2D Full waveform inversion of ground penetrating radar data : towards multiparameter imaging from surface data

François Lavoué

UNIVERSITÉ DE GRENOBLE

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Sciences de la Terre, de l'Univers et de l'Environnement**

Arrêté ministériel : 7 août 2006

Présentée par

François Lavoué

Thèse dirigée par **Stéphane Garambois**
et codirigée par **Jean Virieux**,

préparée au sein de **l'Institut des Sciences de la Terre**
et de **l'école doctorale Terre Univers Environnement**.

2D Full waveform inversion of ground penetrating radar data

Towards multiparameter imaging from surface data

Thèse soutenue publiquement le **9 juillet 2014**,
devant le jury composé de :

Hervé Chauris
Professeur, Ecole des Mines de Paris, France, Rapporteur
Jan van der Kruk
Professeur, FZ Jülich, RWTH Aachen, Germany, Rapporteur
Donatienne Leparoux
Chargée de Recherche, IFSTTAR, Nantes, France, Examinatrice
Philippe Roux
Directeur de Recherche, ISTerre, CNRS, Grenoble, France, Examinateur
Guy Sénéchal
Maître de Conférence, IPRA, Université de Pau et des Pays de l'Adour, France, Examinateur
Stéphane Garambois
Maître de Conférence, ISTerre, Université Joseph Fourier, Grenoble, France, Directeur de thèse
Jean Virieux
Professeur, ISTerre, Université Joseph Fourier, Grenoble, France, Co-Directeur de thèse

A mes grands-pères.
L'un d'eux au moins aura vu l'achèvement de ce travail,
lui qui a tant contribué à éveiller ma curiosité scientifique.

Remerciements

Une thèse, comme tout projet scientifique, ne se mène pas seul. J'aimerais ici remercier celles et ceux qui m'ont aidé à mener à bien ce travail, en demandant à ceux que j'aurais oublié de bien vouloir m'excuser...

Merci tout d'abord aux membres du jury qui ont accepté d'évaluer mon travail. Parmi eux, je remercie tout particulièrement Jan van der Kruk qui a bien voulu rapporter mon manuscrit et faire le voyage jusqu'à Grenoble pour subir ma soutenance en français. J'espère qu'il saura déchiffrer ces quelques lignes de gratitude, lui qui m'a initié au GPR lors de mon stage à Jülich en 2009. Merci également à Hervé Chauris pour avoir accepté d'être mon deuxième rapporteur. J'ai été très agréablement surpris, voire stupéfait, par la justesse de ses commentaires sur mon travail et l'acuité de ses questions : je ne pensais pas pouvoir me faire aussi bien comprendre sur la question délicate du *scaling* à appliquer entre les paramètres et de leurs différents impacts dans les données. Merci à Donatienne Leparoux pour avoir été l'examinatrice de ma soutenance. Je la remercie aussi pour nos discussions lors des congrès IWAGPR et pour m'avoir proposé à plusieurs reprises le jeu de données de F. Lopes, auquel j'ai finalement préféré celui de l'Institut Fresnel pour mon application en milieu contrôlé. J'espère que nous aurons l'occasion d'interagir à l'avenir sur ce jeu de données, ainsi que sur la question du diagramme d'antenne. Je remercie Guy Sénéchal, non seulement pour avoir examiné mon travail mais aussi pour avoir contribué à l'aspect *terrain* de ma thèse, d'abord en participant aux mesures à Celano (Italie), que je ne désespère pas de traiter un jour, puis pour m'avoir fourni le jeu de données acquis par lui-même et Dominique Rousset (que je remercie au passage). Je le remercie pour nos échanges à propos de ces données et j'espère que ces échanges continueront jusqu'à publication de ce travail et au-delà. Enfin, je remercie Philippe Roux pour avoir accepté de présider mon jury de soutenance. Sa vision d'expérimentateur acoustique a apporté un regard extérieur sur mon travail.

Des remerciements tout particuliers vont naturellement à mes directeurs de thèse, Stéphane Garambois et Jean Virieux, qui m'ont d'abord proposé un stage de Master sur l'aspect *problème direct* de ce projet, puis qui ont bien voulu m'accompagner en thèse malgré les quelques travers qui transparaissaient déjà pendant ce stage. Mon indépendance, soulignée pendant ma soutenance, m'a en effet parfois fait perdre un temps précieux en ne profitant pas assez de l'aide que Jean et Stéphane pouvaient m'apporter. J'espère avoir amélioré ce point au cours de la thèse et avoir au final profité pleinement de leurs apports, à défaut de l'avoir fait efficacement. Merci à Stéphane pour ses conseils pragmatiques orientés *données*, que j'aurais sans doute dû plus écouter pour ne pas me perdre dans les *minima locaux*. Merci aussi pour m'avoir fait confiance sur d'autres aspects du métier d'enseignant-chercheur, en me confiant mes premiers articles et rapports de stage d'étudiants à évaluer. Merci à Jean pour son coaching parfois directif mais souvent efficace lorsqu'il s'agissait de me sortir de ces fameux minima locaux. Merci à lui pour m'avoir fait partager un peu de sa vision scientifique, distillée sous forme de sentences plus ou

moins énigmatiques dans nos échanges de mails ou entre deux portes du laboratoire. Sur le plan de l'enseignement, je lui sais gré de m'avoir confié les stages terrain du M2 Pro GER où j'ai eu à faire à des étudiants extrêmement motivés.

Bien que financé sur fonds ministériels, je suis redevable au consortium Seiscope (`http://seiscope2.osug.fr`) de l'environnement de travail privilégié dont j'ai profité, en particulier en ce qui concerne les codes numériques que j'ai utilisés pendant ma thèse. Je me dois donc ici de remercier les sponsors de ce consortium, à savoir BP, CGG, Chevron, Exxon-Mobil, JGI, Petrobras, Saudi Aramco, Schlumberger, Shell, Sinopec, Statoil, Total et Woodside. Au-delà de l'aspect financier, je tiens à remercier les membres du groupe Seiscope qui contribuent à entretenir cet environnement de travail extrêmement enrichissant et stimulant, notamment au travers des visioconférences régulières entre Nice et Grenoble. Je remercie en particulier Romain Brossier qui est à l'origine de la plupart des codes numériques que j'ai utilisés et vers qui je me suis souvent tourné pour mes questions techniques. Bien souvent, cinq minutes de discussion avec Romain me faisait plus progresser que des jours entiers seul face à mon écran. Je ne sais pas si j'atteindrai un jour son efficacité et sa lucidité scientifique. Je remercie aussi Ludovic Métivier pour ses conseils rigoureux sur l'aspect mathématique de l'inversion, en particulier sur la formulation lagrangienne que j'ai dû me faire expliquer plusieurs fois, à l'instar de celui qui comprends vite mais à qui on doit expliquer longtemps... Merci aussi à Stéphane Operto pour avoir partagé un peu de son énorme expérience des différences finies, et d'un point de vue pratique pour m'avoir fourni le code permettant d'optimiser les coefficients du schéma mixte (*mixed-grid stencil*). Cette contribution aurait d'après moi largement justifié de l'intégrer à la liste des auteurs du papier sur les données de Fresnel mais Stéphane est bien trop modeste pour cela et a préféré figurer dans les remerciements, que je réitère donc ici. Merci à Alessandra Ribodetti pour ses conseils sur l'estimation de la source, ainsi que pour ses encouragements et son extrême gentillesse manifestée à chaque rencontre annuelle Seiscope. Enfin, merci aux thésards et post-docs du groupe Seiscope pour leur sympathie et leurs apports scientifiques lors des visioconfs. Un merci tout particulier à Amir pour notre année de co-location du bureau 232. Merci aussi à Clara que j'ai toujours beaucoup de plaisir à croiser aux Annual Meetings pour discuter de régularisation (ou de choses plus joyeuses). Merci enfin à Aurélien, Bastien, François Bretaudeau, Hugo, Isabella, John, Josué, Okba, Paul, Stéphen, Théo, Vadim, les deux Vincent, Wei, Yang et Yaser.

En dehors de ce noyau dur, je remercie les personnes avec qui j'ai interagi de près ou de loin autour de mon sujet de thèse. Ainsi, merci à Ludovic Moreau de m'avoir initié aux équations intégrales en me fournissant les références sur lesquelles me baser pour mon travail sur les données de Fresnel. Je remercie aussi les nombreuses personnes qui ont montré leur intérêt pour mon travail lors des conférences auxquelles je me suis rendu. En particulier, merci à Christelle Eyraud pour nos discussions et son aide sur les données de l'Institut Fresnel que j'ai utilisées pour valider ma méthodologie en milieu contrôlé. Merci aussi aux collègues de Jülich que je retrouve toujours avec plaisir lors des conférences GPR : Frédéric André, K. Z. Jadoon, Anja Klotzsche, Sébastien Lambot et Davood Moghadas. Merci enfin à Raphaël Valensi pour nos longues discussions téléphoniques sur le gradient, la polarisation, les équations intégrales... entre autres choses plus ou moins sérieuses.

J'ai eu la chance pendant ma thèse de pouvoir enseigner la géophysique aux étudiants de Licence et de Master de l'Université Joseph Fourier. J'ai ressenti cette activité d'enseignement comme une ouverture salutaire vers des domaines scientifiques extérieurs à mon sujet de thèse. J'aimerais donc remercier, en plus des étudiants qui ont servi de cobayes à mes premières

expériences pédagogiques, celles et ceux qui m'ont accompagné lors de ces expériences et avec qui je n'aurais pas eu la chance d'interagir dans le seul cadre de mes recherches. Merci à Claire Bouligand pour avoir accepté de superviser mon activité d'enseignement en tant que tutrice, et pour m'avoir confié deux années de suite un groupe de TD que j'ai pu suivre pendant tout un semestre. Merci à Laurent Charlet pour m'avoir confié l'organisation du volet géophysique de son UE *Paléo-environnements* pendant trois années de suite. Dans le cadre de cet UE, merci à Michel Dietrich pour le co-encadrement des étudiants et ses scripts et astuces SU, ainsi qu'à Jean-Pierre Deslandes pour nous avoir prêté main forte sur le terrain. Merci aussi à Frédéric Guiter (IMBE, Aix-Marseille Université) pour nos discussions pendant ces quelques jours passés sur le terrain, ainsi qu'à Sophie Guillon et Eric Pily (ENS, IPGP) pour leur accueil au laboratoire souterrain de Roselend au printemps 2014.

Je tiens également à remercier tous ceux qui ont contribué à faire des quatre années passées à ISTerre une période que je n'oublierai pas. Merci à Bertrand, l'autre co-colocataire du bureau 232, pour m'avoir initié à Awk et m'avoir rappelé, dans les moments de doute, que 8 est bien supérieur à 5. Merci aussi à Christelle, la co-bureau des derniers instants, pour avoir réussi à me faire sortir de mon bureau après le rendu de manuscrit et à me remettre au basket. Merci aux Chambériens, Aurore, Camille, Carlos, Fabien, Hervé, Laure, Léa, Marguerite, Sirel et Thomas, pour leur contribution au pot de thèse qui je crois a laissé des souvenirs. Et merci à tous les autres, croisés au hasard des couloirs, à la cafèt ou au RU: Anne (O. et B.), Aurore L., Gaby, Gwen, Jean, Marie, les Matthieu, Mélanie, les deux Pierre, Rachel, Simon, Vincent...

Parce que la vie d'un doctorant ne se résume pas (toujours) à la vie au labo, je pense aussi aux amis, proches ou lointains, qui ont parfois réussi à me faire parler d'autre chose que de mon sujet de thèse... ou à défaut m'ont patiemment écouté en parler. Merci à la bande des ours, Julie et Flo, Pierre et Estelle, Magalie, Tybain, ainsi qu'à Karim et Anne-So pour les sorties en montagne et les bonnes bouffes. Merci à Yann et François pour leur amitié à toute épreuve et leurs quelques visites à Grenoble. Merci à Catherine, Roman et leur petite Eléa pour leur accueil lors de mes passages éclairs à Paris. Merci enfin à la bande de Lyon pour les bons moments que nous arrivons encore à passer ensemble : Agnès, Baptiste, Camille, John, Laurie, Loraine, Mammouth, Marianna, Mimile, Papy, Pauline, Toto.

Pour finir, il est d'usage de réserver les derniers remerciements aux personnes qui comptent le plus. Je tiens donc ici à remercier ma famille, avec une pensée particulière pour mes grands-parents, pour mes grands-mères bien sûr, qui m'ont toujours témoigné leur fierté, et pour mes grands-pères, qui ne sont malheureusement plus là pour lire ces lignes de gratitude, et à qui je dédicace ce manuscript. Merci à mes parents qui m'ont toujours soutenu dans mes projets. Merci à ma mère pour son amour et son soutien discret mais constant. Merci à mon père, dont le parcours de médecin hospitalier, sans être un modèle puisque je ne l'ai pas suivi, a certainement été un exemple. Merci à mes frères, cousins, tantes et oncles qui ont suivi mon parcours de près ou de loin. Merci aussi à ma belle famille, pour son accueil si chaleureux. Enfin, quelques lignes ne suffiront pas pour exprimer toute ma gratitude envers Anaïs, mon amour, qui a su me supporter pendant ces années de thèse, dans tous les sens du terme. C'est maintenant à son tour de finir la sienne et j'espère pouvoir lui rendre la pareille pour les années à venir.

Résumé

Les premiers mètres à centaines de mètres de la proche surface terrestre sont le siège d'interactions environnementales et d'enjeux sociétaux qui requièrent une compréhension fine de ses processus par le biais d'estimations quantitatives de ses paramètres. Le géoradar est un outil de prospection indirecte de la subsurface à même d'ausculter les milieux naturels ainsi que les matériaux anthropiques, et d'en estimer les propriétés électriques (permittivité et conductivité). Basé sur la propagation d'ondes électromagnétiques, à des fréquences allant du MHz à quelques GHz, le géoradar est utilisé à des échelles et pour des applications variées : estimation de l'épaisseur de glaciers, évaluation de l'endommagement d'ouvrages en béton, ou encore quantification de la teneur en eau de sols agricoles.

Dans ce travail de thèse, je propose une méthode d'imagerie visant à quantifier les paramètres de permittivité et de conductivité sur des sections 2D de la subsurface, à partir de données radar acquises à la surface du sol. La technique mise en oeuvre est l'inversion des formes d'ondes, qui vise à utiliser l'intégralité du champ d'ondes enregistré. Après leur propagation à travers le milieu ausculté, les ondes radar contiennent en effet une information sur ses paramètres qu'il s'agit d'extraire via un processus d'inversion, en comparant les données observées à des simulations.

Dans une première partie, je présente les principes physiques et l'outil de simulation numérique utilisé pour simuler la propagation des ondes électromagnétiques dans les milieux hétérogènes à deux dimensions qui seront la cible de l'imagerie. J'utilise pour cela un algorithme de différences finies en domaine fréquentiel développé dans le cadre des ondes visco-acoustiques, que j'adapte au problème électromagnétique 2D grâce à une analogie mathématique entre les deux systèmes d'équations.

Dans une deuxième partie, le problème d'imagerie est formulé sous la forme d'une optimisation multi-paramètre puis résolu avec l'algorithme de quasi-Newton L-BFGS. Le choix de cet algorithme est motivé par sa capacité à estimer sans surcoût numérique l'effet de la matrice Hessienne, dont le rôle est crucial pour la reconstruction de paramètres de différents types comme c'est le cas du couple permittivité-conductivité. Des tests numériques montrent toutefois que l'algorithme reste sensible aux échelles utilisées pour définir les paramètres à optimiser, soulignant de ce fait les limites de l'approximation L-BFGS. Dans un exemple synthétique représentatif de la proche surface, il est cependant possible de reconstruire des cartes 2D de permittivité et de conductivité à partir de données de surface, en faisant intervenir à la fois un facteur d'échelle et un facteur de régularisation visant à contraindre les paramètres auxquelles l'inversion est la moins sensible. Ces facteurs peuvent être déterminés en analysant uniquement la qualité du fit aux données, sans hypothèse *a priori* autre que la contrainte de lissage introduite par la régularisation.

Dans une dernière partie, la méthode d'imagerie est confrontée à deux jeux de données réelles. Dans un premier temps, l'examen de données expérimentales permet de tester la précision des simulations numériques vis-à-vis de mesures de grande qualité effectuées en environnement contrôlé. La connaissance des cibles à imager permet de valider la méthodologie proposée pour l'imagerie multiparamètre dans des conditions très favorables puisqu'il est possible de calibrer le signal source et de considérer l'espace libre environnant les cibles comme modèle initial pour l'inversion. Dans un deuxième temps, j'envisage le traitement d'un jeu de données radar multi-offsets acquises au sein d'un massif calcaire. L'interprétation de ces données est rendue beaucoup plus difficile par la complexité du milieu géologique environnant, et donc des données elles-mêmes, ainsi que par la méconnaissance d'un certain nombre de facteurs : outre les propriétés de permittivité et de conductivité du milieu, le signal source et les caractéristiques précises des antennes utilisées ne sont pas connus avec précision. L'application de la méthode d'inversion des formes d'ondes à ces données requiert donc une étape préliminaire impliquant une analyse de vitesse plus classique, basée sur les arrivées directes et réfléchies, et des simulations numériques dans des modèles de subsurface hypothétiques à même d'expliquer une partie des données. L'estimation du signal source est effectuée à partir d'arrivées sélectionnées, bien expliquées par ces modèles. Simultanément, des valeurs pour la conductivité et la hauteur des antennes sont évaluées, de façon à reproduire au mieux les amplitudes observées. Malgré cela, les données synthétiques ainsi obtenues ne reproduisent pas l'évolution des amplitudes observées avec l'offset, ce qui suggère un effet du blindage sur le diagramme de rayonnement des antennes. Néanmoins, un premier essai d'inversion montre que l'algorithme est capable d'expliquer les données dans la gamme de fréquences considérée. Une ébauche des principaux réflecteurs est reconstruite, même à basse fréquence et en partant d'un modèle lisse. L'inversion des formes d'ondes a donc bien extrait de l'information des données pour nous livrer une image, très préliminaire, de ce milieu complexe.

Abstract

The quantitative characterisation of the shallow subsurface of the Earth is a critical issue for many environmental and societal challenges. Ground penetrating radar (GPR) is a geophysical method based on the propagation of electromagnetic waves for the prospection of the near subsurface. With central frequencies between 10 MHz and a few GHz, GPR covers a wide range of applications in geology, hydrology and civil engineering. GPR data are sensitive to variations in the electrical properties of the medium which can be related, for instance, to its water content and bring valuable information on hydrological processes. In this work, I develop a quantitative imaging method for the reconstruction of 2D distributions of permittivity and conductivity from GPR data acquired from the ground surface. The method makes use of the full waveform inversion technique (FWI), originating from seismic exploration, which exploits the entire recorded radargrams and has been proved successful in crosshole GPR applications.

In a first part, I present the numerical forward modelling used to simulate the propagation of electromagnetic waves in 2D heterogeneous media and generate the synthetic GPR data that are compared to the recorded radargrams in the inversion process. A frequency-domain finite-difference algorithm originally developed in the visco-acoustic approximation is adapted to the electromagnetic problem in 2D via an acoustic-electromagnetic mathematical analogy.

In a second part, the inversion scheme is formulated as a fully multiparameter optimisation problem which is solved with the quasi-Newton L-BFGS algorithm. In this formulation, the effect of an approximate inverse Hessian is expected to mitigate the trade-off between the impact of permittivity and conductivity on the data. However, numerical tests on a synthetic benchmark of the literature display a large sensitivity of the method with respect to parameter scaling, showing the limits of the L-BFGS approximation. On a realistic subsurface benchmark with surface-to-surface configuration, it has been shown possible to ally parameter scaling and regularisation to reconstruct 2D images of permittivity and conductivity without *a priori* assumptions.

Finally, the imaging method is confronted to two real data sets. The consideration of laboratory-controlled data validates the proposed workflow for multiparameter imaging, as well as the accuracy of the numerical forward solutions. The application to on-ground GPR data acquired in a limestone massif is more challenging and necessitates a thorough investigation involving classical processing techniques and forward simulations. Starting permittivity models are derived from the velocity analysis of the direct arrivals and of the reflected events. The estimation of the source signature is performed together with an evaluation of an average conductivity value and of the unknown antenna height. In spite of this procedure, synthetic data do not reproduce the observed amplitudes, suggesting an effect of the radiation pattern of the shielded antennas. In preliminary tests, the inversion succeeds in fitting the data in the considered frequency range and can reconstruct reflectors from a smooth starting model.

Contents

General introduction

Near surface geophysical imaging and characterisation

The first tens to hundreds metres of the Earth's subsurface is the location of natural phenomena of prior importance for the understanding of environmental interactions, the management of natural resources and the design of anthropic infrastructures. To understand processes such as landslides, aquifer recharge or pavement damaging, a fine characterisation of the near subsurface is crucial. Although near surface environments are close to us, they are not always directly accessible. Drilling a well or digging a trench to access directly the target of interest is time-consuming and expensive, destructive and thus non-repeatable, and it provides generally few local measurements. In near surface applications, it is often desirable to have a global view over large areas, and to preserve the object of study (e.g. archaeological remains or hydrological flows). Near surface geophysics aims at characterising the physical properties of shallow environments with indirect methods which enable a non-invasive prospection.

Geophysical methods are very helpful in many fields of environmental geosciences and geoengineering such as geomorphology (Schrott and Sass, 2008; Kneisel et al., 2008; Jongmans and Garambois, 2007), hydrogeophysics (Rubin and Hubbard, 2005; Vereecken et al., 2006), glaciology (e.g. Vincent et al., 2012), archaeology (Scollar et al., 1990; Hesse, 1999; Gaffney, 2008), and civil engineering (see Metje et al., 2007; or McCann and Forde, 2001, about non-destructive testing). Depending on the properties of interest and on the aim of the survey in terms of targets and penetration depth, various techniques are now available for environmental and engineering geophysicists: gravimetry can detect density anomalies, seismic methods (reflection, refraction, tomography, surface wave analysis) are sensitive to the mechanical properties of the ground, and a wide range of techniques involves its electrical properties (induced or spontaneous polarisation; electrical resistivity tomography, ERT) or its electromagnetic ones (magnetometry; time-domain reflectometry, TDR; electromagnetic induction, EMI; ground-penetrating radar, GPR; radiometry).

First of all, we generally wish to *see* the structures of interest, whether lithological, hydrological, or glaciological ones, which are inaccessible to our eyes: it is the purpose of geophysical imaging. In a second time, further interpretations require to quantify the physical properties involved in the mechanisms of interest – e.g. soil porosity, mechanical resistance or water content – as well as their variation in space (horizontal mapping or imaging at depth) and in time (monitoring). When speaking about imaging, we may thus distinguish between structural images, providing a qualitative view of the underground geometry, and property images, quantifying some physical attributes. This thesis will deal with the quantitative imaging of the subsurface using ground-penetrating radar data.

In most cases, geophysical prospection only provides an indirect view of the targets. In a qualitative approach, the spatial mapping of the measurements can already give a fairly good idea of the geometrical structures in the underground (see Fig. 1). To go further, a quantitative estimation of the property of interest, e.g. water content, can be derived from the measured quantity, e.g. dielectric permittivity from TDR measurements, through theoretical or empirical relations (Topp et al., 1980; Archie, 1942). In many cases, however, the searched parameters cannot be directly derived from the raw measurements. A non-trivial process is often required to infer the subsurface parameters from the acquired data, through the formulation of an *inverse problem*. Inverse problems concern all geophysical methods, and more generally all fields of science and technology where images or parameters must be inferred from observed data (e.g. medical imaging or meteorological prediction). Schematically, solving the inverse problem consists in finding the parameters of the considered system which best explain the data acquired during the observation. The inversion process requires the simulation of the system via a physical modelling. This simulation step is commonly referred to as the *forward problem* and generally involves numerical implementations to treat realistic complex cases. Contrary to the forward problem, which can be solved in a deterministic manner, the inverse problem is generally ill-posed and does not possess unique solutions. The resolution of the inverse problem thus invokes optimisation techniques to adjust the model parameters in order to fit the observed data with the simulated data. My thesis covers these two aspects: simulation and inversion of geophysical data, and more particularly of ground-penetrating radar data.

Ground-Penetrating Radar (GPR)

Principles of GPR measurements

For generalities on GPR, I refer to the book by Jol (2009) and to the recent review of Slob et al. (2010).

Ground-penetrating radar is a non-invasive subsurface prospecting technique based on the propagation of electromagnetic waves. Similarly to seismic reflection experiments, GPR measurements consist in emitting an electromagnetic signal into the ground with a transmitting antenna (or source), and in recording the resulting electromagnetic field after its propagation in the subsurface with a receiving antenna (receiver). This recorded signal contains the signature of the subsurface and analysing this signal provides information on the underlying structures. In geophysical applications, the central frequency of the emitted signal generally ranges between 10 MHz and a few GHz, enabling the investigation of objects at various scales, from hundreds of metres thick polar ice sheets to centimetres thick agricultural soils or concrete structures. The penetration depth of the method is mainly controlled by the central frequency of the antennas (the lower the frequency, the deeper the penetration) and by the electrical conductivity of the sounded medium (electromagnetic waves being rapidly attenuated in conductive media).

GPR measurements can take different configurations. In this work, I do not deal with satellite and airborne radar. Nor do I deal with off-ground GPR, which consists in frequency-domain measurements using vector network analysers, with the antenna placed at a given height above the ground ($\sim$ 1 m, see e.g. Lambot et al., 2006; Yedlin et al., 2009). In my work, I mainly focus on time-domain impulsional on-ground GPR. Most of my implementations can also be applied to crosshole GPR configurations, but I do not consider this kind of data in my

applications. The challenge of my work is precisely to bring full waveform inversion, which has been already applied to crosshole GPR data, to on-ground GPR data.

Despite the recent developments of multi-channel radar systems, most of classical on-ground GPR surveys consist in acquiring data over a profile using a single instrument, containing both the transmitting and the receiving antennas (*monostatic*, or *common-offset*, measurements in a nearby zero-offset configuration, see Fig. 1a). Modern acquisition systems now display in real-time the corresponding GPR common-offset section, also called radargram or B-scan, and even enable to improve the display by applying basic processing steps (filters, gain functions). The amplitude of the recorded signal along time then forms a qualitative image of the buried structures over the profile (see Fig. 1b). More precisely, it represents the reflectivity of the interfaces or obstacles encountered by the electromagnetic waves which propagated through the medium and were reflected or scattered back to the receiving antenna. For instance, in Fig. 1b, the contact between the glacier and the bedrock can be distinguished in the lower and upper parts of the profile. However, some artefacts on this image should not be understood as physical structures: It is the case of the hyperbolae visible around 140 m which result from the diffraction of the signal on punctual obstacles caused by an irregular water pocket. Moreover, such a raw common-offset section should not be interpreted directly as a geometrical representation of the medium, since some of the observed signal might come from out-of-plane reflections. Finally, its vertical axis is expressed in terms of recording time. To convert the axis in units of depth, and make the section representative from the geometry of the medium, the propagation velocity of the GPR signal in the medium must be known.

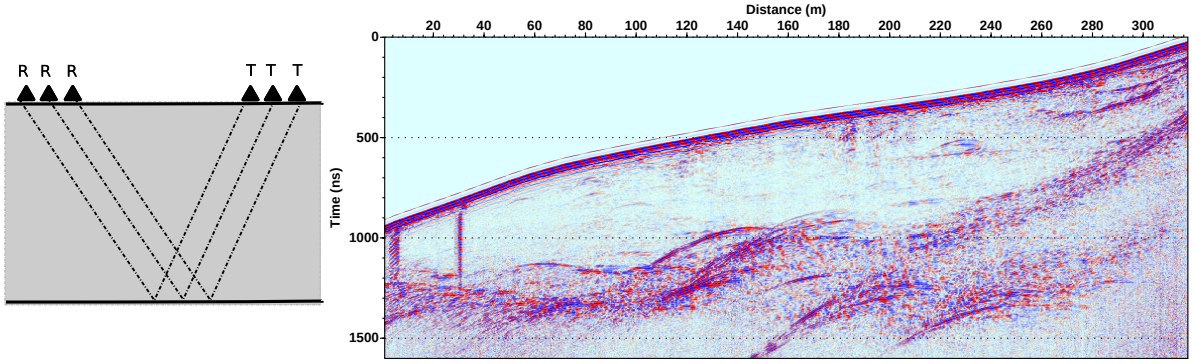


Figure 1: *During a GPR common-offset measurement, the transmitter T and the receiver R are moved together along the acquisition profile, while keeping fixed the offset between them (on the left, from Lopes, 2009). On the right: Example of a GPR common-offset section acquired on the Glacier de Tête Rousse (Mont Blanc massif, France) in May 2010 (Vincent et al., 2010).*

The case of Fig. 1 is rather simple since the glacier can be considered almost homogeneous and the velocity of electromagnetic waves in the ice is well known. Eventually, this velocity could also be determined from the diffraction hyperbolae whose shape depends on velocity. In the general case, however, common-offset GPR measurements are not a robust way to derive quantitative information on the medium because the presence of hyperbolae due to scattering is never guaranteed. A much more systematic method to estimate this velocity is the use of *bistatic*, or *multi-offset*, GPR measurements by varying the distance, or *offset*, between independent transmitting and receiving antennas (see Fig. 2). Such an acquisition can be achieved either by moving the receiver while letting the source fixed (we may then speak about

a *shot gather*, adopting the terminology of seismic reflection), or by moving the source while letting the receiver fixed (we then obtain a *receiver gather*), or by moving both the source and the receiver away from a Common Mid Point (*CMP gather*). Multi-offset measurements can be carried out if two independent GPR antennas are available but they are not routinely operated since they dramatically increase the duration of the survey, and thus decrease the potential of GPR to cover large areas in a reduced time. Though, multi-offset surveys possess the strong advantage of providing a much more quantitative information since they enable the estimation of the *dielectric permittivity* by velocity analysis (hyperbola fitting or semblance analysis) and of the *electrical conductivity* through Amplitude-Versus-Offset analysis (AVO, see e.g. Deeds and Bradford, 2002; Bradford and Deeds, 2006; Deparis and Garambois, 2009).

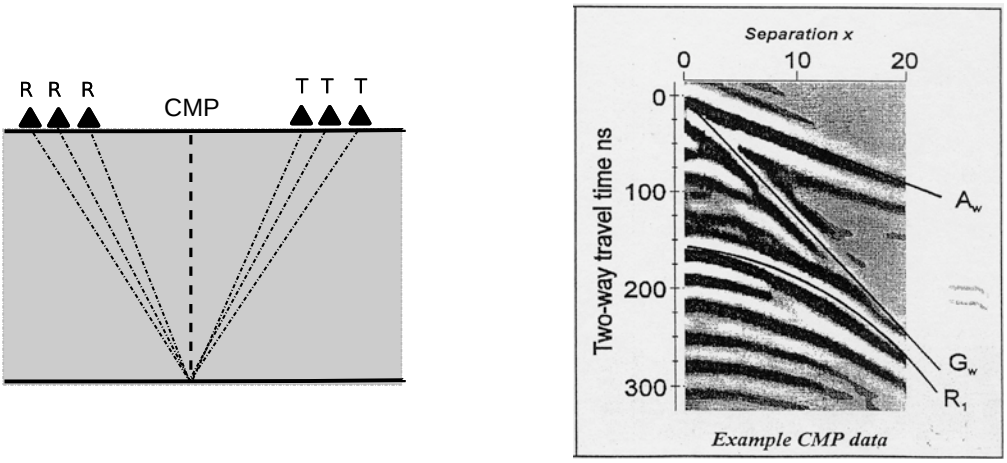


Figure 2: *During a GPR multi-offset measurement, transmitting and receiving antennas are moved away from a common mid point, CMP (on the left, modified from Lopes, 2009). On the right: Example of a CMP gather where the principal events are indicated: the direct air-wave A_W, the direct ground-wave G_W and a reflected wave R_1 (from Annan, 2001).*

GPR applications

Thanks to its multi-scale capability, GPR is widely applied in geosciences and geo-engineering. Historically, GPR was first employed in glaciological applications for estimating the thickness of polar ice sheets (Walford, 1964; Bailey et al., 1964; Bentley, 1964). Since electromagnetic waves propagate very well in the ice, GPR is still intensively used for this purpose (Saintenoy et al., 2013), but also to estimate the snow water equivalent (Bradford et al., 2009b), to monitor the evolution of the snowpack (Heilig et al., 2010), to image under-ice structures (Vincent et al., 2010) or even to detect buried avalanche victims (Instanes et al., 2004; Olhoeft and Modroo, 2006). In the same vein, I can also mention the exotic use of satellite-borne GPR to image the Martian polar ice sheets (Plaut et al., 2007).

In geological surveys, GPR enables to visualise the stratification in sediments (Davis and Annan, 1989; Bristow and Jol, 2003; Neal, 2004, for an exhaustive review), to characterise rock fractures (e.g. Grégoire, 2001; Grégoire and Hollender, 2004; Grasmueck et al., 2005; Jeannin, 2005; Jeannin et al., 2006; Deparis, 2007; Deparis and Garambois, 2009; Dorn et al., 2012), or to image active faults (e.g. McClymont et al., 2008; Beauprêtre, 2013; Beauprêtre et al., 2013). It is also used for mineral resources evaluation (see Francke, 2012, for a review).

Thanks to its great sensitivity to the presence of water, GPR is widely applied for hydrological purposes. In the emerging field of hydrogeophysics, it is particularly appreciated for its capability to prospect quickly at the field scale, bridging the gap between local measurements (core samples; time-domain reflectometry, TDR) and remote sensing at the regional scale (Rubin and Hubbard, 2005; Weihermüller et al., 2007). GPR is thus operated to evaluate the soil water content in the superficial part of agricultural soils (either using the ground-wave method, e.g. Sperl, 1999; Huisman et al., 2003; or more sophisticated inversions, e.g. Lambot et al., 2006), and to estimate the groundwater salinity (al Hagrey and Müller, 2000) or the contamination of the aquifer by organic pollutants (Deeds and Bradford, 2002; Poisson et al., 2009). It is also used to detect the water table level and to monitor water infiltration and flow paths (Doolittle et al., 2006; Saintenoy et al., 2008; Grasmueck et al., 2010; Moysey, 2010), as well as for tracers monitoring (Day-Lewis et al., 2003; Tsoflias and Becker, 2008).

In archaeology, GPR is employed as a quick and simple method to detect suitable locations for excavations (Vaughan, 1986). Today, refined imaging methods provide self-consistent results which make the expensive and destructive excavations less and less required (Conyers, 2007; Forte and Pipan, 2008; Böniger and Tronicke, 2010b).

GPR is also intensively exploited in civil engineering to detect buried anthropic structures in urban contexts (Goodman, 1994; Grandjean et al., 2000; He et al., 2009), to inspect road pavements (Evans et al., 2008), or to estimate the dielectric properties in concrete, which give an idea of their damaging (Bungey, 2004; Ihamouten et al., 2012). Many recent methodological developments have been done in this problematic, involving, among others, full waveform inversion (e.g. Kalogeropoulos et al., 2013).

Finally, GPR is applied as a detection method in forensic applications (see Pringle et al., 2012, for a review). It is for instance very useful for land mine detection (Daniels, 2006, 2008), since GPR can detect mines made of plastic, unlike classical metal detectors. It can also be used in avalanche victims detection (Modroo and Olhoeft, 2004; Olhoeft and Modroo, 2006; Fruehauf et al., 2009). Doppler methods even enable to detect the breathing and heartbeat of human beings under collapsed buildings (Grazzini et al., 2010).

GPR processing and imaging

Thanks to its similarity with seismic reflection, GPR benefits from most of the seismic processing methods developed for exploration geophysics (Yilmaz, 1987). A classical GPR processing flow hence consists in data filtering and/or deconvolution, velocity analysis on the CMP gathers by semblance analysis or hyperbola fitting, and migration (Cassidy, 2009b). GPR interpretation also involves the analysis of signal attributes (Cassidy, 2009b; Sassen and Everett, 2009; Böniger and Tronicke, 2010a) and textures (see Leparoux, 1997, or Moysey et al., 2006 for a tentative to quantify textural analysis).

Nevertheless, GPR also presents specific features due to its electromagnetic background. Cassidy (2009b) points out the following elements:

- electromagnetic waves generally undergo a greater degree of scattering and interference than seismic waves because the size of the heterogeneities in natural materials is closer to the incident wavelength,
- attenuation and dispersion effects are more extreme in GPR than in seismics, and should not be neglected. It is the reason why conductivity is considered in most GPR inversion

schemes, while seismic inversion often rely on lossless approximations. According to Turner and Siggings (1994), the distortion of the GPR signal also explains why the different deconvolution methods used in seismic surveys, based for the most part on the stationary phase principle, are not effective on radar data.

– the source wavelet of time-domain impulsional GPR systems, as well as the spatial distribution of the radiated energy, are more complex than for the seismic active sources used in the oil industry. The source signature and radiation pattern strongly depends on the type of the antenna (Arcone, 1995) but also on antenna-ground coupling which may vary over the profile (Lampe and Holliger, 2003; Maurer et al., 2012; Diamanti et al., 2013),

– depolarisation effects may occur during interface conversions (Lutz et al., 2003), in which case the common approximation of 2D acquisition and propagation is not valid anymore.

These points have important consequences on the design of accurate simulation tools. On the applicative point of view, GPR data processing may also be more site-dependent than seismic interpretation (Cassidy, 2009b).

Among the above mentioned interpretation methods, the migration step should appeal our attention since it is an imaging process in essence. The migration process consists in refocusing the scattered energy on local diffracting points in the medium. To do so, migration techniques rely on imaging conditions which always follow the fundamental principle of *time coincidence of up and down going waves*, enunciated by Claerbout (1971). This principle can be retrieved in many imaging methods, including seismic reverse-time migration but also its acoustic and electromagnetic equivalent of time-reversal mirrors (see e.g. Fink et al., 2000). As presented in Chapter 2, it is also the backbone of full waveform inversion.

However, a good knowledge of the velocity distribution in the medium is required to perform an accurate migration. Otherwise, correlations induce non-constructive interferences between incident and back-propagated fields and energy does not focus accurately at diffracting points. It results in artefacts and in a blurred image. As an example, Fig. 3 shows the section of Fig. 1 after migration using the ice velocity to define a homogeneous velocity model. On this image, hyperbolae on the bottom of the glacier and on the top of the water pocket (between 110 and 170 m) have been refocused such that the glacier-bedrock and glacier-water interfaces

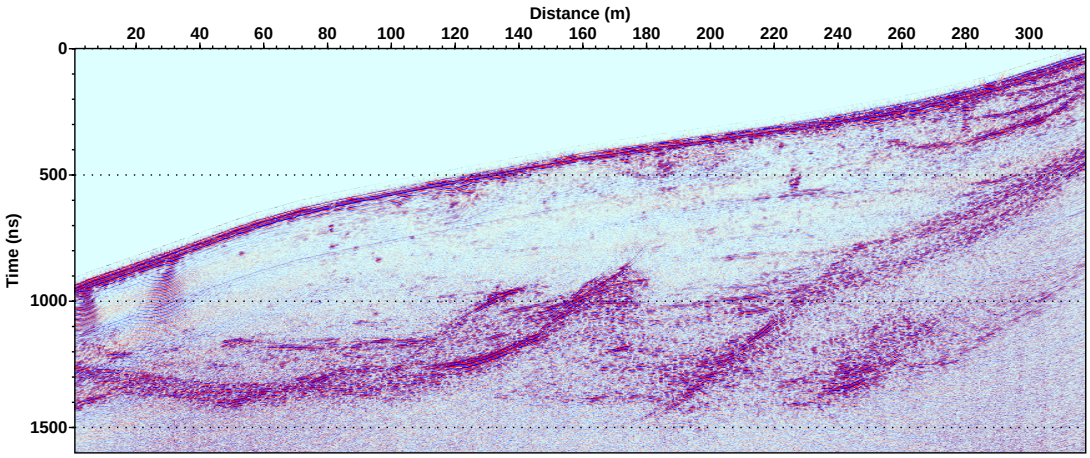


Figure 3: *Migrated GPR section of the Glacier de Tête Rousse.*

can be followed more clearly. Conversely, non-hyperbolic artefacts have been enhanced (see for instance at the beginning of the profile).

While migration constitutes the final step of an accurate *qualitative* imaging process, *quantitative* imaging requires the inversion of the recorded GPR data to infer the values of permittivity and conductivity in the subsurface. This major distinction leads Berkhout (2012) to differentiate *structural* and *property* images. Up to now, quantitative seismic methods based on multi-offset measurements have been adapted to GPR imaging using velocity analysis (e.g. Fischer et al., 1992a), amplitude-versus-offset studies (e.g. Deeds and Bradford, 2002; Deparis and Garambois, 2009) and travel-time and amplitude tomography (Cai et al., 1996; Holliger et al., 2001; Gloaguen et al., 2005; Musil et al., 2006; Hinz and Bradford, 2010). These methods, however, make use of a limited part of the data, as shown in Fig. 4 in the case of the AVO analysis of a reflected event. This limitation results in a limited number of targets, e.g. a given fracture for the AVO analysis of Fig. 4, or in a limited spatial resolution in the case of travel-time tomography. Moreover, first-arrival time tomography is more adapted to crosshole GPR data, using the arrival-time of the signal which is transmitted between the boreholes, than to on-ground, surface data (unlike seismic refraction data, on-ground GPR data do not contain any exploitable refracted event). Very few studies tackle the problem of reflection tomography (Bradford, 2006, 2008; Bradford et al., 2009a; Hinz and Bradford, 2010). Finally, both AVO analysis and travel-time tomography presume that remarkable phases (direct or reflected events) can be clearly identified, which can be difficult in highly heterogeneous media.

Unlike these methods, full waveform inversion is a quantitative imaging technique that aims at exploiting the integrality of the recorded GPR signal.

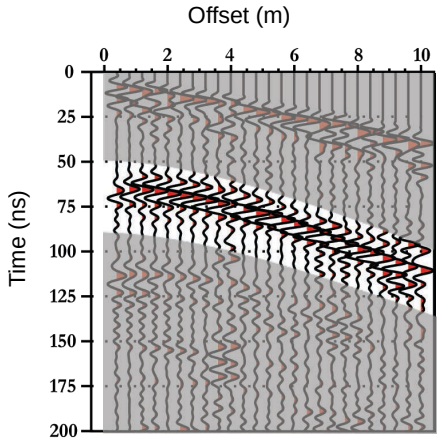


Figure 4: *AVO analysis of a reflected event for the quantitative characterisation of a fracture in terms of depth, aperture and filling properties (from Deparis and Garambois, 2009).*

Full Waveform Inversion (FWI)

Principles, history, and challenges

The full waveform inversion method originates from the time-domain seismic imaging (Lailly, 1983; Tarantola, 1984). It benefits from the adaptation of adjoint methods, which were developed in the meteorological data assimilation community (Lions, 1972) to geophysical problems (Chavent, 1974).

Full waveform inversion is an example of inverse problem where we seek the subsurface parameters which best explain the recorded waveforms. It is generally stated as an optimisation problem through the definition of a misfit function that measures the distance between the observed data and synthetic waveforms computed in a hypothetical subsurface model. In the aim of 2D or 3D imaging, the size of the parameter space generally prevents the use of global optimisation methods and the misfit function is minimised using local optimisation techniques. In complex, heterogeneous media, the local search involved in FWI may be trapped into local minima if a good starting model is not available. To tackle this issue, a frequency-domain formulation of the FWI has been developed (Pratt and Worthington, 1990; Pratt et al., 1998). The frequency-domain approach enables to mitigate the non-linearity of the inverse problem by following a low to high frequency hierarchy (Pratt and Worthington, 1990; Sirgue and Pratt, 2004). It also makes an efficient use of the data by inverting only few frequency components, taking benefit of the data redundancy provided by the acquisition (Sirgue and Pratt, 2004). Frequency-domain FWI is thus able to deal with the huge datasets of the oil industry. Synthetic and real seismic applications of this approach have been very successful (see Virieux and Operto, 2009, for an overview), both for hydrocarbon prospecting and for academic purposes (e.g. Ravaut et al., 2004; Operto et al., 2006; Prieux et al., 2013).

Nowadays, seismic FWI is still a dynamic field of investigations. All these developments are an enormous source of inspiration for adapting the FWI to GPR data.

FWI of GPR data: state of the art

The interest of FWI for GPR data has been first demonstrated for water content estimation in the first centimetres of agricultural soils from off-ground GPR data (e.g. Lambot et al., 2004, 2006; Jadoon et al., 2010; Minet et al., 2011) and for the estimation of permittivity and conductivity in stratified structures such as concrete (Kalogeropoulos et al., 2011, 2013; Patriarca et al., 2011) or layered soils (Busch et al., 2012, 2014). These studies are mostly formulated in the frequency-domain and restrict themselves to 1D geometries (multilayered media). As a consequence, they provide a robust parameter estimation and can account for dispersion, but they do not enable to consider arbitrary heterogeneous distributions of permittivity and conductivity in the subsurface.

A first step towards 2D imaging of permittivity and conductivity by FWI has been performed on GPR crosshole sections by Ernst et al. (2007), in the time-domain. These authors used a time-domain finite-difference modelling (Ernst et al., 2006) which enabled only a scalar representation of the fields. Permittivity and conductivity were reconstructed alternatively, by fixing conductivity when updating permittivity, and *vice-versa*. Their algorithm has then been improved by Meles et al. (2010) who introduce a full vector representation of the fields and a simultaneous reconstruction of permittivity and conductivity. Since then, the algorithm has been applied to various crosshole datasets for the imaging of the vadose zone (Klotzsche et al., 2010, 2012, 2013; Yang et al., 2013).

In parallel, some authors developed frequency-domain FWI schemes of GPR data. For instance, Yang et al. (2012) adapted the algorithm of Meles et al. (2010) in the frequency-domain. El Bouajaji et al. (2011) designed a frequency-domain FWI scheme based on a discontinuous Galerkin formulation of the forward problem and on a conjugate gradient optimisation technique. Ellefsen et al. (2011) propose to invert for the real and imaginary parts of the frequency-domain slowness in an alternating manner. Finally, I should mention the work of Cordua et al.

(2012) in the time-domain, who consider a coarse representation of the subsurface enabling the use of global optimisation methods to explore the parameter space and evaluate uncertainties on the estimated parameters.

Among the existing literature, only Saintenoy (1998), Lopes (2009) and El Bouajaji et al. (2011) tackle the interpretation of surface-based GPR measurements for the quantitative imaging of 2D sections of the medium. However, these authors restrict their investigations to monoparameter inversions, reconstructing only the permittivity distribution. In this work, I propose a FWI method for the simultaneous inversion of permittivity and conductivity in 2D, following the work of Ernst et al. (2007) and Meles et al. (2010), but with a particular interest in data acquired in surface-to-surface multi-offset configuration (on-ground GPR). Surface-based GPR measurements induce a limited illumination of the subsurface which makes the application of FWI particularly challenging.

Objectives of the thesis

The general aim of this work is the development of a full waveform inversion scheme for the quantitative imaging of permittivity and conductivity in 2D vertical sections of the subsurface from on-ground GPR data. To achieve this goal, several intermediate steps must be performed:

1. the accurate and efficient modelling of electromagnetic wave propagation in heterogeneous media,
2. the formulation of the full waveform inversion problem in a multiparameter framework,
3. the adequate data pre-processing and the design of an initial model for starting the FWI process from on-ground GPR data.

Questions related to these points are:

Q1. What numerical method should be used for the resolution of Maxwell's equations?

Q2. What parameters can be reconstructed from on-ground GPR data?

Q3. How to design an accurate starting model from surface-based measurements?

Q4. Which processing should be applied to real field data prior and during inversion?

Outline of the manuscript

The manuscript is organised with regards to the three main parts of my work, which also correspond to the three above-mentioned steps, namely the resolution of the forward problem (**Chapter 1**), the formulation of the multiparameter inverse problem (**Chapter 2**), and the application of the imaging method to real datasets (**Chapter 3**).

The topics related to this work concern a very wide panel, going from electromagnetic theory to GPR data processing, through computational modelling methods and numerical optimisation. Of course, I do not pretend to cover all these fields in an exhaustive way. Instead, I made efforts both to provide short reviews on the related issues but also to concentrate the message on the specific choices I have made in my effective work.

Chapter 2 and Chapter 3 include two articles that I have submitted during my PhD. For consistency, I have included them in the form they were submitted. As a consequence, there

are some redundancies between their introductory sections and other parts of the manuscript. On the other hand, the reader will find more technical details and discussions in the thesis than in these articles. In particular, a special care has been taken throughout the manuscript for validating the methodological tools I have adapted to the GPR FWI problem.

In **Chapter 1**, I begin with a short presentation of electromagnetism theory, with a focus on the electrical properties of natural media. I then introduce the numerical method used for the resolution of the forward problem, which consists in a particular – optimised – finite-difference scheme. I present the specificities of this scheme and validate my numerical implementation, whose accuracy is critical for further applications.

In **Chapter 2**, I consider the inverse problem formulation and its resolution with a focus on quasi-Newton schemes of optimisation which enable to properly address the multiparameter problem of reconstructing both permittivity and conductivity distributions in the subsurface. The main part of this chapter consists in a paper published in the *Geophysical Journal International* (Lavoué et al., 2014). This work presents the proposed methodology for multiparameter imaging. Following the paper, I provide additional details on technical points which could not be developed in the article. In particular, I derive the expression of the gradient of the misfit function and give illustrative examples of its signification.

Chapter 3 presents two applications of the proposed imaging method to real datasets. I first proceed to a validation of the methodology against well-controlled experimental data which were acquired at the Institut Fresnel (Marseille, France). This work, which has been submitted to the journal *Near Surface Geophysics* (Lavoué et al., 2015), is also the opportunity to confront the accuracy of the numerical solutions to real physical data. In a second section, I consider more complex field data for the imaging of a limestone reservoir at the Low Noise Underground Laboratory (LSBB, Rustrel, France). Field data requires first a thorough investigation involving a classical processing workflow (velocity analysis and migration) to design a suitable initial model and estimate the source signature before starting the FWI process. Preliminary results on this data set will illustrate the promise of applying FWI to on-ground GPR data, as well as the specific difficulties encountered to treat such data.

Finally, I conclude on the capacity of FWI to extract reliable information about permittivity and conductivity distributions from on-ground GPR data. This work offers many perspectives to improve the imaging procedure, concerning both the modelling and the inversion points of view.

Chapter 1

The forward problem

Contents

Introduction

In this chapter, I introduce the physical principles and the numerical implementations used to solve the forward problem in the inversion process, i.e. to simulate the propagation of electromagnetic waves in 2D heterogeneous media. As I am interested in quantifying electromagnetic parameters in the subsurface, I first detail how they are involved in electromagnetic equations, and how it is possible to parameterise them. I also highlight the basic assumptions considered in this work, in particular the restriction to a two-dimensional geometry. This assumption enables the use of an acoustic-electromagnetic analogy for adapting a numerical modelling algorithm that was originally developed for seismic purposes. I present the main features of this numerical tool and show validation results for my implementation of electromagnetic parameters.

1.1 Notions of electrodynamics in material media

This section introduces the basic concepts of electromagnetic theory and some specific points involved later in the numerical implementations. In particular, I will explain the implications of some assumptions which are commonly used but rarely fully commented in the literature. This presentation is mainly based on the books by Chew (1995) and Taflove and Hagness (2005), as well as on readings in Landau and Lifchitz (1969) and Feynman et al. (1979).

The equations of electrodynamics and their interpretation are first presented in the time domain, as it is our physical space. I then move to the frequency domain for the rest of the manuscript because, following Yuffa and Scales (2012), it is in the *"non-physical, but highly useful, Fourier domain"* that I will perform my simulations and inversions. The electromagnetic properties of natural media will be discussed, since they are those I want to quantify by inverting GPR data. I thus detail the dielectric response models that provide parameterisations for these properties. In particular, I comment the so-called universal response of Jonscher (1977), which is widely used in geophysical applications but in a truncated form which is rarely discussed. I also comment the consequences of assuming a simpler parameterisation, which will be used further for the inversion.

1.1.1 Maxwell's equations and constitutive relations

The behaviour of electromagnetic waves has been completely and coherently described by Maxwell (1873), based on the previous work of Faraday, Ampère and Gauss. Maxwell's equations can be expressed in a differential form as

$$\nabla \times \mathbf{E}(\mathbf{r},t) = -\frac{\partial \mathbf{B}(\mathbf{r},t)}{\partial t}, \quad \text{Maxwell-Faraday's equation,} \tag{1.1}$$

$$\nabla \times \mathbf{H}(\mathbf{r},t) = \frac{\partial \mathbf{D}(\mathbf{r},t)}{\partial t} + \mathbf{J}(\mathbf{r},t), \quad \text{Maxwell-Ampère's equation,} \tag{1.2}$$

$$\nabla \cdot \mathbf{D}(\mathbf{r},t) = q(\mathbf{r},t), \tag{1.3}$$

$$\nabla \cdot \mathbf{B}(\mathbf{r},t) = 0, \quad \text{Maxwell-Gauss' equations,} \tag{1.4}$$

where $\mathbf{E}$ is the electric field with an amplitude in V/m, $\mathbf{H}$ is the magnetic field in A/m, $\mathbf{D}$ is the electric induction (or electric displacement) in C/m^2, $\mathbf{B}$ is the magnetic induction in T, $\mathbf{J}$ is the conduction current density in A/m^2, and q denotes the electric charge density in C/m^3. The variable $\mathbf{r}$ is the position vector (with coordinates in m) and t denotes time (in s). Following Yuffa and Scales (2012), I stress that all fields and variables are real quantities here.

According to the so-called right-hand rule, Faraday's law indicates that a time-varying magnetic flux $\mathbf{B}$ generates an electric field $\mathbf{E}$ which rotates around $\mathbf{B}$. Similarly, Ampère's equation indicates that a current $\mathbf{J}$ or a time-varying electric flux $\mathbf{D}$ generates a rotating magnetic field $\mathbf{H}$. The physical mean of Maxwell-Gauss' equations is that an electric charge density q is the source of an electric flux $\mathbf{D}$, whereas an equivalent source of magnetic flux does not exist. The reader can refer to Chew (1995, §1.1.2, p. 3) for an integral form of Maxwell's equations and their interpretation.

Time-domain constitutive relations

Maxwell's equations system is very general but under-determined, because its equations are not linearly independent. For instance, assuming fields without constant components, we can derive the second Gauss' law $\nabla \cdot \mathbf{B} = 0$ (1.4) by taking the divergence of Faraday's equation (1.1), because $\nabla \cdot \nabla \times = 0$. Similarly, taking the divergence of Ampère's equation (1.2), and using the conservation law for the electrical charge

$$\nabla \cdot \mathbf{J}(\mathbf{r}, t) + \frac{\partial q(\mathbf{r}, t)}{\partial t} = 0, \tag{1.5}$$

returns the first Gauss' law $\nabla \cdot \mathbf{D} = q$ (1.3). In the following, I thus consider only Faraday's and Ampère's equations (1.1) and (1.2), and I omit the later Gauss' equations (1.3) and (1.4) which can be deduced from the former.

Mathematically, we need additional relations to determine the electromagnetic fields involved in Maxwell's equations. On a physical point of view, we can also notice that Maxwell's equations do not implicate explicitly the properties of the material media we want to investigate with GPR measurements. To represent the response of natural media and solve these equations, the induction vectors are related to the fields vectors through the *constitutive equations*. In vacuum, these relations are simply

$$\mathbf{D}(\mathbf{r}, t) = \varepsilon_o \mathbf{E}(\mathbf{r}, t), \tag{1.6}$$
$$\mathbf{B}(\mathbf{r}, t) = \mu_o \mathbf{H}(\mathbf{r}, t), \tag{1.7}$$

where $\varepsilon_o \simeq 8.85 \times 10^{-12}$ F/m is the dielectric permittivity of vacuum and $\mu_o = 4\pi \times 10^{-7}$ H/m is its magnetic permeability.

In dielectric material media such as rocks and soils, the electromagnetic response is more complex since imposing an electric field $\mathbf{E}$ to the material induces a polarisation, i.e. the orientation of the electrical moments of bounded charges (e.g. electrons linked to their atoms or dipolar molecules such as water) in a given direction. Similarly, applying a magnetic field induces a magnetisation, i.e. the orientation of the magnetic moments of magnetic particles. In linear, isotropic media, polarisation $\mathbf{P}$ and magnetisation $\mathbf{M}$ can be described as the moment vectors

$$\mathbf{P}(\mathbf{r}, t) = \varepsilon_o \chi_e(\mathbf{r}, t) * \mathbf{E}(\mathbf{r}, t), \tag{1.8}$$
$$\mathbf{M}(\mathbf{r}, t) = \mu_o \chi_m(\mathbf{r}, t) * \mathbf{H}(\mathbf{r}, t), \tag{1.9}$$

where $*$ denotes time convolution, χ_e is the dielectric susceptibility of the medium, which describes the capability of molecules to get an electrical polarisation when they are embedded in an electric field $\mathbf{E}$, and χ_m is the magnetic susceptibility, which is the capacity of the particles to get a magnetisation under a magnetic field $\mathbf{H}$. Induction vectors thus results from the response of vacuum (eqs 1.6 and 1.7), plus the *induced* polarisation and magnetisation of the medium, which can be summarised in the following constitutive equations:

$$\mathbf{D}(\mathbf{r}, t) = \varepsilon(\mathbf{r}, t) * \mathbf{E}(\mathbf{r}, t), \tag{1.10}$$
$$\mathbf{B}(\mathbf{r}, t) = \mu(\mathbf{r}, t) * \mathbf{H}(\mathbf{r}, t), \tag{1.11}$$

where $\varepsilon = \varepsilon_o(1 + \chi_e)$ is the dielectric permittivity of the medium (in F/m) and $\mu = \mu_o(1 + \chi_m)$ is its magnetic permeability (in H/m). In a conductive medium, an additional relation comes

from Ohm's law which relates the conductive currents $\mathbf{J}_c$ to the electric field $\mathbf{E}$ via

$$\mathbf{J}_c(\mathbf{r},t) = \sigma(\mathbf{r},t) * \mathbf{E}(\mathbf{r},t), \tag{1.12}$$

where σ is the electrical conductivity (in S/m). The total current $\mathbf{J}$ appearing in Ampère's equation (1.2) is then the sum of the conductive currents $\mathbf{J}_c$ generated by the electric field and of the source current $\mathbf{J}_s$ injected in the GPR antenna:

$$\mathbf{J}(\mathbf{r},t) = \mathbf{J}_c(\mathbf{r},t) + \mathbf{J}_s(\mathbf{r},t). \tag{1.13}$$

In practice, $\mathbf{J}_s(\mathbf{r},t)$ is non-zero only at the transmitting antenna position and during the emission of the GPR pulse.

The constitutive equations and Ohm's law constitute the material's relations which describe the response of the medium to the applied electric and magnetic fields. The electromagnetic parameters $\varepsilon(\mathbf{r},t)$, $\mu(\mathbf{r},t)$ and $\sigma(\mathbf{r},t)$ are thus the response functions of the medium along time, corresponding to dielectric, magnetic, and conductive mechanisms, respectively. I already mention that I consider linear, isotropic media. More fundamentally, I also implicitly assume that the polarisation $\mathbf{P}$ does not depend on the applied magnetic field $\mathbf{H}$ but only reflects the response of the material to the electric field $\mathbf{E}$. Such a dependence of polarisation on the magnetic field can be encountered in the optical frequency range where it implies magneto-optic effects but we will not consider them here. Moreover, I also assume that the polarisation $\mathbf{P}$ is *locally* related to the electric field $\mathbf{E}$ and does not suffer from spatial dispersion. In optics, spatial dispersion can be encountered in materials that are said to be optically active, such as some chiral molecules or crystals. Again, I disregard this possibility. Linearity and isotropy, however, are important assumptions about the properties of the medium and I shall now detail their implications.

Linearity In the constitutive relations (1.10) to (1.12), it is assumed that the medium responds linearly to the applied electric and magnetic fields. Mathematically, it is formalised by the fact that the parameters ε, μ and σ do not depend on the imposed fields $\mathbf{E}$ and $\mathbf{H}$. Physically, it means that the polarisation (/magnetisation) of the particles are proportional to the applied electric (/magnetic) field: polarisation and current vectors are co-linear to the electric field, while magnetic moments are co-linear to the magnetic field. The linearity of the response with respect to the excitation is a property of prior importance and I will often use it in my developments of simulation and inversion algorithms.

Isotropy The constitutive relations also consider an isotropic medium: electromagnetic parameters are scalar values and not tensors, so that the polarisation and magnetisation capability of the particles, as well as their conductivity, does not depend on the direction from which the applied fields are coming. Except in the work of Carcione (1996) and Carcione and Schoenberg (2000), who seek for generality, anisotropy is generally not considered in GPR applications. Indeed, the notion of isotropy depends on the scale at which the medium is described: in this sense, we may distinguish between an intrinsic anisotropy, which is described by a tensor, and a structural anisotropy, that results from the anisotropic arrangement of isotropic materials described by the relations (1.10) to (1.12).

Heterogeneity Electromagnetic parameters ε, μ and σ depend on the position in space $\mathbf{r}$, which simply means that the medium is heterogeneous. It is an obvious but nonetheless important feature to describe natural media, which can be very heterogeneous at the scale of GPR investigations.

Time-dependence These parameters also vary with time: natural media are non-perfect dielectrics, conductors and magnets. As a consequence, they present a *transient* response to the applied electric and magnetic fields. If we consider the polarisation phenomenon, it means that the individual electrical moments of charged particles do not align instantaneously with the vector $\mathbf{E}$ but present some inertia (relaxation process). The resulting macroscopic polarisation $\mathbf{P}$ is well aligned with the electric field, but its intensity varies with time until the orientations of all microscopic moments stabilise. To respect physical causality, i.e. the fact that the response (e.g. polarisation) can not precede the cause (the imposed electric field), the dielectric response function $\varepsilon(t)$ must be zero before the time t_o when the electric field is applied. The magnetic permeability and the electrical conductivity follow the same rule for the establishment of magnetisation and of conductive currents, respectively. The flow of free charges represented by the current $\mathbf{J}_c(\mathbf{r}, t)$ is thus proportional to the applied electric field, but its intensity can vary with time. In the case of electrolytic conduction such as occurring in natural media, a delay in the establishment of the currents can be due to the viscosity of the interstitial fluid carrying the ions.

Arrived to this point, I may specify that my work will focus on the electrical properties of the subsurface, i.e. on the dielectric permittivity ε and on the electrical conductivity σ, since natural media are generally non-magnetic. Formally, the magnetic permeability $\mu(\mathbf{r}, t)$ will depend on space and time in the equations and will be allowed to vary in my numerical implementations. But the reader can keep in mind that in my applications, permeability takes the constant value of vacuum $\mu_o = 4\pi \,.\, 10^{-7}$ H/m. By convention, one commonly refers to the *relative* permeability $\mu_r = \mu/\mu_o = 1$. Similarly, it is usual to speak about the relative permittivity $\varepsilon_r = \varepsilon/\varepsilon_o$.

Time-domain wave equation

In order to show the wave nature of electromagnetic fields, Faraday's equation (1.1) and Ampère's equation (1.2) can be combined with the constitutive relations (1.10) to (1.12). In the simple case of a homogeneous, time-invariant and source-free medium, it yields the following damped wave equation

$$\nabla^2 \mathbf{u}(\mathbf{r}, t) - \underbrace{\varepsilon\mu \frac{\partial^2 \mathbf{u}(\mathbf{r}, t)}{\partial t^2}}_{\text{propagation}} - \underbrace{\sigma\mu \frac{\partial \mathbf{u}(\mathbf{r}, t)}{\partial t}}_{\text{diffusion}} = 0, \tag{1.14}$$

with the field $\mathbf{u}$ being either $\mathbf{E}$ or $\mathbf{H}$. This equation first shows that the electric and magnetic fields are waves propagating at the same velocity $v = 1/\sqrt{\varepsilon\mu}$ and undergoing the same diffusion effects, with a diffusivity $\kappa = 1/(\sigma\mu)$. Considering plane waves of the form $\mathbf{u} \sim e^{\imath \mathbf{k}\cdot\mathbf{r}}$ and recalling Faraday's equation, it is also possible to retrieve the well-known rule that the fields $\mathbf{E}$, $\mathbf{H}$ and the wavenumber vector $\mathbf{k}$ are mutually orthogonal and form a right-handed system (Chew, 1995, §1.2.3, p. 13). More importantly, equation (1.14) gives us a first insight on the effects of permittivity and conductivity on the electromagnetic waves: permittivity controls the wave propagation, whereas conductivity appears in the diffusion term.

To obtain the wave equation (1.14), I assumed a homogeneous, time-invariant medium. It is also possible to derive a wave equation for inhomogeneous media (see Chew, 1995, §1.2.1, p. 11) but the assumption of time-invariance can not be dropped off. In time-varying media,

it is not possible to define a constant velocity v, which is generally difficult to account for in the time-domain. Mathematically, this difficulty comes from the time-convolution products involved in the constitutive relations, which make the material response non-local in time. To consider the effect of the transient electrical response on the electromagnetic waves more easily, it is convenient to switch to the frequency domain.

Frequency domain

Assuming time-harmonic wavefields, or performing a Fourier transformation of Maxwell's system with respect to time[1], enables to express equations (1.1) to (1.4) in the frequency domain:

$$\nabla \times \mathbf{E}(\mathbf{r},\omega) = \imath\omega\mathbf{B}(\mathbf{r},\omega), \tag{1.15}$$
$$\nabla \times \mathbf{H}(\mathbf{r},\omega) = -\imath\omega\mathbf{D}(\mathbf{r},\omega) + \mathbf{J}(\mathbf{r},\omega), \tag{1.16}$$
$$\nabla \cdot \mathbf{D}(\mathbf{r},\omega) = q(\mathbf{r},\omega), \tag{1.17}$$
$$\nabla \cdot \mathbf{B}(\mathbf{r},\omega) = 0, \tag{1.18}$$

where the fields are now complex quantities depending on the angular frequency, or pulsation, ω (in rad/s). By convention, I assume here a time-harmonic dependency in $e^{-\imath\omega t}$. The corresponding Fourier transformations to pass from the time to the frequency domains are

$$f(\omega) = \int_{-\infty}^{+\infty} f(t) e^{\imath\omega t} dt, \tag{1.19}$$
$$\text{and} \quad f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(\omega) e^{-\imath\omega t} d\omega, \tag{1.20}$$

where f denotes the function of interest. Please note that I use the same notation for time-domain and frequency-domain quantities. I shall also underline that this Fourier convention leads to opposite signs for the imaginary parts of complex quantities compared to the notations used, for example, by Hollender and Tillard (1998) and Taflove and Hagness (2005). It is the same convention as in Press et al. (1992), Chew (1995) and Virieux (1996).

In the frequency domain, the constitutive equations and Ohm's law read

$$\mathbf{D}(\mathbf{r},\omega) = \varepsilon(\mathbf{r},\omega)\,\mathbf{E}(\mathbf{r},\omega), \tag{1.21}$$
$$\mathbf{B}(\mathbf{r},\omega) = \mu(\mathbf{r},\omega)\,\mathbf{H}(\mathbf{r},\omega), \tag{1.22}$$
$$\mathbf{J}_c(\mathbf{r},\omega) = \sigma(\mathbf{r},\omega)\,\mathbf{E}(\mathbf{r},\omega), \tag{1.23}$$

where time-convolution products have been replaced by multiplication, which greatly simplify the consideration of the transient response. Since they are the Fourier coefficients of the corresponding real-valued response functions in the time-domain, the electromagnetic properties $\varepsilon(\mathbf{r},\omega)$, $\mu(\mathbf{r},\omega)$ and $\sigma(\mathbf{r},\omega)$ are now frequency-dependent complex quantities. Their imaginary parts account for the energy dissipation occurring during the transient response. Due to their frequency dependency, the medium is said to be *dispersive*. Dispersion and dissipation are

[1] As pointed out by Yuffa and Scales (2012), both are not exactly equivalent. Considering a monochromatic field implicitly assumes that the system reached the steady-state and thus does not enable to study the transient response of the medium. Conversely, passing from time to frequency domains by direct and inverse Fourier transformations does not induce any loss of generality. This subtlety does not have direct consequences on my work but we should be aware of it when working in the frequency domain.

thus indivisible frequency-domain mechanisms describing the time-dependent response of the medium (Toll, 1956; Yuffa and Scales, 2012).

To complete the overview, I have to mention that the real and imaginary parts of the frequency-domain parameters should be linked through the Kramers-Kronig relations in order to verify physical causality, just like time-domain response functions must be zero before the application of any imposed fields (see e.g. Sohl, 2008; Yuffa and Scales, 2012). Strictly speaking, these relations have two consequences. First, it means that the knowledge of one of the two parts (real or imaginary) over the whole frequency spectrum is sufficient to derive the other one from the Kramers-Kronig relations. In practice, however, measurements are limited by the frequency bandwidth of the instrument and this derivation is rarely possible. Secondly, it implies that electromagnetic parameters can not be constant over the whole frequency spectrum, since they are the Fourier transforms of the non-constant time-domain response functions, which have to be zero before t_o to respect causality, and non-zero during the response (except in vacuum). Again, GPR applications generally forget about this physical consideration, due to the limited frequency-bandwidth of the measurements. We shall see in Section 1.1.2.2 in what extent the approximation of constant parameters might be valid in the GPR frequency range.

In the frequency-domain, the propagation and diffusion terms of the wave equation can be gathered into a unique dielectric response, re-writing equation (1.14) under the form of the Helmholtz equation

$$\nabla^2 \mathbf{u}(\mathbf{r},\omega) + \varepsilon_e(\omega)\mu(\omega)\omega^2 \mathbf{u}(\mathbf{r},\omega) = 0, \tag{1.24}$$

where ε_e is an *effective* permittivity which gathers both permittivity and conductivity:

$$\varepsilon_e(\omega) = \varepsilon(\omega) + \frac{\imath\sigma(\omega)}{\omega}. \tag{1.25}$$

Note that, in the frequency domain, the derivation of Helmholtz equation does not require the assumption of a non-dispersive medium as it was the case for the time-domain wave equation (1.14). Frequency-domain wave propagation modelling thus enables to consider dispersion in a straightforward manner, solving Helmholtz equation frequency per frequency. For simplicity, I considered a homogeneous medium to derive equation (1.24), which is in fact a scalar wave equation: In homogeneous media, the field components u_x, u_y and u_z can be considered as independent scalar values and the corresponding Helmholtz equations can be decoupled and solved independently. In heterogeneous media, we can still derive wave equations for the electric and magnetic fields $\mathbf{E}$ and $\mathbf{H}$, but we end up with vector wave equations where field components are not independent anymore (see Chew, 1995, §1.3, p. 17).

Since I am mainly interested in the permittivity and conductivity parameters, and less in the permeability, the question is now how to describe the effective permittivity. In the next section, I develop some models which have been proposed for this purpose. Working in the frequency domain *a priori* enables to finely describe the dielectric response of natural media in the simulations. I should already mention, however, that I will consider frequency-independent, real-valued permittivity and conductivity parameters in the inversion, as done by Meles et al. (2010) in the time domain. Indeed, we shall see in Chapter 2 that the 2D multiparameter imaging of frequency-independent, real-valued permittivity and conductivity distributions is already an ill-posed problem whose resolution is challenging. Accounting for dispersion still increases the number of degrees of freedom and the ill-posedness of the problem. Moreover, GPR data are often mainly sensitive to permittivity, in less extent to conductivity, and in

minor degrees to dispersion. Nevertheless, it is important to understand the general behaviour of dielectric materials to realise the implications of further simplifications.

1.1.2 The electrical response of dielectric solid materials

1.1.2.1 Dielectric response models

As already mentioned, the frequency-domain permittivity and conductivity can be regarded in the case of a linear, isotropic, heterogeneous, dispersive medium as space- and frequency-dependent complex quantities. Thus, we can write them as functions of real quantities as

$$\varepsilon(\mathbf{r},\omega) = \varepsilon'(\mathbf{r},\omega) + \imath\varepsilon''(\mathbf{r},\omega), \tag{1.26}$$

$$\text{and} \quad \sigma(\mathbf{r},\omega) = \sigma'(\mathbf{r},\omega) - \imath\sigma''(\mathbf{r},\omega), \tag{1.27}$$

where the imaginary parts ε'' and σ'' reflect the energy dissipation occurring during the transient response of the medium to the applied electric field. The opposite signs before imaginary parts come from our convention for Fourier transformation (eqs 1.19 and 1.20) and have been chosen such that imaginary parts are positive ($\varepsilon', \varepsilon'', \sigma', \sigma'' \in \mathbb{R}^+$). In the complex plane, dissipation manifests itself by the fact that the response of the medium, e.g. the induction vector $\mathbf{D}$, is not in phase with the applied electric field $\mathbf{E}$.

The effective permittivity can be re-written in terms of the real and imaginary parts of permittivity and conductivity as

$$\varepsilon_e = \underbrace{\left(\varepsilon'(\omega) + \frac{\sigma''(\omega)}{\omega}\right)}_{\varepsilon'_e} + \imath \underbrace{\left(\varepsilon''(\omega) + \frac{\sigma'(\omega)}{\omega}\right)}_{\varepsilon''_e}. \tag{1.28}$$

The total energy losses are then quantified by the loss tangent $\tan\delta = \varepsilon''_e/\varepsilon'_e$. If we neglect the conductive losses for simplicity ($\sigma'' = 0$, I shall justify this hypothesis later on), we have

$$\tan\delta(\omega) = \frac{\varepsilon''(\omega)}{\varepsilon'(\omega)} + \frac{\sigma'(\omega)}{\omega\varepsilon'(\omega)}, \tag{1.29}$$

where the first term corresponds to dielectric losses due to the relaxation of bounded charges and the second term to conductive losses due to the displacement of free charges.

Expression (1.28) has two consequences. First, we can note that the imaginary part of the permittivity plays a similar dissipative role as the real part of the conductivity. Conversely, the influence of the imaginary part of conductivity might be interpreted as a *propagative* effect. Secondly, we can anticipate that this type of parameterisation is largely not optimal in view of the inversion as it would require to estimate four values per frequency (for $\varepsilon'(\omega)$, $\varepsilon''(\omega)$, $\sigma'(\omega)$ and $\sigma''(\omega)$). Moreover, inside the in-phase (ε'_e) and out-of-phase (ε''_e) parts of the dielectric response, we can not discriminate between the contributions of the real and imaginary parts of permittivity and conductivity ($\varepsilon'(\omega)$ vs. $\sigma''(\omega)/\omega$, and $\sigma'(\omega)/\omega$ vs. $\varepsilon''(\omega)$). Indeed, even in a modelling point of view, expression (1.28) is not very adequate. It is certainly exhaustive since all theoretical terms appear explicitly, and therefore it can fit all possible measurements of dielectric relaxation. However, it is not very useful to explain the behaviour of the materials in terms of general processes (Jonscher, 1999).

For these reasons, other models have been proposed for the effective permittivity, involving a limited number of parameters. These models are based on empirical laws derived from dielectric measurements (Jonscher, 1999). Some of them also rely on theoretical assumptions about the underlying relaxation processes (Jonscher, 1981). These dielectric response models take the generic form of

$$\varepsilon_e(\omega) = \varepsilon_\infty + \varepsilon_o \chi(\omega) + \imath \frac{\sigma_{DC}}{\omega}, \tag{1.30}$$

where ε_∞ is an asymptotic limit for permittivity at high frequencies (considered to be real), σ_{DC} is the (real) static conductivity, and $\chi(\omega)$ is a complex electrical susceptibility which characterises dispersion and dissipation due to both dielectric and conductive frequency-dependent phenomena. Note that this frequency-domain susceptibility χ does not strictly correspond to the Fourier transform of the time-domain response χ_e. On one hand, χ does not include the constant asymptotic permittivity ε_∞ ($\neq \varepsilon_o$). On the other hand, it includes all frequency-dependent parts of the conductivity, because these effects can not be distinguished from the dielectric ones[1].

The different models I will now describe only vary by their definition of the electrical susceptibility χ, according to the assumptions done on the underlying polarisation mechanisms. Polarisation in solid materials is generally described as a relaxation process but there are many possible models to formalise it. The following explanations are mostly inspired by the enlightening reviews of Jonscher (1981, 1999).

Debye-type models

A first class of dielectric models assume that the electrical polarisation can be explained by a Debye-type relaxation, with an electrical susceptibility of the form

$$\chi(\omega) = \frac{1}{\varepsilon_o} \frac{\varepsilon_s - \varepsilon_\infty}{1 + \imath\omega\tau}, \tag{1.31}$$

where ε_s is a (real) static permittivity valid at low frequency and τ is the characteristic relaxation time for the considered mechanism. The Debye model is one of the first and simplest attempts to characterise relaxation (Debye, 1929) and also find applications in the mechanics of viscous media. Debye's model considers the individual electrical moments as identical, non-interacting dipoles, having a loss of energy proportional to frequency[2]. Debye's model is particularly suitable for describing the polarisation of dipolar molecules such as water, or the polarisation of ions at interfaces (e.g. between soil particles and water), inducing an interfacial capacitance (Maxwell-Wagner phenomena).

Fig. 1.1 shows the evolution of the real and imaginary parts of the effective permittivity with respect to frequency, for the Debye model of pure water at 25°C (Cassidy, 2009a). The imaginary part of the permittivity ε'' exhibits a clear peak of dissipation at the relaxation frequency $\omega_p = 1/\tau$ corresponding to the characteristic time τ ($\omega_p \simeq 19$ GHz in this case), whereas the real part ε' drops down to its high frequency limit $\varepsilon_{\infty_r} = 5.6$.

[1] The interpretation of the electrical behaviour of the medium in terms of a *dielectric* response is therefore mainly a matter of convention. People interested in the conductive phenomena which dominate at low frequencies may rather choose to express this behaviour in terms of an *electrical* response, involving an effective conductivity $\sigma_e = -\imath\omega\varepsilon_e$.

[2] This can be roughly understood by analogy with a viscous medium where the particles do not have time to vibrate if the excitation varies too quickly. This results in internal frictional losses: the energy of the excitation is not released through a movement but through the dissipation of heat.

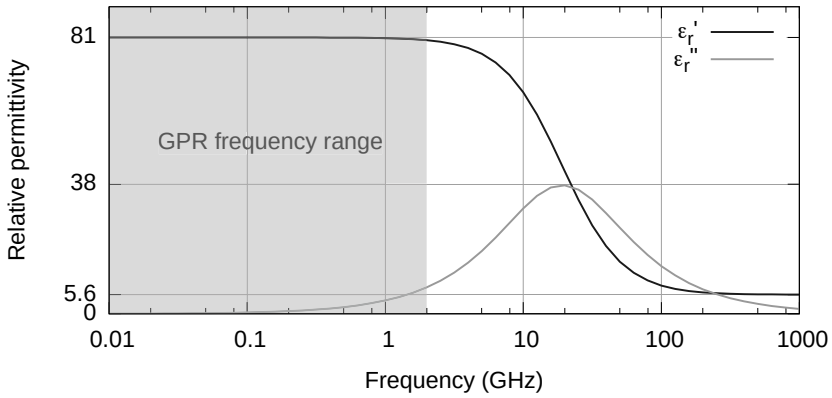


Figure 1.1: *Evolution of the real and imaginary parts of the dielectric permittivity of water according to the Debye model, with* $\varepsilon_{s_r} = 81$, $\varepsilon_{\infty_r} = 5.6$ *and* $\tau = 8.28 \times 10^{-12}$ *s (parameters from Cassidy, 2009a, p. 50).*

According to the considered frequency range, different relaxation mechanisms can be involved (see Fig. 1.2). In the GPR frequency range, dielectric losses at high frequencies are mainly due to dipolar relaxation: it is one of the reasons why GPR data are sensitive to the presence of water in the investigated material (the primary reason being the value $\varepsilon_{s_r} = 81$ itself, much larger than in any other natural media, see Table 1.1). In a smaller extent, we may observe dielectric losses at low frequencies due to Maxwell-Wagner interfacial phenomena (indicated as ionic processes in Fig. 1.2) but these effects are generally dominated by the conductive losses due to the static electrical conductivity of the medium. Actually, most of the

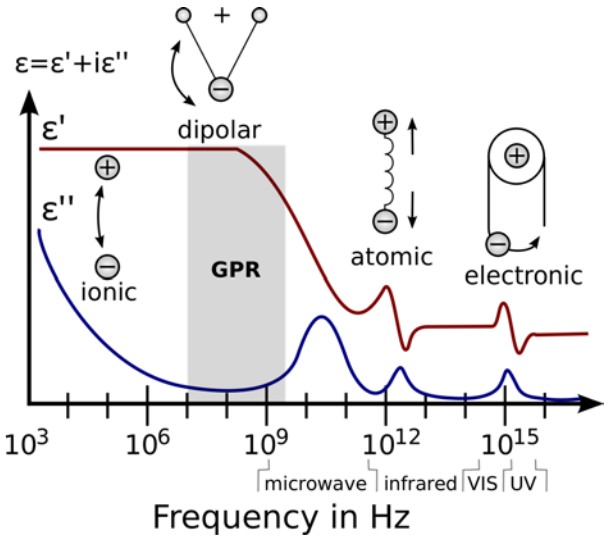


Figure 1.2: *Polarisation processes and dielectric permittivity at different frequencies (adapted from* `http://en.wikipedia.org/wiki/File:Dielectric_responses.svg`*).*

GPR frequency band lies outside any Debye-type relaxation peaks (see Fig. 1.2). And it turns out that, apart from this particular loss peaks, pure Debye's models fail to accurately describe the dielectric behaviour in most of real materials (Jonscher, 1999).

Some authors proposed finer descriptions of the dielectric properties at GPR frequencies involving variants of Debye's model, for instance the Cole-Cole model (Cole and Cole, 1941) or the Cole-Davidson model (Davidson and Cole, 1951). These parameterisations introduce additional exponent factors in the denominator of expression (1.31) that make the dispersion frequency range wider and roughly correspond to giving a *memory* to the relaxation process (Hill and Dissado, 1985). Other authors promote the superposition of several Debye-type peaks (Xu and McMechan, 1997) or the combination of Debye-type models with other relaxation mechanisms (e.g., the Kelvin-Voigt mechanical model, see Carcione, 1996). I will not detail the variety of possible models because the variations between them are not significant for our purposes. As for every empirical law, the preference for a model or another is case-dependent. On a theoretical point of view, Jonscher (1999) further argues that designing sophisticated variants of the Debye model and invoking superposition or combination of different relaxation mechanisms is not satisfactory: of course, it enables to better explain the observed dielectric behaviours (by adjusting more and more tuning parameters) but it fails to provide a unified explanation of the underlying physical processes.

Alternatively, Jonscher (1977) proposed a *universal* response model which is claimed to explain all the observed electrical behaviours with a unified physical mechanism, involving a limited number of independent parameters. Jonscher's model has been promoted by Hollender and Tillard (1998) for geophysical investigations in the GPR frequency band. Since then, it has been used to describe the dielectric properties of rocks (Hollender and Tillard, 1998), sands (Grégoire and Hollender, 2004) or concrete (Bourdi et al., 2008; Ihamouten et al., 2011). Deparis and Garambois (2009) successfully applied it for the characterisation of thin layers in fractured limestones. In the next section, I detail this model, starting from the general form of Jonscher (1981) to understand its meaning and limitations.

Jonscher's model

A general formulation of Jonscher's model is (Jonscher, 1981, eqs 21 to 23):

$$\left.\begin{array}{rcl}\chi'(\omega) & \propto & \chi(0)-\omega^m \\ \chi''(\omega) & \propto & \omega^m\end{array}\right\} \qquad \text{at low frequencies} \quad \omega \ll \omega_p, \tag{1.32}$$

$$\chi(\omega) \propto \frac{1}{1+\imath\omega\tau} \qquad \text{at intermediate frequencies} \quad \omega \simeq \omega_p, \tag{1.33}$$

$$\text{and} \quad \chi(\omega) \propto (\imath\omega)^{n-1} \qquad \text{at high frequencies} \quad \omega \gg \omega_p, \tag{1.34}$$

where $\tau = 2\pi/\omega_p$ is the characteristic time of some Debye-type relaxation mechanism, ω_p denoting the corresponding peak frequency. The exponent factors m and n characterise two regimes of the dielectric response outside the relaxation frequency range, where the susceptibility takes the form of a power-law.

We can notice that Jonscher's model does not evacuate Debye's model but integrates it (eq. 1.33). Jonscher's model implies a Debye's mechanism to describe the dielectric response over the whole spectrum since the power laws (1.32) and (1.34) alone can not reproduce the relaxation peak of dipolar molecules or interfacial polarisation, which are best explained by Debye's model. The improvement of Jonscher's model concern the description of the electrical response outside the frequency range of this Debye-type relaxation, where the power laws of eqs (1.32) and (1.34) have been shown to better fit the observed dielectric measurements than the flat response of Debye's model (Jonscher, 1981, 1999).

It is worth noting that, in their adaptation of Jonscher's model to GPR measurements, Hollender and Tillard (1998) only keep the high frequency regime of the general Jonscher's model. They end up with an expression which verifies equation (1.34), of the form

$$\chi(\omega) = \chi_r \left(\frac{\omega}{\omega_r}\right)^{n-1} \left(1 + \imath \mathrm{cotan}\frac{n\pi}{2}\right), \tag{1.35}$$

where χ_r is this time a real and constant susceptibility parameter and ω_r is a reference frequency which can be defined arbitrarily (it only aims at making the ratio ω/ω_r dimensionless). For GPR applications, it is usual to consider $\omega_r = 2\pi \times 100$ MHz.

Since Hollender and Tillard (1998) discarded the low-frequency part of the law, as well as the Debye-type relaxation peak, it should be stressed that expression (1.35) is only valid in the frequency regions above any Debye-type loss peaks. In particular, it can not describe the polarisation of water molecules. However, as most of the GPR frequency band lies outside any Debye-type peaks (see Fig. 1.2), expression (1.35) provides a good agreement with dielectric measurements at intermediate GPR frequencies, typically between 50 and 300 MHz, in particular in dry rocks (Hollender and Tillard, 1998). But it should be manipulated with cautious when considering conductive materials such as clay at low frequencies (close to the Wagner-Maxwell relaxation peak) or wet rocks and soils at high frequencies (nearby the Debye peak of water relaxation).

Finally, some authors suggest that, using expression (1.35), the static conductivity σ_{DC} may also be included in the electrical susceptibility χ (e.g. Deparis, 2007; Lopes, 2009). This is a pragmatic assumption that aims at decreasing the number of parameters required for the description of the effective permittivity. However, neither Jonscher (1977, 1981, 1999) nor Hollender and Tillard (1998) make this assumption. Hollender and Tillard (1998) simply neglect the contribution of the static conductivity for their applications at GPR frequencies. As for Jonscher (1999, §4.6), he clearly distinguishes between DC conduction and low frequency dispersion. The latter implies a reversible storage of charge in the material and can indeed be described by the universal power-law. Static conductivity, however, consists in flowing charges, without any storage. In the generic expression of the effective permittivity (1.30), I will thus keep the static conductivity decoupled from the frequency-dependent permittivity term:

$$\varepsilon_e(\omega) = \underbrace{\varepsilon_\infty + \varepsilon_o\chi(\omega)}_{\varepsilon(\omega)} + \imath\frac{\sigma_{DC}}{\omega}, \tag{1.30 again}$$

where I remind that the quantities ε_∞ and σ_{DC} are real-valued constants, while the susceptibility $\chi(\omega)$ is a complex variable. Following Jonscher (1981), I then conceive the general, frequency-dependent conductivity as

$$\begin{aligned} \sigma(\omega) &= \sigma_{DC} + \omega\varepsilon''(\omega), & (1.36)\\ &= \sigma_{DC} + \underbrace{\omega\varepsilon_o\chi''(\omega)}_{\sigma_{AC}(\omega)}, & (1.37) \end{aligned}$$

which is now a real quantity (justifying the hypothesis $\sigma'' = 0$ made p. 34), with $\chi''(\omega)$ following the universal response and σ_{DC} being an independent parameter. Again, I shall stress that this decoupling is mainly a matter of interpretation. It only means that I distinguish the flow of (totally) free charges described by the direct current conductivity σ_{DC}, from the complex,

frequency-dependent electrical response $\varepsilon(\omega)$ of (more or less) bounded charges that can give rise to an alternating current (AC). The loss tangent is then simply

$$\tan\delta(\omega) = \frac{\varepsilon_e''(\omega)}{\varepsilon_e'(\omega)} = \frac{\sigma(\omega)}{\omega\varepsilon'(\omega)}. \tag{1.38}$$

Now that I have explained what is exactly an electrical response, i.e. what permittivity $\varepsilon(\omega)$ and conductivity $\sigma(\omega)$ represent, we can have a look at the values of these parameters in natural media. In particular, the values of the dispersive parameters χ_r and n in expression (1.35) will tell us if the assumption of non-dispersive materials is reasonable for considering natural media.

1.1.2.2 Electrical properties of natural media

Table 1.1 presents the electrical properties of some materials, compiled from the PhD thesis of Saintenoy (1998), Girard (2002), Jeannin (2005) and Loeffler (2005). More exhaustive lists of these parameters can be found in these studies, as well as on the internet (e.g. Wikipedia, 2014; Clipper Controls®, 2014). These references give either rough orders of magnitude for the real part of the relative permittivity ε_r' and for conductivity σ in the GPR frequency range, or the corresponding Jonscher parameters. When the Jonscher parameters are given, I derive the corresponding values for $\varepsilon_r'(\omega)$ and $\sigma(\omega)$ for two characteristic frequencies of GPR investigations, namely 100 and 200 MHz. I also derive the ratio $\sigma_{DC}/(\varepsilon'\omega)$ that gives an indication of the proportion of diffusive vs. propagative effects in the behaviour of electromagnetic waves (see eqs 1.14 and 1.29). Small values for this ratio justify to neglect the static conductivity in the description of the medium at GPR frequencies, as argued by some authors (Hollender and Tillard, 1998; Jeannin, 2005). We shall see wether or not this assumption is reasonable.

A first look at Table 1.1 enables to distinguish three classes of materials:

1. low-loss media, such as snow, ice, fresh water, quartzite or dry sand. In these media, electromagnetic waves can propagate over very long distances quasi without attenuation other than geometrical spreading (GPR surveys can thus achieve deep penetration depths). It is then possible to ignore the effect of conductivity. The extreme case being air, that behaves like vacuum.
2. lossy media, such as ion-carrying water, clay, and all rocks and soils containing a significant proportion of interstitial water or clay. In these media, electromagnetic waves are strongly attenuated and it is not possible to neglect the influence of the electrical conductivity. Some of these media also exhibit a strong dispersion, which affects strongly conductivity values, and in a smaller extent permittivity (except for clay).
3. intermediate media, which are weakly dispersive and where the attenuating effect of conductivity is not dramatic but not totally negligible (~ 0.1). In this category, we find in particular samples of limestone, which is the material that I investigate in Section 3.2.

Note that volcanic rocks can present a non-zero magnetic susceptibility, as well as sands and sandstones depending on their detrital origin. I do not indicate these magnetic properties in Table 1.1 because I do not consider magnetic materials in the following.

Medium	ε_{∞_r}	χ_r	n	σ_{DC} (mS/m)	$\varepsilon'_r(\omega)$		$\sigma(\omega)$ (mS/m)		$\sigma_{DC}/(\varepsilon'\omega)$
					100 MHz	200 MHz	100 MHz	200 MHz	(100 MHz)
Air/vacuum	1	0	1	0.	1		0.		0.
Snow (dry) [1]					$\sim$ 1.2 to 1.5		0.		0.
Ice [1̃),(3b̃),(4)]					3.2		$\sim 10^{-3}$ to 10^{-2}		$\sim 10^{-5}$ to 5.10^{-4}
Distilled water [1̃),(4)]					81		$\sim 10^{-3}$ to 10^{-2}		$\sim 10^{-6}$ to 10^{-5}
Fresh water [4]					81		$\sim 10^{-2}$ to 10^{-1}		$\sim 10^{-5}$ to 10^{-4}
Saline water [1̃),(4)]					81		$\sim 10^{2}$ to 10^{3}		$\sim 10^{-1}$ to 2
Peat [1]					$\sim$ 60 to 80		$\sim$ 1 to 10		$\sim 10^{-3}$ to 3.10^{-2}
Sedimentary rocks [1]					$\sim$ 4 to 10		$\sim 10^{-2}$ to 10^{2}		$\sim 10^{-4}$ to 2
Limestone [3a),(4)]					$\sim$ 4 to 9		$\sim 10^{-1}$ to 3.3		$\sim 10^{-4}$ to 10^{-1}
Limestone (dry) [3b]	8.14	0.94	0.82		9.08	8.97	1.5	2.7	$\sim$ 0.03
Limestone [3b] (fresh water saturated)	17.3	1.1	0.33		18.4	18.0	10.7	13.5	$\sim$ 0.10
Gneiss - Schist [1]					$\sim$ 5 to 15		$\sim 10^{-3}$ to 10		$\sim 10^{-5}$ to 0.4
Schist [3a]					$\sim$ 5 to 15		$\sim$ 30 to 3.10^{2}		$\sim 10^{-1}$ to 10
Schist [2]	10.2	13.6	0.66	6.4	23.8	21.0	51	77	$\sim$ 0.2
Clay-rich rocks [1]					$\sim$ 5 to 40		$\sim$ 1 to 10^{3}		$\sim 10^{-2}$ to 10
Clay [3a]					$\sim$ 8 to 12		$\sim$ 10 to 10^{2}		$\sim 10^{-1}$ to 1
Clay [4]					$\sim$ 3 to 60		$\sim 10^{-1}$ to 10^{3}		$\sim 10^{-4}$ to 10
Wet clay [3b]	55	30	0.25		85.0	72.8	403.	479.	$\sim$ 0.7
Shale [3b]	4.3	2.6	0.6		6.9	6.3	10.5	15.9	$\sim$ 0.25
Siltstone [3b]	3.1	4.2	0.68		7.3	6.5	12.8	20.6	$\sim$ 0.3
Sand [1]					$\sim$ 2 to 10		$\sim 10^{-2}$ to 10		$\sim 10^{-4}$ to 0.5
Sand [3a]					$\sim$ 5 to 40		$\sim 10^{-1}$ to 3.3		$\sim 10^{-4}$ to 10^{-1}
Dry sand [4]					$\sim$ 3 to 6		$\sim 10^{-3}$ to 1		$\sim 10^{-5}$ to 10^{-2}
Dry sand [3b]	2.5	0	1		2.5		0.		0.
Sand [4] (water saturated)					$\sim$ 20 to 30		$\sim 10^{-1}$ to 10^{2}		10^{-3} to 0.5
Wet sand [3b]	29	4	0.5		33	31.8	22.2	31.5	$\sim$ 0.1
Quartzite [1]					$\sim$ 3 to 6		$\sim 10^{-3}$ to 10		$\sim 10^{-5}$ to 0.5
Sandstone [3a]					$\sim$ 4 to 5		$\sim$ 1 to 30		$\sim 10^{-2}$ to 1
Sandstone [4]					$\sim$ 4 to 5		$\sim 3.10^{-1}$ to 1		$\sim 10^{-2}$
Granite [3a]					$\sim$ 4 to 6		$\sim 10^{-2}$ to 1		$\sim 10^{-4}$ to 10^{-2}
Granite [2]	5.0	1.10	0.94	0.19	6.10	6.05	0.8	1.3	$\sim 10^{-2}$
Granite (dry) [3b]	4.7	0.7	0.93		5.40	5.37	0.4	0.8	$\sim$ 0.01
Granite [3b] (fresh water saturated)	5.6	0.5	0.51		6.10	6.96	2.7	3.8	$\sim$ 0.7
Volcanic rocks [1]					$\sim$ 4 to 17		$\sim 10^{-3}$ to 1		$\sim 10^{-5}$ to 10^{-2}
Andesite [3b]	3.6	2.1	0.62		5.7	5.2	7.9	12.	$\sim$ 0.2
Gabbro [3b]	5.8	3.4	0.67		9.2	8.5	11.	17.	$\sim$ 0.2

Table 1.1: *Electrical properties of some materials. Selected compilation after* [1] *Saintenoy (1998, tab. 3.1, p. 49),* [2] *Girard (2002),* [3a] *Jeannin (2005, tab. 1.1, p. 41),* [3b] *Jeannin (2005, tab. 2.1, p. 76),* [4] *Loeffler (2005, tab. 1, p. 29). The symbol* [$\tilde{*}$] *means that the values I indicate are consistent but not exactly identical to the ones given by the corresponding reference.*

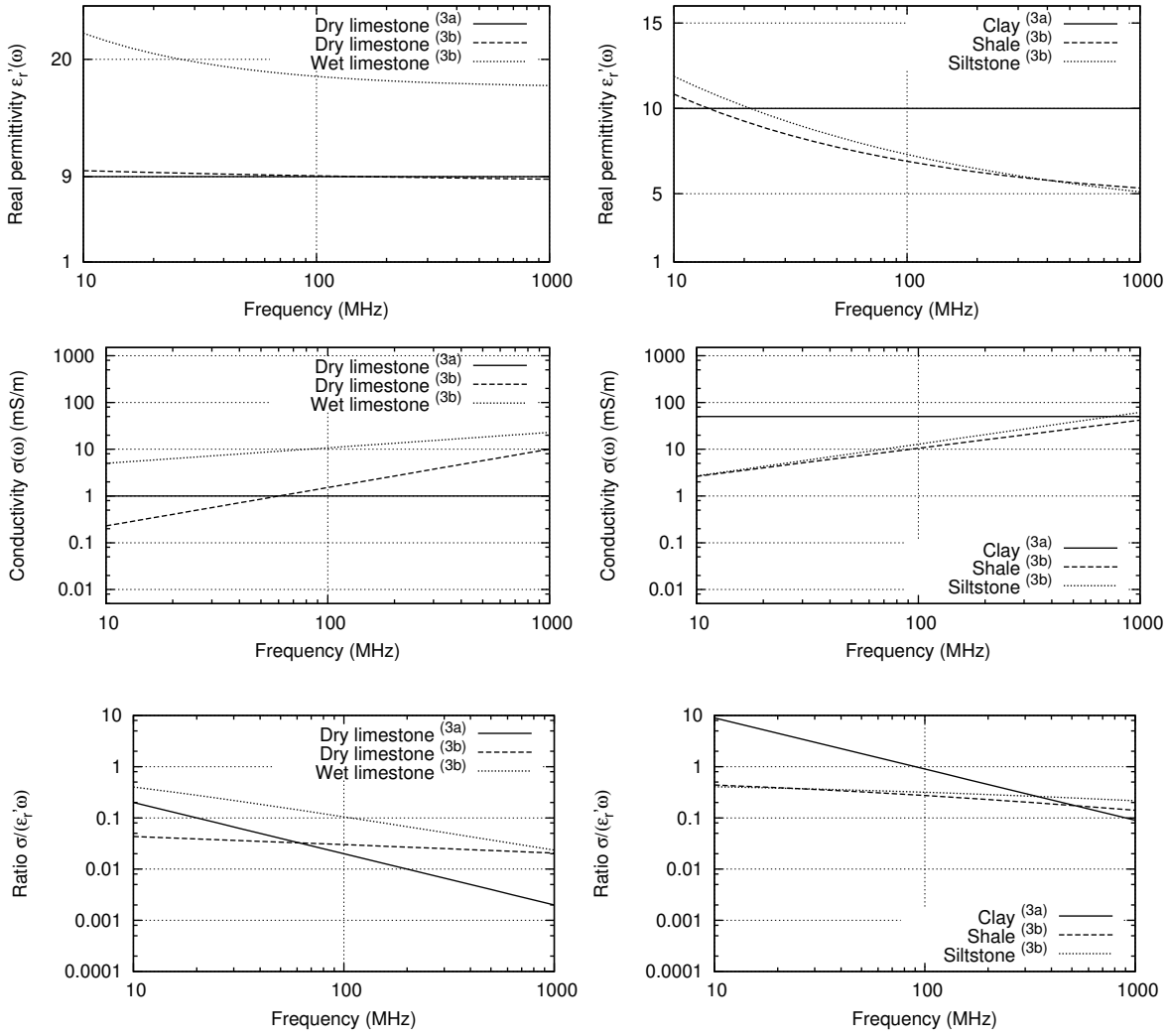
Redundancy in the table is deliberate and aims at showing the variability (or consistency) of these estimations for a given material, which depend not only on the general type of the material, but really on the specific sampling under consideration, and in particular on its water content. For instance, it is not very meaningful to speak about the permittivity of a sand without specifying if it is dry, slightly humid, or water-saturated. Conversely, parameter estimations for dry samples present less variability. Actually, permittivity and conductivity of rocks and soils depend in a large extent on their water content, and this for two reasons. First, water has a remarkably large value of permittivity compared to other natural materials, so that interstitial water in rocks, or melt water in snow and ice, dramatically increase the bulk permittivity of the sample. Secondly, the electrical conductivity of water greatly depend on its chemical composition, more precisely on its content in ions. In rocks and soils, groundwater is not pure and carries ions that cause electrolytic currents, thus increasing the conductivity of the materials. The great sensitivity of electrical properties to the presence of water is also the reason why GPR is so intensively used for hydrological purposes, as mentioned in the general introduction.

Indeed, even if we know that the investigated rock is dry, it is actually quite difficult to infer its nature or composition based on permittivity and conductivity measurements. For instance, a permittivity value of 5 can be attributed to almost all listed rocks, from limestone to volcanic rocks through granite. Similarly, a conductivity value of 10^{-1} mS/m does not bring much information if the nature of the sample is not known at all. Conversely, if we know the nature of the rock (as it is often the case when going on the field), it is possible to infer its water content from permittivity or conductivity measurements. The derivation of water content from electrical measurements (either GPR, TDR, or ERT) is generally based on empirical petrophysical relations such as those proposed by Archie (1942) and Topp et al. (1980). Obviously, permittivity and conductivity also depend on other properties of the investigated material, in particular on its porosity and clay content, such that the application of such empirical laws should be site- and scale-dependent (Chan and Knight, 1999; Moysey and Knight, 2004). Some authors proposed more sophisticated petrophysical models relating electrical properties to water content (e.g. Revil, 2013) or porosity (e.g. Rust et al., 1999). Some others avoid the use of such empirical relations by integrating soil models directly in the inversion process (e.g. Tran et al., 2012).

Another way to look at these properties is to represent the quantities $\varepsilon'_r(\omega)$, $\sigma(\omega)$ and $\tan\delta(\omega) = \sigma(\omega)/(\varepsilon'(\omega)\omega)$ with respect to frequency, for a given material. The variations of $\varepsilon'_r(\omega)$ and $\sigma(\omega)$ vs. frequency will tell us if dispersion is strong and should be considered to describe the material. The variations of $\tan\delta(\omega)$ will indicate if we can neglect the electrical conductivity, and if the losses can be adequately described using simple parameters such as a constant permittivity ε'_r and a constant static conductivity σ_{DC}. To the latter point, I will represent on the same plot the ratio $\sigma(\omega)/(\varepsilon'(\omega)\omega)$ derived from Jonscher parameters and the ratio $\sigma_{DC}/(\varepsilon'\omega)$ given by the rough estimates of constant permittivity and conductivity.

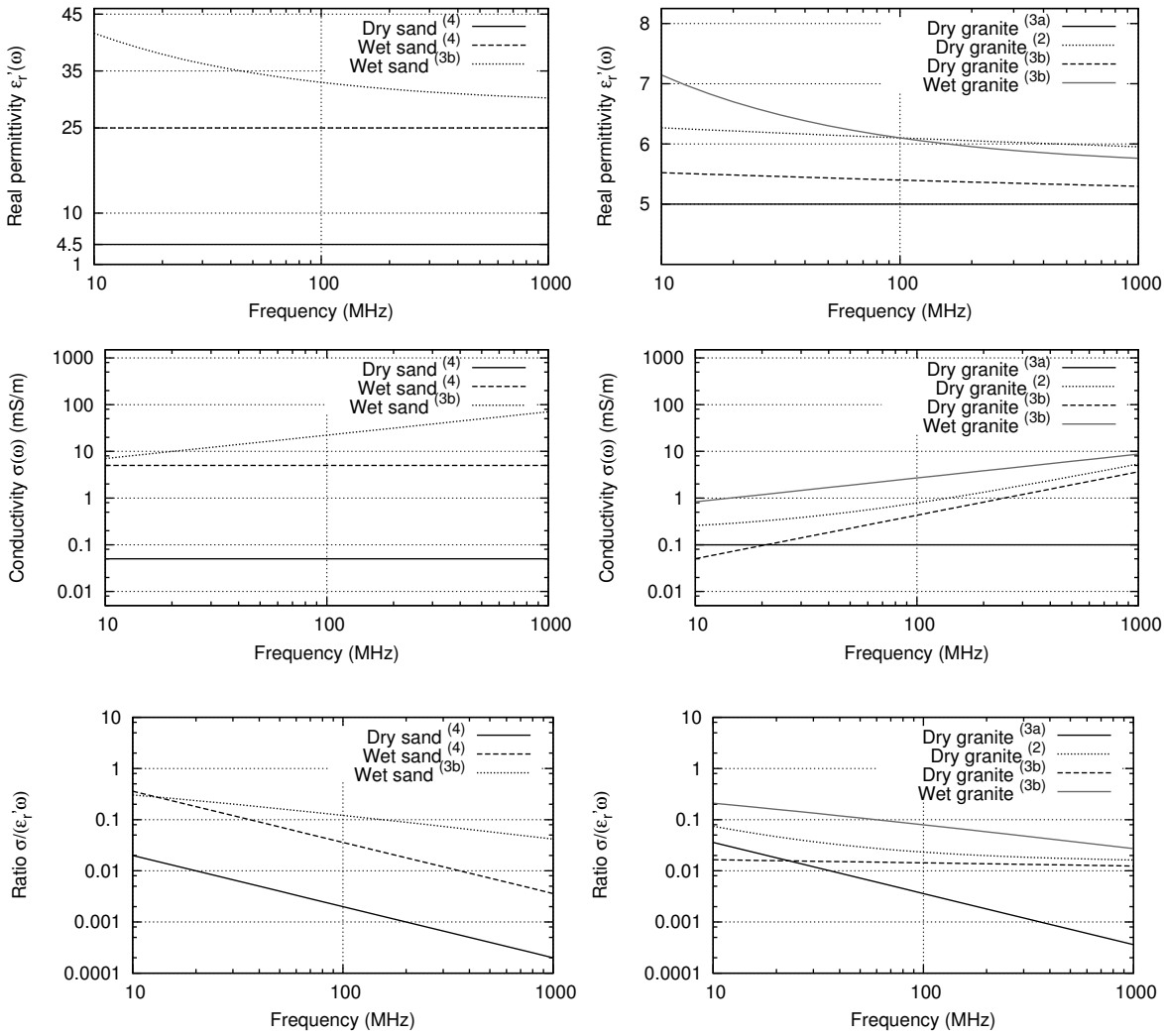
Figs 1.3 and 1.4 show these quantities for limestones, clay-rich rocks, sands and granite. Again, we observe that permittivity varies weakly with frequency in most of the considered samples (even in wet limestone and granite). Significant variations in permittivity can only be observed in water-saturated and clay-rich rocks. Conductivity, however, seems to be much more sensitive to dispersion, since even dry limestone, sand and granite exhibit a conductivity which varies with frequency over one to two decades. This observation seriously compromise the assumption of describing the medium with a constant static conductivity. This point is all

the more important that we can not neglect the effect of the electrical conductivity as soon as the samples are wet, as shown by the values of the loss tangent ($\tan\delta \geq 0.1$). Moreover, we can observe that the slopes of the loss tangents computed with a constant conductivity (samples $^{(3a)}$ and $^{(4)}$) are systematically steeper than the ones computed with frequency-dependent parameters derived from the Jonscher model. It means that assuming a constant conductivity tends to under-estimate the loss of energy with frequency. In some materials, e.g. dry limestone$^{(3b)}$ or dry granite$^{(3b)}$, the electrical losses do not necessarily decrease with frequency but can be constant.



(a) Limestones. **(b)** Clay, shale, siltstone.

Figure 1.3: *Real permittivity (top row), conductivity (middle row), and loss tangent (bottom) of limestone and shale samples. For the limestone $^{(3a)}$, I consider the averaged values $\varepsilon_r = 9$ and $\sigma = 1$ mS/m. For the clay $^{(3a)}$, I consider $\varepsilon_r = 10$ and $\sigma = 50$ mS/m.*



(a) Sand. **(b)** Granite.

Figure 1.4: *Real permittivity (bottom row), conductivity (middle row), and loss tangent (bottom) of sand and granite samples. For the sand* $^{(4)}$*, I consider the averaged values* $\varepsilon_r = 4.5$ *and* $\sigma = 0.05$ *mS/m. For the granite* $^{(3a)}$*, I consider* $\varepsilon_r = 5$ *and* $\sigma = 0.1$ *mS/m.*

As a conclusion concerning the electrical properties of natural media, we can state that

1. The effect of electrical conductivity can generally not be neglected at GPR frequencies, since natural conditions generally implies the presence of water that cause electrolytic currents. Electrical conductivity may be neglected only in dry samples that do not contain clay particles.

2. Dispersion can be observed in most of natural materials. We can notice, however, that it has little effect on permittivity values in dry samples. On the other hand, it induces large variations of conductivity with respect to frequency, which make the assumption of a constant conductivity not satisfactory.

This is the reality of natural media, but I shall now nuance the picture. As we shall see in

Chapter 2, it is actually very difficult to obtain robust reconstructions of dispersive parameters because GPR data are mostly sensitive to permittivity, and much less to conductivity and to Jonscher parameters. To my knowledge, the only studies to tackle the estimation of Jonscher parameters have been performed by Bourdi et al. (2008) and Ihamouten et al. (2011, 2012) for the bulk characterisation of concrete mixtures, and by Deparis and Garambois (2009) for the characterisation of a fracture filled by dispersive material (clay) inside a limestone massif considered as non-dispersive. In multilayered media, and a fortiori in arbitrary heterogeneous media, people generally make the assumption of a constant electrical conductivity.

Concerning 1D applications, it should be noted that van der Kruk et al. (2006, 2009, 2010) are able to reproduce strongly dispersed data without any dispersive parameters. Until now, I only spoke about *intrinsic* dispersion and attenuation, which are caused by the physical properties of the medium (Annan, 1996). But *geometrical* dispersion and attenuation can also occur, due for instance to the scattering of the electromagnetic waves by small heterogeneities. As the GPR pulse is broadband (typically from 50 to 300 MHz for a 100 MHz antenna), heterogeneities of the size of the smallest propagated wavelength will have very few effects on low frequency components and will cause a frequency-dependent scattering and attenuation of the GPR pulse. In the dataset considered by van der Kruk et al. (2006, 2009, 2010), dispersion is due to the presence of thin layers in the subsurface that act as wave guides. Depending on their frequency content, waves do not interact on the same way with the thin layers. In particular, the reflection coefficients that control the energy propagating in the wave guide are frequency-dependent, giving rise to dispersive reflected events recorded at the ground surface.

As a conclusion, I will consider intrinsic dispersion in the implementation of the forward problem, introducing the Jonscher model to define the effective permittivity. This feature enables to study in more details the effect of intrinsically dispersive parameters on the propagation of electromagnetic waves. But in Chapter 2, dedicated to the inverse problem, I will reconstruct non-dispersive parameters, assuming that permittivity and conductivity are real and frequency-independent quantities. Even with this simple parameterisation, it is important to understand the effect of real permittivities and conductivities on the behaviour of the electromagnetic waves, and I will now detail this point.

Considering permittivity and conductivity as real quantities

Here I intend to see the effect of constant, real-valued permittivity and conductivity on the electromagnetic waves propagating in a homogeneous medium. In this case, the solution of Helmholtz equation (1.24) with an impulsive source at $\mathbf{r} = \mathbf{0}$ is a spherical wave of the form

$$\mathbf{u}(\mathbf{r}, \omega) \propto \frac{e^{\imath k r}}{4\pi r}, \tag{1.39}$$

where $r = |\mathbf{r}|$ is the distance to the source and k is the complex wavenumber defined from Helmholtz equation as

$$k = \omega\sqrt{\mu\varepsilon_e} = \beta + i\alpha, \tag{1.40}$$

where

$$\alpha \;=\; \omega\sqrt{\frac{\mu\varepsilon}{2}\left(\sqrt{1+\left(\frac{\sigma}{\varepsilon\omega}\right)^2}-1\right)} \quad \text{is the attenuation coefficient,} \tag{1.41}$$

$$\text{and} \quad \beta = \omega\sqrt{\frac{\mu\varepsilon}{2}\left(\sqrt{1+\left(\frac{\sigma}{\varepsilon\omega}\right)^2}+1\right)} \quad \text{is the propagation wavenumber.} \tag{1.42}$$

Hence, expression (1.39) is equivalent to

$$\mathbf{u}(\mathbf{r},\omega) \propto \frac{e^{\imath\beta r}\, e^{-\alpha r}}{4\pi r}, \tag{1.43}$$

which is composed of an oscillating propagative part $e^{\imath\beta r}$, of an attenuating decay $e^{-\alpha r}$, and of a geometrical spreading $1/r$.

First we can notice that, in the general case, the permittivity and conductivity parameters both play a role in the propagative coefficient β and in the attenuation coefficient α. In particular, it means that the conductivity, even if real and frequency-independent, induces a dispersion of the phase velocity $v(\omega)$ which is expressed as

$$v(\omega) = \frac{\omega}{\beta} = \frac{1}{\sqrt{\mu\varepsilon}}\left[\frac{1}{2}\left(\sqrt{1+\left(\frac{\sigma}{\varepsilon\omega}\right)^2}+1\right)\right]^{-1/2}. \tag{1.44}$$

This velocity coincides with the classical electromagnetic velocity $v_{EM} = 1/\sqrt{\mu\varepsilon}$ only in non-conductive media, or at very high frequencies.

The importance of dispersion depends on the diffusion-over-propagation ratio, or loss-tangent $\tan\delta = \sigma/(\varepsilon\omega)$. In very lossy media where $\tan\delta \gg 1$, the wavenumber β and the attenuation coefficient tends to a common value

$$\alpha \simeq \beta \underset{\tan\delta\gg 1}{\sim} \sqrt{\frac{\mu\sigma\omega}{2}}, \tag{1.45}$$

meaning that the skin depth $\delta = 1/\alpha$ and the wavelength $\lambda = 2\pi/\beta$ increase at the same rate as frequency decreases (behaving like $1/\sqrt{\omega}$, which is typical of diffusive phenomena).

In the frequency range of GPR investigations, we rather encounter low-loss media ($\tan\delta \ll 1$). We then recover the classical electromagnetic velocity, which depends on permittivity:

$$v \underset{\tan\delta\ll 1}{\sim} v_{EM} = \frac{1}{\sqrt{\mu\varepsilon}}, \tag{1.46}$$

whereas the attenuation coefficient tends towards a frequency-independent value, which depends both on permittivity and conductivity:

$$\alpha \underset{\tan\delta\ll 1}{\sim} \frac{\sigma}{2}\sqrt{\frac{\mu}{\varepsilon}} = \frac{\omega}{2vQ}, \tag{1.47}$$

where $Q = 1/\tan\delta$ is the corresponding — frequency-dependent — quality factor.

From these expressions, we can conclude that, at first order, permittivity will mainly control the kinematics of the propagation of electromagnetic waves, while attenuation depends both on permittivity and conductivity through the diffusion-over-propagation ratio. At lower frequencies, conductivity introduces a frequency dependency, both for attenuation and velocity (i.e., dispersion).

Once we have understood these general guiding rules, we can go further and consider the simulation of electromagnetic wave propagation.

1.1.3 Wave propagation in two dimensions

In all the following, I will consider that everything happens in the observation plane $(x0z)$, i.e. I assume that neither the medium properties nor the fields do vary in the y-direction ($\partial_y = 0$). This assumption has numerical motivations since it is much more expensive to perform a numerical simulation of wave propagation in a heterogeneous 3D media than in 2D. The computational over-cost of 3D compared to 2D inversions is all the more important that inversion requires many simulations.

On the other hand, this assumption has important implications: assuming a laterally invariant medium, we will be unable to properly simulate the GPR signals resulting from out-of-plane reflections. Moreover, considering fields invariance in y-direction amounts to assume that the calculated field is generated by an infinite line-source elongated in the y-direction, instead of a finite-size 3D antenna. Restricting the wave propagation to the observation plane has a strong consequence in terms of energy decay, preventing to predict the true amplitudes of the recorded 3D field. Therefore, we will need to apply approximative 3D-to-2D conversions to the simulated data before comparing them with experimental ones (see Chapter 3).

1.1.3.1 TE-TM modes and analogy with the acoustic system

An interesting property of Maxwell's system in 2D is that it reduces to two independent systems, corresponding to two propagation modes called Transverse Electric (TE) and Transverse Magnetic (TM). Developing the field vectors on their three spatial components, we have, in the (xOz)-plane,

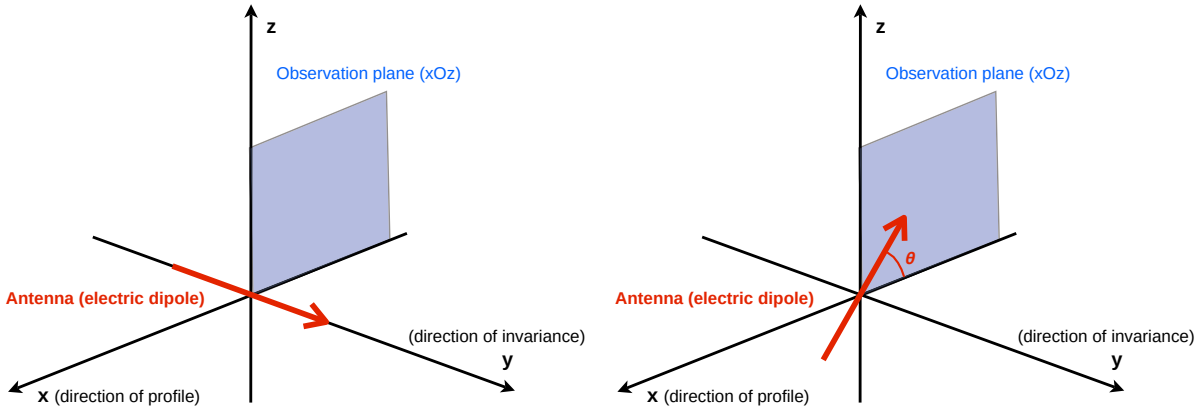
$$\text{TE} \quad \begin{cases} -\imath\omega\mu H_x &= \dfrac{\partial E_y}{\partial z}, \\ \imath\omega\mu H_z &= \dfrac{\partial E_y}{\partial x}, \\ -\imath\omega\varepsilon_e E_y &= \dfrac{\partial H_x}{\partial z} - \dfrac{\partial H_z}{\partial x} - J_{s_y}, \end{cases} \tag{1.48}$$

and

$$\text{TM} \quad \begin{cases} \imath\omega\varepsilon_e E_x &= \dfrac{\partial H_y}{\partial z} + J_{s_x}, \\ -\imath\omega\varepsilon_e E_z &= \dfrac{\partial H_y}{\partial x} - J_{s_z}, \\ -\imath\omega\mu H_y &= \dfrac{\partial E_z}{\partial x} - \dfrac{\partial E_x}{\partial z}, \end{cases} \tag{1.49}$$

where all field components and parameter variables are functions of space coordinates x and z, and of frequency ω.

As shown in Fig. 1.5, these two modes also correspond to two GPR measurement configurations, since the TE mode is excited and recorded by antennas oriented along the y-direction (Fig. 1.5a), whereas the TM mode is excited (/recorded) if the antennas are comprised in the (xOz)-plane (Fig. 1.5b).



(a) TE configuration. **(b)** TM configuration.

Figure 1.5: *Geometry of GPR acquisitions. In theory, the TM mode can be excited and recorded by antennas having an arbitrary orientation θ in the $(x0z)$-plane. In practice, two configurations are usually employed: either vertical antennas ($\theta = 90°$) in borehole acquisitions, or horizontal antennas ($\theta = 0°$) in surface measurements (on-ground GPR).*

Interestingly, the equation systems (1.48) and (1.49) turn out to be mathematically analogous to the acoustic velocity-stress system:

$$\text{Acoustics} \quad \begin{cases} -\imath\omega\rho v_z &= \dfrac{\partial P}{\partial z} + \rho f_z, \\ -\imath\omega\rho v_x &= \dfrac{\partial P}{\partial x} + \rho f_x, \\ -\imath\omega\dfrac{P}{K} &= \dfrac{\partial v_x}{\partial x} + \dfrac{\partial v_z}{\partial z} - \imath\omega\dfrac{P_s}{K}, \end{cases} \tag{1.50}$$

where P is the pressure (in Pa), v_x and v_z are the components of the particle velocity (in m/s), ρ is the mass density (in kg/m^3), K is the bulk modulus (in Pa), f_x and f_z are horizontal and vertical force sources (expressed as accelerations, in m/s^2), and P_s is an explosive source of pressure (in Pa). Again, all these variables depend on space coordinates x and z, and on frequency ω, such that the acoustic-electromagnetic analogy is fully valid for heterogeneous, dispersive media. The correspondence between acoustic and electromagnetic variables is given in Table 1.2.

This mathematical analogy is discussed in the literature by (e.g.) Carcione and Cavallini (1995) and Carcione and Robinson (2002). It is an important property since it allowed me to solve the electromagnetic propagation problem with a numerical modelling tool that was originally developed for the simulation of acoustic waves. I will present this numerical scheme in the following (see Section 1.2.2) but first I would like to derive the analytical solutions to the 2D propagation problem in homogeneous media, which will be used later for the validation of the numerical solutions. Before going further, I insist on the fact that the acoustic-electromagnetic analogy between equation systems (1.48) and (1.49) is only a mathematical trick. It should not be interpreted in terms of physics. In particular, visco-acoustic and electromagnetic fields behave quite differently with regards to attenuation.

TE			**TM**		
E_y	$\leftrightarrow$	P	H_y	$\leftrightarrow$	P
H_z	$\leftrightarrow$	$-v_x$	E_z	$\leftrightarrow$	v_x
H_x	$\leftrightarrow$	v_z	E_x	$\leftrightarrow$	$-v_z$
ε_e	$\leftrightarrow$	$1/K$	ε_e	$\leftrightarrow$	ρ
μ	$\leftrightarrow$	ρ	μ	$\leftrightarrow$	$1/K$
J_{s_y}	$\leftrightarrow$	$\imath\omega P_s/K$	J_{s_x}	$\leftrightarrow$	ρf_z
			J_{s_z}	$\leftrightarrow$	$-\rho f_x$

Table 1.2: *Correspondence between acoustic and electromagnetic variables.*

1.1.3.2 Wave equations and analytical solutions in homogeneous media

Assuming a homogeneous medium, it is useful to derive now the wave equations corresponding to TE and TM modes for further validation of numerical implementations. Derivating the first and second equations of systems (1.48) and (1.49), and injecting the derivatives in their third equation yields the scalar wave equations followed by the electric component E_y in TE mode and by the magnetic component H_y in TM mode:

$$\nabla^2 E_y(x,z,\omega) + \omega^2 \varepsilon_e(\omega)\mu(\omega) E_y(x,z,\omega) = -\imath\omega\mu(\omega) J_{s_y}(x,z,\omega), \tag{1.51}$$

$$\nabla^2 H_y(x,z,\omega) + \omega^2 \varepsilon_e(\omega)\mu(\omega) H_y(x,z,\omega) = \frac{\partial J_{s_z}(x,z,\omega)}{\partial x} - \frac{\partial J_{s_x}(x,z,\omega)}{\partial z}, \tag{1.52}$$

whereas the wave equation corresponding to the acoustic system (1.50) is

$$\nabla^2 P(x,z,\omega) + \omega^2 \frac{\rho(\omega)}{K(\omega)} P(x,z,\omega) = \rho(\omega)\left(\frac{\partial f_x(x,z,\omega)}{\partial x} + \frac{\partial f_z(x,z,\omega)}{\partial z}\right) - \omega^2 \frac{\rho(\omega)}{K(\omega)} P_s(x,z,\omega). \tag{1.53}$$

In the above equations, a particular care has been taken to formulate the source terms as physically homogeneous quantities (electrical currents, mechanical forces and acoustic pressure). Besides, we shall consider the fundamental 2D scalar wave equation, or Helmholtz equation

$$\nabla^2 u(x,z,\omega) + \mathsf{k}^2(\omega)\, u(x,z,\omega) = \delta(r), \tag{1.54}$$

where k is again the complex wavenumber and $r = \sqrt{x^2+z^2}$ is the distance to an impulsive point source located at the origin of the coordinates system and represented by the Dirac delta function, defined as

$$\delta(r) = \begin{cases} 1 & \text{if } r = 0, \\ 0 & \text{otherwise.} \end{cases} \tag{1.55}$$

The solution of equation (1.54) is the 2D Green function, or impulsive response, of the medium:

$$G_{2D}(x,z,\omega) = \frac{\imath}{4} H_o^{(1)}(\mathsf{k}r), \tag{1.56}$$

where $H_o^{(1)}$ is the Hankel function of the first kind at order zero (see e.g. Taflove and Hagness, 2005, §8.2.2, eq. (8.7), p. 332). In the far field, when $|\mathsf{k}r| \to \infty$, we shall consider the following far-field approximation

$$G_{2D}(x,z,\omega) \simeq \frac{\imath}{4}\sqrt{\frac{2}{\pi \mathsf{k} r}} e^{i\mathsf{k}r} e^{-\imath\pi/4}, \tag{1.57}$$

(see Taflove and Hagness, 2005, eq. 8.8). As shown in Fig. 1.6, this far-field approximation is actually valid even at short offsets, as soon as the propagation distance r reaches half a wavelength ($|kr| \geq \pi$). In the following of the manuscript, I sometimes use this far-field expression, in particular to derive formulas for 3D-to-2D conversions (see Chapter 3).

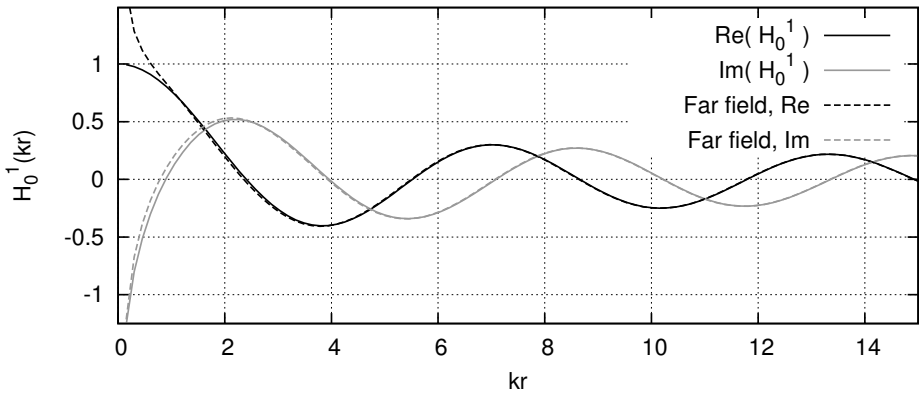


Figure 1.6: *Comparison of the Hankel function $H_o^{(1)}$ with its asymptotic far-field expression.*

TE mode The solutions of equations (1.51) and (1.52) can be derived from the Green function by linearity with respect to the source term. Assuming an impulsive current source $J_{s_y}\delta(r)$, we have for TE mode

$$\begin{aligned} E_y(x,z,\omega) &= -\imath\omega\mu J_{s_y} G_{2D}(x,z,\omega), \\ &= \frac{\omega\mu J_{s_y}}{4} H_o^{(1)}(kr). \end{aligned} \tag{1.58}$$

TM mode In this case, the solution of equation (1.52) is the superposition of the magnetic fields $H_y^{J_{s_z}}$ emitted by a vertical current source J_{s_z}, and $H_y^{J_{s_x}}$ emitted by a horizontal current source J_{s_x}:

$$\begin{aligned} H_y(x,z,\omega) &= H_y^{J_{s_z}}(x,z,\omega) + H_y^{J_{s_x}}(x,z,\omega), \\ &= J_{s_z}\frac{\partial G_{2D}(x,z,\omega)}{\partial x} - J_{s_x}\frac{\partial G_{2D}(x,z,\omega)}{\partial z}. \end{aligned} \tag{1.59}$$

Using the properties of the Hankel function, in particular the fact that $H_o'(x) = -H_1(x)$ (see e.g. Abramowitz and Stegun, 1972, §9.1.27, p. 361), we thus have

$$H_y(x,z,\omega) = \frac{\imath k}{4}\left(\frac{zJ_{s_x}}{r} - \frac{xJ_{s_z}}{r}\right) H_1^{(1)}(kr). \tag{1.60}$$

GPR antennas measure a voltage, i.e. a quantity which is proportional to the electric field. As shown on figure 1.7, a receiver antenna in TM mode with arbitrary orientation θ_r with respect to the x-axis actually records the projection of the total electric field on the antenna, i.e. the value

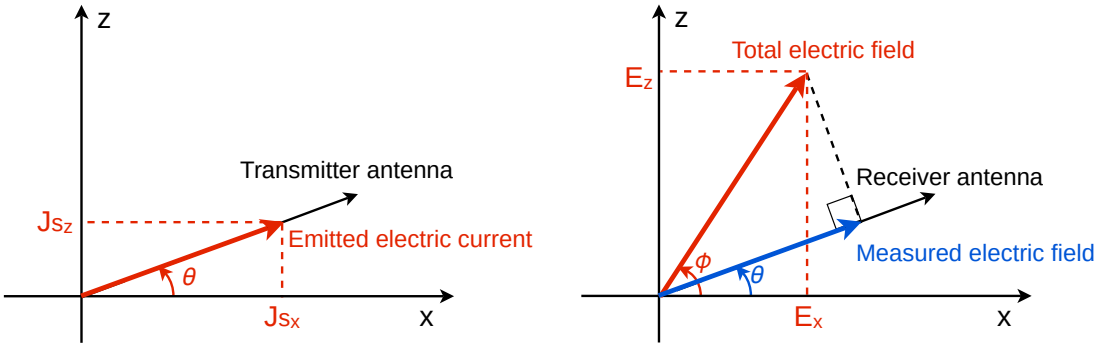
$$E_r(x,z,\omega,\theta_r) = E_x(x,z,\omega)\cos\theta_r + E_z(x,z,\omega)\sin\theta_r, \tag{1.61}$$

where the components of the total field E_x and E_z are deduced from the solution (1.60) using Faraday's law, i.e. the first and second equations in the system (1.49):

$$\begin{aligned} E_x(x,z,\omega) &= \frac{1}{i\omega\varepsilon_e(\omega)}\frac{\partial H_y(x,z,\omega)}{\partial z}, \\ &= \frac{\omega\mu}{4}\left[\frac{J_{s_x}}{kr}H_1^{(1)}(kr) - J_{s_x}\frac{z^2}{r^2}H_2^{(1)}(kr) + J_{s_z}\frac{xz}{r^2}H_2^{(1)}(kr)\right], \end{aligned} \tag{1.62}$$

and

$$\begin{aligned} E_z(x,z,\omega) &= -\frac{1}{i\omega\varepsilon_e(\omega)}\frac{\partial H_y(x,z,\omega)}{\partial x}, \\ &= \frac{\omega\mu}{4}\left[\frac{J_{s_z}}{kr}H_1^{(1)}(kr) + J_{s_x}\frac{xz}{r^2}H_2^{(1)}(kr) - J_{s_z}\frac{x^2}{r^2}H_2^{(1)}(kr)\right]. \end{aligned} \tag{1.63}$$



(a) Transmitter case. **(b)** Receiver case.

Figure 1.7: *Antenna orientation and emitted/recorded electric fields in TM mode.*

Expressions (1.58) and (1.61) to (1.63) are the reference solutions I will use for the validation of my numerical implementations. In the next section, I present the numerical scheme employed in my simulations, after a short overview of the numerical methods generally involved in computational electromagnetics.

1.2 Numerical modelling of electromagnetic waves propagation in 2D heterogeneous media

Natural media, and especially near surface environments, are particularly heterogeneous. They can present strong contrasts in electromagnetic parameters at interfaces between air and soil, in case of cavities, fractures or anthropic structures, or in presence of water. In this context, general solutions can not be found analytically and numerical modelling is an essential tool for quantitative imaging, such as full waveform inversion, but also in order to understand how electromagnetic waves behave in these heterogeneous, complex media.

Simulations for modelling and simulations for imaging have different objectives, and thus different requirements. As we will see in Chapter 2, FWI is an iterative procedure which requires the resolution of many forward problems. Hence, the resolution must be performed very efficiently in terms of computation time and memory requirements. On the other hand, understanding the behaviour of electromagnetic waves in natural environments or complex anthropic structures requires a consistent description of the medium properties. I already mentioned some physical parameterisations enabling to describe dispersion and attenuation. In addition, numerical methods proceed by discretising the medium. How this discretisation is performed is a crucial issue to ensure an accurate description of the medium geometry and macroscopic property.

1.2.1 Introduction: Choice of the numerical method

Integral methods (Chew et al., 2009) are classical in electromagnetics, especially in the domain of telecommunications where waves propagate on large distances in free-space. Numerical techniques derived from integral methods, such as the boundary element methods (Sauter and Schwab, 2011), also known as the method of moments (Harrington, 1993), are particularly efficient for problems with a low surface/volume ratio because they require to compute field values only on the boundary of the domain. However, they require the knowledge of the Green functions of the bounded medium to derive the solution within the domain volume, which is a strong limitation to their application in heterogeneous media. In the following of the manuscript, my numerical implementations involve essentially domain methods. With an exception in Section 3.1, where I will make use of an integral representation of the electric field to restrict the computation domain to the zone of interest.

Among domain methods, the finite-difference time-domain method (FDTD) is probably the most popular. It is a simple, efficient, and robust numerical method for solving electromagnetic problems in heterogeneous media, based on finite-order difference approximations of time and space derivatives. Since the first FDTD scheme for electromagnetic purposes, proposed by Yee (1966), many developments have been achieved (see e.g. Taflove and Hagness, 2005). For instance, Giannopoulos (2005) developed a FDTD-based freely-available program specifically for GPR applications. However, as mentioned in the previous section, the consideration of dispersive properties is not straightforward in the time-domain formulation, since it involves convolution products, requiring the history of the wavefields to be stored. In contrast, frequency-domain finite difference formulations (FDFD) take naturally the dispersion into account. While more memory demanding, FDFD techniques are more efficient for the resolution of problems involving a large number of sources. They are thus suitable for FWI when dense

acquisition arrays are considered. Finally, frequency-domain modelling is justified in the frame of frequency-domain FWI where only few frequency components are considered (Pratt and Worthington, 1990; Pratt, 1999). In my inversions, I make use of the FDFD scheme of Hustedt et al. (2004) which I adapted to the electromagnetic wave propagation problem. In the next section, I present the principles of this scheme, whose main features should be well understood for a correct utilisation in the inversion process.

An other class of wide-spread techniques are finite-volume methods (see e.g. Cioni, 1995; Remaki, 1999) and finite-element methods (see e.g. Cohen and Monk, 1998). Their main advantage consists in the consideration of more complex and flexible meshes. Among these methods, the discontinuous Galerkin method is a method of particular interest in computational electromagnetics since it allows for discontinuities in field values across interfaces between high contrasted media (see e.g. Canouet, 2003; Dolean et al., 2006; or El Bouajaji et al., 2011, for an application to GPR FWI). During my PhD, I also adapted the discontinuous Galerkin algorithm of Brossier et al. (2008) to the electromagnetic problem but I do not detail this part of my work in the manuscript because this implementation will not serve for the inversion due to its higher computational cost.

1.2.2 The frequency-domain finite-difference mixed-grid stencil

The FDFD algorithm of Hustedt et al. (2004) is based on the mixed-grid stencil of Jo et al. (1996). The main feature of this formulation consists in a weighted average of two staggered-grid stencils to decrease the errors due to numerical dispersion and anisotropy. Optimising the weights involved in the averaging, it is possible to decrease the number of required grid points per propagated wavelength (down to four or less) compared to other FDFD schemes, while reaching a quasi-arbitrary accuracy in a homogeneous medium at a given frequency. This property will be fully exploited in the application of the imaging algorithm to the Institut Fresnel data set, which requires a high level of accuracy. In this section, I illustrate qualitatively the principles of the method. For more technical details on the formulation, I refer to the original articles of Jo et al. (1996) and Hustedt et al. (2004).

An optimised FDFD scheme

Fig. 1.8 shows the two finite-difference grids involved in the scheme of Jo et al. (1996): the first one is a classical cartesian grid and the another one a 45°-rotated grid. On the classical staggered grid, pressure and velocity components are not estimated at the same place (hence the qualification of *staggered*): pressure is evaluated at nodes and velocities between nodes. In the classical staggered grid formulation, this enables to mitigate the sub-grid decoupling effect arising when the excitation is implemented as a single point source (see fig. 2 in Hustedt et al., 2004, for an illustration of this effect).

Jo et al. (1996) propose two averaging operations. On one hand, averaging the laplacian term in equation (1.53) leads to less errors due to numerical anisotropy. This is illustrated in Fig. 1.9 which shows the dispersion curves associated to the numerical scheme, i.e. the evolution of the numerical phase velocity as a function of the grid sampling, expressed as the number of grid points per propagated wavelength $G = \lambda/h$. In Fig. 1.9(b), where the laplacian term has been averaged, the dispersion curves vary less with the direction of wave propagation than for a classical staggered grid stencil (Fig. 1.9a). On the other hand, averaging the mass

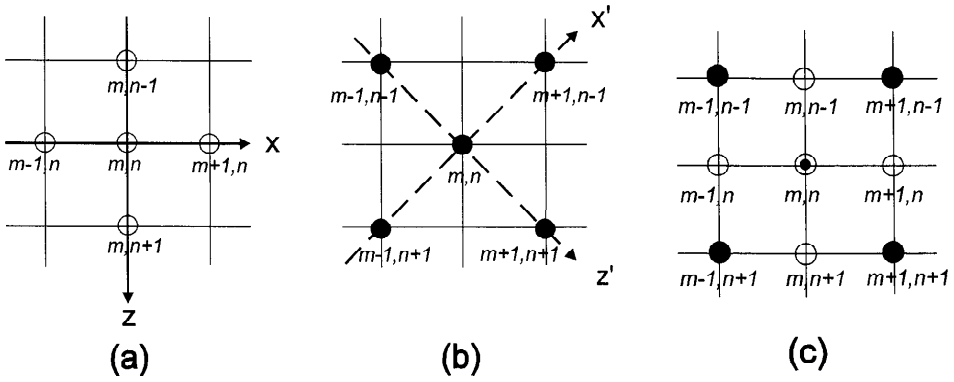
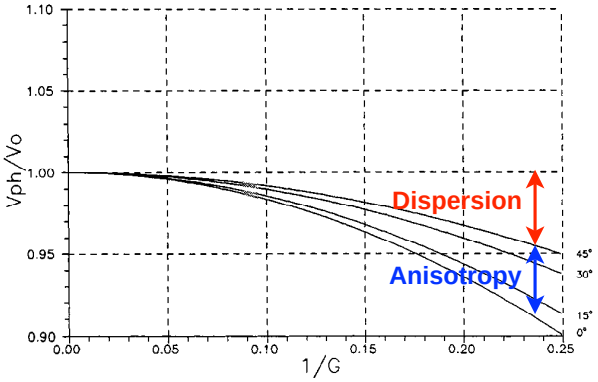
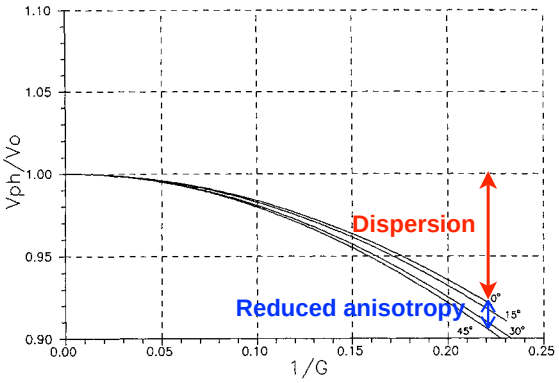


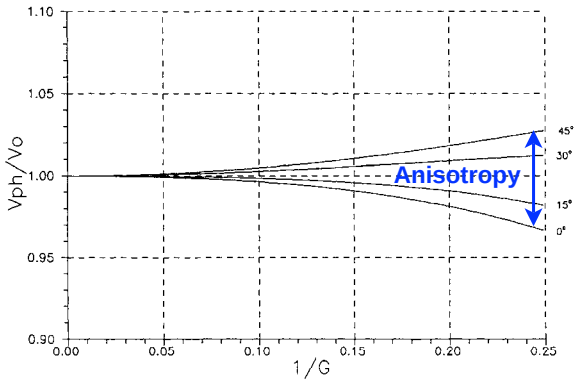
Figure 1.8: *Finite-difference stencils (from Jo et al., 1996). a) Classical five-point stencil. b) Rotated five-point stencil. c) Mixed-grid nine-point stencil.*



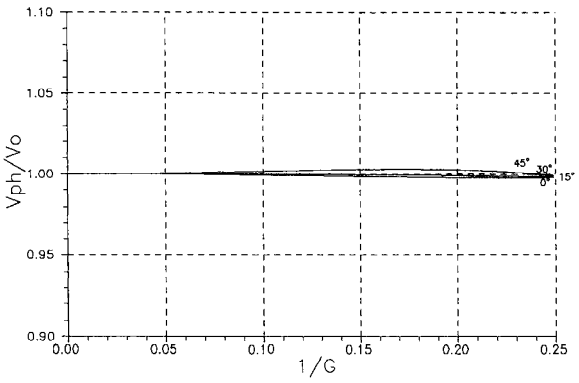
(a) Classical five-point staggered grid stencil.



(b) Mixed-grid nine-point stencil with average of the Laplacian term. Numerical anisotropy is reduced.



(c) Mixed-grid nine-point stencil with average of the mass acceleration term. Numerical dispersion is reduced.



(d) Mixed-grid nine-point stencil with a weighted average of the laplacian and of the mass acceleration terms. Weights are optimised for 4 grid points per wavelength.

Figure 1.9: *Dispersion curves for the numerical phase velocity corresponding to different finite-difference schemes (modified from Jo et al., 1996). Each curve corresponds to a given direction of wave propagation with respect to the cartesian grid (0°, 15°, 30° and 45°).*

acceleration term ($\omega^2 \rho K^{-1} P$ in equation 1.53) leads to less errors due to numerical dispersion (see Fig. 1.9c, where the errors are centred around 0). Performing a weighted average of both the laplacian and of the mass acceleration terms dramatically improves the accuracy of the numerical solutions for a given number of grid points per propagated wavelength, i.e. for a given velocity and a given frequency: see Figs 1.9(d), where the stencil coefficients have been optimised for $G = 4$ grid points per propagated wavelength. In their study, Jo et al. (1996) make use of a steepest descent algorithm for the optimisation of the phase velocity, whose dispersion relation can be derived from a classical plane-wave analysis (see their appendix A).

The mixed-grid stencil formulation ends up with a system which can be recast in matrix form as

$$\mathcal{A}(x, z, \omega)\, \mathbf{u}(x, z, \omega) = \mathbf{s}(x, z, \omega), \tag{1.64}$$

where $\mathcal{A}$ is called the impedance matrix, $\mathbf{u}$ is the wavefield to solve for, and $\mathbf{s}$ is the source term. In the following, I will often refer to this equation as it constitutes the *forward problem* in the inversion process. On the implementation point of view, the adaptation of the algorithm of Hustedt et al. (2004) to the electromagnetic problem only required to substitute the acoustic parameters by the electromagnetic ones, following the acoustic-electromagnetic analogy (see Table 1.2), before the construction of the impedance matrix $\mathcal{A}$.

I now validate this implementation, whose technical realisation and utilisation bring out other problems concerning the accuracy of the solutions. The validation is achieved by considering a homogeneous medium, for which an analytical solution has been derived in Section 1.1.3.2. As we are interested in characterising heterogeneous media, a validation in a heterogeneous medium might have been more relevant, but it is less straightforward since no closed-form solution exists for these problems. Even in the case of a simple interface, resolution involves semi-analytical solutions through a wavenumber decomposition whose discretisation possesses its own source of errors. In Section 3.1, the confrontation of the numerical solutions to physical data generated by a scattering experiment may serve as a demonstration of the validity of the method in heterogeneous media.

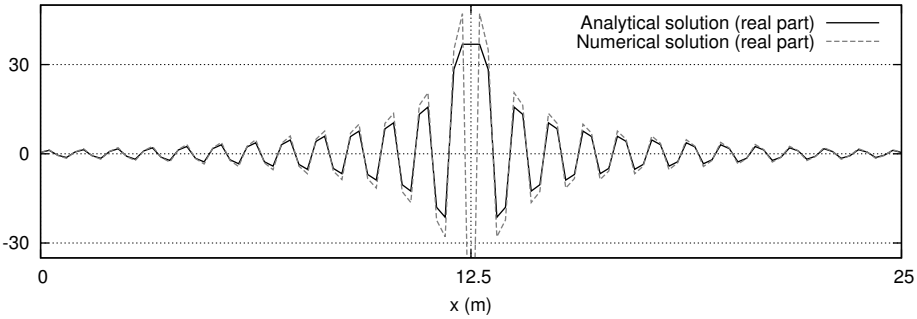
1.2.3 Validation in a homogeneous medium

If we consider a homogeneous medium, the FDFD scheme actually solves the following wave equation:

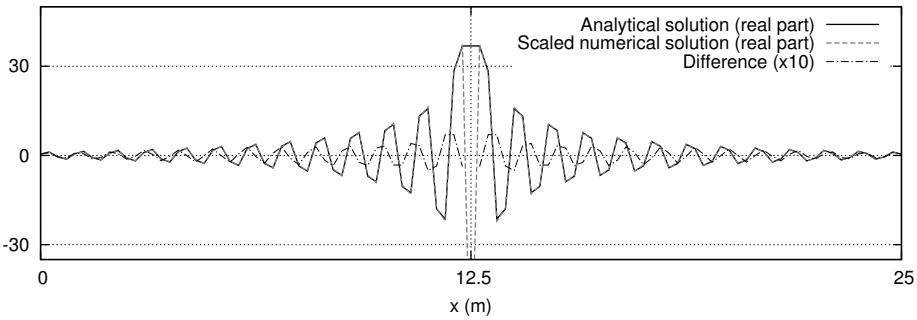
$$\nabla^2 \mathbf{u}(\mathbf{r}, \omega) + \omega^2 \frac{\rho(\omega)}{K(\omega)} \mathbf{u}(\mathbf{r}, \omega) = -\rho(\omega)\, \mathbf{s}(\mathbf{r}, \omega), \tag{1.65}$$

where I keep the acoustic notations that should be substituted by their electromagnetic equivalents in TE or TM mode (see Table 1.2). The expected analytical solutions in TE and TM modes can be derived from the reference solutions presented in the previous section (eqs 1.58 and 1.61 to 1.63) by linearity with respect to the source term.

Fig. 1.10(a) compares the analytical and the numerical solutions computed at a frequency of 100 MHz in a homogeneous medium presenting both attenuation and dispersion, described by the Jonscher's parameters corresponding to limestone$^{(3b)}$ in Table 1.1, with an additional DC conductivity. For this calculation, I use the stencil coefficients of Jo et al. (1996) and a grid step $h = 0.25$ m $\simeq \lambda/4$ (at 100 MHz, $\lambda \simeq 1.0$ m in this medium).



(a) Analytical solution vs. numerical solution (real parts).



(b) Analytical solution vs. scaled numerical solution (real parts).

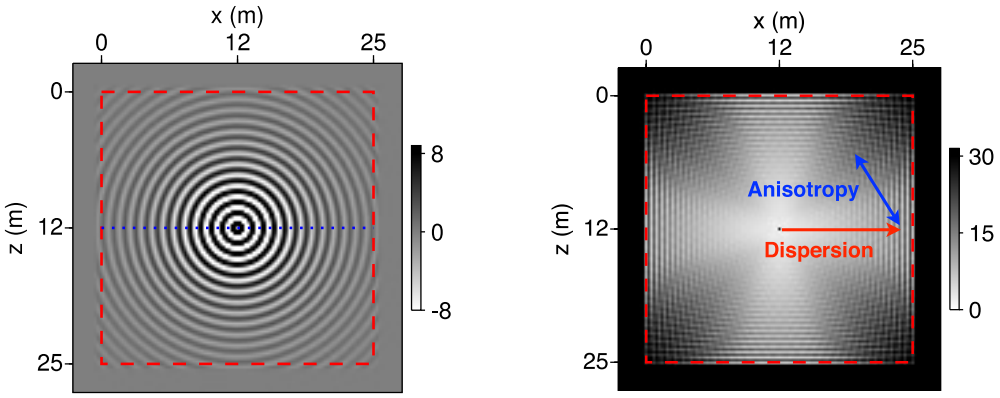
Figure 1.10: *Comparison of the analytical and numerical electric fields for the propagation of TE mode at 100 MHz in a limestone described by the set of Jonscher's parameters* $\varepsilon_{\infty_r} = 8.14$, $\chi_r = 0.94$, $n = 0.82$ *and* $\sigma_{DC} = 1$ *mS/m.*

Effective radiation of hard sources

Significant discrepancies can be observed between the analytical and the numerical solutions in Fig. 1.10(a). In fact, there is a constant ratio between the amplitudes of both solutions. This amplitude ratio is due to the effective radiation of the implemented hard source. Indeed, the numerical source is not as punctual as the theoretical Dirac function. Since the source of eq. (1.65) is implemented as a non-zero value located at one grid point, its effective radiation depends on the size of the grid step used in the FDFD calculations, compared to the propagated wavelength. The observed amplitude discrepancy thus depends on the grid step and on the simulated frequency. It is all the more important in the example of Fig. 1.10(a) that the optimisation of the stencil coefficients enables the use of a grid step which is relatively large compared to the propagated wavelength ($h \simeq \lambda/4$). Using a finer grid step decreases the discrepancy since the source can be considered as more punctual relatively to the wavelength (but doing so, we loose the benefits of the mixed-grid stencil). This effect has been studied in details by Waldschmidt and Taflove (2000) for the FDTD method (see also Taflove and Hagness, 2005, §5.3.2, p. 172). For a better agreement between the analytical and the numerical solutions, these authors recommend to scale the numerical solution by an *ad hoc* factor that fits the analytical amplitudes. It is the approach I adopt here and in the further application to the Institut Fresnel data set (Section 3.1). When performing modelling in the frame of inversion, however, I will simply consider that this scaling factor is absorbed in the estimation of the

source signature.

Fig. 1.10(b) shows the resulting (scaled) numerical values and the difference between numerical and analytical solutions. The agreement is now satisfying, with a relative error of less than 10% along the section. Similarly, Fig. 1.11 shows the scaled numerical wavefield in the entire domain and the relative difference between numerical and analytical solutions. It can be noticed that the relative error can reach up to 30% in the corner of the domain, which suggests a residual numerical anisotropy. This can be understood by looking more precisely on



(a) Numerical solution (scaled, real part).

(b) Relative difference with the analytical solution (%).

Figure 1.11: *Numerical wavefield (TE mode) in a limestone described by the set of Jonscher's parameters* $\varepsilon_{\infty_r} = 8.14$, $\chi_r = 0.94$, $n = 0.82$ *and* $\sigma_{DC} = 1$ *mS/m. The dotted blue line indicates the location where values were extracted for Fig. 1.10. The dashed red line delimits the physical domain, surrounded by PML (Bérenger, 1994).*

the dispersion curves of Fig. 1.9(d), obtained with the coefficients published by Jo et al. (1996): Fig. 1.12(a) presents a zoom on these dispersion curves. On this figure, we can observe that the set of stencil coefficients optimised by Jo et al. (1996) restrict the error on the numerical phase velocity to less than 0.3% using four grid points per wavelength or more ($1/G \geq 0.25$). Nevertheless, this error is not zero for $G \simeq 0.25$ and it is sufficient to generate errors when propagating over many wavelengths in a direction of 45° with respect to the grid.

Optimisation of the stencil coefficients

One possibility to improve the accuracy of the numerical solution is to better optimise the stencil coefficients. In Fig. 1.12(b), I present dispersion curves corresponding to stencil coefficients I have optimised for $G = 4$ using the same VFSA algorithm as Operto et al. (2009)[1]. Using a very low convergence threshold, the VFSA algorithm yields a very low error on phase velocity for the specific value $1/G = 0.25$. It can be noticed, however, that the error increases if smaller grid steps are used (for $0.17 > 1/G > 0.25$, i.e. $4 < G < 5.9$) up to larger errors than for the coefficients of Jo et al. (1996). If it enables to lower the error in a particular homogeneous medium, optimising the coefficients for a specific value of G therefore does not ensure accurate

[1]This algorithm has been kindly provided by Stéphane Operto (Géoazur, CNRS, Univ. Nice-Sophia Antipolis, France).

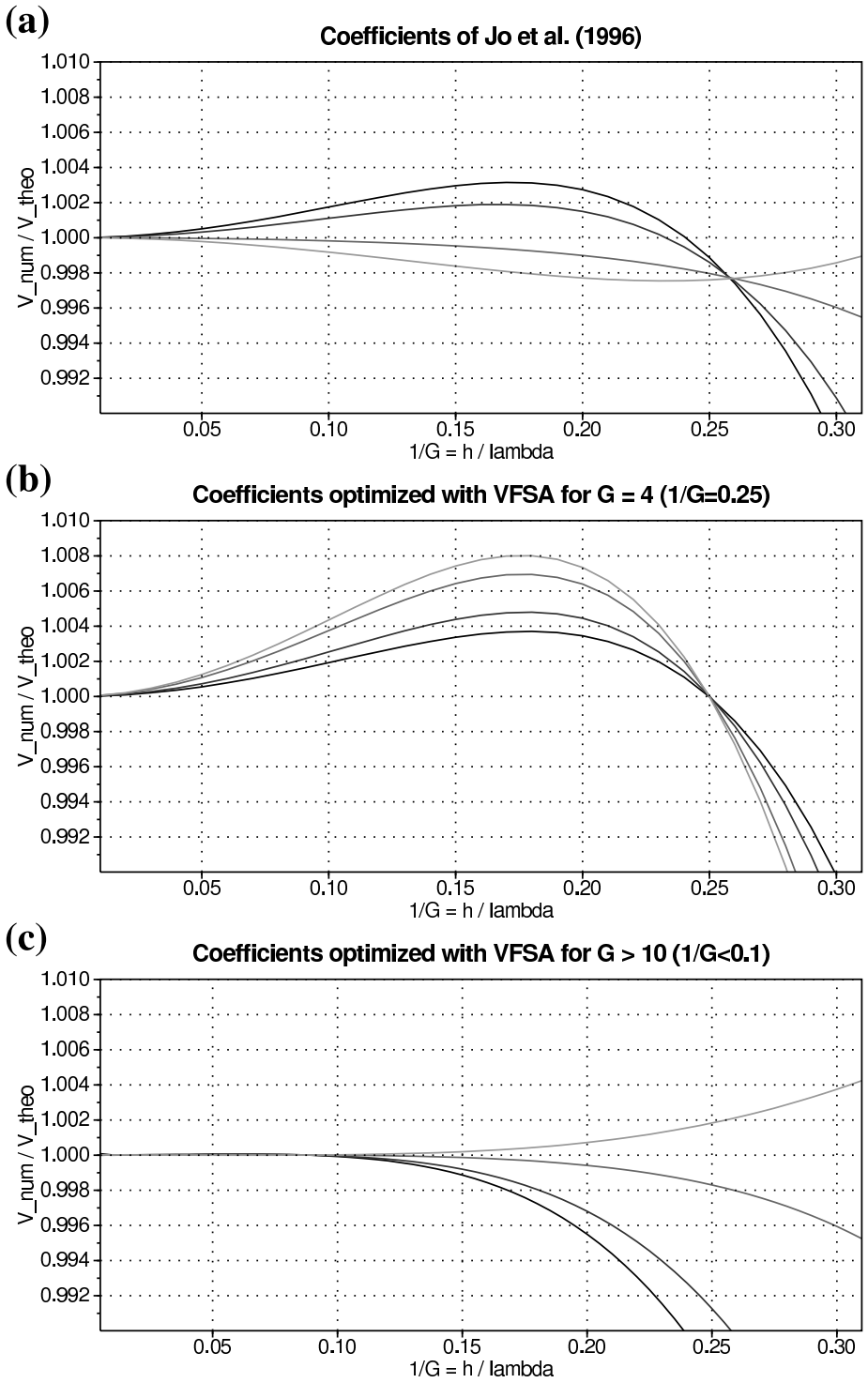


Figure 1.12: *Dispersion curves obtained using (a) the stencil coefficients of Jo et al. (1996), (b) coefficients optimised for $G = 4$ with a VFSA algorithm (Operto et al., 2009; Sen and Stoffa, 1995) and (c) coefficients optimised for $G \geq 10$ using the VFSA algorithm. Again, each curve corresponds to a given direction of wave propagation with respect to the cartesian grid (0°, 15°, 30° and 45° from dark to light grey).*

solutions in heterogeneous media, where the value G varies according to the local propagated wavelength.

A more satisfying strategy consists in optimising the stencil coefficients such that the error on the numerical phase velocity decreases for a range of G values, according to the simulated frequencies and to the range of velocities in the considered heterogeneous medium. Doing so, it appears that the coefficients of Jo et al. (1996) actually yield a good compromise between the error level and their applicability to heterogeneous media. It is not possible to obtain significantly smaller errors without decreasing the minimal number of grid points per wavelength. Unless otherwise stated, I will therefore use these coefficients in my simulations.

On the other hand, if we accept to increase the minimal number of grid points per wavelength at the expense of the computation cost, it is possible to reach a significantly lower error level. For instance, in the application to the Institut Fresnel dataset (Section 3.1), I chose a grid step $h = 1$ mm in order to well describe the geometry of the medium. According to the considered frequencies and to the expected velocities in the medium, the corresponding values for G then range between 10 and 150. The dispersion curves obtained with stencil coefficients optimised in this range of G values are shown in Fig. 1.12(c). The error on the phase velocity is now very low for values of $1/G \leq 0.1$ — in particular, it is significantly lower than using 10 grid points per wavelength with the stencil coefficients of Jo et al. (1996) — which enables to reach the high level of accuracy obtained in Section 3.1.

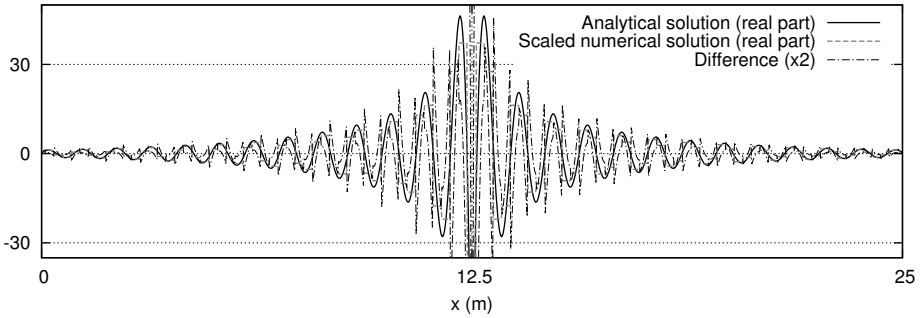
Hicks interpolation at source and receiver positions

Figs 1.10 and 1.11 show the field values computed by the FDFD scheme, and I explained how it is possible to get a good accuracy on these values by optimising the stencil coefficients. During the inversion process, we will need to extract data values from these wavefields at particular positions, where the observed data have been recorded. The fact that source and receiver positions do not necessarily coincide with the cartesian grid points can result in significant errors in the computation of synthetic data.

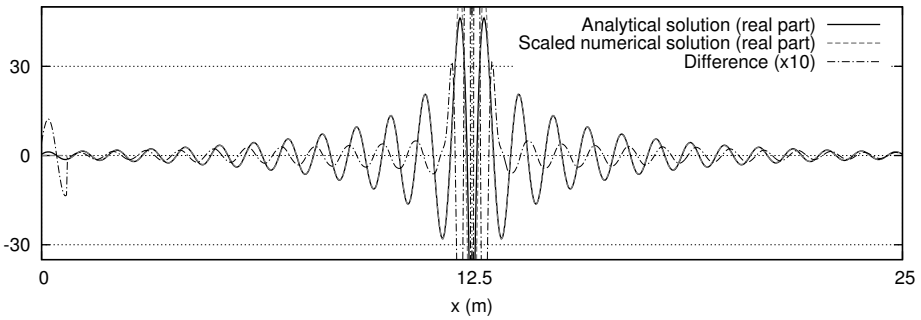
For instance, Fig. 1.13(a) shows the data recorded at receivers located along the blue line of Fig. 1.11(a) every 2.5 cm, i.e. with 10 receivers per grid step. Since the finite-difference solution is considered to be constant within a grid cell, the difference with the analytical solution can be very different for receivers located between two grid points (up to 50%). These errors due to numerical sampling will have a significant impact on the misfit between observed and synthetic data in the inversion. Taking the data as equal to the field value at the closest grid point, we lose all the benefit of having designed an accurate and efficient modelling scheme allowing for a coarse grid. An interpolation method is thus strongly needed to get accurate data values that can be compared with observed data in the inversion process.

This is achieved using Hicks interpolation (Hicks, 2002), which mimics a Dirac spatial sampling by spreading the source and receiver support over several grid nodes. A band-limited Dirac delta function is approximated using Kaiser-windowed sinc functions. Fig. 1.13(b) shows the data obtained using Hicks interpolation at source and receiver locations, spread on a 9×9-points region. The agreement with the analytical solution is now of the same order as for the wavefield (Fig. 1.10), except in the vicinity of the source where virtual receivers interact with the source position, as well as nearby the PML region because the support of the receiver located at the border of the domain overlaps in the PML. We see here the limitation of Hicks

interpolation, which is only valid if the medium is homogeneous around source and receiver locations.



(a) Without Hicks interpolation (data at receiver position is taken as the field value at the closest grid point), relative errors reach 50%.



(b) Using Hicks interpolation at source and receiver positions, the relative errors are decreased down to less than 10%, except in the vicinity of the source and nearby PML.

Figure 1.13: *Comparison of the analytical and numerical data for TE mode in a limestone described by the set of Jonscher's parameters* $\varepsilon_{\infty_r} = 8.14$, $\chi_r = 0.94$, $n = 0.82$ *and* $\sigma_{DC} = 1$ *mS/m, (a) taking the field value at closest grid point, (b) using Hicks interpolation at source and receiver locations. Receivers are located every 2.5 cm (10 receivers per grid step).*

Partial conclusion on the forward problem

In this first chapter, I have described the behaviour of electromagnetic waves in natural media, how they are affected by the electromagnetic properties, and how it was possible to parameterise these properties. To finely describe the propagation of electromagnetic waves in natural media without loss of generality, we need to consider their dispersive nature and working in the frequency domain enables to do it in a straightforward manner. However, we anticipate that considering dispersive parameters in the inversion process will be a difficult task since it increases the degrees of freedom of the problem. We have seen that the restriction to frequency-independent, real-valued permittivity and conductivity might be a reasonable approximation in some materials such as low-saturated limestones.

In a second time, I have presented the numerical direct method I will use for the imaging process, which consists in a frequency-domain finite-difference scheme using an optimised mixed-grid stencil. Optimising the stencil coefficients used in the discretisation of the partial differential equation results in a very flexible method that enables either to use of coarse grid with a reasonable level of errors, or to reach a very high level of accuracy using a finer grid. Identifying the origin of the errors due to hard source implementation, numerical dispersion and anisotropy, and Hicks interpolation is of prior importance for further applications. The properties of this numerical method will be fully exploited in the application to the laboratory controlled data of the Institut Fresnel (Section 3.1).

Finally, I shall recall the main assumption on the basis of my work, namely the restriction to 2D geometries and — more importantly — 2D propagation. This hypothesis results in a low computational cost and allowed me to take benefit from pre-existing modelling tools that could be adapted from acoustics to electromagnetics, but it also induces strong limitations that will have consequences in Chapter 3 for the interpretation of the simulated amplitudes.

Chapter 2

The inverse problem

Contents

Introduction

This chapter is dedicated to the inverse problem, i.e. to the estimation of permittivity and conductivity parameters from GPR measurements.

In Section 2.1, I first introduce the general concepts of inverse problem theory. In this section, I deliberately keep a very general mathematical frame to situate the FWI problem among the wide variety of inverse problems. Understanding the properties of this *ill-posed*, *non-linear*, *large-scale* problem justifies the choice of specific optimisation techniques to solve it. More technical details and physical insights on the imaging process by FWI are provided in subsequent sections.

In Section 2.2, I formulate the multiparameter FWI problem for the reconstruction of 2D distributions of permittivity and conductivity in the subsurface from GPR data. In particular, I investigate the sensitivity of GPR data towards permittivity and conductivity and I propose a methodological workflow for the simultaneous imaging of these parameters based on synthetic examples. This study has been published in the *Geophysical Journal International*.

In Section 2.3, I provide additional details on the formulation, mostly inherited from seismic FWI. In particular, understanding the structure of the gradient gives more insights on the resolution capability of FWI with regards to the illumination of the medium provided by the acquisition, which will be a crucial issue for the application of FWI to surface-based field measurements in Chapter 3.

2.1 Introduction to inverse problems, optimisation and FWI

In this section, I present the general frame of inverse problem theory and some key notions to understand the difficulties raised by the multiparameter FWI problem considered in the following. The reader is referred to Menke (1984, 2012) for a complete overview of linear inverse problems and to Tarantola (2005) for complements on linearised problems and for a more general probabilistic point of view. Nocedal and Wright (2006) provide a solid basis to attack local optimisation problems and I have used Hansen (2010) as an introduction to regularisation methods.

2.1.1 Definition, properties, and resolution of inverse problems in general, and of FWI in particular

Definition of an inverse problem

An inverse problem aims at inferring parameters of a considered system from data recorded on this system. Inverse problems are thus widely spread in many sciences, where collecting measurements aim at understanding the objects of study through the estimation of their properties. Inverse theory provides a mathematical framework to study inverse problems and develop systematic methods for their resolution.

When trying to guess the causes (parameters) from their consequences (data), we implicitly assume a link between both. In geophysical inverse problems, this link is furnished by physical

theories that relate the observed data $\mathbf{d}_{obs}$ to the model parameters $\mathbf{m}^*$ through the so-called *forward problem*

$$\mathbf{d}_{obs} = \mathscr{J}(\mathbf{m}^*), \tag{2.1}$$

where $\mathscr{J}$ is an operator that formalises the physical laws governing the considered process[1]. In Chapter 1, we already encountered an example of forward problem, namely the wave equation, whose resolution provides electric data depending on the electromagnetic parameters of the medium. In equation (2.1), the operator $\mathscr{J}$ can be viewed as the implicit operator that corresponds to solving the wave equation in a specific subsurface model and extracting data from the resulting wavefield. *Inversely*, solving the inverse problem is solving equation (2.1) for the vector of model parameters $\mathbf{m}^*$, knowing the observed data $\mathbf{d}_{obs}$. In our case, it consists in estimating the electromagnetic properties of the subsurface from GPR measurements.

The forward problem is generally deterministic, i.e. the assumption of some model parameters leads to a unique solution for the predicted data (at least in the case of wave propagation). The operator $\mathscr{J}$ is thus an application that associates to any element $\mathbf{m}$ of the model space $\mathbb{M}$ a unique element $\mathbf{d}$ in the data space $\mathbb{D}$:

$$\begin{array}{rccc} \mathscr{J} : & \mathbb{M} & \longrightarrow & \mathbb{D} \\ & \mathbf{m} & \longmapsto & \mathbf{d}. \end{array}$$

On the contrary, inverse problems are generally *ill-posed*, i.e. the *existence* of a solution for the inverse problem is not guaranteed, or the solution can be *non-unique*, and the resolution process is not *stable* because the solution $\mathbf{m}^*$ may depend non-continuously on the input data (Hadamard, 1902).

In addition, while the data space is generally *linear* (remember the linearity of the electromagnetic fields in Chapter 1), the relation $\mathscr{J}$ between model and data is often *non-linear*. In this respect, formulating the forward problem as $\mathbf{d} = \mathscr{J}(\mathbf{m})$ is fundamentally different from the wave equation $\mathcal{A}\mathbf{u} = \mathbf{s}$. The former directly expresses the data as a non-linear function of the model parameters, while the latter is a linear equation in $\mathbf{u}$ where the model parameters are contained implicitly in the impedance matrix $\mathcal{A}(\mathbf{m})$.

The methods involved for solving inverse problems greatly depend on these properties (ill-posedness and non-linearity). I shall now explicit these properties and provide a classification of the inverse problems and of the methods used to solve them, with a particular focus on the specific FWI problem.

Properties of (ill-posed) inverse problems

Non-existence of a strict solution to eq. (2.1) arises in real-world problems because the physical theory often provides a limited description of natural phenomena. The assumptions made in Chapter 1 for modelling the propagation of electromagnetic waves (restriction to a 2D geometry, parameterisation of the electromagnetic properties, reduction of the GPR antenna to an infinitesimal dipole) let us expect that we cannot reproduce exactly the observed data with our simulations. Even in the case of a perfect physical theory, available data are generally noisy and thus cannot be totally explained by the theory.

[1] I do not account here for possible uncertainties in the physical theory, which can be considered with the probabilistic formalism of Tarantola and Valette (1982b) or in the frame of data assimilation (Lahoz et al., 2010).

Therefore, the inverse problem (2.1) is solved in a weak form, as the following optimisation problem

$$\mathbf{m}^* = \min_{\mathbf{m}} \big| \mathbf{d}_{obs} - \mathscr{J}(\mathbf{m}) \big|, \tag{2.2}$$

i.e. we seek for the model parameters $\mathbf{m}^*$ that *best explain* the data, in the sense of some norm $|\cdot|$ which measures the distance between the observed data $\mathbf{d}_{obs}$ and calculated data $\mathbf{d}_{cal} = \mathscr{J}(\mathbf{m})$ predicted by the theory in an hypothetical model. Solving the inverse problem then amounts to minimise a *misfit* (or *objective*) function

$$\mathscr{C}(\mathbf{m}) = \big| \mathbf{d}_{obs} - \mathbf{d}_{cal}(\mathbf{m}) \big|. \tag{2.3}$$

In the following of my work, the misfit between observed and calculated data is generally measured in the least-square sense, using the ℓ_2 norm which corresponds to a classical Euclidian measure of distance. As the ℓ_1 norm is expected to be more robust in presence of noise (Brossier et al., 2009), it will be useful for the treatment of real field data (see Section 3.2).

Non-unicity of the solution arises when the problem is *under-determined*, i.e. that the data we have available are not sufficient to uniquely constrain all model parameters. Then, several set of model parameters can explain the data equivalently well.

The issue of non-unicity of the solution is intimately related to the sensitivity of the data to the model parameters. On-ground GPR probes the subsurface with waves that propagate between antennas located on the ground surface: Therefore we can expect the data to be poorly sensitive to permittivity and conductivity variations in certain portions of the investigated material, especially at large depths, and we cannot hope an accurate estimation of these variations here. To compensate the lack of information in the data, and mitigate the ill-posedness of the problem, there is no other choice than invoking other sources of information (*a priori* information, see e.g. Jackson, 1979) or, for lack of anything better, imposing some regularisation constraints to the model solutions (e.g. Tikhonov and Arsenin, 1977; Hansen, 2010).

In addition, non-unicity can arise because of *trade-offs* between parameters. By trade-off, I mean that a given variation in the data can sometimes be equivalently explained by a variation in one parameter or in another parameter. In reflection seismics, for instance, we are interested in reconstructing both the velocity in the subsurface and in locating the interfaces that create the reflections. If the available data consist only in the travel times of the reflected events recorded at zero offset, then there is a strong trade-off between the determination of the reflector depth and the velocity above this reflector, since a late arrival time can equivalently be explained by a deep reflector in a rapid material or by a shallow reflector in a slow material. It is why velocity analysis of reflection data (and FWI as well) considers multi-offset measurements. Trade-off effects are crucial in FWI, in particular when dealing with multiple parameter classes such as permittivity and conductivity, and I discuss this issue in more details in Section 2.2.

Linear, non-linear and linearised inverse problems The inverse problem is said to be linear when the function $\mathscr{J}$ is linear, i.e. when the data depend linearly on the model parameters. In the discrete case, it can then be represented by the Jacobian matrix $\mathcal{J}$ (or Fréchet derivatives). The study of linear inverse problems therefore benefits from the mathematical tools of linear algebra. In particular, the issue of non-unicity can be investigated by looking at the Kernel (or null space) of the matrix $\mathcal{J}$, e.g. by computing its singular value decomposition

(SVD, see e.g. Jackson, 1972; Hansen, 2010). More generally, linear problems are well understood. In particular, if the misfit function is defined using the ℓ_2 norm, then it is a quadratic form which has the nice property of convexity and a unique minimum. Many efficient methods then exist for finding this minimum, known as the Best Linear Unbiased Operator (BLUE) in the data assimilation community.

Unfortunately, most of inverse problems are non-linear. It is the case of FWI of GPR data since the recorded electric fields depend non-linearly on the electromagnetic properties of the subsurface. On the optimisation point of view, non-linearity manifests itself by the presence of local minima in the shape of the misfit function (see Fig. 2.1). This effect adds to the problem of non-unicity and is of major concern for *local* optimisation methods that may fall in the closest local minimum when solving eq. (2.16), while *global* optimisation methods may regard the solution as the ensemble of acceptable models (e.g. Sambridge, 1999b, 2001).

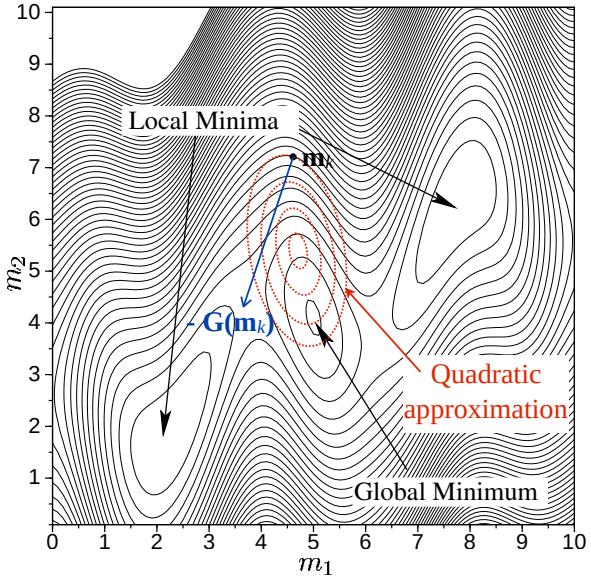


Figure 2.1: *Map of the misfit function in the case of a two-parameter, non-linear problem (modified from Sirgue, 2003). The vector* $-\mathbf{G}(\mathbf{m}_k)$ *indicates the steepest descent direction at point* $\mathbf{m}_k$.

In practice, however, we often consider the problem as linear, to take benefit of the nice properties of linear problems and of the efficient tools developed for solving them. We then speak about *linearised* problems. Linearised problems assume that the misfit function is locally quadratic (see Fig. 2.1). Under this hypothesis, it is possible to decrease the misfit function by applying one of the various methods applicable to linear problems. Non-linearity is accounted by iterating the process until convergence. A consequence of linearisation is the great sensitivity to the choice of the initial model. In Fig. 2.1, it is obvious that we can reach the global minimum only if we start from the good valley of attraction. If we start from another valley, we will fall into a local minimum.

NB: Linearising the inverse problem is *not* equivalent to linearising the forward problem, as does, for example, the Born approximation. In this work, the FWI problem of GPR data will be solved in a linearised approximation, but using the wave propagation modelling presented in Chapter 1 that fully simulates the non-linear effects of the model parameters in the data.

In FWI, non-linearities commonly arise when the misfit function is based on the simple difference between observed and calculated waveforms, because the comparison of oscillating wavelets is ambiguous. This gives rise to the so-called *cycle-skipping* effect, which is illustrated in Fig. 2.2. In frequency-domain FWI, cycle-skipping is due to the fact that monochromatic data can be equivalently matched by data having a phase difference of 2π (i.e. a time shift equal to the period T). Using local optimisation algorithms, cycle-skipping will occur if the initial model generates data that have a time shift larger than half the period of the signal. Cycle-skipping is thus more critical at high frequencies since a given time shift in the initial data is more likely to be greater than $T/2$ for high frequencies than for lower ones (see Fig. 2.2).

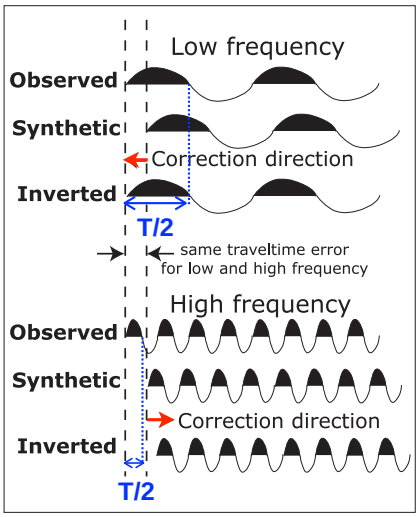


Figure 2.2: *Schematic principle of the cycle-skipping effect on monochromatic data (modified from Yokota and Matsushima, 2004).*

As a consequence, frequency-domain FWI is based on a multi-scale approach, initially promoted by Pratt and Worthington (1990). Inverting first the low frequency components of the data enables to avoid the cycle-skipping problem at the early stage of the inversion, where the initial model is not accurate enough to well predict the arrival times. High-resolution details are then added to the reconstructed image by considering higher frequencies in the inversion. This low to high frequency hierarchy has been adapted to time-domain FWI by Bunks et al. (1995) and recently used by Meles et al. (2011) for the FWI of crosshole GPR data. These authors illustrate that considering low-pass filtered data for the early steps of the inversion makes the misfit function more convex and enables to reach the global minimum when the process is iterated with higher frequencies (see Fig. 2.3).

In the following, I present the methods used for solving inverse problems in general, with a focus on the local descent algorithms used in this work to solve the non-linear FWI problem in the linearised approximation.

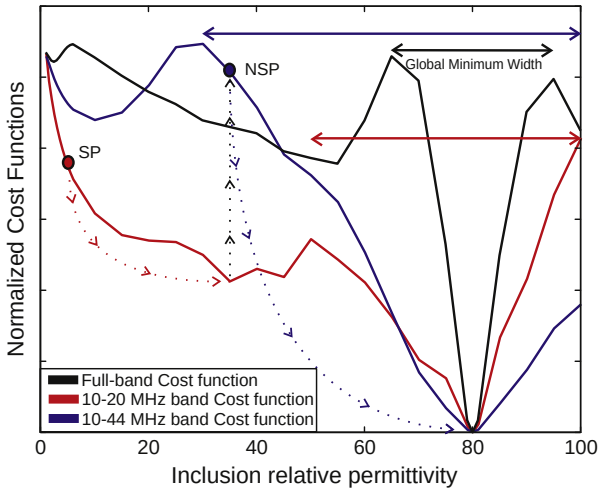


Figure 2.3: *Effect of the hierarchical approach on the convexity of the misfit function (from Meles et al., 2011).*

Methods for solving inverse problems

Direct vs. indirect methods Direct methods[1] are applicable when the inverse operator $\mathscr{J}^{-1}$ can be computed, e.g. using the specific properties of inverse scattering series (e.g. Zhang and Weglein, 2009; Kwon, 2013). Applying this inverse operator to the data then *directly* yields the solution $\mathbf{m}^* = \mathscr{J}^{-1}(\mathbf{d}_{obs})$. Of course, the inverse operator $\mathscr{J}^{-1}$ exists in a strict sense only if the problem is well posed (the function $\mathscr{J}$ must be bijective). But in the case of over-determined, and even under-determined, linear problems, *generalised inverse* operators can be used (e.g. Jackson, 1972). There is then (in the linear case) an equivalence between the generalised inverse solution, the minimisation of the misfit function (2.3) in the least-square sense, and the Bayesian approach with Gaussian statistics (Tarantola and Valette, 1982a).

In practice, however, the inverse operator $\mathscr{J}^{-1}$ can rarely be explicitly computed, either because the theory does not enable it (the applicability of inverse scattering series, for instance, is mainly restricted to problems with simple geometries, typically 1D) or because the size of the problem prevents the use of SVD. Contrary to direct methods, *indirect* methods solve the minimisation problem (2.16) by testing successive models $\mathbf{m}$ until an optimal solution is found. Indirect methods thus fall into the wide class of optimisation techniques.

Global vs. local optimisation When strongly non-linear problems involve a limited number of parameters, it is highly suitable to perform a wide exploration of the parameter space in order to find the *global* minimum of the objective function without having to guess its approximate location through the choice of an initial model (see Fig. 2.1).

If reliable *a priori* uncertainties on the data are available, the probabilistic approach of Tarantola and Valette (1982b) can be fully developed and provides complete *a posteriori* probabilities on the model parameters. Even without consistent *a priori* uncertainties, dealing with

[1]Note that *direct* methods should not be confused with so-called *direct search* methods (e.g. the Nelder-Mead simplex method, Lagarias et al., 1998), which can be indirect but are based only on evaluations of the misfit functions, without requiring to compute its derivatives like gradient-based methods.

a limited number of parameters enables to perform an exhaustive search in the parameter space (grid search) in order to find the global minimum of the misfit function. Analysing the shape of the misfit function in the parameter space enables to quantify the parameter *a posteriori* uncertainties.

Generally, global optimisation methods enable a wide, although not exhaustive, exploration of the parameter space (Sen and Stoffa, 1995). To do so, the wide class of Monte-Carlo methods perform a random sampling of the space (see e.g. Sambridge and Mosegaard, 2002, for a review). To gain computational efficiency, these methods generally try to concentrate on the most plausible zones of the parameter space, where models well explain the data. It is for instance the case for the Markov-chain Monte-Carlo methods (Metropolis and Ulam, 1949), for the simulated annealing algorithm (Kirkpatrick et al., 1983) and for the neighbourhood algorithm (Sambridge, 1999a).

Global optimisation methods are based on a systematic sampling of the parameter space and therefore require many evaluations of the misfit function, i.e. many resolutions of the forward problem. This is unaffordable if the number of searched parameters is too large or if the resolution of the forward problem is computationally expensive. In the following, I consider the imaging of 2D sections of the subsurface, discretised in pixels forming a cartesian grid[1]. For instance, in the case of the cross-shaped benchmark investigated in Section 2.2, the image consists of a cartesian grid with 201×201 pixels. For global optimisation methods, each pixel in the image constitutes a dimension of the model space to be explored, which would require too many simulations to ensure a dense sampling of the parameter space (of the order of 10^5 to 10^6). Even in this small case, where simulations request about 2s of CPU time[2], testing ten different values for each permittivity and conductivity parameter (which is a very coarse sampling) would demand about 450h of CPU time.

Therefore, when the number of model parameters involved is too large, or when the forward problem is computationally expensive, we must opt for *local* optimisation methods. Starting from an initial model $\mathbf{m}_o$, these methods iteratively decrease the misfit function until convergence to a local minimum and require much less computations (typically of the order of 10 to 1000 simulations: in the above-mentioned example, inversions then necessitate less than 12h of CPU time). Among local optimisation methods, the most popular are local descent techniques based on the gradient of the misfit function, which corresponds to the local slope in Fig. 2.1. I now detail their principles since it is the type of method used for my inversions.

2.1.2 Local descent optimisation algorithms

Starting from an initial model $\mathbf{m}_o$, local descent methods create a sequence of models $\mathbf{m}_k$ with, at each iteration k,

$$\begin{aligned} \mathbf{m}_{k+1} &= \mathbf{m}_k + \Delta\mathbf{m}_k, \\ &= \mathbf{m}_k + \alpha_k \mathbf{p}_k, \end{aligned} \tag{2.4}$$

[1]Pixel-based representation of the model space are currently used in FWI, and more largely for tomography and imaging. Of course, we may discuss the relevance of such a representation. Some authors prefer using a coarser representation, enabling the use of global optimisation (e.g. Cordua et al., 2012), but it generally implies some assumptions on the geometry of the medium (layers or blocks). In Section 2.3.3, I show how to invert for a blocky representation of the medium using the linearised FWI developed in this work.

[2]When run in parallel on 8 cores of the R2D2 cluster of the CIMENT platform (Univ. Grenoble).

where $\Delta\mathbf{m}_k$ is a model perturbation that results from a descent step of length α_k in the direction $\mathbf{p}_k$. All the art of local descent methods is to build suitable descent directions and step lengths that enable to decrease the misfit function until convergence.

Line search methods

Let first assume that a descent direction vector $\mathbf{p}_k$ is known (it is generally not difficult to find one, e.g. the steepest descent direction given by the gradient of the misfit function, see Fig. 2.1). Then, the scalar descent step length α_k is generally determined using a line search method[1] that seeks for a step length which decreases the misfit function in the descent direction. Fig. 2.4 illustrates the principle of a line search. It represents the values of the misfit function in the descent direction $\mathbf{p}_k$, as a one-dimensional function of the possible step lengths α. It can be seen as a section of the 2D map of Fig. 2.1, e.g. in the direction of the steepest direction indicated by the gradient $\mathbf{G}(\mathbf{m}_k)$ [2].

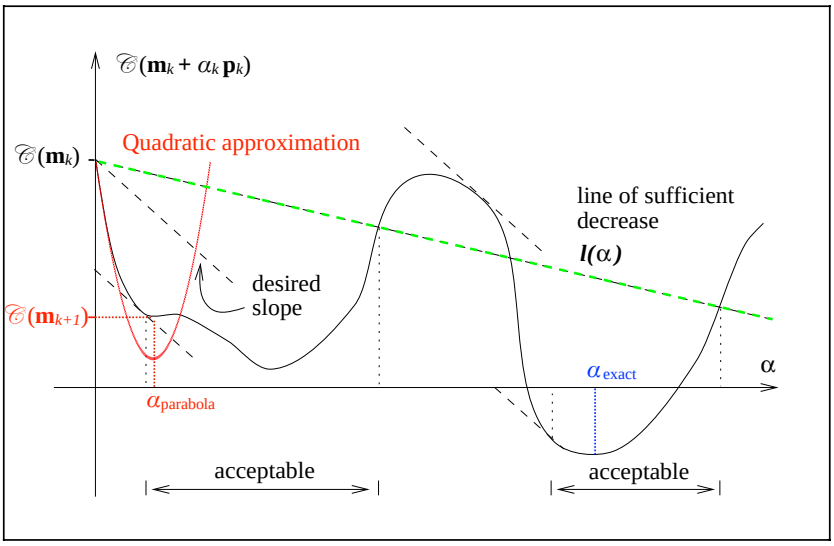


Figure 2.4: *Schematic principle of a local descent algorithm, illustrating the quadratic approximation and the inexact line search for step lengths that satisfy the Wolfe conditions (modified from Nocedal and Wright, 2006).*

A first approach to determine the step length α_k is to perform an *exact* line search by seeking the minimum of the misfit function in the descent direction. The resulting step length is indicated in blue in Fig. 2.4. This strategy, however, is difficult to apply to strongly non-linear problems since it would require the evaluation of the misfit function at many points on the descent direction before finding the minimum.

Therefore, *inexact* line search methods are generally used. A practical example of line search is parabola fitting. This method makes the assumption that the misfit function is locally

[1] Alternatives to line search methods are trust-region methods (Nocedal and Wright, 2006, chap. 4). Instead of defining first the descent direction and then the descent step in this direction using a line search, trust-region methods promote to define first a maximal amplitude for the model perturbation and then the adequate descent direction within this region.

[2] Please note that the comparison is only illustrative, the two examples are different.

quadratic and seek for the minimum of the corresponding parabola (see Fig. 2.4). Building a quadratic approximation requires the evaluation of one more value of the misfit function (three points are needed to define a parabola). If the problem is linear, then this method is an exact line search. However, if the misfit function is strongly non-linear, nothing guarantees that the step length found using parabola fitting effectively decreases the misfit.

Therefore, two conditions are imposed to inexact line search methods to make them efficient:

1. The *sufficient decrease condition* stipulates that a suitable descent step length should decrease the misfit function in a sufficient amount (green dashed line in Fig. 2.4).
2. The *sufficient curvature condition* imposes that the slope at the arriving point $\mathbf{m}_{k+1}$ should be greater than the (negative) slope at point $\mathbf{m}_k$ (black dashed lines in Fig. 2.4). This second condition avoids stopping at points where the slope would enable to further decrease the misfit function. It thus avoids accepting too small step lengths that would make the convergence very slow.

These conditions, known together as the Wolfe conditions (Nocedal and Wright, 2006, §3.1, p. 33), guarantee the convergence of the local descent to a local minimum, provided $\mathbf{p}_k$ is a descent direction. Now I explain how it is possible to find a suitable descent direction $\mathbf{p}_k$.

Steepest-descent method

If we think the representation of the misfit function in Fig. 2.1 as a topographical map, the most intuitive way to decrease the misfit at iteration k is to follow the slope, i.e. to take the steepest descent direction which is by definition the opposite of the gradient of the misfit function. The steepest descent direction provides the most efficient decrease misfit if small step lengths are used. As a drawback, the method converges generally slowly, even in the case of linear problems, as illustrated in Fig. 2.5(a). Moreover, the performance of the steepest descent method is very sensitive to the scaling of the problem (Fig. 2.5b), which is a major issue we will encounter in multiparameter FWI, due to the different units, range of values, and impact of permittivity and conductivity parameters on the data (see Section 2.2).

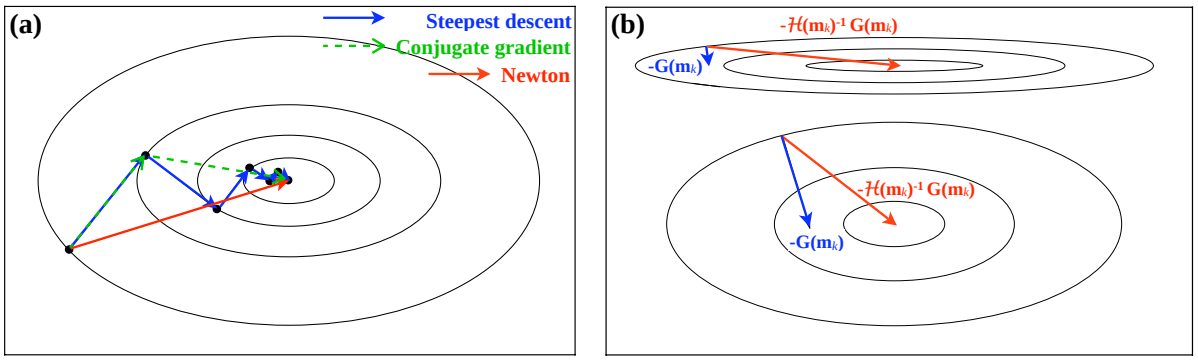


Figure 2.5: *(a) Examples of descent steps using the steepest descent, the conjugate gradient or the Newton methods with an exact line search in a linear case. (b) Performance of the steepest descent and of the Newton methods in the case of poorly scaled (top) and well scaled (bottom) linear problems (modified from Nocedal and Wright, 2006).*

Conjugate gradient methods

Conjugate gradient methods improve the steepest descent method by combining the current steepest descent direction with the previous descent direction:

$$\mathbf{p}_k = -\mathbf{G}(\mathbf{m}_k) + \beta_k \mathbf{p}_{k-1}, \tag{2.5}$$

where β_k is a scalar factor designed such that $\mathbf{p}_k$ is *conjugate* to all previous descent directions (hence the name of the method). This feature confers to conjugate gradient methods the remarkable property to converge in n iterations in the case of a n-parameter linear problem (see Fig. 2.4 for the two-parameter case and Nocedal and Wright, 2006, §5.1, p. 103 for a proof).

As only the gradient and the previous descent direction are used, conjugate gradient methods require a small amount of memory and computation. Consequently, they are widely applied to large scale problems, in particular to design efficient iterative solvers of large linear equations systems (Hestenes and Stiefel, 1952). Many variants of the conjugate gradient method exist and can be applied to non-linear problems in an iterative manner: they mainly vary by their definition of the factor β_k used to combine the descent directions (e.g. Fletcher and Reeves, 1964; Polak and Ribière, 1969).

Newton method

One step further to improve the descent direction can be gained by looking at the Taylor expansion of the misfit function:

$$\begin{aligned} \mathscr{C}(\mathbf{m}_k + \Delta\mathbf{m}_k) &= \mathscr{C}(\mathbf{m}_k) + \frac{\partial \mathscr{C}(\mathbf{m}_k)}{\partial \mathbf{m}}^T \Delta\mathbf{m}_k + \frac{1}{2}\Delta\mathbf{m}_k^T \frac{\partial^2 \mathscr{C}(\mathbf{m}_k)}{\partial \mathbf{m}^2} \Delta\mathbf{m}_k + o(||\Delta\mathbf{m}_k||^2), \\ &= \mathscr{C}(\mathbf{m}_k) + \mathbf{G}(\mathbf{m}_k)^T \Delta\mathbf{m}_k + \frac{1}{2}\Delta\mathbf{m}_k^T \mathcal{H}(\mathbf{m}_k) \Delta\mathbf{m}_k + o(||\Delta\mathbf{m}_k||^2), \end{aligned} \tag{2.6}$$

where $\mathbf{G}$ is the gradient vector (first order derivatives) and $\mathcal{H}$ is the Hessian matrix (second order derivatives) of the misfit function.

Assuming $\mathscr{C}$ is quadratic (i.e. the problem is linear), we can minimise it in one single step, i.e. we can reach the point $\mathbf{m}_k + \Delta\mathbf{m}_k$ where the gradient is null. The condition on the required perturbation to do so is given by setting the derivatives of eq. (2.6) to 0:

$$\mathbf{0} = \mathbf{G}(\mathbf{m}_k + \Delta\mathbf{m}_k) = \mathbf{G}(\mathbf{m}_k) + \mathcal{H}(\mathbf{m}_k)\Delta\mathbf{m}_k. \tag{2.7}$$

We obtain the Newton equation system, also called normal equations:

$$\mathcal{H}(\mathbf{m}_k)\Delta\mathbf{m}_k = -\mathbf{G}(\mathbf{m}_k). \tag{2.8}$$

If the Hessian matrix is invertible, then the system can be solved straightforwardly, giving the suitable perturbation $\Delta\mathbf{m}_k$ required to minimise the quadratic approximation of $\mathscr{C}(\mathbf{m}_k)$:

$$\Delta\mathbf{m}_k = -\mathcal{H}(\mathbf{m}_k)^{-1}\mathbf{G}(\mathbf{m}_k), \tag{2.9}$$

we then arrive directly at the minimum of the quadratic function (see Fig. 2.5b), without having to search for a suitable step length. In particular, the descent direction is not sensitive

anymore to the scaling of the problem: we see here the enormous benefit of using the inverse Hessian in the optimisation, in particular in the frame of multiparameter inversion.

Unfortunately, the exact inverse Hessian is rarely available. First because the Hessian matrix might be non-invertible due to the indetermination of the problem. In the linear two-parameter case of Fig. 2.5(b), indetermination would manifests itself as a *very* poorly scaled problem, represented by an ellipse further elongated in the direction of one parameter: the misfit function would be flat (i.e. insensitive) with respect to this parameter. Indetermination is measured quantitatively by the conditioning of the Hessian matrix, which is by definition the ratio between its higher and smaller eigenvalues. Geometrically, it corresponds to the ratio between the axis lengths of the ellipses of Fig. 2.5. If this ratio is high (if the ellipse is elongated), the Hessian matrix is *ill-conditioned*. It is then very difficult to invert the matrix and to solve the problem.

Secondly, the computation of the Hessian is very expensive: if estimated by finite-differences, it requires at least three computations of the data per parameter. More sophisticated methods exist to achieve the computation more efficiently (e.g., second order adjoint methods) but for large-scale problems, the storage, manipulation and, *a fortiori*, the inversion of this large matrix is prohibitive[1]. Instead, the linear Newton system (2.8) can be solved iteratively using matrix-free approaches, e.g., conjugate gradient methods that are well suited for this purpose (see e.g. the truncated Newton method, Métivier et al., 2013). These methods remain, however, relatively expensive in terms of simulations since an inner iterative loop dedicated to the resolution of the Newton system (2.8) is added to the outer iterative scheme of the local descent optimisation.

Quasi-Newton methods

Quasi-Newton methods attempt to approximate the inverse Hessian in eq. (2.8) without extra computations. The model update then reads

$$\Delta \mathbf{m}_k = -\alpha_k \mathcal{B}_k^{-1} \mathbf{G}(\mathbf{m}_k), \tag{2.10}$$

where $\mathcal{B}_k$ is an approximation to the Hessian $\mathcal{H}(\mathbf{m}_k)$. As an approximation, it does not enable to reach exactly the minimum of a quadratic function and the descent step length α_k must be determined using a line search procedure.

Among quasi-Newton methods, I concentrate on the BFGS formula (invented by Broyden, Fletcher, Goldfarb and Shanno), which constructs the approximate inverse Hessian at each iteration as

$$\mathcal{B}_{k+1}^{-1} = \left(\mathcal{I} - \rho_k \mathbf{s}_k \mathbf{y}_k^T\right) \mathcal{B}_k^{-1} \left(\mathcal{I} - \rho_k \mathbf{s}_k \mathbf{y}_k^T\right) + \rho_k \mathbf{s}_k \mathbf{s}_k^T, \tag{2.11}$$

using only the previous approximate inverse Hessian and the previous model and gradient vector to define the vectors $\mathbf{s}_k = \mathbf{m}_{k+1} - \mathbf{m}_k$ and $\mathbf{y}_k = \mathbf{G}(\mathbf{m}_{k+1}) - \mathbf{G}(\mathbf{m}_k)$, and the scalar $\rho_k = 1/\mathbf{y}_k^T \mathbf{s}_k$.

More precisely, I will use the *limited* BFGS method (L-BFGS). In fact, developing the recursive relation (2.11), we can express the updated inverse matrix $\mathcal{B}_{k+1}^{-1}$ as a function of

[1] To give an order of magnitude, in the *small* case of the cross-shaped benchmark investigated in Section 2.2, discretised on a 201×201 cartesian grid, the size of the Hessian is of the order of $1.6\ 10^6$ elements when a single parameter class is considered, requiring $\simeq 6.5$ Go of memory to store it. These quantities are multiplied by a factor 4 in the multiparameter permittivity-conductivity case.

all previous vectors $\mathbf{s}_o$, $\mathbf{s}_1$, ..., $\mathbf{s}_k$ and $\mathbf{y}_o$, $\mathbf{y}_1$, ..., $\mathbf{y}_k$. Doing so requires the storage of only $2 \times (k+1)$ vectors of size N_M (the number of model parameters), instead of a dense matrix of size $N_M \times N_M$ whose storage and manipulation can be prohibitive for large scale problems. An additional gain is achieved by discarding the oldest model and gradient vectors, keeping only the n_L previous ones: this gives rise to the *limited* BFGS method (L-BFGS) used in the following of my work.

As a conclusion to this introductory section to the FWI problem, the reader should keep in mind that I will perform the resolution of this non-linear problem in a linearised way, using the quasi-Newton L-BFGS algorithm. In the next section, I formulate more explicitly the inverse problem consisting in the reconstruction of 2D distributions of permittivity and conductivity from GPR measurements. The consideration of the effect of an approximate inverse Hessian through the quasi-Newton update is expected to provide a robust multiparameter inversion scheme. In practice, we shall see that the imaging procedure suffers from the non-linearity illustrated in Fig. 2.1 and is very sensitive to the scaling of the problem (Fig. 2.5b).

In Section 2.3, I provide additional details on the FWI formulation, especially on the gradient of the misfit function which is the key ingredient used by the quasi-Newton optimisation. The role of the gradient in the imaging process will be commented to understand the resolution capability of FWI when considering surface-based GPR measurements, compared to crosshole experiments for which the method has already been successful.

2.2 A strategy for multiparameter FWI

This section has been published in the *Geophysical Journal International* (Lavoué et al., 2014). It is presented as it appeared in the journal. Slight modifications only concern the layout and notations, for consistency with the rest of the manuscript. In section 2.3, I detail some methodological points which could not be included in the article for conciseness but which bring additional insights into the multiparameter FWI problem.

Two-dimensional permittivity and conductivity imaging by full waveform inversion of multioffset GPR data: a frequency-domain quasi-Newton approach

F. Lavoué, R. Brossier, L. Métivier, S. Garambois and J. Virieux

Geophysical Journal International, 197 (1), 248-268, *doi:10.1093/gji/ggt528*, 2014.

Abstract

Full waveform inversion of ground-penetrating radar data is an emerging technique for the quantitative, high-resolution imaging of the near subsurface. Here, we present a 2D frequency-domain full waveform inversion for the simultaneous reconstruction of the dielectric permittivity and of the electrical conductivity. The inverse problem is solved with a quasi-Newton optimisation scheme, where the influence of the Hessian is approximated by the L-BFGS-B algorithm. This formulation can be considered to be fully multiparameter since it enables to update permittivity and conductivity values within the same descent step, provided we define scales of measurement through a reference permittivity, a reference conductivity, and an additional scaling factor. Numerical experiments on a benchmark from the literature demonstrate that the inversion is very sensitive to the parameter scaling, despite the consideration of the approximated Hessian that should correct for parameter dimensionalities. A proper scaling should respect the natural sensitivity of the misfit function and give priority to the parameter that has the most impact on the data (the permittivity, in our case). We also investigate the behaviour of the inversion with respect to frequency sampling, considering the selected frequencies either simultaneously or sequentially. As the relative imprint of permittivity and conductivity in the data varies with frequency, the simultaneous reconstruction of both parameters takes a significant benefit from broad frequency bandwidth data, so that simultaneous or cumulative strategies should be favoured. We illustrate our scaling approach with a realistic synthetic example for the imaging of a complex subsurface from on-ground multi-offset data. Considering data acquired only from the ground surface increases the ill-posedness of the inverse problem and leads to a strong indetermination of the less-constrained conductivity parameters. A Tikhonov regularisation can prevent the creation of high-wavenumber artefacts in the conductivity model that compensate for erroneous low-wavenumber structures, thus enabling to select model solutions. We propose a workflow for multiparameter imaging involving both parameter scaling and regularisation. Optimal combinations of scaling factors and regularisation weights can be identified by seeking regularisation levels that exhibit a clear minimum of final data misfit with respect to parameter scaling. We confirm this workflow by inverting noise-contaminated synthetic data. In a surface-to-surface acquisition configuration, we have been able to reconstruct an accurate permittivity structure and a smooth version of

the conductivity distribution, based entirely on the analysis of the data misfit with respect to parameter scaling, for different regularisation levels.

Introduction

Ground-penetrating radar (GPR) is a non-invasive subsurface prospecting technique based on the propagation of electromagnetic waves. Similar in its principle to seismic reflection experiments, GPR imaging took large benefits from seismic processing developments such as migration, so that the method provides today accurate qualitative images of the subsurface from constant offset measurements (e.g. Fischer et al., 1992b; Grasmueck et al., 2005) and more rarely from multi-offset measurements (Fischer et al., 1992a; Greaves et al., 1996; Bradford, 2008; Gerhards et al., 2008). The development of a quantitative imaging that would estimate the electromagnetic properties of the sounded medium — mainly the dielectric permittivity ε [F/m] and the electrical conductivity σ [S/m] — appears as a critical issue for a physical interpretation of the target structures. In particular, geological, hydrological, or geotechnical applications need important informations such as the composition of the material (Deeds and Bradford, 2002; Ihamouten et al., 2012) or its water content (Garambois et al., 2002; Huisman et al., 2003; Day-Lewis et al., 2005; Weihermüller et al., 2007).

Up to now, efforts have been oriented towards quantitative GPR imaging using multi-offset measurements with velocity analysis (e.g. Fischer et al., 1992a), amplitude-versus-offset studies (e.g. Deeds and Bradford, 2002; Deparis and Garambois, 2009), travel-time and amplitude tomography (Cai et al., 1996; Holliger et al., 2001; Gloaguen et al., 2005; Musil et al., 2006), and full waveform inversion (FWI). The latter is one of the most promising techniques for building quantitative, high-resolution images of the subsurface. Contrary to velocity analysis or tomography which exploit a few events in the radargram, FWI takes benefit from the whole recorded signal. Originating from the time-domain seismic imaging (Lailly, 1983; Tarantola, 1984), FWI has then been developed for frequency-domain data (Pratt and Worthington, 1990; Pratt et al., 1998). The frequency-domain approach makes an efficient use of the data by inverting only few frequency components, taking benefit of the data redundancy provided by the acquisition. It also enables to mitigate the non-linearity of the inverse problem by following a low to high frequency hierarchy (Pratt and Worthington, 1990; Sirgue and Pratt, 2004). Synthetic and real seismic applications of this approach have been successful (see Virieux and Operto, 2009, for an overview). The interest of FWI for GPR data has been demonstrated in recent applications for water content estimation in the first centimetres of agricultural soils (Lambot et al., 2006; Minet et al., 2010) and for the estimation of permittivity and conductivity in stratified structures such as concrete (Kalogeropoulos et al., 2011, 2013; Patriarca et al., 2011) or layered soils (Busch et al., 2012). In addition, FWI has been applied to 2D crosshole sections by Ernst et al. (2007), Meles et al. (2010), Klotzsche et al. (2010, 2012) and Cordua et al. (2012) in the time-domain, and by El Bouajaji et al. (2011), Ellefsen et al. (2011), and Yang et al. (2012) in the frequency-domain. Among the existing literature, only Lopes (2009) and El Bouajaji et al. (2011) tackle the interpretation of surface-based GPR measurements for the quantitative imaging of 2D sections of the medium. However, these authors restrict themselves to monoparameter inversions, reconstructing only the permittivity distribution. In this paper, we propose a FWI method for the simultaneous inversion of permittivity and conductivity in 2D, with a particular interest in data acquired in surface-to-surface multi-offset configuration (on-ground GPR).

Compared to crosshole GPR configurations, on-ground GPR measurements provide a reduced coverage of the subsurface at depth, which tends to increase the ill-posedness of the inverse problem (Meles et al., 2012). Moreover, on-ground GPR measurements illuminate the subsurface targets with small reflection angles which may provide a better resolution but a lack of low wavenumbers compared to crosshole experiments (Sirgue and Pratt, 2004), making the design of an adequate initial model to start the FWI process more critical. The reduced illumination may also increase the trade-off between the two parameter types that are permittivity and conductivity (Hak and Mulder, 2010), making the multiparameter imaging more challenging. In addition, crosshole and on-ground GPR data may have different sensitivities to permittivity and conductivity due to the fact that on-ground GPR is mainly based on reflections and diffractions whereas crosshole data mostly contain transmitted signal. In the paper, we spend some time to describe the sensitivity of on-ground GPR data to permittivity and conductivity. Other differences between on-ground and crosshole GPR concern the mode used for the measurement (TE vs. TM, respectively), and the influence of the air-ground interface, which both have an effect on the antenna radiation pattern, but we do not investigate these aspects.

The frequency-domain FWI is formulated as an optimisation problem which consists in minimising a misfit function that measures the distance between observed and calculated data. The minimisation is achieved through a local descent method. We shall focus our attention on the limited Broyden-Fletcher-Goldfarb-Shanno bounded algorithm (L-BFGS-B, Byrd et al., 1995), which belongs to the family of quasi-Newton optimisation schemes. In this algorithm, the effect of the inverse Hessian operator is approximated through previous gradients and updated models, limiting the demand on computer resources. The consideration of the approximated Hessian is expected to improve the convergence of the optimisation process, to partially correct for wave propagation effects such as geometrical spreading and double scattering, and to deconvolve the finite-frequency artefacts due to the limited bandwidth of the source and to the discrete acquisition sampling (Pratt et al., 1998). In a multiparameter framework, the approximated Hessian should also account for differences in sensitivity of the misfit function with respect to different types of parameters, such as permittivity and conductivity. Therefore, an important advantage of our quasi-Newton formulation is that it enables to update permittivity and conductivity simultaneously within the same descent step, and thus to consider the parameter trade-offs (Operto et al., 2013), unlike alternated or decoupled approaches (Ernst et al., 2007; Meles et al., 2010). Besides, the proper consideration of bounds for the possible range of parameters values through active sets in the implementation of L-BFGS-B is of great interest for GPR imaging, where physical limits are often encountered (in the air, for instance).

In the first part of the paper, we begin with a short presentation of the forward problem. We then consider the inverse problem formulation and its resolution with a focus on quasi-Newton schemes of optimisation. In a second part, we highlight the difficulty raised by the simultaneous reconstruction of permittivity and conductivity, due to their different impacts on the data. We illustrate these different sensitivities on a simple synthetic case with perfect illumination inspired from Meles et al. (2011). We will see that the inversion is sensitive to the scale used for the definition of the reconstructed parameters, despite the consideration of the approximated Hessian that should correct for parameter dimensionalities. For a better insight into this problem, we analyse the behaviour of the multiparameter scheme with respect to the parameter scaling and to the frequency sampling strategy. In a third part, we illustrate the proposed imaging method on a more realistic synthetic case with a surface-to-surface acquisition. In this

case, we will show that the parameter scaling must be combined with a regularisation term to prevent the optimisation for over-interpreting the data with undesired structures appearing in the image. Finally, noise will be introduced in the data in order to investigate the feasibility of such approach for future real data inversion.

2.2.1 Forward and inverse problems formulation in the frequency domain

In this section, we first introduce the electromagnetic forward problem in the frequency domain and associated notations. We then formulate the inverse problem and detail the optimisation algorithm used for its resolution.

2.2.1.1 Forward problem

Restricting the Maxwell's equations to a 2D geometry leads to two decoupled systems: the transverse electric mode (TE) and the transverse magnetic mode (TM). In the following, we focus on the TE mode, vibrating in the (xOz) plane, for an electric dipole source oriented along the y-axis. The mathematical analogy between the acoustic system and the TE mode (Carcione and Cavallini, 1995) leads to the following scalar wave equations:

$$\nabla^2 P(\omega,x,z) + \frac{\omega^2}{v_P^2(x,z)} P(\omega,x,z) = \delta(x-x_S)\delta(z-z_S), \tag{2.12}$$

$$\Leftrightarrow \quad \nabla^2 E_y(\omega,x,z) + \varepsilon_e(\omega,x,z)\mu\omega^2 E_y(\omega,x,z) = \delta(x-x_S)\delta(z-z_S), \tag{2.13}$$

where P denotes the acoustic pressure field (in Pa), v_P its velocity (in m/s), E_y the component of the electric field in TE mode (in V/m), μ the magnetic permeability (in H/m), and ε_e a complex-valued effective permittivity, accounting for both propagation and attenuation. The source is located at (x_S, z_S) and the angular frequency is denoted by ω [rad/s]. We use a conventional time-harmonic dependency in $e^{-\imath\omega t}$, denoting $\imath = \sqrt{-1}$ the imaginary unit.

We focus on non-magnetic media, where the permeability μ is homogeneous and equal to $\mu_o = 4\pi.10^{-7}$ H/m (vacuum value). Moreover, we consider the simplest expression for the effective permittivity

$$\varepsilon_e(\omega,x,z) = \varepsilon(x,z) + \imath\sigma(x,z)/\omega, \tag{2.14}$$

where we assume that the dielectric permittivity ε [F/m] and the electrical conductivity σ [S/m] are real quantities. More elaborate parameterisations of the effective permittivity ε_e can be considered to account for dispersive effects (Debye, Cole-Cole, or Jonscher parameterisations, see e.g. Hollender and Tillard, 1998).

The analogy between Maxwell's TE mode and acoustic propagation (equations 2.12 and 2.13) enables us to simulate electromagnetic waves with an optimised finite-difference scheme developed for seismic modelling in the visco-acoustic approximation and introduced by Hustedt et al. (2004). This formulation leads to a linear system of the form

$$\mathcal{A}(\omega,\varepsilon,\sigma)\mathbf{u}(\omega) = \mathbf{s}(\omega), \tag{2.15}$$

where the complex impedance matrix is denoted by $\mathcal{A}$, the vector of the simulated wavefield by $\mathbf{u}$ (E_y in TE mode), and the discrete source term by $\mathbf{s}$. The linear system (2.15) is solved using the direct solver MUMPS (MUMPS-team, 2009). Only one LU factorisation of the

matrix $\mathcal{A}$ is needed for a given medium and frequency. Once the matrix is factorised, the resolution of equation (2.15) for multiple right-hand-side terms can be achieved very efficiently. Computations are performed on a finite-difference grid of size (N_x, N_z) and the order of the matrix $\mathcal{A}$ is $N_M = N_x \times N_z$. Perfectly matched layers (PML) are used to absorb the waves at the boundary of the medium (Bérenger, 1994).

2.2.1.2 Inverse problem

We formulate the FWI problem as the minimisation of a misfit function $\mathscr{C}(\mathbf{m})$:

$$\min_{\mathbf{m}} \mathscr{C}(\mathbf{m}), \tag{2.16}$$

where the model vector $\mathbf{m}$ of size $2 \times N_M$ gathers both permittivity and conductivity values at each point of the finite-difference grid. The misfit function is defined as the fit to the data through the ℓ_2 norm of the residuals $\Delta\mathbf{d}$:

$$\mathscr{C}(\mathbf{m}) = \frac{1}{2} \sum_{i=1}^{N_\omega} \sum_{j=1}^{N_s} \Delta\mathbf{d}(\omega_i, \mathbf{s}_j)^\dagger \Delta\mathbf{d}(\omega_i, \mathbf{s}_j), \tag{2.17}$$

where the symbol † denotes the transpose (^{T}) - conjugate (*) operator. For each of the N_ω frequencies ω_i and each of the N_s sources $\mathbf{s}_j$, the residuals are defined as the difference $\Delta\mathbf{d} = \mathbf{d}_{obs} - \mathbf{d}_{cal}$ between observed data $\mathbf{d}_{obs}$ and calculated data $\mathbf{d}_{cal}$. The calculated data are extracted from the simulated wavefield $\mathbf{u}$ through a projection operator $\mathcal{R}$ to the receiver locations: $\mathbf{d}_{cal} = \mathcal{R}\mathbf{u}$, with $\mathbf{u}$ verifying equation (2.15).

The inverse problem (2.16) is solved through a local descent algorithm: From an initial guess $\mathbf{m}_o$, we build the sequence $\mathbf{m}_k$ such that, for each iteration k,

$$\mathbf{m}_{k+1} = \mathbf{m}_k - \alpha_k \mathcal{B}_k^{-1} \mathbf{G}_k, \tag{2.18}$$

where the scalar α_k denotes the descent step length, the matrix $\mathcal{B}_k$ is an approximation of the Hessian (the second-order derivative of the misfit function with respect to the model parameters) and the vector $\mathbf{G}_k = \mathbf{G}(\mathbf{m}_k)$ is the gradient of the misfit function.

At each grid point $i \in [\![1, N_M]\!]$ in the medium, the gradient value G_i is computed using the adjoint state method (Plessix, 2006) as

$$G_i(\mathbf{m}) = \sum^{N_\omega} \sum^{N_s} \mathcal{R}e \left\{ \mathbf{u}^T \frac{\partial \mathcal{A}}{\partial m_i}^T \mathbf{v}^* \right\}. \tag{2.19}$$

In this formulation, the adjoint wavefield, denoted by $\mathbf{v}$ and verifying the linear system $\mathcal{A}^\dagger \mathbf{v} = \mathcal{R}^\dagger \Delta\mathbf{d}$, corresponds to the back-propagation of the residuals in the medium. In practice, MUMPS software enables to solve for the conjugate equation $\mathcal{A}^T \mathbf{v}^* = \mathcal{R}^T \Delta\mathbf{d}^*$ without computing $\mathcal{A}^T$ and its LU factorisation again. The diffraction matrix $\partial_{m_i}\mathcal{A}$ (or sensitivity kernel) characterises the sensitivity to the parameter m_i, that refers either to the permittivity ε_i or to the conductivity σ_i at grid point i. In the finite-difference scheme of Hustedt et al. (2004), it can be expressed as

$$\frac{\partial \mathcal{A}_{ij}}{\partial \varepsilon_i} = -\omega^2 \delta_{ij}, \quad \text{and} \quad \frac{\partial \mathcal{A}_{ij}}{\partial \sigma_i} = -\imath\omega \delta_{ij}, \tag{2.20}$$

where δ_{ij} is the Kronecker symbol ($\delta_{ij} = 1$ if $i = j$, and 0 otherwise).

In expression (2.18), the model update $\Delta\mathbf{m}_k = -\alpha_k \mathcal{B}_k^{-1}\mathbf{G}_k$ is estimated by the L-BFGS-B algorithm (Byrd et al., 1995). The descent step length α_k is determined using an inexact line search based on the Wolfe conditions (Nocedal and Wright, 2006, p. 33). In practice, $\mathcal{B}_k$ is never built explicitly: The L-BFGS-B algorithm directly builds the matrix-vector product $\mathcal{B}_k^{-1}\mathbf{G}_k$ using a limited number n_L of vectors of the form $\mathbf{s}_l = \mathbf{m}_{l+1} - \mathbf{m}_l$ and $\mathbf{y}_l = \mathbf{G}_{l+1} - \mathbf{G}_l$, with $k - n_L \leq l \leq k - 1$, which limits the storage requirements by making use of the n_L most recent models and gradients only (Nocedal and Wright, 2006, p.177). In our numerical tests, we set $n_L = 5$ as we have found that higher values do not improve the results in the configurations we consider. Since the gradient computation requires one direct and one adjoint simulations, the algorithm needs the resolution of approximately two forward problems per iteration, per source, and per frequency. An over-cost can occur in the line search procedure if the initial step length $\alpha_k = 1$ is not accepted (which is rare). In our experiments, the iterative process stops when the norm of the model update is smaller than 10^4 times the machine precision[1].

The design of a suitable initial model $\mathbf{m}_o$ for starting the full waveform inversion scheme is a crucial point but it is out of the scope of this study. In our numerical tests, we will start either from an obvious background value or from a smooth version of the true model. In the case of real data, an initial permittivity model could be recovered by velocity analysis (hyperbolae fitting or semblance analysis), whereas other geophysical methods can provide a smooth initial model for conductivity (e.g. electromagnetic induction measurements or electrical resistivity tomography). An other important point when dealing with true data is the estimation of the source signature, which is not investigated here (in our tests, we will assume that we know the exact Dirac source in equation 2.13). Usually, the frequency components of the source signal are estimated either by solving an over-determined quadratic problem at each iteration (Pratt and Worthington, 1990) or by including the phase and the amplitude of the source in the parameters to be inverted (Pratt and Worthington, 1990; Busch et al., 2012). Finally, we do not consider the complex radiation pattern of a real antenna, assuming that the source is an infinitesimal electric dipole.

2.2.2 Multiparameter imaging of permittivity and conductivity

To understand how the multiparameter inversion behaves, we first perform numerical experiments on a synthetic example introduced by Meles et al. (2011) (see Fig. 2.6). In these tests, we are interested in reconstructing the two cross-shaped anomalies, starting from the homogeneous background. The step h of the finite-difference grid is taken as $h = 7$ cm to ensure at least four discretisation points per propagated wavelength. This results in $N_M = 201 \times 201 = 40401$ grid points and as many discrete unknowns for permittivity and conductivity. Targets are surrounded by 40 sources and 120 receivers in a perfect illumination configuration, which means that the signal of each source is recorded by all the receivers. Fig. 2.7 shows an example of time-domain shot gathers computed in the true and in the initial models for the source located at $x = 2$ m and $z = 7$ m. Traces number 30 to 60 correspond to the signal recorded by receivers located on the same edge as the source ($x = 2$ m), whereas traces n° 1 to 30 and n° 60 to 90 are recorded by receivers on adjacent edges ($z = 2$ and $z = 12$ m). Traces n° 90

[1] *Erratum to the published version:* The stopping criterion used in L-BFGS-B actually considers the data misfit decrease, and not the norm of the model update (Zhu et al., 1997). The optimisation thus stops when the relative misfit decrease between two iterations becomes smaller than 10^4 times the machine precision.

to 120 correspond to the transmitted signal recorded on the opposite edge ($x = 12$ m). As the initial model is homogeneous (equal to the background model), the initial residuals shown in Fig. 2.7(c) essentially consist in events that are diffracted by the anomalies.

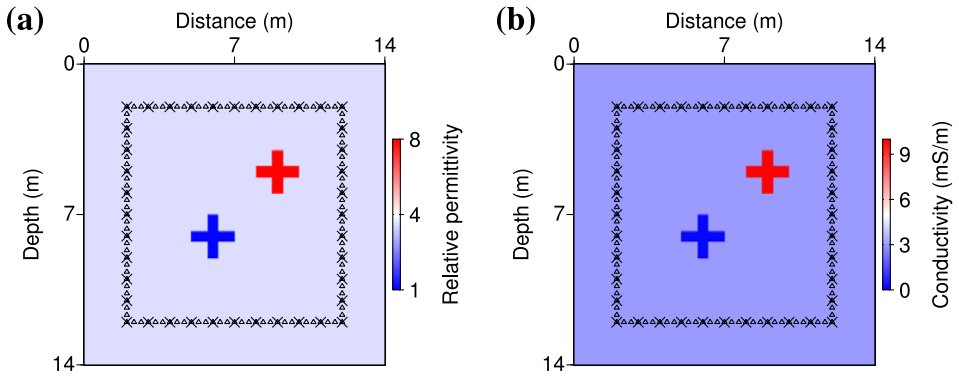


Figure 2.6: *Acquisition setup and true models for permittivity (a) and conductivity (b), after Meles et al. (2011). Black crosses indicate source locations and receiver locations are marked with triangles. Note that we assume the antennas to be perpendicular to the plane of observation (TE mode), whereas Meles et al. (2011) use in-plane antennas (TM mode).*

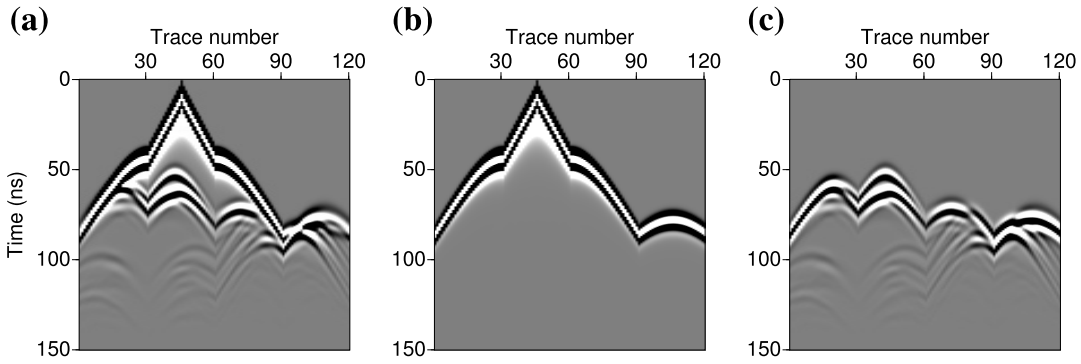


Figure 2.7: *Time-domain shot gathers computed for the cross-shaped benchmark, a) true model, b) initial homogeneous model, c) residuals. Data have been computed in the frequency-domain and convolved with the time-derivative of a Ricker wavelet of central frequency 100 MHz before inverse Fourier transform.*

To perform the inversion, we first compute synthetic *observed* data in the true model of Fig. 2.6 for the seven following frequencies: 50, 60, 70, 80, 100, 150, and 200 MHz, which are consistent with the frequency bandwidth of a 100-MHz antenna. Note that we compute these *observed* data $\mathbf{d}_{obs}$ with the same modelling tool as the one used for computing the calculated data $\mathbf{d}_{cal}$ in the inversion process (inverse crime approach). The irregular frequency sampling is inspired by the strategy of Sirgue and Pratt (2004) who show that the wavenumber coverage increases with frequency (see their fig. 3), so that a fine sampling of high frequencies is not required. We can perform inversion by considering the seven selected frequencies either simultaneously or through a sequential procedure where the initial model for each frequency is the final result of the previous inverted frequency. The sequential strategy based on the low to high frequencies hierarchy has been promoted by Pratt and Worthington (1990) to mitigate

non-linearities such as cycling skipping effects. On the other hand, the strategy of inverting simultaneously all frequencies is more subject to the cycle-skipping problem, depending on the initial model. But if a good initial model is available, it will take benefit from a broadband information. Finally, more elaborated strategies can be used. For instance, we may proceed through a cumulative sequential approach where we keep low-frequency data as we move to high frequencies as suggested by Bunks et al. (1995) for seismics in the time domain, and used by Meles et al. (2011) for GPR data inversion. In the frequency domain, it amounts to invert the following seven groups of cumulative frequencies:

50 MHz,
50 60 MHz,
50 60 70 MHz,
...
50 60 70 80 100 150 200 MHz.

Note that this approach, that we will call the *Bunks' strategy* by analogy with the time domain, implies a consequent computational effort when applied in the frequency-domain. In the following, we will test the three above-mentioned strategies (simultaneous vs. sequential vs. Bunks' strategy) and retain the most robust one for our realistic application.

2.2.2.1 Parameter sensitivity and trade-off

As we are interested in quantifying permittivity and conductivity values, we have to estimate the sensitivity of the data to these parameters. As done by Malinowski et al. (2011) for velocity and attenuation, we can evaluate the impact of both parameters in the data by computing the electric field that is diffracted by anomalies of permittivity and conductivity. Fig. 2.8 shows the real part of such monochromatic scattered fields $\mathbf{u}^{sc}(\mathbf{m}, \delta m_i)$ at a frequency of 100 MHz, computed as the difference between the incident field $\mathbf{u}(\mathbf{m})$ emitted by a source in the homogeneous background $\mathbf{m}$ of Fig. 2.6 and the field emitted by the same source in a perturbed medium $\mathbf{m} + \delta m_i$ where an anomaly δm_i of small amplitude has been added in the centre of the medium. We apply a perturbation amplitude of $\delta p = 5\%$ of the background value, such that $\delta m_i = \delta p \times m_i$, with i the index of the central cell in the finite-difference grid[1].

The scattered wavefield $\mathbf{u}^{sc}(\mathbf{m}, \delta m_i)$ is linked to the partial derivative wavefield $\partial_{m_i}\mathbf{u}$, which in turn can be related to the diffraction matrix by differentiating the forward problem (equation 2.15) with respect to the model parameters, providing at first order

$$\mathbf{u}^{sc}(\mathbf{m}, \delta m_i) \simeq \frac{\partial \mathbf{u}(\mathbf{m})}{\partial m_i} \delta m_i = -\mathcal{A}^{-1} \frac{\partial \mathcal{A}}{\partial m_i} \delta m_i \, \mathbf{u}(\mathbf{m}). \tag{2.21}$$

In equation (2.21), the scattered wavefield can be interpreted as the field emitted by a virtual source $\partial_{m_i}\mathcal{A}\,\delta m_i\,\mathbf{u}$ co-located with the anomaly δm_i and whose signature characterises the data sensitivity to the considered parameter, contained in the diffraction matrix $\partial_{m_i}\mathcal{A}$ (Pratt et al., 1998; Malinowski et al., 2011; Operto et al., 2013). This scattered field is the response of the anomaly to the incident field $\mathbf{u}(\mathbf{m})$ which controls its illumination, depending on the acquisition configuration and on the antenna radiation pattern. Therefore, the shape of the

[1] *Additional note to the published version:* We perturb only the central cell, so the anomaly can be considered as *punctual*. More precisely, it has the size of the grid step $h = 7$ cm, i.e. a relative size compared to the propagated wavelengths of approximately 0.02λ at 50 MHz, 0.05λ at 100 MHz and 0.09λ at 200 MHz.

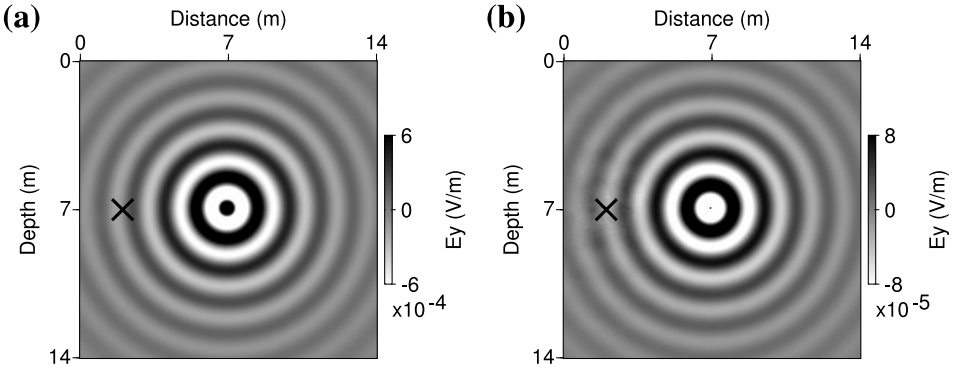


Figure 2.8: *Real part of the monochromatic electric fields diffracted by a permittivity anomaly (a) and a conductivity anomaly (b) at 100 MHz. Anomalies are located at the centre of the medium ($x = 7$ m, $z = 7$ m) and the black cross indicates the source location ($x = 2$ m, $z = 7$ m). Perturbation amplitudes are of 5% of the homogeneous background values ($\delta\varepsilon_r = 0.2$ and $\delta\sigma = 0.15$ mS/m).*

scattered wavefield entirely reflects the response of the anomaly and only depends on the sensitivity kernel $\partial_{m_i}\mathcal{A}$, whereas its amplitude and phase partly come from the incident field, and thus from the GPR antenna.

As shown in Fig. 2.8, the fields diffracted by permittivity and conductivity anomalies in a homogeneous medium are both isotropic but have different amplitudes and phases. In Fig. 2.9, we present the amplitudes and phases of the diffracted fields as a function of the distance to the anomaly for different frequencies. Fig. 2.9(a) shows that the impact of the anomaly of permittivity is larger in amplitude than the one of conductivity in the frequency range we consider for GPR investigation (at least, for equal relative perturbations δp). In the general case, this amplitude ratio depends on the loss-tangent at grid point i, $\tan\delta_i = \sigma_i/(\varepsilon_i\omega)$, that quantifies the energy dissipation in the dielectric lossy background, and on the relative perturbation amplitudes in permittivity δp^ε and conductivity δp^σ. Based on the expressions of the scattered wavefield (equation 2.21) and of the diffraction matrices (equation 2.20), we have[1]

$$\frac{\mathbf{u}^{sc}(\mathbf{m},\delta\sigma_i)}{\mathbf{u}^{sc}(\mathbf{m},\delta\varepsilon_i)} = \frac{\imath\omega\,\delta\sigma_i}{\omega^2\,\delta\varepsilon_i} = \imath\frac{\sigma_i}{\varepsilon_i\omega}\frac{\delta p^\sigma}{\delta p^\varepsilon}. \tag{2.22}$$

In the frequency range of GPR investigations, EM waves encounter rather low-loss media, such that $\tan\delta \ll 1$ (for instance, in the homogeneous background of Fig. 2.6, we have $\tan\delta \simeq 0.14$), so that data are *intrinsically* more sensitive to permittivity than to conductivity. However, it can be compensated by the fact that in natural media, conductivity may present more contrasts than permittivity (conductivity can vary over several order of magnitudes, typically from 10^{-4} to 0.1 S/m, whereas permittivity varies from 1 in air to 81 in water). Consequently, conductivity anomalies can have a significant imprint on the data. Moreover, equation (2.22) and Fig. 2.9(a) show that, in a given medium, the sensitivity of the data to conductivity decreases with frequency relatively to the sensitivity to permittivity. Therefore, high-frequency

[1] *Additional note to the published version:* Eq. (2.22) can be read as

$$\text{scattered amplitude ratio}\left(\frac{\mathbf{u}^{sc}(\mathbf{m},\delta\sigma_i)}{\mathbf{u}^{sc}(\mathbf{m},\delta\varepsilon_i)}\right) = \imath \times \text{loss tangent}\left(\frac{\sigma_i}{\varepsilon_i\omega}\right) \times \text{relative contrasts}\left(\frac{\delta p^\sigma}{\delta p^\varepsilon}\right).$$

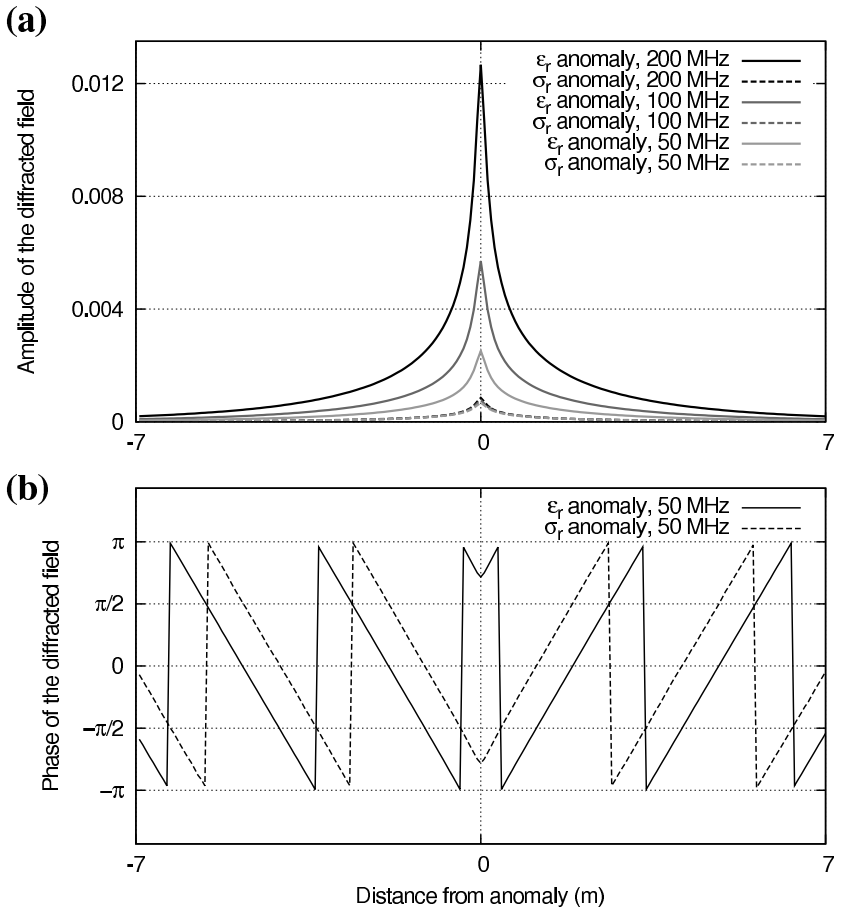


Figure 2.9: *a) Amplitudes of the diffracted fields for three frequencies. b) Phases of the diffracted fields at 50 MHz.*

information about low conductivity contrasts may be hidden below the noise level. Finally, another important feature in Fig. 2.9(b) is that diffractions of permittivity and conductivity present a 90° phase shift, as expected from the presence of the imaginary unit in equation (2.22).

From the differences in frequency-dependency and phase between the diffraction patterns of permittivity and conductivity, we can expect that both parameters could be reconstructed from recorded GPR data, provided that various angles of illumination and a wide frequency bandwidth are available to distinguish their respective signatures (Pratt et al., 1998; Operto et al., 2013). Conversely, a partial illumination and a reduced frequency bandwidth will induce a trade-off between both parameters, meaning that a wave scattered in one direction at one frequency by a permittivity anomaly can be equivalently explained by a conductivity anomaly shifted in space and of stronger amplitude.

To draw a parallel with the reconstruction of seismic velocity and attenuation, let us remark firstly that the imprint of conductivity in GPR data is generally stronger than the effect of the quality factor Q_P in seismic data (see the perturbations applied by Malinowski et al., 2011, and the resulting imprint relatively to velocity). In addition, the attenuation of electromagnetic waves do not suffer from the ambiguity discussed by Mulder and Hak (2009) and Hak and Mulder (2011): Even in low-loss media, the electromagnetic quality factor $Q_{EM} \simeq 1/\tan\delta$

is frequency-dependent. Seismic velocity and attenuation can only be distinguished by their phases, whereas the frequency-dependency of the relative impact of permittivity and conductivity in the data is an additional information that may help to mitigate the trade-off between parameters.

We shall mention that the above discussion on the diffraction patterns of parameter anomalies is particularly adapted to on-ground GPR data, which mostly contain reflections and diffractions. Crosshole GPR data present other sensitivities to permittivity and conductivity. First because crosshole measurements are generally performed in TM mode, for which permittivity and conductivity act differently on the impedance matrix, so that their diffraction patterns are dipolar (independently from the dipolar radiation pattern of finite-line antennas in TM mode). To go further, it is not obvious that this kind of sensitivity analysis based on the diffraction patterns would be consistent when dealing with crosshole data, which mostly contain transmitted signal. In transmission regime, data might be more sensitive to extended anomalies which the waves pass through than to local diffracting points.

The previous remarks about the scattered wavefields have important consequences on the strategy required for multiparameter imaging. In a first approach, reasoning only on the loss tangent in equation (2.22) tends to confirm the common mind that GPR data are mainly sensitive to permittivity, which justifies the use of alternated strategies: fixing the conductivity to an expected value and inverting for the permittivity in a first step, and then proceeding to the conductivity reconstruction with a fixed updated permittivity (Ernst et al., 2007). However, this strategy can fail to retrieve satisfactory models because, as shown in equation (2.22), strong conductivity contrasts may have a significant imprint on the data (both on amplitudes and phases), which may hinder the reconstruction of an accurate permittivity model during the first step. In the second step, the conductivity reconstruction may then suffer from artefacts because it is very sensitive to the kinematic accuracy of the background medium (as is the reconstruction of attenuation in visco-acoustics, see e.g. Kamei and Pratt, 2008, 2013). Because of the trade-offs, errors contained in the previously inverted permittivity, even if small, will systematically map into conductivity artefacts (for an interesting discussion on the effect of the trade-off, see Kamei and Pratt, 2013, §3.3). As a consequence, simultaneous inversion of permittivity and conductivity generally yields better results than alternated or cascaded algorithms (Meles et al., 2010).

More elaborated schemes can consist in a first FWI step for the reconstruction of the permittivity while keeping fixed the conductivity, and in a second FWI step to invert simultaneously for permittivity and conductivity. It is a strategy used for the inversion of velocity v_P and attenuation Q_P in seismics (Kamei and Pratt, 2008; Malinowski et al., 2011; Prieux et al., 2013; Kamei and Pratt, 2013; Operto et al., 2013), where the parameter v_P is first reconstructed with an approximate estimation of a constant Q_P before the simultaneous inversion of the two parameters. This approach yields more satisfactory results because the first step improves the kinematic model for the second step, but a well-designed multiparameter scheme is still needed for the second step (if the initial model is accurate enough, we do not need the first step).

The design of multiparameter FWI is thus crucial but it faces a major problem when dealing with different parameter units and sensitivities (Meles et al., 2012). Meles et al. (2010) tackle this issue by introducing two descent step lengths in their algorithm: one for minimising the misfit function in the gradient direction with respect to permittivity, and another in the direction of conductivity. This approach performs better than the one proposed by Ernst et al. (2007) because permittivity and conductivity models are updated simultaneously at each

iteration (instead of every n iterations). But it amounts to consider the optimisation with respect to permittivity and conductivity as two independent problems, and to neglect the possible trade-off between the two types of parameters. In this study, we propose to investigate a fully multiparameter strategy. The L-BFGS approximation of the inverse Hessian operator in the quasi-Newton scheme (2.18) should take the trade-offs between parameters into account.

2.2.2.2 Parameter scaling

To gather permittivity and conductivity in the model vector $\mathbf{m}$, we have to consider adimensional quantities. It requires to define scales of measurement for the permittivity and the conductivity. The relative permittivity $\varepsilon_r = \varepsilon/\varepsilon_o$ is commonly defined according to the vacuum permittivity $\varepsilon_o \simeq 8.85 \times 10^{-12}$ F/m. In addition, we introduce a relative conductivity $\sigma_r = \sigma/\sigma_o$. By convention, we define the reference conductivity σ_o as the conductivity of a reference medium in which the loss tangent $\tan\delta_o = \sigma_o/(\varepsilon_o\omega_o)$ equals one at a reference frequency $f_o = 100$ MHz. The reference frequency $f_o = \omega_o/(2\pi)$ corresponds to the central frequency of the band we will use in our numerical tests. We shall underline that this arbitrary definition is only a convention used for the optimisation. In particular, the reference medium of permittivity ε_o and conductivity σ_o has nothing to do with the physical medium we want to investigate. This convention leads to the reference value $\sigma_o = \varepsilon_o\omega_o \simeq 5.6$ mS/m. The relative permittivity ε_r and the relative conductivity σ_r constitute the two classes of parameters we will use for the reconstruction.

Note that we could question the choice of ε_r and σ_r as the model parameters m_i to be optimised. Although we are interested in knowing the permittivity and the conductivity in the subsurface because they are meaningful physical quantities, other variables might be considered in the optimisation procedure. We shall mention that we have investigated various couples of parameters (among others, $\sqrt{\varepsilon_r}$, $1/\sqrt{\varepsilon_r}$, $\ln(1+\sigma_r)$, $\tan\delta$, ...) and we did not found significant advantages for using these non-linear parameters compared to the choice of the relative permittivity ε_r and relative conductivity σ_r. On the contrary, inverting for the loss tangent $\tan\delta = \varepsilon/(\sigma\omega)$ should be avoided because it induces a strong coupling between permittivity and conductivity models. In addition, optimising the parameters ε_r and σ_r enables to easily control the relative weight given to permittivity and conductivity in the inversion process, as we will discuss in the following. Thus, we consider linear combinations of the relative permittivity and of the relative conductivity of the form $(\varepsilon_r, \sigma_r/\beta)$, where β is a dimensionless scaling factor which controls the weight of σ_r vs. ε_r in the optimisation process and may compensate for the arbitrary definition of the reference permittivity and reference conductivity. In the following, we refer to the reconstructed parameter σ_r/β as the *scaled conductivity*. We can now give explicit expressions for the model and gradient vectors:

$$\mathbf{m} = \begin{pmatrix} \varepsilon_{\mathbf{r}} \\ \sigma_{\mathbf{r}}/\beta \end{pmatrix}, \quad \text{and} \quad \mathbf{G}(\mathbf{m}) = \begin{pmatrix} \mathbf{G}^{\varepsilon_{\mathbf{r}}}(\mathbf{m}) \\ \beta\mathbf{G}^{\sigma_{\mathbf{r}}}(\mathbf{m}) \end{pmatrix}. \tag{2.23}$$

Here we shall remind that, although the gradient contains two distinct parts related to permittivity and conductivity, the descent direction is computed using the global gradient vector $\mathbf{G}(\mathbf{m}_k)$ and the model update is performed using a unique step length α_k in equation (2.18). The gradient components are deduced from equations (3.18) and (2.20) using the chain rule such that

$$\frac{\partial \mathcal{A}_{ij}}{\partial \varepsilon_{r_i}} = -\varepsilon_o\omega^2\delta_{ij}, \quad \text{and} \quad \frac{\partial \mathcal{A}_{ij}}{\partial \sigma_{r_i}} = -\imath\sigma_o\omega\delta_{ij}. \tag{2.24}$$

Numerical experiments show that the inversion is very sensitive to the respective weights between the relative permittivity and the scaled conductivity in the optimisation process, i.e. to the scales of measurement through the selected scaling factor β. On average, the relative weight of the scaled conductivity vs. permittivity in the gradient is given by

$$\frac{||\beta \mathbf{G}^{\sigma_\mathbf{r}}||}{||\mathbf{G}^{\varepsilon_\mathbf{r}}||} \simeq \beta \frac{||\partial_{\sigma_{r_j}} \mathcal{A}||}{||\partial_{\varepsilon_{r_j}} \mathcal{A}||} \propto \beta \frac{\sigma_o}{\varepsilon_o \omega}, \tag{2.25}$$

which is nothing other than the loss tangent $\tan \delta_o(\omega)$ of the reference medium with permittivity ε_o and conductivity σ_o at frequency ω, scaled by the factor β. According to our definition of the reference conductivity σ_o, the ratio of equation (2.25) will then be about β for a group of frequencies centred around 100 MHz. Note that we recognise an expression for the gradient in equation (2.23) that is similar to the preconditioned gradient proposed by Kamei and Pratt (2013). The difference in our case is that we consider the parameter σ_r/β, whose natural gradient is $\beta \mathbf{G}^{\sigma_\mathbf{r}}$. It is therefore a re-parameterisation, another way to scale the parameter space, and not a preconditioning. However, as in Kamei and Pratt (2013), the factor β will act as an implicit regularisation: We will see that small values of β penalise the conductivity update (at the frequency 100 MHz, it corresponds to values of $\beta < 1$).

In equation (2.18), the descent direction depends not only on the gradient, but also on the L-BFGS approximation of the Hessian $\mathcal{B}$. Although this Hessian approximation is not readily available, understanding the structure of the Hessian through an approximate evaluation should shade light into the optimisation procedure. The true Hessian $\mathcal{H}$ of the misfit function reads

$$\begin{aligned} \mathcal{H} = & \mathcal{R}e \begin{pmatrix} \mathcal{J}^{\varepsilon_\mathbf{r}^\dagger} \mathcal{J}^{\varepsilon_\mathbf{r}} & \beta \mathcal{J}^{\varepsilon_\mathbf{r}^\dagger} \mathcal{J}^{\sigma_\mathbf{r}} \\ \beta \mathcal{J}^{\sigma_\mathbf{r}^\dagger} \mathcal{J}^{\varepsilon_\mathbf{r}} & \beta^2 \mathcal{J}^{\sigma_\mathbf{r}^\dagger} \mathcal{J}^{\sigma_\mathbf{r}} \end{pmatrix} \\ & + \mathcal{R}e \left\{ \sum_{i=1}^{N_D} \Delta d_i \begin{pmatrix} \dfrac{\partial^2 d_{cal_i}}{\partial \varepsilon_{r_j} \partial \varepsilon_{r_l}} & \beta \dfrac{\partial^2 d_{cal_i}}{\partial \varepsilon_{r_j} \partial \sigma_{r_l}} \\ \beta \dfrac{\partial^2 d_{cal_i}}{\partial \sigma_{r_j} \partial \varepsilon_{r_l}} & \beta^2 \dfrac{\partial^2 d_{cal_i}}{\partial \sigma_{r_j} \partial \sigma_{r_l}} \end{pmatrix}^\dagger_{\substack{j \in [\![1, N_M]\!] \\ l \in [\![1, N_M]\!]}} \right\}, \end{aligned} \tag{2.26}$$

In this expression, for each data value i and each grid point j, the elements of the Jacobian matrices $\mathcal{J}$ are defined by

$$\mathcal{J}_{ij}^{\varepsilon_r} = \frac{\partial d_{cal_i}}{\partial \varepsilon_{r_j}} \quad \text{and} \quad \mathcal{J}_{ij}^{\sigma_r} = \frac{\partial d_{cal_i}}{\partial \sigma_{r_j}}. \tag{2.27}$$

In equation (2.26), the first term corresponds to the linear part of the Hessian (Gauss-Newton). It accounts for geometrical spreading and dimensionalities of the parameters (diagonal terms), for the limited bandwidth effects due to the finite-frequency content of the source and to the partial illumination of the medium through the discrete acquisition setup (band-diagonal terms), and for the trade-offs between parameters (off-diagonal blocks). The second term accounts for double-scattered events with second-order derivatives of the wavefield which are neglected in Gauss-Newton approaches. The inverse Hessian approximated in equation (2.18) should correct for all these effects and act as a deconvolution operator on the gradient (Pratt et al., 1998; Operto et al., 2013).

Figs 2.10(a), (b), and (c) show three Hessian matrices $\mathcal{H}^\beta$ for values of $\beta \in \{0.25, 1, 4\}$. These matrices have been computed using a finite-difference approximation around the final

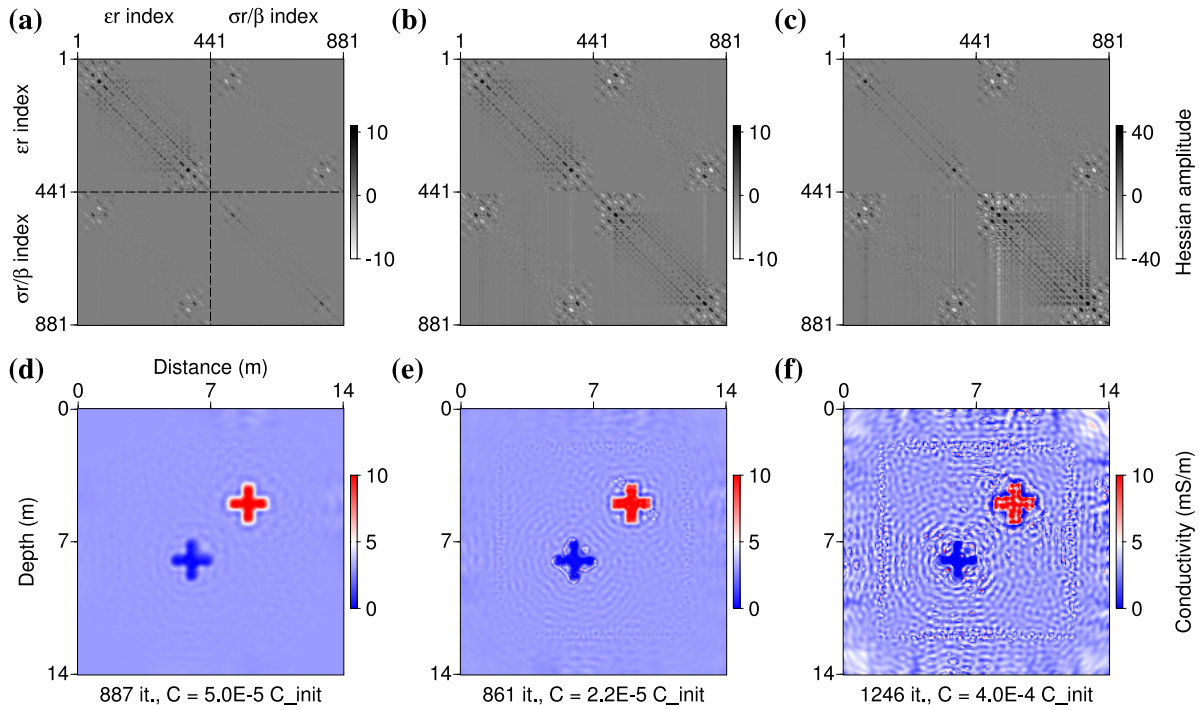


Figure 2.10: *Hessians of the misfit function (a, b, c) and final models of conductivity (d, e, f) for three linear combinations of ε_r and σ_r/β, using a scaling factor $\beta = 0.25$ (a, d), $\beta = 1$ (b, e), and $\beta = 4$ (c, f). The final misfits $\mathscr{C}$ are expressed in fraction of the initial misfit $\mathscr{C}_{init}$.*

reconstructed models $\mathbf{m}^{\beta} = (\varepsilon_{\mathbf{r}}, \sigma_{\mathbf{r}}/\beta)$, which have been obtained inverting the seven frequencies between 50 MHz and 200 MHz simultaneously, starting from the homogeneous background of Fig. 2.6. In these figures, we recognise the expected symmetric structure of the four blocks of the Hessian in equation (2.26). Slight discrepancies in this symmetry are only due to numerical errors in the finite-difference approximation. The correlation between the parameters ε_r and σ_r/β represented in the off-diagonal blocks is not negligible, which justifies their consideration through efficient quasi-Newton methods for solving the multiparameter problem. The amplitude of the trade-off terms is particularly high in the corner of the sub-blocks (i.e. for cells located in the low-illuminated zone outside the acquisition system in Fig. 2.6). This is consistent with the result by Hak and Mulder (2010) that a partial illumination contributes to enhance the ambiguity between the different parameter types.

As expected from equation (2.26), a scaling factor $\beta = 1$ provides four blocks of similar amplitudes (Fig. 2.10b). Alternatively, the value $\beta = 0.25$ penalises the conductivity terms and gives more weight to the sub-blocks associated to permittivity (Fig. 2.10a), whereas the value $\beta = 4$ enhances the blocks related to conductivity (Fig. 2.10c). We present on Figs 2.10(d), (e), and (f) the final conductivity models corresponding to the three values for the scaling factor β. The final images of conductivity are very sensitive to the parameter scaling through the selected factor β. Small values of β provide smooth reconstructions of conductivity (Fig. 2.10d), whereas large values of β enhance the contrasts, but also introduce instabilities (Figs 2.10e and f). Following Kamei and Pratt (2013), we can interpret the artefacts in the conductivity image as unphysical oscillations coming from the trade-off between permittivity and conductivity (monoparameter inversions of conductivity provide images without artefacts when the true

permittivity model is known). In these tests, the reconstruction of the permittivity is less sensitive to the parameter scaling (all three values of β yield nearly identical, satisfactory results that we do not show). Further tests show that very large values of β (> 10) provide smoother images of permittivity (and very high amplitude oscillations in conductivity), whereas very small values of β (< 0.05) may also introduce instabilities in the image of permittivity, whereas the conductivity is not updated at all.

There are two possibilities to mitigate the undesired oscillations in the conductivity image. We can either penalise the relative conductivity with small values of the scaling factor β, as done by Kamei and Pratt (2013), or we can introduce a regularisation term in the misfit function. Actually, both approaches are needed in the multiparameter permittivity-conductivity problem. The use of regularisation may be the first choice in a common optimisation procedure because the high amplitude oscillations in the conductivity model partly comes from the fact that the data are weakly sensitive to conductivity. To constrain the conductivity update, we introduce a Tikhonov regularisation (Tikhonov and Arsenin, 1977) through a model term $\mathscr{C}_M$ in the misfit function:

$$\mathscr{C}(\mathbf{m}) = \mathscr{C}_D(\mathbf{m}) + \lambda \mathscr{C}_M(\mathbf{m}), \tag{2.28}$$

$$\text{with} \quad \mathscr{C}_M(\mathbf{m}) = \frac{1}{2} {\sigma_\mathbf{r}}^T \mathcal{D} \sigma_\mathbf{r}. \tag{2.29}$$

In this new misfit function, the first term $\mathscr{C}_D$ represents the fit to the data according to equation (2.17). In the second term $\mathscr{C}_M$, the matrix $\mathcal{D}$ corresponds to the Laplacian operator, such that the minimisation of the regularisation term forces the conductivity model to be smooth. The hyper-parameter λ is a weighting factor which balances the contribution of the regularisation term with respect to the data term in the misfit function.

The gradient of the new misfit function (2.28) can also be expressed in terms of a data part and of a model part as

$$\mathbf{G}(\mathbf{m}) = \mathbf{G}_D(\mathbf{m}) + \lambda \mathbf{G}_M(\mathbf{m}), \tag{2.30}$$

where the data part $\mathbf{G}_D$ is computed after equations (3.18) and (2.24) and the model part is

$$\mathbf{G}_M = \begin{pmatrix} \mathbf{0} \\ \beta \mathcal{D} \sigma_\mathbf{r} \end{pmatrix}. \tag{2.31}$$

Note that we choose to regularise the relative conductivity σ_r and not the reconstructed parameter σ_r/β. Strictly speaking, both are equivalent because the scaling β would be absorbed in the regularisation weight λ, but regularising σ_r instead of σ_r/β is more convenient for further comparison of solutions obtained with a given regularisation weight λ and different scaling factors β. In this synthetic case, we do not regularise the permittivity which is more constrained by the data. In real data cases, it may be necessary to constrain the permittivity update as well, using a smaller regularisation weight than for conductivity.

Numerical tests involving regularisation show that, in the case of $\beta = 1$, small values for the regularisation weight λ slightly attenuate the very high frequency artefacts in the image of Fig. 2.10(e). However, larger values of λ does not enable to remove entirely the remaining oscillations without degrading the shape of the reconstructed anomalies. In the case of $\beta = 4$, the Tikhonov regularisation cannot both avoid the artefacts and provide a satisfactory reconstruction of the anomalies. The attenuation of the very high amplitude oscillations requires a strong regularisation which prevents the optimisation from finding a satisfying minimisation of

the data misfit. Thus, regularisation alone is not sufficient to design a stable inversion scheme. The parameter scaling through the scaling factor β is crucial both to avoid instabilities and obtain a satisfying resolution in the image of conductivity.

2.2.2.3 Behaviour of the inversion with respect to parameter scaling and frequency sampling

In this section, we try to understand in more details the effect of the scaling parameter β. Once we have selected the optimisation technique, we expect that setting the scaling factor β will depend on the investigated medium, on the initial model, on the acquisition configuration, and on the frequency sampling strategy (because the relative impact of permittivity and conductivity in the data is frequency-dependent). In the following, we investigate the behaviour of the inversion process with respect to the scaling factor and to the frequency sampling strategy in the case of the cross-shaped benchmark with perfect illumination and without regularisation. Although we proceed in the reconstruction of the parameters $[\varepsilon_r, \sigma_r/\beta]$, we shall present results for the quantities $[\varepsilon_r, \sigma]$ as they are those we understand physically.

First we focus on the simultaneous frequency strategy, developing the case of Fig. 2.10. The figure 2.11 shows the evolution of the updated models of permittivity and conductivity, in the cases $\beta = 1$ (Fig. 2.11a) and $\beta = 0.25$ (Fig. 2.11b). We extract updated models at some iterations, corresponding to a given decrease of the misfit function $\mathscr{C}$. In Fig. 2.11(a), we first note that instabilities in the conductivity image appear at early iterations (for $\mathscr{C} = 0.25 \times \mathscr{C}_{init}$ and $\mathscr{C} = 0.1 \times \mathscr{C}_{init}$) and not at the end of the optimisation. As a consequence, they are not due to the fact that the optimisation fits numerical noise and cannot be avoided by stopping the iterative process earlier. It is also the reason why regularisation fails to avoid instabilities when a non-adequate scaling factor is used. It is only when the permittivity is correctly reconstructed, providing a reliable kinematic background, that the image of conductivity converges towards the true one ($\mathscr{C} = 0.01 \times \mathscr{C}_{init}$). Conversely, on Fig. 2.11(b) where more weight is given to permittivity, the permittivity model is reconstructed earlier in the iterations, whereas the reconstruction of conductivity is delayed and, thus, better retrieved when a more reliable kinematic model is available ($\mathscr{C} = 0.01 \times \mathscr{C}_{init}$ and $\mathscr{C} = 0.001 \times \mathscr{C}_{init}$). This numerical test suggests that, even in the frame of a simultaneous reconstruction of permittivity and conductivity, the inversion should be led first by the permittivity update. However, there is a counterpart for this behaviour: As a strong penalisation delays the reconstruction of conductivity, it induces a loss of resolution in the conductivity image for the same misfit decrease. An optimal value for the scaling factor β should both allow to avoid instabilities and to recover an image of conductivity that should be as complete and resolved as possible. The main limitation for achieving this goal is given by the maximal possible decrease of the misfit function (which depends mainly on the signal-over-noise ratio in the real data case).

Up to now, we have seen that the inversion path (Fig. 2.11), and even the final conductivity models (Fig. 2.10), strongly depend on the parameter scaling. This result is quite unexpected and suggests that the L-BFGS approximation of the Hessian correct only partially for parameter dimensionalities. We would like to better appreciate the exact performance of the L-BFGS algorithm. To illustrate the behaviour of the effective descent direction, we downgrade the example of Fig. 2.6 into an optimisation in a two-parameter space. In this very simple problem, permittivity and conductivity in the background and in the blue cross are fixed at their true values and only the values in the red cross are allowed to vary. The anomaly is assumed

to be homogeneous, such that the problem has only two degrees of freedom: ε_r and σ_r/β in the red cross. Fig. 2.12(a) presents a grid analysis of the misfit function for this two-parameter case. The misfit function has been evaluated for ε_r ranging in $[1, 14]$ and σ ranging in $[0, 27]$ mS/m, with the same acquisition setup as presented in Fig. 2.6 and for the seven frequencies simultaneously. We can notice on Fig. 2.12 that the misfit function is convex with a unique minimum. In addition, Fig. 2.12(a) shows the paths followed by the inversion for various values of the scaling factor β. All processes finally reach the global minimum but inversion paths are very sensitive to the parameter scaling as already observed in Fig. 2.11. Intuitively, large scaling factors β tends to orientate the inversion path along the conductivity axis.

Fig. 2.12(b) shows the inversion path in the case of $\beta = 1$, mapped on a scaled view of the misfit function in the parameter space $[\varepsilon_r, \sigma_r/\beta]$. The arrows represent the opposite of

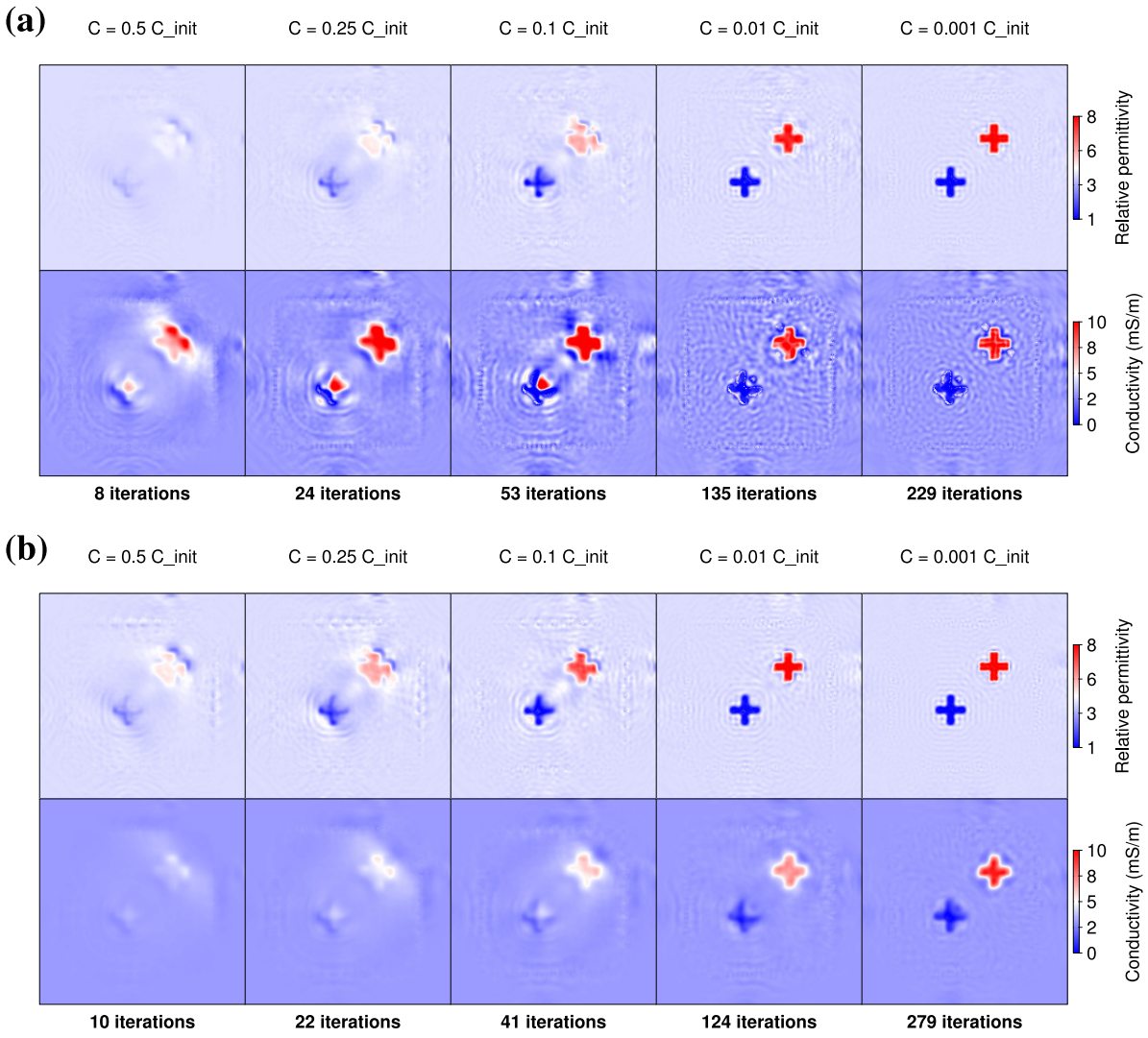


Figure 2.11: *Evolution of permittivity and conductivity models along iterations using a scaling factor $\beta = 1$ (a) and $\beta = 0.25$ (b) when inverting for seven frequencies simultaneously.*

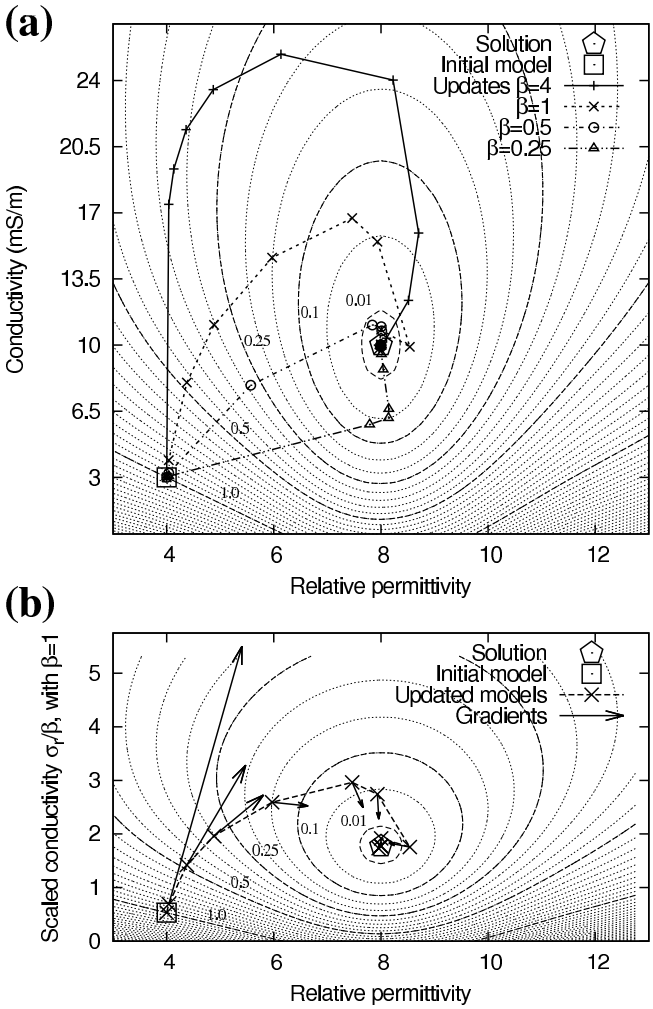


Figure 2.12: *Grid analysis of the misfit function on a simplified two-parameter problem. Dotted contours are spaced every 0.05 while dashed contours indicate particular levels of the misfit function (in fraction of the initial misfit). a) Inversion paths in the physical domain* (ε_r, σ) *using various parameter scalings* β*. b) Inversion path and gradients in the parameter domain* $(\varepsilon_r, \sigma_r/\beta)$ *using a scaling factor* $\beta = 1$*.*

the gradient vectors at each iteration: We can check that they are well orthogonal to the contours of the misfit function, indicating the steepest descent directions. Note that this case corresponds to a nearly circular misfit function in the vicinity of the solution, whereas larger scaling values would elongate the valley in the direction of permittivity, and thus would orientate the gradients in the direction of conductivity. This view of the misfit function in the parameter space, as seen by the optimisation process, helps us to understand the behaviour of the L-BFGS algorithm. For iterations 1, 2, and 3, the descent directions computed by the L-BFGS-B technique (dashed lines between updated models) do not follow the steepest descent directions but are shifted towards the minimum. We can see here the benefit of the quasi-Newton approach which approximates the local curvature. But for later iterations (> 4), the L-BFGS descent directions seem non-optimal compared to the steepest descent directions. This *inertial* effect is partly due to the old curvature information used by the L-BFGS algorithm to

compute the local approximation of the Hessian.

Fig. 2.12 well illustrates that the parameter scaling deforms the parameter space and therefore modifies the descent directions computed by the L-BFGS-B algorithm. Although we could think about better approximations of the Hessian (e.g. preconditioned L-BFGS or truncated Newton methods, see ?Métivier et al., 2013), the optimisation of permittivity and conductivity is likely to remain sensitive to the parameter scaling. As the problem is nonlinear, the local descent directions on Fig. 2.12 cannot point directly to the *global* minimum of the misfit function, even with a good *local* estimation of the curvature (at least before being in the vicinity of the solution, where the quadratic approximation is more valid). In the two-parameter example, the inversion converges towards the unique minimum anyway because the problem is largely over-determined, but we speculate that the parameter scaling is of crucial importance for the high-dimensional case with the additional difficulty of secondary minima.

For this reason, the effect of non-linearity observed at early iterations in Fig. 2.11(a) is strongly dependent on the initial model we have selected, i.e. on how far the initial model is from the validity domain of the quadratic approximation. In particular, if we start from a good kinematic background, the updates of the conductivity model will be improved at early iterations. The initial model can be improved with the low frequency content of the data if we adopt the low to high frequency hierarchy promoted by Pratt and Worthington (1990), inverting the seven frequencies sequentially.

Fig. 2.13 shows the permittivity and conductivity models obtained at the end of each mono-frequency step when inverting the seven frequencies sequentially, each model being the initial model for the next step. Again, we compare the use of a scaling factor equal to $\beta = 1$ (Fig. 2.13a) and $\beta = 0.25$ (Fig. 2.13b). As already observed when inverting the frequencies simultaneously, a low scaling factor $\beta = 0.25$ provides smoother results than the value $\beta = 1$, and the permittivity image is less sensitive to the parameter scaling. Despite the frequency hierarchy, the reconstruction of the conductivity model is not very satisfactory if a scaling factor $\beta = 1$ is used. In particular, the shape of the blue cross is degraded in Fig. 2.13(a), compared to the use of simultaneous frequencies (Fig. 2.10e). We may invoke two reasons to explain this. Firstly, the consideration of simultaneous frequencies might help to better constrain the reconstruction of conductivity, as inferred from the frequency-dependence of its diffraction pattern (Fig. 2.9). Secondly, in the sequential approach, the final reconstructed model results from the inversion of the highest frequency. Even if lower frequencies have been previously inverted, the final model thus contains a stronger finite-frequency imprint than using simultaneous frequencies.

The advantages of the frequency hierarchy and of a large frequency bandwidth can be combined using Bunks' strategy. Further numerical tests involving Bunks' strategy yields similar results as using the simultaneous strategy. Reconstructions are not significantly improved because this simple benchmark do not need a hierarchical approach. Since we choose a lower frequency bandwidth, we do not suffer from the cycle-skipping effect observed by Meles et al. (2011).

As a partial conclusion, we have shown on this perfect illumination case that an *ad hoc* scaling is needed between the relative permittivity and the relative conductivity in our quasi-Newton optimisation scheme. This scaling should play a significant role for surface-to-surface acquisition because partial illumination tends to increase the ill-posedness of the inverse problem (Meles et al., 2012) as well as the trade-off between parameters (Hak and Mulder, 2010).

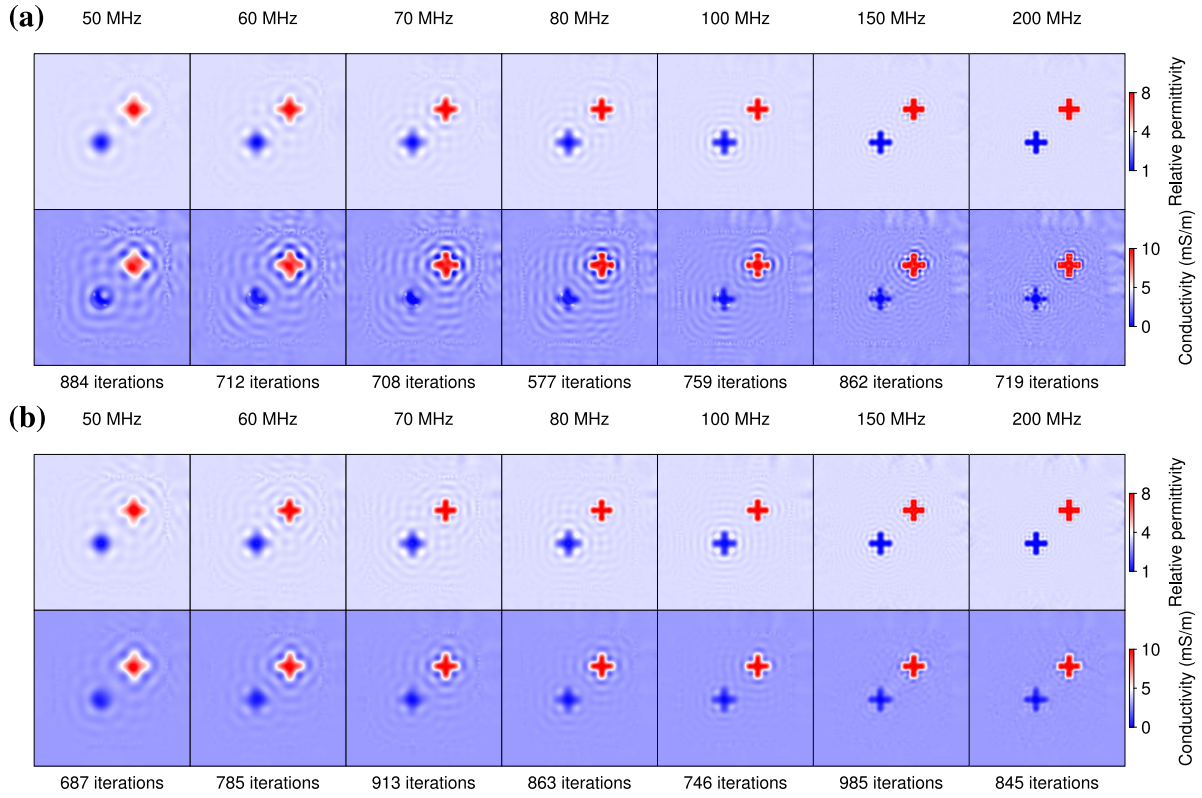


Figure 2.13: *Evolution of the final reconstructed models for each step of the sequential inversion of the seven frequencies using scaling values* $\beta = 1$ *(a) and* $\beta = 0.25$ *(b).*

Up to now, we have performed a quality control of the solution by comparison with the true model in an *inverse crime* way. In the following, we will propose a practical strategy for selecting a reasonable value for the scaling factor using an objective criterion.

2.2.3 A realistic synthetic test

We now introduce a more realistic benchmark for the imaging of complex subsurface structures from multi-offset on-ground GPR data. We first investigate the inversion of noise-free data with respect to the parameter scaling in order to establish a robust criterion for selecting the scaling factor β. In a second time, we add noise to the data to give an insight into the feasibility of the proposed workflow on real data.

2.2.3.1 Benchmark design

A realistic configuration has been designed for a 2D distribution of permittivity (Fig. 2.14a) and conductivity (Fig. 2.14b) according to common-offset GPR profiles, a few common mid-point surveys, and electrical tomography measurements acquired at a test site located on recent fluvial deposits near Grenoble (France). We restrict the zone of interest to a 5-meter-deep and 10-meter-long section of the subsurface. The permittivity and conductivity values are in principle consistent with a silty soil (first layers) including lenses of clay (top right) above a 2-

to 3-meter-thick layer of relatively dry sands. The interface at around 3.8 m depth represents the water table. The 50-cm-thick zone above the zero z-level has values ($\varepsilon_r = 1$, $\sigma = 0$ S/m) of the air. This benchmark displays realistic but challenging sharp and large variations. In the subsurface, permittivity values range from 4 (main part in the middle) to 32 (layers of clay in the top right and at the bottom of the medium), and conductivity values range from 0.1 mS/m (1st layer) to 20 mS/m (bottom). Maximal permittivity contrasts are of 1:10 at the air-ground interface, and of 4:22 in the subsurface (bottom of the main layer at $z \simeq 3$ m). Note the strongly attenuating layer at a depth of $z = 3.5$ m with a conductivity about $\sigma = 10$ mS/m, which may mask the underlying structures.

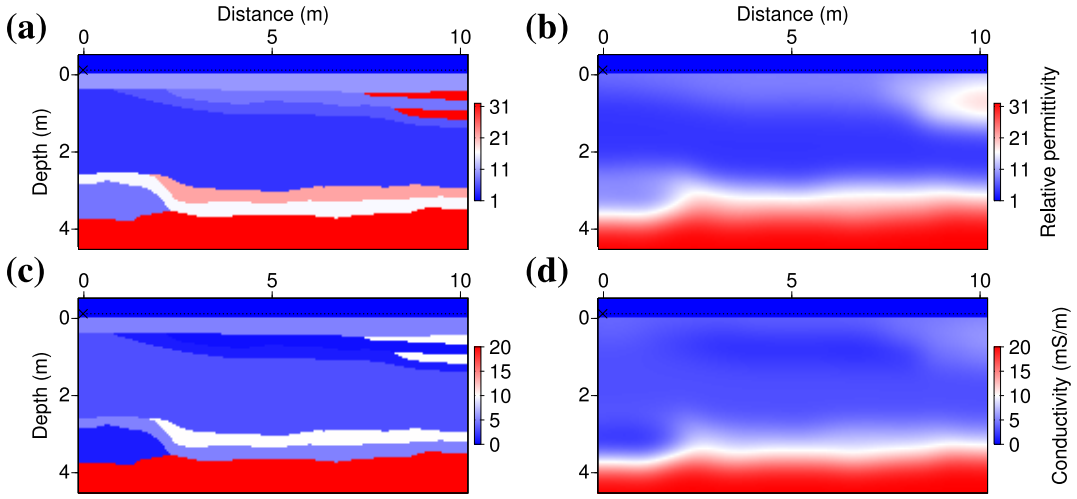


Figure 2.14: *Realistic subsurface benchmark for permittivity (a, b) and conductivity (c, d). (a, c) True models. (b, d) Initial models. Sources and receivers are located 10 cm above the air-ground interface.*

The medium is discretised on a 101×207 grid, with a grid step $h = 5$ cm. This meshing results in 20907 grid points in the finite-difference modelling, and in 18837 unknown values of permittivity and conductivity in the subsurface (values are kept fixed in the air, both to constrain the inversion and to avoid singularities at source and receiver locations). The acquisition setup consists in 41 source locations spaced every 0.25 m and in 101 receiver positions located every 0.1 m at a negative z-level of -0.1 m, i.e. two grid points above the air-ground interface. This setup is consistent with a multi-offset experiment performed on the test site within a day. Initial models for permittivity (Fig. 2.14c) and conductivity (Fig. 2.14d) have been obtained by applying a gaussian smoothing to the true models with a correlation length $\tau = 50$ cm. Only the main trends are depicted in the initial model and all details are erased, in particular the thin lenses in the top right of the medium but also the alternation of high and low values at depth ($z \simeq 3.5$ m). Fig. 2.15 shows the time-domain data computed in the true and initial models, for the first source of the acquisition array (at $x = 0$ m). In Fig. 2.15(a), three major reflections at $t \simeq 20$, 50, and 75 ns can be associated with the interfaces at $z = 0.3$, 2.5, and 3.8 m, while the diffractions at large offsets correspond to the thin lenses of clay. In Fig. 2.15(b), the initial model provides direct arrivals that are kinematically compatible with the observed data but the lack of contrasts does not reproduce the main reflected events and the diffractions are missing. Most of the data are not explained by the initial model and remain in the residuals (Fig. 2.15c).

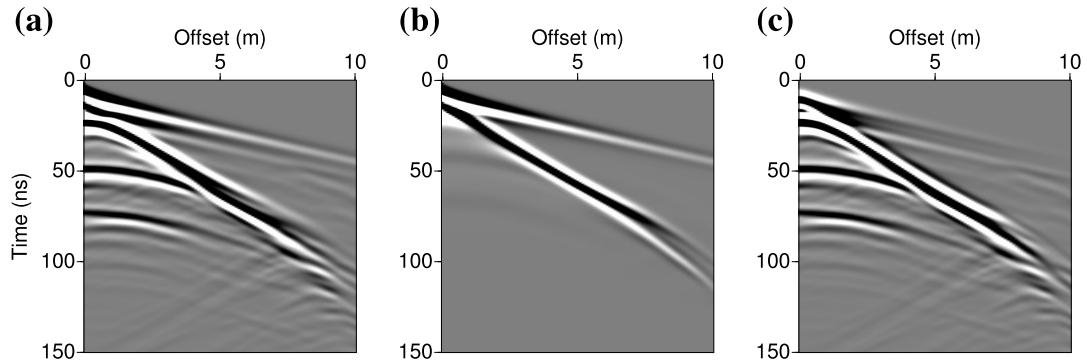


Figure 2.15: *Time-domain shot gathers computed for the subsurface benchmark (Fig. 2.14) for the source located at $x = 0$ m, considering a) the true model, b) the initial model. c) Initial residuals. Data has been computed in the frequency-domain and convolved with the time-derivative of a Ricker wavelet of central frequency 100 MHz before inverse Fourier transform.*

2.2.3.2 Inversion of noise-free data

The frequency sampling is crucial for the imaging of such complex media. Therefore we select a dense frequency sampling with ten frequencies to be inverted: 50, 60, 70, 80, 90, 100, 125, 150, 175, and 200 MHz. Fig. 2.16 shows the final relative misfits obtained by inverting the ten frequencies simultaneously or adopting the so-called Bunks' strategy which results here in ten groups of cumulative frequencies. These two strategies were quite equivalent in the cross-shaped benchmark with perfect illumination. For the imaging of more complex media with surface-to-surface illumination, Bunks' strategy appears as the most efficient approach in terms of misfit decrease and provides more accurate models, because it presents the advantage both to proceed hierarchically and to end up with the inversion of the full frequency band. Again, the sequential strategy yields final reconstructed models that are less satisfactory, confirming the need for keeping the low frequency content in the hierarchical process. Therefore, we will use Bunks' strategy in the following tests.

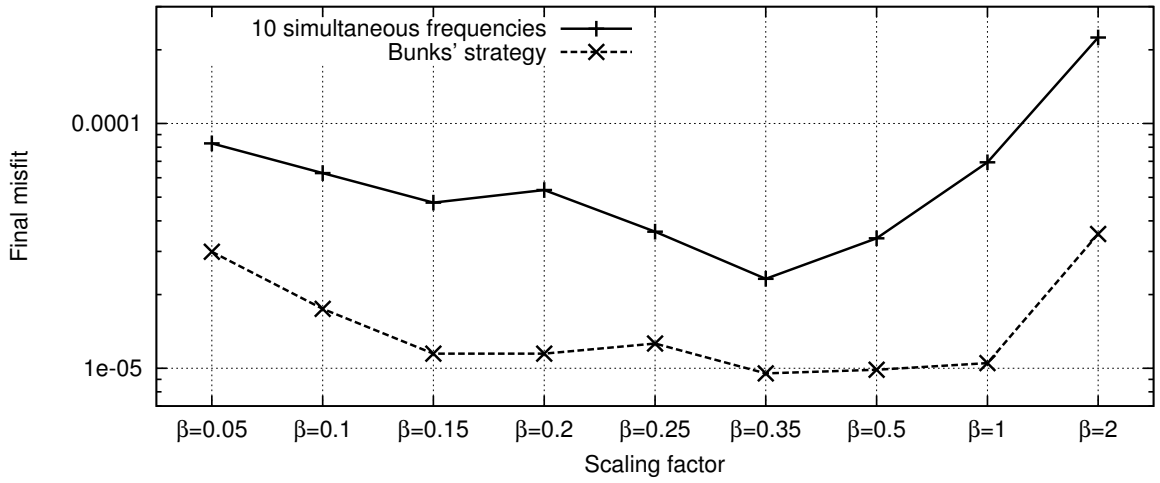


Figure 2.16: *Final data misfit with respect to parameter scaling and frequency sampling strategy.*

In Fig. 2.16, it must be noted that the misfit values reached using the Bunks' strategy are quasi-independent from the parameter scaling in the range $\beta \in [0.15, 1]$ (we do not apply any regularisation here). However, the reconstructed models are very different and remain strongly sensitive to the scaling factor. Fig. 2.17 shows the permittivity and conductivity models obtained with scaling factor values of $\beta = 0.15$ and $\beta = 1$, which both have a misfit very close to 10^{-5} (in fraction of the initial misfit). As in the previous section, the value $\beta = 0.15$ provides a smooth conductivity model (Fig. 2.17c), whereas the value $\beta = 1$ (Fig. 2.17d) introduces instabilities and largely over-estimates the conductivity variations.

The problem we face here is that we cannot discriminate between the two solutions of

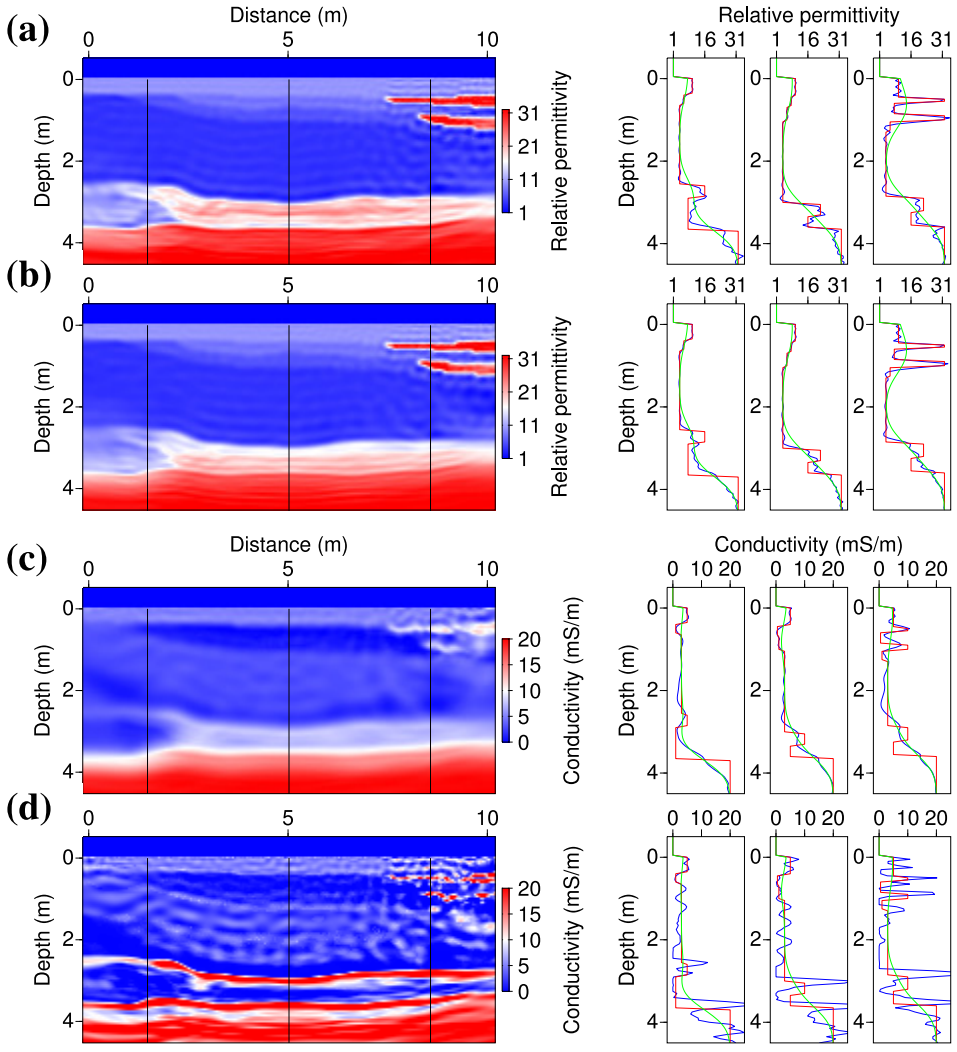


Figure 2.17: *On the left: permittivity (a, b) and conductivity models (c, d) obtained by the inversion of noise-free data using scaling factors $\beta = 0.15$ (a, c) and $\beta = 1$ (b, d). On the right: vertical logs extracted along the black lines indicated on the 2D sections. Red curves denote the true model, blue curves the inverted model, and green curves the initial model. The optimisation required about 350 iterations per frequency group (3448 and 3767 total iterations using $\beta = 0.15$ and $\beta = 1$, respectively).*

Fig. 2.17 with a criterion based only on the data misfit. This could suggest that the variations between the final models obtained with different β values of equivalent misfits all belong to the kernel (or null space) of the misfit function. Then we should conclude that we cannot recover a more precise information about the conductivity from the inverted data.

To have a more detailed insight into this problem, we can try to identify which spectral components differ between the solutions obtained using different scaling factors. To do so, we apply a k_z-transform to the conductivity logs of Figs 2.17(c) and (d). The corresponding amplitude spectra are shown in Fig. 2.18. On these spectra, we observe that the variations in the reconstruction are not restricted to the highest wavenumbers: Low wavenumbers are affected by the scaling as well. This result is quite unexpected because the small eigenvalues of the Hessian, related to the less constrained parameters, are generally related to small scale structures (Hansen, 2010, p.62). Our understanding is that discrepancies in the low wavenumbers should induce a degradation of the misfit but these discrepancies are compensated by high wavenumber structures. If so, removing the high-wavenumber structures should cancel this compensation and lead to a degradation of the data misfit that would enable to distinguish the best models between the different solutions.

Fig. 2.19 shows low-pass filtered versions of the conductivity models of Fig. 2.17, where

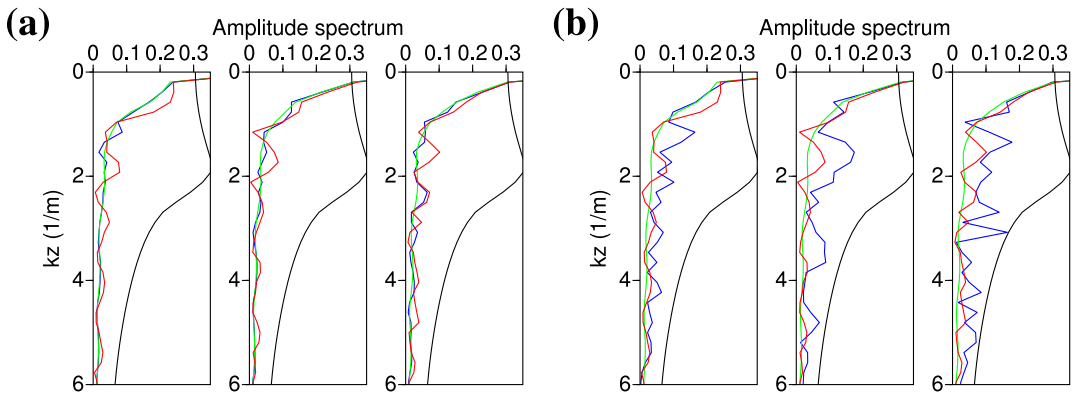


Figure 2.18: *k_z-domain spectra of the vertical conductivity logs of Fig. 2.17 using a scaling factor β = 0.15 (a) and β = 1 (b). The black curves represent the low-pass filter applied both in x- and z-direction to the conductivity models in Fig. 2.19. Red curves denote the true model, blue curves the inverted model, and green curves the initial model.*

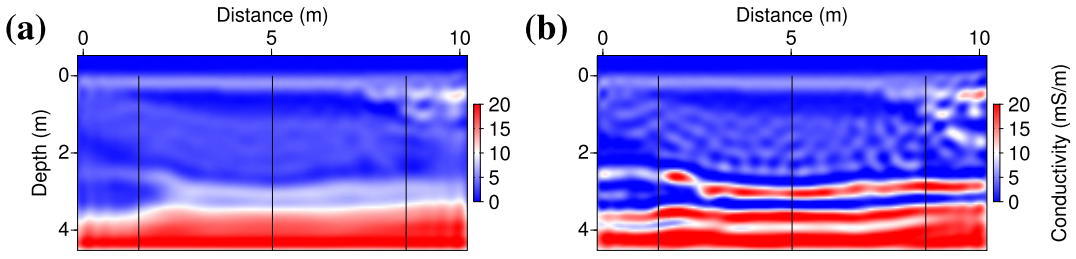


Figure 2.19: *Filtered conductivity models corresponding to β = 0.15 (a) and β = 1 (b). Synthetic data computed in these filtered models yield data misfits $\mathscr{C} = 0.0017 \times \mathscr{C}_{init}$ and $\mathscr{C} = 0.0045 \times \mathscr{C}_{init}$, respectively.*

high wavenumbers $k_x, k_z \geq 2$ are filtered out (both in the x- and z-directions). The threshold $k_{x_{max}}, k_{z_{max}} = 2$ roughly corresponds to the wavelengths propagating in the central part of the medium (where $\varepsilon_r = 4$) at 50 MHz, so it is rather a lower bound of the covered wavenumbers and it induces a quite drastic filtering. As expected from the differences in low-wavenumber contents in Fig. 2.18, the filtered models of Fig. 2.19 are still quite different but, since high wavenumbers have been filtered out, we can now discriminate between them: Computing the corresponding data misfit for these models, we find that the filtered conductivity model obtained with a scaling factor $\beta = 0.15$ (Fig. 2.19a) better explains the data than the one obtained with a value $\beta = 1$ (Fig. 2.19b) by a factor of $\simeq 2.5$.

Consequently, we can expect that introducing a Tikhonov regularisation in the misfit function will enable to identify conductivity solutions, by preventing the creation of the high-wavenumber structures that compensate for erroneous low-wavenumber reconstruction. As in the cross-shaped experiment, we regularise only the conductivity update following equations (2.29) and (2.31). Fig. 2.20 shows the data misfit decrease obtained with the Bunks' strategy when a Tikhonov regularisation is introduced, as a function of the scaling factor β and for different regularisation weights λ. As expected, the use of regularisation makes the final data misfit more sensitive to the parameter scaling. Since it prevents the optimisation to fit the data with high wavenumber structures, we can now discriminate between smooth reconstructed structures that well explain the data and those that do not. Varying the regularisation weight λ, we can see on Fig. 2.20 that the more resolution we allow (with small λ values), the wider is the range of scaling factors β that well explain the data, and the more variability we get in the final conductivity models. Conversely, the more smoothness we impose (with large λ values), the less information we recover. For very large regularisation weights $\lambda > 1$, we recover barely more than the initial model.

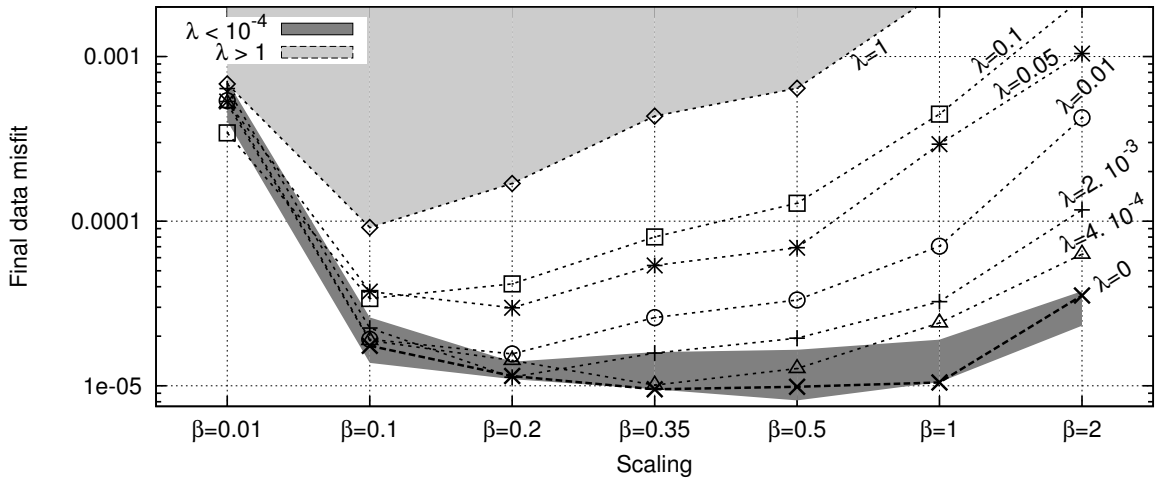


Figure 2.20: *Final data misfit with respect to parameter scaling when a Tikhonov regularisation is applied.*

From Fig. 2.20, a reasonable criterion for selecting an adequate range of values for the scaling factor β and for the regularisation weight λ is to seek for λ values that provide a satisfactory data fit on a small range of β values, i.e. regularisation weights for which it exists a clear minimum of the data misfit with respect to the scaling factors. For instance, the regularisation

weight $\lambda = 0.05$ seems too large as it significantly degrades the data fit. Conversely, the weight $\lambda = 4\,.\,10^{-4}$ is probably too small as it yields good data fits on a wide range of scaling values $\beta \in [0.2, 0.5]$, which may provide dubious models. A reasonable range of values would therefore be $\lambda \in [0.002, 0.01]$ for the regularisation weight and $\beta \in [0.1, 0.35]$ for the scaling factor.

Fig. 2.21 shows the model obtained with a scaling factor $\beta = 0.2$ and a regularisation weight $\lambda = 0.002$. This solution is quite satisfactory when compared with the true model, suggesting that the proposed workflow for selecting the hyper-parameters β and λ is pertinent. In particular, it shows that we can rely on the data misfit (without an arbitrary model criterion) for selecting a reasonable range for the parameter scaling β, in relation with an adequate regularisation level λ for which it exists a clear minimum for the data misfit with respect to parameter scaling. We must underline the lower resolution of the conductivity reconstruction, resulting from the applied regularisation ($\lambda > 0$) and penalisation ($\beta < 1$). However, this solution provides a good compromise in the reconstruction of the high-permittivity, high-conductivity layer at $z = 3$ m, in spite of a slight shift in the conductivity image. The conductivity values in the thin lenses are comparable with those of the true model.

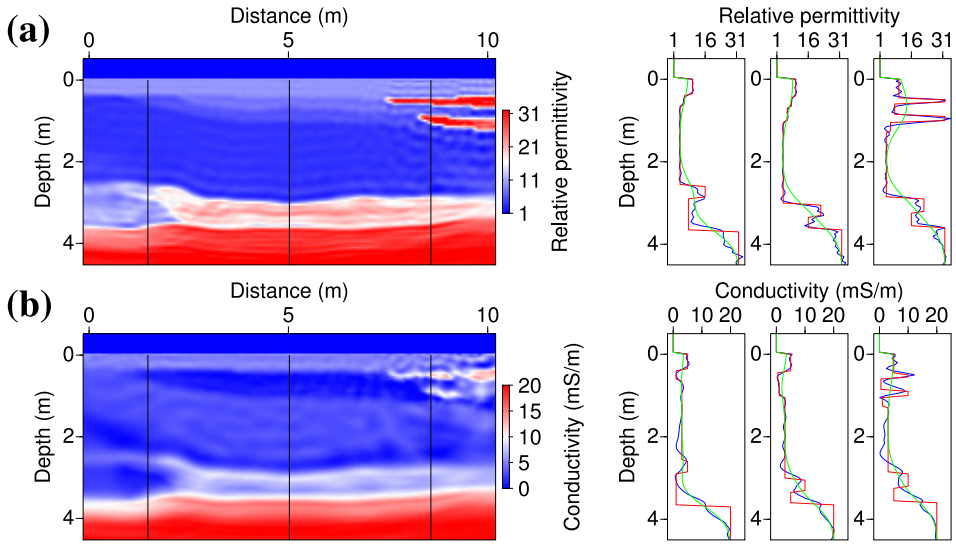


Figure 2.21: *On the left: permittivity (a) and conductivity models (b) obtained by the inversion of noise-free data using a scaling factor* $\beta = 0.2$ *and a regularisation weight* $\lambda = 0.002$. *On the right: vertical logs extracted along the black lines indicated on the 2D sections. Red curves denote the true model, blue curves the inverted model, and green curves the initial model. The optimisation required about 330 iterations per frequency group (3379 total iterations).*

As a partial conclusion, we have shown that the parameter scaling is even more crucial in this realistic case with surface-to-surface illumination than in the previous case with perfect illumination. Due to partial illumination, the inversion is less constrained and various scaling factors β can provide equivalent misfit decreases but very different solutions. Regularisation is necessary to mitigate the ambiguity. By preventing the optimisation to create high-wavenumber structures that artificially explain the data, regularisation makes the final data misfit more sensitive to the scaling factor β. It is thus possible to determine a reasonable range of values both for the regularisation weight λ and for the scaling factor β.

Fig. 2.22 summarises the successive tests we performed in this section. Note that this

diagram does not state for the final workflow that should be applied for multiparameter imaging, but rather as the reasoning flow that leads us to our multiparameter strategy. As indicated in the diagram, if performing regularised FWI still leads to different models of equivalent data misfits, then we can not discriminate between the different solutions based only on the data misfit, and we have to invoke *a priori* information to drive the inversion process towards a unique solution (see e.g. Asnaashari et al., 2013). Alternatively, we could have observed that only high wavenumbers differ between the models reconstructed with different scaling factors. Regularisation would then have avoided the creation of the high-wavenumber artefacts and would probably have yielded similar solutions for the different scaling factors.

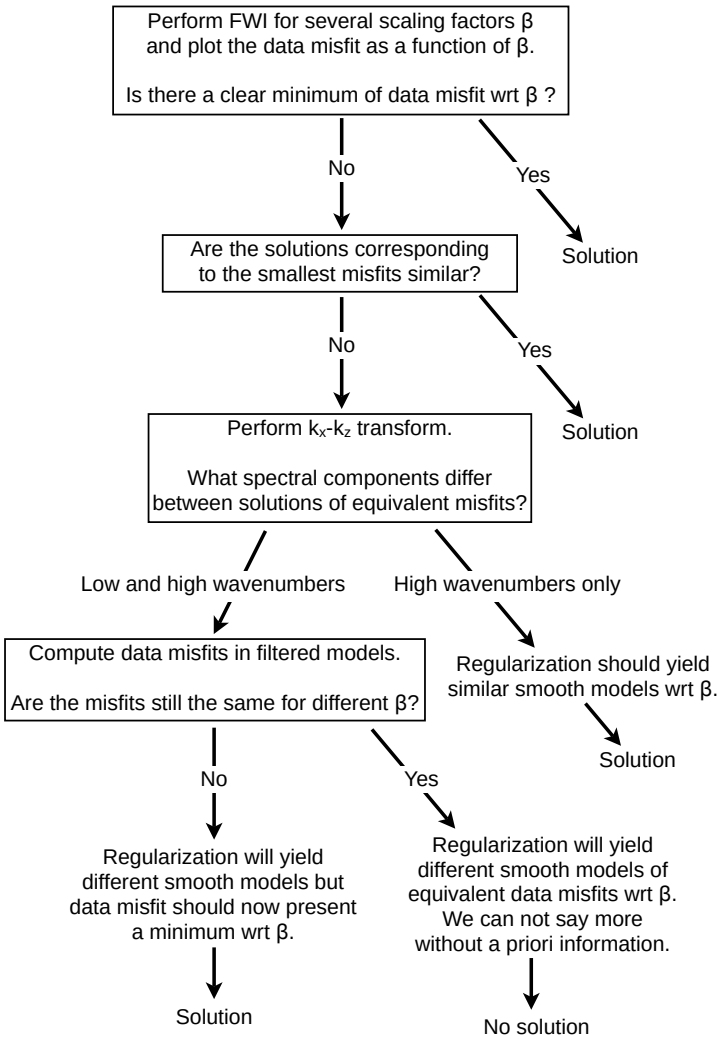


Figure 2.22: *Flow diagram of the successive tests performed in Section 4 and of the related conclusions.*

Finally, the retained workflow for multiparameter imaging would be:

1. Perform FWI with different parameter scalings β and regularisation weights λ.
2. Plot the data misfit as a function of the scaling factor β, for each regularisation weight λ.

3. Identify the regularisation levels that exhibits a clear minimum of the data misfit with respect to parameter scaling. We shall choose the optimal (λ, β) combinations as the smallest λ values for which we can find such a minimum, and β values corresponding to this minimum.

We shall now see whether this workflow can be applied to noisy data, when noise may mask information about conductivity.

2.2.3.3 Inversion of noisy data

In order to tackle a more realistic example, white noise is added to the synthetic frequency data with a signal over noise ratio (SNR) of 25 dB. This noise level is consistent with the noise observed in real GPR data. Fig. 2.23 shows the impact of the noise on a time-domain shot gather. In the noisy data, the main reflected events are still visible but diffractions at large offsets are below the noise level.

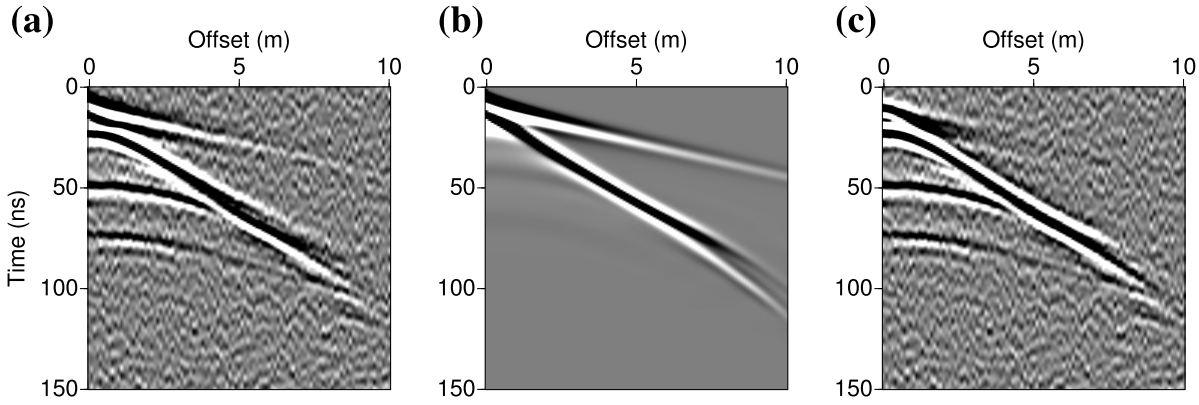


Figure 2.23: *Noisy time-domain shot gathers for the subsurface benchmark (source at $x = 0$ m). a) Observed data, b) data computed in the initial model. c) Initial residuals. Data has been computed in the frequency-domain, then we applied a white noise of SNR = 25 dB in the frequency-domain before the convolution with the time-derivative of a Ricker wavelet of central frequency 100 MHz and inverse Fourier transform.*

In Fig. 2.24, the thick dashed line shows the final data misfit decrease obtained with different values for the scaling factor, without regularisation. The presence of noise in the data prevents to decrease the misfit function below a threshold of about 0.0365 (in fraction of the initial misfit). Again, we reach equivalent data misfits for very different scaling values $\beta \in [0.2, 1]$. Regularisation is needed to constrain the conductivity model and discriminate between the different solutions. The other curves of Fig. 2.24 show the final data misfit reached using the same range of scaling factors β and various regularisation weights λ. It presents the same features as in the case of noise-free data (Fig. 2.20), enabling to determine a range of reasonable values both for the scaling factor β and for the regularisation weight λ on the single criterion of data misfit. Here, the smallest regularisation weights that provide a satisfactory data fit on a small range of scaling factors $\beta \in [0.1, 0.2]$ are $\lambda = 5$ to 10.

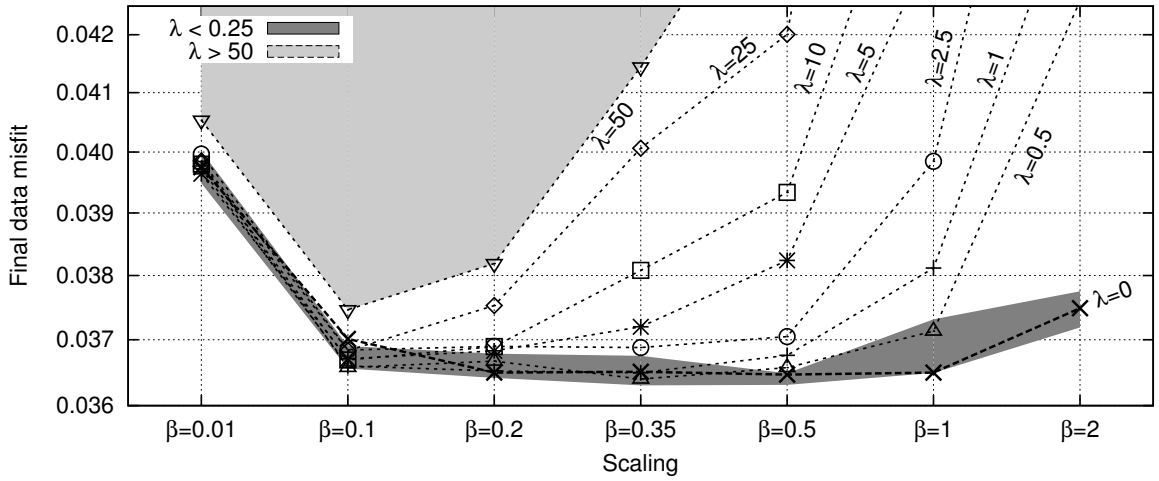


Figure 2.24: *Final data misfit with respect to parameter scaling when a Tikhonov regularisation is introduced in the inversion of noisy data* (SNR = 25 *dB*).

Fig. 2.25 shows the inversion results obtained with a scaling factor $\beta = 0.2$ and a regularisation weight $\lambda = 5$. In the permittivity image, thin superficial layers are still reconstructed quite accurately although the second lens is slightly shifted downwards. Resolution dramatically decreases with depth and the high-permittivity layer does not clearly appear. The strong regularisation of conductivity only provides the main trend of the conductivity structures. The

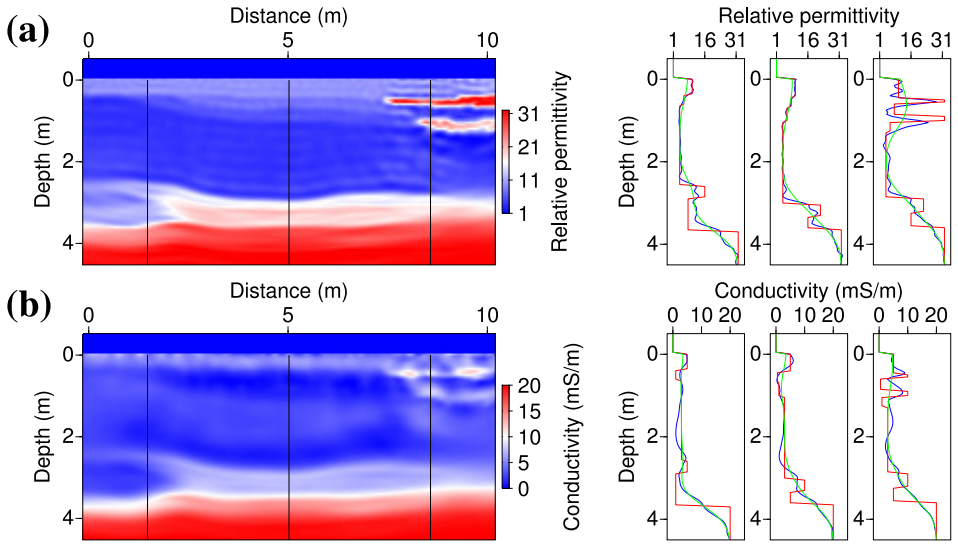


Figure 2.25: *On the left: permittivity (a) and conductivity models (b) obtained by the inversion of noisy data* (SNR = 25 *dB) using a scaling factor* $\beta = 0.2$ *and a regularisation weight* $\lambda = 5$. *On the right: vertical logs extracted along the black lines indicated on the 2D sections. Red curves denote the true model, blue curves the inverted model, and green curves the initial model. The optimisation required about 25 iterations per frequency group (241 total iterations).*

lenses of clay can be distinguished while a blurred image of the alternation of conductivity at depth is obtained. This result may seem disappointing but it should be underlined that the low resolution we obtain is the consequence of the applied regularisation and penalisation, which are necessary to not over-interpret the data.

Finally, Fig. 2.26 compares the time-domain data computed in the inverted model of Fig. 2.25 with the observed noisy data (Fig. 2.26a) and with the observed noise-free data (Fig. 2.26b). For a better visualisation of the signal at late arrival times and large offsets, we apply a time-varying gain and a trace-by-trace normalisation (for each offset, the reference amplitude is the maximum of the observed trace). Every tenth trace is shown. It can be seen that noise is not fitted in the time-domain, although we did not regularise the permittivity update. It suggests that the L-BFGS optimisation is robust with respect to noise, but it also may come from the fact that there does not exist any structure in the model space that could explain the applied noise, which is totally incoherent. We expect to encounter more difficulties when dealing with coherent noise in real data (e.g. ringing effects). In Fig. 2.26(b), it appears that the added noise slightly damaged the fit in some parts of the radargram, especially in the zone related to the thin lenses ($8 \leq x \leq 10$ m). For comparison, it must be mentioned that synthetic data computed in the model of Fig. 2.21, which has been reconstructed by inverting noise-free data, perfectly match the observed data.

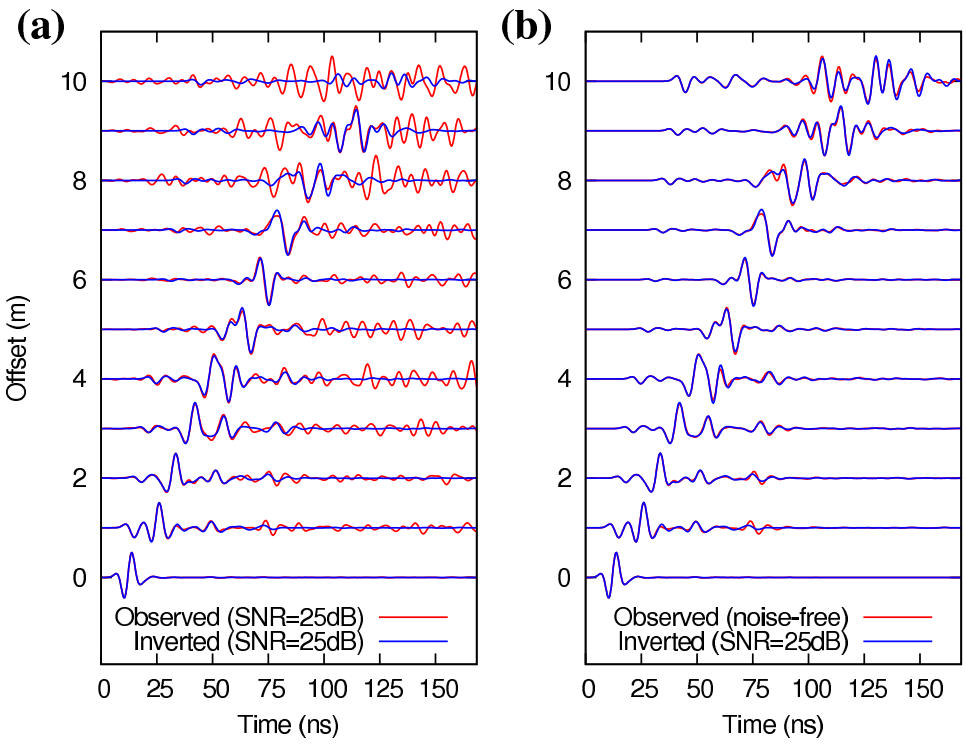


Figure 2.26: *Time data fit. a) Noisy observed data vs. data calculated in the model of figure 2.25 obtained by inverting the noisy data. b) Noise-free observed data vs. the same data calculated in the inverted model of Fig. 2.25. A time-varying gain ($\times t$) and trace-by-trace normalisation have been applied.*

2.2.4 Discussion

We have shown that a robust reconstruction of permittivity and conductivity requires both parameter scaling and regularisation. Here we may comment the similarities and differences between these two ingredients, which may appear redundant but have actually distinct roles.

As shown in Fig. 2.10, small values for the scaling factor β penalise the conductivity updates and provide smooth conductivity models, as does regularisation. Looking at the Hessian matrices, we can also note that penalising the conductivity with a small scaling factor amounts to damp the small singular values of the Hessian, because the misfit function is more sensitive to permittivity than to conductivity. It is also the effect of Tikhonov regularisation (Hansen, 2010, p.62). Finally, parameter scaling and regularisation both modify the shape of the global misfit function, and thus the inversion path, but in different ways we shall describe now.

Penalisation of the conductivity update through the scaling factor β orientates the inversion path in the direction of permittivity, until a satisfying kinematic model is obtained (see Fig. 2.12). The conductivity model is then updated only on the basis of the data misfit decrease, without an explicit smoothness requirement. Potentially, an adequate parameter scaling can guide the inversion on a reasonable path towards the minimum of the data misfit $\mathscr{C}_D$ and this solution can present smooth parts as well as contrasts. Conversely, regularisation attracts the inversion path towards a smooth conductivity model: The minimum of the global misfit function is then shifted towards the minimum of its model term $\mathscr{C}_M$ (in the two-parameter case of Fig. 2.12, it would be a valley located at $\sigma = 3$ mS/m). Consequently, the parameter scaling does not prevent the misfit function to converge, contrary to regularisation which generally results in a lower convergence rate: The optimisation stops when the updated models cannot both minimise the data misfit and satisfy the model smoothness requirement. As a conclusion, the role of the parameter scaling is to guide the inversion on a reasonable path, according to the sensitivity of the data, whereas regularisation constrains the conductivity update, prevents the creation of high-wavenumber structures, and thus enables us to discriminate between low-wavenumber reconstructions that well explain the data or not (Fig. 2.20). Indeed, the inversion results using small values for the scaling factors (e.g. $\beta = 0.2$) are very similar with and without regularisation (e.g. $\lambda = 5$ vs. $\lambda = 0$ in the noisy data case) because the penalisation of conductivity through the scaling factor already acts as a regularisation. The principal role of regularisation in our workflow is to tell us which smooth solution is convenient regarding the fit to the data.

Conclusion

In this study, we have presented a full waveform inversion algorithm of on-ground GPR data for the simultaneous reconstruction of permittivity and conductivity in 2D. The inverse problem has been formulated in the frequency-domain as the minimisation of the misfit to the data in a least-square sense. A model term is added to constrain the inversion with a Tikhonov regularisation. The gradient of the misfit function is defined in the whole parameter space and is computed with the adjoint state method. The optimisation is performed with a quasi-Newton scheme using the L-BFGS-B algorithm to economically estimate the effect of the Hessian on the parameter update.

Tests performed on a synthetic benchmark from the literature shows that the respective weights of permittivity and conductivity in the optimisation process is of prior importance. A

parameter scaling is introduced through a penalisation of the conductivity parameter. Without an adequate value for this scaling factor, regularisation alone cannot provide satisfactory results. The adequate value for the parameter scaling mainly depends on the respective sensitivity of the data to permittivity and conductivity, and to the quality of the initial model. With a weak sensitivity to conductivity and a poor initial model, more weight must be given to the permittivity parameter to give priority to the kinematic reconstruction before reconstructing the conductivity. The sensitivity of the reconstructions to the parameter scaling suggests that the L-BFGS algorithm does not correctly scale the descent direction with respect to different parameter types. We underline the need for investigating more complete approximations of the Hessian (e.g. truncated Newton methods, Métivier et al., 2013) to understand if more information can be extracted from the curvature of the misfit function.

The behaviour of the inversion with respect to frequency sampling has been investigating. As the relative impact of permittivity and conductivity varies with frequency, the reconstruction of both parameters takes a significant benefit from the simultaneous inversion of data with a broad frequency bandwidth. Therefore, simultaneous or cumulative frequency sampling strategies should be favoured, depending on the quality of the initial model.

The algorithm has also been tested on a more realistic benchmark, with a multi-offset, surface-to-surface acquisition configuration. In this case, various parameter scalings can lead to the same misfit decrease but to very different solutions. Regularisation is needed to constrain the conductivity update and reduce the ambiguity. In the synthetic case we investigate, it is possible to find a range of reasonable values for both the scaling factor and the regularisation weight, based only on the data misfit analysis. This workflow can be applied to extract a reliable information about conductivity from noisy data. We shall mention that, in some cases, it could be impossible to constrain the solution based only on the data misfit, and we underline the interest of introducing *a priori* information in the FWI process (Asnaashari et al., 2013).

The proposed workflow implying parameter scaling and regularisation enables us to consider the inversion of real data in the near future. Common obstacles to real data inversion are 3D to 2D conversion, the design of a compatible initial model, and the estimation of the source signal, which must be integrated in the iterative process. Since first travel time and amplitude tomography can not be performed from on-ground GPR data (contrary to crosshole data), the design of initial models for permittivity and conductivity from on-ground GPR data will be particularly challenging.

Acknowledgments

We would like to thank the SEISCOPE consortium for significant contribution (`http://seiscope2.osug.fr`). This work was performed by accessing to the high-performance computing facilities of CIMENT (Université de Grenoble) and of GENCI-CINES under grant 046091. We gratefully acknowledge both of these facilities and the support of their staff. We also thank an anonymous reviewer for its recommendations that helped to improve the clarity of the manuscript.

2.3 Further methodological details

In this section, I develop some points which were not included in the previous article for conciseness. First, I present the complete derivation of the gradient expression (2.19), which is at the core of the FWI. I comment the properties of the gradient in terms of resolution with respect to illumination, which give some insights inherited from seismic FWI. In a second time, I present some protocols used to validate the implementation of the gradient. I also show how to derive the two-parameter case of Fig. 2.12 using a linear change of variables. Finally, I shortly comment the sensitivity of the data to dispersive parameters and justify why I did not consider them in the inversion.

2.3.1 Computation and interpretation of the gradient

In this section, I derive the gradient expression (2.19) using the approach of (e.g.) Pratt et al. (1998) and Shin et al. (2007) which simply consists in differentiating the misfit function and the wave equation with respect to the model parameters. In Appendix A, I present the more general formulation of the adjoint state method (Plessix, 2006) which enables the consideration of other parameterisations and misfit criteria in a more flexible and systematic way.

For simplicity, I will avoid the summation of expression (2.17) by considering data at a given frequency and for one source only[1]. I also consider a more general form for the data misfit function

$$\mathscr{C}(\mathbf{m}) = \frac{1}{2}||\mathcal{W}_d\Delta\mathbf{d}(\mathbf{m})||^2, \tag{2.32}$$

$$= \frac{1}{2}\Delta\mathbf{d}(\mathbf{m})^\dagger\mathcal{W}_d^\dagger\mathcal{W}_d\Delta\mathbf{d}(\mathbf{m}), \tag{2.33}$$

where $\Delta\mathbf{d} = \mathbf{d}_{obs} - \mathbf{d}_{cal}(\mathbf{m})$ are the data residuals and $\mathcal{W}_d$ is a data weighting matrix. If the measurement uncertainties are available, it is consistent to relate this weighting matrix to the data covariance matrix $\mathcal{C}_d$ as

$$\mathcal{W}_d^\dagger\mathcal{W}_d = \mathcal{C}_d^{-1}. \tag{2.34}$$

This approach enables in theory to estimate *a posteriori* uncertainties on the recovered parameters, which are a valuable information (see e.g. Tarantola, 2005, §3.3, p. 70).

In our case, however, it is not obvious to reliably estimate the uncertainty of the GPR measurements. Instead, I will use the operator $\mathcal{W}_d$ as an arbitrary weighting matrix to give more weight to some data compared to others, e.g. to balance the small amplitudes of far offset data. Therefore, the weighting matrix will be a diagonal matrix such that

$$\mathrm{diag}(\mathcal{W}_d) = (w_{d_i})_{i=1}^{N_D}, \tag{2.35}$$

w_{d_i} being the weight given to data component d_{obs_i}.

Even if the same weight is given to each data component, it is consistent to use the weighting matrix $\mathcal{W}_d$ to normalise the misfit function, as done by Yang et al. (2013), using $w_{d_i} = 1/N_D$.

[1] Equivalently, we can consider that the data recorded at different frequencies for different source and receiver positions can be gathered in a unique residual vector $\Delta\mathbf{d}(\mathbf{m}) = \mathbf{d}_{obs} - \mathbf{d}_{cal}(\mathbf{m})$, but the reference to the impedance matrix $\mathcal{A}(\mathbf{m})$ is then less consistent since it should depend on frequency.

Equivalently, a diagonal model weighting matrix $\mathcal{W}_m$ can be used to normalise the model term of the regularised misfit function (**??**), using $w_{m_i} = 1/N_M$.

More generally, it is worth noting that the data and model weighting matrix define norms in the data space and in the model space, respectively (Menke, 2012, §3.9, p. 56), such that their application to the residuals and model vectors can be seen as a change of variables between initial data and model spaces $\mathbb{D}$ and $\mathbb{M}$ to weighted, or normed, spaces.

The gradient of the misfit function is by definition its first derivative with respect to the model parameters, i.e. the vector $\mathbf{G}$ of size N_M whose components are given by

$$\begin{aligned} G_i(\mathbf{m}) &= \frac{\partial \mathscr{C}(\mathbf{m})}{\partial m_i}, && (2.36)\\ &= \frac{1}{2}\frac{\partial \Delta\mathbf{d}^\dagger}{\partial m_i}\mathcal{W}_d^\dagger\mathcal{W}_d\Delta\mathbf{d} + \frac{1}{2}\Delta\mathbf{d}^\dagger\mathcal{W}_d^\dagger\mathcal{W}_d\frac{\partial\Delta\mathbf{d}}{\partial m_i}, \\ &= \Re e\left\{\frac{\partial\Delta\mathbf{d}(\mathbf{m})^\dagger}{\partial m_i}\mathcal{W}_d^\dagger\mathcal{W}_d\Delta\mathbf{d}(\mathbf{m})\right\}. && (2.37)\end{aligned}$$

TE mode

In TE mode, the calculated data $\mathbf{d}_{cal}$ consist in the field $\mathbf{u} = \mathbf{E_y}$ recorded at receiver positions, which can be formalised by

$$\Delta\mathbf{d}(\mathbf{m}) = \mathbf{d}_{obs} - \mathbf{d}_{cal}(\mathbf{m}) = \mathbf{d}_{obs} - \mathcal{R}\mathbf{u}(\mathbf{m}), \tag{2.38}$$

where $\mathcal{R}$ is again the projection operator of the simulated field $\mathbf{u}$ on the receiver positions. So we have

$$\frac{\partial\Delta\mathbf{d}(\mathbf{m})}{\partial m_i} = -\frac{\partial\mathbf{d}_{cal}(\mathbf{m})}{\partial m_i} = -\mathcal{R}\frac{\partial\mathbf{u}(\mathbf{m})}{\partial m_i}. \tag{2.39}$$

Expression (2.39) can be injected in (2.37) to yield

$$G_i = -\Re e\left\{\frac{\partial\mathbf{u}^\dagger}{\partial m_i}\mathcal{R}^\dagger\mathcal{W}_d^\dagger\mathcal{W}_d\Delta\mathbf{d}\right\}. \tag{2.40}$$

On the other hand, derivating the equation of the forward problem $\mathcal{A}(\mathbf{m})\mathbf{u}(\mathbf{m}) = \mathbf{s}$ with respect to the model parameters m_i, we obtain

$$\begin{aligned} \mathcal{A}(\mathbf{m})\frac{\partial\mathbf{u}(\mathbf{m})}{\partial m_i} + \frac{\partial\mathcal{A}(\mathbf{m})}{\partial m_i}\mathbf{u}(\mathbf{m}) &= 0, \\ \text{i.e., formally,}\quad \frac{\partial\mathbf{u}(\mathbf{m})}{\partial m_i} &= -\mathcal{A}(\mathbf{m})^{-1}\frac{\partial\mathcal{A}(\mathbf{m})}{\partial m_i}\mathbf{u}(\mathbf{m}). && (2.41)\end{aligned}$$

Replacing (2.41) in (2.40) leads to the gradient expression of the adjoint state method:

$$G_i(\mathbf{m}) = \Re e\left\{\mathbf{u}(\mathbf{m})^\dagger\frac{\partial\mathcal{A}(\mathbf{m})^\dagger}{\partial m_i}\mathcal{A}(\mathbf{m})^{\dagger^{-1}}\mathcal{R}^\dagger\mathcal{W}_d^\dagger\mathcal{W}_d\Delta\mathbf{d}(\mathbf{m})\right\}, \tag{2.42}$$

or, if we denote $\mathbf{v}$ the adjoint wavefield,

$$G_i(\mathbf{m}) = \Re e \left\{ \mathbf{u}(\mathbf{m})^\dagger \frac{\partial \mathcal{A}(\mathbf{m})^\dagger}{\partial m_i} \mathbf{v}(\mathbf{m}) \right\}, \tag{2.43}$$

which is similar to equation (2.19), given that $\Re e(x) = \Re e(x^*)$ for any complex number x. Note however that defining the adjoint wavefield as

$$\mathbf{v}(\mathbf{m}) = \mathcal{A}(\mathbf{m})^{\dagger^{-1}} \mathcal{R}^\dagger \mathcal{W}_d^\dagger \mathcal{W}_d \Delta\mathbf{d}(\mathbf{m}) \tag{2.44}$$

is only a formal notation since in practice, the inverse matrix $\mathcal{A}^{-1}$ is not explicitly computed. Therefore, it is more rigorous to define the adjoint wavefield $\mathbf{v}$ as the solution of the adjoint equation

$$\mathcal{A}(\mathbf{m})^\dagger \mathbf{v}(\mathbf{m}) = \mathcal{R}^\dagger \mathcal{W}_d^\dagger \mathcal{W}_d \Delta\mathbf{d}(\mathbf{m}), \tag{2.45}$$

which is a wave propagation problem where the sources are located at receiver positions through the operator $\mathcal{R}^\dagger$ and their strength (amplitude and phase) is given by the weighted residuals $\mathcal{W}_d^\dagger \mathcal{W}_d \Delta\mathbf{d}(\mathbf{m})$. Therefore, it is usual to say that the adjoint wavefield results from the back-propagation of the residuals in the medium, since the matrix $\mathcal{A}^\dagger$ is a backward-propagation operator. This property can be more easily seen in the time-domain, where solving the adjoint wave equation amounts to solve the forward wave equation backwards in time (see e.g. Lailly, 1983; Plessix, 2006).

TM mode

The previous derivation stands for the TE mode where the data are values of the electric field recorded at receiver positions, $\mathbf{d} = \mathcal{R}\mathbf{u}$ with $\mathbf{u} = \mathbf{E_y}$. In TM mode, the extraction of the recorded electric field from the simulated magnetic field involves an additional operator $\mathcal{D}$:

$$\mathbf{d}_{cal}(\mathbf{m}) = \mathcal{R}\mathcal{D}\mathbf{u}(\mathbf{m}), \quad \text{with} \quad \mathbf{u} = \mathbf{H_y}, \tag{2.46}$$

$$\text{and} \quad \mathcal{D} = \frac{1}{\imath\omega\varepsilon_e} \left(\cos\theta_r \frac{\partial}{\partial z} - \sin\theta_r \frac{\partial}{\partial x} \right), \tag{2.47}$$

where θ_r is the orientation of the receiving antenna with respect to the x-axis (see equations 1.61 to 1.63 and Fig. 1.7).

Considering the operator $\mathcal{D}$ in the previous development yields the gradient expression for the TM mode:

$$G_i(\mathbf{m}) = \Re e \left\{ \mathbf{u}(\mathbf{m})^\dagger \frac{\partial \mathcal{A}(\mathbf{m})^\dagger}{\partial m_i} \mathcal{A}(\mathbf{m})^{\dagger^{-1}} \mathcal{D}^\dagger \mathcal{R}^\dagger \mathcal{W}_d^\dagger \mathcal{W}_d \Delta\mathbf{d}(\mathbf{m}) \right\}. \tag{2.48}$$

The corresponding adjoint equation is

$$\mathcal{A}(\mathbf{m})^\dagger \mathbf{v}(\mathbf{m}) = \mathcal{D}^\dagger \mathcal{R}^\dagger \mathcal{W}_d^\dagger \mathcal{W}_d \Delta\mathbf{d}(\mathbf{m}), \tag{2.49}$$

where the adjoint operator $\mathcal{D}^\dagger$ must be determined to implement the correct source for the adjoint problem. If we achieve the differentiations of eq. (2.47) with centred finite-differences,

$$\frac{\partial u_{i,j}}{\partial x} \simeq \frac{u_{i+1,j} - u_{i-1,j}}{2h}, \tag{2.50}$$

$$\frac{\partial u_{i,j}}{\partial z} \simeq \frac{u_{i,j+1} - u_{i,j-1}}{2h}, \tag{2.51}$$

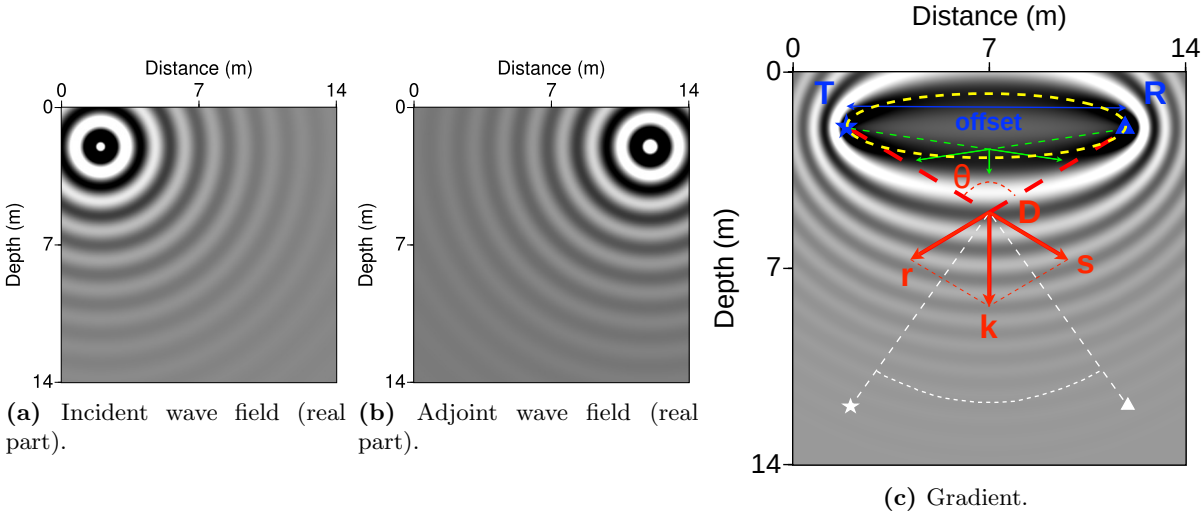
then it turns out that $\mathcal{D}$ is real and anti-symmetric: $\mathcal{D}^\dagger = -\mathcal{D}$.

Interpretation of the gradient in the imaging process

The principles underlying the construction of the gradient in FWI can be encountered in many imaging methods. In fact, it corresponds to the imaging principle of *time-coincidence* I evoked in the general introduction (Claerbout, 1971). In particular, it corresponds to the imaging condition of reverse-time migration, with the notable difference that for migration, not the residuals but the entire recorded data are back-propagated[1], and correlated with the incident field to form the image, which results from the interferences between both wavefields. If an accurate velocity model is available, constructive interferences occur where the wavefields are time-coincident, i.e. where there is a reflector that causes a reflected event in the recorded data[2]. In FWI, the gradient thus act as mapping operator of the residuals in the medium: it will locate model perturbations where physical structures that have caused the un-explained events are missing.

Detailing the construction of the gradient provides useful insight on the resolution capability of the inversion. Fig. 2.27 illustrates this construction for one frequency (100 MHz) and one source-receiver pair in the cross-shaped benchmark case. The computation is performed in the initial homogeneous background for the permittivity component of the gradient. To form the gradient (c), the incident field (a) and the adjoint field (b) are correlated, i.e. multiplied in the frequency domain. The resulting gradient (c) displays a particular interference pattern, commonly referred to as the sensitivity kernel. It consists in elliptical iso-phase contours, i.e. that diffracting points located on these contours would generate diffracted events in the data with the same phase. These iso-phase contours are thus the frequency-domain equivalent of the elliptical iso-travel-time contour in the time-domain, duplicated by the phase ambiguity.

The central ellipse (yellow dashed line) corresponds to the first Fresnel zone, associated to the first arrivals, while the outer fringes are associated to later arrivals (reflected and diffracted



(a) Incident wave field (real part).

(b) Adjoint wave field (real part).

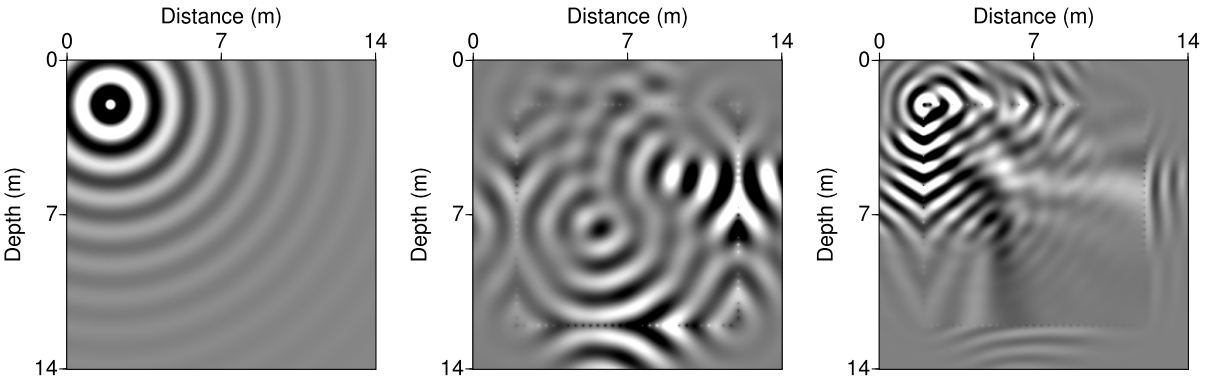
(c) Gradient.

Figure 2.27: *Construction of the gradient in the case of a single source-receiver pair and of a single frequency (100 MHz). In (c), the yellow dashed line delineates the first Fresnel zone.*

[1] As a consequence, a FWI algorithm is also a reverse-time migration algorithm, provided we put the data in the adjoint source instead of the residuals. I will use this property in Chapter 3 for migrating field data.

[2] It is also the principle of time-reversal experiments that consist in refocusing a recorded signal, either acoustic or electromagnetic, at the location of the source (e.g. Fink et al., 2000; Davy, 2010).

events). Imaging the missing diffracting points consists in back-projecting the residuals on the iso-phase contours: residuals associated with first arrivals (i.e., having a time shift of less than $T/2$ with these arrivals, in virtue of the cycle-skipping effect) will be projected into the first Fresnel zone and residuals associated with later arrivals on the outer fringes. Of course, back-projecting the residuals from a single source-receiver pair does not provide a satisfying image since there is an azimuthal ambiguity on the source of the diffractions. Constructive interferences leading to a focused image only arise when multiple receivers are considered (Fig. 2.28), just like the epicentre of earthquakes can be roughly located by triangulation. Finally, the missing crosses can be accurately relocated by superimposing the interference patterns associated to several sources, that provides a more complete illumination of the targets (Fig. 2.29).



(a) Incident wave field. **(b)** Adjoint wave field. **(c)** Gradient.

Figure 2.28: *Construction of the gradient in the case of a single source with multiple receivers at 100 MHz.*

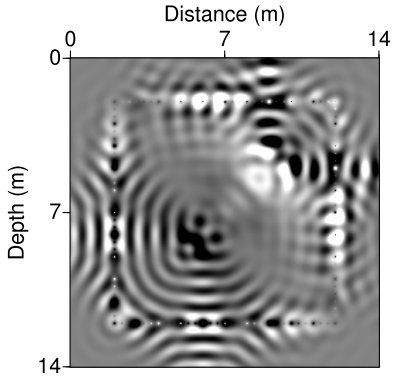


Figure 2.29: *Gradient at 100 MHz resulting from the sum of the interference patterns similar to Fig. 2.28(c) for all sources.*

The question of the illumination of the medium by the acquisition setup, and the related issue of the resolution of the reconstructed image, can be studied in more details thanks to the sensitivity kernel of Fig. 2.27. On this figure, the distance between the iso-phase contours gives an indication on the resolution in the resulting image. This resolution is controlled by the local wavenumber $\mathbf{k}$, which results from the combination of the wavenumber vector $\mathbf{s}$, representing the illumination of the medium by transmitter T, and of the wavenumber vector $\mathbf{r}$, standing for the field recorded by receiver R. Based on a plane wave analysis, Sirgue and Pratt (2004) give an expression for the local wavenumber vector $\mathbf{k}$, as a function of frequency and of the

diffraction angle θ, or aperture of the source-receiver array:

$$\mathbf{k} = \frac{2\omega}{v} \cos\left(\frac{\theta}{2}\right) (\mathbf{s} + \mathbf{r}), \tag{2.52}$$

where v is the electromagnetic velocity in the medium. Equation (2.52) is important for understanding the FWI of surface-based measurements and gives several informations:

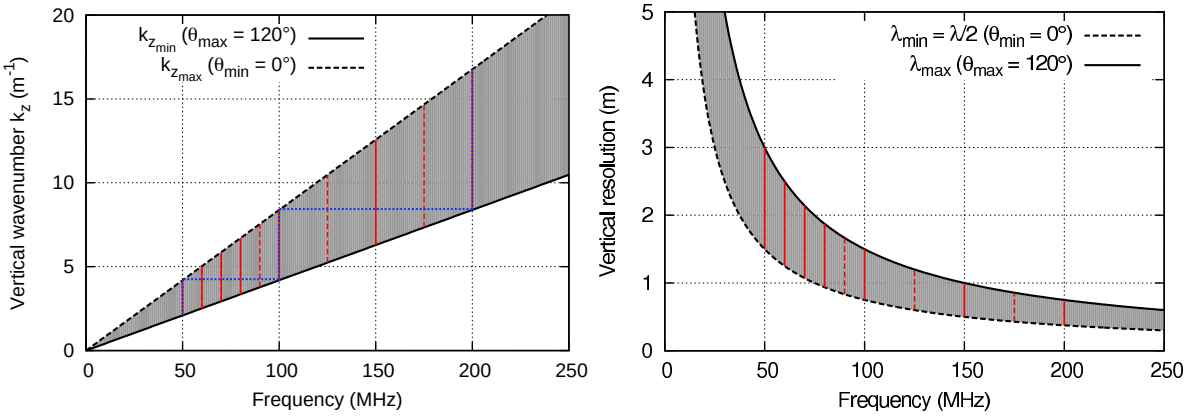
1. The resolution of the reconstructed image both depends on frequency and on the illumination angle, which in turn is related to the offset between source and receiver. More precisely, the resolution increases with increasing frequencies and decreasing diffraction angles (compare the green and the red arrows in Fig. 2.27).
2. The above statement confirms the intuitive interpretation that the low to high frequency hierarchy of Pratt and Worthington (1990) amounts to reconstruct first the low wavenumbers of the image, i.e. the large-scale structures, and then to incorporate small-scale details by inverting higher frequencies.
3. The maximal achievable resolution is obtained from zero-offset data and equals $k_{max} = 2\omega/v$, i.e. $\lambda_{min} = \lambda_{propagated}/2$, in accordance with the diffraction theory.

The first point leads Sirgue and Pratt (2004) to propose a strategy for selecting the adequate frequencies to be considered in the inversion. This strategy is illustrated in Fig. 2.30(a) in the case of the cross-shaped benchmark, for the diffracting point indicated in red in Fig. 2.27, which is located at $x = 7$ m and $z = 3$ m [1]. The left panel of Fig. 2.30(a) follows the 1D analysis of Sirgue and Pratt (2004) and represents only the vertical component of the wavenumber that can be imaged by an acquisition in reflection configuration, with sources and receivers spanned on the top of the model ($z = 2$ m in Fig. 2.6). The right panel of Fig. 2.30(a) displays the corresponding vertical resolution in terms of wavelengths, which have more intuitive units for the interpretation. For each frequency, the minimal and maximal illumination angles provided by the acquisition result in a range of vertical wavenumbers. Two adjacent frequencies therefore provide a large redundancy in wavenumber coverage.

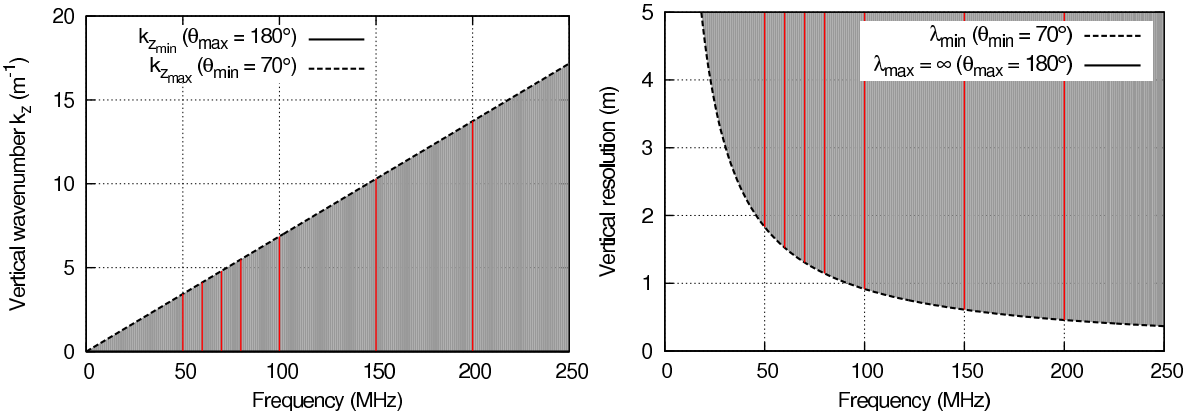
This observation is the motivation for decimating the frequency sampling in frequency-domain FWI: thanks to the redundancy in wavenumber coverage offered by the multi-offset acquisition, we can invert coarsely sampled frequencies without damaging the resolution of the reconstructed image. The dashed blue lines in Fig. 2.30(a) indicate the frequency sampling suggested by Sirgue and Pratt (2004), which is *a priori* sufficient to avoid gaps in the wavenumber coverage. It consists in only three frequencies: 50, 100 and 200 MHz. As this analysis concerns only one point in the medium, and is valid only for the vertical wavenumber, it is often worth over-estimating the required frequency sampling. In the real data case, wavenumber redundancy is even desirable to increase the signal-over-noise ratio. It is why I generally choose a denser frequency sampling than the one proposed by Sirgue and Pratt (2004). As an example, the vertical red lines indicate the frequencies used in the inversions of Section 2.2 (the dashed red lines correspond to the extra frequencies considered in the case of the realistic subsurface benchmark).

The analysis of Sirgue and Pratt (2004) is an interesting tool to understand the difference of illumination provided by surface-based measurements, in comparison with crosshole experiments. I already mentioned this issue in the introduction of Section 2.2 and Meles et al. (2012)

[1] It corresponds to the depth of the red cross of the benchmark. It is also the depth of the high-conductivity layer in the subsurface benchmark of Fig. 2.14.



(a) Reflection configuration (sources and receivers at $z = 2$ m in Fig. 2.6).



(b) Transmission configuration (sources at $x = 2$ m and receivers at $x = 12$ m in Fig. 2.6).

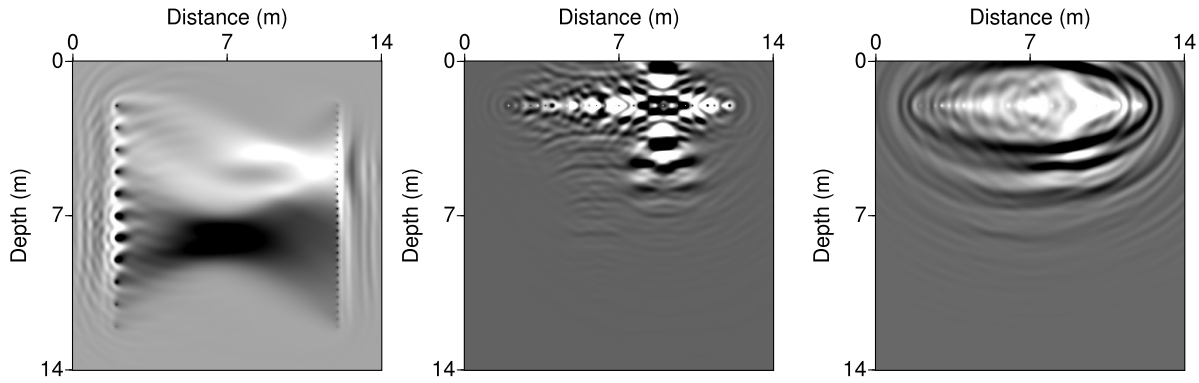
Figure 2.30: *Vertical wavenumbers (left panels) and wavelengths (right panels) imaged in the cross-shaped benchmark, for reduced acquisition geometries and a diffracting point located at 3 m depth.*

provide a detailed discussion of the sensitivity and resolution provided by both configurations, based on the explicit computation of the Jacobian matrix and on the eigenvalue decomposition of the corresponding pseudo-Hessians. In Fig. 2.30(b), I present the vertical wavenumber and resolution provided at the same diffracting point but illuminated by a crosshole acquisition setup, with sources on one side of the model (at $x = 2$ m) and receivers on the opposite side ($x = 12$ m). In this configuration, the minimal illumination angle of the diffracting point is 70° (indicated with white dashed lines in Fig. 2.27) and the maximal diffraction angle is 180° for source and receiver located at the same depth (pure transmission regime).

Eq. (2.52) then predicts a minimal vertical wavenumber of zero, meaning that very large vertical structures can always be imaged, and a maximal wavenumber that is slightly smaller than in the reflection case, due to the larger minimal angle of illumination. Therefore, we can expect a slightly better vertical resolution from reflection measurements than from crosshole experiments ($\simeq$ 75 cm *vs.* $\simeq$ 1 m at 100 MHz, comparing the right panels of Figs 2.30a and b). Conversely, Fig. 2.30(a) suffers from a dramatic lack of low wavenumbers compared to Fig. 2.30(b). As a consequence, in reflection configuration, it will be difficult for the gradient to reconstruct the large-scale structures of the medium. If the low wavenumbers cannot be

imaged from reflection data, then they must be contained in the initial model. Otherwise, the small-scale structures will not be imaged at their correct locations. We can thus anticipate that the design of an accurate initial model of permittivity will be crucial when dealing with on-ground GPR field data, whereas this issue seems less critical in crosshole configuration, according to this analysis.

As a qualitative illustration of the wavenumber coverage provided by reflection and crosshole configurations, Fig. 2.31 presents the gradients corresponding to both configurations. In addition to the incomplete illumination of the medium, already investigated by Meles et al. (2012), it confirms the lack of low wavenumbers in the gradient associated to surface acquisition (b). In panel (c), weights have been applied to increase the contribution of long-offset data, originally of lower amplitudes. It slightly enhances the low wavenumber content but is not sufficient to converge to a satisfactory solution in this specific case because of the incomplete illumination and of the absence of deep reflectors (data weighting will be applied to this purpose on real field data in Chapter 3).



(a) Crosshole configuration. **(b)** Surface configuration. **(c)** Surface configuration with weighted residuals.

Figure 2.31: *Gradients computed in the homogeneous initial background of the cross-shaped benchmark, using various acquisition configurations. In (c), data are weighted such that long offset data have the same amplitudes than short offset ones.*

2.3.2 Validation of the computed gradients

As the gradient is the key ingredient of FWI, it is important to validate the computation performed with the adjoint state method when various parameterisations or misfit functions are tested. To do so, we can make two kinds of tests. Both aim at checking that the computed gradient well corresponds to the first derivatives of the misfit function.

1. We may check that, at each grid point i in a model $\mathbf{m}$,

$$G_i(\mathbf{m}) = \frac{\partial \mathscr{C}}{\partial m_i} \simeq \frac{\mathscr{C}(\mathbf{m} + \delta \mathbf{m_i}) - \mathscr{C}(\mathbf{m})}{\delta m_i}, \quad \forall\, m_i \in \mathbb{M}, \tag{2.53}$$

where the vector $\delta \mathbf{m_i} = (0 \cdots \delta m_i \cdots 0)^T$ has all its components null except the i^{th}. On a practical point of view, it corresponds to construct the gradient in a *naive* way,

without using the adjoint state method, but using a finite difference approximation, i.e. performing one forward problem per parameter to estimate the cost function in a perturbed medium.

However, relation (2.53) is only valid for adequate perturbation amplitudes δm_i, which have to be small enough for verifying the linear approximation, but large enough for not being dominated by rounding errors. An adequate value for δm_i can be estimated by trial and error, but it is sometimes useful to investigate the convergence of the approximation (2.53) in more details.

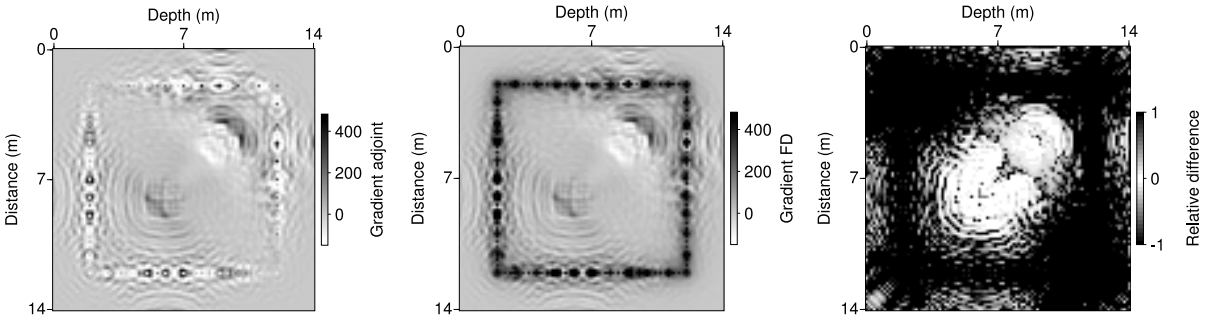
2. Therefore, a more rigorous test is to look at the strict definition of the derivatives, checking that

$$G_i(\mathbf{m}) = \frac{\partial \mathscr{C}}{\partial m_i} = \lim_{\delta m_i \to 0} \frac{\mathscr{C}(\mathbf{m} + \delta \mathbf{m_i}) - \mathscr{C}(\mathbf{m})}{\delta m_i}. \tag{2.54}$$

with decreasing perturbation amplitudes $\delta = \delta m_i / m_i$, such as to obtain a convergence curve as a function of δ. In practice, it amounts to perform the first test (2.53) for decreasing perturbation amplitudes δm_i. Since it requires more computations than the first test, I generally perform this second test at only one point i, such as to obtain a convergence curve for $G_i(\mathbf{m})$.

Below, I present validation tests using the first method for the validation of the permittivity and conductivity gradients derived with the adjoint method. The second method will be used in the next section to validate the gradient in the two-parameter case, where it is not obvious to evaluate the agreement between the adjoint values and the finite-difference approximation.

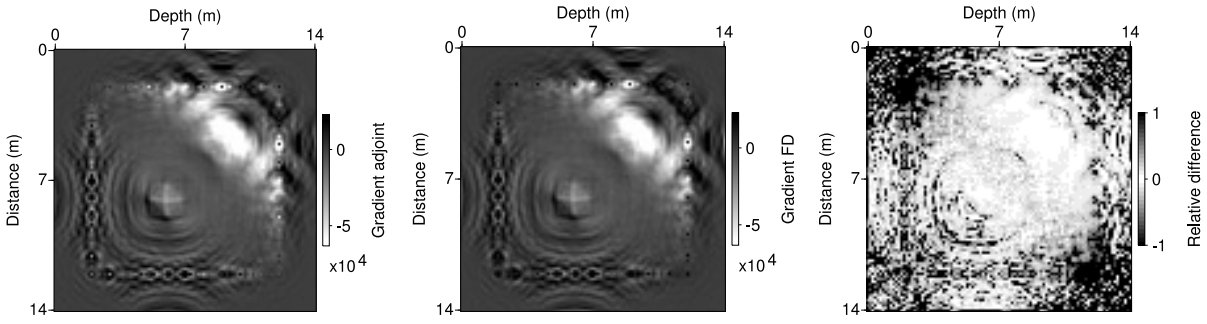
Figs 2.32(a,b) and 2.33(a,b) show the gradients computed using the adjoint state method and the finite-difference approximation in the initial homogeneous background of the cross-shaped benchmark, both for permittivity and conductivity. The qualitative agreement between both computations is very good, except in the immediate vicinity of sources and receivers that cause numerical singularities. Panels (c) present the quantitative relative difference, which is very small at the centre of the medium, in particular in the zone of interest where the crosses are to be reconstructed. In Fig. 2.32(c), however, the discrepancy in the permittivity gradient rapidly increases close to the sources and receivers due to an enhanced sensitivity of the data to this region. Actually, the error resides here in the finite difference computation because



(a) Adjoint. **(b)** Finite-difference. **(c)** Relative difference.

Figure 2.32: *Validation of the permittivity gradient of* $\mathbf{G}^{\varepsilon_\mathbf{r}}$ *in the initial model of the cross-shaped benchmark for TE mode and three simultaneous frequencies (50, 120, and 200 MHz). Finite-differences are computed with perturbations* $\delta = 10\%$.

the applied perturbation amplitude $\delta = 10\%$ does not respect the linear approximation. It has been checked that decreasing the perturbation amplitude δ used to compute the finite-difference gradient enables to decrease the relative error in this region, displaying a convergence curve similar to the one shown in Fig. 2.35(a) in the next section. This effect is less visible in the conductivity gradient (Fig. 2.33c) because of the smaller sensitivity of the data to conductivity. The perturbation amplitude $\delta = 10\%$ thus remains in the validity domain of the linear approximation.



(a) Adjoint. **(b)** Finite-difference. **(c)** Relative difference.

Figure 2.33: *Validation of the conductivity gradient of $\mathbf{G}^{\sigma_\mathbf{r}}$ in the initial model of the cross-shaped benchmark for TE mode and three simultaneous frequencies (50, 120, and 200 MHz). Finite-differences are computed with perturbations $\delta = 10\%$.*

2.3.3 Derivation and validation of the two-parameter problem

In this section, I detail how to derive the two-parameter problem of Fig. 2.12 from the high-dimensional cross-shaped benchmark of Fig. 2.6. By *high-dimensional*, I mean that in Fig. 2.6, the permittivity and conductivity models are represented on a cartesian grid of size $N_x \times N_z = 201 \times 201$. This forward grid is designed to guarantee an accurate solution of the forward problem, whereas an inverse grid aims at representing the resolvable structures of the expected model. Therefore, it may be pertinent in a full waveform inversion to not reconstruct the parameters values at all points of the forward grid. In this section, I derive the computation of the gradient with respect to the parameters defined on a decimated inverse grid. The new parameter set is linked to the forward parameters by a linear change of variable.

General formulation

Let we consider a forward grid of size N_M^{fwd}, and an inverse grid of size N_M^{inv}. As we are interested in reducing the number of parameters to be inverted, I assume $N_M^{inv} < N_M^{fwd}$. The associated vectors for model parameters are $\mathbf{m}^{fwd}$ and $\mathbf{m}^{inv}$, of size $N_p \times N_M^{fwd}$ and $N_p \times N_M^{inv}$, respectively, where N_p is the number of inverted parameter types. In the following, I focus on the monoparameter case ($N_p = 1$) since the multiparameter case is only a matter of size of the operators, and the different parameter types can be considered independently. I define the change of variables $\mathbf{m}^{fwd} \rightarrow \mathbf{m}^{inv}$ through the operator $\mathcal{C}$ such that

$$\mathbf{m}^{fwd} = \mathcal{C}\mathbf{m}^{inv}. \tag{2.55}$$

The matrix $\mathcal{C}$ is then of size (N_M^{fwd}, N_M^{inv}). In the general case, $\mathcal{C}$ will not be square, and therefore not invertible. However, we can obtain the following reciprocal change of variable

$$\mathbf{m}^{inv} = \left(\mathcal{C}^T\mathcal{C}\right)^{-1}\mathcal{C}^T\mathbf{m}^{fwd}, \tag{2.56}$$

if the matrix $\mathcal{C}^T\mathcal{C}$ is invertible (this condition should be verified to apply the workflow for particular cases).

Using the chain rule for the derivation of the misfit function $\mathscr{C}(\mathbf{m}^{inv})$, which is of course the same in both forward and inverse domains ($\mathscr{C}(\mathbf{m}^{inv}) = \mathscr{C}(\mathbf{m}^{fwd})$), we then have

$$\begin{aligned}
\mathbf{G}^{inv} &= \frac{\partial\mathscr{C}(\mathbf{m}^{inv})}{\partial\mathbf{m}^{inv}}, &(2.57)\\
&= \left(\frac{\partial\mathbf{m}^{fwd}}{\partial\mathbf{m}^{inv}}\right)^T\frac{\partial\mathscr{C}(\mathbf{m}^{inv})}{\partial\mathbf{m}^{fwd}} = \left(\frac{\partial\mathbf{m}^{fwd}}{\partial\mathbf{m}^{inv}}\right)^T\frac{\partial\mathscr{C}(\mathbf{m}^{fwd})}{\partial\mathbf{m}^{fwd}}, &(2.58)\\
&= \mathcal{C}^T\mathbf{G}^{fwd}, &(2.59)
\end{aligned}$$

where $\mathbf{G}^{fwd}$ is the gradient computed with the adjoint state method on the forward grid.

Before moving to the case of the cross-shaped benchmark studied in Section 2.2, I want to underline that this derivation, based on the chain-rule, is very general. As soon as it is possible to compute the change of variables and its derivative, it can be applied to a wide range of re-parameterisations. In the following, I consider a blocky description of the medium which makes a strong assumption on its geometry (in particular, I assume that the location and shape of the crosses are known). But it could also be applied to more clever re-parameterisations, e.g. based on the analysis of migrated images (Ma et al., 2012; Zhou et al., 2014).

Example in the case of the cross-shaped benchmark

In this case, we want to retrieve the permittivity and conductivity values in the cross-shaped anomalies, considered as homogeneous blocks in a homogeneous background (the benchmark is shown again in Fig. 2.34). The change of variables then consists to pass from the $N_M^{fwd} = N_x \times N_z$ cells in the FD grid to the blocky representation shown in Fig. 2.34(c), which presents $N_M^{inv} = 3$ sub-blocks. Hereafter, I will denote model and gradient components in the inverse grid as m_j^{inv} and G_j^{inv}, respectively, with $j = 1$ for the first (blue) cross, $j = 2$ for the second (red) cross, and $j = 3$ for the background (following the legend of Fig. 2.34c).

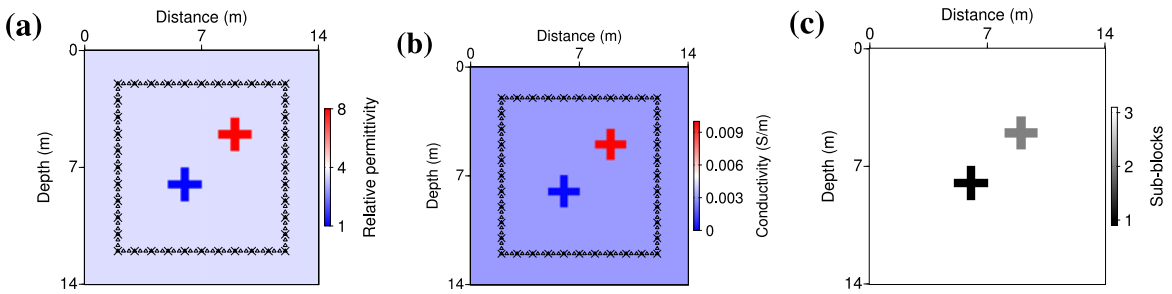


Figure 2.34: *(a,b) Reminder of the true models of the cross-shaped benchmark for permittivity and conductivity. (c) Sub-blocks defining the inverse grid.*

The change of variables $\mathbf{m}^{inv} \rightarrow \mathbf{m}^{fwd}$ then reads

$$\forall i \in [\![1, N_x \times N_z]\!], \quad m_i^{fwd} = \begin{cases} m_1^{inv} & \text{if } i \text{ is a cell of the first cross,} \\ m_2^{inv} & \text{if } i \text{ is a cell of the second cross,} \\ m_3^{inv} & \text{if } i \text{ is a cell of the background,} \end{cases} \tag{2.60}$$

or, in terms of the operator $\mathcal{C}$,

$$\forall i \in [\![1, N_x \times N_z]\!], \quad m_i^{fwd} = \sum_{j=1}^{3} C_{ij} m_j^{inv}, \quad \text{with} \quad C_{ij} = \begin{cases} 1 & \text{if} \quad i \in \mathcal{V}_j, \\ 0 & \text{otherwise,} \end{cases} \tag{2.61}$$

where $\mathcal{V}_j$ is the ensemble of the forward cells belonging to the inverse sub-block j.

In this case, the 3×3 matrix $\mathcal{C}^T\mathcal{C}$ is

$$\mathcal{C}^T\mathcal{C} = \begin{pmatrix} n_1 & 0 & 0 \\ 0 & n_2 & 0 \\ 0 & 0 & n_3 \end{pmatrix}, \tag{2.62}$$

where $n_j = \#\mathcal{V}_j$ is the number of forward cells in sub-block j. The matrix $\mathcal{C}^T\mathcal{C}$ is diagonal (and invertible since $\forall j \in \{1, 2, 3\}$, $n_j \neq 0$), so we can define the reciprocal operator $\mathcal{Q} = \left(\mathcal{C}^T\mathcal{C}\right)^{-1}\mathcal{C}^T$ and the change of variables $\mathbf{m}^{fwd} \rightarrow \mathbf{m}^{inv}$ as

$$\forall j \in \{1, 2, 3\}, \quad m_j^{inv} = \sum_{i=1}^{N_x \times N_z} Q_{ji} m_i^{inv}, \quad \text{with} \quad Q_{ji} = \begin{cases} 1/n_j & \text{if} \quad i \in \mathcal{V}_j, \\ 0 & \text{otherwise.} \end{cases} \tag{2.63}$$

Formally, m_j^{inv} is thus the mean value of m_i^{fwd} in the sub-block j: $m_j^{inv} =< m_i^{fwd} >_{i \in \mathcal{V}_j}$. It becomes here obvious that the change of variables $\mathbf{m}^{fwd} \rightarrow \mathbf{m}^{inv}$ is not injective, as several heterogeneous forward models can result in the same mean values for the inverse model (it was already the case in section 2.3.3, as we assumed that $N_M^{inv} < N_M^{fwd}$). In the case of the blocky representation of the crosses, the ambiguity is removed by the fact that we start with homogeneous sub-blocks, and that the model is updated homogeneously within the sub-blocks.

The gradient in the blocky parameter space is then

$$\forall j \in \{1, 2, 3\}, \quad G_j^{inv} = \sum_{i=1}^{N_x \times N_z} P_{ji} G_i^{fwd}, \tag{2.64}$$

i.e., the sum of the gradient components computed with the adjoint method on the forward grid, related to the cells belonging to the sub-block j.

Validation of the gradient on the inverse grid

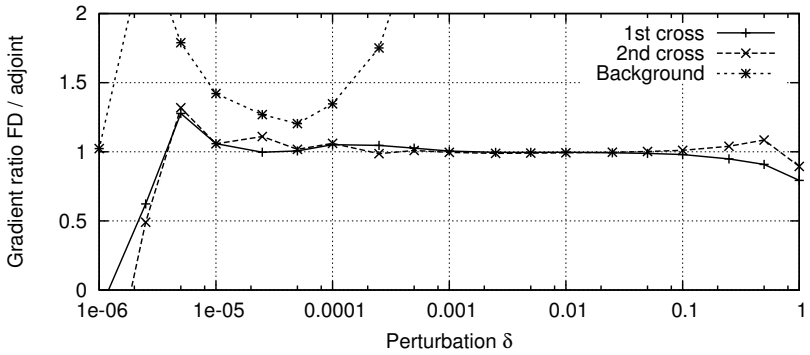
To validate the gradient in the inverse grid $\mathbf{G}^{inv}$, I compare it with a gradient computed by finite-differences $\mathbf{G}^{FD}$, defined as

$$G_j^{FD}(\mathbf{m}^{inv}) = \frac{C(\mathbf{m}^{inv} + \delta m_j) - C(\mathbf{m}^{inv})}{\delta m_j}. \tag{2.65}$$

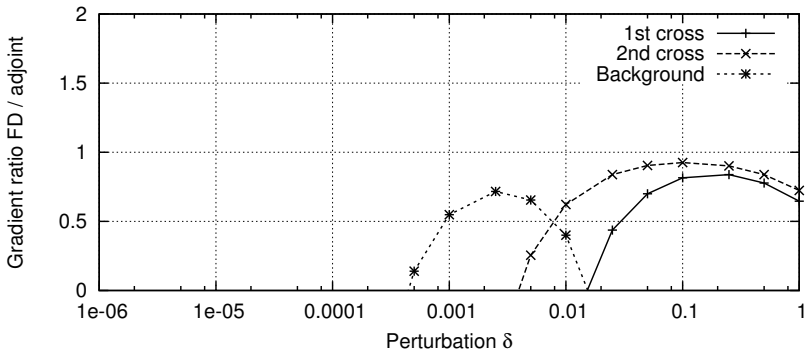
I apply perturbations of the form $\delta m_j = \delta \times m_j$. The convergence of the ratio G_j^{FD}/G_j^{inv} with respect to the perturbation δ is shown for each sub-block j in Fig. 2.35, both for permittivity (a) and conductivity (b).

In Fig. 2.35(a), the ratio G_j^{FD}/G_j^{inv} converges towards 1 for $\delta \in [0.005, 0.1]$ in the two crosses ($j = 1, 2$). In the background, the ratio gets closer to 1 for $\delta \in [10^{-5}, 10^{-4}]$ but never reaches 1, which is probably due to the fact that the acquisition system is comprised in the background, generating singularities both in the adjoint and finite-difference gradients.

In Fig. 2.35(b), the ratio G_j^{FD}/G_j^{inv} gets closer to 1 for higher perturbation amplitudes than in the permittivity case, because the misfit function is less sensitive to conductivity than to permittivity. The ratio never reaches exactly 1. The more reliable gradient is obtained for the second cross. It is why I focused on this second cross (the red one in Fig. 2.34) in Section 2.2. It also allows to avoid the bound $\varepsilon_r = 1$ that makes the first cross less suitable for the grid analysis of Fig. 2.12.



(a) Permittivity gradient $\mathbf{G}^{\varepsilon_\mathrm{r}}$.



(b) Conductivity gradient $\mathbf{G}^{\sigma_\mathrm{r}}$.

Figure 2.35: *Evolution of the ratio between the gradient computed by finite-differences and the adjoint gradient vs. finite-difference perturbation δ. The same parameterisation and frequencies as in Section 2.2 are used (7 frequencies between 50 and 200 MHz and $\sigma_o = \varepsilon_o\omega_o$).*

2.3.4 Sensitivity to dispersive parameters

In Chapter 1, I evacuated the consideration of the dispersive parameters for the inversion, arguing that the data were poorly sensitive to these parameters. Based on the previous discussion on data sensitivity to permittivity and conductivity, I can now justify this statement in more details by computing the diffracted fields corresponding to these parameters.

Fig. 2.36 present the amplitude of the fields diffracted by 10% anomalies of the real and imaginary parts of permittivity and conductivity, as well as of the magnetic permeability. These fields were computed in the same way as in Fig. 2.8, by computing the difference between the field in a homogeneous medium and the field in the same medium where an anomaly of a given parameter has been added (the other parameters remaining homogeneous).

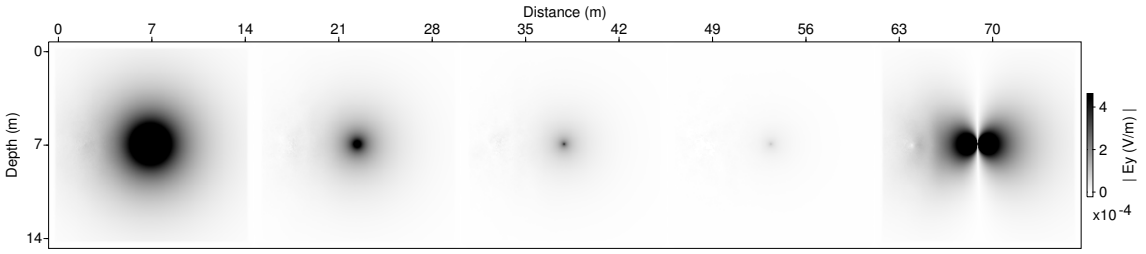


Figure 2.36: *Diffraction patterns of 10% anomalies of ε', σ', ε'', σ'' and μ (from the left to the right) at 50 MHz in TE mode. Background values are $\varepsilon'_r = 4$, $\sigma' = 3$ mS/m, $\varepsilon''_r = 0.4$, $\sigma'' = 0.3$ mS/m and $\mu_r = 1$ ($\simeq$ realistic values).*

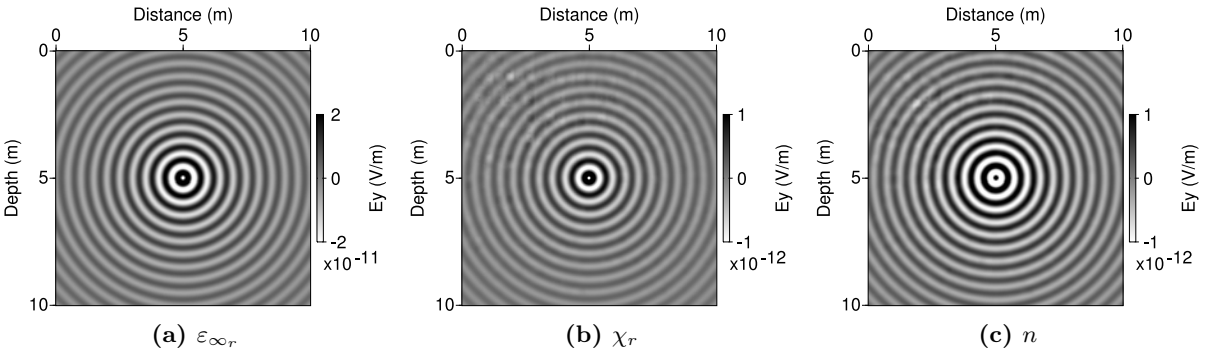
According to Fig. 2.36, the diffracted wavefields display a lower sensitivity to the dispersive parameters ε''_r and σ'' than to the real parts ε'_r and σ. Note the particular diffraction pattern of the magnetic permeability (right panel) which presents an amplitude comparable to the one of the real permittivity and a remarkable dependency to the diffraction angle which may help to mitigate the trade-off effects in multiparameter permittivity-permeability inversions (I might draw the parallel with the inversion of velocity and density in acoustic FWI).

Of course, one may argue that the strength of the diffraction patterns depends on the amplitude of the parameter anomaly. Depending on the goal of the survey, we may not be interested in detecting the same level of parameter perturbations: imaging 20% anomalies in permittivity might be relevant but concerning conductivity, we may be more interested in knowing the order of magnitude, whether it below 1 or more than 100 mS/m (to detect saline water in an aquifer, for instance). In this respect, a logarithmic parameterisation of the conductivity might be more relevant, as suggested by Meles (2011) and Yang et al. (2013). I should mention, however, that a logarithmic parameterisation does not solve the problem of the trade-off. On the contrary, over-estimating the logarithmic conductivity update has a greater impact on the conductivity itself since it acts directly on its order of magnitude.

Finally, the analysis based on the wave scattering by a diffracting point might be non-adapted to study the effect of dispersive parameters, that act on the waveforms when they propagate through extended anomalies. However, the relative sensitivities displayed in Fig. 2.36 are approximately representative of other sensitivity tests involving larger inclusions.

To complete the overview, Fig. 2.37 presents the fields diffracted by anomalies in Jonscher parameters. Again, the data are much less sensitive to the dispersive parameters χ_r and n than to the high-frequency permittivity ε_{∞_r}, by a factor of more than 20. Moreover, to the problem of sensitivity is added the trade-off issue since the difference of phase between the diffracted fields is less pronounced than in the permittivity-conductivity case of Fig. 2.8.

As a conclusion, considering Jonscher parameters — or generally dispersive parameters — amounts to add several poorly constrained parameters to the inversion. I suggest it should be viewed only in a second step, when satisfying models for permittivity and conductivity have



(a) ε_{∞_r} **(b)** χ_r **(c)** n

Figure 2.37: *Fields diffracted by 20% anomalies of the Jonscher parameters at 200 MHz. Background values are* $\varepsilon_{\infty_r} = 8.14$, $\chi_r = 0.94$ *and* $n = 0.82$. *Note the different amplitude scales.*

already been reconstructed. Of course, data sensitivity to these parameters is case-dependent and things might be different in strongly dispersive media where a non-dispersive parameterisation cannot account for the first-order characteristics of the data. Moreover, the reconstruction of dispersive parameters may be envisaged in 1D imaging, which involves fewer parameters and thus better constrains them.

Partial conclusions on the inverse problem

In this chapter, I have presented the FWI problem of GPR data, which is an ill-posed, non-linear problem due to the non-perfect illumination of the subsurface, to the well-known cycle-skipping effect, and to trade-offs between permittivity and conductivity. In addition, the desired 2D image being based on a pixel representation, it makes the problem too large to envisage a systematic search via global optimisation methods. Instead, the problem is linearised and solved iteratively using a local descent optimisation based on the quasi-Newton L-BFGS algorithm.

In spite of the consideration of an approximate inverse Hessian in the quasi-Newton update, the algorithm displays a strong sensitivity to parameter scaling, which suggests that the Hessian approximation is not accurate enough, or that the linearised approximation is not valid. Seeking for the resolution of the complete Newton system through, e.g., truncated Newton methods (Métivier et al., 2013), could enable to mitigate the sensitivity to parameter scaling, and to converge towards satisfying solutions without having to find an *ad hoc* tuning factor β. If not, an optimal parameterisation might be designed to minimise the trade-off between permittivity and conductivity, making the Hessian more diagonal. But this last possibility is still an open issue, since permittivity and conductivity are already physically decoupled parameters, whose influences on the data can be discriminated with regards to their frequency-dependency but not with respect to the illumination angle.

In this section, I have presented a rather classical FWI scheme, based on a ℓ_2-norm misfit function with a Tikhonov regularisation term, and on a linear combination of the physical parameters of interest for the definition of the optimised parameters. Each of these ingredients have many variants. As observables, we may look at only at the phase differences, or only at the amplitudes, via logarithmic residuals (Shin et al., 2007), or even at some attributes (Sassen

and Everett, 2009). As a measure of data misfit, we may use other norms. The ℓ_1-norm is generally more robust with respect to noise (Brossier et al., 2009) because it is less sensitive to high-amplitude outliers in the residuals (it will be used in this purpose in Section 3.2). General ℓ_p norms yield other sensitivities (Kim et al., 2012).

Finally, there are many ways to define the parameters to be optimised. Meles et al. (2010) and Yang et al. (2013) insist on the fact that using logarithmic parameters such as $\ln(\varepsilon/\varepsilon_o)$ and $\ln(\sigma/\sigma_o)$ enables to explore a wider range of magnitudes. More generally, the pixel-based representation of the model can be discussed. Since instabilities arise for parameters that are poorly constrained by the data, an ideal parameterisation should include only well constrained parameters. An optimal parameter set could consist not only in particular pixels in the image but in groups, or patterns of pixels, described through the decomposition of the image in a more suitable space (e.g. using the wavelet transform, Bouchedda et al., 2012; Abubakar et al., 2012, based on some image processing, Ma et al., 2012; Zhou et al., 2014).

In any case, the framework I have presented here is general enough to change any of these ingredients — observables, misfit criterion, parameters or constraints — and it provides guiding rules for computing the corresponding gradient. In Section 2.3.3, I offered a small insight of the flexibility of the formulation by defining another set of parameters corresponding to blocky zones of arbitrary shapes in the medium.

In the next chapter, I investigate how to apply the proposed imaging workflow to real data sets. FWI has already been proved successful for the interpretation of crosshole GPR data (e.g. Ernst et al., 2007; Belina et al., 2009; Klotzsche et al., 2010, 2013) but we expect that its application to on-ground GPR will be challenging due to the reduced illumination of the subsurface and of the lack of low-wavenumber coverage provided by surface-based measurements.

Chapter 3

Application to real data inversion

Contents

Introduction

For geophysicists, numerical methods find their ultimate purpose in the confrontation to real data. In this chapter, I consider two sets of electromagnetic data to test the imaging method proposed in the previous chapter and investigate whether FWI is applicable to on-ground GPR measurements and can bring valuable information about the subsurface.

In section 3.1, I first consider experimental data which were acquired in a well-controlled laboratory environment in the Institut Fresnel (Marseille, France). Knowing the targets used in the experiment and the surrounding medium enables to validate the imaging algorithm proposed in the previous chapter in a favourable situation with regards to starting model and source calibration. It is also the opportunity to test the accuracy of the numerical solutions with respect to physical data.

In section 3.2, I consider on-ground GPR data acquired in a gallery buried inside a limestone massif at the Low Noise Underground Laboratory (LSBB, Rustrel, France). Because of the complex geological environment, applying FWI to this data set is much more challenging. The preliminary study I present here aims mainly at designing an adequate initial model for starting the FWI and at estimating the source signature from the data. It requires first an understanding of the data, and hence of the model, through classical procedures such as velocity analysis and migration.

3.1 Validation of the imaging algorithm against experimental laboratory data

This section has been submitted to the journal *Near Surface Geophysics* (Lavoué et al., 2015) where it has recently been accepted for publication. I present here the revised version, except minor changes in layout, spelling and notations. This leads to some redundancies with previous sections on forward modelling and inverse problem formulation.

This study is a preliminary to the application of the imaging algorithm to real field data. It enables to validate the multi-parameter strategy proposed in (Lavoué et al., 2014, Section 2.2) in a simple case since the targets are located in a known homogeneous environment (namely air), which greatly simplifies the choice of a starting model and the estimation of the source signal. The paper also presents a new modelling feature in my implementations: in order to reduce the computation domain, an integral representation is used to inject the incident field in the zone of interest, and to propagate the scattered field back to receiver positions.

Frequency-domain modelling and inversion of electromagnetic data for 2D permittivity and conductivity imaging: An application to the Institut Fresnel experimental dataset
F. Lavoué, R. Brossier, L. Métivier, S. Garambois and J. Virieux
accepted in *Near Surface Geophysics*.

Abstract

The need for a quantitative imaging of the near subsurface leads to the development of inversion algorithms to infer ground properties from recorded data. The aim of this study is to validate an inversion method recently developed for the simultaneous imaging of dielectric permittivity and electrical conductivity from 2D ground-penetrating radar measurements. The validation is performed using electromagnetic data collected in a well-controlled laboratory environment. In this experiment, the knowledge of the targets enables a quality control of the inversion results. In addition, the free space environment and the measurement of the incident field simplify the choice of a starting model for the inversion, as well as the calibration of the data with respect to the source signature and to the geometrical spreading. To perform accurate and efficient forward simulations, we use a frequency-domain finite-difference scheme whose stencil coefficients can be optimised for each simulated frequency. As the objects of interest are locally concentrated at the centre of the acquisition array, it is possible to restrict the computation domain to a small region enclosing the targets using an integral representation of the analytical incident field coming from the sources and of the scattered field that we analytically propagate towards the receivers. An analysis of the numerical errors done on synthetic data shows that this strategy provides an error level that is low enough to not perturb the inversion, while dramatically decreasing the computational cost compared to a full-domain simulation. The monoparameter reconstruction of a purely dielectric target recovers permittivity values in very good agreement with the expected ones, as well as a very satisfying data fit. We also validate our strategy for multiparameter inversion on targets involving both a purely dielectric cylinder and a purely metallic copper tube, although the optimisation cannot recover the exact conductivity of copper.

Introduction

The quantitative imaging of the near subsurface through non-invasive prospecting techniques appears as a crucial challenge in many fields of application. Among the geophysical methods available to explore the subsurface, ground-penetrating radar (GPR) is particularly interesting for its high resolution properties, despite its depth penetration limits. In geological, hydrological and geotechnical investigations, GPR can provide a qualitative view of the geometry of the sounded medium as well as a quantitative information on the dielectric permittivity ε [F/m] and on the electrical conductivity σ [S/m] in the subsurface. These properties can in turn be interpreted in terms of the composition of the material (Deeds and Bradford, 2002; Ihamouten et al., 2012) or of its water content (Garambois et al., 2002; Huisman et al., 2003; Day-Lewis et al., 2005; Weihermüller et al., 2007).

Recently, Lavoué et al. (2013, 2014) proposed an imaging method based on the full waveform inversion (FWI) of GPR data for the reconstruction of permittivity and conductivity in 2D sections of the subsurface. FWI is a state-of-the-art quantitative imaging that aims at exploiting the whole information of the radargrams. Recent applications of FWI to GPR data have already been efficient for water content estimation in the first centimetres of agricultural soils (e.g. Lambot et al., 2006; Minet et al., 2010) and for the estimation of permittivity and conductivity in stratified structures such as concrete (Kalogeropoulos et al., 2011; Patriarca et al., 2011) or layered soils (Busch et al., 2012). FWI has also been applied to cross-hole radar data for the imaging of permittivity and conductivity in 2D and pseudo-3D (e.g. Ernst et al., 2007; Meles et al., 2010; Ellefsen et al., 2011; Klotzsche et al., 2013). In their previous contributions, Lavoué et al. (2013, 2014) perform numerical analysis on synthetic data to address the problem of multiparameter imaging of permittivity and conductivity by FWI in surface-to-surface acquisition configuration (on-ground GPR) that leads to a decreased illumination of the targets in the subsurface. To tackle the multiparameter FWI problem, they propose a joint optimisation in the permittivity-conductivity parameter space. In this approach, permittivity and conductivity models are updated using a quasi-Newton scheme that considers the effect of an approximated Hessian matrix on the descent direction. The Hessian matrix (second-order derivatives of the misfit function) is expected to be crucial for multiparameter problems as it accounts for parameter dimensionalities and for possible trade-offs between different parameter types (Pratt et al., 1998; Operto et al., 2013). The strategy of Lavoué et al. (2013, 2014) involves parameter scaling and regularisation factors to determine optimal solutions for the permittivity and conductivity distributions, based on the analysis of the data fit. In the present study, we would like to validate this strategy for multiparameter imaging by considering the inversion of well-controlled data acquired in a laboratory environment that provides a complete illumination of the targets.

The experimental scattering database collected by the Institut Fresnel (Marseille, France) is an interesting tool for testing and validating inversion algorithms on real physical data (see the dedicated special sections Belkebir and Saillard, 2001, 2005; Litman and Crocco, 2009). In this work, we focus on a dataset acquired on inhomogeneous targets (Belkebir and Saillard, 2005), with a particular interest for those involving both dielectric and metallic objects which enable to address the problem of multiparameter reconstruction. The targets are cylinders located in free space (i.e., surrounded by air) and elongated perpendicularly to the observation plane, so that the problem can be considered to be 2D (see Geffrin et al., 2005, for the description of the experimental setup). The use of such experimental data greatly simplifies some issues

that usually constitute major obstacles in FWI applications. First, the free space environment eliminates the crucial question of designing a suitable initial model. Second, the measurement of the incident field (in the absence of the targets) enables to accurately characterise the source signature and to calibrate the recorded data.

The present paper is organised in two parts, related to the two components of our imaging algorithm, namely the forward and the inverse problems. A preliminary but important aspect of the validation task concerns the design of the forward modelling that should accurately and efficiently reproduce the observed data. Our frequency-domain finite-difference (FDFD) modelling uses a mixed-grid stencil with optimised coefficients (Jo et al., 1996; Hustedt et al., 2004) that can simulate the total field in the entire domain without significant errors due to numerical dispersion and numerical anisotropy, but it implies a significant computational effort. In a free space environment, it is highly desirable to restrict the computational domain to a small region enclosing the targets. Following the work of Wilcox and Velichko (2010), we use an integral representation of the fields to inject an analytical incident field in the reduced computation domain where FDFD computations are performed. A similar integral can be used to propagate analytically the recorded scattered field back to the real receiver locations. However, the discretisation of the integral leads to numerical errors whose magnitude must be compared with the dispersion errors of the pure numerical approach to assess the advantage of one strategy over the other. In a second part of the paper, we challenge our imaging algorithm by reconstructing first a purely dielectric target (monoparameter inversion) and secondly a target containing both dielectric and metallic objects (multiparameter inversion).

3.1.1 Presentation of the data (Institut Fresnel, Marseille, France)

Fig. 3.1(a) shows the configuration of the experiment in the (xy)-plane (see Geffrin et al., 2005, their fig. 1, for a 3D representation). The targets, located at the center of the acquisition array, consist of cylinders elongated in the z-direction. Several experiments were performed

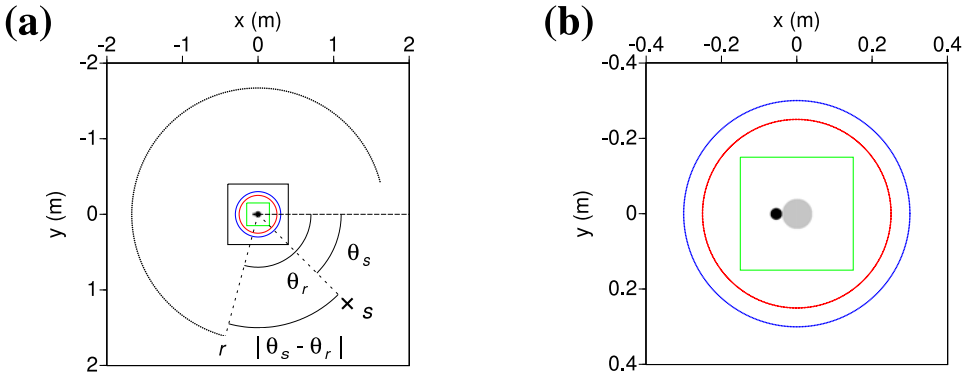


Figure 3.1: *(a) Acquisition setup and computation domain. The cross located at $\theta_s = 45^o$ indicates a source and the dots between $\theta_r = 105^o$ and $\theta_r = 345^o$ the corresponding receiver locations. The 80-cm-by-80-cm black box corresponds to the reduced domain for FDFD computation. (b) Zoom on the computation domain. The 30-cm-by-30-cm green box delimits the zone reconstructed during the inversion. The red circle indicates the loading contour (393 virtual sources) and the blue one the monitoring contour (472 virtual receivers).*

for different targets (Geffrin et al., 2005). For each experiment, the transmitting and receiving antennas are moved on a circular frame of radius $d_e = 1.67$ m around the targets. Fig. 3.1(a) exemplifies one source location (black cross) and the corresponding receiver locations (black dots), sweeping successively every 1^o. Following Geffrin et al. (2005), we denote θ_s and θ_r the angular positions of sources and receivers, respectively, and we call *angle between source and receiver* the angle $|\theta_s - \theta_r|$ formed by the source, the target and the receiver (see Fig. 3.1a). During the experiment, the array is rotated to acquire data at different source positions, the angle between source and receiver always being comprised between 60^o and 300^o. For each source and receiver position, two measurements are performed: one for the incident field u_{inc} in free space (without the target) and one for the total field u_{tot} (with the target), enabling to deduce the field scattered by the targets $u_{sc} = u_{tot} - u_{inc}$. Finally, measurements for the two modes TE and TM are available. In this study, we consider only the measurements performed in TM mode which involves the electric field component E_z polarized perpendicularly to the observation plane (xy). Please note that we follow the same terminology as Geffrin et al. (2005) by adopting the convention of the electromagnetic community for the TM mode (it is the TM_z mode defined in Taflove and Hagness, 2005, p. 55). It corresponds to the TE mode for the geophysical community (see e.g. Jol, 2009, fig. 1.8, p. 13).

Fig. 3.2 shows synthetic models corresponding to the targets we will consider in the following, labelled as FoamDielExt (Fig. 3.2a) and FoamMetExt (Fig. 3.2b,c) by Geffrin et al. (2005). The target FoamDielExt consists of two purely dielectric cylinders made of foam ($\varepsilon_r = 1.45 \pm 0.15$) and of plastic ($\varepsilon_r = 3 \pm 0.3$), of diameters 8 cm and 3.1 cm, respectively. Their conductivity is considered to be zero. Geffrin et al. (2005) perform numerical modelling on these synthetic models using a method of moments and domain integral formulations, so that we can compare our simulation results. Regarding the inversion, we will use the experimental dataset FoamDielExt in order to validate the monoparameter reconstruction of permittivity. The target FoamMetExt consists of the same purely dielectric foam cylinder ($\varepsilon_r = 1.45 \pm 0.15$) and of a purely metallic copper tube ($\varepsilon = 1$, $\sigma \gg 1$ S/m) of diameter 2.85 cm and of thickness 2 mm. We will use this dataset to challenge our strategy for multiparameter inversion. Of course, we cannot hope to reconstruct the exact amplitude of the conductivity anomaly (which is around 10^6 S/m for pure copper). But we can analyse the trade-off effect between permittivity and conductivity, i.e. the trend to recover erroneous non-zero conductivity values for the foam cylinder, or permittivity values $\varepsilon_r \neq 1$ for the copper.

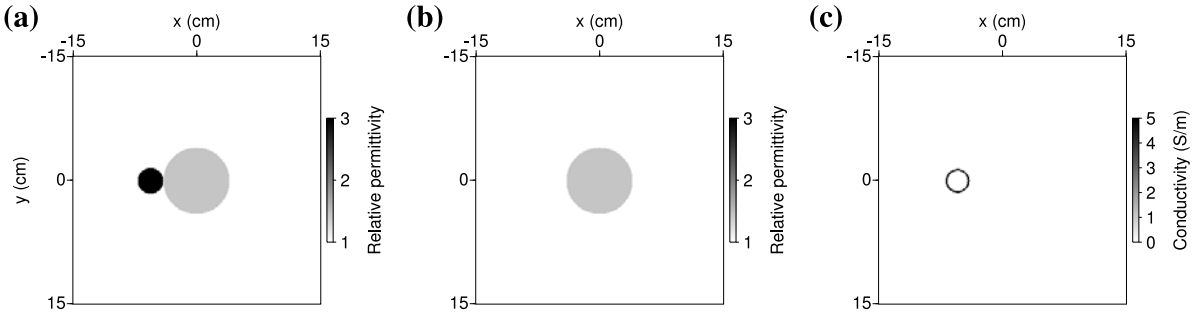


Figure 3.2: *Synthetic models corresponding to the experimental targets. (a) FoamDielExt (permittivity model). (b,c) FoamMetExt (b, permittivity; c, conductivity). In (a) and (b), the central foam cylinder has a permittivity* $\varepsilon_r = 1.45$.

Table 3.1 sums up the measurement properties. For the target FoamDielExt, measurements are performed for 9 frequencies between 2 and 10 GHz, and repeated over 8 source positions with an interval of 45^o. For the target FoamMetExt, measurements are performed for 17 frequencies between 2 and 18 GHz, and repeated over 18 source positions with an interval of 20^o. In the latter case, the extension of the frequency bandwidth and the refinement of the source sampling aim at better constraining the multiparameter inversion for the reconstruction of the complex target FoamMetExt (Geffrin et al., 2005).

Dataset	**Frequencies**			**Sources**		**Receivers**	
	Nb	Min-max	Interval	Nb	Interval	Nb	Interval
FoamDielExt	9	2-10 GHz	1 GHz	8	45°	241	1°
FoamMetExt	17	2-18 GHz	1 GHz	18	20°	241	1°

Table 3.1: *Measurements properties (after Geffrin et al., 2005, their table 1).*

In the frequency range we consider, target sizes are of the order of — or even below — the smallest wavelength in free space, while the propagation distance goes beyond 100 wavelengths (see Table 3.2). Two aspects may require our attention. First, we cannot expect to recover the exact thickness of the copper tube which is of the order of $\lambda/8$ for the highest available frequency, given that the maximal resolution of our imaging algorithm, based on the diffraction principle, is of the order of $\lambda/2$. Secondly, we can expect to face a numerical challenge to simulate the wave propagation over 100 wavelengths between the target and the receivers (even over 200 wavelengths on the path source-target-receiver).

Objects/lengths	**(cm)**	**2 GHz** $\lambda = 15$ cm	**10 GHz** $\lambda = 3$ cm	**18 GHz** $\lambda = 1.7$ cm
Grid step h	0.1	$\lambda/150$	$\lambda/30$	$\lambda/16$
Copper thickness	0.2	$\lambda/75$	$\lambda/15$	$\lambda/8$
Copper diameter	2.85	$\simeq \lambda/5$	$\simeq \lambda$	$\simeq 1.8\lambda$
Plastic diameter	3.1			
Foam diameter	8.0	$\lambda/2$	2.7λ	5λ
Distance target-receiver	167	11λ	56λ	100λ
Dist. source-target-receiver	334	22λ	112λ	200λ

Table 3.2: *Target sizes and propagation distances in terms of wavelengths λ propagated in free space.*

3.1.2 Forward problem

3.1.2.1 Numerical strategy

A first challenge consists in accurately and efficiently performing synthetic simulations to be compared with experimental data in the inversion process. In this study, we use a FDFD scheme based on the optimised mixed-grid stencil of Hustedt et al. (2004), which leads to the linear system

$$\mathcal{A}(\omega, \mathbf{m})\, \mathbf{u}(\omega) = \mathbf{S}(\omega), \tag{3.1}$$

where ω is the simulated angular frequency (in rad/s), $\mathbf{m}$ represents the physical model, $\mathcal{A}$ is the impedance matrix resulting from the FDFD scheme, $\mathbf{u}$ is the simulated wavefield (i.e. the

component E_z in TM mode) and $\mathbf{S}$ denotes the source term. We solve the linear system (3.1) for each frequency by LU factorization using the direct solver MUMPS (MUMPS-team, 2011).

In the finite-difference scheme, the optimisation of the weighting coefficients of the stencil for a given λ/h ratio (with h the grid step) enables to minimize the errors due to numerical dispersion and numerical anisotropy (Jo et al., 1996; Hustedt et al., 2004). It is thus possible to achieve an error of less than 1.5% on the incident field recorded by the receivers. However, the computation in the entire domain of Fig. 3.1(a) implies a significant computational effort, both in terms of CPU time and of memory requirement, especially because the large number of degrees of freedom requires the use of double precision arithmetic to avoid instabilities in MUMPS. In view of the inversion, the efficiency of the forward problem is crucial. In a free space environment, it is highly desirable to restrict the computational domain to a small region enclosing the target, as shown in Fig. 3.1(b).

To perform the inversion in this reduced domain, we must be able:

1. to inject in the domain an arbitrary incoming wavefield emitted from a remote source of the real acquisition array located outside the domain,
2. to propagate analytically the field scattered by the target and computed numerically in the reduced domain back to the real receiver locations outside the domain.

To do so, we use an integral representation of the fields, following the work of Wilcox and Velichko (2010) and Velichko and Wilcox (2010). Integral representations are usual in non-destructive testing applications (NDT). For instance, it has been recently used by Zhao et al. (2013) for GPR data redatuming in a tunnel grouting test context. In our case, it is particularly suited because we know the exact analytical solution for the wave propagation in free space between the acquisition array and the target region. The Helmholtz-Kirchhoff integral provides an expression for the scalar wavefield u at any point $\mathbf{r}$ within a closed contour $\mathcal{C}$, as a function of an arbitrary field u_{inc} on the boundary:

$$u(\mathbf{r}) = \int_{\mathcal{C}} \left[G(\mathbf{r},\mathbf{s})\, \nabla u_{inc}(\mathbf{s}) - u_{inc}(\mathbf{s})\, \nabla G(\mathbf{r},\mathbf{s}) \right] \cdot \mathbf{n} dS, \tag{3.2}$$

where $\mathbf{n}$ is the outward normal-contour vector and $G(\mathbf{r},\mathbf{s})$ denotes the Green function, i.e. the field value at point $\mathbf{r}$, resulting from a source located at $\mathbf{s}$. The vector $\mathbf{r}$ locates a point inside the contour $\mathcal{C}$, while the vector $\mathbf{s}$ points on the contour. In other words, we can deduce the field u inside the contour from the superposition of fields generated by monopole sources (G terms) and dipole sources (∇G terms) located on the contour. In the following, we will refer to these sources as *virtual* sources (in opposition with real ones, located on the acquisition array outside the computation domain, see Fig. 3.1a). In expression (3.2), the Green function $G(\mathbf{r},\mathbf{s})$ is computed numerically with the FDFD scheme. The values of the incident field on the contour $u_{inc}(\mathbf{s})$ are known analytically as the solution of 2D wave propagation in free space between the real and the virtual source arrays (see e.g. Taflove and Hagness, 2005, §5.3.1, p.172).

A similar integral representation can be used to describe the scattered field $u_{sc}(\mathbf{r})$ outside the contour $\mathcal{C}$ if the scattered field on the contour $u_{sc}(\mathbf{s})$ is known (Velichko and Wilcox, 2010):

$$u_{sc}(\mathbf{r}) = -\int_{\mathcal{C}} \left[G(\mathbf{r},\mathbf{s})\, \nabla u_{sc}(\mathbf{s}) - u_{sc}(\mathbf{s})\, \nabla G(\mathbf{r},\mathbf{s}) \right] \cdot \mathbf{n} dS, \tag{3.3}$$

where the vector $\mathbf{r}$ locates now a point outside the contour. In eq. (3.3), the scattered field $u_{sc}(\mathbf{s})$ on the contour is computed numerically with the FDFD scheme (it corresponds in fact

to the field $u(\mathbf{r})$ in eq. 3.2), while the propagator $G(\mathbf{r},\mathbf{s})$ and its gradient across the contour $\nabla G(\mathbf{r},\mathbf{s})$ are computed analytically. The minus sign before the integral is due to the orientation convention of the outward-pointing vector $\mathbf{n}$.

We use the windowed Sinc interpolation proposed by Hicks (2002) to locate accurately the virtual sources on the integration contour, that does not coincide with the finite-difference grid. This interpolation consists in spreading the source excitation on several grid nodes around the true source location to mimic a bandlimited version of a monopole or dipole point source. Because the implementation of point sources creates singularities in the numerical solution, it is not possible to record the scattered field $u_{sc}(\mathbf{s})$ of eq. (3.3) in the neighbourhood of nodes where the incident field $u_{inc}(\mathbf{s})$ is injected (eq. 3.2). Following Velichko and Wilcox (2010), we therefore implement two distinct contours: a *loading* contour $\mathcal{C}_l$ to inject the incident field and a *monitoring* contour $\mathcal{C}_m$, located outside the loading contour, to record the scattered field and apply eq. (3.3) to propagate it towards the receiver in the far-field region (see Fig. 3.1).

The discretisation of the loading contour $\mathcal{C}_l$ in N_l segments of length $\delta\mathcal{C}$ then leads eq. (3.2) to become

$$u(\mathbf{r}) = \delta\mathcal{C}\sum_{i=1}^{N_l} \quad \left[\cos(\theta_i)\frac{\partial u_{inc}}{\partial x}(\mathbf{s}_i) + \sin(\theta_i)\frac{\partial u_{inc}}{\partial z}(\mathbf{s}_i)\right] G(\mathbf{r},\mathbf{s}_i) \quad \text{(monopole)}$$
$$-u_{inc}(\mathbf{s}_i)\left[\cos(\theta_i)D_x + \sin(\theta_i)D_z\right] G(\mathbf{r},\mathbf{s}_i), \qquad \text{(dipole)} \tag{3.4}$$

where the operators D_x and D_z correspond to x- and z-oriented dipoles using the windowed Sinc interpolation and θ_i is the angular position of the i^{th} virtual source. The partial derivatives of the incident field u_{inc} are computed analytically.

Conversely, when implementing eq. (3.3), the windowed Sinc interpolation is used to extract values of the scattered field at virtual receivers on the monitoring contour $\mathcal{C}_m$ and the derivatives of the propagator G are computed analytically:

$$u_{sc}(\mathbf{r}) = -\delta\mathcal{C}\sum_{i=1}^{N_m} \qquad \left[\cos(\theta_i)D_x + \sin(\theta_i)D_z\right] u_{inc}(\mathbf{s}_i)\, G(\mathbf{r},\mathbf{s}_i) \qquad \text{(monopole)}$$
$$-u_{inc}(\mathbf{s}_i)\left[\cos(\theta_i)\frac{\partial G}{\partial x}(\mathbf{r},\mathbf{s}_i) + \sin(\theta_i)\frac{\partial G}{\partial z}(\mathbf{r},\mathbf{s}_i)\right]. \quad \text{(dipole)} \tag{3.5}$$

Figs 3.3(a) and (c) show the incident fields resulting from eq. (3.4) at frequencies 2 and 18 GHz respectively, using a contour element $\delta\mathcal{C} = 4$ mm ($= \lambda/4$ at 18 GHz). Perfectly Matched Layers (PMLs, Bérenger, 1994) are used to absorb the waves on the edges of the computation domain (not shown). In Figs 3.3(b) and (d), the relative errors against analytical solution are shown. Using the integral representation to inject the incident field in the reduced domain, we achieve an error level which is slightly higher than when computing the field in the entire domain ($\simeq 0.5\%$ vs. $\simeq 0.2\%$ respectively), but it still remains of the same order of magnitude and we consider it as satisfactory.

An important property of the integral representation, eq. (3.2), is that the incident field synthesized by the virtual array exists only inside the contour $\mathcal{C}$, where eq. (3.2) is valid (see Fig. 3.3). However, if a scatterer is placed inside the contour, then the scattered field generated by the incident field of eq. (3.2) exists both inside and outside the contour (because the scatterer acts as an internal source). This property is illustrated in Fig. 3.4, where we inject the incident

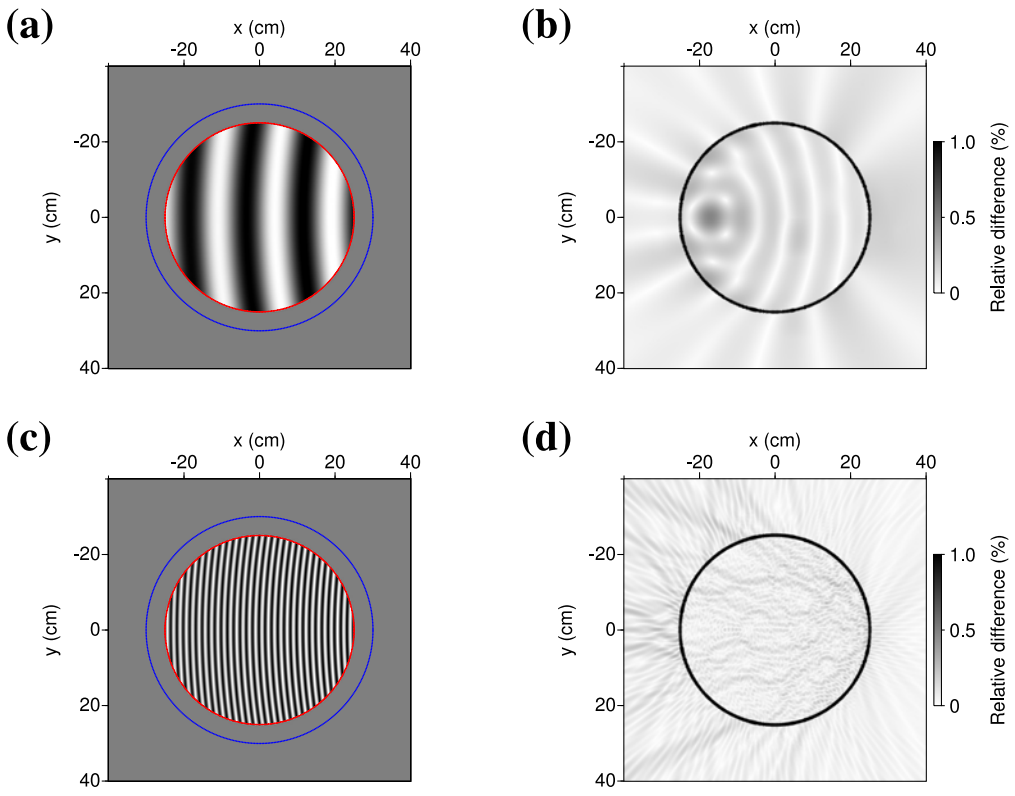


Figure 3.3: *(a,c) Wavefields emitted by the virtual source array (loading contour, in red). The blue circle indicates the monitoring contour. (b,d) Relative difference with the corresponding analytical wavefields. (a,b) 2 GHz, (c,d) 18 GHz. Excepted at nodes where sources are injected, the errors are below 0.47% at 2 GHz (b) and below 0.29% at 18 GHz (d) (to be compared with the errors done when performing the FDFD computation in the entire domain, which are below 0.19% and 0.14%, respectively). Note that the error ϵ is not computed in the same way inside and outside the loading contour: inside the contour, $\epsilon = |u_{numerical} - u_{analytical}|/|u_{analytical}|$, and outside, $\epsilon = |u_{numerical}|/|u_{analytical}|$ (because $u_{numerical}$ should be zero).*

field of Fig. 3.3(a) in a domain that contains the target FoamDielExt of Fig. 3.2(a). The scattered field is then analytically propagated using eq. (3.3) from the monitoring contour (blue circle in Fig. 3.4) up to the receiver positions on the acquisition array (see Fig. 3.1a). Note that if the monitoring contour were placed inside the loading contour, recording the total field, we would still propagate only the scattered field outwards the contour using eq. (3.3) (this is a property of the Helmholtz-Kirchhoff integral). We place the monitoring contour outside the loading contour to avoid the cumulative errors of both integrals on the propagation of the incident field.

In order to evaluate the numerical error done on the scattered field computed in the reduced domain and propagated towards the receivers, we perform the following simple test: We put a point source at the center of the acquisition frame (instead of the target), we record the resulting field on the monitoring contour, and we propagate it analytically towards the receivers on the acquisition array. We then compare the values obtained at receivers with the analytical solution and we find an error below 0.3%. Note that, when an incident field is injected on the loading

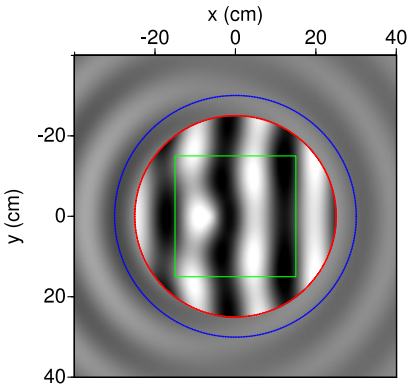


Figure 3.4: *Example of wavefield scattered by the synthetic target FoamDielExt of Fig. 3.2(a), illuminated by the incident field of Fig. 3.3(a), at 2 GHz. Inside the loading contour (red circle), the total field is computed, whereas only the diffracted field propagates outside and is recorded by the monitoring contour (blue circle).*

contour and impinges a target, the error on the scattered field resulting from eq. (3.3) should also include the error done on the injected incident field ($\simeq 0.5\%$ after Fig. 3.3). So we can estimate the global error done on the scattered field to approximately 0.8%. To compare this error with the full-domain case, we have to consider that, when computing the scattered field in the entire domain, the solution suffers from errors due to numerical dispersion and numerical anisotropy on the entire travel path source-target-receiver (of length $\simeq 2d_e$). As a result, in the full-domain case, we consider that the error on the incident data recorded in transmission configuration (that we can compute against analytical solution) is also representative for the error done on the scattered data.

These results are reported in Table 3.3, together with the associated computational requirements. Regarding the errors done on the incident and scattered fields, the restriction of the computation in the reduced domain using the integral representation appears as a satisfying alternative to the full-domain computation, decreasing the computational time by a factor of more than 20, while dividing the memory requirement of the LU factorization by a factor of 40.

	Full domain	**Reduced domain**
Number of grid points (including PMLs)	4051×4051	851×851
Elapsed time for LU factorization	71 s	0.67 s
Time for FoamMetExt simulation (18 src, 17 freq.)	2100 s	98 s
MUMPS memory requirement (LU)	59 Go	1.5 Go
Error on the incident field in the reduced domain	$< 0.2\%$	$< 0.5\%$
Error on the scattered field at receiver positions	$< 1.5\%$	$< 0.3\%$ (0.8%)

Table 3.3: *Computation cost and accuracy of full-domain vs. reduced-domain FDFD simulations. The LU factorization and the resolution of the linear system are performed by MUMPS in parallel on 16 MPI processes. We report elapsed times using one node of 16 cores of the Froggy cluster from the CIMENT platform (Univ. Grenoble). Note that the full-domain computation is performed with double-precision arithmetic while the reduced-domain computation uses single precision to achieve the level of accuracy indicated in the table.*

3.1.2.2 Data pre-processing

To be able to compare our synthetic data with the observed ones, we have to consider that observed and calculated data differ by three major aspects:

1. Observed data are 3D whereas our modelling is 2D.
2. The antennas used in the experiment have a given frequency signature, whereas the simulated source is a Dirac in time (unity source in the frequency domain).
3. The experimental setup use ridged horn antennas with a given radiation pattern whereas the modelling assumes an elementary dipole oriented in the z-direction which radiates isotropically in the (xy)-plane.

These characteristics have to be accounted for to convert the observed data in a value that we can compare with the synthetic one.

Usually, the estimation of the source complex spectrum is performed as a part of the frequency-domain inversion process by linear estimation (Pratt, 1999). In our case, however, it is highly beneficial to apply this estimation to the incident field measured in free space, which enables us both to characterize the antennas and to perform a simple 3D-to-2D conversion. To do so, let us write the observed data d_{inc}^{obs} resulting from the measurements of the incident field, for each angular frequency ω, each source located at angle θ_s, and each receiver at θ_r, as

$$d_{inc}^{obs}(\omega, \theta_s, \theta_r) = S_{obs}(\omega)\, R_{obs}(\omega, \theta_s - \theta_r)\, G_{3D}(\omega, d_{sr}), \tag{3.6}$$

where $S_{obs}(\omega)$ denotes the source complex spectrum, $R_{obs}(\omega, \theta_s - \theta_r)$ accounts for the spatial radiation pattern of the antennas, and G_{3D} is the Green function of electromagnetic wave propagation in 3D, that varies with the distance d_{sr} travelled by the wave from source to receiver.

If we adopt the convention that, for each angular frequency ω, the radiation pattern coefficients are 1 in pure transmission regime (i.e., $R_{obs}(\omega, \theta_o) = 1$ for an angle between source and receiver of $\theta_o = |\theta_r - \theta_r| = 180^o$), then we can deduce the source spectrum of the antennas from the incident field recorded in transmission configuration:

$$S_{obs}(\omega) = \frac{d_{inc}^{obs}(\omega, \theta_s, \theta_r = \theta_s - \theta_o)}{G_{3D}(\omega, d_{rs} = 2d_e)}. \tag{3.7}$$

Fig. 3.5 shows the obtained source spectrum, both in amplitude and phase. Computing the spectrum for all source positions, we find that its variations with respect to source location θ_s are negligible (of the order of 0.01%), showing the quality of these highly reproducible data.

Once we know the spectrum $S_{obs}(\omega)$ from the measurements in pure transmission configuration, we could deduce the radiation pattern of the antennas from the incident field recorded at the other incidence angles. However, we do not need it for our present study because we only consider the scattered field for the inversion. As the scattered field results from an incident field impinging the target at zero incidence (from the source point of view), and as it is itself recorded at zero incidence by the receiver, we can reasonably neglect the effect of the radiation pattern on the scattered data.

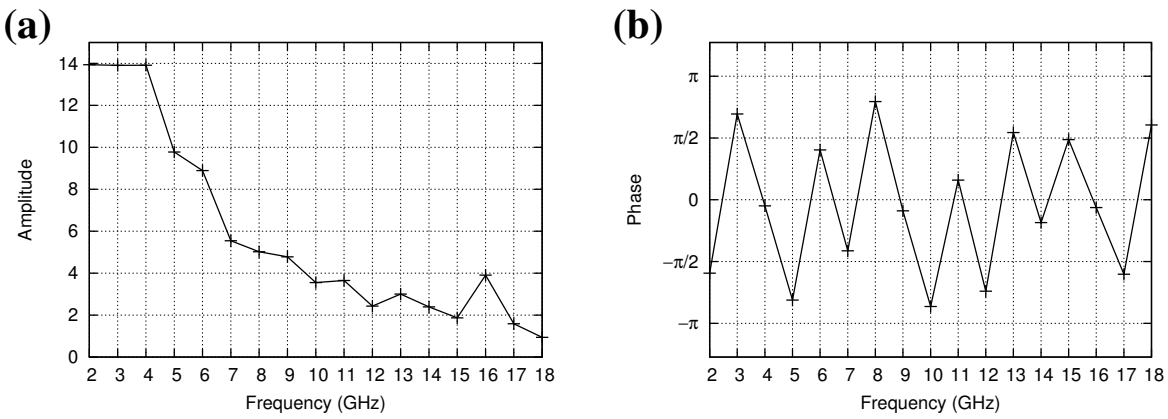


Figure 3.5: *(a) Amplitude and (b) phase of the antenna complex spectrum $S_{obs}(\omega)$.*

A last step to make the observed and calculated fields comparable is to correct for the 3D geometrical expansion, in order to build *2D observed* data. By analogy with expression (3.6), we assume that the observed and synthetic scattered data can be written as

$$d_{sc}^{obs}(\omega, \theta_s, \theta_r) = S_{obs}(\omega)\, T(\omega, \theta_s, \theta_r)\, G_{3D}(\omega, d_{str}), \tag{3.8}$$

$$d_{sc}^{cal}(\omega, \theta_s, \theta_r) = S_{cal}(\omega)\, T(\omega, \theta_s, \theta_r)\, G_{2D}(\omega, d_{str}), \tag{3.9}$$

where $S_{cal}(\omega)$ is a known factor depending on our FDFD implementation, G_{2D} is the 2D Green function which is known analytically, d_{str} is the distance travelled by the scattered wave from source to target and from target to receiver, which can be approximated by $d_{str} \simeq 2d_e$, and T is the response of the targets, which can be considered identical in 2D and 3D since the targets have a 2D geometry. Thus, the transformation to be applied to the observed data to compare them to the synthetic ones is

$$d_{sc}^{obs,\,2D}(\omega, \theta_s, \theta_r) = \frac{S_{cal}(\omega)}{S_{obs}(\omega)} \frac{G_{2D}(\omega, 2d_e)}{G_{3D}(\omega, 2d_e)}\, d_{sc}^{obs}(\omega, \theta_s, \theta_r), \tag{3.10}$$

where the 3D-to-2D conversion factor $k_{3D\to 2D} = G_{2D}/G_{3D}$ can be expressed by using the far-field approximations of the Green functions (see Taflove and Hagness, 2005, §8.2.2, p.332),

$$k_{3D\to 2D} = 2\sqrt{\frac{\pi d_e v_o}{\omega}}\, e^{\imath\pi/4}, \tag{3.11}$$

with v_o the velocity of light in free space and $\imath$ the imaginary unit.

3.1.2.3 Simulation of synthetic data

Simulation results performed in the synthetic models of Fig. 3.2 are shown in Figs 3.6 and 3.7 for the targets FoamDielExt and FoamMetExt, respectively. We choose the same frequencies and source positions as shown by Geffrin et al. (2005, their figs. 8 and 11) such that we can compare our results. In view of the inversion, we also perform numerical simulations in modified versions of the synthetic models of Fig. 3.2, in order to investigate the sensitivity of the data to model variations.

For the target FoamDieExt, our simulation results are very similar to those of Geffrin et al. (2005). As suggested by these authors, we also perform a simulation in a modified version of the model shown in Fig. 3.2(a), where we translated the targets by $dx = 2$ mm and $dy = 1$ mm, and we replaced the permittivity value of the plastic cylinder by 3.3 instead of 3 to obtain a better match (blue dashed line in Fig. 3.6).

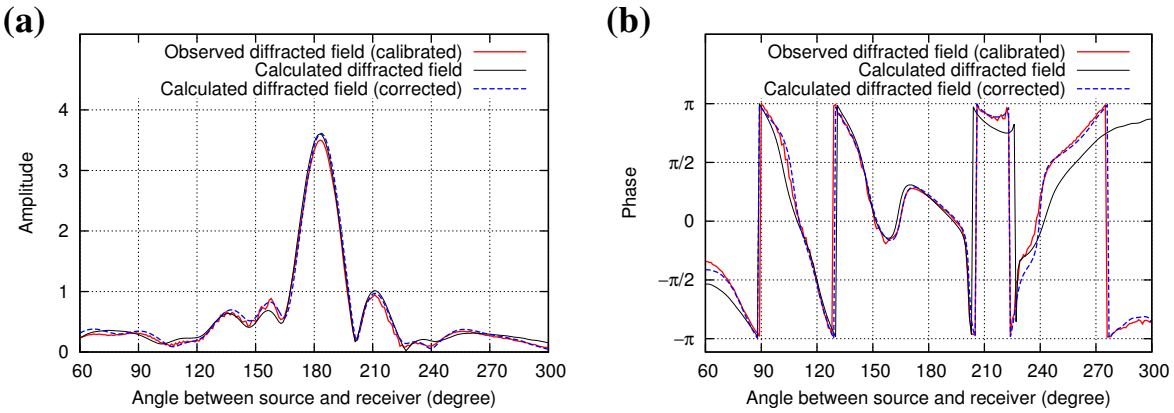


Figure 3.6: *(a) Amplitude and (b) phase of synthetic vs. observed data for the target FoamDielExt, for frequency 8 GHz and the source at $\theta_s = 270^o$. The so-called corrected field (blue dashed line) corresponds to the synthetic field computed in a modified version of the model shown in Fig. 3.2(a) where we translated the targets by $dx = 2$ mm and $dy = 1$ mm and we replaced the permittivity value of the plastic cylinder by 3.3 instead of 3. The match between the observed diffracted field and the synthetic one is similar as in Geffrin et al. (2005, their fig. 8).*

For the target FoamMetExt (Figs 3.2b,c), simulation results at 18 GHz are shown in Fig. 3.7 (black line). In this case, our simulation results differ slightly from those presented by Geffrin et al. (2005, their fig. 11): Contrary to these authors, who calibrate the amplitude of their numerical solution such as to match the central diffraction peak (at 180^o), we do not recover the same amplitude for the observed and for the synthetic data at 180^o. Nonetheless, we obtain a similar match as Geffrin et al. (2005) for the other angles. These differences in amplitude with respect to angle are probably due to our 3D-to-2D conversion which is based on the incident field in free space (without the target): It may be thus poorly valid to convert the signal recorded at 180^o, which is transmitted through the target. Again, we perform simulations in a modified version of the FoamMetExt models of Figs 3.2(b,c), where we replace the conductivity value of the copper tube by 100 S/m instead of 5 S/m (blue dashed line). Doing so, we obtain a better match at extreme angles. It shows that the data are well sensitive to conductivity variations in the copper tube and thus suggests that a quantitative reconstruction can be attempted, even if we cannot expect to recover the true value of conductivity in copper ($\simeq 10^6$ S/m). As another sensitivity test, we also compared the synthetic data of Fig 3.7 with data computed in a model where we replaced the empty copper tube of Fig. 3.2(c) by a full copper cylinder of same diameter and properties. An important result is that the full cylinder provides exactly the same scattered field as the empty tube when a high conductivity value is used for copper ($\sigma = 100$ S/m), suggesting that any information inside the real tube will be out of reach. Assuming a lower (non-physical) conductivity value for copper ($\sigma = 5$ S/m), data are slightly sensitive to the filling of the tube, which will have consequences in the reconstructions.

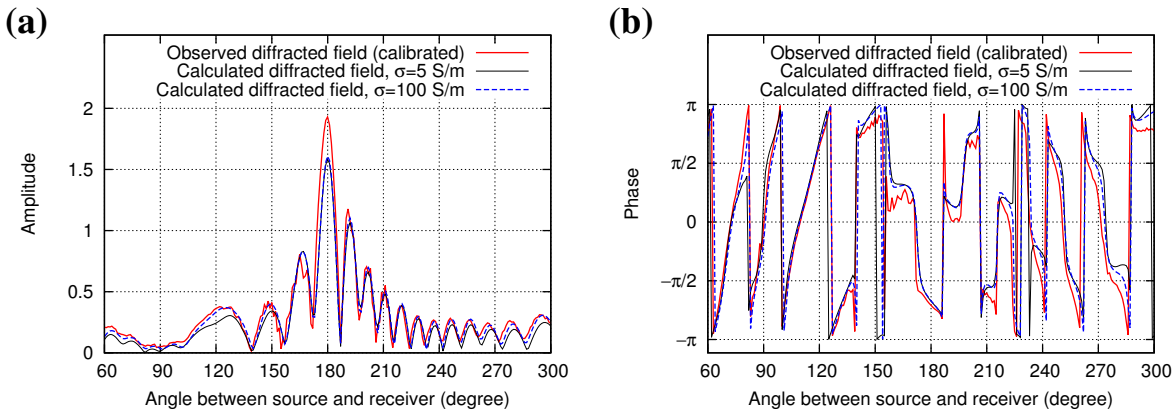


Figure 3.7: *(a) Amplitude and (b) phase of synthetic vs. observed data for the target FoamMetExt, for frequency 18 GHz and the source at $\theta_s = 240^o$. The match obtained with $\sigma = 5$ S/m for copper (black line) is similar as in Geffrin et al. (2005, their fig. 11) for extreme angles but we do not recover the same amplitude as these authors at 180^o. A better match is obtained with $\sigma = 100$ S/m (blue dashed line).*

In Fig. 3.8, we compare in the case FoamDielExt the magnitude of the data residuals, i.e. the difference between observed and synthetic data, with the numerical errors, computed as the difference between the analytical solution of 2D wave propagation in free space and the synthetic incident field computed in the entire domain. According to the numerical analysis of the previous section, this error is an upper bound for the error done on the synthetic scattered field that will be involved in the inversion (see Table 3.3). For information, we also indicate the magnitude of the total and scattered fields: As underlined by Geffrin et al. (2005), the high ratio between the magnitude of the scattered and total fields makes the need for an accurate modelling critical. In Fig. 3.8, the error level is significantly lower than the data residuals for most of the source-receiver pairs (it is also true for other sources and frequencies, not

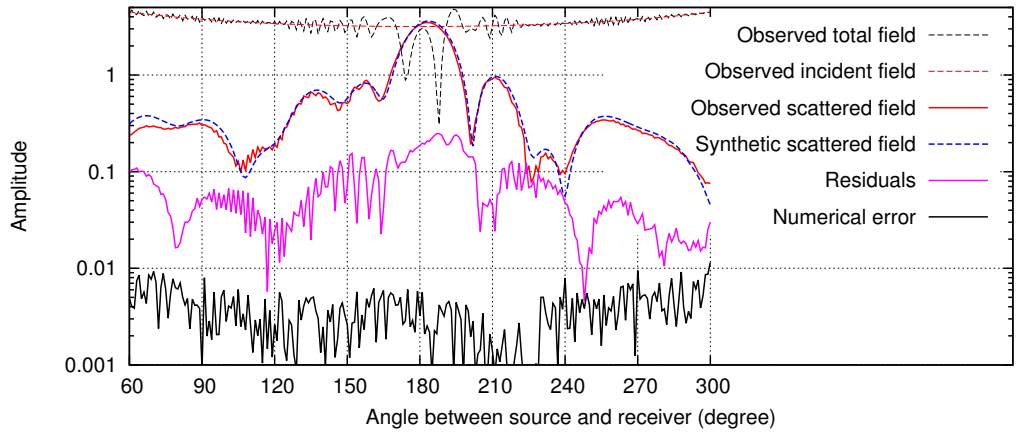


Figure 3.8: *Comparison of the magnitude of the total and scattered fields vs. residuals and numerical errors (frequency 8 GHz and source at $\theta = 270^o$). The synthetic scattered field is computed in the modified version of Fig. 3.2(a). The numerical error is obtained by difference between the analytical solution and the total incident field computed in the entire domain.*

shown in Fig. 3.8), which means that most of the residuals are due to differences between the synthetic model and the reality. This result suggests that our modelling tool fills the accuracy requirement: The inversion should not be affected by the numerical errors. The very good agreement between observed and synthetic data in Fig. 3.6, as well as the low error level compared to the residuals (Fig. 3.8), allow us to envisage the inversion of these data with some confidence.

3.1.3 Data inversion

3.1.3.1 Inverse problem formulation

Here we recall the basic ingredients of the method proposed by Lavoué et al. (2014). At each iteration k of the optimisation process, we aim at minimizing a misfit function that we define as the sum over frequencies of a data part $\mathscr{C}_D$ and of a model part $\mathscr{C}_M$:

$$\mathscr{C}(\mathbf{m}^k) = \mathscr{C}_D(\mathbf{m}^k) + \lambda \mathscr{C}_M(\mathbf{m}^k), \tag{3.12}$$

$$= \frac{1}{2}\sum_{\omega} \Delta\mathbf{d}(\omega, \mathbf{m}^k)^\dagger \Delta\mathbf{d}(\omega, \mathbf{m}^k) + \frac{\lambda}{2}{\mathbf{m}^k}^T \mathcal{D}\mathbf{m}^k. \tag{3.13}$$

The first term $\mathscr{C}_D$ accounts for the misfit between the observed and the synthetic data through the ℓ_2-norm of the residuals $\Delta\mathbf{d}(\omega, \mathbf{m}^k) = d^{obs}(\omega) - d^{cal}(\omega, \mathbf{m}^k)$ that measure the difference between the observed data d^{obs} and the synthetic data d^{cal} computed in the current model $\mathbf{m}^k$. Usually, the residuals involve the total simulated wavefield u_{tot}. In our case, however, it is much more convenient to use the scattered field $u_{sc} = u_{tot} - u_{inc}$ since

1. it can be computed using the measurements of the incident field u_{inc},
2. it is insensitive to antenna radiation pattern,
3. it is the field that is analytically propagated toward the receivers by the integral representation (see eq. 3.3 and Fig. 3.4).

Assuming that the incident field can be accurately simulated (which has been demonstrated by our numerical analysis, see Figs 3.3 and 3.8), we thus have

$$\Delta\mathbf{d}(\omega, \mathbf{m}^k) = \mathbf{d}^{obs}_{tot}(\omega) - \mathcal{R}\,\mathbf{u}^{cal}_{tot}(\omega, \mathbf{m}^k), \tag{3.14}$$

$$= \mathbf{d}^{obs}_{sc}(\omega) - \mathcal{R}\,\mathbf{u}^{cal}_{sc}(\omega, \mathbf{m}^k), \tag{3.15}$$

where $\mathcal{R}$ is the projection operator of the wavefield on the receiver location, which takes the integral representation and the windowed Sinc interpolation into account.

The second term $\mathscr{C}_M$ of the misfit function (3.13) introduces a Tikhonov regularisation (Tikhonov and Arsenin, 1977). The operator $\mathcal{D}$ corresponds to the Laplacian so that its minimization tends to provide smooth solutions. The hyperparameter λ is a regularisation weight that balances the importance given to model smoothness relatively to the data misfit. In eq. (3.13), the symbol ^{T} denotes transposition and the notation † corresponds to the transpose^T-conjugate* operator.

To minimize the misfit function, we use a quasi-Newton optimisation method (L-BFGS-B algorithm, Byrd et al., 1995) which performs a local descent based on the gradient of the misfit function $\mathbf{G}(\mathbf{m}^k) = \nabla_{\mathbf{m}^k}\mathscr{C}(\mathbf{m}^k)$. The gradient of the model term is straightforwardly

computed by finite differences on the model $\mathbf{m}^k$ whereas the gradient of the data term $\mathbf{G}_D(\mathbf{m}^k)$ is computed via the adjoint state method (Plessix, 2006). Differentiating eqs (3.13) and (3.14) with respect to the model parameters, we have

$$\begin{aligned} G_{D_i}(\mathbf{m}^k) &= \frac{\partial \mathscr{C}_D(\mathbf{m}^k)}{\partial m_i}, \\ &= \sum_\omega \mathcal{R}e \left\{ \frac{\partial \Delta \mathbf{d}(\omega, \mathbf{m}^k)}{\partial m_i}^\dagger \Delta \mathbf{d}(\omega, \mathbf{m}^k) \right\}, \\ &= -\sum_\omega \mathcal{R}e \left\{ \frac{\partial \mathbf{u}_{tot}^{cal}(\omega, \mathbf{m}^k)}{\partial m_i}^\dagger \mathcal{R}^\dagger \Delta \mathbf{d}(\omega, \mathbf{m}^k) \right\}, \end{aligned} \tag{3.16}$$

where u_{tot}^{cal} verifies the equation of the forward problem (3.1), which gives by differentiation

$$\frac{\partial \mathbf{u}_{tot}^{cal}(\omega, \mathbf{m}^k)}{\partial m_i} = -\mathcal{A}(\omega, \mathbf{m}^k)^{-1} \frac{\partial \mathcal{A}(\omega, \mathbf{m}^k)}{\partial m_i} \mathbf{u}_{tot}^{cal}(\omega, \mathbf{m}^k). \tag{3.17}$$

Injecting eq. (3.17) in eq. (3.16) yields the expression of the gradient

$$G_{D_i}(\mathbf{m}) = \sum_\omega \mathcal{R}e \left\{ \mathbf{u}_{tot}^T \frac{\partial \mathcal{A}}{\partial m_i}^T \mathbf{v}^* \right\}, \tag{3.18}$$

where $\mathbf{v}$ is the adjoint wavefield, that verifies the linear system $\mathcal{A}^\dagger \mathbf{v} = \mathcal{R}^\dagger \Delta \mathbf{d}$ and corresponds to the back-propagation of the residuals in the medium (Plessix, 2006).

In expression (3.18), the diffraction matrix $\partial_{m_i} \mathcal{A}$ characterizes the sensitivity of the data to the parameter m_i, that refers either to the permittivity ε_i or to the conductivity σ_i at grid point i. Lavoué et al. (2014) promote to consider dimensionless parameters that can be gathered in the same model vector $\mathbf{m}$, thus enabling to perform the optimisation in the joint permittivity-conductivity parameter space (in opposition to cascaded or decoupled approaches as used for instance by Ernst et al., 2007; Meles et al., 2010; Ellefsen et al., 2011). Thus, we consider a relative permittivity $\varepsilon_r = \varepsilon/\varepsilon_o$ (with $\varepsilon_o \simeq 8.85 \times 10^{-12}$ F/m the dielectric constant) and a relative conductivity $\sigma_r = \sigma/\sigma_o$. The reference conductivity σ_o can be defined as $\sigma_o = \varepsilon_o \omega_o$, by analogy with the contrast function used in the inverse scattering community (see e.g. Abubakar et al., 2005), with ω_o a reference angular frequency that we take as the mean frequency of the measurements (i.e., we consider $\omega_o = 2\pi \times 6$ GHz for the dataset FoamDielExt and $\omega_o = 2\pi \times 10$ GHz for FoamMetExt). However, these definitions are arbitrary (they are only conventions) and we pointed out in Lavoué et al. (2014) that the inversion is very sensitive to the priority given to the permittivity or to the conductivity update via the parameter scaling. Therefore, we introduce an additional scaling factor β and we consider the couple of parameters $(\varepsilon_r, \sigma_r/\beta)$ for the optimisation. As detailed in Lavoué et al. (2014), a small value for the factor β will give priority to the reconstruction of permittivity, whereas a large scaling factor will enhance the conductivity update.

On synthetic experiments, Lavoué et al. (2014) propose a workflow to choose an adequate value for the scaling factor β, in conjunction with an appropriate regularisation weight λ, based on the data misfit analysis. In the following, we aim at validating this workflow on the experimental data provided by the Institut Fresnel. In all the following inversion tests, we stop the optimisation process when the relative misfit decrease between two iterations becomes smaller than 10^4 times the machine precision (default stopping criterion in the L-BFGS-B algorithm, Zhu et al., 1997).

3.1.3.2 Monoparameter inversion of the dataset FoamDielExt

First of all, we wish to validate the quantitative imaging of permittivity through monoparameter inversion, that does not involve any parameter scaling. The target FoamDielExt is particularly suited for this purpose as the objects are supposed to be purely dielectric and have well constrained permittivity values. Fig. 3.9 shows the inversion result when inverting the 9 frequencies between 2 and 10 GHz simultaneously, without any regularisation ($\lambda = 0$). During the inversion, the misfit function has been decreased by 98.8% within 32 iterations. We recover permittivity values of 1.42 ± 0.05 for the foam cylinder and $\varepsilon_r = 3.10 \pm 0.42$ for the plastic one (mean and variance are computed within the red dashed circles shown in Fig. 3.9a). These recovered values (in blue on the log in Fig. 3.9b) are in very good agreement with the expected ones (in red in Fig. 3.9b). On the other hand, the data fit is also very satisfying, as shown in Fig. 3.10 which compares the observed data with synthetic data computed in the reconstructed model of Fig. 3.9 (data are shown for the same frequency 8 GHz and source position $\theta_s = 270^o$ as in Fig. 3.6).

It is interesting to note in Fig. 3.9(a) that the image presents a particular geometrical pattern: oscillations with a radial symmetry that manifest themselves as fluctuations on the blue curve in Fig. 3.9(b). These fluctuations are absent in the background due to the bound constraint $\varepsilon_r \geq 1$ considered in the L-BFGS-B algorithm. In the target, the oscillating pattern is due to the limited wavenumber coverage of our imaging technique, that depends both on the finite-frequency content of the data and on the discrete spatial sampling of the measurements by the acquisition array (Sirgue and Pratt, 2004). For comparison, we indicate in Fig. 3.9(b) low-pass-filtered versions of the expected (red) log, considering the full frequency bandwidth from 2 to 10 GHz (in black), or only the lowest frequency of 2 GHz (in green). For filtering, we use the maximal reconstructed wavenumber given by Sirgue and Pratt (2004) as a function

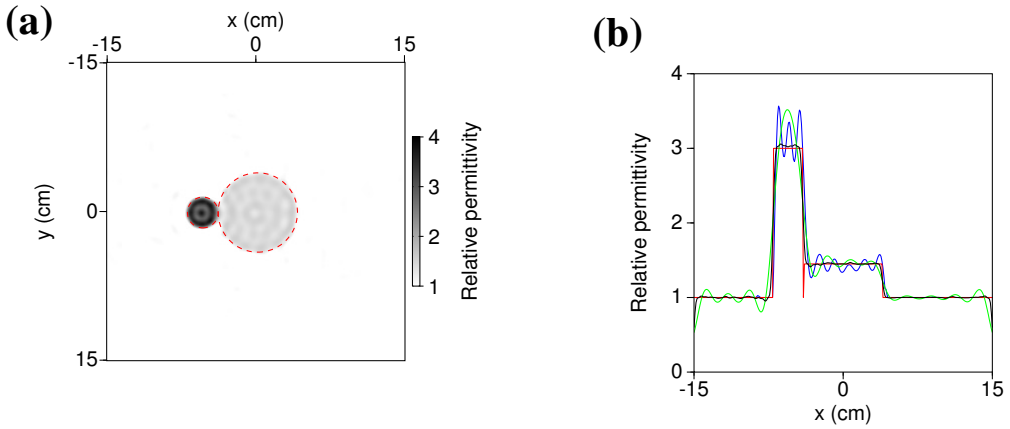


Figure 3.9: *Permittivity reconstruction of the target FoamDielExt by monoparameter inversion. (a) 2D model. The red dashed circles indicate the expected contours of the objects, after the translation by $dx = 2$ mm, $dy = 1$ mm suggested by Geffrin et al. (2005). (b) Logs extracted from the 2D models along the horizontal line $y = 0$. The reconstructed values (blue line) are compared to the synthetic model of Fig. 3.2(a) (red line) and to low-pass-filtered versions of this model, considering the full frequency bandwidth (black line) or only the lowest frequency (green line).*

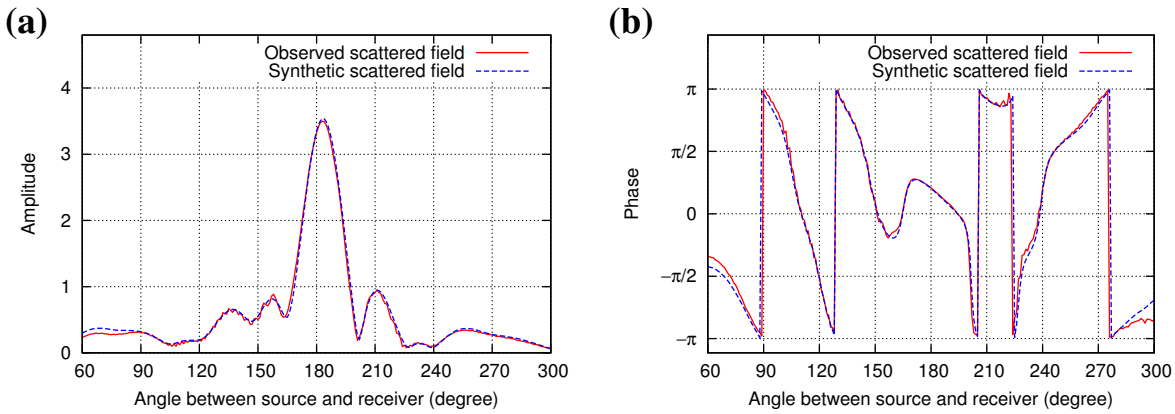


Figure 3.10: *(a) Amplitude and (b) phase of observed vs. synthetic data computed in the final reconstructed model of Fig. 3.9, for frequency 8 GHz and the source at* $\theta_s = 270^o$.

of frequency and of the minimal illumination angle:

$$k_{max} = \frac{2\omega}{v_o} \cos\left(\frac{|\theta_s - \theta_r|_{min} = 60^o}{2}\right). \tag{3.19}$$

As this filter assumes a continuous coverage of the wavenumbers up to k_{max} (i.e. an infinitesimal spatial sampling), the full-bandwidth low-pass-filtered log over-estimates the accuracy of the reconstruction and the actual reconstructed values exhibit a lower wavenumber content: Although very dense, the acquisition sampling still has an imprint in the final image. In particular, the source intervals are responsible for the radial symmetry observed in Fig. 3.9(a).

The image of Fig. 3.9(a) can thus be seen as the convolution of the real model with the resolution operator of the imaging technique. Knowing the frequency content of the source and the illumination of the target by the acquisition array, we could deconvolve the reconstructed image of Fig. 3.9(a) in an impulsive image that would better reflect the reality, as proposed by Ribodetti et al. (2000). This finite-frequency effect also explains why, when the initial model is good enough, inverting all available frequencies simultaneously yields smoother images than inverting them sequentially as proposed by Pratt and Worthington (1990). Using a sequential strategy, the final reconstructed model results from the inversion of the highest frequency and is subject to mono-frequency fluctuations, whereas the simultaneous frequency strategy benefits from a broadband frequency content (Lavoué et al., 2014).

3.1.3.3 Multiparameter inversion of the dataset FoamMetExt

To invert the dataset FoamMetExt, we apply the methodology proposed by Lavoué et al. (2014): As the choice of an adequate scaling factor β is not straightforward, we first perform several multiparameter inversions independently, using various scaling values, without regularisation. Fig. 3.11 shows the final misfits obtained with the different scaling factors. Based on the data misfit, a scaling factor $\beta = 10$ seems to be the most adequate value but we can observe that misfits of the same order of magnitude can be obtained for lower values, down to $\beta = 0.5$.

As already observed by Lavoué et al. (2014) on synthetic data, the reconstructed models corresponding to roughly equivalent misfits can be quite different. For instance, Figs 3.12

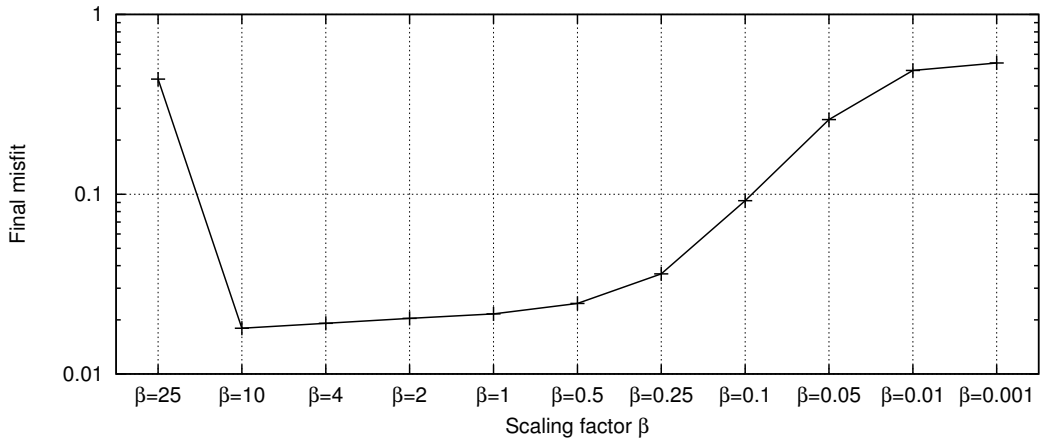


Figure 3.11: *Final misfit decrease as a function of the scaling factor* β *used in the multiparameter inversion of the dataset FoamMetExt.*

and 3.13 show the reconstructed models obtained using scaling factors $\beta = 10$ and $\beta = 0.5$, respectively. The permittivity solutions appear to be particularly sensitive to the choice of the parameter scaling, because the imprint of the foam cylinder on the data is less important than the one of the strongly diffracting copper tube[1].

The main difference between the two solutions is that, using a scaling factor $\beta = 0.5$, the optimisation artificially creates an erroneous permittivity structure inside the copper tube. Based on our forward simulations, we know that this structure is not reliable because the measured data are not sensitive to the inner filling of the highly-conductive copper tube ($\sigma \gg 100$ S/m). Synthetic data, however, may be sensitive to this structure since the reconstructed conductivity value is lower than the actual one ($\sigma \simeq 5$ S/m), in particular in the early iterations of the inversion. The erroneous permittivity reconstruction inside the tube is thus a nice example of the trade-off between permittivity and conductivity: comparing Figs 3.12 and 3.13, it can be observed that both permittivity and conductivity reconstructions vary inside the tube, indicating that variations of one parameter compensate the variations of the other regarding the data misfit. These trade-off effects justify the need for considering the Hessian matrix in multiparameter optimisation schemes (Operto et al., 2013; Lavoué et al., 2014).

Apart from the artifact inside the copper tube, the two solutions also differ by the amplitude of the fluctuations of permittivity values in the foam cylinder (again, the absence of fluctuations in the background is imposed by the bound constraints $\varepsilon_r \geq 1$ and $\sigma \geq 0$ given to the L-BFGS-B algorithm). Finally, the solutions present a thin circular artifact around the copper tube, a region to which the data should be sensitive. These effects can be mitigated by introducing a Tikhonov regularisation in the optimisation, in order to obtain smoother permittivity models.

[1] *Additional note to the published version:* The relative sensitivities of these laboratory data to permittivity and conductivity are reversed compared to the previous study (Section 2.2, Lavoué et al., 2014) because the synthetic subsurface model of Section 2.2 was inspired by natural media where conductivity has generally a weaker signature than permittivity.

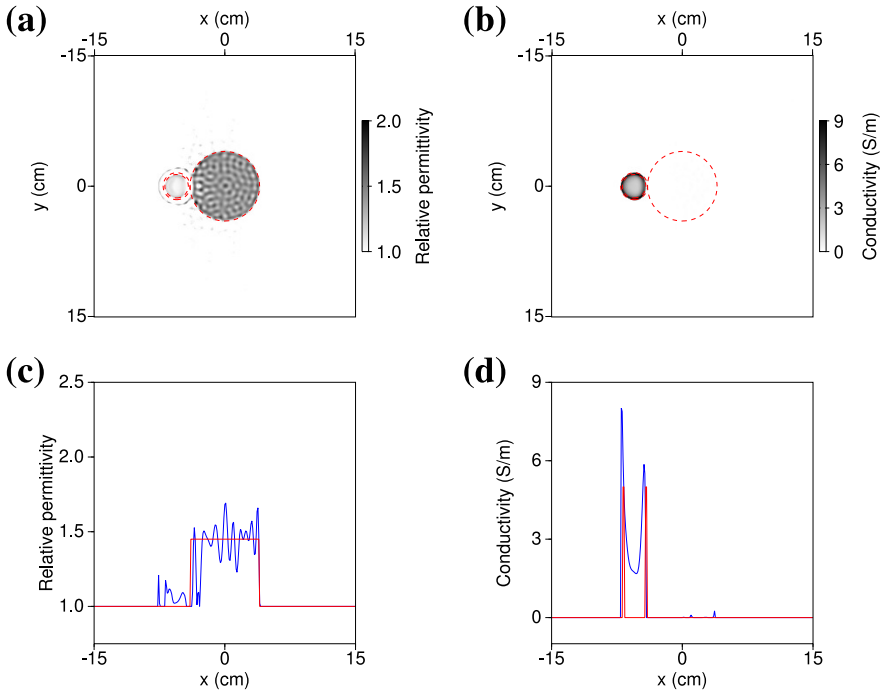


Figure 3.12: *(a,c) Permittivity and (b,d) conductivity models obtained by multiparameter inversion of the dataset FoamMetExt, without regularisation and using a scaling factor* $\beta = 10$. *The misfit function has been decreased by 98.2% in 120 iterations.*

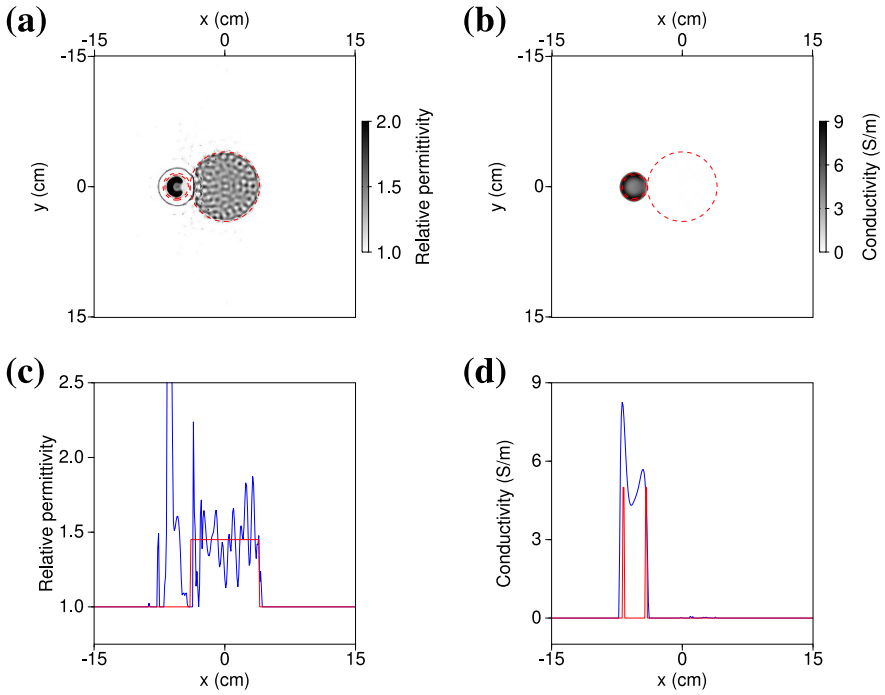


Figure 3.13: *(a,c) Permittivity and (b,d) conductivity models obtained by multiparameter inversion of the dataset FoamMetExt, without regularisation and using a scaling factor* $\beta = 0.5$. *The misfit function has been decreased by 97.5% in 61 iterations. The logs of panels (c) and (d) are extracted along the line* $y = 0$ *in panels (a) and (b), respectively.*

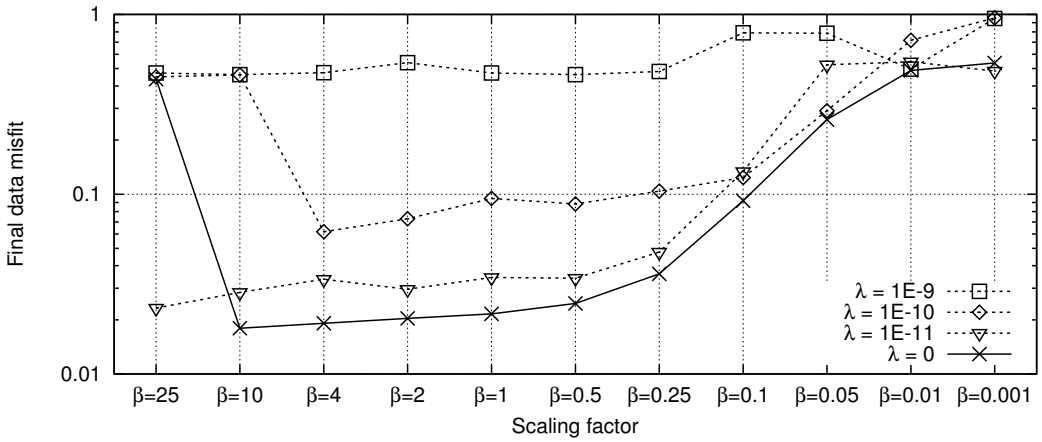


Figure 3.14: *Final data misfits as a function of the scaling factor* β *and of the regularisation weight* λ *used in the multiparameter inversion of the dataset FoamMetExt.*

Fig. 3.14 shows the final misfits that we obtain when performing the multiparameter inversion with various scaling factors β and regularisation weights λ. As artifacts arise mainly in the permittivity image, whereas conductivity is better constrained by the data, we only apply regularisation on the permittivity model (a finer implementation might involve regularisation weights adapted to each parameter type, depending on their imprint on the data, but this fine tuning is not critical for our purpose). Contrary to the results obtained by Lavoué et al. (2014), the final data misfits are not more sensitive to the parameter scaling with regularisation than without, so the regularisation does not enable to distinguish between the solutions obtained with different scaling factors. However, looking at the final reconstructed models, we can observe that the permittivity solutions obtained with regularisation are now quite similar (see Figs 3.15 and 3.16). The use of regularisation erases the permittivity artifacts around the copper tube, yielding satisfactory results. Adequate values for the parameter scaling and for the regularisation weight λ can be selected based on the data misfit (Fig. 3.14), given that applying an adequate regularisation, we obtain similar results for all scaling factors that display a good data misfit. Note that the Tikhonov regularisation does not induce a dramatic smoothing of the solutions, although a multiplicative regularisation may better preserve the contrasts in the present case of well-delimited, piecewise-constant targets (Abubakar et al., 2005).

Finally, Fig. 3.17 compares the observed data with the synthetic data computed in the final reconstructed models of Figs 3.15 and 3.16. The fit to the data is naturally dominated by the main peak amplitude of the signal at angles around 180^o, while the data at extreme angles are not well fitted. Applying an angle-dependent weighting to the data through a data covariance matrix in the misfit function could enable to better fit the data presenting minor amplitudes.

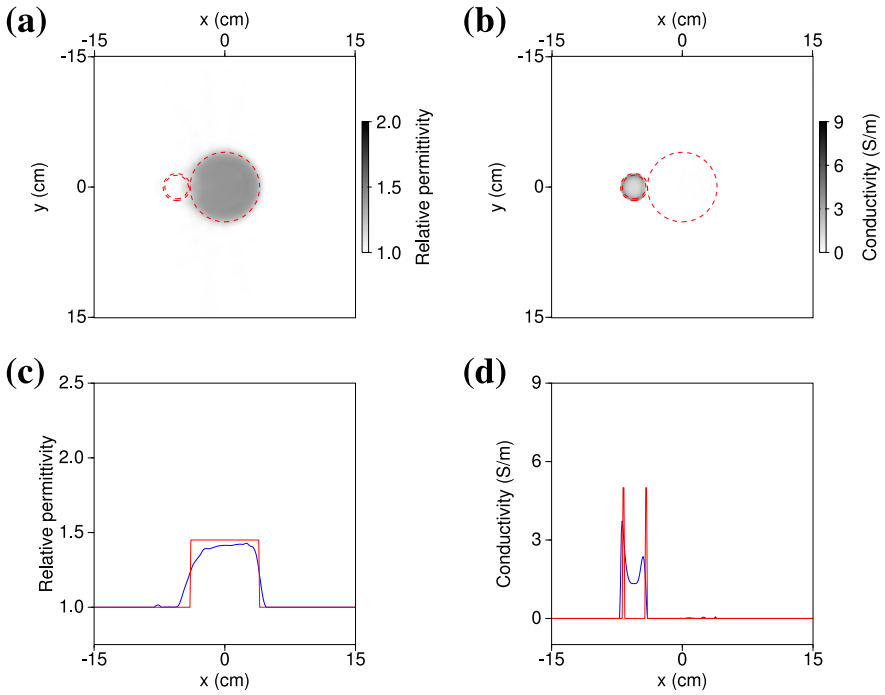


Figure 3.15: *(a,c) Permittivity and (b,d) conductivity models obtained by multiparameter inversion of the dataset FoamMetExt, using a regularisation weight* $\lambda = 10^{-11}$ *and a scaling factor* $\beta = 10$*. The misfit function has been decreased by 97.1% in 59 iterations. The logs of panels (c) and (d) are extracted along the line* $y = 0$ *in panels (a) and (b), respectively.*

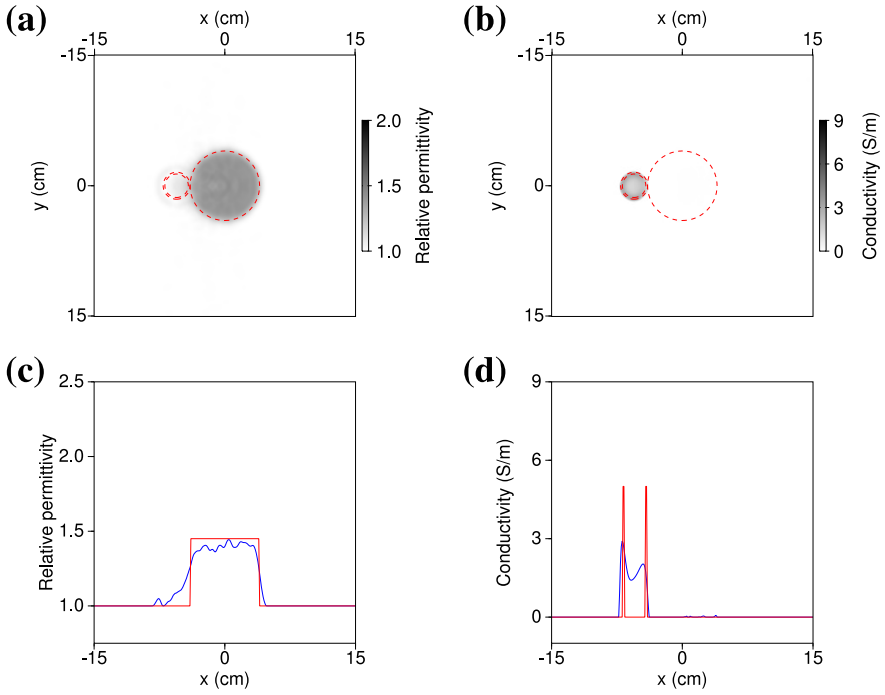


Figure 3.16: *(a,c) Permittivity and (b,d) conductivity models obtained by multiparameter inversion of the dataset FoamMetExt, using a regularisation weight* $\lambda = 10^{-11}$ *and a scaling factor* $\beta = 0.5$*. The misfit function has been decreased by 96.5% in 101 iterations. The logs of panels (c) and (d) are extracted along the line* $y = 0$ *in panels (a) and (b), respectively.*

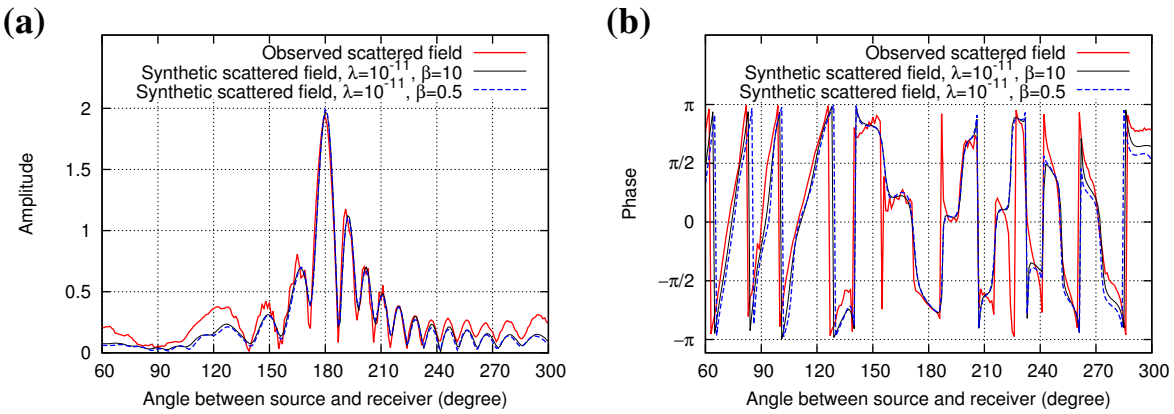


Figure 3.17: *(a) Amplitude and (b) phase of observed vs. synthetic data computed in the final reconstructed model of Figs 3.15 and 3.16, for frequency 18 GHz and the source at* $\theta_s = 240^o$.

Conclusions

In this study, we have presented a comprehensive view of the inversion of electromagnetic data collected during a well-controlled laboratory experiment. For an accurate and efficient resolution of the forward problem, we use a FDFD scheme where the stencil coefficients are optimised to each simulated frequency. An integral representation of the fields enables to reduce the computation domain to a small zone enclosing the targets and has been shown to be an accurate, efficient, and elegant alternative to brute force calculations in the entire domain. An important ingredient of the modelling consists in accurately positioning virtual sources and receivers on the integration contours using a windowed Sinc interpolation.

Scattered data produced by the purely dielectric target FoamDielExt allowed us to validate our algorithm regarding monoparameter inversion. On the reconstructed permittivity image, the targets are well-delimited and the recovered values are very close to the expected ones. Observed and synthetic data are in very good agreement. We also confirm our strategy for multiparameter inversion on the dataset FoamMetExt. As already observed on synthetic data, the reconstructed models are sensitive to the scaling applied between different parameters types, especially for parameters that are less constrained by the data. These model variations can be barely visible in the data misfit, leading to an ambiguity between the different solutions. An adequate regularisation weight enables to mitigate the artifacts, so that satisfactory models can be obtained and, more importantly, identified on the basis of the data misfit analysis. Besides, forward simulations on synthetic models are of great help to estimate the sensitivity of the data to model variations, and hence to evaluate the reliability of the inversion results. Synthetic data are almost insensitive to some parts of the model where our data-driven technique is not able to deduce any valuable information without introducing some prior information in the misfit function[1].

[1] *Additional note to the published version:* We may point out that our imaging method is not totally exempted from any prior information. For instance, lower limits for permittivity and conductivity in free space ($\varepsilon_r = 1$, $\sigma = 0$ S/m) are explicitly given to the bounded L-BFGS optimisation algorithm. Of course these bounds are very reasonable hypothesis since they are physical limits, but they still have a strong effect on the results: It is why the reconstructed images suffer from so few artifacts in the neighbourhood of the targets. The choice of the discretisation step h in the FD grid can also be regarded as an implicit prior information: It has been partly defined knowing the sizes of the structures to be imaged.

The success of inverting these experimental data lets us envisage the inversion of on-ground GPR field data in a near future. The present work made us free from the inverse crime approach and required to accurately simulate the observed data. When dealing with GPR field data, we expect to face other major obstacles. In particular, we may encounter difficulties for estimating the GPR source signature, and for designing a suitable initial model for starting the full waveform inversion process. The acquisition configuration is also very different in on-ground GPR applications where data are acquired only from the surface, and it has a strong effect on the ability of the imaging technique to recover the subsurface targets.

In particular, in the frame of multiparameter inversion, a partial illumination tends to enhance the trade-off between parameters (Hak and Mulder, 2010). Our previous study on synthetic data (Lavoué et al., 2014) suggests that multiparameter FWI can be performed from surface data if the information contained in the Hessian of the misfit function is taken into account through quasi-Newton methods. However, an illustration on real field data still has to be performed. In a preliminary stage, the Institut Fresnel database may constitute an interesting tool to test the sensitivity of the inversion with respect to the acquisition configuration, as well as to investigate optimisation methods that better take the Hessian information into account (e.g. the truncated Newton method, Métivier et al., 2013; **?**).

Acknowledgments

This work was performed using the high-performance computing facilities of the mesocentre CIMENT (Université de Grenoble) and the national facility GENCI (Grand Equipement National de Calcul Intensif) under grant 046091. It benefited from a close collaboration with the SEISCOPE consortium (`http://seiscope2.osug.fr`). The authors would like to thank the Institut Fresnel for freely providing their high-quality data (`http://iopscience.iop.org/0266-5611/21/6/S09/media`). F. Lavoué is particularly grateful to Christelle Eyraud (Institut Fresnel) for useful informations about the database, to Ludovic Moreau (ISTerre) for his advices on fields integral representation, and to Stéphane Operto (Géoazur, Nice-Sophia Antipolis) for sharing his experience on FDFD schemes and for his careful reading of the manuscript. We also thank two anonymous reviewers for their corrections and X. Dérobert (IFSTTAR, Nantes) for his invitation to publish in this special issue of *Near Surface Geophysics*.

3.2 Imaging a limestone reservoir from on-ground GPR data

3.2.1 Introduction: Context and aim of the study

In this section, I consider on-ground GPR data which were acquired by G. Sénéchal[1] and D. Rousset[1] on the site of the inter-Disciplinary Underground Science and Technology laboratory (or *Laboratoire Souterrain à Bas Bruit*, LSBB[2], see Fig. 3.18 for a regional localisation). The site of the LSBB is a former military facility of the french nuclear missile system consisting in a 1.5-km-long tunnel that leads to a shielded launching control room buried at 500 m depth under the Grande Montagne (Fig. 3.19). Decommissioned in 1997, the LSBB provides now a unique opportunity to access directly the interior of the carbonate massif and to study its unsaturated zone (Fig. 3.20).

Karstic limestone massifs are geologic formations of major importance since they contain a significant part of drinking water resources throughout the world. They also present an interest as geological archives of sedimentary platforms and as analogues of deep hydrocarbon reservoirs. However, the understanding of karstic structures is often made difficult by the complexity of their evolution. In particular, the multi-scale heterogeneities of these formations have a great influence on the reservoir properties (Leonide et al., 2012). On the long-term, GPR imaging at LSBB therefore aims at characterising the limestone massif in the context of karstic hydrogeology and at producing small-scale reservoir models.

In the frame of my PhD, my personal interest is to determine whether FWI can contribute to this characterisation, by providing high-resolution, quantitative, and reliable images of the distribution of permittivity (and eventually conductivity) in the carbonates. In particular, we shall address the question to know if FWI can bring more information than a classical processing workflow implying velocity analysis and migration. Classical approaches are mainly based on the assumption of layered media and therefore present limitations in complex media, where the recorded signal is the superimposition of several waves. We can expect FWI to overcome these limitations since it *a priori* enables to reproduce complex wave propagation.

Of course, a short term pre-requisite to answer these interesting questions is that FWI must simply *work* with surface GPR data acquired on the field. Indeed, the preliminary study I present here is only focused on this aspect: how can we apply 2D FWI to on-ground GPR data? To answer this question, we have to deal with the following points:

1. First of all, FWI needs a good starting model to avoid to fall into local minima. The specificity of on-ground GPR compared to crosshole GPR or refraction seismics resides in the fact that we cannot invoke first arrival time tomography to build this initial model. Instead, I shall rely on a classical velocity analysis based on reflected events.

2. FWI also requires a correct estimation of the source signature, which is a difficult task since there is a strong trade-off between this estimation and the one of the model parameters. When a good initial model is available, source estimation is usually performed in the frequency domain within the inversion process. If we do not trust the current model, it might be more relevant to estimate the source in the time domain, by selecting some specific signal.

[1]IPRA, Université de Pau et des Pays de l'Adour, France.
[2]Rustrel, France, `http://www.lsbb.eu`.

3. Before applying FWI, the acquired data must be properly pre-processed so that they can be interpreted by our 2D numerical simulations. This requires to eliminate out-of-plane signals that can not be reproduced, and to perform a 3D-to-2D transformation of the data to well explain the amplitudes.

The outline of the section is the following: First, I show that applying a classical processing workflow — involving a systematic semblance analysis, NMO correction, and post- or pre-stack migration — difficultly leads to satisfactory images because of local inaccuracies of the velocity model, which points the need for a more accurate velocity estimation through an inversion procedure. Nevertheless, semblance analysis and hyperbola picking provides a first NMO model that can be considered for starting the FWI process.

In a second time, I go back to a simpler approach leading to a blocky description of the medium and show that it can explain a significant part of the complex observed data. It also provides a simple medium where I can estimate the source wavelet.

Then I present the pre-processing steps applied to the data which are finally transformed into the frequency domain. An analysis of the characteristics of the data in the frequency domain, partly based on the approach of Sirgue and Pratt (2004) presented in Chapter 2, guides the choice of the frequencies considered for inversion.

Finally, I present some preliminary inversion results. These attempts are preliminary in two respects. First, I concentrate on the reconstruction of the permittivity model because it is expected that most of the observed reflections and diffractions are due to permittivity contrasts in the limestone, which may not present high conductivity values nor variations (at least at first order). As discussed in Chapter 2, this preliminary monoparameter step is important to improve the permittivity model before envisaging the reconstruction of conductivity in further multiparameter stages. Secondly, I restrict the inversion to a low frequency bandwidth, i.e. to the early stage of a low-to-high-frequency hierarchy following the Bunks' strategy. At this stage, my aim is not to get a final interpretable image, but only to see if FWI can converge using the proposed workflow, and what features it can reconstruct depending on the initial model and on the source signal we consider.

NB: The data set is quite large and in the following, I do not show all data at each step of interpretation to avoid too many figures, but only selected common-offset sections or CMP gathers. A more complete overview of the data is provided in Appendix B.

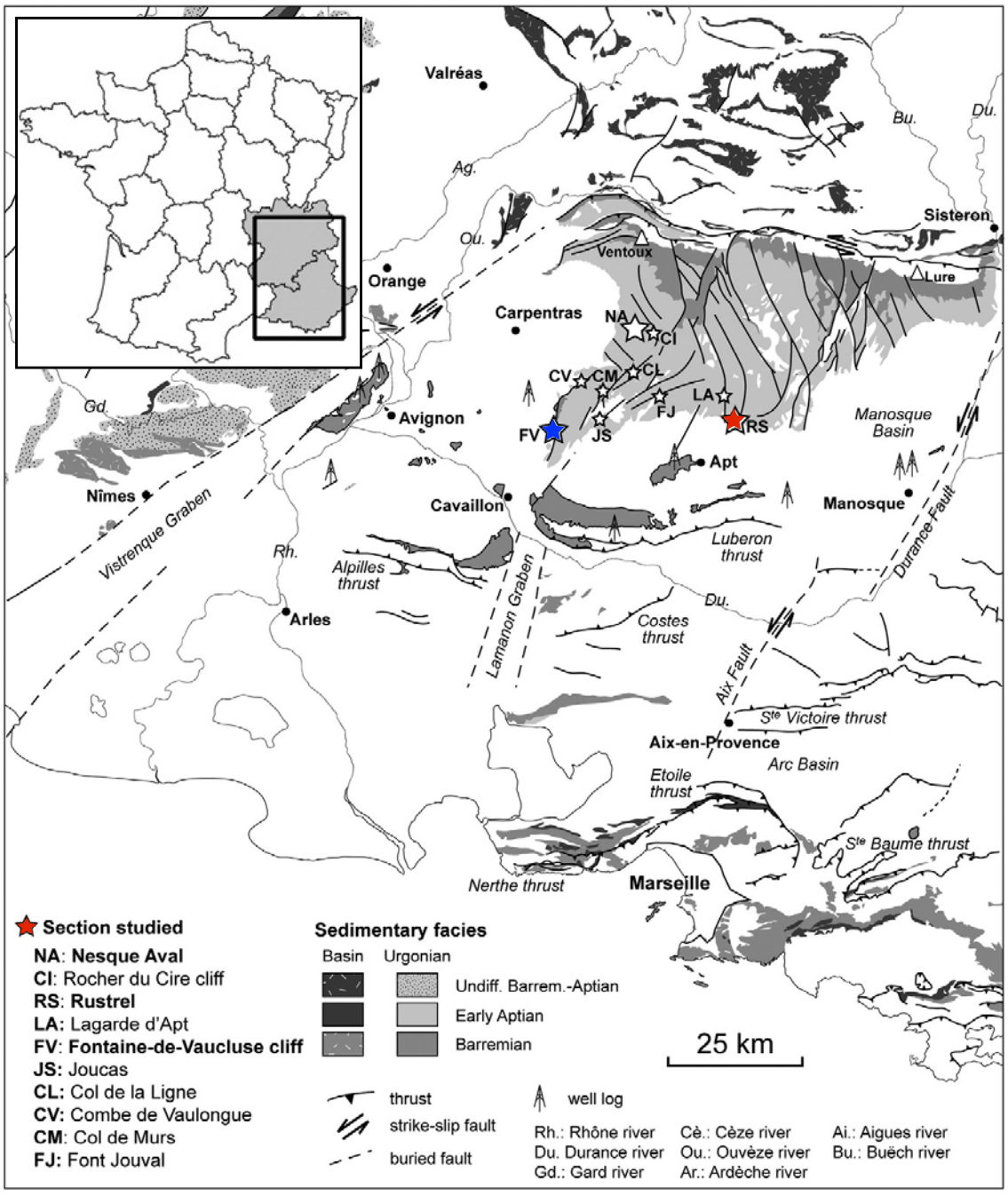
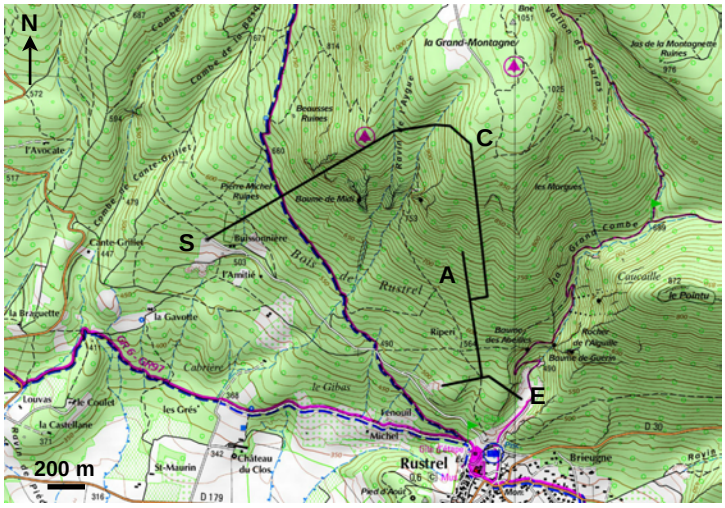
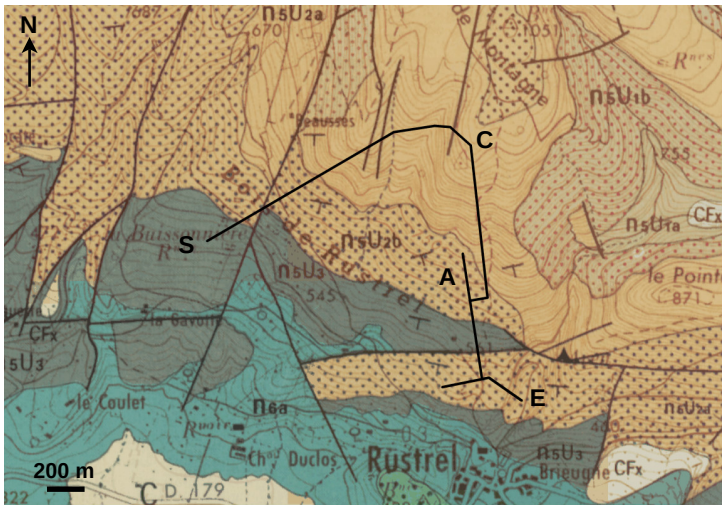


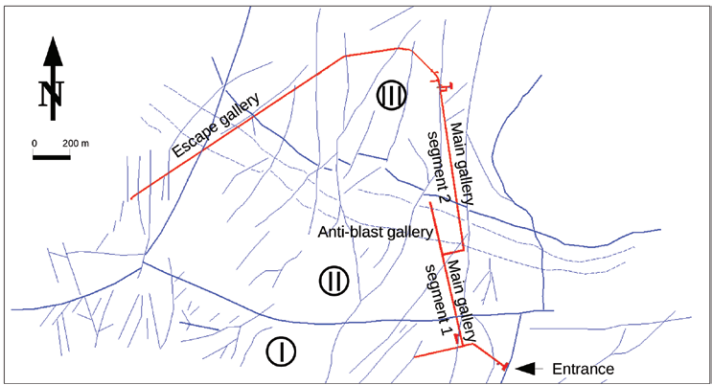
Figure 3.18: *Geology of SE France and location of the site of study (Rustrel,* ***RS****, modified from Leonide et al., 2012). The considered carbonate massif belongs to the Urgonian platform that was deposited on the border of the Vocontian basin during the Lower Cretaceous (Bedoulian, Aptian, ca. 125 My). Note the proximity of Rustrel with the Fontaine de Vaucluse (****FV****), which is the principal hydrological outlet of the Monts de Vaucluse massif (south-western part of the light grey area).*



(a) Topographic map (IGN 1/25000, edited from `http://www.geoportail.gouv.fr`) and location of the LSBB galleries (black line). E denotes the main entrance, A the anti-blast gallery where the data were acquired, C the shielded capsule (former launching control room, 500 m under the ground surface), and S the emergency exit.



(b) Geological map (BRGM 1/50000, edited from `http://www.geoportail.gouv.fr`). The units n5U correspond to the Bedoulian limestones (Urgonian platform) and n6 to the Gargasian marls. Note the stratigraphic dip of the surface outcrops, S-SE to S-SW, compared to the S-SE direction of the anti-blast gallery (A).



(c) Structural map (from Sénéchal et al., 2013, modified from Thiébaud, 2003, with contribution from Matonti et al., 2011). Roman numbers correspond to the geological compartments of Fig. 3.20 and blue lines indicate the fracture network. Eventual minor fractures (dashed blue line) can cross the anti-blast gallery and might affect the profile (see also Derode et al., 2013, concerning fractures in the anti-blast gallery).

Figure 3.19: *Topography and geology of the Grande Montagne and location of LSBB galleries.*

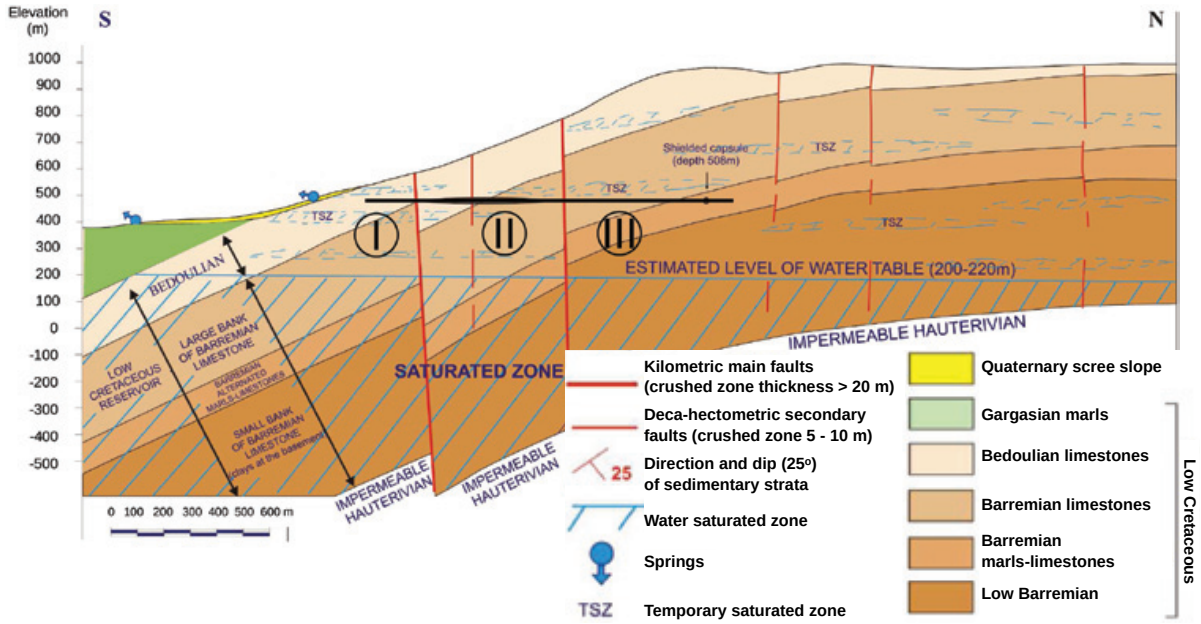


Figure 3.20: *Schematic S-N geological section of the Grande Montagne (from Sénéchal et al., 2013, modified from Thiébaud, 2003). The horizontal black line indicates the level of the galleries. The anti-blast gallery is located in compartment II where a transition between Barremian and Bedoulian limestones is expected.*

Previous geophysical experiments and location of the GPR profile

Since its decommissioning, the LSBB has been the location of numerous hydrogeological studies (see e.g. Thiébaud, 2003; Garry, 2007; Blondel, 2008; Perineau, 2013), based on observations at the surface and in the tunnel, as well as on hydraulic tracer monitoring. Unfortunately, most of the galleries are cemented which makes the direct observations seldom. There is thus a strong need for imaging the interior of the massif.

In order to fill this gap, geophysical measurements were performed to have an insight into the limestone reservoir. For instance, Maufroy et al. (2014) performed a seismic travel-time tomography to image the massif between the ground surface and the tunnel level. At a smaller scale, Sénéchal et al. (2004) performed seismic experiments between the main tunnel and the anti-blast gallery[1] (see Figs 3.19 and 3.20 for a localisation). These authors could estimate the P-wave velocity in the limestone but seismics does not have enough resolution to perform a high-resolution of the karstic structures. GPR is a good candidate for achieving such a high-resolution imaging in carbonates, which present a low conductivity and thus enable a deep penetration of electromagnetic waves. GPR measurements were achieved by Carrière et al. (2013) on the plateau and by Sénéchal et al. (2013) in the main tunnel and in the anti-blast gallery. Finally, Van Vorst et al. (2010) performed a travel-time tomography from crosshole GPR measurements acquired in the anti-blast gallery.

[1]This gallery owes its name to the fact that it was originally designed to prevent the blast of an eventual bombing attack from propagating into the tunnel: the launching control room should remain operational under any circumstance.

The GPR data I consider here were also acquired in the anti-blast gallery, which is the only place in the LSBB tunnel where the walls are not consolidated with concrete, allowing an easier access to the outcrops (at least partially, because the walls are still coated in some places with a thin layer of cement). The data were acquired along a 50-meter-long profile using 250-MHz shielded Ramac antennas (Malå Geosciences) in common-offset configuration. Measurements were carried out for 15 offsets between 1.24 and 8.34 m, with an offset interval of $\simeq$ 50 cm. Traces were recorded every 10 cm along the profile, which results in 501 CMP locations for each offset and 7515 traces in total. Acquisition triggering and trace positioning were performed using a calibrated wheel whose accuracy can be estimated to a few cm. Distance stretching was corrected *a posteriori* by repositioning the traces regularly between 0 and 50 m.

Unfortunately, the profile does not comprise the section imaged by Van Vorst et al. (2010): our profile starts at $x = 70$ m from the entrance of the gallery, whereas the boreholes of Van Vorst et al. (2010) are located around $x = 40$ m (Guy Sénéchal, personal communication). Nevertheless, the tomography of Van Vorst et al. (2010) gives us a first *a priori* on the expected relative permittivities in the medium, comprised between 8 and 12.5, while Sénéchal et al. (2013) estimate velocities about 9 cm/ns, i.e. permittivities around 11, by semblance analysis of a CMP gather acquired in the anti-blast gallery. The latter authors also mention a stratigraphic dip of about 25°. Contrary to what suggests the surface dip indicated on the geological map (Fig. 3.19b), the tunnel (and thus the profile) is not parallel to this stratigraphic dip. After time-to-depth conversion of their common-offset section, Sénéchal et al. (2013) measure an apparent dip of 16°. A more precise indication on the stratigraphic dip and on fractures in the anti-blast gallery is given by Derode et al. (2013). These authors mention N120-20°SW limestone layers, cut by a family of fractures of azimuth N30° with dip angles of 60-80°W. I underline, however, that the precise location of our profile with respect to the above-mentioned fractured zone is not exactly known and should be specified by *in situ* observations.

Time-domain common-offset sections

Fig. 3.21 shows common-offset sections of the profile, for the offsets 1.24 m and 8.34 m (the complete data set is presented in Appendix B.1). Applying few processing to the data (constant component removal and amplitude saturation), it is possible to visualise the main structures of the medium. In particular, at offset 1.24 m, three main reflectors are clearly visible, which correspond to dipping stratification planes (denoted as RW in Fig. 3.21). At second order, we can also remark some discontinuities of these reflectors, as well as numerous diffraction hyperbolae (D). We may wonder if the discontinuities are caused by fractures in the limestone or if they are artefacts due to velocity variations. Similarly, it would be interesting to explain the origin of the diffractions: are they cavities partially filled by air or water, or heterogeneities such as calcite seals? Finally, the section presents lateral amplitude variations that can be due either to the transparency of the medium, or on the contrary to its attenuation, or even to lateral variations in the antenna-ground coupling. Variations in the amplitude of the reflections might also be due to the varying thickness of the reflectors that can act as thin layers in the considered frequency range. Understanding this effect is important since thin layers can have a strong reflectivity: an open question in the interpretation of this dataset is to know if the observed reflections are due to an impedance contrast between the layers of limestone, or to the reflectivity of stratigraphic interfaces acting as thin layers within a homogeneous limestone massif.

At offset 8.34 m, the interpretation is made slightly more difficult since the structures are deformed during a longer travel-time. However, it is still possible to recognise the main reflectors (RW). We also remark a high amplitude aerial event arriving at about 50 ns (RAW). CMP gathers suggest that this signal is a reflection on one of the sides of the tunnel and should be discarded for the inversion because the modelling will not reproduce this feature. At large offsets, we also observe enhanced multiples (MRW) and aerial dipping events corresponding to reflected-refracted waves (RRW). Moreover, the arrival times of the direct ground-wave (GW) present strong lateral variations (dashed red lines), which indicates lateral variations of the velocity in the superficial part of the medium. This point partially answers the previous question of knowing whether the observed reflections are due to a contrast between limestone layers or to thin-layer effects. The linear move-out of the ground-wave suggests that there is a

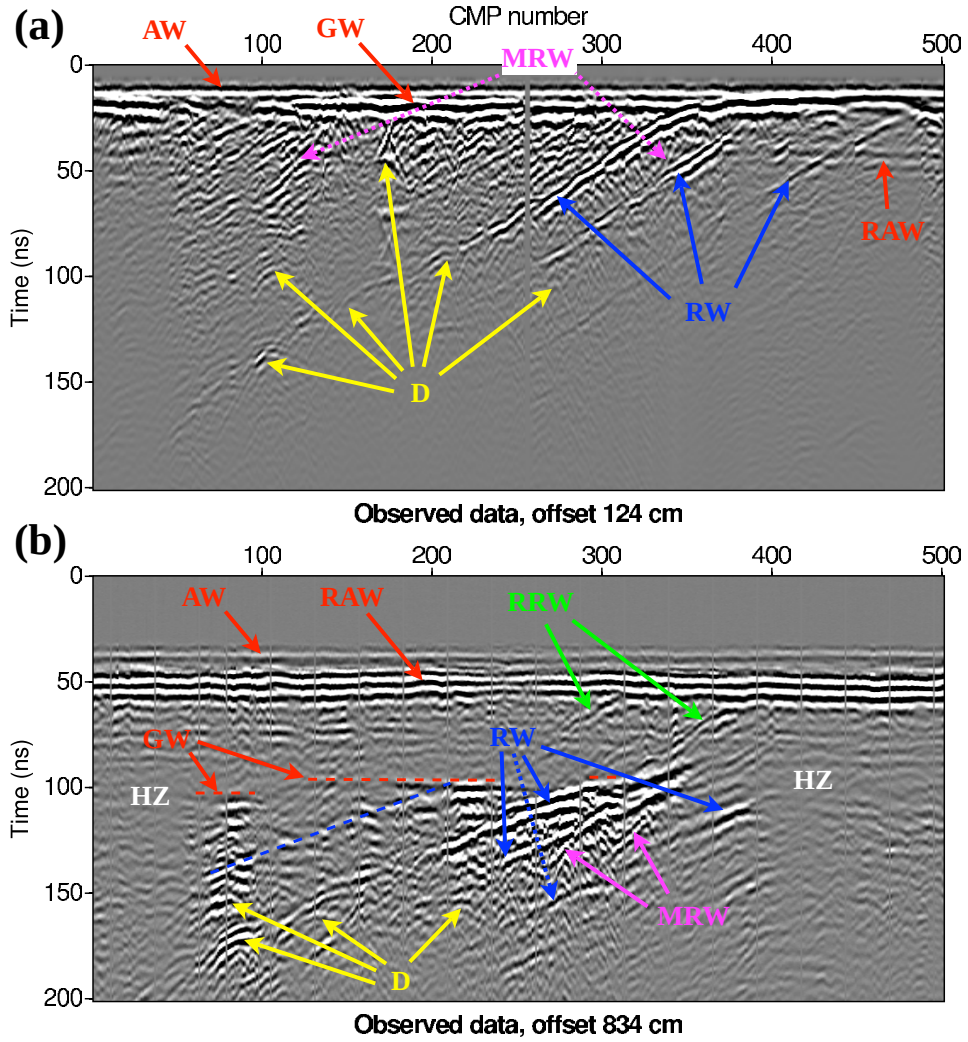


Figure 3.21: *Common-offset sections for the minimal and maximal acquired offsets: (a) 124 cm, (b) 834 cm. AW = direct air-wave, GW = direct ground-wave, RAW = reflected air-wave, RW = reflected waves, MRW = multiply-reflected waves, RRW = reflected-refracted waves, D = diffractions, HZ = hidden zones. Processing consisted only in t_o-correction, constant component removal using* `sugain mbal=1` *and amplitude saturation (*`perc=97`*).*

velocity contrast between some limestone layers, consistently with the results of Van Vorst et al. (2010). Some other reflections, however, are not associated with a significant shift of the direct ground-wave arrival (dashed blue line) and might be interpreted as a thin-layer reflectivity. Finally, the amplitude of the direct ground-wave presents strong lateral variations, indicated as *hidden zones* (HZ) in Fig. 3.21, which are difficult to understand at this stage of interpretation. I will explain them later, based on numerical simulations.

In the following, the quantitative analysis of the data aims at improving this first qualitative overview.

3.2.2 Classical processing: velocity analysis, migration, forward modelling

3.2.2.1 Semblance analysis, direct wave and hyperbola picking

A first way to derive a quantitative information from the data is to perform a systematic velocity analysis, by semblance analysis of the CMP gathers. Fig. 3.22 shows an example of semblance panel with the corresponding CMP gather where I represent the hyperbolae that corresponds to the picked velocities, accounting for the observed normal move-out (NMO). It should be underlined that the value of picked NMO velocity does not always result in a relevant hyperbola. When the hyperbolae deduced from the NMO velocities are not satisfying, I rather pick the hyperbolae directly on the CMP and perform a $t^2 - x^2$ regression to obtain the NMO

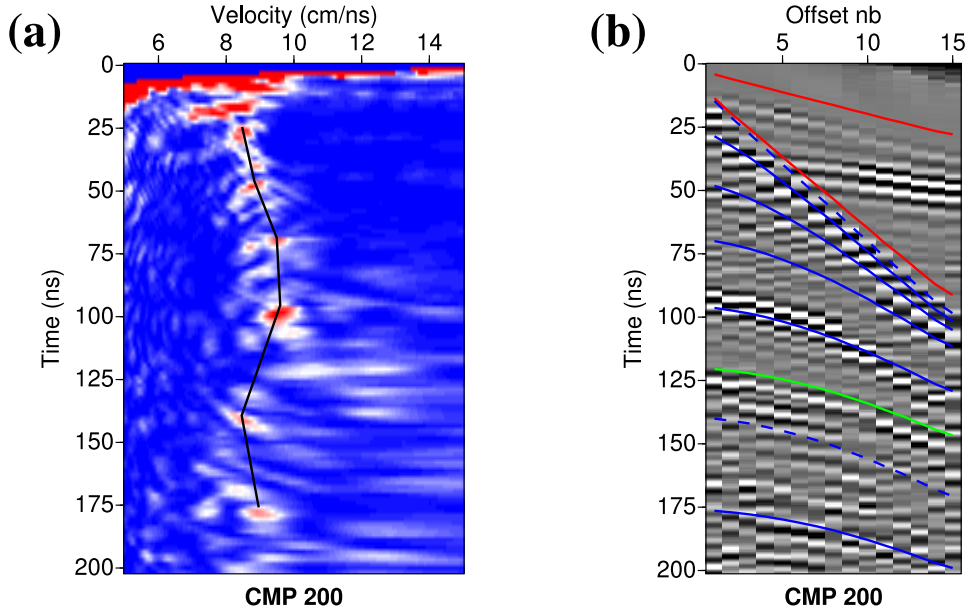


Figure 3.22: *(a) Semblance panel and picked NMO velocities (broken black line). (b) Corresponding CMP gather with the hyperbolae corresponding to the picked NMO velocities (in blue). Red straight lines correspond to the theoretical direct air-wave and to the picked direct ground-wave. The green curve is an example of hyperbola I prefer picking rather than relying on the picked NMO velocity from the semblance panel at $\simeq 140$ ns, which does not correspond clearly to any hyperbola. NMO velocities associated to dashed blue lines are discarded for building the interval velocity model.*

velocities (dashed blue line in Fig. 3.22(b). When possible, I also pick the direct ground-wave (red line in Fig. 3.22b).

Repeated for each CMP gather (every 5 CMP in practice, i.e. every 50 cm), semblance analysis and hyperbolae picking enable to derive NMO velocities as a function of time and of distance along the profile. Interval velocities are then derived from NMO velocities using Dix formula (Dix, 1955). This approach, however, assumes a horizontally layered medium and is not relevant in the case of dipping interfaces. To derive interval velocities, I applied a correction for the dip θ ($v_{int}^{corr} = v_{int} \times \cos\theta$, see Levin, 1971).

Fig. 3.23(a) shows the permittivity model derived from the interval velocities computed from the NMO velocities assuming a dip value $\theta = 13.5°$ (this value is justified later). As the velocity analysis has been performed every 50 cm, this model is very heterogeneous and not realistic. It is also due to the fact that deriving interval velocities from NMO velocities with the iterative Dix formula is not a stable process and results in over-estimated velocity variations at depth if the superficial velocities are not correct. However, the evolution of permittivity with distance and depth nearly reproduce the dip of the main reflectors ($\simeq 12°$ to $16°$). Fig. 3.23(b) presents a smooth version of the NMO model, which has been obtained using a gaussian smoothing in the direction of the dip $\theta = 13.5°$. This smooth model is our first candidate to be used as an initial model for starting the FWI process. Although this permittivity model has been derived from interval velocities, it will be referred as the *smooth NMO model* in the following.

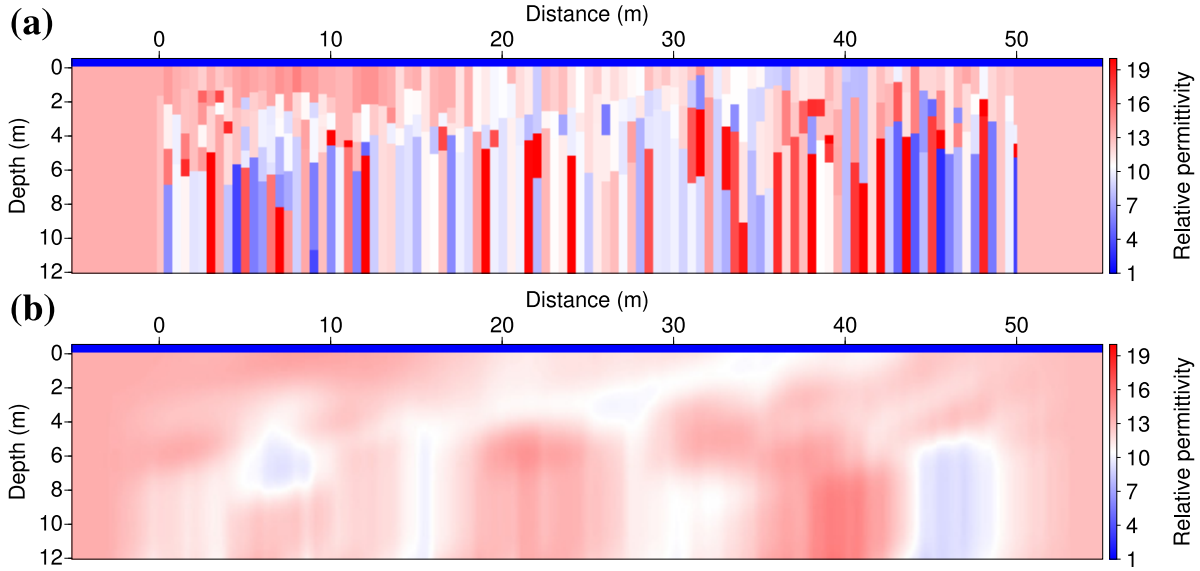


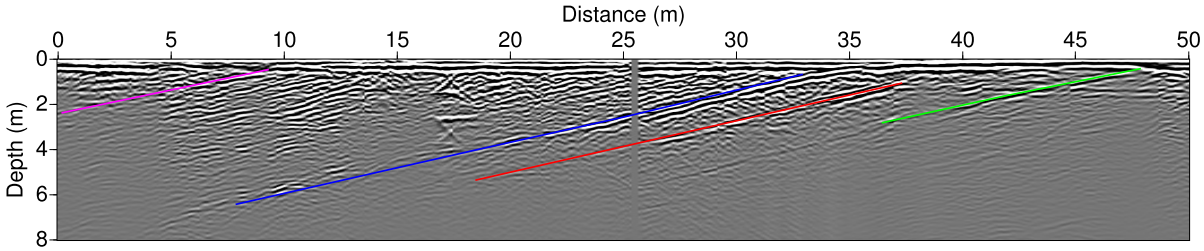
Figure 3.23: *Permittivity models derived from semblance analysis every 50 cm. (a) Raw model from the interval velocities derived from v_{NMO} assuming a dip of 13.5°. (b) Smooth version obtained with a gaussian smoothing along the dip ($\tau_{\|} = 3$ m, $\tau_{\perp} = 1$ m).*

3.2.2.2 NMO correction and stack, reverse-time migration

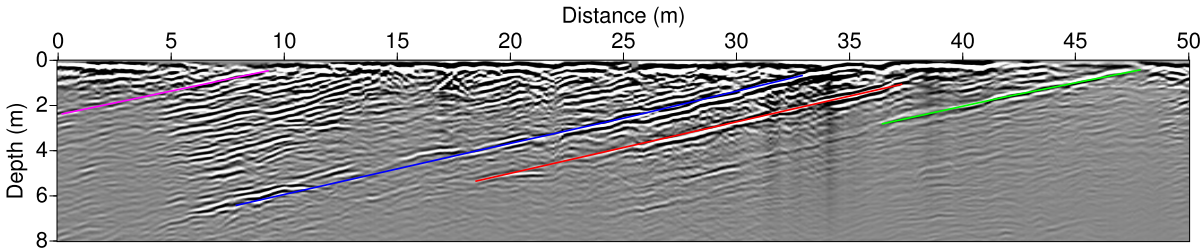
Using the NMO velocities (those derived from semblance analysis and hyperbola picking, not the ones of Fig. 3.23), it is possible to correct the CMP gathers for normal move-out. In a classical processing workflow, it is usual to stack the resulting zero-offset sections to increase

the signal-over-noise ratio, taking advantage of the data redundancy. Fig. 3.24 shows the result of such a stack and compares the depth-converted stacked section with the raw common-offset section of Fig. 3.21, after delay correction (to make it comparable with zero-offset sections). In Fig. 3.24(b), the relative amplitudes of the common-offset sections have been kept unchanged before stacking, such that the resulting image is dominated by short-offset signal and resembles the raw common-offset section of Fig. 3.24(a). Some reflectors are slightly more pronounced due to the stack but some hyperbolae are still visible. It is not obvious that this image really helps for improving our interpretation of the data, compared to the raw data. In Fig. 3.24(c), amplitudes have been normalised before stacking, such that all common-offset panels have similar contribution in the sum. Because of the variability of the NMO velocities from a CMP to another, it results in a quite noisy image where more artefacts have been added than valuable information, compared to the non-corrected common-offset section. A fortiori, post-stack migration is not successful and results in many artefacts that make the image not interpretable. These observations motivate the need for an inversion process that could improve the velocity model and enable a more accurate migration and interpretation of the image.

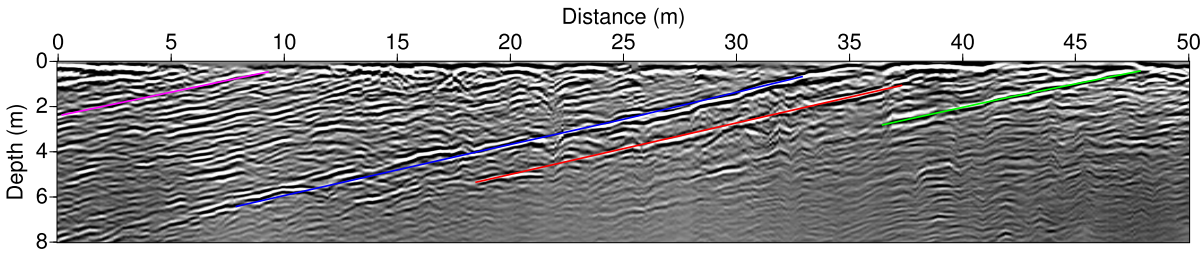
If migrated images are not readily interpretable, they can still be exploited to test if the



(a) Common-offset section of Fig. 3.21 at 124 cm, corrected for arrival-time and converted to depth.



(b) Stacked NMO zero-offset section (true amplitudes).



(c) Stacked NMO zero-offset section (amplitudes normalised vs. offset).

Figure 3.24: *Common-offset section at 124 cm and zero-offset stacked sections after NMO correction. Time-to-depth conversion assumes an average velocity of 9 cm/ns ($\varepsilon_r \simeq 11$). The mean value of interface dip is $\simeq 13.5^o$.*

permittivity model of Fig. 3.23(b) derived from semblance analysis and hyperbola picking can be used as an initial model for starting the FWI. Of course, a more direct way to test the kinematic compatibility of this model is simply to perform forward simulations and compare the calculated data with the observed ones. The resulting comparison is shown later in Fig. 3.32 and confirms that the model well explain the arrival time of the direct ground-wave. But, as the model of Fig. 3.23(b) is smooth, it does not reproduce the observed reflected events and it is not possible to know if they can predict the right arrival times for the reflections.

To do so, I compare the individual NMO-corrected sections to see if the reflectors have been moved to the same place with respect to offset: this gives an indication of the kinematic compatibility of the NMO velocity model with the observed reflections. Converting the time axis to depth with a constant velocity of 9 cm/ns (which is quite valid above the main reflector), it is also possible to estimate the apparent dip of the reflectors, which display an average of 13.5° (hence the value used to convert NMO velocities to interval velocities in the previous paragraph).

In the same idea, I also performed reverse-time migrations (RTM) using the smooth NMO model of Fig. 3.23(b). As mentioned in Chapter 2, RTM can be performed straightforwardly using the FWI algorithm if the entire recorded data are injected in the adjoint source instead of the residuals. Correlating the incident and the adjoint wavefields for each frequency component of the temporal signal amounts to the time-domain imaging condition and produces the migrated image. As I did not estimate the source signature yet, I do not expect to replace the reflector at the right depth. But looking if they are refocused at the same depth with respect to offset gives an indication on the capability of the velocity model to explain the observed arrival times. Comparing the migrated images with respect to offset (see Fig. 3.25) suggests that the smooth velocity model of Fig. 3.23(b) is well compatible with the arrival times of the reflected events. Therefore, we can expect to reconstruct the reflector at the right place during the inversion, provided we correctly estimate the source signature.

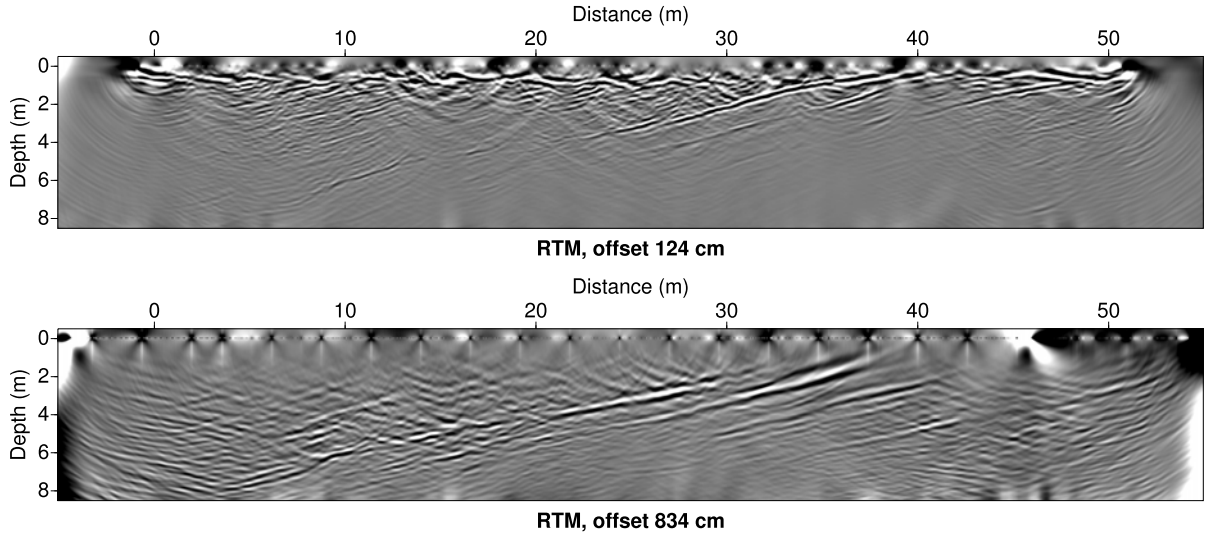


Figure 3.25: *RTM images using the smoothed NMO model for the offsets 124 and 834 cm.*

3.2.2.3 Forward modelling in a blocky model

Up to now, we have seen that a systematic velocity estimation based on semblance analysis and arrival-time picking provides velocity models that are kinematically compatible with the observed data, but that do not enable an accurate imaging and an improved interpretation. In this paragraph, I present a simpler interpretation of the data, based on the study of the linear move-out of the direct ground-wave with respect to offset. As seen in Fig. 3.22(b), the direct ground-wave can sometimes be picked on the CMP gathers, but it is rare. Here, I look directly on the common-offset sections, and I consider the medium as made of blocks delimited by the reflectors previously identified on the NMO-corrected sections of Fig. 3.24. Picking the mean arrival-time of the direct ground-wave within these blocks enables to derive a blocky permittivity model shown in Fig. 3.26.

Fig. 3.27 recaps the different estimations of permittivity and velocity derived up to now from semblance analysis, hyperbola fitting, direct wave fitting, and average linear move-out

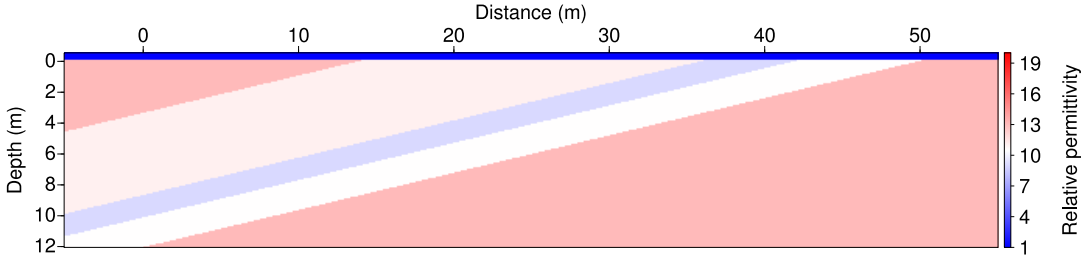


Figure 3.26: *Blocky model derived from direct ground-wave picking. The relative permittivity of the blocks are, from the left to the right, 13, 11, 9, 10.5 and 13. The dip of the interfaces is derived from the zero-offset staked section (*13.5^o*, see Fig. 3.24).*

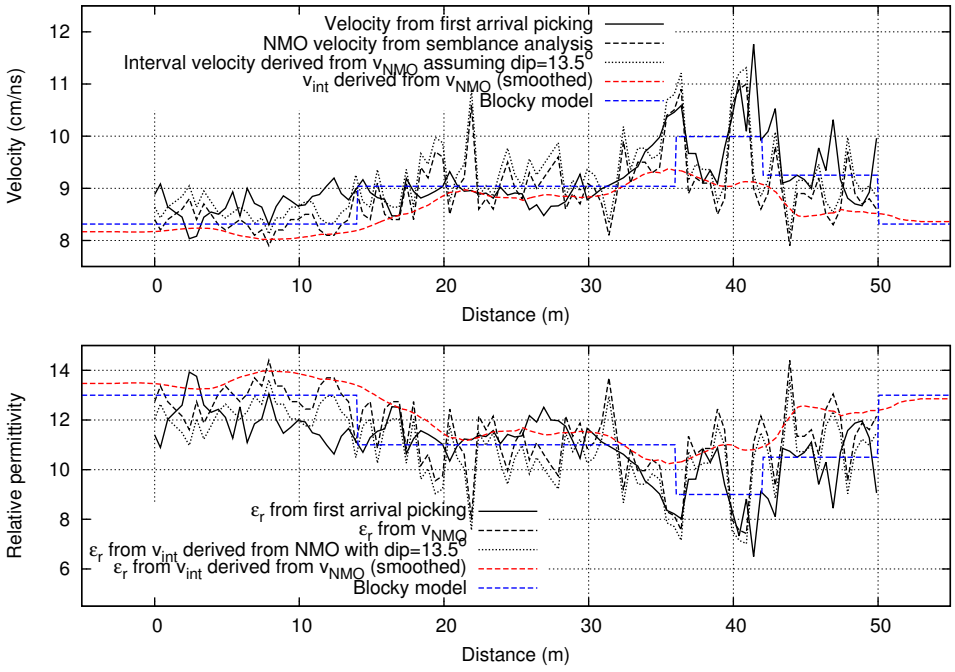


Figure 3.27: *Superficial velocities and permittivities (recap of the different estimations).*

within the blocks. The motivation of the blocky model of Fig. 3.26 is trying to explain the complex observed data with a simple model of the subsurface. Indeed, such a simple model already enables to explain a significant part of the data complexity.

To illustrate the complexity that can arise with such a blocky model and understand in more details the observed data and further simulation results, Fig. 3.28 presents time-domain simulations in a model comprising only one dipping interface. Because of the interaction of the waves with the edge between the interface and the ground surface (Fig. 3.28a), the corresponding common-offset section shown in Fig. 3.28(b) displays a complex pattern. In addition to the expected air-wave, ground-wave and primary reflected waves, we can observe the multiples of the reflected wave, as well as primary and secondary reflected-refracted waves. This simulation has been used to identify the events observed in the data (Fig. 3.21).

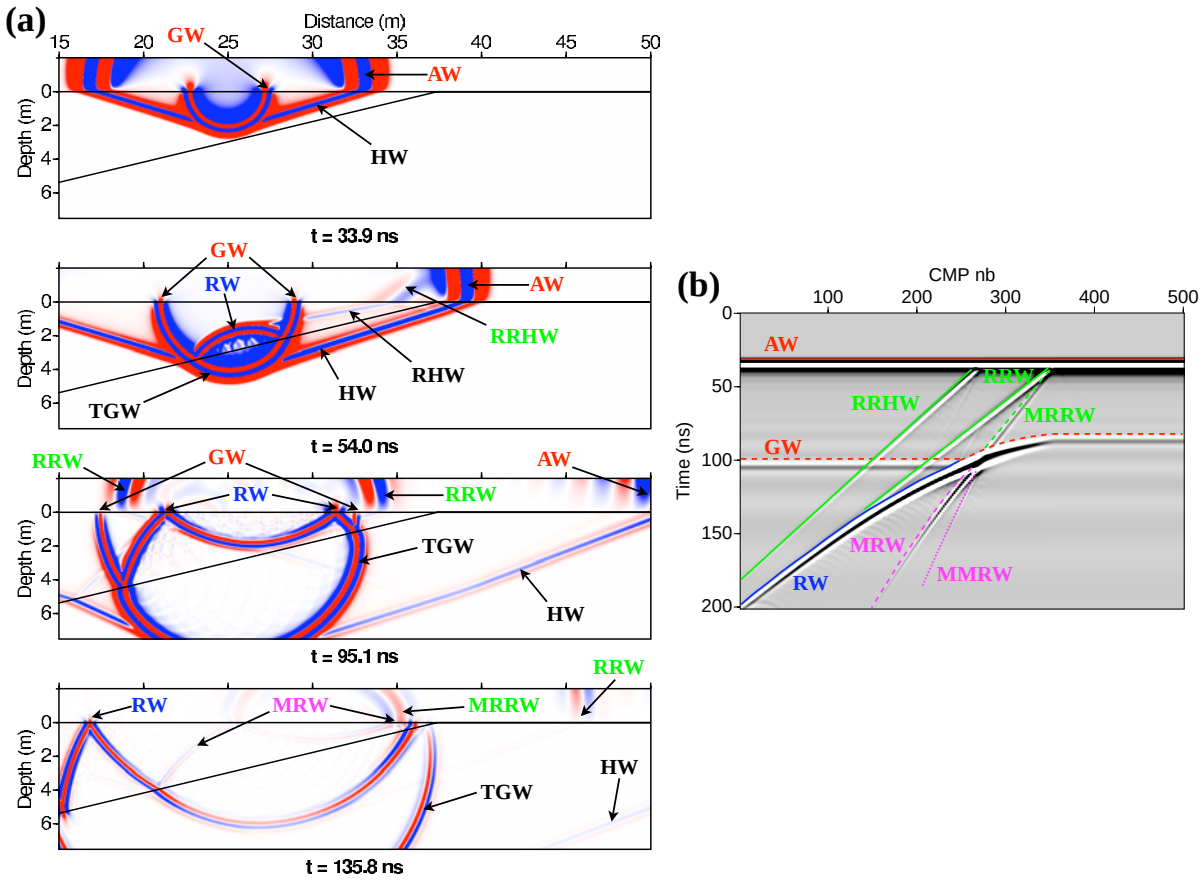


Figure 3.28: *Time-domain simulation in a blocky model presenting one dipping interface. (a) Snapshots of the wavefield emitted by a source at $x = 25$ m. (b) Common-offset section for offset 834 cm. AW = direct air-wave, GW = direct ground-wave, TGW = transmitted ground-wave, HW = head wave, RHW = reflected head wave, RRHW = reflected-refracted head wave, RW = reflected wave, MRW = first multiple of RW, MMRW = second multiple of RW, RRW = reflected-refracted wave, MRRW = multiple of RRW.*

Fig. 3.29 now presents common-offset sections computed in the blocky model of Fig. 3.26, presenting several interfaces. These synthetic data consist in a superimposition of several patterns similar to the one of Fig. 3.28(b) and can explain a significant part of the observed

events. In addition to the primary and secondary main reflections, we can notice that the alternation of blocks with different velocities does reproduce the observed lateral variations in the amplitude of the direct ground-wave. In particular, we retrieve hidden zones, where the direct ground-wave disappears, when passing from a high-velocity block to a low-velocity block on the right of the profile (CMP $\simeq$ 380 and 450 in Fig. 3.29). A similar alternation might be invoked to explain the hidden zone on the left of the profile (CMP $<$ 75). It is quite remarkable that a simple blocky model can explain these amplitude variations, although it has been designed on the only basis of the kinematics. This observation mitigates our *a priori* concern about laterally varying attenuation or antenna-ground coupling, and confirms that most features of the observed data can be explained by permittivity variations.

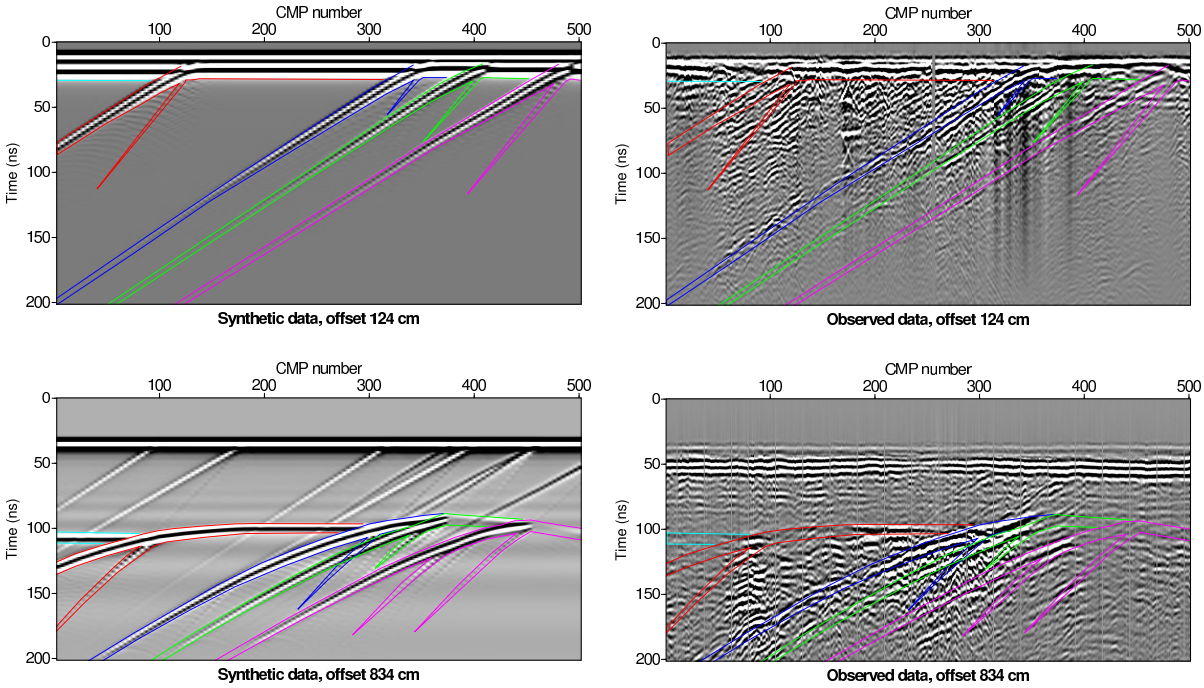


Figure 3.29: *Left: Simulations in the blocky model of Fig. 3.26 at offsets 124 and 834 cm. Right: Corresponding observed data. Each coloured curve corresponds to the pattern of a specific dipping interface, as observed in Fig. 3.28(b), underlining the direct ground-wave, and the primary and secondary reflected events. These curves are identical on the left and right sides, indicating the good kinematic agreement of the synthetic data with the observed ones (except on the extreme left and right parts of the profile where the ground-wave disappears at large offset, making the comparison difficult). Please note that the simulation reproduces the observed hidden zones on the right of the profile.*

Partial conclusion on the velocity analysis

As a partial conclusion, we now have two permittivity models as candidates for starting the FWI: the smooth NMO model of Fig. 3.23(b) and the blocky model of Fig. 3.26. Both models are kinematically consistent with the observed data, in the sense that they well explain some of the observed arrivals (direct ground-wave and main reflections). The major difference between the smooth permittivity model derived from semblance analysis and the blocky model resides

in the fact that the blocky model generates reflections, whereas the smooth NMO model only explains the direct arrivals, due to the lack of sharp contrasts.

Consequently, my goals will be different when trying to invert the data by starting from one of these two models: using the smooth NMO model, I am interested to see whether the inversion can reproduce the main reflections which are initially lacking; whereas using the blocky model, I want to see what features the inversion does add to the pre-existing reflectors (and if it does not degrade them). I may also address the question to know whether the inversion does need pre-existing reflectors in the starting model to converge towards a satisfactory solution.

In the following, I will try to perform FWI to update these permittivity models. This requires first to pre-process the data and to estimate the source signature to make the observed waveforms comparable to the synthetic ones.

3.2.3 Pre-processing steps towards FWI

3.2.3.1 Data pre-processing: Mute and 3D-to-2D conversion

FWI relies on the comparison between observed and synthetic data. A first processing step towards inversion is therefore to prepare the observed data to make them comparable with the synthetic ones. The processing steps I present here consist in three operations:

1. elimination of the signals that can not be reproduced by the simulations,
2. 3D-to-2D transformation,
3. source estimation.

In the following, I present these steps in a quite linear way but, in practice, these steps were applied iteratively in order to converge towards cleaned, virtually 2D, observed data that could be compared to the synthetic ones.

Muting out-of-plane events

In Fig. 3.21, we have identified a high-amplitude signal as an out-of-plane aerial reflection, that I will call the reflected air-wave (RAW) in the following. As the simulations will not reproduce this event, and since it does not bring any valuable information about the underground, the reflected air-wave should be removed from the data.

To do so, I simply applied a mute in the time-CMP domain[1], for the offsets where it could be clearly distinguished from other relevant signals. At short offsets (≤ 280 cm), the reflected air-wave arrives later than the direct ground-wave, together with reflected events. In this case, it has been muted only where it was clearly identifiable, and where it displayed high amplitudes that could disturb the inversion[2] (see Fig. 3.30a, at CMP ≥ 450 and $t \simeq 50$ ns). At intermediate offsets (between 336 and 438 cm), it was not possible to distinguish the reflected air-wave from the direct ground-wave, so I do not use these offsets for the inversion. At large offsets (≥ 488 cm), it was possible to mute this parasite signal along the whole profile. For

[1]Looking at the data in other domains, e.g. $\tau - p$ or $\omega - k$, does not enable to better isolate the reflected air-wave from the events of interest for the conflicted offsets 336 to 438

[2]If not muted, these disturbances could be evidenced in RTM images as horizontal artefacts (not visible in Fig. 3.25 because I used muted data to perform the RTM).

these offsets, I also mute the direct air-wave which tends to overlap with the reflected air-wave. To replace the muted direct air-wave, a synthetic direct air-wave will be added after source estimation. On the other hand, I kept the signals that I identified as reflected-refracted waves and that can be reproduced by the simulations as shown in Fig. 3.29.

Fig. 3.30 shows the resulting muted common-offset sections. The complete muted dataset is shown in Appendix B.2.

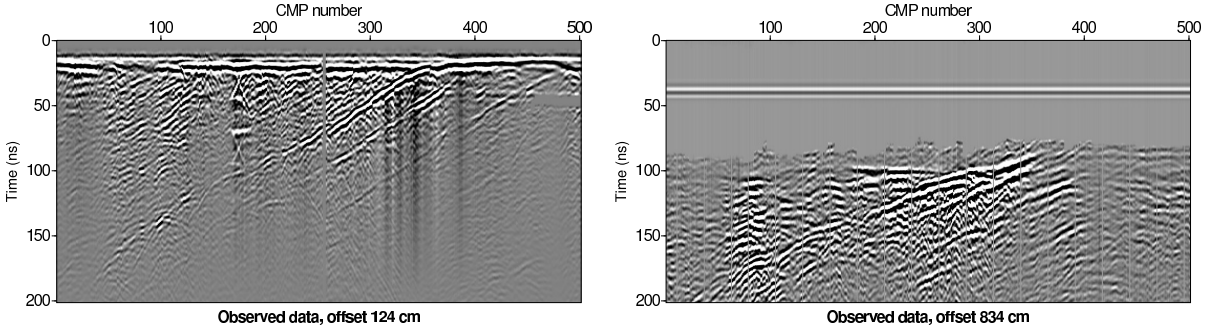


Figure 3.30: *Muted data for offsets 124 and 834 cm. At offset 124 cm, the reflected air-wave has been muted only for CMP* > 450*. The vertical drips at* $300 <$ *CMP* < 400 *are very low-frequency artefacts due to the 3D-to-2D conversion (they do not have consequences on the frequency band used for inversion). At offset 834 cm, note the synthetic direct air-wave and the remaining aerial signals, associated to reflected-refracted waves which the simulation should be able to reproduce (see Fig. 3.28).*

3D-to-2D conversion

For lack of anything better, I apply here again a 3D-to-2D transformation inspired from the wave solution in a homogeneous medium:

$$d_{obs}^{2D}(t) = \sqrt{\frac{2\pi v(t)^2 t}{\omega}} e^{\imath\pi/4}\, d_{obs}^{3D}(t) \tag{3.20}$$

which corresponds to eq. (21) in Bleistein (1986), with a ray parameter $\sigma = v^2 t$. Note however that I take into account the velocity contrast at the air-ground interface, using a time-varying velocity $v(t)$. Otherwise, it is not possible to account for the respective amplitudes of the direct air-wave and ground-wave, i.e. for the energy distribution between air and ground. Since this energy distribution also strongly depends on non-controlled parameters as the antenna height above ground, the suited 3D-to-2D conversion is intimately related to the estimation of the source, that I detail in the next paragraph. To consider the velocity contrast at the air-ground interface, the velocity $v(t)$ in eq. (3.20) is defined as

$$v(t) = \begin{cases} v_o & \text{if } t < t_{AW} + \Delta t_{AW}, \\ v_1 + (v_o - v_1)\cos^2\left(\pi \frac{t-(t_{AW}(x)+\Delta t_{AW})}{2\tau}\right) & \text{if } t_{AW} + \Delta t_{AW} \le t \le t_{AW} + \Delta t_{AW} + \tau, \\ v_1 & \text{if } t > t_{AW} + \Delta t_{AW} + \tau, \end{cases} \tag{3.21}$$

where v_o is the velocity in the air, v_1 the velocity in the ground, t_{AW} the arrival time of the direct air-wave (which depends on the offset x), $\Delta t_{AW} = 20$ ns the duration of the pulse and $\tau = 4$ ns

a tapper window. This definition formalises the fact that I consider a transition between air and ground just after the direct air-wave. Note that I could also consider a transition just before the arrival of the direct ground-wave but doing so increases the amplitudes of the reflected-refracted waves that I left during the muting step. As these events might be polluted with remaining aerial reflections, it is not desirable to enhance them. Moreover, the arrival time of the direct ground-wave t_{GW} is less straightforward to determinate than the one of the direct air-wave t_{AW}. For simplicity, I consider a homogeneous ground velocity $v_1 = 9$ cm/ns ($\varepsilon_r \simeq 11$). Of course, this value should be refined but in a first approach, the amplitude ratio between the aerial and the ground signals is mainly driven by the velocity contrast between air and ground.

3.2.3.2 Estimation of a source wavelet

In frequency-domain FWI schemes (Pratt, 1999), the source signature is classically estimated through a linear inversion as

$$S(\omega) = \frac{\mathbf{d}_{cal}(\omega)^\dagger \mathbf{d}_{obs}^{2D}(\omega)}{\mathbf{d}_{cal}(\omega)^\dagger \mathbf{d}_{cal}(\omega)}, \tag{3.22}$$

where $S(\omega)$ is the complex frequency spectrum of the source wavelet, $\mathbf{d}_{cal}(\omega)$ are frequency-domain synthetic data computed in some subsurface model and $\mathbf{d}_{obs}^{2D}(\omega)$ are the observed data in the frequency domain, after 3D-to-2D transformation. This estimation is either performed in the initial model and kept fixed during the inversion (Pratt, 1999), or re-iterated during the inversion process (Belina et al., 2012a), or even included in the reconstructed parameters (Busch et al., 2012). Of course, there is a trade-off between the quality of the model where synthetics are computed and the resulting source estimation: if the model does not reproduce the observed data, the source estimation will be biased.

Since the initial models we have are not accurate enough to reproduce all the observed events and estimate a satisfactory source wavelet, I rather chose to estimate the source wavelet from selected traces where the blocky model accurately predicts the arrival time of the direct ground-wave, i.e. for CMP number 248 to 292 (see Fig. 3.29). Since the late arrivals are poorly predicted, I use an exponential time damping to damp the data and select only the direct arrivals (air-wave and ground-wave) and I restrict the data to the offsets 124 to 283 cm to not include other aerial events in the observed data:

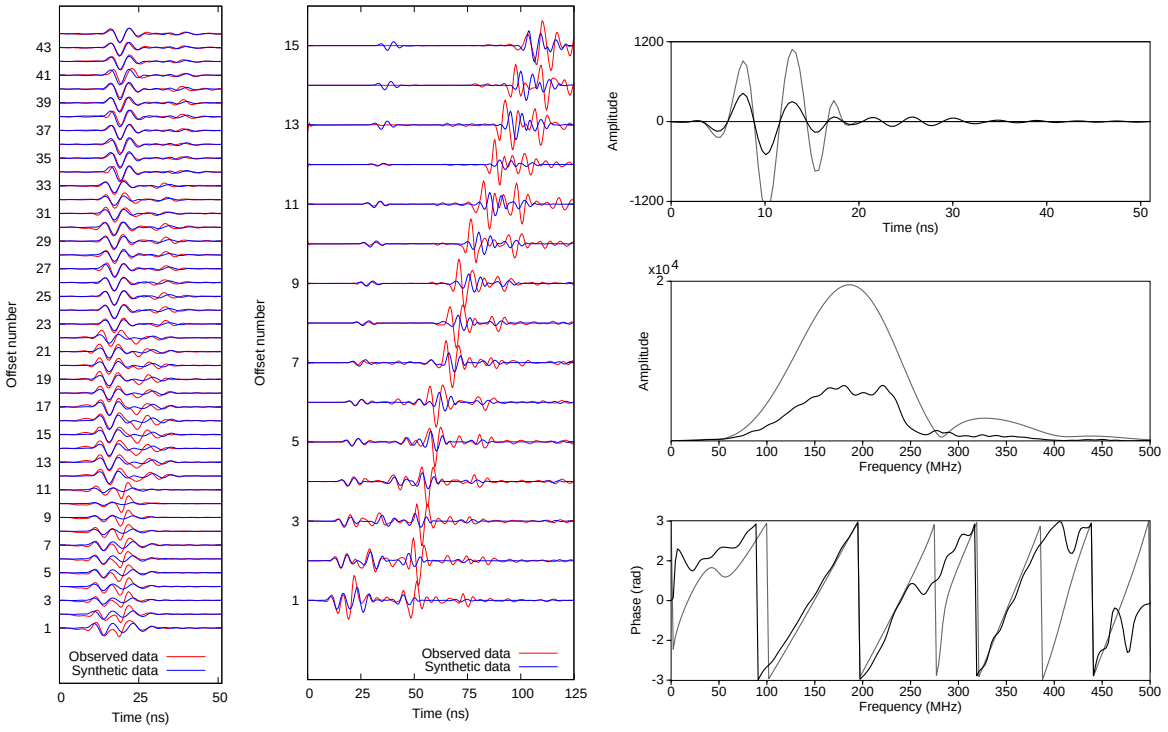
$$d_{obs/cal}^{used}(x,t) = e^{-\left(t - t_{AW}(x)\right)/\tau} d_{obs/cal}(x,t), \tag{3.23}$$

with $\tau = 10$ ns. Using such a time damping is more convenient for selecting the direct arrivals than a simple mute because the end of the ground-wave signal often overlaps with other events and thus can not be easily determined. On the other hand, the drawback of this time damping is that the estimation is mainly based on the air-wave, in less extent on the damped ground-wave, and not at all on the reflected waves. We can wonder if this estimation is relevant in the case of shielded antennas, where the shielding is expected to induce a distortion of the signal emitted at 90° (parallel to the air-ground interface) compared to the signal emitted vertically in the ground (Diamanti et al., 2013). But estimating a source signal from a reflection would require a more accurate model where the reflectivity of the interface is well known. It also poses the problem of properly isolating the reflected event amongst the surrounding diffractions.

Finally, I already mentioned that the estimation of the source also depends on several other non-controlled parameters, such as the vertical position of the effective antenna in the

shielding (which is a black box for the user) and as the conductivity value in the ground. As a consequence, I perform the estimation of the source in a variety of configurations, with conductivity values σ ranging from 1 to 6 mS/m and antenna heights z_{src} varying from 0 to 5 cm above ground. Since the ultimate aim of this processing is to make the synthetic data comparable to the observed ones, I then select the configuration of conductivity, antenna height and source estimation that provides the best match between the observed data and the synthetic data computed in the blocky model.

A good compromise has been found for the combination $z_{src} = 0$ and $\sigma = 4$ mS/m: I thus retain these values for the inversion. The corresponding estimated source is shown in Fig. 3.31, where it can be seen than the synthetic air-wave and ground-wave well match the observed ones for offsets number 2 to 7 (180 to 438 cm). At offset number 1 (124 cm), we observe a discrepancy in the amplitude of the ground-wave, which is probably due to the applied 3D-to-2D transformation that over-amplifies the observed ground-wave (because it overlaps with the end of the air-wave, see eq. 3.20). The amplitude of the reflected events are not well reproduce, which can be due to the fact that the blocky model does not present the right reflectivity, or to the effect of the antenna shielding. In the first hypothesis, we might expect the inversion to retrieve the right reflectivity when updating the model.



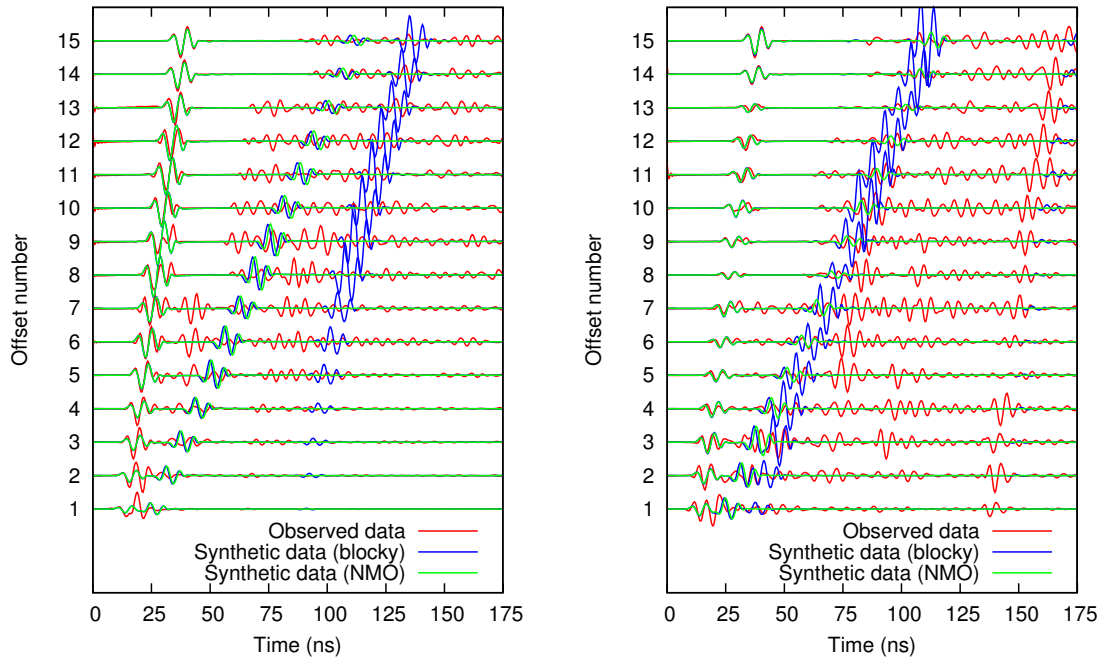
(a) Data fit (damped). **(b)** CMP 300. (c) Source wavelet and spectrum.

Figure 3.31: *Estimation of a time-domain wavelet using selected, damped traces. (a) Fit of damped data (86 traces in total, one out of two is shown). (b) CMP 300 in the blocky model, using $\sigma = 4$ mS/m and $z_{src} = 0$. (c) Source wavelet and its spectrum (black: estimated damped source, grey: re-inflated source after mute at 20 ns). In (a) and (b), data are normalised trace-by-trace with the observed amplitudes as reference (observed and synthetic amplitudes are comparable within a trace).*

Observed vs. synthetic time-domain data in the initial models

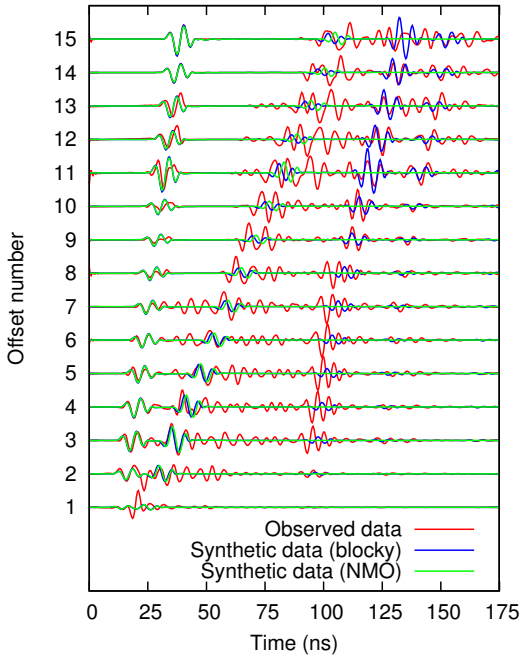
Once we have estimated a source wavelet, we can rigorously compare the time-domain observed data to synthetic data computed in the NMO and the blocky permittivity models of Figs 3.23(b) and 3.26, which are our likely candidates as initial models for the FWI. Such a comparison is presented in Fig. 3.32 for the CMP numbers 1, 100, 200, 300, 400 and 500 in order to roughly span the whole profile. Both models well explain the arrival time of the direct ground-wave, which suggests that their superficial velocities is correct. The blocky model well explains the arrival times of the main reflections at CMP 200, 300 and 400. The reflected amplitudes well match for CMP 200 and 300 at large offsets, but are over-estimated at short offsets and for other CMP gathers, suggesting an effect of the antenna radiation pattern and/or lateral variations in the ground conductivity.

Nevertheless, the good kinematic compatibility between observed and synthetic data in both models confirms that these models can be used for starting the FWI process. Amplitudes are expected to be matched in the inversion by adapting the reflectivity.

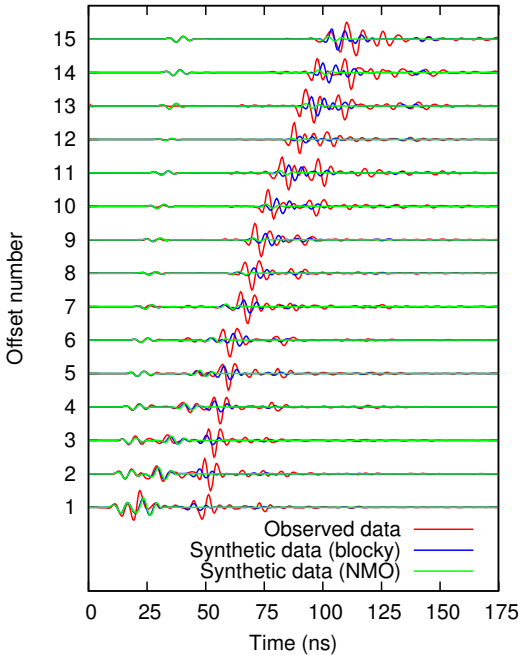


(a) CMP 1. **(b)** CMP 100.

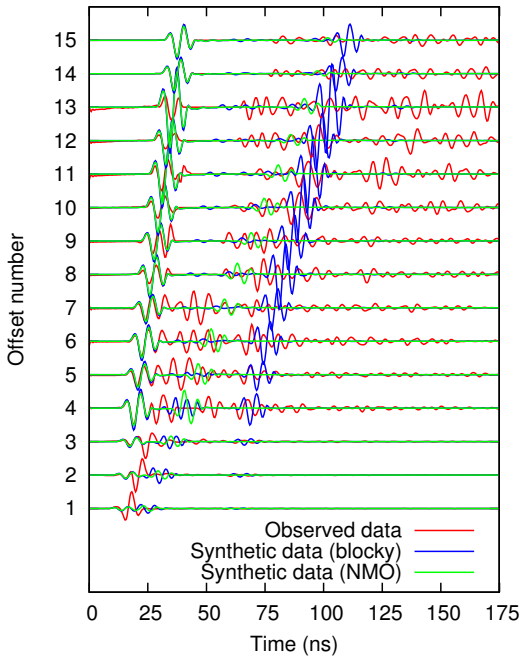
Figure 3.32: *Observed vs. synthetic radargrams in the smooth NMO and in the blocky models (normalised trace-by-trace with observed amplitudes as reference, amplitudes are comparable).*



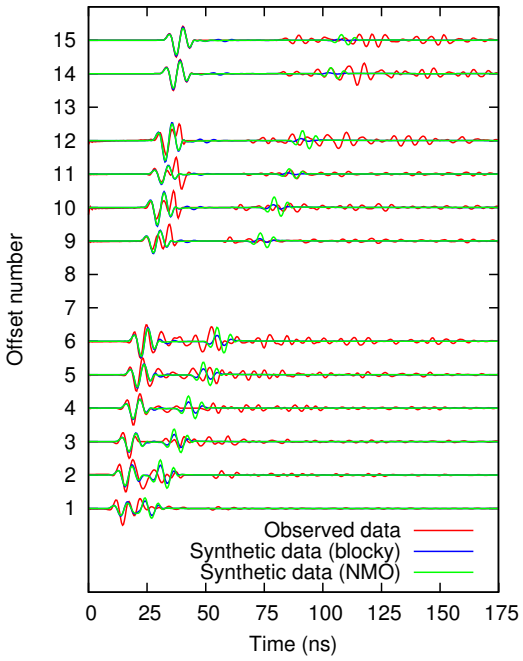
(c) CMP 200.



(d) CMP 300.



(e) CMP 400.



(f) CMP 500.

Figure 3.32: (continuation) *Observed vs. synthetic radargrams in the smooth NMO and in the blocky models (normalised trace-by-trace with observed amplitudes as reference, amplitudes are comparable).*

Partial conclusion on the pre-processing steps: Sources of error on the amplitudes

Concerning the reflected waves:

- Reflected amplitudes are not matched, suggesting that the permittivity contrasts are not adequate, or that the interfaces act as thin layers displaying a strong reflectivity without permittivity contrasts at the scale of the wavelength. Permittivity contrasts are expected to be reconstructed by the inversion. The match of the reflected amplitudes should be controlled after inversion to discriminate between both effects.

Concerning the direct (and reflected) waves:

- 3D-to-2D transformation: The consideration of the velocity contrast at the air-ground interface is critical to explain the relative amplitudes of the direct air-wave and ground-wave. The applied transformation is still highly improvable. In particular, lateral velocity variations should be considered in the transfer function.
- Unknown conductivity and antenna height: I choose the values that provide the best fit when comparing observed (corrected) data and synthetic data in the blocky model convolved with the estimated source, but a homogeneous conductivity does not account for the lateral amplitude variations observed in the data.
- Radiation pattern of the shielded antennas: The importance of this effect is difficult to determined because there is a strong trade-off between the effect of the shielding and i) the AVO response due to conductivity, on one hand, and ii) the reflectivity, on the other hand. It seems however that the amplitudes of the reflected waves are better fitted at long offsets than at short offsets, where they are under-estimated (see e.g. CMP 200 in Fig. 3.32c). This is an indication that the shielding has a significant effect on the recorded AVO of the signal. Indeed, this observation can be explained in two ways:
  1. My source estimation under-evaluates the energy radiated into the ground but the conductivity value is actually larger than I estimated. So the synthetic amplitudes, initially under-estimated, are under-attenuated during propagation and can fit the observed data at large offsets.
  2. My source estimation under-evaluates the energy radiated *vertically* into the ground because it is based on the direct waves (air-wave and ground-wave), that propagate at nearly 90° along the air-ground interface. Thus, synthetic data convolved with the estimated source better fit the data at long offsets because they correspond to large emission/reception angles. But the amplitudes of signal emitted/recorded at short angles are not well reproduced because the shielding introduces a bias in the recorded AVO of the signal.

  The fact that the conductivity value I estimated roughly explain the AVO of the direct ground wave, as well as the one of superficial reflected events (see Fig. 3.32a,c,d,f), suggests that the second hypothesis is the most plausible. Inversion tests will tell us if the observed reflectivity can be explained by updating the permittivity model, or whether a refined estimation of the conductivity and/or a more detailed description of the source radiation pattern are required.

3.2.3.3 Frequency-domain analysis

I now present the data in the frequency domain, since it is in the frequency domain that I will invert them. A first question concerns the adequate frequencies to consider for inversion. In the frame of frequency-domain FWI, it is usual to start from low frequency components and to introduce progressively higher frequencies in successive inversion steps (Pratt and Worthington, 1990; Pratt, 1999). In Section 2.2, we have seen that GPR data inversion takes benefit of considering several frequencies simultaneously, and that the strategy of Bunks et al. (1995) was the most robust treating surface data. The strategy I apply here thus considers successive groups of cumulative frequencies, whose bounds and sampling must be defined.

First of all, we can look at the spectrum of the data, which is presented in Fig. 3.33, both for raw 3D data and for processed, virtually 2D data. 3D data exhibit a peak frequency at 168 MHz. 2D data exhibit a peak at 162 MHz and a low-frequency artefact below 20 MHz. For both data sets, most of the energy is comprised between 50 and 350 MHz. However, looking at these spectra does not indicate precisely what should be considered as the minimal frequency for the inversion.

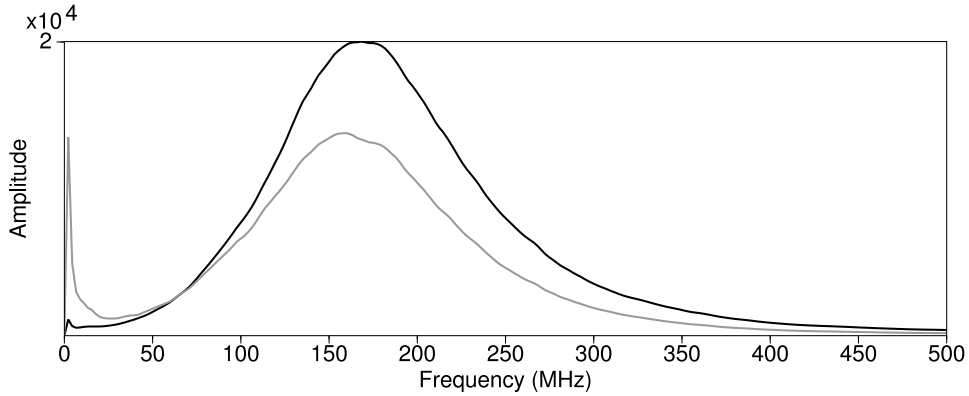
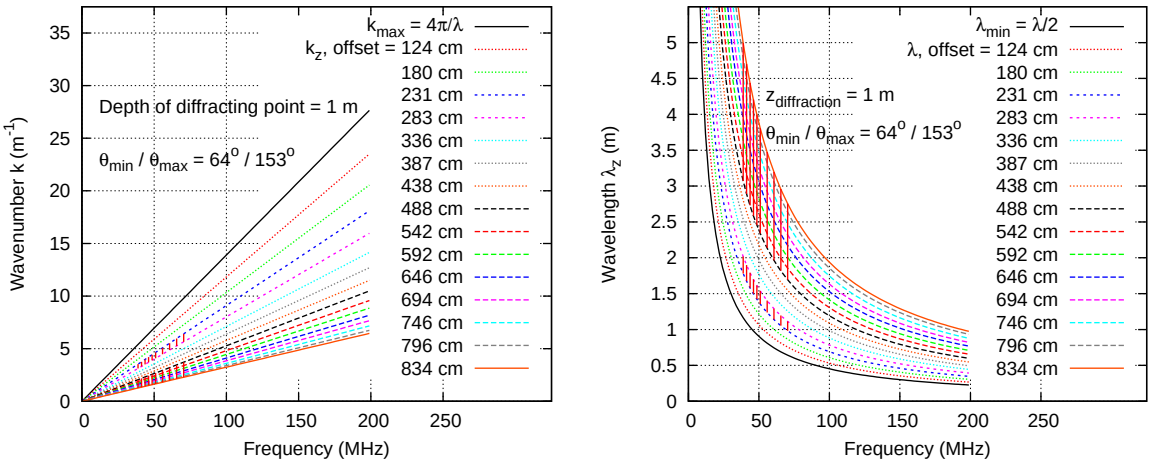
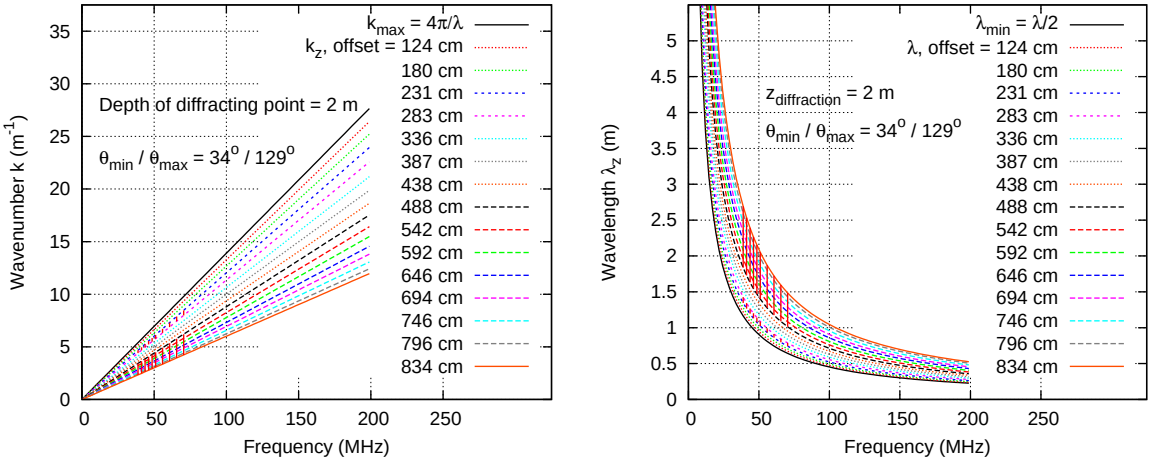


Figure 3.33: *Mean amplitude spectra of 3D (black line) vs. 2D-converted data (grey line).*

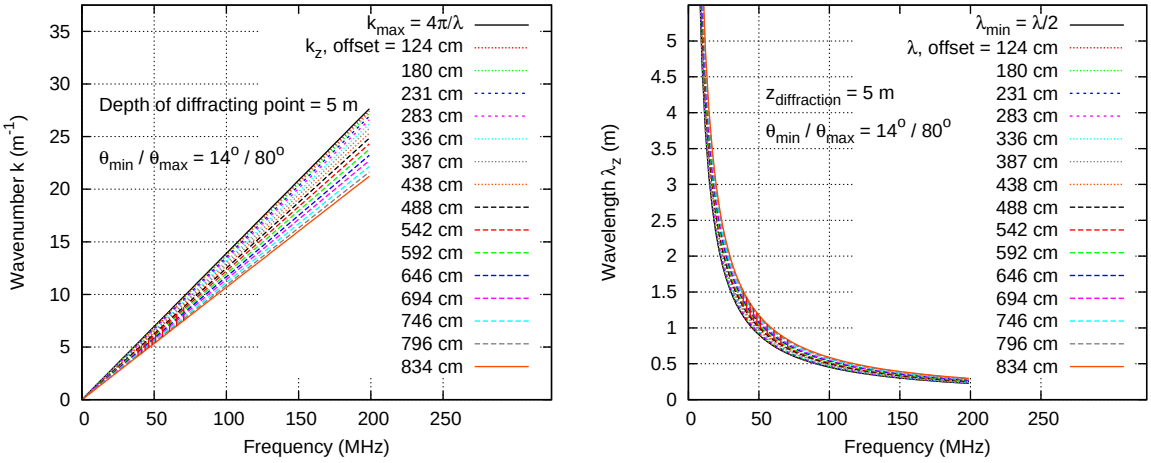
As already discussed in Section 2.3.1, a general guiding rule to select the inverted frequencies has been given by Sirgue and Pratt (2004), based on the wavenumbers that can be imaged at a given frequency, according to the acquisition setup. In Fig. 3.34, I present the frequency-wavenumber diagrams corresponding to the acquisition setup, considering different depths of diffracting points (1, 2 and 5 m). I express it both in terms of wavenumbers and of wavelengths, the latter being more intuitive. In dotted lines, I have represented the offsets that could not be used in the inversion because of the overlapping of the reflected air-wave with the direct ground-wave signal. As this analysis is based on a planar assumption, I should underline that it provides only a crude approximation of the expected wavenumber coverage and vertical resolution. Nevertheless, it gives an idea of the frequency sampling needed to ensure a correct wavenumber coverage in the reconstructed image. It can be observed in Fig. 3.34 that the acquisition redundancy is quite low: To avoid gaps in the covered wavenumbers, frequency sampling must be very dense, and we have interest in beginning the inversion with the lowest possible frequency.



(a) Diffracting point at depth $z = 1$ m.



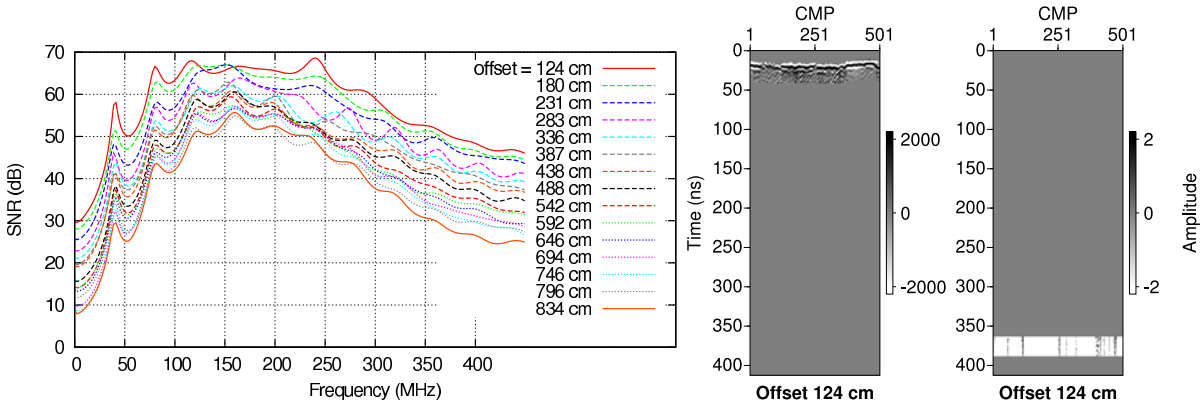
(b) Diffracting point at depth $z = 2$ m.



(c) Diffracting point at depth $z = 5$ m.

Figure 3.34: *Imaged vertical wavenumber k_z (left) and vertical resolution λ (right) as a function of frequency and offset, for different depths of diffracting points. Dotted lines correspond to unusable offsets due to the overlap of the parasite reflected air-wave with the direct ground-wave.*

To determine this lowest possible frequency (in particular, to see if it is possible to consider frequencies below 50 MHz), Fig. 3.35 presents the signal-over-noise ratio of the data with respect to offset and frequency. According to this figure, the SNR is similar at 40 and 60 MHz, so we can reasonably choose 40 MHz as the minimal frequency to be considered in the inversion.



(a) SNR $= 10\log(E_{signal}/E_{noise}) = 20\log(A_{signal}/A_{noise})$. **(b)** Selected signal. **(c)** Selected noise.

Figure 3.35: *(a) SNR as a function of offset and frequency, obtained by computing the energy for each frequency of the common-offset sections (stacked over CMP) in 25-ns-long time windows (b,c). "Signal" is chosen as the 25-ns-long window following the direct ground-wave arrival. As noise, I choose the first-to-last 25 ns (the end of the trace, before non-physical noise due to t_o-correction).*

Fig. 3.36 shows common-offset sections filtered in the frequency range 40-70 MHz, which is the lowest bandwidth I envisage for inversion. At short offsets ($\leq$ 180 cm), the filtered data seem to contain very few information of the subsurface: they contain mainly horizontal signal (ringing), probably enhanced by the inaccurate 3D-to-2D conversion (overlap of the direct ground-wave and air-wave). Consequently, I will not consider these offsets for inversion. In

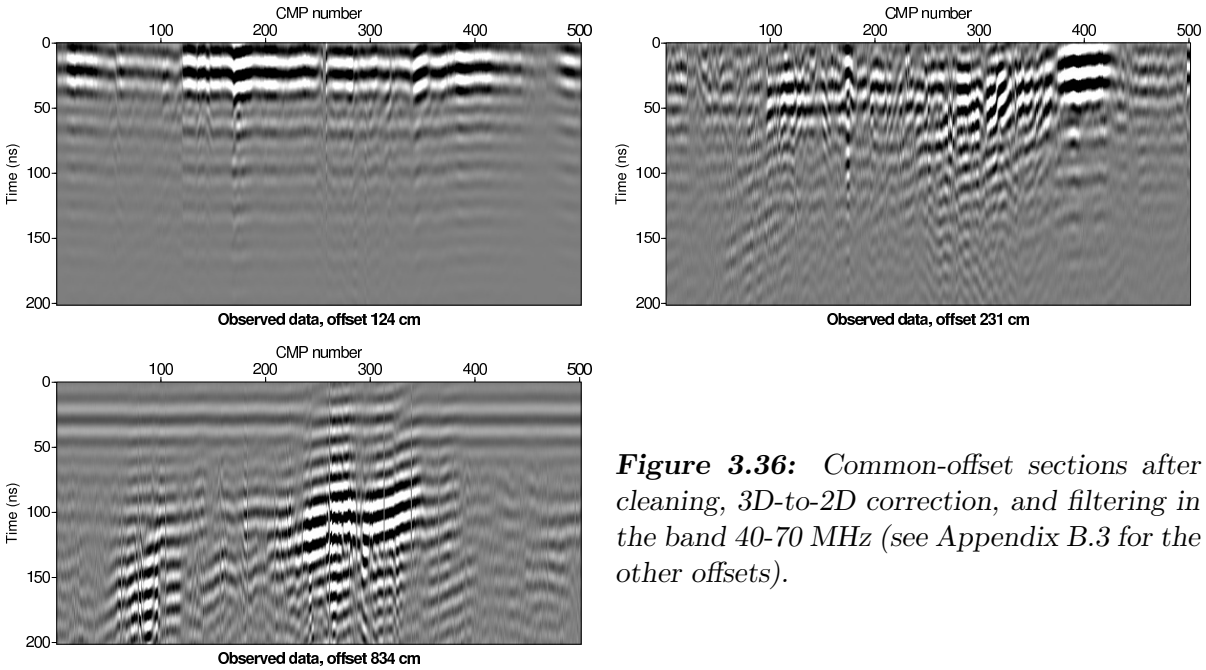
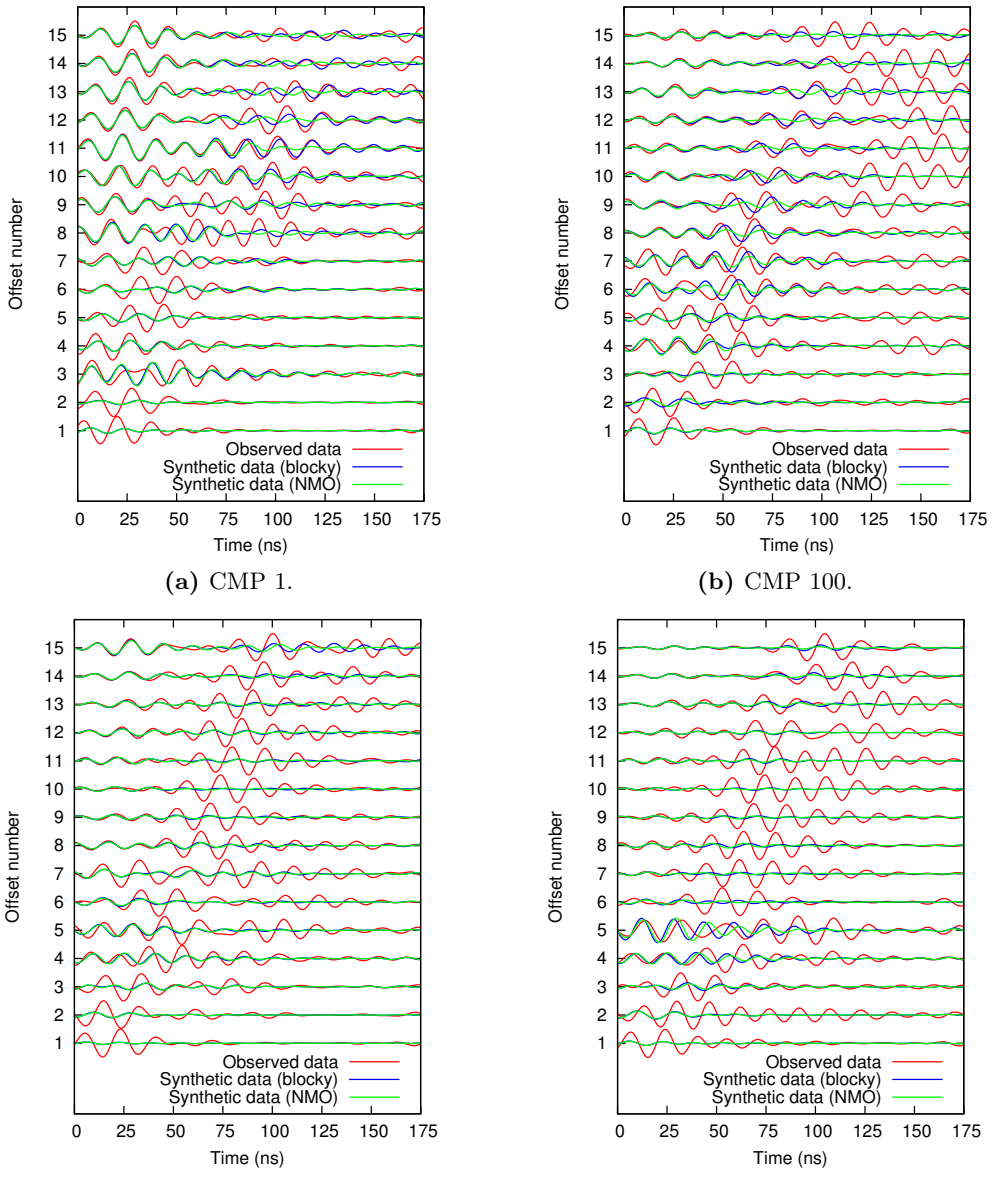


Figure 3.36: *Common-offset sections after cleaning, 3D-to-2D correction, and filtering in the band 40-70 MHz (see Appendix B.3 for the other offsets).*

my inversion tests, I could verify that considering these short offsets introduces non-physical horizontal structures in the reconstructed image. As the direct ground-wave is artificially over-amplified by the 3D-to-2D conversion, keeping these offsets for inversion also disturb the source estimation when optimising the source signature in the initial model by using all data, which further enhances the creation of horizontal artefacts in the image.

In Fig. 3.37, I present a comparison of observed and synthetic data after filtering in the selected frequency bandwidth 40-70 MHz. As in Fig. 3.32, synthetic data are computed either in the smooth NMO model of Fig. 3.23(b) or in the blocky model of Fig. 3.26. From Fig. 3.37,



(a) CMP 1. **(b)** CMP 100.

(c) CMP 200. **(d)** CMP 300.

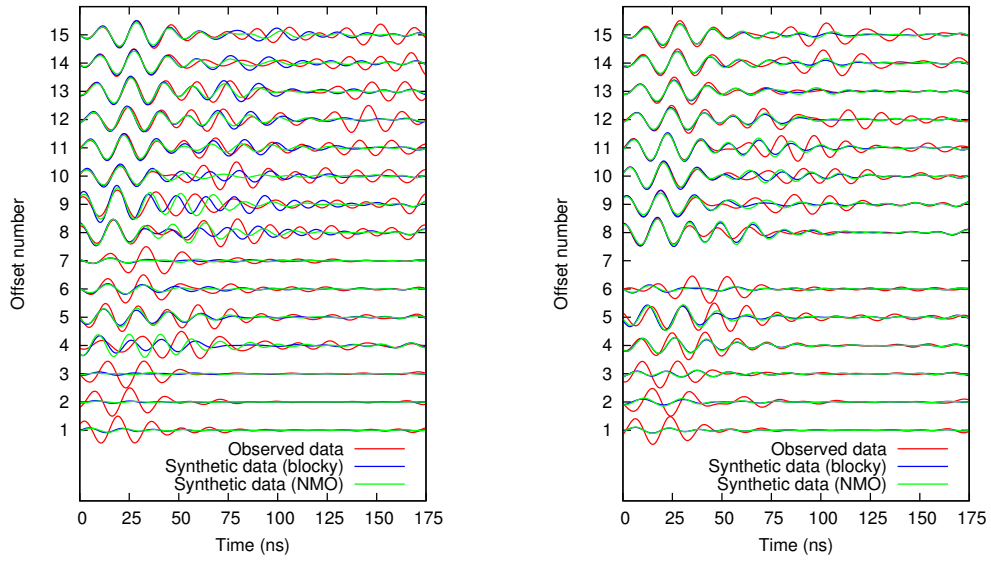
Figure 3.37: *Observed vs. synthetic radargrams in the smooth NMO and in the blocky models, filtered between 40 and 70 MHz (normalised trace-by-trace with observed amplitudes as reference, amplitudes are comparable).*

it is clear that the initial models are not accurate enough and we may expect cycle-skipping to occur in some parts of the data (e.g. CMP 400 offsets 8-10). On the other hand, synthetic data display a good match with the observed ones in other portions (e.g. CMP 100, CMP 1 offsets 10-11, CMP 400 offsets 11-13). The blocky model presents generally a better match than the smooth NMO model (CMP 1 and 100).

Finally, Figs 3.38 and 3.39 present the observed data in the frequency-domain, at 40 and 70 MHz, respectively (more precisely, 38.8 and 70.4 MHz, the frequency step related to the temporal sampling being 2.43 MHz). Again, observed data are compared with synthetic data computed in the blocky and in the smooth NMO models. Data at 40 MHz are quite noisy but some trends can be observed with respect to offset and CMP, in particular where the main reflector arises (CMP $\simeq$ 350). This trend can also be seen in the synthetic phase computed in the initial blocky model but it is far less visible in the smooth NMO model. The same observations can be made at 70 MHz (Fig. 3.39), for which the blocky model seems to well match the pattern of the observed phases, in spite of some phase difference. As the initial models do not predict all the observed events, their frequency-domain amplitudes are lower than the observed ones (see the amplitude scales).

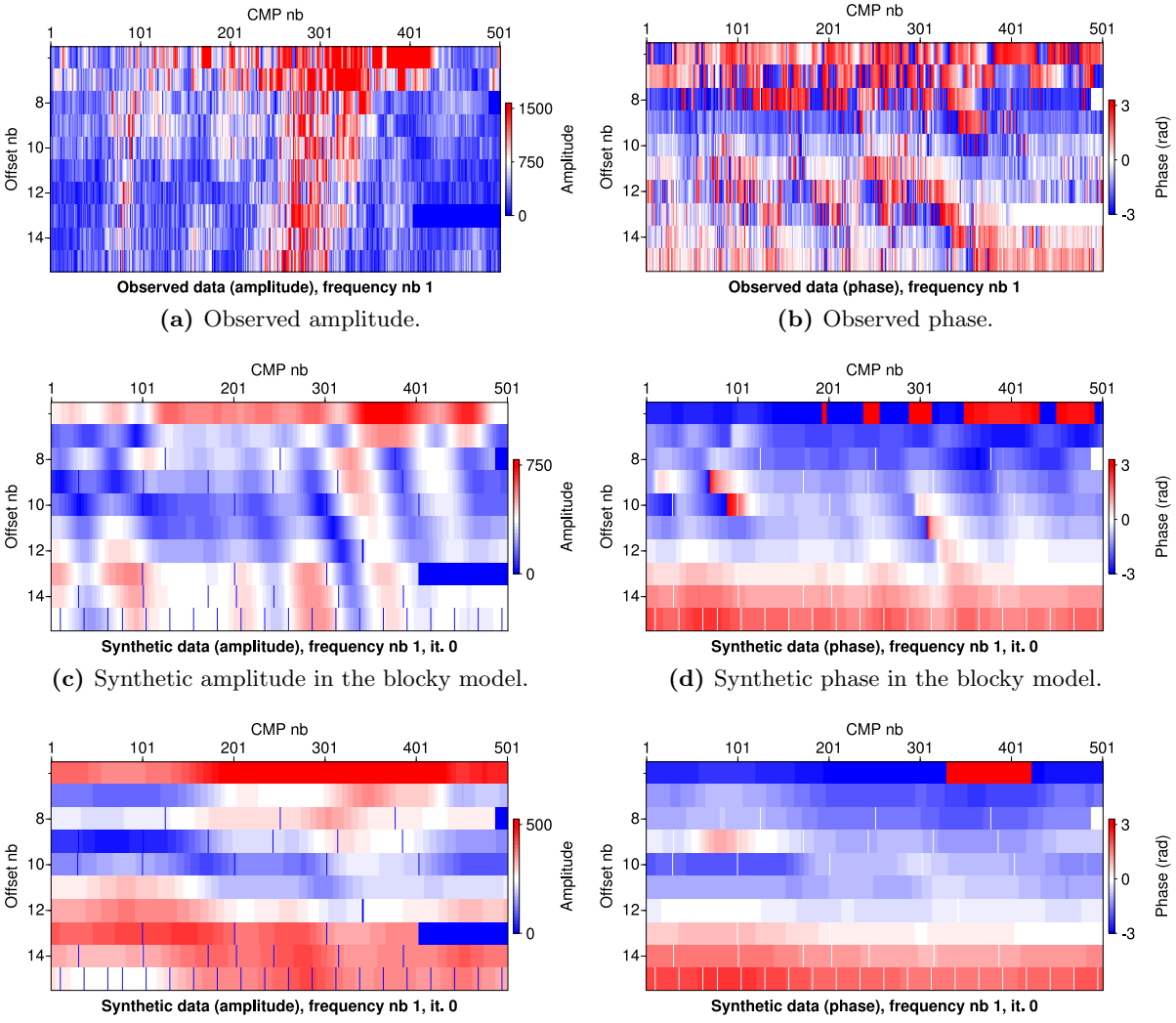
Note that Figs 3.38 and 3.39 present the frequency-domain data as they are used for inversion. In particular, I discarded the undesired offsets (124, 180, 336, 387 and 438 cm). The first two offsets shown are thus 231 and 283 cm, followed by offsets number 8 to 15, as indicated, i.e. 488 to 834 cm. In addition, I apply a data weighting matrix such that each filtered common-offset section has the same RMS amplitude. Relative amplitudes within a common-offset section are conserved to not amplify noise due to low-amplitude traces. I apply the same weighting in the following inversions.

For information, I provide in Appendices B.3 and B.4 the complete filtered data set and frequency-domain data in the whole frequency bandwidth (40 to 300 MHz), respectively.



(e) CMP 400. **(f)** CMP 500.

Figure 3.37: (continuation) *Observed vs. synthetic radargrams in the smooth NMO and in the blocky models, filtered between 40 and 70 MHz (normalised trace-by-trace with observed amplitudes as reference, amplitudes are comparable).*



(a) Observed amplitude.

(b) Observed phase.

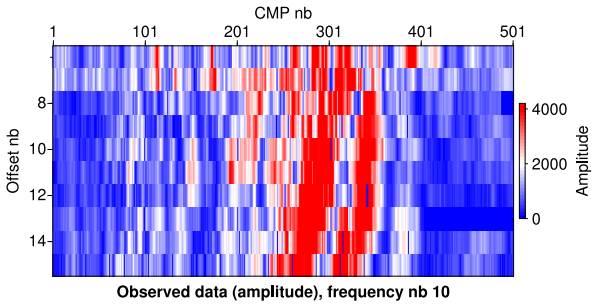
(c) Synthetic amplitude in the blocky model.

(d) Synthetic phase in the blocky model.

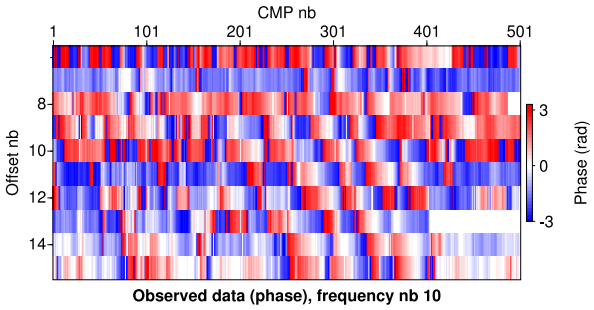
(e) Synthetic amplitude in the smooth NMO model.

(f) Synthetic phase in the smooth NMO model.

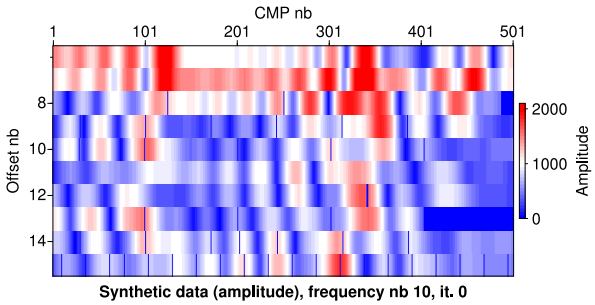
Figure 3.38: *Observed vs. synthetic data in the initial models, at frequency 40 MHz. Amplitudes are balanced with offset. Note the different amplitude scales.*



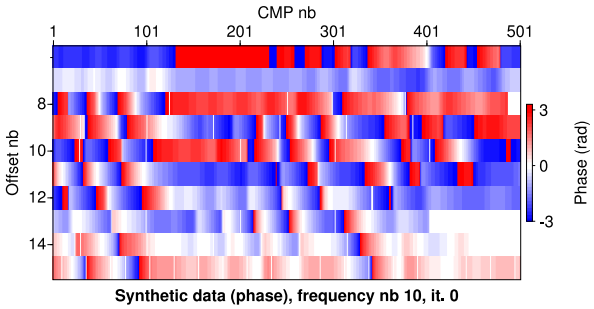
(a) Observed amplitude.



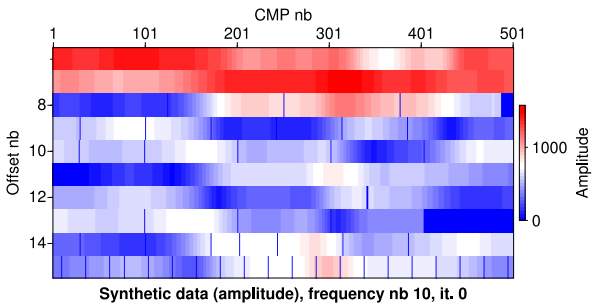
(b) Observed phase.



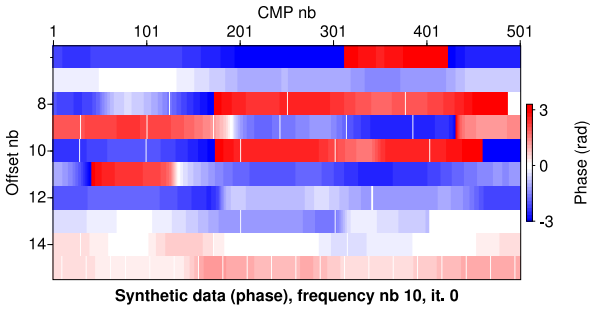
(c) Synthetic amplitude in the blocky model.



(d) Synthetic phase in the blocky model.



(e) Synthetic amplitude in the smooth NMO model.



(f) Synthetic phase in the smooth NMO model.

Figure 3.39: *Observed vs. synthetic data in the initial models, at frequency 70 MHz. Amplitudes are balanced with offset. Note the different amplitude scales.*

3.2.4 Preliminary FWI results

Table 3.4 recaps the setup of the inversions I will now perform. I comment these settings in the following.

Frequencies (MHz)	
1^{st} group (10 freq.)	38.8, 41.3, 43.7, 46.1, 48.6, 51.0, 55.8, 60.7, 65.5, 70.4
2^{nd} group (13 freq.)	41.3, 46.1, 51.0, 55.8, 60.7, 65.5, 70.4, 75.3, 80.1, 85.0, 89.8, 94.7, 99.5
3^{rd} group (18 freq.)	41.3, ... , 99.5, 109, 119, 129, 138, 148
Offsets (cm)	231, 283, 488, 542, 592, 646, 694, 746, 796, 834
Total number of traces	4852 out of 7515
Total nb of source locations	3503 out of 5053

Table 3.4: *Inversion settings.*

Frequency sampling For the inversion, I consider 3 groups of cumulative frequencies. The first frequency bandwidth ranges between 40 and 70 MHz. It is then extended to 100 MHz and to 150 MHz, which is the peak frequency in the data spectrum (see Fig. 3.33). The selected frequencies are not regularly spaced. Following the FK-diagrams of Fig. 3.34, I choose a denser sampling for low frequencies than for high frequencies to ensure a good coverage of low wavenumbers. Moreover, the dense sampling of low frequencies partially compensate for the fact that they have less power in the spectrum than high frequencies (more rigorously, the data spectrum should be whitened).

Data selection I remind that I discard the shortest offsets 124 and 180 cm, which present a strong ringing and few valuable signal, as well as the offsets 336, 387 and 438 m, where the energetic reflected-refracted air-wave overlaps with the direct ground-wave. I thus retain 10 acquired offsets out of 15 in total. Further cleaning of individual traces leads to 4852 inverted traces out of 7515 in total.

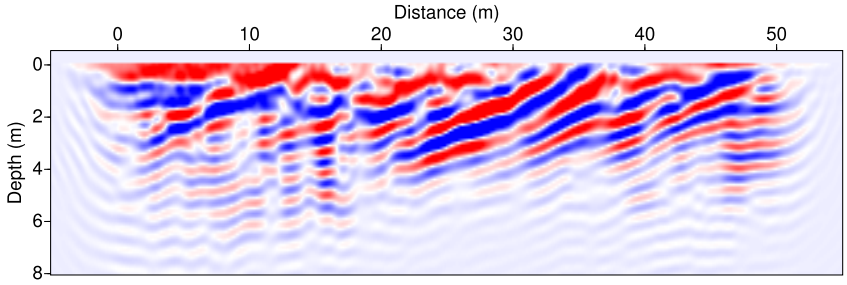
Misfit function As seen on Figs 3.38 and 3.39, frequency-domain data are noisy. Consequently, I choose a ℓ_1-norm misfit function which is supposed to be more robust with respect to noise, by giving less weight to high-amplitude outliers in the residuals (Brossier et al., 2009).

Data weighting Finally, I apply a data weighting to balance the amplitudes of the data with respect to offset, as already shown in Figs 3.38 and 3.39. Giving more weight to large offsets is expected to help for recovering low wavenumbers (but may also enhance the imprint of noise).

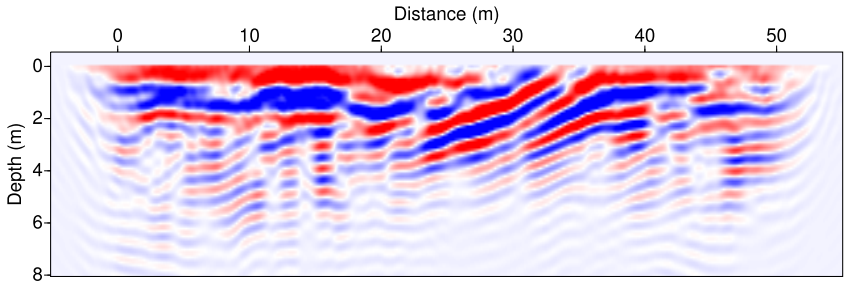
Source The source wavelet estimated in Section 3.2.3.2 (see Fig. 3.31) is fixed for the inversion. Alternatively, I also perform inversions by estimating the source at the first iteration using all inverted data. I may comment these results but I do not show them (they are quite similar as the following ones).

In the following, I detail the inversion results for the first frequency bandwidth (40 to 70 MHz). Then I present the results for the subsequent frequency groups.

As a first quality control on the inversion process, Fig. 3.40 shows the gradients computed in the starting models, using either the smooth NMO model or the blocky model. In Fig. 3.40, both gradients present some discontinuities, indicating that the inversion will locally suffer



(a) Gradient in the initial blocky model.



(b) Gradient in the smooth NMO model.

Figure 3.40: *Gradients in the initial blocky model (a) and in the smooth NMO model (b), for the first frequency group (40 to 70 MHz). Note the discontinuities due to cycle-skipping at* $10 < x < 20$ *m,* $z > 4$ *m and at* $x \simeq 42$ *m in (b). The z-axis is exaggerated twice.*

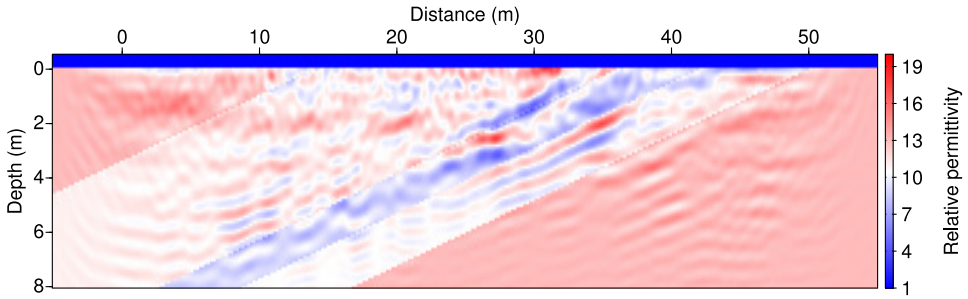
from cycle-skipping. They also present similar dipping structures, corresponding to the main reflectors. Note however that the gradient computed in the blocky model presents strong anomalies at the position of pre-existing interfaces, indicating that they are not well placed.

Fig. 3.41 shows the permittivity models reconstructed after 10 iterations of the inversion process for the first frequency group (40 to 70 MHz). As a complement, to better visualise the modifications added by the inversion to the initial model, Fig. 3.42 shows the difference between the permittivity reconstructions of Fig. 3.41 and the initial blocky and smooth NMO models. The reconstructed permittivity models obtained from the blocky and from the smooth NMO models are quite similar. In particular, the inversion was able to add the main reflectors to the smooth NMO model at the same locations as in the blocky model. This tends to confirm the good kinematic compatibility of these models but also suggests that FWI does not need pre-existing reflectors in the starting model. On the contrary, designing a blocky initial model might introduce biases in the inversion if the reflectors are not well placed (it is the case on the left of Fig. 3.41a), although FWI seems also able to shift mislocated interfaces, at least in a certain extent (see Fig. 3.41a, at $x = 30$ m, $z \simeq 2$ m). Finally, pre-existing reflectors keep an imprint in the final image, which makes its interpretation more difficult.

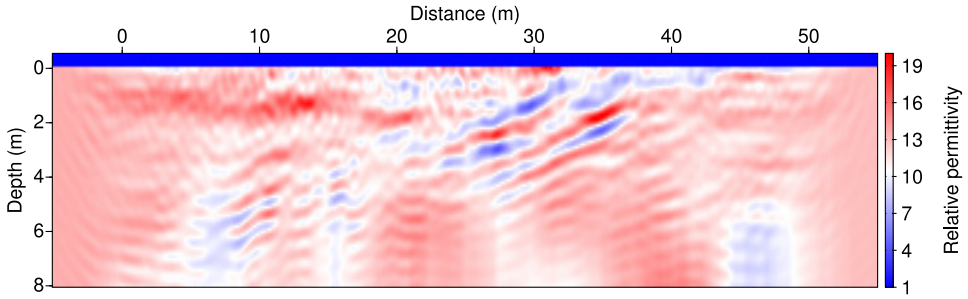
As a complement, it should be mentioned that performing the inversion by estimating the source from the inverted data in the initial model yields nearly similar images, suggesting that the considered pre-processing (including data exclusion) enables a consistent comparison between observed and synthetic data since few structures are put in the source.

Figs 3.43 and 3.44 show the data fit reached at iteration 10 for the inverted frequencies 40 and 70 MHz. Consistently with the similar permittivity reconstructions, the smooth NMO and the blocky starting models provide very similar final data fits, although the smooth NMO model produces initial data much farther from the observed ones (remember Figs 3.38 and 3.39).

Data fit is particularly satisfying at frequency 70 MHz because it has a larger weight in the optimisation (since data are not whitened). At 70 MHz, both amplitudes and phases display a satisfactory match.

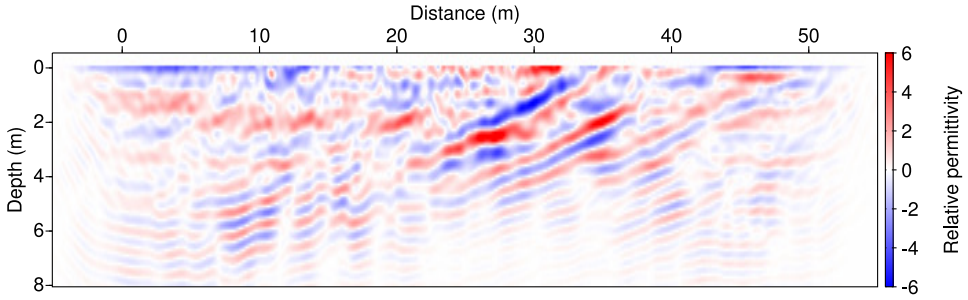


(a) Starting from the blocky model. Misfit decrease of 29%.

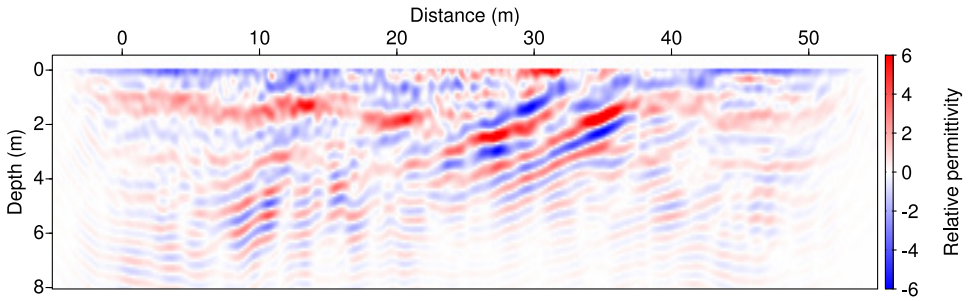


(b) Starting from the smooth NMO model. Misfit decrease of 30%.

Figure 3.41: *Permittivity reconstructions after 10 iterations (first frequency group).*

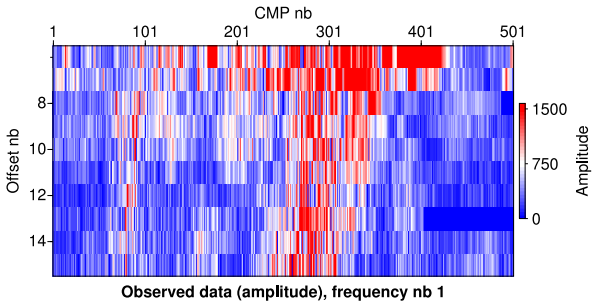


(a) Starting from the blocky model. Misfit decrease 29%.

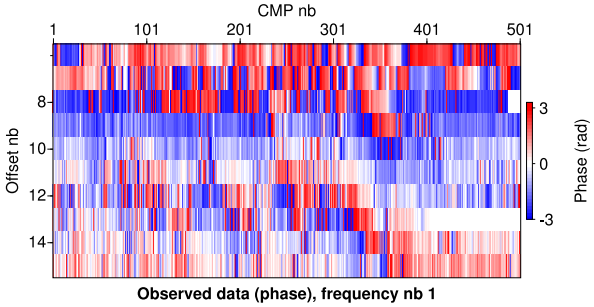


(b) Starting from the smooth NMO model. Misfit decrease 30%.

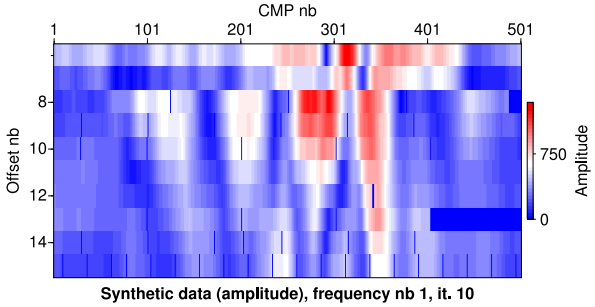
Figure 3.42: *Cumulative perturbations added to the initial models after 10 iterations (first frequency group, 40 to 70 MHz).*



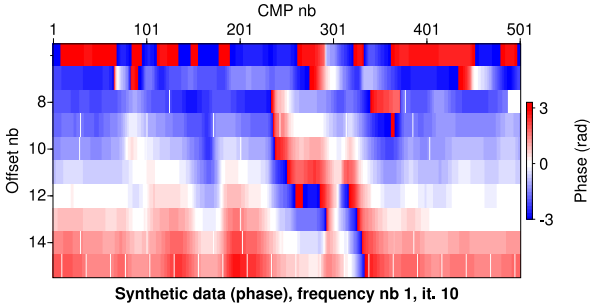
(a) Observed amplitudes.



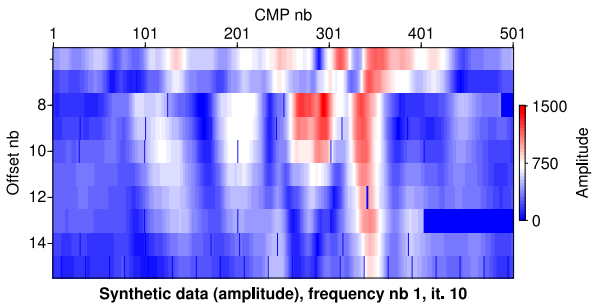
(b) Observed phases.



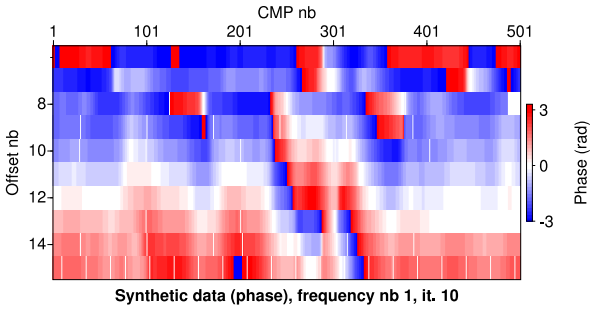
(c) Final amplitudes, starting from the blocky model.



(d) Final phases, starting from the blocky model.

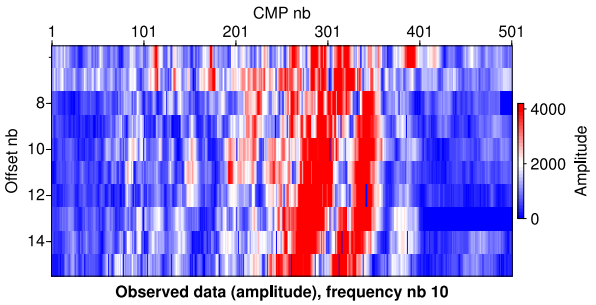


(e) Final amplitudes, starting from NMO model.

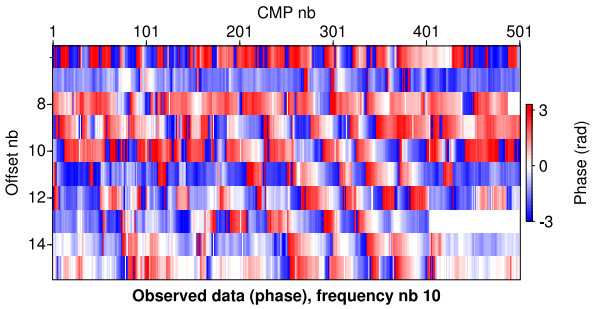


(f) Final phases, starting from NMO model.

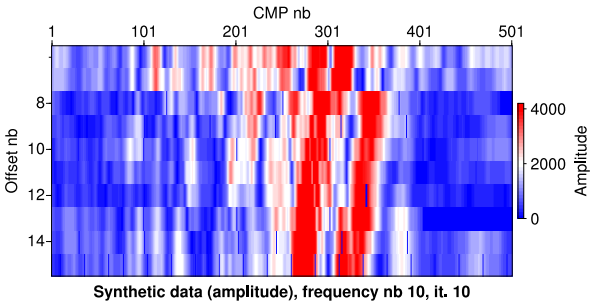
Figure 3.43: *Observed vs. synthetic data in the reconstructed models at frequency 40 MHz, after the inversion of the 1st frequency group (40 to 70 MHz).*



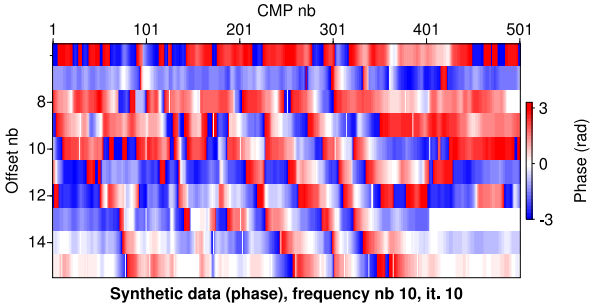
(a) Observed amplitudes.



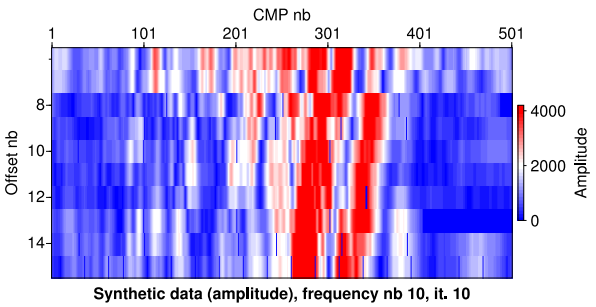
(b) Observed phases.



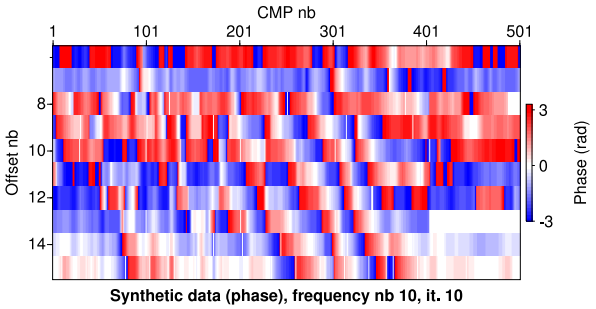
(c) Final amplitudes, starting from the blocky model.



(d) Final phases, starting from the blocky model.



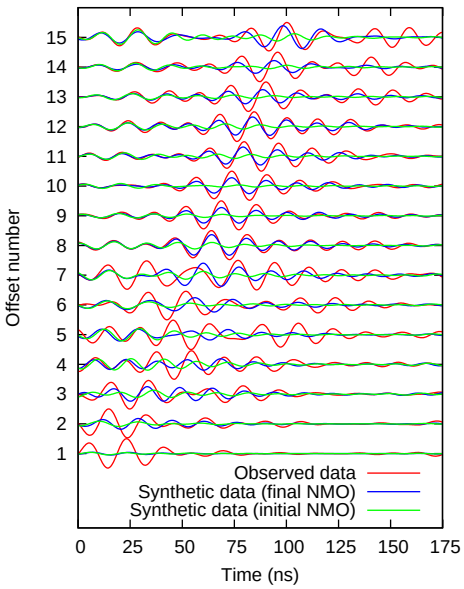
(e) Final amplitudes, starting from NMO model.



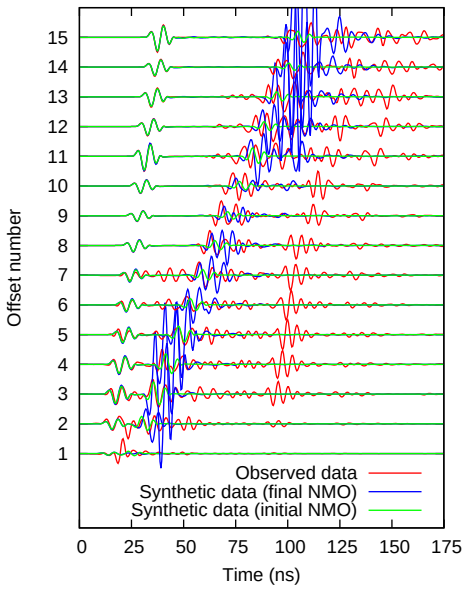
(f) Final phases, starting from NMO model.

Figure 3.44: *Observed vs. synthetic data in the reconstructed models at frequency 70 MHz, after the inversion of the 1st frequency group (40 to 70 MHz).*

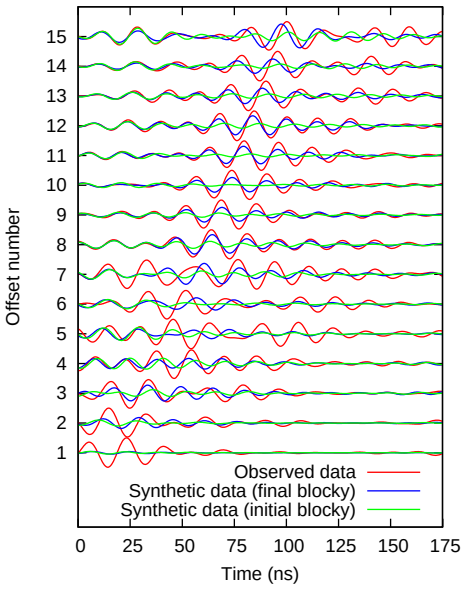
Finally, Fig. 3.45 shows the data fit in the time-domain between observed and synthetic data computed both in the initial and in the reconstructed models after inversion of the first frequency group, using either the NMO model or the blocky model. CMP number 200 is shown. Time-domain data fit for others CMP gathers are presented in Appendix B.5. Consistently with the frequency-domain data fit, a good agreement between observed and synthetic data



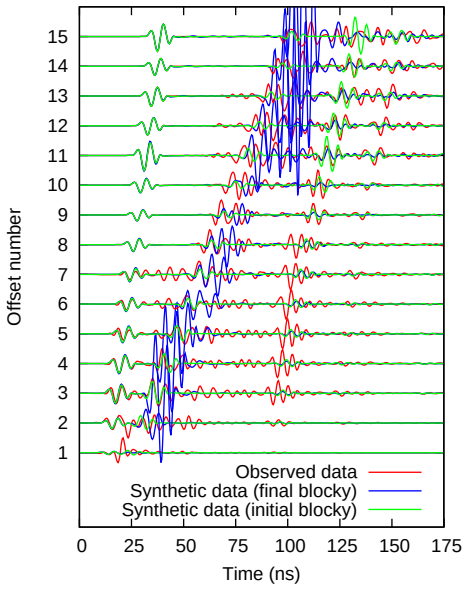
(a) CMP 200, NMO model (40-70 MHz).



(b) CMP 200, NMO model (no filter).



(c) CMP 200, blocky model (40-70 MHz).

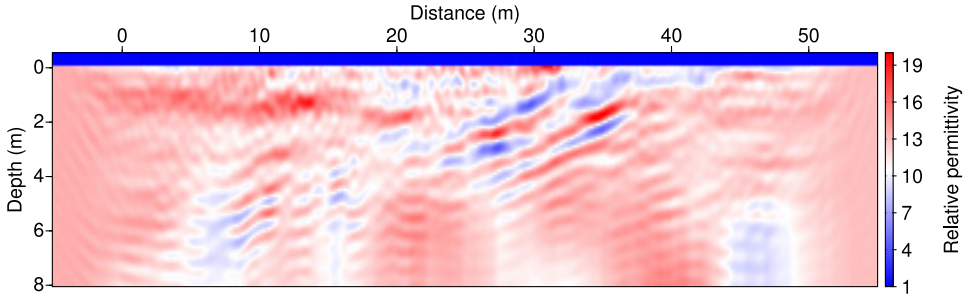


(d) CMP 200, blocky model (no filter).

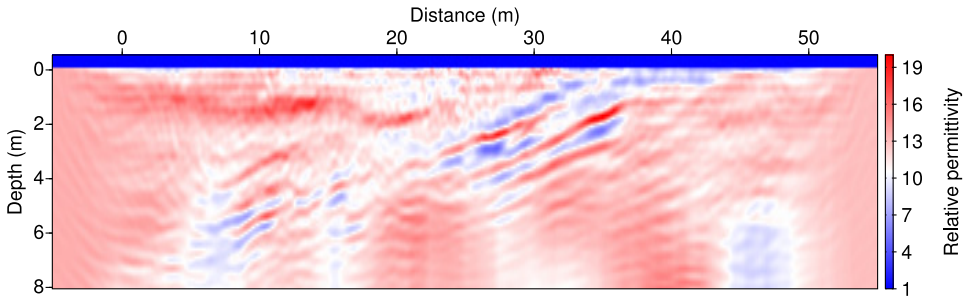
Figure 3.45: *Observed vs. synthetic data at CMP 200, after 10 iterations in the first frequency group (40 to 70 MHz).*

can be observed in the inverted frequency range (Figs 3.45a and c). More precisely, the fit in the reconstructed models is significantly better than in the initial models, indicating that the inversion did well extract information from the data. It can be noticed, however, that a good fit at low frequencies does not imply a good fit at higher frequencies (see Figs 3.45b and d). During the inversion of the low frequency components, nothing prevents the creation of heterogeneities that generate high-amplitude diffractions of the higher components that are not considered. The inversion must thus be continued at higher frequencies to improve the fit over the whole frequency band, and better constrain these heterogeneities.

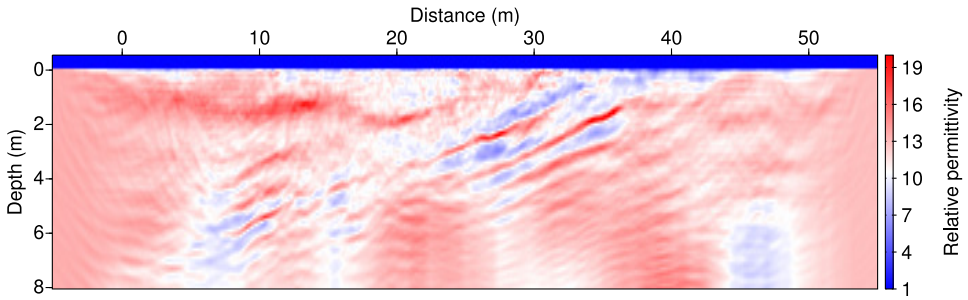
Given that the pre-existence of reflectors in the initial model has been found to be unnecessary, I now continue the inversion started from the smooth NMO model, taking the reconstructed model of Fig. 3.41(b) as a starting model for inverting the second frequency group (40 to 100 MHz). After 10 iterations, the updated model is in turn used as a starting model for inverting the third frequency group (40 to 150 MHz). The corresponding results are shown in Fig. 3.46. When extending the inverted bandwidth, the inversion introduces more details in



(a) 1st frequency group (40 to 70 MHz). Misfit decrease of 30%. *(reminder of Fig. 3.41b)*



(b) 2nd frequency group (40 to 100 MHz). Misfit decrease of 48%.

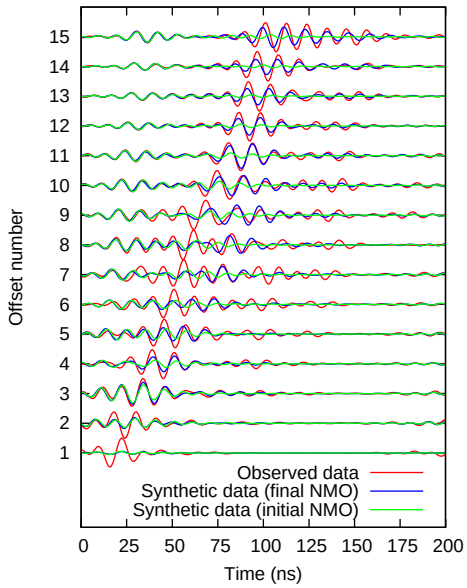


(c) 3rd frequency group (40 to 150 MHz). Misfit decrease of 43%.

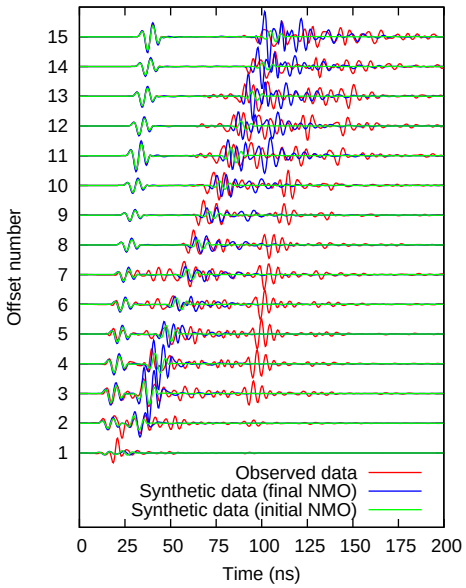
Figure 3.46: *Permittivity reconstructions after 10 iterations. Each reconstructed model serves as an initial model for the subsequent frequency group.*

the image. In particular, the delineation of the interfaces becomes sharper.

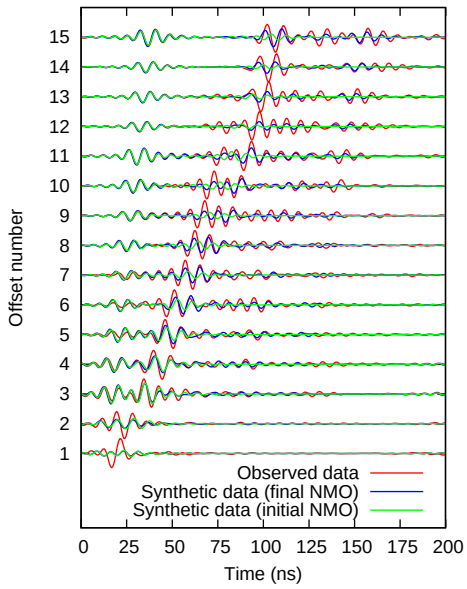
Fig. 3.47 shows the time-domain data fit corresponding to the reconstructed models of Fig. 3.46. The fit is satisfying in the inverted frequency bandwidths (Figs 3.47a and c). Of course, the fit in the entire bandwidth improves with the extension of the inverted bandwidth (Figs 3.47b and d). Note however that a major reflection at $t_o \simeq 100$ ns is not reproduced. Extending the bandwidth beyond 150 MHz does not improve the fit, nor introduce significant



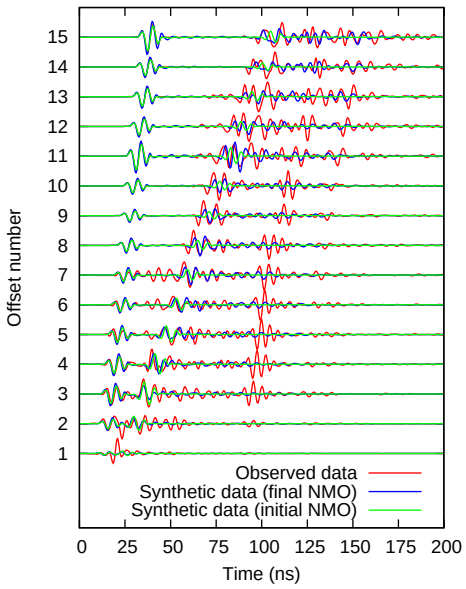
(a) CMP 200 after inversion of the 2nd frequency group (filtered 40-100 MHz).



(b) CMP 200 after inversion of the 2nd frequency group (no filter).



(c) CMP 200 after inversion of the 3rd frequency group (filtered 40-150 MHz).



(d) CMP 200 after inversion of the 3rd frequency group (no filter).

Figure 3.47: *Observed vs. synthetic data after 10 iterations of the inversion process for the 2nd frequency group (a,b) and the 3rd frequency group (c,d).*

changes in the models, because the bandwidth considered in the third frequency group (40 to 150 MHz) already contains most of the energy of the data. The data spectrum should be whitened before continuing the inversion at higher frequencies.

Of course, the geological interpretation of the reconstructed images shown in Fig. 3.46 should be achieved with cautious. I remind that they are not final interpretable images since they contain a strong finite-frequency imprint which causes non-physical oscillations at the positions of interfaces. As a consequence, the apparent width of the interfaces should not be interpreted as the thickness of the seals between two limestone strata (these seals are actually smaller). Similarly, the apparent permittivity of these interfaces (red delineations) are probably not characteristic values for their filling material. We can state, however, that the inversion confirms and enhances the trend which we could already observed in the smooth initial model of Fig. 3.23(b), with high-permittivity blocks at the beginning and at the end of the profile, and one or several low-permittivity layers at the centre, which could indicate a (dry) material with more fractures.

Knowing the frequency content of the source used for the inversion, some deconvolution should be performed before further interpretation (but it is a tricky task because the oscillations also contain the imprint of the acquisition). Alternatively, we may introduce some regularisation in the inversion in order to obtain more interpretable images. Regularisation based on total variation, for instance, might mitigate the oscillating appearance of the image and render a blocky reconstruction more compatible with a geological interpretation (Castellanos-Lopez, 2014). Finally, forward simulations in hypothetical models could also help interpretation. In particular, the response of water-, air- or clay-filled thin layers should be investigated to better characterise the nature of the interfaces.

Conclusions on the LSBB data set

Classical processing: Qualitative understanding of the data and quantitative constraints on the model

Full waveform inversion is not a first approach technique. Its application to real field data requires a good knowledge of the data characteristics, as well as *a priori* information on the quantitative properties of the underlying medium. A first and important interpretation step thus consists in acquiring a qualitative understanding of the data and some quantitative constraints on the subsurface model, using classical processing methods.

The acquired single-offset profiles provide a first qualitative view of the main structures (dipping interfaces), but also suggest more complex heterogeneities (reflector discontinuities, diffractions). A systematic semblance analysis yields a 2D velocity model that has been proved to be kinematically consistent with the observed data (with few local exceptions), and that finally leads to satisfactory inversion results. In itself, however, this model is not of great help for a fine interpretation of the karstic structures, because the derivation of interval velocities from NMO velocities is not a stable process. Instead, the main trend of the velocity model can be recast in a simpler model, which describes the subsurface as a succession of blocks separated by dipping interfaces. Though simplistic, this blocky model well explains the complex pattern of the recorded signal due to the interaction of the waves with the dipping reflectors and the air-ground interface — including direct waves, reflected waves, multiples, reflected-refracted waves and hidden zones. On the other hand, such a simple model fails to describe punctual diffractions and, in its current form, the detailed undulations and discontinuities of the interfaces, which causes cycle-skipping when this blocky model is used to start the inversion.

Processing steps towards FWI

Full waveform inversion being based on the comparison between observed and synthetic data, several pre-processing steps must be performed to make this comparison possible, given that the assumptions underlying our forward modelling tool do not allow to reproduce exactly the observed data.

First, I have identified and eliminated out-of-plane events that could not be reproduce in our 2D simulations (and that do not bring valuable information about the subsurface targets). Applying a mute in the offset-time domain has been found to be the most convenient approach for this elimination but was not possible on some offsets that must be ignored for the inversion. Secondly, the data that we recorded in the true 3D world have been transformed to virtually 2D data that could be fitted by the synthetic data. It has been found that the velocity contrast at the air-ground interface should be considered in the correction to explain the relative amplitudes of the direct air-wave and direct ground-wave, i.e. the energy distribution of the radiated field between air and ground. I may underline, however, that the correction I applied remains a crude approximation that could and should be improved. In particular, a collateral damage of this approximate correction is that the amplitude of the direct ground-wave is erroneously over-amplified at short offsets, which makes these data not usable for inversion.

The estimation of the source signature is known to be a major issue in FWI applications to GPR data (Busch et al., 2012; Belina et al., 2012a,b). In the case of the LSBB data set, there are two main obstacles to a simple source estimation:

1. The available initial models are not accurate enough to predict all the observed events. As a consequence, frequency-domain data computed in the initial models are far from the observed ones, which results in a bias in the source estimation. To overcome this point, I have performed the source estimation from selected traces that I have identified as kinematically compatible with the observed data. In addition, I applied an exponential time damping to select the early arrivals, direct air-wave and ground-wave, which were well predicted by the synthetic model.
2. The distribution of the energy radiated into the ground greatly depends on the antenna height above ground (Jiao et al., 2000; Diamanti and Annan, 2013). Moreover, the unknown ground conductivity has an effect on the AVO of the signal, and therefore on the source estimation: over-estimating the conductivity yields an over-estimated source amplitude since the synthetic data are over-attenuated. As a result, I have estimated suitable values for the antenna height and the mean conductivity simultaneously with the source signature. I have retained the combination of ground conductivity, antenna height and source wavelet that best fitted (qualitatively) the observed data, both in terms of waveform and of amplitude vs. offset.

Preliminary FWI results and perspectives

A frequency-domain analysis of the data has been performed to select the adequate bandwidths to be considered in the inversion. The plane wave analysis of Sirgue and Pratt (2004) presented in Section 2.3.1 has been used to evaluate the illumination of the subsurface and the expected wavenumber coverage. Though simplistic (1D approximation), this analysis suggests that a fine frequency sampling is required to compensate the low redundancy provided by the acquisition. It also motivates the need for starting at a very low frequency. Based on this analysis and of the SNR, I choose a first frequency bandwidth between 40 MHz and 70 MHz which I extend in a cumulative manner up to 150 MHz.

The results demonstrate that the inversion succeeds in fitting the data in the considered frequency range. The obtained permittivity distributions are in accordance with the expected subsurface structures. The two to three main reflectors are sketched out, in spite of the lack of high frequencies in the inverted range. We obtain similar reconstructions and a similar data fit when using the blocky model or the smooth NMO model, indicating that the inversion is able to reconstruct the main reflectors from a smooth background. Sharp pre-existing reflectors are thus not needed in the starting model. They may even mislead the inversion if they are not well located.

Before continuing the inversion at higher frequencies, several aspects should be considered. I list them below in priority order:

1. Data spectrum should be whitened such that all frequencies have the same weight in the inversion process.
2. Although the obtained reconstructions seem qualitatively satisfying, they still display some non-physical discontinuities due to cycle-skipping effects, in particular at depth ($z \geq 4$ m). Clearly, we are in a local minimum here. I suggest that the preliminary low-frequency steps should be processed again using improved initial models, in order to avoid this local minima. Then, further high-frequency steps could be performed, restarting from models that do not induce cycle-skipping.

3. From the pre-processing steps, it is obvious that efforts must be achieved in order to better account for the observed amplitudes. This includes the following requirements:
   (a) A more accurate 3D-to-2D conversion: the currently applied transformation is suspected to degrade the quality of the data at short offsets (124 and 180 cm) by over-amplifying the direct ground-wave.
   (b) A refined estimation of the (effective) conductivity, eventually including lateral variations or a blocky representation. By *effective*, I mean that the estimated conductivity aims in a first time only at fitting the observed amplitudes, and should not be interpreted as a physical conductivity value because of the trade-off between the effects of the conductivity, of the source estimation, and eventually of the antenna radiation pattern on the amplitudes.
   (c) An evaluation of the impact of the antenna radiation pattern: because of the above-mentioned trade-off, it is difficult to state the importance of this impact now. A detailed analysis of the amplitudes vs. offset (i.e. emission/reception angles) should be performed. Numerical simulations involving a detailed shielded antenna would also provide additional insights (e.g. Warren and Giannopoulos, 2011; Diamanti and Annan, 2013; Diamanti et al., 2013; Sagnard et al., 2013).
4. The estimation of the source signature is related to the problem of amplitude matching. The source wavelet may be re-evaluated while the initial model, especially the conductivity values, is improved. We might think about a more systematic way to estimate the source signature, together with the values of conductivity and antenna height, eventually using a global optimisation method (Sen and Stoffa, 1995).
5. In the previous results, I have avoided artefacts by stopping the inversion at an early stage of convergence (10 iterations). To go further without introducing high-wavenumber artefacts, a regularisation term should be introduced, which requires to find an adequate regularisation weight λ. In particular, a regularisation term based on the total variation of the model might attenuate the finite-frequency imprint and thus render more interpretable images (Castellanos-Lopez, 2014). Alternatively, a model term could also be introduced to add *a priori* information and guide the inversion process.

Partial conclusion on real data inversion

In this chapter, I have presented two applications of the imaging algorithm to real datasets.

In Section 3.1, the consideration of laboratory data is the occasion to confront the accuracy of the numerical method presented in Chapter 1 to physical data, as well as to validate the workflow for multiparameter imaging proposed in Chapter 2. The well-controlled environment of the experiment greatly simplifies the practical application of FWI to this dataset since the initial model can be considered as homogeneous and equal to free space, and since the data can be calibrated with regards to geometrical expansion and source signature using measurements of the incident field. In contrast, when dealing with the field data acquired at the LSBB, the design of the initial model and the estimation of the source signature is made difficult by the complexity of the geological environment and by the influence of several non-controlled parameters (conductivity and antenna height, essentially).

In both cases, it appears that a pre-requisite to data inversion is the fine understanding of the data. This understanding is mainly attained through forward simulations in hypothetical models. In the case of the Institut Fresnel data set, forward modelling enabled to validate the accuracy of our modelling tool, to choose an adequate grid step to describe the targets, and to test some hypothesis. In particular, it was possible to test the data sensitivity to the inner filling of the copper tube, and thus to not over-interpret the inversion results in this zone. In the case of the LSBB data set, forward modelling in a blocky model allowed to understand the main features of the observed data, caused by wave interferences at the intersection between the dipping interfaces and the ground surface. It also enables to test several combinations of conductivity values and antenna height to estimate the source signature and understand how to match the observed amplitudes.

The reader might have noticed the big gap between the methodological developments of Chapter 2 based on synthetics and their applicability to real field data. Actually, the study of the LSBB data set is preliminary in two respects. In the view of a multiparameter imaging seeking for both permittivity and conductivity distributions to understand the hydrogeology of the limestone massif, this study constitutes a first monoparameter step to improve the permittivity model, whose accuracy is critical for a correct reconstruction of conductivity in the subsequent multiparameter stage (see Chapter 2). Within the monoparameter step, the work presented above considers only the starting phase of FWI, i.e. the design of the initial model, the estimation of the source signature, and the inversion of a few groups of low to medium frequencies. This preliminary processing should be improved before continuing towards inversion of higher frequencies and multiparameter reconstruction. But it already provides promising results in the sense that FWI is able to fit the data in the considered frequency bandwidth, and to reconstruct reflectors from a smooth initial model. A more accurate understanding of the separate effects of the air-ground interface and of the antenna characteristics on the observed amplitudes is required to envisage a quantitative estimation of conductivity.

Conclusions and perspectives

Conclusive sum-up

The aim of my work was to develop a quantitative imaging method based on the full waveform inversion of GPR data. I dedicated the first part of this thesis to the forward problem, which consists in simulating the propagation of electromagnetic waves in 2D heterogeneous natural media. In the second part, I addressed the inverse problem, i.e. the reconstruction of 2D distributions of permittivity and conductivity from multi-offset GPR measurements. Finally, I was able to validate the proposed methodology against laboratory-controlled data and I addressed the issue of its application to real field data.

Forward problem

In Chapter 1, I first presented the system of Maxwell's equations that governs the propagation of electromagnetic waves. Based on these equations and on the constitutive relations that describe the electrical response of dielectric conductive materials, I drew the general behaviour of EM waves in natural non-magnetic media. At first order, electromagnetic wave velocity is controlled by the dielectric permittivity while the static conductivity governs diffusive processes and has mainly an attenuation effect. At low frequencies, static conductivity also induces a dispersion of the propagation velocity. At a second order, dispersion occurs in natural media due to their transient response to the electric excitation. This effect can be naturally taken into account when the computation is performed in the frequency domain, using frequency-dependent electromagnetic parameters in the constitutive relations. This hierarchy of the electromagnetic processes in natural materials has driven the choice of the parameters considered for inversion, which were restricted to real-valued, frequency-independent permittivity and conductivity, such as to explain the first order response of the investigated media.

In a second section of Chapter 1, I presented the numerical method used for the simulation of electromagnetic wave propagation in 2D heterogeneous media. The method makes use of a frequency-domain finite-difference algorithm (FDFD) initially developed for the simulation of seismic waves in the visco-acoustic approximation (Hustedt et al., 2004; Operto et al., 2009). With hindsights, the motivation for working in the frequency-domain is rather numerical than physical, and I discuss this point in the following. The modelling algorithm is based on a mixed grid stencil whose coefficients can be tuned to minimise the errors due to numerical dispersion and anisotropy (Jo et al., 1996). I studied this property in details to ensure a good utilisation of the method and validated my implementation against analytical solutions in a homogeneous medium. In Chapter 3, the confrontation of the numerical solutions to laboratory-controlled

experimental data demonstrated the accuracy of the simulations in presence of heterogeneous targets.

The visco-acoustic FDFD modelling could be adapted to the electromagnetic propagation problem using a mathematical analogy between the first-order systems of electromagnetic and acoustic equations, provided that a two-dimensional geometry is assumed for the medium and for the propagation. The 2D approximation is further motivated by the reduced computational cost of 2D simulations compared to 3D modelling, and corresponds to the current acquisition capability of GPR measurements (performing dense multi-offset measurements along a profile for 2D imaging is already a tedious task). On the other hand, assuming a 2D propagation induces strong limitations on the interpretation of the simulated amplitudes. In a well-controlled environment, amplitudes can be corrected to match physical data (see Section 3.1). In the case of real field data, however, amplitude matching requires the application of *ad hoc* 3D-to-2D conversions which are far from satisfying.

Inverse problem

In Chapter 2, I provided a short overview of inverse problem theory and detailed the key properties of the specific FWI problem, which is an *ill-posed*, *non-linear*, *large-scale* problem. In virtue of these characteristics, the FWI of GPR data for the imaging of permittivity and conductivity has been formulated as a local optimisation problem and solved using the quasi-Newton L-BFGS algorithm. In this algorithm, the consideration of the effect of an approximate inverse Hessian on the model update should partially correct for the different dimensionalities of the parameters and for the different sensitivities of the data with respect to permittivity and conductivity. In practice, however, numerical tests on a cross-shaped benchmark originally proposed by Meles et al. (2011) display a great sensitivity to parameter scaling, suggesting that the L-BFGS approximation of the inverse Hessian is not appropriate, and/or that the problem does not honour the quadratic approximation. On a more realistic subsurface benchmark, it has been shown that allying parameter scaling and regularisation allows one to select reasonable model solutions on the basis on the data misfit only, analysed for a range of scaling factors and regularisation weights. The imaging algorithm was then able to reconstruct 2D permittivity and conductivity distributions from data acquired in a surface-to-surface configuration. Intuitively, the selected tuning factors provide a well resolved image of the well constrained parameter (the permittivity in the synthetic subsurface benchmark) and a smooth reconstruction of the poor constrained parameter (conductivity in the synthetic example).

Following the presentation of this workflow, I supplied additional insights into some key features of FWI. In particular, the gradient of the misfit function has been derived and interpreted in terms of its resolution capability in the imaging process. A short discussion based on the analysis of Sirgue and Pratt (2004) further illustrated the fundamental difference between surface and crosshole GPR experiments, as a complement of the sensitivity analysis provided by Meles et al. (2012).

Application to real data inversion

In Chapter 3, I confronted the imaging algorithm to two real data sets. The experimental data from the Institut Fresnel first enabled to validate the proposed workflow for mono-parameter and multi-parameter imaging. The well controlled laboratory environment of this experiment

offers a favourable situation for the application of FWI. In particular, the initial model and the source signature were either known or accurately calibrated. Actually, this application appeared less critical than the synthetic examples since regularisation alone was able to mitigate the indetermination and to erase the artefacts arising in the image of the less constrained parameter (the permittivity, in this case). On the modelling point of view, it has been the occasion to challenge the accuracy of the numerical simulations against high-quality physical data. It was also the opportunity of a small excursion into the field of integral methods which offered an elegant alternative to brute force FDFD calculations by restricting the computation domain to the zone of interest.

The application of FWI to on-ground GPR field data is much more challenging because of several non-controlled parameters: the searched permittivity and conductivity distributions, of course, but also the source signature, and even the precise antenna height, that controls the distribution of the radiated energy between air and ground. As a consequence, the treatment of the LSBB data that has been presented here is only a prologue to the interpretation of this rich data set. First, it is restricted to a preliminary mono-parameter estimation of the permittivity model, which is expected to explain most of the observed data and without which a subsequent multi-parameter inversion cannot be envisaged. Secondly, I considered the inversion of a few groups of low to medium frequencies, the early stages in a Bunks' approach. Finally, I principally investigated how to design a suitable initial model and estimate the source signature for starting the FWI.

It has been found that the consideration of the velocity contrast at the air-ground interface in the 3D-to-2D conversion was required to explain the relative amplitudes of the direct air-wave and ground-wave. A procedure has been designed to estimate the source wavelet from selected observed traces where the direct air-wave and ground-wave could be well matched by synthetic ones. Performing the source estimation for various combinations of antenna heights and conductivity values enabled to deduced rough estimates for these non-controlled parameters, together with the source signature. However, an amplitude mismatch could be observed between short and long offsets, suggesting an effect of the antenna shielding on the radiation pattern, resulting in a biased AVO of the signal.

A frequency-domain analysis of the data based on the approach of Sirgue and Pratt (2004) served as a guiding rule for choosing the first frequency bandwidth to be inverted. The limited redundancy provided by the acquisition required a fine frequency sampling and a low starting frequency. The inversion could fit the data in the inverted frequency band, starting either from the first-order blocky model or from a smooth permittivity model derived from velocity analysis. Nevertheless, cycle-skipping occurred locally, indicating the need for a refined starting model before continuing the inversion at higher frequencies. The reconstructed permittivity models and the final data fit obtained using the smooth NMO model and the blocky model are very similar, indicating that the inversion does not need pre-existing reflectors in the starting model but can create them by fitting the data. Continuing the inversion process up to the peak frequency of the data spectrum enables to sharpen the contrasts in the reconstructed image but a pre-whitening is required for exploiting the full frequency bandwidth.

Perspectives

Forward perspectives: 3D modelling and simulation of the antenna radiation pattern

A direct consequence of the difficulties encountered with the LSBB data set to match the observed amplitudes is the motivation for a 3D modelling that would directly simulate real-world data. This is a classical perspective given to most 2D studies but I specify that I do not envisage 3D modelling for 3D imaging, but primarily to ensure a consistent simulation of the amplitudes. Indeed, full 3D FWI of GPR data seems quite far today, not only because of the computational burden of simulating and inverting data in a 3D model, but also due to data acquisition. GPR does not benefit of the same acquisition capabilities as the oil industry and acquiring dense multi-offset GPR data along a single profile for a correct 2D imaging seems already a challenging-enough task.

Another point that may be developed to further improve the simulation and inversion of real field data is the consideration of the antenna radiation pattern. Before going in this direction, however, the real impact of antenna shielding on multi-offset data and on their inversion should be carefully investigated. This probably requires a multi-method approach, involving detailed numerical simulations of realistic antennas (e.g. Lampe and Holliger, 2005; Warren and Giannopoulos, 2011; Diamanti and Annan, 2013; Sagnard et al., 2013), accurate laboratory measurements of real antennas characteristics (Lutz, 2002; Sagnard et al., 2013) and *in situ* measurements to study the effect of antenna ground-coupling and of laterally variable media (Jiao et al., 2000). Finally, if the effect of the antenna radiation pattern is found to be critical for the FWI to match the recorded amplitudes, it should be investigated how it is possible to include the complex radiation pattern of real antennas in the forward modelling. Ideally, an *effective* source should be implemented, that would mimic the radiation of a real antennas without having to detail its structure as part of the model. Some food for thought are given by e.g. Carcione (1998) and Lambot et al. (2004, 2010) to reach this goal.

Inverse perspectives: optimisation issues, extension to other parameters and other data

Towards better estimations of the Hessian

On the methodological point of view, the large sensitivity of the optimisation to the parameter scaling is not satisfying. It suggest that the L-BFGS algorithm poorly approximates the effect of the inverse Hessian on the descent direction. To overcome this limitation, we might opt for more complete approximation of the Newton method, e.g. the truncated Newton method (Métivier et al., 2013) that solves the Newton system at each non-linear iteration of the optimisation, using a conjugate gradient method. The question of knowing if it will erase the sensitivity to the scaling factor β is still open. I expect that it will improve the situation, but not totally solve it if the initial model is too far from the domain of validity of the quadratic approximation.

Furthermore, an important aspect in optimisation techniques that I did not investigate is preconditioning. Preconditioning can be seen as a change of variables that aims at rendering the Hessian matrix more diagonal. Hence, preconditioning also contributes to mitigate the trade-off between permittivity and conductivity and might be investigated to solve this problem.

Reconstruction of dispersive parameters

In my synthetic tests, data displayed a very low sensitivity to the parameters n and χ_r of the Jonscher's parameterisation (see Section 2.3.4). Therefore, I am quite pessimistic concerning the reconstruction of these parameters in 2D. However, data sensitivity of course depends on the considered dataset, and thus on the investigated medium. In some cases (e.g. a stratified strongly dispersive medium such as clay, silt, or schist layers), it may be possible to retrieve these parameters through a 1D inversion with the methodology of van der Kruk et al. (2006) or through a 1D FWI following the approach of Kalogeropoulos et al. (2011); Busch et al. (2012).

Joint inversion

Following Meles et al. (2013), I should indicate the possibility of separately or jointly invert data acquired in TE and TM configurations. I should mention, however, that preliminary tests on the realistic subsurface benchmark of Section 2.2 do not exhibit significant differences between TE and TM inversions. The benefit of joint TE-TM inversion probably depends on the dataset, i.e. on the investigated medium. For instance, the presence of thin layers acting as wave guides can result in TE and TM data having a different frequency content (van der Kruk et al., 2006, 2012) due to the different reflection coefficients for TE and TM. The difference in reflection coefficients should also results in different sensitivities to permittivity and conductivity.

Beyond the TE-TM problematic, FWI of GPR data could be extended to other components of the electric field, in particular to cross-polarised components. Multicomponent data are expected to bring much more information in 3D media where depolarisation effects can happen (Lutz et al., 2003). But this extension is not possible within the 2D approximation: it is another motivation for a full 3D modelling.

Finally, we could think about combining GPR data with other geophysical measurements. In a first time, electromagnetic induction (EMI) or electrical resistivity tomography could easily furnish an estimation of smooth initial models for the conductivity. To go further, the methodological tools developed in this thesis could allow one to envisage the FWI of multi-offset EMI data for getting a smooth model of conductivity that could be used as a starting model for the FWI of GPR data. Finally, we might attempt on the long-term at jointly inverting GPR and EMI data, but it would require the simulation of electromagnetic fields in very different frequency ranges, which might be problematic on the modelling point of view.

I have a more reserved opinion on *joint* inversion schemes that mix data coming from different physical processes, such as GPR and seismics (e.g. Tronicke et al., 2011; Rumpf and Tronicke, 2014). In my understanding, when considering data (e.g. GPR and seismic traveltimes) that are sensitive to different physical parameter types (e.g. permittivity and P-wave velocity), the additional constraints on the reconstruction of each parameter type come principally from the underlying geometrical assumptions (i.e. the hypothesis that both parameter types vary simultaneously from a layer to another).

Applicative perspectives: starting model, source estimation and acquisition design

Building the starting model: Reflection tomography and FWI

GPR data contain mostly reflections that bring few low wavenumbers to the FWI process. The design of an accurate initial model is thus crucial: the initial model should contain the missing low wavenumbers to enable the correct relocation of the reflectors. In Section 3.2, I performed a classical velocity analysis, based on semblance analysis and event picking, and I sometimes corrected the resulting velocity models iteratively by comparing synthetic data generated in these models with the observed data to check their kinematic compatibility. We could think about more systematic ways to design the initial model.

On-ground GPR data generally present a clear hierarchy of direct and reflected arrivals which can be (almost) unambiguously separated. GPR data are therefore well suited for the application of seismic reflection methods, as reflection tomography (Bradford, 2006, 2008), as well as of emerging reflection FWI methods that were recently developed to circumvent the problem of poor starting models (Brossier et al., 2014). These methods are based on the decoupling between a smooth background velocity model, which is reconstructed first, and a high-wavenumber reflectivity model, introduced via sharp discontinuities in density which are reconstructed in an iterative way when the background velocity enables to relocate them accurately. However, the adaptation of these methods to the GPR case is not straightforward since the electromagnetic impedance mainly depends on the (effective) permittivity and on the magnetic permeability, i.e. on the effective permittivity only in non-magnetic media.

Source estimation

Similarly, we could think about more systematic ways of estimating the source signature, using e.g. global optimisation methods (Sen and Stoffa, 1995), given that it strongly depends on the unknown antenna height and on the ground conductivity.

Moreover, efforts might be made for estimating the source signature from reflected events, which are more representative of the signal injected into the ground than the direct air and ground waves. It requires, however, to have a good *a priori* knowledge on the reflectivity of some reflectors. This knowledge could result from a classical velocity analysis leading to a blocky model, as done in Section 3.2.2.3, or preferentially from a reflection tomography.

Adequate acquisition design for FWI of on-ground GPR data

The application to the field data set acquired at LSBB gives rise to important conclusions concerning the acquisition setup required for FWI. If new data should be acquired at the LSBB, I would strongly recommend to perform more common-offset profiles, or alternatively, to perform the acquisition in a shot-gather configuration, which offers several key advantages. First, shot-gather acquisitions ensure a finer offset sampling, which is critical in FWI for a correct wavenumber coverage (see Fig. 3.34) but also in the preliminary study for velocity analysis based on the CMP gathers. It makes also possible to acquire data for larger offsets without limitations coming from the link between antennas. Indeed, data acquired at LSBB with a maximal offset of 8.34 m still contain a lot of valuable signals: larger offsets would bring

an important low wavenumber information on the subsurface. Secondly, acquiring data in shot gathers would enable to estimate one source signature per transmitter locations to account for lateral variations of ground properties and of antenna-ground coupling[1]. Finally, on a computational point of view, inverting shot-gathered data would be more economical since it reduces the number of source positions to be simulated. The only drawback of acquiring data in shot-gather configuration resides in a long and tedious acquisition since numerous source positions must be successively acquired to achieve the same lateral sampling as with common-offset acquisitions. Nevertheless, it is certainly worth spending time to acquire reliable data if it facilitates their processing and interpretation.

Discussion: Time vs. frequency domain FWI of GPR data

In this thesis, I worked in the frequency domain rather than in the time domain, contrary to the studies by Ernst et al. (2007) and Meles et al. (2010). I may now comment this basic assumption.

Retrospectively, it appears that the arguments for working in the frequency domain fall mainly under computational considerations. On the forward modelling point of view, it takes benefit from the capability of direct solvers like MUMPS to treat efficiently multiple sources. Concerning inversion, frequency-domain FWI has been promoted by Pratt and Worthington (1990) to take benefit of the data redundancy provided by dense seismic acquisitions. Of course, frequency-domain modelling and inversion are particularly suited when considering frequency-domain measurements, such as the Institut Fresnel data set (section 3.1). For the consideration of time-domain GPR data, however, frequency-domain inversion seems less attractive.

First, GPR acquisitions generally provide data which are far less redundant that seismic data and which can not be dramatically decimated: performing frequency-domain FWI is definitely less justified if the inversion of each frequency component of the time-domain signal is required (which is almost the case in Section 3.2, though in a reduced frequency bandwidth). This remark, however, also claims for denser GPR acquisitions.

A second point concerns the description of dispersive materials, which is the main physical motivation for working in the frequency domain. The introduction of dispersive parameters in the forward modelling is certainly more satisfying to describe the natural media, but the reconstruction of their 2D distributions through an inversion procedure seems for now hardly feasible, given the ill-posedness of multi-parameter inversion when considering only two parameter classes (Section 2.2) and the small sensitivity of the data to dispersive parameters (Section 2.3.4). Thus, I have to recognise that all my developments about multi-parameter imaging could have been done in the time-domain, since I consider constant, real-valued permittivity and conductivity parameters.

[1]Note however that the problem of *source estimation* is actually symmetric for the transmitting and receiving GPR antennas. There is no particular reason to estimate the source signature rather than the receiver response, but we cannot estimate a transfer coefficient for each source-receiver pair since it would amount to fit the data without inverting for a subsurface model (in fact, when optimising a source coefficient, we include the receiver response in the source signature). Moreover, Belina et al. (2012b) suggest that a global source estimation is a reasonable approximation in the case of crosshole GPR data acquired in weakly to moderately heterogeneous media. It should be investigated whether this conclusion applies in the case of on-ground GPR, where the variation of antenna-ground coupling might be more pronounced.

Finally, frequency-domain FWI leaves very little flexibility in the data processing step compared to time-domain FWI where identified events can be selected and fitted separately. This lack of flexibility makes an important difference when dealing with complex data sets that require successive processing and inversion steps to ensure a hierarchical treatment of the information.

Before claiming the superiority of time-domain FWI for GPR data, however, it is worth noting firstly that I did not investigate the question in details, and secondly that Meles et al. (2013), who are deeply involved in time-domain FWI, now promote to move to frequency-domain FWI in their recent review. Among their arguments to do so, we can find the consideration of dispersive parameters and the decimation of the inverted frequencies using the multi-scale approach, two points that my previous remarks have partially attenuated (partially only because the multi-scale approach remains more efficient in the frequency domain than in the time domain, where all data must be computed before throwing the high frequency part of the spectrum).

Meles et al. (2013) also point out the need for modelling realistic GPR sources. In this respect, frequency-domain modelling probably offers more flexibility than time-domain simulations. For instance, the problem of hard *vs.* soft sources mentioned by these authors, which refers to the order of the time-derivative of the input signal (electric field or current) is totally transparent in the frequency-domain where the source can have an arbitrary complex spectrum. Similarly, building an effective source that radiates like a realistic GPR antennas might be more straightforward to implement in the frequency-domain, where hard sources can be easily superimposed, than in the time-domain, where injecting an imposed solution in the grid creates point sources that act as diffracting points for back-scattered waves. In the time-domain, the integral representation presented in Section 3.1 can be used to perform a near-to-far field transformation of scattered fields after Fourier transformation (Taflove and Hagness, 2005, chap. 8) but it cannot be used to inject the incident wave coming from the far field in the zone of interest. To do so, time-domain numerical schemes must be modified using (e.g.) the total-field/scattered-field technique (see Taflove and Hagness, 2005, §5.6, p. 186).

Apart from the multi-scale approach, that should remain efficient for densely acquired data sets, the easy management of hard sources is probably the main advantage of working in the frequency domain, and could enable in a near future the modelling of effective sources to simulate the complex radiation pattern of realistic GPR antennas. In 3D, frequency-domain FWI could also avoid complex and expensive checkpointing techniques required for computing the gradient by convolution of attenuated time-domain wavefields in the entire domain.

Appendix A

Adjoint state method using a Lagrangian formulation

From Chapter 2, the reader has retained that the full waveform inversion problem can be formulated as the following least-square minimisation problem

$$\min_{\mathbf{m}} \quad \frac{1}{2}\big\|\mathbf{d}_{obs} - \mathbf{d}_{cal}(\mathbf{m})\big\|. \tag{A.1}$$

In this section, I explain how this problem can be seen as a *constrained* optimisation problem. This point of view constitutes the adjoint state formulation *sensu stricto* (Plessix, 2006) and offers a more general framework for the derivation of further developments.

In addition to the requirement that the solution model $\mathbf{m}$ should explain the observed data (eq. A.1), constraint is given by the fact that the data result from the measurement of a wavefield $\mathbf{u}(\mathbf{m})$. This wavefield must verify the wave equation, which is represented by the forward problem

$$\mathcal{A}(\mathbf{m})\mathbf{u}(\mathbf{m}) = \mathbf{s}. \tag{A.2}$$

For completeness, an additional constraint comes from the relation between the simulated field $\mathbf{u}(\mathbf{m})$ and the data used in the inversion $\mathbf{d}_{cal}(\mathbf{m})$, which can be expressed as

$$\begin{aligned} \mathbf{d}_{cal}(\mathbf{m}) = \mathcal{P}\mathbf{u}(\mathbf{m}) &= \begin{cases} \mathcal{R}\mathbf{u}(\mathbf{m}) & \text{with } \mathbf{u} = \mathbf{E_y} \text{ in TE mode,} \\ \mathcal{R}\mathcal{D}\mathbf{u}(\mathbf{m}) & \text{with } \mathbf{u} = \mathbf{H_y} \text{ in TM mode,} \end{cases} \\ \text{where} \quad \mathcal{D} &= \frac{1}{\imath\omega\varepsilon_e}\left(\cos\theta_r\frac{\partial}{\partial z} - \sin\theta_r\frac{\partial}{\partial x}\right). \end{aligned} \tag{A.3}$$

In TE mode, the operator $\mathcal{P}$ is simply the projection $\mathcal{R}$ of the wavefield onto the receiver locations. In TM mode, this operator also includes the derivation $\mathcal{D}$ of the recorded electric field from the simulated magnetic field (see Section 1.1.3.2).

To represent these requirements, a Lagrangian function can be defined as

$$\mathscr{L}(\mathbf{m}, \boldsymbol{d}, \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = \frac{1}{2}\big\|\mathbf{d}_{obs} - \boldsymbol{d}\big\|^2 + \Re e\big\langle \boldsymbol{d} - \mathcal{P}\boldsymbol{u}, \boldsymbol{w}\big\rangle + \Re e\big\langle \mathcal{A}(\mathbf{m})\boldsymbol{u} - \mathbf{s}, \boldsymbol{v}\big\rangle, \tag{A.4}$$

where $\langle \mathbf{x}, \mathbf{y}\rangle = \mathbf{x}^*\mathbf{y}$ denotes the Hermitian scalar product for any vector $\mathbf{x}$ and $\mathbf{y}$ in the complex space $\mathbb{C}^n$, n being either N_M or N_D, depending on whether the scalar product is achieved in the space of the wavefields or in the data space, respectively.

Each term of the Lagrangian function corresponds to a given requirement and involves Lagrangian multipliers $\boldsymbol{v}$ and $\boldsymbol{w}$, also called adjoint variables. Hence, the Lagrangian function recaps the different ingredients of the inverse problem. An intuitive reading of the function would be, from the left to the right,

1. we seek to minimise the distance between observed and calculated data,
2. the calculated data $\boldsymbol{d}$ are extracted from a wavefield $\boldsymbol{u}$,
3. the wavefield $\boldsymbol{u}$ should verified the wave equation in a given model $\mathbf{m}$.

Note that the arguments $\mathbf{m}$, $\boldsymbol{d}$, $\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{w}$ are considered as independent variables. In particular, $\boldsymbol{u}$ and $\boldsymbol{v}$ denote arbitrary wavefields, *a priori* different from the particular forward and adjoint fields $\mathbf{u}$ and $\mathbf{v}$ we have seen until now, and which were related to the model $\mathbf{m}$ through the forward and the adjoint problems. In the following, we shall see that the relations between these variables naturally arise from the Lagrangian formulation.

Solving the optimisation problem amounts to find the saddle points of the Lagrangian function, which are given by

$$\begin{aligned}
\frac{\partial \mathscr{L}}{\partial \mathbf{m}}(\mathbf{m},\boldsymbol{d},\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}) = \mathbf{0} &= \frac{\partial}{\partial \mathbf{m}}\Big(\Re e\big\langle \mathcal{A}(\mathbf{m})\boldsymbol{u}-\mathbf{s},\boldsymbol{v}\big\rangle\Big), && \text{(A.5)}\\
&= \Re e\left\langle \frac{\partial \mathcal{A}(\mathbf{m})}{\partial \mathbf{m}}\mathbf{u},\mathbf{v}\right\rangle,\\
\frac{\partial \mathscr{L}}{\partial \boldsymbol{v}}(\mathbf{m},\boldsymbol{d},\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}) = \mathbf{0} &= \frac{\partial}{\partial \boldsymbol{v}}\Big(\Re e\big\langle \mathcal{A}(\mathbf{m})\boldsymbol{u}-\mathbf{s},\boldsymbol{v}\big\rangle\Big), && \text{(A.6)}\\
&= \Re e\big\{\mathcal{A}(\mathbf{m})\boldsymbol{u}-\mathbf{s}\big\},\\
\frac{\partial \mathscr{L}}{\partial \boldsymbol{w}}(\mathbf{m},\boldsymbol{d},\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}) = \mathbf{0} &= \frac{\partial}{\partial \boldsymbol{w}}\Big(\Re e\big\langle \boldsymbol{d}-\mathcal{P}\boldsymbol{u},\boldsymbol{w}\big\rangle\Big), && \text{(A.7)}\\
&= \Re e\big\{\boldsymbol{d}-\mathcal{P}\boldsymbol{u}\big\},\\
\frac{\partial \mathscr{L}}{\partial \boldsymbol{u}}(\mathbf{m},\boldsymbol{d},\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}) = \mathbf{0} &= \frac{\partial}{\partial \boldsymbol{u}}\Big(\Re e\big\langle \mathcal{A}(\mathbf{m})\boldsymbol{u}-\mathbf{s},\boldsymbol{v}\big\rangle\Big) + \frac{\partial}{\partial \boldsymbol{u}}\Big(\Re e\big\langle \boldsymbol{d}-\mathcal{P}\boldsymbol{u},\boldsymbol{w}\big\rangle\Big), && \text{(A.8)}\\
\frac{\partial \mathscr{L}}{\partial \boldsymbol{d}}(\mathbf{m},\boldsymbol{d},\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}) = \mathbf{0} &= \frac{\partial}{\partial \boldsymbol{d}}\left(\frac{1}{2}\big\|\mathbf{d}_{obs}-\boldsymbol{d}\big\|^2\right) + \frac{\partial}{\partial \boldsymbol{d}}\Big(\Re e\big\langle \boldsymbol{d}-\mathcal{P}\boldsymbol{u},\boldsymbol{w}\big\rangle\Big). && \text{(A.9)}
\end{aligned}$$

Considering equation (A.6), we retrieve an expression similar as the gradient of Chapter 2 (eq. 2.43):

$$G_i(\mathbf{m}) = \frac{\partial \mathscr{L}}{\partial m_i} = \Re e\left\{\boldsymbol{u}^\dagger \frac{\partial \mathcal{A}(\mathbf{m})^\dagger}{\partial m_i}\boldsymbol{v}\right\}, \qquad \text{(A.10)}$$

but the variables $\boldsymbol{u}$ and $\boldsymbol{v}$ still have to be determined. In the following, I show that this expression effectively corresponds to the gradient of the misfit function presented in Chapter 2, provided that $\boldsymbol{u}$ and $\boldsymbol{v}$ verify the forward and the adjoint equation, respectively.

Indeed, looking at the second saddle point (A.7), we see immediately that it is has for solution the field $\mathbf{u}(\mathbf{m})$ verifying the forward problem:

$$\text{if} \quad \boldsymbol{u}=\mathbf{u}(\mathbf{m}) \quad \text{such that} \quad \mathcal{A}(\mathbf{m})\mathbf{u}(\mathbf{m})=\mathbf{s}, \quad \text{then} \quad \Re e\big\{\mathcal{A}(\mathbf{m})\boldsymbol{u}-\mathbf{s}\big\}=0. \qquad \text{(A.11)}$$

Similarly, the third saddle point (A.8) yields the relation between the inverted data and the simulated field, $\boldsymbol{d}=\mathcal{P}\boldsymbol{u}$, so we have well

$$\boldsymbol{d}=\mathbf{d}_{cal}(\mathbf{m})=\mathcal{P}\mathbf{u}(\mathbf{m}). \qquad \text{(A.3 again)}$$

It remains to determine the adjoint variable $\boldsymbol{v}$ in the gradient expression (A.10). It is given by eq. (A.8). Actually, this equation poses a problem since it is not trivial to compute derivatives with respect to the complex variable $\boldsymbol{u}$ [1]. Following the definition of Fréchet derivatives,

$$\frac{\partial \mathscr{L}}{\partial \boldsymbol{u}}(\mathbf{m},\boldsymbol{d},\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}) = \lim_{\|\delta\mathbf{u}\|\to\mathbf{0}} \frac{\mathscr{L}(\mathbf{m},\boldsymbol{d},\boldsymbol{u}+\delta\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}) - \mathscr{L}(\mathbf{m},\boldsymbol{d},\boldsymbol{u},\boldsymbol{v},\boldsymbol{w})}{\|\delta\boldsymbol{u}\|}, \tag{A.12}$$

we can state that a sufficient condition for satisfying the saddle point (A.8) is to verify

$$\mathscr{L}(\mathbf{m},\boldsymbol{d},\boldsymbol{u}+\delta\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}) - \mathscr{L}(\mathbf{m},\boldsymbol{d},\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}) = 0, \quad \forall\, \delta\mathbf{u}\to\mathbf{0}, \tag{A.13}$$

where

$$\begin{aligned}
\mathscr{L}(\mathbf{m},\boldsymbol{d},\boldsymbol{u}+\delta\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}) &= \Re e\big\langle \mathcal{A}(\mathbf{m})(\boldsymbol{u}+\delta\boldsymbol{u}) - \mathbf{s},\boldsymbol{v}\big\rangle + \Re e\big\langle \boldsymbol{d} - \mathcal{P}(\boldsymbol{u}+\delta\boldsymbol{u}),\boldsymbol{w}\big\rangle,\\
&= \Re e\big\langle \mathcal{A}(\mathbf{m})\boldsymbol{u} - \mathbf{s},\boldsymbol{v}\big\rangle + \Re e\big\langle \boldsymbol{d} - \mathcal{P}\boldsymbol{u},\boldsymbol{w}\big\rangle\\
&\quad + \Re e\big\langle \mathcal{A}(\mathbf{m})\delta\boldsymbol{u},\boldsymbol{v}\big\rangle + \Re e\big\langle \mathcal{P}\delta\boldsymbol{u},\boldsymbol{w}\big\rangle.
\end{aligned}$$

So, we have

$$\begin{aligned}
&\mathscr{L}(\mathbf{m},\boldsymbol{d},\boldsymbol{u}+\delta\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}) - \mathscr{L}(\mathbf{m},\boldsymbol{d},\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}) \qquad (A.14)\\
&= \Re e\big\langle \mathcal{A}(\mathbf{m})\delta\boldsymbol{u},\boldsymbol{v}\big\rangle + \Re e\big\langle \mathcal{P}\delta\boldsymbol{u},\boldsymbol{w}\big\rangle,\\
&= \Re e\big\langle \delta\boldsymbol{u},\mathcal{A}(\mathbf{m})^\dagger\boldsymbol{v}\big\rangle + \Re e\big\langle \delta\boldsymbol{u},\mathcal{P}^\dagger\boldsymbol{w}\big\rangle, \quad \text{(by definition of the adjoint, } \langle\mathcal{M}\mathbf{x},\mathbf{y}\rangle = \langle\mathbf{x},\mathcal{M}^\dagger\mathbf{y}\rangle)\\
&= \Re e\big\langle \delta\boldsymbol{u},\mathcal{A}(\mathbf{m})^\dagger\boldsymbol{v} + \mathcal{P}^\dagger\boldsymbol{w}\big\rangle. \qquad \text{(by sesquilinearity of the Hermitian scalar product)}
\end{aligned}$$

Since eq. (A.13) must be verified $\forall\, \delta\boldsymbol{u}\to\mathbf{0}$, we must have

$$\mathcal{A}(\mathbf{m})^\dagger\boldsymbol{v} + \mathcal{P}^\dagger\boldsymbol{w} = \mathbf{0},$$

which is nothing but the adjoint equation

$$\mathcal{A}(\mathbf{m})^\dagger\boldsymbol{v} = -\mathcal{P}^\dagger\boldsymbol{w}, \tag{A.15}$$

where the adjoint source $-\mathcal{P}^\dagger\boldsymbol{w}$ must still be determined, using the last saddle point (A.9). The derivative with respect to the complex variable $\boldsymbol{d}$ is treated as the previous one, by computing

$$\begin{aligned}
\mathscr{L}(\mathbf{m},\boldsymbol{d}+\delta\boldsymbol{d},\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}) &= \frac{1}{2}\big\|\mathbf{d}_{obs} - \boldsymbol{d} - \delta\boldsymbol{d}\big\|^2 + \Re e\big\langle \boldsymbol{d}+\delta\boldsymbol{d} - \mathcal{P}\boldsymbol{u},\boldsymbol{w}\big\rangle,\\
&= \frac{1}{2}\big\|\mathbf{d}_{obs} - \boldsymbol{d}\big\|^2 - \Re e\big\langle \mathbf{d}_{obs} - \boldsymbol{d},\delta\boldsymbol{d}\big\rangle + \frac{1}{2}\big\|\delta\boldsymbol{d}\big\|^2\\
&\quad + \Re e\big\langle \boldsymbol{d} - \mathcal{P}\boldsymbol{u},\boldsymbol{w}\big\rangle + \Re e\big\langle \delta\boldsymbol{d},\boldsymbol{w}\big\rangle.
\end{aligned}$$

So, using the same manipulations as for eq. (A.14), we have

$$\begin{aligned}
\mathscr{L}(\mathbf{m},\boldsymbol{d}+\delta\boldsymbol{d},\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}) - \mathscr{L}(\mathbf{m},\boldsymbol{d},\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}) &= -\Re e\big\langle \mathbf{d}_{obs} - \boldsymbol{d},\delta\boldsymbol{d}\big\rangle + \Re e\big\langle \delta\boldsymbol{d},\boldsymbol{w}\big\rangle + \frac{1}{2}\big\|\delta\boldsymbol{d}\big\|^2,\\
&= -\Re e\big\langle \mathbf{d}_{obs} - \boldsymbol{d} + \boldsymbol{w},\delta\boldsymbol{d}\big\rangle + \frac{1}{2}\big\|\delta\boldsymbol{d}\big\|^2,
\end{aligned}$$

[1] I did not raise this problem for eqs (A.7) and (A.8) because the relation $\partial_z z = 1$ is quite intuitive, and mathematically correct, even in the complex space.

which must be zero $\forall\, \delta \boldsymbol{d} \rightarrow \mathbf{0}$. So we deduce the expression of the adjoint variable $\boldsymbol{w}$, which is nothing but the data residuals:

$$\boldsymbol{w} = \boldsymbol{d} - \mathbf{d}_{obs}. \tag{A.16}$$

Replacing the variables of the Lagrangian function with their actual expressions, we well retrieve the same adjoint equation as in Chapter 2:

$$\mathcal{A}(\mathbf{m})^{\dagger}\mathbf{v}(\mathbf{m}) = \mathcal{P}^{\dagger}(\mathbf{d}_{obs} - \mathbf{d}_{cal}(\mathbf{m})), \tag{A.17}$$

and the resulting adjoint wavefield $\mathbf{v}(\mathbf{m})$ can now be injected in the gradient expression (A.10):

$$G_i(\mathbf{m}) = \Re e \left\{ \mathbf{u}(\mathbf{m})^{\dagger} \frac{\partial \mathcal{A}(\mathbf{m})^{\dagger}}{\partial m_i} \mathbf{v}(\mathbf{m}) \right\}. \tag{A.10 again}$$

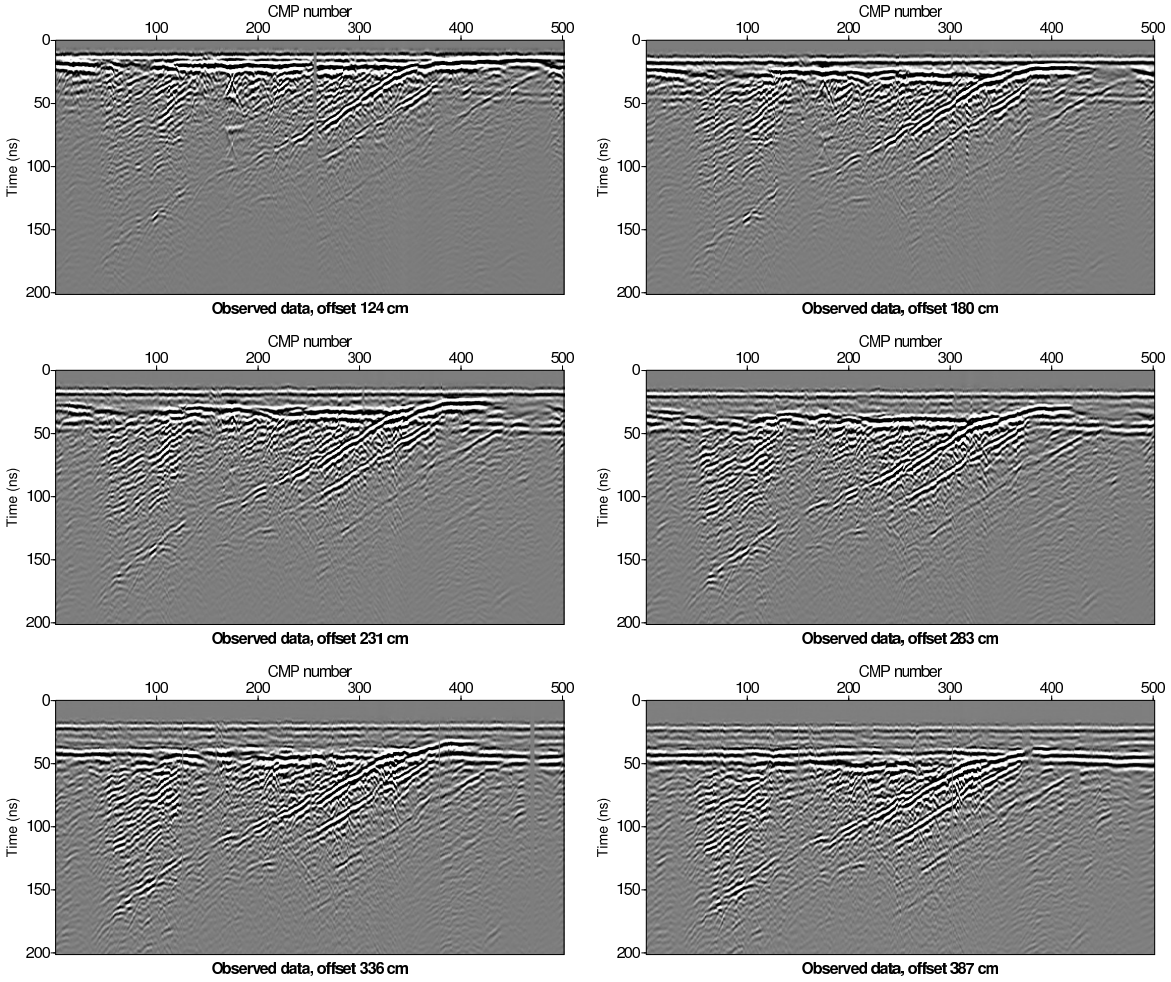
The development of the adjoint state method using Lagrangian multipliers may appear quite artificial here because the gradient (A.10) can be computed using the simple derivation of Section 2.3.1. But it is a more general formulation since it enables to include further constraints to the optimisation. It is also very flexible and provides a systematic method for computing the gradient for various parameterisations, observables or misfit criteria, while the derivation presented in Section 2.3.1 does not always authorise to differentiate the misfit function with respect to the wavefield $\mathbf{u}(\mathbf{m})$ in a straightforward way. Finally, this formulation makes possible the development of second-order adjoint methods that permit an efficient resolution of the Newton system (Métivier et al., 2013).

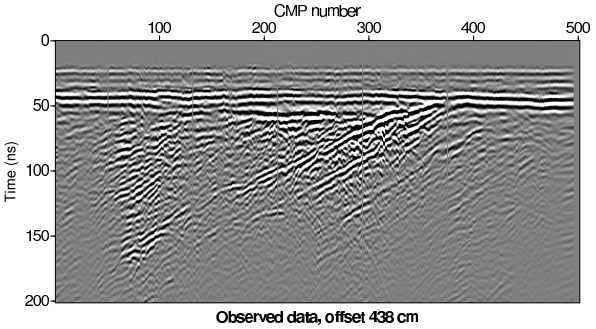
Appendix B

Complete LSBB data set

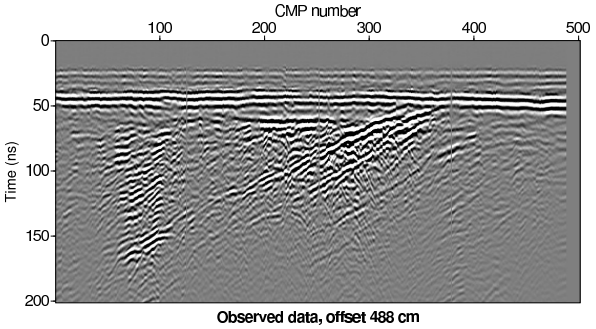
B.1 Raw common-offset sections

Here I present the complete LSBB data set. Correction for t_o and constant component removal have been applied, as well as an amplitude saturation using `perc=97`.

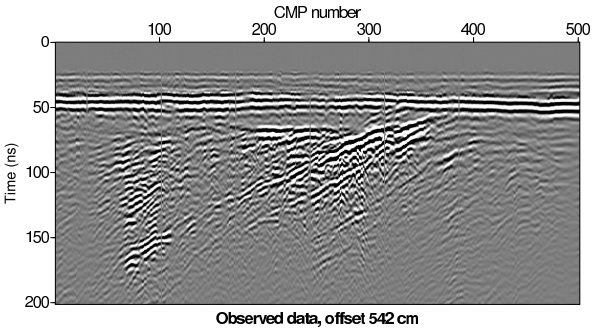




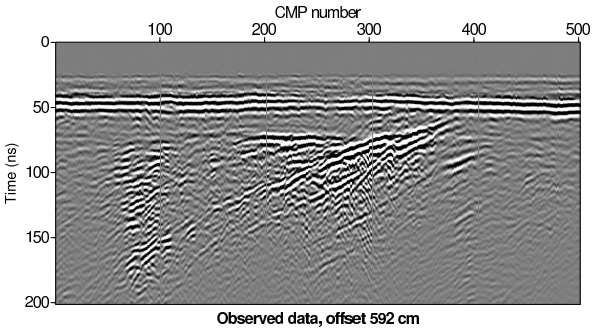
Observed data, offset 438 cm



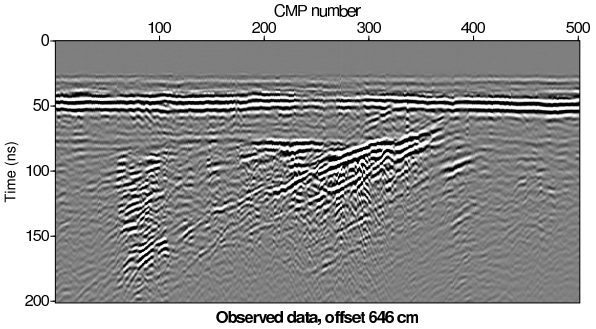
Observed data, offset 488 cm



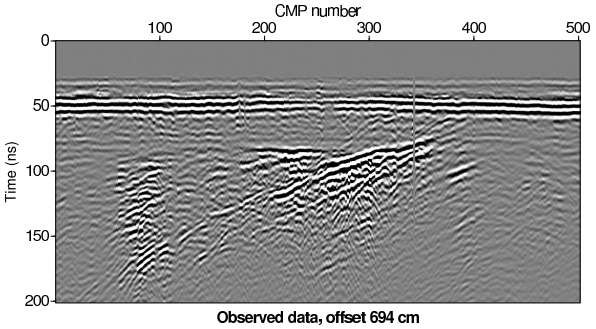
Observed data, offset 542 cm



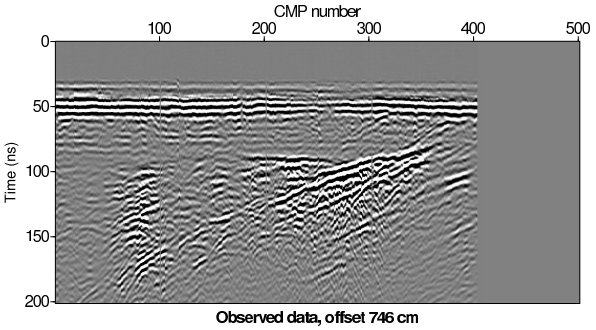
Observed data, offset 592 cm



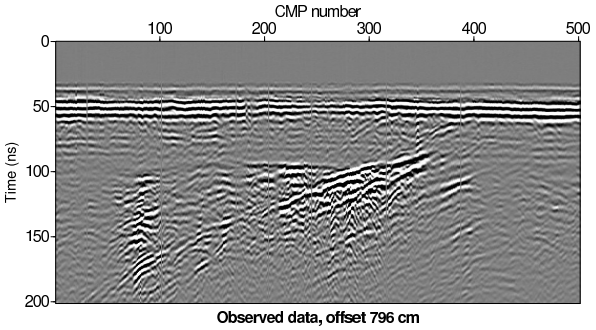
Observed data, offset 646 cm



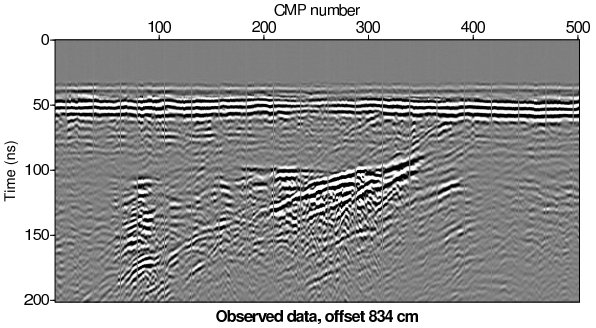
Observed data, offset 694 cm



Observed data, offset 746 cm



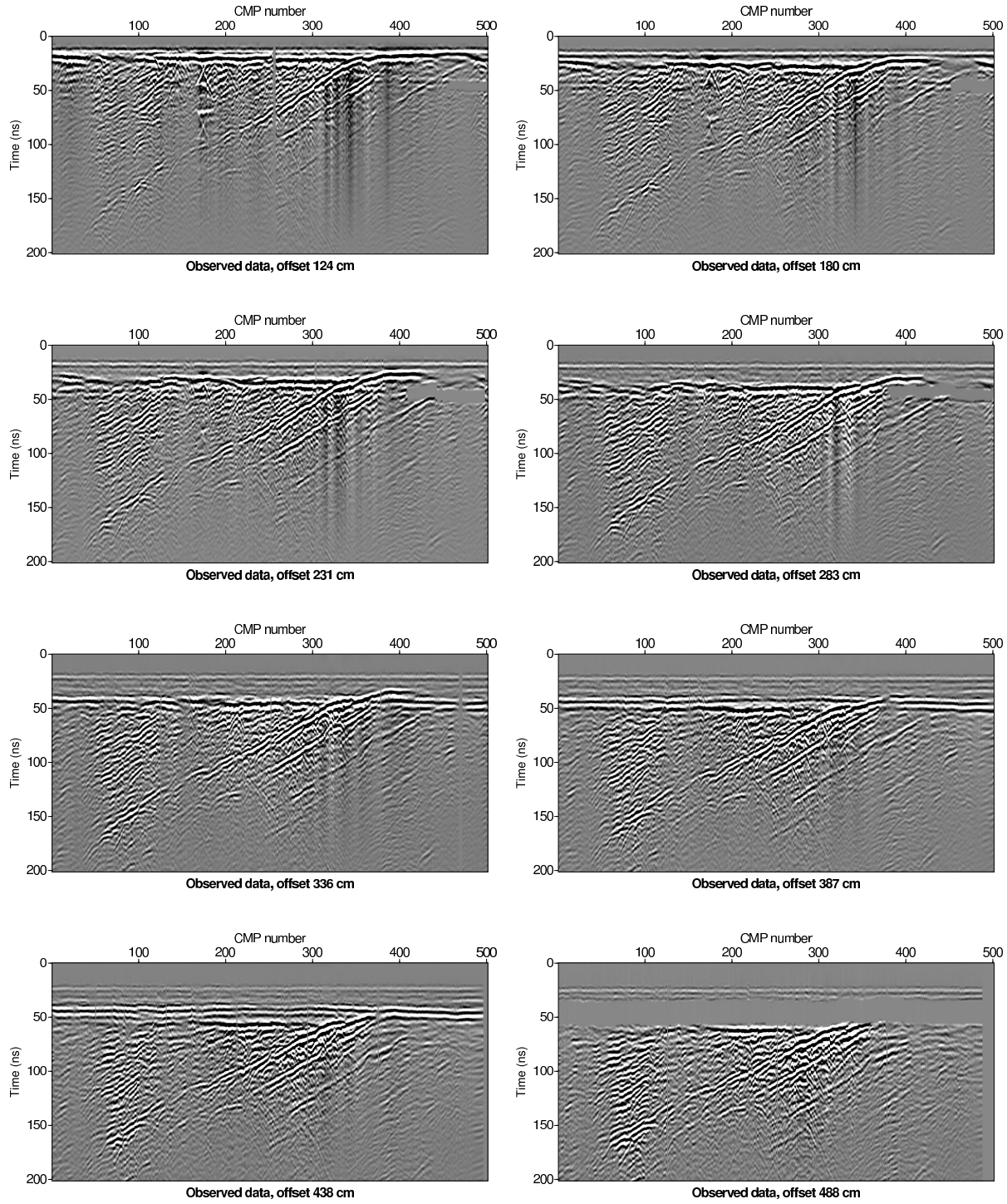
Observed data, offset 796 cm

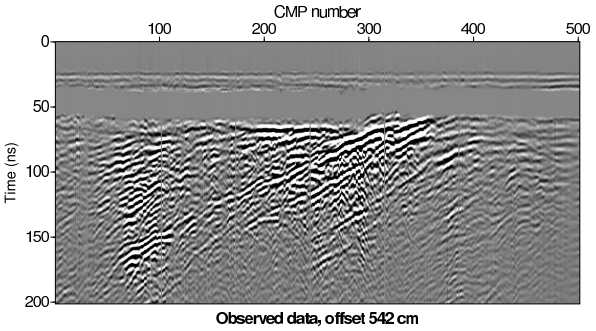


Observed data, offset 834 cm

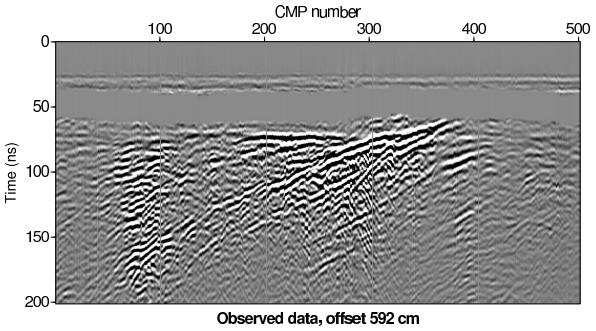
B.2 Processed common-offset sections

Here are the pre-processed data. Processing consisted in 3D-to-2D conversion and in muting the parasite reflected air-wave (RAW in Fig. 3.21). Again, amplitudes are saturated using `perc=97`.

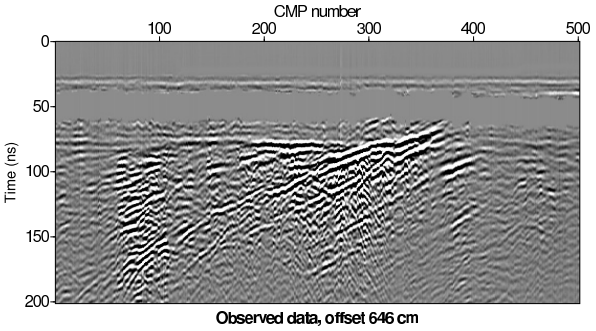




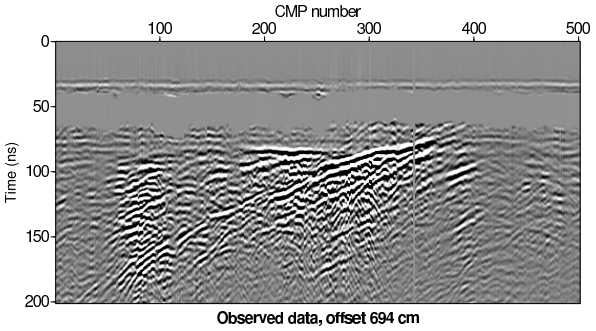
Observed data, offset 542 cm



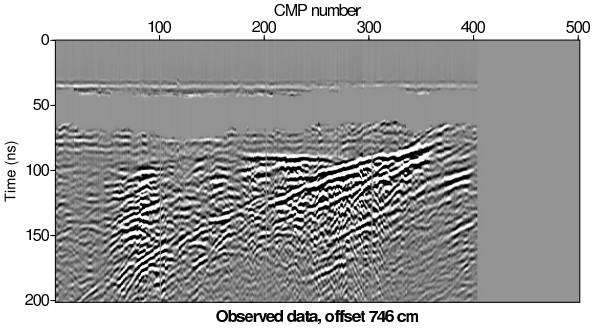
Observed data, offset 592 cm



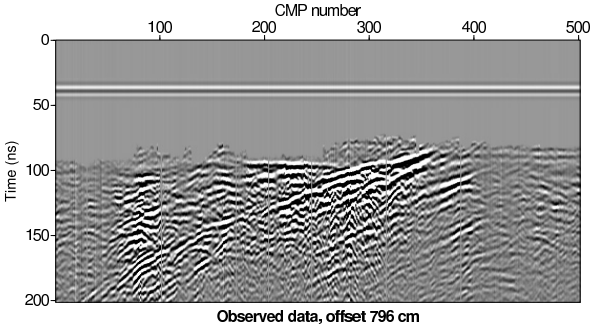
Observed data, offset 646 cm



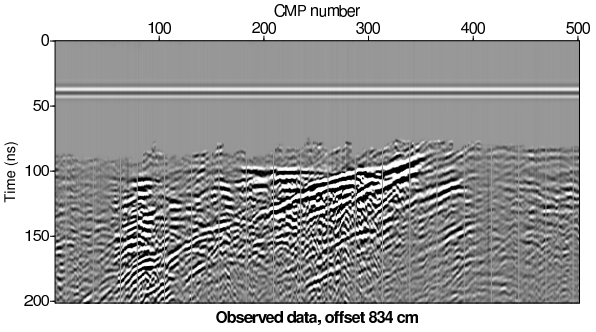
Observed data, offset 694 cm



Observed data, offset 746 cm



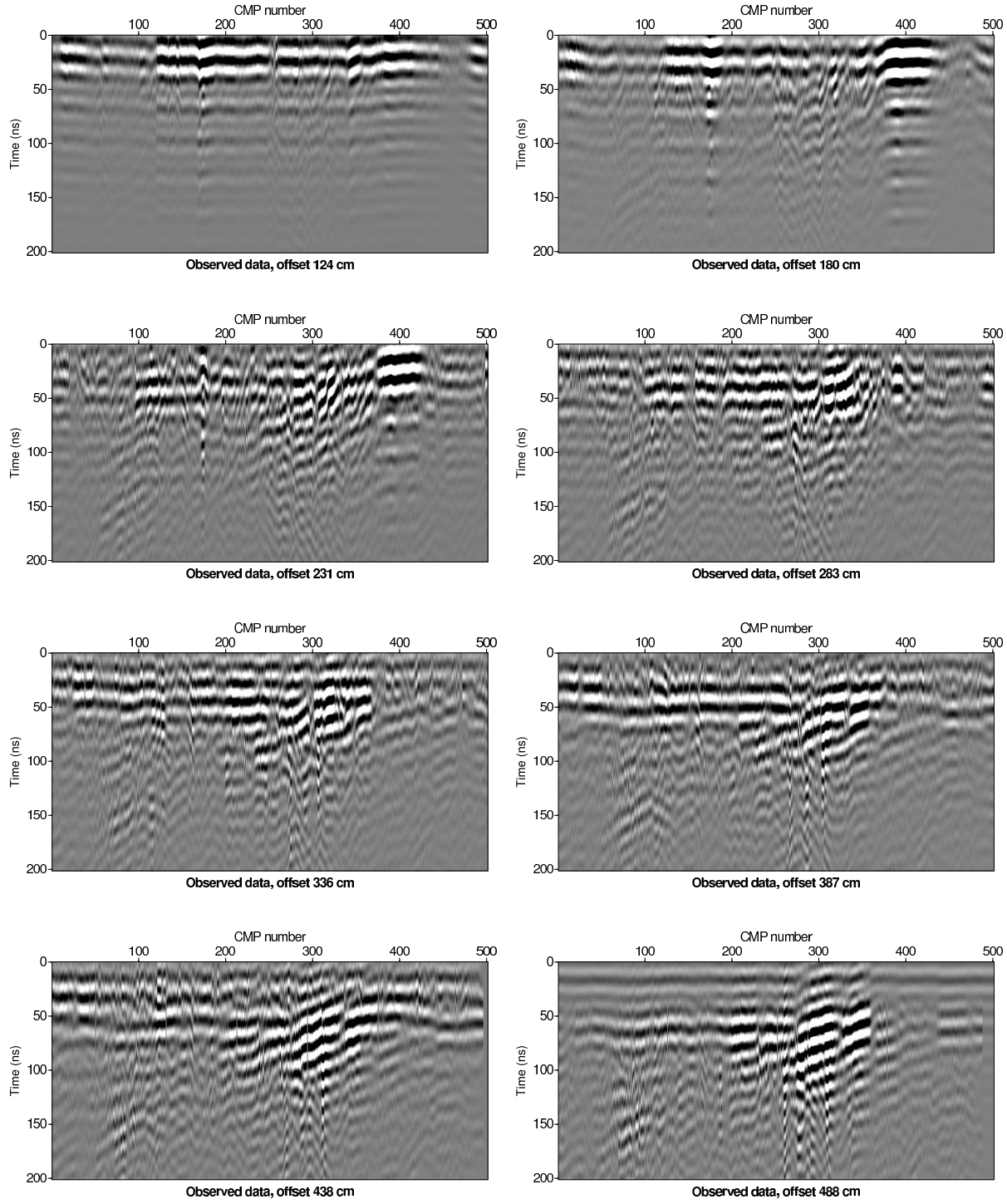
Observed data, offset 796 cm

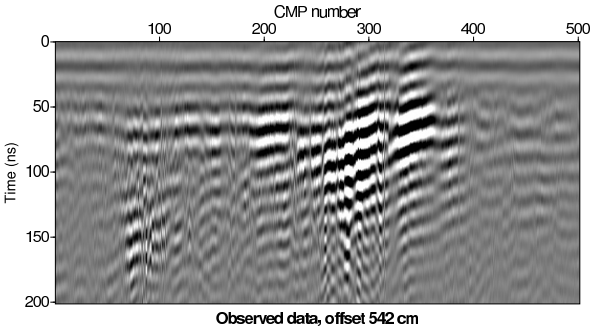


Observed data, offset 834 cm

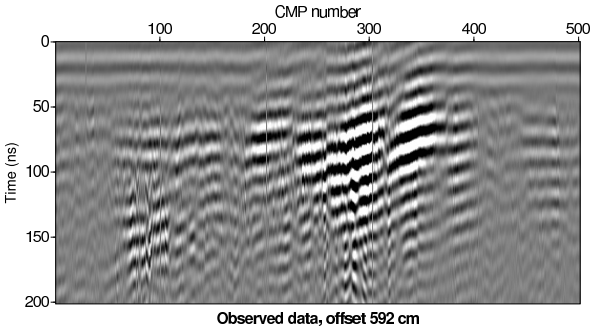
B.3 Filtered common-offset sections

Filtering is applied using `sufilter f=38,40,69,71 MHz`, which makes use of a zero-phase, sine-squared tapered filter (Cohen and Stockwell, 2008).

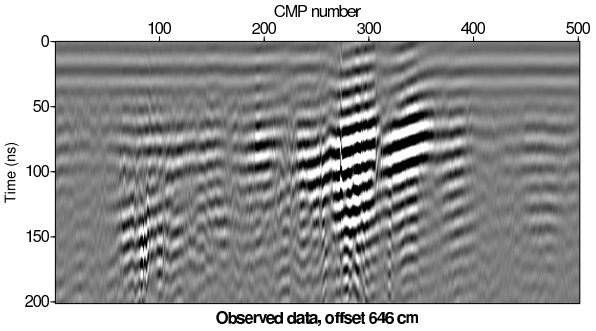




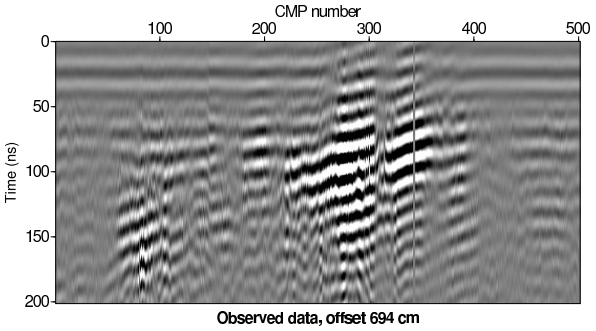
Observed data, offset 542 cm



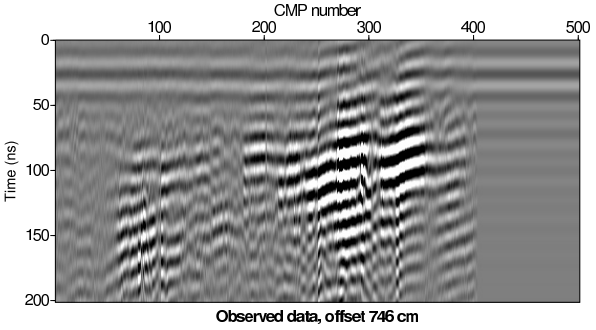
Observed data, offset 592 cm



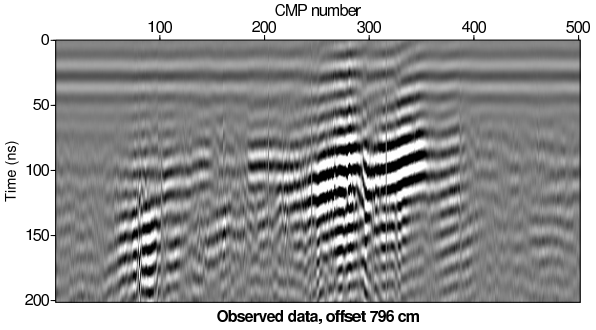
Observed data, offset 646 cm



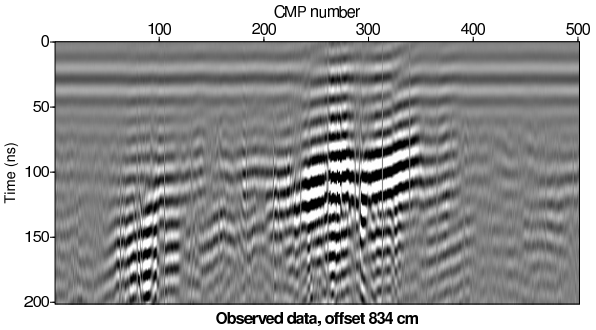
Observed data, offset 694 cm



Observed data, offset 746 cm



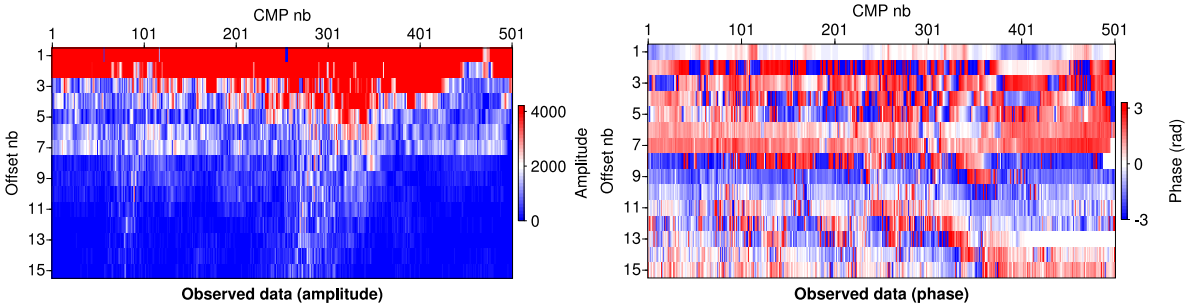
Observed data, offset 796 cm



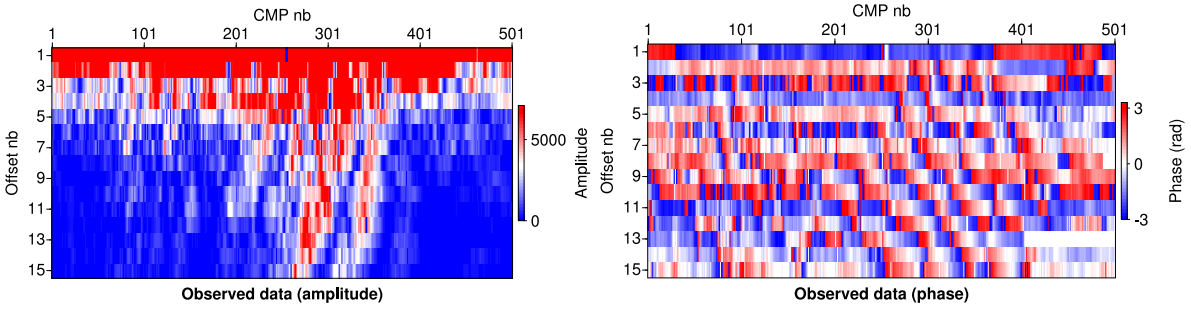
Observed data, offset 834 cm

B.4 Frequency-domain data

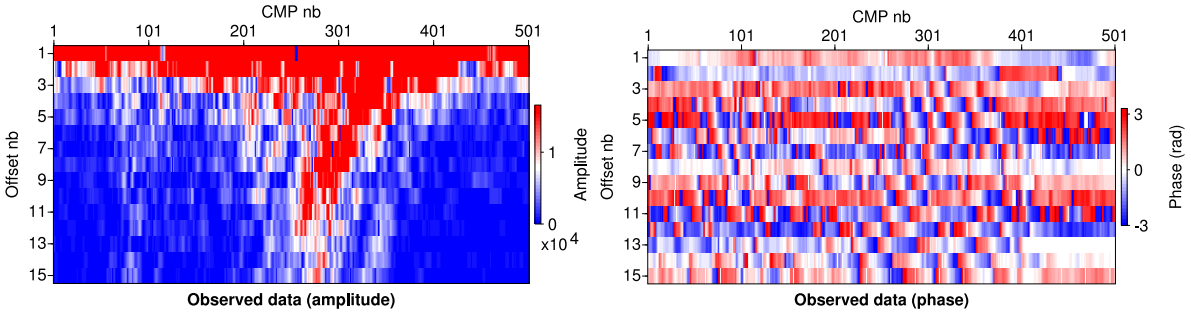
Here I represent the observed data in the frequency-domain, after 3D-to-2D conversion and with their true amplitudes. Contrary to Figs 3.38, 3.39, 3.43 and 3.44, no data weighting vs. offset has been applied. All the offsets are shown, including those that are not inverted.



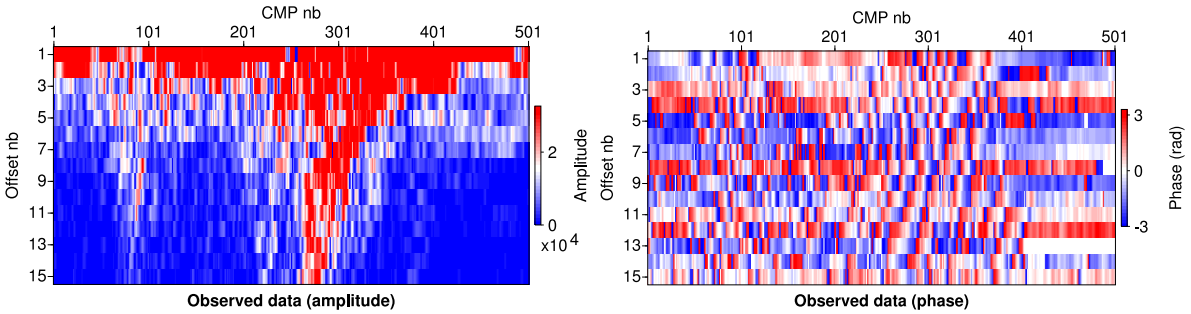
(a) Frequency 40 MHz.



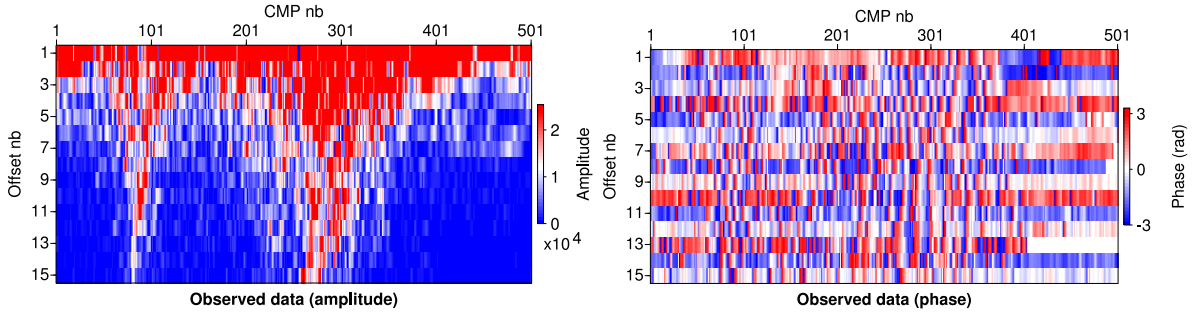
(b) Frequency 70 MHz.



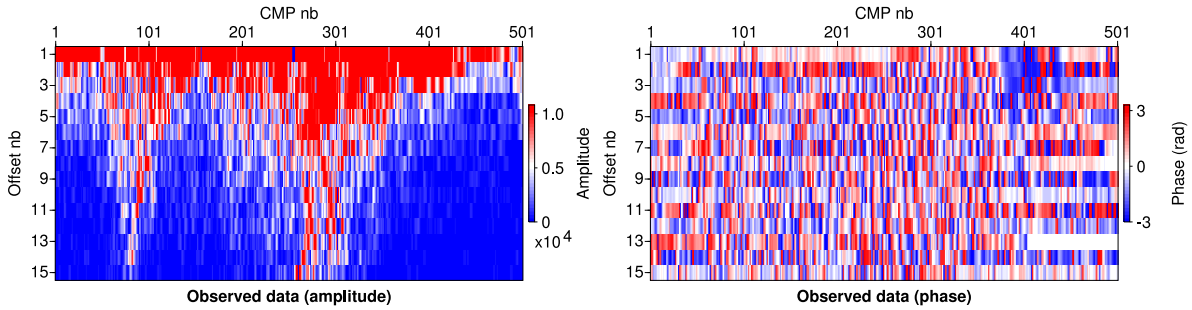
(c) Frequency 100 MHz.



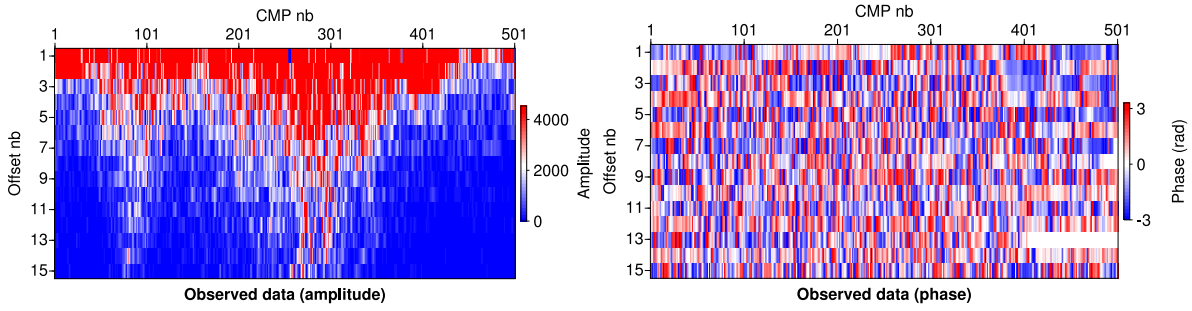
(d) Frequency 150 MHz.



(e) Frequency 200 MHz.



(f) Frequency 250 MHz.

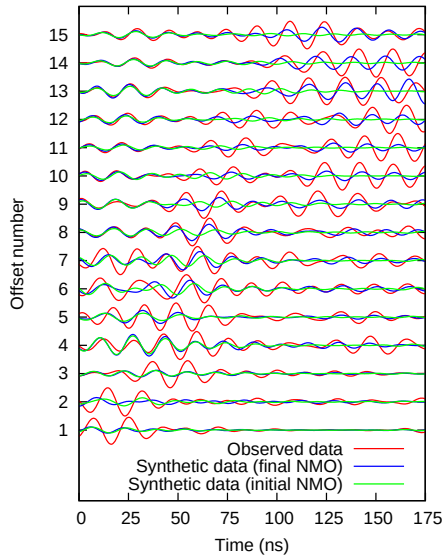


(g) Frequency 300 MHz.

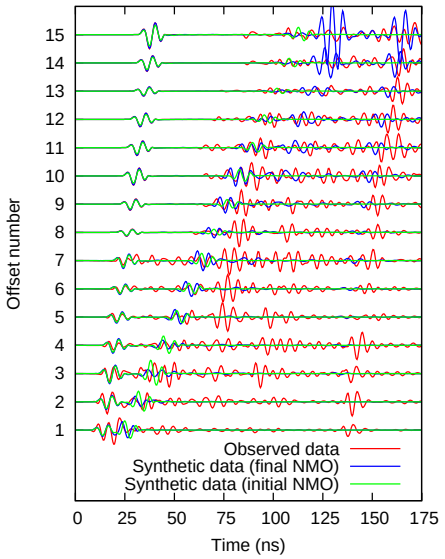
B.5 Time-domain data fit

For each figure, panels (a) and (b) present synthetic data computed in the initial NMO model of Fig. 3.23(b) (green line) and in the corresponding final reconstructed model of Fig. 3.41(b) (blue line). In panels (c,d), synthetics are computed in the initial blocky model of Fig. 3.26 (green) and in the corresponding final reconstructed model of Fig. 3.41(a) (blue). Panels (a,c) present data filtered in the frequency range considered for inversion (40-70 MHz) whereas panels (b,d) show the time-domain traces in the full frequency band.

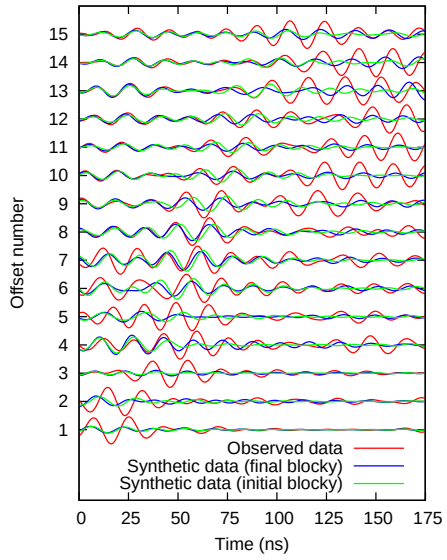
I remind that offsets number 1, 2, 5, 6 and 7 are not considered in the inversion. Data are normalised trace-by-trace with the observed amplitudes as reference, so that amplitudes are comparable between the data sets.



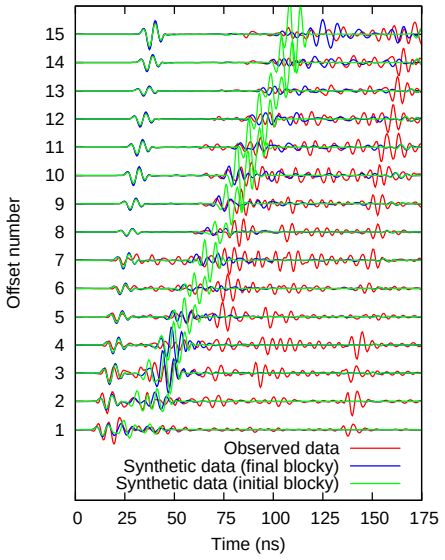
(a) CMP 100, NMO model (40-70 MHz).



(b) CMP 100, NMO model (no filter).

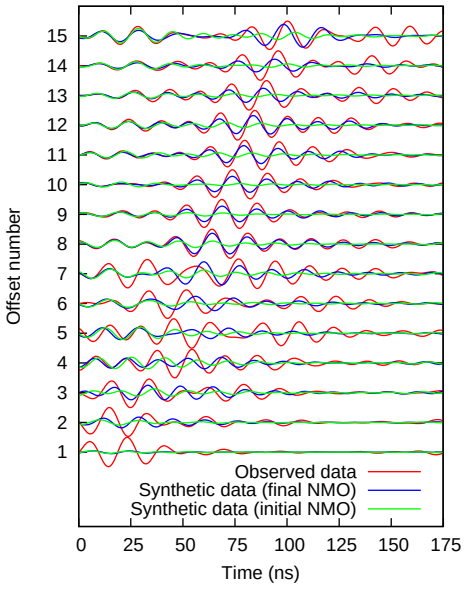


(c) CMP 100, blocky model (40-70 MHz).

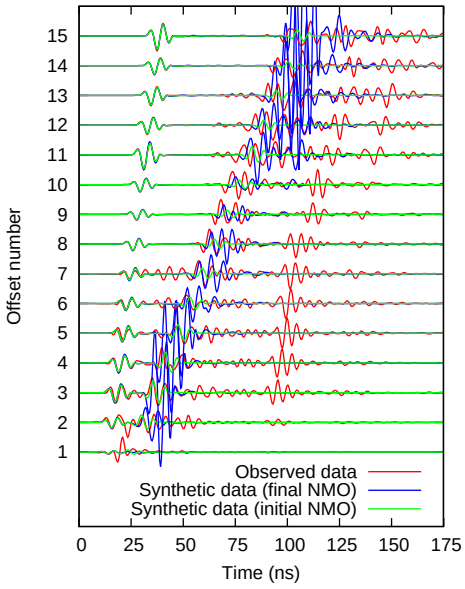


(d) CMP 100, blocky model (no filter).

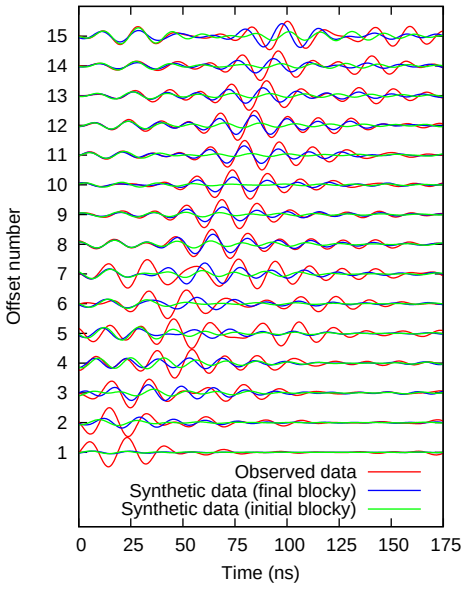
Figure B.1: *Observed vs. synthetic data at CMP 100.*



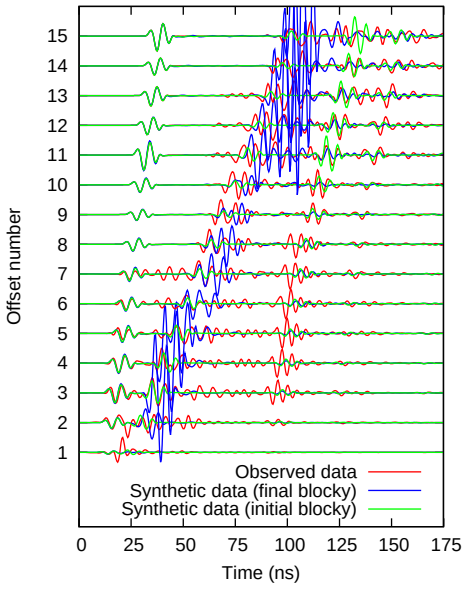
(a) CMP 200, NMO model (40-70 MHz).



(b) CMP 200, NMO model (no filter).

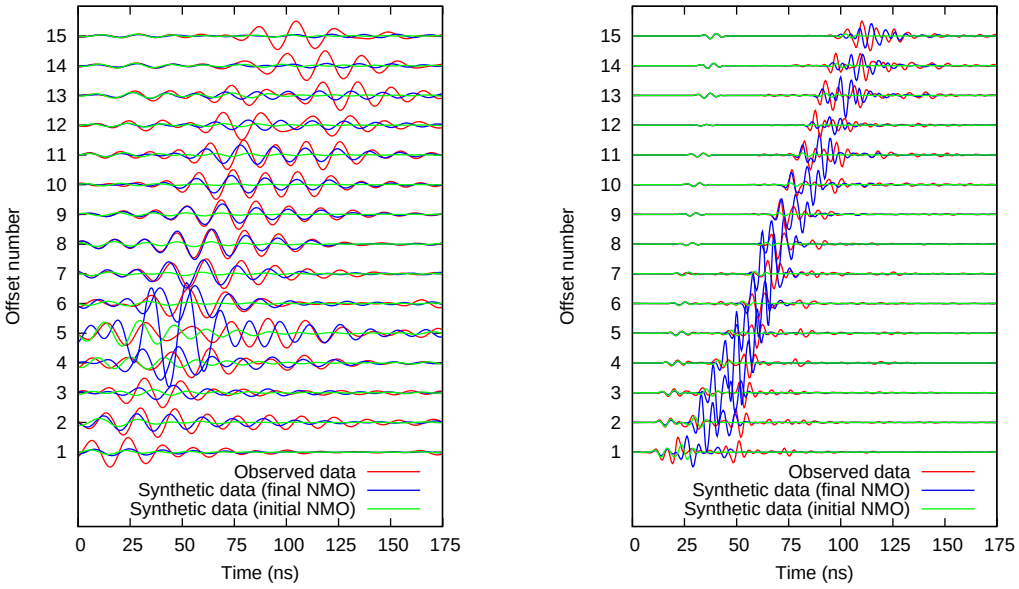


(c) CMP 200, blocky model (40-70 MHz).



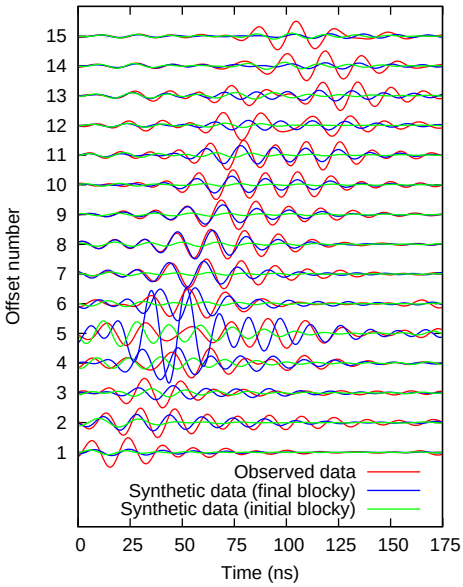
(d) CMP 200, blocky model (no filter).

Figure B.2: *Observed vs. synthetic data at CMP 200.*

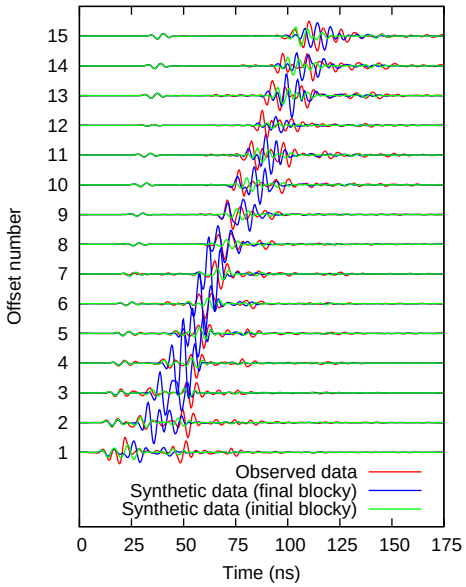


(a) CMP 300, NMO model (40-70 MHz).

(b) CMP 300, NMO model (no filter).

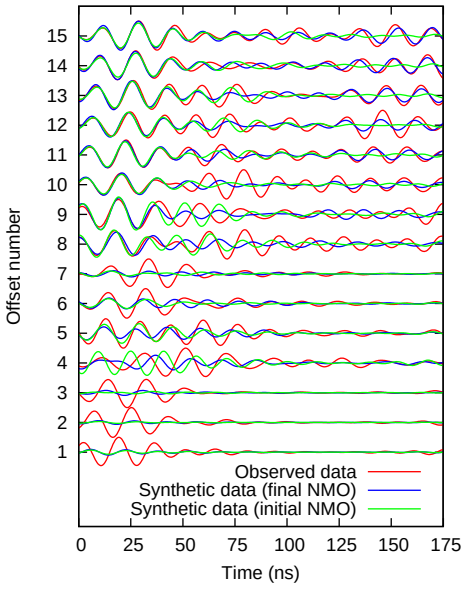


(c) CMP 300, blocky model (40-70 MHz).

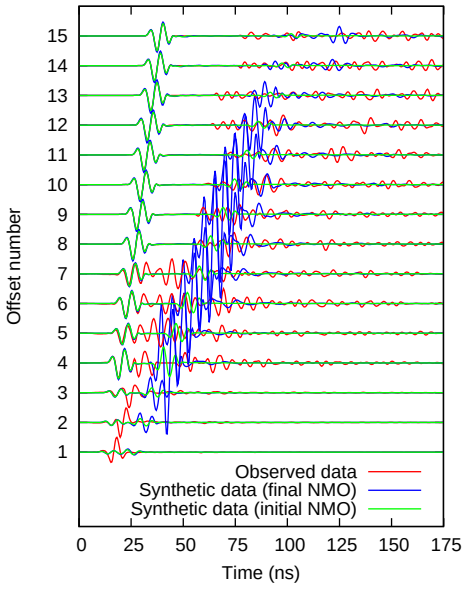


(d) CMP 300, blocky model (no filter).

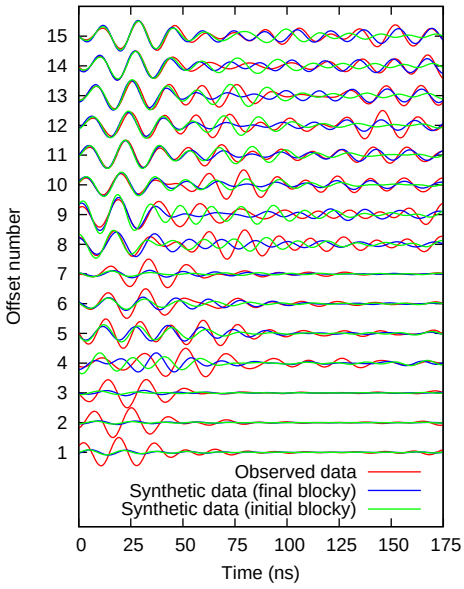
Figure B.3: *Observed vs. synthetic data at CMP 300.*



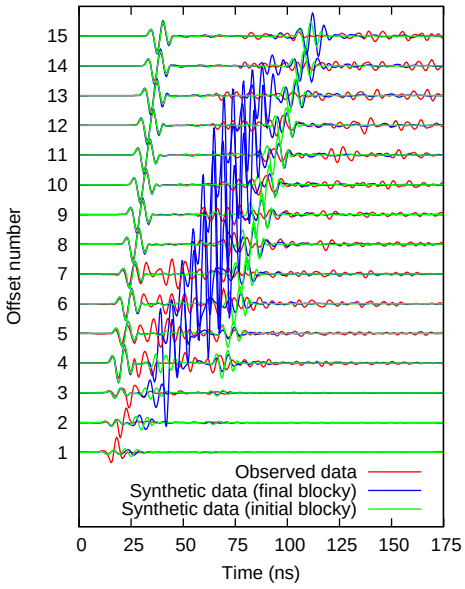
(a) CMP 400, NMO model (40-70 MHz).



(b) CMP 400, NMO model (no filter).

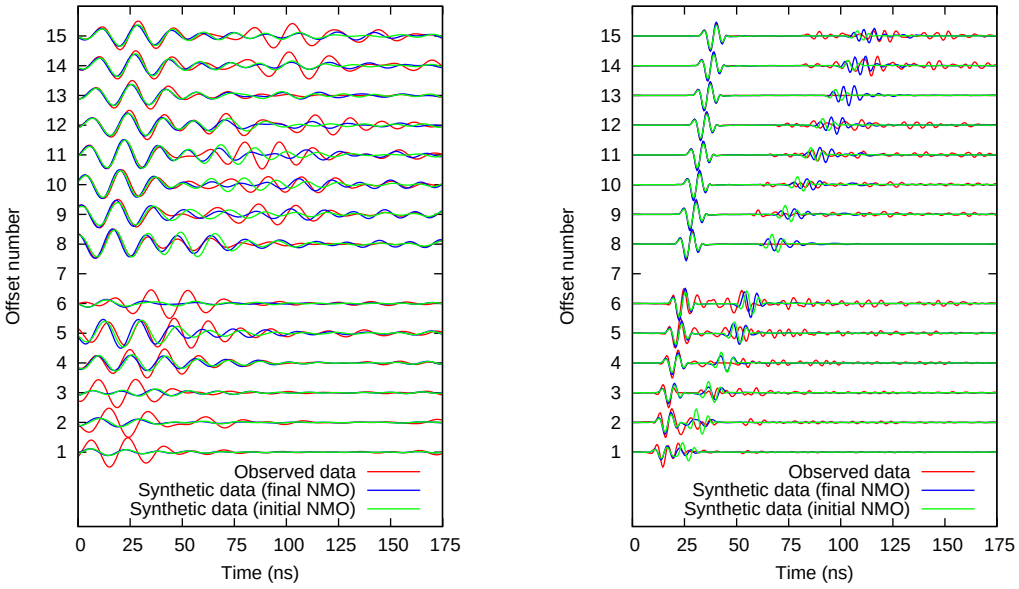


(c) CMP 400, blocky model (40-70 MHz).



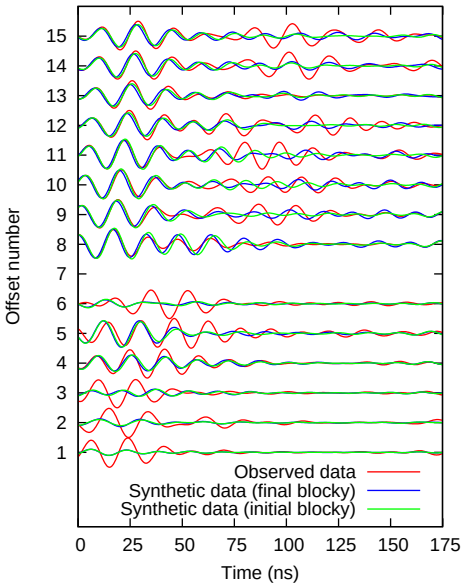
(d) CMP 400, blocky model (no filter).

Figure B.4: *Observed vs. synthetic data at CMP 400.*

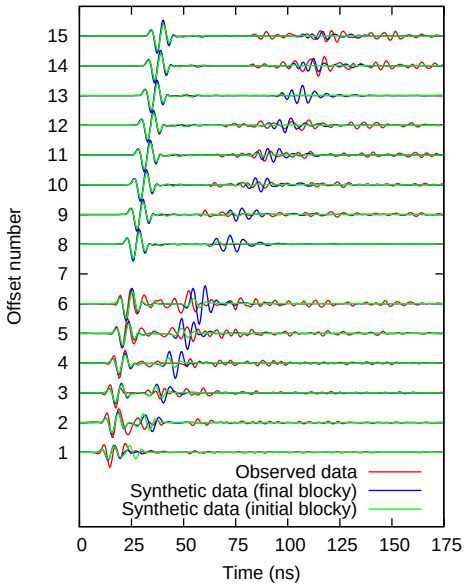


(a) CMP 500 (40-70 MHz).

(b) CMP 500 (no filter).



(c) CMP 500, blocky model (40-70 MHz).



(d) CMP 500, blocky model (no filter).

Figure B.5: *Observed vs. synthetic data at CMP 500.*

Bibliography

Abramowitz, M. and Stegun, I. A. (1972). *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. Dover publications, Inc., New York.

Abubakar, A., Habashy, T. M., Lin, Y., and Li, M. (2012). A model-compression scheme for nonlinear electromagnetic inversions. *Geophysics*, 77(5):E379–E389.

Abubakar, A., van den Berg, P. M., and Habashy, T. M. (2005). Application of the multiplicative regularized contrast source inversion method on TM- and TE-polarized experimental Fresnel data. *Inverse Problems*, 21:S5–S13.

al Hagrey, S. A. and Müller, C. (2000). GPR study of pore water content and salinity in sand. *Geophysical Prospecting*, 48:63–85.

Annan, A. P. (1996). Transmission Dispersion and GPR. *Journal of Environmental and Engineering Geophysics*, 1(B):125–136.

Annan, A. P. (2001). Ground penetrating radar - Worshop notes: Sensors and Softwares. Technical report, Mississauga, Ontario.

Archie, G. E. (1942). The electrical resistivity log as an aid in determining some reservoir characteristics. *Transactions of the American Intitute of Mining and Metallurgical Engineers*, 146:55–62.

Arcone, S. A. (1995). Numerical studies of the radiation pattern of resistively loaded dipoles. *Applied Geophysics*, 33:39–52.

Asnaashari, A., Brossier, R., Garambois, S., Audebert, F., Thore, P., and Virieux, J. (2013). Regularized seismic full waveform inversion with prior model information. *Geophysics*, 78(2):R25–R36.

Bailey, J. T., Evans, S., and Robin, G. Q. (1964). Radio echo sounding of polar ice sheets. *Nature*, 204(4957):420–421.

Beauprêtre, S. (2013). *Développement d'une approche de paléosismologie géophysique par imagerie Géoradar. Applications aux failles décrochantes actives de Nouvelle Zélande*. PhD thesis, Université de Grenoble.

Beauprêtre, S., Manighetti, I., Garambois, S., Malavieille, J., and Dominguez, S. (2013). Stratigraphic architecture and fault offsets of alluvial terraces at Te Marua, Wellington fault, New Zealand, revealed by pseudo-3D GPR investigation. *Journal of Geophysical Research: Solid Earth*, 118(8):4564–4585.

Belina, F., Irving, J., Ernst, J., and Holliger, K. (2012a). Analysis of an iterative deconvolution approach for estimating the source wavelet during waveform inversion of crosshole georadar data. *Journal of Applied Geophysics*, 78:20–30.

Belina, F., Irving, J., Ernst, J., and Holliger, K. (2012b). Waveform inversion of crosshole georadar data: Influence of source wavelet variability and the suitability of a single wavelet assumption. *IEEE Transactions on Geoscience and Remote Sensing*, 50(11):4610–4625.

Belina, F. A., Ernst, J. R., and Holliger, K. (2009). Inversion of crosshole seismic data in heterogeneous environments: Comparison of waveform and ray-based approaches. *Journal of Applied Geophysics*, 68:85–94.

Belkebir, K. and Saillard, M. (2001). Special section: Testing inversion algorithms against experimental data. *Inverse Problems*, 17:1565–1571.

Belkebir, K. and Saillard, M. (2005). Guest Editors' introduction: Testing inversion algorithms against experimental data: inhomogeneous targets. *Inverse Problems*, 21:1–3.

Bentley, C. R. (1964). The structure of Antarctica and its ice cover. In Odishaw, H., editor, *Research in Geophysics, v.2 Solid Earth and Interface Phenomena*, pages 335–389. Technology Press of Massachusetts Institute of Technology.

Bérenger, J.-P. (1994). A perfectly matched layer for absorption of electromagnetic waves. *Journal of Computational Physics*, 114:185–200.

Berkhout, A. J. (2012). Combining full wavefield migration and full waveform inversion, a glance into the future of seismic imaging. *Geophysics*, 77(2):S43–S50.

Bleistein, N. (1986). Two-and-one-half dimensional in-plane wave propagation. *Geophysical Prospecting*, 34:686–703.

Blondel, T. (2008). *Traçage spatial et temporel des eaux souterraines dans les hydrosystèmes karstiques par les matières organiques dissoutes: expérimentation et application sur les sites du Laboratoire Souterrain à Bas Bruit (LSBB) de Rustrel-Pays d'Apt et de Fontaine de Vaucluse.* PhD thesis, Université d'Avignon.

Böniger, U. and Tronicke, J. (2010a). Improving the interpretability of 3D GPR data using target-specific attributes: application to tomb detection. *Journal of Archaeological Science*, 37:672–679.

Böniger, U. and Tronicke, J. (2010b). Integrated data analysis at an archaeological site: A case study using 3D GPR, magnetic, and high-resolution topographic data. *Geophysics*, 75(4):B169–B176.

Bouchedda, A., Chouteau, M., Binley, A., and Giroux, B. (2012). 2-D joint structural inversion of cross-hole electrical resistance and ground penetrating radar data. *Journal of Applied Geophysics*, 78:52–67.

Bourdi, T., Rhazi, J. E., Boone, F., and Ballivy, G. (2008). Application of Jonscher model for the characterization of the dielectric permittivity of concrete. *Journal of Physics D: Applied Physics*, 41(20):205410.

Bradford, J. H. (2006). Applying reflection tomography in the postmigration domain to multifold ground-penetrating radar data. *Geophysics*, 71(1):K1–K8.

Bradford, J. H. (2008). Measuring water content heterogeneity using multifold GPR with reflection tomography. *Vadose Zone Journal*, 7(1):184–193.

Bradford, J. H., Clement, W. P., and Barrash, W. (2009a). Estimating porosity with ground-penetrating radar reflection tomography: A controlled 3-D experiment at the Boise hydrogeophysical research site. *Water Resources Research*, 45:W00D26.

Bradford, J. H. and Deeds, J. C. (2006). Ground-penetrating radar theory and application of thin-bed offset dependent reflectivity. *Geophysics*, 71(3):K47–K57.

Bradford, J. H., Harper, J. T., and Brown, J. (2009b). Complex dielectric permittivity measurements from ground-penetrating radar data to estimate snow liquid water content in the pendular regime. *Water Resources Research*, 45:W08403.

Bristow, C. S. and Jol, H. M., editors (2003). *Ground Penetrating Radar in Sediments*. Geological Society, London.

Brossier, R., Operto, S., and Virieux, J. (2009). Robust elastic frequency-domain full-waveform inversion using the l_1 norm. *Geophysical Research Letters*, 36:L20310.

Brossier, R., Operto, S., and Virieux, J. (2014). Velocity model building from seismic reflection data by full waveform inversion. *Geophysical Prospecting*, in press.

Brossier, R., Virieux, J., and Operto, S. (2008). Parsimonious finite-volume frequency-domain method for 2-D P-SV-wave modelling. *Geophysical Journal International*, 175(2):541–559.

Bungey, J. H. (2004). Sub-surface radar testing of concrete: a review. *Construction and Building Materials*, 18:1–8.

Bunks, C., Salek, F. M., Zaleski, S., and Chavent, G. (1995). Multiscale seismic waveform inversion. *Geophysics*, 60(5):1457–1473.

Busch, S., van der Kruk, J., Bikowski, J., and Vereecken, H. (2012). Quantitative conductivity and permittivity estimation using full-waveform inversion of on-ground GPR data. *Geophysics*, 77(6):H79–H91.

Busch, S., van der Kruk, J., and Vereecken, H. (2014). Improved characterization of fine-texture soils using on-ground GPR full-waveform inversion. *IEEE Transactions on Geoscience and Remote Sensing*, (accepted).

Byrd, R. H., Lu, P., and Nocedal, J. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific and Statistical Computing*, 16:1190–1208.

Cai, W., Qin, F., and Schuster, G. T. (1996). Electromagnetic velocity inversion using 2-D Maxwell's equations. *Geophysics*, 61(4):1007–1021.

Canouet, N. (2003). *Méthodes de Galerkin Discontinu pour la résolution du système de Maxwell sur des maillages localement raffinés non-conformes*. PhD thesis, Ecole Nationale des Ponts et Chaussées.

Carcione, J. M. (1996). Ground-penetrating radar: Wave theory and numerical simulation in lossy anisotropic media. *Geophysics*, 61(6):1664–1677.

Carcione, J. M. (1998). Radiation patterns for 2-D GPR forward modeling. *Geophysics*, 63(2):424–430.

Carcione, J. M. and Cavallini, F. (1995). On the acoustic-electromagnetic analogy. *Wave motion*, 21:149–162.

Carcione, J. M. and Robinson, E. M. (2002). On the acoustic-electromagnetic analogy for the reflection-refraction problem. *Studia Geophysica et Geodaetica*, 46:321–346.

Carcione, J. M. and Schoenberg, M. A. (2000). 3-D ground-penetrating radar simulation and plane-wave theory in anisotropic media. *Geophysics*, 65(5):1527–1541.

Carrière, S. D., Chalikakis, K., Sénéchal, G., Danquigny, C., and Emblanch, C. (2013). Combining Electrical Resistivity Tomography and Ground Penetrating Radar to study geological structuring of karst Unsaturated Zone. *Journal of Applied Geophysics*, 94:31–41.

Cassidy, N. J. (2009a). Electrical and magnetic properties of rocks, soils and fluids. In Jol, H. M., editor, *Ground Penetrating Radar: Theory and Applications*, chapter 2, pages 41–72. Elsevier.

Cassidy, N. J. (2009b). Ground penetrating radar data processing, modelling and analysis. In Jol, H. M., editor, *Ground Penetrating Radar: Theory and Applications*, chapter 5, pages 141–176. Elsevier.

Castellanos-Lopez, C. (2014). *Speed-up and regularization techniques for seismic full waveform inversion*. PhD thesis, Université de Nice-Sophia Antipolis.

Chan, C. Y. and Knight, R. J. (1999). Determining water content and saturation from dielectric measurements in layered materials. *Water Resources Research*, 35(1):85–93.

Chavent, G. (1974). Identification of parameter distributed systems. In Goodson, R. and Polis, M., editors, *Identification of function parameters in partial differential equations*, pages 31–48. American Society of Mechanical Engineers, New York.

Chew, W. C. (1995). *Waves and Fields in Inhomogeneous Media*. IEEE Press, New York.

Chew, W. C., Tong, M. S., and Hu, B. (2009). *Integral equation methods for electromagnetic and elastic waves*, volume 3 of *Synthesis Lectures on Computational Electromagnetics*. Morgan & Claypool Publishers.

Cioni, J.-P. (1995). *Résolution numérique des équations de Maxwell instationnaires par une méthode de volumes finis*. PhD thesis, Université de Nice-Sophia Antipolis.

Claerbout, J. (1971). Towards a unified theory of reflector mapping. *Geophysics*, 36:467–481.

Clipper Controls® (2014). Dielectric constants of various materials. `http://www.clippercontrols.com/pages/Dielectric-Constant-Values.html#C`. Accessed 18 January 2014.

Cohen, G. and Monk, P. (1998). Gauss point mass lumping schemes for Maxwell's equations. *Numerical Methods for Partial Differential Equations*, 14:63–88.

Cohen, J. K. and Stockwell, J. J. W. (2008). CWP/SU: Seismic Unix release No. 41: an open source software package for seismic research and processing. *Center for Wave Phenomena, Colorado School of Mines*.

Cole, K. S. and Cole, R. H. (1941). Dispersion and absorption in dielectrics. *Journal of Chemical Physics*, 9:341–351.

Conyers, L. B. (2007). Ground-penetrating radar for archaeological mapping. In Wiseman, J. and El-Baz, F., editors, *Remote Sensing in Archaeology*, pages 329–344. Springer.

Cordua, K. S., Hansen, T. M., and Mosegaard, K. (2012). Monte Carlo full-waveform inversion of crosshole GPR data using multiple-point geostatistical a priori information. *Geophysics*, 77(2):H19–H31.

Daniels, D. J. (2006). A review of GPR for landmine detection. *Sensing and Imaging: An International Journal*, 7(3):90–123.

Daniels, D. J. (2008). A review of landmine detection using GPR. In *5th European Radar Conference, Amsterdam (The Netherlands)*, pages 280–283.

Davidson, D. W. and Cole, R. H. (1951). Dielectric relaxation in glycerol, propylene glycol, and *n*-propanol. *The Journal of Chemical Physics*, 19(12):1484–1490.

Davis, J. L. and Annan, A. P. (1989). Ground penetrating radar for high-resolution mapping of soil and rock stratigraphy. *Geophysical Prospecting*, 37:531–551.

Davy, M. (2010). *Application du retournement temporel en micro-ondes à l'amplification d'impulsions et l'imagerie*. PhD thesis, Université Denis Diderot - Paris 7.

Day-Lewis, F. D., Lane Jr., J. W. J., Harris, J. M., and Gorelick, S. M. (2003). Time-lapse imaging of saline-tracer transport in fractured rock using difference-attenuation radar tomography. *Water Resources Research*, 39(10):1290.

Day-Lewis, F. D., Singha, K., and Binley, A. M. (2005). Applying petrophysical models to radar travel time and electrical resistivity tomograms: Resolution-dependent limitations. *Journal of Geophysical Research*, 110(B8).

Debye, P. (1929). Polar molecules. *Chemical Catalogue Co.*

Deeds, J. and Bradford, J. (2002). Characterization of an aquitard and direct detection of LNAPL at Hill Air Force base using GPR AVO and migration velocity analyses. In *9th International Conference on Ground Penetrating Radar (GPR 2002), Santa Barbara, California (USA)*, volume 4758 of *SPIE proceedings series*, pages 323–329.

Deparis, J. (2007). *Etude des éboulements rocheux par méthodes géophysiques.* PhD thesis, Université Joseph Fourier, Grenoble 1.

Deparis, J. and Garambois, S. (2009). On the use of dispersive APVO GPR curves for thin-bed properties estimation: Theory and application to fracture characterization. *Geophysics*, 74(1):J1–J12.

Derode, B., Cappa, F., Guglielmi, Y., and Rutqvist, J. (2013). Coupled seismo-hydromechanical monitoring of inelastic effects on injection-induced fracture permeability. *International Journal of Rock Mechanics and Mining Sciences*, 61:266–274.

Diamanti, N. and Annan, A. P. (2013). Characterizing the energy distribution around GPR antennas. *Journal of Applied Geophysics*, 99:83–90.

Diamanti, N., Annan, P. A., and Redman, D. (2013). Quantifying GPR transient waveforms in the intermediate zone. In *7th International Workshop on Advanced Ground-Penetrating Radar (IWAGPR 2013), Nantes (France).*

Dix, C. H. (1955). Seismic velocities from surface measurements. *Geophysics*, 20:68–86.

Dolean, V., Fol, H., Lanteri, S., and Piperno, S. (2006). Méthodes de type Galerkin discontinu pour la résolution numérique des équations de Maxwell en régime fréquentiel. Technical Report 5904, INRIA.

Doolittle, J. A., Jenkinson, B., Hopkins, D., Ulmer, M., and Tuttle, W. (2006). Hydropedological investigations with ground-penetrating radar (GPR): Estimating water-table depths and local groundwater flow pattern in areas of coarse-textured soils. *Geoderma*, 131:317–329.

Dorn, C., Linde, N., Doetsch, J., Borgne, T. L., and Bour, O. (2012). Fracture imaging within a granitic rock aquifer using multiple-offset single-hole and cross-hole GPR reflection data. *Journal of Applied Geophysics*, 78:123–132.

El Bouajaji, M., Lanteri, S., and Yedlin, M. (2011). Discontinuous Galerkin frequency domain forward modelling for the inversion of electric permittivity in the 2D case. *Geophysical Prospecting*, 59(5):920–933.

Ellefsen, K. J., Mazzella, A. T., Horton, R. J., and McKenna, J. R. (2011). Phase and amplitude inversion of crosswell radar data. *Geophysics*, 76(3):J1–J12.

Ernst, J. R., Green, A. G., Maurer, H., and Holliger, K. (2007). Application of a new 2D time-domain full-waveform inversion scheme to crosshole radar data. *Geophysics*, 72(5):J53–J64.

Ernst, J. R., Holliger, K., Maurer, H., and Green, A. G. (2006). Realistic FDTD modelling of borehole georadar antenna radiation: methodology and application. *Near Surface Geophysics*, pages 19–30.

Evans, R., Frost, M., Stonecliffe-Jones, M., and Dixon, N. (2008). A review of pavement assessment using ground penetrating radar (GPR). In *12th International Conference on Ground Penetrating Radar (GPR 2008), Birmingham (UK).*

Feynman, R. P., Leighton, R. B., and Sands, M. (1979). *Le cours de physique de Feynman - Electromagnétisme 1.* InterEditions, Paris.

Fink, M., Cassereau, D., Derode, A., Prada, C., Roux, P., Tanter, M., Thomas, J.-L., and Wu, F. (2000). Time-reverse acoustics. *Reports on Progress in Physics*, 63:1933–1995.

Fischer, E., McMechan, G. A., and Annan, A. P. (1992a). Acquisition and processing of wide-aperture ground-penetrating radar data. *Geophysics*, 57(3):495–504.

Fischer, E., McMechan, G. A., Annan, A. P., and Cosway, S. W. (1992b). Examples of reverse-time migration of single-channel, ground-penetrating radar profiles. *Geophysics*, 57(4):577–586.

Fletcher, R. and Reeves, C. M. (1964). Function minimization by conjugate gradient. *Computer Journal*, 7:149–154.

Forte, E. and Pipan, M. (2008). Integrated seismic tomography and ground-penetrating radar (GPR) for the high-resolution study of burial mounds. *Journal of Archaeological Science*, 35:2614–2623.

Francke, J. (2012). A review of selected ground penetrating radar applications to mineral resource evaluations. *Journal of Applied Geophysics*, 81:29–37.

Fruehauf, F., Heilig, A., Schneebeli, M., Fellin, W., and Scherzer, O. (2009). Experiments and algorithms to detect snow avalanche victims using airborne ground-penetrating radar. *IEEE Transactions on Geoscience and Remote Sensing*, 47(7):2240–2251.

Gaffney, C. (2008). Detecting trends in the prediction of the buried past: A review of geophysical techniques in archaeology. *Archaeometry*, 50(2):313–336.

Garambois, S., Sénéchal, P., and Perroud, H. (2002). On the use of combined geophysical methods to assess water content and water conductivity of near-surface formations. *Journal of Hydrology*, 259:32–48.

Garry, B. (2007). *Etude des processus d'écoulements de la zone non saturée pour la modélisation des aquifères karstiques: expérimentation hydrodynamique et hydrochimique sur les sites du laboratoire souterrain à bas bruits (LSBB) de Rustrel et de Fontaine de Vaucluse*. PhD thesis, Université d'Avignon.

Geffrin, J.-M., Sabouroux, P., and Eyraud, C. (2005). Free space experimental scattering database continuation: experimental set-up and measurement precision. *Inverse Problems*, 21:S117–S130.

Gerhards, H., Wollschläger, U., Yu, Q., Schiwek, P., Pan, X., and Roth, K. (2008). Continuous and simultaneous measurement of reflector depth and average soil-water content with multichannel ground-penetrating radar. *Geophysics*, 73(4):J15–J23.

Giannopoulos, A. (2005). Modelling ground penetrating radar by GprMax. *Construction and Building Materials*, 19:755–762.

Girard, J.-F. (2002). *Imagerie géoradar et modélisation des diffractions multiples*. PhD thesis, Université Louis Pasteur, Strasbourg 1.

Gloaguen, E., Marcotte, D., Chouteau, M., and Perroud, H. (2005). Borehole radar velocity inversion using cokriging and cosimulation. *Journal of Applied Geophysics*, 57:242–259.

Goodman, D. (1994). Ground-Penetrating Radar simulation in engineering and archaeology. *Geophysics*, 59:224–232.

Grandjean, G., Gourry, J.-C., and Bitri, A. (2000). Evaluation of GPR techniques for civil-engineering applications: study on a test site. *Journal of Applied Geophysics*, 45:141–156.

Grasmueck, M., Marchesini, P., Eberli, G. P., Zeller, M., and Van Dam, R. L. (2010). 4D GPR tracking of water infiltration in fractured high-porosity limestone. In *13th International Conference on Ground Penetrating Radar (GPR 2010), Lecce (Italy)*.

Grasmueck, M., Weger, R., and Horstmeyer, H. (2005). Full-resolution 3D GPR imaging. *Geophysics*, 70(1):K12–K19.

Grazzini, G., Pieraccini, M., Parrini, F., Spinetti, A., Macaluso, G., Dei, D., and Atzeni, C. (2010). An Ultra-Wideband High-Dynamic Range GPR for detecting buried people after collapse of buildings. In *10th International Conference on Ground Penetrating Radar (GPR 2004), Delft (The Nederlands)*.

Greaves, R. J., Lesmes, D. P., Lee, J. M., and Toksöz, M. N. (1996). Velocity variations and water content estimated from multi-offset, ground-penetrating radar. *Geophysics*, 61(3):683–695.

Grégoire, C. (2001). *Fracture characterization by Ground Penetrating Radar*. PhD thesis, Katholieke Universiteit Leuven.

Grégoire, C. and Hollender, F. (2004). Discontinuity characterization by the inversion of the spectral content of ground-penetrating radar (GPR) reflections—Application of the Jonscher model. *Geophysics*, 69(6):1414–1424.

Hadamard, J. (1902). Sur les problèmes aux dérivés partielles et leur signification physique. *Princeton University Bulletin*, pages 49–52.

Hak, B. and Mulder, W. A. (2010). Migration for velocity and attenuation perturbations. *Geophysical Prospecting*, 58:939–951.

Hak, B. and Mulder, W. A. (2011). Seismic attenuation imaging with causality. *Geophysical Journal International*, 184(1):439–451.

Hansen, P. C. (2010). *Discrete Inverse Problems: Insight and Algorithms*. SIAM.

Harrington, R. F. (1993). *Field Computation by Moment Methods*. IEEE Press.

He, X.-Q., Zhu, Z.-Q., Liu, Q.-Y., and Lu, G.-Y. (2009). Review of GPR rebar detection. In *PIERS Proceedings, Beijing (China)*.

Heilig, A., Eisen, O., and Schneebeli, M. (2010). Temporal observations of a seasonal snowpack using upward-looking GPR. *Hydrological Processes*, 24:3133–3145.

Hesse, A. (1999). Multi-parametric survey for archaeology: how and why, or how and why not? *Journal of Applied Geophysics*, 41:157–168.

Hestenes, M. R. and Stiefel, E. (1952). Methods of conjugate gradients for solving linear systems. *Journal of Research of the National Bureau of Standards*, 49(6):409–436.

Hicks, G. J. (2002). Arbitrary source and receiver positioning in finite-difference schemes using kaiser windowed sinc functions. *Geophysics*, 67:156–166.

Hill, R. M. and Dissado, L. A. (1985). Debye and non-Debye relaxation. *Journal of Physics C: Solid State Physics*, 18:3829–3836.

Hinz, E. A. and Bradford, J. H. (2010). Ground-penetrating-radar reflection attenuation tomography with an adaptive mesh. *Geophysics*, 75(4):WA251–WA261.

Hollender, F. and Tillard, S. (1998). Modeling ground-penetrating radar wave propagation and reflection with the Jonscher parameterization. *Geophysics*, 63(6):1933–1942.

Holliger, K., Musil, M., and Maurer, H. (2001). Ray-based amplitude tomography for crosshole georadar data: a numerical assessment. *Journal of Applied Geophysics*, 47:285–298.

Huisman, J. A., Hubbard, S. S., Redman, J. D., and Annan, A. P. (2003). Measuring soil water content with ground penetrating radar: A review. *Vadose Zone Journal*, 2:476–491.

Hustedt, B., Operto, S., and Virieux, J. (2004). Mixed-grid and staggered-grid finite difference methods for frequency domain acoustic wave modelling. *Geophysical Journal International*, 157:1269–1296.

Ihamouten, A., Chahine, K., Baltazart, V., Villain, G., and Derobert, X. (2011). On variants of the frequency power law for the electromagnetic characterization of hydraulic concrete. *IEEE Transactions on Instrumentation and Measurement*, 60(11):3658–3668.

Ihamouten, A., Villain, G., and Dérobert, X. (2012). Complex permittivity frequency variations from multioffset GPR data: Hydraulic concrete characterization. *IEEE Transactions on Geoscience and Remote Sensing*, 61(6):1636–1648.

Instanes, A., Lønne, I., and Sandaker, K. (2004). Location of avalanche victims with ground-penetrating radar. *Cold regions science and technology*, 38:55–61.

Jackson, D. D. (1972). Interpretation of inaccurate, insufficient and inconsistent data. *Geophysical Journal of the Royal Astronomical Society*, 28:97–109.

Jackson, D. D. (1979). The use of a priori data to resolve non-uniqueness in linear inversion. *Geophysical Journal of the Royal Astronomical Society*, 57(1):137–157.

Jadoon, K. Z., Lambot, S., Scharnagl, B., van der Kruk, J., Slob, E., and Vereecken, H. (2010). Quantifying field-scale surface soil water content from proximal GPR signal inversion in the time domain. *Near Surface Geophysics*, 8(6):483–491.

Jeannin, M. (2005). *Etude des processus d'instabilités des versants rocheux par prospection géophysique - Apport du radar géologique*. PhD thesis, Université Joseph Fourier, Grenoble 1.

Jeannin, M., Garambois, S., Grégoire, C., and Jongmans, D. (2006). Multiconfiguration GPR measurements for geometric fracture characterization in limestone cliffs (Alps). *Geophysics*, 71(3):B85–B92.

Jiao, Y., McMechan, G. A., and Pettinelli, E. (2000). In situ 2-D and 3-D measurements of radiation patterns of half-wave dipole GPR antennas. *Journal of Applied Geophysics*, 43:69–89.

Jo, C. H., Shin, C., and Suh, J. H. (1996). An optimal 9-point, finite-difference, frequency-space 2D scalar extrapolator. *Geophysics*, 61:529–537.

Jol, H. M. (2009). *Ground Penetrating Radar: Theory and Applications*. Elsevier.

Jongmans, D. and Garambois, S. (2007). Geophysical investigation of landslides : a review. *Bulletin de la Société Géologique de France*, 178(2):101–112.

Jonscher, A. (1977). The universal dielectric response. *Nature*, 267(5613):673–679.

Jonscher, A. K. (1981). A new understanding of the dielectric relaxation of solids. *Journal of Materials Science*, 16:2037–2060.

Jonscher, A. K. (1999). Dielectric relaxation in solids. *Journal of Physics D: Applied Physics*, 32:R57–70.

Kalogeropoulos, A., van der Kruk, J., Hugenschmidt, J., Bikowski, J., and Brühwiler, E. (2013). Full-waveform GPR inversion to assess chloride gradients in concrete. *NDT&E International*, 57:74–84.

Kalogeropoulos, A., van der Kruk, J., Hugenschmidt, J., Busch, S., and Merz, K. (2011). Chlorides and moisture assessment in concrete by GPR full waveform inversion. *Near Surface Geophysics*, 9(3):277–285.

Kamei, R. and Pratt, R. G. (2008). Waveform tomography strategies for imaging attenuation structure for cross-hole data. In *Extended Abstracts*, page F019.

Kamei, R. and Pratt, R. G. (2013). Inversion strategies for visco-acoustic waveform inversion. *Geophysical Journal International*, 194:859–894.

Kim, J.-H., Supper, R., Tsourlos, P., and Yi, M.-J. (2012). 4D Inversion of resistivity monitoring data through lp norm minimizations. In *Near Surface Geoscience 2012 – 18th European Meeting of Environmental and Engineering Geophysics, Paris (France)*.

Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, 220(4598):671–680.

Klotzsche, A., van der Kruk, J., Linde, N., Doetsch, J., and Vereecken, H. (2013). 3-D characterization of high-permeability zones in a gravel aquifer using 2-D crosshole GPR full-waveform inversion and waveguide detection. *Geophysical Journal International*, 195(2):932–944.

Klotzsche, A., van der Kruk, J., Meles, G. A., Doetsch, J. A., Maurer, H., and Linde, N. (2010). Full-waveform inversion of cross-hole ground-penetrating radar data to characterize a gravel aquifer close to the Thur River, Switzerland. *Near Surface Geophysics*, 8.

Klotzsche, A., van der Kruk, J., Meles, G. A., and Vereecken, H. (2012). Crosshole GPR full-waveform inversion of waveguides acting as preferential flow paths within aquifer systems. *Geophysics*, 77(4):H57–H62.

Kneisel, C., Hauck, C., Fortier, R., and Moorman, B. (2008). Advances in geophysical methods for permafrost investigations. *Permafrost and Periglacial Processes*, 19(2):157–178.

Kwon, M. J. (2013). Two unconventional approaches to electromagnetic inversion: Hierarchical Bayesian Inversion and Inverse Scattering Series. Master's thesis, Colorado School of Mines.

Lagarias, J. C., Reeds, J. A., Wright, M. H., and Wright, P. E. (1998). Convergence properties of the Nelder-Mead simplex method in low dimensions. *SIAM Journal on Optimization*, 9(1):112–147.

Lahoz, W., Khattatov, B., and Ménard, R., editors (2010). *Data Assimilation*. Springer.

Lailly, P. (1983). The seismic inverse problem as a sequence of before stack migrations. In Bednar, R. and Weglein, editors, *Conference on Inverse Scattering, Theory and application, Society for Industrial and Applied Mathematics, Philadelphia*, pages 206–220.

Lambot, S., André, F., Jadoon, K. Z., Slob, E. C., and Vereecken, H. (2010). Full-waveform modeling of ground-coupled GPR antennas for wave propagation in multilayered media: the problem solved? In *13th International Conference on Ground Penetrating Radar (GPR 2010), Lecce (Italy)*.

Lambot, S., Slob, E. C., van den Bosch, I., Stockbroeckx, B., and Vanclooster, M. (2004). Modeling of ground-penetrating radar for accurate characterization of subsurface electric properties. *IEEE Transactions on Geoscience and Remote Sensing*, 42(11):2555–2568.

Lambot, S., Weihermüller, L., Huisman, J. A., Vereecken, H., Vanclooster, M., and Slob, E. C. (2006). Analysis of air-launched ground-penetrating radar techniques to measure the soil surface water content. *Water Resources Research*, 42:W11403.

Lampe, B. and Holliger, K. (2003). Effects of fractal fluctuations in topographic relief, permittivity and conductivity on ground-penetrating radar antenna radiation. *Geophysics*, 68(6):1934–1944.

Lampe, B. and Holliger, K. (2005). Resistively loaded antennas for ground-penetrating radar: A modeling approach. *Geophysics*, 70(3):K23–K32.

Landau, L. D. and Lifchitz, E. M. (1969). *Electrodynamique des milieux continus*. Physique Théorique, tome 8. Editions MIR.

Lavoué, F., Brossier, R., Métivier, L., Garambois, S., and Virieux, J. (2013). 2D full waveform inversion of GPR surface data: permittivity and conductivity imaging. In *7th International Workshop on Advanced Ground-Penetrating Radar (IWAGPR 2013), Nantes (France)*.

Lavoué, F., Brossier, R., Métivier, L., Garambois, S., and Virieux, J. (2014). Two-dimensional permittivity and conductivity imaging by full waveform inversion of multioffset GPR data: a frequency-domain quasi-Newton approach. *Geophysical Journal International*, 197(1):248–268.

Lavoué, F., Brossier, R., Métivier, L., Garambois, S., and Virieux, J. (2015). Frequency-domain modelling and inversion of electromagnetic data for 2D permittivity and conductivity imaging: An application to the Institut Fresnel experimental dataset. *Near Surface Geophysics*, (accepted).

Leonide, P., Borgomano, J., Masse, J.-P., and Doublet, S. (2012). Relation between stratigraphic architecture and multi-scale heterogeneities in carbonate platforms: The Barremian-lower Aptian of the Monts de Vaucluse, SE France. *Sedimentary Geology*, 265:87–109.

Leparoux, D. (1997). *Mise au point de méthodes radar pour l'auscultation structurale et texturale de milieux géologiques très hétérogènes*. PhD thesis, Université de Rennes 1.

Levin, F. K. (1971). Apparent velocity from dipping interface reflections. *Geophysics*, 36(3):510–516.

Lions, J. (1972). *Nonhomogeneous boundary value problems and applications*. Springer Verlag, Berlin.

Litman, A. and Crocco, L. (2009). Testing inversion algorithms against experimental data: 3D targets. *Inverse Problems*, 25(2).

Loeffler, O. (2005). *Modélisation géoradar de la proche surface, estimation de la teneur en eau et influence d'un polluant*. PhD thesis, Université Louis Pasteur, Strasbourg 1.

Lopes, F. (2009). *Inversion des formes d'ondes électromagnétiques de données radar (GPR) multioffsets*. PhD thesis, Université Denis Diderot, Paris 7.

Lutz, P. (2002). *Acquisitions multi-modes en radar géologique de surface*. PhD thesis, Université de Pau et des Pays de l'Adour.

Lutz, P., Garambois, S., and Perroud, H. (2003). Influence of antenna configurations for GPR survey: information from polarization and amplitude versus offset measurements. *Geological Society, London, Special Publications*, 211(1):299–313.

Ma, Y., Hale, D., Gong, B., and Meng, Z. J. (2012). Image-guided sparse-model full waveform inversion. *Geophysics*, 77(4):R189–R198.

Malinowski, M., Operto, S., and Ribodetti, A. (2011). High-resolution seismic attenuation imaging from wide-aperture onshore data by visco-acoustic frequency-domain full waveform inversion. *Geophysical Journal International*, 186(3):1179–1204.

Matonti, C., Viseur, S., and Boyer, D. (2011). Modèle GOCAD du LSBB, Rustrel, France. Technical report, GSRC & LSBB.

Maufroy, E., Gaffet, S., Operto, S., Guglielmi, Y., and Boyer, D. (2014). Travel time inversion from ground level to gallery: protocol for the characterization of P-wave signature in a fractured-porous Urgonian platform at hectometric scale. *Near surface Geophysics*, 12:doi:10.3997/1873–0604.2014025.

Maurer, H., Greenhalgh, S. A., Manukyan, E., Marelli, S., and Green, A. G. (2012). Receiver-coupling effects in seismic waveform inversions. *Geophysics*, 77(1):R57–R63.

Maxwell, J. C. (1873). *A Treatise on Electricity and Magnetism*. Clarendon Press, Oxford.

McCann, D. M. and Forde, M. C. (2001). Review of NDT methods in the assessment of concrete and masonry structures. *NDT&E International*, 34:71–84.

McClymont, A. F., Green, A. G., Streich, R., Horstmeyer, H., Tronicke, J., Nobes, D. C., Pettinga, J., Campbell, J., and Langridge, R. (2008). Visualization of active faults using geometric attributes of 3D GPR data: An example from the Alpine Fault Zone, New Zealand. *Geophysics*, 73(2):B11–B23.

Meles, G., Greenhalgh, S., Maurer, H., and Green, A. (2013). Some ideas yet unattempted in Georadar full waveform inversion. In *PIERS Proceedings, Stockholm (Sweden)*, pages 714–720.

Meles, G. A. (2011). *New Developments in Full Waveform Inversion of GPR Data*. PhD thesis, Swiss Federal Institute of Technology Zürich (ETHZ).

Meles, G. A., Greenhalgh, S., van der Kruk, J., Green, A. G., and Maurer, H. (2011). Taming the non-linearity problem in GPR full-waveform inversion for high contrast media. *Journal of Applied Geophysics*, 73:174–186.

Meles, G. A., Greenhalgh, S. A., Green, A. G., Maurer, H., and van der Kruk, J. (2012). GPR full-waveform sensitivity and resolution analysis using an FDTD adjoint method. *IEEE Transactions on Geoscience and Remote Sensing*, 50(5):1881–1896.

Meles, G. A., van der Kruk, J., Greenhalgh, S. A., Ernst, J. R., Maurer, H., and Green, A. G. (2010). A new vector waveform inversion algorithm for simultaneous updating of conductivity and permittivity parameters from combination crosshole/borehole-to-surface GPR data. *IEEE Transactions on Geoscience and Remote Sensing*, 48:3391–3407.

Menke, W. (1984). *Geophysical Data Analysis: Discrete Inverse Theory*. Academic Press, Inc., Orlando, USA.

Menke, W. (2012). *Geophysical Data Analysis: Discrete Inverse Theory*. Academic Press, 3rd edition.

Métivier, L., Brossier, R., Virieux, J., and Operto, S. (2013). Full Waveform Inversion and the truncated Newton method. *SIAM Journal On Scientific Computing*, 35(2):B401–B437.

Metje, N., Atkins, P. R., Brennan, M. J., Chapman, D. N., Lim, H. M., Machell, J., Muggleton, J. M., Pennock, S., Ratcliffe, J., Redfern, M., Rogers, C. D. F., Saul, A. J., Shan, Q., Swingler, S., and Thomas, A. M. (2007). Mapping the Underworld – State-of-the-art review. *Tunnelling and Underground Space Technology*, 22:568–586.

Metropolis, N. and Ulam, S. (1949). The Monte Carlo Method. *Journal of the American Statistical Association*, 44(247):335–341.

Minet, J., Lambot, S., Slob, E. C., and Vanclooster, M. (2010). Soil surface water content estimation by full-waveform GPR signal inversion in the presence of thin layers. *IEEE Transactions on Geoscience and Remote Sensing*, 48(3):1138–1150.

Minet, J., Wahyudi, A., Bogaert, P., Vanclooster, M., and Lambot, S. (2011). Mapping shallow soil moisture profiles at the field scale using full-waveform inversion of ground penetrating radar data. *Geoderma*, 161(3-4):225–237.

Modroo, J. J. and Olhoeft, G. R. (2004). Avalanche rescue using ground penetrating radar. In *10th International Conference on Ground Penetrating Radar (GPR 2004), Delft (The Nederlands)*.

Moysey, S. and Knight, R. J. (2004). Modeling the field-scale relationship between dielectric constant and water content in heterogeneous systems. *Water Resources Research*, 40:W03510.

Moysey, S., Knight, R. J., and Jol, H. M. (2006). Texture-based classification of ground-penetrating radar images. *Geophysics*, 71(6):K111–K118.

Moysey, S. M. J. (2010). Hydrologic trajectories in transient ground-penetrating-radar reflection data. *Geophysics*, 75(4):WA211–WA219.

Mulder, W. A. and Hak, B. (2009). An ambiguity in attenuation scattering imaging. *Geophysical Journal International*, 178:1614–1624.

MUMPS-team (2009). *MUMPS - MUltifrontal Massively Parallel Solver users' guide - version 4.9.2 (November 5, 2009)*. ENSEEIHT-ENS Lyon, http://mumps.enseeiht.fr or http://graal.ens-lyon.fr/MUMPS. Accessed 15 January 2014.

MUMPS-team (2011). *MUMPS - MUltifrontal Massively Parallel Solver users' guide - version 4.10.0 (May 10, 2011).* ENSEEIHT-ENS Lyon, http://mumps.enseeiht.fr or http://graal.ens-lyon.fr/MUMPS. Accessed 15 January 2014.

Musil, M., Maurer, H., Holliger, K., and Green, A. G. (2006). Internal structure of an alpine rock glacier based on crosshole georadar traveltimes and amplitudes. *Geophysical Prospecting*, 54:273–285.

Neal, A. (2004). Ground-penetrating radar and its use in sedimentology: principles, problems and progress. *Earth-Science Reviews*, 66:261–330.

Nocedal, J. and Wright, S. J. (2006). *Numerical Optimization*. Springer, 2nd edition.

Olhoeft, G. R. and Modroo, J. J. (2006). Locating and identfying avalanche victims with GPR. *The Leading Edge*, 25(3):306–308.

Operto, S., Brossier, R., Gholami, Y., Métivier, L., Prieux, V., Ribodetti, A., and Virieux, J. (2013). A guided tour of multiparameter full waveform inversion for multicomponent data: from theory to practice. *The Leading Edge*, Special section Full Waveform Inversion(September):1040–1054.

Operto, S., Virieux, J., Dessa, J. X., and Pascal, G. (2006). Crustal imaging from multifold ocean bottom seismometers data by frequency-domain full-waveform tomography: application to the eastern Nankai trough. *Journal of Geophysical Research*, 111(B09306):doi:10.1029/2005JB003835.

Operto, S., Virieux, J., Ribodetti, A., and Anderson, J. E. (2009). Finite-difference frequency-domain modeling of visco-acoustic wave propagation in two-dimensional TTI media. *Geophysics*, 74 (5):T75–T95.

Patriarca, C., Lambot, S., Mahmoudzadeh, M. R., Minet, J., and Slob, E. (2011). Reconstruction of sub-wavelength fractures and physical properties of masonry media using full-waveform inversion of proximal penetrating radar. *Journal of Applied Geophysics*, 74:26–37.

Perineau, A. (2013). *Caractérisation directe du rôle et du fonctionnement de la zone non saturée des aquifères karstiques pour le développement d'un modèle de gestion active de ces aquifères en région méditerranéenne : Application au Laboratoire Souterrain Bas Bruit de Rustrel (LSBB) et au bassin versant expérimental de la Fontaine de Vaucluse*. PhD thesis, Université d'Avignon.

Plaut, J. J., Picardi, G., Safaeinili, A., Ivanov, A. B., Milkovich, S. M., Cicchetti, A., Kofman, W., Mouginot, J., Farrell, W. M., Phillips, R. J., Clifford, S. M., Frigeri, A., Orosei, R., Federico, C., Williams, I. P., Gurnett, D. A., Nielsen, E., Hagfors, T., Heggy, E., Stofan, E. R., Plettemeier, D., Watters, T. R., Leuschen, C. J., and Edenhofer, P. (2007). Subsurface radar sounding of the south polar layered deposits of Mars. *Science*, 316:92–95.

Plessix, R. E. (2006). A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International*, 167(2):495–503.

Poisson, J., Chouteau, M., Aubertin, M., and Campos, D. (2009). Geophysical experiments to image the shallow internal structure and the moisture distribution of a mine waste rock pile. *Journal of Applied Geophysics*, 67(2):179–192.

Polak, E. and Ribière, G. (1969). Note sur la convergence de méthodes de directions conjuguées. *Revue Française d'Informatique et de Recherche Opérationnelle*, 16:35–43.

Pratt, R. G. (1999). Seismic waveform inversion in the frequency domain, part I : theory and verification in a physic scale model. *Geophysics*, 64:888–901.

Pratt, R. G., Shin, C., and Hicks, G. J. (1998). Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion. *Geophysical Journal International*, 133:341–362.

Pratt, R. G. and Worthington, M. H. (1990). Inverse theory applied to multi-source cross-hole tomography. Part 1: Acoustic wave-equation method. *Geophysical Prospecting*, 38:287–310.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1992). *Numerical recipes in FORTRAN : the art of scientific computing; second edition*. Cambridge University press.

Prieux, V., Brossier, R., Operto, S., and Virieux, J. (2013). Multiparameter full waveform inversion of multicomponent OBC data from valhall. Part 1: imaging compressional wavespeed, density and attenuation. *Geophysical Journal International*, 194(3):1640–1664.

Pringle, J. K., Ruffell, A., Jervis, J. R., Donnelly, L., McKinley, J., Hansen, J., Morgan, R., Pirrie, D., and Harrison, M. (2012). The use of geoscience methods for terrestrial forensic searches. *Earth-Science Reviews*, 114:108–123.

Ravaut, C., Operto, S., Improta, L., Virieux, J., Herrero, A., and dell'Aversana, P. (2004). Multi-scale imaging of complex structures from multi-fold wide-aperture seismic data by frequency-domain full-wavefield inversions: application to a thrust belt. *Geophysical Journal International*, 159:1032–1056.

Remaki, M. (1999). *Méthodes numériques pour les équations de Maxwell instationnaires en milieu hétérogène*. PhD thesis, Ecole Nationale des Ponts et Chaussées.

Revil, A. (2013). Effective conductivity and permittivity of unsaturated porous materials in the frequency range 1 mHz-1GHz. *Water Resources Research*, 49:306–327.

Ribodetti, A., Operto, S., Virieux, J., Lambaré, G., Valéro, H.-P., and Gibert, D. (2000). Asymptotic viscoacoustic diffraction tomography of ultrasonic laboratory data : a tool for rock properties analysis. *Geophysical Journal International*, 140:324–340.

Rubin, Y. and Hubbard, S. S., editors (2005). *Hydrogeophysics*. Springer.

Rumpf, M. and Tronicke, J. (2014). Predicting 2D geotechnical parameter fields in near-surface sedimentary environments. *Journal of Applied Geophysics*, 101:95–107.

Rust, A. C., Russell, J. K., and Knight, R. J. (1999). Dielectric constant as a predictor of porosity in dry volcanic rocks. *Journal of Volcanology and Geothermal Research*, 91:79–96.

Sagnard, F., Tebchrany, E., and Baltazart, V. (2013). Evaluation of an UWB ground-coupled radar in the detection of discontinuities using polarization diversity: FDTD modeling and experiments. In *7th International Workshop on Advanced Ground-Penetrating Radar (IWAGPR 2013), Nantes (France)*, pages 153–158.

Saintenoy, A. (1998). *Radar géologique : acquisition de données multi-déports pour une mesure multi-paramètres*. PhD thesis, Université Denis Diderot, Paris 7.

Saintenoy, A., Friedt, J.-M., Booth, A. D., Tolle, F., Bernard, E., Laffly, D., Marlin, C., and Griselin, M. (2013). Deriving ice thickness, glacier volume and bedrock morphology of Austre Lovénbreen (Svalbard) using GPR. *Near Surface Geophysics*, 11(2):253–261.

Saintenoy, A., Schneider, S., and Tucholka, P. (2008). Evaluating Ground Penetrating Radar use for water infiltration monitoring. *Vadose Zone Journal*, 7(1):208–214.

Sambridge, M. (2001). Finding acceptable models in nonlinear inverse problems using a neighbourhood algorithm. *Inverse Problems*, 17:387–403.

Sambridge, M. and Mosegaard, K. (2002). Monte Carlo methods in geophysical inverse problems. *Reviews of Geophysics*, 40(3):1–29.

Sambridge, M. S. (1999a). Geophysical inversion with a neighbourhood algorithm - I. Searching a parameter space. *Geophysical Journal International*, 138:479–494.

Sambridge, M. S. (1999b). Geophysical inversion with a neighbourhood algorithm - II. Appraising the ensemble. *Geophysical Journal International*, 138:727–746.

Sassen, D. S. and Everett, M. E. (2009). 3D polarimetric GPR coherency attributes and full-waveform inversion of transmission data for characterizing fractured rock. *Geophysics*, 74(3):J23–J34.

Sauter, S. A. and Schwab, C. (2011). *Boundary Element Methods*. Springer.

Schrott, L. and Sass, O. (2008). Application of field geophysics in geomorphology: Advances and limitations exemplified by case studies. *Geomorphology*, 93(1):55–73.

Scollar, I., Tabbagh, A., Hesse, A., and Herzog, I. (1990). *Archaeological Prospection and Remote Sensing*. Topics in remote sensing. Cambridge University Press.

Sen, M. K. and Stoffa, P. L. (1995). *Global Optimization Methods in Geophysical Inversion*. Elsevier Science Publishing Co.

Sénéchal, G., Rousset, D., and Gaffet, S. (2013). Ground-penetrating radar investigation inside a karstified limestone reservoir. *Near Surface Geophysics*, 11:283–291.

Sénéchal, G., Rousset, D., Gloaguen, E., and Lebourg, T. (2004). High resolution seismic experiment within the low-noise underground laboratory (LSBB) of Rustrel (France). In *10th European Meeting of Environmental and Engineering Geophysics*.

Shin, C., Pyun, S., and Bednar, J. B. (2007). Comparison of waveform inversion, part 1: conventional wavefield vs logarithmic wavefield. *Geophysical Prospecting*, 55(4):449–464.

Sirgue, L. (2003). *Inversion de la forme d'onde dans le domaine fréquentiel de données sismiques grand offset*. PhD thesis, Université Paris 11, France - Queen's University, Canada.

Sirgue, L. and Pratt, R. G. (2004). Efficient waveform inversion and imaging : a strategy for selecting temporal frequencies. *Geophysics*, 69(1):231–248.

Slob, E., Sato, M., and Olhoeft, G. R. (2010). Surface and borehole ground-penetrating-radar developments. *Geophysics*, 75(5):75A103–75A120.

Sohl, C. (2008). *Dispersion relations in scattering and antenna problems*. PhD thesis, Lund University.

Sperl, C. (1999). *Determination of spatial and temporal variation of the soil water content in an agro-ecosystem with ground-penetrating radar*. PhD thesis, Technische Universität München, München, Germany.

Taflove, A. and Hagness, S. C. (2005). *Computational Electrodynamics: The Finite-Difference Time-Domain Method*. Artech House, 3rd edition.

Tarantola, A. (1984). Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, 49(8):1259–1266.

Tarantola, A. (2005). *Inverse Problem theory and methods for model parameter estimation*. Society for Industrial and Applied Mathematics, Philadelphia.

Tarantola, A. and Valette, B. (1982a). Generalized nonlinear inverse problems solved using the least square criterion. *Reviews of Geophysical and Space Physics*, 20:219–232.

Tarantola, A. and Valette, B. (1982b). Inverse problems = quest for information. *Journal of Geophysics*, 50:159–170.

Thiébaud, E. (2003). Etude structurale et hydrogéologique du site du Laboratoire Souterrain à Bas Bruit de Rustrel. Master's thesis, Université de Franche-Comté.

Tikhonov, A. and Arsenin, V. (1977). *Solution of ill-posed problems*. Winston, Washington, DC.

Toll, J. S. (1956). Causality and dispersion relation: Logical foundations. *Physical Review*, 104(6):1760–1770.

Topp, G. C., Davis, J. L., and Annan, A. P. (1980). Electromagnetic determination of soil water content: measurements in coaxial transmission lines. *Water Resources Research*, 16(3):574–582.

Tran, A. P., Ardekani, M. R. M., and Lambot, S. (2012). Coupling of dielectric mixing models with full-wave ground-penetrating radar signal inversion for sandy-soil-moisture estimation. *Geophysics*, 77(3):H33–H44.

Tronicke, J., Paasche, H., and Böniger, U. (2011). Joint global inversion of GPR and P-wave seismic traveltimes using particle swarm optimization. In *6th International Workshop on Advanced Ground-Penetrating Radar (IWAGPR 2011), Aachen (Germany).*

Tsoflias, G. P. and Becker, M. W. (2008). Ground-penetrating-radar response to fracture-fluid salinity: Why lower frequencies are favorable for resolving salinity changes. *Geophysics*, 73(5):J25–J30.

Turner, G. and Siggings, A. F. (1994). Constant Q attenuation of subsurface radar pulses. *Geophysics*, 59:1192–1200.

van der Kruk, J., Diamanti, N., Giannopoulos, A., and Vereecken, H. (2012). Inversion of dispersive GPR pulse propagation in waveguides with heterogeneities and rough and dipping interfaces. *Journal of Applied Geophysics*, 81:88–96.

van der Kruk, J., Jacob, R. W., and Vereecken, H. (2010). Properties of precipitation-induced multi-layer surface waveguides derived from inversion of dispersive TE and TM GPR data. *Geophysics*, 75(4):WA263–WA273.

van der Kruk, J., Streich, R., and Green, A. G. (2006). Properties of surface waveguides derived from separate and joint inversion of dispersive TE and TM GPR data. *Geophysics*, 71(1):K19–K29.

van der Kruk, J., Vereecken, H., and Jacob, R. W. (2009). Identifying dispersive GPR signals and inverting for surface wave-guide properties. *The Leading Edge*, 28(10):1234–1239.

Van Vorst, D., Yedlin, M., Guglielmi, Y., Cappa, F., Gaffet, S., Maxwell, M., and Oldenburg, D. (2010). GPR imaging of a fracture zone in the Vaucluse Karst Aquifer using 2D eikonal inversion. In *3rd inter-Disciplinary Underground Science & Technology Conference (i-DUST 2010), Rustrel (France).*

Vaughan, C. J. (1986). Ground-penetrating radar surveys used in archaeological investigations. *Geophysics*, 51(3):595–604.

Velichko, A. and Wilcox, P. D. (2010). A generalized approach for efficient finite element modeling of elastodynamic scattering in two and three dimensions. *Journal of Acoustical Society of America*, 128(3):1004–1014.

Vereecken, H., Binley, A., Cassiani, G., Revil, A., and Titov, K., editors (2006). *Applied Hydrogeophysics*. Springer.

Vincent, C., Descloitres, M., Garambois, S., Legchenko, A., Guyard, H., and Gilbert, A. (2012). Detection of a subglacial lake in Glacier de Tête Rousse (Mont Blanc area, France). *Journal of Glaciology*, 58(211):866–878.

Vincent, C., Garambois, S., Thibert, E., Lefèbvre, E., Meur, E. L., and Six, D. (2010). Origin of the outburst flood from Glacier de Tête Rousse in 1892 (Mont Blanc area, France). *Journal of Glaciology*, 56(198):688–698.

Virieux, J. (1996). Seismic ray tracing. In Boschi, E., Ekström, G., and Morelli, A., editors, *Seismic modelling of Earth structure*, pages 223–304. Editrice Compositori, Bologna.

Virieux, J. and Operto, S. (2009). An overview of full waveform inversion in exploration geophysics. *Geophysics*, 74(6):WCC1–WCC26.

Waldschmidt, G. and Taflove, A. (2000). The determination of the effective radius of a filamentary source in the FDTD mesh. *IEEE Microwave and Guided Wave Letters*, 10(6):217–219.

Walford, M. E. R. (1964). Radio echo sounding through an ice shelf. *Nature*, 204(4956):317–319.

Warren, C. and Giannopoulos, A. (2011). Creating finite-difference time-domain models of commercial ground-penetrating radar antennas using Taguchi's optimization method. *Geophysics*, 76(2):G37–G47.

Weihermüller, L., Huisman, J. A., Lambot, S., Herbst, M., and Vereecken, H. (2007). Mapping the spatial variation of soil water content at the field scale with different ground penetrating radar techniques. *Journal of Hydrology*, 340:205–216.

Wikipedia (2014). Electrical resistivity and conductivity - Wikipedia, the free encyclopedia. `http://en.wikipedia.org/wiki/Electrical_conductivity`. Accessed 18 January 2014.

Wilcox, P. D. and Velichko, A. (2010). Efficient frequency-domain finite element modeling of two-dimensional elastodynamic scattering. *Journal of Acoustical Society of America*, 127(1):155–165.

Xu, T. and McMechan, G. (1997). GPR attenuation and its numerical simulation in 2.5 dimensions. *Geophysics*, 62(2):403–414.

Yang, X., Klotzsche, A., Meles, G., Vereecken, H., and van der Kruk, J. (2013). Improvements in cross-hole GPR full-waveform inversion and application on data measured at the Boise Hydrogeophysics Research site. *Journal of Applied Geophysics*, 99:114–124.

Yang, X., van der Kruk, J., Bikowski, J., Kumbhar, P., Vereecken, H., and Meles, G. (2012). Full-waveform inversion of GPR data in frequency-domain. In *14th International Conference on Ground Penetrating Radar (GPR 2012), Shanghai (China)*, pages 324–328.

Yedlin, M., Cresp, A., Pichot, C., Aliferis, I., Dauvignac, J.-Y., Gaffet, S., and Sénéchal, G. (2009). Ultra-wideband microwave imaging of heterogeneities. *Journal of Applied Geophysics*, 68:17–25.

Yee, K. S. (1966). Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media. *IEEE Transactions on Antennas and Propagation*, 14:302–307.

Yilmaz, Ö. (1987). *Seismic data processing*. Society of Exploration Geophyscists.

Yokota, T. and Matsushima, J. (2004). Seismic waveform tomography in the frequency-space domain: selection of the optimal temporal frequency for inversion. *Exploration Geophysics*, 35(1):19–24.

Yuffa, A. J. and Scales, J. A. (2012). Linear response laws and causality in electrodynamics. *European Journal of Physics*, 33:1635–1650.

Zhang, H. and Weglein, A. (2009). Direct nonlinear inversion of 1D acoustic media using inverse scattering subseries. *Geophysics*, 74(6):WCD29–WCD39.

Zhao, Y., Xie, X., Wu, J., Chen, J., and Ge, S. (2013). Maxwell curl equation datuming for GPR based on the Kirchhoff integral solution and application in a tunnel grouting test. *Near Surface Geophysics*, 11:211–219.

Zhou, J., Revil, A., Karaoulis, M., Hale, D., Doetsch, J., and Cuttler, S. (2014). Image-guided inversion of electrical resistivity data. *Geophysical Journal International*, 197(1):292–309.

Zhu, C., Byrd, R. H., and Nocedal, J. (1997). Algorithm 778: L-BFGS-B, FORTRAN routines for large scale bound constrained optimization. *ACM Transactions on Mathematical Software*, 23(4):550–560.

List of notations

Hereafter is a list of some notations and abbreviations which the reader can find throughout the manuscript and which are generally defined only at their first use. This list might be not exhaustive and does not include notations which appear locally in the text.

*	complex conjugate operator		
†	adjoint operator		
$\sim$	*of the order of*		
$\simeq$	*approximately equal to*		
$\propto$	*proportional to*		
		$\mathbf{E}$	electric field vector (in V/m)
α	attenuation coefficient (in m^{-1})	eq.	equation
$\boldsymbol{\alpha}$	descent step length		
A	Ampère units	F	Farad units
$\mathcal{A}$	impedance matrix	f	frequency (Hz)
AVO	amplitude-vs-offset	f_x, f_z	mechanical forces
AW	(direct) air-wave	Fig.	Figure
		freq.	frequency
β	propagation wavenumber (in m^{-1})	FWI	full waveform inversion
$\boldsymbol{\beta}$	parameter scaling factor		
$\mathbf{B}$	magnetic induction vector (in T)	$\mathbf{G}$	gradient vector
$\mathcal{B}$	approximate Hessian matrix	GPR	ground-penetrating radar
		GW	(direct) ground-wave
C	Coulomb units		
$\mathscr{C}$	misfit function	H	Hankel function
$\mathcal{C}_d$	data covariance matrix	$\mathbf{H}$	magnetic field vector (in A/m)
$\mathbb{C}$	ensemble of complex numbers	$\mathcal{H}$	Hessian matrix
CMP	common mid-point	Hz	Hertz units
δ	perturbation	$\imath$	imaginary unit
$\delta(x)$	Dirac delta function	$\Im m$	imaginary part
$\tan\delta$	loss tangent	it.	iteration(s)
$\boldsymbol{\Delta}\mathbf{m}$	model perturbation vector		
$\mathbf{D}$	electrical induction vector (in $\mathrm{C/m}^2$)	$\mathbf{J}$	current density vector (in $\mathrm{A/m}^2$)
$\mathbb{D}$	data space ($= \mathbb{C}^{N_D}$)	$\mathcal{J}$	Jacobian matrix
$\mathbf{d}$	data vector		
d_{obs}	observed data	k	iteration number
d_{cal}	synthetic data	k	complex wavenumber
		K	bulk modulus (in Pa)
ε	dielectric permittivity (in F/m)		
ε_o	dielectric constant in vacuum	λ	wavelength (in m)
	($\varepsilon_o = 1/(\mu_o v_o^2) \simeq 8.85 \times 10^{-12}$ F/m)	λ	Tikhonov regularisation weight
$\varepsilon_r = \varepsilon/\varepsilon_o$	relative permittivity	$\mathscr{L}$	Lagrangian functional

L-BFGS-B	Limited Broyden-Fletcher-Goldfarb-Shanno Bounded algorithm
LSBB	Laboratoire Souterrain à Bas Bruit (Low Noise Underground Lab)
LU factorisation	Lower Upper factorisation
μ	magnetic permeability (in H/m; in vacuum, $\mu_o = 4\pi \times 10^{-7}$ H/m)
$\mathbf{m}$	model vector
$\mathbf{m}_o$	starting model
$\mathbf{m}_k$	current model at iteration k
m_i	model parameter at grid point i
$\mathbb{M}$	model space ($= \mathbb{R}^{N_M}$)
MRW	multiply-reflected wave
N_D	number of data
N_M	number of model parameters
N_s	number of sources
N_ω	number of inverted frequencies
NMO	normal move-out
P	acoustic pressure (in Pa)
Pa	Pascal units
q	electric charge density (in C/m^3)
ρ	mass density (in kg/m^3)
$\mathbf{r}$	position vector
$\mathcal{R}$	restriction operator to receivers
$\mathbb{R}$	ensemble of real numbers
$\Re e$	real part
RW	reflected wave
RAW	reflected air-wave
RRW	reflected-refracted wave
S	Siemens units
σ	electrical conductivity (in S/m)
σ_o	reference conductivity (in S/m)
$\sigma_r = \sigma/\sigma_o$	relative conductivity
SNR	signal over noise ratio
src	source
SVD	singular value decomposition
t	time (in s)
T	Tesla units
^{T}	transposition operator
TE mode	Transverse Electric mode
TM mode	Transverse Magnetic mode
$\mathbf{u}$	wavefield vector
u_{inc}	incident wavefield
u_{sc}	scattered wavefield
$u_{tot} = u_{inc} + u_{sc}$	total wavefield
v	wave propagation velocity (in m/s)
v_o	light velocity in free space (v_o = 299,792,458 m/s)
v_x, v_z	acoustic particle velocities (in m/s)
$\mathbf{v}$	adjoint wavefield
V	Volt units
ω	angular frequency (in rad/s)
$\mathcal{W}_d$	data weighting matrix
χ	susceptibility