = \sum_{i=1}^{N} n_i \tag{3.30}$$

Field et Welsh, 2007 ont étudié différents schémas de ré-échantillonnage de données hiérarchiques, avec remplacement à un seul niveau ou plusieurs. Un des schémas les plus performants est celui du remplacement au niveau le plus élevé qui correspond ici à celui des événements géomagnétiques. Dans ce type de bootstrap, les échantillons bootstrap $\mathbf{w}^b$ n'ont pas tous le même nombre de coefficients d'ondelettes, mais le nombre d'événements reste identique dans chaque échantillon.

Pour résumer, les estimations des fonctions de transfert se font donc sur la population des multiplets $\mathbf{w}$. Chaque échantillon bootstrap $\mathbf{w}^b$ est défini par ré-échantillonnage des multiplets w_i avec remplacement. Les fonctions de réponse MT sont calculées pour chaque échantillon bootstrap, et l'estimation de l'écart type est donné par la même relation (3.27) que pour le bootstrap conventionnel.

3.4.3 Applications

3.4.3.1 Eclairs atmospherics

Pour introduire l'apport du bootstrap sur la détermination des fonctions de réponse MT, nous comparons ici les distributions bootstrap sur le même exemple présenté à la Figure 3.5. Les résultats sont présentés à la Figure 3.6. Il est clair que les distributions bootstrap présentent une dispersion bien moins importante que celles définies par le calcul par combinaisons. A la Figure 3.7, nous présentons les sondages obtenus par les différentes inversions des fonctions de réponse MT. Excepté dans l'Audio-MT *dead band*, les sondages sont très comparables pour toutes les techniques. En revanche, plus le nombre d'événements utilisés dans l'inversion est important, plus la stabilité de la courbe est améliorée.

3.4.3.2 Article GJI - Détermination de la réponse MT par bootstrap

Résumé en français Les paramètres physiques de la sub-surface à des profondeurs entre 50 mètres et plusieurs centaines de mètres, sont d'importance capitale pour la compréhension de divers processus dans domaines variés, industriels et environnementaux, tels que la compréhension de la zone critique en hydro-géophysique. La méthode Audio-MT, basée sur la mesure passive des champs électrique et magnétique naturels, est l'une des quelques techniques permettant d'apporter des informations à cette échelle de profondeur. Cependant, elle est sujette à la variabilité des sources EM naturelles dans l'Audio-MT *dead band*, intervalle de fréquence compris entre 1 et 5 kHz.

Nous montrons que la caractérisation des fonctions de réponse MT, à partir desquelles la distribution de conductivité dans le sous-sol est déterminée, peut être significativement améliorée dans l'Audio-MT *dead band*, en utilisant la transformée en ondelette continue. Cette analyse, comparée à des procédures standard de traitement de données utilisant la transformée de Fourier, permet de sélectionner les sources naturelles dont le contenu fréquentiel dans l'Audio-MT *dead band* est significatif, améliorant ainsi le ratio signal-sur-bruit. Les fonctions de réponse MT et leurs incertitudes associées sont calculées en utilisant des procédures de statistiques robustes et de bootstrap.

Nous illustrons cette procédure sur des données acquises en Mars 2015 à deux stations situées près de l'observatoire magnétique de Chambon-La-Forêt, en France. Les séries temporelles présentent de forts niveaux de bruit anthropiques. Dans l'Audio-MT *dead band*, les fonctions de réponse MT déterminées par les méthodes conventionnelles présentent de large intervalles de fréquences où aucune interprétation n'est possible. La procédure basée sur les ondelettes, appliquée sur des séries temporelles de 15 minutes permet la détection d'événements géomagnétiques

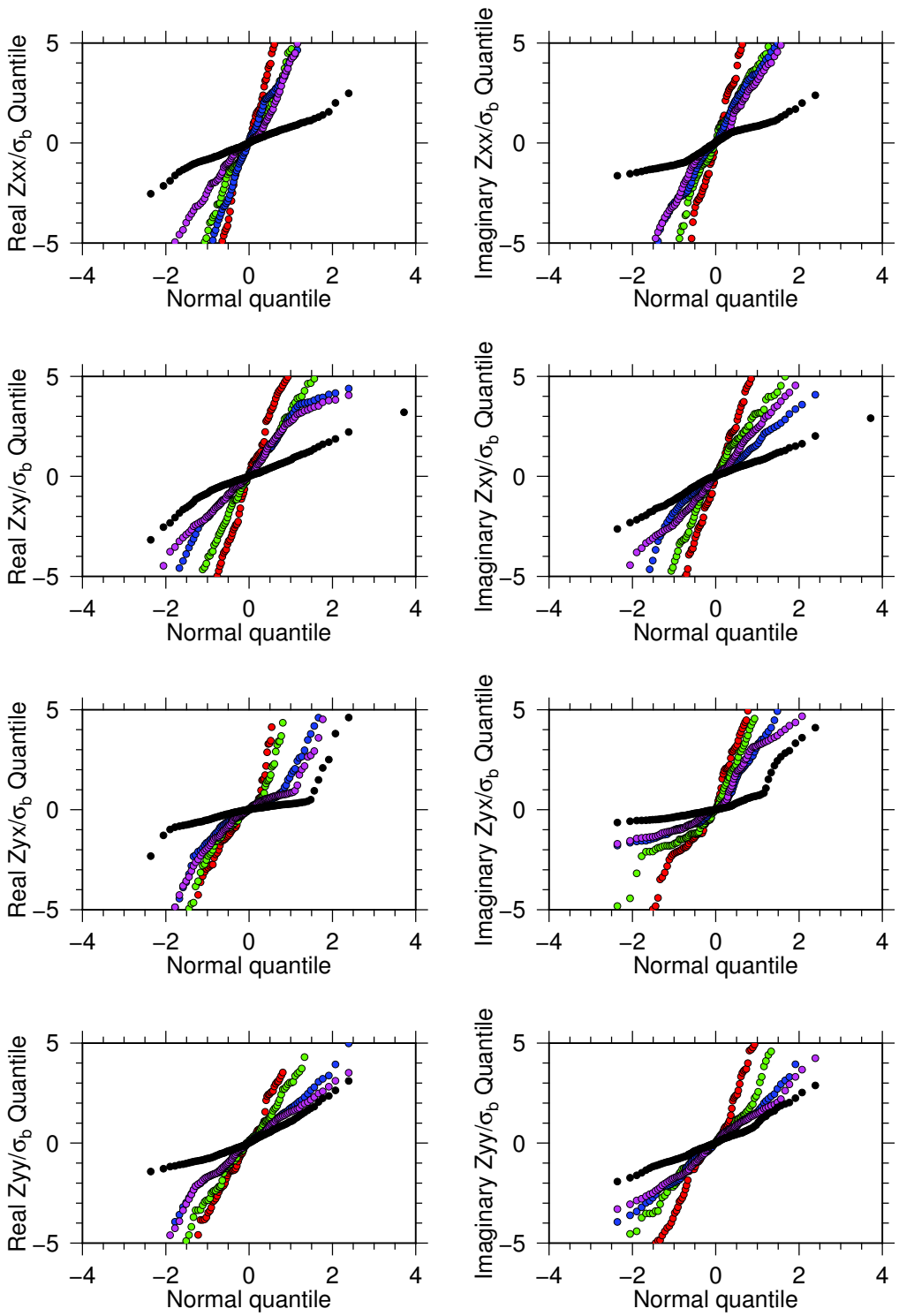


FIGURE 3.6: Distribution des événements inversés par la méthode utilisant les combinaisons pour une fréquence de 107 Hz. Les distributions ont été ramenés à une médiane nulle et divisées par l'écart type de la distribution Bootstrap. Rouge : Combinaisons de 2 événements. Vert : Combinaisons de 3 événements. Bleu : Combinaisons de 4 événements. Magenta : Combinaisons de 5 événements.

FIGURE 3.6: *Impedance tensor distribution at 107 Hz frequency by inverting combinations of events wavelet coefficients. Medians of each distribution have been set to 0 and distributions have been divided by the bootstrap distribution standard deviation. Red : Combinations of two events. Green : Combinations of three events. Blue : Combinations of four events. Pink : Combinations of five events. Black : Bootstrap inversion.*

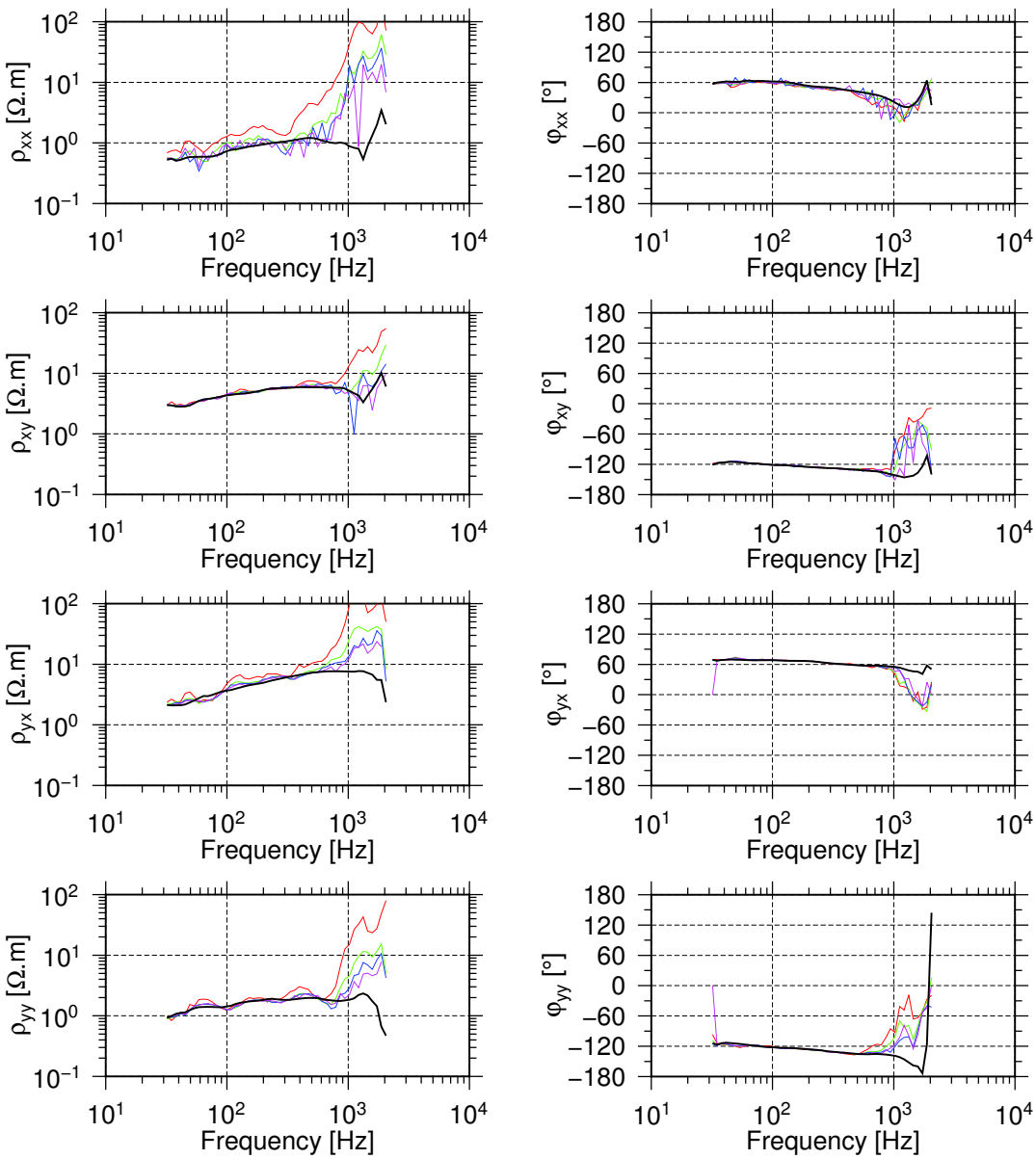


FIGURE 3.7: Distributions des inversions bootstrap pour une fréquence de 107 Hz.
FIGURE 3.7: *Bootstrap distributions for 107 Hz frequency*

avec un large rapport signal-sur-bruit, réduisant ainsi la largeur de la *dead band* après processing. Selon les conditions ionosphériques, la *dead band* peut être réduite jusqu'à 1-2 kHz quand des *whistlers* sont présents, et 1-5 kHz quand ils sont absents. Cette méthodologie permet une interprétation des sondages plus précise. La transformée en ondelettes apparaît ainsi comme un outil essentiel pour accroître les applications géophysiques de la méthode Audio-MT.

Cet article, en préparation, est prévu pour soumission à *Geophysical Journal International*.

submitted to *Geophys. J. Int.*

Determination of MT response function using hierarchical bootstrapping of geomagnetic events

H. Larnier[1], P. Sailhac[1], A. Chambodut[1], F. Perrier[2]

[1] *Institut de Physique du Globe de Strasbourg, CNRS UMR-7516, Université de Strasbourg, France*

Received ? Month? ?; in original form ? Month? ?

SUMMARY

The physical condition of the subsurface at depth between 50 m to several hundred meters, which hosts the surface aquifer-bearing formations, is of tremendous importance for the understanding of the critical zone and for environmental and industrial applications. Audio-Magneto-Telluric (AMT) sounding, based on passive joint determination of electrical and geomagnetic fields in the audio-frequency band, is one of the few available techniques able to provide information in this depth range. However, it suffers from the scarcity of natural sources in the so called AMT dead band (from 1 kHz to 5 kHz), We show that the determination of magnetotelluric impedance, from which the electrical conductivity structure is derived, can be significantly improved in the frequency range 200 Hz to 5-7 kHz using a wavelet based processing. Indeed, Continuous Wavelet Transform (CWT), by contrast with the standard robust processing based on Fourier transform, allows to select in the time series relevant sources with frequency content penetrating the AMT dead band. As previously shown, the CWT improves the signal over noise ratio. Now we determine response functions and their accuracy by using robust statistics and a bootstrap approach. This is illustrated using time series acquired in March 2015 at two locations near the French Magnetic Observatory in Chambon-La-Forêt, France. Time series recorded in this area display large noise levels. In the AMT dead band, robust processing outputs rough response functions with large confidence intervals where no

interpretation is possible. The wavelet-based processing applied to 15 minutes time series allows detection of large signal-to-noise ratio geomagnetic events to reduce the width of the dead band. Depending on ionospheric conditions, the dead band is narrowed down to 1 kHz up to 2 kHz when significative whistlers waves occur, or 3-5 kHz when no whistlers waves are identifiable. Using this methodology, AMT soundings interpretation is eased. CWT thus appears as an essential tool to increase the range of applications of the AMT method in geophysical problems.

Key words: Atmospherics, Magnetotelluric response function, Continuous Wavelet Transform, Bootstrap.

1 INTRODUCTION

Natural geomagnetic sources in the frequency range 1 Hz-10 kHz mostly originate from global lightning activity (Constable, 2016). This activity gives rise to two types of sources. A low signal-to-noise ratio atmospheric noise and large signal-to-noise ratio transient events (Rakov & Uman, 2003). Both sources allow to infer informations about the state of the ionosphere (Hughes & Pappert, 1975) or atmospheric processes (Williams, 1992).

Natural geomagnetic sources such as lightning emitted waves also act as inductive sources as they induce electric currents in the earth. This is studied in the framework of audio-magnetotellurics (AMT) (Strangway *et al.*, 1973) to provide information about the electric state of the subsurface (Cagniard, 1953). Among electromagnetic (EM) methods, AMT can reach depths of several hundreds meters. Larger depths can be studied using geomagnetic sources at longer periods such as geomagnetic pulsations (Chave & Jones, 2012) in a magnetotelluric (MT) framework.

Two mains difficulties arise in AMT. The magnetic field has natural low amplitude in the so-called AMT dead band frequency, between 1 kHz and 5 kHz (Garcia & Jones, 2002). Only a few natural events have sufficient signal-to-noise ratio in this frequency band (Larnier *et al.*, 2016b). The other major issue is the contamination of time series by anthropogeneous noise. Depending on the site location, noise can be non-stationnary or above the signal level of MT signals. One issue in this case is the use of a remote station in processing schemes (Gamble *et al.*, 1979) to adress noise

on the magnetic field. Robust statistics are also necessary in this processing (Jones *et al.*, 1989). These issues have serious consequences on AMT interpretation of data acquired near housing or simply because of the reliability of AMT soundings in the AMT dead band.

One approach has been proposed by several authors in the last decades. They offer to study the magnetic sources in time series prior to process MT data. Escalas *et al.* (2013) also applied CWT on MT time series to study and characterize cultural noise sources. Larnier *et al.* (2016a) demonstrated using continuous wavelet transform that focusing on large signal-to-noise ratio allows to drastically reduce noise bias in single site estimation of MT response functions.

We propose here to extend the work of Zhang & Paulson (1997). They used continuous wavelet transform (CWT) to automatically detect atmospheric sources using thresholding. However, two major limitations arise in their methodology. On the one hand, they do not distinguish the several types of waves emitted by lightning strikes, such as atmospherics, slow tails and whistlers (Larnier *et al.*, 2016b). On the other hand, only least-squares inversion of MT response functions was performed.

In this paper, we recall a methodology previously described in Larnier *et al.*, (2016b) allowing automatic detection and characterisation of high-frequency EM sources. Using informations of geomagnetic sources, we have developed a wavelet-based approach to recover response functions and their uncertainty. This scheme relies on robust statistics and hierarchical bootstrapping of wavelet coefficients. We apply this methodology on a dataset acquired near the french magnetic observatory of Chambon-La-Forêt (CLF). The dataset presents large level of noise compared to natural sources levels in standard electric and magnetic time series. We finally demonstrate the improvement compared to standard robust processing algorithms. The quality of final Audio-MT soundings is greatly improved using the wavelet-based approach.

2 MATERIALS AND METHODS

2.1 Data sets

We conduced an AMT campaign on March 2015 near the magnetic observatory of Chambon-La-Forêt (CLF), France (see Figure 1). We focus on 3 stations B2, B4 and C5 stations (Figure 1).

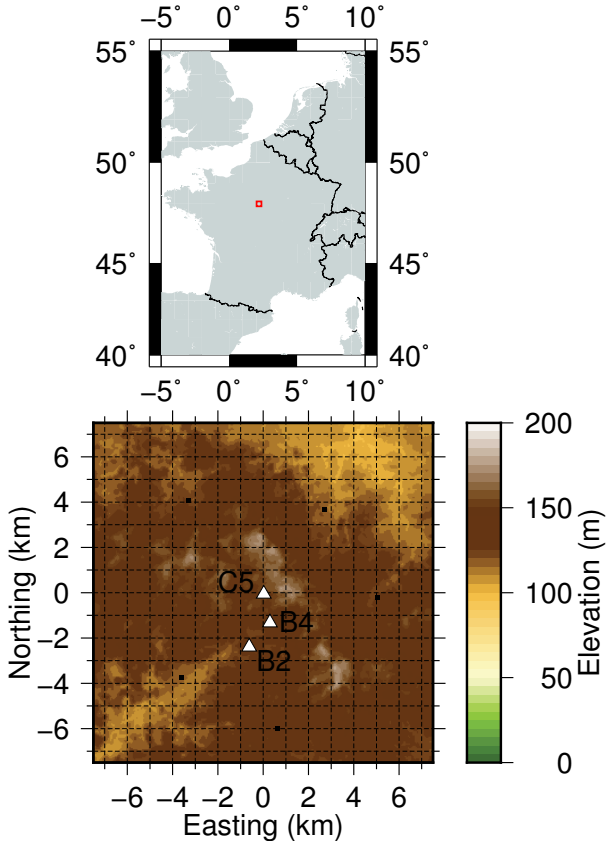


Figure 1. Map of AMT soundings made in march 2015 campaign. Top: Location in France. Bottom: Precise positioning with topography. White triangles indicate AMT stations, open black square indicates the position of CLF observatory and black squares the nearest towns in the area.

AMT time series were acquired at high frequency sampling: 30 min at 4096 Hz and 5 min at 65536 Hz. All stations use MFS07e coils from Metronix-Geophysics with Chopper Off to record the magnetic field and 30 meters dipoles with steel electrodes to measure electric fields.

As commonly observed in Europe, a lot of man-made EM sources affects EM records from 10 Hz to more than 10 kHz. Figure 2 shows such typical noise on the electric field time series at B2. It appears necessary to filter out anthropic noise before AMT interpretation. We used notch filters to get rid of as much spectral rays as possible. An high-pass filter has also been applied to reduce the amplitude of frequencies below 1 Hz. Time series are decimated down to 32768 Hz to avoid very low frequency transmitters with frequencies over 16 kHz. On the raw time series, large 50 Hz power line noise is noticeable, hiding most of EM waves. For some AMT sites, broadband noise

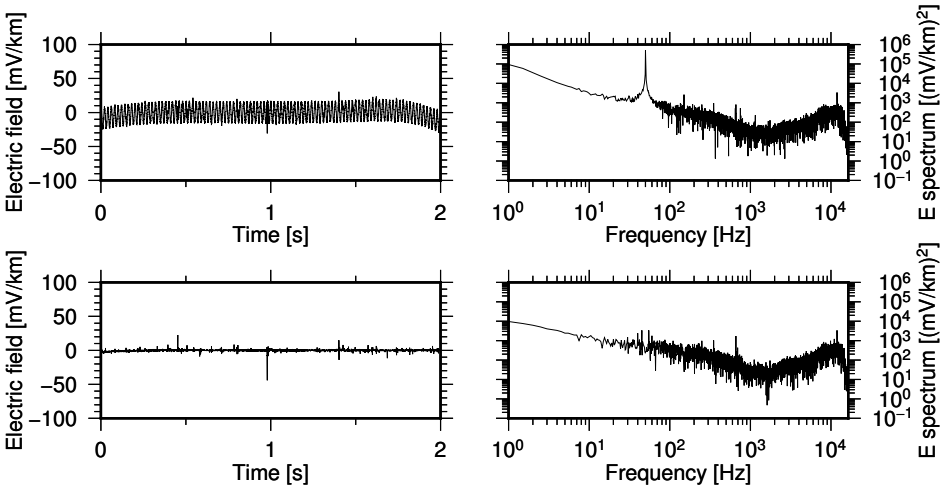


Figure 2. Difference between raw (top) and filtered (bottom) spectra for B2 station.

occurs and remains difficult to remove. On Figure 3, we have represented typical data acquired in the area. The top panel represent clean data where EM waves are easily identifiable on the time series. Still, numerous spectral rays are noticeable on the fourier spectrum. Bottom panel represent typical noisy data in the area. Many transient signals are present but not in the simultaneous clean dataset indicating large local signals. Fourier spectrum also present large noise level. The usual shape of natural electromagnetic spectrum (Constable, 2016) is not recognizable. For these sites, low signal-to-noise ratio is observed on time series.

2.2 Classical robust processing of MT time series

Magnetotellurics aims to recover the electric distribution of the subsurface through the so-called impedance response 2×2 tensor Z relating induced horizontal electric field E to inducing horizontal magnetic field H in frequency domain:

$$E = ZH \tag{1}$$

The complex quantity Z can be transformed into two more interpretable quantities, apparent resistivity ρ and phase ϕ with:

$$\begin{aligned} \rho &= |Z|^2/(\omega\mu) \\ \phi &= \arctan(\Im(Z)/\Re(Z)) \end{aligned} \tag{2}$$

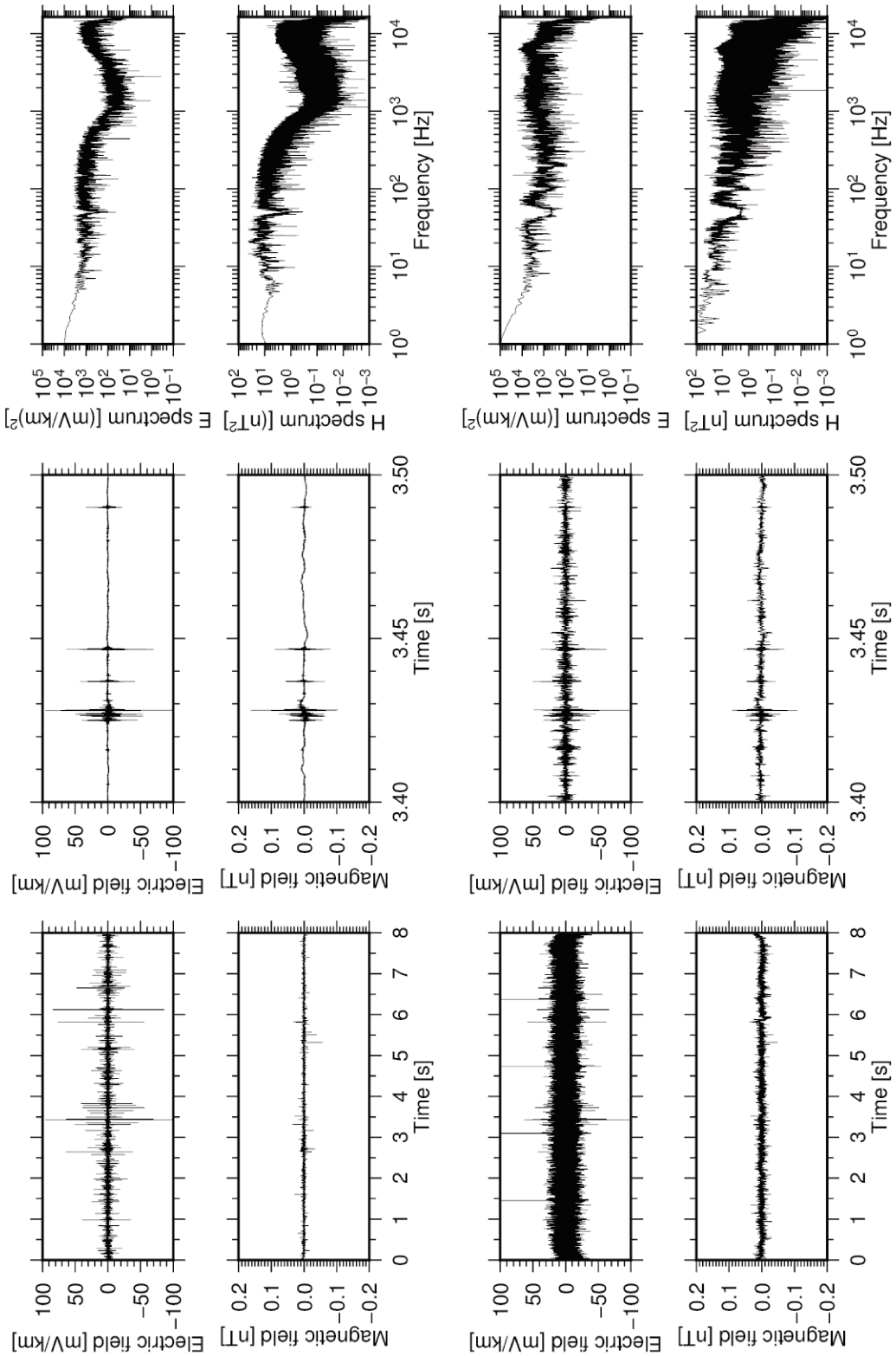


Figure 3. Difference between B2 and C5 time series. First line: Electric time series at B2 station. Second line: Magnetic time series at B2 station. Third line: Electric time series at C5 station. Fourth line: Magnetic time series at C5 station. Left: 8 seconds of time series. Middle: Zoom of 0.1 seconds. Right: Fourier spectra of the 8 seconds of data. All presented time series at B2 station and C5 are acquired simultaneously.

Most processing schemes (Chave & Thomson, 2004; Egbert & Booker, 1986) use Fourier transform and robust regression (Huber, 1973) to recover the response functions. Presence of noise in EM time series induces downward bias in MT responses. It is therefore necessary to include a reference station (Gamble *et al.*, 1979) to remove uncorrelated noise through correlation of fields of local and reference stations (hereafter referred to as R_x and R_y). In this campaign, processing of AMT stations is performed using C5 station as reference station. C5 station can be processed using any mobile station as reference. We choose B2 as reference to process C5. In this paper, we use the code BIRRP (*Bounded Influence Remote Reference Processing*) from the work of Alan D. Chave (Chave *et al.*, 1987; Chave & Thomson, 2004) to calculate reference MT soundings to compare with the new methodology.

2.3 Wavelet analysis

We decompose EM time series on a time-frequency representation using the continuous wavelet transform (CWT), based on the correlation of the 1D signal with a class of functions called wavelets (Holschneider, 1995). These functions are localized in time and frequency simultaneously. Each wavelet $\psi_{a,b}$ is derived from a function called mother wavelet ψ by dilatation factor a and parameter b:

$$\psi_{a,b}(t) = \psi\left(\frac{t-b}{a}\right) \tag{3}$$

The factor a acts as the frequency and b the time. Wavelet coefficients are derived from:

$$\mathcal{W}_\psi(a,b)[s] = \int_{-\infty}^{+\infty} dt \frac{1}{a} \psi^*_{a,b}(t) s(t) \tag{4}$$

In this paper, we use the Morlet wavelet (Goupillaud *et al.*, 1984) defined in time domain by:

$$\psi_m(at) = \pi^{-1/4} e^{i\omega_0 at} e^{-(at)^2/2}, \tag{5}$$

and in frequency domain by:

$$\widehat{\psi}_m(a\omega) = \pi^{-1/4} H(\omega) e^{-(a\omega-\omega_0)^2/2}, \tag{6}$$

where H is the Heaviside step function and ω_0 the wavelet parameter. ω_0 can be modified to adjust the time-frequency resolution of the analyzing wavelet.

This decomposition has first been used in magnetotelluric applications by Zhang & Paulson (1997). They used a signal-to-noise ratio threshold in the time-frequency domain to detect atmospherics in AMT data before using least-squares determination of response functions. Escalas *et al.* (2013) used wavelet analysis to study polarisation attributes of cultural noise sources in CSAMT experiments. Larnier *et al.* (2016a) extended Zhang & Paulson (1997) work by showing that high signal-to-noise ratio events help to reduce impedance bias estimates in single site configuration. On Figure 4, we present one excerpt of B2 time series at 32768 Hz along with corresponding magnetic remote station time series R_x and R_y. Electric time series are characterized by a relatively higher noise level compared to magnetic time series and their spectrograms. Remote time series are affected with spikes. Large signal-to-noise ratio events are visible in both electric and magnetic local time series. Only one event at 2.75 seconds is common between local and remote magnetic time series due to the large noise level at the remote AMT station. Corresponding spectrogram display three different types of EM waves, namely atmospherics, whistlers and slow tails described in Larnier *et al.* (2016b). In local magnetic time series, atmospherics display coefficients with large amplitude from more than 10 kHz down to 7-8 kHz. Whistlers display large signal-to-noise ratio wavelet coefficients from 10 kHz down to 2 kHz. Finally, slow tails display large amplitude from 100 Hz up to 1 kHz for the largest ones.

The spikes in remote magnetic time series are characterized by large amplitude coefficients for frequencies from 700 Hz up to 2-3 kHz. Only the events with the largest amplitude are recognizable in the time-frequency plane such as the slow tail at 2.75 seconds and the atmospheric at 2.64 seconds.

On Figure 5, we present an excerpt of the time series acquired at the C5 station at 4096 Hz along with their corresponding spectrograms. This station, located near the noise sources is badly tainted with noise signals such as at 45.75 and 47.4 seconds. These signal are not present in remote time series, and thus do not fulfill the quasi-uniform source hypothesis required to apply magnetotellurics and they probably originate from a nearby source. The local transient sources of noise are characterized by large amplitude wavelet coefficients for frequencies between 50 Hz and 1 kHz

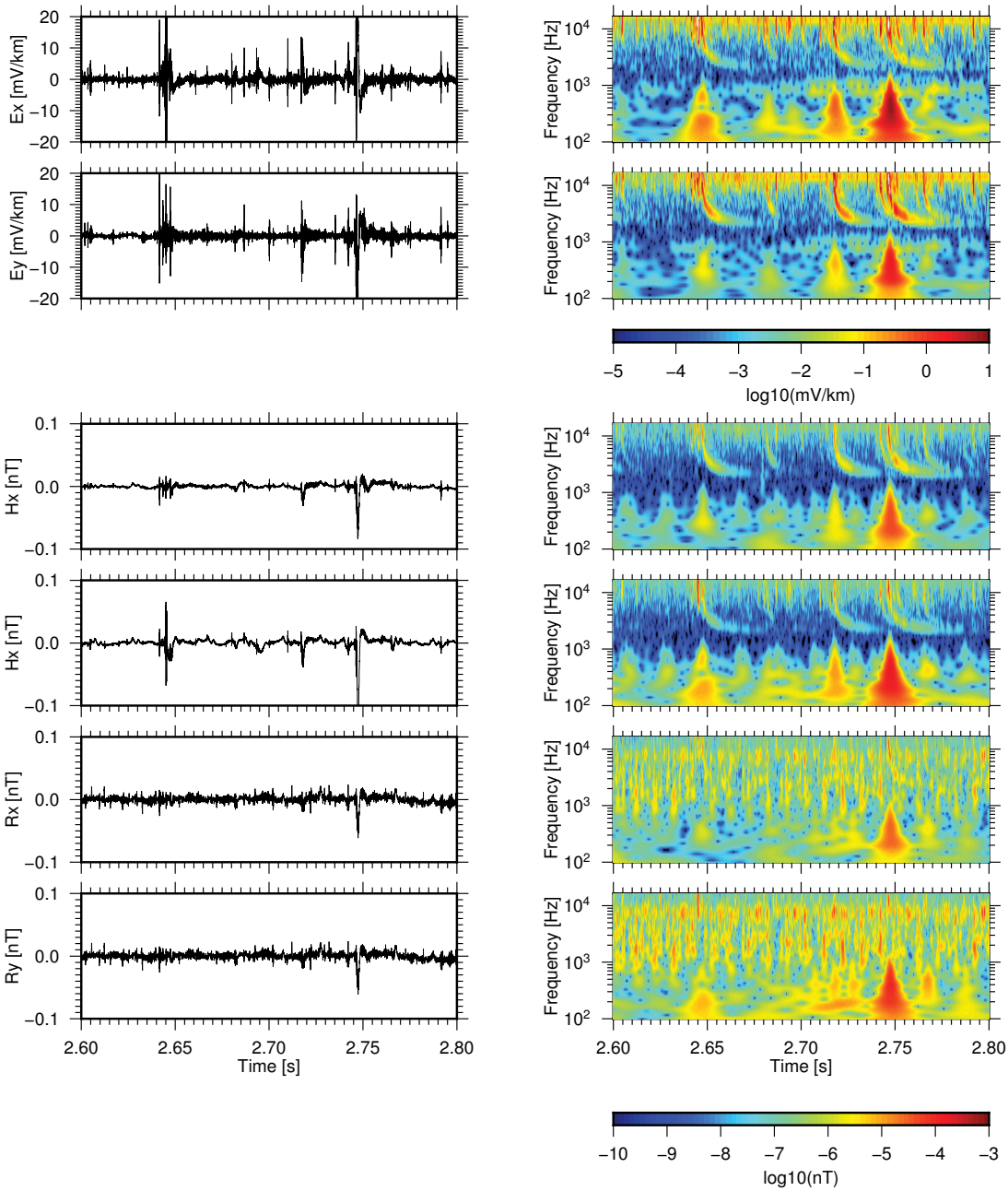


Figure 4. Extract of 4096 Hz EM time series at B2 station. From top to bottom: E_x, E_y, H_x, H_y, R_x, R_y.

but are not present at the remote station. The only significant event common in both stations is the large amplitude slow tail at 46.8 seconds.

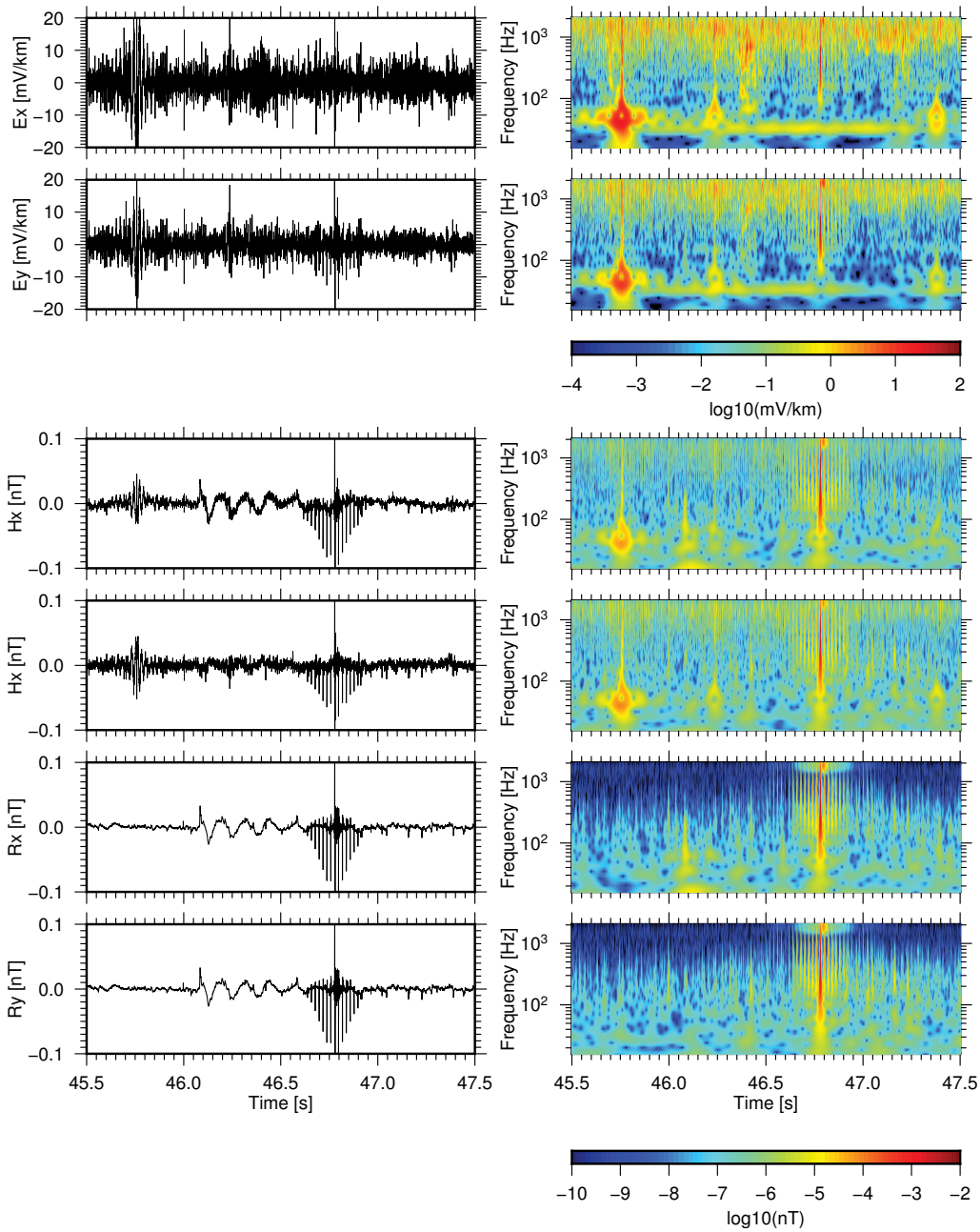


Figure 5. Extract of 4096 Hz EM time series at C5 station. From top to bottom: E_x, E_y, H_x, H_y, R_x, R_y.

2.4 Wavelet-based detection of EM waves

First step of the processing scheme follows the automatic detection and characterization of lightning-based EM waves as detailed in (Larnier *et al.*, 2016b). This methodology relies on the specific shape of EM waves in the time-frequency plane.

The sought EM waves have two distinct characteristics:

- Large signal-to-noise ratio.
- Time-frequency pattern.

We briefly recall in Appendix B the principles of Larnier *et al.*, (2016b) method of detecting EM waves in each channel. The MT system (1) can be resolved separately for each couple of response functions (z_{xx}, z_{xy}) or (z_{yx}, z_{yy}). In the following we treat each electric channel e_x and e_y separately and call e any of the processed electric channel, e_x or e_y.

For each channel e, h_x and h_y, application of the detection scheme on each channel gives a set of maxima chains. The MT system (1) can be resolved only if proper signal (i.e. following the quasi-uniform source approximation) is present in all channels. Therefore only common maxima chains in the three channels are kept. If a remote station is available, we also add the constraint of keeping maxima chains if they are present in all channels in (e, h_x, h_y, r_x, r_y).

After detection, let N be the number of maxima chains (i.e. events) output by the detection process. Each event i is characterized by a maxima chain c_i. This chain is described by an extent in frequency (and scale) given by two indices $j_{i,1}$ and $j_{i,2}$ and a set of time values $b_{i,j}$ such that:

$$c_i = \{b_{i,j},\ j_{i,1} < j < j_{i,2}\} \tag{7}$$

In a way similar to the characterisation of polarization attributes in Larnier *et al.*, (2016b), we describe the amplitude and phase of each events by a set of wavelet coefficients. At each scale j, we define a correlation length d_j based on the size of Morlet kernel above a critical value (Maraun *et al.*, 2007). The event i is then characterized for each channel s by w_{ij}, a subset of neighbour coefficients defined by:

$$w_{ij}[s] = \{\mathcal{W}_\psi(a_j, b)[s], |b - b_{i,j}| < d_j\}, \tag{8}$$

The event i is then described by $\mathcal{W}_i$, the collection of w_{ij} at each scale. Each subset w_{ij} represents the wavelet coefficients stored at scale j for geomagnetic event i. Then:

$$\#\mathcal{W}_i = \sum_{j=1}^{n} \#w_{ij} \tag{9}$$

At each scale, the set of selected coefficients among all events is:

$$\#w_j = \sum_{i=1}^{N} \#w_{ij} \tag{10}$$

2.5 Wavelet-based response functions estimation

To recover the estimation of impedance response functions at each frequency, we follow the methodology of Zhang & Paulson (1997). The authors have derived the expression of the MT system in time-frequency domain:

$$\begin{aligned} e_x(a,b) &= Z_{xx}(a)h_x(a,b) + Z_{xy}(a)h_y(a,b) \\ e_y(a,b) &= Z_{yx}(a)h_x(a,b) + Z_{yy}(a)h_y(a,b) \end{aligned} \tag{11}$$

However, they used least-squares to resolve this system on wavelet coefficients of only two events. Classical least-squares has been proven biased in the presence of noise. We choosed to use robust statistics to reduce noise influence in the determination of Z. Each equation is solved at each scale $a = a_j$ following the robust procedure of Huber weights (Huber, 1973). Robust statistics were introduced in MT to allow for outliers in the computation of the impedance response function (Chave *et al.*, 1987). We use the most common robust method, which involves M-estimators to deal with the previous system as univariate regression problems written as follows:

$$\mathcal{W}_\psi(a_j,\cdot)[e] = \mathbf{z} \times \mathcal{W}_\psi(a_j,\cdot)[h] + \varepsilon, \tag{12}$$

where $\mathcal{W}_\psi(a_j,\cdot)[e]$ a $(1, \#w_j)$ vector, $\mathcal{W}_\psi(a_j,\cdot)[h]$ a $(2, \#w_j)$ vector, and $\mathbf{z}$ is a $(1,2)$ tensor. $\times$ is the matrix product and ε are the residuals on the electric field. The remote reference is easily included in this scheme as described by Chave & Thomson (2004).

2.6 Wavelet-based hierarchical bootstrap standard error

One effective method to recover the estimation of standard deviation of a given dataset is the bootstrap approach (Efron & Tibshirani, 1993). Let $\mathbf{x}$ be the considered dataset of N samples identically and independently distributed:

$$\mathbf{x} = \{x_1, x_2, ..., x_n\} \tag{13}$$

The original dataset $\mathbf{x}$ has a statistic of interest $s(\mathbf{x})$. This statistic could be the sample mean or in our case the impedance response functions Z.

The bootstrap approach principle is to create B bootstrap samples $\mathbf{x}^{*\mathbf{b}}$ by using replacement in the original dataset. Each new bootstrap sample gives one bootstrap replication $\hat{\theta}^*(b) = s(\mathbf{x}^{*\mathbf{b}})$ with s being the same statistic applied to $\mathbf{x}$. After estimating all bootstrap replications, the bootstrap standard error is given by:

$$se_F = \left\{ \sum_{b=1}^{B} [\hat{\theta}^*(b) - \bar{\theta}^*]^2/(B-1) \right\}^{1/2}, \tag{14}$$

with $\bar{\theta}^* = \sum_{b=1}^{B} \hat{\theta}^*(b)/B$ being the mean of the bootstrap dataset. This approach relies on extensive computations but is easy to implement.

In our case, the statistic of interest is the robust regression defined previously and the dataset is constituted of wavelet coefficients of electric and magnetic time series. However, close wavelet coefficients are by definition correlated through the wavelet kernel. Direct bootstrapping of wavelet coefficients is therefore forbidden due to the independence condition of samples in the dataset. Still, wavelet coefficients can be sorted through the sets of coefficients w_{ij} belonging to each geomagnetic event. Each set w_{ij} is independent of the others.

One way to bootstrap such hierarchical data is to bootstrap the sets of dependent coefficients. Two levels have to be distinguished. The lowest represents the coefficients themselves. The second one is the level of clusters of dependent coefficients w_{ij}. We use the approach of hierarchical bootstrap (Field & Welsch, 2007) based on the resampling of clusters instead of samples inside clusters. Replacement is then made at the cluster level.

At each scale j, the dataset of available events is:

$$\mathbf{w_j} = \{w_{1j}, w_{2j}, ..., w_{Nj}\} \tag{15}$$

Each bootstrap sample is then:

$$\mathbf{w_j^b} = \{w_{b_1 j}, w_{b_2 j}, ..., w_{b_N j}\} \tag{16}$$

Each bootstrap sample gives an expectation value through the robust determination of impedance tensor:

$$\hat{\theta}_j^*(b) = z_j^b \tag{17}$$

The standard error associated to each element of the impedance tensor at scale j is then given by:

$$se_{z_j} = \left\{ \sum_{b=1}^{B} [\hat{\theta}_j^*(b) - \bar{\theta}_j^*]^2/(B-1) \right\}^{1/2}, \tag{18}$$

with $\bar{\theta}_j^* = \sum_{b=1}^{B} \hat{\theta}_j^*(b)/B$.

To avoid error transmission, we directly compute standard errors on phase and resistivity through the bootstrap samples estimations $\hat{\theta}_{\ln(\rho_j)}^*(b) = \ln(\rho_j^b)$ and $\hat{\theta}_{\phi_j}^*(b) = \arctan(\Im(z_j)^b/\Re(z_j)^b)$ Assuming a log-normal distribution for resistivity, the standard error for $\ln(\rho)$ is given by:

$$se_{\ln(\rho_j)} = \left\{ \sum_{b=1}^{B} [\hat{\theta}_{\ln(\rho_j)}^*(b) - \bar{\theta}_{\ln(\rho_j)}^*]^2/(B-1) \right\}^{1/2}, \tag{19}$$

with $\bar{\theta}_{\ln(\rho_j)}^* = \sum_{b=1}^{B} \hat{\theta}_{\ln(\rho_j)}^*(b)/B$. The variance of ρ_j is then given by:

$$\sigma_{\rho_j} = (e^{se_{\ln(\rho_j)}} - 1)\exp(e^{\bar{\theta}_{\ln(\rho_j)}^*} + se_{\ln(\rho_j)}/2) \tag{20}$$

The phase standard error is given by application of circular statistics on the bootstrap samples $\hat{\theta}_{\phi_j}^*(b)$ (Fisher, 1995). This methodology relies on the sum of vectors defined by the angles $\hat{\theta}_{\phi_j}^*(b)$. Let C et S be defined by:

$$\begin{aligned} C &= \sum_{i=1}^{B} \cos(\hat{\theta}_{\phi_j}^*(b)) \\ S &= \sum_{i=1}^{B} \sin(\hat{\theta}_{\phi_j}^*(b)) \end{aligned} \tag{21}$$

The bootstrap phase mean $\bar{\theta}_{\phi_j}^*(b)$ is given by:

$$\bar{\theta}_{\phi_j}^*(b) = \arctan(S/C) \tag{22}$$

Time series frequency	32768 Hz	4096 Hz
Frequency band	2 - 16 kHz	32 Hz - 2 kHz
Considered waves	Atmospherics, whistlers	Slow tails
Window length: (duration)	short: 16 seconds long: 2 minutes	short: 2 minutes long: 16 minutes
Window length: (number of samples)	short: 2 * 262144 long: 15 * 262144	short: 2 * 262144 long: 15 * 262144

Table 1. Description of processing methodology.

The bootstrap phase standard error is given by:

$$se_{\phi_j} = \left\{ \sum_{b=1}^{B} [d(\hat{\theta}^*_{\phi_j}(b), \bar{\theta}^*_{\phi_j})]^2/(B-1) \right\}^{1/2}, \tag{23}$$

where $d(\phi_1, \phi_2)$ is the distance in radians between two angles ϕ_1 and ϕ_2 on the unit circle.

ELF time series						
Station	B4		B2		C5	
Length	Short	Long	Short	Long	Short	Long
K	0.95	0.95	0.97	0.90	0.95	0.98
α	0.90	0.90	0.90	0.99	0.90	0.90
$[f_{min}, f_{max}]$ (Hz)	[64, 500]	[64, 900]	[64, 400]	[128, 900]	[64, 300]	[128, 700]
VLF time series						
Station	B4		B2		C5	
Length	Short	Long	Short	Long	Short	Long
K	0.99	0.99	0.80	0.80	0.99	0.99
α	0.99	0.99	0.90	0.90	0.99	0.99
$[f_{min}, f_{max}]$ (kHz)	[7.5, 15]	[6.5, 15]	[5, 12]	[4, 12]	[8, 15]	[7, 15]

Table 2. Wavelet-based processing parameters used in the processing of AMT time series.

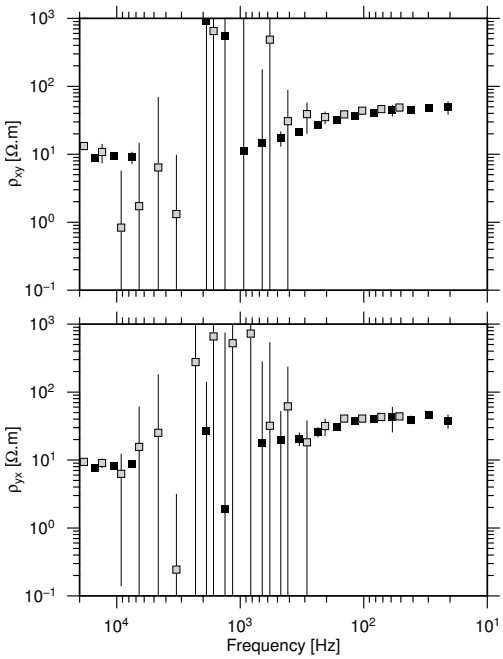


Figure 6. Comparison between robust (grey squares) and wavelet based processing (black squares) for station B4 using the short time series. Top: ρ_{xy}. Bottom: ρ_{yx}. Robust results are shifted to the left for visual clarity.

3 APPLICATION

We illustrate the present methodology on three AMT stations, B4, B2, and C5. These stations were chosen because of their noise level from light (B4), to moderate (B2) and high (C5). To illustrate the influence of geomagnetic activity, we will illustrate the present methodology on short and long time series. The description of the processing scheme of time series for robust and present methodology is described Table 1.

3.1 Comparison with standard processing

We apply the previously introduced methodology on time series presented in section 2.1. The number of bootstrap iterations was set to $B = 1000$.

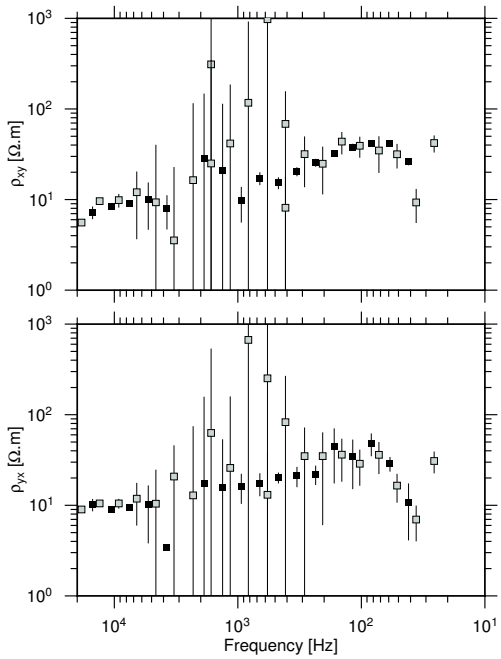


Figure 7. Comparison between robust (grey squares) and wavelet based processing (black squares) for station B2 using the short time series. Top: ρ_{xy}. Bottom: ρ_{yx}. Robust results are shifted to the left for visual clarity.

3.1.1 Short time series

On Figure 6, 7 and 8 we present the results obtained on the 4096 Hz and 32768 Hz short time series for B4, B2 and C5 station respectively. The processing parameters specified are summarized Table 2. Classical robust processing and the present methodology agree very well for frequencies below 200 Hz and above 10 kHz with small confidence intervals for the three stations. Wavelet-based results present smooth sounding and relatively narrow confidence invervals for frequencies below 1 kHz and above 6 kHz for B4 station, below 1 kHz and above 4 kHz for B2 station and below 500 Hz and above 7 kHz for C5 station. The width of the AMT dead band is then reduced compared to robust response functions. Still, a large frequency gap may be observed from 3-5 kHz down to 500 Hz - 1 kHz depending on the station.

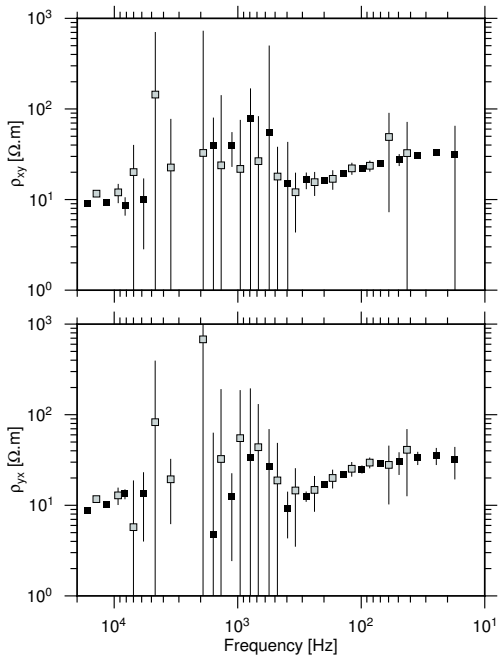


Figure 8. Comparison between robust (grey squares) and wavelet based processing (black squares) for station C5 using the short time series. Top: ρ_{xy}. Bottom: ρ_{yx}. Robust results are shifted to the left for visual clarity.

3.1.2 Long time series

Larnier *et al.*, (2016b) showed that large events penetrating the AMT dead band have a frequency of occurence much lower than most of transient events. Recording time therefore needs to be set to longer duration to collect such events. We then applied the same methodology on a much longer time seriesto keep only events penetrating the AMT dead band. The difference between sought events in short time series and long time series are described in Table 2. Processing results are showed Figure 9, 10 and 11 for B2, B4 and C5 AMT station respectively. Robust processing results are improved for frequencies below 400 Hz and above 5 kHz but are still erratic from 400 Hz to 5 kHz with large confidence intervals. Wavelet-based processing show improvement for the three stations: B4 and B2 stations have a reduced dead band from 1.5 kHz up to 3-4 kHz. C5 station response functions have the AMT dead band narrowed down 1 kHz - 6 kHz. This makes the interpretation much more reliable and easier on wavelet based response functions than using the reference robust sounding obtained using BIRRP code. For the three AMT stations, setting

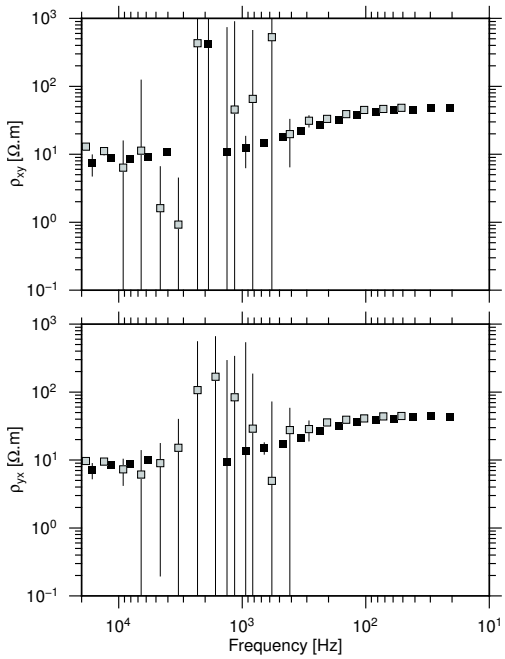


Figure 9. Comparison between robust (grey squares) and wavelet based processing (black squares) for station B4 using the long time series. Top: ρ_{xy}. Bottom: ρ_{yx}. Robust results are shifted to the left for visual clarity.

the algorithm to detect events with frequency content larger than described in Table 2 led to no detected events at all. This results in no determination of response functions.

3.1.3 Accuracy of wavelet-based soundings

Discriminating between two propositions of MT soundings remains a complicated task. One way to rule out a sounding is to examine its smoothness. Indeed, physics of EM induction impose smoothness of MT response functions. Using this criteria, wavelet-based results may be considered as significantly better than robust processing results. However, smooth response functions is not a proof of correctness. Another criterion of choice is to consider local geology to be related to the MT sounding through realistic resistivity structures.In this area, we have examined local geology using the BRGM database obtained through the website Infoterre (http://infoterre.brgm.fr/). The area of Chambon-La-Forêt magnetic observatory in the Paris basin magnetic C5 has been scouted for oil exploration in 1983. Numerous shallow wells were dug and

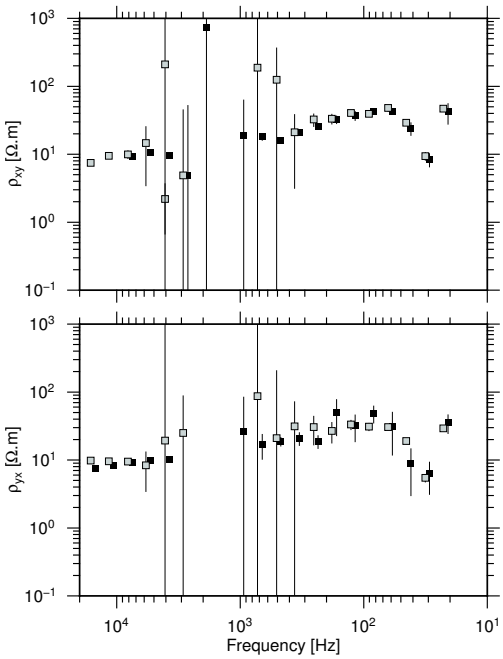


Figure 10. Comparison between robust (grey squares) and wavelet based processing (black squares) for station B2 using the long time series. Top: ρ_{xy}. Bottom: ρ_{yx}. Robust results are shifted to the left for visual clarity.

geological logs were kept. Among all of geological logs, we choosed as relevant those close to our sites (03642X0069/VT-286 and 03642X0085/LOI018). We represent on Figure (12, top) a typical geological log obtained in the area. Resistivity values were chosen from EM and Schumberger sounding results at a close site by Valla (1991). The apparent resistivity value corresponding to the highest frequency was chosen as those of the substratum, the closest layer to the surface. Using these assumptions, we obtain 1D synthetic soundings given in Figure 12 bottom. Local wells were performed more than 500 meters from our soundings location, their main features are in agreement with wavelet-based results. Indeed the decrease of apparent resistivity from 10 kHz to 1000 kHz is observable on the C5 sounding. The increase from 1 kHz to 100 Hz is coherent with all soundings and the stabilization at 100 Hz also. We can therefore consider that determination of the MT response functions from wavelet coefficients are correct.

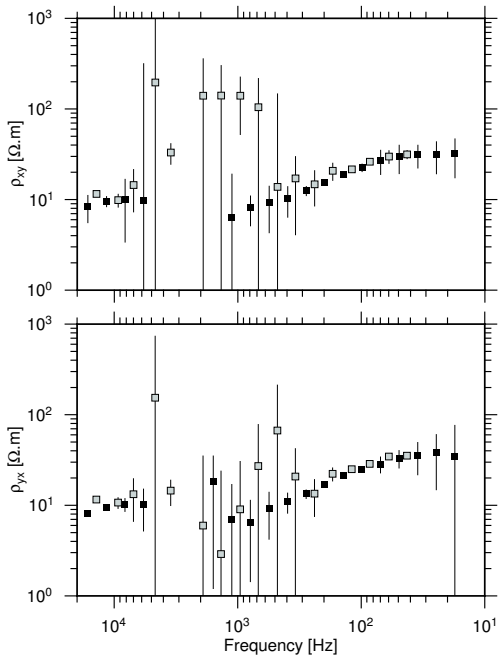


Figure 11. Comparison between robust (grey squares) and wavelet based processing (black squares) for station C5 using the long time series. Top: ρ_{xy}. Bottom: ρ_{yx}. Robust results are shifted to the left for visual clarity.

3.2 Residuals

Residuals are an important quantity to assess the quality of the output response function. Residuals in the regression problem involving wavelet coefficients are computed at each scale a_j using equation (24).

$$r = \mathcal{W}_\psi(a_j, \cdot)[e] - \mathbf{z} \times \mathcal{W}_\psi(a_j, \cdot)[h] \tag{24}$$

Following Chave & Thomson, (2004), we assume through the regression model that real and imaginary parts of residuals follow normal distributions so their absolute value should be Rayleigh distributed. To check this assumption, we represent the residuals against a standardized Rayleigh distribution through Quantile-Quantile (QQ) plots. If residuals distribution are Rayleigh distributed, such plots appear as straight lines. Two distributions are compared in the QQ-plots. The first one corresponds to residuals calculated on all wavelet coefficients at scale a_j, the other is the residuals distribution corresponding only to selected coefficients through the procedure described above.

Figure (13) illustrates two QQ plots at two different scales for the processing of B2 station us-

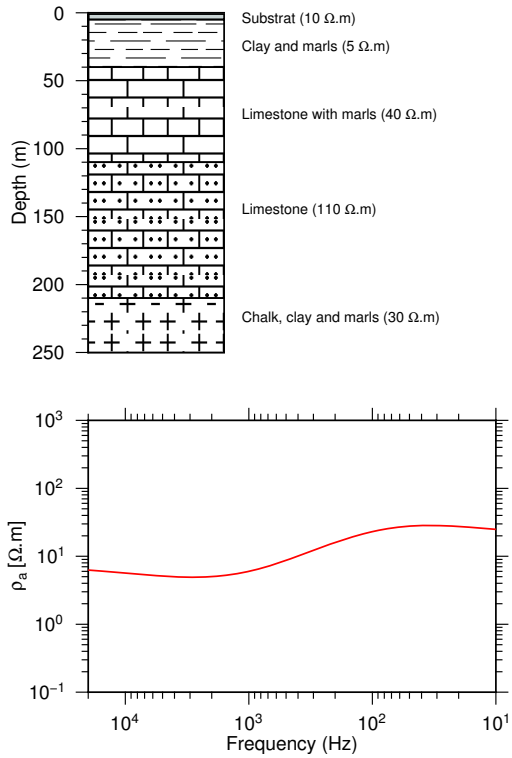


Figure 12. Top: Representative geology of CLF area. Bottom: 1D MT associated MT sounding.

ing long time series. One striking difference between the two distributions is the amplitude scale. Keeping all wavelet coefficients leads to larger residuals than pre-selecting a set of wavelet coefficients before computing the response functions. Distributions for selected coefficients appear irregular due to the low number of selected wavelet coefficients but are still consistent with Rayleigh distributions. Residuals for the whole time series appear unconsistent with Rayleigh distributions.

3.3 Influence of the number of bootstrap iterations

We also check if the number of bootstrap iterations is critical in the final estimation of standard error. We perform 6 determinations of standard deviations of response functions by varying the number of bootstrap iterations from 10 to 1000. Results are shown Figure 14. 10 iterations is clearly unstable with large difference with the other curves. However, from 50 iterations, results

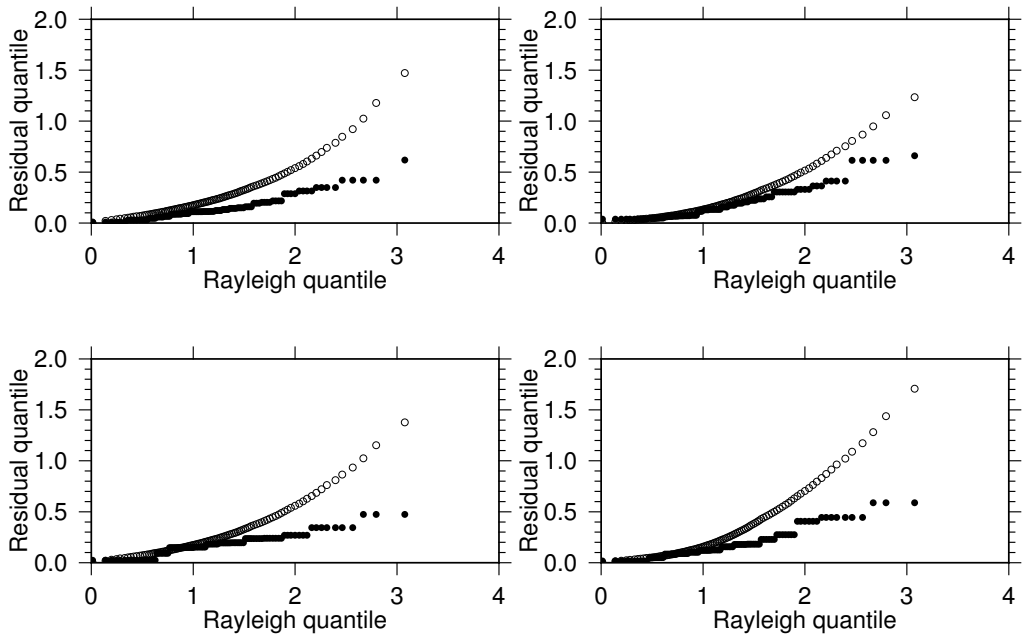


Figure 13. Quantile-Quantile plots of residual distributions against standardized Rayleigh distribution. Top: North-South component residuals. Bottom: East-West component residuals. Left: Distribution at scale corresponding to 512 Hz. Right: Distribution at scale corresponding to 140 Hz. Open circles corresponds to the residuals of the whole time series, filled corresponds to the selected events residuals.

are very similar for frequencies below 300 Hz. Above 300 Hz, decreasing the number of iterations tends to underestimate the standard deviation.

4 DISCUSSION

Even though MT data have been collected in an unfavorable period for AMT soundings (Garcia & Jones, 2002), we have shown that large frequency band sferics allow to recover the impedance response functions in most of the AMT frequency band. The drawback is that AMT soundings have to be set for longer time periods to collect events reaching in the AMT dead band. Such pre-processing step was found critical before using standard robust statistics in the determination of MT response functions. Indeed, classical processing results are not as significantly improved as wavelet processing when using longer time series, meaning that it is necessary to adapt the methodology to such transients events. Combined with the information about seasonal variation in

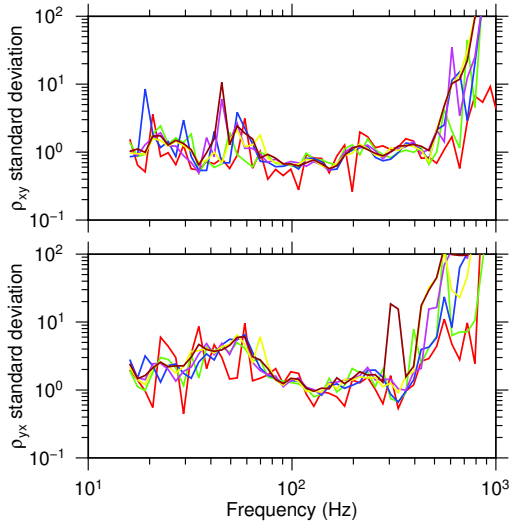


Figure 14. Variation of final standard deviation with the number of bootstrap iterations. Red: 10 bootstrap iterations. Green: 50 bootstrap iterations. Blue: 100 bootstrap iterations. Magenta: 200 bootstrap iterations. Yellow: 500 bootstrap iterations. Brown: 1000 bootstrap iterations.

the AMT dead band (Garcia & Jones, 2002), this methodology could lead to a great improvement in the planification and processing of AMT campaigns.

The other main application of this work is the ability to process AMT data in noisy environnement. In the area, several types of noise sources were present whether they are transient or stationary. Identifying geomagnetic sources in time series was critical to the processing. This was eased by the use of a remote station. Indeed, the property of quasi-uniformity of MT sources allows to discriminate between local noise and MT sources.

Finally, Chave (2014) demonstrated that residuals for low-frequency geomagnetic sources appear to be stably distributed. Preliminary results from this study suggests that high-frequency sources related to atmospheric activity may not need to rely on such assumption for residual distributions. This need to be further examined in future work.

5 CONCLUSION

In this paper, we have described a methodology to process AMT time series based on the continuous wavelet transform. This technique relies on the detection of transient signals in AMT time

series that uses time-frequency characteristics of EM waves emitted from lightning strikes. The selection step allows to select relevant wavelet coefficients in the the time-frequency domain before performing determination of MT response functions. This step is performed through robust inversion of response functions, associated to a boostrap approach allowing uncertainties to be derived. We have applied this methodology on a dataset from Chambon-La-Forêt magnetic area. Compared to a standard robust processing technique, this methodology allows to recover smoother results on larger frequency bands. Results were found in good agreement with local geology and thus can be considered as better than standard robust processing results.

We also emphasize that although processing method based on the use of transients cannot solve all issues related to the AMT dead band, we nevertheless have shown that some signals such as whistlers may help to reduce the uncertainty even if the signal-to-noise ratio of VLF waves remains low for a large frequency range. These signals are rare but should be shought.

In addition to the processing of AMT or MT survey data, the wavelet transform MT processing technique has also potential applications to MT monitoring with control on time-frequency resolution and accuracy, but this is beyond the scope of this paper. Such techniques will allow the AMT sounding method to be used now in urbanized areas or in the vicinity of industrial zones to address pollution issues or to contribute to the preservation of groundwater resources.

REFERENCES

Cagniard, L., 1953, Basic theory of the magnetotelluric method of geophysical prospecting, *Geophysics*, **18**, 605–635.

Chave, A. D., Thomson, D. J., Ander, M. E., 1987, On the robust estimation of power spectra, coherences and transfer functions, *J. geophys. Res.*, **92**(B1), 633-648.

Chave, A. D., Thomson, D. J., 2004, Bounded influence magnetotelluric response function estimation, *Geophys. J Int.*, **130**(2), 475–496.

Chave, A.D., Jones, A. G., 2012, The Magnetotelluric Method: Theory And Practice, *Cambridge University Press*, Cambridge

Chave, A. D., 2014, Magnetotelluric data, stable distributions and impropriety: an existential combination, *Geophys. J Int.*, , **198**, 622-636.

Constable, C., 2016, Earth's electromagnetic environment, *Survey in Geophysics*, **37**, 27–45.

Efron, B, & Thishirani, R. J., 1993, An introduction to the bootstrap, *Chapman & Hall/CRC*, 436pp.

Egbert, G. D., Booker, J. R., 1986, Robust estimation of geomagnetic transfer functions, *Geophys. J. R. astr. Soc.*, **87**, 173–194.

Escalas, M., P. Queralt, J. Ledo, A. Marcuello, Polarisation analysis of magnetotelluric time series using a wavelet-based scheme: A method for detection and characterisation of cultural noise sources, *Physics of the Earth and Planetary Interiors*, **218**, 31–50.

Farge M., 1992, Wavelet transforms and their applications to turbulence. *Annu. Rev. Fluid Mech.*, **24**, 395–457

Field, C. A., Welsh, A. H., 2007, Bootstrapping clustered data, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **69**, 369–390.

Fisher, N. I., 1995, Statistical analysis of circular data, *Cambridge University Press*, Cambdridge, 277pp.

Gamble, T.D., Goubau W.M., Clarke J., 1979, Magnetotellurics with a remote magnetic reference, *Geophysics*, **44**, 53–68.

Garcia, X., Jones, A. G., 2002, Atmospheric sources for audio-magnetotelluric (AMT) sounding, *Geophysics*, **67**(2), 448–458.

Goupillaud, P., Grossmann, A., Morlet, J., 1984, Cycle-octave and related transforms in seismic signal analysis, *Geoexploration*, **23**, 85–102.

Helliwell, R. A., 1965. Whistlers and related ionospheric phenomena, *Stanford University Press*.

Holschneider, M., 1995, Wavelets, an analysis tool, *Oxford Science Publications*, Oxford.

Huber, P. J, 1973, Robust regression: asymptotics, conjectures and Monte Carlo, *The Annals of Statistics*, **1**, 799–821.

Hughes, H. G. & Pappert, R. A., 1975, Propagation prediction model selection using VLF atmospherics. *Geophys. Res. Lett.*, **2**, 96–98

Jones, A. G., Chave, A. D., Auld, D., Bahr, K., Egbert, G., 1989. A comparison of techniques for magnetotelluric response function estimation, *J. geophys. Res.*, *Solid Earth*, **94**(14), 14 201–14 213.

Larnier, H., Sailhac, P., Chambodut, A., 2016, New application of wavelets in magnetotelluric data processing: Reducing impedance bias, *Earth Planets Space*,**68**(70).

Larnier, H., Sailhac, P., Chambodut, A., 2016, Detection and characterisation of lightning-based sources using continuous wavelet transform: application to audiomagnetotellurics, *Geophys. J Int.*, ***to be submitted***

Maraun, D., Kurths, J., Holschneider, M., 2007. Nonstationary gaussian processes in wavelet domain: Synthesis, estimation, and significance testing, *Physical Review E*, **75**.

Nabighian, M. N., 1991, Electromagnetic Methods in Applied Geophysics: Volume 2, Application, Parts A and B, *Society of Exploration Geophysicists*, 991pp.

Rakov, V. A. & Uman, M. A., 2003, Lightning: Physics and Effects, *Cambridge University Press*, Cam-

bridge.

Strangway, D. W., Swift, C. M., Holmer, R. C., 1973, The application of audio frequency magnetotellurics (AMT) to mineral exploration, *Geophysics*, **38**, 1159–1175.

Torrence, C., Compo, G. P., 1998, A practical guide to wavelet analysis, *Bulletin of the American Meterological Society*, **79**:1, 61–78.

Valla, P., 1991, BRGM Melis multifrequency EM system. Electromagnetic methods in applied geophysics, **2**, 417–421.

Williams, E. R., 1992, The Schumann Resonance: A Global Tropical Thermometer, *Science*, **256**, 1184–1187.

Zhang, Y., Paulson, K. V., 1997, Enhancement of signal-to-noise ratio in natural-source transient MT data with wavelet transform, *Pure and Applied Geophysics*, **149**, 405–419.

APPENDIX A: WAVELET KERNEL AND CORRELATION LENGTH

The wavelet kernel K_ψ of the wavelet ψ characterizes the correlation of two differents coefficients at positions (a_1, b_1) and (a_2, b_2) in the scale-time domain (Farge, 1992; Holschneider, 1995). This correlation is defined by:

$$K_\psi(a_1, a_2, b_1, b_2) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \frac{1}{a_1} \psi\left(\frac{t - b_1}{a_1}\right) \frac{1}{a_2} \psi\left(\frac{t - b_2}{a_2}\right) dt \tag{A.1}$$

where C_ψ is the following normalizing factor defined from the Fourier transform of the wavelet ψ:

$$C_\psi = \int_{-\infty}^{\infty} |\hat{\psi}_0(\omega)| \frac{d\omega}{|\omega|} \tag{A.2}$$

Maraun *et al.* (2007) derived the Morlet wavelet kernel:

$$\begin{aligned} K_\psi(a_1, a_2, b_1, b_2) = & \frac{2a_1 a_2}{a_1^2 + a_2^2} \exp\left(i\omega_0 \frac{a_1 + a_2}{a_1^2 + a_2^2}(b_2 - b_1)\right) \\ & \exp\left(-\frac{1}{2} \frac{(b_2 - b_1)^2 + \omega_0 (a_2 - a_1)^2}{a_1^2 + a_2^2}\right) \end{aligned} \tag{A.3}$$

This kernel is used to defined a time resolution. By choosing a critical correlation amplitude c, we define the resolution time limit at scale a by:

$$\begin{aligned} l_{\psi,c}(a) = & \max(b_2 | K_\psi(a, a_2, b_1, b_2) > c, \forall a_2, \forall b_2) \\ & - \min(b_2 | K_\psi(a, a_2, b_1, b_2) > c, \forall a_2, \forall b_2) \end{aligned} \tag{A.4}$$

APPENDIX B: DETECTION OF EM WAVES

The detection of lightning-based EM waves is based on two main properties:

- High amplitude of wavelets coefficients.
- Dispersion patterns in time-frequency plane.

B1 Methodology of significant coefficients

To recover large amplitude coefficients, we use the methodology of significant coefficients of Torrence & Compo (1998). They showed that for white or red noise of variance σ^2, the distribution $D(a_j) = \{|\mathcal{W}_\psi(a_j, b_i)[s]|^2,\ i = 1...N\}$ at each scale a_k follows a χ^2 distribution:

$$D(a_j) \rightarrow \frac{1}{2} P_j \sigma^2 \chi_2^2, \tag{B.1}$$

where P_j is the Fourier background spectrum corresponding to scale a_j. We then use the assumption that the global wavelet spectrum defined by the equation (B.2) is a good approximation of the background spectrum $P_j\sigma^2$.

$$G_\psi^2(a)[s] = \int_{-\infty}^{+\infty} db |\mathcal{W}_\psi(a, b)[s]|^2, \tag{B.2}$$

Then, using a confidence level α, one can derive a threshold η at each scale defined by:

$$\eta(a_j, \alpha) = \frac{1}{2} G_\psi^2(a_j)[s] \chi_2^2(\alpha) \tag{B.3}$$

All coefficients above this threshold are considered as significant.

B2 Maxima chains

In the time-frequency plane, a local maxima (a_j, τ_i) is defined at each scale a_j by:

$$\begin{aligned} |\mathcal{W}_\psi(a_j, \tau_i)| &> |\mathcal{W}_\psi(a_j, \tau_{i+1})| \\ |\mathcal{W}_\psi(a_j, \tau_i)| &> |\mathcal{W}_\psi(a_j, \tau_{i-1})| \end{aligned} \tag{B.4}$$

Local maxima at each scales are chained using the following chain rule: We follow the dispersion rule by Helliwell, (1965),

$$\Delta t = D f^{-1/2}. \tag{B.5}$$

For slow tails, D is zero (no apparent dispersion). Each scale a_j corresponds to a specific frequency, thus a time shift $\Delta_{t,j}$. Let f_{min} and f_{max} be the minimum and maximum frequencies expected for each event. For example 3 kHz and 10 kHz respectively for atmospherics or 64 Hz and 1 kHz for slow tails. We start from a specific scale a_α defined as the mean (in logarithm) of f_{min} and f_{max}.

(i) At starting scale a_α, we seek every local maxima. Let M_α be the set of all local maxima at scale a_α.

(ii) For every $\tau_{i_\alpha} \in M_\alpha$, we connect the local maxima $(a_\alpha, \tau_{i_\alpha})$ to a maximum on larger scale if:

$$\exists \tau_{i_{\alpha+1}} \begin{cases} \tau_{i_{\alpha+1}} \in M_{\alpha+1} \\ \tau_{i_{\alpha+1}} \in [\tau_{i_\alpha} + \Delta_{t,\alpha} - l_{\psi,c}(a_{\alpha+1})/2 : \\ \qquad \tau_{i_\alpha} + \Delta_{t,\alpha} + l_{\psi,c}(a_{\alpha+1})/2] \end{cases} \tag{B.6}$$

To connect it to a maximum on smaller scale:

$$\exists \tau_{i_{\alpha-1}} \begin{cases} \tau_{i_{\alpha-1}} \in M_{\alpha-1} \\ \tau_{i_{\alpha-1}} \in [\tau_{i_\alpha} - \Delta_{t,\alpha} - l_{\psi,c}(a_{\alpha-1})/2 : \\ \qquad \tau_{i_\alpha} + \Delta_{t,\alpha} + l_{\psi,c}(a_{\alpha-1})/2] \end{cases} \tag{B.7}$$

Flexibility in the search of local maxima is allowed through the time resolution $l_{\psi,c}(a_{\alpha-1})$. Only maxima chains continuous from at least f_{min} to f_{max} are kept in AMT processing.

Depending on the frequency limits, the confidence level α, and the kernel critical value, the number of detected events varies (Larnier *et al.*, 2016b).

APPENDIX C: FIGURES

3.4.3.3 Pulsations géomagnétiques

L'application de la segmentation des coefficients en groupes d'événements géomagnétiques devient problématique dans le cas des pulsations géomagnétiques. Les ondes EM issues d'éclairs atmosphériques étant impulsives, il est aisé de les délimiter dans le temps. Les pulsations du type continues sont entretenues, il est ainsi délicat de définir avec exactitude un début et une fin d'événement. Nous présentons dans cette section l'application du bootstrap par événements aux pulsations géomagnétiques, en l'illustrant sur un cas de données MT.

Etude du cas des données MT de Schwabwiller - Données peu bruitées Pour illustrer l'utilisation de la transformée en ondelettes continues et des données d'observatoires dans le traitement des données MT, nous prenons le cas de données acquises à Schwabwiller en Alsace (Figure 3.8) dans le cadre d'un projet de monitoring MT (Abdelfettah et al., 2014). Les stations MT utilisent des dipôles électriques de 100 mètres dans les directions Nord-Sud et Est-Ouest, avec des bobines d'induction MFS07e pour enregistrer les variations horizontales et verticale du champ magnétique.

La série temporelle MT utilisée débute le 02/12/2014 à 18h, et dure 24h. L'observatoire utilisé en tant que station de référence pour le calcul des estimations MT est celui de Chambon-La-Forêt (CLF), situé au sud de Paris en France. L'échantillonnage initial des données MT étant à 512 Hz, la série temporelle MT à été décimée à 1 Hz pour avoir la même fréquence d'échantillonnage que les données d'observatoire. Seules les fréquences inférieures à 1000 secondes étant d'intérêt pour cette étude, nous avons appliqué un filtre passe-haut avec une fréquence de coupure à 0.0005 Hz sur les deux types de séries temporelles (Figure 3.9).

Parmi les indicateurs permettant de calculer la similarité de deux séries temporelles x et y, la cohérence C définie en domaine spectral par :

$$C_{xy}(f) = \frac{|G_{xy}(f)|^2}{G_{xx}(f)G_{yy}(f)}, \tag{3.31}$$

où G_{xy} est la densité spectrale croisée entre les signaux x et y, et G_{xx}, G_{yy} sont les autocorrélations spectrales des signaux x et y respectivement. Une valeur proche de 1 indique une forte similarité entre les signaux x et y. Inversement, une valeur proche de 0 indique deux signaux complètement différents à la fréquence f. Calculée sur la journée entière (Figure 3.10), la cohérence est en moyenne au dessus de 0.7 pour les fréquences inférieures à 0.1 Hz, avec une forte décroissance vers 0.005 Hz.

Calculée sur des plages d'une heure, la cohérence montre une forte variabilité en fonction de l'heure considérée (Figure 3.11). Entre 16h et 17h le 03/12/2014, les cohérences NS et EW sont inférieures à 0.5, alors qu'elles sont proches de 1 entre 2h et 3h du matin à toutes les fréquences.

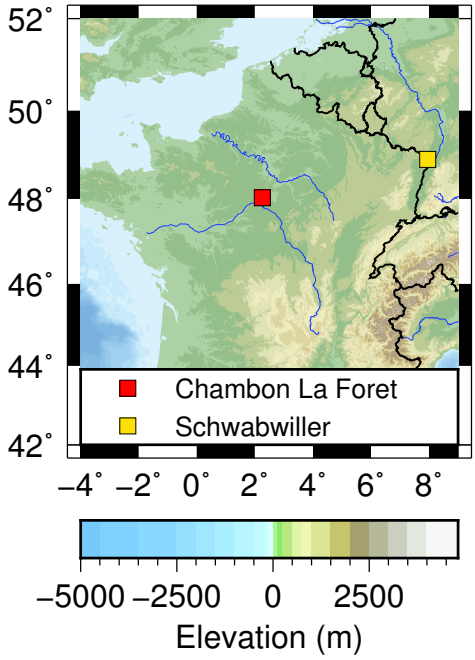


FIGURE 3.8: Carte indiquant la position de la station MT et de l'observatoire magnétique de CLF.

FIGURE 3.8: *Map of MT station and CLF magnetic observatory.*

Ces constatations sont à rapprocher du caractère intermittent des pulsations géomagnétiques qui ont des plages horaires où l'amplitude est plus importante (Section 2.3.1). L'application de la méthodologie de détection des échelles significatives par simulations de surrogates (section 2.3.2) peut alors s'appliquer ici.

Contrairement au cas de la détection des échelles significatives sur un seul observatoire (Section 2.3.2), nous nous intéressons ici au cas d'échelles significatives communes à deux séries temporelles. La quantité testée par rapport à la distribution des fonctions surrogates est définie par :

$$\mathcal{W}_{\psi}^{J_i}(a,b) = |\mathcal{W}_{\psi}^{XX_i'}(a,b)| + |\mathcal{W}_{\psi}^{XY_i'}(a,b)| + |\mathcal{W}_{\psi}^{YX_i'}(a,b)| + |\mathcal{W}_{\psi}^{YY_i'}(a,b)|, \tag{3.32}$$

où X et Y indiquent les séries temporelles MT locales et X_i, Y_i les séries de l'observatoire i dans le cas où plusieurs observatoires sont utilisés simultanément.

Dans le cas de séries temporelles extrêmement bruitées comme présentées Figure 2.20, ce type de procédure ne permet pas de récupérer les échelles significatives car le bruit sera aussi significatif. Il sera donc nécessaire de supprimer les parties de la série temporelle fortement contaminées par du bruit anthropique.

Après application des critères d'amplitude et de taille des régions significatives, la matrice des coefficients significatives est constituée de groupes connectés. Nous utilisons la méthodologie de He et al. (2009) permettant une segmentation rapide de grandes matrices peuplées d'indices 0 et

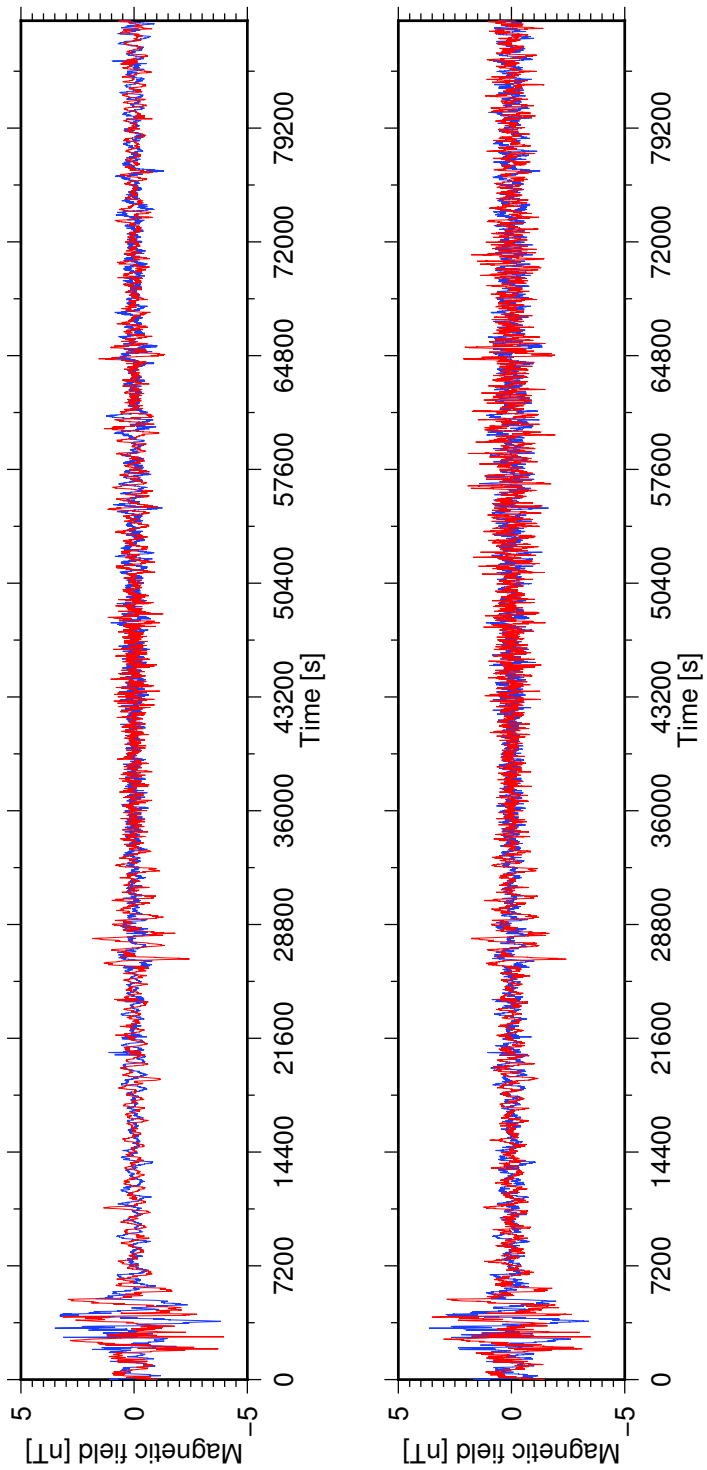


FIGURE 3.9: Séries temporelles à Schwabwiller et à l'observatoire magnétique de Chambon-La-Forêt pour le 03/12/2014.

FIGURE 3.9: *MT (top) and observatory (bottom) time series on the 02/12/2014. Time series start at 18 :00 :00 PM.*

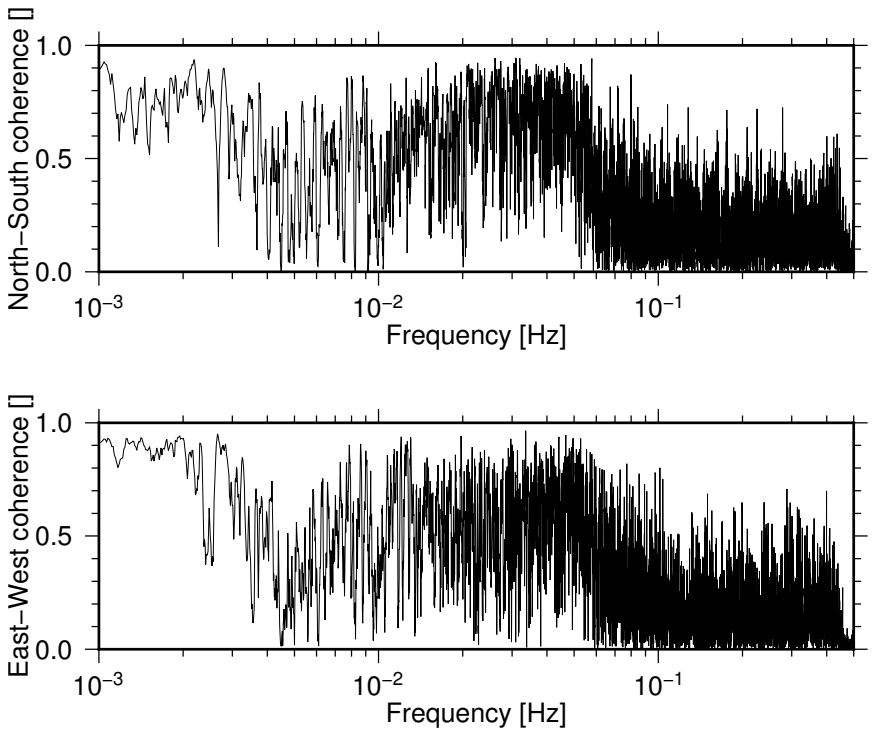


FIGURE 3.10: Cohérence sur 24h entre les séries MT et d'observatoire.
FIGURE 3.10: *24h coherency between MT and observatory time series.*

1 pour attribuer à chaque groupe un indice. La Figure 3.12 illustre le principe de la méthode. La Figure 3.13 montre le résultat de l'application de cette méthode aux séries temporelles étudiées.

Dans chaque groupe, les coefficients sont connectés, mais les groupes ne sont pas connectés entre eux. Dans la méthodologie d'inversion des fonctions de réponse MT présentée précédemment, chaque groupe correspond à un événement géomagnétique. Ainsi, nous pouvons appliquer directement la procédure utilisant le bootstrap hiérarchique présentée précédemment.

Le résultat de la procédure bootstrap est présenté Figure 3.14 par comparaison avec le résultat robuste. 100 itérations bootstrap ont été utilisés dans le calcul des incertitudes. Les deux procédures délivrent des résultats comparables dans la gamme de période 10 - 512 secondes, avec des incertitudes légèrement plus faibles pour le résultat par ondelettes.

Nous avons examiné les résidus sur le champ électrique après détermination du tenseur d'impédance. Deux ensembles de résidus ont été calculés, ceux sur les coefficients sélectionnés et ceux sur l'ensemble de la série temporelle. Nécessairement, les résidus les plus important seront dans l'ensemble contenant l'ensemble de la série temporelle, mais l'idée est de vérifier si la sélection des coefficients permet aussi la réduction des résidus sur le champ électrique. A la Figure 3.15, nous présentons les résidus maximaux à chaque échelle pour les deux ensembles. Les résidus maximaux sont de 4 à 10 fois plus important sur l'ensemble des coefficients que sur ceux sélectionnés. Il est intéressant aussi de remarquer que seule une légère différence existe entre les

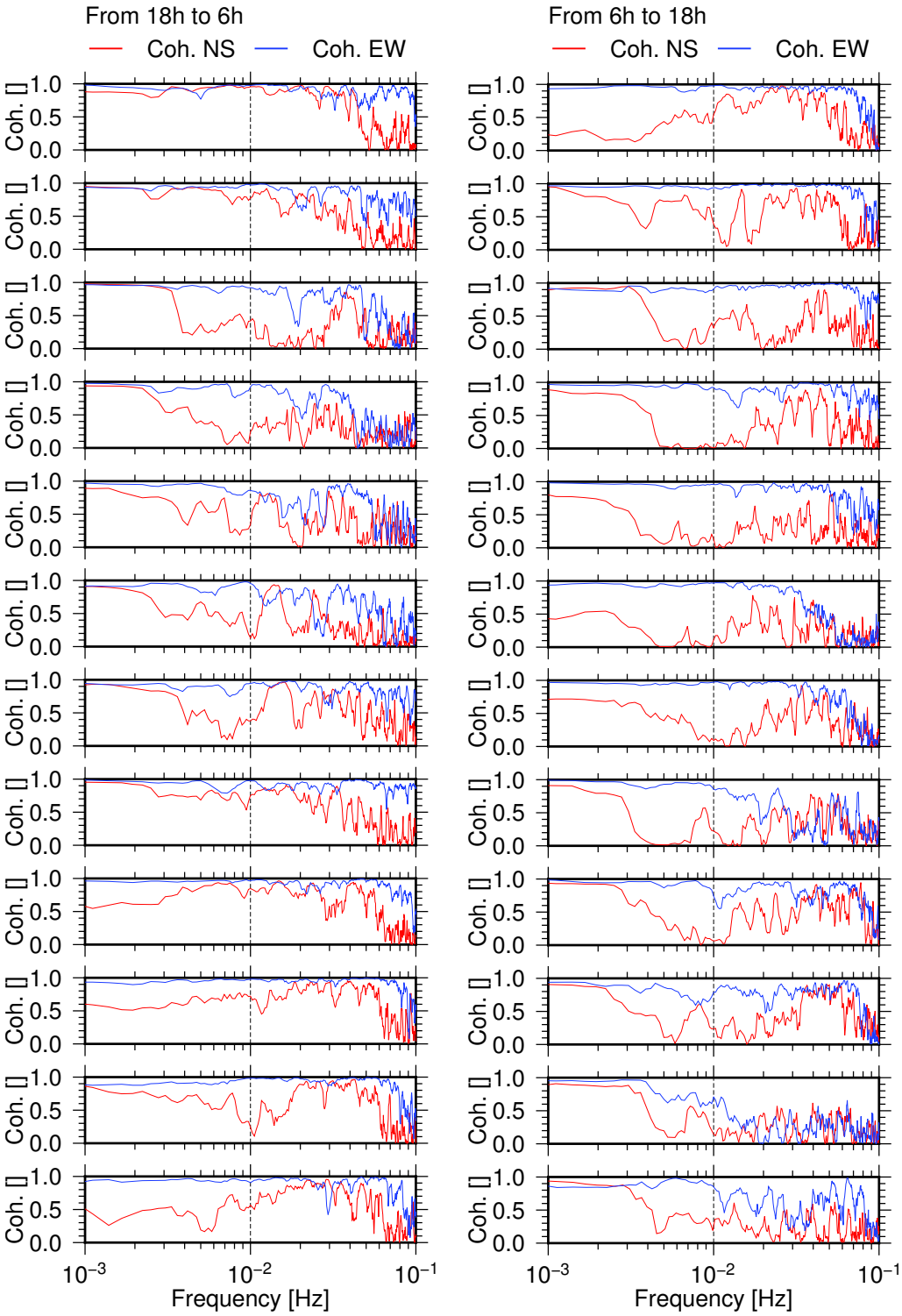


FIGURE 3.11: Cohérence sur 1h entre les séries MT et d'observatoire.
FIGURE 3.11: *1h coherency between MT and observatory time series.*

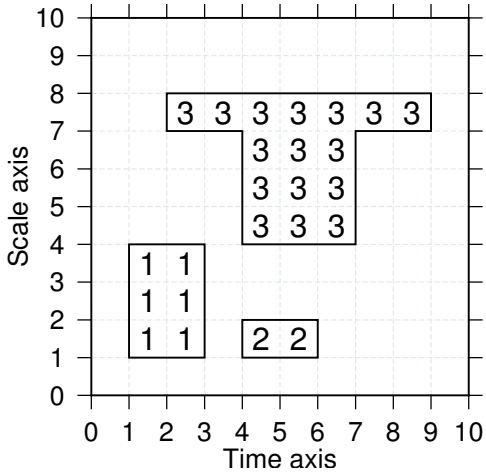


FIGURE 3.12: Illustration de la connectivité des pixels et de l'assignation d'un label à chaque ensemble.

FIGURE 3.12: *Illustration of the label assignement to groups of significant coefficients.*

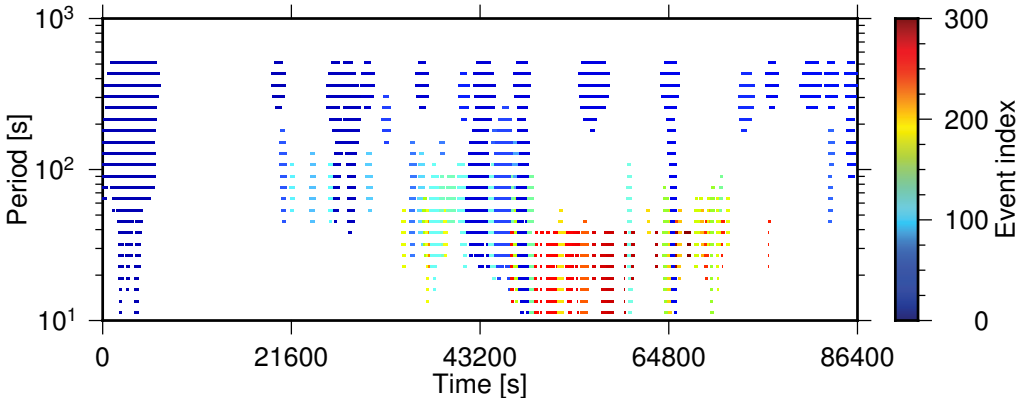


FIGURE 3.13: Indices des événements après application de la méthodologie de segmentation de He et al. (2009).

FIGURE 3.13: *Events labels after application of segmentation methodology of* He *et al. (2009).*

résidus sur E_x et E_y pour les coefficients sélectionnés alors qu'un facteur 2 sépare les résidus maximaux sur l'ensemble des coefficients aux longues périodes.

Les résidus maximaux ne constituant pas une indication suffisante quand à la différence entre les deux types de traitement, nous avons représenté les QQ-plots des distributions de résidus par rapport à ceux de la distribution de Rayleigh normalisée (Figure 3.15). Les deux distributions sont très comparables, seul le dernier centile constituant la queue des résidus présente une différence significative. La sélection des coefficients par analyse en ondelette permet donc la réduction des résidus importants sur le champ électrique. Cette propriété est donc valable pour les pulsations géomagnétiques et les éclairs atmosphériques.

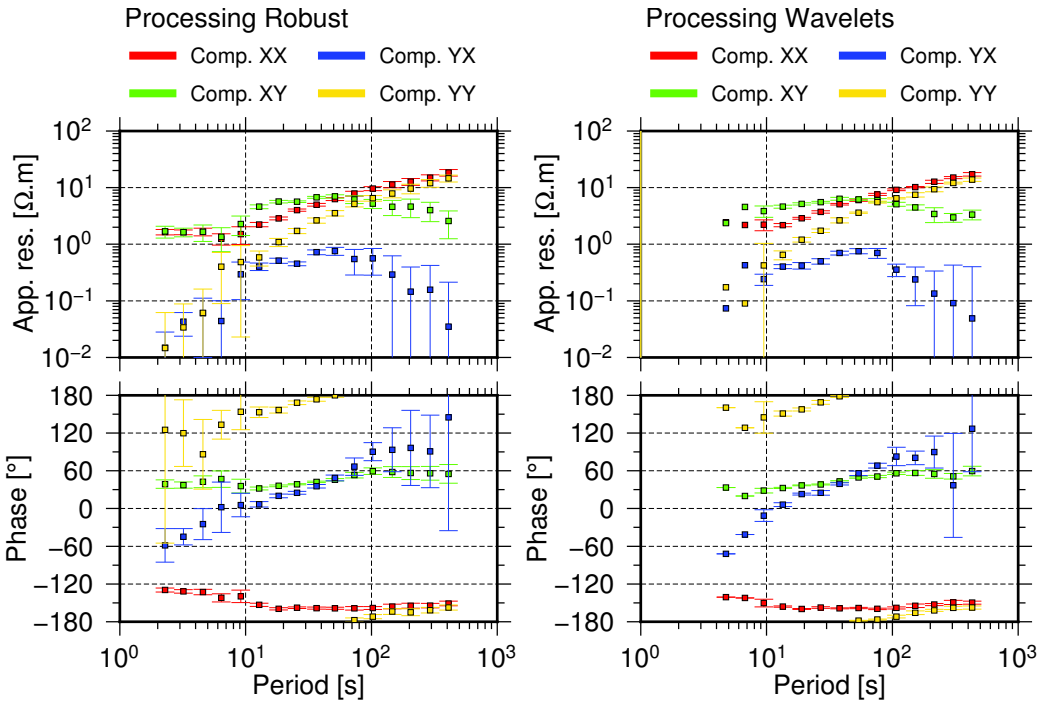


FIGURE 3.14: Sondages MT obtenus par application du code BIRRP (gauche) et du bootstrap (droite).

FIGURE 3.14: *MT soundings obtained through BIRRP processing (left) and bootstrap processing (right).*

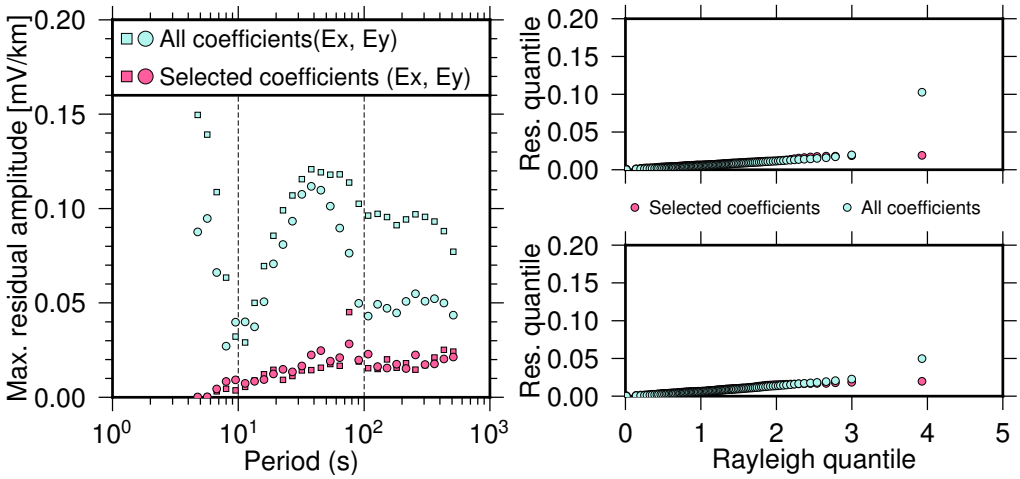


FIGURE 3.15: Gauche : Résidu maximal obtenu en prenant toute la série temporelle ou seulement la sélection des coefficients. Carrés : Résidus sur E_x, Cercles : Résidus sur E_y. Droite : QQ-plots des résidus par rapport à la distribution de Rayleigh. Haut : Résidus sur E_x, Bas : Résidus sur E_y.

FIGURE 3.15: *Left : maximum residual using all wavelet coefficients or only selected coefficients. Squares : E_x residus, Circles : E_y residus. Left : Residuals QQ-plots against Rayleigh distribution. Top : E_x residuals. Bottom : E_y residuals.*

3.5 Conclusions et perspectives

Dans ce chapitre, nous avons étudié les sources électromagnétiques naturelles utilisées en prospection magnétotellurique pour des périodes comprises entre 0.0001 et 1000 secondes. Réutilisant les méthodes d'inférence statistique introduites en magnétotelluriques dans les années 1980 et les procédures de détection d'événements géomagnétiques et atmosphériques présentées précédemment, nous avons mis en place une méthodologie de calcul des fonctions de réponse MT.

Sans utiliser tous les raffinements tels que l'utilisation de la matrice de projection, nous avons montré que la sélection des coefficients d'ondelette correspondant aux périodes d'activité géomagnétique ou atmosphérique permettait d'améliorer significativement les sondages MT. Ainsi, de manière simple il est possible de réduire le biais magnétotellurique observé en station simple, et de diminuer la largeur fréquentielle de l'Audio-MT *dead band* par pré-sélection des meilleurs événements. Les résidus sur le champ électrique sont significativement réduits, ainsi le bootstrap hiérarchique permet aussi de réduire les incertitudes associés aux sondages MT.

Une grande faiblesse des traitements présentés ici est liée au temps de calcul. En effet, dans les procédures robustes décrites précédemment, les poids calculés dans la régression robuste sont réutilisés pour le calcul des incertitudes Jackknife. Nous avons choisi de recalculer à chaque itération bootstrap les poids de Huber pour optimiser la détermination des fonctions de réponse MT, entraînant une augmentation significative du temps de calcul.

Parmi les axes d'amélioration de ces méthodes, nous n'avons pas étudié l'utilisation de la polarisation dans le calcul statistique de l'impédance, autre que celui du rejet des éclairs atmosphériques proches. L'ensemble des améliorations statistiques telles que l'utilisation de la matrice de projection dans la détection des valeurs influentes dans la régression ou l'utilisation d'une autre distribution dans l'ajustement des résidus peuvent être appliquées aux développements méthodologiques présentés dans cette thèse. La question des sources entretenues dans la MT *dead band* se pose aussi car même si aucune source n'est visible dans les spectrogrammes d'ondelette, les procédures de traitement robuste permettent de déterminer des fonctions de réponse MT.

Troisième partie

Interprétation et inversion de sondages magnétotelluriques

Chapitre 4

Chambon La Forêt

Ce chapitre présente la campagne de mesures Audio-MT et MT effectuée autour de l'observatoire magnétique de Chambon-La-Forêt en collaboration avec l'équipe de l'observatoire magnétique et Prof. Frédéric Perrier de l'Institut de Physique du Globe de Paris (IPGP).

4.1 Situation géographique de l'observatoire

4.1.1 Observatoire magnétique national de Chambon-La-Forêt (CLF)

L'observatoire magnétique de Chambon-La-Forêt (département du Loiret ; Région Centre-Val de Loire ; Figure 4.1) est considéré comme l'observatoire magnétique national français. Il est géré par l'IPGP. Créé en 1883 (IPGP Website) à Saint-Maur-des-Fossés (Val de Marne), il a subi plusieurs déménagements du fait de l'accroissement de l'activité anthropique aux alentours. Il a d'abord été transféré sur le site de Val-Joyeux en 1901, et enfin à Chambon-La-Forêt en 1936. Les coordonnées de l'observatoire sont indiquées dans la Table 4.1 pour une altitude de 145 mètres.

Coordonnées	Latitude (degrés)	Longitude (degrés)
Géographiques	48.025N	2.26E
Géomagnétiques	49.55N	85.73E
Géomagnétiques corrigées	43.21N	79.05E

Table 4.1: Coordonnées de l'observatoire magnétique de Chambon-La-Forêt. Les coordonnées géomagnétiques corrigées sont calculées à partir du modèle IGRF 2010 pour l'année 2011.
TABLE 4.1: *Geomagnetic coordinates of the magnetic observatory of Chambon-La-Forêt. Corrected geomagnetic coordinates are computed from the IGRF-10 model of year 2011.*

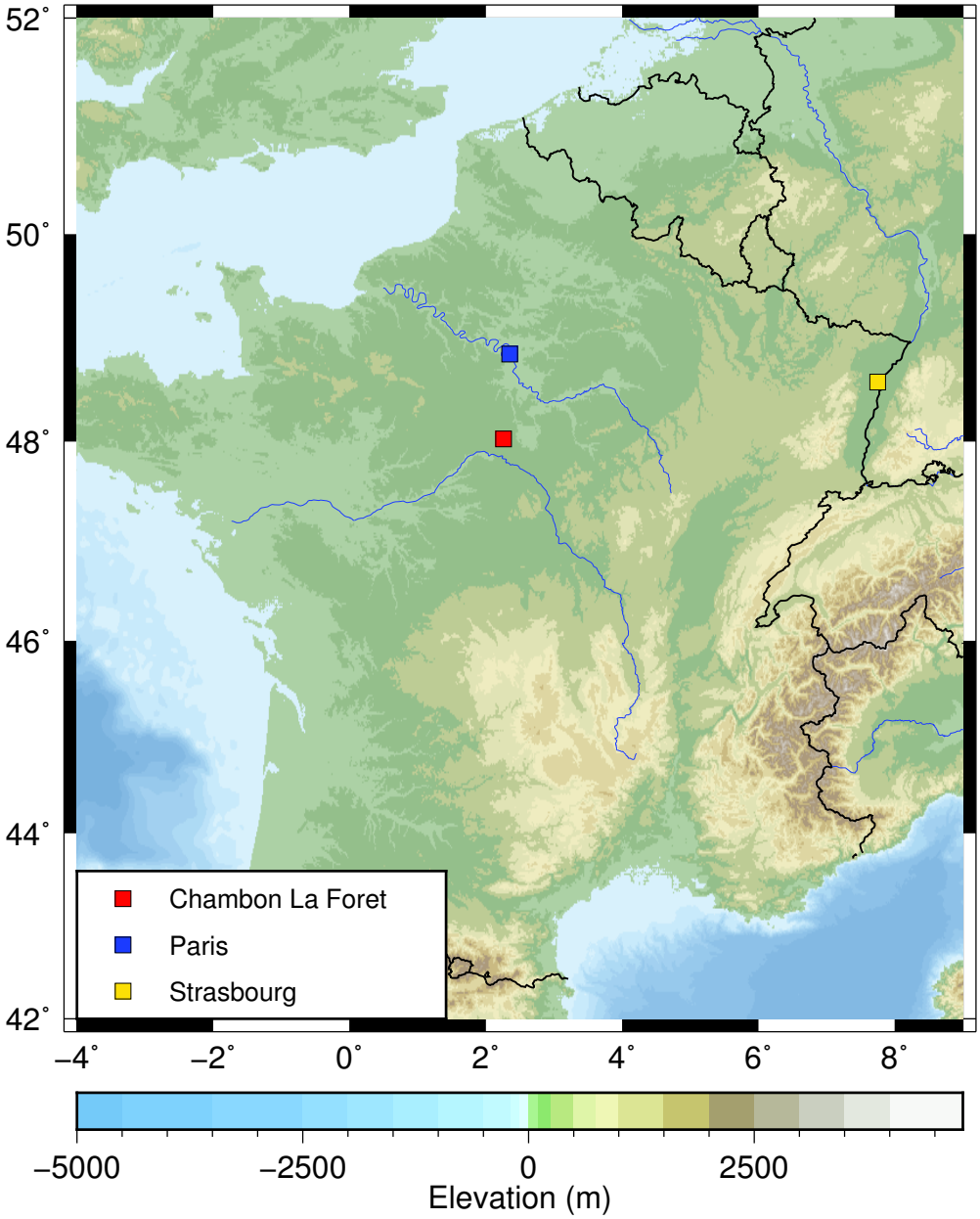


Figure 4.1: Carte de la France avec la localisation de la campagne Audio-MT, MT. La topographie est extraite du modèle ETOPO1 (Amante et Eakins, 2009).
FIGURE 4.1: *Map of France with location of Audio-MT, MT campaign. Topography is from the ETOPO1 model (*Amante *et* Eakins, *2009).*

Les données délivrées par l'observatoire sont de type seconde depuis le 01 Janvier 2011 (Site du Bureau Central de Magnétisme Terrestre, BCMT). Les données minutes sont disponibles depuis le 01 Janvier 1981. CLF participe également au calcul des indices géomagnétiques *am* et *ASY* (ISGI Website).

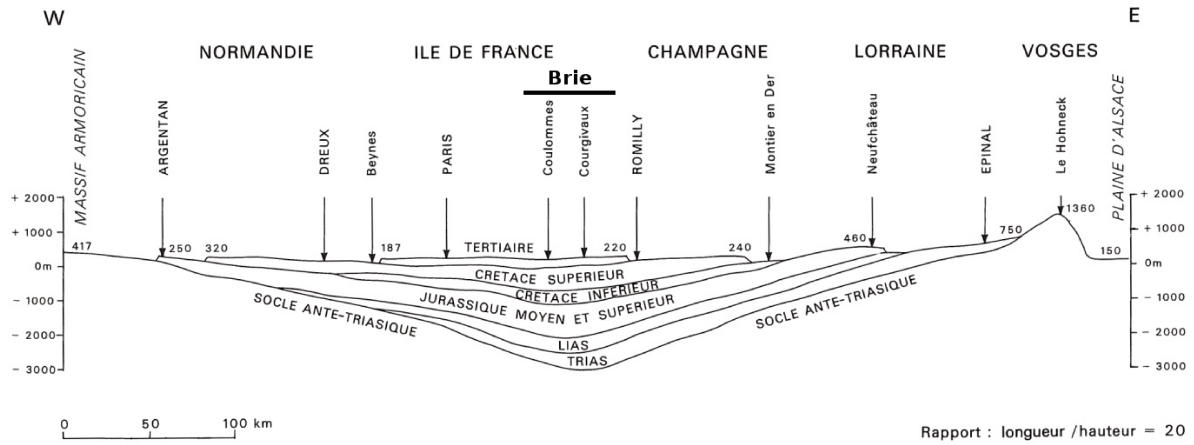


FIGURE 4.2: Coupe géologique Est-Ouest du bassin Parisien. L'échelle verticale est amplifiée. D'après LORENZ (1980).
FIGURE 4.2: *East-West geological cross section of the Paris basin. Note the vertical exaggeration From* LORENZ *(1980)*

4.1.2 Géologie du bassin parisien

L'observatoire magnétique de CLF se situe au coeur du bassin sédimentaire de Paris. Ce bassin est délimité par d'anciens massifs hercyniens érodés tels que le massif Armoricain, les Ardennes, les Vosges et le massif du Morvan. Dans cette région, le socle, défini comme les couches sédimentaires antérieures au Mésozoïque, est atteint à des profondeurs variables. Il affleure dans certaines régions aux bords du bassin, et peut atteindre la profondeur de 3000 mètres au niveau de la région de la Brie, maximum de la subsidence du bassin de Paris (Figure 4.2). Au niveau de la région de la Forêt d'Orléans où se situe CLF, il est possible d'observer toutes les principales formations sédimentaires secondaires post-Paléozoïque, du Triassique au Crétacé (LORENZ, 1980).

4.1.3 Etudes géophysiques précédentes dans la région de CLF

Différentes études géophysiques ont été menées dans la région du bassin Parisien. Les études sismiques, combinées aux forages pétroliers ont principalement permis de déterminer la géométrie des failles et des formations sédimentaires dans le cadre de l'exploration pétrolière. Les principales études gravimétriques, magnétiques et magnétotelluriques dans la région ont majoritairement été concentrées sur l'étude de l'Anomalie Magnétique du Bassin Parisien (AMBP) (GERARD et WEBER, 1971 ; DUPIS et al., 1990 ; PHAM, BOYER et LE MOUËL, 2000 ; MARTELET et al., 2013), mais d'autres études ont été menées plus largement à des fins géothermiques (DEBEGLIA, 2005).

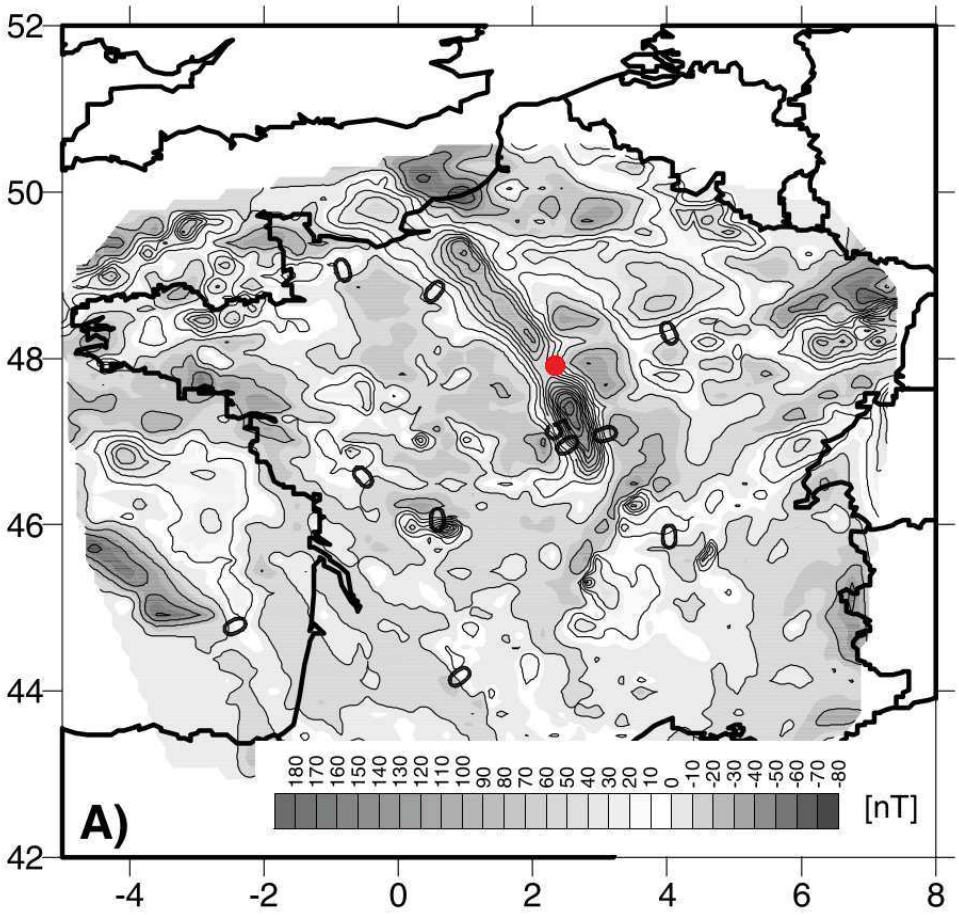


FIGURE 4.3: Intensité de l'anomalie magnétique (non réduite aux pôles) obtenue en France. Le cercle rouge indique la position de l'observatoire magnétique. Modifié d'après THÉBAULT, MANDEA et SCHOTT (2006).
FIGURE 4.3: *Magnetic anomaly (without reduction to the pole) intensity over France. The red circle indicates the location of CLF magnetic observatory. Modified from* THÉBAULT, MANDEA *et* SCHOTT *(2006)*

L'AMPB s'étend de Nevers jusqu'à la Manche (Figure 4.3). L'intensité de cette anomalie à la surface de la Terre peut atteindre plus de 200 nT (DEBEGLIA, 2005 ; THÉBAULT, MANDEA et SCHOTT, 2006) et sa géométrie évolue du Sud-Est au Nord-Ouest. Dans certaines régions, notamment à son extrémité sud-est, l'anomalie gravimétrique se superpose à l'anomalie magnétique (BAYER, GUILLEN et WEBER, 1987). Inversement, en allant vers la Manche, l'anomalie gravimétrique ne correspond plus à la forme de l'anomalie magnétique (AUTRAN et al., 1986).

De nombreuses études ont cherché à déterminer la source de cette anomalie (DUPIS et al., 1990 ; PHAM, BOYER et LE MOUËL, 2000), mais son origine est encore discutée (THÉBAULT, MANDEA et SCHOTT, 2006). PHAM, BOYER et LE MOUËL (2000) ont interprété et inversé des mesures MT acquises le long du profil EUCORS (BOIS et al., 1986), sans pouvoir parfaitement déterminer une origine à l'AMBP dans le modèle de résistivité. DUPIS et al. (1990) ont utilisé la

MT dans le cadre de l'implantation du sondage profond de Couy, mais n'ont pas permis non plus de donner d'information sur l'origine de l'AMBP. Une étude (Autran et al., 1986), à l'aide de modélisations de données magnétiques, a montré que le toit de l'anomalie au niveau de Sancerre pouvait être situé à seulement 2.3 km de profondeur. En revanche, dans cette étude, l'utilisation des données gravimétriques en plus de magnétiques permet de donner deux solutions acceptables, l'une dont le toit est à 1,8 km de profondeur, l'autre à 4,5 km. Ces modélisations ont permis à Autran et al. (1986) de proposer une origine géologique à l'AMBP, en situant son origine dans l'ouverture d'un rift au Dévonien, qui se serait refermé au Carbonifère.

4.2 Présentation de la campagne

4.2.1 Mesures utilisées

4.2.1.1 Dipôles électriques permanents

F. Perrier (IPGP) et l'équipe de l'observatoire magnétique ont installé en septembre 2012 six dipôles électriques permanents sur le site de l'observatoire (F. Perrier et al., 2014). Deux types de mesures de différences de potentiel ont été effectués, à dix mètres de profondeur, et à un mètre de profondeur. Les deux types d'installations utilisent des électrodes de type Petiau de seconde génération (Petiau, 2000). L'acquisition est échantillonnée à 0.5 Hz de septembre 2012 à septembre 2013. A partir de septembre 2013, l'acquisition est effectuée à un échantillonnage de 0.1 Hz.

Ces mesures ont permis à F. Perrier et al. (2014) de tester la différence de stabilité entre les électrodes situées en profondeur et en surface, et de détecter des effets de marées océaniques uniquement polarisés selon la direction Nord-Sud du champ électrique à CLF .

4.2.1.2 Mesures MT

Nous avons effectué des mesures Audio-MT et MT du 2 au 7 mars 2015. 3 stations MT et 13 stations Audio-MT (Figure 4.4) ont été installées grâce à la collaboration avec le personnel de l'observatoire magnétique. Pour une description détaillée des mesures effectuées, voir l'Annexe E. Les objectifs de cette campagne de mesure étaient multiples :

- Etudier le bruit électromagnétique dans la région de l'observatoire.
- Permettre la caractérisation de la structure électrique du sous-sol.
- Analyser son influence sur les mesures de l'observatoire.

Les deux derniers objectifs étant reliés, nous les traiterons dans une même section.

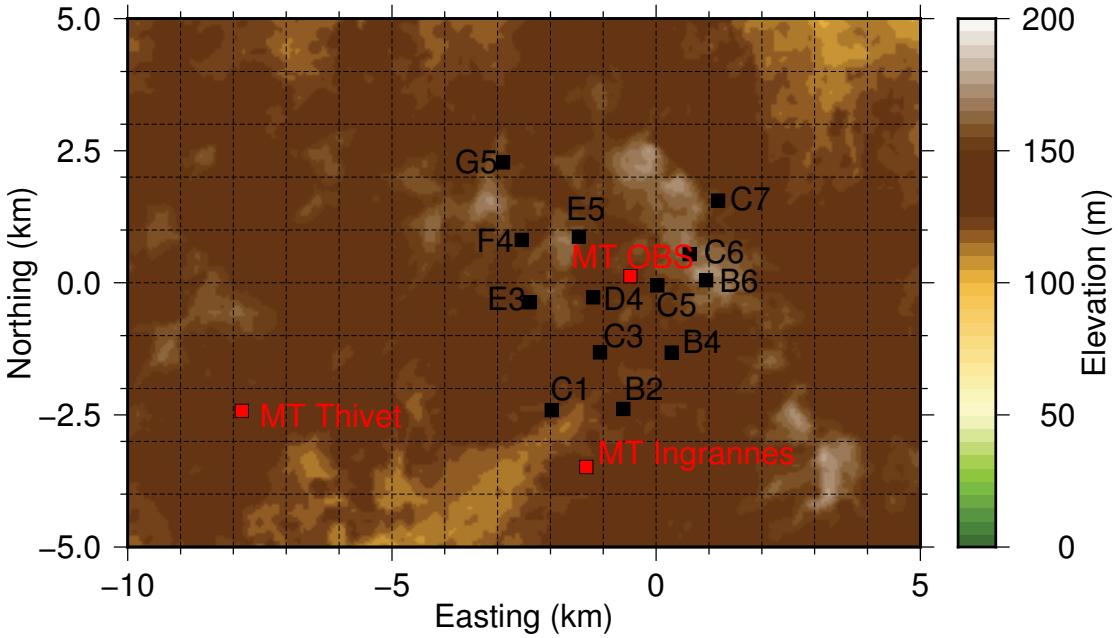


FIGURE 4.4: Position des stations Audio-MT et MT installées en mars 2016 relativement à l'observatoire située en (0, 0). Projection UTM20N.
FIGURE 4.4: *Position of Audio-MT and MT stations relative to the observatory located at position (0, 0). UTM 20N projection.*

4.3 Bruit anthropique autour de l'observatoire

CLF est situé dans la forêt d'Orléans, à 4.5 kilomètres de la ville de Chambon-La-Forêt. Les principaux bruits anthropiques observés étant des bruits de fréquences supérieures au Hz, nous représentons par la suite les séries temporelles quasi-brutes. Les seules transformations appliqués sont la calibration des champs EM et une décimation appliquées aux plus hautes fréquences pour se limiter aux fréquences inférieures à celles des émetteurs VLF de fréquences supérieures à 15 kHz. Les Figures 4.5, 4.6, et 4.7 présentent la comparaison des séries temporelles et les spectres en puissance du champ magnétique pour la station localisée à l'observatoire et des stations mobiles. Une des caractéristiques directement visible est l'amplitude importante des harmoniques du 50 Hz au niveau de l'observatoire, même pour des fréquences supérieures au kHz. Les enregistrements simultanés aux stations mobiles présentent invariablement des amplitudes inférieures pour ces harmoniques. L'amplitude du bruit au niveau de l'observatoire est tel que le spectre en puissance ne présente plus les caractéristiques attendues d'un spectre de variations naturelles. En domaine temporel, ceci se traduit par une impossibilité d'identifier la plupart des événements naturels atmosphériques, excepté ceux ayant une amplitude très importante (voir section 3.4.3.2). Pour les stations mobiles, l'amplitude associé aux harmoniques du 50 Hz est nettement inférieure. Un simple filtrage du seul 50 Hz permet d'identifier les événements atmosphériques.

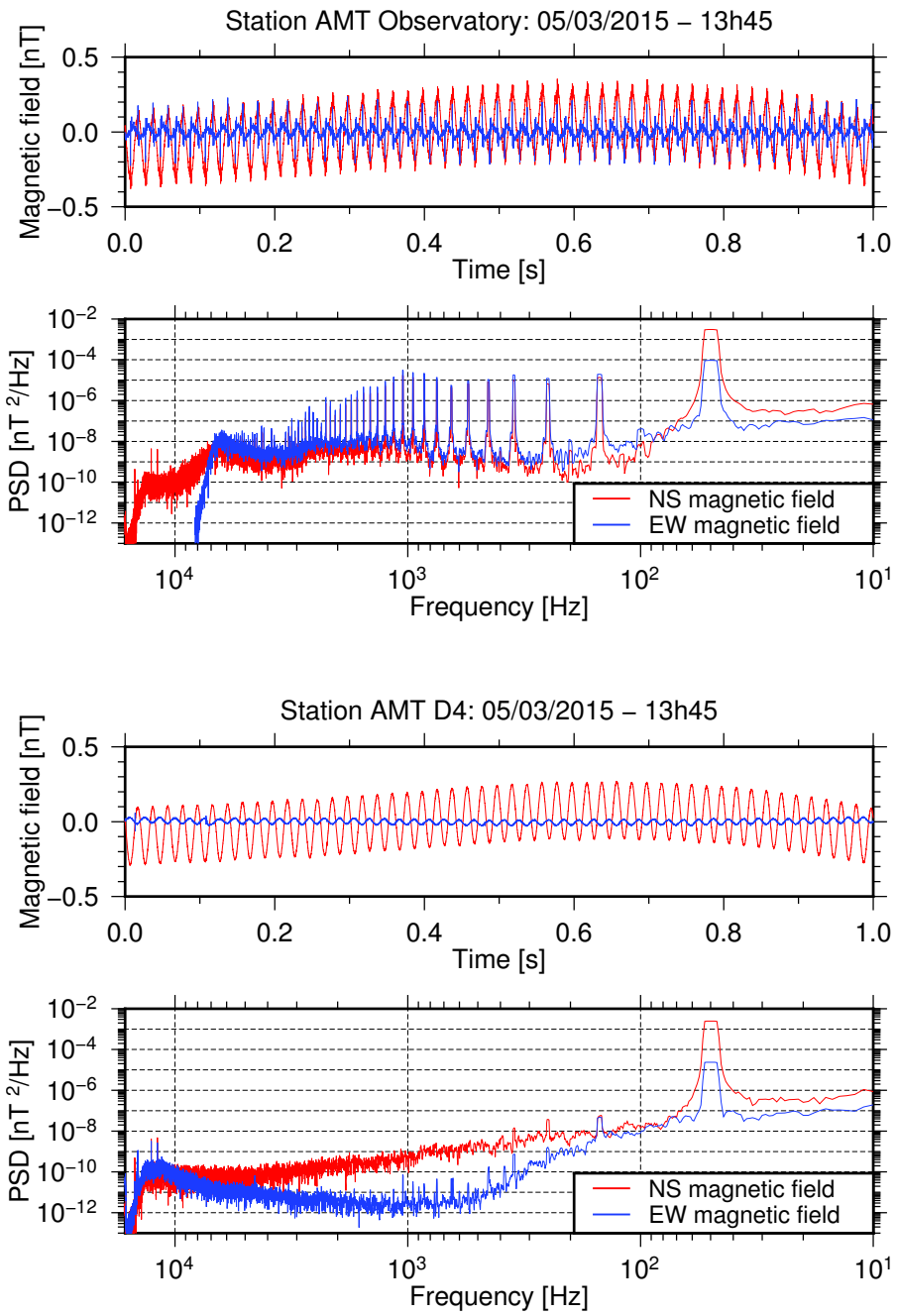


FIGURE 4.5: Comparaison des séries temporelles entre la station Observatoire et D4. L'échantillonnage des séries temporelles est à 32768 Hz.

FIGURE 4.5: *Comparison between D4 and Observatory station time series. Sampling frequency is 32,768 Hz.*

Le bruit à 50 Hz à proximité de l'observatoire présente de plus une variation de son intensité avec le temps, variation beaucoup moins présente sur les sites mobiles (Figure 4.7). En résumé, sur toutes les stations étudiées, l'amplitude maximale de bruit observé se situe à proximité de l'observatoire magnétique.

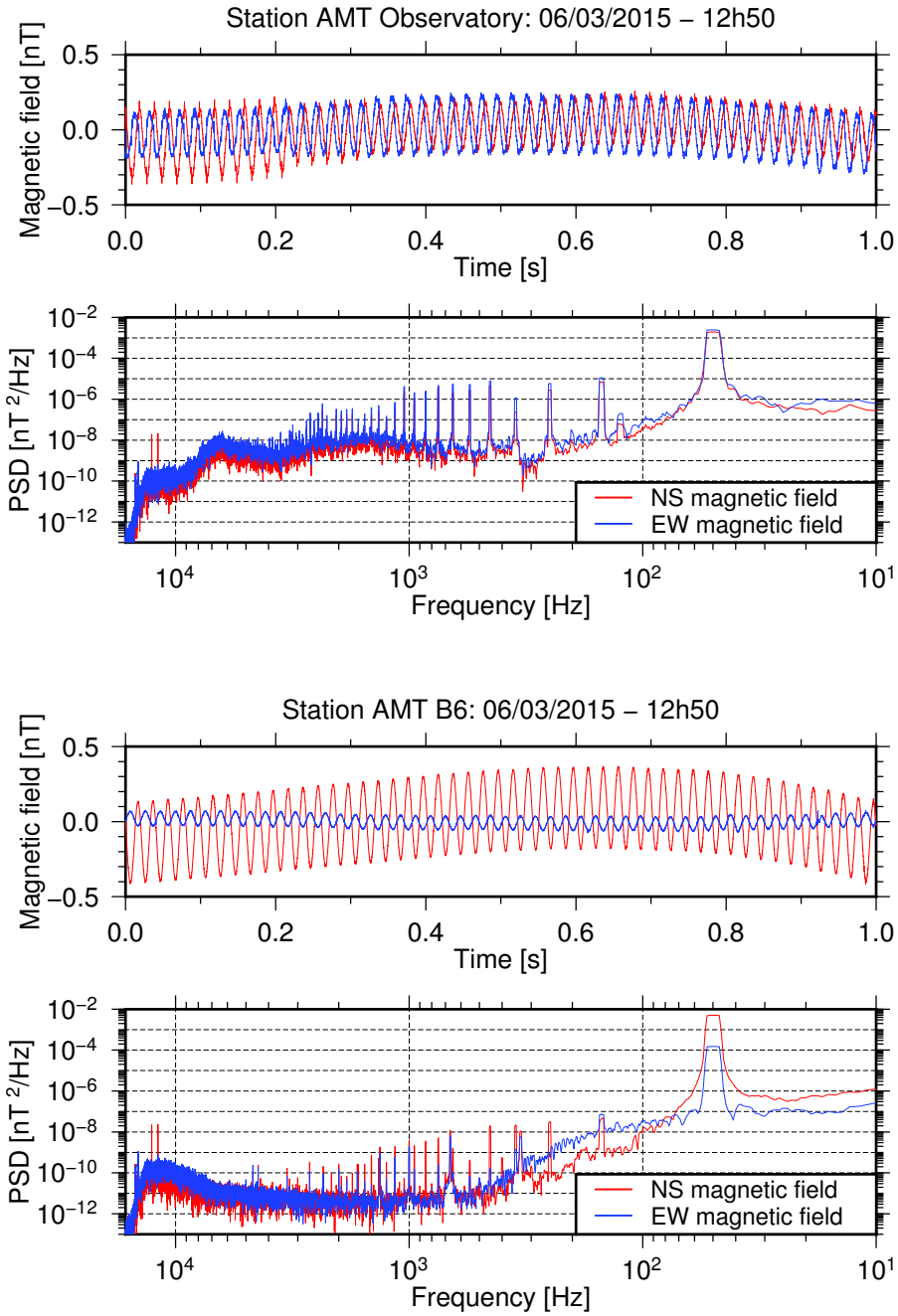


FIGURE 4.6: Comparaison des séries temporelles entre la station Observatoire et B6. L'échantillonnage des séries temporelles est à 32,768 Hz.

FIGURE 4.6: *Comparison between B6 and Observatory station time series. Sampling frequency is 32,768 Hz.*

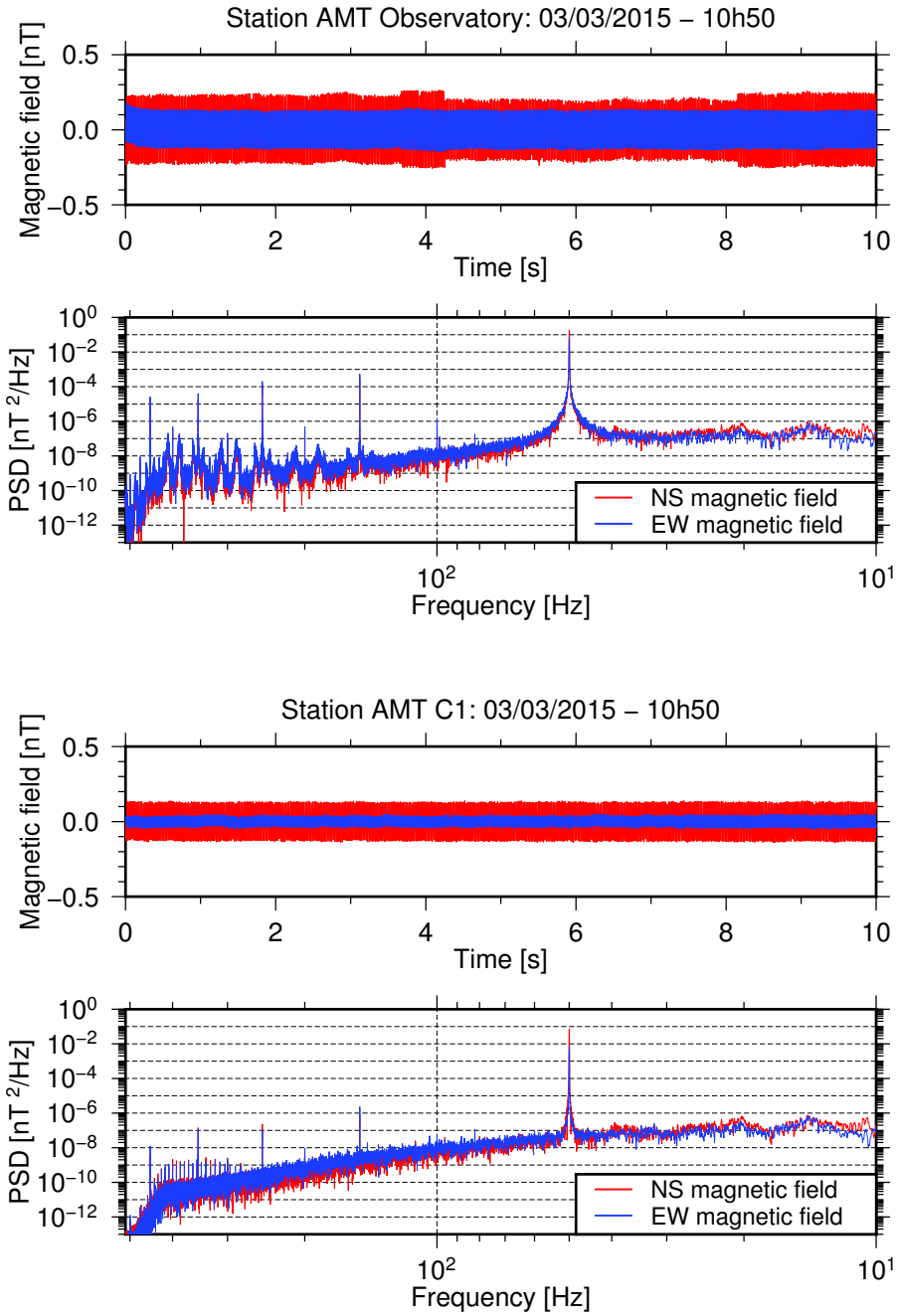


FIGURE 4.7: Comparaison des séries temporelles entre la station Observatoire et C1. L'échantillonnage des séries temporelles est à 1024 Hz.

FIGURE 4.7: *Comparison between B6 and Observatory station time series. Sampling frequency is 1,024 Hz.*

4.4 Traitement et interprétation des données MT

4.4.1 Traitement des données MT

Le bruit électromagnétique sur les données acquises étant important, la méthodologie de traitement de données par ondelettes continues, présentée à la partie II, a été utilisée pour obtenir les fonctions de réponses MT pour les fréquences supérieures à 10 Hz. Les fonctions de réponses MT pour les fréquences inférieures à 10 Hz ont été obtenues par application du code BIRRP (Alan D. Chave et David J. Thomson, 2004). Dans les traitements MT, les mesures hautes fréquences utilisent celles acquises à l'observatoire magnétique en tant que mesures *remote*. Les sites MT utilisent les données magnétiques acquises à la station Géophysique du Welschbruch située dans les Vosges. Enfin le traitement des dipôles d'observatoire a été effectué en prenant les mesures de l'observatoire d'Ebro (EBR, Espagne) comme mesures de référence.

Pour les sites Audio-MT, les réponses MT ont été obtenues pour des fréquences entre 16 kHz et 10 Hz. Des estimations peu fiables sont observables dans les intervalles de fréquences compris entre 1 et 4 kHz pour les sites les moins bruités, et entre 600 Hz et 5 kHz pour les sites très bruités. Les réponses pour les stations MT et les dipôles d'observatoire sont disponibles pour des fréquences comprises entre 100 et 0.005 Hz et entre 0.1 et 0.00001 Hz respectivement.

4.4.2 Dimensionnalité des sondages MT

Les sondages MT ont été analysés selon les techniques de décomposition de Groom-Bailey (Groom et Bailey, 1989) et du tenseur de phase (Caldwell, Bibby et Brown, 2004). Ces méthodes sont décrites en annexe D.

Les ellipses du tenseur de phase (Figure 4.8) indiquent un milieu 1D, dont la résistivité ne varie qu'avec la profondeur pour des fréquences supérieures au Hz, avec le paramètre λ majoritairement inférieur à 0.2 et le paramètre β proche de 0. Pour les fréquences inférieures au Hz, les ellipses du tenseur de phase deviennent elliptiques, indiquant un début de distorsion du champ électrique caractérisé par le paramètres β. λ augmente aussi pour atteindre des valeurs supérieures à 0.2 et inférieures à 0.5. Toutes les ellipses s'alignent entre 0.1 Hz et 0.001 Hz, indiquant un milieu à deux dimensions. Enfin à longues périodes, le paramètre λ devient supérieur à 0.5, indication du passage à un milieu 2D/3D (A. D. Chave et A. Jones, 2012).

La décomposition du tenseur par les paramètres de Groom-Bailey est en accord avec l'interprétation donnée par le tenseur de phase (Figure 4.9). Aux hautes fréquences, le *twist* et le *shear* restent proches de 0, indication d'une absence de distorsion du champ électrique. Inférieur au Hz, les valeurs du *twist* et du *shear* augmentent légèrement, indication d'une distorsion significative d'un milieu ayant au minimum 2 dimensions. Cette augmentation continue jusqu'aux longues

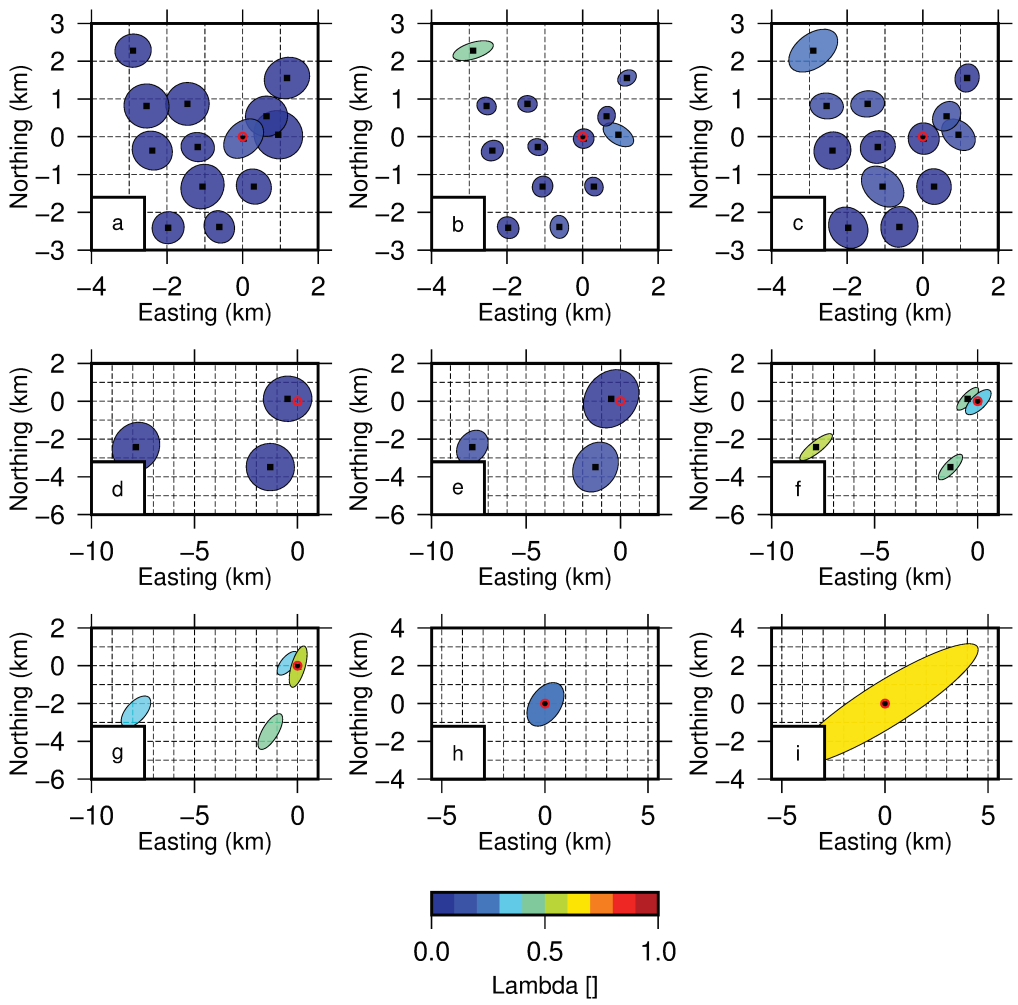


FIGURE 4.8: Ellipses du tenseur de phase aux fréquences : a) 6 kHz b) 279 Hz c) 34 Hz d) 7 Hz e) 0.5 Hz f) 0.039062 Hz g) 0.007812 Hz h) 0.000521 Hz i) 6.5e-5 Hz. La couleur de l'ellipse indique la valeur du paramètre λ.

FIGURE 4.8: *Phase tensor ellipses at frequencies : a) 6 kHz b) 279 Hz c) 34 Hz d) 7 Hz e) 0.5 Hz f) 0.039062 Hz g) 0.007812 Hz h) 0.000521 Hz i) 6.5e-5 Hz. Ellipse color indicates the value of λ parameter.*

périodes. A toutes les fréquences, le test d'adéquation d'un modèle de distorsion géologique par une hétérogénéité 3D de Groom et Bailey (1989) est validé par le test du χ^2 (Alan D Chave et J. T. Smith, 1994), les valeurs du test étant majoritairement inférieures à 1. La structure régionale peut donc être assimilée à une structure 2D.

Enfin, un autre indicateur de la présence d'une variation latérale de conductivité se trouve dans les fonctions de transfert verticales (Figure 4.10). Aux courtes périodes, l'amplitude des fonctions de transfert verticales est faible (< 0.1), indication d'une variation latérale très faible de la conductivité, en cohérence avec les indicateurs précédents. L'amplitude des flèches d'induction augmente au dessus de 0.1 pour les fréquences inférieures à 0.1 Hz et leurs directions indiquent

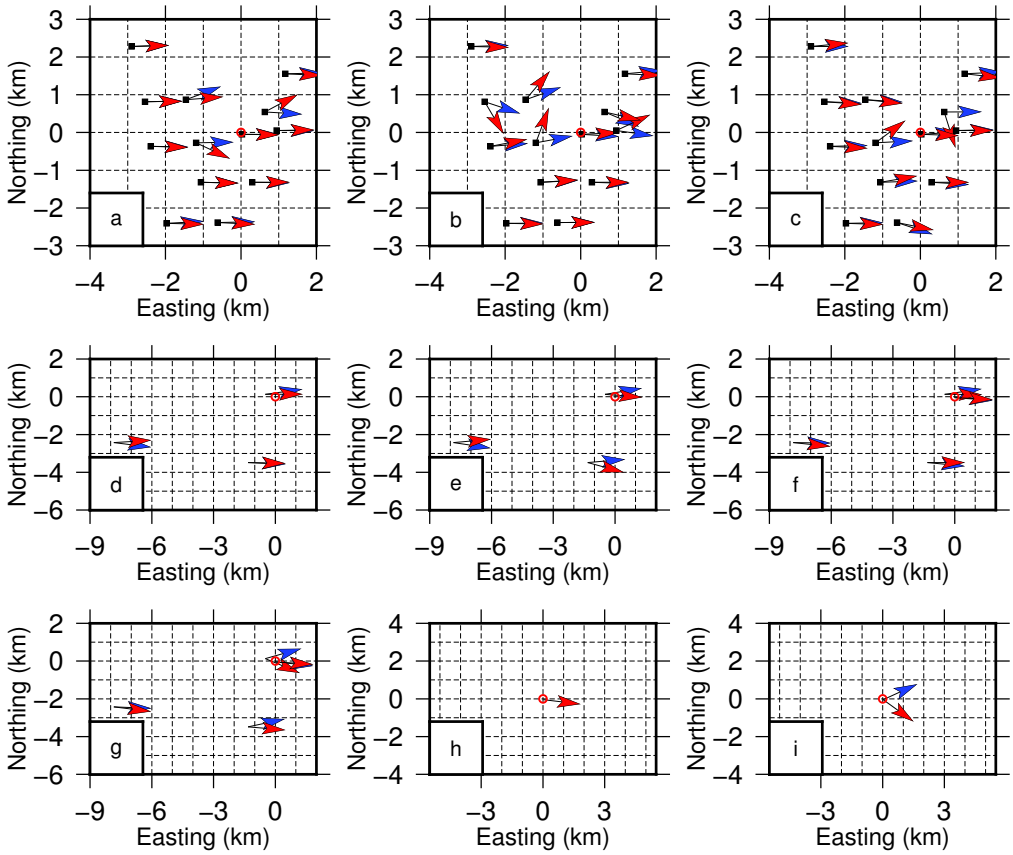


FIGURE 4.9: *Twist* et *Shear* de la décomposition de Groom-Bailey aux fréquences : a) 6 kHz b) 279 Hz c) 34 Hz d) 7 Hz e) 0.5 Hz f) 0.039062 Hz g) 0.007812 Hz h) 0.000521 Hz i) 6.5e-5 Hz. Rouge : *Twist*. Blue : *Shear*

FIGURE 4.9: *Twist and shear of Groom-Bailey decompositions at frequencies : a) 6 kHz b) 279 Hz c) 34 Hz d) 7 Hz e) 0.5 Hz f) 0.039062 Hz g) 0.007812 Hz h) 0.000521 Hz i) 6.5e-5 Hz. Red : Twist, Blue : Shear.*

la zone conductrice dans la direction Ouest. Aux très longues périodes (Figure 4.10, Panneau f), les flèches d'induction indiquent la direction Nord-Ouest.

Les sondages indiquent donc une dimensionnalité 1D au niveau des points de mesure pour les fréquences supérieures au Hz. Or, la carte géologique (Figure 4.11) montre une distribution de matériaux en surface variant fortement dans la zone étudiée. Les valeurs de résistivités apparentes des sondages aux plus hautes fréquences montrent une légère variation selon les sondages sans relation avec la carte géologique. La profondeur de peau pour un milieu d'environ 10 Ω.m à 10 kHz étant de 22.5 mètres, les valeurs des sondages MT à haute fréquence indiquent déjà une résistivité intégrée des premières couches et peuvent donc être différentes de celles de la roche en surface indiquée par la carte géologique. En conclusion, nous pouvons dire à partir de l'analyse de la dimensionnalité des sondages que la distribution de résistivité sous l'observatoire est localement 1D pour les fréquences supérieures au Hz, puis 2D aux fréquences entre le Hz et 0.0001 Hz, où le milieu devient 3D. Au Hz, la résistivité apparente du sondage MT à l'observatoire

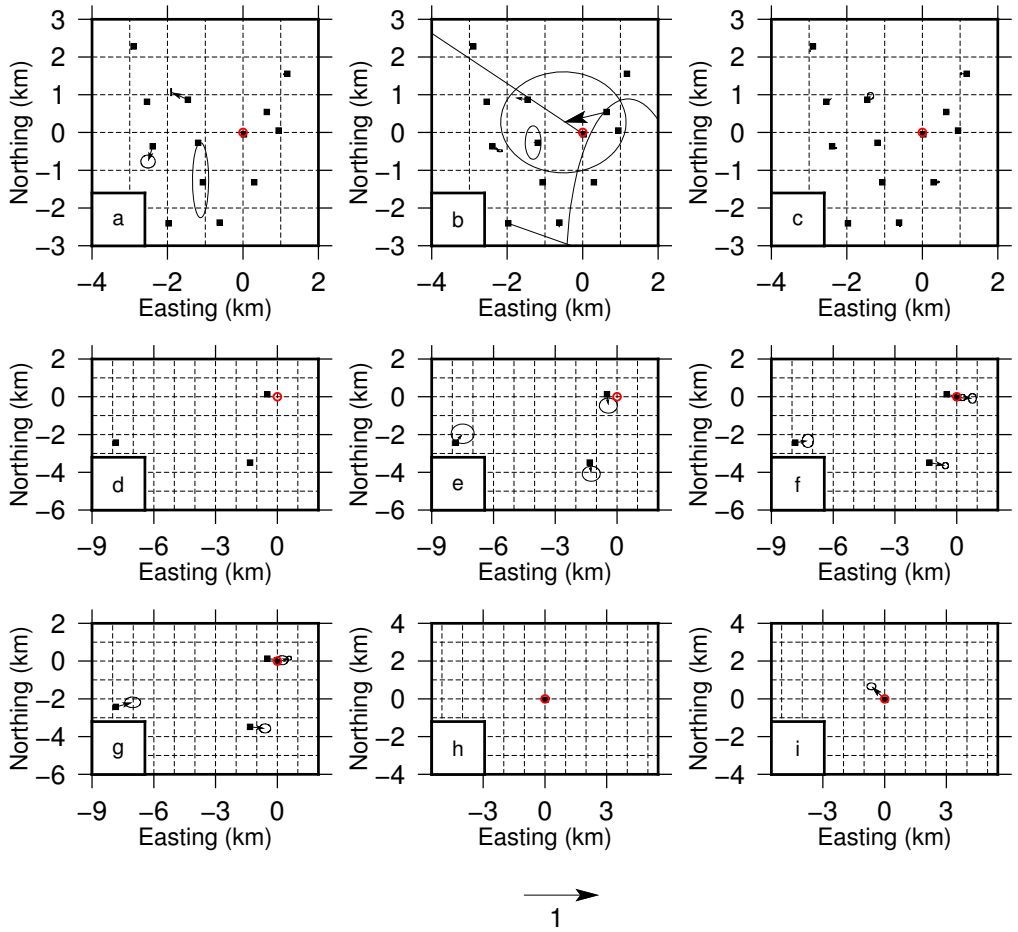


FIGURE 4.10: Partie réelle des flèches d'induction en convention de Parkinson aux fréquences : a) 6 kHz b) 279 Hz c) 34 Hz d) 7 Hz e) 0.5 Hz f) 0.039062 Hz g) 0.007812 Hz h) 0.000521 Hz i) 6.5e-5 Hz.

FIGURE 4.10: *Real part of inductions arrows in Parkinson convention at frequencies : a) 6 kHz b) 279 Hz c) 34 Hz d) 7 Hz e) 0.5 Hz f) 0.039062 Hz g) 0.007812 Hz h) 0.000521 Hz i) 6.5e-5 Hz.*

est de 20 Ω.m. La profondeur de peau correspondante est de 2249 mètres. Le toit de la structure géologique responsable de la distorsion 2D se situe donc aux environs de cette profondeur.

La direction de cette structure 2D est indiquée par le strike des sondages MT (Figure 4.12). Entre 1 et 0.01 Hz, le strike du tenseur de phase pour les trois sites varie fortement entre 30 et 60 degrés. La décomposition de Groom-Bailey permet de donner une estimation plus stable du strike de cette structure. La décomposition jointe des 3 sites MT aux fréquences inférieures à 1.5 Hz donne un strike égal à 47 degrés et la décomposition du sondage longues périodes donne un strike égal à 44.5 degrés.

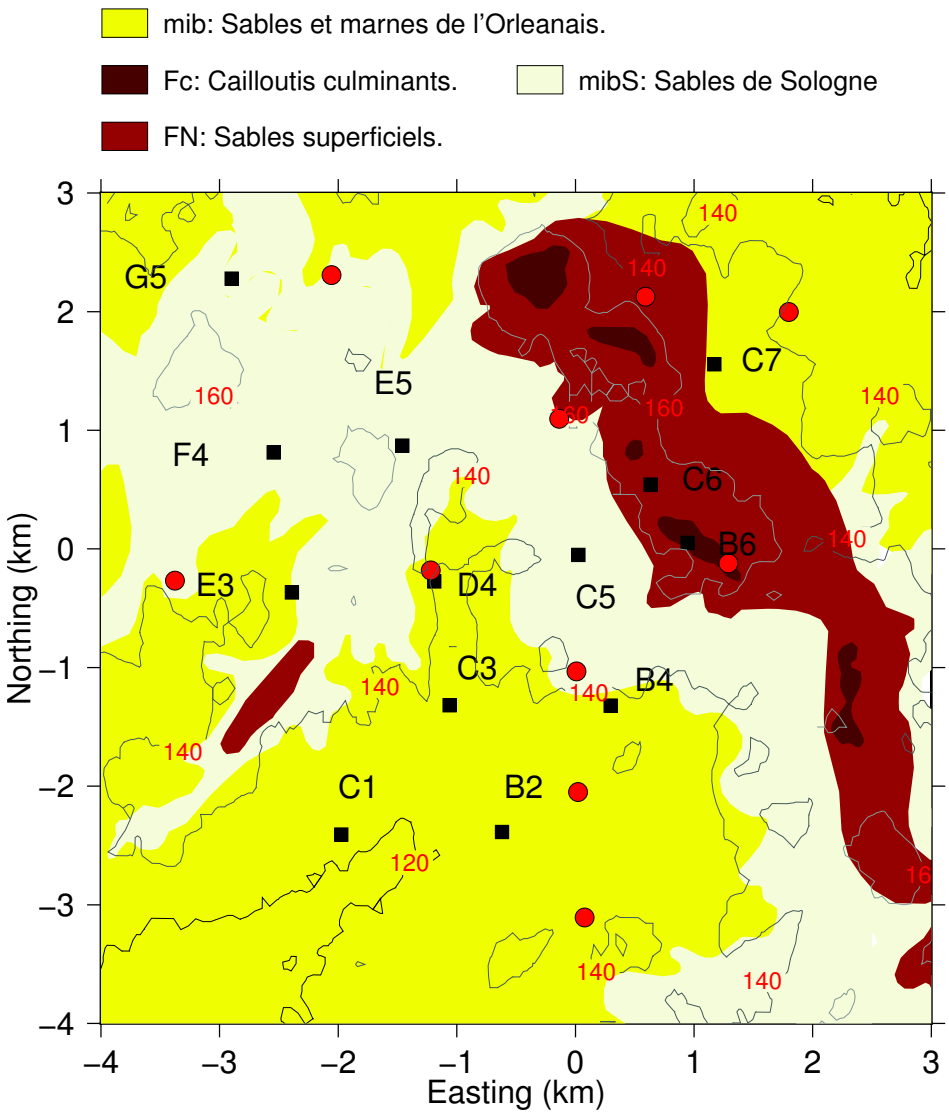


FIGURE 4.11: Carte géologique reproduite à partir de la carte au 1/50 000 fournie par le BRGM. Les points rouges indiquent la position des forages proche-surface.
FIGURE 4.11: *Géological map of the area reproduced from the 1/50 000 map available on BRGM website. Red dots indicates near-surface wells.*

4.4.3 Modélisation

4.4.3.1 1D

Dans les années 1980, des forages ont été effectués dans l'environnement immédiat de l'observatoire magnétique. Les profondeurs atteintes varient entre 50 mètres jusqu'à 200 mètres pour les forages dont le détail de la stratigraphie est disponible sur le site du Bureau de Recherches Géologiques et Minières (BRGM). Pour des profondeurs faibles, les forages présentent principalement les caractéristiques suivantes (Figure 4.13) :

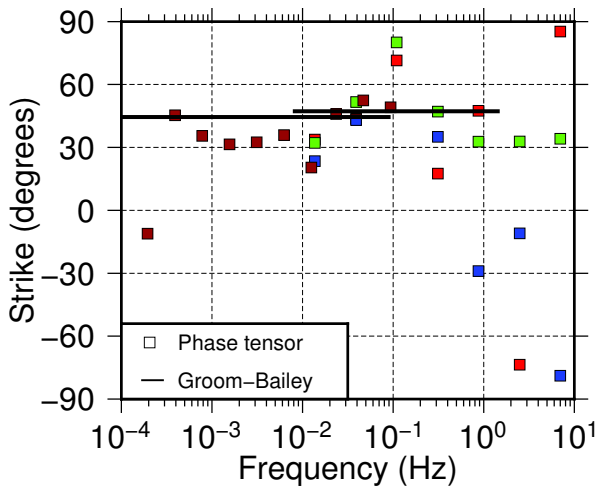


FIGURE 4.12: Strike obtenus par les différentes décompositions. Rouge : Sondage MT Observatoire. Vert : Sondage MT Thivet. Bleu : Sondage MT Ingrannes. Marron : Sondage dipôles d'observatoire.
FIGURE 4.12: *Strike obtained through the different decompositions. Red : Observatory MT sounding. Green : Thivet MT sounding. Blue : Ingrannes MT sounding. Brown : Observatory electric dipoles sounding.*

- Une couche superficielle de sables ou de mélange sable/argile, d'épaisseur variant entre 2 et 20 mètres.
- Une couche de Marnes et sables de l'Orléanais, d'épaisseur variant entre 19 et et 40 mètres.
- Une couche de calcaire complétant jusqu'à la profondeur maximale du forage.
- Une couche de craie pour le forage le plus profond, atteint à la côte de 200 mètres environ.

Les formations géologiques rencontrées correspondent aux valeurs de résistivité apparente observées sur les sondages Audio-MT, caractéristiques de formations sédimentaires conductrices. Ces variations latérales autour de l'observatoire de Chambon-La-Forêt, plus l'information sur la géologie de surface (Figure 4.11) expliquent la variation observée entre les différents sondages Audio-MT. D'autres sondages pétroliers ont été forés aux villes de Coudray, Châtenoy, Vitry-Aux-Loges, Saint-Denis-De-L'Hôtel, Juranville, Sully-La-Chapelle et Boiscommun. La cote maximale atteinte est de 1730 mètres pour le site de Boiscommun. Une interprétation géologique de ces différents forages a permis de proposer une carte du toit des différentes formations géologiques rencontrés dans les forages. La coupe géologique montrée Figure 4.13 mesure environ 13 km. Les sites Audio-MT sont situés au milieu de la coupe, là où les variations du toit des différentes couches sont faibles, en accord avec le caractère localement 1D des sondages Audio-MT. Pour obtenir un sondage représentatif de la distribution 1D de conductivité observé dans la région, nous avons modélisé par méthode d'essai-erreur le sondage Audio-MT de la station installée à l'observatoire magnétique en prenant pour principe de n'utiliser que le nombre minimal de couches possible. Le modèle obtenu contient 5 couches. La première couche correspond à la géologie de surface, correspondant aux sables de Sologne, de résistivité 8 $\Omega.m$, et d'épaisseur 5

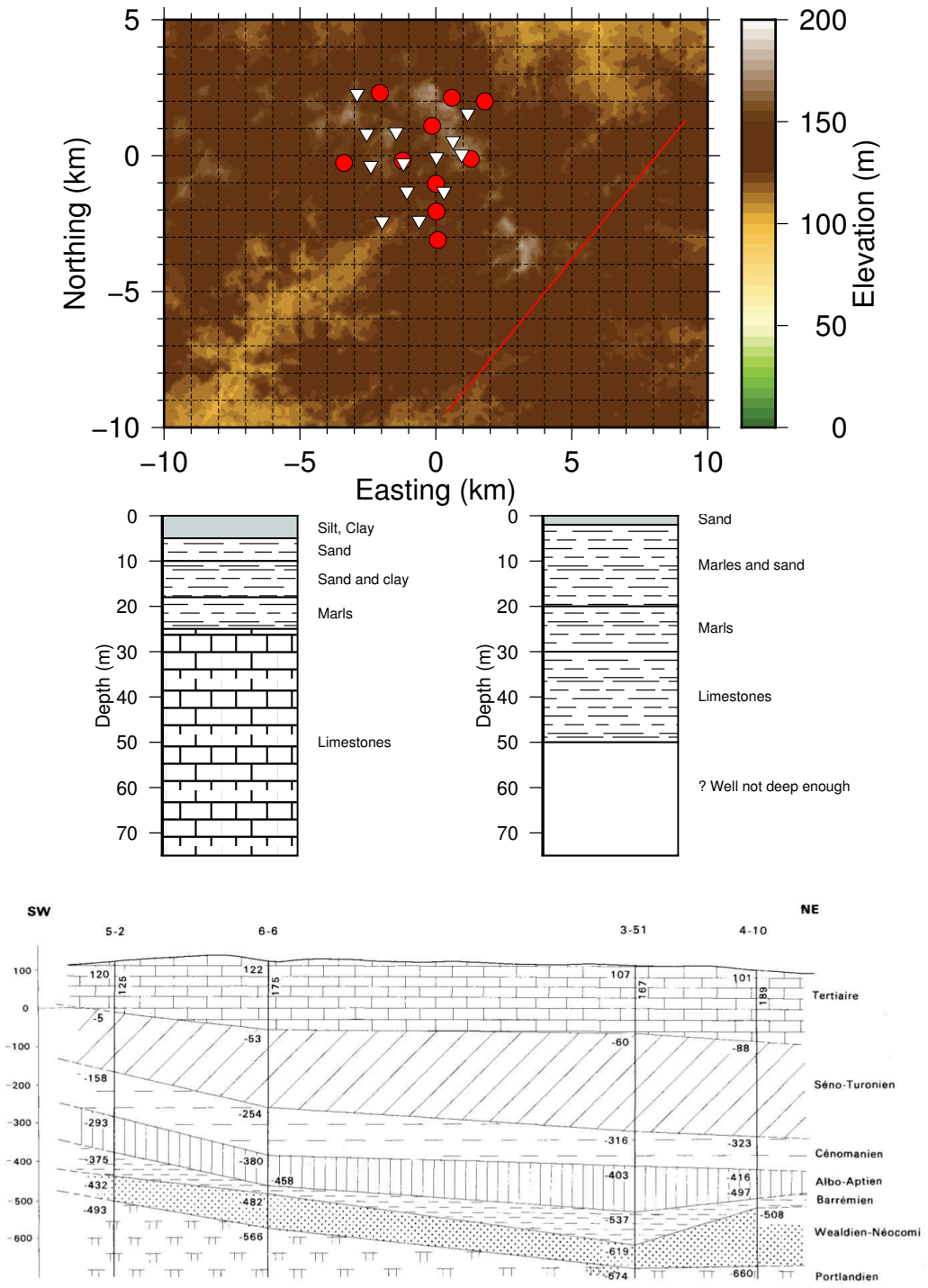


FIGURE 4.13: Haut : Position des sites Audio-MT (triangles blancs), des forages (points rouges) et de la coupe géologique (ligne rouge). Milieu : Exemple de sondages obtenus sur le site du BRGM. Gauche : 03642X0066. Droite : 03642X0085. Bas : Coupe géologique obtenue à partir de données de forage. D'après GIGOT, 1984.

FIGURE 4.13: *Top : Location of Audio-MT sites (white triangles), wells (red dots) and geological cross-section (red line). Middle : Example of wells, Left : 03642X0066, Right : 03642X0085. Bottom : Geological cross-section from drilling information. From* GIGOT *(1984).*

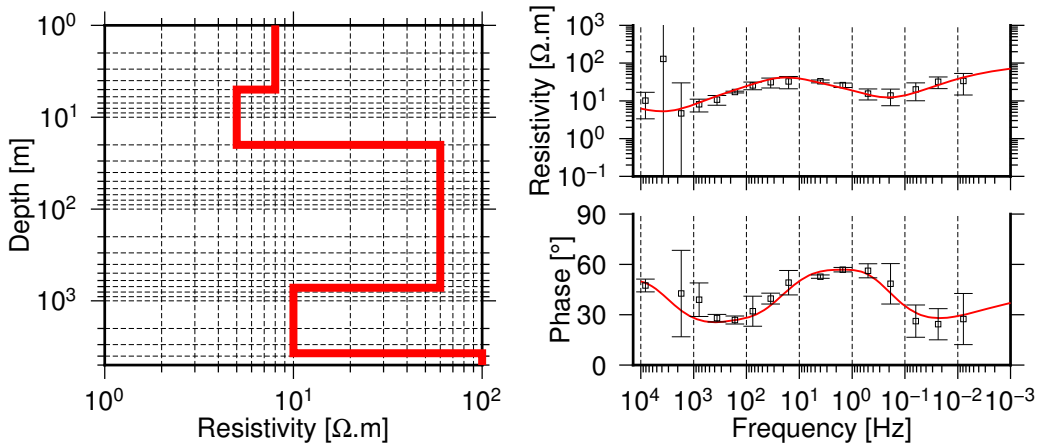


FIGURE 4.14: Gauche : Modèle 1D de résistivité utilisé dans les modélisations. Droite : Comparaison entre le sondage correspondant au modèle 1D utilisé à gauche dans les modélisations et le sondage MT obtenu à l'observatoire.
FIGURE 4.14: *Left : 1D resistivity model used in the modeling. Right : Comparison between the modeling of the mean 1D model and the Observatory MT sounding.*

m. La deuxième couche, plus conductrice de résistivité 5 $\Omega.m$, et d'épaisseur 15 m, correspond aux marnes et sables de l'Orléanais. Nous avons ensuite utilisé une couche de 700 m d'épaisseur et de résistivité 60 $\Omega.m$ pour représenter les différentes alternances de calcaires et craies, formations plus compactes et plus résistantes. Une couche de 3000 m d'épaisseur et de 10 $\Omega.m$ et un demi-espace de résistivité 100 $\Omega.m$ complètent le modèle 1D. La Figure 4.14 montre le profil du modèle de résistivité vraie et la comparaison entre les données calculées et observées.

4.4.3.2 3D

Effet topographique La topographie étant relativement faible dans la région, nous sommes partis d'un modèle 1D moyen et avons modélisé l'effet de distorsion du à la topographie par rapport à un modèle sans topographie. Le modèle 1D utilisé est celui décrit à la Figure 4.14. Nous avons utilisé le code ModEM pour effectuer la modélisation du milieu 3D. Le modèle de topographie utilisé est le modèle SRTM (Farr et al., 2007), qui a une résolution d'environ 90 mètres dans la région. Le modèle utilisé est de 70 × 70 × 110 cellules, pour une taille de 135 × 135 × 65 km. La distribution de résistivité est parallèle à la topographie. Pour chaque fréquence et chaque composante anti-diagonale, nous avons calculé l'écart défini par :

$$\Delta\rho = \frac{|\rho_{topo} - \rho_{notopo}|}{\rho_{topo}} \tag{4.1}$$

$$\Delta\phi = d(\varphi_{topo}, \varphi_{notopo}), \tag{4.2}$$

avec $d(\varphi_1, \varphi_2)$ la distance sur le cercle trigonométrique.

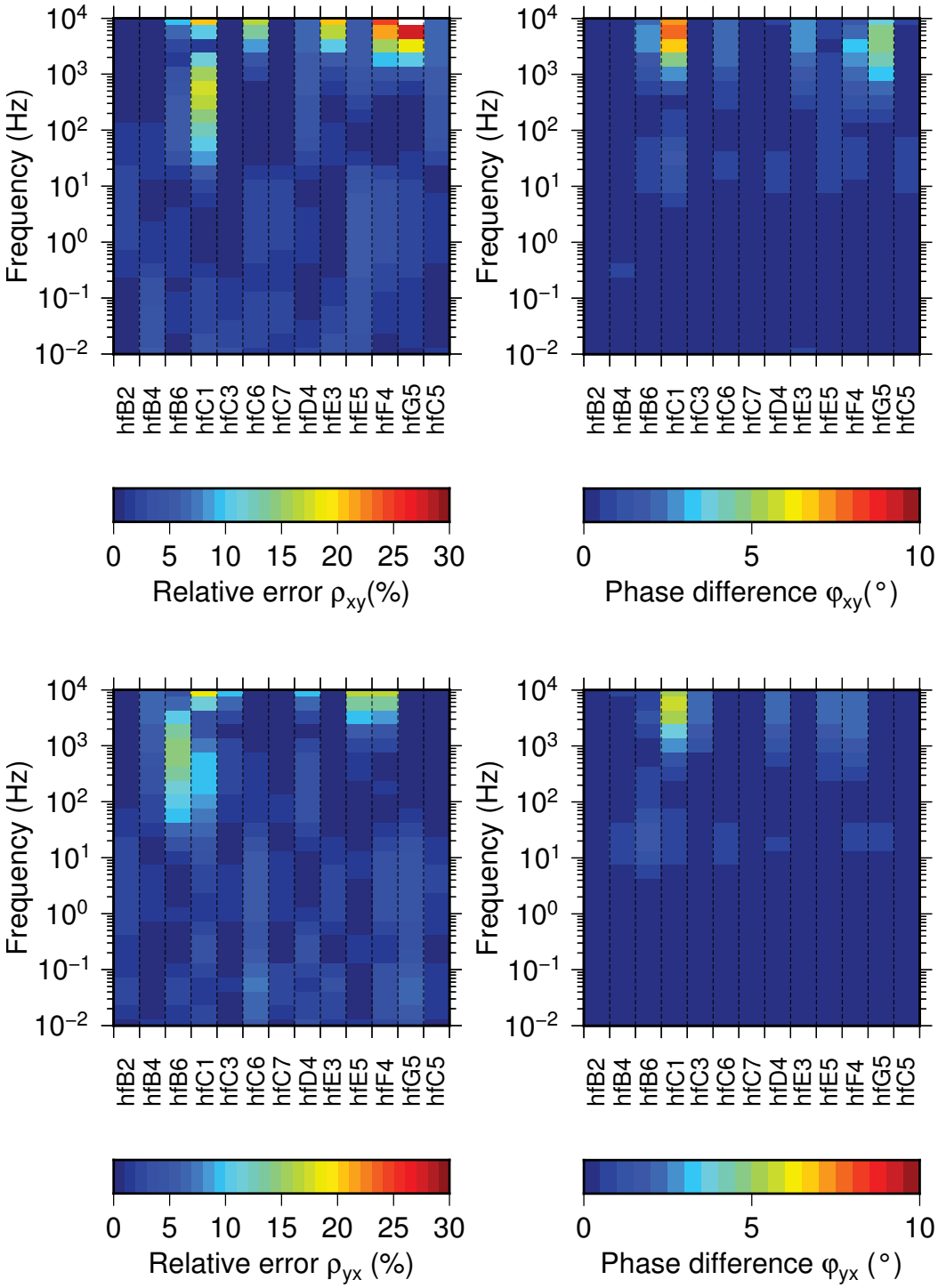


FIGURE 4.15: Comparaison entre les modélisations de sondage MT pour des modèles avec ou sans topographie aux positions de la campagne.

FIGURE 4.15: *Comparison between the modeling of MT soundings for 3D models with and without topography.*

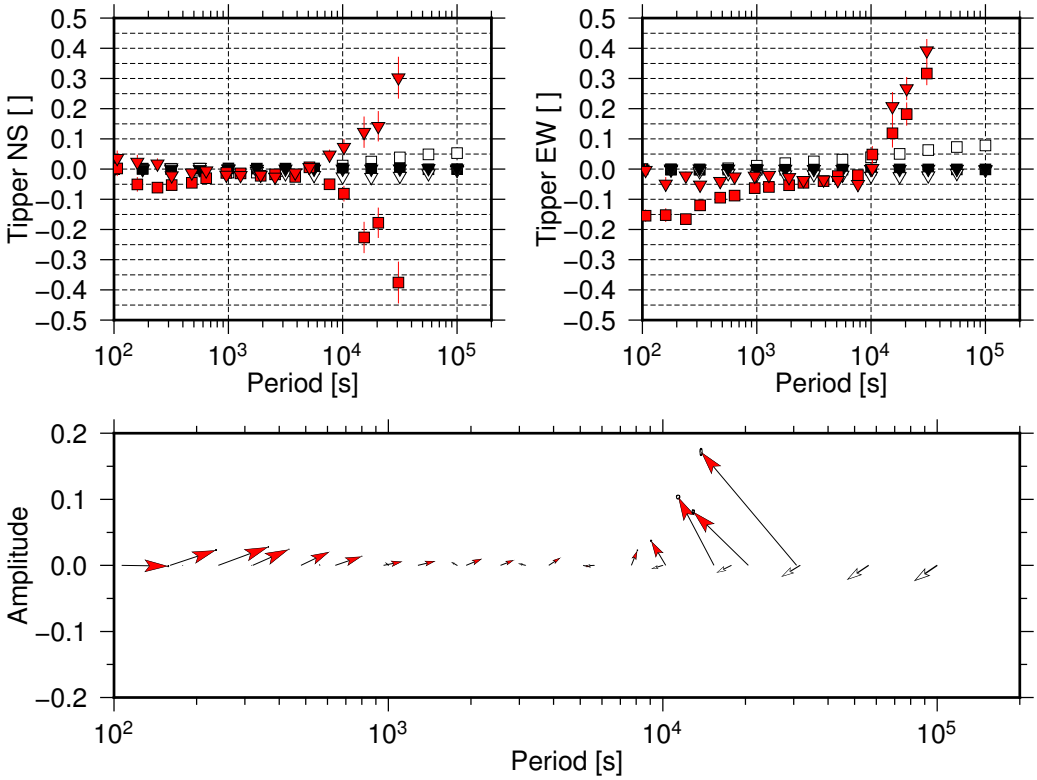


FIGURE 4.16: Comparaison entre les modélisations de sondage MT pour des modèles avec ou sans océans aux positions de la campagne. Haut : Fonctions de transfert verticales réelles et imaginaires. Bas : Direction des flèches d'induction réelles. Rouge : Données. Blanc : Modèle avec océan. Noir : Modèle sans océan.
FIGURE 4.16: *Comparison between the modeling of MT soundings for 3D models with and without oceans. Top : Vertical transfer function. Bottom : Direction of induction arrows. Red : Data. White : Ocean model. Black : Land model.*

L'effet topographique (Figure 4.15), comme prévu, est relativement limité aussi bien dans le contenu fréquentiel que dans l'espace. Seules quelques stations comme C1 ou E5 présentent une distorsion pour des fréquences supérieures à 1 kHz. Aux fréquences inférieures au kHz, la différence entre le modèle avec topographie et celui sans topographie est très faible. la distorsion observée n'est pas de l'ordre du simple *static shift* puisque la phase est aussi affectée.

Effet de l'océan Un autre effet notable sur le champ magnétique est l'apparition d'une composante verticale du champ magnétique par induction de courants électriques dans les structures conductrices. Un des effets couramment observé est l'induction de courants électriques dans les océans. La mer la plus proche se situe à environ 230 kilomètres au Nord-Ouest avec la mer Manche. Nous avons donc créé un modèle de 110 × 110 × 50 cellules couvrant un volume de 1150 × 1150 × 720 km. Ainsi, l'océan Atlantique et la Manche se retrouvent aux bordures du modèle. Deux modèles sont créés, l'un avec les cellules correspondant à celles de l'océan avec une conductivité de 3 S.m, l'autre en remplaçant la résistivité de ces cellules par celles du milieu 1D

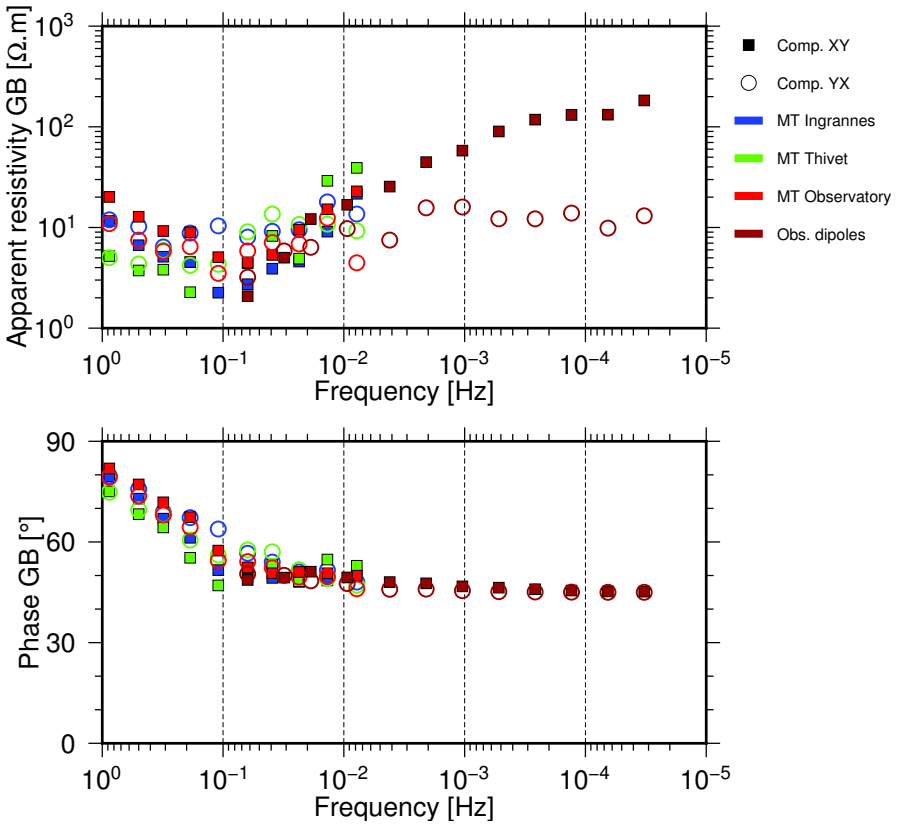


FIGURE 4.17: Sondages MT obtenus par la décomposition de Groom-Bailey avec une direction de strike 45°.

FIGURE 4.17: *Groom-Bailey MT soundings with a strike direction of 45°.*

correspondant. La modélisation des fonctions de transfert verticales (Figure 4.16) montre que l'océan a une influence non négligeable au fur et à mesure que la fréquence diminue. La partie réelle des flèches d'induction pointe d'abord vers la direction NO, vers la manche, puis vers l'Est, vers l'Océan Atlantique.

En revanche, même si l'amplitude des fonctions de réponse verticales dues aux mers ne sont pas négligeables, elles sont d'environ un ordre de grandeur inférieure à l'amplitude observée du *tipper*. Ainsi, la composante verticale observée provient en majorité de l'anomalie 2D observée.

4.4.3.3 2D

Pour étudier la structure 2D régionale, nous avons décomposé les sondages MT des stations mobiles et celui calculé avec les dipôles d'observatoire selon la méthode de Groom et Bailey (1989) en prenant un strike de 45 degrés. La Figure 4.17 montre le résultat de la décomposition. Les quatre sondages, du à leur proximité, ne présentent pas de différences significatives. La caractéristique principale de ce sondage décomposé est la séparation des composantes principales à partir de 0.05 Hz. La composante xy augmente fortement pour atteindre des valeurs supérieures

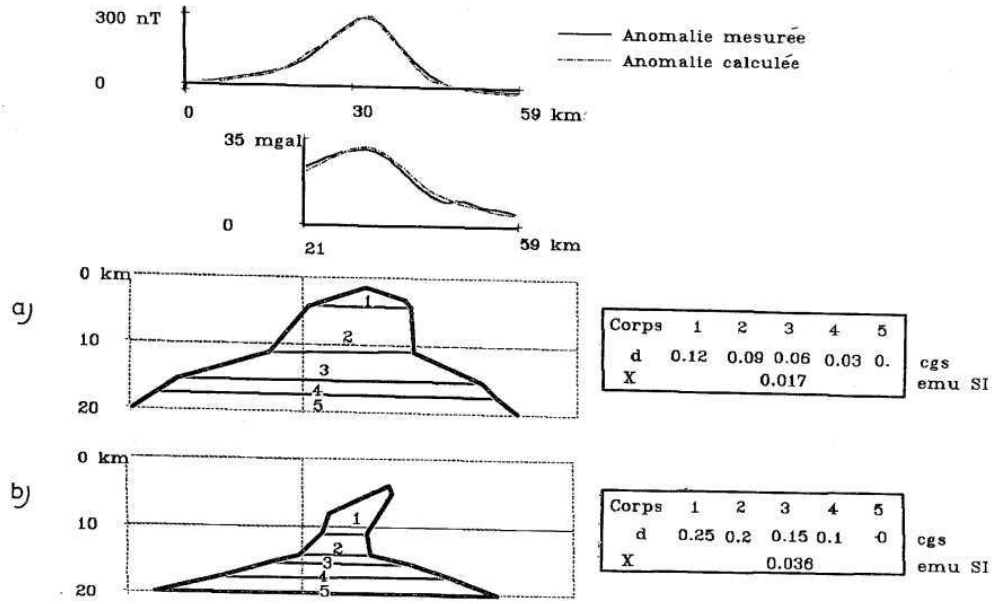


FIGURE 4.18: Modèles d'origine de l'anomalie magnétique du bassin parisien proposés par AUTRAN et al. (1986).
FIGURE 4.18: *Models for the origin of the Paris basin magnetic anomaly. From* AUTRAN *et al. (1986)*

à 100 $\Omega.m$ pour des fréquences inférieures à 0.0001 Hz. En revanche, la composante yx n'évolue que peu pour atteindre des valeurs légèrement supérieures à 10 $\Omega.m$ pour des fréquences inférieures à 0.0001 Hz.

Avec si peu de sondages MT, et si rapprochés, il est en revanche délicat d'effectuer une inversion des sondages MT. Nous avons cherché dans la littérature les différents modèles représentant l'anomalie du bassin de Paris pour déterminer si un modèle pouvait correspondre aux sondages observés. Le modèle le plus détaillé trouvé est celui décrit dans AUTRAN et al. (1986) représenté Figure 4.18. Nous avons incorporé ce modèle dans différents tests de modélisations en supposant l'anomalie résistance ou conductrice, sans nécessairement chercher à reproduire les courbes de résistivité apparente, mais plutôt à retrouver les caractéristiques principales des sondages décomposés observés. Le modèle 1D de référence utilise celui décrit Figure 4.14, sans utiliser les couches superficielles de résistivités 8 et 5 $\Omega.m$ et d'épaisseurs respectives 5 et 15 mètres. L'anomalie est à 1000 $\Omega.m$ pour le cas résistant et à 10 $\Omega.m$ dans le cas conducteur. La Figure 4.19 montre les deux modèles.

Aucun sondage du modèle avec anomalie résistante ne présente le comportement observé sur les sondages régionaux (Figure 4.20). Pour les sondages correspondants aux stations situées au dessus de l'anomalie, les courbes de résistivité apparente présentent le comportement contraire à celui observé, avec la composante yx qui devient d'amplitude supérieure à xy.

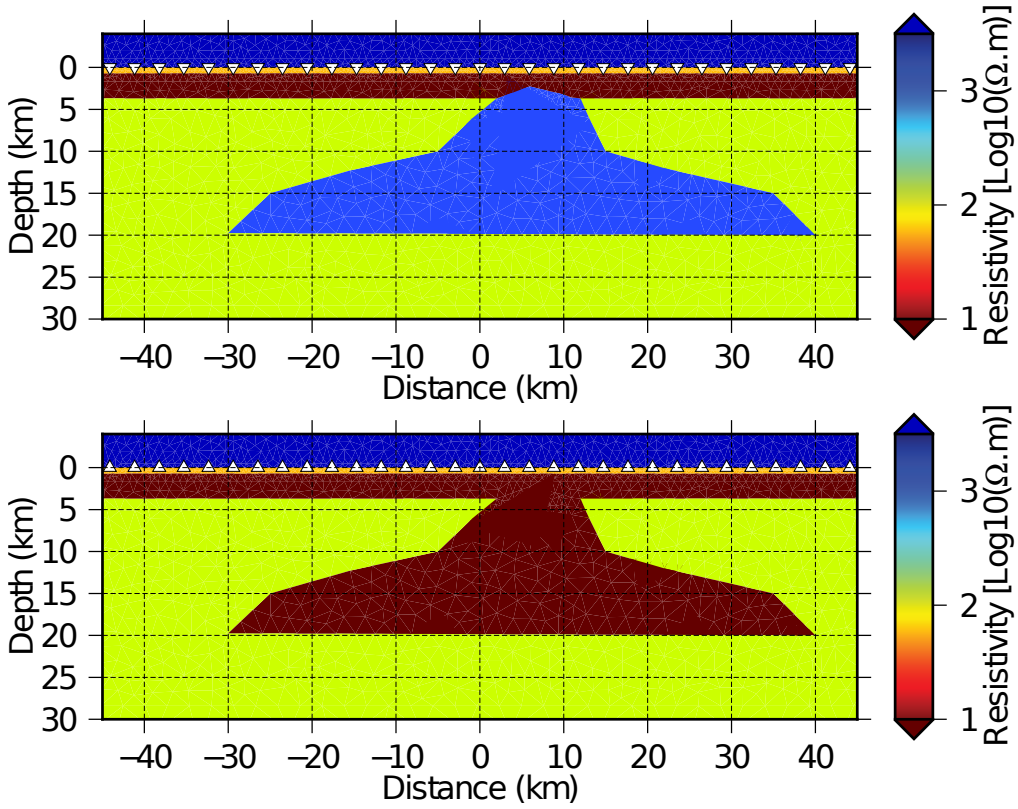


FIGURE 4.19: Modèles de conductivité 2D utilisés pour le calcul des réponses MT. Haut : Modèle avec une anomalie résistante. Bas : Modèle avec une anomalie conductrice. Les triangles blancs indiquent la position des stations MT.
FIGURE 4.19: *2D conductivity models for the computation of MT response functions. White triangles indicates the location of MT stations.*

Pour le modèle conducteur, la situation est inverse. Certains sondages présentent une augmentation de la composante xy avec une stagnation de la composante yx (Figure 4.20). Cependant, cette variation intervient à des fréquences inférieures à celles observées. Nous avons également modélisé cette anomalie pour une résistivité inférieure à 5 $\Omega.m$, sans variation significative du comportement si ce n'est une diminution de la résistivité apparente. Ainsi, la théorie d'une anomalie conductrice semble plus probable que celle d'une anomalie résistante. Mais encore une fois, au vu des études précédentes, la dimension de l'anomalie est probablement 3D et nécessite une étude magnétotellurique d'envergure pour prétendre pouvoir la délimiter correctement.

4.5 Conclusion

Dans ce chapitre, nous avons étudié les résultats d'une première campagne d'acquisition MT autour de l'observatoire de Chambon-La-Forêt. Deux principaux résultats sont à mettre en relation vis-à-vis du principe d'un observatoire magnétique.

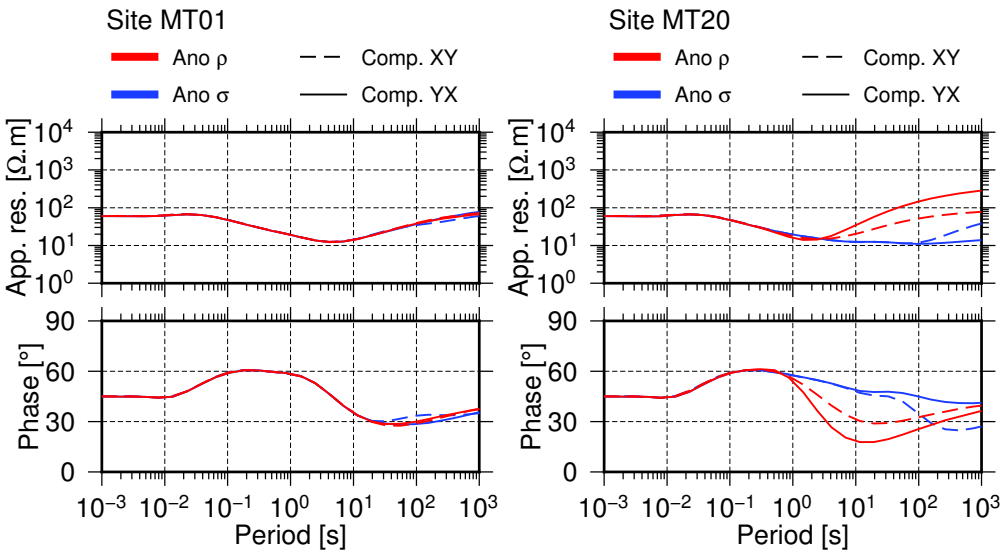


FIGURE 4.20: Sondages modélisés aux stations situées à -50 km (gauche) et à 5.8 km (droite).
FIGURE 4.20: *Response functions modelling for MT stations located at -50 km (left) and 5.8 km (right).*

Tout d'abord, les mesures EM effectuées aux alentours de l'observatoire montrent un niveau de bruit très important, aussi bien dans les données hautes-fréquences que basses-fréquences.

En revanche, en considérant la distribution de résistivité, l'observatoire est idéalement situé en un lieu où les distorsions du champ magnétique sont quasi-inexistantes, nonobstant la création d'une composante verticale par variation latérale de résistivité. L'analyse des sondages MT montre que la source de cette variation latérale de conductivité se trouve dans la même direction que l'anomalie magnétique du bassin parisien. En revanche, avec le peu de sites basses-fréquences dont nous disposons il n'est pas possible de préciser les limites spatiales de cette anomalie.

Chapitre 5

Interprétation de données MT au Népal

Ce chapitre présente l'interprétation de données MT acquises au Népal pendant la campagne effectuée en mars 2016, en collaboration avec le *Department of Mines and Geology* (DMG) de Kathmandu, Népal et l'*Earth Observatory of Singapore* (EOS). Les travaux présentés ici font suite à ceux effectués par Carole Lemonnier pendant sa thèse à l'Institut de Physique du Globe de Strasbourg entre 1996 et 2000.

Dans ce chapitre, nous présentons brièvement le contexte géologique du Népal et les travaux antérieurs effectués dans la région étudiée. Nous décrivons ensuite l'acquisition et le traitement des données MT effectués. Les nouveaux sondages MT obtenus sont ensuite comparés aux anciens sondages obtenus lors des campagnes 1996, 1999 et 2000. L'analyse des distorsions est ensuite effectuée et l'influence de la topographie est discutée. Enfin, l'inversion 2D des données est présentée.

5.1 Introduction

5.1.1 Contexte géologique

5.1.1.1 Collision Inde-Eurasie

Le Népal se situe sur la chaine Himalayenne, en bordure des hauts massifs tels que l'Everest ou les Annapurnas. L'Himalaya est le résultat de la collision continentale entre l'Inde et l'Eurasie (Royer et Patriat, 2002), collision initiée il y a environ 250 Ma par la dislocation du super-continent initial, la Pangée. A cette époque, l'ensemble des continents forme deux grandes masses continentales, le Gondwana (Amérique du Sud, Afrique, Arabie, Inde, Madagascar, Antarctique, Australie), et la Laurasie (Amérique du Nord, Groenland, Europe et Sibérie). Ces deux continents sont séparés par une mer appelée Téthys. A environ 210 Ma, le Gondwana se disloque, par la

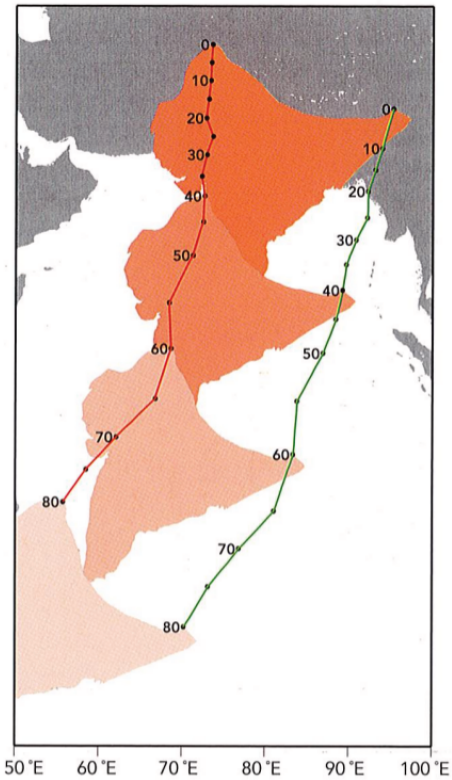


FIGURE 5.1: Mouvement de l'Inde par rapport à la plaque eurasienne. D'après ROYER et PATRIAT (2002).
FIGURE 5.1: *Displacement of Indian plate relative to Eurasia plate. From* ROYER *et* PATRIAT *(2002).*

perte d'un petit continent qui migre vers le Nord en direction de la Leurasie, séparant la Téthys en deux mers ; Néotéthys au sud, Paléotéthys au nord. Le Gondwana rentre en collision avec la Leurasie, fermant la Paléotéthys. C'est la collision Cimmérienne. Les deux continents Laurasie et Gondwana se séparent par la création d'un rift aboutissant à l'ouverture de l'Océan Atlantique entre 180 et 80 Ma. Les deux continents se réorganisent et d'autres rifts se créent permettant l'apparition de l'Océan Atlantique Sud et Indien. Vers 80 Ma, l'Inde se sépare du Gondwana et entame une progression vers le Nord en direction de l'Eurasie (Figure 5.1) à la vitesse de 17 cm/an pour rentrer en collision avec celle-ci entre 60 et 50 Ma en fermant la Néotéthys au niveau de l'ancienne zone de subduction.

5.1.1.2 Structure de la chaîne Himalayenne

Structuralement parlant, la chaîne Himalayenne se situe au sud de la zone de suture créée lors de la disparition de la Néotéthys. Cette chaîne, d'environ 300 km de largeur, se développe du nord de l'Inde jusqu'aux hauts plateaux tibétains suivant une direction moyenne SSW-NNE. De

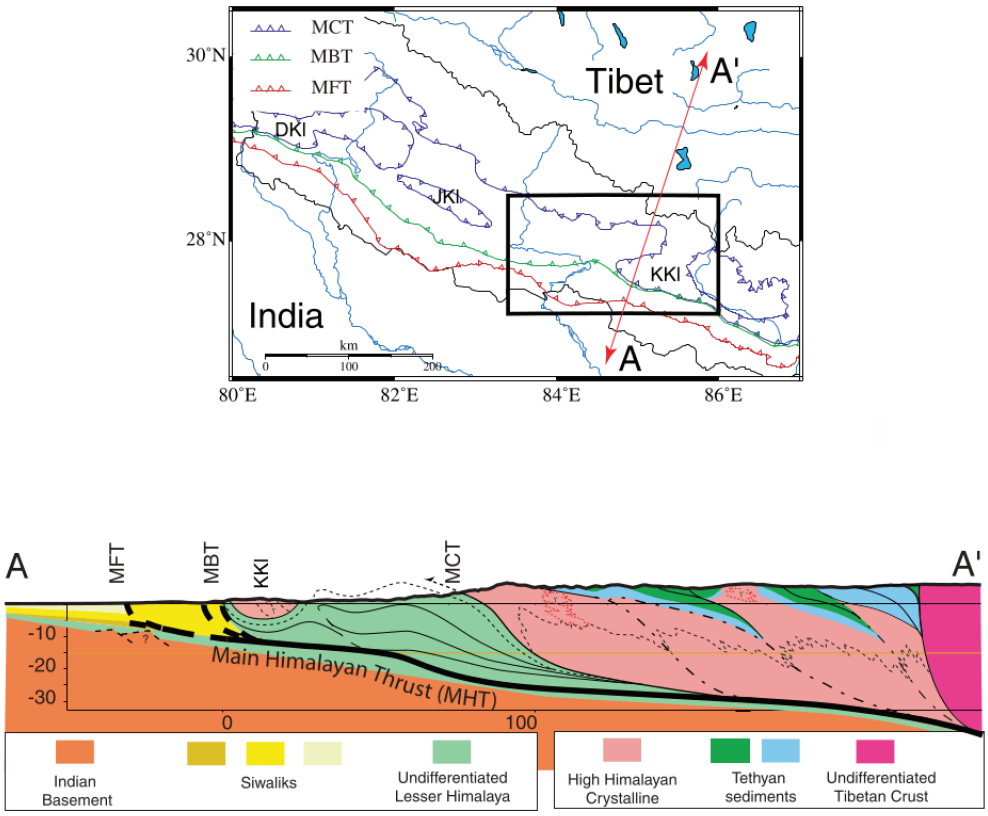


FIGURE 5.2: Haut : Position de la coupe géologique traversant la chaîne Himalayenne. Bas : Schéma des principales formations géologiques de la chaîne Himalayenne. DKI : Dadeldhura klippe. JKI : Jajarkot klippe. KKI : Katmandhu klippe. D'après BOLLINGER et al. (2004). Voir le texte pour la description des formations géologiques.

FIGURE 5.2: *Top : Position of geological cross-section across the Himalayan chain. Bottom : Sketch of main geological formations of Himalayan chain. DKI : Dadeldhura klippe. JKI : Jajarkot klippe. KKI : Katmandhu klippe. From* BOLLINGER *et al. (2004). See text for a description of geological structures.*

nombreuses études géologiques et géophysiques ont permis de déterminer la position des structures principales constituant la chaîne Himalayenne. Cinq grandes régions sont principalement distinguées (Figure 5.2) :

- La plaine du Gange.
- Le domaine des Siwalik.
- Le Moyen-Himalaya.
- Le Haut-Himalaya.
- Le plateau tibétain.

Chaque région est séparée par des failles structurelles majeures appelées chevauchements. Les failles majeures situées dans la chaîne Himalayenne sont le *Main Frontal Thrust* (MFT), le *Main Boundary Thrust* (MBT) et le *Main Central Thrust* (MCT). Ces trois failles majeures permettent le chevauchement des structures du Nord au Sud (Figure 5.2). Ainsi, le domaine

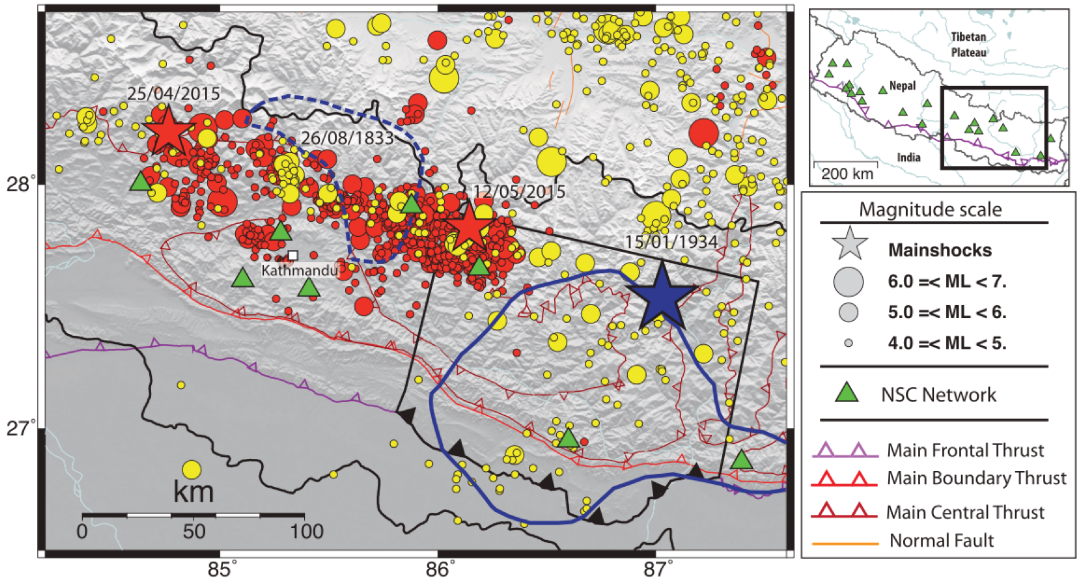


FIGURE 5.3: Carte des répliques du séisme de Gorka au niveau de l'Himalaya. KKI : Katmandhu klippe. D'après ADHIKARI et al. (2015).
FIGURE 5.3: *Map of Gorka earthquake aftershocks. KKI : Katmandhu klippe. From* ADHIKARI *et al. (2015).*

des Siwalik chevauche la plaine du Gange au niveau du MFT, le Moyen-Himalaya chevauche les Siwalik par le MBT, et le Haut-Himalaya chevauche le Moyen-Himalaya par le MCT. L'ensemble de ces chevauchements se rejoignent en profondeur au niveau du *Main Himalayan Thrust* (MHT).

5.1.2 Le séisme de Gorkha

Le 25 Avril 2015 à 11h56 heure locale, un séisme de magnitude M_w 7.8 (M_l 7.6) (ADHIKARI et al., 2015) a frappé la partie centrale du Népal causant plus de 8700 décès et laissant des dizaines de milliers de personnes sans foyer dans la région de Kathmandu jusqu'au pied du Haut-Himalaya. L'épicentre du séisme a été localisé dans la région de Gorkha, à l'est de Kathmandu à environ 15 km de profondeur (AVOUAC et al., 2015). Cette profondeur a conduit à l'hypothèse de la rupture d'une partie du MHT. 553 répliques de magnitude supérieure à 4 ont été enregistrées dans la région pendant les 45 premiers jours suivant le séisme (Figure 5.3).

L'étude de ce séisme a permis d'établir qu'il a débloqué une portion inférieure du MHT (AVOUAC et al., 2015), et de préciser la géométrie des failles actives en profondeur (HUBBARD et al., 2016).

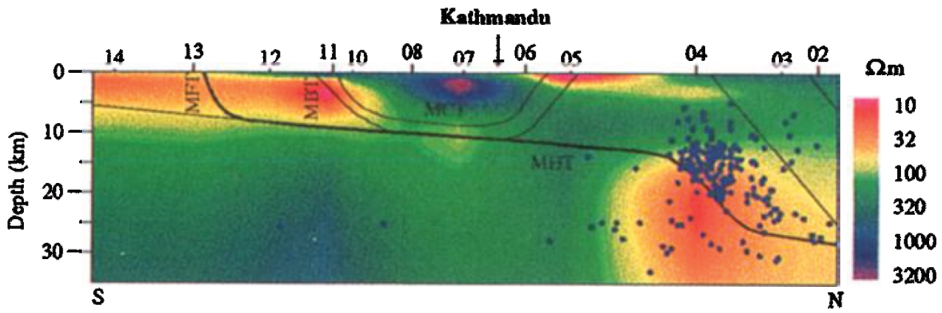


FIGURE 5.4: Modèle de conductivité électrique dans la vallée de Katmandou. D'après Carole LEMONNIER et al. (1999).
FIGURE 5.4: *Electrical conductivity model in Katmandu valley. From Carole* LEMONNIER *et al. (1999).*

5.1.3 Travaux antérieurs en magnétotellurique sur la chaîne Himalayenne

La méthode magnétotellurique, de part son aptitude à imager les structures profondes a été largement utilisée dans l'étude de la chaine Himalayenne, de la jonction entre la plaine du Gange aux hauts plateaux tibétains (J. E. SPRATT et al., 2005 ; M. UNSWORTH et al., 2005 ; LE PAPE et al., 2012).

Les premiers travaux en magnétotellurique dans la région centrale du Népal dans la partie inférieure de l'Himalaya ont été effectués par Carole LEMONNIER et al. (1999). 12 stations MT ont été installées dans la vallée de Katmandou et un premier profil de conductivité électrique a été établi. Ce profil, obtenu par l'inversion des stations MT pour des périodes entre 0.001 et 1000 secondes par la méthode de *Rapid Relaxation Inversion (RRI)* (J. T. SMITH et John R BOOKER, 1991), postule l'existence d'une structure conductrice sur le MHT (Figure 5.4). La structure conductrice obtenue, coïncidant avec une zone sismiquement active, a été interprétée comme une zone où les contraintes tectoniques augmentent permettant une augmentation de la connectivité des roches plus une possibilité de fusion partielle, entraînant ainsi une augmentation de la conductivité.

5.2 Campagne magnétotellurique 2016

La campagne magnétotellurique 2016 s'est effectuée du 03 au 15 mars 2016. 3 sites ont d'abord été acquis aux environs de Katmandu (Tokha, Kakani et Patle) puis nous nous sommes installés à Dhunche pour acquérir les 6 derniers points dans la vallée de Trisuli (Figures 5.5 et 5.6). Le détail de l'acquisition est donné en annexe. Les principales difficultés rencontrées pendant le déroulement de la campagne sont principalement reliées à la localisation de sites suffisamment plans et éloignés de sources de bruits anthropiques.

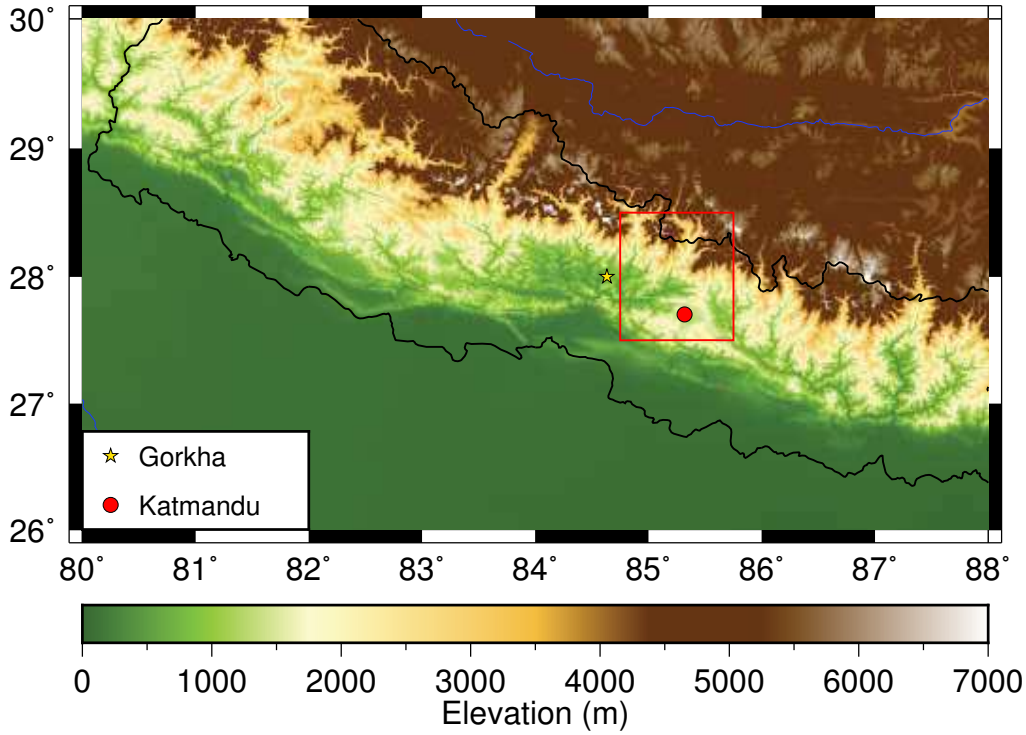


FIGURE 5.5: Localisation de la zone étudiée. Modèle topographique extrait du modèle ETOPO1 (AMANTE et EAKINS, 2009).
FIGURE 5.5: *Location of studied area. Topographic data is taken from ETOPO1 model (AMANTE et EAKINS, 2009).*

Pour obtenir un rapport signal-bruit suffisant sur les séries temporelles permettant d'obtenir des réponses MT de 0.01 à 1000 secondes, les tailles de dipôles recommandées sont d'environ 50 mètres minimum à installer à la surface du sol, à l'horizontale. Les cultures en montagne au Népal utilisent le principe des terrasses pour optimiser la surface des cultures. La taille horizontale des terrains nous a donc forcé à installer des stations en montagne où la taille des dipôles est inférieure à 30 mètres. La direction Est-Ouest du site de Patle mesure par exemple 27 mètres. De plus, l'accès et l'installation des points MT n'a pu se faire que le long de la route de la vallée de Trisuli, route qui suit le trajet des pylônes électriques permettant l'approvisionnement

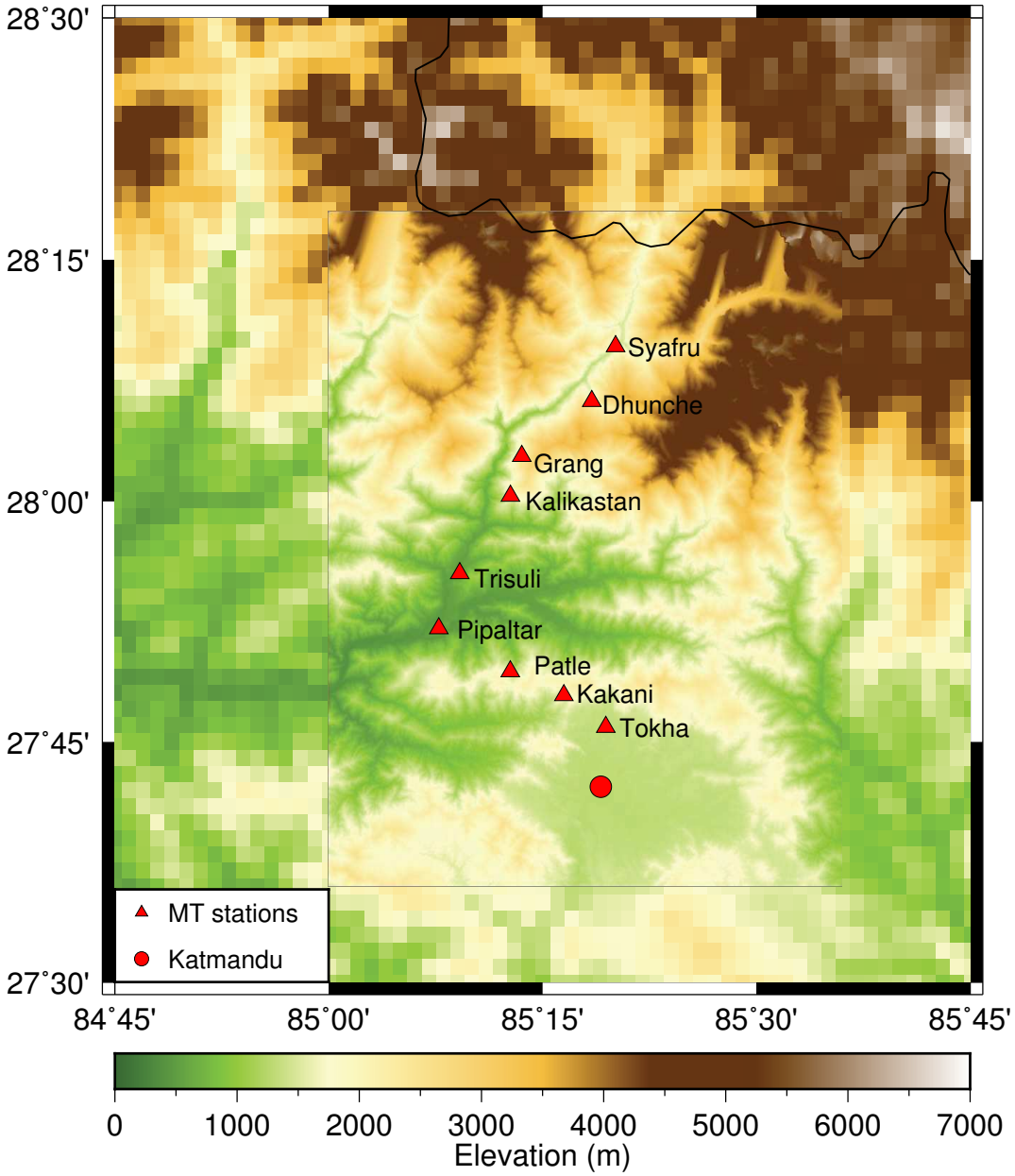


FIGURE 5.6: Plan des stations magnétotelluriques installées. Topographie combinée des modèles ETOPO1 (AMANTE et EAKINS, 2009) et SRTM (FARR et al., 2007).
FIGURE 5.6: *Map of magnetotelluric stations. Topography extracted from ETOPO1 (AMANTE et EAKINS, 2009) and SRTM (FARR et al., 2007) models.*

en électricité des villages situés dans la vallée. Ainsi, les séries temporelles enregistrées peuvent présenter de larges niveaux de bruit.

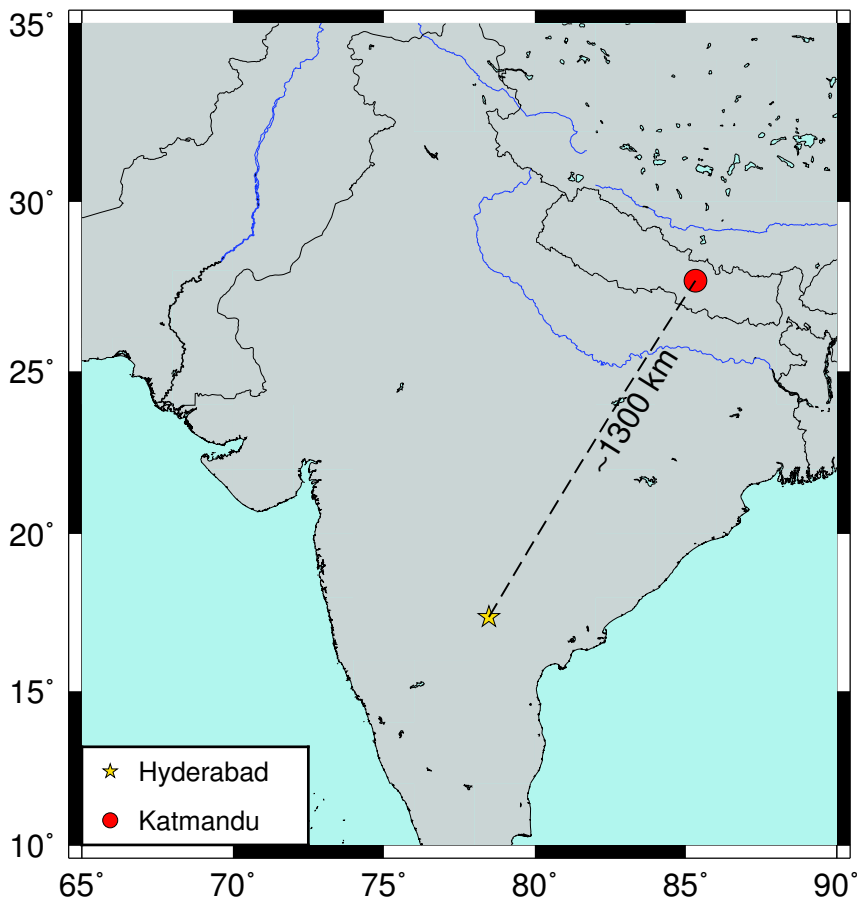


FIGURE 5.7: Carte indiquant la position de la campagne MT et de l'observatoire magnétique utilisée comme station de référence.
FIGURE 5.7: *Map of local MT stations in Nepal and HYB magnetic observatory India used as reference station.*

5.3 Traitement des données

Les données MT ont été traitées en utilisant l'algorithme BIRRP dont le principe a été décrit au chapitre 3. Les périodes inférieures à 32 secondes ont été traitées en utilisant une autre station MT en acquisition simultanée (voir le planning de l'acquisition décrit en annexe). Pour les périodes à la dizaine de secondes, nous avons utilisé les mesures magnétiques de l'observatoire d'Hyderabad (HYB, Inde) en tant que station de référence. Cet observatoire est situé à environ 1300 km de la campagne de mesure au Népal (Figure 5.7). Comme décrit au Chapitre 2, les variations du champ magnétique externe sont observables et cohérentes sur plusieurs centaines de kilomètres, ainsi, l'utilisation d'une station lointaine ne pose pas nécessairement de problème dans les procédures de statistiques robustes. A la Figure 5.8, nous représentons les séries temporelles magnétiques communes entre l'observatoire magnétique HYB et la station MT de Grang le 14 Mars 2016. La Figure 5.8 montre la comparaison entre les séries temporelles de ces deux

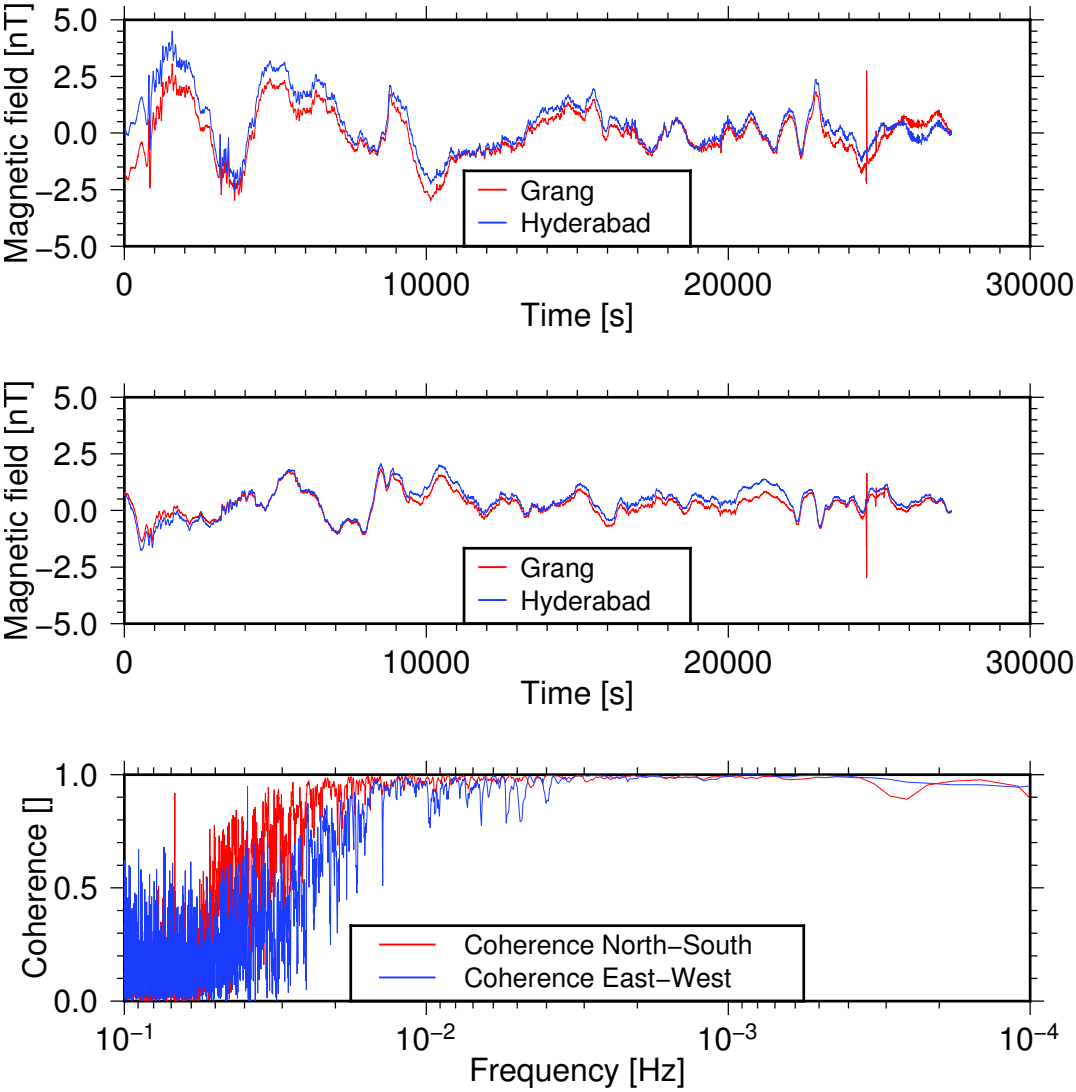


FIGURE 5.8: Haut : Comparaison des composantes magnétiques Nord-Sud entre locale et *remote*. Milieu : Comparaison des composantes magnétiques Est-Ouest entre locale et remote. Bas : Cohérence entre les composantes locales et *remotes* pour les mêmes directions.
FIGURE 5.8: *Top : Comparison between North-South magnetic components at local and remote station. Top : Comparison between East-West magnetic components at local and remote station. Bottom : Coherency between local and remote stations.*

enregistrements, ainsi qu'un calcul de cohérence entre les voies Nord-Sud et Est-Ouest pour s'assurer que l'observatoire magnétique peut être utilisé comme station de référence.

Les cohérences ainsi observées montrent une très forte similitude (> 0.7) entre les séries temporelles de l'observatoire magnétique et celles des stations MT pour des périodes supérieures à la trentaine de secondes. L'utilisation d'une *remote* de ce type nous permet également de s'affranchir complètement des effets de distorsion locale ou de bruits électromagnétiques locaux qu'il est courant d'observer au Népal.

Ainsi, les sondages obtenus nous permettent d'avoir des estimations de très bonne qualité pour des périodes comprises entre 0.005 et 1000 secondes pour la majeure partie stations du profil Nord-Sud. Seule une voie électrique de la station de Trisuli est contaminée par un niveau de bruit important, empêchant d'obtenir de bonnes estimations des fonctions de réponses MT pour les composantes Z_{yx} et Z_{yy}.

5.4 Analyse des sondages

5.4.1 Comparaison avec les anciennes campagnes

La première étape de notre travail consiste à comparer les données acquises en 1996, 1999 et 2000 à celles de la nouvelle campagne de mesures pour déterminer si une variation importante des sondages était déjà observable. Nous avons donc récupéré les sondages calculés en 1996, sans refaire le calcul des fonctions de réponse MT sur les séries temporelles brutes. Le traitement des données MT était fait avec l'ancienne version du code BIRRP. Les comparaisons pour les 6 stations du profil nord-Sud sont présentées aux Figures 5.9, 5.10, 5.11 5.37, 5.38, et 5.39.

La différence principale entre les deux campagnes s'observe principalement sur la qualité des sondages MT obtenus. En effet, certains sites comme Kalikastan (Figure 5.10) présentent une forte variabilité des sondages MT en 1996 qui incite à la prudence quand à leur interprétation. Pour la station de Pipaltar (Figure 5.11), un décalage d'amplitude est même observable sur les quatre composantes du tenseur MT entre les deux campagnes MT. En revanche, pour les stations telles que Grang, Dhunche ou Syafru, les résistivités apparentes déterminées sont du même ordre de grandeur. En règle générale, les sondages obtenus en 2016 présentent une qualité nettement supérieure à ceux obtenus en 1996. Cette différence peut principalement s'expliquer par le matériel utilisé à l'époque (impossibilité de faire une remote réellement lointaine) et l'amélioration des procédures de traitements robustes. La différence observable sur les sondages de bonne qualité tels que Grang s'explique aussi par l'impossibilité de pouvoir reproduire les mesures aux mêmes points qu'en 1996. La topographie étant imposante dans la région, des effets de distorsion électriques locaux aux sites peuvent considérablement impacter les sondages MT même pour des sites distants de 500 mètres.

Nous avons également comparé les flèches d'induction entre les deux campagnes (Figure 5.12). Dans l'ensemble, les vecteurs sont très cohérents pour les périodes supérieures à la seconde. Pour les périodes inférieures à la seconde, des différences significatives sont observables entre les deux campagnes. Les stations situées au nord du profil sont cohérentes contrairement à celles situées au sud du profil. Comme pour les fonctions de réponse MT, ces différences peuvent s'expliquer par l'amélioration du matériel et des techniques de traitement statistiques des séries temporelles.

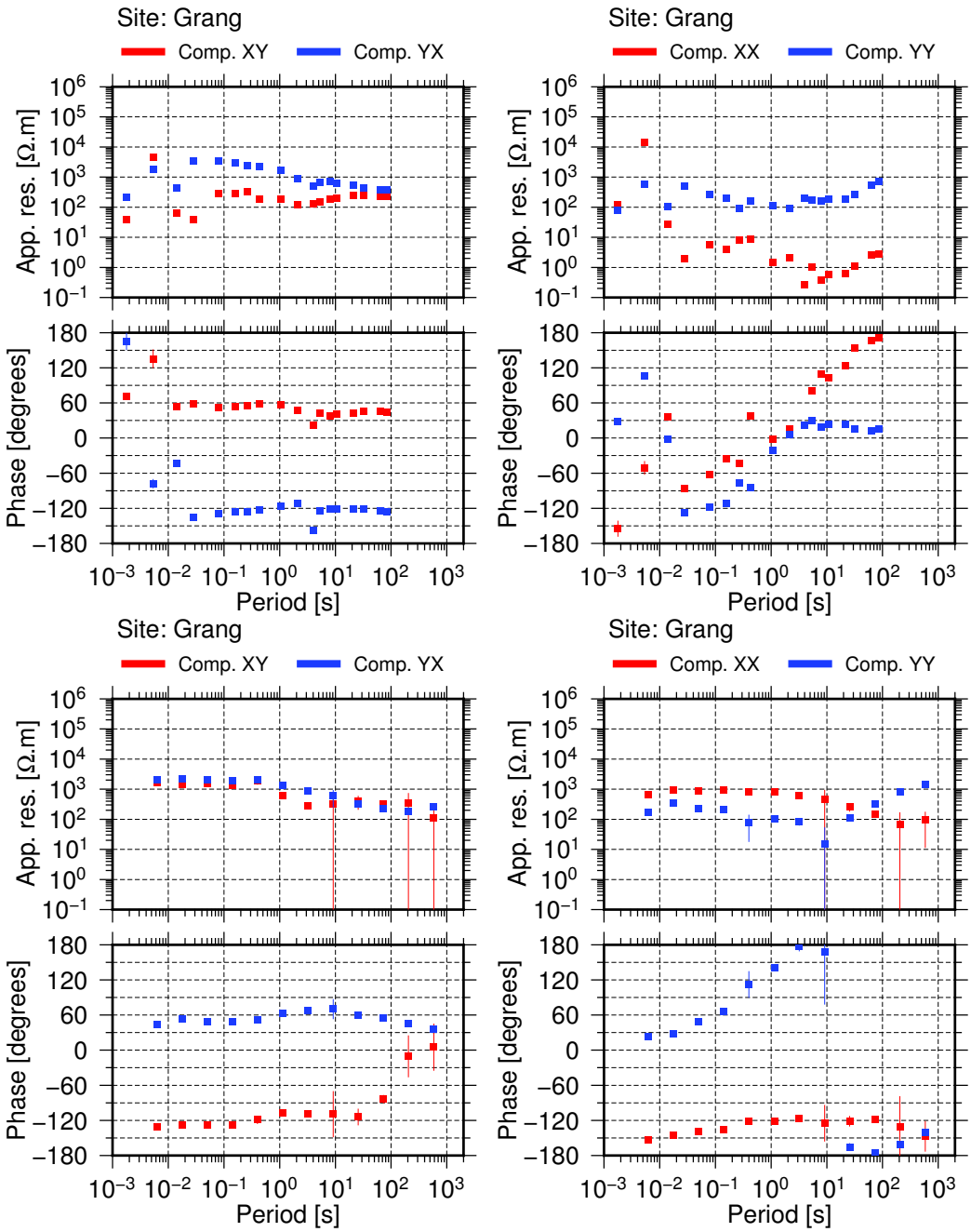


FIGURE 5.9: Comparaison entre les sondages obtenus en 1996 (haut) et 2016 (bas) pour la station de Grang.

FIGURE 5.9: *Comparison between 1996 MT soundings (top) and 2016 MT soundings (bottom) for Grang station.*

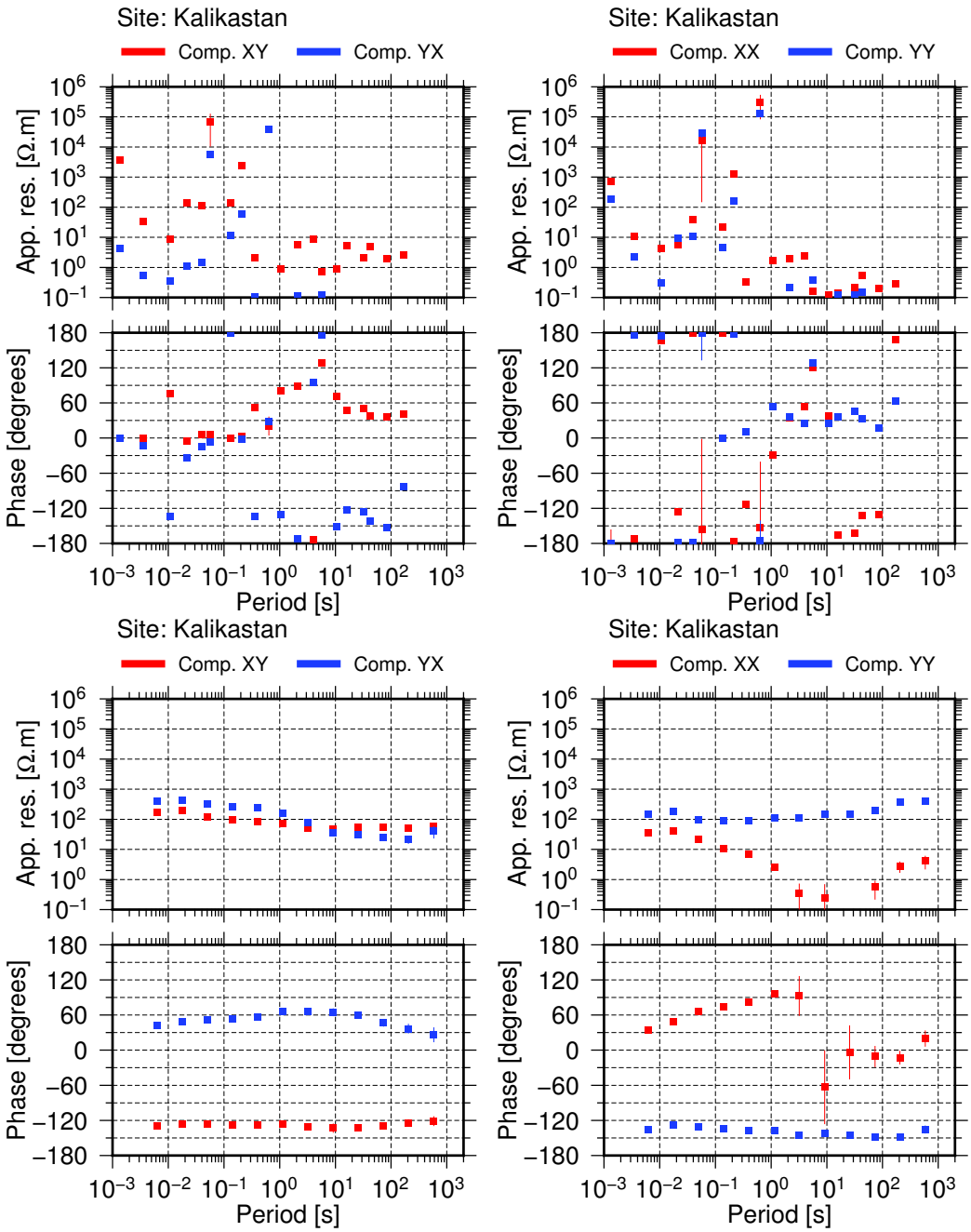


FIGURE 5.10: Comparaison entre les sondages obtenus en 1996 (haut) et 2016 (bas) pour la station de Kalikastan.

FIGURE 5.10: *Comparison between 1996 MT soundings (top) and 2016 MT soundings (bottom) for Kalikastan station.*

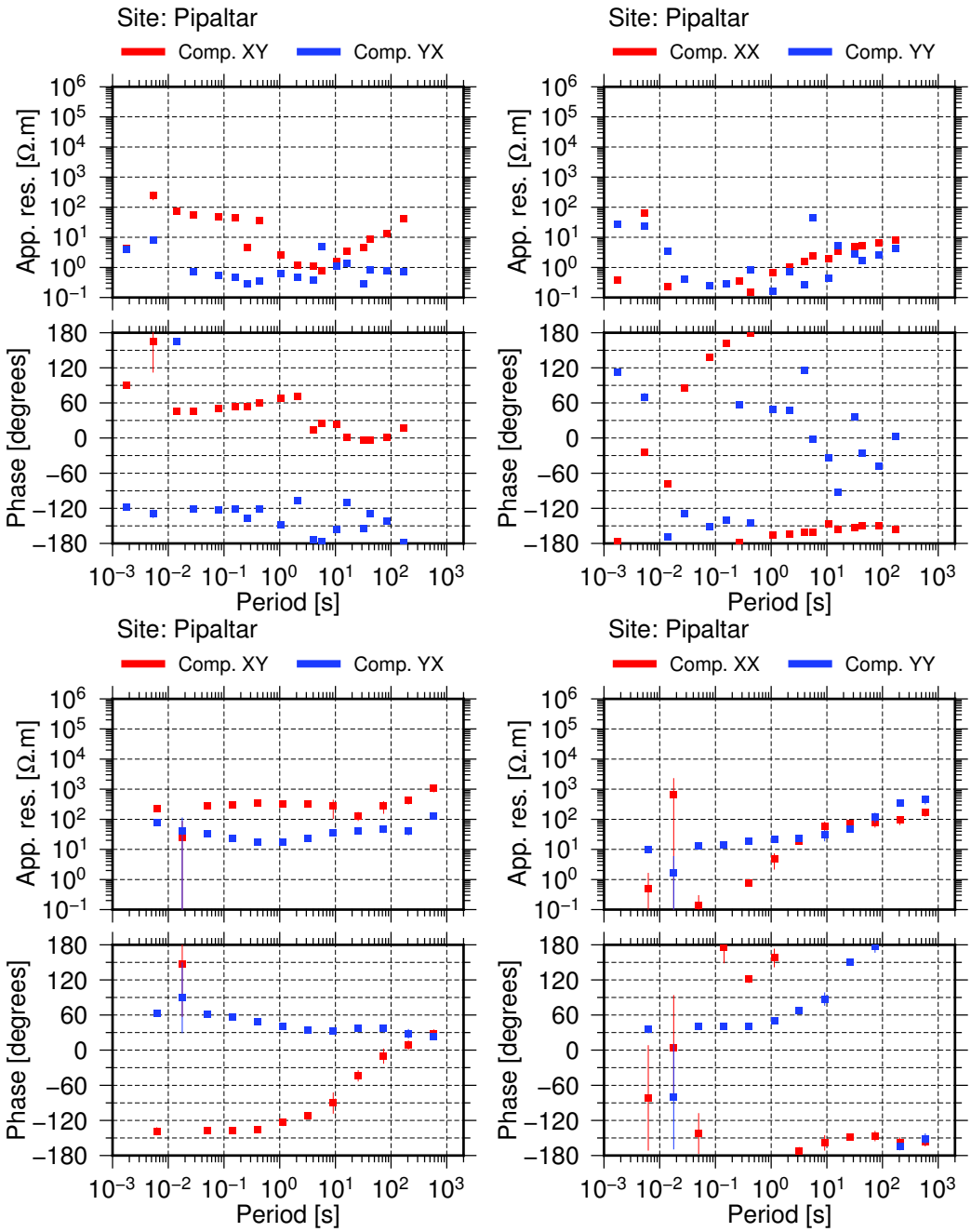


FIGURE 5.11: Comparaison entre les sondages obtenus en 1996 (haut) et 2016 (bas) pour la station de Pipaltar.

FIGURE 5.11: *Comparison between 1996 MT soundings (top) and 2016 MT soundings (bottom) for Pipaltar station.*

C. Lemonnier (2000) a conclu à la présence d'un conducteur profond au nord du profil par l'interprétation des flèches d'inductions. Ces conclusions semblent toujours valides par l'analyse des nouvelles données.

5.4.2 Distorsions du champ électrique et dimensionnalité

Nous avons étudié les distorsions du champ électrique et la dimensionnalité du milieu avec les deux outils majoritairement utilisés, le tenseur de phase (Caldwell, Bibby et Brown, 2004) et la décomposition de Groom-Bailey (Groom et Bailey, 1989) (voir Annexe D).

Les indicateurs de dimensionnalité du tenseur de phase sont indiqués à la Figure 5.13. Les indicateurs montrent une dimensionnalité 2D des tenseurs d'impédance aux périodes inférieures à la seconde, avec $\lambda > 0.1$ et β proche de zéro pour la majorité des stations. Aux périodes supérieures à la seconde, le paramètre β augmente pour l'ensemble de stations et le milieu peut être considéré comme 3D. Le strike $\alpha - \beta$ (Figure 5.14) du tenseur de phase n'indique pas une valeur cohérente entre tous les sites MT pour les périodes inférieures à la seconde, mais tend à se stabiliser entre 0 et 30 degrés pour les périodes supérieures à la seconde, ce qui est cohérent avec les différentes études géologiques menées dans la région et la direction de la collision Inde-Népal.

Comme indiqué par Chave & Jones (2012), les décompositions de Groom-Bailey sont les plus résistantes à une légère introduction de bruit aléatoire dans les données. Nous avons donc utilisé la méthodologie de Chave & Smith (1993) pour décomposer les données. Nous avons effectués plusieurs décompositions en testant des décompositions par fréquence et par station, multi-stations, et en imposant le strike à celui N20W, valeur déterminée par Carole Lemonnier et al. (1999).

Les tenseurs d'impédance régionaux pour les six stations sont représentés à l'annexe 5.9.2. Nous représentons à la Figure 5.15 les paramètres de distorsion pour les décompositions multi-sites et en imposant le strike à N20W. Les deux modèles présentent des paramètres de distorsion équivalents, ce qui permet d'avoir confiance dans le strike déterminé par la décomposition multi-site. Un autre paramètre important dans la décomposition Groom-Bailey est le test du χ^2 qui permet de déterminer si le modèle géologique sous-jacent peut être assimilé à celui d'un modèle régional 2D dont une hétérogénéité 3D provoque la distorsion du champ électrique.

Le test du χ^2 est présenté Figure 5.16. Le nombre de degrés de liberté est différent pour chaque décomposition. Pour une décomposition par fréquence et par site, 8 données sont disponibles (parties réelles et imaginaires du tenseur d'impédance), et 7 paramètres sont à déterminer (paramètres de twist, shear et strike, ainsi que les parties réelles et imaginaires du tenseur d'impédance régional antidiagonal). Ainsi, le nombre de degrés de liberté est 1. La *p-value* pour un

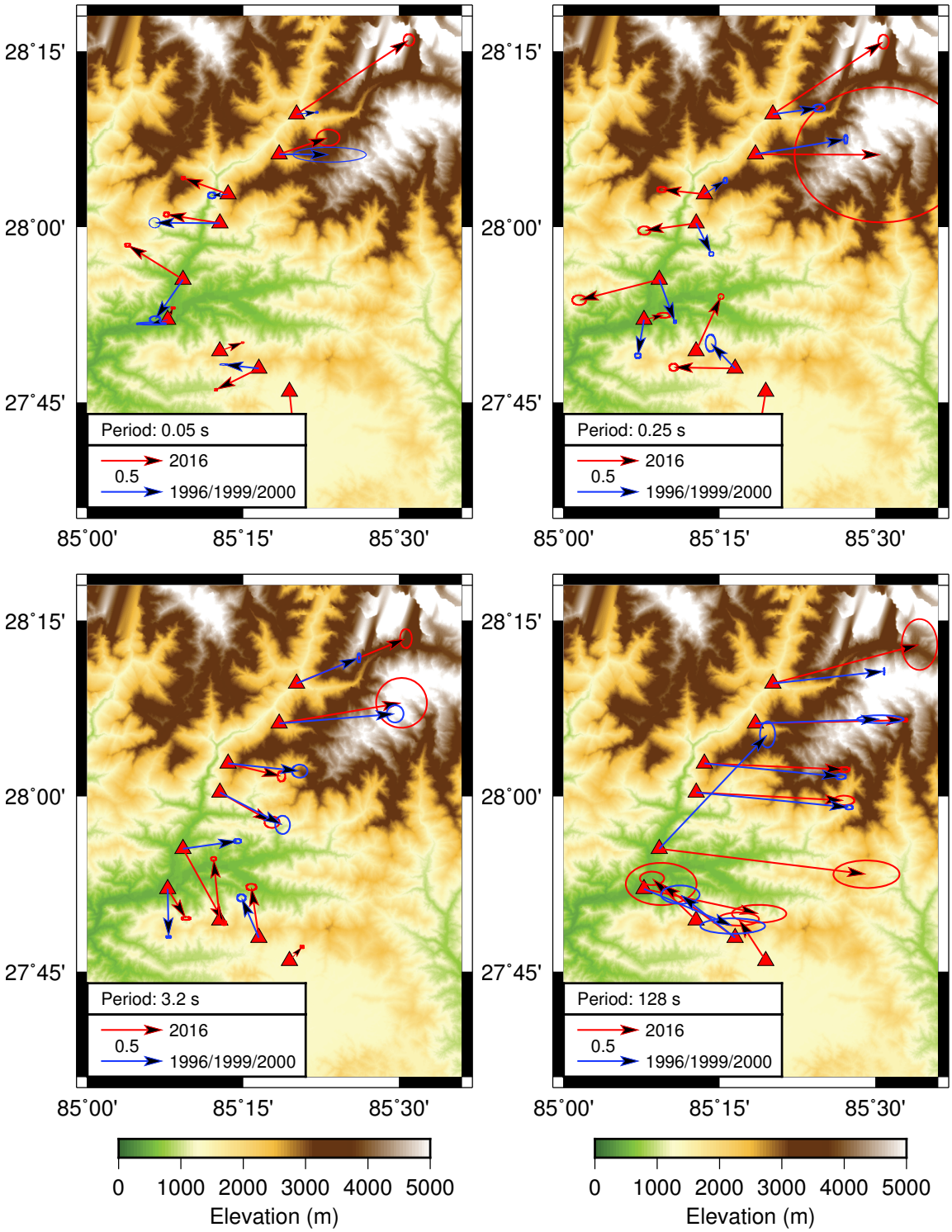


FIGURE 5.12: Comparaison des vecteurs d'induction obtenus aux campagnes de 1996 et 2016.
FIGURE 5.12: *Comparison of induction arrows between 1996 and 2016 MT campaigns.*

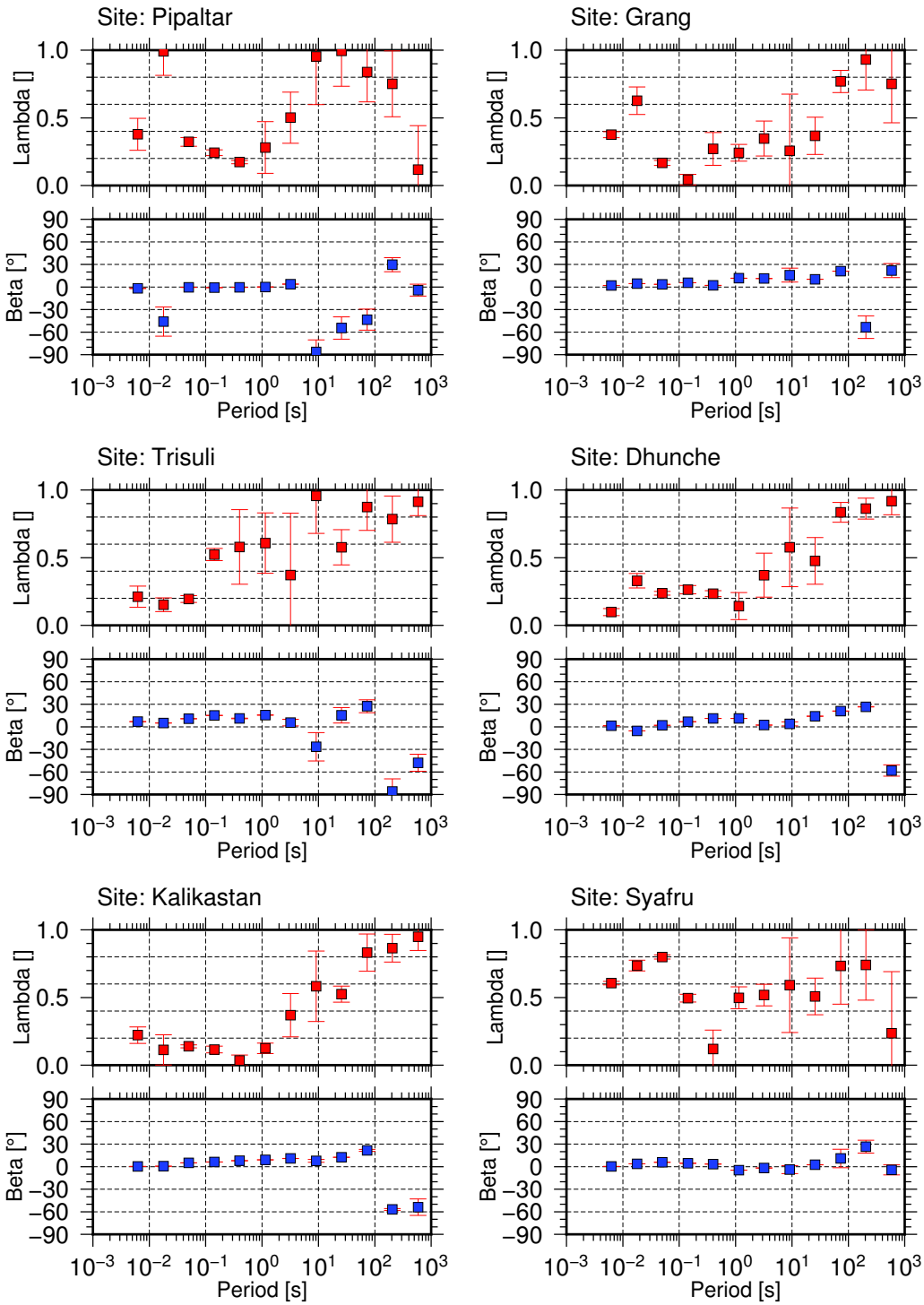


FIGURE 5.13: Indicateurs de dimensionnalité du tenseur de phase λ et β pour les 6 stations MT.

FIGURE 5.13: *Dimensionality indicators of phase tensor λ and β for the 6 MT stations.*

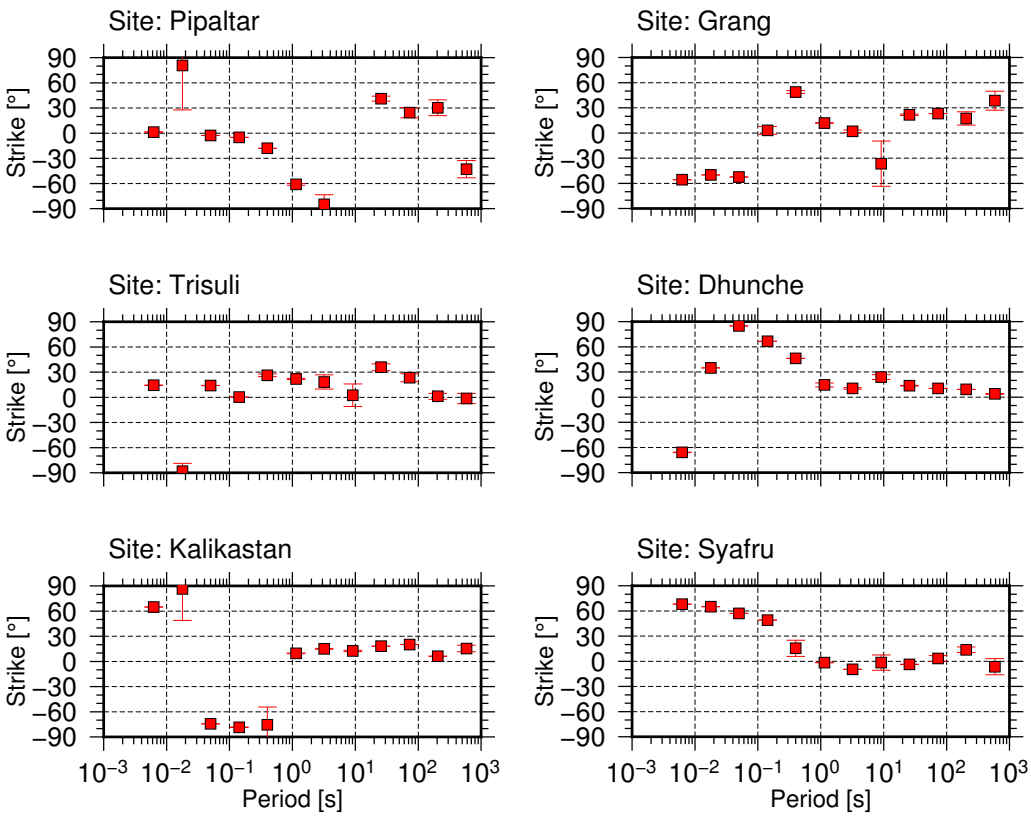


FIGURE 5.14: Strike du tenseur de phase $\alpha - \beta$ pour les 6 stations MT.
FIGURE 5.14: *Phase tensor strike $\alpha - \beta$ for the 6 MT stations.*

seuil de confiance de 95% à un degré de liberté est de 3.84. Pour une décomposition en fixant le strike à N10W, 8 données sont disponibles et 6 paramètres sont à déterminer à chaque fréquence (paramètres de twist, shear et tenseur d'impédance régional). Le nombre de degrés de liberté est donc 2. La *p-value* pour un seuil de confiance de 95% à deux degrés de liberté est de 5.99. Enfin, pour la décomposition multi-sites en contraignant les paramètres de twist, shear et strike à être constants sur l'ensemble de l'intervalle de périodes, 8 données sont toujours disponibles pour la détermination du modèle, mais seulement 4 paramètres sont à contraindre par fréquence, 2 par site (twist et shear), et un pour toutes les stations (strike). Le nombre de degrés de liberté est donc légèrement inférieur à 4. La *p-value* pour un seuil de confiance de 95% à quatre degré de liberté est de 11.07. Les trois décompositions dépassent fréquemment le seuil de confiance de 95 % et peuvent atteindre des valeurs extrêmement élevées pour l'ensemble des fréquences et des sites considérés.

Ainsi, l'hypothèse d'un milieu 2D régional est peu convaincante au regard de la décomposition de Groom-Bailey et du tenseur de phase. Cette propriété du milieu devra donc être prise en compte dans le cas d'une interprétation 2D des données.

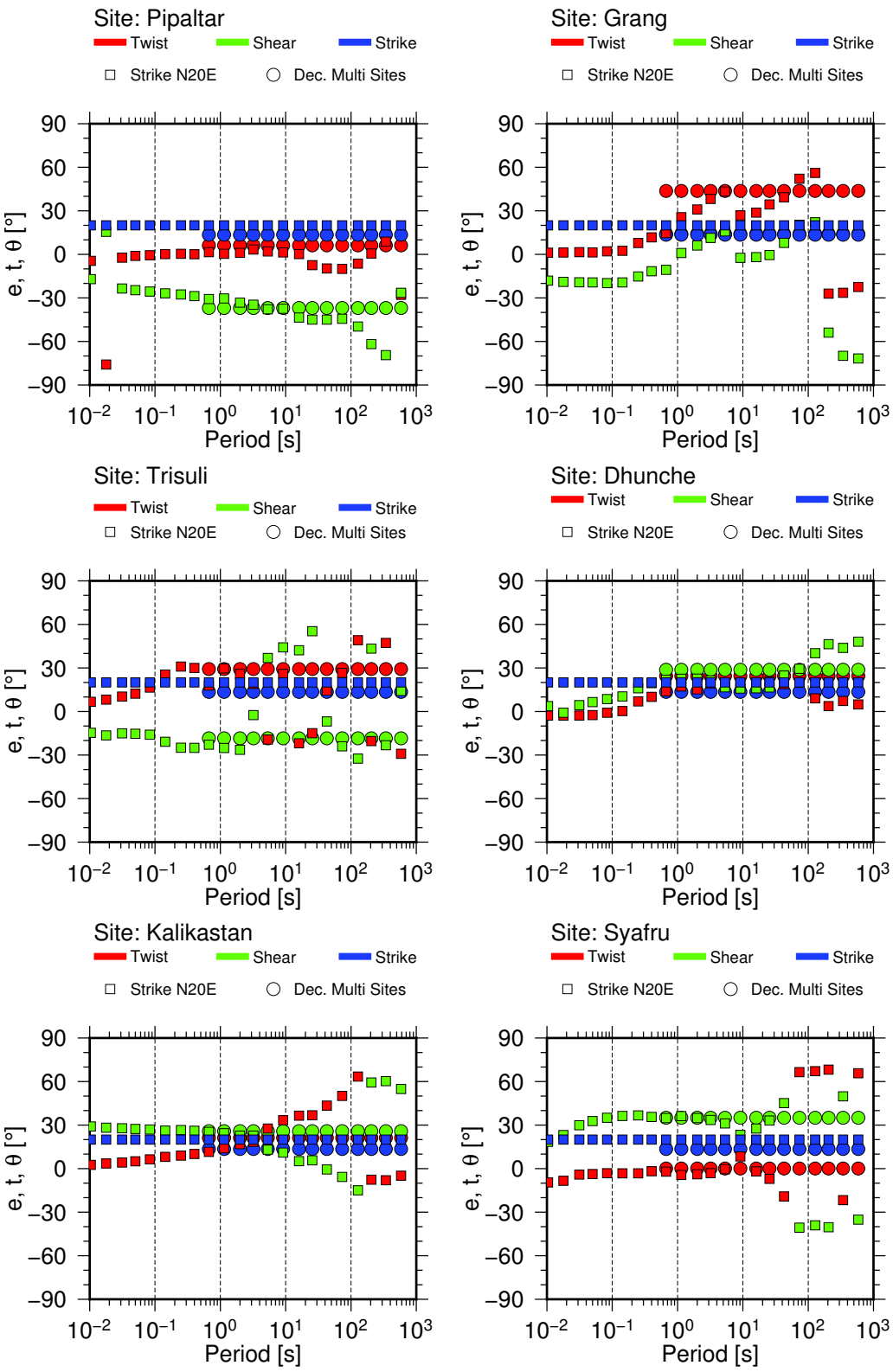


FIGURE 5.15: Comparaison des paramètres de distorsion pour une décomposition avec un strike de N20E et une décomposition multi-sites pour les 6 stations MT.

FIGURE 5.15: *Comparison of distortion parameters for single-site single frequency decomposition and multi-sites decomposition of the 6 MT stations.*

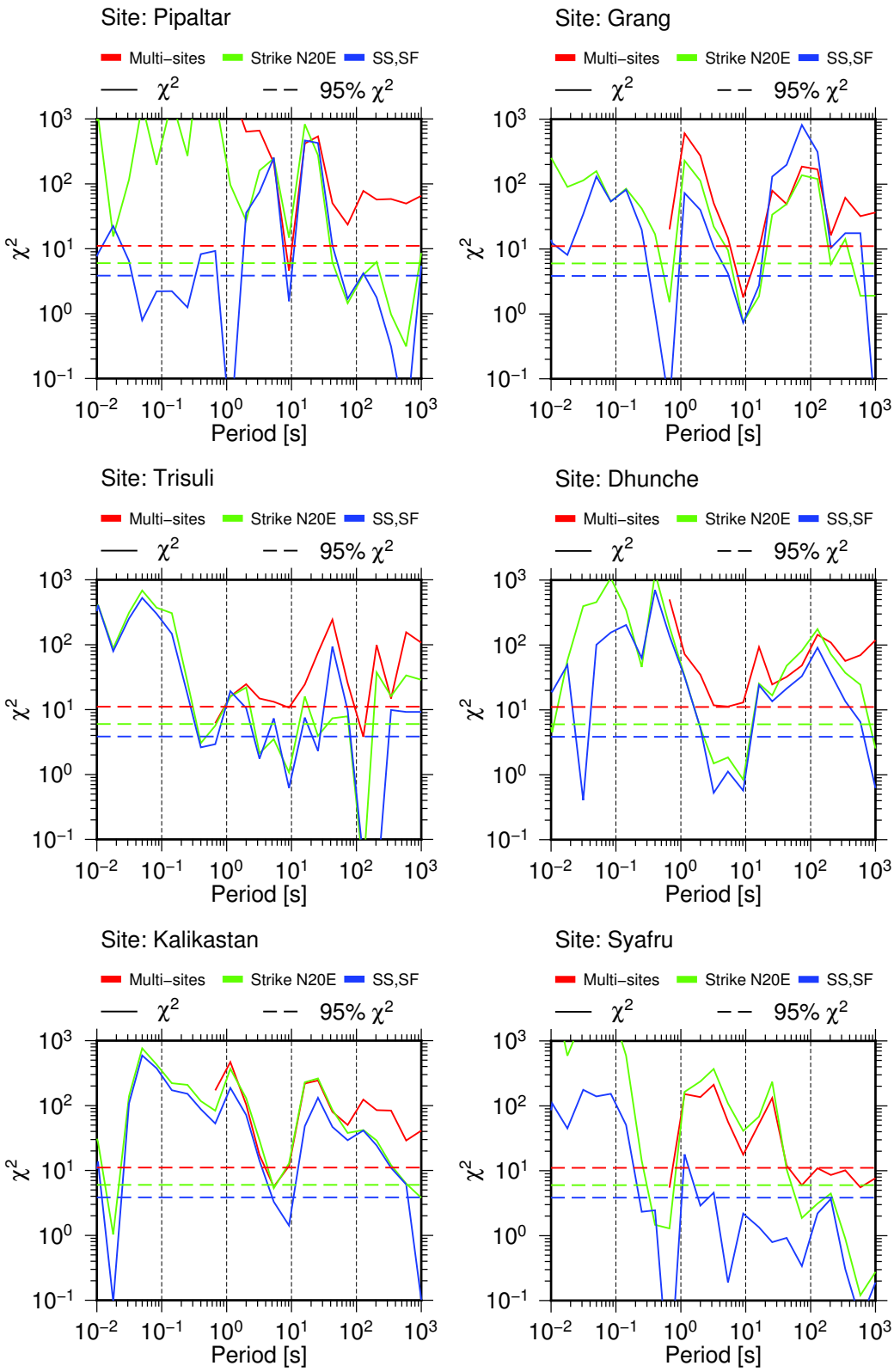


FIGURE 5.16: Test du χ^2 appliqué aux différentes décompositions. Multi-sites : Décomposition simultanée des sites du profil. Strike 10W : Décomposition en fixant le strike à N10W. SS,SF : Décomposition par site et par fréquence.

FIGURE 5.16: *χ^2 test applied to the different decompositions. Multi-sites : Simultaneous decomposition of all MT stations. Strike 10W : Decomposition by setting the strike at N10W. SS,SF : Single site, single frequency decomposition.*

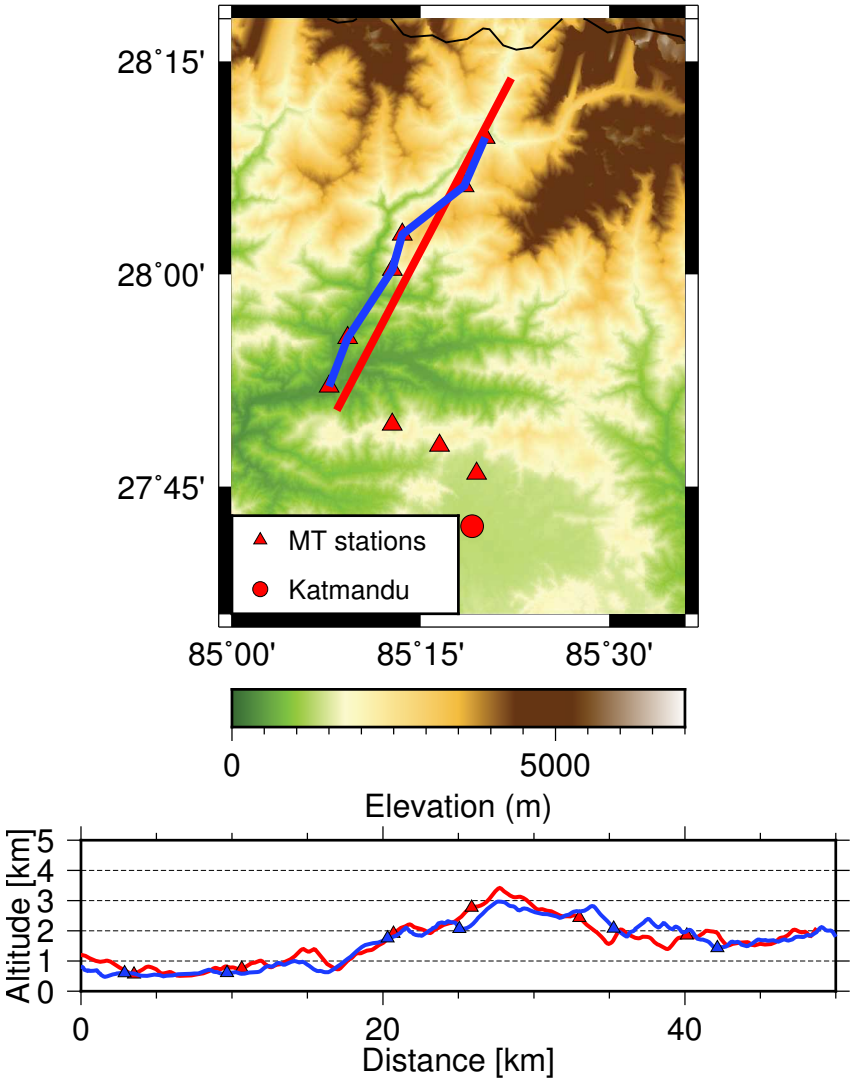


FIGURE 5.17: Haut : Profils 2D utilisés. Bas : Dénivelé le long des profils et position des stations après projection orthogonale.
FIGURE 5.17: *Top : 2D profil used in the modeling. Bottom : Topography along profil with MT stations positions after orthogonal projection.*

5.5 Modélisation des effets topographiques

5.5.1 Modélisation et inversion d'effets 2D

Pour évaluer l'impact de la topographie sur les sondages magnétotelluriques obtenus, nous avons effectué une modélisation 2D d'un milieu homogène en considérant la topographie observée dans la vallée de Trisuli. Deux profils ont été considérés, l'un droit entre les points de coordonnées (85.1389, 27.84) et (85.37, 28.23). Les stations ont ensuite été projetées orthogonalement sur le profil. La difficulté majeure d'utiliser une telle méthode réside dans la forte variabilité de topographie. En effet, la station de Grang a été acquise sur les flancs d'un massif culminant à

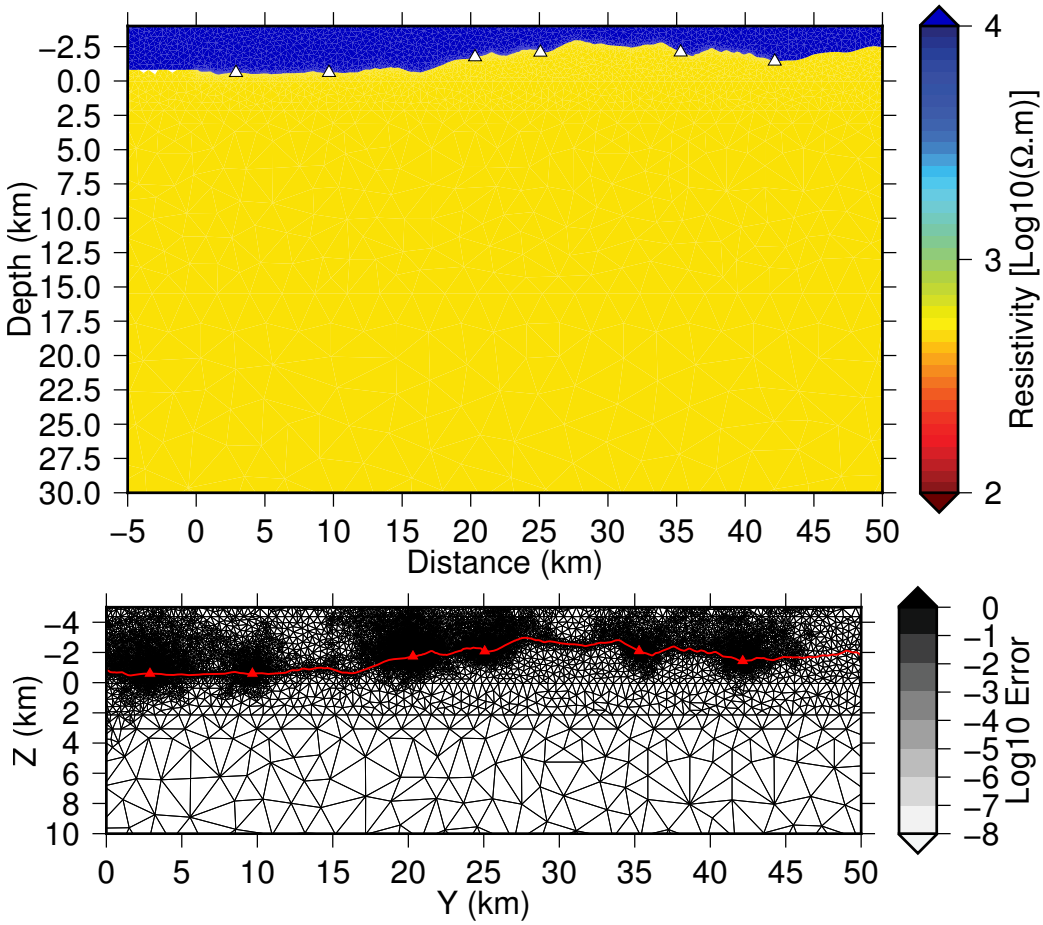


FIGURE 5.18: Haut : Maillage du modèle en éléments finis pour le calcul de l'effet topographique. Bas : Maillage final pour le calcul du modèle direct.
FIGURE 5.18: *Top : Finite element mesh for topography effect computation. Bottom : Final mesh used in forward modeling.*

environ 3000 mètres croisé par le profil. La projection orthogonale attribue donc la station Grang à une altitude bien supérieure à son altitude réelle (profil rouge sur la Figure 5.17). Pour éviter ce problème, nous avons effectué un profil suivant approximativement la route utilisée qui suit les vallées. De légers décalages sont observables sur les stations, mais les altitudes sont justes par rapport à celles mesurées. Les altitudes du profil sont calculées à partir des données de la mission *Shuttle Radar Topography Mission (SRTM)* (Farr et al., 2007).

Carole Lemonnier et al. (1999) ont étudié l'influence de la topographie sur les flèches d'induction avant l'inversion les sondages MT. Ils ont montré, par modélisation 3D utilisant l'algorithme de Mackie, J. T. Smith et T. R. Madden (1994), que la topographie avait un effet négligeable sur ces fonctions de transfert à basses fréquences. La topographie n'a donc pas été incluse dans l'inversion des fonctions de réponse MT. Or, il est connu depuis longtemps (Jiracek, 1990) que la topographie peut créer des distorsions sur le champ électrique. L'inversion sans topographie peut ainsi induire des artefacts lors de l'inversion de fonctions de réponse MT.

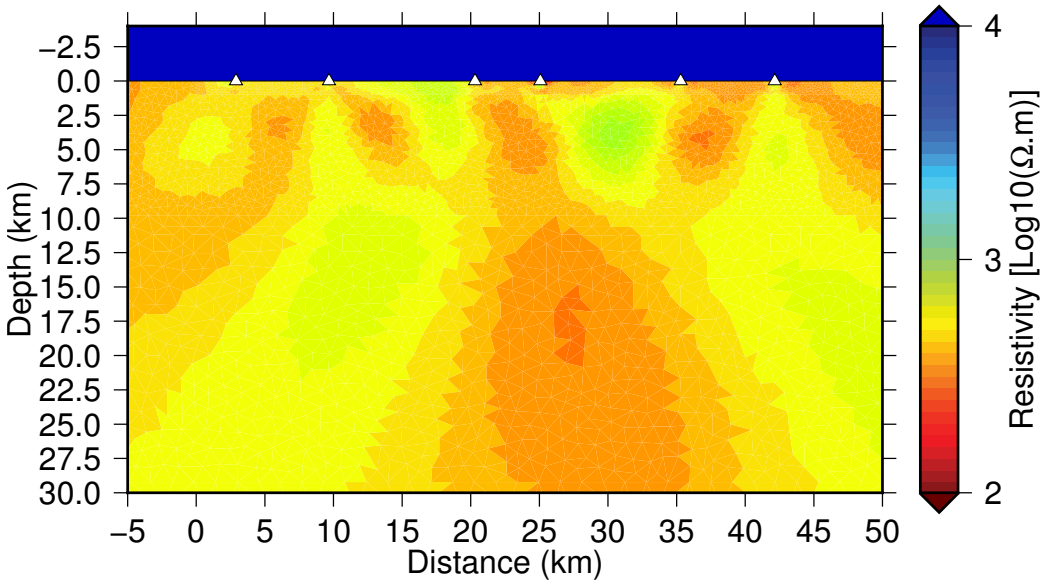


FIGURE 5.19: Modèle obtenu après inversion sans topographie des données modélisées avec topographie.
FIGURE 5.19: *Inversion without topography of synthetic soundings from model with topography.*

Pour déterminer quelle influence cette hypothèse a pu avoir sur l'inversion effectuée en 1999, nous avons modélisé les fonctions de réponse MT pour un milieu homogène avec topographie (Figure 5.19). Pour modéliser le milieu 2D avec topographie, nous avons utilisé le code MARE2DEM (Key, 2016). La topographie et les positions des stations MT sont celles indiquées Figure 5.17.

Le maillage utilisé pour la modélisation directe est présenté Figure 5.18 avec le maillage final utilisé dans le calcul direct. Le modèle est celui d'un milieu homogène à 500 Ω.m, valeur moyenne obtenue lors de l'inversion par Carole Lemonnier et al. (1999). Il s'étend à -200 km au sud du profil, 300 km au nord, -100 km dans l'air, 300 km en profondeur et contient 16 167 régions. La résistivité de l'air est fixée à 10^{13} $\Omega.m$. L'avantage du code Mare2DEM réside dans la forme triangulaire des mailles utilisées, permettant une description fine de la topographie. Le maillage final pour le calcul direct contient 66 742 régions. L'algorithme raffine principalement les mailles situées au niveau des stations et le long de l'interface sol-air pour accommoder le contraste de résistivité.

Ces fonctions ont ensuite été inversées pour un modèle ne comportant pas de topographie. 9 itérations ont été nécessaires pour la convergence du modèle. Le *misfit* final est égal à 1.001. Le modèle issu de l'inversion est présenté Figure 5.19. La modélisation des sondages par le modèle issu de l'inversion est présentée Figure 5.20. Ceux-ci indiquent une forte distorsion du champ électrique, même à basses fréquences ce qui contredit l'assertion utilisée dans la thèse de C. Lemonnier (effet négligeable de la topographie). L'effet le plus présent dans les sondages est une diminution de la résistivité apparente par rapport à la résistivité du milieu homogène. Lors

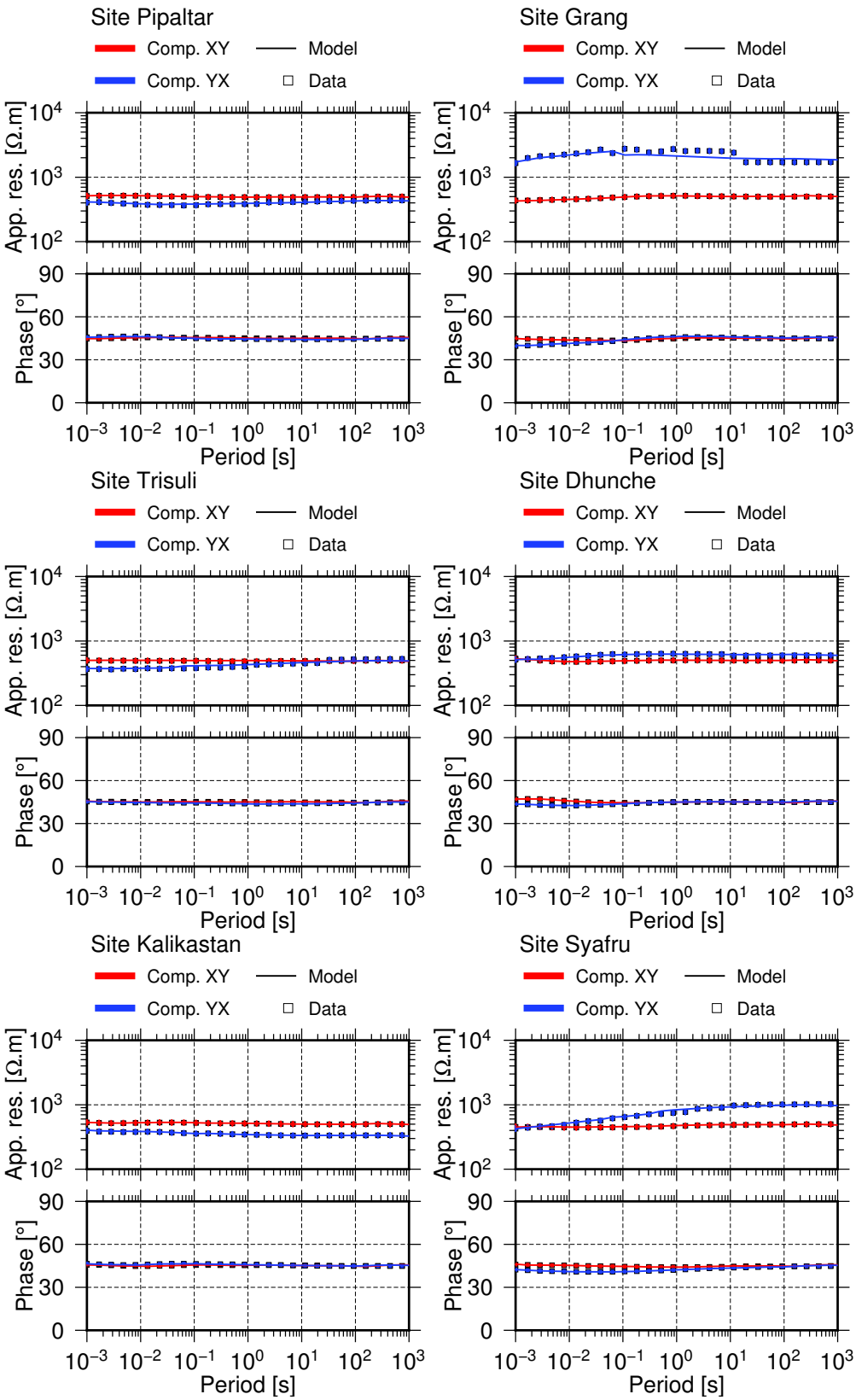


FIGURE 5.20: Comparaison entre données et résultat de l'inversion sans topographie.
FIGURE 5.20: *Comparison between data and inversion result without topography included.*

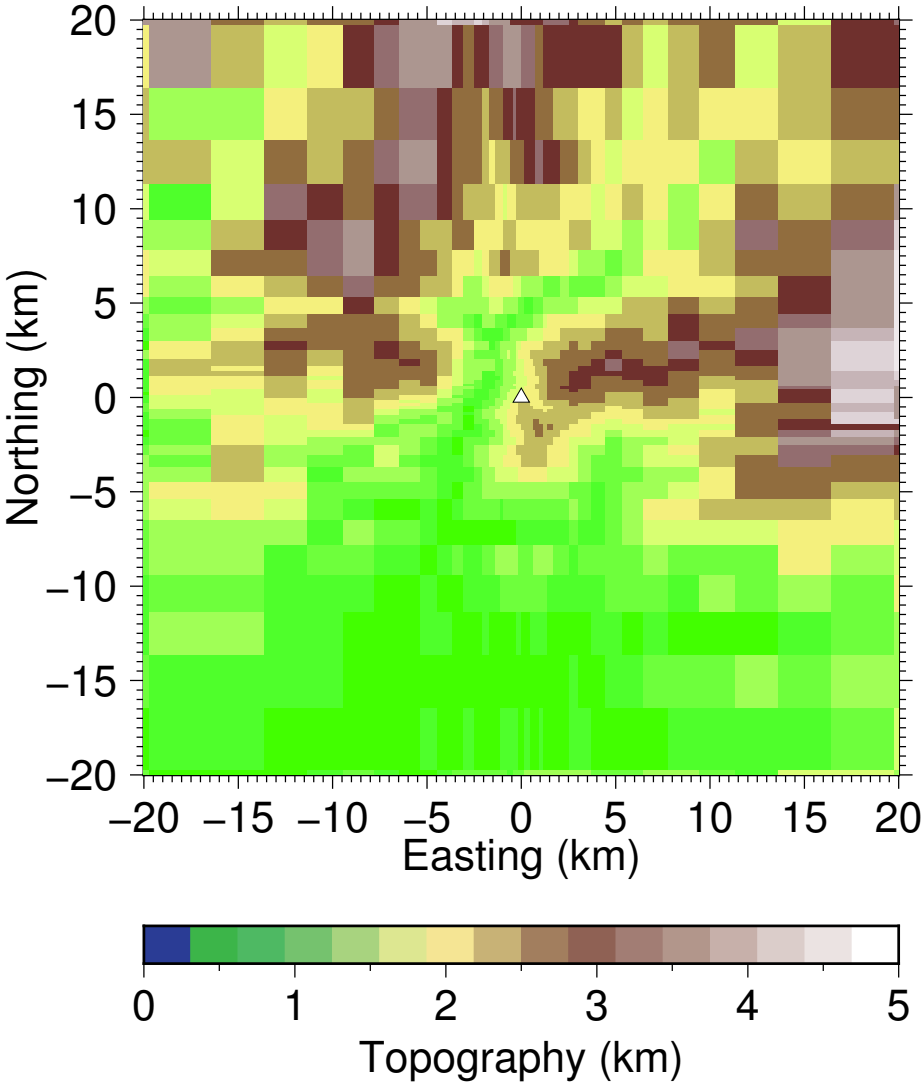


FIGURE 5.21: Modèle de topographie 3D pour le site de Grang
FIGURE 5.21: *3D topography model for Grang MT site..*

du processus d'inversion sans topographie, ceci force l'algorithme d'inversion à introduire des corps géologiques non réalistes et notamment un conducteur sous le maximum de topographie. Mis en relation avec le conducteur déterminé par Carole LEMONNIER et al. (1999), ceci pose la question de la validité du modèle de conductivité trouvé. En effet, un corps conducteur artificiel sous l'hypothèse de l'effet négligeable de la topographie a pu être inventé par l'inversion. Moins critique et déjà contesté par Carole LEMONNIER et al. (1999), le conducteur trouvé sous le site de Naubise, site situé au sud du profil actuel où la topographie redevient importante, pourrait aussi venir de la topographie négligée.

5.5.2 Modélisation 3D

Pour tester les éventuels effets de la topographie en 3D sur les tenseurs d'impédance, nous avons créé des modèles 3D homogènes pour chaque station. La topographie étant accidentée dans la région, nous avons optimisé les calculs en définissant un maillage par station. Chaque maillage utilisé est composé de 72 × 72 × 120 cellules. La taille totale du modèle est de 907.84 x 907.84 x 856.486 km. La topographie est interpolée à partir du modèle SRTM (Farr et al., 2007). Autour du site, les cellules horizontales mesurent 150 mètres et 50 mètres en vertical. La Figure 5.21 montre le modèle de topographie pour le site de Grang.

Le résultat de la modélisation 3D pour les sites de Kalikastan, Dhunche, Grang et Syafru sont présentés à la Figure 5.22. La présence d'un effet topographique à toutes les fréquences est aussi observable comme dans le cas 2D, aussi bien sur la phase que la résistivité apparente. Une décroissance de la résistivité apparente est notamment observée sur les composantes anti-diagonales pour les sites de Kalikastan et Grang, purement liée à la topographie et non à la présence d'un conducteur en profondeur. Sur certains sites, les composantes diagonales ne sont inférieures que d'un seul ordre de grandeur par rapport aux composantes anti-diagonales, renforçant l'hypothèse d'une dimensionnalité 3D au niveau des sites de mesures.

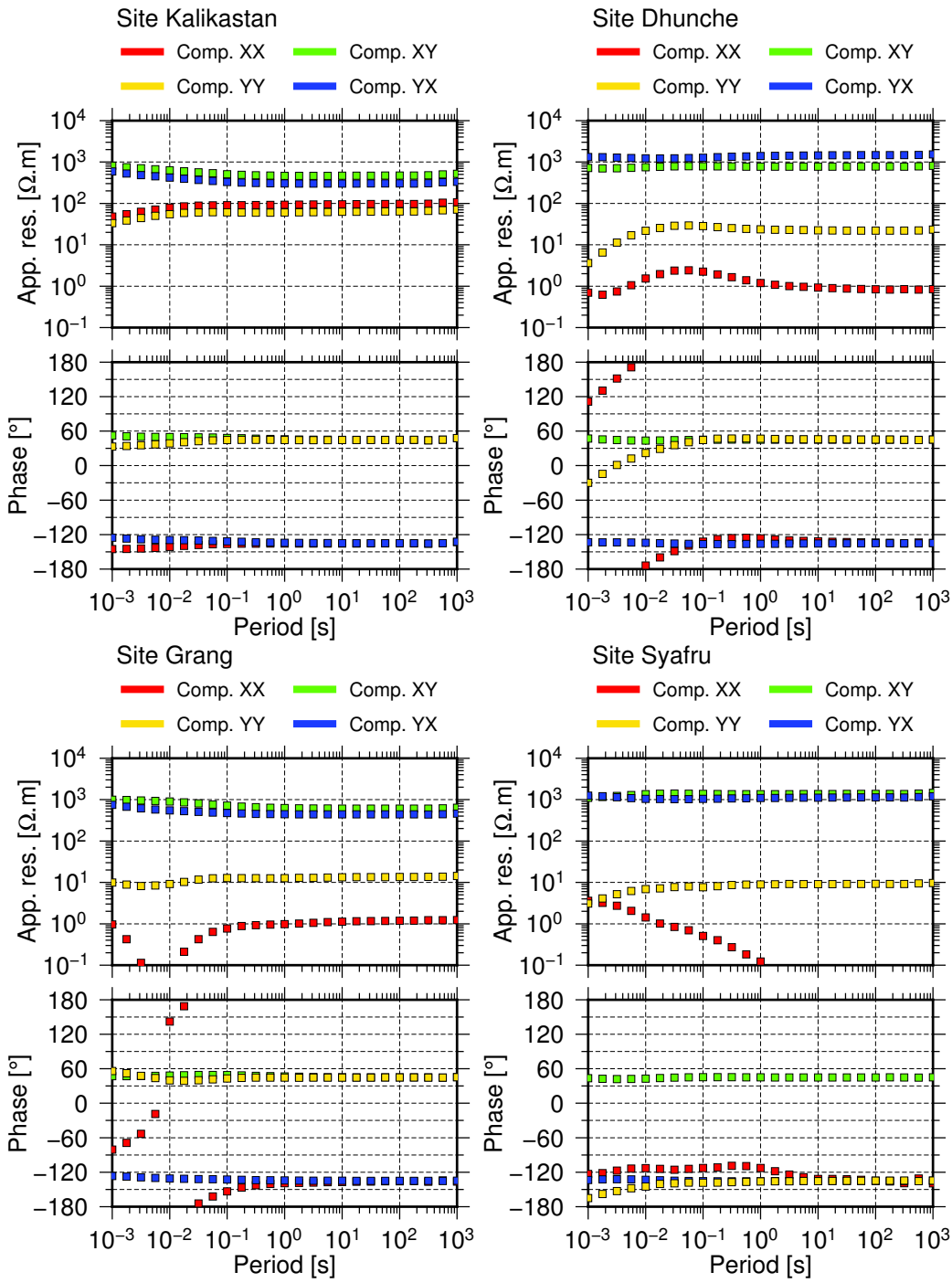


FIGURE 5.22: Sondages 3D pour un modèle homogène de terre à 1000 $\Omega.m$.
FIGURE 5.22: *3D MT soundings for an homogenous earth with a resistivity of 1000* $\Omega.m$.

5.6 Inversion du profil 2D

5.6.1 Présentation de l'ancien profil

Nous avons récupéré le modèle et les données issus du traitement et de l'inversion de 1999. Le modèle de conductivité non interpolé est présenté Figure 5.23. La comparaison entre les données réelles et modélisées est présentée aux Figures 5.24 et 5.25. La modélisation des sondages MT pour les sites situés au sud du profil (Simra, Pantale) est en accord avec les données mesurées. La modélisation des sites de Daman et Kakani montre que la tendance des données est correcte, mais ces deux stations présentent un static shift non modélisé. Au nord du profil, les sites de Pipaltar, Dhunche et Ramche présentent aussi des écarts de type static shift entre la modélisation et les données. Enfin la modélisation du site de Syafru n'est pas en accord avec les données mesurées.

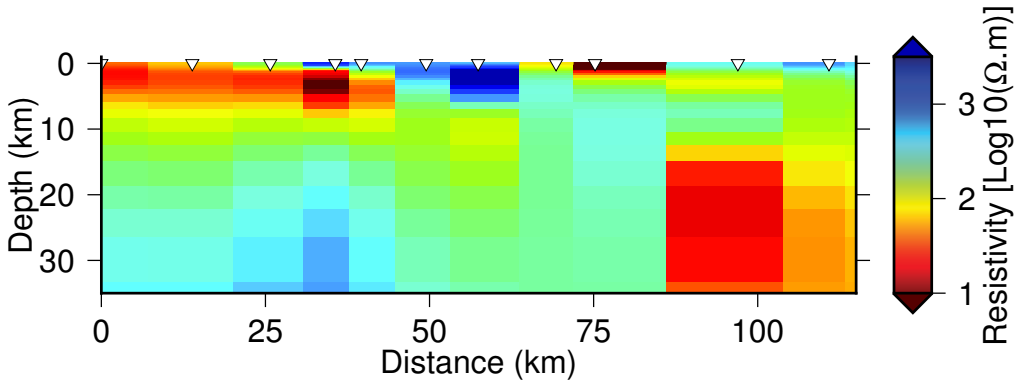


FIGURE 5.23: Modèle de Carole LEMONNIER et al. (1999) non interpolé.
FIGURE 5.23: *Final resistivity model from Carole* LEMONNIER *et al. (1999) without interpolation.*

Les sites où la modélisation de la distribution de conductivité inversée présente le plus d'écart avec les données mesurées sont ceux où la topographie présente le plus de variation (Kakani, Ramche, Dhunche, Syafru). A l'inverse, les données proches de la plaine du Gange (Simra, Pantale) où la topographie ne présente pas de fortes variations sont bien modélisés. Étant donné l'effet topographique attendu au nord du profil, il apparaît nécessaire d'inverser les données en prenant en compte la topographie pour minimiser les risques d'introduire de fausses structures dans le processus d'inversion.

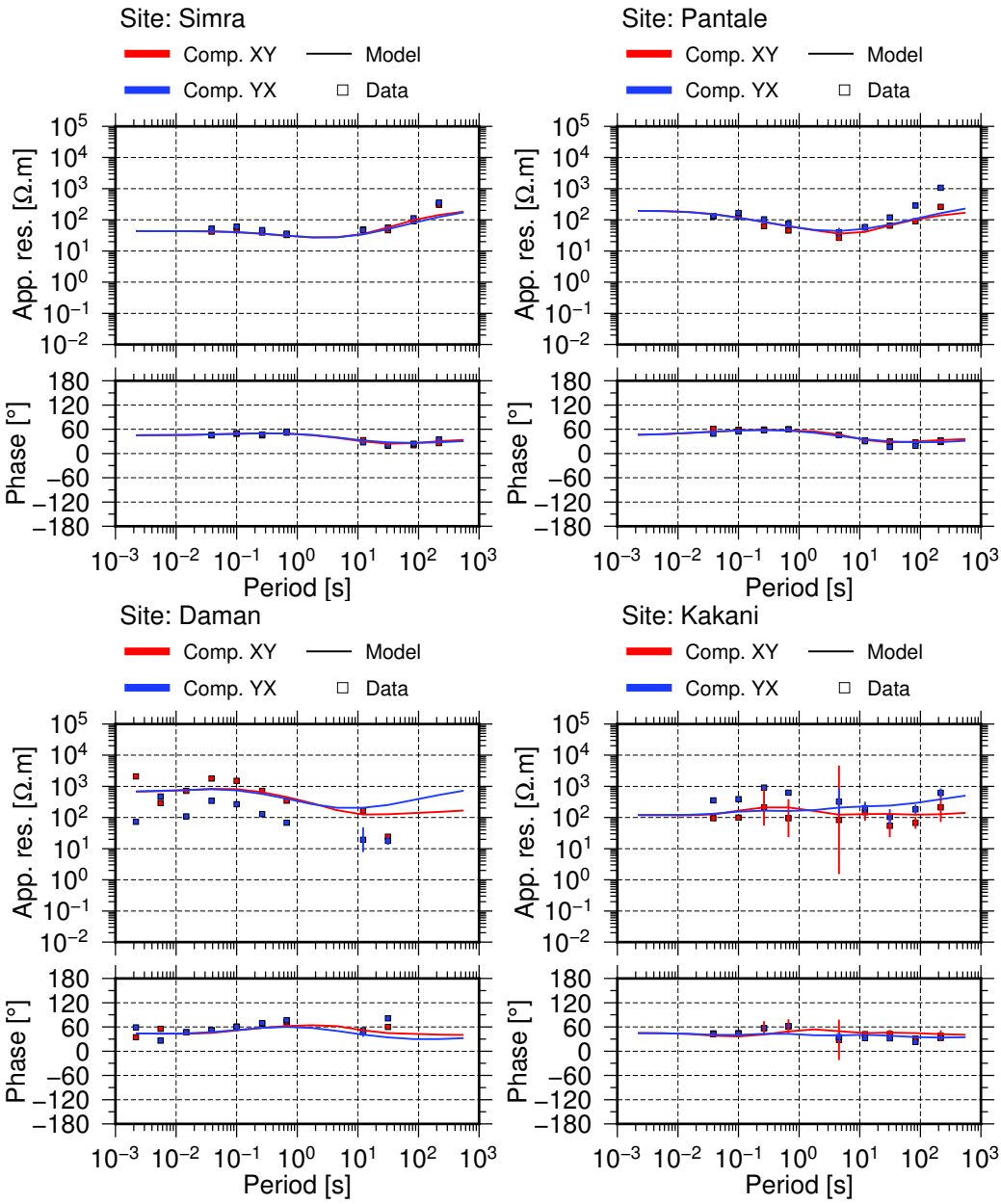


FIGURE 5.24: Données réelles et modélisées pour les stations de Simra, Pantale, Daman et Kakani d'après l'inversion de 1999.

FIGURE 5.24: *Real data and modeling of 1999 conductivity model for Simra, Pantale, Daman and Kakani MT stations.*

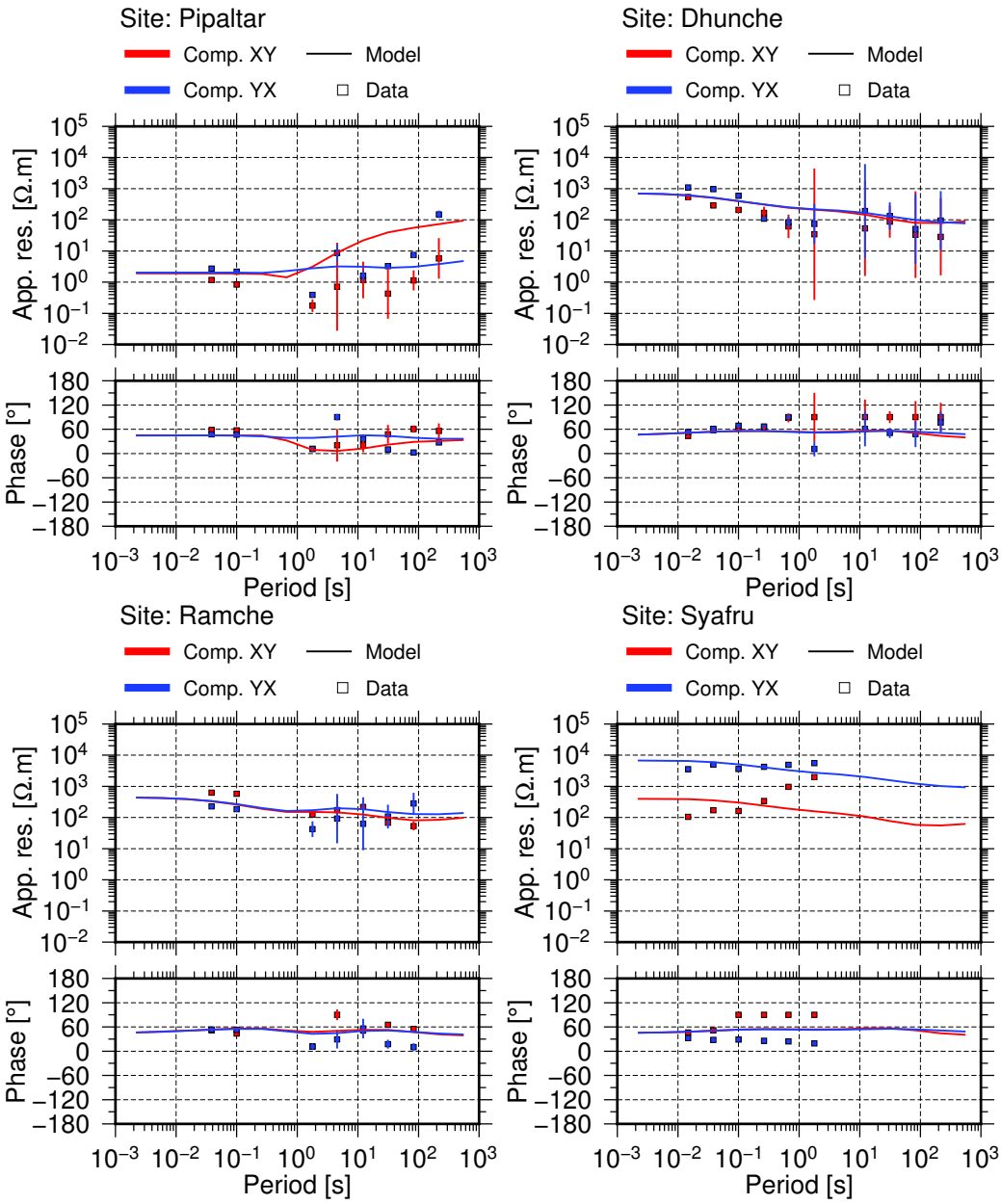


FIGURE 5.25: Données réelles et modélisées pour les stations de Pipaltar, Dhunche, Ramche, Syafru d'après l'inversion de 1999.

FIGURE 5.25: *Real data and modeling of 1999 conductivity model for Pipaltar, Dhunche, Ramche and Syafru MT stations.*

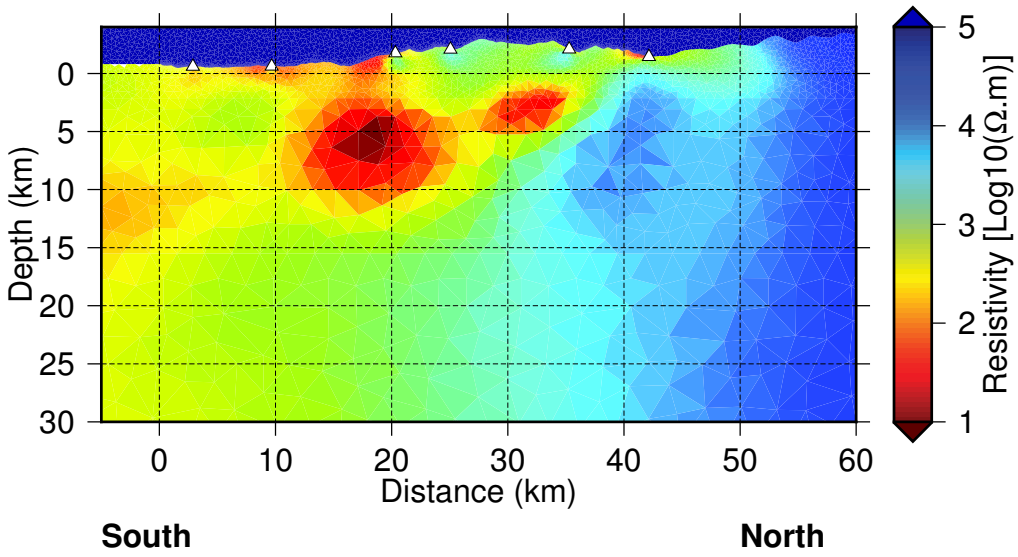


FIGURE 5.26: Modèle de résistivité issue de l'inversion à partir des données de 2016.
FIGURE 5.26: *Resistivity modeling from 2016 dataset.*

5.6.2 Inversion des données de la campagne 2016

Pour interpréter les données 2D, nous avons décomposé les données en utilisant la décomposition de Groom-Bailey en fixant le strike à N20W. Le maillage utilisé pour l'inversion contient 16167 mailles, dont 9470 utilisées pour l'inversion du profil. Le reste des cellules est dédiée à la modélisation de l'air dont la résistivité est fixée à $10^{13}\Omega.m$. Les erreurs attribuées au données ont été attribués vont de 5 % pour celles de meilleures qualités, jusqu'à 20% pour les valeurs de fonctions de réponse MT les moins fiables.

9 itérations ont été nécessaires pour que l'algorithme converge vers une valeur de *misfit* stable. Le *misfit* atteint est de 11.7. Le *misfit* est calculé en utilisant une norme L_2. Cette norme est peu robuste, et sa valeur augmente rapidement si une petite fraction de données modélisées diffère des données observées. La Figure 5.26 montre le modèle issu de l'inversion des données MT de 2016. Le Figure 5.27 montre l'adéquation entre les données à inverser et les sondages modélisés.

Dans l'ensemble, la modélisation de la distribution de conductivité inversée permet de correctement reproduire les sondages MT. Des différences notables sont observables. L'augmentation rapide de la phase de la composante xy de la station Pipaltar n'est pas correctement reproduite. La composante TE observée à Syafru n'est elle aussi pas fidèlement décrite par la distribution de conductivité inversée. En revanche, les tendances pour le reste des stations sont adéquatement reproduites par le modèle inversé.

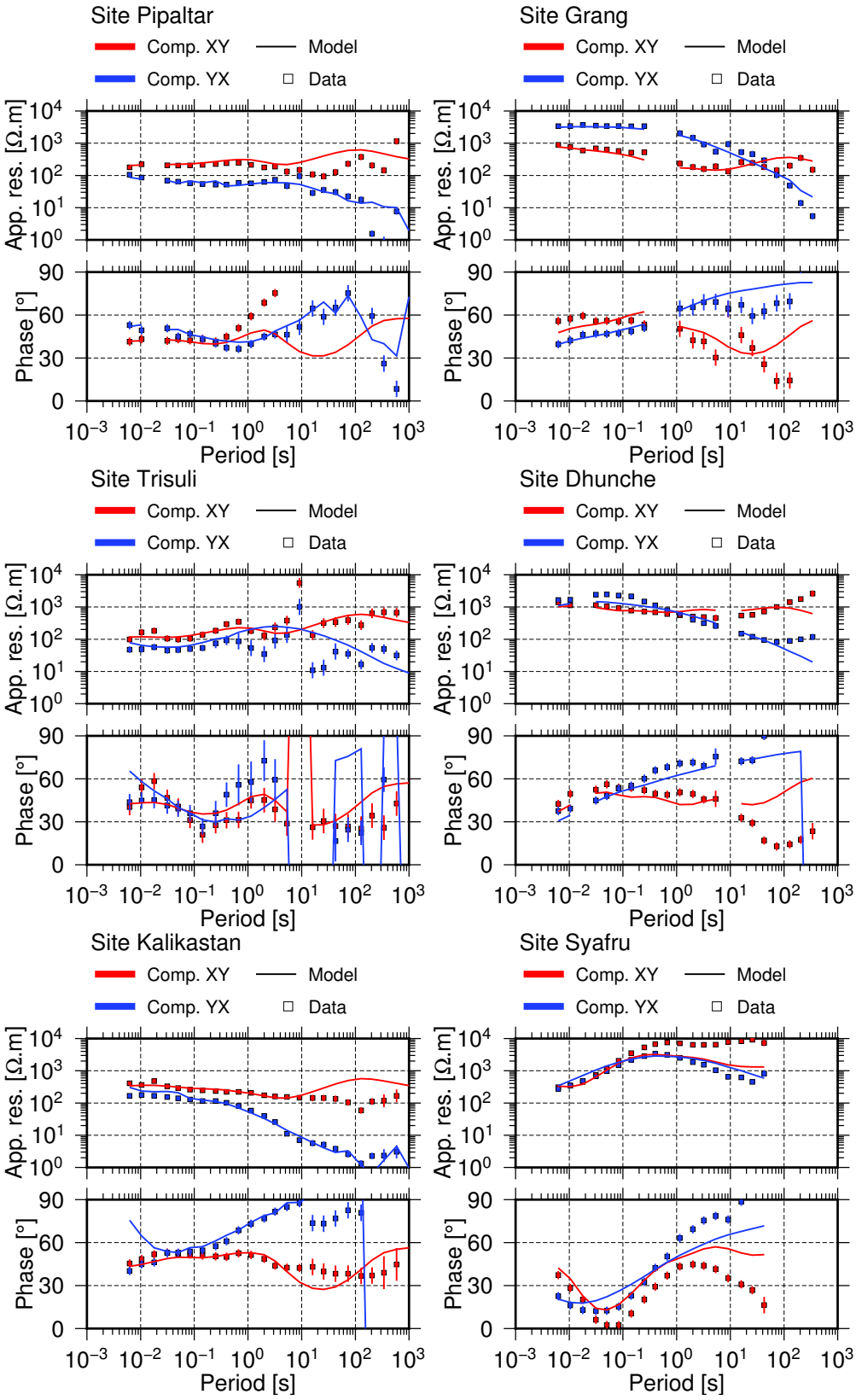


FIGURE 5.27: Comparaison entre données et résultat de l'inversion à partir des données de 2016.

FIGURE 5.27: *Comparison between data and inversion results from 2016 dataset.*

5.6.3 Tests de sensibilité

5.6.3.1 Aux stations

Une méthode permettant de déterminer si des structures de conductivité inversées sont robustes ou non, est de tester la sensibilité du modèle inversé après une modification des données. Il est ainsi possible de changer l'échantillonnage en fréquence des sondages MT, de changer l'échantillonnage spatial des stations inversées si le nombre de stations disponibles est suffisant, de retirer certaines stations... Avec le peu de stations dont nous disposons sur le profil, il apparaît compliqué de retirer plus d'une station par test.

Nous avons donc inversé les données en retirant à chaque fois une station sur le profil pour les stations de Syafru, Dhunche, Grang, Kalikastan et Trisuli. Les modèles après convergence de l'inversion sont présentés aux Figures 5.28 et 5.29. Les *misfits*, entre 10.9 et 12.4, sont comparables à celui de l'inversion utilisant toutes les stations. La comparaison des sondages modélisés à partir des distributions de conductivité inversées avec les données est présentée Figures 5.30 et 5.31.

Tous les tests de sensibilité présentent des sondages MT modélisés cohérents avec les sondages inversés. La caractéristique majeure observée sur toutes les distributions de résistivités est la séparation du profil en deux grandes structures, résistant au nord et conducteur au sud. Le corps conducteur situé entre 15 et 20 km le long du profil est aussi constamment retrouvé sur tous les modèles inversés, en mettant une nuance dans le cas de l'inversion sans le site de Kalikastan où le corps conducteur est moins défini. Les structures conductrices proches surfaces au niveau des sites de Pipaltar, Trisuli et Kalikastan sont aussi retrouvées sur les différents modèles.

Le corps conducteur situé sous la station de Dhunche, entre 0 et 5 km de profondeur apparaît contrôlé par l'inversion du site de Dhunche. Le modèle inversé sans Dhunche présente un sondage modélisé à Dhunche complètement différent de celui observé. L'ensemble des autres modèles présente ce corps conducteur et reproduit les données de Dhunche.

La structure conductrice au nord du site de Syafru apparaît contrôlée par ce site. Etant donné le manque de sensibilité en dehors du profil il est cependant délicat de juger de son existence réelle ou non. La connectivité des corps conducteurs au milieu du profil n'est pas clair sur l'ensemble des modèles inversés.

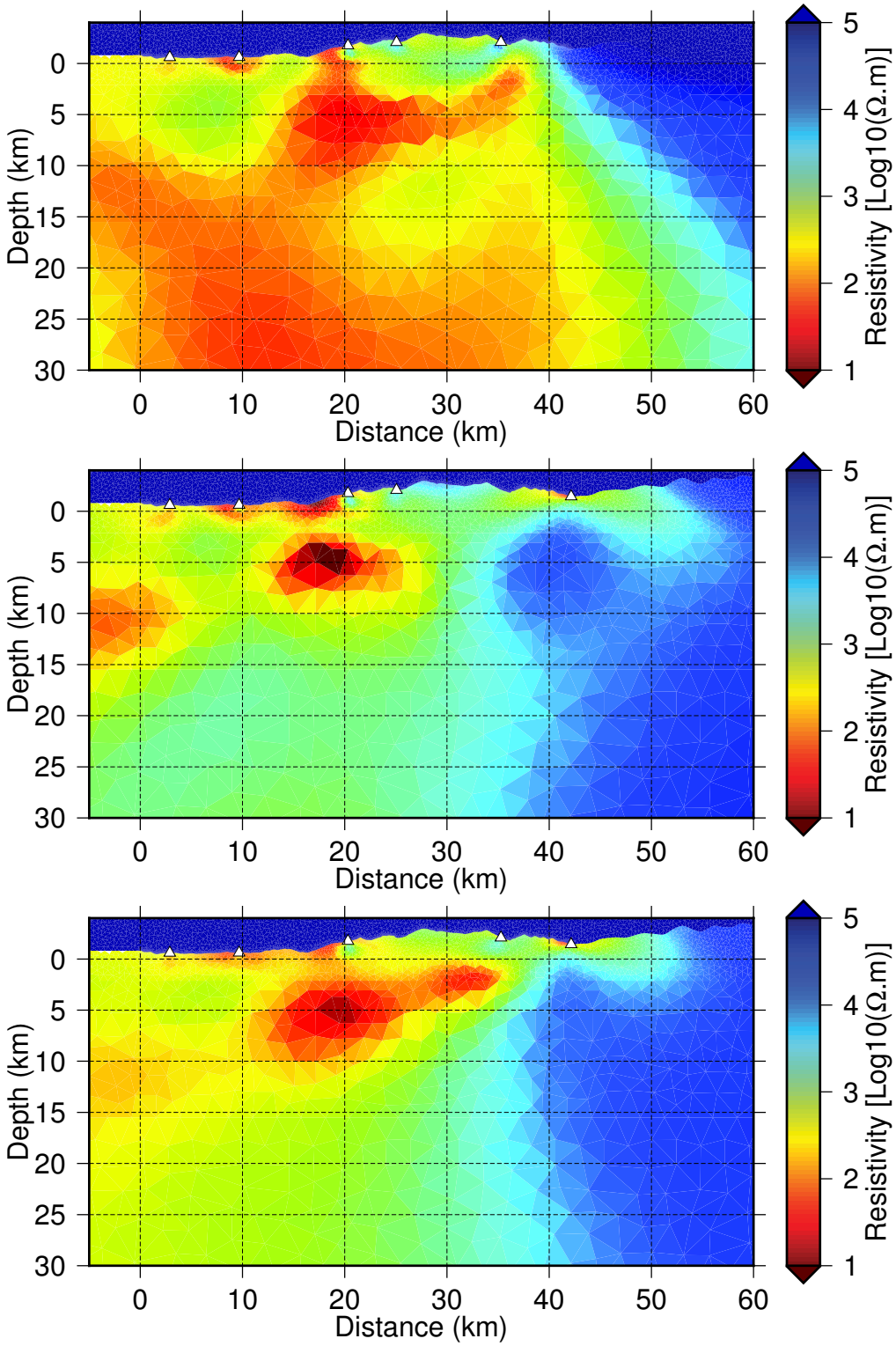


FIGURE 5.28: Modèle de résistivité issue de l'inversion du profil en retirant Syafru (haut), Dhunche (milieu) et Grang (bas). Les triangles blancs indiquent la position des stations utilisées dans l'inversion.

FIGURE 5.28: *Resistivity model from inversion without Syafru MT station (top), without Dhunche MT station (middle), and without Grang MT station (bottom). White triangles indicate MT stations used in inversion.*

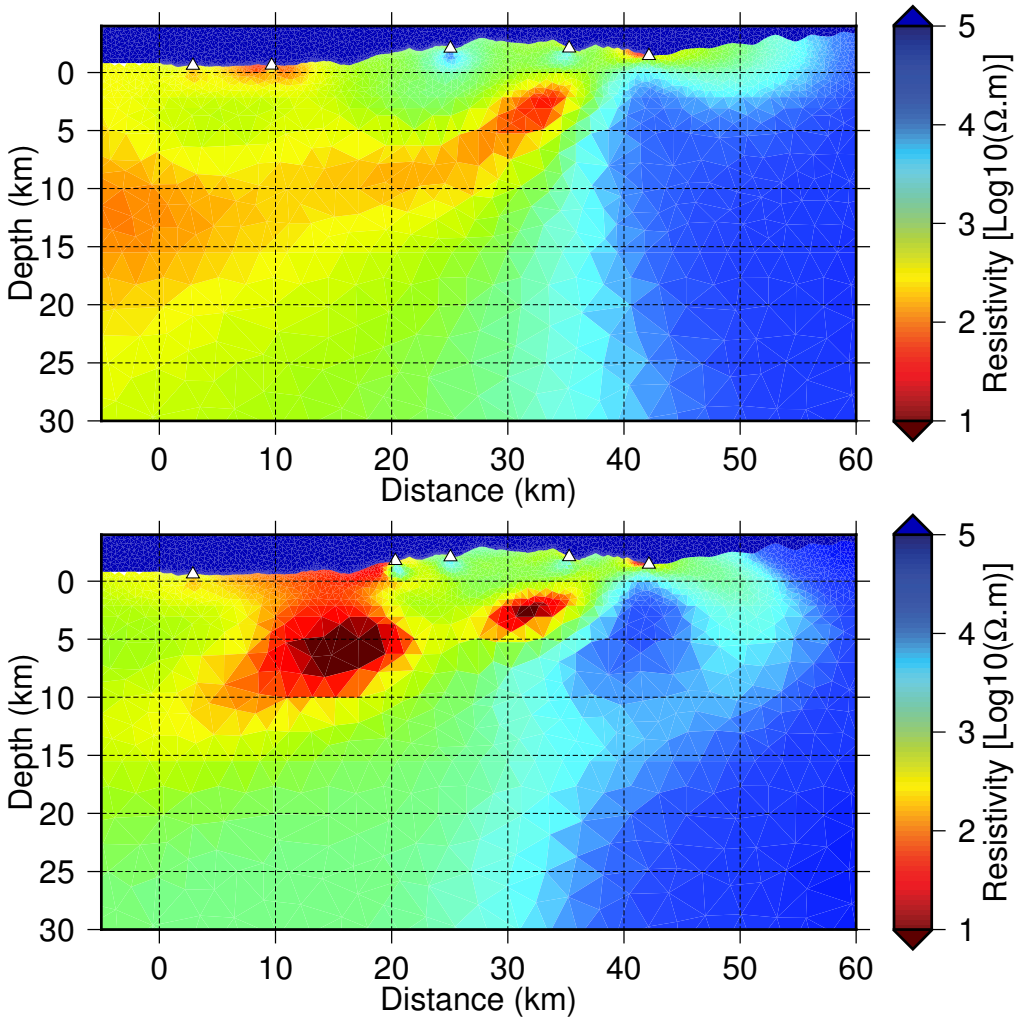


FIGURE 5.29: Modèle de résistivité issue de l'inversion du profil en retirant Kalikastan (haut) et Trisuli (bas). Les triangles blancs indiquent la position des stations utilisées dans l'inversion.
FIGURE 5.29: *Resistivity model from inversion without Kalikastan MT station (top), without Trisuli MT station (bottom). White triangles indicate MT stations used in inversion.*

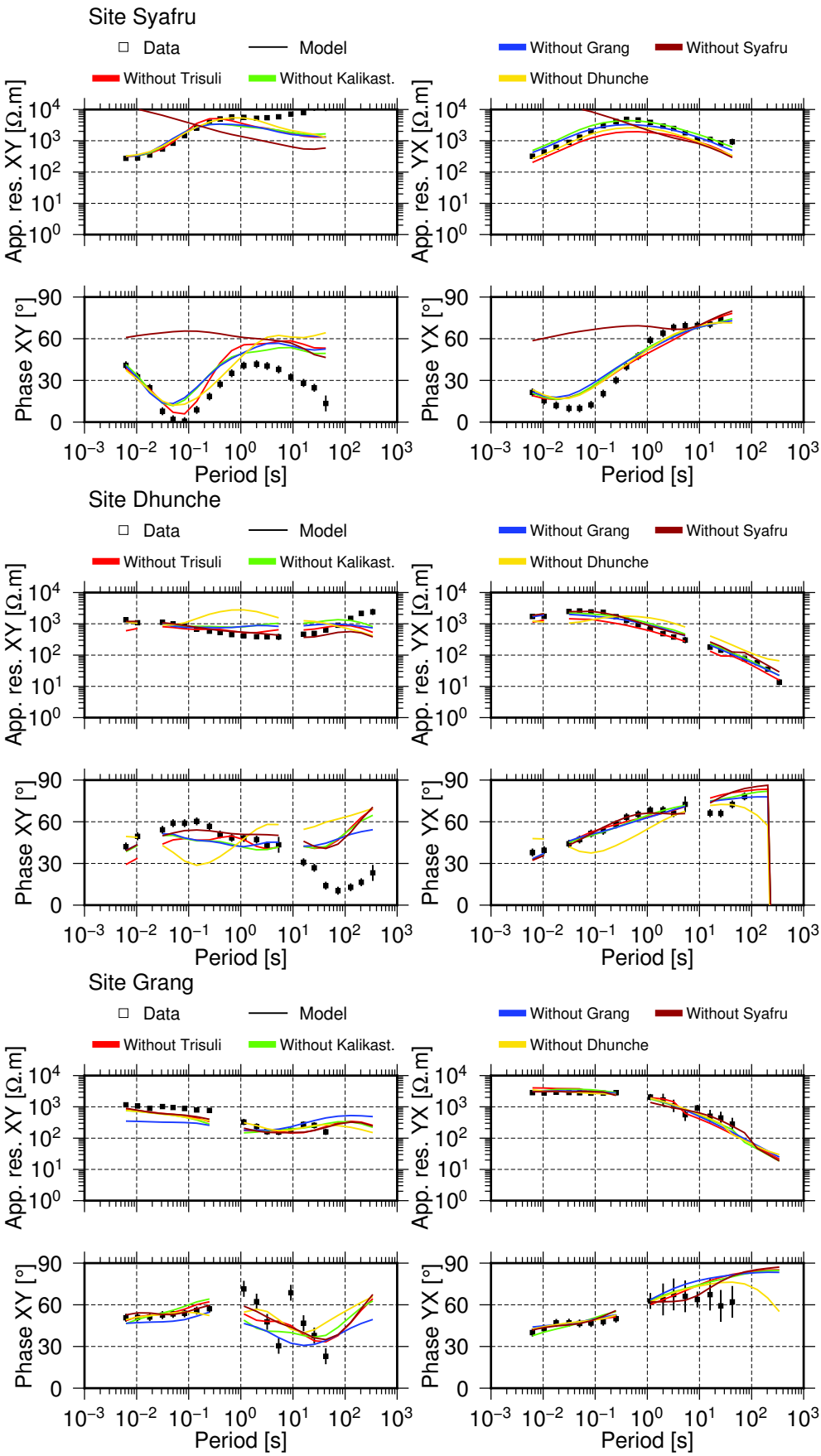


FIGURE 5.30: Sondages obtenus par le retrait des stations MT pour les stations de Syafru (top), Dhunche (middle) and Grang (bottom).

FIGURE 5.30: *Sondages obtained from forward modeling with removal of MT stations for Syafru (top), Dhunche (middle), Grang (bottom).*

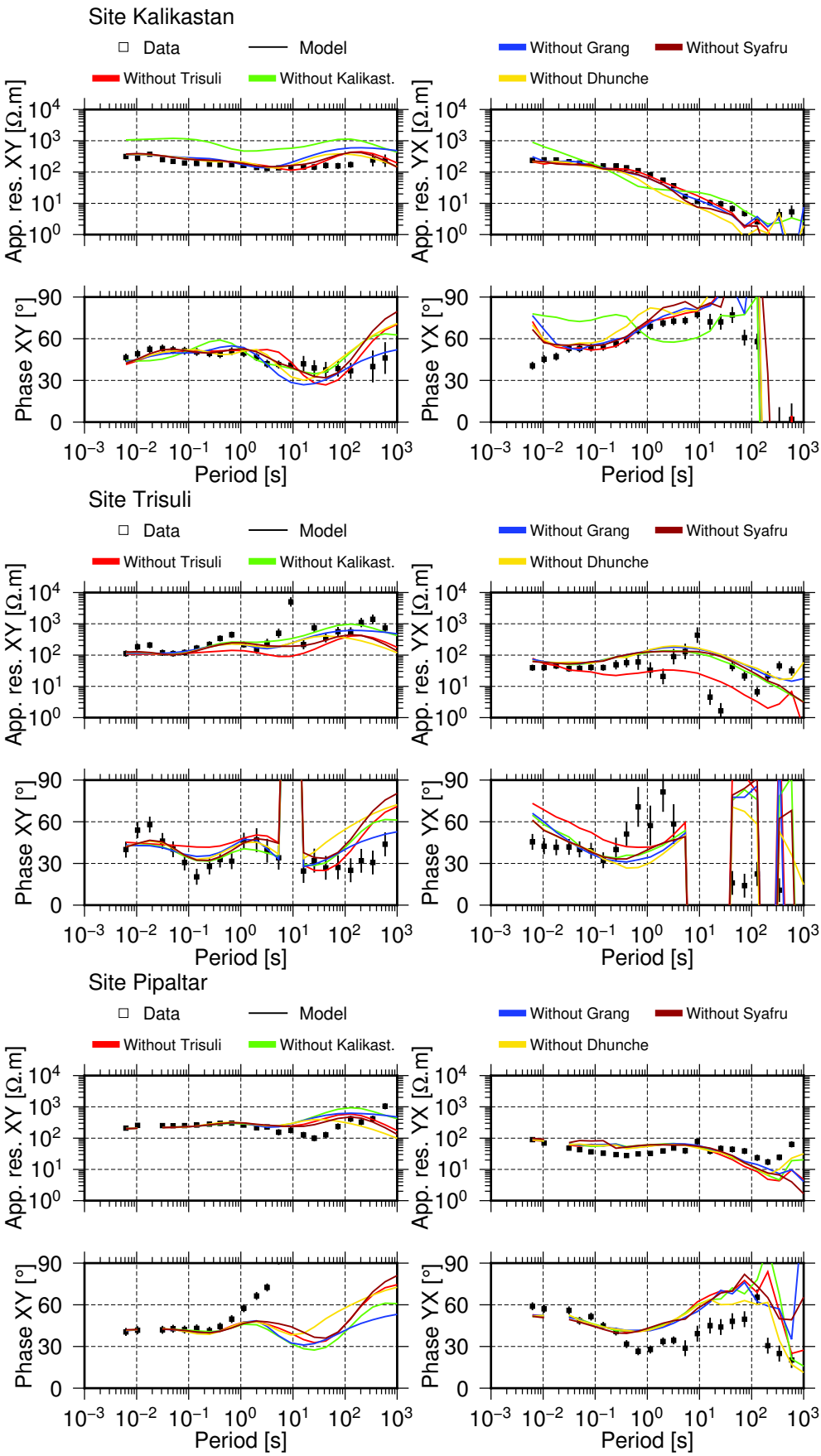


FIGURE 5.31: Sondages obtenus par le retrait des stations MT pour les stations de Kalikastan (top), Trisuli (middle) and Pipaltar (bottom).

FIGURE 5.31: *Sondages obtained from forward modeling with removal of MT stations for Kalikastan (top), Trisuli (middle), Pipaltar (bottom).*

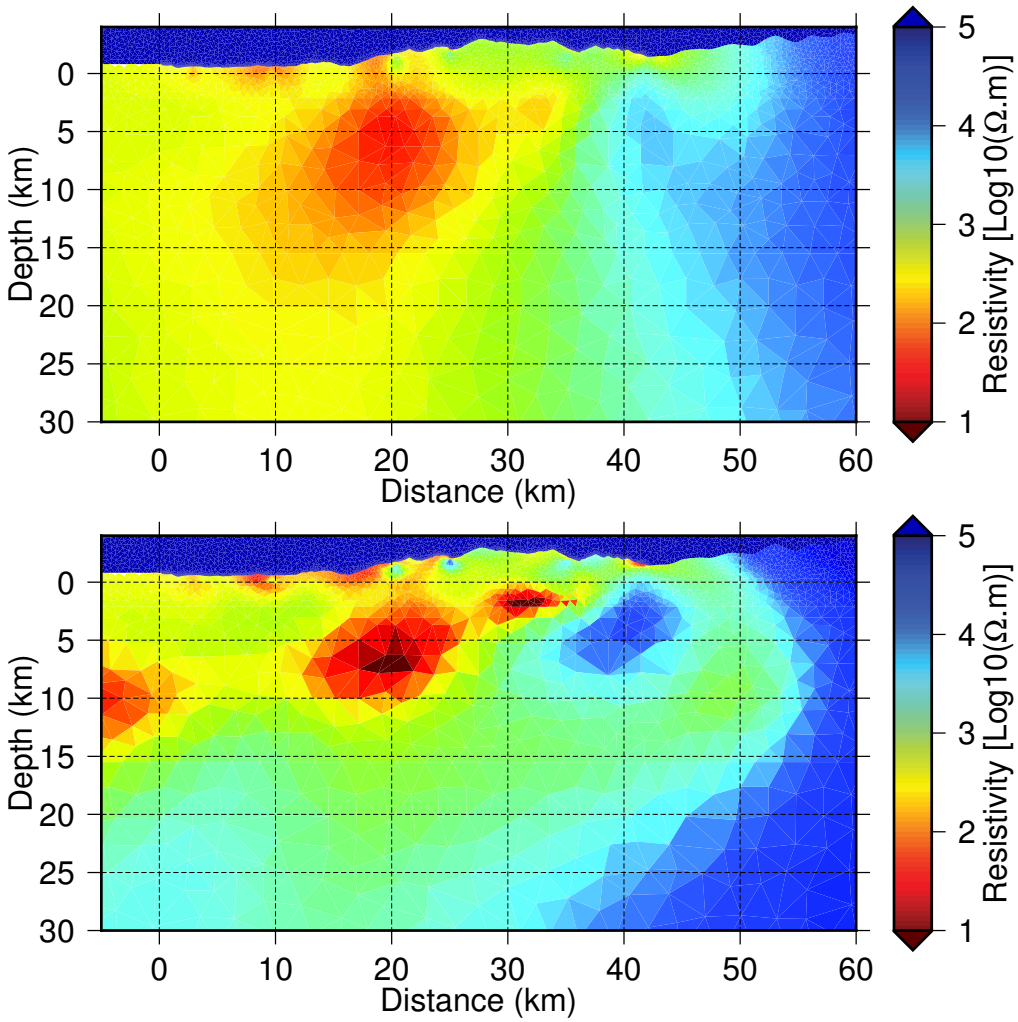


FIGURE 5.32: Modèles de résistivité obtenus en changeant le paramètre de régularisation réglant la résolution horizontale et verticale. Haut : Rapport horizontal/vertical 1/3. Bas : Rapport horizontal/vertical 1. Les triangles blancs indiquent la position des stations utilisées dans l'inversion.

FIGURE 5.32: *Resistivity model obtained by changing the regularization parameter between horizontal and vertical resolution. Top : Ratio horizontal/vertical 1/3. Bottom : Ratio horizontal/vertical 1. White triangles indicate MT stations used in inversion.*

5.6.3.2 Variation de la résolution verticale et horizontale

Un autre test pour évaluer la robustesse de la distribution de résistivité obtenue par inversion est de modifier la régularisation appliquée dans la procédure d'inversion. Par défaut, Mare2DEM utilise un facteur 3 entre la régularisation horizontale et verticale, privilégiant la résolution horizontale à la résolution verticale. La Figure 5.32 montre le résultat de l'inversion pour un facteur 3 entre la régularisation verticale et horizontales, et pour une égalité entre les régularisations horizontales et verticales. Pour une valeur supérieure de la régularisation verticale, l'inversion a convergé pour un *misfit* plus important que l'ensemble des autres tests, avec un *misfit* final à

13.33. Le modèle final est de fait moins résolu, mais permet de toujours distinguer les principales structures déterminées précédemment.
L'inversion avec égalité entre les résolutions horizontales et verticales est très similaire, mais ajoute une structure au nord du profil qui apparaît comme une connexion entre la structure conductrice au nord de Syafru et celles du sud du profil. Ètant donné la sensibilité des données MT en dehors du profil, il serait très imprudent d'interpréter cette structure qui n'a pas de contrainte réelle. Le reste du profil est cohérent avec les autres inversions.

5.7 Interprétation

5.7.1 Comparaison entre les différents modèles obtenus et les données

La Figure 5.33 montre les trois modèles obtenus jusqu'à présent à partir des données MT acquises dans la vallée de Trisuli (Carole Lemonnier et al. (1999), C. Lemonnier (2000), et le modèle présenté ici). La caractéristique principale des modèles présentés en 1999 et 2000 est la présence du corps conducteur situé à environ 15-20 km de profondeur au nord du profil. Ce corps conducteur est totalement absent du nouveau profil. Celui-ci présente des corps conducteurs situés sous le maximum de topographie, situés plus au sud, et moins profondément, entre 5 et 10 km sous le niveau de la mer.

Au sud, à -10 km dans la direction du profil, la structure conductrice située entre 10 et 15 kilomètres de profondeur apparaît cohérente entre les trois inversions. Il est cependant délicat de vouloir confirmer cette structure pour le nouveau modèle du fait du manque de sensibilité en dehors du profil de stations MT.

Pour comparer les différents modèles entre eux, nous avons calculé le *misfit* sur le déterminant du tenseur. Pour un modèle 2D dans la direction de la structure régionale, le déterminant est simplement $Z_{det} = \sqrt{Z_{TE} Z_{TM}}$. Ainsi :

$$\begin{aligned} \rho_{det} = & \quad (\rho_{TE} \times \rho_{TM})^{1/2} \\ \phi_{det} = & \quad (\varphi_{TE} + \varphi_{TM})/2 \end{aligned} \tag{5.1}$$

La RMS est calculée par :

$$\begin{aligned} RMS_{\rho} = & \quad 100 \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\rho^i_{data} - \rho^i_{model})^2}}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\rho^i_{data})^2}} \\ RMS_{\phi} = & \quad 100 \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N} d(\phi^i_{data}, \phi^i_{model})^2}}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\phi^i_{data})^2}} \end{aligned} \tag{5.2}$$

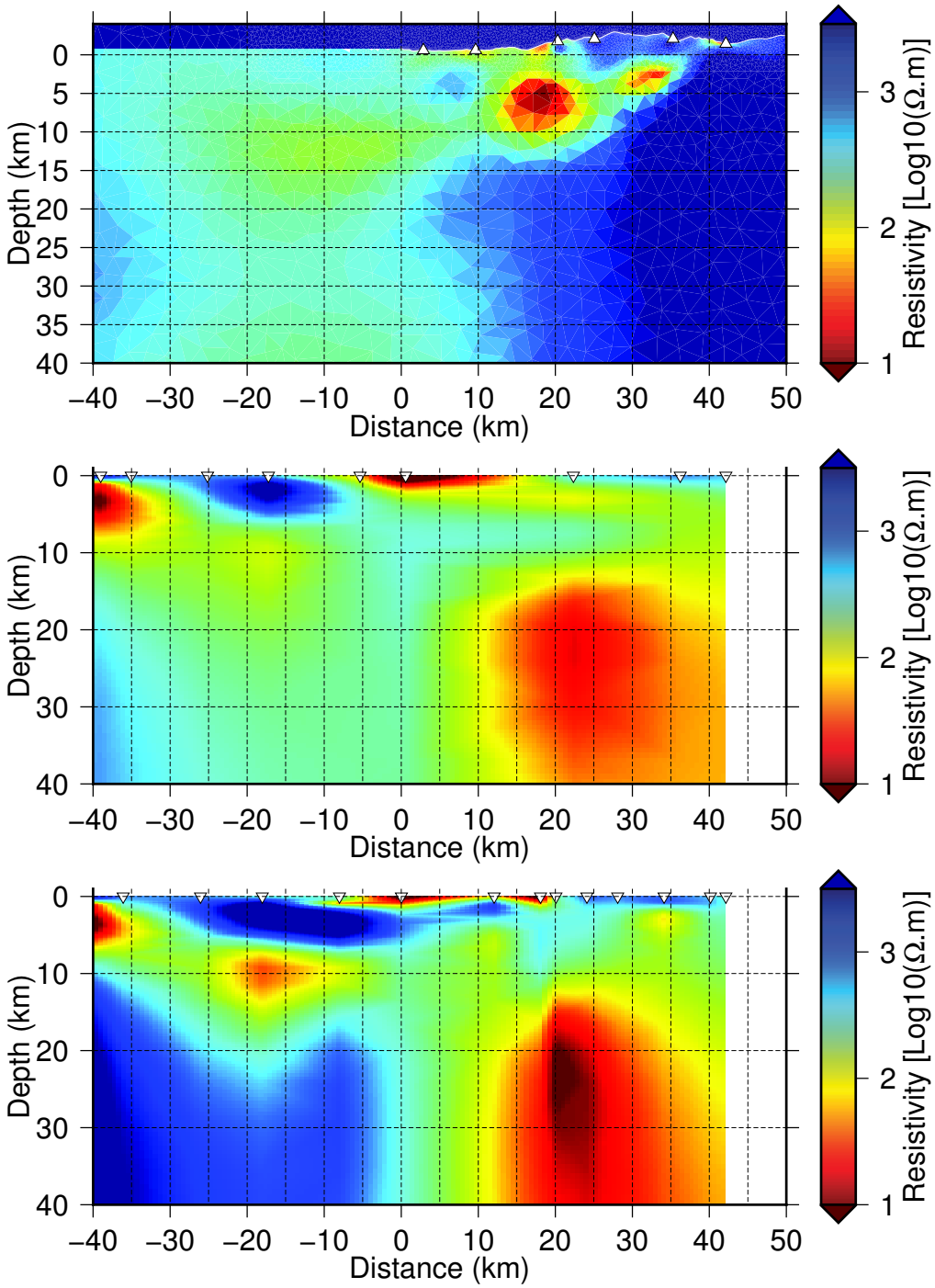


FIGURE 5.33: Modèles de conductivité obtenus par l'inversion des données de 2016 (haut), de 1996 (milieu) et des campagnes de 1996, 1999, 2000 (bas). Les triangles blancs indiquent la position des stations utilisées dans l'inversion.

FIGURE 5.33: *Conductivity distributions from the inversion of 2016 data (top), 1996 data (middle) and 1996, 1999, 2000 data (bottom). White triangles indicate MT stations used in inversion.*

Pour permettre le calcul de la RMS pour les anciens modèles à la position des nouveaux sites, nous avons interpolé les anciens modèles dans un maillage en éléments finis. A partir de ces

interpolations, nous avons modélisé les nouveaux sondages MT avec Mare2DEM. Les anciens sondages pour des sites comparables tels que Dhunche ou Syafru ont aussi été interpolés pour permettre le calcul de la RMS aux nouvelles fréquences. La Figure 5.34 montre le calcul des RMS.

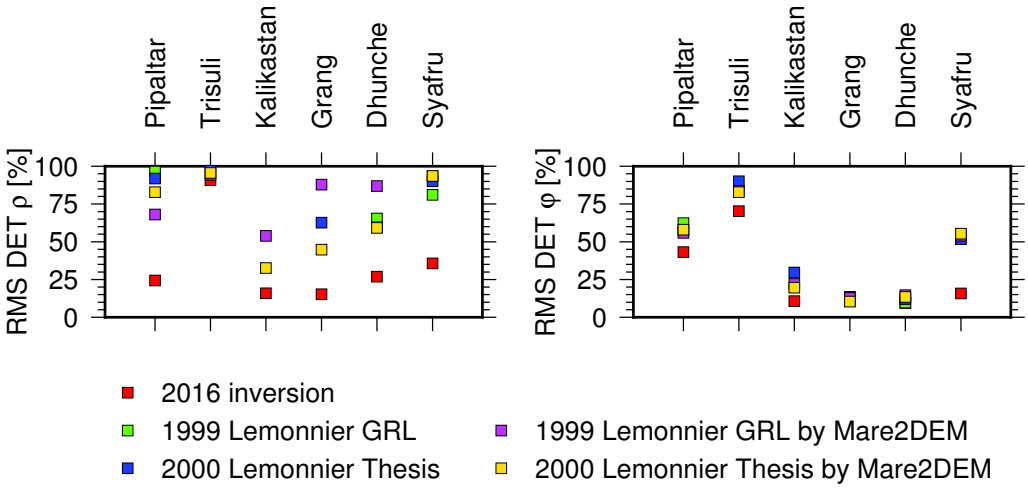


FIGURE 5.34: RMS mesurées entre les sondages MT 2016 et les sondages modélisés par les différentes distributions de conductivité.
FIGURE 5.34: *Difference between 2016 MT soundings and the modeling of conductivity distributions.*

Les RMS du nouveau modèle sont significativement inférieures à celles des anciens modèles. Trisuli présente un grand écart du à la composante TM extrêmement bruitée. La RMS de Pipaltar concernant la phase est aussi importante due à la composante TE dont la phase augmente de façon importante pour les périodes supérieures à la seconde. Le site de Syafru est cependant bien mieux modélisé qu'auparavant. L'incorporation de la topographie change ici de façon importante la distribution de conductivité électrique inversée.

5.7.2 Lien avec d'autres études géophysiques

Girault et Frédéric Perrier (2014) ont mesuré les flux de radons et de CO_2 au niveau du village de Syafru. Ils ont proposé trois modèles permettant d'expliquer l'origine des flux. La Figure 5.35 présente les trois modèles. Le scénario M3 propose une origine en profondeur des flux, par des failles se propageant jusqu'à la surface. Le modèle de conductivité proposé ici montre une connectivité des différentes anomalies conductrices, partant de l'anomalie située entre 5 et 10 km de profondeur, se continuant jusqu'aux zones conductrices observées à Syafru. La position de ces zones conductrices est partiellement en accord avec la coupe géologique proposée par Hubbard et al. (2016) montrée Figure 5.36. Elles correspondent à la remontée de la faille MBT actuellement inactive. En revanche, la flexion du MBT qui lui permet de replonger en

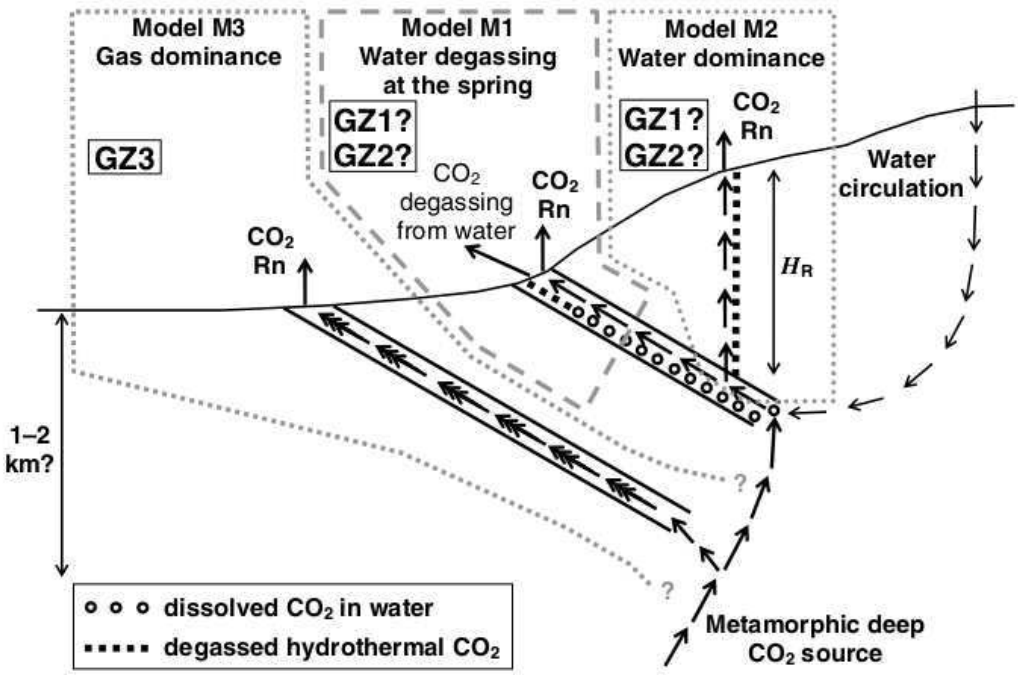


FIGURE 5.35: Différents modèles d'origine des flux de CO_2 et de radon à la source hydrothermale de Syabru-Bensi. D'après GIRAULT et Frédéric PERRIER (2014).
FIGURE 5.35: *Sketch of possible origine of CO_2 and radon fluxes at the Syabru-Bensi hydrothermal system. From* GIRAULT *et Frédéric* PERRIER *(2014).*

profondeur n'est pas cohérente avec l'image MT observée. Le modèle proposé par GIRAULT et Frédéric PERRIER (2014) décrivant la source du CO_2 en profondeur, amené en surface par un système de failles, permettrait d'apporter des informations sur la phase de nucléation des séismes. La nouvelle section MT ne permet en revanche pas de mettre en relation les foyers de sismicité des répliques observées jusqu'à 45 jours après le séisme de Gorkha. La sismicité est localisée entre 10 et 15 kilomètres de profondeur (LETORT et al., 2016).

5.8 Conclusion

Dans ce chapitre, nous avons présenté la nouvelle campagne d'acquisition MT au Népal et les enseignement que nous pouvons en tirer.

Les sondages MT montrent une nette amélioration de la qualité par rapport aux campagnes effectuées à la fin des années 1990. Les fonctions de transfert verticales sont cohérentes entre les deux campagnes et permettent de confirmer les interprétations reliées aux flèches d'inductions.

Nous avons montré que l'utilisation de la topographie dans les modélisations modifie considérablement la distribution de conductivité inversée. Ainsi, le modèle inversé ne présente non

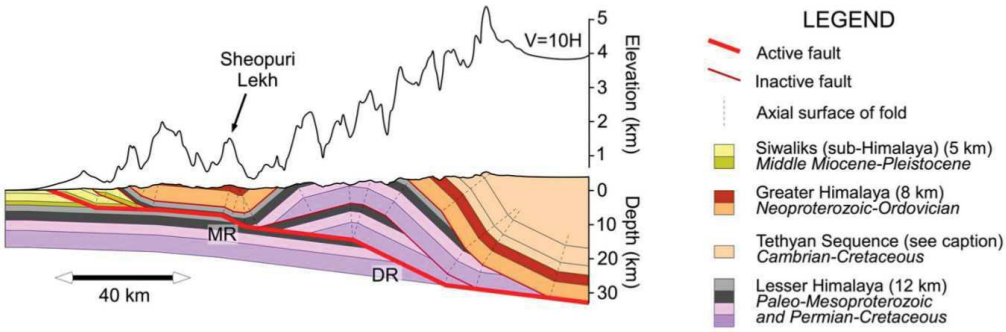


FIGURE 5.36: Coupe géologique de la vallée de Trisuli au Népal. D'après HUBBARD et al. (2016).
FIGURE 5.36: *Geological cross-section of Trisuli valley in Nepal. From* HUBBARD *et al. (2016).*

plus une anomalie de conductivité profonde, mais située plus en surface. Cette anomalie apparaît située au niveau de l'ancienne faille du MBT et non plus au niveau de la faille du MHT actuellement active, reliée au séisme de Gorkha. La plupart des structures inversées peuvent être considérées comme robustes suite aux différents tests de sensibilité que nous avons effectué. Le nouveau modèle proposé peut notamment être mis en relation avec d'autres études géophysiques effectuées dans la vallée de Trisuli. Un travail d'interprétation géologique est actuellement en cours pour pouvoir intégrer ce modèle de résistivité dans les modèles tectoniques et en déterminer les éventuelles implications.

Les résultats obtenus sont cependant à considérer avec précaution, compte tenu des informations sur la dimensionnalité des structures, obtenues à partir de l'analyse des sondages MT. En effet, nous avons montré par une modélisation 3D que pour pouvoir interpréter correctement les sondages MT au Népal, il est en revanche nécessaire de pouvoir modéliser les effets topographiques 3D. La modélisation de la topographie doit ainsi être extrêmement fine ou faite autrement que par des mailles rectangulaires pour pouvoir accommoder des variations de topographie importantes telles qu'observées sur le terrain. Les codes de modélisation 3D actuels ne permettent pas cette démarche.

5.9 Annexes

5.9.1 Comparaison entre les campagnes de 1996 et 2016

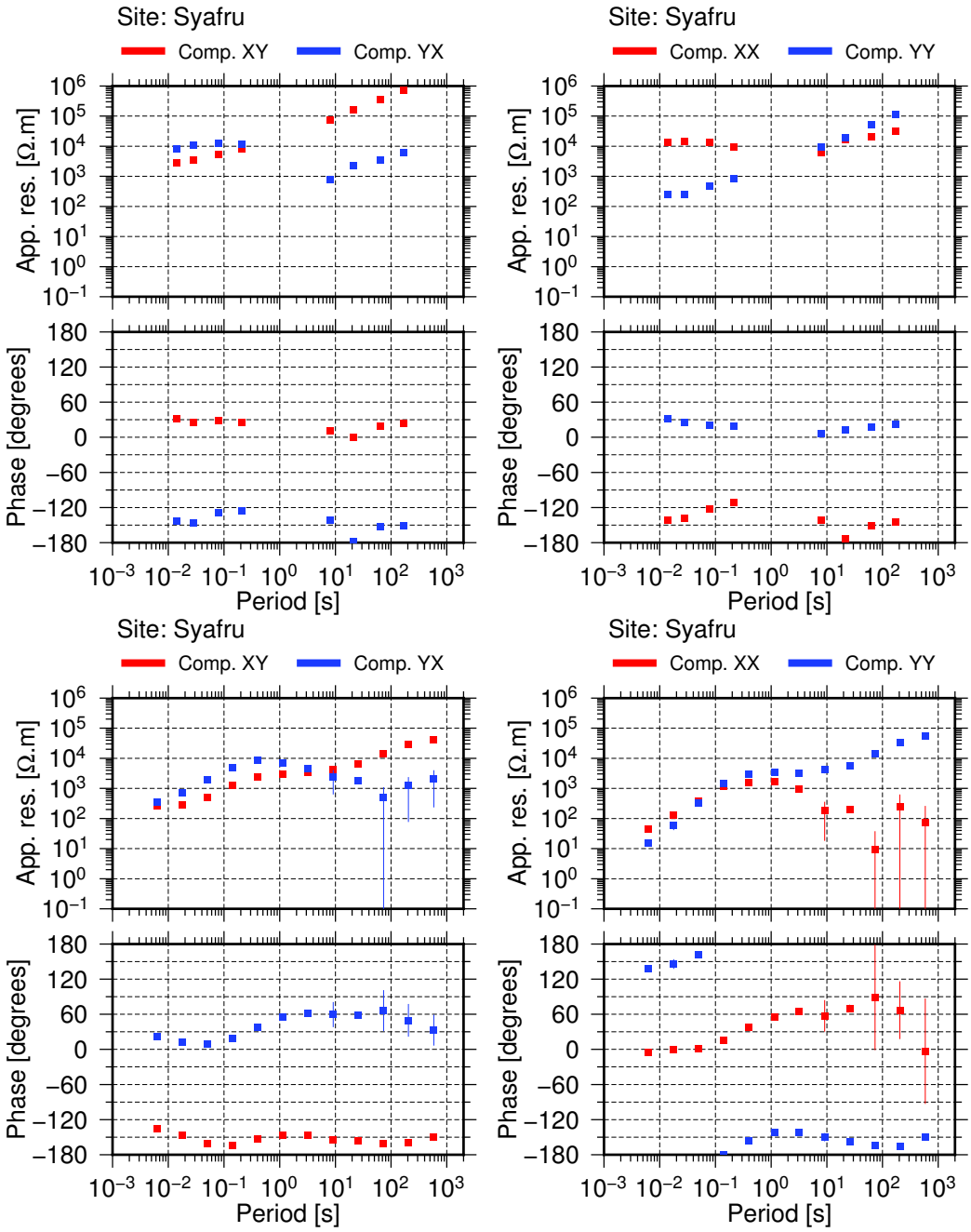


FIGURE 5.37: Comparaison entre les sondages obtenus en 1996 (haut) et 2016 (bas) pour la station de Syafru.

FIGURE 5.37: *Comparison between 1996 MT soundings (top) and 2016 MT soundings (bottom) for Syafru station.*

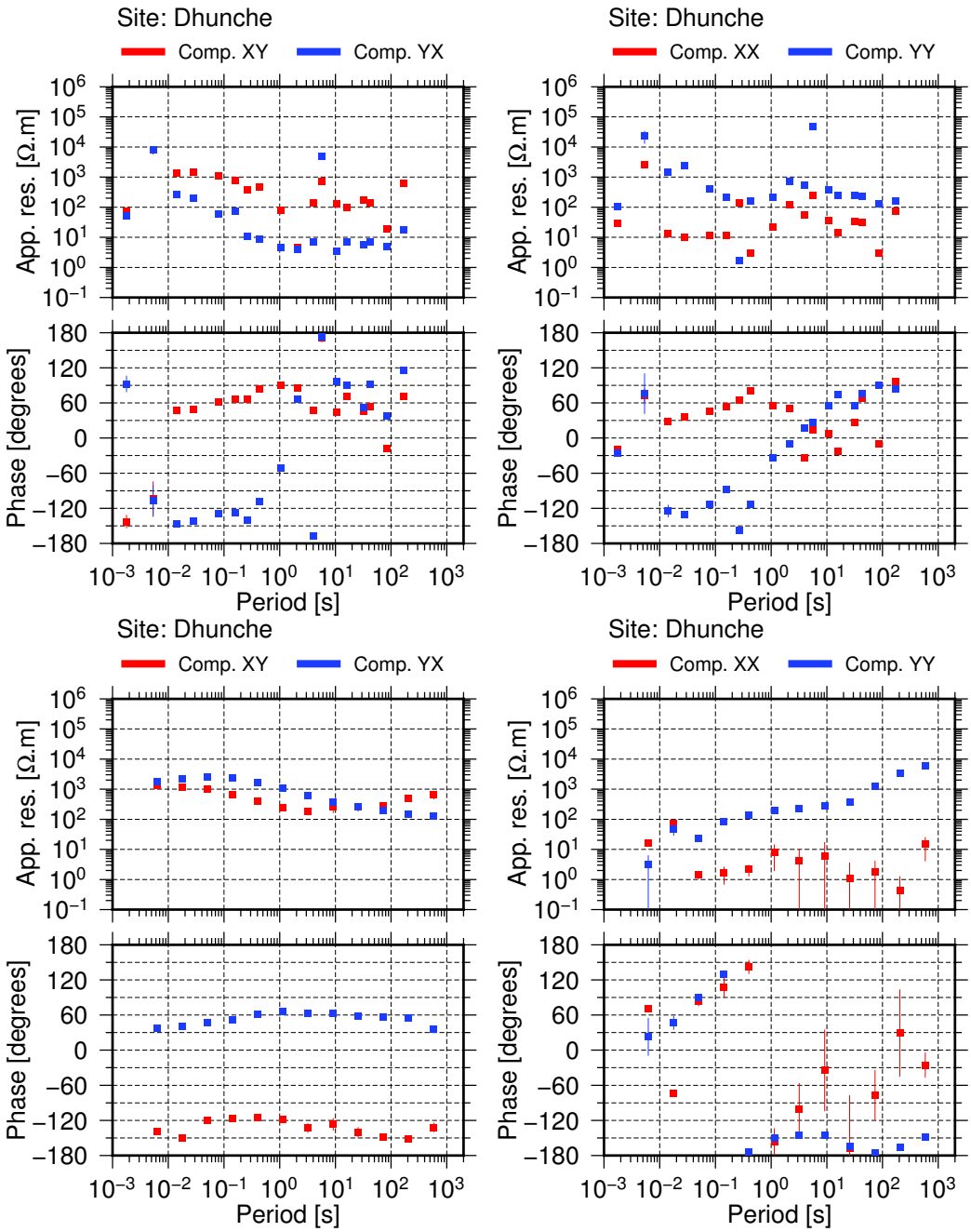


FIGURE 5.38: Comparaison entre les sondages obtenus en 1996 (haut) et 2016 (bas) pour la station de Dhunche.

FIGURE 5.38: *Comparison between 1996 MT soundings (top) and 2016 MT soundings (bottom) for Dhunche station.*

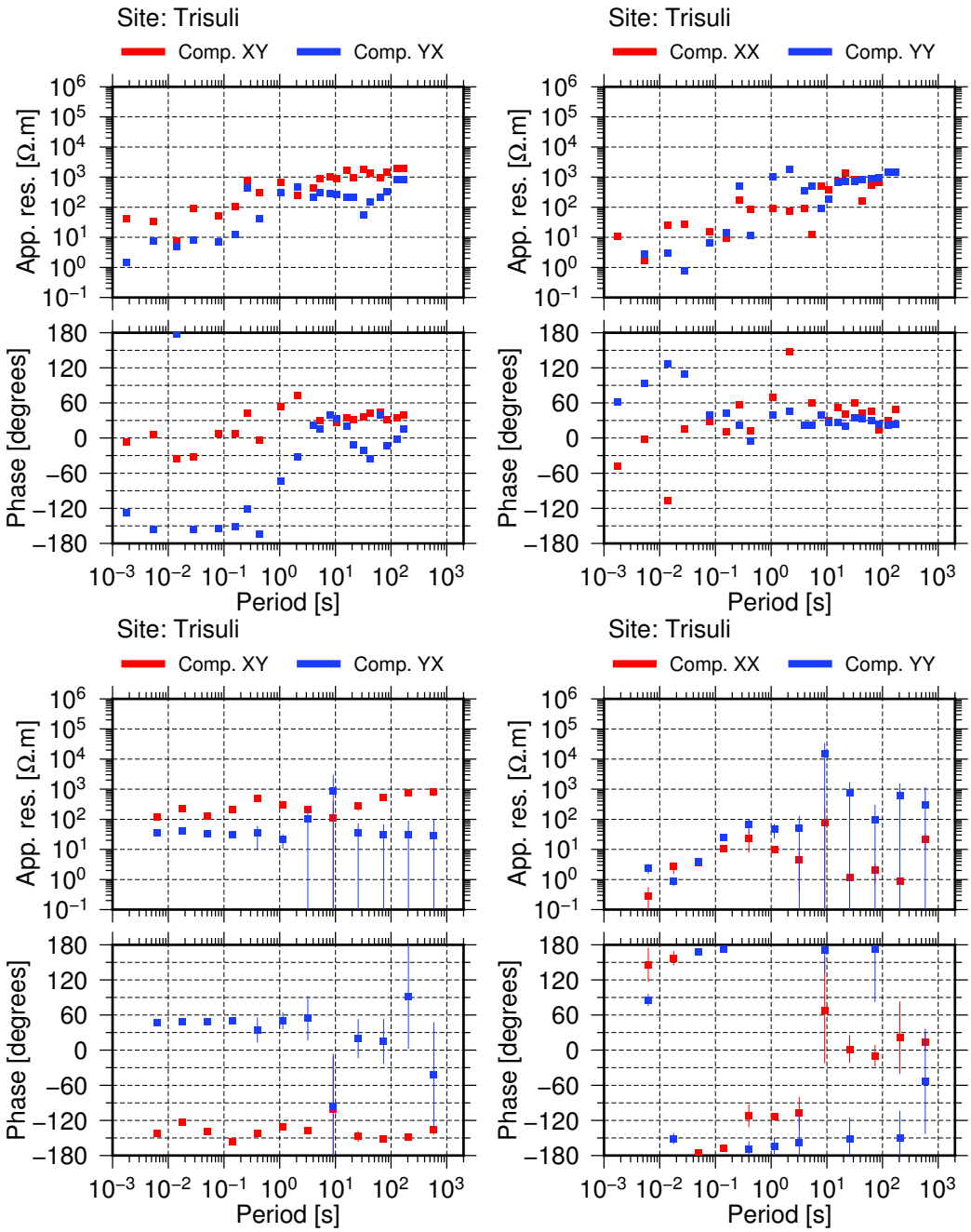


FIGURE 5.39: Comparaison entre les sondages obtenus en 1996 (haut) et 2016 (bas) pour la station de Trisuli.

FIGURE 5.39: *Comparison between 1996 MT soundings (top) and 2016 MT soundings (bottom) for Trisuli station.*

5.9.2 Décompositions de Groom-Bailey

5.9.2.1 Station simple, par période

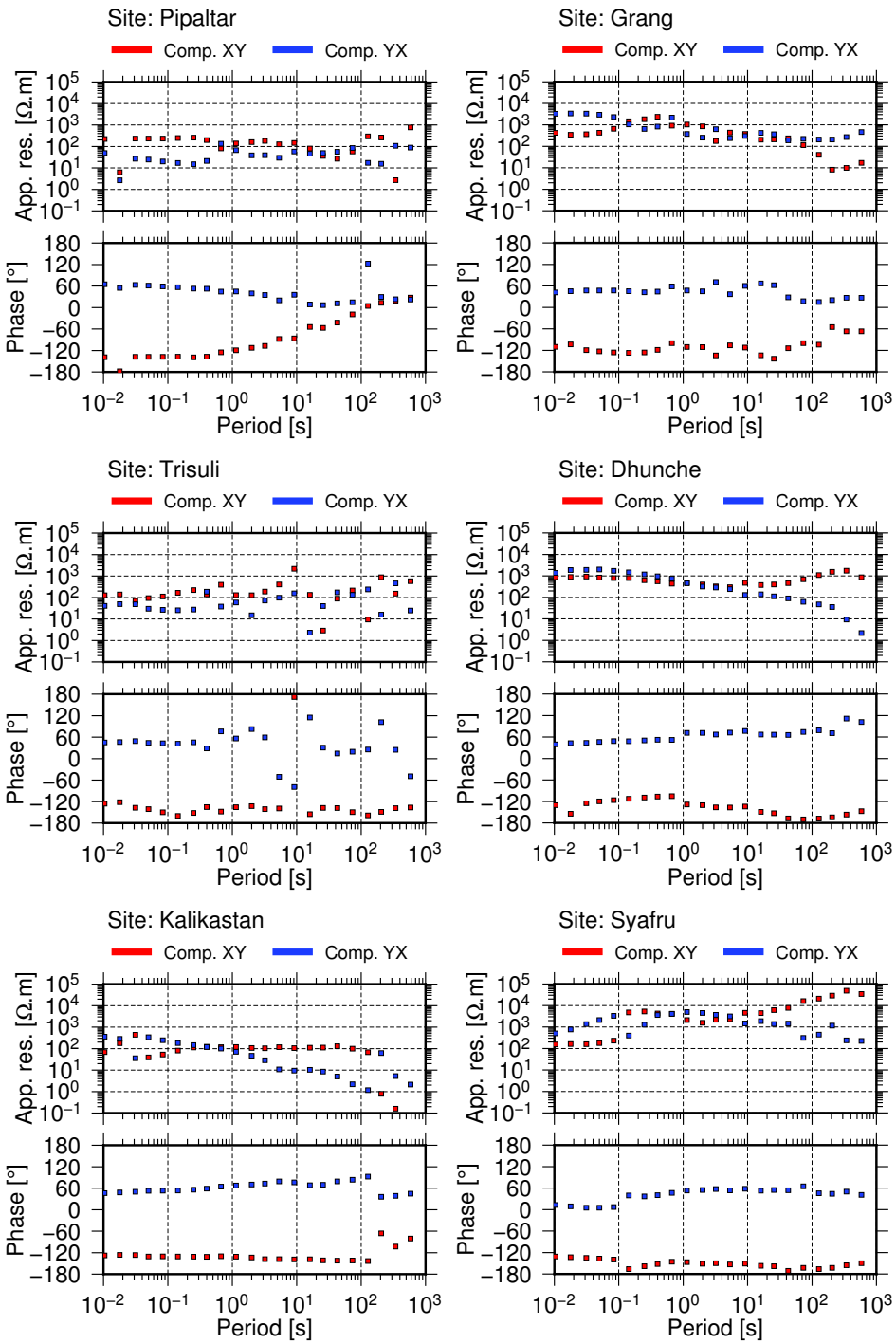


FIGURE 5.40: Décomposition de Groom-Bailey par fréquence et par site.
FIGURE 5.40: *Single site, single frequency Groom-Bailey decomposition..*

5.9.2.2 Station simple, par période, strike fixé à N20E

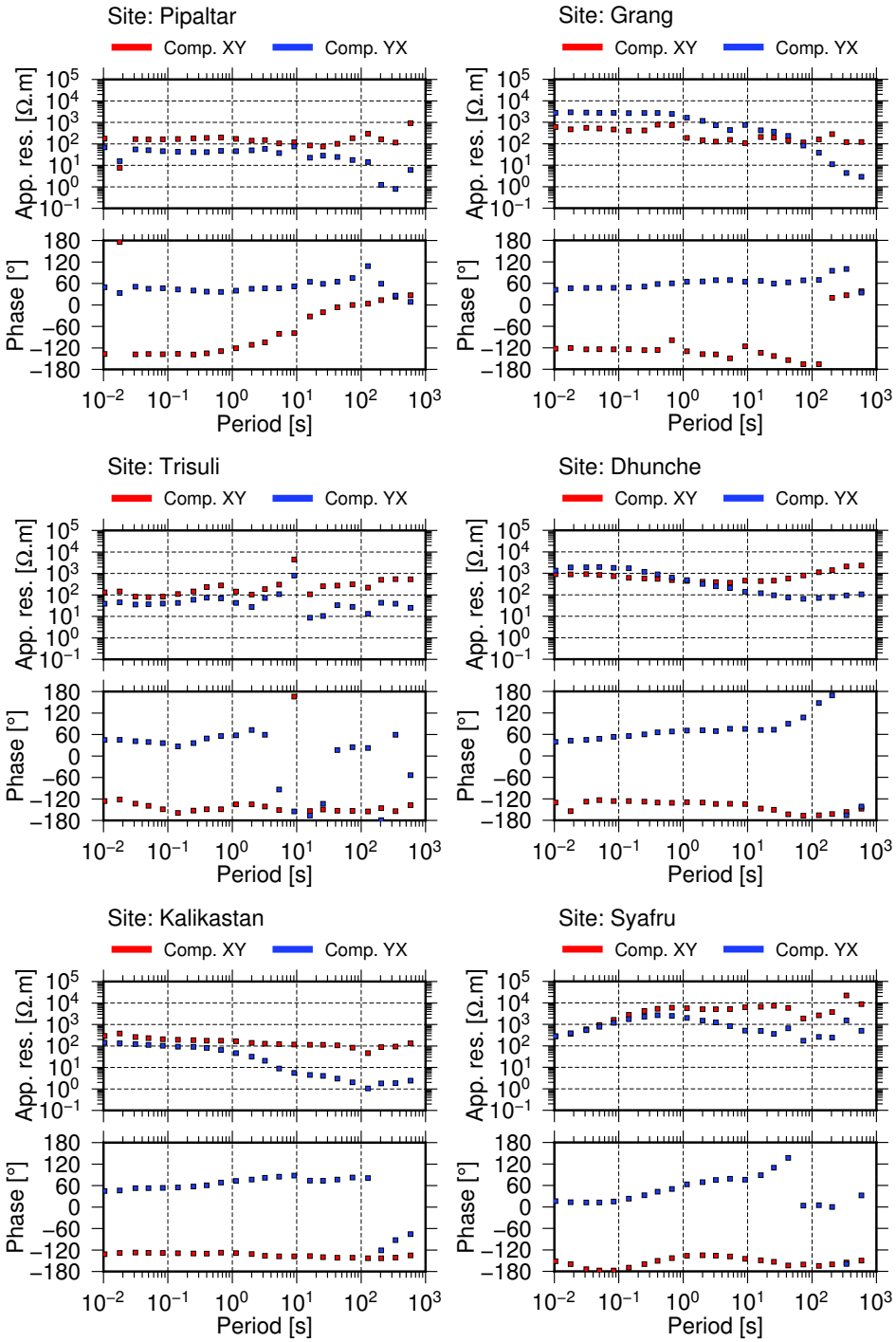


FIGURE 5.41: Décompositions de Groom-Bailey pour un strike fixé à N20E.
FIGURE 5.41: *Groom-Bailey decomposition for a N20E strike.*

5.9.2.3 Multi-station, pour les périodes > 1s

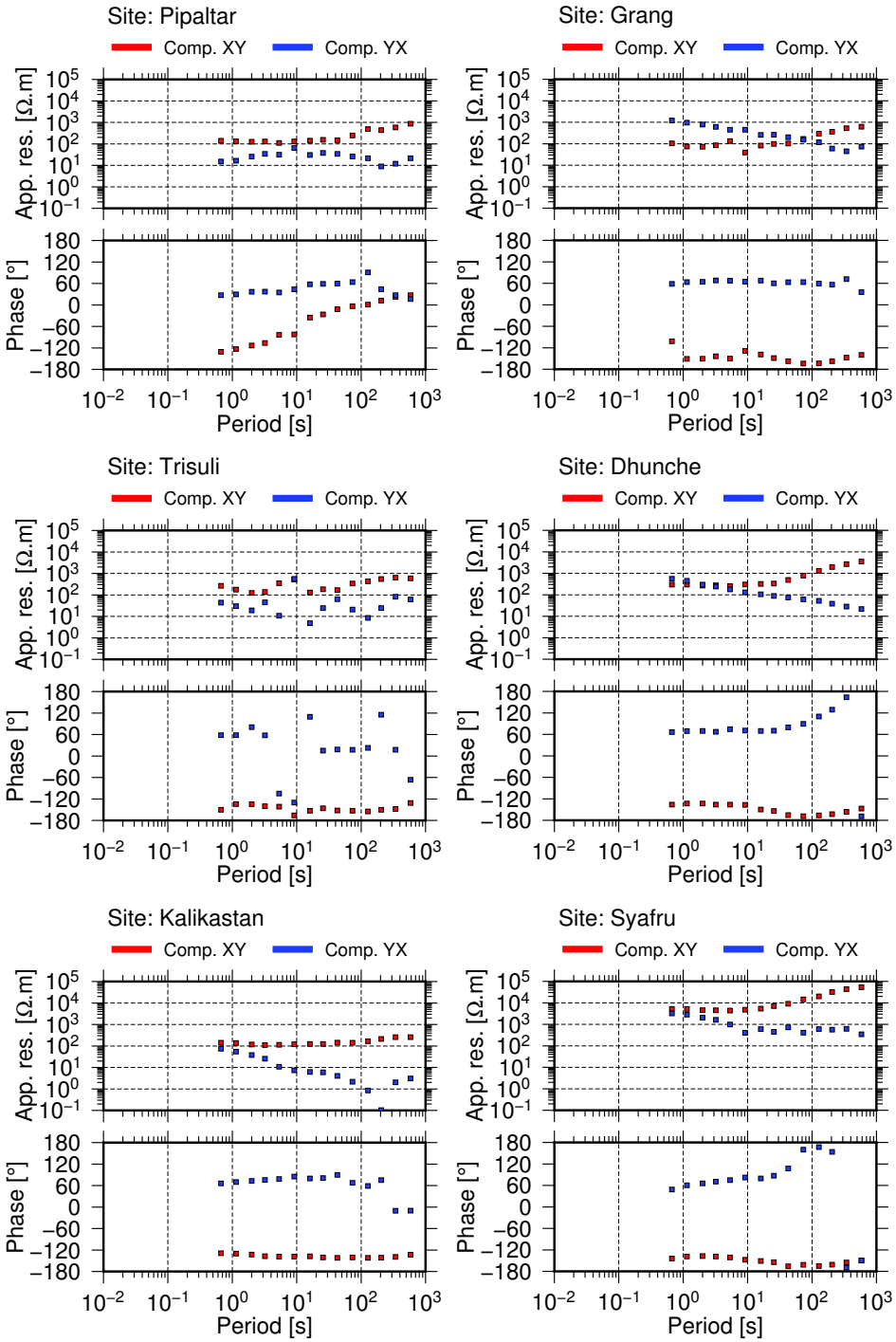


FIGURE 5.42: Décompositions de Groom-Bailey multi-sites.
FIGURE 5.42: *Multi-sites Groom-Bailey decomposition.*

Quatrième partie

Conclusion

Chapitre 6

Conclusion générale

Traitement de données MT

Les sources géomagnétiques et atmosphériques ont chacune une variabilité qui a de fortes répercussions sur le résultat des traitements de séries MT. Ainsi, avant même de se poser la question du traitement statistique, il est nécessaire d'avoir des outils permettant de vérifier si une série temporelle contient véritablement des sources naturelles d'induction. Le bruit d'origine anthropique étant cohérent, les outils de cohérence entre séries magnétiques et électriques ne peuvent donc pas être utilisés seuls.
Les outils que nous avons développé permettent de déterminer la présence ou l'absence de sources MT, ainsi que leurs contenus fréquentiels. Un diagnostic peut alors être effectué à posteriori sur la qualité du sondage MT, ou, en "temps réél", sur la gestion même des campagnes de mesures MT. La surveillance des événements géomagnétiques par des stations, telles les observatoires magnétiques, permet de juger si une station temporaire MT a besoin d'être laissée plus longtemps en acquisition pour obtenir plus de signal naturel lors de son traitement en laboratoire. La fréquence d'échantillonnage des données des observatoires magnétiques (1 Hz pour INTERMAGNET, 10 Hz dans de rares cas) limite leur utilisation aux sources basses fréquences. A l'heure actuelle, il est donc impossible d'utiliser les observatoires en tant que stations de référence pour les sources atmosphériques. Certaines démarches visent à pallier à cette limitation. Ainsi, localement, à la station géophysique du Welschbruch, nous avons installé une station MT enregistrant à plus haute fréquence. A une plus grande échelle, un réseau tel que le *National Lightning Detection Network* (NLDN), installé sur le territoire des Etats-Unis d'Amérique, permet de localiser les événements atmosphériques. Ce type de réseaux pourrait avoir avoir des implications majeures dans les traitements MT. Evidemment, un tel réseau transposé en France ou en Europe présenterait des limitations structurelles et logistiques. En effet, pour une étude optimale de l'ensemble des signaux atmosphériques en MT, les acquisitions nécessiteraient d'être mise en place à la

fréquence de 32 kHz. Cela entrainerait des besoins conséquents en moyens de stockage. A titre d'exmple, actuellement, une journée d'acquisition des voies magnétiques horizontales installées au Welschbruch à la fréquence de 512 Hz fait une taille de 338 Mo. L'équivalent 32 kHz ferait ainsi 21 Go pour 24h d'enregistrement pour une seule station.

Nous avons montré qu'une meilleure connaissance des sources MT permet également d'améliorer le sondage MT. Le signal MT subit des variations importantes dans certaines bandes de fréquences, comme l'Audio-MT *dead band*. Or, les traitements de données MT sont majoritairement basés sur l'hypothèse de source quasi-stationnaire. Les fluctuations observées affectent donc les résidus du champ électrique pour ces intervalles de fréquences. D'après les derniers développements disponibles dans la littérature, la distribution des résidus sur le champ électrique présenterait une forme de distribution stable. Nous avons montré que la sélection des coefficients permet aux résidus de mieux s'ajuster à des distributions gaussiennes, aussi bien pour les sources hautes fréquences que basses fréquences. Les distributions observées à queue lourde traduisent-elles l'absence de certaines sources ?
Enfin, nous nous posons la question des sources identifiables dans la MT *dead band*. Pour certaines séries temporelles, le calcul classique des fonctions de réponse MT donne un résultat statistiquement correct, même en l'absence de sources repérables dans les coefficients d'ondelettes.
Pour établir ces résultats, nous avons utilisé la transformée en ondelettes continue (TO). Cette technique est une des nombreuses représentations temps-fréquence possible d'un signal. La TO demeure, dans certains cas, limitée pour isoler des composantes qui se superposent. Ainsi, le cas d'un orage magnétique avec superposition de plusieurs modes de résonances du champ magnétique pose problème dans l'application de la transformée en ondelette car celle-ci ne permet pas nécessairement de distinguer les différents modes. Dans ce cas précis, la transformée de Hilbert-Huang a donné d'excellents résultats dans la séparation de plusieurs pulsations se superposant à la suite d'un orage magnétique. La TO n'apparait alors sans doute pas la méthode unique à privilégier pour l'analyse des séries temporelles. Un couplage entre ces deux techniques est donc une piste prometteuse pour une détection et une caractérisation optimale des pulsations géomagnétiques.

Interprétation de sondages MT

Les études menées dans ce manuscrit ont menés à deux interprétations importantes, en France sur le site de l'observatoire magnétique National de Chambon-La-Forêt, et au Népal, dans une région actuellement active sismiquement.

D'abord, l'analyse des séries temporelles et des sondages MT au niveau du site de l'observatoire magnétique National CLF, montre un niveau de bruit anthropique surprenant. Le bruit observé est principalement haute-fréquence, il apparait peu probable qu'aucune incidence n'existe sur les données de l'observatoire magnétique. Le placement de l'observatoire est cependant idéal sur une telle distribution de résistivité ne présentant que très peu de gradients latéraux de conductivité. En effet, à cette localisation, les phénomènes d'induction magnétique sont limités et les variations du champ magnétique enregistrées sont représentatives des variations du champ principal et des champs externes avec des effets très réduits des phénomènes d'inductions dans la croûte. Il est à noter qu'au niveau de la composante verticale, une légère distorsion apparait. D'après les premières analyses, cette distortion proviendrait d'une large structure correspondant à l'anomalie magnétique du bassin parisien.
Pour parfaitement décrire la situation de l'observatoire magnétique, il serait ainsi nécessaire de poursuivre cette analyse. Ainsi, en déployant des stations MT sur une zone plus importante, nous pourrions accéder à plus d'informations sur la structure 2D observée dans les sondages MT. La comparaison des éventuels nouveaux résultats au grand nombre d'études disponibles en méthodes potentielles sur l'analyse du bassin Parisien permettrait de confirmer l'origine de la distortion que nous observons et éventuellement d'en préciser les contours géologiques.

Ensuite, l'analyse des données de la campagne d'acquisition MT au Népal montre que la zone demeure un object géophysique complexe. Les analyses de dimensionnalité montrent qu'une simple analyse 2D reste insuffisante pour parfaitement décrire le comportement des sondages MT. Seule une campagne 3D, couplée à des codes permettant la modélisation d'une topographie précise (reliefs très accidentés au Népal) permettront de déterminer la distribution de conductivité dans le sous-sol. Nous pourrions ainsi considérablement contribuer à l'amélioration des connaissances sur les processus géologiques de cette zone.
Cependant, une telle campagne 3D se heurte à un problème logistique compliqué, considérant la distribution des accès routiers dans la région. Une grille de stations MT parfaitement régulière telle que celle du projet Earthscope aux Etats-Unis d'Amérique est inenvisageable. Les schémas d'inversion doivent donc également permettre la prise en compte de ce type d'échantillonnage irrégulier des stations.
Nous avons néanmoins pu exploiter les données récoltées durant la campagne de mesures de Mars 2016. Nous avons ainsi montré que, la simple incorportation de la topographie dans les modèles, changeait drastiquement l'interprétation obtenue après les processus d'inversion. Ainsi, des structures factices peuvent apparaître aussi bien à faible qu'à forte profondeur. L'inversion 2D des nouvelles données MT récoltées a permis d'ajuster, données et résultat de l'inversion. Le modèle obtenu, réactualisation de celui de 1996, montre une distribution de conductivité totalement différente de l'ancien modèle.

Bibliographie

Abdelfettah, Y. et al. (2014). "Preliminary magnetotelluric monitoring results at Rittershoffen". In : *3rd European Geothermal Workshop, Karlsruhe, 15-16 October 2014*.

Abramowitz, M. et R. Stegun (1964). *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. Dover Publications.

Adhikari, LB et al. (2015). "The aftershock sequence of the 2015 April 25 Gorkha–Nepal earthquake". In : *Geophysical Journal International* 203.3, p. 2119–2124.

Amante, Christopher et Barry W Eakins (2009). *ETOPO1 1 arc-minute global relief model : procedures, data sources and analysis*. US Department of Commerce, National Oceanic et al.

Antoine, J-P., R. Murenzi et P. Vandergheynst (1999). "Directional wavelet revisited : Cauchy wavelets and symmetry detection in patterns". In : *Applied and computational harmonic analysis* 6, p. 314–345.

Autran, A. et al. (1986). "Nouvelles contraintes géophysiques et géodynamiques pour l'interprétation de l'anomalie magnétique du bassin de Paris : hypothèse d'un rift paléozoique refermé au Carbonifere". In : *Bull. Soc. géol. Fr* 8, p. 125–141.

Avouac, Jean-Philippe et al. (2015). "Lower edge of locked Main Himalayan Thrust unzipped by the 2015 Gorkha earthquake". In : *Nature Geoscience*.

Axford, WI (1969). "Magnetospheric convection". In : *Reviews of Geophysics* 7.1-2, p. 421–459.

Barkhausen, Heinrich (1930). "Whistling tones from the Earth". In : *Proceedings of the Institute of Radio Engineers* 18.7, p. 1155–1159.

Bayer, R, A Guillen et C Weber (1987). "Investigation par forage profond de l'anomalie magnétique du Bassin de Paris (région de Sancerre) : modélisation préalable à partir des données magnétiques et gravimétriques". In : *Comptes rendus de l'Académie des sciences. Série 2, Mécanique, Physique, Chimie, Sciences de l'univers, Sciences de la Terre* 305.3, p. 203–208.

Benioff, Hugo (1960). "Observations of geomagnetic fluctuations in the period range 0.3 to 120 seconds". In : *Journal of Geophysical Research* 65.5, p. 1413–1422.

Boccippio, Dennis J et al. (1995). "Sprites, ELF transients, and positive ground strokes". In : *Science* 269.5227, p. 1088.

Bois, Ch et al. (1986). "Deep seismic profiling of the crust in northern France : The ECORS project". In : *Reflection seismology : A global perspective*, p. 21–29.

Bollinger, L et al. (2004). "Thermal structure and exhumation history of the Lesser Himalaya in central Nepal". In : *Tectonics* 23.5.

Bostick, F. X. (1977). "A simple almost exact method of MT analysis". In : *Workshop on electrical Methods in Geothermal Exploration, U. S. Geological Survey, Contract No. 14080001-8-359.*

Butler, Karl E et R Don Russell (1993). "Subtraction of powerline harmonics from geophysical records". In : *Geophysics* 58.6, p. 898–903.

Cagniard, L. (1953). "Basic theory of the magnetotelluric method of geophysical prospecting". In : *Geophysics* 18, p. 605–635.

Caldwell, T Grant, Hugh M Bibby et Colin Brown (2004). "The magnetotelluric phase tensor". In : *Geophysical Journal International* 158.2, p. 457–469.

Cantwell, T. (1960). "Detection and analysis of low frequency magnetotelluric signals". Thèse de doct. Massachusetts Institute of Technology.

Carpenter, DL et al. (1981). "A comparison of equatorial electron densities measured by whistlers and by a satellite radio technique". In : *Geophysical Research Letters* 8.10, p. 1107.

Chand, R, M Israil et J Rai (2009). "Schumann resonance frequency variations observed in magnetotelluric data recorded from Garhwal Himalayan region India". In : *Annales geophysicae : atmospheres, hydrospheres and space sciences.* T. 27. 9, p. 3497.

Chapman, Sydney et Vincent CA Ferraro (1931). "A new theory of magnetic storms". In : *Terrestrial Magnetism and Atmospheric Electricity* 36.2, p. 77–97.

Chave, A. D. et A. Jones (2012). *The Magnetotelluric method : Theory and Practice.* Cambridge : Cambridge University Press.

Chave, Alan D (2014). "Magnetotelluric data, stable distributions and impropriety : an existential combination". In : *Geophysical Journal International*, ggu121.

Chave, Alan D et Pamela Lezaeta (2007). "The statistical distribution of magnetotelluric apparent resistivity and phase". In : *Geophysical Journal International* 171.1, p. 127–132.

Chave, Alan D et J Torquil Smith (1994). "On electric and magnetic galvanic distortion tensor decompositions". In : *Journal of Geophysical Research : Solid Earth* 99.B3, p. 4669–4682.

Chave, Alan D. et David J. Thomson (1987). "On the Robust Estimation of Power Spectra, Coherences, and Transfer Functions". In : *Journal of Geophysical Research* 92.B1, p. 633–648.

— (1989). "Some comments on magnetotelluric response function estimation". In : *J. geophys. Res* 94.14, p. 215–14.

— (2004). "Bounded influence magnetotelluric response function estimation". In : *Geophysical Journal International* 157, p. 988–1006.

CHEN, Liu et Akira HASEGAWA (1974). "A theory of long-period magnetic pulsations : 1. Steady state excitation of field line resonance". In : *Journal of Geophysical Research* 79.7, p. 1024–1032.

CHRISSAN, DA et AC FRASER-SMITH (1996). *Seasonal variations of globally measured ELF/VLF radio noise*. Rapp. tech. DTIC Document.

COMMITTEE, Intermagnet Operations et Executive COUNCIL (2012). "Intermagnet technical reference manual". In :

CONSTABLE, Catherine (2016). "Earth's Electromagnetic Environment". In : *Surveys in Geophysics* 37.1, p. 27–45.

CONSTABLE, Steven C, Robert L PARKER et Catherine G CONSTABLE (1987). "Occam's inversion : A practical algorithm for generating smooth models from electromagnetic sounding data". In : *Geophysics* 52.3, p. 289–300.

COWLEY, Stanley WH (1995). "The Earth's magnetosphere : A brief beginner's guide". In : *Eos, Transactions American Geophysical Union* 76.51, p. 525–529.

— (2007). "Magnetosphere of the Earth". In : *Encyclopedia of Geomagnetism and Paleomagnetism*. Springer, p. 656–664.

DANIELS, Richard W (1974). *Approximation methods for electronic filter design : with applications to passive, active, and digital networks*. McGraw-Hill Companies.

DEBEGLIA, N. (2005). "Réinterprétation géophysiques dans le centre du bassin de Paris pour l'évaluation du potentiel géothermique (COPGEN 2005)". In : *BRGM/RP-54248-FR*, p. 55.

DIALLO, M. et al. (2006). "Characterization of polarization attributes of seismic waves using con- tinuous wavelet transforms". In : *Geophysics* 71, p. 67–77.

DU, Hua-Kun, Zheng-Yong REN et Jing-Tian TANG (2016). "A finite-volume approach for 2D magnetotellurics modeling with arbitrary topographies". In : *Studia Geophysica et Geodaetica* 60.2, p. 332–347.

DUNGEY, James W (1961). "Interplanetary magnetic field and the auroral zones". In : *Physical Review Letters* 6.2, p. 47.

DUPIS, A. et al. (1990). "Contribution de la magnéto-tellurique et de la sismique à l'étude de l'anomalie magnétique du Bassin parisien". In : *Bull. Soc. géol. Fl* 990, p. 5.

DZIEWONSKI, Adam M et Don L ANDERSON (1981). "Preliminary reference Earth model". In : *Physics of the earth and planetary interiors* 25.4, p. 297–356.

EASTWOOD, JP et al. (2011). "Transient Pc3 wave activity generated by a hot flow anomaly : Cluster, Rosetta, and ground-based observations". In : *Journal of Geophysical Research : Space Physics* 116.A8.

EFRON, Bradley et Robert J TIBSHIRANI (1994). *An introduction to the bootstrap*. CRC press.

EGBERT, G. D. (1997). "Robust multiple-station magnetotelluric data processing". In : *Geophysical Journal International* 130, p. 475–496.

EGBERT, G. D. et J. R. BOOKER (1986). "Robust estimation of geomagnetic transfer functions". In : *Geophysical Journal International* 87.1, p. 173–194.

EGBERT, Gary D et Anna KELBERT (2012). "Computational recipes for electromagnetic inverse problems". In : *Geophysical Journal International* 189.1, p. 251–267.

ESCALAS, M. et al. (2013). "Polarization analysis of magnetotelluric time series using a wavelet-based scheme : A method for detection and characterisation of cultural noise sources". In : *Physics of the Earth and Planetary Interiors* 218, p. 31–50.

FARR, Tom G et al. (2007). "The shuttle radar topography mission". In : *Reviews of geophysics* 45.2.

FEJER, JA (1964). "Hydromagnetic stability at a fluid velocity discontinuity between compressible fluids". In : *Physics of Fluids (1958-1988)* 7.4, p. 499–503.

FIELD, Christopher A et Alan H WELSH (2007). "Bootstrapping clustered data". In : *Journal of the Royal Statistical Society : Series B (Statistical Methodology)* 69.3, p. 369–390.

FOWLER, RA, BJ KOTICK et RD ELLIOTT (1967). "Polarization analysis of natural and artificially induced geomagnetic micropulsations". In : *Journal of Geophysical Research* 72.11, p. 2871–2883.

GAILLOT, P. (2000). "Ondelettes continues en Science de la Terre - Méthodes et applications". Thèse de doct. Université de Toulouse III.

GAMBLE, T. D., W. M. GOUBEAU et J. CLARKE (1979). "Magnetotellurics with a remote magnetic reference". In : *Geophysics* 44.1, p. 53–68.

GARCIA, X et A. G. JONES (2002). "Atmospheric sources for audio-magnetotelluric (AMT) sounding". In : *Geophysics* 67.2, p. 448–458.

GARCIA, X. et A. G. JONES (2008). "Robust processing of magnetotelluric data in the AMT dead band using the continuous wavelet transform". In : *Geophysics* 73.6, p. 223–234.

GERARD, A et C WEBER (1971). "L'anomalie magnétique du Bassin de Paris interprétée comme élément structural majeur dans l'histoire géologique de la France". In : *CR Acad. Sci. Paris, Série D* 272, p. 921–923.

GIGOT, C (1984). "Notice explicative, carte géologique de France à 1/5 0 000, feuille Bellegarde-du-Loiret (n 364)". In : *Bureau de Recherches Géologiques et Minières, Orléans.*

GIRAULT, Frédéric et Frédéric PERRIER (2014). "The Syabru-Bensi hydrothermal system in central Nepal : 2. Modeling and significance of the radon signature". In : *Journal of Geophysical Research : Solid Earth* 119.5, p. 4056–4089.

GOUPILLAUD, P., A. GROSSMANN et J. MORLET (1984). "Cycle-octave and related transforms in seismic signal analysis". In : *Seismic signal analysis and discrimination III* 23.1, p. 85–102.

GROOM, Ross W et Richard C BAILEY (1989). "Decomposition of magnetotelluric impedance tensors in the presence of local three-dimensional galvanic distortion". In : *Journal of Geophysical Research : Solid Earth* 94.B2, p. 1913–1925.

GUBBINS, David et Emilio HERRERO-BERVERA (2007). *Encyclopedia of geomagnetism and paleomagnetism.* Springer Science & Business Media.

HE, L. et al. (2009). "Fast connected-component labeling". In : *Pattern Recognition* 42, p. 1977–1987.

HELLIWELL, Robert A (1965). *Whistlers and related ionospheric phenomena.* T. 1. 1. Stanford University Press Stanford.

HOLSCHNEIDER, M. (1995). *Wavelets : An Analysis Tool.* Oxford : Oxford Science Publications.

HUBBARD, Judith et al. (2016). "Structural segmentation controlled the 2015 Mw 7.8 Gorkha earthquake rupture in Nepal". In : *Geology* 44.8, p. 639–642.

HUBER, Peter J (1981). "Robust Statistics". In :

HUGHES, H. G. et R. A. PAPERT (1975). "Propagation prediction model selection using VLF atmospherics". In : *Geophysical Research Letters* 2, p. 96–98.

HURSAN, Gabor et Michael S ZHDANOV (2002). "Contraction integral equation method in three-dimensional electromagnetic modeling". In : *Radio Science* 37.6.

HWANG, K-J et al. (2012). "The first in situ observation of Kelvin-Helmholtz waves at high-latitude magnetopause during strongly dawnward interplanetary magnetic field conditions". In : *Journal of Geophysical Research : Space Physics* 117.A8.

JACOBS, J. A. et al. (1964). "Classification of geomagnetic micropulsations". In : *Journal of Geophysical Research* 69.1, p. 180–181.

JACOBS, JA et K SINNO (1960). "World-wide characteristics of geomagnetic micropulsations". In : *Geophysical Journal International* 3.3, p. 333–353.

JIRACEK, George R (1990). "Near-surface and topographic distortions in electromagnetic induction". In : *Surveys in Geophysics* 11.2-3, p. 163–203.

JONES, A. G. (1983). "On the equivalence of the "Niblett" and "Bostick" transformations in the magnetotelluric method". In : *Journal of Geophysics* 53, p. 72–73.

JONES, A. G. et al. (1989). "A comparison of techniques for magnetotelluric response function estimation". In : *Journal of Geophysical Research : Solid Earth* 94.B10, p. 14201–14213.

KANGAS, Jorma, A GUGLIELMI et O POKHOTELOV (1998). "Morphology and physics of short-period magnetic pulsations". In : *Space Science Reviews* 83.3-4, p. 435–512.

KATO, Y (1962). "Geomagnetic micropulsations". In : *Australian Journal of Physics* 15, p. 70.

KATO, Y et T SAITO (1962). "Morphological study of geomagnetic pulsations". In : *Journal of the Physical Society of Japan Supplement* 17, p. 34.

KAVOSI, Shiva et Joachim RAEDER (2015). "Ubiquity of Kelvin-Helmholtz waves at Earth/'s magnetopause". In : *Nature communications* 6.

KAWAMURA, M (1969a). "Behavior of the Horizontal Disturbing Vector of the Pearl-Type Micropulsation in Lower Latitude". In : *National ELF and ULF Symposium Held at Toyokawa, Japan, on Jan.* T. 26.

— (1969b). "On the Diurnal Variation in the Pc 3 Period at Kakioka". In : *National ELF and ULF Symposium Held at Toyokawa, Japan, on Jan.* T. 26.

KELBERT, Anna et al. (2014). "ModEM : a modular system for inversion of electromagnetic geophysical data". In : *Computers & Geosciences* 66, p. 40–53.

KEPKO, L, Harlan E SPENCE et HJ SINGER (2002). "ULF waves in the solar wind as direct drivers of magnetospheric pulsations". In : *Geophysical research letters* 29.8.

KEY, Kerry (2016). "MARE2DEM : a 2-D inversion code for controlled-source electromagnetic and magnetotelluric data". In : *Geophysical Journal International* 207.1, p. 571–588.

KEY, Kerry et Jeffrey OVALL (2011). "A parallel goal-oriented adaptive finite element method for 2.5-D electromagnetic modelling". In : *Geophysical Journal International* 186.1, p. 137–154.

KULESH, M. et al. (2007). "Polarization analysis of a Pi2 pulsation using continuous wavelet transform". In : *Earth Planet and Space* 59, p. 961–970.

LARNIER, Hugo, Pascal SAILHAC et Aude CHAMBODUT (2016). "New application of wavelets in magnetotelluric data processing : reducing impedance bias". In : *Earth, Planets and Space* 68.1, p. 1–11.

LE PAPE, Florian et al. (2012). "Penetration of crustal melt beyond the Kunlun Fault into northern Tibet". In : *Nature Geoscience* 5.5, p. 330–335.

LEE, LC et YC KWOK (1984). "A mechanism for the IPDP pulsations". In : *Journal of Geophysical Research : Space Physics* 89.A2, p. 877–882.

LEMONNIER, C. (2000). "Imagerie Electromagnétique de la Convergence Himalayenne au Népal". Thèse de doct. Université Strasbourg I Louis Pasteur.

LEMONNIER, Carole et al. (1999). "Electrical structure of the Himalaya of Central Nepal : high conductivity around the mid-crustal ramp along the MHT". In : *Geophysical Research Letters* 26.21, p. 3261–3264.

LETORT, J. et al. (2016). "Teleseismic depth estimation of the 2015 Gorkha- Nepal aftershocks". In : *Geophysical Journal International* 207.3, p. 1584–1595.

LI, Yuguo et Josef PEK (2008). "Adaptive finite element modelling of two-dimensional magnetotelluric fields in general anisotropic media". In : *Geophysical Journal International* 175.3, p. 942–954.

LORENZ, Claude (1980). *Géologie des pays européens : France, Belgique, Luxembourg.* T. 1. Dunod.

LOVE, Jeffrey J et Arnaud CHULLIAT (2013). "An international network of magnetic observatories". In : *Eos, Transactions American Geophysical Union* 94.42, p. 373–374.

MACKAY, C. et A. C FRASER-SMITH (2010). "Lightning location using the slow tails of sferics". In : *Radio Science* 45.

MACKIE, Randall L, J Torquil SMITH et Theodore R MADDEN (1994). "Three-dimensional electromagnetic modeling using finite difference equations : The magnetotelluric example". In : *Radio Science* 29.4, p. 923–935.

MADDEN, T et W THOMPSON (1965). "Low-frequency electromagnetic oscillations of the earth-ionosphere cavity". In : *Reviews of Geophysics* 3.2, p. 211–254.

MARAUN, Douglas, J KURTHS et M HOLSCHNEIDER (2007). "Nonstationary Gaussian processes in wavelet domain : synthesis, estimation, and significance testing". In : *Physical Review E* 75.1, p. 016707.

MARTELET, Guillaume et al. (2013). "Fast mapping of magnetic basement depth, structure and nature using aeromagnetic and gravity data : combined methods and their application in the Paris Basin". In : *Geophysical prospecting* 61.4, p. 857–873.

MATSUSHITA, S (1967). "Physics of geomagnetic phenomena". In :

MAXWELL, J. C. (1865). "A Dynamical Theory of the Electromagnetic Field". In : *Royal Society Transactions* 155, p. 459–512.

MCNEICE, Gary W et Alan G JONES (2001). "Multisite, multifrequency tensor decomposition of magnetotelluric data". In : *Geophysics* 66.1, p. 158–173.

MCPHERSON, Robert L. (2005). "Magnetic Pulsations : Their Sources and relation to solar wind and geomagnetic activity". In : *Surveys in Geophysics* 26.5, p. 545–592.

MISITI, M. et al. (2007). *Wavelets and their applications.* ISTE Ltd.

MOREAU, F. (1995). "Méthodes de traitement de données géophysique par transformée en ondelettes". Thèse de doct. Université de Rennes I.

NIBLETT, E. R. et V. SAYN-WITTGENSTEIN (1960). "Variation of electrical conductivity with depths by the magnetotelluric method". In : *Geophysics* 25, p. 998–1008.

OGAWA, Toshio et Masayuki KOMATSU (2007). "Analysis of Q burst waveforms". In : *Radio Science* 42.2.

OGAWA, Toshio, Yoshikazu TANAKA et al. (1966). "Observations of natural ELF and VLF electromagnetic noises by using ball antennas". In : *Journal of geomagnetism and geoelectricity* 18.4, p. 443–454.

OLSEN, N. (2007). "Natural sources for electromagnetic induction studies". In : *In : Gubbins D., Herrero-Bervera E.(eds) Encyclopedia of Geomagnetism and Paleomagnetism. Springer*, p. 696–700.

PARKER, Eugene N (1958). "Dynamics of the Interplanetary Gas and Magnetic Fields". In : *The Astrophysical Journal* 128, p. 664.

PARKINSON, W. D. (1959). "Directions of Rapid geomagnetic fluctuations". In : *Geophysical Journal of the Royal Astronomical Society* 2.1, p. 1–14.

— (1962). "The influence of continents and oceans on geomagnetic variations". In : *Geophysical Journal of the Royal Astronomical Society* 6.4, p. 441–449.

PÉREZ, J., R. CARLES et R. FLECKINGER (2009). *Electromagnétisme : Fondements et applications, 4ème édition.* Paris : Dunot.

PERRIER, F. et al. (2014). "Permanent electric dipoles in the magnetic observatory of Chambon-la-Forêt, France : preliminary results." In : *EMIW 2014 Workshop - Extended abstract.*

PETIAU, GILBERT (2000). "Second generation of lead-lead chloride electrodes for geophysical applications". In : *Pure and Applied Geophysics* 157.3, p. 357–382.

PHAM, V. N., D. BOYER et J.-L. LE MOUËL (2000). "Nouveaux arguments sur l'origine de l'anomalie magnétique du Bassin parisien (AMBP) d'après les propriétés électriques de la croûte". In : *Comptes Rendus de l'Académie des Sciences-Series IIA-Earth and Planetary Science* 331.7, p. 443–449.

PILIPENKO, V et al. (2010). "Generation of magnetic and particle Pc5 pulsations during the recovery phase of strong magnetic storms". In : *Proceedings of the Royal Society of London A : Mathematical, Physical and Engineering Sciences.* T. 466. 2123. The Royal Society, p. 3363–3390.

RAKOV, Vladimir A et Martin A UMAN (2003). *Lightning : physics and effects.* Cambridge University Press.

RASSON, Jean L (2007). "Observatories, instrumentation". In : *Encyclopedia of Geomagnetism and Paleomagnetism*, p. 711–713.

REUVENI, Yuval et al. (2010). "Natural atmospheric noise statistics from VLF measurements in the eastern Mediterranean". In : *Radio Science* 45.5.

RICE, John (2006). *Mathematical statistics and data analysis.* Nelson Education.

RICHMOND, Arthur D (2007). "Ionosphere". In : *Encyclopedia of Geomagnetism and Paleomagnetism*, p. 452–454.

RIKITAKE, T. (1948). "Notes on electromagnetic induction within the Earth". In : *Bulletin of Earthquake research institute* 24, p. 1–9.

ROUYER, T. et al. (2008). "Analysing multiple time series and extending significance testing in wavelet analysis". In : *Marine Ecology Progress Series* 359, p. 11–23.

ROYER, JY et P PATRIAT (2002). "L'inde part à la dérive". In : *Himalaya-Tibet, Le choc des continents*, p. 25–31.

SAILHAC, P. (1999). "Analyse multiéchelle et inversion de données géophysiques en Guyane Française". Thèse de doct. Institut de Physique du Globe de Paris.

SAILHAC, P. et al. (2016). "Processing and modeling magnetotelluric data at La Soufrière of Guadeloupe lava dome (France)". In : *EMIW 2016 Chiang Mai Extended Abstract.*

SAITO, T et S MATSUSHITA (1967). "Variations in the Period of Geomagnetic Pulsation Pc4 during One Recent Solar Cycle'". In : *Trans. Amer. Geophys. Union* 48, p. 70.

SAITO, Takao (1964). "Mechanisms of geomagnetic continuous pulsations and physical states of the exosphere". In : *Journal of geomagnetism and geoelectricity* 16.2, p. 115–151.

— (1969). "Geomagnetic Pulsations". In : *Space Science Reviews* 10, p. 319–412.

SAMSON, JC, JA JACOBS et G ROSTOKER (1971). "Latitude-dependent characteristics of long-period geomagnetic micropulsations". In : *Journal of Geophysical Research* 76.16, p. 3675–3683.

SIMS, W. E., F. X. Jr. BOSTICK et H. W. SMITH (1971). "The estimation of magnetotelluric impedance tensor elements from measured data". In : *Geophysics* 36.5, p. 938–942.

SIRIPUNVARAPORN, Weerachai, Gary EGBERT et Yongwimon LENBURY (2002). "Numerical accuracy of magnetotelluric modeling : a comparison of finite difference approximations". In : *Earth, planets and space* 54.6, p. 721–725.

SLAWINSKI, R et al. (1988). "Transmission of solar wind hydromagnetic energy into the terrestrial magnetosphere". In : *Geophysical research letters* 15.11, p. 1275–1278.

SLEPIAN, David (1978). "Prolate spheroidal wave functions, Fourier analysis, and uncertainty-V : The discrete case". In : *Bell System Technical Journal* 57.5, p. 1371–1430.

SMITH, J Torquil et John R BOOKER (1991). "Rapid inversion of two-and three-dimensional magnetotelluric data". In : *Journal of Geophysical Research : Solid Earth* 96.B3, p. 3905–3922.

SOUTHWOOD, DJ et WJ HUGHES (1983). "Theory of hydromagnetic waves in the magnetosphere". In : *Space Science Reviews* 35.4, p. 301–366.

SPRATT, Jessica E et al. (2005). "Crustal structure of the India–Asia collision zone, southern Tibet, from INDEPTH MT investigations". In : *Physics of the Earth and Planetary Interiors* 150.1, p. 227–237.

STEWART, B. (1861). "On the Great Magnetic Disturbance Which Extended from August 28 to September 7, 1859, as Recorded by Photography at the Kew Observatory". In : *Philosophical Transactions of the Royal Society of London* 151, p. 423–430.

STRANGWAY, DW, CM SWIFT JR et RC HOLMER (1973). "The application of audio-frequency magnetotellurics (AMT) to mineral exploration". In : *Geophysics* 38.6, p. 1159–1175.

SWIFT, C. M. (1967). "A magnetotelluric investigation of an electrical conductivity anomaly in the south-western United States". Thèse de doct. Massachusetts Institute of Technology.

THÉBAULT, E, M MANDEA et JJ SCHOTT (2006). "Modeling the lithospheric magnetic field over France by means of revised spherical cap harmonic analysis (R-SCHA)". In : *Journal of Geophysical Research : Solid Earth* 111.B5.

THOMSON, David J (1982). "Spectrum estimation and harmonic analysis". In : *Proceedings of the IEEE* 70.9, p. 1055–1096.

TIKHONOV, A. N. (1950). "On determining electrical characteristics of deep layers of the Earth's crust". In : *Doklady* 73, p. 295–297.

TING, Sam C et Gerald W HOHMANN (1981). "Integral equation modeling of three-dimensional magnetotelluric response". In : *Geophysics* 46.2, p. 182–197.

TORRENCE, C. et G. P. COMPO (1998). "A practical guide to wavelet analysis". In : *Bulletin of the American Meteorological Society* 79.1, p. 61–78.

TROITSKAYA, VA (1967). "Micropulsations and the state of the magnetosphere". In : *Solar-Terrestrial Physics*, p. 213.

TZANIS, A et D BEAMISH (1987). "Audiomagnetotelluric sounding using the Schumann resonances". In : *Journal of Geophysics* 61, p. 97–109.

UNSWORTH, MJ et al. (2005). "Crustal rheology of the Himalaya and Southern Tibet inferred from magnetotelluric data". In : *Nature* 438.7064, p. 78–81.

VELLANTE, M et al. (2004). "Ground/satellite signatures of field line resonance : A test of theoretical predictions". In : *Journal of Geophysical Research : Space Physics* 109.A6.

WAIT, J. R. (1954). "On the relation between telluric currents and the Earth's magnetic field". In : *Geophysics* 19, p. 281–289.

WAIT, J. R (1960a). "Mode theory and the propagation of ELF radio waves". In : *J. Res. Natl. Bur. Stand. D* 64.4, p. 387–404.

— (1960b). "Terrestrial propagation of very-low-frequency radio waves, a theoretical investigation". In : *J. Res. Nat. Bureau Stand* 64, p. 153–204.

WIESE, H (1962). "Geomagnetische Tiefentellurik Teil II : die Streichrichtung der Untergrundstrukturen des elektrischen Widerstandes, erschlossen aus geomagnetischen Variationen". In : *Geofisica pura e applicata* 52.1, p. 83–103.

WILLIAMS, E. R. (1992). "The Schumann resonance : a global tropical thermometer". In : *Science* 256, p. 1184–1187.

WINCH, Denis (2007). "Periodic External fields". In : *Encyclopedia of Geomagnetism and Paleomagnetism.* Springer, p. 809–816.

YEOMAN, Tim K et DM WRIGHT (2001). "ULF waves with drift resonance and drift-bounce resonance energy sources as observed in artificially-induced HF radar backscatter". In : *Annales Geophysicae.*

YUMOTO, Kiyohumi et Takao SAITO (1980). "Hydromagnetic waves driven by velocity shear instability in the magnetospheric boundary layer". In : *Planetary and Space Science* 28.8, p. 789–798.

ZHANG, Yuanchou et K. V. PAULSON (1997). "Enhancement of Signal-to-noise Ratio in Natural-source Transient Magnetotelluric Data with Wavelet Transform". In : *Pure and Applied Geophysics* 149, p. 405–419.

ZLOTNICKI, J. et al. (2005). "Automatic fuzzy-logic recognition of anomalous activity on long geophysical records : Application to electric signals associated with the volcanic activity of La Fournaise volcano (Reunion Island)". In : *Earth and Planetary Science Letters* 234, p. 261–278.

Annexe A

Données utilisées et traitements appliqués

A.1 Données d'observatoires

A.1.1 Présentation

Les observatoires magnétiques sont des stations permettant l'enregistrement continu du champ magnétique terrestre en un point sur de longues périodes de temps. Elles permettent notamment l'étude des variations lentes du champ magnétique et ainsi une meilleure compréhension des phénomènes associés.

La distribution géographique des observatoires est aléatoire et dépend principalement des moyens financiers alloués aux instituts responsables et à la présence de terrains permettant de telles installations. Ainsi seules les terres émergées permettent l'accueil d'observatoires magnétiques laissant de grandes zones géographiques non couvertes en données magnétiques de surface (ce qui est contrebalancé par des données satellitaires ayant une excellente répartition spatiale mais pauvre en temps). Une autre condition nécessaire à l'installation d'un observatoire est une faible activité anthropique que ce soit de la part de l'observatoire ou du milieu environnant, ainsi qu'un faible gradient spatial du champ magnétique sur la zone de l'observatoire (Rasson, 2007).

Actuellement, environ 180 observatoires sont en activité dans le monde, avec de grandes disparités géographiques. Une grande densité d'observatoires sont disponibles en Europe, moins dans le continent Nord Américain et en Asie, et peu en Afrique, Amérique du Sud, Russie.

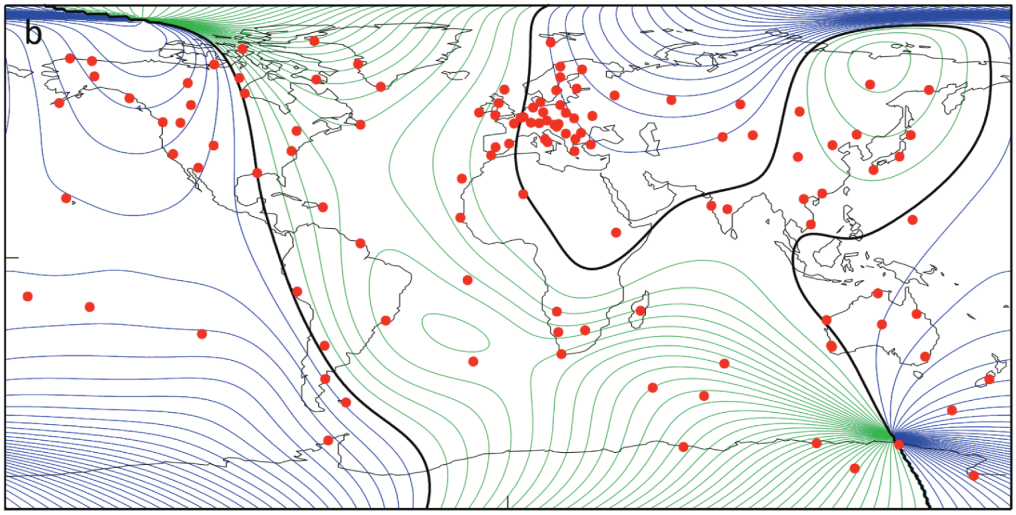


FIGURE A.1: Carte des observatoires magnétiques dans le monde en 2013. La déclinaison magnétique est superposée à la carte. D'après LOVE et CHULLIAT, 2013.
FIGURE A.1: *Map of worldwide magnetic observatories in 2013. Magnetic declination is overlaid on the map. From* LOVE *et* CHULLIAT *(2013)*

A.1.2 Matériel

Les observatoires appartenant au réseau INTERMAGNET doivent suivre des standards décrits dans le manuel de référence technique (COMMITTEE et COUNCIL, 2012). Le matériel permettant l'acquisition de la valeur absolue du champ magnétique ne peuvent actuellement effectuer des mesures à haute fréquence. Deux types d'instruments sont alors utilisés pour obtenir la valeur du champ magnétique :

- Des variomètres enregistrant les variations continues du champ magnétique à la fréquence d'acquisition requise. Ces instruments sont cependant soumis à des dérives à long terme ainsi qu'à des erreurs de mesure (sauts, etc...), et doivent être corrigés par des mesures absolues. Les variomètres sont actuellement principalement des capteurs de type *fluxgate* (RASSON, 2007).
- Les mesures absolues du champ magnétique sont effectuées à intervalles réguliers par des instruments du type DI-Flux (pour Déclinaison, Inclinaison et Fluxgate). Ils permettent la détermination de la déclinaison et de l'inclinaison du champ magnétique, mais un opérateur humain est nécessaire pour cette opération. Des appareils de mesure automatique de ces angles sont actuellement en développement ou disponibles et pourront permettre l'instrumentation de sites où la présence continue de personnel humain est compliquée comme le continent Antarctique.

Les deux types de données sont ensuite fusionnés pour aboutir aux valeurs absolues du champ magnétique terrestre. Les mesures absolues sont utilisées dans le calcul d'une ligne de calibration, appelée *baseline*, qui permet la livraison de données dites "définitives".

A.1.3 Type de données

Les données d'observatoire du réseau INTERMAGNET sont délivrées sous la forme de fichier du type IAGA-2002. Les données passent par quatre stades de traitement :

- Reported : Données brutes issues de l'observatoire.
- Adjusted : Données du type *reported* sur lesquelles des traitements ont été appliquées du type correction de décalages temporels, suppression de données aberrantes.
- Quasi-définitive : Données du type *adjusted* corrigées à l'aide de baselines provisoires.
- Définitive : Données finales délivrées par l'institut responsable de l'observatoire magnétique. Le manuel technique d'INTERMAGNET précise que la différence entre les données quasi-définitives et définitives doit être inférieure à 2%.

Les données finales selon le matériel utilisé à l'observatoire sont soit de type minute ou seconde.

L'utilisation des données d'observatoire dans les procédures de traitement MT ne nécessite pas d'avoir la valeur vraie du champ magnétique, mais seulement ses variations. Ainsi, seules les corrections de décalages temporels et le retrait des données aberrantes sont suffisants pour les développements méthodologiques présentés dans ce manuscrit.

A.2 Données MT

A.2.1 Matériel

Une mesure MT consiste en l'enregistrement simultané des composantes électriques et magnétiques horizontales avec si possible celui de la composante magnétique verticale. Un dispositif MT est composée d'une unité d'acquisition permettant l'enregistrement et la numérisation des signaux mesurés par les capteurs mesurant le champ électrique et magnétique. Cette acquisition est ensuite reliée à une batterie pour l'alimentation et à un GPS pour la synchronisation des mesures avec d'autres appareils si besoin.

A.2.1.1 Mesure du champ électrique

La mesure électrique se base sur le calcul de la différence de potentiel V_{AB} entre deux points A et B. La relation entre le potentiel V et le champ électrique étant $\mathbf{E} = -\nabla V$, V_{AB} est égal à :

$$V_{AB} = \int_B^A \mathbf{E} \cdot \mathbf{r} \tag{A.1}$$

Le potentiel ne dépendant que des points de départ et d'arrivée, si l est la distance en ligne droite entre A et B et en supposant que le champ électrique est uniforme entre les points A et

B et horizontal dans la direction AB, le potentiel est donc égal à :

$$V_{AB} = El \tag{A.2}$$

Pour mesurer le champ électrique horizontal, il est donc nécessaire d'installer deux dipôles horizontaux orthogonaux. La mesure de la différence de potentiel divisée par la longueur du dipôle permet de déterminer le champ électrique.

Deux types d'électrodes sont couramment utilisées, des électrodes simples comme en acquisition DC, ou des électrodes impolarisables. Les électrodes impolarisables sont notamment utilisées dans le cas où l'acquisition est suffisamment longue pour que des phénomènes de polarisation soient perceptibles et susceptibles d'affecter les valeurs de potentiels mesurées.

A.2.1.2 Mesure du champ magnétique

La mesure du champ magnétique en MT est principalement effectuées à l'aide de bobines d'induction. Le principe physique est celui de l'induction d'une force électromotrice dans un circuit selon la loi de Maxwell-Faraday. La bobine est composée de spires dont l'axe est parallèle à l'axe de la bobine. Une variation de champ magnétique créé une force électromotrice e dans les spires dont la fonction de calibration permettant de transformer e en champ d'induction est connue. Seule la variation du champ magnétique dans la direction de la bobine est connue, il est donc nécessaire d'orienter trois bobines de façon orthogonale pour obtenir la variation du champ magnétique dans l'espace.

La mesure brute délivre donc des mesures en Volts. Pour obtenir les séries temporelles utilisées dans le traitement MT, il faut calibrer les réponses brutes.
Pour obtenir les réponses magnétiques, il est nécessaire de multiplier les réponses en fréquence par une fonction de calibration propre à chaque sonde et qui caractérise sa réponse pour chaque fréquence. La fonction de calibration est exprimée en T/m.

Dans cette thèse, le matériel utilisé est celui de la société Metronix Geophysics. Deux types de sondes ont été utilisées, les MFS07e et MFS06e. La figure A.2 montre les fonctions de calibration associées à ces deux types de sondes. D'autres types de capteurs permettent d'obtenir la valeur du champ magnétique tels que les capteurs Fluxgate ou les *superconducting quantum interference device* (SQUID), mais n'ont pas été utilisés dans le cadre de cette thèse.

A.2.1.3 Procédures d'acquisition

Nos protocoles d'acquisition sont basés sur l'enregistrement de séries à la fréquence d'échantillonnage de 512 Hz pour obtenir des séries basses fréquences avec les sondes MFS06/7. Pour

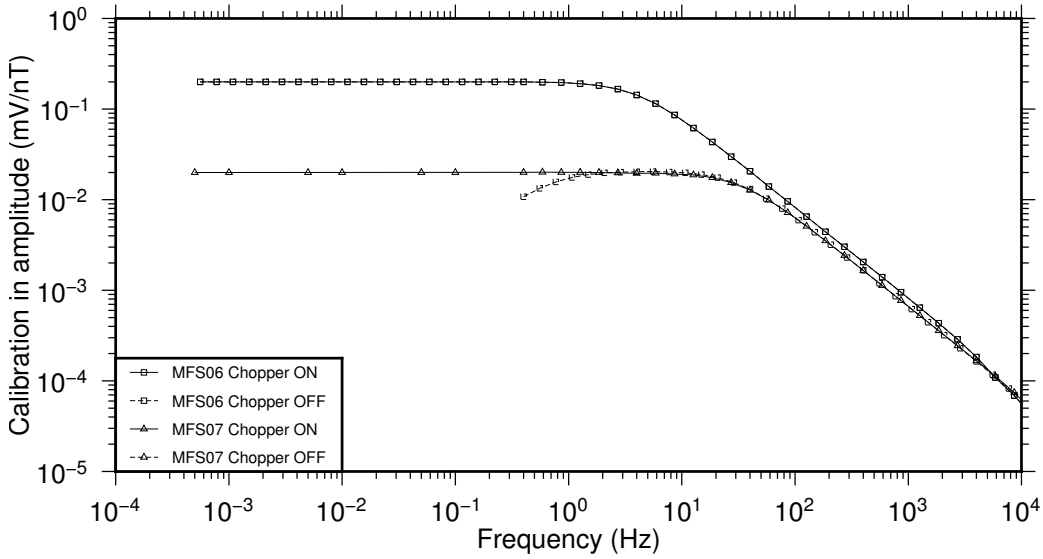


FIGURE A.2: Fonctions de calibration des sondes MFS06 et MFS07 pour leurs deux modes de fonctionnement.

FIGURE A.2: *MFS06e and MFS07e calibration function for both acquisition modes.*

obtenir des longues séries nécessaires au traitement des fréquences inférieures au Hertz, nous devons donc décimer les séries temporelles pour éviter de traiter des fichiers de taille trop importante (1 journée de données à 512 Hz représente plus de 44 millions d'échantillons). Pour décimer les fichiers, nous avons utilisés des procédures écrites dans les librairies Python existantes et basées sur l'application de filtre passe-bas (Daniels, 1974) avant décimation .

A.3 Filtrage de données MT

Après calibration, à part pour quelques exceptions, il est nécessaire de filtrer les séries temporelles pour réduire l'influence du bruit anthropique avant calcul des fonctions de réponse MT. Les principaux bruits à traiter ont été :

- Le réseau électrique domestique à 50 Hz et ses harmoniques.
- Le réseau ferré allemand à 50/3 Hz et ses harmoniques pour les stations installées en Alsace.
- Emetteurs VLF autour de 20 kHz pour obtenir les estimations MT supérieures à 10 kHz.

D'autres raies spectrales apparaissent dans certaines séries temporelles, mais les fréquences identifiées ne sont pas attribuées à des signaux connus.

Pour réduire ces signaux, deux principaux filtres ont été utilisés :

- Filtre de Butterworth pour les filtres passe-haut et passe bas utilisant les routines Python définies dans *Scipy*.

- Filtre de Butler et Russel (Butler et Russell, 1993) pour le retrait des réseaux électriques et de chemin de fer. Ce filtre découpe le signal en portions sur lesquelles il minimise l'écart entre une sinusoïde de fréquence celle à filtrer et la portion. Le principal écueil à éviter dans l'utilisation de ce filtre étant de filtrer les basses fréquences pour éviter des décalages entre les portions filtrées. L'implémentation en Python a été écrite à partir de Butler et Russell (1993).
- Filtre FIR pour les filtre coupe-bande utilisant les routines Python définies dans *Scipy*.

Annexe B

Transformée en ondelettes

Dans cette annexe, nous présentons les principales caractéristiques et quantités utilisées dans ce manuscrit. Pour de plus amples détails sur l'analyse en ondelettes avec des exemples, nous invitons le lecteur à se reporter aux références suivantes :

- M. Holschneider (1995) pour une description mathématique de l'analyse en ondelette.
- Moreau (1995), Gaillot (2000) et Escalas et al. (2013) pour des applications en géophysiques et contenant un grand nombre d'exemples d'analyse sur des signaux synthétiques.

B.1 Introduction sur le temps-fréquence

B.1.1 Représentations classiques

Les représentations classiques de signaux en géophysique sont principalement de deux types, temporelle et fréquentielle.

La représentation temporelle consiste en la décomposition d'un signal selon la base des distributions de Dirac δ.

$$f(t) = \int_{-\infty}^{\infty} d\tau f(\tau)\delta(t-\tau) \tag{B.1}$$

La représentation fréquentielle décompose le signal selon une base de fonctions sinusoïdales telles que :

$$F(\nu) = \int_{-\infty}^{\infty} d\tau f(\tau)e^{-i2\pi\nu\tau} \tag{B.2}$$

Ces deux représentations sont complémentaires et isométriques par le théorème de Parseval, qui décrit la conservation de l'énergie du signal E entre représentation temporelle et fréquentielle :

$$E = \int_{-\infty}^{+\infty} |f(t)|^2 dt = \int_{-\infty}^{+\infty} |F(\nu)|^2 d\nu \tag{B.3}$$

La représentation temporelle permet de décrire la distribution temporelle de l'énergie du signal, la résolution fréquentielle permettant de décrire la distribution fréquentielle de l'énergie du signal.

En revanche, aucune de ces deux représentations n'est capable de donner d'information en temps et fréquence simultanément. Une façon de créer une représentation temps-fréquence de signaux réside dans l'application de la transformée de Fourier sur des sections de signaux, appelée Transformée de Fourier à Court Terme (TFCT, ou STFT en anglais). La taille de fenêtre étant constante pour toutes les fréquences, cette représentation a cependant le défaut de ne pas pouvoir adapter la taille de la fenêtre aux fréquences considérées.

L'analyse en ondelettes a été développée pour palier aux limitations des représentations précédentes en se basant sur une famille de fonctions dépendant à la fois de la fréquence à analyser et du temps. Cette analyse a été pour la première fois introduite en géophysique par Goupillaud, Grossmann et Morlet (1984), et s'est répandue dans de nombreux domaines d'applications : biologie, médecine, informatique, astrophysique, géophysique. Un développement sur les nombreuses applications des ondelettes est notamment disponible au chapitre 9 de Misiti et al. (2007).

B.2 Transformée en ondelettes continues - 1D

B.2.1 Fonctions d'ondelettes

L'analyse en ondelette consiste en la décomposition d'un signal sur une base de fonctions appelées ondelettes. Ces fonctions doivent vérifier deux propriétés (M. Holschneider, 1995) :

- L'ondelette ψ doit être de moyenne nulle :

$$\int_{-\infty}^{+\infty} \psi(t) dt = 0 \tag{B.4}$$

  Cette condition permet d'obtenir une ondelette localisée en fréquence.
- ψ doit être localisée en temps et en fréquence.

$$C_\psi = \int_0^{+\infty} d\omega \frac{|\hat{\psi}(\omega)|^2}{\omega} < \infty \tag{B.5}$$

L'analyse temporelle utilise la base de Dirac qui utilise des distributions localisées en temps, mais distribuées uniformément dans le domaine fréquentiel. Inversement, l'analyse de Fourier utilise

une base de fonctions localisées de façon précise en fréquence, mais distribuées uniformément dans le temps. Les ondelettes sont un compromis entre les deux représentations, elles sont moins précises en fréquence que les fonctions composant la base de Fourier, et moins précises que les distributions de Dirac, mais elles permettent de donner des informations sur les deux domaines.

B.2.2 Ondelettes mère

Les bases de fonctions permettant l'analyse en ondelettes sont nombreuses. Toute fonction vérifiant les conditions citées précédemment permettent une analyse en ondelette. Chaque base de fonction est dérivée à partir d'une ondelette appelée *ondelette mère*. Le choix de l'ondelette mère dépend de la nature des signaux à analyser. Nous présentons ci-dessous les deux ondelettes utilisées dans cette thèse.

Ondelette de Morlet

L'ondelette de Morlet est la première a avoir été introduite en géophysique par Goupillaud, Grossmann et Morlet (1984). Cette ondelette est définie comme la translation d'une gaussienne dans le domaine des pulsations d'ondelettes ω (sans dimension). ω_0 est le paramètre d'ondelette, H la fonction de Heaviside (Abramowitz et Stegun, 1964).

$$\widehat{\psi}_m(\omega) = \pi^{-1/4} H(\omega) e^{-(\omega-\omega_0)^2/2} \tag{B.6}$$

En domaine temporel :

$$\psi_m(\eta) = \pi^{-1/4} e^{i\omega_0\eta} e^{-\eta^2/2} \tag{B.7}$$

Une représentation dans les deux domaines est illustrée Figure B.1.

En théorie, cette fonction n'est pas une ondelette car sa moyenne n'est pas nulle. Si ω_0 est suffisamment haut (en général, la valeur $\omega_0 > 5$ est considérée suffisante), alors la moyenne de la fonction devient négligeable et celle-ci peut être considérée comme une ondelette admissible.

Le paramètre ω_0 permet l'ajustement de la balance entre la résolution temporelle et la résolution fréquentielle (Figure B.2). Augmenter ω_0 revient à augmenter le nombre d'oscillations dans l'ondelette dans le domaine temporel, augmentant son support temporel et diminuant son support fréquentiel.

La plupart des applications se basent sur cette ondelette pour ses propriétés de bon compromis entre résolution temporelle et fréquentielle (Y. Zhang et Paulson, 1997 ; Diallo et al., 2006 ; Kulesh et al., 2007 ; Rouyer et al., 2008).

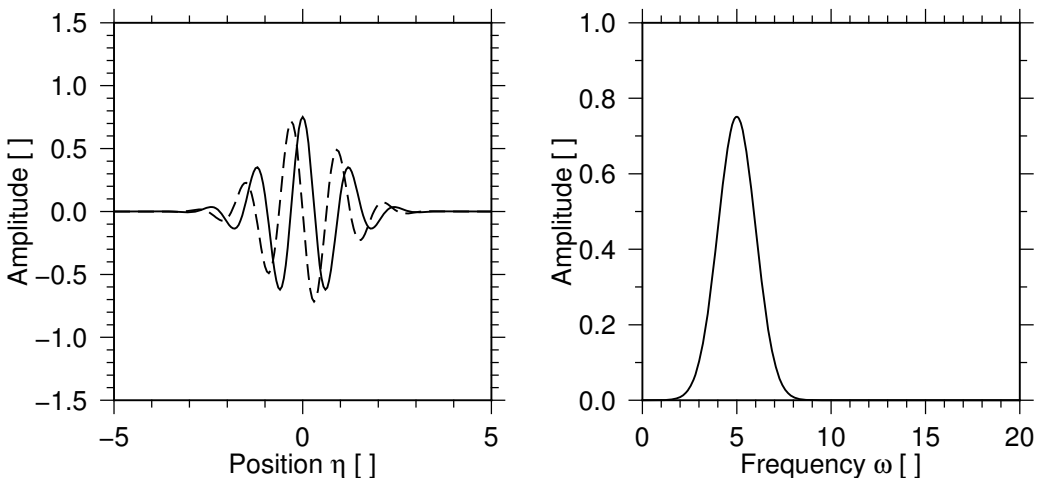


FIGURE B.1: Ondelette mère de Morlet avec le paramètre $\omega_0 = 5$. Gauche : Représentation temporelle. Trait plein : Partie réelle. Pointillé : Partie imaginaire. Droite : Module de la transformée de Fourier.
FIGURE B.1: *Morlet mother wavelet with parameter* $\omega_0 = 5$. *Left : Temporal representation. Full line : Real line. Dashed line : Imaginary line. Right : Amplitude of Fourier transform.*

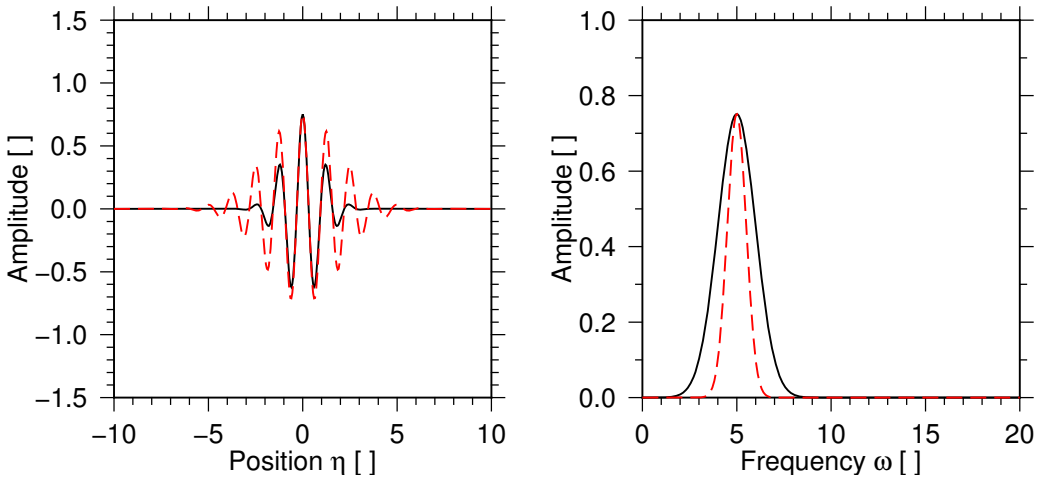


FIGURE B.2: Comparaison entre deux ondelettes mère de Morlet. Gauche : Partie réelle en domaine temporel. Droite : Domaine fréquentiel. Trait plein noir : Ondelette mère avec $\omega_0 = 5$. Trait pointillé rouge : Ondelette mère avec $\omega_0 = 10$.
FIGURE B.2: *Comparison between two Morlet mother wavelet. Left : Real part of temporal representation. Right : Amplitude of Fourier transform. Full line :* $\omega_0 = 5$. *Red dashed line :* $\omega_0 = 10$.

Ondelettes de Paul

L'ondelette de Paul (Antoine, Murenzi et Vandergheynst, 1999) (ou ondelette de Cauchy suivant les auteurs, Figure B.3), est une ondelette avec un support temporel plus petit. Elle a notamment été utilisée dans l'analyse de champs de potentiels par Moreau (1995) et P. Sailhac (1999). L'ondelette est décrite en domaine temporel par la relation (Torrence et Compo,

1998) :

$$\psi_c(\eta) = \frac{2^m i^m m!}{\sqrt{\pi(2m)!)}}(1 - i\eta)^{-(m+1)} \tag{B.8}$$

En domaine fréquentiel, l'ondelette est donnée par :

$$\widehat{\psi}_c(a\omega) = \frac{2^m}{\sqrt{m\Gamma(2m)}} H(\omega)(a\omega)^m e^{-a\omega} \tag{B.9}$$

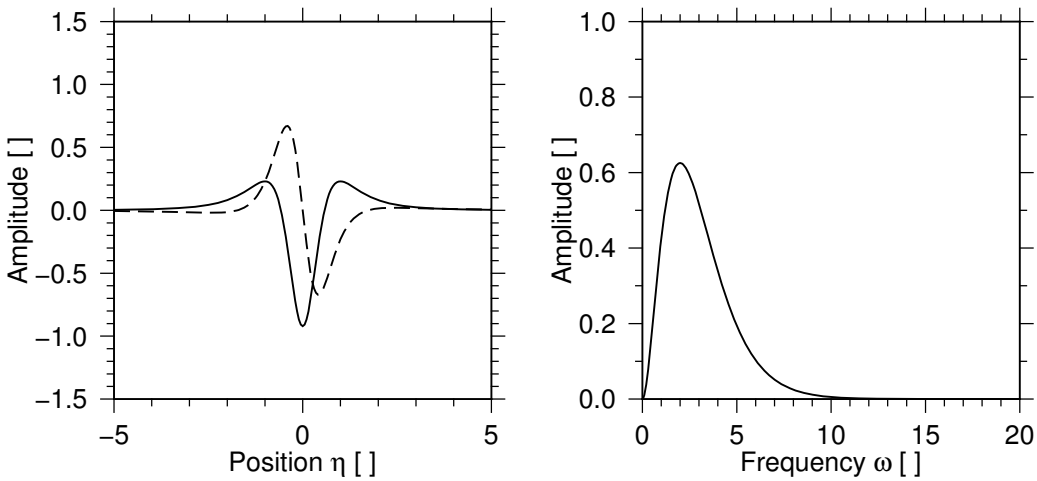


FIGURE B.3: Ondelette mère de Paul d'ordre $m = 2$. Gauche : Représentation temporelle. Trait plein : Partie réelle. Pointillé : Partie imaginaire. Droite : Module de la transformée de Fourier.

FIGURE B.3: *Paul mother wavelet of order m. Left : Temporal representation. Full line : Real line. Dashed line : Imaginary line. Right : Amplitude of Fourier transform.*

B.2.3 Ondelettes filles

Les ondelettes filles sont calculées à partir de l'ondelette mère en utilisant deux paramètres : a et τ.

$$\psi_{a,\tau}(t) = \psi\left(\frac{t-\tau}{a}\right) \tag{B.10}$$

a est appelé paramètre de dilatation, il permet de transformer l'ondelette mère en changeant sa fréquence principale par dilatation ($a > 1$) ou contraction ($a < 1$) (Figure B.4). Le paramètre τ permet la translation en temps de l'ondelette fille.

La relation entre échelle a et fréquence f pour l'ondelette de Morlet s'obtient en étudiant l'annulation de $\partial\widehat{\psi}_m(\omega)/\partial\omega$ est donnée par la relation :

$$a = \frac{\omega_0\sqrt{\omega_0^2 + 2}}{4\pi f} \tag{B.11}$$

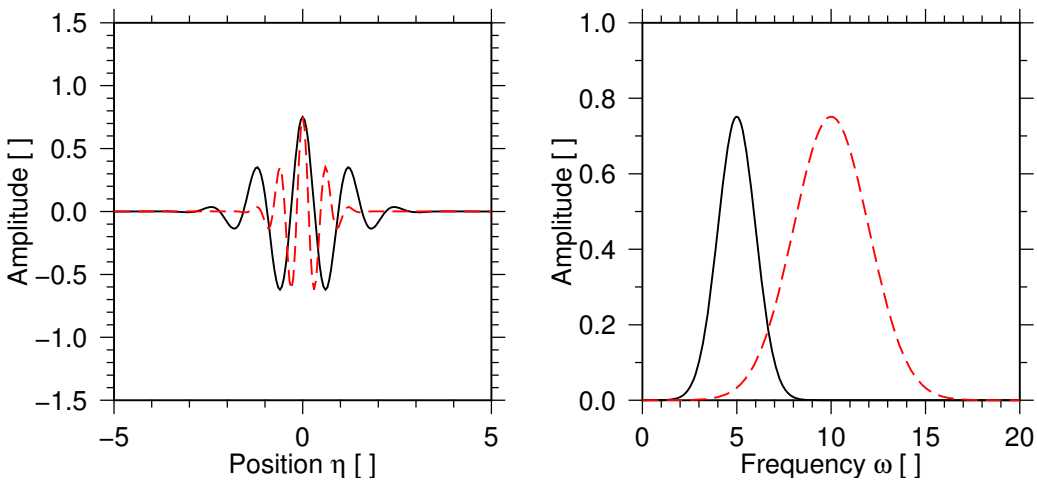


FIGURE B.4: Comparaison entre l'ondelette mère et l'ondelette fille de Morlet par dilatation de facteur $a = 1/2$. Gauche : Partie réelle en domaine temporel. Droite : Domaine fréquentiel. Trait plein noir : Ondelette mère. Trait pointillé rouge : Ondelette fille.
FIGURE B.4: *Comparison between mother and dilated Morlet wavelet. Left : Real part of temporal representation. Right : Amplitude of Fourier transform. Full line : Mother wavelet. Red dashed line : Dilated wavelet.*

La relation entre l'échelle a et fréquence f pour l'ondelette de Paul est donnée par la relation :

$$a = \frac{2m+1}{4\pi f} \tag{B.12}$$

B.2.4 Calcul des coefficients

Soit s le signal à analyser, ψ l'ondelette mère choisie. Le calcul des coefficients d'ondelettes s'effectue par le produit de corrélation entre le signal s et une ondelette fille $\psi_{a,\tau}$:

$$\mathcal{W}_\psi(a,\tau)[s] = \int_{-\infty}^{+\infty} dt \frac{1}{a} \psi^*_{a,\tau}(t) s(t) \tag{B.13}$$

L'énergie se déduit des coefficients par :

$$E = \int_{-\infty}^{+\infty} |f(t)|^2 dt = \frac{1}{C_\psi} \int_0^{+\infty} \frac{da}{a} \int_{-\infty}^{+\infty} db |\mathcal{W}_\psi(a,\tau)[s]|^2 \tag{B.14}$$

La représentation des coefficients d'ondelette s'effectue soit par le tracé des fonctions 2D représentant la partie réelle et imaginaire des coefficients d'ondelettes, soit par la fonction 2D $f(a,\tau) = |\mathcal{W}_\psi(a,\tau)[s]|^2$ appelée *scalogramme*. La quantité $|\mathcal{W}_\psi(a,\tau)|^2/a$ représente la *densité d'énergie locale* associée à l'échelle a et au temps τ. Pour des raisons de visualisations, il est pratique de représenter les coefficients $|\mathcal{W}_\psi(a,\tau)[s]|^2$, et non pas $|\mathcal{W}_\psi(a,\tau)[s]|^2/a$. Dans le deuxième

cas, les coefficients sont ramenés à l'échelle considérée, ce qui provoque une diminution de l'amplitude des coefficients pour les grandes échelles rendant la comparaison entre petites et grandes échelles compliquées.

Les coefficients peuvent aussi servir à la reconstruction du signal par l'opération de transformée inverse :

$$f(t) = \frac{1}{C_\psi} \int_{a=0}^{a=+\infty} \int_{\tau=-\infty}^{\tau=+\infty} \frac{1}{a} \mathcal{W}_\psi(a,\tau)[s]\psi_{a,\tau}(t) da d\tau \tag{B.15}$$

B.3 Noyau d'ondelette et distance de corrélation

Par définition, la transformée en ondelettes continues produit des coefficients corrélés. La corrélation entre deux coefficients est défini par le noyau d'ondelette K (Maraun, Kurths et M Holschneider, 2007). Ce noyau est défini par :

$$K_\psi(a_1, a_2, b_1, b_2) = \frac{1}{C_\psi} \mathcal{W}_\psi \left(\frac{a_2}{a_1}, \frac{b_2 - b_1}{a_1} \right) \tag{B.16}$$

$$\begin{aligned} K_\psi(a_1, a_2, b_1, b_2) = & \frac{2a_1a_2}{a_1^2 + a_2^2} \exp\left(i\omega_0 \frac{a_1 + a_2}{a_1^2 + a_2^2}(b_2 - b_1) \right) \\ & \exp\left(-\frac{1}{2} \frac{(b_2 - b_1)^2 + \omega_0(a_2 - a_1)^2}{a_1^2 + a_2^2} \right) \end{aligned} \tag{B.17}$$

B.4 Autres quantités définies à partir des coefficients

Dans cette thèse, d'autres quantitées définies à partir des coefficients d'ondelette sont utilisées :

B.4.1 Coefficients de corrélation croisée

Soit deux signaux s_1 et s_2, dont l'analyse en ondelette fournit les coefficients $\mathcal{W}_\psi(a,b)[s_1]$ et $\mathcal{W}_\psi(a,b)[s_2]$ respectivement. Le spectrogramme croisé entre s_1 et s_2 est défini par (Torrence et Compo, 1998) :

$$\mathcal{W}_\psi(a,b)[s_1; s_2] = \mathcal{W}_\psi(a,b)[s_1]\mathcal{W}_\psi^*(a,b)[s_2], \tag{B.18}$$

avec $*$ l'opérateur complexe conjugué. Cette quantité traduit l'interaction à chaque échelle entre les deux processus s_1 et s_2. Si un processus apparait simultanément dans s_1 et s_2, le spectrogramme, alors l'amplitude du spectrogramme croisé à l'échelle correspondant sera importante.

B.4.2 Spectre global d'ondelette

Le spectre global d'ondelette est définie par :

$$G_\psi[s](a) = \int_{-\infty}^{+\infty} dt |\mathcal{W}_\psi(a, dt)[s]|^2 \tag{B.19}$$

Cette quantité est notamment utilisée dans la détection des échelles significatives dans l'étude des signaux non stationnaires (Torrence et Compo, 1998).

Annexe C

Polarisation

C.1 Attributs de polarisation

C.1.1 Domaine temporel

La polarisation d'une onde est définie comme l'orientation des oscillations des différentes composantes de celle-ci. En MT, les champs considérés sont horizontaux, ils sont donc décomposables selon deux orientations perpendiculaires. Traditionnellement les directions Nord/Sud et Est/Ouest sont utilisées. En considérant une onde monochromatique de pulsation ω, ces composantes peuvent s'écrire sous la forme :

$$\begin{aligned} f_x(t) =& A_x \cos(\omega t) \\ f_y(t) =& A_y \cos(\omega t + \varphi) \end{aligned} \tag{C.1}$$

A_x et A_y sont les amplitudes sur les composantes Est/Ouest et Nord/Sud respectives. φ représente le déphasage entre les composantes f_x et f_y.
En maniant ces expressions, l'équation d'une ellipse apparait :

$$\begin{aligned} \frac{f_y}{A_y} =& \cos(\omega t)\cos(\varphi) + \sin(\omega t)\sin(\varphi) \\ =& \frac{f_x}{A_x}\cos(\varphi) + \left(1 - \frac{f_x}{A_x}\right)\sin(\varphi) \end{aligned} \tag{C.2}$$

En prenant le carré de cette expression, on obtient (Pérez, Carles et Fleckinger, 2009) :

$$\frac{f_x^2}{A_x^2} + \frac{f_y^2}{A_y^2} - 2\frac{f_x f_y}{A_x A_y}\cos(\varphi) = \sin^2(\varphi) \tag{C.3}$$

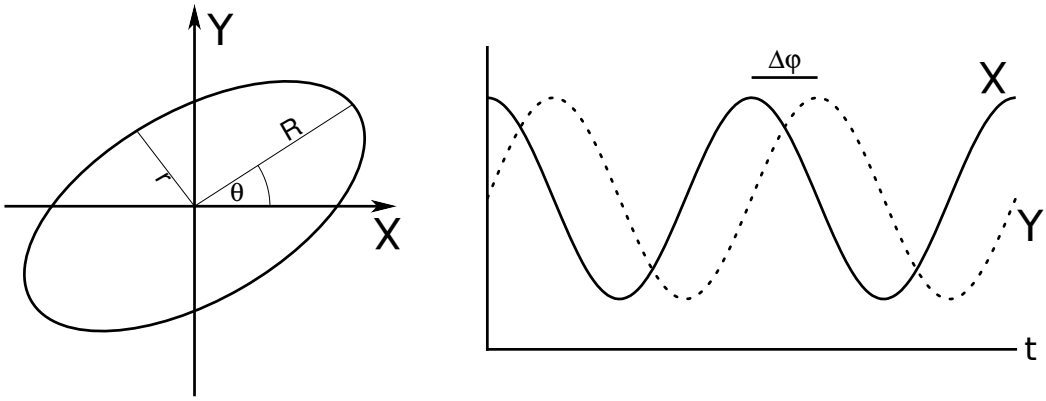


FIGURE C.1: Illustration des attributs de polarisation dans le cas général d'une polarisation elliptique.

FIGURE C.1: *Illustration of polarisation attributes in the case of elliptic polarisation.*

Les trois attributs de polarisation sont (Figure C.1) :

$\Delta\varphi$ **: Déphasage.** Le déphasage, représente la différence de phase entre les deux composantes de l'onde. Ce paramètre est exprimé en radians et prend des valeurs comprises entre $-\pi$ et π.

θ **: Angle de polarisation.** L'angle de polarisation caractérise l'angle que fait l'axe principal de l'ellipse avec la direction EW. Ce paramètre est exprimé en radians et prend des valeurs comprises entre 0 et π.

e **: Ellipticité.** Le rapport entre l'axe mineur et l'axe majeur de l'ellipse est appelé ellipticité. Ce paramètre prend des valeurs comprises entre -1 et 1.

Trois états de polarisation principaux sont définis :

Polarisation linéaire : . Cet état est décrit quand le déphasage est multiple de π. Ainsi, l'ellipticité e est égale à 0, et l'ellipse de polarisation devient une ligne. L'angle de polarisation est déterminé par l'angle entre la direction EW et la ligne de polarisation.

Polarisation circulaire : . Cet état est décrit quand l'ellipticité e est égale à $\pm$ 1 et le déphasage égale à $\pi/2$ modulo 2π. Dans cet état, l'ellipse de polarisation devient un cercle, et l'angle de polarisation est indéfini.

Polarisation elliptique : Dans tous les autres cas, ($e \neq 0$ et $\Delta\varphi \neq 0$ modulo 2π.), la polarisation est elliptique, et le vecteur décrit l'ellipse générale.

Un hodogramme est une représentation illustrant la variation temporelle d'une composante en fonction d'une autre. Les différents attributs de polarisation sont directement lisibles sur un hodogramme pour des signaux simples tels que : (i.e. Fig. C.2)

$$\begin{aligned} f_x(t) &= 0.5\cos\left(\frac{2\pi t}{50}\right) \\ f_y(t) &= \cos\left(\frac{2\pi t}{50} + \frac{\pi}{3}\right) \end{aligned} \tag{C.4}$$

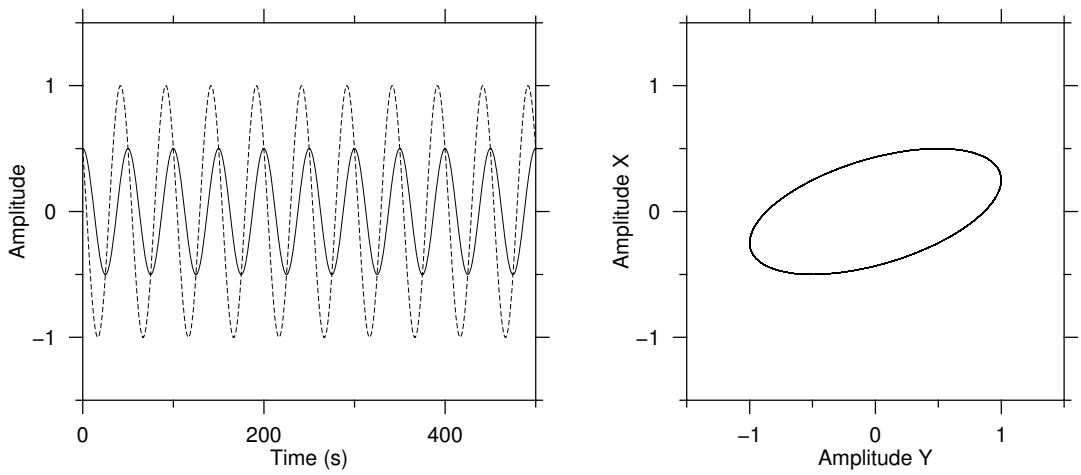


FIGURE C.2: Gauche : Signal monochromatique décrit équation C.4. Trait plein : f_x. Trait pointillé : f_y. Droite : Hodogramme des signaux.
FIGURE C.2: *Left : Monochromatic signal from equation (C.4). Full line : f_x. Dashed line : f_y. Right : Signal hodogram.*

En revanche, dès que le signal devient compliqué, cette représentation ne permet pas une détermination claire des attributs de polarisation (**FOW1967**). Par exemple, si deux signaux monochromatiques de fréquence différentes sont simultanément analysés ; tel que celui décrit équation C.5 ; les hodogrammes se superposent et la détermination devient impossible (i.e. Fig C.3).

$$\begin{aligned} f_x(t) &= 0.5\cos\left(\frac{2\pi t}{50}\right) + \cos\left(\frac{2\pi t}{100}\right) \\ f_y(t) &= f_y(t) = \cos\left(\frac{2\pi t}{50} + \frac{\pi}{3}\right) + 0.5\cos\left(\frac{2\pi t}{100} + \frac{\pi}{7}\right) \end{aligned} \tag{C.5}$$

En MT, même si une seule onde naturelle monochromatique est présente à un moment donnée dans la série temporelle, celles-ci sont contaminées par du bruit et des composantes plus hautes et basses fréquences. Ces composantes vont aussi gêner la détermination des attributs de polarisation par lecture sur l'hodogramme. L'équation (C.6) est un exemple de signal monochromatique contaminé par du bruit uniforme. La Figure C.4 montre l'hodogramme de ce signal.

$$\begin{aligned} f_x(t) &= 0.5\cos\left(\frac{2\pi t}{50}\right) + \cos\left(\frac{2\pi t}{1000}\right) + \frac{1}{5}U(0,1) \\ f_y(t) &= \cos\left(\frac{2\pi t}{50} + \frac{\pi}{3}\right) + 0.5\cos\left(\frac{2\pi t}{1000}\right) + \frac{1}{5}U(0,1), \end{aligned} \tag{C.6}$$

avec $U(0,1)$ la loi uniforme standard.

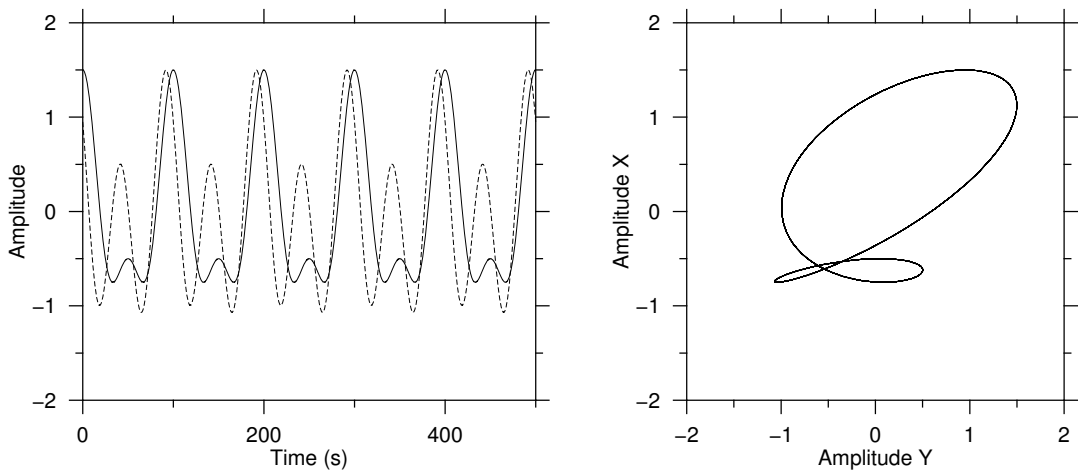


FIGURE C.3: Gauche : Signal monochromatique décrit équation C.5. Trait plein : f_x. Trait pointillé : f_y. Droite : Hodogramme des signaux.
FIGURE C.3: *Left : Monochromatic signal from equation (C.5). Full line : f_x. Dashed line : f_y. Right : Signal hodogram.*

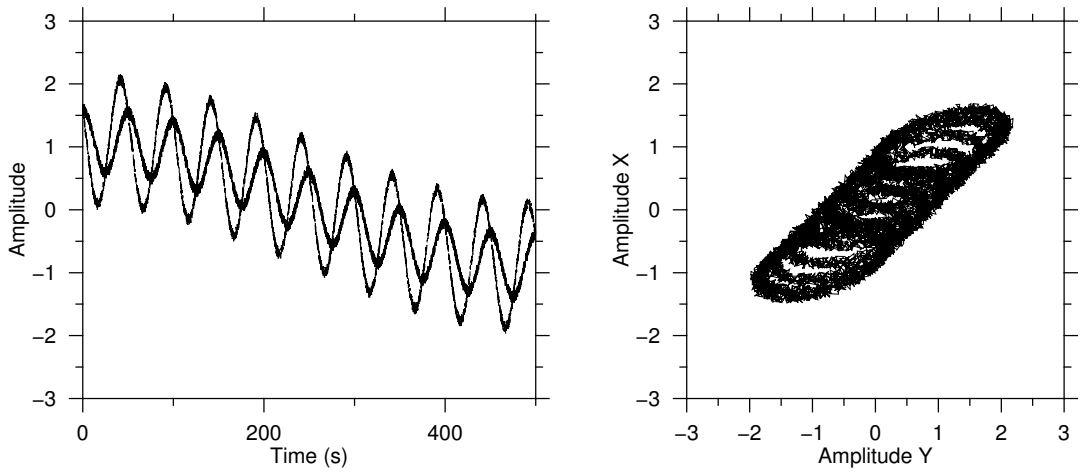


FIGURE C.4: Gauche : Signal monochromatique décrit équation C.6. Trait plein : f_x. Trait pointillé : f_y. Droite : Hodogramme des signaux.
FIGURE C.4: *Left : Monochromatic signal from equation (C.6). Full line : f_x. Dashed line : f_y. Right : Signal hodogram.*

C.1.2 Domaine de Fourier

Dans le cadre de l'étude des pulsations géomagnétiques, **FOW1967** a introduit une méthode de calcul utilisant la transformée de Fourier outrepassant les limitations des hodogrammes. La détermination des attributs se base sur le calcul de la matrice de cohérence définie par :

$$J = \begin{pmatrix} \langle H_x H_x^* \rangle & \langle H_x H_y^* \rangle \\ \langle H_y H_x^* \rangle & \langle H_y H_y^* \rangle, \end{pmatrix} \tag{C.7}$$

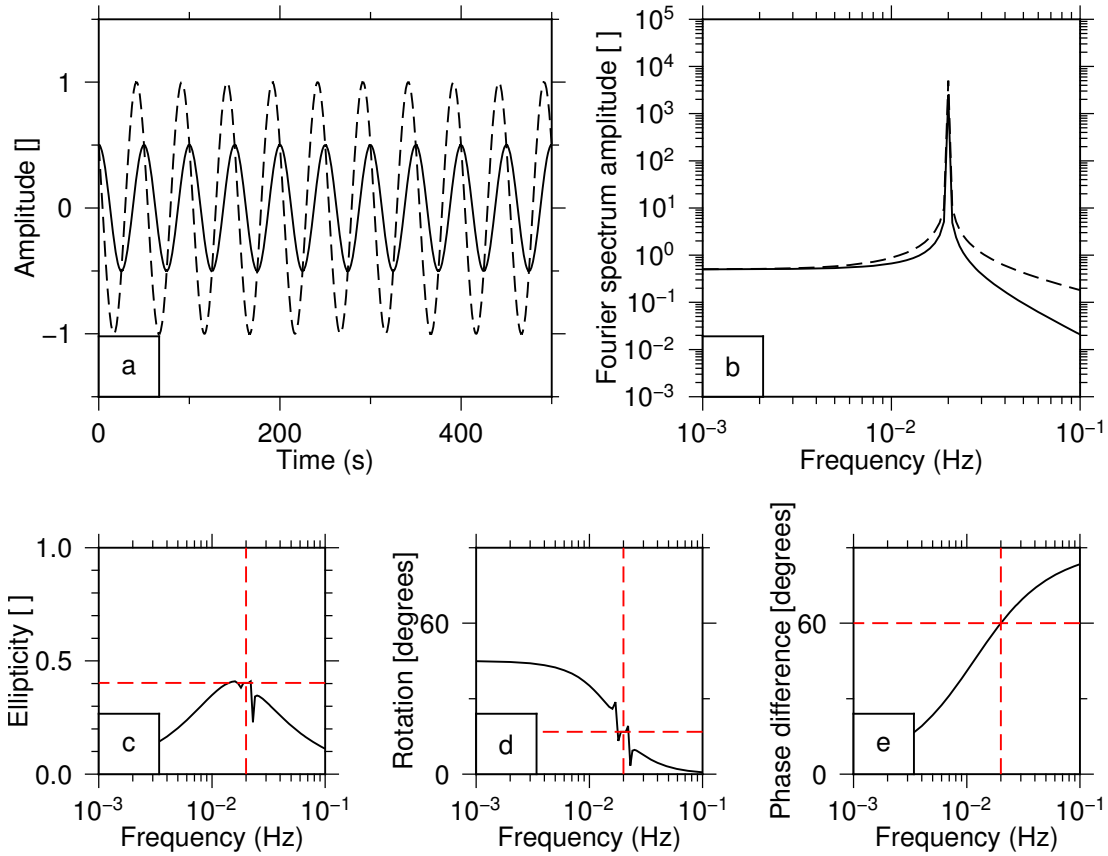


FIGURE C.5: Haut, Gauche : Signal monochromatique décrit équation (C.4). Trait plein : f_x. Trait pointillé : f_y. Haut, Droite : Spectre d'amplitude du signal décrit équation (C.4). Bas : Attributs de polarisation en fonction de la fréquence.

FIGURE C.5: *Top, Left : Monochromatic signal from equation (C.4). Top, Right : Amplitude spectra of the signal from equation (C.4). Bottom : Fourier polarisation attributes versus frequency.*

avec * l'opération conjugué et $\langle\cdot\rangle$ l'opération de moyenne sur plusieurs échantillons adjacents. A partir de l'expression des composantes définies à l'équation (C.1), la matrice de cohérence s'écrit :

$$J = \begin{pmatrix} A_x^2 & A_x A_y e^{i\varphi} \\ A_x A_y e^{-i\varphi} & A_y^2 \end{pmatrix} \tag{C.8}$$

Pour une onde monochromatique complétement polarisée, les attributs de polarisation s'obtiennent par les relations suivantes :

$$\sin(2\beta) = \frac{2\Im(J_{xy})}{J_{xx} + J_{yy}} \tag{C.9}$$

$$\tan(2\theta) = \frac{2\Re(J_{xy})}{J_{xx} - J_{yy}}. \tag{C.10}$$

Pour une superposition d'ondes monochromatiques se propageant dans la même direction, la

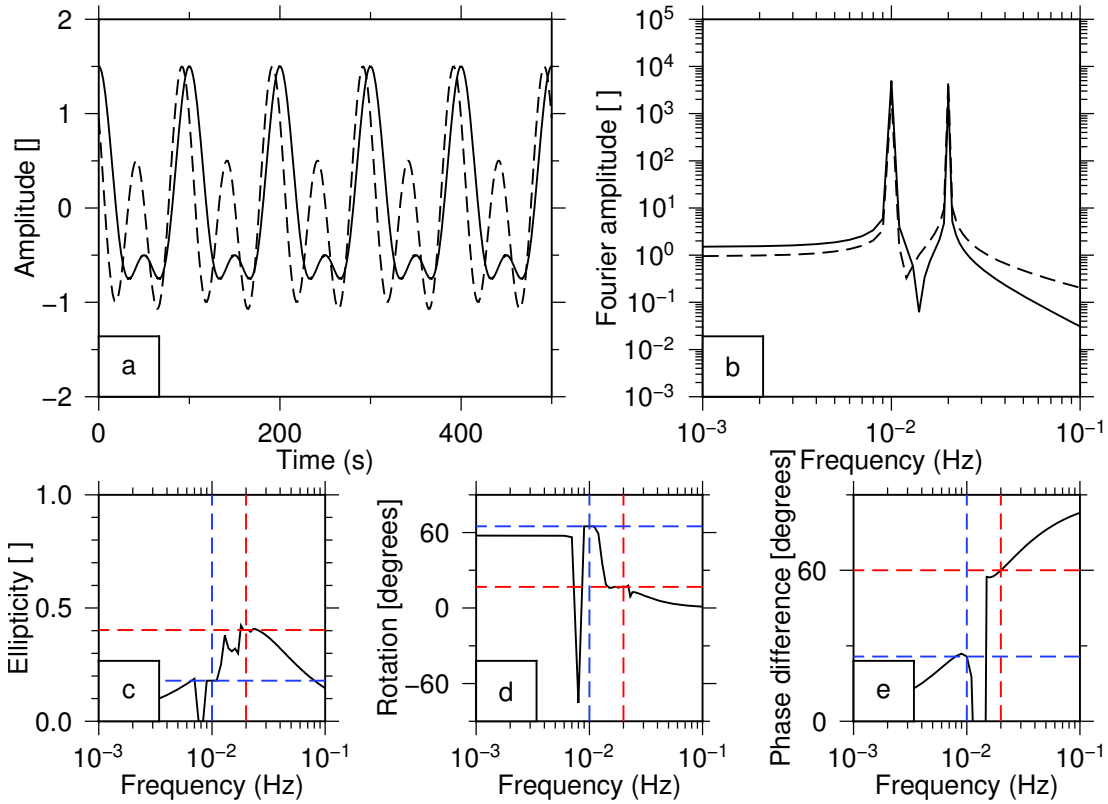


FIGURE C.6: a) Signal monochromatique décrit équation (C.5). Trait plein : f_x. Trait pointillé : f_y. b) Spectre d'amplitude du signal décrit équation (C.5). c) d) e) Attributs de polarisation en fonction de la fréquence.

FIGURE C.6: *a) Monochromatic signal from equation (C.5). b) Amplitude spectra of the signal from equation (C.5). c) d) e) Fourier polarisation attributes versus frequency.*

matrice de cohérence peut alors s'écrire sous la forme d'une matrice P représentant le signal polarisé et d'une autre matrice U représentant la partie non polarisée du signal :

$$J = P + U \tag{C.11}$$

Dans ce cas, la matrice U peut s'écrire comme une matrice diagonale dont tous les éléments sont égaux à un coefficient D tel que :

$$D = \frac{1}{2}(J_{xx} + J_{yy}) - \frac{1}{2}\left[(J_{xx} + J_{yy})^2 - 4|J|\right]^{1/2} \tag{C.12}$$

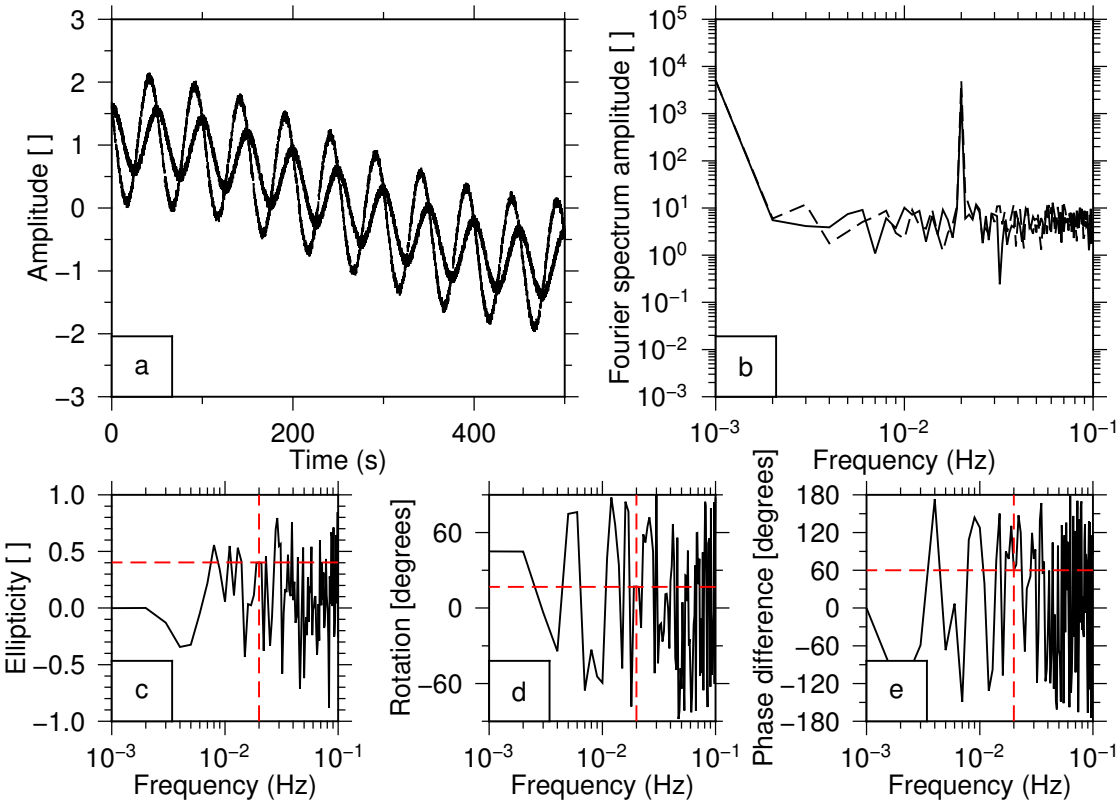


FIGURE C.7: a) Signal monochromatique décrit équation (C.6). Trait plein : f_x. Trait pointillé : f_y. b) Spectre d'amplitude du signal décrit équation (C.6). c) d) e) Attributs de polarisation en fonction de la fréquence.

FIGURE C.7: *a) Monochromatic signal from equation (C.6). b) Amplitude spectra of the signal from equation (C.6). c) d) e) Fourier polarisation attributes versus frequency.*

Les attributs de polarisation peuvent ensuite se déduire de la matrice P tels que :

$$\sin(2\beta) = \frac{2\Im(J_{xy})}{J_{xx} + J_{yy}} \tag{C.13}$$

$$\tan(2\theta) = \frac{2\Re(J_{xy})}{J_{xx} - J_{yy}}. \tag{C.14}$$

$$R = \left(1 - \frac{4|J|}{\mathrm{Tr}^2(J)}\right)^{1/2} \tag{C.15}$$

Dans le cas d'un signal monochromatique, la détermination des attributs de polarisation est directe par leur lecture en domaine fréquentiel (Figure C.5). Pour des signaux dont les hodogrammes ne permettaient pas la détermination des attributs de polarisation (superposition de fréquence, signal avec bruit), la méthode par analyse de Fourier permet la détermination des attributs de polarisations, en connaissant la fréquence du signal monochromatique (Figures C.6 et C.7).

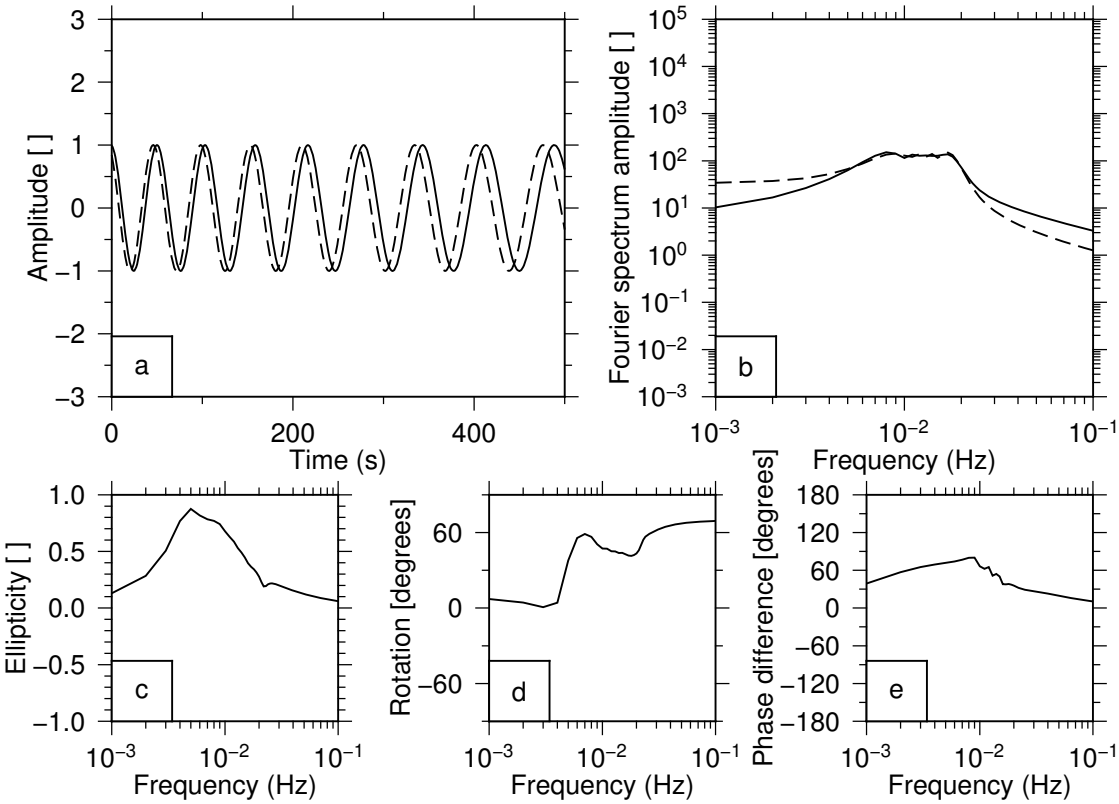


FIGURE C.8: a) Signal monochromatique décrit équation (C.16). Trait plein : f_x. Trait pointillé : f_y. b) Spectre d'amplitude du signal décrit équation (C.16). c) d) e) Attributs de polarisation en fonction de la fréquence.
FIGURE C.8: *a) Monochromatic signal from equation (C.16). b) Amplitude spectra of the signal from equation (C.16). c) d) e) Fourier polarisation attributes versus frequency.*

En revanche, cette méthode ne permet pas de rendre compte efficacement de phénomènes présentant une évolution temporelle des attributions de polarisation des signaux étudiés ou de la fréquence instantanée.

Considérons par exemple le signal suivant défini par l'équation C.16. Ce signal est sinusoïdal, et le déphasage augmente avec le temps. Le déphasage entre les composantes augmentant, l'ellipticité va aussi être modifiée.

$$\begin{aligned} x(t) &= \cos\left(\frac{2\pi}{50} + 2\pi\left[\frac{1}{50}t + \left(\frac{1}{200} - \frac{1}{50}\right)\frac{t^2}{2T}\right]\right) \\ y(t) &= \cos\left(\frac{2\pi}{50} + 2\pi\left[\frac{1}{50}t + \left(\frac{1}{200} - \frac{1}{50}\right)\frac{t^2}{2T}\right] + \phi_{inst}\right) \\ \text{avec : } \phi_{inst} &= \left(\frac{1}{50} + \left(\frac{1}{200} - \frac{1}{50}\right)\frac{t}{T}\right)^{-1} \end{aligned} \tag{C.16}$$

La figure C.8 présente la détermination des attributs de polarisation de ce signal. Contrairement aux signaux précédents, le signal ne possède pas de composantes clairement définies par

leur spectre d'amplitude. La lecture des attributs de polarisation ne permet pas de décrire la variation des attributs de polarisation avec le temps.

C.1.3 Coefficients de polarisation instantanée

René et al. (1986) ont développé une méthode de calcul des attributs de polarisation en domaine temporel basé sur l'analyse d'un signal complexe défini par :

$$c(t) = s_x(t) + is_y(t) \tag{C.17}$$

Ce signal peut s'écrire sous la forme d'une ellipse dont l'axe principal tourne dans le temps :

$$c(t) = A^+ e^{i\omega^+ t} + A^- e^{i\omega^- t} + \tag{C.18}$$

En utilisant la décomposition en ondelettes, tout signal peut s'écrire comme la somme d'une partie progressive et d'une partie régressive telle que :

$$\mathcal{W}_\psi(a,\tau)[c] = \mathcal{W}_\psi^+(a,\tau)[c] + \mathcal{W}_\psi^-(a,\tau)[c] \tag{C.19}$$

Ainsi, **DIA2006** ont montré qu'en tout instant, la transformée en ondelette pouvait s'approximer par une ellipse telle que :

$$\mathcal{W}_\psi(a,\tau+d\tau)[c] = \mathcal{W}_\psi^+(a,\tau)[c]e^{i\omega^+(a,\tau)d\tau} + \mathcal{W}_\psi^-(a,\tau)[c]e^{i\omega^-(a,\tau)d\tau} \tag{C.20}$$

A partir de cette ellipse, les attributs de polarisation sont calculés avec les relations suivantes :

$$e(a,\tau) = \frac{\left||\mathcal{W}_\psi^+(a,\tau)[c]| - |\mathcal{W}_\psi^-(a,\tau)[c]|\right|}{|\mathcal{W}_\psi^+(a,\tau)[c]| - |\mathcal{W}_\psi^-(a,\tau)[c]|} \tag{C.21}$$

$$\theta(a,\tau) = \frac{1}{2}\arg(\mathcal{W}_\psi^+(a,\tau)[c] \otimes \mathcal{W}_\psi^-(a,\tau)[c]) \tag{C.22}$$

$$\Delta\phi(a,\tau) = \arg\left(\frac{\mathcal{W}_\psi^+(a,\tau)[c] + (\mathcal{W}_\psi^-(a,\tau)[c])^*}{\mathcal{W}_\psi^+(a,\tau)[c] - (\mathcal{W}_\psi^-(a,\tau)[c])^*}\right) + \frac{\pi}{2} \tag{C.23}$$

Les attributs de polarisation calculés de cette façon sont des attributs instantanés. **DIA2006** et **KUL2007** mettent en garde contre l'interprétation d'un coefficient seul et recommandent l'interprétation d'un ensemble de coefficients. La figure C.9 présente les coefficients d'ondelettes et les coefficients de polarisation instantanés du signal présenté à l'équation C.16.

En utilisant un simple critère de rapport d'amplitude à chaque indice de temps, nous ne gardons que les coefficients dont l'amplitude est supérieure à un dixième de l'amplitude maximale. Les attributs de polarisation présentés à la figure C.10 sont calculés par simple moyenne

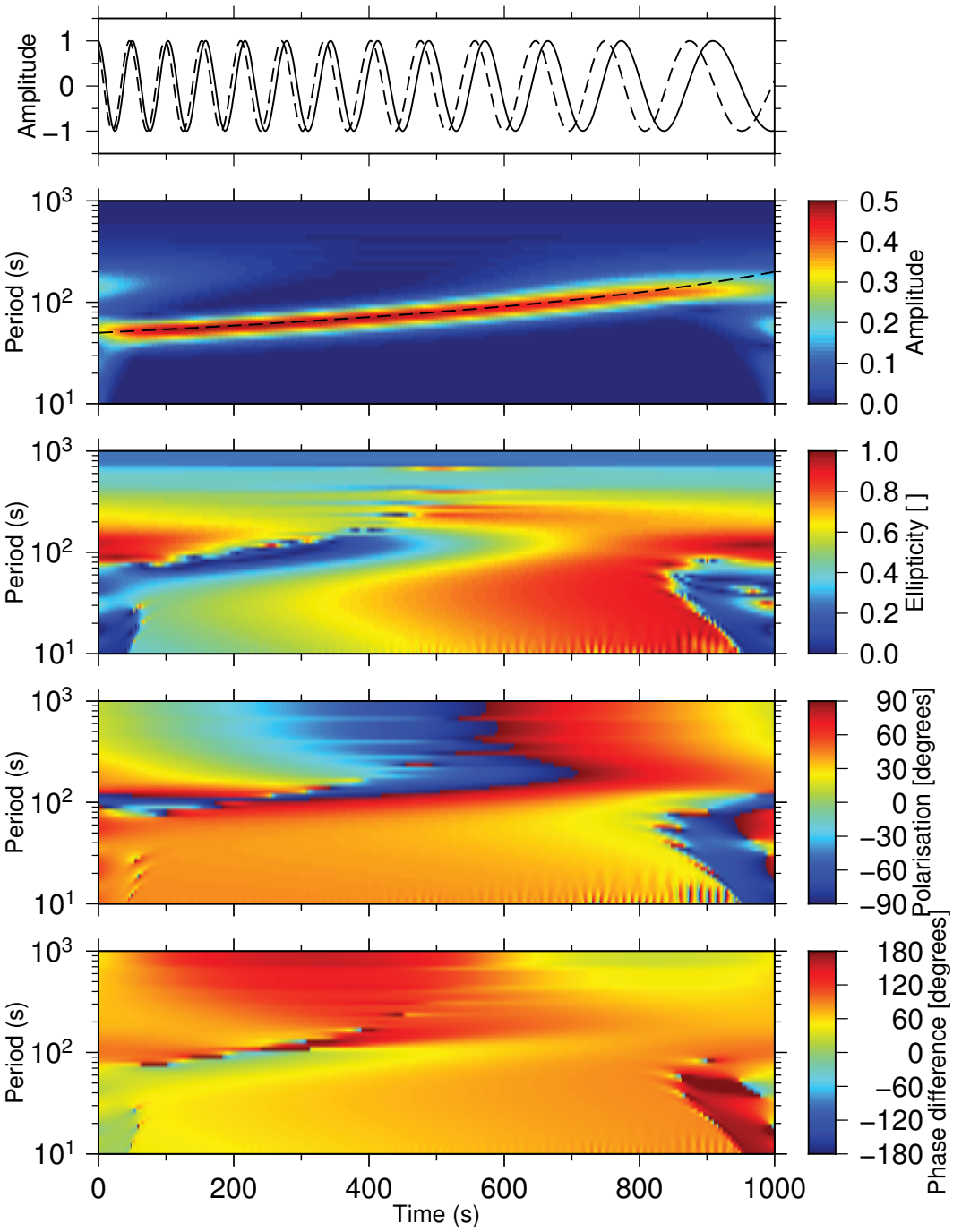


FIGURE C.9: De haut en bas : Signal monochromatique décrit équation (C.16) ; Amplitude des coefficients d'ondelettes ; Coefficients d'ellipticité instantanée ; Coefficients d'angle de polarisation instantanée ; Coefficients de différence de phase instantanée.

FIGURE C.9: *From top to bottom : Monochromatic signal from equation (C.16) ; Wavelet coefficients amplitude ; Instantaneous ellipticity coefficients ; Instantaneous polarisation angle coefficients ; Instantaneous phase difference coefficients.*

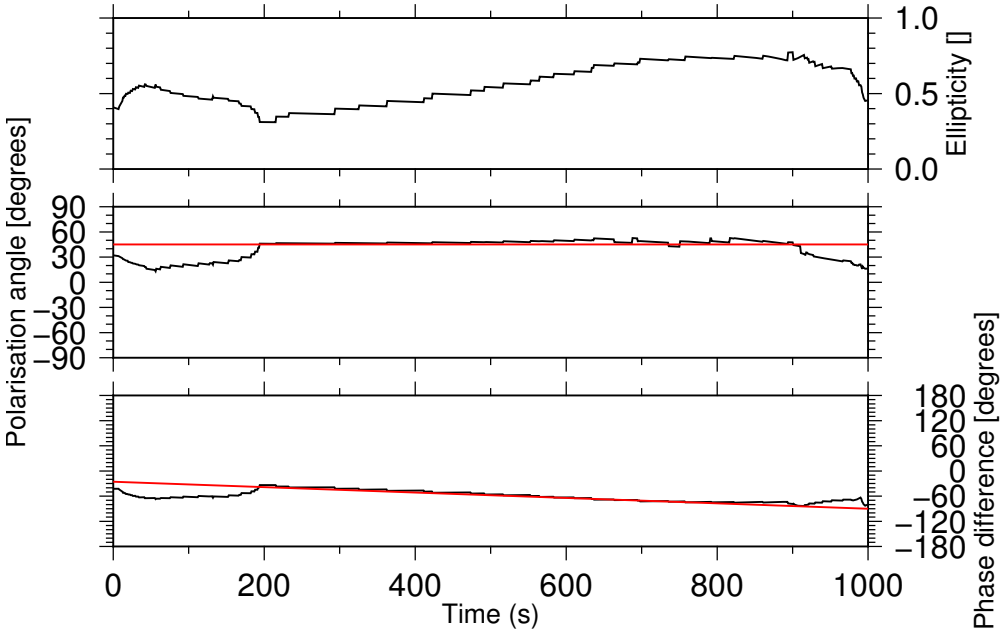


FIGURE C.10: Détermination des attributs de polarisation par la méthode d'analyse en ondelettes. Noir : Moyenne arithmétique des coefficients de polarisation instantanés. Rouge : Modèle imposé.

FIGURE C.10: *Determination of polarisation attributes using continuous wavelet transform. Black : Arithmetic mean of instantanous polarisation coefficients. Red : Model.*

arithmétique des coefficients ainsi conservés. Un simple critère d'amplitude permet de récupérer les attributs de polarisation variant avec le temps. Les variations avec le modèle sur le bord du signal sont dus aux effets de bords de l'ondelette mère.

Pour des exemples d'application de cette méthode, nous renvoyons le lecteur aux publications principales traitant de cette méthode, **DIA2006 KUL2007 ESC2013** ainsi qu'à la thèse de M. Escalas qui regorge abondamment d'applications aussi bien en méthode avec source artificielle ou naturelle.

Annexe D

Outils d'interprétation

D.1 Etude des distortions

Dans les études de distortion, le champ électrique $\mathbf{E}$ peut être supposé comme la somme d'un champ régional $\mathbf{E_R}$ et d'un champ $\mathbf{E_S}$ résultant de l'action du champ régional sur une hétérogenéité locale. Ainsi, $\mathbf{E} = \mathbf{E_R} + \mathbf{E_S}$. En supposant que l'hétérogénéité est suffisamment petite pour que $\mathbf{E_R}$ ne varie pas sur son étendue, alors on peut supposer que $\mathbf{E_S}$ est proportionnel à $\mathbf{E_R}$ (A. D. Chave et A. Jones, 2012). Ainsi, il existe une matrice de distortion $\mathbf{D}$ de rang 2, réelle, telle que :

$$\mathbf{E} = \mathbf{D}\mathbf{E_R} \tag{D.1}$$

En MT, les régions d'intérêt ne sont pas les hétérogenéités locales, mais les résultantes du champ régional. Différents outils permettent d'étudier ce champ régional. Certains sont basés sur une manipulation des éléments du tenseur (e.g. Caldwell, Bibby et Brown (2004)), d'autres sur un ajustement à un modèle prédéfini (e.g. Groom et Bailey (1989)).

Dans ce manuscrit, nous avons utilisé les deux outils principaux d'interprétation du tenseur d'impédance MT, le tenseur de phase (Caldwell, Bibby et Brown, 2004), et la décomposition de Groom-Bailey (Groom et Bailey, 1989). Ces deux techniques ont été démontrées comme parmi les plus robustes à l'introduction de bruit dans les sondages par A. D. Chave et A. Jones (2012). Elles permettent toutes les deux de donner des informations sur les distortions sur le champ électriques provoquées par des structures dans le sous-sol.

D.1.1 Tenseur de phase

En nommant $\mathbf{X}$ et $\mathbf{Y}$ les parties réelles et imaginaires du tenseur d'impédance $\mathbf{Z}$, alors le tenseur de phase est donné par :

$$\mathbf{\Phi} = \mathbf{X}^{-1}\mathbf{Y} \tag{D.2}$$

L'intérêt du tenseur de phase est son insensibilité aux distortions locales. Ainsi, si le tenseur d'impédance mesuré $\mathbf{Z}$ est donné par $\mathbf{Z} = \mathbf{D}\mathbf{Z_R}$, alors cette relation est aussi valable pour les parties réelles et imaginaires par propriété de distribution du produit matriciel. Le tenseur de phase mesuré est donc égal au tenseur de phase régional par :

$$\begin{aligned}
\mathbf{\Phi} &= \mathbf{X}^{-1}\mathbf{Y} && \text{(D.3)}\\
&= (\mathbf{D}\mathbf{X_R})^{-1}\mathbf{D}\mathbf{Y_R} && \text{(D.4)}\\
&= \mathbf{X_R}^{-1}\mathbf{D}^{-1}\mathbf{D}\mathbf{Y_R} && \text{(D.5)}\\
&= \mathbf{X_R}^{-1}\mathbf{Y_R} && \text{(D.6)}\\
&= \mathbf{\Phi_R} && \text{(D.7)}
\end{aligned}$$

La représentation du tenseur de phase se fait majoritairement au travers d'ellipses calculées à partir d'invariants du tenseur de phase :

$$\begin{aligned}
\Phi_1 &= \mathrm{tr}(\mathbf{\Phi}/2 = (\Phi_{xx} + \Phi_{yy})/2 && \text{(D.8)}\\
\det(\mathbf{\Phi}) &= \det(\mathbf{Y})/\det(\mathbf{X}) = \Phi_{xx}\Phi_{yy} - \Phi_{xy}\Phi_{yx} && \text{(D.9)}\\
\Phi_1 &= \mathrm{sk}(\mathbf{\Phi}/2 = (\Phi_{xy} + \Phi_{yx})/2 && \text{(D.10)}
\end{aligned}$$

Les paramètres utilisés sont :

$$\begin{aligned}
\Phi_{max} &= (\Phi_1^2 + \Phi_3^2)^{1/2} + (\Phi_1^2 + \Phi_3^2 - \det(\mathbf{\Phi}))^{1/2} && \text{(D.11)}\\
\Phi_{min} &= (\Phi_1^2 + \Phi_3^2)^{1/2} - (\Phi_1^2 + \Phi_3^2 - \det(\mathbf{\Phi}))^{1/2} && \text{(D.12)}\\
\beta &= \frac{1}{2}\tan^{-1}\left(\frac{\Phi_3}{\Phi_1}\right) = \frac{1}{2}\tan^{-1}\left(\frac{\Phi_{xy} - \Phi_{yx}}{\Phi_{xx} + \Phi_{yy}}\right) && \text{(D.13)}\\
\alpha &= \frac{1}{2}\tan^{-1}\left(\frac{\Phi_{xy} + \Phi_{yx}}{\Phi_{xx} - \Phi_{yy}}\right) && \text{(D.14)}\\
\lambda &= \frac{\left[(\Phi_{xx} - \Phi_{yy})^2 + (\Phi_{xy} + \Phi_{yx})^2\right]}{\left[(\Phi_{xx} + \Phi_{yy})^2 + (\Phi_{xy} - \Phi_{yx})^2\right]} && \text{(D.15)}
\end{aligned}$$

Une représentation graphique du tenseur de phase se trouve à la figure D.1.

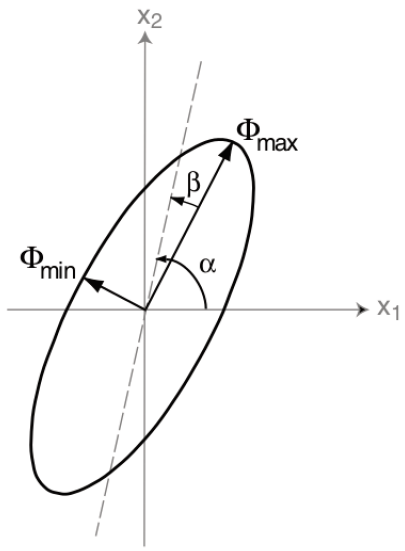


FIGURE D.1: Représentation graphique du tenseur de phase. D'après CALDWELL, BIBBY et BROWN (2004).
FIGURE D.1: *Graphical representation of the phase tensor. From* CALDWELL, BIBBY *et* BROWN *(2004).*

D.1.2 Décomposition de Groom-Bailey

La décomposition de Groom-Bailey (GROOM et BAILEY, 1989), au lieu de se baser sur un calcul à partir des éléments du tenseur d'impédance, cherche à déterminer un modèle de distortion prédéfini correspondant le mieux aux tenseur mesuré. Ce modèle prédéfini est celui d'une distortion électrique d'un milieu régional 2D par une petite hétérogénéité 3D. Selon GROOM et BAILEY (1989), l'effet de l'hétérogénéité sur le tenseur d'impédance régional peut se décomposer en quatre éléments :

- g, une amplification.
- $\mathbf{T}$, un *twist* (ou rotation) qui effectue une rotation du champ électrique horizontal.
- $\mathbf{S}$, un cisaillement.
- $\mathbf{A}$, une anisotropie.

Leur effet est illustré par la Figure D.2. Ainsi, le tenseur mesuré dans la direction de la structure régional peut s'écrire :

$$\mathbf{Z}' = g\mathbf{TSAZ_R} \tag{D.16}$$

Avec :

$$\mathbf{Z_R} = \begin{pmatrix} 0 & a \\ -b & 0 \end{pmatrix} \tag{D.17}$$

Enfin, le tenseur mesuré dans les directions EW et NS s'écrit :

$$\mathbf{Z} = \mathbf{R}(\theta)\mathbf{Z}'\mathbf{R}^{\dagger}(\theta) \tag{D.18}$$

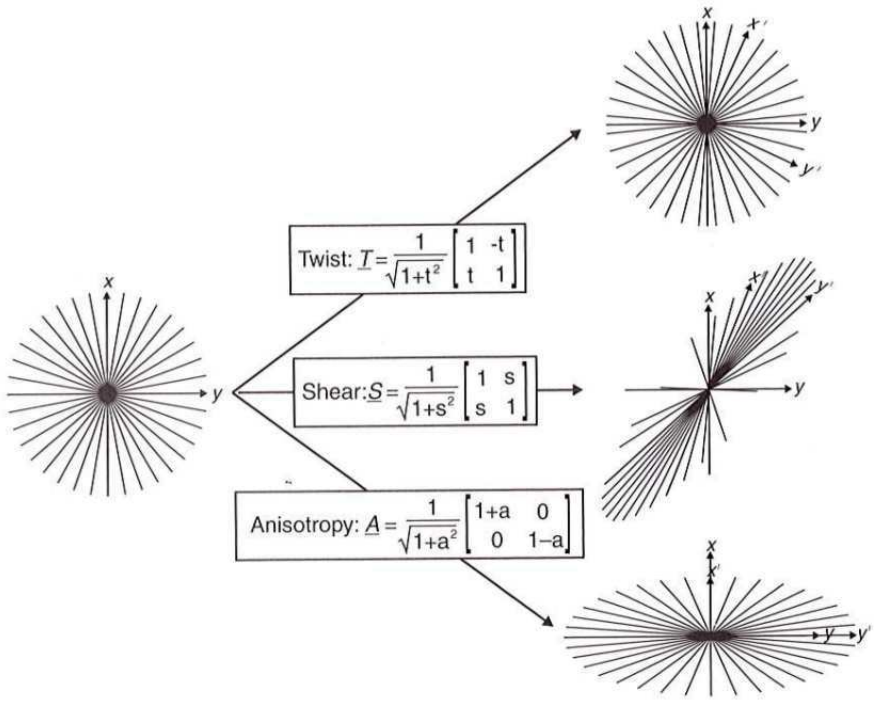


FIGURE D.2: Illustration de l'effet des matrices de distortion de Groom-Bailey. D'après A. D. CHAVE et A. JONES (2012).
FIGURE D.2: *Illustration of the effect of Groom-Bailey distortion matrix. From A. D.* CHAVE *et A.* JONES *(2012).*

La détermination des paramètres de distortion est obtenu par la résolution d'un système d'équations non linéaires :

$$\begin{aligned}
\alpha_0 &= t\sigma + e\delta \\
\alpha_1 &= (\delta - et\sigma)\cos(2\theta) - (t\delta + e\sigma)\sin(2\theta) \\
\alpha_2 &= -\sigma + et\delta \\
\alpha_3 &= -(t\delta + e\sigma)\cos(2\theta) - (\delta - et\sigma)\sin(2\theta)
\end{aligned} \tag{D.19}$$

où :

$$\alpha_0 = Z_{xx} + Z_{yy} \tag{D.20}$$

$$\alpha_1 = Z_{xy} + Z_{yx} \tag{D.21}$$

$$\alpha_2 = Z_{yx} - Z_{xy} \tag{D.22}$$

$$\alpha_3 = Z_{xx} - Z_{yy} \tag{D.23}$$

$$\sigma = a + b \tag{D.24}$$

$$\delta = a - b \tag{D.25}$$

Le système (D.19) contient 8 équations (4 pour les parties réelles, 4 pour les parties imaginaires), pour 7 inconnues (4 pour les paramètres du tenseur régional 2D, plus le *twist*, le *shear* et le *strike*). Il est donc résolvable par n'importe quelle procédure de minimisation de systèmes non linéaire.

La méthodologie initiale présentée par Groom et Bailey (1989) résout le système pour chaque fréquence et pour chaque site, et uniquement pour le cas de la distortion électrique. Alan D Chave et J. T. Smith (1994) ont complété le modèle de distortion par celui d'une distortion du champ magnétique. McNeice et Alan G Jones (2001) ont étendu l'inversion en permettant de fixer les paramètres de distortion en les supposant communs à plusieurs sites et sur plusieurs intervalles de fréquence dans le cas où la distortion est spatialement cohérente sur un nombre important de stations. Cette procédure permet de stabiliser le calcul des paramètres de distortion.

D.2 Codes de modélisation et d'inversion

Deux codes de modélisations de champ EM ont été utilisés dans ce manuscrit, l'un permettant la modélisation des champs EM en 3D, ModEM, l'autre en 2D, Mare2DEM. Le principe des codes de modélisation MT se basent tous sur la résolution des équations de Maxwell. En magnétotellurique, les sources considérées sont en dehors des volumes inversés. Ainsi, l'équation à résoudre est :

$$\nabla \times \nabla \times \mathbf{E} + i\omega\mu\sigma\mathbf{E} = 0, \tag{D.26}$$

où $\mathbf{E}$ est le champ électrique considéré à une fréquence donnée. Pour la résolution du champ magnétique, l'équation est :

$$\nabla \times \rho\nabla \times \mathbf{H} + i\omega\mu\mathbf{H} = 0, \tag{D.27}$$

où $\mathbf{E}$ est le champ électrique considéré à une fréquence donnée. Plusieurs méthodes permettent de résoudre numériquement cette équation. Les méthodes numériques les plus utilisées sont :

- Les différences finies (*Finite Difference, FD*) (Mackie, J. T. Smith et T. R. Madden, 1994 ; Siripunvaraporn, G. Egbert et Lenbury, 2002 ; Gary D Egbert et Kelbert, 2012).
- Les éléments finis (*Finite Elements, FE*) (Li et Pek, 2008 ; Key et Ovall, 2011 ; Key, 2016).
- La méthode des équations intégrales (*Integral Equations, IE*) (Ting et Hohmann, 1981 ; Hursan et Zhdanov, 2002).
- Les volumes finis (*Finite Volumes, FV*) (Du, Ren et Tang, 2016).

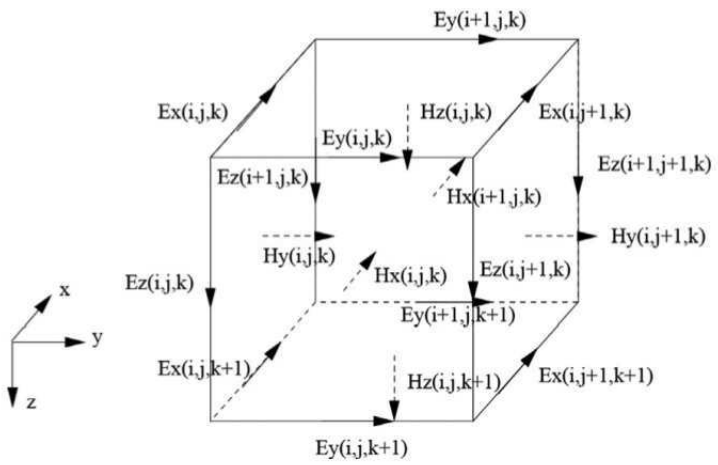


FIGURE D.3: Grille en différence finies pour le calcul des champs électromagnétiques. D'après KELBERT et al. (2014)
FIGURE D.3: *Finite-difference grid for the computation of EM fields. From* KELBERT *et al. (2014).*

D.2.1 ModEM

Le code ModEM (pour *MOdular system for ElectroMagnetic inversion* est un ensemble de codes permettant la modélisation et l'inversion de données électromagnétiques fréquentielles, non limitées à la MT (KELBERT et al., 2014). La partie MT qui est actuellement distribuée par Oregon State University (USA) est basé sur le calcul des champs électromagnétiques en différences finies (Figure D.3).

Dans les méthodes de modélisation 3D en différences finies, le calcul des champs EM se fait sur un maillage rectangulaire. Pour la résolution de l'équation (D.26), le champ électrique est calculé sur l'arête des cubes, et le champ magnétique sur la face des cubes (Figure D.3). Dans chaque cas, le champ secondaire ($\mathbf{H}$ si l'équation D.26 est résolue, et inversement) est résolu par application des équations de Maxwell. La résolution de l'équation (D.27) inverse la position des champs sur la grille de calcul. Les conditions aux limites de la grille sont définies soient par une condition de champ constant, de modélisation 1D pour l'interface inférieure, ou 2D pour les bords latéraux du modèle. Le détail de l'implémentation de la discrétisation du calcul direct des champs EM développé dans ModEM est précisé dans Gary D EGBERT et KELBERT (2012).

L'inversion des paramètres de conductivité du modèle est effectuée par une procédure *Non Linear Conjugate Gradients* (NLCG). La fonction objectif Φ minimisée est donnée par :

$$\Phi(\mathbf{m},\mathbf{d}) = (\mathbf{d}-\mathbf{f}(\mathbf{m}))^T\mathbf{C}_d^{-1}(\mathbf{d}-\mathbf{f}(\mathbf{m})) + \nu(\mathbf{m}-\mathbf{m}_0)^T\mathbf{C}_m^{-1}(\mathbf{m}-\mathbf{m}_0), \tag{D.28}$$

où $\mathbf{m}$ contient les paramètres du modèle, $\mathbf{d}$ les données inversées, $\mathbf{f}$ est l'opérateur de modélisation, $\mathbf{C}_d$ contient les erreurs sur les données, $\nu\mathbf{C}_m$ contient la covariance du modèle, $\mathbf{m}_0$ est le modèle initial.

Les modules sont actuellement implémentés en Fortran 95 et parallélisé avec la librairie *Message Passing Interface (MPI)*.

D.2.2 Mare2DEM

Mare2DEM (pour *Modeling with Adaptively Refined Elements for 2-D EM*) (Key, 2016) est un code de modélisation et d'inversion de données MT et *Controlled-Source EM* (CSEM) utilisant un maillage triangulaire permettant une discrétisation des interfaces plus précise que pour un maillage rectangulaire. La procédure d'inversion des données MT et CSEM utilise le principe de l'inversion Occam (S. C. Constable, R. L. Parker et C. G. Constable, 1987).

La modélisation des champs EM pour un modèle conductivité anisotrope triaxial pour des éléments finis est décrit par Key et Ovall (2011).

Dans la procédure d'inversion Occam, la fonction objectif à miniser est définie par :

$$U = ||\mathbf{Rm}||^2 + ||\mathbf{P}(\mathbf{m} - \mathbf{m}_*)||^2 + \mu^{-1}||\mathbf{W}(d - \mathcal{F}(\mathbf{m}))||^2 \tag{D.29}$$

Le premier terme $\mathbf{Rm}$ concerne la rugosité du modèle $\mathbf{m}$. L'algorithme va privilégier les modèles avec le moins de variations possibles de conductivité. Le deuxième terme décrit l'ajustement à un modèle de référence $\mathbf{m}_*$. La matrice $\mathbf{P}$ contient les poids attribués au modèle à priori $\mathbf{m}_*$. Le troisième terme est celui de l'ajustement des prédictions $\mathcal{F}(\mathbf{m})$ aux données $\mathbf{d}$, éventuellement pondérés par un modèle d'erreur contenu dans $\mathbf{W}$. Le terme de rugosité décrit la variation dans le modèle pour des cellules adjacentes :

$$||\mathbf{Rm}||^2 = \int_\Omega \nabla m \cdot \nabla m \; d\Omega, \tag{D.30}$$

où Ω décrit l'ensemble du modèle $\mathbf{m}$. Pour un maillage triangulaire, le calcul de la rugosité s'effectue en considérant l'ensemble des triangles adjacents et en attribuant des poids dépendants de la taille des cellules. La description du calcul des rugosités est effectuée dans Key (2016). L'introduction de frontières où la rugosité n'est pas calculée est possible, permettant l'utilisation d'informations à priori provenant d'autres méthodes pour l'inversion du modèle.

L'algorithme d'inversion Occam (S. C. Constable, R. L. Parker et C. G. Constable, 1987) est dérivé d'une minimisation par procédure de Gauss-Newton non linéaire. A partir d'un

modèle initial $\mathbf{m}_k$, le modèle suivant est obtenu par :

$$\mathbf{m}_{k+1} = [\mu(\mathbf{R}^T\mathbf{R} + \mathbf{P}^T\mathbf{P}) + (\mathbf{W}\mathbf{J}_k)^T\mathbf{W}\mathbf{J}_k]^{-1} \ \times \ [(\mathbf{W}\mathbf{J}_k)^T\mathbf{W}\hat{\mathbf{d}} + \mu\mathbf{P}^T\mathbf{m}_*] \tag{D.31}$$

où $\hat{\mathbf{d}}$ est :

$$\hat{\mathbf{d}} = \mathbf{d} - \mathcal{F}(\mathbf{m}_k) + \mathbf{J}_k\mathbf{m}_k \tag{D.32}$$

$\mathbf{J}_k$ est le jacobien du modèle, calculé par :

$$J_{ij} = \frac{\partial \mathcal{F}_i(\mathbf{m}_k)}{\partial m_j} = -\frac{\ln(10)}{\rho_j}\frac{\partial \mathcal{F}_i(\mathbf{m}_k)}{\partial \sigma_j} \tag{D.33}$$

L'algorithme Occam effectue une procédure de *line search* en testant plusieurs valeurs de μ en cherchant à obtenir un modèle $\mathbf{m}_{k+1}$ ayant un *misfit* inférieur au modèle $\mathbf{m}_k$. En pratique, cette procédure nécessite plusieurs appels à la procédure de calcul direct. Dans l'algorithme d'inversion utilisé par Mare2DEM, si un modèle est trouvé et présente une réduction du *misfit* suffisante (fixée à 15% dans Mare2DEM), ce modèle est choisi pour passer à l'itération suivante.

Le code Mare2DEM est développé en Fortran 90 et parallélisé pour l'utilisation de mémoire distribuée en utilisant le standard *Message Passing Interface* (MPI). L'utilisation de ce type de parallélisation est nécessaire pour permettre l'inversion de modèles contenant plusieurs dizaines de milliers de paramètres.

Annexe E

Détail des données

E.1 Mission Népal

E.1.1 Calendrier d'acquisition

Dates (Mars 2015)	03	04	05	06	07	08	09	10	11	12	13	14	15
Tokha													
Okharpauwa													
Patle													
Pipaltar													
Syafru													
Dhunche													
Kalikastan													
Grang													
Trisuli													

E.1.2 Détail des mesures

03 Mars 2016

- Tokha - Intallation de la station.
- Tokha - 08h 15m 00s → 08h 20m 00s. Test de fonctionnement.
- Tokha - 09h 00m 00s → 10h 00m 00s. PST à 512 Hz avec les 3 ADU.
- Tokha - 11h 00m 00s → 00h 00m 00s. Mesure MT.

04 Mars 2016

- Okharpauwa - Installation de la station.
- Okharpauwa - 09h 38m 00s → 09h 43m 00s. Test de fonctionnement.
- Okharpauwa - 10h 30m 00s → 23h 59m 00s. Mesure MT.

05 Mars 2016

- Patle - Installation de la station.
- Patle - 10h 38m 37s → 10h 43m 37s. Test de fonctionnement.
- Patle - 11h 00m 00s → 23h 59m 00s. Mesure MT
- Tokha - 00h 00m 00s → 00m 00h 00s. Mesure MT
- Okharpauwa - 00h 00m 00s → 00m 00h 00s. Mesure MT.
  Arrêt en cours de mesure.

06 Mars 2016

- Tokha - 00h 00m 01s → 10m 12h 19s. Arrêt de la mesure. Démontage de la station.
- Patle - 00h 00m 00s → 23h 59m 00s. Mesure MT

07 Mars 2016

- Pipaltar - Installation de la station.
- Pipaltar - 08h 48m 00s → 08h 53m 00. Test de fonctionnement.
- Pipaltar - 09h 30m 00s → 23h 59m 00. Mesure MT.
- Okharpauwa - Démontage de la station.
- Patle - 00h 00m 00s → 23h 59m 00s. Mesure MT.

08 Mars 2016

- Patle - Démontage de la station.
- Pipaltar - 00h 00m 00s → 23h 59m 00. Mesure MT.

09 Mars 2016

- Syafru - Installation de la station.
- Syafru - 04h 45m 00s → 04h 50m 00. Test de fonctionnement.
- Syafru - 06h 00m 00s → 23h 59m 00. Mesure MT.
- Dhunche - Installation de la station.
- Dhunche - 11h 45m 27s → 11h 50m 27s. Test de fonctionnement.

- Dhunche - 12h 00m 00s → 23h 59m 00s. Mesure MT.
- Pipaltar - 00h 00m 00s → 22h 51m 35. Mesure MT.
  Arrêt dû à la batterie vide.

10 Mars 2016

- Syafru - 00h 00m 00s → 23h 59m 00. Mesure MT.
- Dhunche - 00h 00m 00s → 23h 59m 00s. Mesure MT.
- Pipaltar - Démontage de la station.

11 Mars 2016

- Kalikastan - Installation de la station.
- Kalikastan - 06h 54m 48s → 06h 59m 48s. Test de fonctionnement.
- Kalikastan - 07h 30m 00s → 23h 59m 00s. Mesure MT.
- Syafru - 00h 00m 00s → 10h 51m 39. Arrêt de la mesure. Démontage de la station.
- Dhunche - 00h 00m 00s → 23h 59m 00s. Mesure MT.

12 Mars 2016

- Grang - Installation de la station.
- Grang - 05h 46m 17s → 05h 51m 17s. Test de fonctionnement.
- Grang - 06h 30m 00s → 23h 59m 00s. Mesure MT.
- Dhunche - 00h 00m 00s → 09h 15m 46s. Arrêt de la mesure. Démontage de la station.
- Kalikastan - 00h 00m 00s → 23h 59m 00s. Mesure MT.

13 Mars 2016

- Trisuli - Installation de la station.
- Trisuli - 07h 37m 04s → 07h 42m 04s. Test de fonctionnement.
- Trisuli - 08h 15m 00s → 23h 59m 00s. Mesure MT.
- Kalikastan - 00h 00m 00s → 23h 59m 00s. Mesure MT.
- Grang - 00h 00m 00s → 23h 59m 00s. Mesure MT.

14 Mars 2016

- Kalikastan - 00h 00m 00s → 05h 07m 14s. Arrêt de la mesure. Démontage de la station.
- Grang - 00h 00m 00s → 07h 01m 10s. Arrêt de la mesure.
- Grang - 07h 20m 00s → 23h 59m 10s. Reprise de la mesure.
- Trisuli - 00h 00m 00s → 23h 59m 00s. Mesure MT.

15 Mars 2016

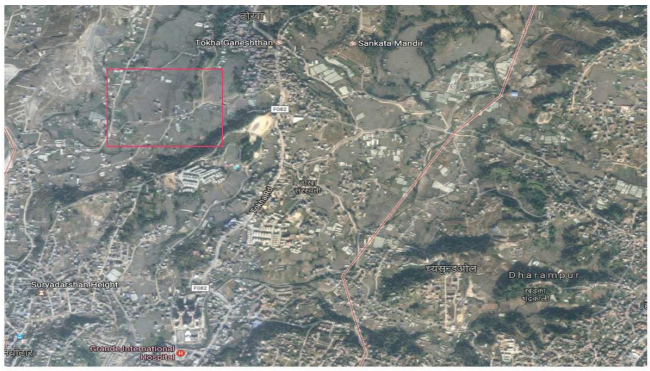
- Grang - 00h 00m 00s → 04h 51m 50s. Arrêt de la mesure. Démontage de la station.
- Trisuli - 00h 00m 00s → 06h 51m 44s. Arrêt de la mesure. Démontage de la station.

E.1.3 Détail des points mesurés

Stations	**Tokha**	**Kakani**	**Patle**	**Pipaltar**	**Syabru**	**Dhunche**	**Kalikastan**	**Grang**	**Trisuli**
Latitude (^{o})	27.76550	27.79846	27.82319	27.86796	28.16041	28.10368	28.00546	28.04673	27.92510
Longitude (^{o})	85.32405	85.27490	85.21229	85.12866	85.33551	85.30749	85.21255	85.22569	85.15329
Altitude (m)	1335	1890	1465	597	1457	2062	1769	2066	600
ADU	29	391	28	29	28	391	29	28	391
Sondes magnétiques									
H_x	MFS06 #267	MFS06 #263	MFS06 #341	MFS06 # 267	MFS06 #341	MFS06 #263	MFS06 #267	MFS06 #341	MFS06 #163
H_y	MFS06 # 276	MFS06 #264	MFS06 #342	MFS06 # 276	MFS06 #342	MFS06 #264	MFS06 #276	MFS06 #342	MFS06 #264
H_z	CM13 #103	CM13 #101	CM13 #102	CM13 #103	CM13 #102	CM13 #101	CM13 #103	CM13 #102	CM13 #101
Dipôle NS									
Centre dipôle (m)	50	25	38	46	48	8.5	21	20	51
Position N (m)	1	0	0	0	0	0	0	0	0
Position S (m)	89	48	48.5	84	64	27.5	75	58	74.5
Dipôle EW									
Centre dipôle (m)	50	21	18	46	36.5	39.5	31	28.5	50
Position E (m)	4	0	0	0	0	0	0	0	0
Position W (m)	100	38.5	27	90	60	71	54	55.5	95
Résistances (*selftest*)									
E_x ($k.\Omega$)	0.96	10	0.55	3.01	0.25	4.46	16.5	10	0.85
E_y ($k.\Omega$)	0.75	20	0.57	3.51	0.29	10.7	13.8	10	1.16

Table E.1: Détail des configurations AMT

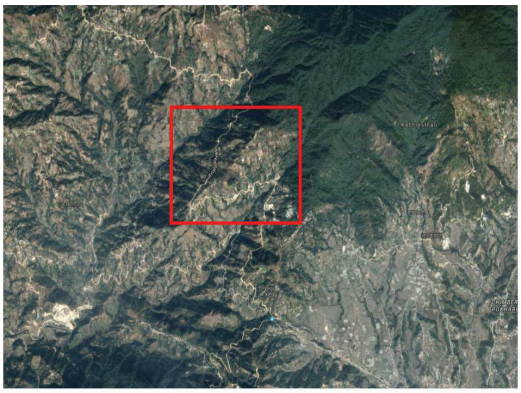
Tokha



Remarques :

- Vers l'ouest, à environ 1 km, présence d'une colline avec beaucoup de travaux (tracteurs, pelleteuses...).
- Ce site à servi pour la réalisation du test des sondes magnétiques (*Parallel Sensor Test*).

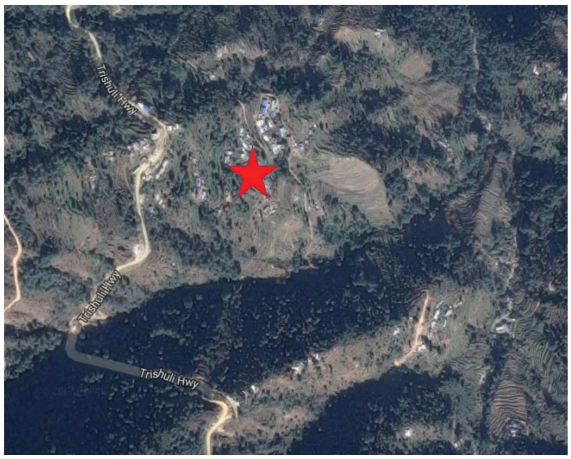
Kakani (Okharpauwa)



Remarques :

- Longue barre métallique à environ 20 m du centre du dispositif (25 m de la sonde la plus proche – H_x).
- Conduite d'eau sur le sol traversant le dispositif.
- Après l'installation du site, des enfants sont venus jouer avec des vélos autour du site.
- Le site Kakani effectué en 1996 est visible, mais de l'autre côté de la vallée.
- Site difficile d'accès.

Patle

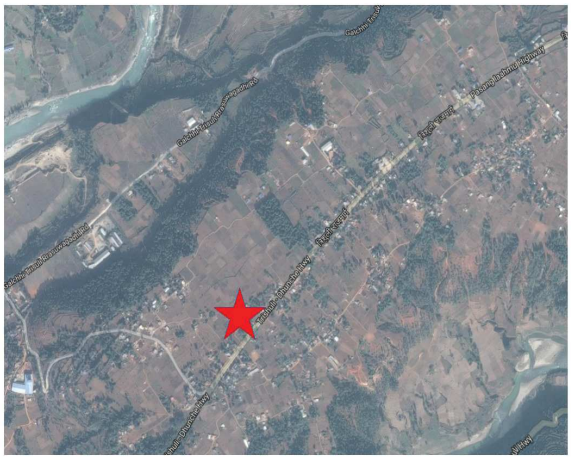


Remarques :

- Modification des gains par la commande d'*Autogain*.
  Avant utilisation : tous les gains à 1.
  Après utilisation : 4, 8, 1, 8, 1.
  Avant le lancement des tests, les gains ont été fixés à : 4, 4, 1, 1, 1. Ces valeurs ont été utilisées pour les mesures.
- Au début des séries, beaucoup d'enfants et des travaux de désherbage à la machette près du site.

- Petit poteau électrique près de l'électrode N.

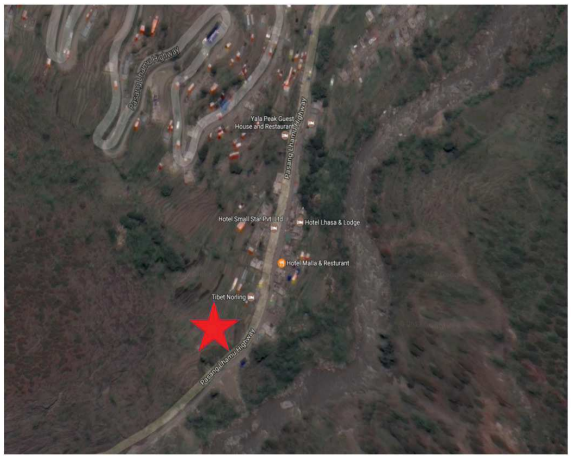
Pipaltar



Remarques :

- Site à environ 100 m de la route.
- Site avec une petite ligne électrique à environ 70 m.
- Beaucoup d'enfants.
- Une batterie à été déconnectée et l'acquisition s'est arrêtée quelques heures avant le démontage de la station.

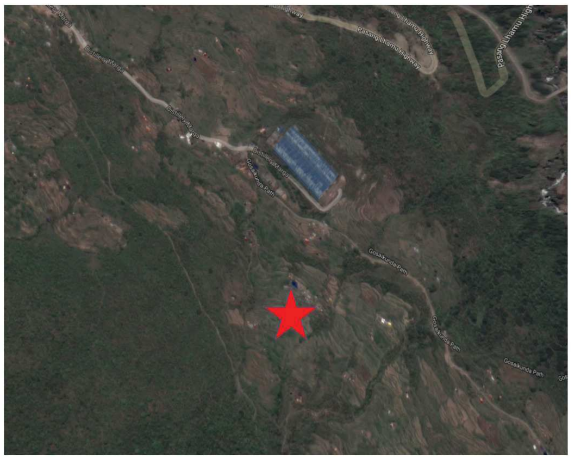
Syafrubensi (Syabru)



Remarques :

- Léger dénivelé du terrain :
  - 11 terrasses d'environ 50 cm de hauteur entre les électrodes N et S, avec $z_N > z_S$.
  - 3 terrasses d'environ 50 cm de hauteur entre les électrodes E et W, avec $z_N > z_S$.
- Transformateur à environ 1 km de distance.

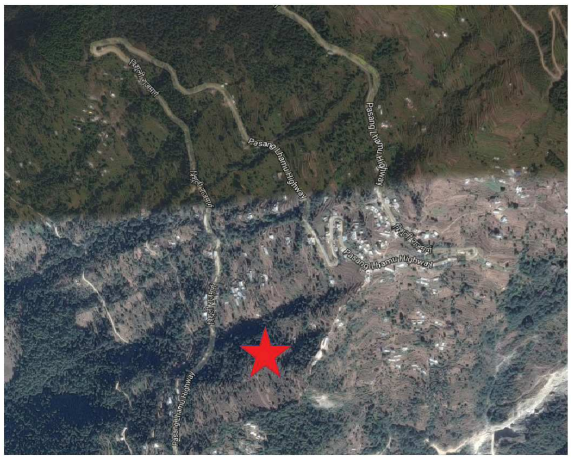
Dhunche



Remarques :

- Dipôle E_y plus horizontal que E_x (quelques mètres de dénivelé entre Nord et Sud, avec $z_S > z_N$).

Kalikastan



Remarques :

- Terrain en pente :
  $z_S > z_N$, d'environ 3 mètres (2 terrasses d'environ 1.5 m).
  $z_E > z_W$, d'environ 7.5 mètres (5 terrasses d'environ 1.5 m).
- Poteau électrique près de l'électrode sud.
- Site calme (excepté bruit électrique).

Grang

Remarques :

- Site surélevé (exploitation d'herbes médicinales).
- Route en bas du site (à environ 50 mètres).
- Pente sur le site :

  $z_S > z_N$ d'environ 5 m.

  $z_E > z_W$ d'environ 5 m.

Trisuli

Remarques :

- Ligne électrique à environ 50 m vers l'est.
- Route proche à environ 5 m de l'électrode ouest.
- Présence d'animaux broutant autour et dans le dispositif. Une électrode à notamment été arrachée le matin du dernier jour sur place (quelques heures de données perdues).

E.2 Chambon-La-Forêt

E.2.1 Détail des mesures

Stations	**Ingrannes**	**Thivet**	**Observatoire**
Latitude (^{o})	47.99352	48.002468	48.026089
Longitude (^{o})	2.242743	2.1552133	2.253436
Altitude (m)	130.50	134.31	129.62
ADU	29	29	28
Sondes magnétiques			
H_x	MFS06 # 341	MFS07 # 217	MFS06 # 264
H_y	MFS06 # 341	MFS07 # 218	MFS06 # 267
H_z	MFS06 # 263	MFS07 # 222	CMA
Dipôle NS			
Centre dipôle (m)	50	40	50
Position N (m)	0	0	0
Position S (m)	100	80	97.5
Dipôle EW			
Centre dipôle (m)	45	41	50
Position E (m)	0	0	0
Position W (m)	95	81	91.5
Résistances (*selftest*)			
E_x (Ω)	10834.5	719.937	3067.94
E_y (Ω)	5472.79	713.279	4673.5

Table E.2: Détail des configurations MT

Stations	**B2**	**B4**	**B6**	**C1**	**C3**	**C6**	**C7**
Latitude (^{o})	48.0034783	48.0131449	48.025545	48.004425	48.013043	48.030506	48.038808
Longitude (^{o})	2.25200833	2.264153	2.272658	2.2345133	2.24608	2.2696	2.275408
Altitude (m)	126.13	137.38	163.09	103.67	130.94	142.05	134.13
ADU	238	238	238	238	238	238	238
Sondes magnétiques							
H_x	MFS07 # 217	MFS07 # 217	MFS07 # 217	MFS07 # 217	MFS07 # 217	MFS07 # 217	MFS07 # 217
H_y	MFS07 # 218	MFS07 # 218	MFS07 # 218	MFS07 # 218	MFS07 # 218	MFS07 # 218	MFS07 # 218
H_z	MFS07 # 222	MFS07 # 222	MFS07 # 222	MFS07 # 222	MFS07 # 222	MFS07 # 222	MFS07 # 222
Dipôle NS							
Centre dipôle (m)	15	15	15	15	15	15	15
Position N (m)	0	0	0	0	0	0	0
Position S (m)	30	30	30	30	30	32	30
Dipôle EW							
Centre dipôle (m)	15	15	15	15	15	16	16
Position E (m)	0	0	0	0	0	0	0
Position W (m)	30	30	30	30	31	32	31
Résistances (*selftest*)							
E_x (Ω)	184.862	185.119	185.665	185.078	185.306	185.144	184.842
E_y (Ω)	185.284	185.556	185.753	184.878	185.653	185.764	184.838

Table E.3: Détail des configurations AMT

Stations	**D4**	**E3**	**E5**	**F4**	**G5**	**Obs**
Latitude (^{o})	48.022434	48.021488	48.03382	48.03203	48.04523	48.02458
Longitude (^{o})	2.244091	2.227945	2.241436	2.2257483	2.220825	2.260291
Altitude (m)	129.72	134.31	140.63	137.70	133.57	148.93
ADU	238	238	238	238	238	238
Sondes magnétiques						
H_x	MFS07 # 217	MFS07 # 217	MFS07 # 217	MFS07 # 217	MFS07 # 217	MFS07 # 213
H_y	MFS07 # 218	MFS07 # 218	MFS07 # 218	MFS07 # 218	MFS07 # 218	MFS07 # 214
H_z	MFS07 # 222	MFS07 # 222	MFS07 # 222	MFS07 # 222	MFS07 # 222	MFS07 # 220
Dipôle NS						
Centre dipôle (m)	15	15	15	15	15	15
Position N (m)	0	0	0	0	0	0
Position S (m)	30	30	30	30	30	30
Dipôle EW						
Centre dipôle (m)	15	15	16	16	17	15
Position E (m)	0	0	0	0	0	0
Position W (m)	30	30	31	31	30	30
Résistances (*selftest*)						
E_x (Ω)	190.812	181.768	185.477	184.855	185.446	185.31
E_y (Ω)	167.846	185.511	187.956	185.536	185.565	185.84

Table E.4: Détail des configurations AMT

UNIVERSITÉ DE STRASBOURG

Hugo LARNIER

Ecole et Observatoire des Sciences de la Terre

Intégration des données d'observatoires magnétiques dans l'interprétation de sondages magnétotelluriques :

Acquisition, Traitement, Interprétation

Résumé

Dans ce manuscrit, nous développons des méthodologies de détection et caractérisation de sources géomagnétiques et atmosphériques en se basant sur la transformée en ondelettes continues. Les techniques introduites se basent sur les caractéristiques temps-fréquence des ondes observées dans les séries temporelles magnétotelluriques (MT). A partir de ces procédures de détection, nous détaillons l'implémentation d'une stratégie de détermination des fonctions de réponse MT basée sur les statistiques robustes, et du bootstrap hiérarchique pour le calcul des incertitudes.

Deux études MT sont également détaillées. La première étude MT concerne la caractérisation de la structure géoélectrique situé sous l'observatoire magnétique de Chambon-La-Forêt, France. La seconde étude concerne des mesures effectuées dans la vallée de Trisuli au Népal en mars 2016. L'objectif de cette campagne est la comparaison avec une étude effectuée en 1996. Nous discutons des effets topographiques sur les sondages MT. Nous présentons également une nouvelle interprétation de la distribution de conductivité dans le sous-sol de vallée de Trisuli.

Mots clefs : Magnétotellurique, Géomagnétisme, Observatoires magnétiques, Transformée en ondelette continue, Statistiques robustes, Chambon-La-Forêt, Népal.

Résumé en anglais

In this manuscript, we detail the application of continuous wavelet transform to processing schemes for the detection and the characterisation of geomagnetic and atmospheric sources. Presented techniques are based on time-frequency properties of electromagnetic (EM) waves observed in magnetotellurics (MT) time series. We detail the application of these detection procedures in a MT processing scheme. To recover MT response functions, we use robust statistics and a hierarchical bootstrap approach for uncertainties determination.

Interpretation of two datasets are also presented. The first MT study deals with the caracterisation of the resistivity distribution below the French National magnetic observatory of Chambol-la-For\^et. The second study details the interpretation of new MT soundings acquired in March 2016 in the Trisuli valley, Nepal. The main objective of this campaign was to compare the new soundings with an old campaign in 1996. We discuss topography effects on MT soundings and their implication on the resistivity distribution. We also introduce a new interpretation of the resistivity distribution in Trisuli valley.

Keywords : Magnetotellurics, Geomagnetism, Magnetic Observatories, Continuous wavelet transform, Robust statistics, Chambon-La-Forêt, Népal.